On Periodic Correlation of Binary Sequences

by

Nam Yul Yu

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A (periodic) correlation function is an important measure to evaluate the effectiveness of pseudorandom sequences. In practice, the sequences are required to have the impulse-like autocorrelation function. Also, crosscorrelation magnitudes of the distinct sequences must be as low as possible. Theoretically, the correlation of sequences has a strong connection with problems of algebraic coding and combinatorics. Namely, the correlation distribution of sequences is related to the weight distribution of codewords in algebraic codes. Furthermore, binary sequences with ideal two-level autocorrelation are equivalent to cyclic Hadamard difference sets in combinatorics. Therefore, a good knowledge of coding theory and combinatorics is helpful for a study of sequences with low correlation.

In this thesis, the correlation of binary sequences is studied by the aid of fruitful results of coding theory and combinatorics. From this study, some interesting properties are presented on the correlation of binary sequences, and several new binary sequences with low correlation are discovered. The outline of thesis is as follows.

First of all, crosscorrelation properties of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation are studied for odd n. As a basic approach, the Hadamard equivalence of binary sequences is exploited. The 3- and 5-valued crosscorrelation properties of several classes of binary sequences with ideal two-level autocorrelation are completely determined by either theoretical proofs or conjectures. Second, new binary sequences of period $4(2^m - 1)$ for even $m \geq 4$ are proposed by making use of a $4 \times (2^m - 1)$ interleaved structure of sequences. The sequences are almost balanced and have three-valued out-of-phase autocorrelation, i.e., $\{0, \pm 4\}$, which is optimal with respect to autocorrelation magnitude. Third, a new binary sequence family with low correlation and a large family size is proposed by generalizing the Gold-like and the Udaya's sequences. The weight distribution of subcodes of the second order Reed Muller codes is discussed and studied to derive the correlation distribution of the sequence family. With respect to maximum correlation, a family size, and a linear complexity, it is shown that the sequence family of period $2^n - 1$ for odd n is competitive among all known binary sequence families. Finally,

the construction of all quadratic bent functions represented by a polynomial form of $f(x) = \sum_{i=1}^{n/2-1} c_i Tr(x^{1+2^i}) + c_{n/2} Tr_1^{n/2}(x^{1+2^{n/2}}), x \in GF(2^n), c_i \in GF(2)$, is presented for special values of n by providing a necessary and sufficient condition on c_i's. The resulting quadratic bent functions can be employed for the linear feedback shift register (LFSR) implementation of a family of bent sequences, which is a potential candidate for future code-division multiple access (CDMA) system.

Acknowledgements

I would like to thank my supervisor, Professor Guang Gong, for her guidance and encouragement. I highly appreciate her passion and time for training me, and her constructive comments and criticism for leading me to the right way. I appreciate Professor Lloyd R. Welch at the University of Southern California for his valuable comments on my thesis. I also appreciate the other members of my examining committee, Professor Weihua Zhuang, Professor Mohamed Oussama Damen, and Professor R. Bruce Richter at the University of Waterloo for helping me improve the quality of my thesis by their insightful comments. It is my privilege to have such a great committee.

This research was supported by funds from the Natural Sciences and Engineering Research Council of Canada (NSERC) which I thank for the financial support of my research. I would like thank my lab colleagues, Katrin Hoepper, Yassir Nawaz, Xinxin Fan, and Zhijun Li for their kindness and supports. I also thank my Korean friend Jaewook Chung for helping me settle down in this place.

I would like to thank my parents and parents-in-law for their attention and supports to my study. Most of all, I am thankful to my lovely wife, Eun Jung. Without her sacrifice, I would have never been able to make this achievement. I appreciate her dedication of the best days of her life to the support of my Ph. D. study. Also, I would like to thank my son, Ji Min, for growing up very well in this new place. As a token of my gratitude, I dedicate my thesis to my wife and son.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Pseudorandom sequences are generated by deterministic methods, which are periodic or aperiodic with certain randomness properties. In this chapter, the randomness properties of binary sequences are introduced. For practical applications, sequences for communication and cryptographic systems are then reviewed. It is convinced from the review that sequences with low auto and crosscorrelation are essential and challenging for practical systems. Then, we presented a thesis overview by giving an outline of the works in this thesis. We concluded this chapter by providing basic definitions and concepts which are necessary for understanding next chapters.

1.1 Randomness Properties of Binary Sequences

Historically, Golomb [32] firstly established some statistical properties which should be satisfied by pseudorandom sequences generated by linear feedback shift registers (LFSRs). The *Golomb's randomness postulates* are necessary conditions for a binary periodic sequence to look random.

- Balance property: The number of 0's is nearly equal to the number of 1's in a period of a binary sequence.
- Run property: For a period of a binary sequence, k consecutive ones (or zeros)

preceded and followed by zeros (or ones) are called a run of ones (or zeros) of length k. In a period of a sequence, there are equally many numbers of runs of 0's and 1's for each length of runs.

– Ideal two-level autocorrelation: Every out-of-phase autocorrelation of a binary sequence is -1.

Including the Golomb's randomness postulates, there are several other randomness criteria which should be satisfied by pseudorandom sequences [33].

– Large period.

– Ideal k-tuple distribution: In a period N of a sequence, each k-tuple occurs almost equally many times for $1 \leq k \leq \log_2 N$.

– Low correlation: A family of sequences must have low crosscorrelation between distinct sequences as well as low out-of-phase autocorrelation of a sequence.

– Large linear complexity: Sequences must be generated by a large number of LFSRs to avoid the reconstruction from the Berlekamp-Massey algorithm [10] [62].

However, it is not easy to design binary pseudorandom sequences satisfying all of the above randomness criteria. Hence, we should design sequences so that they have the properties essential for their specific applications. From next two sections, it is shown that sequences with low auto- and crosscorrelation are basically required for both communication and cryptographic systems.

1.2 Sequences for Communication Systems

1.2.1 Principles of CDMA Systems and Sequences

Pseudorandom sequences are widely used for communication systems. In particular, the sequences play crucial roles in spread spectrum and code division multiple access (CDMA) communication systems.

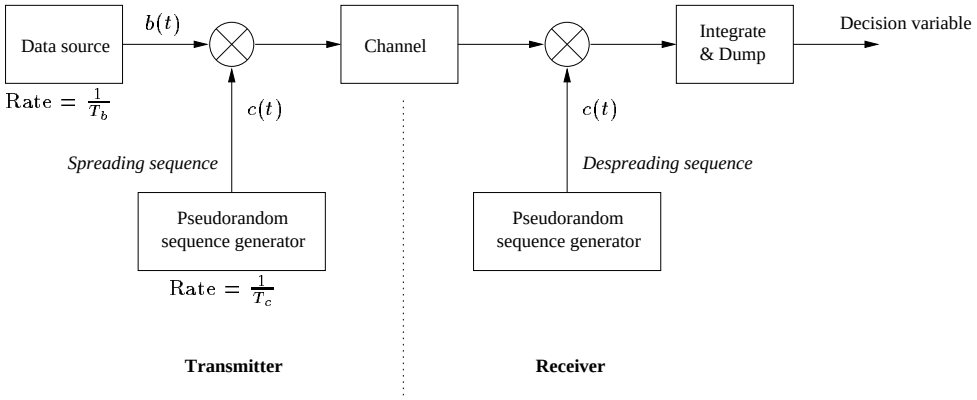


Figure 1.1: Baseband model of a DS-SS system

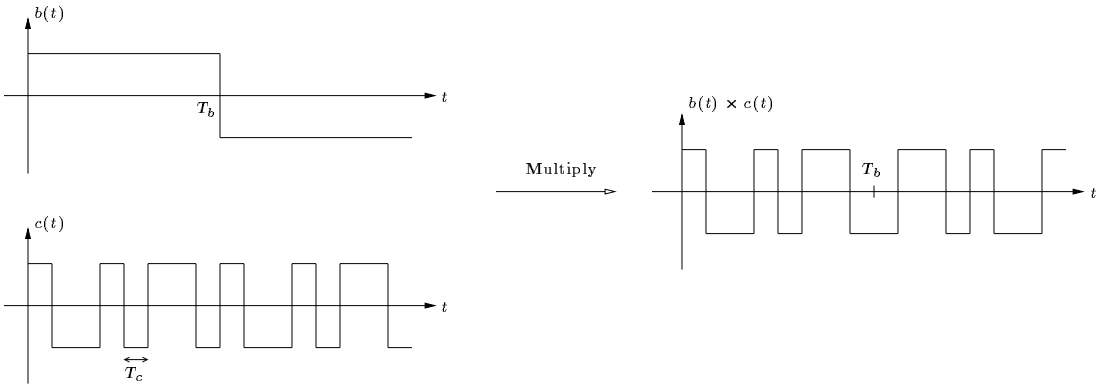


Figure 1.2: Data and spreading sequences in time domain

Figure 1.1 shows a conceptual baseband model of direct-sequence spread spectrum (DS-SS) systems [73]. In Figure 1.1, $b(t)$ is a data signal of a bit duration T_b, and $c(t)$ is a spreading or chip sequence of a chip duration $T_c = 1/f_c$. The ratio T_b/T_c is called a *processing gain* or *spreading factor*. Figure 1.2 shows the rectangular waveforms of $b(t)$ and $c(t)$, respectively.

In time domain, the data signal $b(t)$ is modulated by the spreading sequence $c(t)$ from the direct multiplication of $b(t)c(t)$ as shown in Figure 1.2. A pseudorandom sequence with the randomness properties described in Section 1.1 is used as the spreading sequence $c(t)$. The modulation by the pseudorandom sequence spreads the power spectrum of a data signal in frequency domain.

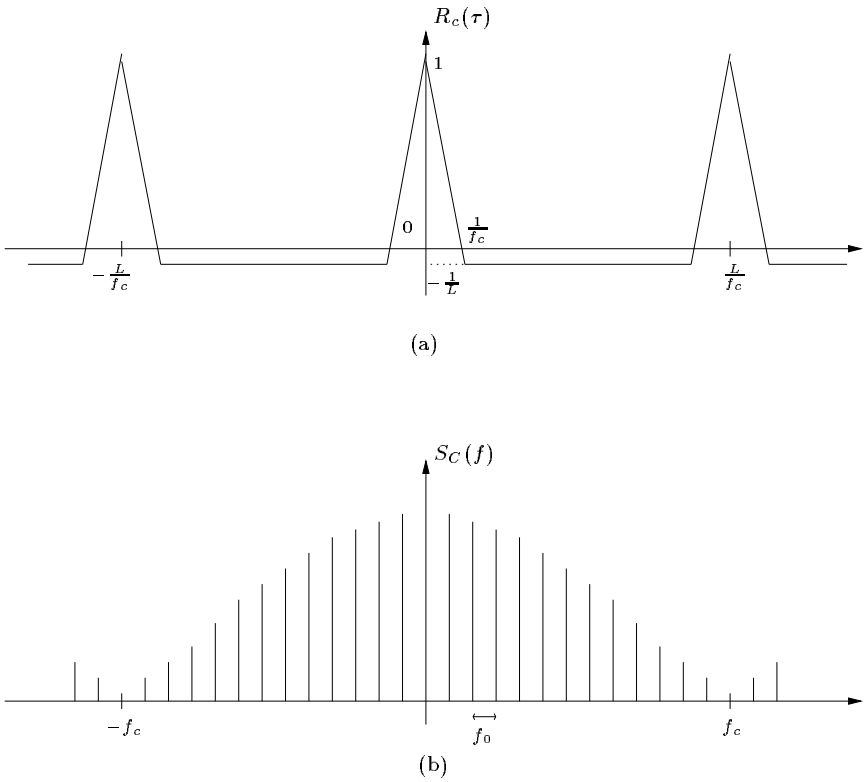


Figure 1.3: Autocorrelation function and power spectral density of $c(t)$

Since the power spectrum of the pseudorandom sequence waveform $c(t)$ is determined by the Fourier transform of its autocorrelation function, the sequence $c(t)$ is required to have a good autocorrelation function such that the power spectrum has a desired shape similar to that of a pure random sequence. If the pseudorandom sequence $c(t)$ has the ideal two-level autocorrelation, then the code waveform has the normalized autocorrelation function $R_c(\tau)$ shown in Figure 1.3-(a). Then, the power spectrum is given by [73]

$$S_C(f) = \frac{L+1}{L}\left(\frac{\sin \pi f/f_c}{\pi f/f_c}\right)^2 \sum_{m=-\infty}^{\infty} \delta(f - mf_0) + \frac{1}{L^2}\delta(f)$$

where $f_0 = f_c/L$ and L is a period of $c(t)$. $S_C(f)$ gives the line spectrum shown

in Figure 1.3-(b). If L is very large, the spectral line gets closer, and the spectrum can be viewed as a continuous spread spectrum desired for practical purposes.

Hence, the autocorrelation function of a pseudorandom sequence shown in Figure 1.3-(a) is important for spread spectrum communications. Also, the impulse-like autocorrelation function is required for a synchronization process where the autocorrelation of a pseudorandom sequence is measured to acquire the accurate timing information. For more details of acquisition and tracking, see [87] and [90].

In direct-sequence CDMA (DS-CDMA) communication systems, each user has its own spread spectrum modulation where each user's data signal is modulated by its own spreading sequence. Then, the modulated signal is multiplexed to others and transmitted through a channel. In a receiver, the received signal goes through the matched filtering by each despreading sequence and an integrator.

Figure 1.4 describes a baseband DS-CDMA system model. In Figure 1.4, $b_k(t)$ and $c_k(t)$, $1 \leq k \leq K$, are the kth user's data signal and its spreading sequence, respectively, where $b_k(t), c_k(t) \in \{-1, +1\}$ and K is the number of users in the system. Without loss of generality, a time delay τ_k, $1 \leq k \leq K$ that each user experiences through a channel can be integrated into a transmitter in Figure 1.4. For synchronous CDMA systems, $\tau_k = 0$ for all k's. On the other hand, it is randomly distributed over $0 \leq \tau_k \leq T_b$ for asynchronous CDMA systems. In this section, the asynchronous case is considered in general. In the receiver, $r(t)$ is a received signal through the additive white Gaussian noise (AWGN) channel $n(t)$, and $\widehat{b_i}(t)$ is the estimate of $b_i(t)$ in the ith receiver. In the baseband model, the perfect carrier recovery for each user is assumed, i.e., a carrier phase offset is equal to zero for all users. The received signal $r(t)$ at the ith user's receiver is given by

$$r(t) = \sum_{k=1}^{K} c_k(t - \tau_k) b_k(t - \tau_k) + n(t)$$

where a scaling energy factor is ignored.

In order to emphasize correlation effects to interference, we analyzed the receiver processing in a discrete time domain by setting $t = jT_c$, $0 \leq j \leq L-1$ and $T_b = LT_c$. For simplicity, it is also assumed that $\tau_k = l_k T_c$, $0 \leq l_k \leq L-1$. (In fact, τ_k can be

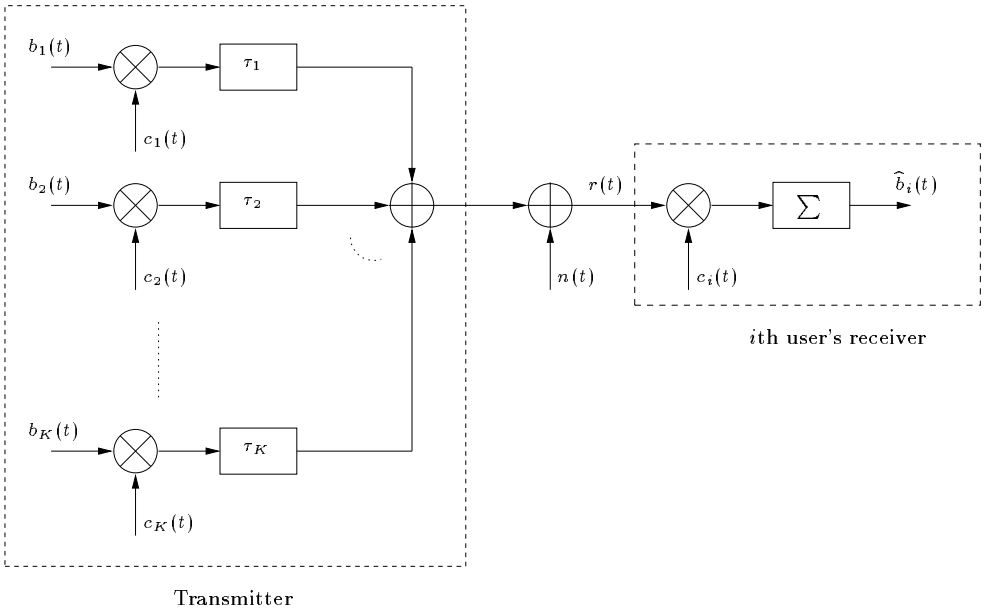


Figure 1.4: Baseband DS-CDMA system model

a value in $l_k T_c \leq \tau_k \leq (l_k + 1)T_c$. However, we assume only a multiple chip delay is allowed. For more detailed analysis, see [76].) In each discrete time signal, T_c is omitted from its notation, e.g., $r(jT_c) = r(j)$. Then, the discrete time received signal is represented by

$$r(j) = \sum_{k=1}^{K} c_k(j - l_k) b_k(j - l_k) + n(j), \quad 0 \leq j \leq L - 1.$$

At $t = T_b$, the correlator output of the ith receiver is $\widehat{b_i}(0)$, which is given by

$$\begin{aligned} \widehat{b_i}(0) &= \sum_{j=0}^{L-1} r(j) c_i(j) \\ &= \sum_{k=1}^{K} \sum_{j=0}^{L-1} c_k(j - l_k) c_i(j) b_k(j - l_k) + \sum_{j=0}^{L-1} n(j) c_i(j) \end{aligned}$$

and let $\sum_{j=0}^{L-1} n(j) c_i(j) = \eta$. In the ith user's receiver, we can assume $\tau_i = 0$ and

$l_i = 0$ without loss of generality. Then,

$$\begin{aligned}\widehat{b_i}(0) &= \sum_{k=1,k\neq i}^{K}\sum_{j=0}^{L-1} c_k(j-l_k)c_i(j)b_k(j-l_k) + \sum_{j=0}^{L-1} c_i(j)^2 b_i(j) + \eta \\ &= \sum_{k=1,k\neq i}^{K}\left(\sum_{j=0}^{l_k-1} c_k(L-l_k+j)c_i(j)b_k(-1) + \sum_{j=l_k}^{L-1} c_k(j-l_k)c_i(j)b_k(0)\right) \\ &\quad + L\cdot b_i(0) + \eta\end{aligned}$$

where $b_i(0)$ is the desired signal component of $b_i(t)$ for $0 \le t \le T_b$ and η is a noise component. Note that $j - l_k \equiv L - l_k + j \pmod{L}$ for $j < l_k$ because the spreading sequence $c_k(j)$ has a period L. In summary, the estimate $\widehat{b_i}(0)$ at time $t = T_b$ in the ith user's receiver is represented by

$$\widehat{b_i}(0) = L\cdot b_i(0) + \sum_{k=1,k\neq i}^{K}\left(b_k(-1)A_{i,k}(l_k) + b_k(0)\widehat{A}_{i,k}(l_k)\right) + \eta \tag{1.1}$$

where $A_{i,k}(l)$ and $\widehat{A}_{i,k}(l)$ are the aperiodic correlation functions defined by

$$A_{i,k}(l) = \sum_{j=0}^{l-1} c_k(L-l+j)c_i(j), \quad \widehat{A}_{i,k}(l) = \sum_{j=l}^{L-1} c_k(j-l)c_i(j).$$

The second hand term on the right-hand side of (1.1) is the interference contributed by other users' data and spreading sequences, called the *multiple access interference (MAI)*. As the MAI increases, each user's signal-to-interference ratio (SIR) decreases, which degrades the bit error rate (BER) performance of the user. To keep the acceptable performance of each user in the system, therefore, the MAI is required to be minimized. Since the MAI is determined by the aperiodic correlation functions in (1.1), spreading sequences should have low aperiodic correlation functions between them. With the low aperiodic correlation functions, the number of users that the system is able to support, or the *CDMA capacity*, can be increased within a certain level of MAI where the acceptable performance of each user is

guaranteed.

Let $C_{i,k}(l)$ be the periodic correlation function of $c_k(j)$ and $c_i(j)$. Then, it is defined by

$$\begin{aligned} C_{i,k}(l) &= \sum_{j=0}^{L-1} c_k(j-l)c_i(j) = \sum_{j=0}^{l-1} c_k(L-l+j)c_i(j) + \sum_{j=l}^{L-1} c_k(j-l)c_i(j) \\ &= A_{i,k}(l) + \widehat{A}_{i,k}(l). \end{aligned} \tag{1.2}$$

Specifically, $C_{i,k}(l)$ is called *even correlation. Odd correlation* is also defined by

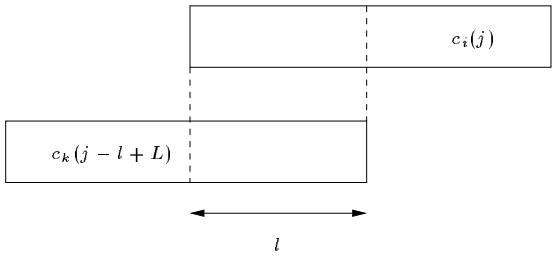
$$\widehat{C}_{i,k} = A_{i,k}(l) - \widehat{A}_{i,k}(l). \tag{1.3}$$

Since $b_k(-1), b_k(0) \in \{-1, +1\}$, the MAI in (1.1) is determined by the even and odd correlation functions from (1.2) and (1.3). Finally, the low even and odd correlation functions are essential for increasing the capacity of CDMA systems by minimizing the MAI.

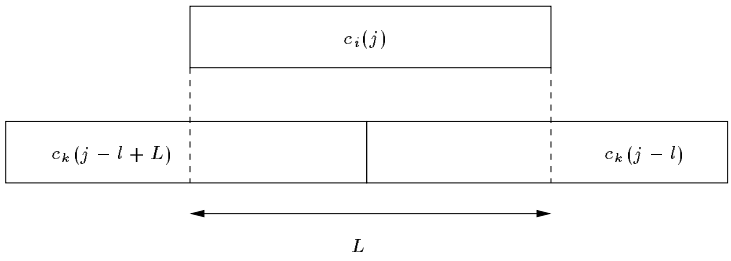
Fig. 1.5 illustrates the computations of the aperiodic correlation $A_{i,k}(l)$ and periodic correlation $C_{i,k}(l)$, respectively. In $A_{i,k}(l)$, the correlation window through which the correlation is computed, is l. Thus, the correlation window is changed according to l, so it is hard to design sequences with good aperiodic correlation whose structure can be handled by analytical way. In fact, sequence design with good aperiodic correlation is known to be a notoriously difficult mathematical problem [44].

In $C_{i,k}(l)$, on the other hand, the correlation window is always L. Hence, the periodic correlation is handled by the structure of linear cyclic codes of length L and their weight distributions.

For a long time, therefore, many researchers have made efforts to the design of sequences with good periodic correlation rather than good aperiodic correlation. In this thesis, *binary sequences with low even periodic correlation* are mainly studied as well for the design of sequences applied for CDMA systems.



(a) Aperiodic correlation $A_{i,k}(l)$



(b) Periodic correlation $C_{i,k}(l)$

Figure 1.5: Aperiodic and periodic correlation

1.2.2 Application of Sequences for CDMA Systems

From the previous section, sequences with low auto and crosscorrelation play important roles in wireless communication systems. There are several different research areas on the sequences according to specific applications of communication systems.

Orthogonal Codes: In CDMA communication systems, all active users or physical channels occupy the same frequency band at the same time. They are spread and distinguished by different code assignments. To reduce the mutual interference between distinct users or channels in synchronous CDMA systems, the codes assigned to each user or channel should be mutually orthogonal, which means that in-phase crosscorrelation of any pair of distinct codes is zero.

As the orthogonal codes, the Walsh codes and the OVSF (orthogonal variable spreading factor) codes are adopted in the third generation wireless communication

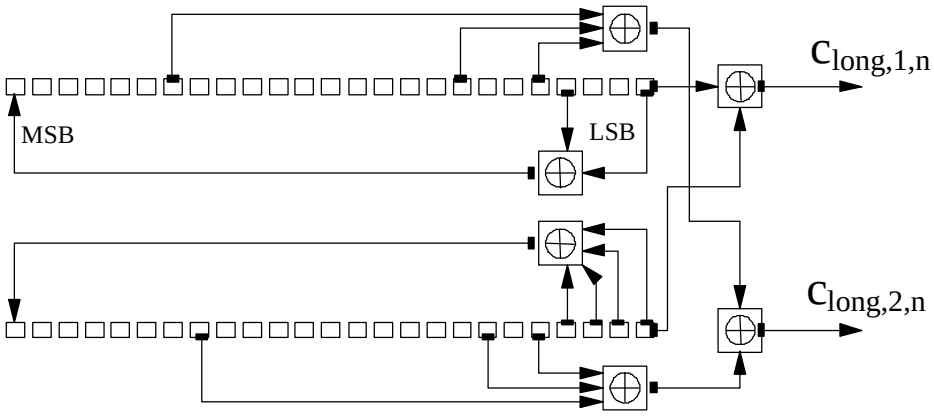


Figure 1.6: UMTS uplink long scrambling code generator

standards of the cdma2000 [4] and the UMTS [2], respectively. Each Walsh or OVSF code of length n corresponds to each row of an $n \times n$ Hadamard matrix [83]. The orthogonal codes are used for channelization, spreading, and orthogonal modulation in uplink and downlink transmissions of the cdma2000 and the UMTS. In Figure 1.2, a data signal is spread by a 8 bit Walsh code, where $T_b = 8T_c$, i.e., a spreading factor (SF) is 8.

Sequences with Low Correlation: Sequences with low auto and crosscorrelation are used for the identification of users and base stations in multiple access communication systems. Furthermore, such sequences play a role of data scrambling.

There are several sequence families proposed for these purposes; The Gold sequences [31], the Kasami sequences [46], the Gold-like sequences [11], the interleaved sequences [35], and the Z_4 sequences [51]. In practice, the Gold sequences are used to generate long scrambling codes in the UMTS [2], where the sequences are constructed by combining two cyclically distinct binary m-sequences [31]. The Gold sequences have optimal correlation achieving the Sidelnikov bound [86]. Also, the quaternary Z_4 sequences [51] are adopted as short scrambling codes for uplink transmission in the UMTS.

Figure 1.6 describes a long scrambling code generator for the uplink UMTS where $c_{long,1,n}$ and $c_{long,2,n}$ are binary Gold sequences.

Sequences with Ideal Two-level Autocorrelation: Sequences with optimal autocorrelation have very important roles in synchronization and positioning processes of CDMA and radar communication systems. Moreover, since a Hadamard matrix can be constructed from a binary sequence with ideal two-level autocorrelation, they are basic sequences for orthogonal codes which are used as spreading codes.

The m-sequences [32], the GMW sequences [37], the Legendre sequences [43], and the Hall's sextic residue sequences [38] are well known sequences of period 2^n-1 with ideal two-level autocorrelation. Since 1997, several new binary sequences of period 2^n-1 with ideal two-level autocorrelation have been discovered together with the construction of corresponding new cyclic difference sets; the Kasami power function (KPF) sequences [23], the Welch-Gong (WG) sequences [70] [69], and the hyperoval sequences [61]. Also, all binary sequences of period 2^n-1 with ideal two-level autocorrelation, which are not m-sequences, can be applied again to construct new binary sequences with ideal two-level autocorrelation, so called the generalized GMW sequences [33], using the subfields of the finite field if n is a composite number.

In the cdma2000 standard, a binary m-sequence of period $2^{42}-1$ is employed as a long code for identification, data scrambling, and power control bit randomization [4]. A pair of binary m-sequences are also used for quadrature spreading in the cdma2000 after being modified to the de-Bruijn sequences [32]. Figure 1.7 shows a long code generator of the cdma2000 standard which is based on a binary m-sequence of period $2^{42}-1$.

Sequences with Good Aperiodic Autocorrelation: Recently, multicarrier transmission techniques such as orthogonal frequency division multiplexing (OFDM) and multicarrier CDMA (MC-CDMA) have attracted much attention for future communication systems.

The Golay complementary sequences [30], already adopted as primary synchronization codes (PSC) in the UMTS [2], are one of the most interesting sequences for these applications. By definition, the sum of out-of-phase aperiodic autocorrelation of a pair of Golay sequences is zero, which leads to the fact that a peak power of the

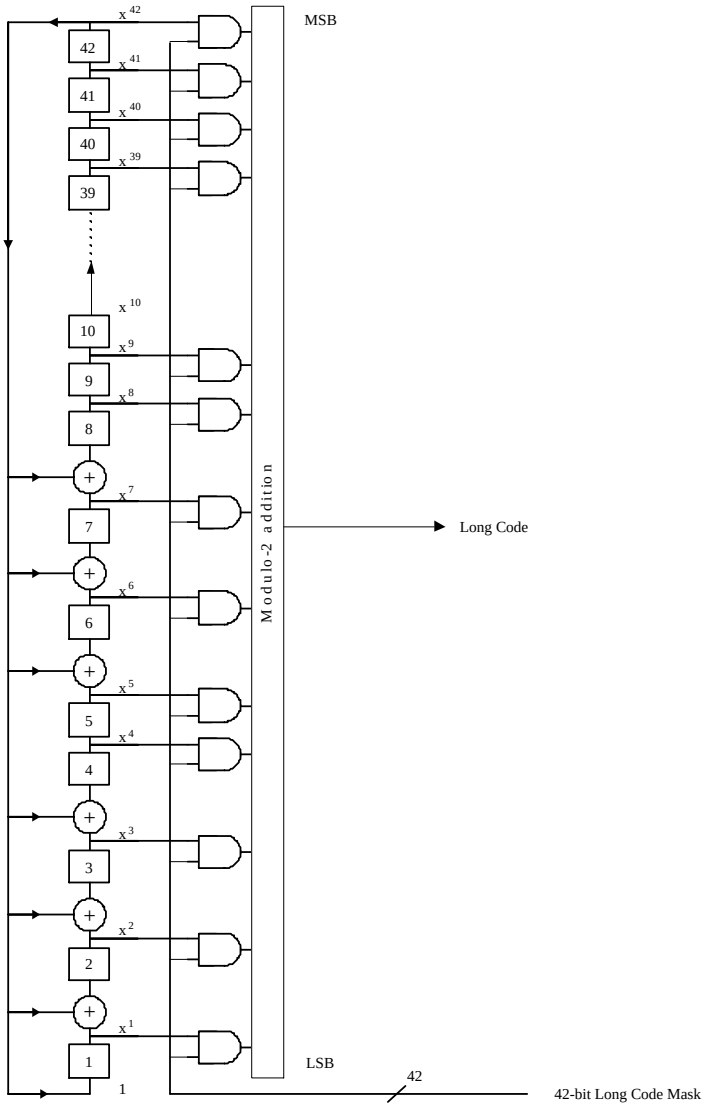


Figure 1.7: cdma2000 long code generator

sequences is always bounded by a constant [74]. Hence, the Golay complementary sequences are considered to be a good option for solving the peak-to-average power (PAPR) problem of multicarrier transmissions.

The Barker sequences [7] are binary sequences with optimal aperiodic autocorrelation; All out-of-phase aperiodic autocorrelation magnitudes are at most 1. Unfortunately, however, all known Barker sequences have a length of at most 13, and it seems that there exist no other Barker sequences [40] [88]. In practice, the 11 bit Barker sequences are used as spreading sequences of direct-sequence spread spectrum physical layer of IEEE 802.11 wireless local area network (WLAN) standard [3].

1.3 Sequences for Cryptographic Systems

1.3.1 Principles of Security and Sequences

The security of a cryptographic system depends on pseudorandom numbers or sequences used in the system. For the robustness to cryptanalyses from malicious attackers, the sequences must not be predicted by certain deterministic algorithms within a polynomial time [9] [65]. For the purpose, pseudorandom sequences used for the keystream generation are required to have cryptographically secure properties. Among the properties, high nonlinearity and large linear complexities of the pseudorandom sequences are necessary for the security of a system using the sequences. Every secure cryptographic system should possess the high nonlinearity and the large linear complexity even if the properties do not necessarily guarantee the full security of the system.

Let $\mathbb{F}_2^n$ be an n-dimensional vector space over $\mathbb{F}_2 = \{0, 1\}$, and $f(\mathbf{x})$ and $g(\mathbf{x})$ be Boolean functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. Let $\mathbf{f}$ and $\mathbf{g}$ be vectors of length 2^n determined by $f(\mathbf{x})$ and $g(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + c$ for all $\mathbf{x} \in \mathbb{F}_2^n$, respectively, where $\mathbf{w} \in \mathbb{F}_2^n$ and $c \in \mathbb{F}_2$. Then, the nonlinearity of a Boolean function $f(\mathbf{x})$ is defined by [33]

$$N_f = \min_{\mathbf{w} \in \mathbb{F}_2^n, c \in \mathbb{F}_2} d(\mathbf{f}, \mathbf{g})$$

where $d(\mathbf{f}, \mathbf{g})$ denotes the Hamming distance between two vectors $\mathbf{f}$ and $\mathbf{g}$. That is, the nonlinearity of a Boolean function is defined by the minimum Hamming distance of the function from all affine functions.

The nonlinearity of a Boolean function $f(\mathbf{x})$ can be characterized by the Walsh transform of $f(\mathbf{x})$ which will be defined in Section 1.5. In other words,

$$N_f = 2^{n-1} - \frac{1}{2} \max_{\mathbf{w} \in \mathbb{F}_2^n} |\widehat{f}(\mathbf{w})|.$$

In addition, the Walsh transform of a Boolean function $f(\mathbf{x})$ is equivalent to the Hadamard transform of a polynomial function $f(x)$ which has the one-to-one correspondence to $f(\mathbf{x})$. Therefore, the nonlinearity is also defined by

$$N_f = 2^{n-1} - \frac{1}{2} \max_{\lambda \in \mathbb{F}_{2^n}} |\widehat{f}(\lambda)|$$

where $\mathbb{F}_{2^n}$ is a finite field with elements 2^n. The Hadamard transform $\widehat{f}(\lambda)$ will be defined in Section 1.5. The Hadamard transform of a polynomial function $f(x)$ is also equivalent to the correlation function between a sequence represented by $f(x)$ and a binary m-sequence. In sequence aspects, therefore, the high nonlinearity corresponds to low crosscorrelation of a binary sequence with a linear m-sequence. Consequently, the design of a Boolean function with high nonlinearity is equivalent to the design of a sequence with low correlation. A known Boolean function with the highest nonlinearity is the *bent function* [78].

In addition to the nonlinearity, a Boolean function is also expected to have another essential secure property for cryptographic applications. In other words, it should be difficult to predict the whole outputs of a Boolean function by observing a segment of the outputs. For this purpose, the sequence corresponding to the Boolean function should have a large linear complexity so that it cannot be reconstructed by the *Berlekamp-Massey algorithm* [10] [62]. If the linear complexity of a sequence is n, for example, then the Berlekamp-Massy algorithm efficiently enables one to obtain the minimal polynomial of the sequence only with the observation of $2n$ bits. Thus, the linear complexity of sequences should be as large as possible to

avoid the reconstruction of the sequences from the Berlekamp-Massey algorithm. For more details of the algorithm, see [10] and [62].

In summary, sequences used for cryptographic systems are first required to have low correlation for the high nonlinearity. Moreover, the sequences should have the large linear complexity to avoid the cryptanalysis from deterministic algorithms within a polynomial time. Therefore, the low correlation and the large linear complexity are necessary conditions that pseudorandom sequences possess for the security of cryptographic systems.

1.3.2 Application of Sequences for Cryptographic Systems

From the previous section, it is shown that pseudorandom sequences with low correlation and large linear complexity play significant roles in cryptographic systems. In practice, the sequences generated by linear feedback shift registers (LFSRs) have the other advantages of simple implementation and fast processing. Therefore, the stream cipher based on sequences has many cryptographic applications.

In a stream cipher, a key stream bit is added to a plaintext bit by the exclusive-OR (XOR) operation to generate a ciphertext bit. In practice, the keystreams of a stream cipher are generated by the combination of linear feedback shift registers (LFSRs) and other nonlinear functions. Figure 1.8 shows the basic structures of a pseudorandom keystream generator of a stream cipher. In Figure 1.8 (a) and (c), nonlinear Boolean functions are combined to one or more LFSRs for the nonlinearity of keystreams. In Figure 1.8 (b), on the other hand, a nonlinear logic is employed for controlling one or more LFSRs.

Since the LFSR based keystream generators are well suited for fast hardware implementation, the stream cipher is attractive for the security of wireless communication systems. Currently, it is used for encryption and decryption of plaintext data in several wireless communication systems; A5/1 [12] in the GSM, E0 [1] in Bluetooth, and RC4 [77] in Wired Equivalent Privacy (WEP) in IEEE 802.11 WLAN standard.

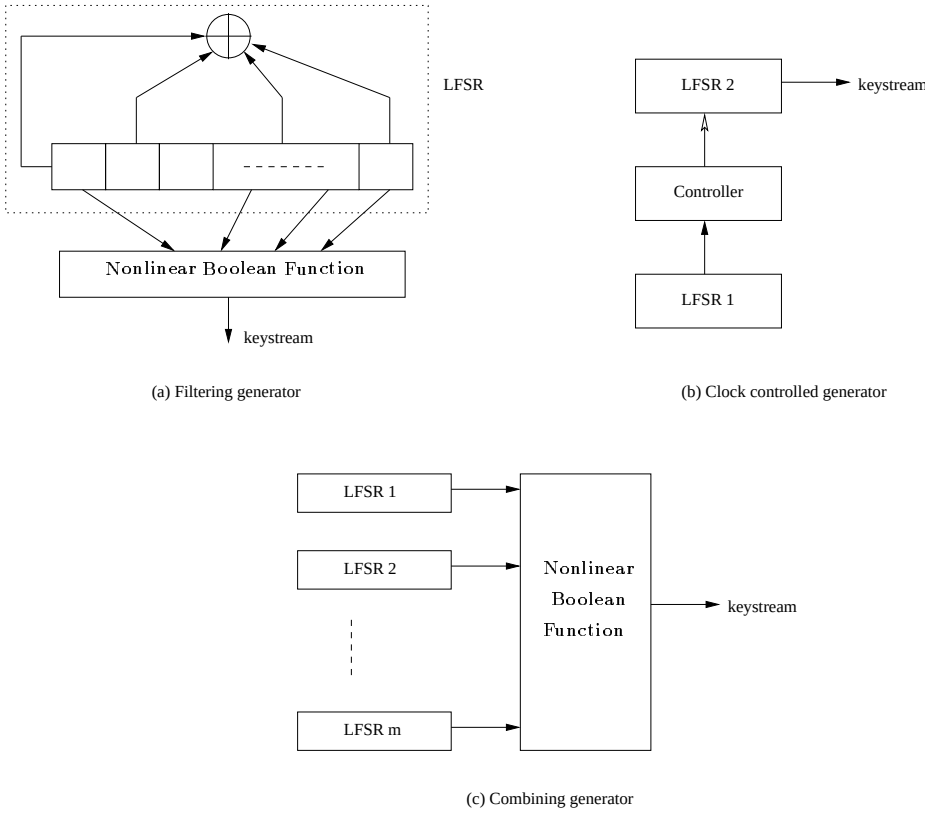


Figure 1.8: Basic structures of a keystream generator for a stream cipher

1.4 Thesis Overview

From Sections 1.2 and 1.3, it is convinced that sequences with low correlation are essential for both communication and cryptographic systems. In this thesis, (periodic) correlation of binary sequences is mainly studied. Theoretically, the correlation of sequences has a strong connection with the problems of algebraic coding and combinatorics. Namely, the correlation distribution of sequences is related to the weight distribution of codewords in algebraic codes [60]. Furthermore, binary sequences with ideal two-level autocorrelation are equivalent to cyclic difference sets in combinatorics [8]. By the aid of fruitful results of coding theory and combinatorics, some interesting properties are presented on the correlation of binary sequences, and several new binary sequences with low correlation are proposed in

this thesis. The outline of thesis is as follows.

Since 1997, several new classes of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation have been discovered together with the construction of corresponding new cyclic difference sets; the Kasami power function (KPF) sequences, the Welch-Gong (WG) sequences, and the hyperoval sequences. In Chapter 2, the sequences are first introduced by the corresponding cyclic difference sets, and then the sequences are studied for odd n in terms of the crosscorrelation properties between them. In addition, the crosscorrelation of the sequences and binary m-sequences is also examined. As a basic approach, the Hadamard equivalence of binary sequences is exploited. The relation of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation is discussed in terms of the 3- and 5-valued crosscorrelation, which are theoretically proved or conjectured.

In Chapter 3, binary sequences with optimal autocorrelation are classified in terms of its period, and the known constructions are reviewed. Then, new binary sequences of period $N = 4(2^m - 1)$ for even $m \geq 4$ are proposed by making use of a $4 \times (2^m - 1)$ interleaved structure of sequences. The sequences are almost balanced and have three-valued out-of-phase autocorrelation, i.e., $\{0, \pm 4\}$, which is optimal with respect to autocorrelation magnitude. In addition, it is shown that the sequences have large linear complexity from the implementation with linear feedback shift registers (LFSRs) and a simple logic.

In Chapter 4, known binary sequence families with low correlation are first reviewed, and then a new binary sequence family with low correlation and a large family size is proposed by generalizing the Gold-like [11] and the Udaya's [89] sequences. The weight distribution of subcodes of the second order Reed Muller code is discussed and studied to derive the correlation distribution of the sequence family. According to the flexible parameters of maximum correlation and a family size, the sequence family provides various sub-sequence families. With respect to maximum correlation, a family size, and a linear complexity, it is shown that the sequence family of period $2^n - 1$ for odd n is competitive among all known binary sequence families.

In Chapter 5, the bent functions represented by a polynomial form are in-

troduced and the known constructions are first reviewed. Then, the constructions of all quadratic bent functions represented by a polynomial form of $f(x) = \sum_{i=1}^{n/2-1} c_i Tr(x^{1+2^i}) + c_{n/2} Tr_1^{n/2}(x^{1+2^{n/2}}), x \in GF(2^n), c_i \in GF(2)$, are presented for special values of n by providing a necessary and sufficient condition on c_i's. Factorization of the cyclotomic polynomials [10] [56] is considered to derive the condition. The resulting quadratic bent functions can be used for the LFSR implementation of a family of bent sequences [71], which is a potential candidate for future CDMA systems.

Finally, concluding remarks are presented in Chapter 6.

1.5 Definitions and Concepts

In this section, preliminary definitions and concepts are given to describe the works of this thesis. Since binary sequences and their correlation are main research topics, every definition and concept are described from the viewpoint of binary sequences. The following notations will be used throughout this thesis. Most of basic concepts and properties about sequences we introduce in this section can be found in [33].

- $\mathbb{Z}_m$ is a ring of integers modulo m and $\mathbb{Z}_m^+ = \{r \in \mathbb{Z}_m \mid r \neq 0\}$.
- $\mathbb{F}_q = GF(q)$ is a finite field with q elements and $\mathbb{F}_q^*$ is a multiplicative group of $\mathbb{F}_q$.
- $\mathbb{F}_2^n$ is a vector space over $\mathbb{F}_2 = \{0, 1\}$ with a set of all binary n-tuples.
- For a binary sequence $\mathbf{a} = \{a_t\}$, $a_t \in \{0, 1\}$. Also, $\overline{\mathbf{a}}$ is its complement or $\overline{\mathbf{a}} = \{a_t + 1\}$ where the addition is computed modulo 2.
- For a sequence $\mathbf{a} = \{a_t\}$ over $\mathbb{Z}_q$ and an integer g, $\mathbf{a} + g = \{a_t + g\}$ where the addition is computed modulo q.
- $\mathrm{ord}_p(2)$ is the order of 2 modulo p, i.e., the smallest integer s such that $2^s \equiv 1 \pmod{p}$.

– For positive integers n and m, let $m|n$. A trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$ is denoted by $Tr_m^n(x)$, i.e.,

$$Tr_m^n(x) = x + x^{2^m} + \cdots + x^{2^{m(\frac{n}{m}-1)}}, \quad x \in \mathbb{F}_{2^n},$$

or simply as $Tr(x)$ if $m = 1$ and the context is clear.

Boolean Functions: Let $\mathbf{x} = (x_1, \cdots, x_n)$ be a vector in $\mathbb{F}_2^n$ with $x_i \in \mathbb{F}_2$. A function $f(\mathbf{x})$ from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ which takes on values 0 or 1 is called a *Boolean function* [60]. A Boolean function consists of a sum of all possible products of x_{i_j}'s with coefficients 0 or 1, i.e.,

$$f(\mathbf{x}) = f(x_1, \cdots, x_n) = c_0 + \sum_{1 \le j \le n} c_{i_1 i_2 \cdots i_j} x_{i_1} x_{i_2} \cdots x_{i_j}, \quad c_0, c_{i_1 i_2 \cdots i_j} \in \mathbb{F}_2 \qquad (1.4)$$

where the sum runs through all non-empty subsets of $\{i_1, \cdots, i_j\} \subset \{1, \cdots, n\}$. The maximum value of j with nonzero $c_{i_1 i_2 \cdots i_j}$ is called the *degree* of the Boolean function $f(\mathbf{x})$. (1.4) is called the *algebraic normal form* of a Boolean function [33].

Polynomial Functions: A function $f(x)$ from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ can be represented as

$$f(x) = \sum_{i \in I} Tr_1^{n_i}(A_i x^i) + A_{2^n-1} x^{2^n-1}, \quad A_i \in \mathbb{F}_{2^{n_i}}, A_{2^n-1} \in \mathbb{F}_2 \qquad (1.5)$$

where I is a subset of $\mathbb{Z}_{2^n-1}$ containing all coset leaders modulo $2^n - 1$ and $n_i|n$ is a size of the cyclotomic coset containing i [40] [33]. (1.5) is called a *polynomial* or *trace representation* of $f(x)$.

With respect to a basis $\{\beta_1, \cdots, \beta_n\}$ of $\mathbb{F}_{2^n}$, any element $x \in \mathbb{F}_{2^n}$ has the expansion of $x = \sum_{i=1}^n x_i \beta_i$ with $x_i \in \mathbb{F}_2$. Applying this expansion to (1.5), a polynomial function is converted into the equivalent Boolean function of (1.4).

Periodic Sequences and Polynomial Functions: Let $\mathcal{S}$ be a set of all binary sequences with period $N|(2^n - 1)$ and $\mathcal{F}$ be a set of all polynomial functions from

$\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. For any $x \in \mathbb{F}_{2^n}^*$, $x^{2^n-1} = 1$ in (1.5), and thus $f(x)$ is represented by

$$f(x) = A_0 + \sum_{i \in I, i \neq 0} Tr_1^{n_i}(A_i x^i) + A_{2^n-1}, \quad x \in \mathbb{F}_{2^n}^*. \tag{1.6}$$

Since A_0 and A_{2^n-1} are in $\mathbb{F}_2$, $A_{2^n-1} = 0$ in (1.6) without loss of generality. In terms of the relation of sequences and polynomial functions, therefore, it is sufficient to consider a polynomial function in $\mathcal{F}$ represented by

$$f(x) = \sum_{i \in I} Tr_1^{n_i}(A_i x^i), \quad A_i \in \mathbb{F}_{2^{n_i}}. \tag{1.7}$$

For any sequence $\mathbf{a} = \{a_t\} \in \mathcal{S}$, there exists $f(x)$ of (1.7) such that

$$a_t = f(\alpha^t), \quad t = 0, 1, \cdots,$$

where α is a primitive element of $\mathbb{F}_{2^n}$. Then, $f(x)$ is called the *trace representation* of $\mathbf{a}$. The notation of $\mathbf{a} \leftrightarrow f(x)$ is used to represent a binary sequence and its trace representation. In particular, if $f(x)$ is a single term trace function, i.e.,

$$a_t = Tr_1^n(\beta \alpha^t), \quad t = 0, 1, \cdots, \quad \beta \in \mathbb{F}_{2^n}^*, \tag{1.8}$$

then $\mathbf{a}$ is a binary m-sequence of period $2^n - 1$.

Again, $f(x)$ in (1.7) is represented by

$$f(x) = A_0 + \sum_{i \in I, i \neq 0} Tr_1^{n_i}(A_i x^i), \quad A_0 \in \mathbb{F}_2, A_i \in \mathbb{F}_{2^{n_i}}.$$

Let $\mathbf{a} \leftrightarrow f(x)$ with $A_0 = 0$, and $\mathbf{b} \leftrightarrow f(x)$ with $A_0 = 1$. Then, $\mathbf{a}$ and $\mathbf{b}$ are complement to each other. In correlation aspects, it is clear that $\mathbf{a}$ and $\mathbf{b}$ show the equivalent correlation properties. In this thesis, therefore, we only consider $\mathbf{a}$ with $A_0 = 0$. In summary, the trace representation $f(x)$ of a binary sequence is given by

$$f(x) = \sum_{i \in I, i \neq 0} Tr_1^{n_i}(A_i x^i), \quad A_i \in \mathbb{F}_{2^{n_i}}$$

without loss of generality. Therefore, $f(0) = 0$ for the trace representation of any binary sequence.

Example 1 *Consider a finite field* $\mathbb{F}_{2^3}$ *generated by a primitive polynomial* $x^3 + x + 1$. *Then, the primitive element* α *is a root of the primitive polynomial, and* $\alpha^7 = 1,\ \alpha^3 + \alpha + 1 = 0$. *Let* $Tr(x)$ *be a trace function from* $\mathbb{F}_{2^3}$ *to* $\mathbb{F}_2$. *Then,*

$$\begin{aligned}
&Tr(1) = 1 + 1 + 1 = 1,\\
&Tr(\alpha) = \alpha + \alpha^2 + \alpha^4 = 0,\ Tr(\alpha^2) = Tr(\alpha^4) = Tr(\alpha) = 0,\\
&Tr(\alpha^3) = \alpha^3 + \alpha^6 + \alpha^5 = 1,\ Tr(\alpha^6) = Tr(\alpha^5) = Tr(\alpha^3) = 1
\end{aligned}$$

Hence, a binary sequence represented by $a_t = Tr(\alpha^t),\ t = 0, 1, \cdots, 6$, *is* $\mathbf{a} = (1, 0, 0, 1, 0, 1, 1)$, *which is a binary m-sequence of period* 7. *Consequently,* $Tr(x)$ *is a trace representation of a binary m-sequence* $\mathbf{a}$ *of period* 7.

Equivalence Relation of Sequences: Let $\mathbf{a} = \{a_t\}$ and $\mathbf{b} = \{b_t\}$ be two periodic sequences. Then, they are called *cyclically equivalent* [33] if there exists an integer k such that

$$a_t = b_{t+k} \text{ for all } t \geq 0$$

denoted by $\mathbf{a} = L^k(\mathbf{b})$. Otherwise, they are called *cyclically distinct.*

Balance and Almost Balance Properties: Let $\mathbf{a} = \{a_t\}$ be a binary sequence of period N. Then $\mathbf{a}$ is called *balanced* [32] if the number of zeros is nearly equal to the number of ones in a period, i.e.,

$$S = \left| \sum_{t=0}^{N-1} (-1)^{a_t} \right| \leq 1$$

where S denotes a difference between the numbers of zeros and ones of a binary sequence in a period. For odd N, $\mathbf{a}$ is balanced if and only if $S = 1$, and for even N, it is balanced if and only if $S = 0$. On the other hand, if N is even and $S = 2$, then $\mathbf{a}$ is called *almost balanced* [68].

Constant-on-cosets Property: Let $\mathbf{a} = \{a_t\}$ be a binary sequence of period $2^n - 1$. If there exists some τ such that $a_{\tau+2i} = a_{\tau+i}$, $i = 0, 1, \cdots$, then $\mathbf{a}$ is said to have the *constant-on-cosets* property [33]. In other words, $L^\tau(\mathbf{a})$ takes a fixed value on every cyclotomic coset modulo $2^n - 1$.

Let α be a primitive element of $\mathbb{F}_{2^n}$. Then, a binary m-sequence $\mathbf{a} = \{a_t\}$ defined by (1.8) is constant-on-coset if and only if $\beta = 1$.

Example 2 *Consider* $\mathbf{a} = (a_0, a_1, a_2, a_3, a_4, a_5, a_6) = (1, 0, 0, 1, 0, 1, 1)$, *a binary m-sequence of period* 7. *The cyclotomic coset modulo* 7 *is given by* $C_1 = \{1, 2, 4\}$, $C_3 = \{3, 6, 5\}$. *In* $\mathbf{a}$, $a_1 = a_2 = a_4 = 0$ *where indices belong to* C_1. *Also,* $a_3 = a_6 = a_5 = 1$ *where indices belong to* C_3. *Therefore, elements of* $\mathbf{a}$ *is constant on the cosets, i.e.,* $\mathbf{a}$ *has the constant-on-cosets property.*

Decimation of Periodic Sequences: Let $\mathbf{a}$ be a binary sequence of period $N|(2^n - 1)$ and $f(x)$ be the trace representation of $\mathbf{a}$. Let $0 < s < N$. Then a sequence $\mathbf{b} = \{b_t\}$ whose elements are given by

$$b_t = a_{st}, \quad t = 0, 1, \cdots,$$

is said to be an *s-decimation* of $\mathbf{a}$, denoted by $\mathbf{a}^{(s)}$. The trace representation of $\mathbf{a}^{(s)}$ is $f(x^s)$. That is,

$$\mathbf{a} \longleftrightarrow \quad f(x), \quad \mathbf{a}^{(s)} \longleftrightarrow \quad f(x^s).$$

If $\mathbf{a}$ is a binary m-sequence of period $2^n - 1$ and $\gcd(s, 2^n - 1) = 1$, then $\mathbf{a}^{(s)}$, the s-decimation of $\mathbf{a}$, is also a binary m-sequence of period $2^n - 1$. Similarly, if $\mathbf{a}$ is a binary sequence of period $2^n - 1$ with ideal two-level autocorrelation and $\gcd(s, 2^n - 1) = 1$, then $\mathbf{a}^{(s)}$ is also a binary sequence with ideal two-level autocorrelation. From now on, it is called that $\mathbf{a}^{(s)}$ is *decimation-equivalent* to $\mathbf{a}$, and all decimation equivalent sequences of $\mathbf{a}$ constitute a *(decimation-equivalent) class of the sequence.*

Example 3 *In Examples 1 and 2,* $\mathbf{a}^{(3)}$*, a* 3*-decimation of* $\mathbf{a}$*, is given by*

$$\mathbf{a}^{(3)} = (1, 1, 1, 0, 1, 0, 0)$$

It is easily known that $\mathbf{a}^{(3)}$ *also has the constant-on-cosets property. By some manipulations, we can also check that the trace representation of* $\mathbf{a}^{(3)}$ *is* $Tr(x^3)$*.*

Autocorrelation of Binary Sequences: A (periodic) autocorrelation function of a binary sequence $\mathbf{a} = \{a_t\}$ of period N is defined by

$$C_{\mathbf{a}}(\tau) = \sum_{t=0}^{N-1} (-1)^{a_{t+\tau}+a_t}, \quad 0 \le \tau \le N-1 \tag{1.9}$$

where τ is a phase shift of $\mathbf{a}$ and the indices are computed modulo N. For a sequence $\mathbf{a}$ of period N, it is implied that $C_{\mathbf{a}}(\tau) = N$ occurs only at $\tau \equiv 0 \pmod{N}$. In particular, if $\mathbf{a}$ has the autocorrelation function of case 1), i.e.,

$$C_{\mathbf{a}}(\tau) = \begin{cases} -1, & \text{if } \tau \not\equiv 0 \bmod N \\ N, & \text{if } \tau \equiv 0 \bmod N, \end{cases}$$

then the sequence is said to have the *(ideal) two-level autocorrelation function.* A binary sequence with ideal two-level autocorrelation corresponds to the *cyclic Hadamard difference set* [8].

Perfect Sequences: Let $\mathbf{a}$ be a binary sequence of period N. If its autocorrelation $C_{\mathbf{a}}(\tau)$ is equal to 0 for all $\tau \not\equiv 0 \pmod{N}$, i.e.,

$$C_{\mathbf{a}}(\tau) = \begin{cases} 0, & \text{if } \tau \not\equiv 0 \bmod N \\ N, & \text{if } \tau \equiv 0 \bmod N, \end{cases}$$

then $\mathbf{a}$ is called the *perfect sequence.* We also define the perfect sequence for a nonbinary sequence by extending the definition of its autocorrelation [40]. For nonbinary cases, a few polyphase perfect sequences are known in [19] and [28]. However, the

only known perfect binary sequence is $\mathbf{a} = (0, 1, 1, 1)$ or its complement [8]. For a period of $4 < N < 108900$, no perfect binary sequences are found [80], and it is conjectured in [45] that no other perfect binary sequences exist except for $N = 4$.

Orthogonal Functions: Let $f(x)$ be a polynomial function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ with $f(0) = 0$. If we have

$$C_f(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(\lambda x)+f(x)} = \begin{cases} 0, & \text{if } \lambda \neq 1 \\ 2^n, & \text{if } \lambda = 1 \end{cases}$$

for $\lambda \in \mathbb{F}_{2^n}$, then we say that $f(x)$ is *orthogonal* over $\mathbb{F}$. If $f(x)$ is the trace representation of a binary sequence $\mathbf{a}$ of period $2^n - 1$ and $C_{\mathbf{a}}(\tau)$ is the autocorrelation function of $\mathbf{a}$ defined in (1.9), then

$$C_{\mathbf{a}}(\tau) = -1 + C_f(\lambda), \quad \lambda = \alpha^\tau \in \mathbb{F}_{2^n}^* \tag{1.10}$$

where α is a primitive element of $\mathbb{F}_{2^n}$. Thus, $C_f(\lambda)$ has the one-to-one correspondence with the autocorrelation function $C_{\mathbf{a}}(\tau)$. Furthermore, the trace representation of a binary sequence of period $2^n - 1$ with ideal two-level autocorrelation is an orthogonal function from (1.10).

Crosscorrelation of Binary Sequences: A (periodic) crosscorrelation function of binary sequences $\mathbf{a}$ and $\mathbf{b}$ of period N is defined by

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{t=0}^{N-1} (-1)^{a_{t+\tau}+b_t}, \quad 0 \leq \tau \leq N-1$$

where τ is a phase shift of the sequence $\mathbf{a}$ and the indices are computed modulo N. Let $N = 2^n - 1$. If $f(x)$ and $g(x)$ are the trace representations of $\mathbf{a}$ and $\mathbf{b}$ with $f(0) = g(0) = 0$, respectively, then the crosscorrelation is given by

$$C_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{t=0}^{2^n-2} (-1)^{a_{t+\tau}+b_t} = -1 + \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(\lambda x)+g(x)} = -1 + C_{f,g}(\lambda) \tag{1.11}$$

Table 1.1: Crosscorrelations $C_{\mathbf{a},\mathbf{b}}(\tau)$ and $C_{f,g}(\lambda)$

τ	$\mathbf{b}$	$L^{(\tau)}(\mathbf{a})$	$C_{\mathbf{a},\mathbf{b}}(\tau)$	λ	$C_{f,g}(\lambda)$
0	1110100	1001011	-5	1	-4
1	1110100	0010111	-1	α	0
2	1110100	0101110	-1	α^2	0
3	1110100	1011100	$+3$	α^3	$+4$
4	1110100	0111001	-1	α^4	0
5	1110100	1110010	$+3$	α^5	$+4$
6	1110100	1100101	$+3$	α^6	$+4$

where $\lambda = \alpha^\tau$ with $0 \le \tau \le 2^n - 2$, τ is a phase shift of the sequence $\mathbf{a}$, and α is a primitive element of $\mathbb{F}_{2^n}$. From (1.11), the crosscorrelation function $C_{f,g}(\lambda)$ of trace representations has the one-to-one correspondence with the crosscorrelation function $C_{\mathbf{a},\mathbf{b}}(\tau)$ of the corresponding sequences.

Example 4 *In Example 3, let* $\mathbf{a} = (1,0,0,1,0,1,1)$ *and* $\mathbf{b} = \mathbf{a}^{(3)} = (1,1,1,0,1,0,0)$. *Then, their trace representations are given by* $f(x) = Tr(x)$, $g(x) = Tr(x^3)$, *respectively. By computation, the crosscorrelations* $C_{\mathbf{a},\mathbf{b}}(\tau)$, $\tau = 0,1,\cdots,6$ *and* $C_{f,g}(\lambda)$, $\lambda = \alpha^\tau$ *are shown in Table 1.1.*

Hamming Weight and Exponential Sum: Let $\mathbf{a} = (a_0, a_1, \cdots, a_{N-1})$ be a binary codeword of length N, or a binary sequence of period N. The number of 1's in the codeword is called the *(Hamming) weight*. Clearly, the total sum of the additive characters $(-1)^{a_t}$ of a codeword or sequence $\mathbf{a}$ with weight w is given by

$$\sum_{t=0}^{N-1} (-1)^{a_t} = N - 2w.$$

If $N = 2^n - 1$ and $\mathbf{a}$ is represented by the trace representation $f(x)$, then the *exponential sum of* $f(x)$ *over* $\mathbb{F}_{2^n}$ is given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = 1 + \sum_{t=0}^{N-1} (-1)^{a_t} = 2^n - 2w$$

with $f(0) = 0$. Hence, the exponential sum of a polynomial function $f(x)$ has the one-to-one correspondence with the weight of a codeword or a sequence represented by $f(x)$.

Family of Binary Sequences: For r binary cyclically distinct sequences of period N, i.e., $\mathbf{s}_j = (s_{j,0}, \cdots, s_{j,N-1})$ with $0 \leq j < r$, let $\mathcal{S} = \{\mathbf{s}_0, \cdots, \mathbf{s}_{r-1}\}$ and

$$C_{\max} = \max |C_{\mathbf{s}_i,\mathbf{s}_j}(\tau)| \text{ for any } 0 \leq \tau < N,\ 0 \leq i, j < r$$

where $\tau \neq 0$ if $i = j$. Clearly, $C_{\max}$ is the maximum of all nontrivial auto and crosscorrelation of the sequences in $\mathcal{S}$. The set $\mathcal{S}$ is called a $(N, r, C_{\max})$ *signal set* or *family of sequences*, where r is the *set size* or *family size*, and $C_{\max}$ is the *maximum correlation magnitude* of $\mathcal{S}$.

Hadamard Transform: Let $f(x)$ be a polynomial function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. With a trace function $Tr(x)$ from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$, the Hadamard transform of $f(x)$ is defined by

$$\widehat{f}(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+f(x)}, \quad \lambda \in \mathbb{F}_{2^n}.$$

The inverse formula is given by

$$(-1)^{f(\lambda)} = \frac{1}{2^n} \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)} \widehat{f}(x), \quad \lambda \in \mathbb{F}_{2^n}.$$

The Hadamard transform of $f(x)$ is generalized by replacing $Tr(x)$ by another orthogonal function $h(x)$ from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. Then, the *generalized* Hadamard transform is defined by

$$\widehat{f}_h(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{h(\lambda x)+f(x)}, \quad \lambda \in \mathbb{F}_{2^n}. \tag{1.12}$$

The inverse formula of the generalized Hadamard transform is also defined by

$$(-1)^{f(\lambda)} = \frac{1}{2^n} \sum_{x \in \mathbb{F}_{2^n}} (-1)^{h(\lambda x)} \widehat{f}_h(x), \quad \lambda \in \mathbb{F}_{2^n}.$$

Hadamard Equivalence: Let $\widehat{f}(\lambda)$ and $\widehat{g}(\lambda)$ be the Hadamard transforms of polynomial functions $f(x)$ and $g(x)$, respectively. Let $\mathbf{a}$ and $\mathbf{b}$ be binary sequences represented by $f(x)$ and $g(x)$, respectively. For a positive integer s with $\gcd(s, 2^n - 1) = 1$, if we have

$$\widehat{f}(\lambda) = \widehat{g}(\lambda^s), \tag{1.13}$$

then we say that $f(x)$ is *Hadamard equivalent* to $g(x)$ [23]. In sequence aspects, we also say that $\mathbf{a}$ is Hadamard equivalent to $\mathbf{b}$.

Walsh Transform: Let $\mathbf{x} = (x_1, \cdots, x_n)$ be a binary vector in $\mathbb{F}_2^n$ and $f(\mathbf{x})$ a Boolean function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. The Walsh transform of a Boolean function $f(\mathbf{x})$ is defined by

$$\widehat{f}(\mathbf{w}) = \sum_{x \in \mathbb{F}_2^n} (-1)^{\mathbf{w} \cdot \mathbf{x} + f(\mathbf{x})}, \quad \mathbf{w} \in \mathbb{F}_2^n.$$

From the equivalence relation of polynomial and Boolean functions, the Hadamard transform of a polynomial function corresponds to the Walsh transform of the equivalent Boolean function [33].

Hadamard Transform and Correlation: Let $f(x)$ and $g(x)$ be polynomial functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$, and $\mathbf{a}$ and $\mathbf{b}$ be binary sequences of period $2^n - 1$ represented by $f(x)$ and $g(x)$, respectively. From (1.11) and (1.12),

$$C_{\mathbf{a},\mathbf{b}}(\tau) = -1 + C_{f,g}(\lambda) = -1 + \widehat{g}_f(\lambda), \quad \lambda = \alpha^\tau$$

where τ, $0 \le \tau \le 2^n - 2$, is a phase shift of a sequence and α is a primitive element of $\mathbb{F}_{2^n}$. In other words, the correlation of a pair of binary sequences corresponds to the (generalized) Hadamard transform of a pair of the corresponding trace representations.

Example 5 *In Example 4,* $C_{f,g}(\lambda) = C_{Tr,Tr^{(3)}}(\lambda)$ *corresponds to the Hadamard transform of* $g(x) = Tr(x^3)$*, i.e.,* $C_{Tr,Tr^{(3)}}(\lambda) = \widehat{g}(\lambda)$.

Parseval's Equation: Let $f(x)$ and $g(x)$ be polynomial functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$, respectively. Then, the *Parseval's equation* [33] is given by

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{g(x)+f(x)} = \frac{1}{2^n} \sum_{x\in\mathbb{F}_{2^n}} \widehat{g}(x)\widehat{f}(x)$$

where $\widehat{f}(x) = \sum_{y\in\mathbb{F}_{2^n}} (-1)^{Tr(xy)+f(y)}$ is the Hadamard transform of $f(y)$.

Let $h(x)$ be an orthogonal function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. From the generalized Hadamard transform in (1.12), the Parseval's equation is generalized by

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{g(x)+f(x)} = \frac{1}{2^n} \sum_{x\in\mathbb{F}_{2^n}} \widehat{g_h}(x)\widehat{f_h}(x). \tag{1.14}$$

where $\widehat{f_h}(x) = \sum_{y\in\mathbb{F}_{2^n}} (-1)^{h(xy)+f(y)}$ is the generalized Hadamard transform of $f(y)$.

Interleaved Sequences: Let $\mathbf{u} = \{u_t\}$ be a binary sequence of period vw where both v and w are not equal to 1. Then, we can arrange $\mathbf{u}$ by a $v \times w$ matrix, i.e.,

$$U = (\mathbf{u}_0, \mathbf{u}_1, \cdots, \mathbf{u}_{w-1}) = \begin{bmatrix} u_0 & u_1 & \cdots & u_{w-1} \\ u_w & u_{w+1} & \cdots & u_{2w-1} \\ \vdots & \vdots & \cdots & \vdots \\ u_{(v-1)w} & u_{(v-1)w+1} & \cdots & u_{vw-1} \end{bmatrix}.$$

If each column $\mathbf{u}_j$ is either a cyclic shift of a binary sequence $\mathbf{a}$ of period v or a zero sequence, then $\mathbf{u}$ is called the binary (v, w) *interleaved sequence* [34]. According to the definition, $\mathbf{u}_j = L^{e_j}(\mathbf{a})$, $0 \le j \le w-1$ where L^{e_j} denotes a cyclic e_j left shift operation, and $e_j \in \mathbb{Z}_v$ or $e_j = \infty$ if $\mathbf{u}_j = (0, 0, \cdots, 0)$. Here, a transpose notation is omitted because we consider it as a sequence. In the interleaved sequence, $\mathbf{a} = (a_0, a_1, \cdots, a_{v-1})$ and $\mathbf{e} = (e_0, e_1, \cdots, e_{w-1})$ are called the *base sequence* and the *shift sequence* of $\mathbf{u}$, respectively. With the base and the shift sequences, $\mathbf{u}$ is denoted

by $\mathbf{u} = A(\mathbf{a}, \mathbf{e})$, and the matrix U is also used for the array form of $\mathbf{u}$, i.e., $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) = U$.

In $\mathbf{w} = \mathbf{u} + \mathbf{b} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$, the (v, w) interleaved structure is preserved by adding a binary sequence $\mathbf{b}$ of period w. Let $W = (\mathbf{w}_0, \cdots, \mathbf{w}_{w-1}) = \mathbf{w}$. In W, $\mathbf{b} = \{b_j \mid 0 \leq j \leq w-1\}$ is used as the *indicator sequence* where $\mathbf{w}_j = \mathbf{u}_j$ if $b_j = 0$, or $\mathbf{w}_j = \overline{\mathbf{u}}_j$ otherwise.

From next chapters, we will show the research results on the correlation of binary sequences. First, we will investigate the crosscorrelation properties of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation for odd n. Then, we will construct new binary sequences of period of a multiple of 4 with optimal autocorrelation magnitude by applying the interleaved structure of binary sequences. Third, a new binary sequence family with low correlation and a large family size will be proposed, and its asymptotic optimality will be examined. Finally, we will construct quadratic bent functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ of a polynomial form by giving their necessary and sufficient conditions for special values of n.

Chapter 2

Crosscorrelation Properties of Binary Sequences with Ideal Two-level Autocorrelation

In this chapter, crosscorrelation properties of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation are investigated [93]. First, the related works are introduced after literatures survey, and the motivation is given. Then, several classes of recently discovered binary sequences of period $2^n - 1$ with ideal two-level autocorrelation are briefly reviewed by the corresponding cyclic difference sets. Based on the Niho's dissertation [67], at most 5-valued crosscorrelation properties of binary m-sequences are recalled. As main results of this chapter, several theorems and conjectures are established for the 3- and 5-valued crosscorrelation of a pair of binary sequences with ideal two-level autocorrelation.

2.1 Related Works and Motivation

Since 1997, several new classes of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation have been discovered; the Kasami power function (KPF) sequences [23], the Welch-Gong (WG) sequences [70] [69], and the hyperoval sequences [61]. Together with traditionally well known m-sequences, the Gordon-

Mills-Welch (GMW) sequences [37], the generalized GMW sequences, the quadratic residue (QR) sequences, and the Hall's sextic residue sequences, these new classes of sequences constitute all known binary two-level autocorrelation sequences of period $2^n - 1$ [23].

For theory and practice of sequences, it would be interesting to study the crosscorrelation of a pair of binary two-level autocorrelation sequences of period $2^n - 1$. For odd n, the crosscorrelation has been investigated for the following pairs of binary sequences of period $2^n - 1$.

- An m-sequence and its decimations [31] [46] [67] [40] (with the Gold, Kasami, Welch, Niho, and some conjectured exponents)
- An m-sequence and a GMW sequence with the same primitive polynomials [29], and a pair of GMW sequences [5] (The crosscorrelation is reduced to the crosscorrelation of m-sequences)
- An m-sequence and a decimated KPF sequence with one particular exponent [23]
- An m-sequence and a WG sequence without decimation [36]
- An m-sequence and a hyperoval sequence without decimation [22]
- A pair of KPF sequences without decimation [41]

Throughout this chapter, the crosscorrelation of a pair of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation is presented by the equivalent crosscorrelation of trace representations of the pair. For a trace representation of a binary sequence, see Section 1.5. If the maximum crosscorrelation of a pair of binary sequences of period $2^n - 1$ is much larger than its optimum value achieving the Welch [91] or the Sidelnikov bound [86], then the pair is not so attractive for communication and cryptographic applications. For odd n, therefore, the 3-valued crosscorrelation, i.e., $\{0, \pm 2^{\frac{n+1}{2}}\}$, has been intensively studied by many researchers. In this work, we are also interested in the 5-valued crosscorrelation, i.e., $\{0, \pm 2^{\frac{n+1}{2}}, \pm 2^{\frac{n+3}{2}}\}$, which might be suboptimal for some applications.

In this chapter, the 3- or 5-valued crosscorrelation of a pair of binary two-level autocorrelation sequences of period $2^n - 1$ is studied for odd n, excluding the (generalized) GMW, QR, and Hall's sextic residue sequences. First, at most 5-valued crosscorrelation of m-sequences is discussed, which is linked to crosscorrelation of some other sequences. Then, the 3- or 5-valued crosscorrelation of the following pairs is investigated.

- A 5-term KPF sequence and a decimated WG sequence with one new exponent
- An m-sequence and a decimated WG sequence with one new exponent
- An m-sequence and a decimated hyperoval sequence with several new exponents
- An m-sequence and a decimated 3-term KPF sequence with one new exponent

With the new and the already known results, the relation of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation is summarized with respect to the 3- and 5-valued crosscorrelation. From our experiments for $n = 13, 15, 17$, and 19, it is observed that every 3- and 5-valued crosscorrelation of a pair of binary two-level autocorrelation sequences considered in this work is completely described by the new and the already known results listed above unless both are m-sequences.

2.2 Recently Constructed Binary Two-level Autocorrelation Sequences

This section briefly introduces three classes of binary two-level autocorrelation sequences of period $2^n - 1$ which have been constructed recently. For ideal two-level autocorrelation, recall the definition in Section 1.5.

Kasami Power Function (KPF) Sequences: Let k be an integer of $1 \leq k < \lfloor \frac{n}{2} \rfloor$

with $\gcd(k, n) = 1$. For $d = 2^{2k} - 2^k + 1$, consider a set

$$B_k = \{(x+1)^d + x^d + 1 \mid x \in \mathbb{F}_{2^n}\}.$$

Then, its characteristic sequence given by

$$a_i = \begin{cases} 0, & \text{if } \alpha^i \in B_k \\ 1, & \text{if } \alpha^i \notin B_k \end{cases}$$

has the ideal two-level autocorrelation, where the sequence is called the *Kasami power function (KPF) sequence* [23]. According to k with $\gcd(k, n) = 1$, there exist $\frac{\phi(n)}{2}$ inequivalent KPF sequences of period $2^n - 1$, where $\phi(\cdot)$ is the Euler-totient function. If $k = 1$, in particular, the KPF sequence is identical to an m-sequence.

Let $b_k(x)$ be the trace representation of the KPF sequence. For odd n, the KPF sequence has the Hadamard equivalence [23] given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x) + b_k(x^{2^k+1})} = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{\frac{2^k+1}{3}} x) + Tr(x^3)} = H_3(\lambda^{\frac{2^k+1}{3}}) \tag{2.1}$$

where $H_d(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x) + Tr(x^d)}$. Since '3' is the Gold exponent [31], (2.1) is 3-valued. For the definition of Hadamard equivalence, see Section 1.5.

Welch-Gong (WG) Sequences: For $n = 3k \pm 1$ and $d = 2^{2k} - 2^k + 1$, consider a map $\delta_k(x) = (x+1)^d + x^d$ and a set

$$W_k = \begin{cases} \delta_k(x), & \text{if } n \text{ is even} \\ \mathbb{F}_{2^n} \setminus \delta_k(x), & \text{if } n \text{ is odd.} \end{cases}$$

Then, its characteristic sequence given by

$$a_i = \begin{cases} 0, & \text{if } \alpha^i \in W_k \\ 1, & \text{if } \alpha^i \notin W_k \end{cases}$$

has the ideal two-level autocorrelation [69]. This sequence is identical to the *Welch-*

Gong sequence, which is obtained from the Welch-Gong transformation of the 5-term sequences [70].

Let $w_k(x)$ be the trace representation of the WG sequence. For odd n, the WG sequence has the Hadamard equivalence [36] given by

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+w_k(x)} = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{d^{-1}}x)+Tr(x^{2^k+1})} = H_{2^k+1}(\lambda^{d^{-1}}) \qquad (2.2)$$

which is also 3-valued from the Gold exponent 2^k+1.

Hyperoval Sequences: For odd n, consider a set

$$M_k = \{x + x^k \mid x \in \mathbb{F}_{2^n}\}$$

where k is given as follows [61].

i) Singer type: $k = 2$, Segre type: $k = 6$.

ii) Glynn type I: $k = 2^\sigma + 2^\tau$ where $\sigma = \frac{n+1}{2}$ and $4\tau \equiv 1 \pmod{n}$.

iii) Glynn type II: $k = 3 \cdot 2^\sigma + 4$ with $\sigma = \frac{n+1}{2}$.

Then, the characteristic sequence of M_k given by

$$a_i = \begin{cases} 0, & \text{if } \alpha^i \in M_k \\ 1, & \text{if } \alpha^i \notin M_k \end{cases}$$

has the ideal two-level autocorrelation, where the sequence is called the *hyperoval sequence*. In this chapter, we are only interested in the Glynn type I and II hyperoval sequences because the Singer and the Segre type hyperoval sequences are identical to m-sequences and the KPF sequences represented by $b_2(x)$, respectively [23].

Let $h_k(x)$ be the trace representation of the hyperoval sequence. For odd n, Dillon derived the Hadamard equivalence of the hyperoval sequence [22], i.e.,

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+h_k(x)} = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{\frac{k-1}{k}}x)+Tr(x^k)} = H_k(\lambda^{\frac{k-1}{k}}). \qquad (2.3)$$

Table 2.1: Exponents d for 3-valued $H_d(\lambda)$ for odd n

	d	k
Gold	2^k+1	$\gcd(k,n)=1,\ 1\le k\le \frac{n-1}{2}$
Kasami	$2^{2k}-2^k+1$	$\gcd(k,n)=1,\ 1\le k\le \frac{n-1}{2}$
Welch	$2^{\frac{n-1}{2}}+3$	
Niho	$2^{2k}+2^k-1$	$k=\begin{cases}\frac{n-1}{4}, & \text{if } n\equiv 1 \pmod 4\\ \frac{3n-1}{4}, & \text{if } n\equiv 3 \pmod 4\end{cases}$

If k is the Glynn type I exponent in ii), then (2.3) is 3-valued because k is quadratic. If k is the Glynn type II exponent in iii), on the other hand, then (2.3) is conjectured to be at most 5-valued because $k=3\cdot 2^{\frac{n+1}{2}}+4\equiv 2^{\frac{n-1}{2}}+2^{\frac{n-3}{2}}+1 \pmod{2^n-1}$ is equivalently an inverse of the exponent of Conjecture 4-6 (1) in [67], where $H_{k^{-1}}(\lambda)$ is conjectured to be at most 5-valued by Niho. This will be restated in Conjecture 2 of this chapter.

In Appendix A, the trace representations of the KPF, WG, and hyperoval sequences of period 2^n-1 are given for $n=13,15,17$, and 19 in Tables A.1 – A.5. For such sequences with $n\le 11$, see [33]. In each table, the KPF sequence corresponding to B_1 and the Segre type hyperoval sequence are identical to an m-sequence and the KPF sequence corresponding to B_2, respectively.

2.3 Some Observations of Crosscorrelation of Binary m-sequences

This section recalls at most 5-valued crosscorrelation of a binary m-sequence and its d-decimation, i.e., $H_d(\lambda)=\sum_{x\in\mathbb{F}_{2^n}}(-1)^{Tr(\lambda x)+Tr(x^d)}$. In terms of the 3-valued $H_d(\lambda)$, many exponents d are known, i.e., the Gold [31], Kasami [46], Welch, Niho [67] exponents, and their respective inverses. Table 2.1 lists the known exponents d for 3-valued $H_d(\lambda)$.

In terms of the 5-valued $H_d(\lambda)$, on the other hand, we need to clarify the known results.

Proposition 1 *Let n be odd, t be a positive integer of $1\le t\le\frac{n-1}{2}$, and $e=$*

$\gcd(n,t)$ *with* $n/e \geq 4$. *Let* $d(k,l) = (1+2^k)/(1+2^l)$ *with positive integers* k *and* l $(k \neq l)$. *Then,* $H_{d(k,l)}(\lambda)$ *belongs to* $\{0, \pm 2^{(n+e)/2}, \pm 2^{(n+3e)/2}\}$ *if a pair* (k,l) *is one of the following three cases*

$$(a)\ (k,l) = (5t,t), \quad (b)\ (k,l) = (5t,3t), \quad (c)\ (k,l) = (2t,t)$$

where the multiplication is computed modulo n. *If* $e = 1$, *in particular,* $H_{d(k,l)}(\lambda)$ *is at most 5-valued, i.e.,* $\{0, \pm 2^{(n+1)/2}, \pm 2^{(n+3)/2}\}$.

Proposition 1-(a) has been proven by Niho (Lemma 4-1 in [67]). Although he had never stated Proposition 1-(b) and (c) in [67], it is believed that those have been implicitly known to many coding and sequence experts. In literatures, however, an explicit proof for (b) and (c), which is not trivial, could not be found. So, this section presents it because the result is linked to crosscorrelation of some other binary two-level autocorrelation sequences. In order to prove Proposition 1, the Kasami's Theorem on the weight distribution of subcodes of the second order Reed-Muller codes is needed, which was partly used by Niho to prove Proposition 1-(a). In the following, the odd case of Kasami's original theorem in [46] is considered.

Fact 1 (Kasami [46]) *For odd* n, *let* t *and* u *be positive integers with* $1 \leq t \leq \frac{n-1}{2}$ *and* $1 \leq u \leq \lfloor \frac{n}{2e} \rfloor + 1$ *where* $e = \gcd(n,t)$. *Let* $A_t(u)$ *be a binary cyclic code of length* $2^n - 1$ *whose generator polynomial is given by* $g_a(x) = \prod_{i=0}^{u-1} m_{1+2^{ti}}(x)$ *where* $m_i(x)$ *is a minimal polynomial of* α^i *and* α *is a primitive element of* $\mathbb{F}_{2^n}$. *Similarly, let* $F_t(u)$ *be a binary cyclic code of length* $2^n - 1$ *whose generator polynomial is given by* $g_f(x) = \prod_{i=0}^{u-1} m_{1+2^{t(2i+1)}}(x)$. *Dual codes of* $A_t(u)$ *and* $F_t(u)$ *are denoted by* $A_t(u)^\perp$ *and* $F_t(u)^\perp$, *respectively. Then,* $A_t(u)^\perp$ *and* $F_t(u)^\perp$ *have the same weight distribution as those of* $A_e(u)^\perp$ *whose distinct weights are given by*

$$\{0, 2^{n-1}, 2^{n-1} \pm 2^{(n-e)/2+ie-1}\} \text{ for } 1 \leq i \leq u-1.$$

Using Fact 1, Proposition 1 can be proved.

Proof of Proposition 1. In (a) and (b), $H_{d(k,l)}(\lambda)$ is represented by

$$H_{d(k,l)}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x + x^{\frac{1+2^k}{1+2^l}})} = \begin{cases} \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x^{1+2^t} + x^{1+2^{5t}})} & \text{for } (a) \\ \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x^{1+2^{3t}} + x^{1+2^{5t}})} & \text{for } (b). \end{cases}$$

Then, the linear cyclic codes $\mathcal{R}_5$ and $\mathcal{R}_{5/3}$ given by

$$\mathcal{R}_5 = \{Tr(\eta x^{1+2^t} + \beta x^{1+2^{5t}}) \mid \eta, \beta \in \mathbb{F}_{2^n}\},$$
$$\mathcal{R}_{5/3} = \{Tr(\gamma x^{1+2^{3t}} + \delta x^{1+2^{5t}}) \mid \gamma, \delta \in \mathbb{F}_{2^n}\}$$

are subcodes of $F_t(u)^\perp$ for $u = 3$ where $F_t(u)$ has zeros $\alpha^{1+2^{t(2i+1)}}, i = 0, 1, 2$. For any t of $1 \le t \le \frac{n-1}{2}$, therefore, the weight distributions of $\mathcal{R}_5$ and $\mathcal{R}_{5/3}$ are immediate from Fact 1, and consequently $H_{d(k,l)}(\lambda)$ belongs to $\{0, \pm 2^{\frac{n+e}{2}}, \pm 2^{\frac{n+3e}{2}}\}$ for both (a) and (b).

In (c), on the other hand, $A_t(3)$ generated by $g_a(x)$ has zeros $\{\alpha^2, \alpha^{1+2^t}, \alpha^{1+2^{2t}}\}$, so the linear cyclic code $\mathcal{R}_2$ given by

$$\mathcal{R}_2 = \{Tr(\zeta x^{1+2^t} + \theta x^{1+2^{2t}}) \mid \zeta, \theta \in \mathbb{F}_{2^n}\}$$

is also a subcode of $A_t(3)^\perp$. From Fact 1, therefore, it is clear that $H_{d(k,l)}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x^{1+2^t} + x^{1+2^{2t}})}$ belongs to $\{0, \pm 2^{\frac{n+e}{2}}, \pm 2^{\frac{n+3e}{2}}\}$. □

Lemma 1 *For odd n, let k and l be positive integers of $1 \le k, l \le \frac{n-1}{2}$ $(k \ne l)$, and $d(k,l) = \frac{1+2^k}{1+2^l}$. Then, $H_{d(n-k,l)}(\lambda), H_{d(k,n-l)}(\lambda)$, and $H_{d(n-k,n-l)}(\lambda)$ have the same correlation spectrum as $H_{d(k,l)}(\lambda)$. Furthermore, $H_{d(l,k)}(\lambda)$ also belongs to the same correlation spectrum as $H_{d(k,l)}(\lambda)$.*

Proof. Note that $H_{d\cdot 2^j}(\lambda) = H_d(\lambda)$ for any integer j [40]. Since $2^{n-k} \cdot (1 + 2^k) = 2^{n-k} + 2^n \equiv 2^{n-k} + 1 \pmod{2^n - 1}$, it is clear that $1 + 2^k$ and $1 + 2^{n-k}$ belong to the same cyclotomic coset. Hence, $d(k,l)$ belongs to the same cyclotomic coset as $d(n-k,l)$. Therefore, $H_{d(k,l)}(\lambda)$ and $H_{d(n-k,l)}(\lambda)$ have the same correlation distribution. By the similar way, the cases of $H_{d(k,n-l)}(\lambda)$ and $H_{d(n-k,n-l)}(\lambda)$ are simply proved. From $d(l,k) = d(k,l)^{-1}$, furthermore, it is immediate that $H_{d(k,l)}(\lambda)$

Table 2.2: (k,l) pairs and corresponding $d(k,l)$'s for 5-valued crosscorrelation of $Tr(x)$ and $Tr(x^{d(k,l)})$

n	(k,l)	$d(k,l)$	n	(k,l)	$d(k,l)$	n	(k,l)	$d(k,l)$	n	(k,l)	$d(k,l)$
9	$(2,1)^{*,o}$	43	13	$(4,2)^{+,o}$	1645	15	$(7,4)^{o}$	2895	17	$(4,3)^{*}$	14571
9	$(4,1)^{*,o}$	11	13	$(5,3)^{+,o}$	1367	15	$(6,5)^{+}$	1119	17	$(5,3)^{+}$	21847
9	$(4,2)^{*,o}$	109	13	$(6,3)^{+,o}$	939	15	$(7,5)^{*}$	3229	17	$(6,3)^{o}$	14679
11	$(2,1)^{*,+,o}$	171	13	$(5,4)^{+,o}$	1461	17	$(2,1)^{o}$	10923	17	$(7,3)^{o}$	15019
11	$(5,1)^{*,+,o}$	11	13	$(6,4)^{*}$	497	17	$(4,1)^{+}$	2731	17	$(6,4)^{*}$	11567
11	$(4,2)^{*,+,o}$	423	15	$(2,1)^{o}$	2731	17	$(5,1)^{*}$	11	17	$(8,4)^{o}$	7831
11	$(4,3)^{*,+,o}$	235	15	$(5,1)^{*}$	11	17	$(7,1)^{*}$	43	17	$(6,5)^{o}$	12909
11	$(5,3)^{*,+,o}$	343	15	$(7,1)^{o}$	43	17	$(8,1)^{o}$	171	17	$(7,5)^{o}$	13917
13	$(2,1)^{+,o}$	683	15	$(4,2)^{o}$	6567	17	$(3,2)^{*}$	3277	17	$(8,5)^{*}$	4003
13	$(5,1)^{*}$	11	15	$(5,2)^{*}$	205	17	$(4,2)^{o}$	26221	17	$(7,6)^{+}$	10587
13	$(6,1)^{+,o}$	43	15	$(5,3)^{+}$	5463	17	$(7,2)^{*}$	205	17	$(8,6)^{*}$	2143
13	$(3,2)^{*}$	205	15	$(5,4)^{*}$	1943	17	$(8,2)^{+}$	26317			

and $H_{d(l,k)}(\lambda)$ belong to the same correlation spectrum. □

Table 2.2 shows (k,l) pairs and $d(k,l) = \frac{1+2^k}{1+2^l}$ corresponding to the 5-valued $H_{d(k,l)}(\lambda)$ from computer experiments. It presents only the pairs of $1 \le l < k \le \frac{n-1}{2}$ which are sufficient to cover the other possible pairs from Lemma 1. Each pair of '*' is due to (a), '+' due to (b), and 'o' due to (c) in Proposition 1, respectively. For odd n, $9 \le n \le 17$, Proposition 1 is verified from the experiments.

2.4 Crosscorrelation of A Pair of Binary Two-level Autocorrelation Sequences

2.4.1 A Pair of KPF Sequences

In [41], Hertel investigated the crosscorrelation of two distinct KPF sequences for odd n. (She called the sequences as the Dillon-Dobbertin (DD) sequences named after their discoverers.)

Fact 2 (Hertel [41]) *For odd n, let k and l be distinct positive integers with* $\gcd(n,k) = \gcd(n,l) = 1$. *Let $b_k(x)$ and $b_l(x)$ be the trace representations of two*

distinct KPF sequences, respectively. Then,

$$C_{b_k,b_l}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{b_k(\lambda x)+b_l(x)} = H_{d(k,l)}(\lambda^{\frac{1}{1+2^k}}), \quad \lambda \in \mathbb{F}_{2^n}$$

where $d(k,l) = \frac{1+2^k}{1+2^l}$. *If* $(k,l) = (3t,t)$, *in particular,* $C_{b_k,b_l}(\lambda)$ *is 3-valued, i.e.,* $\{0, \pm 2^{\frac{n+1}{2}}\}$.

Corollary 1 *With the notation of Proposition 1 and Fact 2, if a pair* (k,l) *is one of the pairs in Proposition 1, then* $C_{b_k,b_l}(\lambda)$ *is at most 5-valued, i.e.,* $\{0, \pm 2^{\frac{n+1}{2}}, \pm 2^{\frac{n+3}{2}}\}$.

Proof. Corollary 1 is immediate by combining Proposition 1 and Fact 2. □

From Corollary 1, it is obvious that the crosscorrelation of $b_k(x)$ and $b_l(x)$ with a (k,l) pair in Table 2.2 is 5-valued.

2.4.2 5-term KPF Sequences and Welch-Gong (WG) Sequences

The WG sequences are obtained from the Welch-Gong transformation of the KPF sequences for $k = \frac{n\pm1}{3}$, where the KPF sequences always have five trace terms [23] [70]. By the Parseval's equation exploited in [41], one theorem on the crosscorrelation of the 5-term KPF and the WG sequences is derived. For the exact formula of the Parseval's equation, see Section 1.5.

Theorem 1 *Let* n *be odd and* $n = 3k \pm 1$. *Let* $b_k(x)$ *and* $w_k(x)$ *be the trace representations of the KPF sequences and the WG sequences, respectively. For* $s = \frac{1}{2^k+1}$, *the crosscorrelation of the two sequences given by*

$$C_{b_k,w_k^{(s)}}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{b_k(\lambda x)+w_k(x^s)} = H_{\frac{2^k+1}{3}}(\lambda)$$

is 3-valued, i.e., $\{0, \pm 2^{\frac{n+1}{2}}\}$.

Proof. Applying the Parseval's equation,

$$\begin{aligned} C_{b_k,w_k^{(s)}}(\lambda) =& \frac{1}{2^n}\sum_{x\in\mathbb{F}_{2^n}}\sum_{y\in\mathbb{F}_{2^n}}(-1)^{b_k(\lambda y)+Tr(xy^{\frac{1}{2^k+1}})}\sum_{z\in\mathbb{F}_{2^n}}(-1)^{w_k(z^s)+Tr(xz^{\frac{1}{2^k+1}})} \\ =& \frac{1}{2^{2n}}\sum_{x,z\in\mathbb{F}_{2^n}}\sum_{y\in\mathbb{F}_{2^n}}(-1)^{b_k(y)+Tr(x\lambda^{-\frac{1}{2^k+1}}y^{\frac{1}{2^k+1}})} \\ & \cdot\sum_{u\in\mathbb{F}_{2^n}}(-1)^{w_k(u^s)+Tr(zu^s)}\sum_{v\in\mathbb{F}_{2^n}}(-1)^{Tr(xv^{\frac{1}{2^k+1}})+Tr(zv^s)}. \end{aligned} \tag{2.4}$$

If $s=\frac{1}{2^k+1}$, then

$$\sum_{v\in\mathbb{F}_{2^n}}(-1)^{Tr(xv^{\frac{1}{2^k+1}})+Tr(zv^s)}=\sum_{v\in\mathbb{F}_{2^n}}(-1)^{Tr((x+z)v^{\frac{1}{2^k+1}})}=\begin{cases}2^n, & \text{if } x=z\\ 0, & \text{if } x\neq z.\end{cases}$$

If the Hadamard equivalences (2.1) and (2.2) are applied to (2.4), then we have

$$\begin{aligned} C_{b_k,w_k^{(s)}}(\lambda) =& \frac{1}{2^n}\sum_{x\in\mathbb{F}_{2^n}}\sum_{y\in\mathbb{F}_{2^n}}(-1)^{Tr(\lambda^{-\frac{1}{3}}x^{\frac{2^k+1}{3}}y)+Tr(y^3)}\sum_{u\in\mathbb{F}_{2^n}}(-1)^{Tr(x^bu)+Tr(u^a)} \\ =& \frac{1}{2^n}\sum_{y\in\mathbb{F}_{2^n}}\sum_{u\in\mathbb{F}_{2^n}}(-1)^{Tr(y^3)+Tr(u^a)}\sum_{x\in\mathbb{F}_{2^n}}(-1)^{Tr(\lambda^{-\frac{1}{3}}x^{\frac{2^k+1}{3}}y)+Tr(ux^b)} \end{aligned}$$

where $a=2^k+1$ and $b=(2^{2k}-2^k+1)^{-1}$. If $3k=n+1$, it is clear that $b^{-1}\cdot\frac{2^k+1}{3}=\frac{2^{3k}+1}{3}\equiv 1 \pmod{2^n-1}$. If $3k=n-1$, then $b^{-1}\cdot\frac{2^k+1}{3}=\frac{2^{3k}+1}{3}=\frac{2^{n-1}+1}{3}$. On the other hand, $2\cdot\frac{2^{n-1}+1}{3}\equiv 1 \pmod{2^n-1}$. Since both belong to the same cyclotomic coset, we can choose $b\equiv\frac{2^k+1}{3} \pmod{2^n-1}$. Consequently,

$$\begin{aligned} C_{b_k,w_k^{(s)}}(\lambda) =& \frac{1}{2^n}\sum_{y\in\mathbb{F}_{2^n}}\sum_{u\in\mathbb{F}_{2^n}}(-1)^{Tr(y^3)+Tr(u^a)}\sum_{x\in\mathbb{F}_{2^n}}(-1)^{Tr((\lambda^{-\frac{1}{3}}y+u)x^{\frac{2^k+1}{3}})} \\ =& \sum_{u\in\mathbb{F}_{2^n}}(-1)^{Tr(\lambda u^3+u^a)}=\sum_{u\in\mathbb{F}_{2^n}}(-1)^{Tr(\lambda u)+Tr(u^{\frac{2^k+1}{3}})}=H_{\frac{2^k+1}{3}}(\lambda) \end{aligned} \tag{2.5}$$

where $y = \lambda^{\frac{1}{3}}u$. In (2.5), $\frac{2^k+1}{3} \equiv b = (2^{2k} - 2^k + 1)^{-1}$. Since it is the inverse of the Kasami exponent with $\gcd(n, k) = 1$, it is obvious that $H_{\frac{2^k+1}{3}}(\lambda)$ is 3-valued and so is $C_{b_k, w_k^{(s)}}(\lambda)$. □

2.4.3 m-sequences and Welch-Gong (WG) Sequences

In an effort to search for new two-level autocorrelation sequences, Gong and Golomb proposed the *decimation-Hadamard transform (DHT)* in [36] by generalizing the Hadamard equivalence developed in [23]. With respect to orthogonal functions $f(x)$, $g(x)$, and $h(x)$, they defined a *realizable pair* (v, t) of a *realization* $g(x)$ in the DHT where

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{h(\lambda^t x) + f(x^v)} = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{h(\lambda x) + g(x)}. \tag{2.6}$$

They also showed that there exist at most 6 realizable pairs for the realization. Among them, if (v, t) is a realizable pair of $g(x)$, then $(t, -(vt)^{-1})$ is also a realizable pair of $g(x^{(vt)^{-1}})$ [36] from which we have

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{h(\lambda^{-(vt)^{-1}} x) + f(x^t)} = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{h(\lambda x) + g(x^{(vt)^{-1}})}. \tag{2.7}$$

Using this, one theorem on the crosscorrelation of m-sequences and the WG sequences is established.

Theorem 2 *Let $n = 3k \pm 1$ be odd and $d = 2^{2k} - 2^k + 1$. Let $w_k(x)$ be the trace representation of the WG sequences. For $s = \frac{d}{2^k+1}$, the crosscorrelation of m-sequences and the WG sequences given by*

$$C_{Tr, w_k^{(s)}}(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{Tr(\lambda x) + w_k(x^s)} = H_{d^{-1}}(\lambda^{-s})$$

is 3-valued, i.e., $\{0, \pm 2^{\frac{n+1}{2}}\}$.

Proof. From the Hadamard equivalence of (2.2), we have a realizable pair $(v, t) = (2^k + 1, d^{-1})$ in (2.6) where $f(x) = h(x) = Tr(x)$ and $g(x) = w_k(x)$. From (2.7),

therefore,

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{-\frac{d}{2^k+1}}x)+Tr(x^{d^{-1}})} = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+w_k(x^{\frac{d}{2^k+1}})}.$$

Thus, $C_{Tr,w_k^{(s)}}(\lambda) = H_{d^{-1}}(\lambda^{-s})$ for $s = \frac{d}{2^k+1}$. Since d is the Kasami exponent with $\gcd(n,k) = 1$, $H_{d^{-1}}(\lambda^{-s})$ is 3-valued and so is $C_{Tr,w_k^{(s)}}(\lambda)$. □

2.4.4 m-sequences and Hyperoval Sequences

Applying (2.7) to hyperoval sequences with the Hadamard equivalence of (2.3), another Hadamard equivalence can be derived, i.e.,

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+h_k(x^{\frac{1}{k-1}})} = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{-\frac{1}{k-1}}x)+Tr(x^{\frac{k-1}{k}})} = H_{\frac{k-1}{k}}(\lambda^{-\frac{1}{k-1}}). \quad (2.8)$$

From (2.8), one theorem for the Glynn type II hyperoval sequences is derived.

Theorem 3 *Let n be odd and $k = 3\cdot 2^\sigma + 4$ where $\sigma = \frac{n+1}{2}$. Let $h_k(x)$ be the trace representation of the Glynn type II hyperoval sequences. For $s = \frac{1}{k-1}$, the crosscorrelation of m-sequences and the Glynn type II hyperoval sequences given by*

$$C_{Tr,h_k^{(s)}}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+h_k(x^s)} = H_{\frac{k-1}{k}}(\lambda^{-s}) \quad (2.9)$$

is at most 5-valued, i.e., $\{0, \pm 2^{\frac{n+1}{2}}, \pm 2^{\frac{n+3}{2}}\}$.

Proof. From (2.8), the correlation spectrum $C_{Tr,h_k^{(s)}}(\lambda)$ is determined by a decimation factor $\frac{k-1}{k}$ of a trace function. Note that the cyclotomic coset that $\frac{k-1}{k}$ belongs to does not change by multiplying its numerator and denominator by $2^{\frac{n-1}{2}}$ and

$2^{\frac{n-3}{2}}$, respectively. Then,

$$\frac{k-1}{k} \equiv \frac{2^{\frac{n-1}{2}}}{2^{\frac{n-3}{2}}} \cdot \frac{(k-1)}{k} = \frac{2^{\frac{n-1}{2}}}{2^{\frac{n-3}{2}}} \cdot \frac{3 \cdot (1+2^{\frac{n+1}{2}})}{(2^{\frac{n+3}{2}} + 2^{\frac{n+1}{2}} + 4)} \equiv \frac{3 \cdot (1+2^{\frac{n-1}{2}})}{(1+2^{\frac{n-1}{2}})^2}$$
$$= \frac{1+2}{1+2^{\frac{n-1}{2}}} \pmod{2^n - 1}.$$

Hence, $\frac{k-1}{k} \equiv \frac{1+2^{\mu}}{1+2^{\nu}} = d(\mu, \nu)$ in Proposition 1 where $\mu = 1$ and $\nu = \frac{n-1}{2}$. Since $2\nu = n - \mu$, we have $(n-\mu, \nu) = (2t, t)$ with $t = \frac{n-1}{2}$, a pair of Proposition 1-(c). From $e = \gcd(n, t) = \gcd(n, \frac{n-1}{2}) = 1$, it is clear that $H_{d(n-\mu,\nu)}(\lambda)$ is at most 5-valued and so is $H_{d(\mu,\nu)}(\lambda)$ from Lemma 1, which completes the proof. □

In terms of the Glynn type I hyperoval sequences, on the other hand, $k = 2^{\sigma} + 2^{\tau}$ where $\sigma = \frac{n+1}{2}$ and $\tau = \frac{n+1}{4}$ or $\tau = \frac{3n+1}{4}$ such that $4\tau \equiv 1 \pmod{n}$. Using the similar approach to the proof of Theorem 3, the following equivalence of $\frac{k-1}{k}$ is established.

$$\frac{k-1}{k} \equiv \begin{cases} 2^{\frac{n-1}{2}} - 2^{\frac{n+1}{4}} + 1 = r_1, & \text{if } \tau = \frac{n+1}{4} \\ 2^{\frac{n+1}{2}} - 2^{\frac{n+3}{4}} + 1 = r_2, & \text{if } \tau = \frac{3n+1}{4}. \end{cases} \tag{2.10}$$

In (2.10), the decimation factors r_1 and r_2 are identical to the ones in Conjecture 4-6 (3) and (4) of [67], respectively, where $H_{r_1}(\lambda)$ and $H_{r_2}(\lambda)$ are conjectured to be at most 5-valued by Niho in 1972. From the Hadamard equivalence (2.8), the following conjecture is established.

Conjecture 1 *Let n be odd and $k = 2^{\sigma} + 2^{\tau}$ where $\sigma = \frac{n+1}{2}$ and $4\tau \equiv 1 \pmod{n}$. Let $h_k(x)$ be the trace representation of the Glynn type I hyperoval sequences. For $s = \frac{1}{k-1}$, the crosscorrelation of m-sequences and the Glynn type I hyperoval sequences given by $C_{Tr,h_k^{(s)}}(\lambda)$ in (2.9) is at most 5-valued, i.e., $\{0, \pm 2^{\frac{n+1}{2}}, \pm 2^{\frac{n+3}{2}}\}$.*

With respect to the crosscorrelation of m-sequences and the Glynn type II hyperoval sequences, another exponent corresponding to at most 5-valued crosscorrelation is also observed. In (2.3), furthermore, the Glynn type II exponent k is conjectured to be at most 5-valued because $k = 3 \cdot 2^{\frac{n+1}{2}} + 4 \equiv 2^{\frac{n-1}{2}} + 2^{\frac{n-3}{2}} + 1$

(mod 2^n-1) is equivalently an inverse of the exponent of Conjecture 4-6 (1) in [67], where $H_{k^{-1}}(\lambda)$ is conjectured to be at most 5-valued by Niho. With these results, Conjecture 2 is established.

Conjecture 2 *For odd n, let $h_k(x)$ be the trace representation of the Glynn type II hyperoval sequences. For $s = 1$ or $\frac{1}{3}$, the crosscorrelation of m-sequences and the Glynn type II hyperoval sequences given by $C_{Tr,h_k^{(s)}}(\lambda)$ is at most 5-valued, i.e., $\{0, \pm 2^{\frac{n+1}{2}}, \pm 2^{\frac{n+3}{2}}\}$.*

Conjectures 1 and 2 have been verified for odd n, $9 \leq n \leq 21$, through computer experiments. We believe that Conjectures 1 and 2 are hard to prove because both are equivalent to the Niho's conjectures [67] in 1972, which have been unsolved for more than 30 years.

2.4.5 m-sequences and 3-term KPF Sequences

In [23], the 3-term KPF sequences are represented by

$$b_k(x) = Tr(x + x^{2^k+1} + x^{2^k-1}), \quad k = \frac{n+1}{2}$$

where n is odd. On the other hand, *T3 sequences*, or the 3-term sequences with ideal two-level autocorrelation which had been conjectured in [70] are represented by

$$T_3(x) = Tr(x + x^r + x^{r^2}), \quad r = 2^{\frac{n-1}{2}} + 1.$$

With the equivalence under modulo $2^n - 1$, it is clear that the T3 sequences are the decimation of the 3-term KPF sequences, i.e., $T_3(x) = b_k(x^{2^k+1})$ where $k = \frac{n+1}{2}$. Using this relation, the following theorem is established.

Theorem 4 *Let n be odd and $k = \frac{n+1}{2}$. Let $b_k(x)$ be the trace representation of the 3-term KPF sequences. For $s = 2^k - 1$, the crosscorrelation of m-sequences and the*

3-term KPF sequences given by

$$C_{Tr,b_k^{(s)}}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+b_k(x^s)}$$

is at most 5-valued, i.e., $\{0, \pm 2^{\frac{n+1}{2}}, \pm 2^{\frac{n+3}{2}}\}$.

Proof. In [17], Chang *et al.* showed that the binary cyclic codes represented by

$$\mathcal{T} = \{Tr(ax + bx^r + cx^{r^2}) \mid a, b, c \in \mathbb{F}_{2^n},\ r = 2^{\frac{n-1}{2}} + 1\}$$

have five nonzero distinct weights as a dual of triple error correcting cyclic codes. Then, the crosscorrelation of m-sequences and the T3 sequences given by $C_{Tr,T_3}(\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda x)+T_3(x)}$ is at most 5-valued - in fact, 3-valued - because the exponent of (-1) in the summation is a codeword of $\mathcal{T}$.

In the following, another at most 5-valued crosscorrelation $C_{Tr^{(r^2)},T_3}(\lambda)$ can be considered. With $r = 2^{\frac{n-1}{2}}+1$, note that $2^{\frac{n+1}{2}} \cdot r = 2^n + 2^{\frac{n+1}{2}} \equiv 1 + 2^k \pmod{2^n - 1}$, and thus $r \equiv 2^k + 1$. Therefore, $T_3(x) = b_k(x^r)$ where $k = \frac{n+1}{2}$. Then,

$$\begin{aligned} C_{Tr^{(r^2)},T_3}(\lambda) &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{r^2}x^{r^2})+T_3(x)} = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{r^2}x)+T_3(x^{r^{-2}})} \\ &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{Tr(\lambda^{r^2}x)+b_k(x^{r^{-1}})} = C_{Tr,b_k^{(s)}}(\lambda^{r^2}) \end{aligned}$$

where $s = r^{-1} = (2^k + 1)^{-1} \equiv 2^k - 1 \pmod{2^n - 1}$. Since the exponent of (-1) in $C_{Tr^{(r^2)},T_3}(\lambda)$ is also a codeword of $\mathcal{T}$, $C_{Tr^{(r^2)},T_3}(\lambda)$ is at most 5-valued and so is $C_{Tr,b_k^{(s)}}(\lambda)$. □

2.5 Conclusion

In this chapter, the 3- and 5-valued crosscorrelation of a pair of binary sequences of period $2^n - 1$ with ideal two-level autocorrelation has been studied, which is given

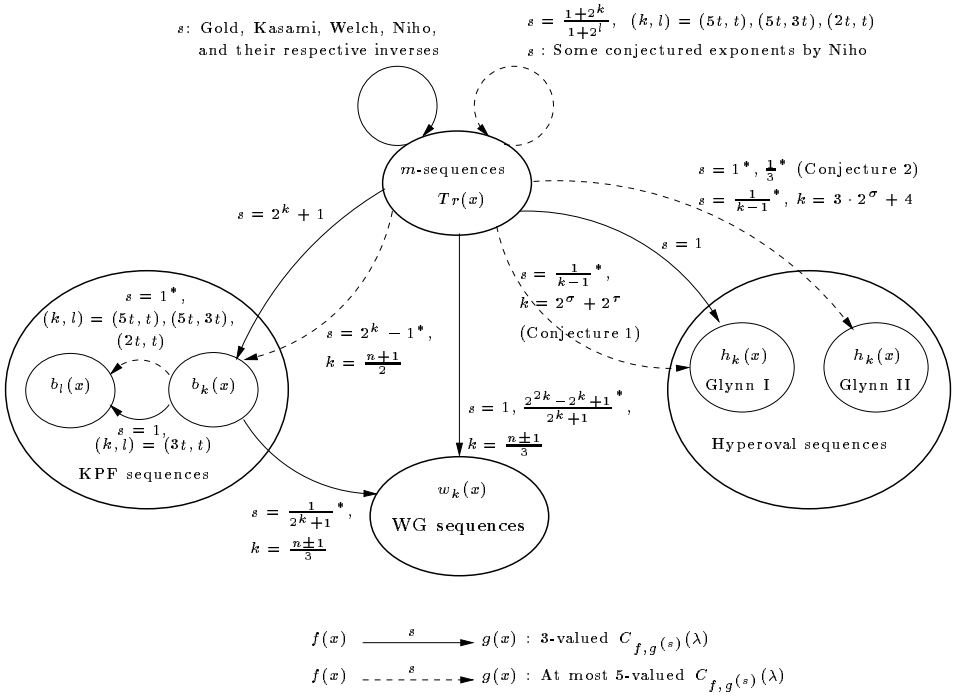


Figure 2.1: Relation of binary two-level autocorrelation sequences of period $2^n - 1$ with respect to 3- or 5-valued crosscorrelation ($\gcd(n,t) = 1$, $\sigma = \frac{n+1}{2}$, and $4\tau \equiv 1 \pmod{n}$).

by

$$C_{f,g^{(s)}}(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(\lambda x) + g(x^s)}$$

where n is odd, and $f(x)$ and $g(x)$ are the trace representations of the pair, excluding the (generalized) GMW, QR, and Hall's sextic residue sequences.

If $f(x) = g(x) = Tr(x)$, all known exponents s's of the 3- or 5-valued $C_{f,g^{(s)}}(\lambda)$ are (a) the Gold, Kasami, Welch, Niho exponents, and their respective inverses; (b) the exponents of Proposition 1; (c) the other exponents conjectured by Niho [67] and their inverses. Otherwise, all known exponents s's of the 3- or 5-valued $C_{f,g^{(s)}}(\lambda)$ for the corresponding $f(x)$ and $g(x)$ are (a) $s = 2^k + 1$ from (2.1), or $s = 1$ from (2.2) and (2.3); (b) s's from Fact 2 and Theorems $1 - 4$; (c) s's from Conjectures 1 and 2. With the classification, the relation of binary two-level autocorrelation

sequences of period $2^n - 1$ with respect to 3- or 5-valued crosscorrelation can be summarized by Figure 2.1, where a solid line is for exactly 3-valued crosscorrelation and a dotted line for at most 5-valued crosscorrelation. (For small n, it may be 3-valued in some cases.) In Figure 2.1, the crosscorrelation corresponding to the exponents s with '$*$' is proved or conjectured in [93].

From the observation of our experiments for $n = 13, 15, 17,$ and 19, it is interesting that the relation of Figure 2.1 completely describes all the 3- or 5-valued crosscorrelation of binary two-level autocorrelation sequences studied in this chapter unless both are m-sequences.

Chapter 3

New Binary Sequences with Optimal Autocorrelation Magnitude

In this chapter, new binary sequences with optimal autocorrelation magnitude are studied. First, the related works are reviewed after literatures survey and the motivation of this work is presented. Then, binary sequences with optimal autocorrelation are classified by their periods, and the known constructions are discussed. The interleaved structure of binary sequences is recalled, and some classes of binary sequences with optimal autocorrelation are analyzed by the interleaved structure. As main results of this chapter, we construct new binary sequences of period $N = 4(2^m - 1)$ for even $m \geq 4$ by exploiting a $4 \times (2^m - 1)$ interleaved structure, and mathematically derive the complete autocorrelation distribution [94]. In addition, it is shown that the sequences have large linear complexities from the implementation with linear feedback shift registers (LFSRs) and a simple logic [96].

3.1 Related Works and Motivation

Binary pseudorandom sequences with optimal autocorrelation play important roles in many areas of communication and cryptography. In code-division multiple access

(CDMA) communication systems, the sequences are needed to acquire the accurate timing information of received signals. In cryptography, on the other hand, the sequences are employed to generate key streams in stream cipher encryptions.

For binary sequences of period $N = 2^m - 1$, binary m-sequences are traditionally well-known sequences with ideal two-level autocorrelation of $\{N, -1\}$. Due to their good randomness properties and simple implementation [32] [33], m-sequences have been widely used for communication systems. Besides m-sequences, several other inequivalent classes of binary sequences with ideal two-level autocorrelation have been constructed and discovered [23] [37] [61] [70] [38].

For binary sequences of even period N, on the other hand, Lempel, Cohn, and Eastman [53] showed that 1) autocorrelation must have at least two distinct out-of-phase values and 2) a difference between any two autocorrelation values is divisible by 4. Therefore, the optimal autocorrelation is given by $\{N, 2, -2\}$ if $N \equiv 2 \pmod 4$, and $\{N, 0, -4\}$ or $\{N, 0, 4\}$ if $N \equiv 0 \pmod 4$. Several classes of binary sequences of even period with optimal autocorrelation are known. Initially, Lempel, Cohn, and Eastman [53] presented a class of the balanced binary sequences of period $N = p^m - 1$ for odd prime p. (It is known that this has been already described in [85] and called as the *Sidelnikov sequences.*) Then, No, Chung, Song, Yang, Lee, and Helleseth [68] gave another class of the binary sequences of period $N = p^m - 1$ for odd prime p using a polynomial $(z+1)^d + az^d + b$ over a finite field. From group division structure by the Chinese Remainder theorem, Ding, Helleseth, and Martinsen [26] also presented several families of the binary sequences of period $N = 2p$ for odd prime $p \equiv 5 \pmod 8$, which correspond to almost difference sets. Using known cyclic difference sets, Arasu, Ding, Helleseth, Kumar, and Martinsen [6] constructed four classes of almost difference sets which give inequivalent classes of the binary sequences of period $N \equiv 0 \pmod 4$. These sequences generally contain the binary sequences of period $N \equiv 0 \pmod 4$ constructed from the product method in [57]. Recently, Zhang, Lei, and Zhang [99] presented an almost difference set corresponding to the binary sequence of period $N \equiv 0 \pmod 4$ by adding two indices to the one class of almost difference sets in [6], where the corresponding cyclic difference set is from the Legendre sequences.

For a period $N \equiv 0 \pmod 4$, the autocorrelation $\{N, 0, -4\}$ or $\{N, 0, 4\}$ is optimal from the Lempel, Cohn, and Eastman's assertion in the sense that it has two out-of-phase values with the smallest magnitudes. If we allow three out-of-phase values with the smallest magnitudes, on the other hand, then optimal autocorrelation should be $\{N, 0, \pm 4\}$, where the autocorrelation is *optimal with respect to its magnitude*. In practical applications, it has the same meaning as conventional optimal autocorrelation. Consequently, the autocorrelation of $\{N, 0, \pm 4\}$ is also considered as optimal in this chapter.

In [34], Gong introduced the interleaved structure of sequences which is indeed a good method not only for understanding a sequence structure, but also for constructing new sequences of an interleaved form [34] [35]. In this chapter, it is shown that the binary sequence of period $4(2^m - 1)$ with optimal autocorrelation shown in [6] can be represented by a $(2^m - 1) \times 4$ interleaved structure. It is also shown that the binary product sequence [57] of period $4(2^m - 1)$ with optimal autocorrelation can be represented by a $4 \times (2^m - 1)$ interleaved structure. Inspired by these interpretations, a new construction of binary sequences of period $N = 4(2^m - 1)$ with autocorrelation $\{N, 0, \pm 4\}$ is discovered by the interleaved method. In details, a $4 \times (2^m - 1)$ interleaved structure defined by a perfect binary sequence of period 4 and a binary m-sequence of period $2^m - 1$ is used. In the interleaved structure, a sequence defined over $\mathbb{Z}_4$ is used as a shift sequence. The new sequences are almost balanced, i.e., a difference between the numbers of zeros and ones in a period is 2 [68], and optimal with respect to autocorrelation magnitude. The complete autocorrelation distribution and the exact linear complexity of the sequences are mathematically derived. From the simple implementation with a small number of shift registers and a connector, the sequences have a benefit of obtaining the large linear complexity.

3.2 Optimal Autocorrelation

This section recalls the classification of optimal autocorrelation in Chapter 1. (Periodic) autocorrelation of a binary sequence $\mathbf{a} = \{a_t\}$ of period N is defined by

(1.9). For a sequence $\mathbf{a}$ of period N, it is implied that $C_{\mathbf{a}}(\tau) = N$ occurs only at $\tau \equiv 0 \pmod{N}$. $C_{\mathbf{a}}(\tau)$ is called *optimal autocorrelation* [6] if it satisfies

1) $C_{\mathbf{a}}(\tau) \in \{N, -1\}$ if $N \equiv 3 \pmod 4$, or

2) $C_{\mathbf{a}}(\tau) \in \{N, 1, -3\}$ if $N \equiv 1 \pmod 4$, or

3) $C_{\mathbf{a}}(\tau) \in \{N, 2, -2\}$ if $N \equiv 2 \pmod 4$, or

4) $C_{\mathbf{a}}(\tau) \in \{N, 0, -4\}$ or $\{N, 0, 4\}$ if $N \equiv 0 \pmod 4$

for all τ. In particular, case 1) is called the *ideal two-level autocorrelation* and binary sequences of case 1) correspond to *cyclic difference sets* [8] [40]. Complete classes of all known inequivalent binary sequences of period $N = 2^m - 1$ with ideal two-level autocorrelation are summarized in [23] along with the corresponding cyclic difference sets. On the other hand, binary sequences of cases 2) – 4) correspond to *almost difference sets*. The almost difference sets are the generalization of two different types of almost difference sets independently defined by Ding [24] and Davis [20]. The generalization was done by Ding, Helleseth, and Martinsen in [26].

Several classes of binary sequences of cases 2) and 3) are described by the corresponding almost difference sets in [25] and [26], respectively. Four classes of binary sequences of case 4) and the corresponding almost difference sets are presented in [6], where the sequences generally contain the binary sequences constructed from the product method in [57]. Another almost difference set corresponding to a binary sequence of case 4) is presented in [99] where two indices are added to the one class of almost difference sets in [6]. Meanwhile, from a finite field approach, binary sequences of period $N = p^m - 1$ for odd prime p corresponding to cases 3) and 4) are also described in [53] and [68], respectively. For a survey of binary and quadriphase sequences with optimal autocorrelation, see [58].

In this chapter, if $C_{\mathbf{a}}(\tau) \in \{N, 0, \pm 4\}$ for $N \equiv 0 \pmod 4$, it is also considered as optimal in the sense that *its autocorrelation magnitude is identical to that of case 4).*

3.3 Interleaved Structures of Known Binary Sequences with Optimal Autocorrelation

In this section, we first review the interleaved structure of binary m-sequences of period $2^m - 1$ with the constant-on-cosets property. Then, two other classes of known binary sequences with optimal autocorrelation are investigated by the interleaved structures. For the interleaved structure and the constant-on-cosets property of binary sequences, recall the definitions in Section 1.5.

3.3.1 Binary m-sequences of Period $2^m - 1$

Let $m = 2k$ and $\mathbf{u}$ be a binary m-sequence of period $2^m - 1$ represented by $u_t = Tr_1^m(\alpha^t)$, $0 \le t \le 2^m - 2$ where α is a primitive element of $\mathbb{F}_{2^m}$. (As a binary m-sequence, we consider a sequence satisfying the *constant-on-cosets* property [33].) Then, $\mathbf{u}$ is represented by a $(2^k-1, 2^k+1)$ interleaved sequence [34], i.e., $\mathbf{u} = A(\mathbf{a}, \mathbf{e})$. In its array form $U = (\mathbf{u}_0, \cdots, \mathbf{u}_{2^k})$, the base sequence $\mathbf{a}$ is a binary m-sequence of period $2^k - 1$ represented by $a_i = Tr_1^k(\beta^i)$, $0 \le i \le 2^k - 2$ where $\beta = \alpha^{2^k+1}$ is a primitive element of $\mathbb{F}_{2^k}$. Also, the shift sequence $\mathbf{e}$ is given by

$$e_j = \begin{cases} \infty, & j = 0 \\ Tr_k^m(\alpha^j), & 1 \le j \le 2^k. \end{cases}$$

In other words, $\mathbf{u}_0$ is a zero sequence of length $2^k - 1$, and $\mathbf{u}_j$, $j \neq 0$ is a cyclic e_j shift of a binary m-sequence $\mathbf{a}$ of period $2^k - 1$.

Example 6 *For $m = 4$, a binary m-sequence of period 15 with the constant-on-cosets property is given by $\mathbf{u} = (0,0,0,1,0,0,1,1,0,1,0,1,1,1,1)$. Then, it is represented by a 3×5 interleaved sequence, i.e.,*

$$\mathbf{u} = A(\mathbf{a}, \mathbf{e}) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

where $\mathbf{a} = (0, 1, 1)$*, a binary m-sequence of period* 3*, and* $\mathbf{e} = (\infty, 0, 0, 2, 0)$.

3.3.2 The Almost Difference Set (ADS) Sequences of Period $N \equiv 0 \pmod 4$

In [6], Arasu, Ding, Helleseth, Kumar, and Martinsen presented binary sequences of period $N \equiv 0 \pmod 4$ with optimal autocorrelation. Let $\mathbf{a} = \{a_t\}$ be a binary sequence of period v with ideal two-level autocorrelation. A matrix $G = (g_{x,y}), 0 \leq x \leq 3, 0 \leq y \leq v-1$, is defined by

$$G = \begin{bmatrix} a_0 & a_1 & \cdots & a_{v-1} \\ \overline{a}_\eta & \overline{a}_{\eta+1} & \cdots & \overline{a}_{v-1+\eta} \\ \overline{a}_0 & \overline{a}_1 & \cdots & \overline{a}_{v-1} \\ \overline{a}_\eta & \overline{a}_{\eta+1} & \cdots & \overline{a}_{v-1+\eta} \end{bmatrix} \tag{3.1}$$

where η is any integer in $0 \leq \eta \leq v-1$ and the indices are computed modulo v. A binary sequence $\mathbf{s} = \{s_t\}$ of period $N = 4v$ is defined by

$$s_t = g_{x,y} \quad \text{where } x \equiv t \pmod 4 \text{ and } y \equiv t \pmod v. \tag{3.2}$$

In other words, $\mathbf{s}$ is generated by reading the matrix G in a diagonal way. Then $\mathbf{s}$ has optimal autocorrelation of $C_{\mathbf{s}}(\tau) \in \{N, 0, -4\}$ for every $\eta, 0 \leq \eta \leq v-1$ by providing the corresponding almost difference set [6]. Throughout this chapter, $\mathbf{s}$ is called the *ADS sequence*.

In the ADS sequence $\mathbf{s}$ defined by (3.2), let $v = 2^m - 1$ and $\mathbf{a}$ be a binary sequence of period $2^m - 1$ with ideal two-level autocorrelation and the constant-on-cosets property. If $\mathbf{s}$ is represented by a $(2^m - 1) \times 4$ interleaved structure, then

$$\mathbf{s} = \begin{bmatrix} a_0 & \overline{a}_{1+\eta} & \overline{a}_2 & \overline{a}_{3+\eta} \\ a_4 & \overline{a}_{5+\eta} & \overline{a}_6 & \overline{a}_{7+\eta} \\ \vdots & \vdots & \vdots & \vdots \\ a_{4(2^m-1)-4} & \overline{a}_{4(2^m-1)-3+\eta} & \overline{a}_{4(2^m-1)-2} & \overline{a}_{4(2^m-1)-1+\eta} \end{bmatrix} = (\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3)$$

where $\mathbf{s}_j = \{s_{i,j}\}, 0 \le i \le 2^m - 2, 0 \le j \le 3$, and η is an integer, $0 \le \eta \le 2^m - 2$. From the constant-on-cosets property of $\mathbf{a}$ and $4^{-1} \equiv 2^{m-2} \pmod{2^m - 1}$, we have

1) $s_{i,0} = a_{4i} = a_i$,

2) $s_{i,1} = \overline{a}_{4i+1+\eta} = \overline{a}_{4(i+4^{-1}(1+\eta))} = \overline{a}_{4(i+2^{m-2}(1+\eta))} = \overline{a}_{i+2^{m-2}(1+\eta)}$,

3) $s_{i,2} = \overline{a}_{4i+2} = \overline{a}_{4(i+4^{-1}\cdot 2)} = \overline{a}_{4(i+2^{m-1})} = \overline{a}_{i+2^{m-1}}$,

4) $s_{i,3} = \overline{a}_{4i+3+\eta} = \overline{a}_{4(i+4^{-1}(3+\eta))} = \overline{a}_{4(i+2^{m-2}(3+\eta))} = \overline{a}_{i+2^{m-2}(3+\eta)}$

where the indices are computed modulo $2^m - 1$. From this, $\mathbf{s}$ has the following interleaved structure. For the definition of perfect sequences, see Section 1.5.

Property 1 *Let* $\mathbf{s}$ *be the ADS sequence of period* $4(2^m - 1)$. *Then,* $\mathbf{s}$ *has a* $(2^m - 1) \times 4$ *interleaved structure, i.e.,* $\mathbf{s} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ *where*

1) $\mathbf{a}$ *is a binary sequence of period* $2^m - 1$ *with ideal two-level autocorrelation and the constant-on-cosets property,*

2) $\mathbf{e} = (e_0, e_1, e_2, e_3) = (0, 2^{m-2}(\eta + 1), 2^{m-1}, 2^{m-2}(\eta + 3))$ *is a shift sequence,*

3) $\mathbf{b} = (0, 1, 1, 1)$ *is a perfect binary sequence*

where η *is an integer,* $0 \le \eta \le 2^m - 2$.

From the interleaved structure, $\mathbf{s}$ is cyclically distinct for each $\eta, 0 \le \eta \le 2^m - 2$. If $\eta = 0$, in particular, it is pointed out in [6] that $\mathbf{s}$ is equivalent to the product sequence of $\mathbf{a}$ and $\mathbf{b}$, i.e., $\mathbf{s} = \mathbf{a} + \mathbf{b}$. Thus, the product sequence of period $4(2^m - 1)$ with optimal autocorrelation is a special case of the ADS sequence.

Example 7 *For* $m = 3$, *let* $\mathbf{a}$ *be a binary m-sequence of period 7, i.e.,* $\mathbf{a} = (1, 0, 0, 1, 0, 1, 1)$. *If* $\eta = 1$, *then* $\mathbf{e} = (0, 4, 4, 1)$. *Then, the corresponding ADS*

sequence $\mathbf{s}$ *is given by*

$$\mathbf{s} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = (1,1,1,1,0,0,0,1,0,0,0,0,1,0,0,1,0,1,1,0,1,1,1,0,1,0,0,0).$$

3.3.3 Product Sequences

Let $\mathbf{a}$ and $\mathbf{b}$ be binary sequences of periods N_1 and N_2, respectively, where $\gcd(N_1, N_2) = 1$. Then the *product sequence* [57] $\mathbf{p} = \mathbf{a} + \mathbf{b} = (p_0, p_1, \cdots, p_{N-1})$ of period $N = N_1 N_2$ is defined by the component-wise addition of $p_t = a_t + b_t,\ 0 \le t \le N-1$ where the addition is computed modulo 2. Autocorrelation of the product sequence is given by

$$\begin{aligned} C_{\mathbf{p}}(\tau) &= \sum_{t=0}^{N-1} (-1)^{p_{t+\tau}+p_t} = \left[\sum_{t_1=0}^{N_1-1} (-1)^{a_{t_1+\tau}+a_{t_1}}\right] \cdot \left[\sum_{t_2=0}^{N_2-1} (-1)^{b_{t_2+\tau}+b_{t_2}}\right] \\ &= C_{\mathbf{a}}(\tau) \cdot C_{\mathbf{b}}(\tau), \quad 0 \le \tau \le N-1 \end{aligned} \tag{3.3}$$

where the indices of a sequence are computed modulo its own period [57].

In Section 3.3.2, the product sequence of period $4(2^m - 1)$ with optimal autocorrelation is represented by a $(2^m - 1) \times 4$ interleaved structure as a special case of the ADS sequence ($\eta = 0$). Here, it is also represented by a $4 \times (2^m - 1)$ interleaved structure with different base and shift sequences.

Let $\mathbf{p} = \mathbf{a} + \mathbf{b}$ be a binary product sequence of period $4(2^m - 1)$ where $\mathbf{a} = (0,1,1,1)$ is a perfect binary sequence of period 4 and $\mathbf{b}$ a binary sequence of period $2^m - 1$ with ideal two-level autocorrelation. Then $\mathbf{p}$ has optimal autocorrelation from (3.3). Note that $\gcd(4, 2^m - 1) = 1$ and $2^m - 1 \equiv 3 \pmod 4$ for any integer $m > 1$. First, a $4 \times (2^m - 1)$ interleaved structure $M = (\mathbf{m}_0, \mathbf{m}_1, \cdots, \mathbf{m}_{2^m-2})$

associated with $\mathbf{a}$ of length $4(2^m-1)$ is considered. Then, $\mathbf{m}_j$ is given by

$$\mathbf{m}_j = \{a_{(2^m-1)i+j}\}, \quad 0 \le i \le 3,\ 0 \le j \le 2^m-2$$

where the index is computed modulo 4. From $3^{-1} \equiv 3 \pmod 4$ and $a_{3i} = a_i$,

$$a_{(2^m-1)i+j} = a_{3i+j} = a_{3(i+3^{-1}j)} = a_{3(i+3j)} = a_{i+3j}. \tag{3.4}$$

Thus, the product sequence $\mathbf{p}$ has the following interleaved structure.

Property 2 *Let $\mathbf{p} = \mathbf{a} + \mathbf{b}$ be a binary product sequence of period $4(2^m-1)$ with optimal autocorrelation. Then, it has a $4 \times (2^m-1)$ interleaved structure, i.e., $\mathbf{p} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ where*

1) *$\mathbf{a} = (0,1,1,1)$ is a perfect binary sequence,*

2) *$\mathbf{e} = (e_0, e_1, \cdots, e_{2^m-2})$ is a shift sequence defined over $\mathbb{Z}_4$ where the j-th element is given by*

$$e_j \equiv 3j \pmod 4, \quad 0 \le j \le 2^m-2, \tag{3.5}$$

3) *$\mathbf{b}$ is a binary sequence of period 2^m-1 with ideal two-level autocorrelation.*

In the interleaved structure of $\mathbf{p} = (\mathbf{p}_0, \cdots, \mathbf{p}_{2^m-2})$, its j-th column is given by $\mathbf{p}_j = L^{e_j}(\mathbf{a})$ if $b_j = 0$, or $\mathbf{p}_j = L^{e_j}(\overline{\mathbf{a}})$ otherwise.

Example 8 *For $m = 4$, a product sequence of period $N = 4 \times 15 = 60$ with optimal autocorrelation is given by $\mathbf{p} = \mathbf{a} + \mathbf{b}$ where $\mathbf{a} = (0,1,1,1)$ is a perfect binary sequence of period 4, and $\mathbf{b} = (0,0,0,1,0,0,1,1,0,1,0,1,1,1,1)$ is a binary m-sequence of period 15. From (3.5), $\mathbf{e}$ is defined by*

$$\mathbf{e} = (0,3,2,1,0,3,2,1,0,3,2,1,0,3,2).$$

Then, $\mathbf{p}$ *is given by*

$$\mathbf{p} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Now, we focus on the shift sequence over $\mathbb{Z}_4$ of a product sequence of period $N = 4(2^m - 1), m = 2k$ for an integer $k > 1$ with optimal autocorrelation. For $m = 2k$, the shift sequence shown in (3.5) can be also represented by a $(2^k - 1) \times (2^k + 1)$ interleaved structure, i.e.,

$$\mathbf{e} = (e_0, e_1, \cdots, e_{2^m-2}) = \begin{bmatrix} e_0 & e_1 & \cdots & e_{2^k} \\ e_{2^k+1} & e_{2^k+2} & \cdots & e_{2\cdot 2^k+1} \\ \vdots & \vdots & \cdots & \vdots \\ e_{(2^k-2)(2^k+1)} & e_{(2^k-2)(2^k+1)+1} & \cdots & e_{2^m-2} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 3 & 2 & 1 & 0 & 3 & 2 & 1 & \cdots & 0 \\ 3 & 2 & 1 & 0 & 3 & 2 & 1 & 0 & \cdots & 3 \\ 2 & 1 & 0 & 3 & 2 & 1 & 0 & 3 & \cdots & 2 \\ \vdots & & & & \vdots & & & & \cdots & \vdots \\ 2 & 1 & 0 & 3 & 2 & 1 & 0 & 3 & \cdots & 2 \end{bmatrix}.$$

Interestingly, the $(2^k - 1) \times (2^k + 1)$ interleaved structure of $\mathbf{e}$ is given by

$$\begin{aligned} &\mathbf{e} = (\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_{2^k}) \\ &\text{where } \mathbf{e}_j = \{e_{i,j}\},\ e_{i,j} = 3(i + j) \pmod 4,\ 0 \le i \le 2^k - 2,\ 0 \le j \le 2^k. \end{aligned} \tag{3.6}$$

Note that $e_{i,j} = e_{i(2^k+1)+j}$, and both expressions are used throughout this chapter.

In next section, we will present a new construction of binary sequences of period $N = 4(2^m - 1)$ for even $m \ge 4$ with optimal four-valued autocorrelation by modifying the shift sequence of the interleaved structure of a product sequence of period N.

3.4 New Binary Sequences with Optimal Four-Valued Autocorrelation

This section presents a new construction of binary sequences of period $N = 4(2^m-1)$ for even $m \geq 4$ with optimal four-valued autocorrelation, i.e., $C_{\mathbf{u}}(\tau) \in \{N, 0, \pm 4\}$ for any τ.

3.4.1 Construction

Construction 1 *Let $k > 1$ be a positive integer. A new binary sequence $\mathbf{u}$ of period $N = 4(2^m - 1), m = 2k$ is defined by a $4 \times (2^m - 1)$ interleaved structure of $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ where*

1) $\mathbf{a} = (0, 1, 1, 1)$ *is a perfect binary sequence of period* 4,

2) $\mathbf{b}$ *is a binary m-sequence of period $2^m - 1$ with the constant-on-cosets property,*

3) $\mathbf{e}$ *is a sequence defined over $\mathbb{Z}_4$ of length $2^m - 1$, and represented by a $(2^k - 1) \times (2^k + 1)$ interleaved structure, i.e.,*

$$\mathbf{e} = (\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_{2^k})$$
$$\text{where } \mathbf{e}_j = \{e_{i,j}\},\ e_{i,j} = \begin{cases} 3i + \delta \pmod 4, & \text{if } j = 0 \\ 3(i+j) \pmod 4, & \text{if } 1 \leq j \leq 2^k \end{cases} \tag{3.7}$$

where $0 \leq i \leq 2^k - 2$ and $\delta = 1$ or -1.

In fact, we obtain the new sequence $\mathbf{u}$ by modifying the shift sequence of (3.6) in the interleaved structure of a product sequence.

Remark 1 With $\delta = \pm 1$ and cyclically distinct binary m-sequences of $\mathbf{b}$, Construction 1 gives $\frac{2\phi(2^m-1)}{m}$ cyclically distinct binary sequences $\mathbf{u}$, where $\phi(\cdot)$ is the Euler-totient function.

In the next theorem, we prove that the new sequence from Construction 1 is almost balanced. For the almost balancedness, recall Section 1.5.

Theorem 5 *Let* $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ *be the binary sequence from Construction 1. Then* $\mathbf{u}$ *is almost balanced.*

Proof. In the $4 \times (2^m - 1)$ interleaved structure of $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$, note that the l-th column is a cyclic shift of either $\mathbf{a} = (0, 1, 1, 1)$ or $\overline{\mathbf{a}}$, which is determined by $\mathbf{b} = \{b_l\}$. In other words, each column of $\mathbf{u}$ has 3 ones if $b_l = 0$, or 1 ones if $b_l = 1$. Since $\mathbf{b}$ is a binary m-sequence of period $2^m - 1$ with the balance property [33], the number of ones in $\mathbf{u}$ is given by

$$3(2^{m-1} - 1) + 2^{m-1} = 2^{m+1} - 3 = \frac{N}{2} - 1.$$

Hence, a difference between the numbers of zeros and ones in a period is 2, i.e., $\mathbf{u}$ is almost balanced. □

3.4.2 Autocorrelation

To compute the autocorrelation function of $\mathbf{u}$, Proposition 2 is first considered.

Proposition 2 *Let* $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ *be the binary sequence from Construction 1. Let* $\tau = r(2^m - 1) + s$, $0 \le r \le 3$, $0 \le s \le 2^m - 2$. *Then the autocorrelation function* $C_{\mathbf{u}}(\tau)$ *is given by*

$$C_{\mathbf{u}}(\tau) = \sum_{t=0}^{N-1} (-1)^{u_t + u_{t+\tau}} = \sum_{l=0}^{2^m - 2} (-1)^{d_l} C_{\mathbf{a}}(t_l) \tag{3.8}$$

where $d_l \equiv b_l - b_{l+s} \pmod 2$, $t_l \equiv e_{l+s} - e_l + r \pmod 4$, *and* $C_{\mathbf{a}}(t_l)$ *is the autocorrelation of a base sequence* $\mathbf{a}$. *In* b_{l+s}, *the index is computed modulo* $2^m - 1$. *On the other hand, as in equation (12) of [35],*

$$e_{l+s} \equiv e_{l+s-(2^m-1)} + 1 \pmod 4 \text{ if } l + s \ge 2^m - 1. \tag{3.9}$$

Proof. From $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$, it is immediate from Lemma 2 of [35]. □

In Construction 1, note that the indicator sequence $\mathbf{b}$ and the shift sequence $\mathbf{e}$ have $(2^k-1)\times(2^k+1)$ interleaved structures. Then, $\mathbf{d}=\{d_l \mid 0\le l\le 2^m-2\}$ in Proposition 2 is represented by a difference array $D=B-B_s$ where $(2^k-1)\times(2^k+1)$ arrays B and B_s represent $\mathbf{b}=\{b_l\}$ and $L^s(\mathbf{b})=\{b_{l+s}\}$, respectively. From the shift-and-add property [33] of the binary m-sequence $\mathbf{b}$, it is clear that $D=(\mathbf{d}_0,\cdots,\mathbf{d}_{2^k})$ also represents a binary m-sequence of period 2^m-1 for $s\neq 0$. Also, the j-th column $\mathbf{d}_j, 0\le j\le 2^k$ is either a cyclic shift of a binary m-sequence of period 2^k-1 or a zero sequence from the interleaved structure of binary m-sequences. In the following lemma, the structure of array D is further studied. From now on, the following notation will be used in all lemmas and theorems in this section.

$$\begin{aligned} \tau &= r(2^m-1)+s, \quad 0\le r\le 3,\ 0\le s\le 2^m-2 \\ \text{where } s &= x(2^k+1)+y, \quad 0\le x\le 2^k-2,\ 0\le y\le 2^k \end{aligned} \tag{3.10}$$

Lemma 2 *Let B and B_s be the $(2^k-1)\times(2^k+1)$ arrays of $\mathbf{b}$ and $L^s(\mathbf{b})$, respectively, where $s\neq 0$. In a difference array $D=B-B_s=(\mathbf{d}_0,\cdots,\mathbf{d}_{2^k})$, the j-th column $\mathbf{d}_j$ has the following properties.*

1) *If $y=0$, then a zero column $\mathbf{d}_j$ exists only at $j=0$.*

2) *If $y\neq 0$, then a zero column $\mathbf{d}_j$ exists at exactly one j for $0\le j\le 2^k$ with $j\neq 0$ and $j\neq -y$, where '$-y$' is computed modulo 2^k+1.*

Proof. In $B=(\mathbf{b}_0,\cdots,\mathbf{b}_{2^k})$, $\mathbf{b}_0$ is a zero column and $\mathbf{b}_j, j\neq 0$ is a cyclic shift of a binary m-sequence. In another array $B_s=(\mathbf{s}_0,\cdots,\mathbf{s}_{2^k})$ of $L^s(\mathbf{b})$, on the other hand, $\mathbf{s}_{-y}$ is a zero column and $\mathbf{s}_j, j\neq -y$ is a cyclic shift of a binary m-sequence. In the difference $D=B-B_s$, therefore, $\mathbf{d}_0$ is still a zero column if $y=0$. If $y\neq 0$, on the other hand, neither $\mathbf{d}_0=\mathbf{b}_0-\mathbf{s}_0$ nor $\mathbf{d}_{-y}=\mathbf{b}_{-y}-\mathbf{s}_{-y}$ can be a zero column because both $\mathbf{s}_0$ and $\mathbf{b}_{-y}$ are nonzero columns. Instead, a zero column $\mathbf{d}_j$ exists at another column index j with $j\neq 0$ and $j\neq -y$ because there should be exactly one zero column in the difference array D which represents a binary m-sequence. This completes the proof of Lemma 2. □

In Proposition 2, $\mathbf{t} = \{t_l \mid 0 \le l \le 2^m - 2\}$ can be also represented by a $(2^k - 1) \times (2^k + 1)$ interleaved structure T. To obtain T, the following two lemmas are needed.

Lemma 3 *In the array structure of* $\mathbf{e}$ *in (3.7), an s-shift* $L^s(\mathbf{e})$ *of* $\mathbf{e}$ *is given by*

$$L^s(\mathbf{e}) = (\mathbf{e}_y - x, \mathbf{e}_{1+y} - x, \cdots, \mathbf{e}_{2^k+y} - x) \tag{3.11}$$

where $\mathbf{e}_j$ *is defined in Construction 1 and extended to* $\mathbf{e}_j = \mathbf{e}_{j-(2^k+1)} + 3 \pmod 4$ *for* $j \ge 2^k + 1$.

Proof. If $L^s(\mathbf{e})$ is arranged in an array form, then

$$L^s(\mathbf{e}) = \begin{bmatrix} e_{x,y} & \cdots & e_{x,2^k} & e_{x+1,0} & \cdots & e_{x+1,y-1} \\ e_{x+1,y} & \cdots & e_{x+1,2^k} & e_{x+2,0} & \cdots & e_{x+2,y-1} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ e_{2^k-3,y} & \cdots & e_{2^k-3,2^k} & e_{2^k-2,0} & \cdots & e_{2^k-2,y-1} \\ e_{2^k-2,y} & \cdots & e_{2^k-2,2^k} & e_{0,0}+1 & \cdots & e_{0,y-1}+1 \\ e_{0,y}+1 & \cdots & e_{0,2^k}+1 & e_{1,0}+1 & \cdots & e_{1,y-1}+1 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ e_{x-1,y}+1 & \cdots & e_{x-1,2^k}+1 & e_{x,0}+1 & \cdots & e_{x,y-1}+1 \end{bmatrix} \tag{3.12}$$

where the '+1' addition is from (3.9) and computed modulo 4. From the definition of $e_{i,j}$ in (3.7),

$$e_{i+1,j} - e_{i,j} \equiv 3 \pmod 4, \quad 0 \le i \le 2^k - 3. \tag{3.13}$$

Also,

$$e_{0,j} + 1 - e_{2^k-2,j} = 3j + 1 + \Delta - 3(2^k - 2 + j) - \Delta \equiv 3 \pmod 4, \quad 0 \le j \le 2^k \tag{3.14}$$

where $\Delta = \delta = \pm 1$ if $j = 0$, or $\Delta = 0$ otherwise. From (3.13) and (3.14), we see that in (3.12), every element in a difference vector between the $(i+1)$-th row and the i-th row for $0 \le i \le 2^k - 3$ is 3 (mod 4). With $e_{i,j} = e_{0,j} + 3i \pmod 4$ from

(3.7), therefore, (3.12) can be written as

$$L^s(\mathbf{e}) = (\mathbf{e}_y + 3x, \cdots, \mathbf{e}_{2^k} + 3x, \mathbf{e}_0 + 3x + 3, \cdots, \mathbf{e}_{y-1} + 3x + 3).$$

For $j \geq 2^k + 1$, if $\mathbf{e}_j$ is defined by $\mathbf{e}_j = \mathbf{e}_{j-(2^k+1)} + 3 \pmod 4$, then $L^s(\mathbf{e})$ is given by (3.11) from $3x \equiv -x \pmod 4$. □

Lemma 4 *With the notation of Proposition 2 and Lemma 3, let* $l = i(2^k + 1) + j$ *where* $0 \leq l \leq 2^m - 2$, $0 \leq i \leq 2^k - 2$, *and* $0 \leq j \leq 2^k$. *Then,* $t_l \equiv e_{l+s} - e_l + r \pmod 4$ *is given by*

$$t_l = \begin{cases} 3y - x + r, & \text{if } j \neq 0 \text{ and } j + y \not\equiv 0 \pmod{2^k+1}, \\ 3y - x + r + \delta, & \text{if } j \neq 0 \text{ and } j + y \equiv 0 \pmod{2^k+1}, \\ 3y - x + r - \delta, & \text{if } j = 0 \text{ and } j + y \not\equiv 0 \pmod{2^k+1}, \\ -x + r, & \text{if } j = 0 \text{ and } j + y \equiv 0 \pmod{2^k+1}. \end{cases} \tag{3.15}$$

In a $(2^k - 1) \times (2^k + 1)$ *array* $T = (\mathbf{t}_0, \cdots, \mathbf{t}_{2^k})$ *of* $\mathbf{t} = \{t_l\}$, *therefore, each column has constant elements given as follows.*

1) *If* $y = 0$, *then*

$$\mathbf{t}_j = \{t_{i,j} = -x + r \mid 0 \leq i \leq 2^k - 2\} \text{ for all } 0 \leq j \leq 2^k. \tag{3.16}$$

2) *If* $y \neq 0$, *then*

$$\begin{aligned} \mathbf{t}_0 &= \{t_{i,0} = 3y - x + r - \delta \mid 0 \leq i \leq 2^k - 2\}, \\ \mathbf{t}_{-y} &= \{t_{i,-y} = 3y - x + r + \delta \mid 0 \leq i \leq 2^k - 2\}, \\ \mathbf{t}_j &= \{t_{i,j} = 3y - x + r \mid 0 \leq i \leq 2^k - 2\} \text{ for } j \neq 0 \text{ and } j + y \not\equiv 0 \pmod{2^k+1} \end{aligned} \tag{3.17}$$

where '$-y$*' is computed modulo* $2^k + 1$.

Proof. The $(2^k - 1) \times (2^k + 1)$ array structure of $\mathbf{t}$ is given by

$$T = L^s(\mathbf{e}) - \mathbf{e} + r$$

where r is added to all elements of the array. From Lemma 3, the j-th column vector of T is given by

$$\mathbf{t}_j = \mathbf{e}_{j+y} - \mathbf{e}_j - x + r = \{e_{i,j+y} - e_{i,j} - x + r \mid 0 \le i \le 2^k - 2\}, \quad 0 \le j \le 2^k.$$

From $l = i(2^k + 1) + j$, it is clear that t_l is the i-th element of $\mathbf{t}_j$, i.e.,

$$t_l = e_{l+s} - e_l + r = e_{i,j+y} - e_{i,j} - x + r. \tag{3.18}$$

Together with (3.18) and (3.7), (3.15) follows immediately. The assertions of (3.16) and (3.17) are from (3.15). □

Now, we are ready to compute $C_{\mathbf{u}}(\tau)$.

Theorem 6 *Let* $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ *be the binary sequence of period* $N = 4(2^m - 1), m = 2k$ *for an integer* $k > 1$ *from Construction 1. Then it has the four-valued optimal autocorrelation, i.e.,* $C_{\mathbf{u}}(\tau) \in \{N, 0, \pm 4\}$ *for any* τ. *Precisely, its complete autocorrelation is given by*

$$C_{\mathbf{u}}(\tau) = \begin{cases} 4(2^m - 1), & \text{if } \tau = 0 \\ 0, & \text{if } (\tau \neq 0 \text{ and } s = 0) \text{ or } (s, y, v) = (\sigma, 0, \nu) \text{ or } (y, v) = (\psi, 2) \\ -4, & \text{if } (s, y, v) = (\sigma, 0, 0) \text{ or } (y, v) = (\psi, 1) \text{ or } (y, v) = (\psi, 3) \\ +4, & \text{if } (y, v) = (\psi, 0) \end{cases}$$

where $s \equiv \tau \pmod{2^m - 1}$, $y \equiv \tau \pmod{2^k + 1}$, *and* $v \equiv \tau \pmod 4$. *Also,* $\sigma \in \mathbb{Z}^+_{2^m-1}$, $\psi \in \mathbb{Z}^+_{2^k+1}$, *and* $\nu \in \mathbb{Z}^+_4$.

Proof. To compute $C_{\mathbf{u}}(\tau)$, (3.8) in Proposition 2 is used. Since $C_{\mathbf{a}}(t_l) = 0$ at any nonzero t_l, $C_{\mathbf{u}}(\tau)$ is determined by cases of $t_l = 0$ in (3.8). For the following three cases, recall (3.10) and the arrays D and T of $\mathbf{d} = \{d_l\}$ and $\mathbf{t} = \{t_l\}$ in Lemmas 2 and 4, respectively.

Case 1. $s = 0$: If $r = 0$ in this case, then $\tau = 0$ and $C_{\mathbf{u}}(\tau) = 4(2^m - 1)$, a trivial in-phase autocorrelation. If $r \neq 0$, on the other hand, $t_l = e_{l+s} - e_l + r = r \neq 0$

and then $C_{\mathbf{u}}(\tau) = 0$ in (3.8) because $C_{\mathbf{a}}(t_l) = 0$ for all nonzero t_l's. Thus, we have

$$C_{\mathbf{u}}(\tau) = \begin{cases} 4(2^m - 1), & \text{if } \tau = 0, \\ 0, & \text{if } \tau \neq 0 \text{ and } s = 0. \end{cases} \tag{3.19}$$

Case 2. $s \neq 0$ and $y = 0$: From (3.16), T is a constant array where each element is $-x + r$, and D is balanced from the difference property of m-sequences, i.e., $\sum_{l=0}^{2^m-2}(-1)^{d_l} = -1$. If $r = x$, then $t_l = 0$ for all $0 \leq l \leq 2^m - 2$. Hence, $C_{\mathbf{u}}(\tau) = -1 \cdot C_{\mathbf{a}}(0) = -4$ from (3.8). If $r \neq x$, on the other hand, $t_l \neq 0$ and $C_{\mathbf{a}}(t_l) = 0$ for all $0 \leq l \leq 2^m - 2$, and hence $C_{\mathbf{u}}(\tau) = 0$. From (3.10), note that if $y = 0$, then $v \equiv \tau \mod 4 \equiv -r + x$. Therefore,

$$C_{\mathbf{u}}(\tau) = \begin{cases} -4, & \text{if } (s, y, v) = (\sigma, 0, 0) \\ 0, & \text{if } (s, y, v) = (\sigma, 0, \nu). \end{cases} \tag{3.20}$$

where σ and ν are elements in $\mathbb{Z}^+_{2^m-1}$ and $\mathbb{Z}^+_4$, respectively.

Case 3. $s \neq 0$ and $y \neq 0$: Let $l = i(2^k + 1) + j$. From Lemma 4, there are three distinct t_l's in T, i.e., $t_l = h, h \pm 1$ where $h = 3y - x + r$. (Note that $\delta = \pm 1$ in (3.7).) $t_l = h$ corresponds to $\mathbf{t}_j$'s of $0 \leq j \leq 2^k$ with $j \neq 0$ and $j \neq -y$. For such j's, $\mathbf{t}_j$ is a constant column of h and there exists one j such that $\mathbf{d}_j$ is a zero column from Lemma 2. On the other hand, $t_l = h \pm 1$ corresponds to $\mathbf{t}_0$ and $\mathbf{t}_{-y}$, respectively, and both $\mathbf{d}_0$ and $\mathbf{d}_{-y}$ are nonzero m-sequences in D.

1) $h = 0$: In this case, all $\mathbf{t}_j$'s of $0 \leq j \leq 2^k$ with $j \neq 0$ and $j \neq -y$ are zero columns. On the other hand, $\mathbf{t}_0$ and $\mathbf{t}_{-y}$ are nonzero. Let n_0 and n_1 be the numbers of zeros and ones in $\mathbf{d}_j$'s for $0 \leq j \leq 2^k$ with $j \neq 0$ and $j \neq -y$. Then,

$$n_0 = (2^k - 2)(2^{k-1} - 1) + 2^k - 1 = 2^{2k-1} - 2^k + 1$$
$$n_1 = (2^k - 2)2^{k-1} = 2^{2k-1} - 2^k = n_0 - 1.$$

From (3.8), $C_{\mathbf{u}}(\tau) = (n_0 \cdot 1 + n_1 \cdot (-1)) \cdot C_{\mathbf{a}}(0) = 1 \cdot C_{\mathbf{a}}(0) = 4$.

2) $h = \pm 1$: In this case, either $\mathbf{t}_0$ or $\mathbf{t}_{-y}$ is a zero column for given h and δ.

On the other hand, all the other columns are nonzero. If n_0 and n_1 are the numbers of zeros and ones in $\mathbf{d}_0$ or $\mathbf{d}_{-y}$, then

$$n_0 = 2^{k-1} - 1, \quad n_1 = 2^{k-1} = n_0 + 1$$

Thus, $C_{\mathbf{u}}(\tau) = (n_0 \cdot 1 + n_1 \cdot (-1)) \cdot C_{\mathbf{a}}(0) = (-1) \cdot C_{\mathbf{a}}(0) = -4$.

3) $h = 2$: In this case, no elements are zero in T. Thus, $C_{\mathbf{u}}(\tau) = 0$ from $C_{\mathbf{a}}(t_l) = 0$.

From (3.10), note that $v \equiv \tau \mod 4 \equiv 3r + x + y \equiv 3 \cdot (3y - x + r) \equiv -h$. Combining 1), 2) and 3),

$$C_{\mathbf{u}}(\tau) = \begin{cases} +4, & \text{if } (y, v) = (\psi, 0) \\ -4, & \text{if } (y, v) = (\psi, 1) \text{ or } (y, v) = (\psi, 3) \\ 0, & \text{if } (y, v) = (\psi, 2) \end{cases} \tag{3.21}$$

where ψ is an element in $\mathbb{Z}^+_{2^k+1}$. In (3.21), note that $y \neq 0$ implies $s \neq 0$.

If we combine (3.19), (3.20), and (3.21), then the proof is completed. □

Remark 2 In Remark 1, there are many cyclically distinct sequences of $\mathbf{u}$ according to δ and $\mathbf{b}$. In Theorem 6, however, the distinction disappears regarding their autocorrelation, and consequently all the sequences from Construction 1 have the identical autocorrelation distribution of Theorem 6 regardless of δ and $\mathbf{b}$.

Theorem 7 *With the notation in Theorem 6, the complete distribution of* $C_{\mathbf{u}}(\tau)$ *is given by*

$$C_{\mathbf{u}}(\tau) = \begin{cases} 4(2^m - 1), & 1 \textit{ time} \\ 0, & 2^{2k} + 2^{k+1} - 3 \textit{ times} \\ -4, & 2^{2k+1} - 2^k - 2 \textit{ times} \\ +4, & 2^{2k} - 2^k \textit{ times.} \end{cases}$$

Proof. From Theorem 6, the trivial in-phase autocorrelation occurs only once. Hence, the other exclusive cases of Theorem 6 are counted.

Case 1. $C_{\mathbf{u}}(\tau) = -4$:

1) $(s, y, v) = (\sigma, 0, 0)$: In this case, the possible τ's are $4(2^k+1), 8(2^k+1) \cdots, 4(2^k-2)(2^k+1)$. Thus its number of occurrences is $w_0 = 2^k - 2$.

2) $(y, v) = (\psi, 1)$ or $(\psi, 3)$: Note that $\gcd(2^k + 1, 4) = 1$. By the Chinese Remainder theorem, there exists a unique solution of τ for $(y, v) = (\psi, 1)$ with given $\psi \in \mathbb{Z}^+_{2^k+1}$. For all $\psi \in \mathbb{Z}^+_{2^k+1}$, the number of distinct solutions of τ in $\mathbb{Z}^+_{4(2^k+1)}$ for $(y, v) = (\psi, 1)$ is 2^k. From the isomorphism $\mathbb{Z}_{4(2^m-1)} \cong \mathbb{Z}_{4(2^k+1)} \times \mathbb{Z}_{2^k-1}$, the number of τ's in $\mathbb{Z}_{4(2^m-1)}$ for $(y, v) = (\psi, 1)$ is $2^k(2^k - 1)$. Considering the exclusive cases of $(y, v) = (\psi, 1)$ and $(\psi, 3)$, the number of such τ's is $w_1 = 2^{k+1}(2^k - 1)$.

Combining 1) and 2), the number of occurrences of $C_{\mathbf{u}}(\tau) = -4$ is $\lambda_0 = w_0 + w_1 = 2^{2k+1} - 2^k - 2$.

Case 2. $C_{\mathbf{u}}(\tau) = +4$: This corresponds to $(y, v) = (\psi, 0)$. By the similar approach to Case 1-2), the number of such τ's is equal to $\lambda_1 = w_1/2 = 2^k(2^k - 1)$.

Case 3. $C_{\mathbf{u}}(\tau) = 0$: The number of such τ's is $\lambda_2 = 4(2^m - 1) - (1 + \lambda_0 + \lambda_1) = 2^{2k} + 2^{k+1} - 3$.

From Cases $1 - 3$, the proof is completed. □

Example 9 *For $m = 2k = 4$, consider a new sequence* $\mathbf{u}$ *in Construction 1 with $\delta = 1$. In its interleaved structure* $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$, *a base sequence is* $\mathbf{a} = (0, 1, 1, 1)$ *and an indicator sequence is* $\mathbf{b} = (0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1)$, *a binary m-sequence of period 15. A shift sequence* $\mathbf{e}$ *is defined by (3.7), and* $\mathbf{b}$ *and* $\mathbf{e}$ *are represented by 3×5 arrays, respectively, i.e.,*

$$\mathbf{b} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} 1 & 3 & 2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 3 \\ 3 & 1 & 0 & 3 & 2 \end{bmatrix}.$$

Table 3.1: Autocorrelation values of $C_{\mathbf{u}}(\tau)$ in Example 3

τ	$C_{\mathbf{u}}(\tau)$														
$0-14$	60	-4	0	-4	4	0	0	-4	4	-4	0	-4	4	-4	0
$15-29$	0	4	-4	0	-4	-4	-4	0	-4	4	0	0	-4	4	-4
$30-44$	0	-4	4	-4	0	0	4	-4	0	-4	-4	-4	0	-4	4
$45-59$	0	0	-4	4	-4	0	-4	4	-4	0	0	4	-4	0	-4

Then, a new sequence $\mathbf{u}$ *of period* $60 = 4 \times 15$ *is given by*

$$\mathbf{u} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Consider $C_{\mathbf{u}}(1)$. *From* $\tau = 1$, *we have* $r = 0, s = 1, x = 0$, *and* $y = 1$ *from (3.10). D and T are given by*

$$D = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} 2 & 3 & 3 & 3 & 0 \\ 2 & 3 & 3 & 3 & 0 \\ 2 & 3 & 3 & 3 & 0 \end{bmatrix}.$$

For the last zero column of T, *the number of* 1*'s is one greater than the number of* 0*'s in the corresponding last column of* D. *Thus, the contribution of* D *and* T *to* $C_{\mathbf{u}}(1)$ *is* -1, *and finally* $C_{\mathbf{u}}(1) = -4$.

The autocorrelation function from Theorem 6 is shown in Table 3.1. From Theorem 7, the complete autocorrelation distribution is given by

$$C_{\mathbf{u}}(\tau) = \begin{cases} 60, & 1 \text{ } time \\ 0, & 21 \text{ } times \\ -4, & 26 \text{ } times \\ +4, & 12 \text{ } times. \end{cases}$$

Both the autocorrelation function and the distribution are verified from computer experiments.

3.5 Other Aspects of New Binary Sequences

In this section, the exact linear complexity of the new binary sequences is mathematically derived. Implementation of the sequences requires only a small number of shift registers and a simple logic.

3.5.1 Linear Complexity

The linear complexity of a sequence is defined as the shortest length of a shift register which generates the sequence, or equivalently a degree of the minimal polynomial of the sequence [33]. Before examining the linear complexity of the sequence **u** from Construction 1, the following lemmas are considered.

Lemma 5 *Let* $\mathbf{z} = \{z_t\} = (1,1,0,0)$ *or* $(1,0,0,1)$. *Let* $\mathbf{c} = \{c_t\}$ *be a binary sequence of period* $n = 4(2^k+1)$ *such that*

$$c_t = \begin{cases} 0, & \text{if } t \neq t'(2^k+1), \\ z_{t'}, & \text{if } t = t'(2^k+1) \end{cases} \tag{3.22}$$

where k *is a positive integer and* t' *is an integer,* $0 \leq t' \leq 3$. *Then, the minimal polynomial of* $\mathbf{c}$ *is* $m_c(x) = (x^{2^k+1}+1)^3$.

Proof. Let $C(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1}x^{n-1}$. From (3.22), $C(x)$ is given by

$$C(x) = \begin{cases} 1 + x^{2^k+1}, & \text{if } \mathbf{z} = (1,1,0,0), \\ 1 + x^{3(2^k+1)}, & \text{if } \mathbf{z} = (1,0,0,1). \end{cases}$$

From [56], the minimal polynomial of $\mathbf{c}$ is given by

$$\begin{aligned} m_c(x) &= \frac{x^n+1}{\gcd(x^n+1, C(x))} = \frac{x^{4(2^k+1)}+1}{\gcd(x^{4(2^k+1)}+1, C(x))} = \frac{(x^{2^k+1}+1)^4}{x^{2^k+1}+1} \\ &= (x^{2^k+1}+1)^3 \end{aligned}$$

for both cases of $\mathbf{z}$. □

Lemma 6 *Let* $\mathbf{u} = A(\mathbf{a}, \mathbf{e}) + \mathbf{b}$ *be the binary sequence from Construction 1. Then,* $\mathbf{u}$ *is represented by* $\mathbf{u} = \mathbf{a} + \mathbf{b} + \mathbf{c}$*, i.e.,*

$$u_t = a_t + b_t + c_t, \quad 0 \le t \le 4(2^m - 1) - 1 \tag{3.23}$$

where $\mathbf{c} = \{c_t\}$ *is the binary sequence from Lemma 5.*

Proof. Assume that $\mathbf{u} = \mathbf{a} + \mathbf{b} + \mathbf{f}$ where $\mathbf{f} = \{f_t\}$. Since $u_t = a_t + b_t$ at $t \not\equiv 0 \pmod{2^k + 1}$, it is clear that

$$f_t = 0 \text{ for } t \not\equiv 0 \pmod{2^k + 1}. \tag{3.24}$$

At $t = t'(2^k + 1)$ with $0 \le t' \le 4(2^k - 1) - 1$, on the other hand, u_t is represented as

$$u_t = a_t + b_t + f_t = a_{i+3j} + b_t + f_t \tag{3.25}$$

from the interleaved structure of $\mathbf{a}$ in (3.4). From Construction 1, u_t is also represented in a different way, i.e.,

$$u_t = a_{i+3j+\delta} + b_t \tag{3.26}$$

at $t = t'(2^k + 1)$. For $t = t'(2^k + 1)$,

$$t = t'(2^k + 1) = (2^m - 1)i + j$$

where i and j are the row and column indices of the interleaved structure of $\mathbf{u}$, respectively. Thus, $t \equiv t' \equiv 3i + j \pmod 4$. Also, $3^{-1} \equiv 3 \pmod 4$ and $a_{3i} = a_i$. Then, from (3.25) and (3.26),

$$f_t = f_{t'(2^k+1)} = a_{i+3j+\delta} + a_{i+3j} = a_{3i+j+3\delta} + a_{3i+j} = a_{t'+3\delta} + a_{t'} = h_{t'} \tag{3.27}$$

where $\delta = \pm 1$ and the indices are computed modulo 4. Let $\mathbf{h} = \{h_{t'}\}$. Since $\mathbf{a}$ has a period 4, $\mathbf{h}$ and $\mathbf{f}$ have periods 4 and $4(2^k+1)$ from (3.27), respectively. Also, it is easily known from (3.27) that $\mathbf{h} = (1, 1, 0, 0)$ if $\delta = 1$, or $\mathbf{h} = (1, 0, 0, 1)$ if $\delta = -1$,

which is identical to **z** from Lemma 5. Thus,

$$f_t = z_{t'} \text{ for } t = t'(2^k + 1), \quad 0 \le t' \le 3. \tag{3.28}$$

From (3.24) and (3.28), $\mathbf{f} = \mathbf{c}$ in (3.22), and hence (3.23) is true. □

The linear complexity of the binary sequences from Construction 1 is presented by Theorem 8.

Theorem 8 *Let* **u** *be the binary sequence of period* $N = 4(2^m - 1), m = 2k$ *for an integer* $k > 1$ *from Construction 1. Then, the linear complexity of* **u** *is given by*

$$L_c = 3(1 + 2^k) + 1 + 2k.$$

Proof. By the definition of minimal polynomials, $m_u(x)$, the minimal polynomial of **u** is determined by the least common multiple (lcm) of the minimal polynomials of **a**, **b**, and **c** in Lemma 6. Let $m_a(x), m_b(x)$, and $m_c(x)$ be the minimal polynomials of **a**, **b**, and **c**, respectively. It is easily known that $m_a(x) = (x+1)^4$ and $m_b(x)$ is a primitive polynomial of degree $m = 2k$. From Lemma 5,

$$\begin{aligned} lcm(m_a(x), m_c(x)) &= lcm((x+1)^4, (x^{2^k+1}+1)^3) \\ &= lcm((x+1)^4, (x+1)^3(x^{2^k} + x^{2^k-1} + \cdots + x + 1)^3) \\ &= (x+1)(x^{2^k+1}+1)^3. \end{aligned}$$

Note that $\gcd(m_b(x), x+1) = 1$ and if $k > 1$, $\gcd(m_b(x), x^{2^k+1}+1) = 1$. Finally,

$$\begin{aligned} m_u(x) &= lcm(lcm(m_a(x), m_c(x)), m_b(x)) = lcm(m_a(x), m_c(x)) \cdot m_b(x) \\ &= (x+1)(x^{2^k+1}+1)^3 \cdot m_b(x) \end{aligned}$$

where a degree of $m_u(x)$ or the linear complexity of **u** is given by

$$L_c = 3(2^k + 1) + 1 + m = 3(2^k + 1) + 1 + 2k$$

which completes the proof. □

Remark 3 In Theorem 8, $k > 1$ is requested. For the case of $k = 1$, the new binary sequence $\mathbf{u}$ of period 12 also has optimal autocorrelation $C_{\mathbf{u}}(\tau) \in \{12, 0, \pm 4\}$ for any τ. However, its linear complexity is $L_c = 3(2^k+1)+1 = 10$. It is because in the proof of Theorem 8, $m_b(x) = x^2+x+1$ is a factor of $m_c(x) = (x^3+1)^3$, and thus the minimal polynomial $m_u(x)$ of $\mathbf{u}$ is given by $lcm(m_a(x), m_c(x)) = (x+1)(x^3+1)^3$.

To compare the linear complexities of binary sequences with optimal autocorrelation, the following two lemmas are shown on the linear complexities of the product and the ADS sequences of period $4(2^m - 1)$.

Lemma 7 *Let* $\mathbf{p} = \mathbf{a} + \mathbf{b}$ *be the product sequence of period* $4(2^m - 1)$ *with optimal autocorrelation, where* $\mathbf{a}$ *is a perfect binary sequence of period* 4 *and* $\mathbf{b}$ *a binary sequence of period* $2^m - 1$ *with ideal two-level autocorrelation. Then, its linear complexity is given by* $L_c = l + 4$ *where* l *is the linear complexity of* $\mathbf{b}$.

Proof. The minimal polynomial of a perfect binary sequence $\mathbf{a}$ is $m_a(x) = (x+1)^4$. Let $m_b(x)$ be the minimal polynomial of $\mathbf{b}$. In general, $m_b(x) = \prod_i r_i(x)$ where $r_i(x)$ is a primitive polynomial over $\mathbb{F}_2$. We assume $\gcd(m_b(x), x+1) = 1$ without loss of generality. Then the minimal polynomial of $\mathbf{p}$ is given by $m_p(x) = lcm(m_a(x), m_b(x)) = (x+1)^4 \cdot m_b(x)$. Therefore, a degree of $m_p(x)$ or the linear complexity of $\mathbf{p}$ is given by $L_c = l + 4$ where l is a degree of $m_b(x)$ or the linear complexity of $\mathbf{b}$. □

Lemma 8 *Let* $\mathbf{s}$ *be the ADS sequence of period* $4(2^m - 1)$ *defined by (3.2) with a binary two-level autocorrelation sequence* $\mathbf{a}$ *of period* $2^m - 1$ *and a matrix* G *in (3.1). If the linear complexity of* $\mathbf{a}$ *is* l*, then the linear complexity* L_c *of* $\mathbf{s}$ *is at most* $2l + 4$*, i.e.,* $L_c \leq 2l + 4$*. In particular, the equality is achieved if* $\mathbf{a}$ *is a binary m-sequence and* $\eta \neq 0$.

Proof. From (3.1) and (3.2), $\mathbf{s} = \{s_t\}$ is represented by

$$\mathbf{s} = \mathbf{w} + \mathbf{b} \text{ where } s_t = w_t + b_t,\ 0 \leq t \leq 4(2^m - 1) - 1 \tag{3.29}$$

where $\mathbf{b} = \{b_t\}$ is a perfect binary sequence of period 4 and $\mathbf{w} = \{w_t\}$ is a binary sequence defined by

$$w_t = \begin{cases} a_t, & \text{if } t \equiv 0 \pmod 2 \\ a_{t+\eta}, & \text{if } t \equiv 1 \pmod 2 \end{cases} \tag{3.30}$$

where $\mathbf{a} = \{a_t\}$ is a binary sequence of period $2^m - 1$ with ideal two-level autocorrelation. If $\eta = 0$, then $\mathbf{w} = \mathbf{a}$ and $\mathbf{s} = \mathbf{a} + \mathbf{b}$. Thus, the linear complexity of $\mathbf{s}$ is given by $L_c = l + 4$ from Lemma 7.

For a nontrivial ADS sequence with $\eta \neq 0$, it is clear that $\mathbf{w}$ is a binary sequence of period $2(2^m - 1)$. From (3.30), $\mathbf{w}$ is represented by a $(2^m - 1) \times 2$ interleaved structure where its first column is $(a_0, a_2, a_4, \cdots)$ and the second column is $(a_{1+\eta}, a_{3+\eta}, a_{5+\eta}, \cdots)$. Since each column sequence is a form of a shift and $2-$decimation of $\mathbf{a}$, its minimal polynomial is identical to $m_a(x)$, the minimal polynomial of $\mathbf{a}$. Let $m_w(x)$ be the minimal polynomial of $\mathbf{w}$. From the interleaved structure of $\mathbf{w}$,

$$m_w(x) | m_a(x^2) = (m_a(x))^2 \tag{3.31}$$

from Lemma 1 of [34]. Similar to the proof of Lemma 7, we assume $\gcd(m_a(x), x + 1) = 1$ and thus $\gcd(m_w(x), x+1) = 1$. From (3.29), the minimal polynomial $m_s(x)$ of $\mathbf{s}$ is given by

$$m_s(x) = lcm(m_w(x), m_b(x)) = lcm(m_w(x), (x+1)^4) = m_w(x)(x+1)^4 \tag{3.32}$$

where $m_b(x) = (x + 1)^4$ is the minimal polynomial of a perfect binary sequence $\mathbf{b}$. From (3.31) and (3.32), $L_c = l_w + 4 \leq 2l + 4$ where l_w is a degree of $m_w(x)$.

In particular, if $\mathbf{a}$ is a binary m-sequence, then $m_a(x)$ is a primitive polynomial of degree l and thus $m_w(x) = m_a(x)$ or $(m_a(x))^2$ from (3.31). If $\eta \neq 0$, then $\mathbf{w} \neq \mathbf{a}$ and thus $m_w(x) \neq m_a(x)$. Hence, $m_w(x) = (m_a(x))^2$ and consequently, $L_c = 2l + 4$ from (3.32). □

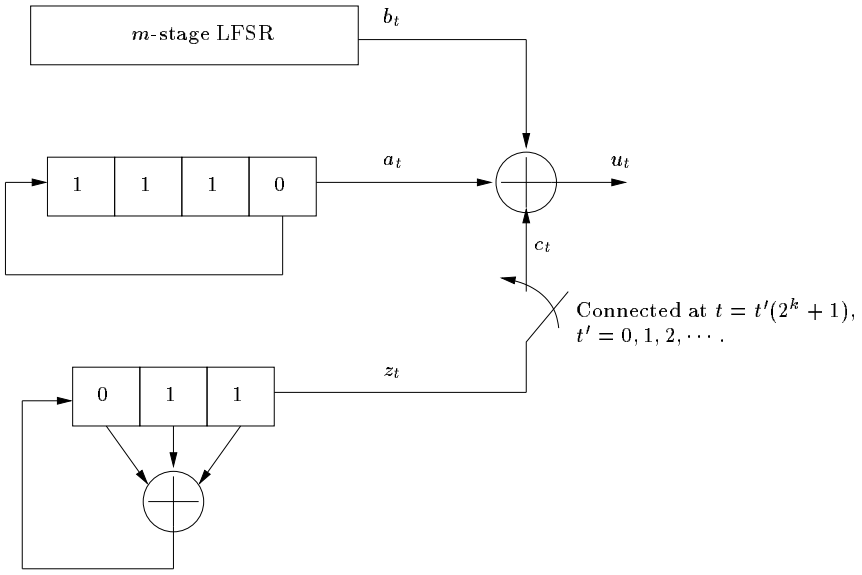


Figure 3.1: Implementation of a new binary sequence of period $4(2^m - 1), m = 2k, k > 1$ with optimal autocorrelation ($\delta = 1$)

3.5.2 Implementation

Figure 3.1 shows the implementation of the new binary sequence $\mathbf{u}$ of period $4(2^m - 1)$ with $\delta = 1$. (If $\delta = -1$, we only need to change the initial state of the 3-stage LFSR from $(0, 1, 1)$ to $(0, 0, 1)$.) In Figure 3.1, the 3-stage linear feedback shift register (LFSR) is enabled and connected to other LFSRs only at time $t = t'(2^k + 1), t' = 0, 1, 2, \cdots$. In the implementation, $\mathbf{u} = \{u_t\}$ is generated by combining only $(m + 7)$ shift registers and a simple connector. From Theorem 8, however, its actual linear complexity is given by $L_c = 3(2^k + 1) + 1 + 2k \gg 2k + 7$ for $k = m/2$, where the large linear complexity is obtained by the low implementation cost.

Remark 4 Let l be the linear complexity of a binary sequence of period $2^m - 1$ with ideal two-level autocorrelation, which is used as an indicator or a base sequence of the product or the ADS sequences. Then, the linear complexities of the product and the ADS sequences are given by $l + 4$ and at most $2l + 4$ from Lemmas 7 and 8, respectively. If $l \gg m$, then the linear complexities can be larger than that of the new sequences from Construction 1. In this case, however, as many numbers of

Table 3.2: Linear complexities of binary sequences of period $4(2^m-1)$ with optimal autocorrelation adopting binary m-sequences

m	Period	Product Sequence	ADS Sequence	New Sequence
4	60	8	12	20
6	252	10	16	34
8	1020	12	20	60
10	4092	14	24	110
12	16380	16	28	208
14	65532	18	32	402
16	262140	20	36	788
18	1048572	22	40	1558

shift registers as the linear complexities are needed for their implementation, which requires the larger implementation cost.

3.5.3 Comparison of Binary Sequences of Period $N \equiv 0 \pmod 4$

In this subsection, we compare three different binary sequences of period $4(2^m-1)$ with optimal autocorrelation – the product, the ADS, and the new sequences. In practical applications, the three sequences will show the same performance in terms of autocorrelation. Thus, we compare them with respect to their linear complexities.

Table 3.2 shows the linear complexities of the three different binary sequences of period $4(2^m-1)$ with optimal autocorrelation. For the product and the ADS sequences adopting m-sequences of period 2^m-1 as indicator and base sequences in their interleaved structures, their linear complexities are given by $m+4$ and $2m+4$ from Lemmas 7 and 8, respectively. In Table 3.2, the new binary sequences from Construction 1 provides much larger linear complexity than the other two classes of sequences. The linear complexities of Table 3.2 are confirmed by computer experiments using the Berlekamp-Massey algorithm [10] [62]. In conclusion, the linear complexity of the new binary sequences is much larger that the others if similar costs are considered for their implementations.

3.6 Conclusion

From a $4 \times (2^m - 1)$ interleaved structure, new binary sequences of period $N = 4(2^m - 1)$ for even $m \geq 4$ with four-valued autocorrelation $\{N, 0, \pm 4\}$ have been constructed, which are optimal with respect to autocorrelation magnitude. The complete autocorrelation distribution and the exact linear complexity of the sequences have been mathematically derived. Only with $(m + 7)$ shift registers and a simple connector, the sequences are implemented to give the large linear complexity.

Chapter 4

New Binary Sequence Family with Low Correlation and A Large Size

In this chapter, a new binary sequence family with low correlation and a large size is proposed [95]. The related works are first reviewed and the motivation is given. As a preliminary study, the Reed-Muller (RM) codes are introduced and the weight distribution of subcodes of the second order RM codes is reviewed. As a main result of this chapter, a new binary sequence family of period $2^n - 1$ is constructed for both odd and even n. The sequence family has several subsequence families induced by a variable factor ρ. In particular, the complete correlation distribution of the sequence family for odd n and $\rho = 2$ is mathematically derived. In addition, the asymptotes of the Welch [91], Sidelnikov [86], and Levenshtein [54] bounds are studied, and the asymptotic optimality of the new binary sequence family is examined by the asymptotic bounds [97].

4.1 Related Works and Motivation

In code division multiple access (CDMA) communication systems, pseudo-noise sequences are assigned to distinct users existing in a common channel at the same time [87]. To distinguish each user and minimize the mutual interference, we must have low crosscorrelation between the distinct sequences. Furthermore, we must

also have low autocorrelation between a sequence and its time shifted version in order to acquire the accurate timing information at a receiver. We can increase the capacity of CDMA systems by obtaining an increased number of sequences which support a larger number of distinct users. Consequently, a family of sequences with low correlation and a large family size plays important roles in CDMA communication systems.

The construction of a family of binary sequences of period $2^n - 1$ is based on a combination of a binary m-sequence and its decimations such that the resulting sequences have low correlation achieving the well known lower bounds derived by Welch [91], Sidelnikov [86], and Levenshtein [54]. For odd n, the Gold sequences [31] constitute one of the families with optimal correlation achieving the Sidelnikov bound. A family of the Kasami (small set) sequences [46] gives sequences with optimal correlation achieving the Welch's lower bound for even n.

In order to obtain binary sequences with large linear spans as well as low correlation, Boztas and Kumar constructed a new family of binary sequences for odd n, so called the *Gold-like sequences* [11]. It has the same period, family size, and maximum correlation as those of the family of Gold sequences, but larger linear spans giving a better potential cryptographic property. Similarly, Udaya constructed a new family of binary sequences with low correlation but large linear spans for even n [89]. In [49], Kim and No generalized these two constructions at the price of the decrease of maximum linear spans and the increase of maximum correlation.

Other known approaches are summarized as follows. In [17], Chang *et al.* showed that a binary cyclic code based on the three-term sequences [70] has the five-valued nonzero weight distribution, which is identical to that of a dual code of the triple error correcting BCH code. From these cyclic codes, equivalent families of binary sequences with six-valued correlation of maximum $1 + 2^{\frac{n+3}{2}}$, a family size 2^{2n}, and a maximum linear span $3n$, can be constructed. In [66], it is shown that the Z_4-linear binary codes become nonlinear cyclic codes after a proper permutation. From these codes, Shanbhag, Kumar, and Helleseth presented a new generalized construction of binary sequence families in [84], including the Kerdock and the Delsarte-Goethals sequences in [40].

In this chapter, a new family $\mathcal{S}_o(\rho)$ of binary sequences of period $2^n - 1$ is constructed for odd $n = 2l + 1$ and an integer ρ with $1 \leq \rho \leq l$. For a given ρ, the maximum correlation of sequences in $\mathcal{S}_o(\rho)$ is $1 + 2^{\frac{n+2\rho-1}{2}}$, and its family size is $2^{n\rho}$. Similarly, a new family $\mathcal{S}_e(\rho)$ of binary sequences of period $2^n - 1$ is also presented for even $n = 2l$ and an integer ρ with $1 \leq \rho < l$, where the maximum correlation and the family size are $1 + 2^{\frac{n}{2}+\rho}$ and $2^{n\rho}$, respectively. Maximum and minimum linear spans of sequences in both $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are $\frac{n(n+1)}{2}$ and $\frac{n(n-2\rho+1)}{2}$, respectively. Here, $\mathcal{S}_o(1)$ is equivalent to the family of Gold-like sequences constructed by Boztas and Kumar, and $\mathcal{S}_e(1)$ is equivalent to the one constructed by Udaya, where m-sequences are excluded in both constructions.

According to specific applications, a proper value ρ and the corresponding family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) can be chosen. For example, a small value of ρ is chosen if low correlation is more crucial than a large family size in the application. If a large family size is more important, on the other hand, we can choose a large value of ρ. The flexibility due to ρ allows the new sequence family to possess the adaptive family size and the maximum correlation for practical applications. Furthermore, the new sequence family has a good potential cryptographic property by the large linear spans. Implementation of $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) is extremely easy by summing up linear feedback shift register (LFSR) outputs. The family $\mathcal{S}_o(2)$ with maximum correlation $1 + 2^{\frac{n+3}{2}}$ and a family size 2^{2n} is a good example for compromise between correlation and a family size.

4.2 Reed-Muller Codes and The Weight Distribution

In this section, the rth order Reed-Muller codes are first introduced. Then, the weight distribution of subcodes of the second order RM codes is explained, which is closely related to the correlation of binary sequence families [46] [60].

4.2.1 Reed-Muller Codes

For $0 \leq r \leq n$, the *rth order Reed-Muller (RM) code* $R(r,n)$ of length $N = 2^n$ is defined [60] by a set of all vectors $F(\mathbf{y}_1, \cdots, \mathbf{y}_n)$ of length N given by

$$F(\mathbf{y}_1, \cdots, \mathbf{y}_n) = c_0\underline{\mathbf{1}} + \sum_{1 \leq j \leq r} c_{i_1 i_2 \cdots i_j} \mathbf{y}_{i_1} \cdot \mathbf{y}_{i_2} \cdots \mathbf{y}_{i_j}, \quad c_0, c_{i_1 i_2 \cdots i_j} \in \mathbb{F}_2 \tag{4.1}$$

where $\underline{\mathbf{1}} = (1, \cdots, 1)$ of length N, and the sum runs through all non-empty subsets of $\{i_1, \cdots, i_j\} \subset \{1, 2, \cdots, n\}$. In (4.1), $\mathbf{y}_1, \cdots, \mathbf{y}_n$ are basis vectors of length N for $R(1,n)$, and $\mathbf{y}_i \cdot \mathbf{y}_j$ is a vector whose kth element is a product of the kth elements of $\mathbf{y}_i$ and $\mathbf{y}_j$, respectively. Indeed, $R(0,n)$ is composed of all zero and all one vectors of length N. $R(1,n)$ is always the dual of an extended Hamming code, and it is also obtained from the Sylvester-type Hadamard matrix [60]. In (4.1), the kth element of a codeword in $R(r,n)$ is given by a Boolean function $f(y_{1,k}, \cdots, y_{n,k})$ with a degree of at most r, where $y_{1,k}, \cdots, y_{n,k}$ are the kth elements of $\mathbf{y}_1, \cdots, \mathbf{y}_n$, respectively. If the kth components corresponding to $y_{1,k} = \cdots = y_{n,k} = 0$ are removed, then a *punctured* RM code $R(r,n)^*$ for $0 \leq r \leq n-1$ is obtained, where each codeword has a length $2^n - 1$.

4.2.2 Weight Distribution of Linear Subcodes of $R(2,n)^*$

Throughout this chapter, the component-wise sum of a pair of binary sequences is considered, where the sum is equivalent to a codeword of a linear subcode of the second order Reed-Muller code. Thus, the weight distribution of the subcode can be applied for investigating the distribution of correlation values of the sequences. In this chapter, the codewords of a linear subcode of the second order RM code are described by the Boolean and polynomial functions. For more details of the functions, recall their definitions in Section 1.5. Also, recall the relation of Hamming weights and exponential sums described in Section 1.5.

Weight distribution of a codeword set with rank $2h$: For odd $n = 2l + 1$,

consider a codeword given by

$$f(x) = Tr(\eta_0 x) + \sum_{j=1}^{l} Tr(\eta_j x^{1+2^j}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.2}$$

where each η_j with $0 \leq j \leq l$ is an element in $\mathbb{F}_{2^n}$. For even $n = 2l$, on the other hand, consider a codeword given by

$$f(x) = Tr(\eta_0 x) + \sum_{j=1}^{l-1} Tr(\eta_j x^{1+2^j}) + Tr_1^l(\eta_l x^{1+2^l}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.3}$$

where each $\eta_j \in \mathbb{F}_{2^n}$ for $0 \leq j \leq l-1$, and $\eta_l \in \mathbb{F}_{2^l}$. With respect to a basis $\{\beta_1, \cdots, \beta_n\}$ of $\mathbb{F}_{2^n}$, x has an expansion of $x = \sum_{i=1}^{n} x_i \beta_i$ with $x_i \in \mathbb{F}_2$ for all i's. Applying this expansion to (4.2) or (4.3), $f(x) = f\left(\sum_{i=1}^{n} x_i \beta_i\right)$ is equivalent to a Boolean function of degree less than or equal to 2, and it may be written as follows [40].

$$f(\mathbf{x}) = f(x_1, \cdots, x_n) = \mathbf{x}\mathbf{Q}\mathbf{x}^T + \mathbf{L} \cdot \mathbf{x}^T, \quad \mathbf{x} \in \mathbb{F}_2^n \setminus \{\underline{\mathbf{0}}\} \tag{4.4}$$

where $\mathbf{Q}$ is an $n \times n$ binary upper triangular matrix, $\mathbf{L}$ is a binary vector in $\mathbb{F}_2^n$, and $\underline{\mathbf{0}} = (0, \cdots, 0)$ of length n. Obviously, $\mathbf{Q}$ is determined by η_j's for nonzero j, and $\mathbf{L}$ by η_0. While $\mathbf{x} = (x_1, \cdots, x_n)$ runs through each nonzero binary n-tuple in $\mathbb{F}_2^n$, $f(\mathbf{x})$ produces each element of a codeword of length $2^n - 1$ in $R(2,n)^*$. Equivalently, $f(x)$ forms a codeword in $R(2,n)^*$ for $x \in \mathbb{F}_{2^n}^*$.

For a given nonzero $\mathbf{Q}$ and all $\mathbf{L}$'s, it is well known that the weight distribution of a set of codewords of the *quadratic Boolean function* $f(\mathbf{x})$ is determined by a rank of the *symplectic matrix* $\mathbf{B} = \mathbf{Q} + \mathbf{Q}^T$ [60]. Equivalent to $\mathbf{B}$, the *symplectic form* $B_f(x,z) = f(x) + f(x+z) + f(z)$ associated with $f(x)$ [40] can be used for given η_j's with $1 \leq j \leq l$ and all η_0's in order to derive the weight distribution of a set of codewords given by $f(x)$. The following fact explains the distribution of exponential sums of $f(x)$, which corresponds to the weight distribution of $f(x)$.

Fact 3 (Theorem 6.2 in [40]) *Let η_j's of $f(x)$ in (4.2) or (4.3) be given such*

Table 4.1: Weight distribution of C_G by $f(x) = Tr(ax) + Tr(bx^{2^k+1})$

Weight	Exponential Sum	Distribution
0	2^n	1
$2^{n-1} - 2^{n-h-1}$	2^{n-h}	$(2^n - 1)(2^{2h-1} + 2^{h-1})$
2^{n-1}	0	$(2^n - 1)(2^n - 2^{2h} + 1)$
$2^{n-1} + 2^{n-h-1}$	-2^{n-h}	$(2^n - 1)(2^{2h-1} - 2^{h-1})$

Table 4.2: Weight distribution of C_T by $f(x) = Tr(ax) + Tr(bx^3) + Tr(cx^5)$

Weight	Exponential Sum	Distribution
0	2^n	1
$2^{n-1} \pm 2^{\frac{n+1}{2}}$	$\mp 2^{\frac{n+3}{2}}$	$\frac{1}{3}(2^n - 1) \cdot 2^{\frac{n-5}{2}} \cdot (2^{\frac{n-3}{2}} \mp 1) \cdot (2^{n-1} - 1)$
2^{n-1}	0	$(2^n - 1) \cdot (9 \cdot 2^{2n-4} + 3 \cdot 2^{n-3} + 1)$
$2^{n-1} \pm 2^{\frac{n-1}{2}}$	$\mp 2^{\frac{n+1}{2}}$	$\frac{1}{3}(2^n - 1) \cdot 2^{\frac{n-3}{2}} \cdot (2^{\frac{n-1}{2}} \mp 1) \cdot (5 \cdot 2^{n-1} + 4)$

that at least one η_j is nonzero for $1 \leq j \leq l$, where $l = \lfloor \frac{n}{2} \rfloor$. For an integer h with $1 \leq h \leq l$, if $B_f(x, z)$ has a rank $2h$, or equivalently $B_f(x, z) = 0$ has 2^{n-2h} solutions in $x \in \mathbb{F}_{2^n}$ for all $z \in \mathbb{F}_{2^n}^*$, then the exponential sum of $f(x)$ takes on values of 0 and $\pm 2^{n-h}$ for all $\eta_0 \in \mathbb{F}_{2^n}$, and its distribution is given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = \begin{cases} 0, & 2^n - 2^{2h} \text{ times} \\ +2^{n-h}, & 2^{2h-1} + 2^{h-1} \text{ times} \\ -2^{n-h}, & 2^{2h-1} - 2^{h-1} \text{ times} \end{cases}$$

with $f(0) = 0$.

Weight distribution of a linear subcode with multiple ranks: For a set of distinct nonzero $\mathbf{Q}$'s, a set of codewords given by $f(\mathbf{x})$ in (4.4) can be considered for all $\mathbf{L}$'s. Equivalently, a set of codewords given by $f(x)$ can be considered for distinct sets of η_j's such that at least one η_j is nonzero in each set of η_j's for $1 \leq j \leq l$. Then, $f(x)$ may have distinct multiple ranks each of which corresponds to each set of given η_j's. Consequently, the exponential sums of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ from Fact 3 for each possible h. If $f(x)$ further constitutes a linear subcode for the sets of η_j's, the weight distribution of the subcode may be used in order to investigate the distribution of exponential sums of $f(x)$.

This section is concluded by specifying the known weight distributions of two linear cyclic subcodes of $R(2,n)^*$ for odd n, which will be used in a later section for determining the correlation distribution of a new family of sequences. For a, b in $\mathbb{F}_{2^n}$ and k with $\gcd(k,n) = 1$ for odd n, a linear cyclic subcode C_G given by $f(x) = Tr(ax) + Tr(bx^{2^k+1})$ for $x \in \mathbb{F}_{2^n}^*$ has the distribution of weights and the corresponding exponential sums in Table 4.1, where $h = \frac{n-1}{2}$. For a, b, c in $\mathbb{F}_{2^n}$ for odd n, on the other hand, a linear cyclic subcode C_T given by $f(x) = Tr(ax) + Tr(bx^3) + Tr(cx^5)$ for $x \in \mathbb{F}_{2^n}^*$ has the distribution of weights and the corresponding exponential sums in Table 4.2. In both tables, the exponential sum means $\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)}$ with $f(0) = 0$. There are several different ways to establish the validity of the weight distribution of C_G, for example, see [31] and [60]. For C_T, see [60].

4.3 New Binary Sequence Family with Low Correlation and A Large Size

This section presents a new family of binary sequences with a large family size as well as a large linear span for odd and even n. For a family of binary sequences, recall the definition in Section 1.5.

4.3.1 Construction of $\mathcal{S}_o(\rho)$ for Odd n

Construction 2 *For odd $n = 2l + 1$ and an integer ρ with $1 \le \rho \le l$, a family $\mathcal{S}_o(\rho)$ of binary sequences is defined by*

$$\mathcal{S}_o(\rho) = \{\mathbf{s}^{\mathbf{\Lambda}} \mid \mathbf{\Lambda} = (\lambda_0, \cdots, \lambda_{\rho-1}),\ \lambda_i \in \mathbb{F}_{2^n}\}$$

where $\mathbf{s}^{\mathbf{\Lambda}} = \{s_0^{\mathbf{\Lambda}}, s_1^{\mathbf{\Lambda}}, \cdots, s_{2^n-2}^{\mathbf{\Lambda}}\}$ is a binary sequence of period $2^n - 1$ with $s_t^{\mathbf{\Lambda}} = s_{\mathbf{\Lambda}}(\alpha^t)$ for a primitive element α of $\mathbb{F}_{2^n}$, where $s_{\mathbf{\Lambda}}(x)$, the trace representation of

$s_t^{\mathbf{\Lambda}}$, *is given by*

$$\begin{aligned} s_{\mathbf{\Lambda}}(x) &= s_{\lambda_0,\cdots,\lambda_{\rho-1}}(x) \\ &= Tr(\lambda_0 x) + \sum_{i=1}^{\rho-1} Tr(\lambda_i x^{1+2^i}) + \sum_{i=\rho}^{l} Tr(x^{1+2^i}) \end{aligned} \tag{4.5}$$

for $x \in \mathbb{F}_{2^n}^*$.

The parameters of a new signal set $\mathcal{S}_o(\rho)$ are determined by the following theorem.

Theorem 9 *For odd* $n = 2l+1$ *and an integer* ρ *with* $1 \le \rho \le l$, *the family* $\mathcal{S}_o(\rho)$ *has* $2^{n\rho}$ *cyclically distinct binary sequences of period* $2^n - 1$. *The correlation of sequences is* $(2\rho+2)$*-valued and the maximum correlation is* $1+2^{\frac{n+2\rho-1}{2}}$. *Therefore,* $\mathcal{S}_o(\rho)$ *constitutes a* $(2^n-1, 2^{n\rho}, 1+2^{\frac{n+2\rho-1}{2}})$ *signal set.*

In order to prove Theorem 9, i.e., to determine the family size and the correlation of $\mathcal{S}_o(\rho)$, the following lemmas are needed. For the cyclic equivalence and distinctness of binary sequences, recall their definitions in Section 1.5.

Lemma 9 *All sequences in* $\mathcal{S}_o(\rho)$ *are cyclically distinct. Thus, the family size of* $\mathcal{S}_o(\rho)$ *is* $2^{n\rho}$.

Proof. Consider a time shifted version of a sequence in $\mathcal{S}_o(\rho)$ represented by

$$s_{\mathbf{\Theta}}(\delta x) = Tr(\theta_0 \delta x) + \sum_{i=1}^{\rho-1} Tr(\theta_i \delta^{1+2^i} x^{1+2^i}) + \sum_{i=\rho}^{l} Tr(\delta^{1+2^i} x^{1+2^i})$$

for $\mathbf{\Theta} = (\theta_0, \cdots, \theta_{\rho-1})$, $\theta_i \in \mathbb{F}_{2^n}$, and $\delta \in \mathbb{F}_{2^n}^*$. It is identical to another sequence of (4.5), i.e., $s_{\mathbf{\Lambda}}(x) = s_{\mathbf{\Theta}}(\delta x)$ for all $x \in \mathbb{F}_{2^n}^*$ if and only if

$$\lambda_0 = \theta_0 \delta,\ \lambda_i = \theta_i \delta^{1+2^i} \text{ for } 1 \le i < \rho, \text{ and } \delta^{1+2^i} = 1 \text{ for } \rho \le i \le l. \tag{4.6}$$

Since $\gcd(2^n - 1, 1 + 2^l) = 1$ for odd n, a unique solution achieving $\delta^{1+2^l} = 1$ is $\delta = 1$. If $\delta = 1$ in (4.6), it gives only a trivial solution of $\lambda_i = \theta_i$ for $0 \le i < \rho$.

Thus, the sequences in $\mathcal{S}_o(\rho)$ represented by $s_{\mathbf{\Lambda}}(x)$ for any λ_i in $\mathbb{F}_{2^n}$ with $0 \le i < \rho$, are cyclically distinct. □

Recall the definition of crosscorrelation of binary sequences described in Section 1.5. Then, the crosscorrelation of two sequences $\mathbf{s}^{\mathbf{\Lambda}}$ and $\mathbf{s}^{\mathbf{\Theta}}$ in $\mathcal{S}_o(\rho)$ is given by

$$C_{\mathbf{s}^{\mathbf{\Lambda}},\mathbf{s}^{\mathbf{\Theta}}}(\tau) = -1 + \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)}$$

where

$$f(x) = Tr(\eta_0 x) + \sum_{i=1}^{l} Tr(\eta_i x^{1+2^i}) \tag{4.7}$$

where

$$\eta_i = \begin{cases} \lambda_0 + \theta_0\delta, & i = 0 \\ \lambda_i + \theta_i\delta^{1+2^i}, & 1 \le i < \rho \\ 1 + \delta^{1+2^i}, & \rho \le i \le l \end{cases} \tag{4.8}$$

for $\lambda_i, \theta_i \in \mathbb{F}_{2^n}$ with $0 \le i < \rho$ and $\delta = \alpha^\tau \in \mathbb{F}_{2^n}^*$ where α is a primitive element of $\mathbb{F}_{2^n}$. In other words, the sum of $\mathbf{s}^{\mathbf{\Lambda}}$ and a τ-shifted version of $\mathbf{s}^{\mathbf{\Theta}}$ can be considered as a codeword given by $f(x)$ in (4.7). Thus, we need to investigate the exponential sum of $f(x)$ for the correlation of the pair of sequences. In the following, the exponential sum is classified by η_i's.

Case 1. $\eta_i = 0$ for $0 \le i \le l$. In this case, $f(x) = 0$. This corresponds to a trivial exponential sum corresponding to in-phase autocorrelation of a sequence. Thus, its exponential sum is 2^n.

Case 2. $\eta_0 \neq 0$ and $\eta_i = 0$ for $1 \le i \le l$. In this case, $f(x) = Tr(\eta_0 x)$. Hence, the exponential sum is 0 for any $\eta_0 \in \mathbb{F}_{2^n}^*$ from the orthogonality of a trace function [33].

Case 3. At least one $\eta_i \neq 0$ for $1 \le i \le l$. In this case, $f(x)$ is equivalent to a quadratic Boolean function. Thus, we need to investigate the number of roots of its symplectic form $B_f(x, z)$ in order to apply Fact 3 for determining the distribution of the exponential sums of $f(x)$.

Lemma 10 examines the number of roots of $B_f(x,z)$ associated with $f(x)$ in case 3.

Lemma 10 *For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, let η_i's of $f(x)$ in (4.7) be given such that at least one η_i is nonzero for $1 \le i \le l$. Then, the symplectic form $B_f(x,z)$ associated with $f(x)$ has at most $2^{2\rho-1}$ roots in $x \in \mathbb{F}_{2^n}$ for all $z \in \mathbb{F}_{2^n}^*$.*

Proof. For given η_i's, the symplectic form $B_f(x,z)$ associated with $f(x)$ is given by

$$\begin{aligned} B_f(x,z) &= f(x) + f(x+z) + f(z) \\ &= \sum_{i=1}^{l} Tr\left(\eta_i (x z^{2^i} + x^{2^i} z)\right) = Tr\left(z \sum_{i=1}^{l} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i})\right) \\ &= Tr\,(zL(x)) \end{aligned} \tag{4.9}$$

where $L(x) = \sum_{i=1}^{l} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i})$. $B_f(x,z) = 0$ for all $z \in \mathbb{F}_{2^n}^*$ if and only if $L(x) = 0$. From (4.8),

$$\begin{aligned} L(x) &= \sum_{i=1}^{\rho-1} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i}) + \sum_{i=\rho}^{l} ((1+\delta^{1+2^{-i}}) x^{2^{-i}} + (1+\delta^{1+2^i}) x^{2^i}) \\ &= \sum_{i=1}^{\rho-1} (\eta_i^{2^{-i}} x^{2^{-i}} + \eta_i x^{2^i}) + \sum_{i=1}^{\rho-1} ((1+\delta^{1+2^{-i}}) x^{2^{-i}} + (1+\delta^{1+2^i}) x^{2^i}) \\ &\quad + \sum_{i=1}^{l} ((1+\delta^{1+2^{-i}}) x^{2^{-i}} + (1+\delta^{1+2^i}) x^{2^i}) \\ &= \sum_{i=1}^{\rho-1} ((\eta_i^{2^{-i}} + 1 + \delta^{1+2^{-i}}) x^{2^{-i}} + (\eta_i + 1 + \delta^{1+2^i}) x^{2^i}) + \sum_{i=1}^{n-1} (1+\delta^{1+2^i}) x^{2^i}. \end{aligned}$$

Note that

$$\begin{aligned}\sum_{i=1}^{n-1}(1+\delta^{1+2^i})x^{2^i} &= \sum_{i=1}^{n-1}x^{2^i}+\delta\sum_{i=1}^{n-1}(\delta x)^{2^i}\\ &= x+Tr(x)+\delta(Tr(\delta x)+\delta x)\\ &= (1+\delta^2)x+Tr(x)+\delta Tr(\delta x).\end{aligned} \tag{4.10}$$

Let

$$\gamma_i = \eta_i + 1 + \delta^{1+2^i}, \quad 1 \le i \le \rho-1. \tag{4.11}$$

Then, $\gamma_i^{2^{-i}} = \eta_i^{2^{-i}} + 1 + \delta^{1+2^{-i}}$. Together with (4.10),

$$L(x) = q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) + Tr(x) + \delta Tr(\delta x) \tag{4.12}$$

where

$$q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) = (1+\delta^2)x + \sum_{i=1}^{\rho-1}(\gamma_i^{2^{-i}}x^{2^{-i}} + \gamma_i x^{2^i}). \tag{4.13}$$

For $L(x)=0$, we have to count the number of solutions in the equation

$$q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) + Tr(x) + \delta Tr(\delta x) = 0 \tag{4.14}$$

for given γ_i's in $\mathbb{F}_{2^n}$ and δ in $\mathbb{F}_{2^n}^*$.

First, it is verified that $q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x)$ is not a constant polynomial of x (i.e., a trivial polynomial). If $\delta=1$, then $\gamma_i=\eta_i$ for $1\le i\le\rho-1$ from (4.11), and $\eta_i=0$ for $\rho\le i\le l$ from (4.8). Then, at least one γ_i is nonzero for $1\le i\le\rho-1$ because at least one η_i is nonzero. If $\delta\neq 1$, on the other hand, $(1+\delta^2)$ cannot be zero although γ_i may be zero for all $1\le i\le\rho-1$. Therefore, $q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x)$ is a polynomial of x with at least one nonzero coefficient of x^{2^i} for $-\rho<i<\rho$.

For the nontrivial polynomial $q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x)$, the equation in (4.14) can be divided into four classes.

i) $Tr(x)=0$ and $Tr(\delta x)=0 \;\Rightarrow\; q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x)=0$,

ii) $Tr(x)=0$ and $Tr(\delta x)=1 \;\Rightarrow\; q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x)+\delta=0$,

iii) $Tr(x) = 1$ and $Tr(\delta x) = 0 \;\Rightarrow\; q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) + 1 = 0$,

iv) $Tr(x) = 1$ and $Tr(\delta x) = 1 \;\Rightarrow\; q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) + 1 + \delta = 0$.

Thus, the left-hand side of (4.14) can be presented by

$$A_a(x) = q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) + a = a + (1+\delta^2)x + \sum_{i=1}^{\rho-1}(\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i})$$

for $a \in \{0, 1, \delta, 1+\delta\}$. Then,

$$\begin{aligned} A_a^{2^{\rho-1}}(x) &= \left[q_{\gamma_1,\cdots,\gamma_{\rho-1},\delta}(x) + a\right]^{2^{\rho-1}} \\ &= a^{2^{\rho-1}} + (1+\delta^2)^{2^{\rho-1}} x^{2^{\rho-1}} + \sum_{i=1}^{\rho-1}(\gamma_i^{2^{\rho-1-i}} x^{2^{\rho-1-i}} + \gamma_i^{2^{\rho-1}} x^{2^{\rho-1+i}}) \end{aligned}$$

and

$$A_a^{2^{\rho-1}}(x) = 0 \iff A_a(x) = 0.$$

Thus, solutions for $A_a^{2^{\rho-1}}(x) = 0$ are all solutions for $A_a(x) = 0$, and vice versa. Therefore, the number of solutions for $A_a(x) = 0$ is equal to that for $A_a^{2^{\rho-1}}(x) = 0$. Since the maximum degree of $A_a^{2^{\rho-1}}(x)$ is $2^{2(\rho-1)}$, $A_a^{2^{\rho-1}}(x) = 0$ has at most $2^{2(\rho-1)}$ solutions. If $\delta \neq 1$, the solutions of $A_a^{2^{\rho-1}}(x) = 0$ are disjoint for different $a \in \{0, 1, \delta, 1+\delta\}$, so the total number of solutions of (4.14) is at most $2^{2(\rho-1)} \cdot 4 = 2^{2\rho}$. If $\delta = 1$, on the other hand, the total number of solutions of (4.14) is at most $2^{2\rho-2}$. Meanwhile, from Fact 3, the possible number of roots of $B_f(x, z)$ is 2^{n-2h} where $n - 2h$ is a positive odd integer for odd n. Therefore, the maximum number of solutions of $B_f(x, z) = 0$ is $2^{2\rho-1}$. □

From Lemma 10, $B_f(x, z) = 0$ has 2^{n-2h} solutions for an integer h where $n - 2h$ is a positive odd integer less than or equal to $2\rho - 1$. Applying this to Fact 3, for all $\eta_0 \in \mathbb{F}_{2^n}$, the exponential sums of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ for an integer h where $n - 2h$ is a positive odd integer less than or equal to $2\rho - 1$. At this point, we need to show that they can take on values of 0 and $\pm 2^{n-h}$ for all h's such that $n - 2h$ is *every* positive odd integer less than or equal to $2\rho - 1$. In order

to do so, the following fact is needed from [60]. (Note: The original representation of the result in [60] is slightly changed in Fact 4.)

Fact 4 ([60], page 454) *Let ζ be a set of the symplectic forms of (4.9) for any $\eta_i \in \mathbb{F}_{2^n}$. For some fixed integer d with $1 \le d \le \lfloor \frac{n}{2} \rfloor$, assume that the rank of every nonzero form in ζ is at least $2d$. Then, the maximum size of ζ is given by*

$$|\zeta|_{\max} = \begin{cases} 2^{n(\frac{n+1}{2}-d)} & \text{for odd } n \\ 2^{(n-1)(\frac{n+2}{2}-d)} & \text{for even } n. \end{cases}$$

In the following, a specific linearized polynomial is given, and a rank of a symplectic form corresponding to the polynomial is investigated.

Lemma 11 *For an integer $\rho > 1$, consider a symplectic form $B(x,z) = Tr(zU(x))$ where the linearized polynomial $U(x)$ is given by*

$$U(x) = \sum_{i=1}^{\rho-1} (\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i})$$

where γ_i can be any element in $\mathbb{F}_{2^n}$. For odd n, the rank of $B(x,z)$ is at least $n - 2\rho + 3$ and every possible rank $2h$ for $n - 2\rho + 3 \le 2h \le n - 1$ occurs at least once while γ_i runs through $\mathbb{F}_{2^n}$. For even n, on the other hand, the rank is at least $n - 2\rho + 2$ and every possible rank $2h$ for $n - 2\rho + 2 \le 2h \le n - 2$ occurs at least once while γ_i runs through $\mathbb{F}_{2^n}$.

Proof. Note that the odd case in Lemma 11 is implicitly known from Theorem 16 and Corollary 17 of Chapter 15 in [60]. Here, it is reproduced for the completeness.

Let k be an integer with $2 \le k \le \rho$. By $\gamma_k = \gamma_{k+1} = \cdots = \gamma_{\rho-1} = 0$, we have

$$B_k(x,z) = Tr(zU_k(x)) \text{ where } U_k(x) = \sum_{i=1}^{k-1} (\gamma_i^{2^{-i}} x^{2^{-i}} + \gamma_i x^{2^i}).$$

Then, the maximum degree of $(U_k(x))^{2^{(k-1)}}$ is $2^{2(k-1)}$ and thus the rank of $B_k(x,z)$ is at least $n - 2k + 2$ (If n is odd, the rank is at least $n - 2k + 3$ because it should

be even).

For odd n, assume that the rank $n-2k+3$ never occurs for all γ_i's with $1 \le i \le k-1$, and thus the rank of $B_k(x,z)$ is at least $n-2k+5$. Then, from Fact 4, the maximum size of a set ζ of such $B_k(x,z)$'s is $|\zeta|_{\max} = 2^{n(k-2)}$. However, its actual size is $2^{n(k-1)}$ when γ_i's with $1 \le i \le k-1$ run through $\mathbb{F}_{2^n}$, which is greater than $|\zeta|_{\max}$. Therefore, the rank of $n-2k+3$ occurs at least once while γ_i's run through $\mathbb{F}_{2^n}$. For even n, if the rank $n-2k+2$ never occurs, then the minimum rank is $n-2k+4$. Similarly, the maximum size of a set ζ of such $B_k(x,z)$'s is $|\zeta|_{\max} = 2^{(n-1)(k-1)}$, which is smaller than its actual size $2^{n(k-1)}$. Therefore, the rank of $n-2k+2$ occurs at least once while γ_i's run through $\mathbb{F}_{2^n}$.

Since $U(x)$ contains every $U_k(x)$'s for $2 \le k \le \rho$ while γ_i's with $1 \le i \le \rho-1$ run through $\mathbb{F}_{2^n}$, the assertion of Lemma 11 follows. □

Using Lemma 11, we have the following result.

Lemma 12 *For odd $n = 2l+1$ and an integer ρ with $1 \le \rho \le l$, let η_i's of $f(x)$ in (4.7) be given such that at least one η_i is nonzero for $1 \le i \le l$, and $\eta_0 \in \mathbb{F}_{2^n}$. Then, the exponential sums of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ for integers h where $n-2h$ is every positive odd integer less than or equal to $2\rho-1$.*

Proof. For $f(x)$, if $\delta = 1$, the linearized polynomial given in (4.12) is of a form of $U(x)$ in Lemma 11. From Lemma 11, therefore, the exponential sum of $f(x)$ for $\delta = 1$ is equal to 0 or $\pm 2^{n-h}$ for all h's such that $n-2h = 1, 3, \cdots, 2\rho-3$. Thus, we can say that the exponential sums of $f(x)$ take on values of 0 and $\pm 2^{n-h}$ for every integer h such that $n-2h = 1, 3, \cdots, 2\rho-3$.

Now, it suffices to show that the exponential sum takes on values of $\pm 2^{n-h}$ at least once when $n-2h = 2\rho-1$. Or equivalently, we need to show the symplectic form $B_f(x,z)$ in (4.9) has a rank of $n-2\rho+1$ at least once for some η_i's. Assume the contradiction that this rank never occurs for all η_i's. Then, the rank of $B_f(x,z)$ is at least $2d = n-2\rho+3$. From Fact 4, the maximum size of a set ζ of such $B_f(x,z)$'s is $|\zeta|_{\max} = 2^{n(\rho-1)}$. However, from (4.8) and (4.9), its actual size is $2^{n\rho}$, which is greater than $|\zeta|_{\max}$. Therefore, the rank of $n-2\rho+1$ occurs at least once while η_i's run through $\mathbb{F}_{2^n}$. This completes the proof of Lemma 12. □

Combining the cases 1, 2, and 3, the following result is obtained on the correlation of sequences in $\mathcal{S}_o(\rho)$.

Lemma 13 *The correlation of binary sequences in* $\mathcal{S}_o(\rho)$ *is* $(2\rho+2)$*-valued and the maximum correlation* $C_{\max}$ *is* $1+2^{\frac{n+2\rho-1}{2}}$.

Proof. For the correlation of sequences in $\mathcal{S}_o(\rho)$, the exponential sum of $f(x)$ in (4.7) is considered. In cases 1 and 2, the exponential sum takes on values of 2^n and 0, respectively. For all η_i's in case 3, the exponential sum takes on 0 and $\pm 2^{n-h}$ for each integer h such that $n-2h = 1, 3, \cdots, 2\rho-1$ from Lemma 12, where the exponential sum takes on 2ρ nonzero distinct values. Including 2^n and 0, therefore, the overall exponential sum is $(2\rho+2)$-valued. Equivalently, the correlation of sequences in $\mathcal{S}_o(\rho)$ is $(2\rho+2)$-valued. Since the maximum value is determined by $n-2h = 2\rho-1$ from Lemma 12, $C_{\max} = |-1-2^{n-h}| = 1+2^{\frac{n+2\rho-1}{2}}$. □

Proof of Theorem 9. The results follow directly from Lemmas 9 and 13. □

Remark 5 From Fact 4, a maximum size of a set of symplectic forms with a rank of at least $2d$ is $2^{n\left(\frac{n+1}{2}-d\right)}$ for odd n. In $\mathcal{S}_o(\rho)$, therefore, $2d = n-2\rho+1$ and a maximum size of the set is $2^{n\rho}$, which is exactly identical to the family size of $\mathcal{S}_o(\rho)$. With the given maximum correlation, therefore, $\mathcal{S}_o(\rho)$ has the largest family size of all binary sequences constructed from subcodes of the second order Reed-Muller code. In other words, $\mathcal{S}_o(\rho)$ is optimum among all *quadratic-form* binary sequence families with respect to maximum correlation and a family size.

Remark 6 If $\rho = 1$,

$$s_{\lambda_0}(x) = Tr(\lambda_0 x) + \sum_{i=1}^{l} Tr(x^{1+2^i}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.15}$$

which represents the Gold-like sequences introduced by Boztas and Kumar [11]. From Lemma 12, a positive odd integer less than 2ρ is only 1, so $n-2h = 1$. Hence, $h = \frac{n-1}{2}$. Finally, the Gold-like sequences given by (4.15) have four-valued correlation, or $\{2^n-1, -1, -1 \pm 2^{\frac{n+1}{2}}\}$ for all $\lambda_0 \in \mathbb{F}_{2^n}$.

Example 10 *If* $\rho = 2$,

$$s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + \sum_{i=2}^{l} Tr(x^{1+2^i}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.16}$$

for odd $n \geq 5$. *From Lemma 12,* $n - 2h = 1$ *and* 3. *Thus,* $h = \frac{n-1}{2}$ *and* $\frac{n-3}{2}$. *From Lemma 13, sequences given by (4.16) have six-valued correlation, or* $\{2^n - 1, -1, -1 \pm 2^{\frac{n+1}{2}}, -1 \pm 2^{\frac{n+3}{2}}\}$ *for all* $\lambda_0, \lambda_1 \in \mathbb{F}_{2^n}$. *For the sequences represented by* $s_{\lambda_0,\lambda_1}(x)$ *and* $s_{\theta_0,\theta_1}(\delta x)$, *the corresponding* $f(x)$ *in (4.7) constitutes a linear subcode with five nonzero distinct weights if we assume that* δ *can be any element in* $\mathbb{F}_{2^n}$. *In fact, the correlation distribution of* $\mathcal{S}_o(2)$ *can be derived from Table 4.2 by the linear subcode structure. The details will be discussed in Section 4.4.*

Example 11 *If* $\rho = 3$,

$$s_{\lambda_0,\lambda_1,\lambda_2}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + Tr(\lambda_2 x^5) + \sum_{i=3}^{l} Tr(x^{1+2^i}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.17}$$

for odd $n \geq 7$. *Similarly,* $h = \frac{n-1}{2}, \frac{n-3}{2}$, *and* $\frac{n-5}{2}$. *From Lemma 13, the sequences given by (4.17) have eight-valued correlation, or* $\{2^n - 1, -1, -1 \pm 2^{\frac{n+1}{2}}, -1 \pm 2^{\frac{n+3}{2}}, -1 \pm 2^{\frac{n+5}{2}}\}$ *for all* $\lambda_0, \lambda_1, \lambda_2 \in \mathbb{F}_{2^n}$.

4.3.2 Construction of $\mathcal{S}_e(\rho)$ for Even n

Construction 3 *For even* $n = 2l$ *and an integer* ρ *with* $1 \leq \rho < l$, *a family* $\mathcal{S}_e(\rho)$ *of binary sequences is defined by*

$$\mathcal{S}_e(\rho) = \{\mathbf{s}^{\mathbf{\Lambda}} \mid \mathbf{\Lambda} = (\lambda_0, \cdots, \lambda_{\rho-1}),\ \lambda_i \in \mathbb{F}_{2^n}\}$$

where $\mathbf{s}^{\mathbf{\Lambda}} = \{s_0^{\mathbf{\Lambda}}, s_1^{\mathbf{\Lambda}}, \cdots, s_{2^n-2}^{\mathbf{\Lambda}}\}$ *is a binary sequence of period* $2^n - 1$ *with* $s_i^{\mathbf{\Lambda}} = s_{\mathbf{\Lambda}}(\alpha^t)$ *for a primitive element* α *of* $\mathbb{F}_{2^n}$, *where* $s_{\mathbf{\Lambda}}(x)$, *the trace representation of*

$s_t^{\mathbf{\Lambda}}$, *is given by*

$$s_{\mathbf{\Lambda}}(x) = Tr(\lambda_0 x) + \sum_{i=1}^{\rho-1} Tr(\lambda_i x^{1+2^i}) + \sum_{i=\rho}^{l-1} Tr(x^{1+2^i}) + Tr_1^l(x^{1+2^l}) \tag{4.18}$$

for $x \in \mathbb{F}_{2^n}^*$.

Theorem 10 *For even* $n = 2l$ *and an integer* ρ *with* $1 \le \rho < l$, *the family* $\mathcal{S}_e(\rho)$ *has* $2^{n\rho}$ *cyclically distinct binary sequences of period* $2^n - 1$. *The correlation of sequences is* $(2\rho+4)$*-valued and maximum correlation is* $1+2^{\frac{n}{2}+\rho}$. *Therefore,* $\mathcal{S}_e(\rho)$ *constitutes a* $(2^n - 1, 2^{n\rho}, 1 + 2^{\frac{n}{2}+\rho})$ *signal set.*

In order to prove Theorem 10, the following lemmas are needed. Since their proofs are similar to those for $\mathcal{S}_o(\rho)$, the details are omitted.

Lemma 14 *All sequences in* $\mathcal{S}_e(\rho)$ *are cyclically distinct. Hence, the family size of* $\mathcal{S}_e(\rho)$ *is* $2^{n\rho}$.

Proof. Similar to the proof of Lemma 9, $s_{\mathbf{\Lambda}}(x) = s_{\mathbf{\Theta}}(\delta x)$ for all x in $\mathbb{F}_{2^n}^*$ if and only if (4.6) is achieved. If $\gcd(1 + 2^i, 2^n - 1) = d > 1$, then d is not a factor of $\gcd(1+2^{i-1}, 2^n - 1)$ since $\gcd(1+2^i, 1+2^{i-1}) = 1$ for any integer i. Hence, if $\delta_1 \neq 1$ is a solution of $\delta^{1+2^i} = 1$, then it cannot be a solution of $\delta^{1+2^{i-1}} = 1$. Meanwhile, there exist at least two equations of $\delta^{1+2^i} = 1$ for $\rho \le i \le l$ in (4.6) because $\rho < l$ for even n. Thus, (4.6) has a unique solution given by $\delta = 1$. Hence, all sequences in $\mathcal{S}_e(\rho)$ are cyclically distinct. □

To investigate the correlation of sequences in $\mathcal{S}_e(\rho)$, the corresponding exponential sum $f(x)$ is given by

$$f(x) = Tr(\eta_0 x) + \sum_{i=1}^{l-1} Tr(\eta_i x^{1+2^i}) + Tr_1^l(\eta_l x^{1+2^l}) \tag{4.19}$$

where

$$\eta_i = \begin{cases} \lambda_0 + \theta_0\delta, & i = 0 \\ \lambda_i + \theta_i\delta^{1+2^i}, & 1 \le i < \rho \\ 1 + \delta^{1+2^i}, & \rho \le i \le l \end{cases} \tag{4.20}$$

for $\lambda_i, \theta_i \in \mathbb{F}_{2^n}$ with $0 \le i < \rho$ and $\delta \in \mathbb{F}_{2^n}^*$. If $\eta_i = 0$ for all $1 \le i \le l$, the exponential sum of $f(x)$ has a value of 0 or 2^n. Otherwise, we have to consider the number of solutions of $B_f(x,z) = 0$ in order to derive the distribution of exponential sums of $f(x)$ in (4.19).

Lemma 15 *For even $n = 2l$ and an integer ρ with $1 \le \rho < l$, let η_i's of $f(x)$ in (4.19) be given such that at least one η_i is nonzero for $1 \le i \le l$. Then, the symplectic form $B_f(x,z)$ associated with $f(x)$ has at most $2^{2\rho}$ roots in $x \in \mathbb{F}_{2^n}$ for all $z \in \mathbb{F}_{2^n}^*$.*

Proof. The symplectic form associated with $f(x)$ is given by

$$\begin{aligned} B_f(x,z) &= Tr\left(z\left(\sum_{i=1}^{l-1}(\eta_i^{2^{-i}}x^{2^{-i}} + \eta_i x^{2^i}) + \eta_l x^{2^l}\right)\right) \\ &= Tr\left(zL(x)\right). \end{aligned}$$

Thus, $B_f(x,z) = 0$ for all $z \in \mathbb{F}_{2^n}^*$ if and only if $L(x) = 0$. Using the same approach as in the proof of Lemma 10, $L(x)$ identical to (4.12) is obtained and the equation (4.14) is derived. Following the same steps as in the proof of Lemma 10, it is proved that the number of solutions of $B_f(x,z) = 0$ is at most $2^{2\rho}$. □

Lemma 16 *For even $n = 2l$ and an integer ρ with $1 \le \rho < l$, let η_i's of $f(x)$ in (4.19) be given such that at least one η_i is nonzero for $1 \le i \le l$, and $\eta_0 \in \mathbb{F}_{2^n}$. Then, the exponential sums of $f(x)$ can take on values of 0 and $\pm 2^{n-h}$ for integers h where $n-2h$ is zero or every positive even integer less than or equal to 2ρ. Hence, the correlation of binary sequences in $\mathcal{S}_e(\rho)$ is $(2\rho+4)$-valued and the maximum correlation $C_{\max}$ is $1 + 2^{\frac{n}{2}+\rho}$.*

Proof. While each η_i with $1 \le i < \rho$ runs through $\mathbb{F}_{2^n}$, each γ_i of $L(x)$ in (4.12) runs through $\mathbb{F}_{2^n}$. From Lemma 11, therefore, the exponential sums of $f(x)$ take

on values of 0 and $\pm 2^{n-h}$ for every integer h such that $n-2h = 2, 4, \cdots, 2\rho-2$.

Next, it is shown that the cases $n-2h=0$ and $n-2h=2\rho$ also occur. First, if all γ_i's of $L(x)$ in (4.12) are zero, then $L(x)=0$ has at most four solutions, i.e., $x = 0, \frac{\delta}{1+\delta^2}, \frac{1}{1+\delta^2}$, and $\frac{1}{1+\delta}$, where the last three solutions are valid only if $Tr\left(\frac{1}{1+\delta}\right)=1$. If all γ_i's are zero and $Tr\left(\frac{1}{1+\delta}\right)=0$, therefore, $L(x)$ in (4.12) has a unique solution $x=0$. Hence, the case $n-2h=0$ occurs at least once. Second, applying Fact 4, the exponential sum of $f(x)$ can take on values of $\pm 2^{n-h}$ at least once such that $n-2h=2\rho$ in the similar way to the proof of Lemma 12.

Consequently, the exponential sums of $f(x)$ take on values of 0 and $\pm 2^{n-h}$ for integers h where $n-2h = 0, 2, \cdots, 2\rho$. Including 0 and 2^n, the correlation is $(2\rho+4)$-valued. Furthermore, $C_{\max} = 1 + 2^{\frac{n}{2}+\rho}$ for $h = \frac{n}{2} - \rho$. □

Proof of Theorem 10. The results follow directly from Lemmas 14 and 16. □

Remark 7 If $\rho = 1$,

$$s_{\lambda_0}(x) = Tr(\lambda_0 x) + \sum_{i=1}^{l-1} Tr(x^{1+2^i}) + Tr_1^l(x^{1+2^l}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.21}$$

which represents the sequences constructed by Udaya [89]. From Lemma 16, $n-2h=0$ and 2. Hence, $h = \frac{n}{2}$ and $\frac{n}{2}-1$. Finally, the sequences given by (4.21) have six-valued correlation, or $\{2^n-1, -1, -1\pm 2^{\frac{n}{2}}, -1\pm 2^{\frac{n}{2}+1}\}$ for all $\lambda_0 \in \mathbb{F}_{2^n}$.

Example 12 *If $\rho = 2$,*

$$s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + \sum_{i=2}^{l} Tr(x^{1+2^i}) + Tr_1^l(x^{1+2^l}), \quad x \in \mathbb{F}_{2^n}^* \tag{4.22}$$

for even $n \geq 6$. From Lemma 16, $n-2h = 0, 2, 4$ and thus, $h = \frac{n}{2}, \frac{n}{2}-1, \frac{n}{2}-2$. Hence, the correlation of sequences given by (4.22) belongs to $\{2^n-1, -1, -1\pm 2^{\frac{n}{2}}, -1\pm 2^{\frac{n}{2}+1}, -1\pm 2^{\frac{n}{2}+2}\}$ for all $\lambda_0, \lambda_1 \in \mathbb{F}_{2^n}$.

Remark 8 Contrary to odd n, $f(x)$ in (4.19) does not constitute a linear subcode because η_i with $\rho \leq i < l$ may not run through all elements in $\mathbb{F}_{2^n}$. Therefore, the

Table 4.3: Linear span and the number of sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$)

$LS(\rho)$	Number of sequences
$\frac{n(n+1)}{2}$	$(2^n-1)^\rho$
$\frac{n(n-1)}{2}$	$\binom{\rho}{1}\cdot(2^n-1)^{\rho-1}$
$\frac{n(n-3)}{2}$	$\binom{\rho}{2}\cdot(2^n-1)^{\rho-2}$
$\vdots$	$\vdots$
$\frac{n(n-2\rho+1)}{2}$	1

weight distribution of linear subcodes cannot be used to investigate the correlation distribution of sequences in $\mathcal{S}_e(2)$.

4.3.3 Linear Spans of $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$

Linear spans of binary sequences in $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are determined by the number of nonzero λ_i's with $0 \le i < \rho$.

Theorem 11 *In the family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$), consider a sequence represented by $s_{\mathbf{\Lambda}}(x)$ where j λ_i's in $\mathbf{\Lambda} = (\lambda_0, \cdots, \lambda_{\rho-1})$ are equal to 0. Let $LS_j(\rho)$ be the linear span of the sequence. Then,*

$$LS_j(\rho) = \frac{n(n-2j+1)}{2}, \quad 0 \le j \le \rho$$

and there are $\binom{\rho}{j}\cdot(2^n-1)^{\rho-j}$ sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) having the linear span $LS_j(\rho)$. From this result, the linear spans of sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) and their distribution are shown in Table 4.3.

Proof. First, consider the linear spans of sequences in $\mathcal{S}_o(\rho)$. In Construction 2, a sequence represented by $s_{\mathbf{\Lambda}}(x)$ has a total $l+1 = \frac{n+1}{2}$ trace terms and each trace term has a linear span of n. If j λ_i's of the sequence are equal to 0, it has $(\frac{n+1}{2} - j)$ nonzero trace terms and the corresponding linear span of the sequence is given by

$$LS_j(\rho) = \left(\frac{n+1}{2} - j\right)\cdot n = \frac{n(n-2j+1)}{2}, \quad 0 \le j \le \rho.$$

Table 4.4: Comparison of the families of binary sequences with low correlation

Family of Sequences	Period	Family Size	$C_{\max}$	Linear Span (Maximum, Minimum)	n
Gold [31]	2^n-1	2^n+1	$1+2^{\frac{n+1}{2}}$	$(2n,n)$	odd
Kasami (Small Set) [46]	2^n-1	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$(\frac{3n}{2},n)$	even
Kasami (Large Set) [46]	2^n-1	$2^{\frac{n}{2}}(2^n+1)-1$ or $2^{\frac{n}{2}}(2^n+1)$	$1+2^{\frac{n}{2}+1}$	$(\frac{5n}{2},n)$	even
Bent [71]	2^n-1	$2^{\frac{n}{2}}$	$1+2^{\frac{n}{2}}$	$(l_{\max},l_{\min})$ 1)	even
Gold-like [11]	2^n-1	2^n+1	$1+2^{\frac{n+1}{2}}$	$\left(\frac{n(n+1)}{2},n\right)$	odd
Udaya [89]	2^n-1	2^n+1	$1+2^{\frac{n}{2}+1}$	$\left(\frac{n(n+1)}{2},n\right)$	even
Chang *et al.* [17] or Dual of BCH [60]	2^n-1	2^{2n}	$1+2^{\frac{n+3}{2}}$	$(3n,n)$	odd
Rothaus [79]	2^n-1	$2^{2n}+2^n+1$	$1+2^{\frac{n+3}{2}}$	$(3n,n)$	odd
Kerdock [84]	$2(2^n-1)$	2^n	$2+2^{\frac{n+1}{2}}$	– 2)	odd
Delsarte-Goethals [84]	$2(2^n-1)$	2^{2n}	$2+2^{\frac{n+3}{2}}$	– 2)	odd
New Family $\mathcal{S}_o(2)$	2^n-1	2^{2n}	$1+2^{\frac{n+3}{2}}$	$\left(\frac{n(n+1)}{2},\frac{n(n-3)}{2}\right)$	odd
New Family $\mathcal{S}_e(2)$	2^n-1	2^{2n}	$1+2^{\frac{n}{2}+2}$	$\left(\frac{n(n+1)}{2},\frac{n(n-3)}{2}\right)$	even
New Family $\mathcal{S}_o(\rho)$ 3)	2^n-1	$2^{n\rho}$	$1+2^{\frac{n+2\rho-1}{2}}$	$\left(\frac{n(n+1)}{2},\frac{n(n-2\rho+1)}{2}\right)$	odd
New Family $\mathcal{S}_e(\rho)$ 3)	2^n-1	$2^{n\rho}$	$1+2^{\frac{n}{2}+\rho}$	$\left(\frac{n(n+1)}{2},\frac{n(n-2\rho+1)}{2}\right)$	even

1) $l_{\max}=\sum_{i=1}^{l-1}\binom{n}{i}+\binom{m}{l}2^l-\sum_{i=1}^{\lfloor\frac{l-1}{2}\rfloor}\binom{m}{i}$, $l_{\min}=\binom{m}{l}2^l+\sum_{i=2}^{l-1}\binom{m}{i}2^{i-1}+n$, $n=2m=4l$. See [52].
2) The linear spans of Kerdock and Delsarte-Goethals sequences are not given in [84].
3) $1<\rho\le\lfloor\frac{n-1}{2}\rfloor$.

Since j λ_i's are 0 and $(\rho-j)$ λ_i's are nonzero, the number of corresponding sequences given by $s_{\mathbf{\Lambda}}(x)$ is $\binom{\rho}{j}\cdot(2^n-1)^{\rho-j}$. Applying this result to each j with $0\le j\le\rho$, the linear spans $LS(\rho)$ of binary sequences in $\mathcal{S}_o(\rho)$ have the distribution shown in Table 4.3. Using the similar approach to odd case, it is verified that the linear spans of sequences in $\mathcal{S}_e(\rho)$ are the same as those of $\mathcal{S}_o(\rho)$. □

Corollary 2 *Maximum and minimum linear spans of sequences in* $\mathcal{S}_o(\rho)$ *and* $\mathcal{S}_e(\rho)$ *are* $\frac{n(n+1)}{2}$ *and* $\frac{n(n-2\rho+1)}{2}$, *respectively.*

4.3.4 Comparison of Binary Sequence Families

Table 4.4 compares $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ with the well known families of binary sequences with low correlation.

In terms of periods and maximum linear spans, new families $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are identical to the families of Gold-like sequences and Udaya's, respectively. At

the expense of maximum correlation $C_{\max}$, however, $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$, $\rho > 1$, gain larger family sizes than those families. Furthermore, we can choose ρ and the corresponding family $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) for its specific application. For example, if low correlation is more crucial than a large family size in the application, then a small value of ρ is chosen. Otherwise, a large value of ρ is chosen in order to get a large family size.

In $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$, m-sequences are not included for increasing their minimum linear spans. Thus, they have larger minimum linear spans than any other binary sequence families in Table 4.4 except for the bent sequences, which will be a good property for potential cryptographic applications.

4.4 Correlation Distribution of $\mathcal{S}_o(2)$

In many applications, a family of binary sequences with both low correlation and a large family size is required. Thus, we need to consider a family of sequences which can be a good compromise between correlation and a family size. As a good candidate for such a family, a new sequence family $\mathcal{S}_o(2)$ is intensively studied by deriving its correlation distribution.

For odd $n = 2l + 1$, a family $\mathcal{S}_o(2)$ of binary sequences is given by

$$\mathcal{S}_o(2) = \{\mathbf{s}^{\mathbf{\Lambda}} \mid \mathbf{\Lambda} = (\lambda_0, \lambda_1),\ \lambda_i \in \mathbb{F}_{2^n}\}$$

where $\mathbf{s}^{\mathbf{\Lambda}} = \{s_0^{\mathbf{\Lambda}}, s_1^{\mathbf{\Lambda}}, \cdots, s_{2^n-2}^{\mathbf{\Lambda}}\}$ is a binary sequence of period $2^n - 1$ with $s_t^{\mathbf{\Lambda}} = s_{\lambda_0,\lambda_1}(\alpha^t)$ for a primitive element α of $\mathbb{F}_{2^n}$, where $s_{\lambda_0,\lambda_1}(x)$ is given by

$$s_{\lambda_0,\lambda_1}(x) = Tr(\lambda_0 x) + Tr(\lambda_1 x^3) + \sum_{i=2}^{l} Tr(x^{1+2^i})$$

for $x \in \mathbb{F}_{2^n}^*$.

From Theorem 9, the family $\mathcal{S}_o(2)$ has 2^{2n} cyclically distinct binary sequences of period $2^n - 1$. The correlation of sequences in $\mathcal{S}_o(2)$ is six-valued and its maximum is $1 + 2^{\frac{n+3}{2}}$. Consequently, $\mathcal{S}_o(2)$ constitutes a $(2^n - 1, 2^{2n}, 1 + 2^{\frac{n+3}{2}})$ signal set. This

data of $\mathcal{S}_o(2)$ is listed in Table 4.4.

Next, the distribution of correlation values of sequences in $\mathcal{S}_o(2)$ is investigated. Correlation of a pair of sequences in $\mathcal{S}_o(2)$ is derived from the exponential sum of

$$f(x) = Tr((\lambda_0 + \theta_0\delta)x) + Tr((\lambda_1 + \theta_1\delta^3)x^3) + \sum_{i=2}^{l} Tr((1 + \delta^{1+2^i})x^{1+2^i})$$

for $\lambda_0, \lambda_1, \theta_0, \theta_1 \in \mathbb{F}_{2^n}$ and $\delta \in \mathbb{F}_{2^n}^*$. With $a = \lambda_0 + \theta_0\delta$, $b = \lambda_1 + \theta_1\delta^3$, and $c_i = 1 + \delta^{1+2^i}$ for $2 \le i \le l$, $f(x)$ has a form of

$$f(x) = Tr(ax) + Tr(bx^3) + \sum_{i=2}^{l} Tr(c_i x^{1+2^i}) \tag{4.23}$$

for $a, b \in \mathbb{F}_{2^n}$ and $c_i \in \mathbb{F}_{2^n} \setminus \{1\}$. According to c_i's, the exponential sum of $f(x)$ can be classified into two exclusive cases. To facilitate the analysis, assume that c_i can be any element in $\mathbb{F}_{2^n}$.

Case 1. $c_i = 0$ for $2 \le i \le l$. In this case, $f(x) = Tr(ax) + Tr(bx^3)$ constitutes a linear cyclic subcode of $R(2,n)^*$ for $a, b \in \mathbb{F}_{2^n}$. Thus, the distribution of exponential sums of $f(x)$ is given by Table 4.1 with $h = h_1 = \frac{n-1}{2}$.

Case 2. At least one $c_i \neq 0$ for $2 \le i \le l$. In this case, the distribution of exponential sums of $f(x)$ is determined by the following lemmas.

Lemma 17 *Let d_1 and d_2 be positive integers such that $d_1 + d_2 = 2^n$. For given c_i's with at least one nonzero c_i, $f(x)$ in (4.23) has the five-valued exponential sums for all a, b in $\mathbb{F}_{2^n}$, and the distribution is given by*

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = \begin{cases} \pm 2^{n-h_1}, & d_1(2^{2h_1-1} \pm 2^{h_1-1}) \text{ times} \\ 0, & d_1(2^n - 2^{2h_1}) + d_2(2^n - 2^{2h_2}) \text{ times} \\ \pm 2^{n-h_2}, & d_2(2^{2h_2-1} \pm 2^{h_2-1}) \text{ times} \end{cases} \tag{4.24}$$

where $h_1 = \frac{n-1}{2}, h_2 = \frac{n-3}{2}$.

Proof. If b is fixed to an element in $\mathbb{F}_{2^n}$, the exponential sum of $f(x)$ follows the

distribution in Fact 3 depending on the rank of its symplectic form. Since $\rho = 2$, it is obvious from Lemma 12 that the symplectic form of $f(x)$ with at least one nonzero c_i may have a pair of distinct ranks $2h_1$ and $2h_2$ according to b, where $n - 2h_1 = 1$ and $n - 2h_2 = 3$, respectively. For a subset Ω of $\mathbb{F}_{2^n}$, assume that if $b \in \Omega$, then $f(x)$ has the rank $2h_1$ and otherwise, $f(x)$ has the rank $2h_2$, where $|\Omega| = d_1$ and $|\Omega^c| = d_2 = 2^n - d_1$. If b runs through all elements in $\mathbb{F}_{2^n}$, the overall distribution of the exponential sums becomes (4.24) after summing up the distributions in Fact 3 for h_1 and h_2. □

By combining both cases 1 and 2, $f(x)$ has the six-valued exponential sums for any c_i in $\mathbb{F}_{2^n}$. Indeed, $f(x)$, $x \in \mathbb{F}_{2^n}^*$ equivalently constitutes a linear subcode of $R(2,n)^*$ with five nonzero distinct weights for a, b, and c_i's in $\mathbb{F}_{2^n}$ with the assumption that each c_i determined by δ can be any element in $\mathbb{F}_{2^n}$. Hence, each codeword in the subcode has a length $2^n - 1$ and its dimension is $3n$ due to a, b, and δ in $\mathbb{F}_{2^n}$. Since the subcode contains the dual of the double error correcting BCH code, the BCH code contains the dual of our subcode. Therefore, it is clear that the minimum distance of the dual of our subcode is at least 5.

Meanwhile, if the number of nonzero distinct weights of a code is less than or equal to a minimum distance of its dual code, then its weight distribution is an explicit function of its codeword length, dimension, and distinct weights (Chapter 6, Theorem 2 in [60]). Hence, if this is applied to our subcode, its weight distribution is determined by the codeword length, dimension, and distinct weights because it has five nonzero distinct weights which are less than or equal to the minimum distance of its dual code. Moreover, the subcode has the same weight distribution as in Table 4.2 because it has the same codeword length, dimension, and weights as those of C_T. Now, the values of d_1 and d_2 in Lemma 17 are derived.

Lemma 18 *For d_1 and d_2 in Lemma 17,*

$$d_1 = \frac{1}{3}(5 \cdot 2^{n-1} + 1) \text{ and } d_2 = \frac{1}{3}(2^{n-1} - 1).$$

Proof. For given c_i's with at least one nonzero c_i, the exponential sums of $f(x)$ have the distribution of (4.24). If $c_i = 0$ for $2 \le i \le l$, on the other hand,

$f(x) = Tr(ax) + Tr(bx^3)$, where the exponential sums of $f(x)$ have the distribution in Table 4.1 for $h = h_1 = \frac{n-1}{2}$. By summing up the distributions of both cases, the overall distribution of exponential sums of $f(x)$ is given by

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)} = \begin{cases} 2^n, & 1 \text{ time} \\ \pm 2^{n-h_1}, & (2^n-1)(d_1+1)(2^{2h_1-1} \pm 2^{h_1-1}) \text{ times} \\ 0, & (2^n-1)(d_1(2^n - 2^{2h_1}) \\ & +d_2(2^n - 2^{2h_2}) + 2^n - 2^{2h_1} + 1) \text{ times} \\ \pm 2^{n-h_2}, & (2^n-1)d_2(2^{2h_2-1} \pm 2^{h_2-1}) \text{ times} \end{cases} \tag{4.25}$$

for all sets of c_i's in $\mathbb{F}_{2^n}$. Moreover, this distribution should be identical to the weight distribution in Table 4.2. Note that $h_1 = \frac{n-1}{2}$ and $h_2 = \frac{n-3}{2}$. From the comparison of (4.25) and Table 4.2, the values of d_1 and d_2 follow immediately. □

So far, we assumed that c_i can be any element in $\mathbb{F}_{2^n}$. In sequences' aspect, however, c_i's in (4.23) cannot be 1 with nonzero δ. Thus, this case should be removed from the distribution in (4.25) in order to obtain the distribution of correlation values of sequences in $\mathcal{S}_o(2)$.

Theorem 12 *The complete distribution of correlation values of all pairs of binary sequences in $\mathcal{S}_o(2)$ is as follows.*

$$C_{\mathbf{s}^{\mathbf{\Lambda}}, \mathbf{s}^{\mathbf{\Lambda}'}} = \begin{cases} 2^n - 1, & 2^{2n} \textit{ times} \\ -1 \pm 2^{\frac{n+1}{2}}, & \frac{1}{3} \cdot 2^{2n} 2^{\frac{n-3}{2}} (2^{\frac{n-1}{2}} \pm 1) \cdot (5 \cdot 2^{2n-1} - 2^n - 5) \textit{ times} \\ -1, & 2^{2n}(9 \cdot 2^{3n-4} - 3 \cdot 2^{2n-2} + 3 \cdot 2^{n-2} - 1) \textit{ times} \\ -1 \pm 2^{\frac{n+3}{2}}, & \frac{1}{3} \cdot 2^{2n} 2^{\frac{n-3}{2}} (2^{\frac{n-3}{2}} \pm 1) \cdot (2^{n-1} - 1)^2 \textit{ times} \end{cases} \tag{4.26}$$

Proof. (4.25) shows the distribution of exponential sums of $f(x)$ for all a, b, and c_i's in $\mathbb{F}_{2^n}$. To investigate the distribution of correlation values of sequences in $\mathcal{S}_o(2)$, we need to consider the distribution of exponential sums of $f(x)$ without $c_i = 1$ ($\delta = 0$), which means to remove the distribution of (4.24) from (4.25). Moreover, note that with respect to the correlation of sequences, a and b run through all elements in

$\mathbb{F}_{2^n}$ by θ_0 and θ_1 as well as λ_0 and λ_1, respectively. By additionally multiplying each distribution of exponential sums by 2^{2n}, therefore, we get the distribution of correlation values of (4.26). □

From Theorem 11, the linear spans of sequences in $\mathcal{S}_o(2)$ have the following distribution.

$$LS(2) = \begin{cases} \frac{n(n+1)}{2}, & (2^n-1)^2 \text{ times} \\ \frac{n(n-1)}{2}, & 2(2^n-1) \text{ times} \\ \frac{n(n-3)}{2}. & 1 \text{ time} \end{cases}$$

4.5 Example and Implementation

In this section, an example of sequences in $\mathcal{S}_o(2)$ is given for $n = 7$. Also, the implementation of sequences in $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) is presented by linear feedback shift registers (LFSRs).

4.5.1 Example of $\mathcal{S}_o(2)$

Consider a finite field $\mathbb{F}_{2^7}$ generated by a primitive element α satisfying $\alpha^7+\alpha+1 = 0$. In $\mathcal{S}_o(2)$, the kth user's sequence $\mathbf{s}^k = \{s_t^k\}$ is given by an evaluation of (4.16) at $x = \alpha^t, 0 \le t \le 126$, i.e.,

$$s_t^k = \begin{cases} Tr(\alpha^{i_0}\alpha^t + \alpha^{i_1}\alpha^{3t} + \alpha^{5t} + \alpha^{9t}), & i_0 \neq 127, i_1 \neq 127 \\ Tr(\alpha^{i_1}\alpha^{3t} + \alpha^{5t} + \alpha^{9t}), & i_0 = 127, i_1 \neq 127 \\ Tr(\alpha^{i_0}\alpha^t + \alpha^{5t} + \alpha^{9t}), & i_0 \neq 127, i_1 = 127 \\ Tr(\alpha^{5t} + \alpha^{9t}), & i_0 = i_1 = 127 \end{cases} \tag{4.27}$$

where $k = i_0 + 128 i_1$ for $0 \le i_0, i_1 \le 127$. There are 16384 cyclically distinct binary sequences in $\mathcal{S}_o(2)$.

Let $\mathbf{c} = \{c_t\}$ where $c_t = Tr(\alpha^t)$. Then, $\mathbf{c}$ is given by

$$\begin{aligned}\{c_t\} =&1000000100000110000101000111100 1\\ &0001011001110101001111101000011 1\\ &0001001001101101011011110110001 1\\ &0100101110111001100101010111111.\end{aligned} \tag{4.28}$$

Let $\mathbf{c}^{(j)} = \{c_{jt}\}$ be the j-decimation sequence from $\mathbf{c}$. From (4.27), $\{s_t^k\}$ is obtained from the linear combination of m-sequences $\mathbf{c}, \mathbf{c}^{(3)}, \mathbf{c}^{(5)}, \mathbf{c}^{(9)}$, or their shifts. In detail, for $0 \le i_0, i_1 \le 126$, $s_t^k = c_{t+i_0} + c_{3t+i_1} + c_{5t} + c_{9t}$. If either $i_0 = 127$ or $i_1 = 127$, $s_t^k = c_{3t+i_1} + c_{5t} + c_{9t}$ or $s_t^k = c_{t+i_0} + c_{5t} + c_{9t}$. If $i_0 = i_1 = 127$, $s_t^{16383} = c_{5t} + c_{9t}$.

From Theorem 12, the correlation distribution of sequences in $\mathcal{S}_o(2)$ for $n = 7$ is given by

$$C_{k,k'}(\tau) = \begin{cases} 127, & 16384 \text{ times} \\ 15, & 8026914816 \text{ times} \\ -17, & 6243155968 \text{ times} \\ -1, & 19127582720 \text{ times} \\ 31, & 433520640 \text{ times} \\ -33, & 260112384 \text{ times} \end{cases}$$

which is also verified from computer experiments.

4.5.2 LFSR Implementation

From the previous example, the process for obtaining a sequence in $\mathcal{S}_o(\rho)$ is understood by adding $(l+1)$ m-sequences with different feedback polynomials and different initial states. Hence, $\mathcal{S}_o(\rho)$ is easy to implement by summing up the LFSR outputs just like the Gold sequences.

For $n = 2l+1$, $(l+1)$ n-stage LFSRs are required to implement the sequences in $\mathcal{S}_o(\rho)$, where the LFSRs have different characteristic polynomials for generating cyclically distinct m-sequences. Specifically, let $g(x)$ be a primitive polynomial over $\mathbb{F}_2$ of degree n and α be a root of $g(x)$ in $\mathbb{F}_{2^n}$. $g_i(x)$ is a minimal polynomial of α^{1+2^i} over $\mathbb{F}_2$ with $1 \le i \le l$ (for more details about computation of minimal polynomials,

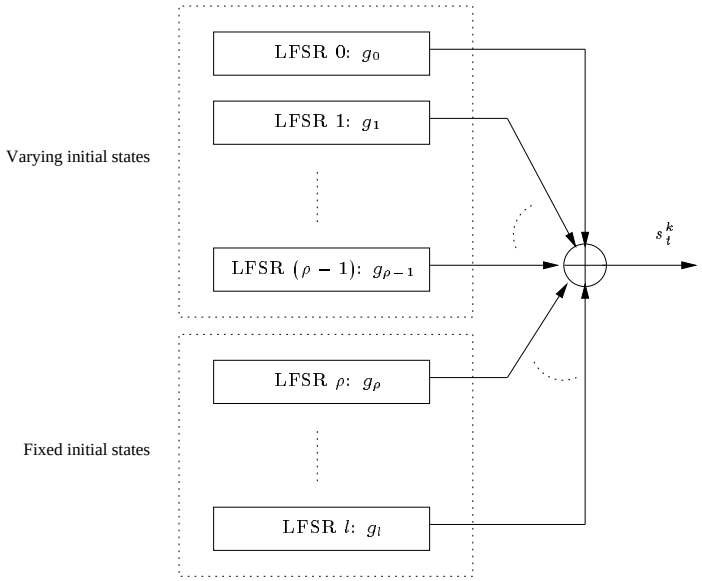


Figure 4.1: LFSR implementation of $\mathcal{S}_o(\rho)$. $k = \sum_{j=0}^{\rho-1} i_j 2^{nj}$, $0 \le i_j \le 2^n - 1$. If $i_j = 2^n - 1$, then the initial state of LFSR j is set as zero.

see [33]). Using $g_i(x)$ as the characteristic polynomial of the ith LFSR for $0 \le i \le l$ (set $g_0(x) = g(x)$), the initial states of LFSR i for $0 \le i \le \rho - 1$ can be arbitrary including zero. For $\rho \le i \le l$, on the other hand, the initial state of LFSR i is given by $\{Tr(\alpha^{(1+2^i)j})\}$ with $0 \le j \le n-1$, which is fixed for all users. The generic description of the LFSR implementation of $\mathcal{S}_o(\rho)$ is shown in Figure 4.1.

For even $n = 2l$, the sequence family $\mathcal{S}_e(\rho)$ can be implemented in the similar way to $\mathcal{S}_o(\rho)$, which is omitted here. Indeed, the implementation of $\mathcal{S}_e(\rho)$ is similar to Figure 4.1 except that the LFSR l has a size of $\frac{n}{2}$ with a fixed initial state.

Figure 4.2 shows the LFSR implementation of $\mathcal{S}_o(2)$ for $n = 7$ in the previous example. In Figure 4.2, each 7-stage LFSR generates an m-sequence. The initial states of the upper two LFSRs are differently loaded according to i_0 and i_1 for a user index $k = i_0 + 128 i_1$. On the other hand, the initial states of the lower two LFSRs are fixed for all users as shown in the figure. The initial states of LFSR 2 and LFSR 3 are given by $(c_0, c_5, c_{10}, c_{15}, c_{20}, c_{25}, c_{30}) = (1,0,0,0,0,1,0)$ and $(c_0, c_9, c_{18}, c_{27}, c_{36}, c_{45}, c_{54}) = (1,0,0,1,0,1,1)$, respectively, which can be obtained

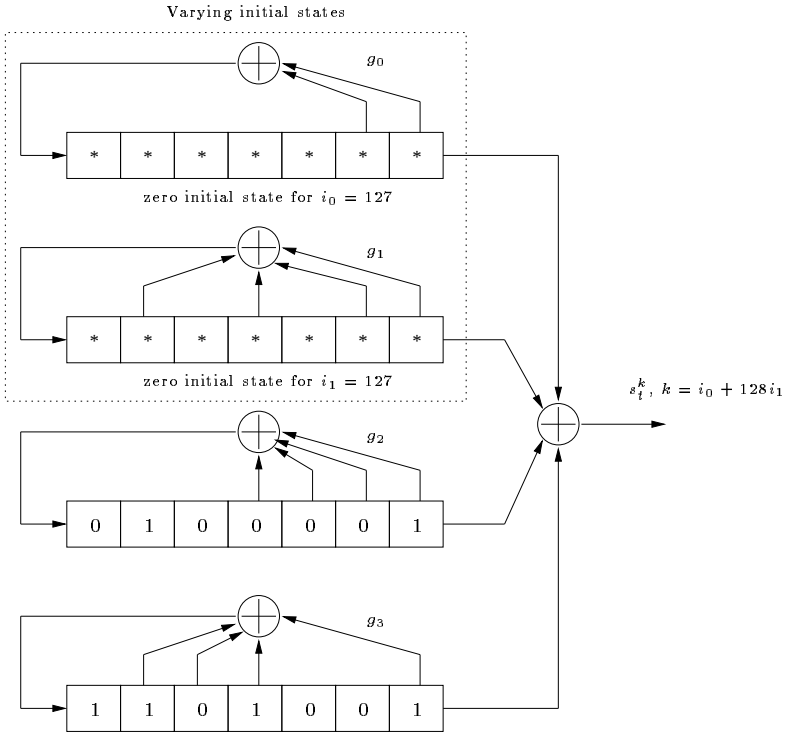


Figure 4.2: LFSR implementation of $\mathcal{S}_o(2)$ for $n = 7$ with characteristic polynomials $g_0(x) = x^7 + x + 1, g_1(x) = x^7 + x^5 + x^3 + x + 1$, $g_2(x) = x^7 + x^3 + x^2 + x + 1$, and $g_3(x) = x^7 + x^5 + x^4 + x^3 + 1$.

from (4.28). According to the different initial states of upper two LFSRs, 16384 cyclically distinct binary sequences are generated to support as many different users.

4.6 Asymptotic Optimality of Binary Sequence Families

On maximum correlation of sequences, Welch [91] presented its lower bound by investigating the even moments of the correlation. Sidelnikov [86] also derived another lower bound from the ratio of successive even moments of the correlation. In [54], Levenshtein presented a lower bound on the correlation of sequences, which is known to be the best of all known bounds [40].

With respect to maximum correlation, a few binary sequence families are indeed optimum in the sense that their maximum correlation asymptotically achieves the equality of the above lower bounds. On the asymptotic optimality of binary sequence families, it is important to consider maximum correlation and a family size simultaneously. It is because a binary sequence family with a larger family size is preferred for CDMA communication systems if the same maximum correlation is given.

In this section, the lower bounds – the Welch, Sidelnikov, and Levenshtein bounds – are first recalled on maximum correlation of a binary sequence family of size $M = L^l$ with a positive real value l, and their asymptotic behaviors are described. Then, the asymptotic parameters of maximum correlation and family sizes of several known binary sequence families including a new sequence family of this chapter are given and compared with the asymptotic bounds. From the comparison, the asymptotic optimality of binary sequence families is investigated in terms of both maximum correlation and a family size.

4.6.1 Asymptotic Lower Bounds

For asymptotic behaviors of lower bounds, we assume that a family size M of a binary sequence family is approximately a power of period L of the sequence, i.e., $M = L^l$ where l is a positive real value. Then, the asymptotic lower bounds – the Welch, Sidelnikov, and Levenshtein bounds – are presented in terms of a variable $l = \log_L M$. The asymptotic bounds play a role of measures to compare binary sequence families in terms of both maximum correlation and a family size. For the asymptotic representation, L is assumed to be sufficiently large.

The Welch bound: Let $C_{\max}$ be the maximum correlation of a binary sequence family. For a positive integer k, the Welch bound [91] gives an inequality of $C_{\max}^{2k}$, i.e.,

$$C_{\max}^{2k} \geq \frac{1}{ML-1}\left\{\frac{ML^{2k+1}}{\binom{k+L-1}{k}} - L^{2k}\right\}.$$

Applying $M = L^l$, we have

$$\begin{aligned} C_{\max}^{2k} &> \frac{L^{2k}}{\binom{k+L-1}{k}} - \frac{L^{2k-1}}{M} = \frac{L^{2k}}{\prod_{i=0}^{k-1}\left(\frac{L+k-1-i}{k-i}\right)} - L^{2k-l-1} \\ &> \frac{k!L^{2k}}{(L+k-1)^k} - L^{2k-l-1} \\ &= k!\left(\frac{L^2}{L+k-1}\right)^k - L^{2k-l-1} \\ &\approx k!L^k - L^{2k-l-1} \end{aligned} \tag{4.29}$$

where we assume $k \ll L$. If $k > l+1$, the right-hand side of (4.29) may be negative, which makes the bound trivial. Since k is a positive integer, let $k = \lfloor l+1 \rfloor$. Then,

$$C_{\max}^{2k}/L^k > (\lfloor l+1 \rfloor)! - L^{\lfloor l+1 \rfloor - (l+1)}. \tag{4.30}$$

In order to remove the dependency on L of the right-hand side of (4.30), consider t such that

$$L^x < (x+1)^t = T(x,t), \quad -1 < x \leq 0 \tag{4.31}$$

for a sufficiently large L. Figure 4.3 shows $y = L^x$ and $y = T(x,t) = (x+1)^t$ for various L's and $t = 6$. From Figure 4.3, $t = 6$ can be chosen for (4.31) if $L > 10^3$.

If $x = \lfloor l+1 \rfloor - (l+1)$, then $-1 < x \leq 0$, and from (4.30) and (4.31),

$$L^{\lfloor l+1 \rfloor - (l+1)} < (\lfloor l+1 \rfloor - l)^t$$

and

$$\frac{C_{\max}}{\sqrt{L}} > \left((\lfloor \log_L M + 1 \rfloor)! - (\lfloor \log_L M + 1 \rfloor - \log_L M)^t\right)^{\frac{1}{2\lfloor \log_L M + 1 \rfloor}}. \tag{4.32}$$

From Figure 4.3, it is clear that a larger t can be chosen for more accurate asymptotic Welch bound of (4.32) for a larger L. As a matter of fact, Welch stated the asymptotic behavior of his bound for a positive integer $\log_L M$ in [91]. However, (4.32) is more general in the sense that it contains a case of a positive *real* value

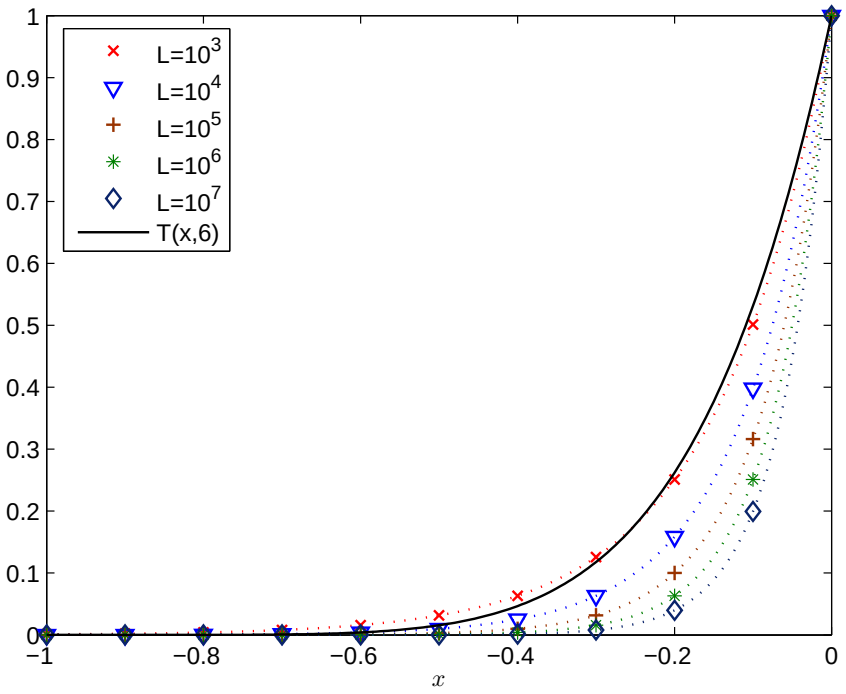


Figure 4.3: $y = L^x$ and $y = T(x, 6) = (x + 1)^6$

$l = \log_L M$.

Figure 4.4 shows the original and asymptotic Welch bounds on the normalized maximum correlation for $L = 2^n$ and $t = 6$. Note that $k = \lfloor l + 1 \rfloor$ is applied for the original Welch bound as well as its asymptotic bound. From Figure 4.4, the asymptotic bound is quite close to the original Welch bound for sufficiently large L's.

The Sidelnikov bound: In the Sidelnikov bound [86], $C^2_{\max}$ has the following

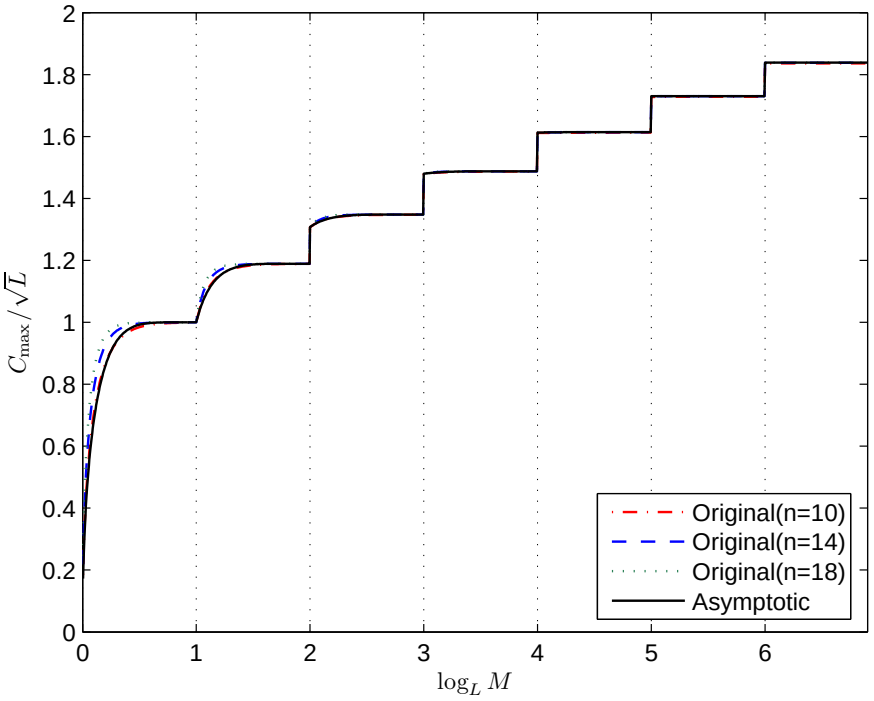


Figure 4.4: The Welch bound on normalized maximum correlation ($t = 6$)

inequality according to an integer k, $0 \leq k < 2L/5$.

$$\begin{aligned} C_{\max}^2 &\geq (2k+1)(L-k) + \frac{k(k+1)}{2} - \frac{2^k L^{2k+2}}{ML(2k)!\binom{L}{k}} \\ &= (2k+1)L - \frac{k(3k+1)}{2} - \frac{2^k L^{2k+2}}{ML(2k)!\binom{L}{k}}. \end{aligned} \tag{4.33}$$

Assume k is set to be sufficiently small such that $k \ll L$. Then, $\frac{k(3k+1)}{2}$ is negligible

in the right-hand side of (4.33). With $M = L^l$,

$$\begin{aligned} C_{\max}^2 &\geq (2k+1)L - \frac{k!2^k L^{2k+2}}{ML(2k)!\prod_{i=0}^{k-1}(L-i)} \\ &> (2k+1)L - \frac{k!2^k}{(2k)!} \cdot \frac{L^{2k-l+1}}{(L-k+1)^k} \\ &= (2k+1)L - \frac{2^k}{\prod_{i=0}^{k-1}(2k-i)} \cdot \frac{L^{2k-l+1}}{(L-k+1)^k} \\ &> (2k+1)L - \frac{2^k}{(k+1)^k} \cdot \frac{L^{2k-l+1}}{(L-k+1)^k} \\ &= (2k+1)L - \left(\frac{2L}{(k+1)(L-k+1)}\right)^k \cdot L^{k-l+1} \\ &\approx (2k+1)L - \left(\frac{2}{k+1}\right)^k L^{k-l+1}. \end{aligned} \tag{4.34}$$

The last approximation is from $L \gg k$. If $k - l > 0$, the right-hand side of (4.34) may be negative, and then the bound becomes trivial. Since k is a positive integer, let $k = \lfloor l \rfloor$. Then,

$$C_{\max}^2/L > 2\lfloor l \rfloor + 1 - \left(\frac{2}{\lfloor l \rfloor + 1}\right)^{\lfloor l \rfloor} L^{\lfloor l \rfloor - l}. \tag{4.35}$$

Similar to the asymptotic Welch bound, (4.31) is applied to (4.35), and then

$$L^{\lfloor l \rfloor - l} < (\lfloor l \rfloor - l + 1)^t$$

for sufficiently large L's, and finally

$$\frac{C_{\max}}{\sqrt{L}} > \sqrt{2\lfloor \log_L M \rfloor + 1 - \left(\frac{2}{\lfloor \log_L M \rfloor + 1}\right)^{\lfloor \log_L M \rfloor} (\lfloor \log_L M \rfloor - \log_L M + 1)^t}. \tag{4.36}$$

From Figure 4.3, it is also clear that a larger t can be chosen for more accurate asymptotic Sidelnikov bound of (4.36) for a larger L.

Figure 4.5 shows the original and asymptotic Sidelnikov bounds on the normal-

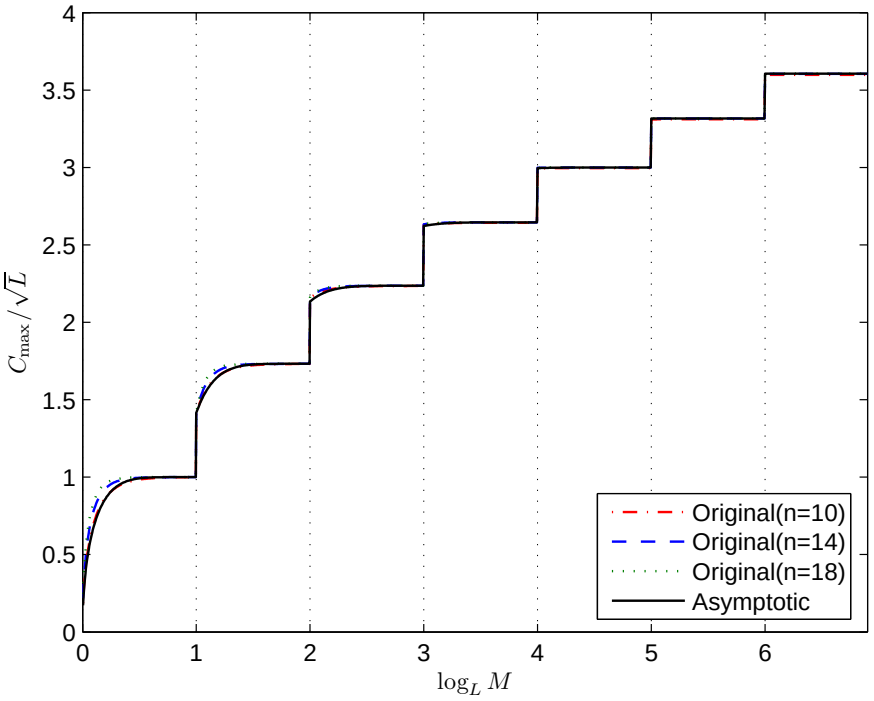


Figure 4.5: The Sidelnikov bound on normalized maximum correlation ($t = 6$)

ized maximum correlation for $L = 2^n$ and $t = 6$. Note that $k = \lfloor l \rfloor$ is applied for the original Sidelnikov bound as well as its asymptotic bound. Similar to the asymptotic Welch bound, we have the good asymptotic Sidelnikov bound for sufficiently large L's.

The Levenshtein bound: For $M = L^l$ with a positive integer l, Levenshtein presented the following asymptotic bound [55].

$$C_{\max}^2/L > c(l)$$

where $c(l)$ is a constant of Table 1 in [55]. To derive its asymptotic bound for a positive real value l, it should be noted that the Levenshtein bound coincides with the Welch bound for $0 \leq l < 1$, and also coincides with the Sidelnikov bound for

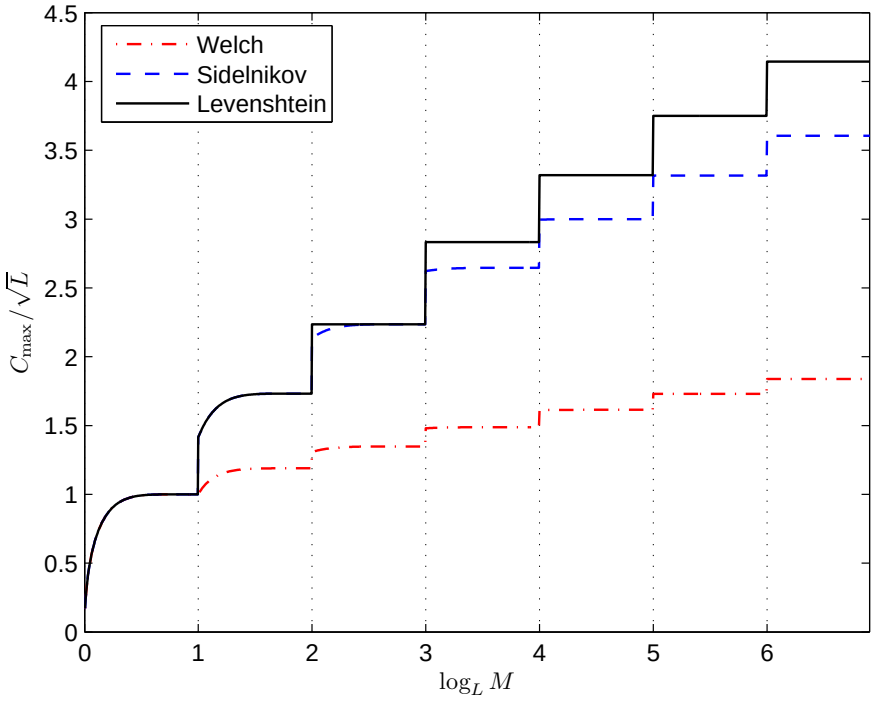


Figure 4.6: Asymptotic bounds on normalized maximum correlation

$1 \leq l < 2$ [55] [40]. For $0 \leq l < 2$, therefore, we may use the asymptotic Welch and Sidelnikov bounds given by (4.32) and (4.36), respectively. If $l \geq 2$, on the other hand, we choose simply $C_{\max}^2/L > c(\lfloor l \rfloor)$. In fact, it looks tricky but it is always true because the original Levenshtein bound is an increasing function over l [55]. Then, the asymptotic Levenshtein bound is given by

$$C_{\max}/\sqrt{L} > \begin{cases} \text{right-hand side of (4.32)}, & \text{if } 0 \leq \log_L M < 1 \\ \text{right-hand side of (4.36)}, & \text{if } 1 \leq \log_L M < 2 \\ \sqrt{c(\lfloor \log_L M \rfloor)}, & \text{if } \log_L M \geq 2 \end{cases} \tag{4.37}$$

Figure 4.6 shows the asymptotic bounds of the Welch, Sidelnikov, and Levenshtein bounds. In Figure 4.6, $t = 6$ for (4.32) and (4.36). Asymptotically, the Levenshtein bound is also the best of all the three bounds.

4.6.2 Asymptotic Parameters of Binary Sequence Families

This section lists several binary sequence families by giving their asymptotic parameters – periods, family sizes, and maximum correlation. First, two binary sequence families $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) and $\mathcal{Z}(n-1, D)$ [84] are recalled and their asymptotic parameters are discussed in detail. Then, the asymptotic parameters of other well known binary sequence families are also given.

$\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$: According to a positive integer ρ, $1 \le \rho \le \lfloor \frac{n-1}{2} \rfloor$, the families have the following parameters.

$$L = 2^n - 1, \quad M = 2^{n\rho}, \quad C_{\max} = \begin{cases} 1 + 2^{\frac{n+2\rho-1}{2}} & \text{for } \mathcal{S}_o(\rho) \\ 1 + 2^{\frac{n}{2}+\rho} & \text{for } \mathcal{S}_e(\rho) \end{cases}$$

For their asymptotic behaviors, let $L \approx 2^n$. Then, $\log_L M = \rho$ and $C_{\max}$ is asymptotically given by

$$C_{\max} \approx \begin{cases} 2^{\rho-\frac{1}{2}}\sqrt{L} & \text{for } \mathcal{S}_o(\rho) \\ 2^{\rho}\sqrt{L} & \text{for } \mathcal{S}_e(\rho) \end{cases}$$

Z_4-linear binary sequence family: The Z_4-linear binary sequence family, denoted by $\mathcal{Z}(n-1, D)$, was introduced by Shanbhag, Kumar, and Helleseth [84], where the notation is slightly changed for period $L = 2^n - 2$. For a positive integer $D \ge 2$, it has the following parameters [84].

$$L = 2^n - 2, \quad M \ge \left(\frac{L}{2}\right)^{(D-\lfloor \frac{D}{4} \rfloor - 1)}, \quad C_{\max} \le \begin{cases} (D-1)\sqrt{2L}, & \text{for odd } n \\ (D-1)\sqrt{L}, & \text{for even } n \end{cases} \tag{4.38}$$

(In the original construction in [84], the family size was a half of M. In [40], however, it is pointed out that the size can be doubled.) In this chapter, we assume that the inequalities in (4.38) can be replaced by equalities. From $M = \left(\frac{L}{2}\right)^{(D-\lfloor \frac{D}{4} \rfloor - 1)}$, we have

$$\log_L M = \left(D - \left\lfloor \frac{D}{4} \right\rfloor - 1\right)(1 - \log_L 2) = \left(D - \left\lfloor \frac{D}{4} \right\rfloor - 1\right)\left(\frac{n-1}{n}\right)$$

where $L \approx 2^n$. Thus,

$$D - \left\lfloor \frac{D}{4} \right\rfloor - 1 = \frac{n}{n-1} \log_L M \triangleq \eta. \tag{4.39}$$

Since D is a positive integer, note that η is an integer while $\log_L M$ is a real value for $n > 2$. From (4.39),

$$D = \eta + 1 + \left\lfloor \frac{\eta}{3} \right\rfloor.$$

(In (4.39), two different D's may exist for the same η, where the smaller one is chosen for D [84].) Finally, the asymptotic parameters $\log_L M$ and $C_{\max}$ are represented by an integer η, i.e.,

$$\log_L M = \left(\frac{n-1}{n}\right)\eta, \quad C_{\max} = \begin{cases} \left(\eta + \left\lfloor \frac{\eta}{3} \right\rfloor\right)\sqrt{2L}, & \text{for odd } n \\ \left(\eta + \left\lfloor \frac{\eta}{3} \right\rfloor\right)\sqrt{L}, & \text{for even } n. \end{cases}$$

For simplicity, we use notations $\mathcal{Z}_o(D)$ for odd n and $\mathcal{Z}_e(D)$ for even n to denote the Z_4-linear binary sequence family.

For the asymptotic period $L = 2^n$ of a binary sequence, Table 4.5 shows the asymptotic parameters $\log_L M$ and $C_{\max}$ for several well known binary sequence families. In Table 4.5, the family sizes of the Rothaus [79] and the Kasami (large set) sequences [46] are asymptotically assumed as $M = 2^{2n}$ and $2^{\frac{3n}{2}}$, respectively. The Gold-like [11] and the Udaya's [89] sequences are considered as subsequence classes of $\mathcal{S}_o(1)$ and $\mathcal{S}_e(1)$, respectively.

4.6.3 Asymptotic Optimality of Binary Sequence Families

This section compares the asymptotic parameters of normalized maximum correlation and family sizes of binary sequence families. As references, the asymptotic bounds derived in Section 4.6.1 are shown together. Then the asymptotic optimality of the families is discussed. It is shown that $\mathcal{S}_o(\rho)$ of $\rho \leq 2$ and $\mathcal{Z}_e(D)$ are the best known ones in terms of both maximum correlation and family sizes.

Figure 4.7 shows the normalized maximum correlation of several binary sequence families and its asymptotic bounds ($t = 6$). Since $\mathcal{S}_e(\rho)$ and $\mathcal{Z}_o(D)$ have relatively

Table 4.5: Asymptotic parameters of several binary sequence families with given period L

Sequence	$\log_L M$	$C_{\max}$	n
Gold [31]	1	$\sqrt{2}\sqrt{L}$	odd
Rothaus [79]	2	$2\sqrt{2}\sqrt{L}$	odd
Kasami (small set) [46]	1/2	$\sqrt{L}$	even
Kasami (large set) [46]	3/2	$2\sqrt{L}$	even
Bent [71]	1/2	$\sqrt{L}$	even, multiple of 4
$\mathcal{S}_o(\rho)$	ρ	$2^{\rho-\frac{1}{2}}\sqrt{L}$	odd
$\mathcal{S}_e(\rho)$	ρ	$2^{\rho}\sqrt{L}$	even
$\mathcal{Z}_o(D)$	$\left(\frac{n-1}{n}\right)\eta$	$\left(\eta+\lfloor\frac{\eta}{3}\rfloor\right)\sqrt{2}\sqrt{L}$	odd
$\mathcal{Z}_e(D)$	$\left(\frac{n-1}{n}\right)\eta$	$\left(\eta+\lfloor\frac{\eta}{3}\rfloor\right)\sqrt{L}$	even

worse parameters than $\mathcal{S}_o(\rho)$ and $\mathcal{Z}_e(D)$, respectively, only the latter families are shown. Note that $\mathcal{S}_o(\rho)$ and $\mathcal{Z}_e(D)$ contain their own subsequence classes according to ρ and D, respectively.

In Figure 4.7, the Gold-like sequences (or $\mathcal{S}_o(1)$) asymptotically achieve the Sidelnikov and the Levenshtein bounds. On the other hand, $\mathcal{S}_o(2)$ does not achieve the asymptotic bounds, but provide a larger family size for all n's than any other known binary sequence families with the given normalized maximum correlation. The Gold and the Rothaus sequences also present the same parameters as $\mathcal{S}_o(1)$ and $\mathcal{S}_o(2)$, respectively, but they provide much smaller linear spans. The Kasami (small set) and the bent sequences asymptotically achieve the Welch, Sidelnikov, and Levenshtein bounds. However, their family sizes are rather small ($\sqrt{L}$). $\mathcal{Z}_e(2)$, on the other hand, not only achieves all the asymptotic bounds, but also provides the family size as large as possible for sufficiently large n. Therefore, $\mathcal{Z}_e(2)$ or the *Kerdock* sequences [40] seem to be the best known binary sequences achieving all the three asymptotic bounds for sufficiently large n. $\mathcal{Z}_e(3)$ or the *Delsarte-Goethals* sequences [40] does not achieve any bounds on the normalized maximum correlation, but the family size is larger than any other known binary sequence families including the Kasami (large set) sequences. If $D \geq 5$, then $\mathcal{Z}_e(D)$ gets far from the asymptotic bounds.

In conclusion, $\mathcal{S}_o(\rho)$'s of $\rho = 1$ and 2 provide the largest family sizes and large linear spans for all n's with the given normalized maximum correlation. For

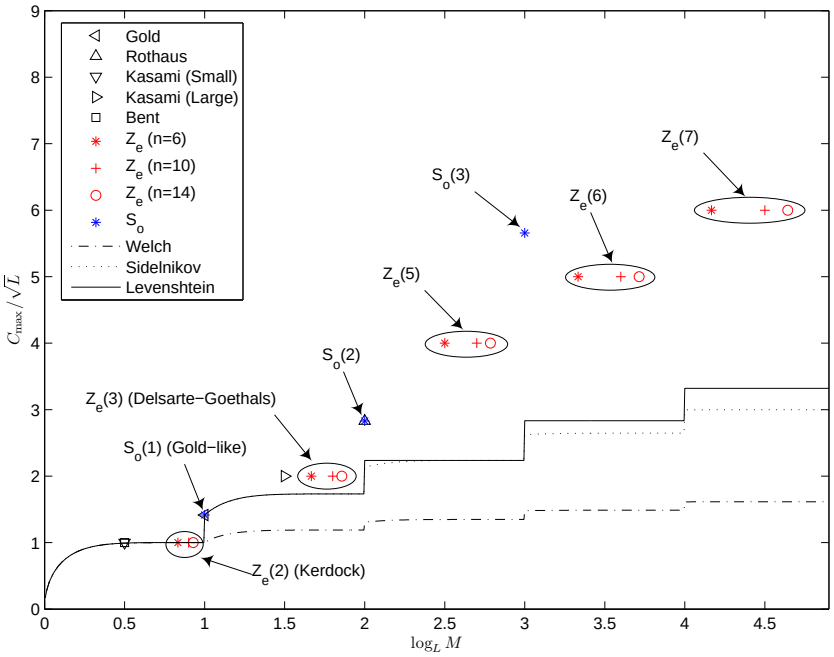


Figure 4.7: Asymptotic normalized maximum correlation and the bounds

$\log_L M = 1$, in particular, the Gold-like sequences (or $\mathcal{S}_o(1)$) even achieve the asymptotic Sidelnikov and Levenshtein bounds. For sufficiently large n's, on the other hand, $\mathcal{Z}_e(D)$ is also the best known binary sequence family for all regions of $\log_L M$ in terms of maximum correlation and family sizes. In particular, the Kerdock sequences give the largest family size for sufficiently large n's asymptotically achieving all the three bounds.

4.7 Conclusion

A new family of binary sequences of period $2^n - 1$ has been presented for both odd and even n in this chapter. For odd $n = 2l + 1$ and an integer ρ with $1 \leq \rho \leq l$, a new family $\mathcal{S}_o(\rho)$ has a family size of $2^{n\rho}$ and maximum correlation of $1 + 2^{\frac{n+2\rho-1}{2}}$. For even $n = 2l$ and an integer ρ with $1 \leq \rho < l$, on the other hand, a new family

$\mathcal{S}_e(\rho)$ has a family size of $2^{n\rho}$ and maximum correlation of $1+2^{\frac{n}{2}+\rho}$. Maximum and minimum linear spans of both $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ are $\frac{n(n+1)}{2}$ and $\frac{n(n-2\rho+1)}{2}$, respectively. From the flexibility due to ρ, the new sequence family has the adaptive family size and maximum correlation. The large linear spans of $\mathcal{S}_o(\rho)$ and $\mathcal{S}_e(\rho)$ imply that they have a good potential cryptographic property.

The new family contains the Boztas and Kumar's case corresponding to $\mathcal{S}_o(1)$ (or Udaya's case corresponding to $\mathcal{S}_e(1)$) as a subset if m-sequences are excluded from both cases. The family $\mathcal{S}_o(2)$ is considered as a good candidate with a large family size as well as low correlation. Thus, the correlation distribution of the sequences in $\mathcal{S}_o(2)$ has been further derived. At the end, the LFSR implementation of $\mathcal{S}_o(\rho)$ (or $\mathcal{S}_e(\rho)$) has been presented.

In addition, the asymptotic Welch, Sidelnikov, and Levenshtein bounds have been studied for a binary sequence family of size $M = L^l$ with a positive real value l. The bounds have been compared with the asymptotic parameters of several known binary sequence families. From the comparison, the asymptotic optimality of the families have been discussed. In conclusion, it is shown that $\mathcal{S}_o(\rho)$, $\rho \leq 2$ and $Z_e(D)$ are the best known binary sequence families in terms of both maximum correlation and family sizes.

Chapter 5

Quadratic Bent Functions of A Polynomial Form

This chapter presents the constructions and enumerations of all bent functions represented by a polynomial form of $f(x) = \sum_{i=1}^{n/2-1} c_i Tr(x^{1+2^i}) + c_{n/2} Tr_1^{n/2}(x^{1+2^{n/2}}), c_i \in \mathbb{F}_2$, by deriving necessary and sufficient conditions on c_i's for special cases of n [92]. The related works for bent functions are first reviewed and the motivation of this work is given. Then, the cyclotomic polynomials are introduced, which plays a crucial role in deriving the necessary and sufficient conditions. To describe a characterization of bent functions of a polynomial form, a criterion for bent functions with quadratic exponents is reviewed. As main results of this chapter, the constructions and enumerations of quadratic bent functions of a polynomial form are presented for special values of n. In addition, the implementation of bent sequences employing the constructed quadratic bent function is considered, and the exact linear complexity of the bent sequences is computed [98].

5.1 Related Works and Motivation

A bent function is a Boolean function with even number of variables whose Walsh transform has a constant magnitude [33], where the Walsh transform is defined in Section 1.5. In coding context, it is a coset of the first order Reed-Muller code with

the largest minimum Hamming weight [60]. In other words, a bent function has a maximum distance from a linear function, so it is called *maximally nonlinear*. Bent functions with the maximum nonlinearity have been paid much attention to by researchers for cryptographic applications [18] [13] [14] [16] and [15]. Moreover, the maximum nonlinearity of bent functions corresponds to minimized maximum correlation between the function and a trace function. Thus, bent functions also have many applications in algebraic coding and sequence design [60] [71].

In [48] and [47], Khoo, Gong, and Stinson investigated the following sum of monomial trace terms with quadratic exponents where each exponent of variables has Hamming weight 2. For odd n,

$$f(x) = \sum_{i=1}^{\frac{n-1}{2}} c_i Tr(x^{1+2^i}), \quad c_i \in \mathbb{F}_2, \ x \in \mathbb{F}_{2^n}$$

where $Tr(x)$ is a trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. Recall the definition of the Hadamard transform described in Section 1.5. If the Hadamard transform of $f(x)$ only takes the three values of $\{0, \pm 2^{\frac{n+1}{2}}\}$, then $f(x)$ is called a *semi-bent function* for odd n [48]. Let $c(x) = \sum_{i=1}^{\frac{n-1}{2}} c_i(x^i + x^{n-i})$. Khoo, Gong, and Stinson derived a necessary and sufficient condition for a semi-bent function, i.e., $f(x)$ is semi-bent if and only if

$$\gcd(c(x), x^n + 1) = x + 1.$$

Following this work, Charpin, Pasalic, and Tavernier [18] considered

$$f(x) = \begin{cases} \sum_{i=1}^{\frac{n-1}{2}} c_i Tr(x^{1+2^i}), & c_i \in \mathbb{F}_2, \text{ odd } n, \\ \sum_{i=1}^{\frac{n-2}{2}} c_i Tr(x^{1+2^i}), & c_i \in \mathbb{F}_2, \text{ even } n. \end{cases} \tag{5.1}$$

For even n, $f(x)$ is called a semi-bent function in [18] if the Hadamard transform of $f(x)$ belongs to $\{0, \pm 2^{\frac{n+2}{2}}\}$. They showed that for even n, $f(x)$ is semi-bent if and only if $\gcd(c(x), x^n + 1) = x^2 + 1$. For odd n, on the other hand, they derived some conditions where $f(x)$ with three or four trace terms is semi-bent. Then, they derived the construction of semi-bent functions of odd n with higher degree from

semi-bent functions of even n in (5.1), and also derived the construction of bent functions with higher degree from semi-bent functions of odd n.

Applying the techniques developed in [48], Ma, Lee, and Zhang [59] showed that a necessary condition for the bent functions with such a representation is as follows.

$$f(x) = \sum_{i=1}^{n/2-1} c_i Tr(x^{1+2^i}) + Tr_1^{n/2}(x^{1+2^{n/2}}), \quad c_i \in \mathbb{F}_2,\ x \in \mathbb{F}_{2^n} \tag{5.2}$$

where n is even and $Tr_1^{n/2}(x)$ is a trace function from $\mathbb{F}_{2^{n/2}}$ to $\mathbb{F}_2$. Equivalently, the monomial trace term $Tr_1^{n/2}(x^{1+2^{n/2}})$ has to be presented in the representation of a bent function $f(x)$. Furthermore, a necessary and sufficient condition for $f(x)$ given by (5.2) to be bent is $\gcd(c(x), x^n + 1) = 1$ [59] where

$$c(x) = \sum_{i=1}^{n/2-1} c_i(x^i + x^{n-i}) + x^{n/2}. \tag{5.3}$$

For the quadratic bent functions represented by a polynomial form (5.2), the known cases are: (a) all c_i's are zero corresponding to the Kasami (small) signal set [46], or all c_i's are one giving the Udaya's construction [89], or ones of c_i's are distributed by equal distance of i's giving the Kim and No's signal set [49], respectively; (b) all choices of c_i's for $n = 2^v$ [59]. For the polynomial construction of non-quadratic bent functions, on the other hand, the known cases are monomial trace functions with the Kasami [23] and the Dillon exponents [21], and a sum of trace functions with the Niho exponents [27].

It is worthwhile to point out that all quadratic bent functions of a Boolean form are known, which can be obtained by applying an affine transform to $x_0, x_1, \cdots, x_{n-1}$ in $\sum_{i=0}^{n/2-1} x_{2i}x_{2i+1}$ [60]. The *Maiorana-McFarland's* construction [64] for quadratic bent functions belongs to this class. However, the construction of quadratic bent functions of a polynomial form is not a trivial work. Therefore, this chapter studies how to construct quadratic bent functions of a polynomial form. Precisely, the construction of all the quadratic bent functions represented by a polynomial form (5.2) is presented by necessary and sufficient conditions on c_i's for special cases of

n.

5.2 Cyclotomic Polynomials

A polynomial whose roots are the field elements of order d is called the *dth cyclotomic polynomial* [56], denoted by $Q_d(x)$. $Q_d(x)$ is a monic polynomial of order d and degree $\phi(d)$, where $\phi(d)$ is the Euler-totient function, defined as the number of integers k of $1 \leq k \leq d$ with $\gcd(k, d) = 1$ [56]. $Q_d(x)$ has the following basic properties [10] [63].

Property 3 *Let $Q_d(x)$ be the dth cyclotomic polynomial.*

(a) $x^m - 1 = \prod_{d|m} Q_d(x)$. *In particular,* $x^{p^r} - 1 = Q_1(x)Q_p(x)\cdots Q_{p^{r-1}}(x)Q_{p^r}(x)$ *where p is prime.*

(b) *For $d \geq 2$,* $x^{\phi(d)}Q_d(x^{-1}) = Q_d(x)$. *In other words, $Q_d(x)$ is self-reciprocal. (A polynomial $g(x)$ with degree m is called* self-reciprocal *if* $x^m g(x^{-1}) = g(x)$.)

(c) *For prime p,* $Q_p(x) = \sum_{i=1}^{p} x^{p-i} = x^{p-1} + x^{p-2} + \cdots + x + 1$.

(d) *For prime p,* $Q_{p^k}(x) = Q_p(x^{p^{k-1}})$.

The cyclotomic polynomial is irreducible over the integer ring $\mathbb{Z}$, but it may not be irreducible over $\mathbb{F}_2$. Throughout this chapter, the cyclotomic polynomials are considered over $\mathbb{F}_2$. Several useful properties on the factorization of $Q_d(x)$ over $\mathbb{F}_2$ are listed without proofs. For more details, see [10], [56], and [63].

Property 4 *$Q_d(x)$ is irreducible over $\mathbb{F}_2$ if and only if* $\mathrm{ord}_d(2)$ *is $\phi(d)$.*

Property 5 *For prime p, let $Q_p(x) = g_1(x)g_2(x)\cdots g_t(x)$ where $g_i(x)$ for $1 \leq i \leq t$ is irreducible over $\mathbb{F}_2$. Then, the degree and the order of each $g_i(x)$ for $1 \leq i \leq t$ are given by $\frac{\phi(p)}{t} = \frac{p-1}{t}$ and p, respectively.*

Property 6 *For prime p, let $f_1(x), f_2(x), \cdots$, and $f_t(x)$ be distinct monic irreducible polynomials over $\mathbb{F}_2$ of degree $\frac{\phi(p)}{t}$ and order p, and let $s = p^{k-1}$. Then, $f_1(x^s), f_2(x^s), \cdots$, and $f_t(x^s)$ are distinct monic irreducible polynomials over $\mathbb{F}_2$ of degree $\frac{\phi(p)p^{k-1}}{t}$ and order p^k.*

5.3 A Criterion for Bent Functions with Quadratic Exponents

For odd n, Khoo, Gong, and Stinson showed a necessary and sufficient condition for a semi-bent function with quadratic exponents [48] [47]. Similarly, a necessary and sufficient condition for a bent function with quadratic exponents can be directly resulted from the techniques developed in [48] and [47]. The following fact appears in [59].

Fact 5 *For even n, let*

$$f(x) = \sum_{i=1}^{n/2-1} c_i Tr(x^{1+2^i}) + c_{n/2} Tr_1^{n/2}(x^{1+2^{n/2}}), \quad c_i \in \mathbb{F}_2, \ x \in \mathbb{F}_{2^n}.$$

Then, $f(x)$ is bent if and only if $\gcd(c(x), x^n + 1) = 1$, where

$$c(x) = \sum_{i=1}^{n/2-1} c_i(x^i + x^{n-i}) + c_{n/2}x^{n/2}. \tag{5.4}$$

In addition, $c_{n/2} = 1$ if $f(x)$ is bent.

Corollary 3 *For $n = 2^v m$ with $v \geq 1$ and odd m, $f(x)$ given by (5.2) is bent if and only if $\gcd(c(x), x^m + 1) = 1$.*

Proof. From $x^n + 1 = (x^m + 1)^{2^v}$, $\gcd(c(x), x^n + 1) = 1$ if and only if $\gcd(c(x), x^m + 1) = 1$. Hence, Corollary 3 is true. □

For $n = 2^v$ with $v \geq 2$, it is immediate from Fact 5 and Corollary 3 that $f(x)$ given by (5.2) is a bent function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ for any choices of c_i's [59]. Thus,

the total number of such bent functions for $n = 2^v$ is equal to $2^{\frac{n}{2}-1}$.

5.4 Construction and Enumeration for $n = 2^v p$

In this section, we construct and enumerate all bent functions given by

$$f(x) = \sum_{i=1}^{n/2-1} c_i Tr(x^{1+2^i}) + Tr_1^{n/2}(x^{1+2^{n/2}}), \quad c_i \in \mathbb{F}_2 \tag{5.5}$$

for $n = 2^v p$ with $v \geq 1$, where p is odd prime with $\mathrm{ord}_p(2) = p-1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ with odd s.

Before the construction and enumeration are presented, some preparations are needed on the greatest common divisor of $c(x)$ given by (5.3) and the pth cyclotomic polynomial $Q_p(x)$. In the following, it is implied that a root of a polynomial over $\mathbb{F}_2$ belongs to some extension field of $\mathbb{F}_2$.

Lemma 19 *Let $\bar{c}(x)$ be $c(x)$ reduced modulo x^p+1, and z be a root of x^p+1 with $z \neq 1$. Then,*

(a)

$$\begin{aligned} \bar{c}(x) \equiv c(x) \pmod{x^p+1} &= \sum_{i=1}^{p-1} w_i(x^i + x^{p-i}) + 1 \\ &= \sum_{i=1}^{\frac{p-1}{2}} (w_i + w_{p-i})(x^i + x^{p-i}) + 1 \end{aligned} \tag{5.6}$$

where $w_i = c_i + c_{i+p} + \cdots + c_{i+\frac{n}{2}-p} = \sum_{l=0}^{\frac{n}{2p}-1} c_{i+lp}$.

(b) Let p be odd prime with $\mathrm{ord}_p(2) = p-1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ where s is odd. Then, $\gcd(c(x), x^p+1) \neq 1$ if and only if $\bar{c}(x) = Q_p(x)$ which is equivalent to

$$w_i \neq w_{p-i} \text{ for all } 1 \leq i \leq \frac{p-1}{2}. \tag{5.7}$$

Proof. (a) $c(x)$ can be rewritten as

$$c(x) = \sum_{i=1}^{p-1} c_i(x^i + x^{n-i}) + c_p(x^p + x^{n-p}) + \sum_{i=p+1}^{2p-1} c_i(x^i + x^{n-i}) + c_{2p}(x^{2p} + x^{n-2p}) + \cdots + \sum_{i=(\frac{n}{2p}-1)p+1}^{\frac{n}{2}-1} c_i(x^i + x^{n-i}) + x^{\frac{n}{2}}. \tag{5.8}$$

Hence, (5.6) is obtained from (5.8) modulo $x^p + 1$.

(b) $x^p + 1 = Q_1(x)Q_p(x)$ for a prime p, where $Q_1(x) = x + 1$. Obviously, $\gcd(c(x), Q_1(x)) = 1$, and $Q_p(x) = \sum_{i=1}^{\frac{p-1}{2}}(x^i + x^{p-i}) + 1$.

If $\mathrm{ord}_p(2) = p - 1$, then $Q_p(x)$ is irreducible over $\mathbb{F}_2$. Hence,

$$\begin{aligned} \gcd(c(x), x^p + 1) \neq 1 &\Longleftrightarrow \gcd(c(x), x^p + 1) = Q_p(x) \\ &\Longleftrightarrow \bar{c}(x) = Q_p(x). \end{aligned} \tag{5.9}$$

The last equivalence is from the Euclidean algorithm. By comparing $\bar{c}(x)$ and $Q_p(x)$, (5.7) is obtained.

If $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ where s is odd, on the other hand, $Q_p(x) = g_1(x)g_2(x)$ where $g_1(x)$ and $g_2(x)$ are irreducible over $\mathbb{F}_2$ with $g_2(x) = x^s g_1(x^{-1})$ [18]. Thus, if z is a root of $g_1(x)$, then z^{-1} is a root of $g_2(x)$, and vice versa. (Note that if z is a root of $g_1(x)$, then z^{-1} cannot be its root [18].) Since z is a root of $Q_p(x)$, the order of z is p, i.e., $z^p = z^n = 1$. From (5.3), therefore,

$$c(z) = \sum_{i=1}^{n/2-1} c_i(z^i + z^{n-i}) + z^{n/2} = \sum_{i=1}^{n/2-1} c_i(z^i + z^{-i}) + z^{-n/2}$$

and thus z^{-1} is a root of $c(x)$ if z is its root. If $c(x)$ has the irreducible factor $g_1(x)$, therefore, it simultaneously has the other irreducible factor $g_2(x)$, and vice versa. Hence, (5.9) is also true in this case. Similar to the case of $\mathrm{ord}_p(2) = p - 1$, (5.7) is obtained. □

Theorem 13 *Let $n = 2^v p$ with $v \geq 1$ and p be odd prime with $\mathrm{ord}_p(2) = p - 1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ where s is odd. Then, $f(x)$ given by (5.5) is bent if and only if there exists at least one i for $1 \leq i \leq \frac{p-1}{2}$ such that*

$$w_i = w_{p-i} \tag{5.10}$$

where $w_i = \sum_{l=0}^{\frac{n}{2p}-1} c_{i+lp}$. The number of bent functions of $f(x)$, denoted by N_b, is given by

$$N_b = 2^{\frac{n}{2}-1} - 2^{\frac{n-1-p}{2}}.$$

Note that $N_{nb} = 2^{\frac{n-1-p}{2}}$ is the number of non-bent functions of $f(x)$.

Proof. From Fact 5 and Corollary 3, $f(x)$ is non-bent if and only if $\gcd(c(x), x^p + 1) \neq 1$. Applying Lemma 19-($b$), $f(x)$ is non-bent if and only if (5.7) is achieved. Therefore, $f(x)$ is bent if and only if there exists at least one i for $1 \leq i \leq \frac{p-1}{2}$ such that (5.10) is true, which completes the first part of the proof.

Next, we enumerate vectors $\mathbf{c} = (c_1, c_2, \cdots, c_{n/2-1})$ which satisfy (5.10). The elements of $\mathbf{c}$ can be arranged into an $\frac{n}{2p} \times p$ matrix as follows.

$$M = \begin{bmatrix} c_0 & c_1 & \cdots & c_{p-2} & c_{p-1} \\ c_p & c_{p+1} & \cdots & c_{2p-2} & c_{2p-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ c_{(\frac{n}{2p}-1)p} & c_{(\frac{n}{2p}-1)p+1} & \cdots & c_{\frac{n}{2}-2} & c_{\frac{n}{2}-1} \end{bmatrix}$$

where $c_0 = 0$. Then, w_i is equal to the sum of entries in the ith column of M for $1 \leq i \leq p-1$. Thus, all j's of c_j occurring in w_i and w_{p-i} in (5.10) are distinct. In the following, the number of non-bent functions, i.e., the number of vectors $\mathbf{c}$ satisfying (5.7) will be first counted. The condition of (5.7) is equivalent to

$$w_i + w_{p-i} = 1 \text{ for all } 1 \leq i \leq \frac{p-1}{2},$$

i.e.,

$$\begin{aligned} & c_i + c_{i+p} + \cdots + c_{i+(\frac{n}{2p}-1)p} + c_{p-i} + c_{p-i+p} + \cdots + c_{p-i+(\frac{n}{2p}-1)p} \\ & = 1 \text{ for all } 1 \le i \le \frac{p-1}{2}. \end{aligned} \tag{5.11}$$

For each i, there are $\frac{n}{p} = 2^v$ distinct c_j's in (5.11) where $j \in \{i, i+p, \cdots, i+(2^{v-1}-1)p, p-i, p-i+p, \cdots, p-i+(2^{v-1}-1)p\}$. For (5.11), the number of c_j's which take on the value 1 should be odd for each i. Therefore, there are $\binom{2^v}{1} + \binom{2^v}{3} + \cdots + \binom{2^v}{2^v-1} = \sum_{k=0}^{2^{v-1}-1} \binom{2^v}{2k+1}$ choices of such c_j's for each i. Since there are $\frac{p-1}{2}$ choices of i, the number of c_j's satisfying (5.11) is given by

$$A = \left[\sum_{k=0}^{2^{v-1}-1} \binom{2^v}{2k+1} \right]^{\frac{p-1}{2}} = 2^{(2^v-1)(\frac{p-1}{2})}.$$

Meanwhile, there are no conditions on c_j's for $j \in \{p, 2p, \cdots, (\frac{n}{2p}-1)p\}$. Thus, the number of choices of such c_j's is given by

$$B = 2^{\frac{n}{2p}-1} = 2^{2^{v-1}-1}.$$

Consequently, the number of vectors $\mathbf{c}$ which satisfy (5.7) is given by

$$N_{nb} = A \cdot B = 2^{\frac{n-1-p}{2}}$$

which is equal to the number of vectors $\mathbf{c}$ producing the non-bent functions. Therefore, the number of vectors $\mathbf{c}$ producing the bent functions is given by $N_b = 2^{\frac{n}{2}-1} - N_{nb}$. □

Remark 9 The first few primes p of Theorem 13 are

$$3, 5, 7, 11, 13, 19, 23, 29, 37, 47, 53, 59, 61, 71, \cdots .$$

The matrix M in the proof of Theorem 13 is useful for understanding the construction. Let $M = [\mathbf{V}_0, \mathbf{V}_1, \cdots, \mathbf{V}_{p-1}]$, where $\mathbf{V}_i$, $0 \le i \le p-1$ is the ith column

Table 5.1: Bent functions with quadratic exponents for $n = 12$ (65 corresponds to $Tr_1^6(x^{65})$)

$(c_1c_2c_3c_4c_5)$	Trace Exponents	$(c_1c_2c_3c_4c_5)$	Trace Exponents
(00000)	65	(01010)	5, 17, 65
(00100)	9, 65	(01110)	5, 9, 17, 65
(00011)	17, 33, 65	(10010)	3, 17, 65
(00111)	9, 17, 33, 65	(10110)	3, 9, 17, 65
(01001)	5, 33, 65	(11000)	3, 5, 65
(01101)	5, 9, 33, 65	(11100)	3, 5, 9, 65
(10001)	3, 33, 65	(11011)	3, 5, 17, 33, 65
(10101)	3, 9, 33, 65	(11111)	3, 5, 9, 17, 33, 65

vector of M. Then, it is clear that the sum of elements of a pair of column vectors $\mathbf{V}_i$ and $\mathbf{V}_{p-i}$ is identical to $w_i + w_{p-i}$ for $1 \leq i \leq \frac{p-1}{2}$. Thus, Theorem 13 means that $f(x)$ is bent if and only if there exists at least a pair of column vectors $\mathbf{V}_i$ and $\mathbf{V}_{p-i}$ for $1 \leq i \leq \frac{p-1}{2}$ such that the sum of elements in the pair is equal to 0.

Example 13 *For $n = 12 = 2^2 \times 3$, the matrix of coefficients is given by*

$$M = \begin{bmatrix} 0 & c_1 & c_2 \\ c_3 & c_4 & c_5 \end{bmatrix}.$$

Thus, $f(x) = \sum_{i=1}^{5} c_i Tr_1^{12}(x^{1+2^i}) + Tr_1^6(x^{65})$ is bent if and only if

$$c_1 + c_4 + c_2 + c_5 = 0$$

and c_3 can be free to choose. Also, the number of bent functions is given by

$$N_b = 2^5 - 2^{\frac{12-1-3}{2}} = 32 - 16 = 16.$$

All the possible 16 bent functions are listed in Table 5.1.

Note. For $n = 2p$, we have $w_i = c_i$. Thus, $f(x)$ is bent if and only if there exists at least one i with $1 \leq i \leq \frac{p-1}{2}$ such that $c_i = c_{p-i}$. In Theorem 2 of [59], the authors attempted to state the sufficient condition of this result. However, the assertion that appeared there is not in a clear way.

5.5 Construction for $n = 2^v p^r$

This section presents a necessary and sufficient condition on c_i's that $f(x)$ given by (5.5) is bent for $n = 2^v p^r$ with $v \geq 1$ and $r \geq 2$, where p is odd prime with $\mathrm{ord}_p(2) = p - 1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ with odd s. This section starts from the following definition and lemma.

Definition 1 *Let $t(x) = \sum_{i=1}^{m} t_i x^i + 1$ where $t_i \in \mathbb{F}_2$ and m is even. Then, $t(x)$ is called* circular symmetric with m *if*

$$t_{m+1-i} = t_i, \quad i = 1, 2, \cdots, m/2.$$

Lemma 20 *Let $m = p^k - 1$ for odd prime p and an integer $k > 1$, and $t(x)$ be circular symmetric with m. Assume that there exists a polynomial $h(x) \in \mathbb{F}_2[x]$ such that $t(x) = Q_{p^k}(x)h(x)$, where $Q_{p^k}(x)$ is the p^kth cyclotomic polynomial. Then,*

(a) $h(x)$ is circular symmetric with a degree of $\deg(h) \leq p^{k-1} - 1$.

(b) If $h(x) = \sum_{i=1}^{p^{k-1}-1} h_i x^i + 1$, $h_i \in \mathbb{F}_2$, then

$$t(x) = Q_{p^k}(x) + \sum_{j=0}^{p-1} \sum_{i=1}^{p^{k-1}-1} h_i x^{i+jp^{k-1}}.$$

In other words, $t(x)$ contains all monomial terms of $Q_{p^k}(x)$.

Proof. (a) A circular symmetric polynomial $t(x)$ with $m = p^k - 1$ has the following property.

$$x^{p^k} t(x^{-1}) = \sum_{i=1}^{p^k-1} t_i x^{p^k-i} + x^{p^k} = \sum_{i=1}^{p^k-1} t_{p^k-i} x^i + x^{p^k} = t(x) + x^{p^k} + 1. \qquad (5.12)$$

From $t(x) = Q_{p^k}(x)h(x)$ and the self-reciprocity of $Q_{p^k}(x)$ (from Property 3), we also have

$$\begin{aligned} x^{p^k} t(x^{-1}) &= x^{p^k} Q_{p^k}(x^{-1})h(x^{-1}) = x^{\phi(p^k)} Q_{p^k}(x^{-1}) \cdot x^{p^k - \phi(p^k)} h(x^{-1}) \\ &= Q_{p^k}(x) \cdot x^{p^k - \phi(p^k)} h(x^{-1}). \end{aligned} \tag{5.13}$$

From (5.12) and (5.13),

$$x^{p^k} t(x^{-1}) + t(x) = Q_{p^k}(x)(x^{p^k - \phi(p^k)} h(x^{-1}) + h(x)) = x^{p^k} + 1 = \prod_{d|p^k} Q_d(x).$$

Therefore,

$$\begin{aligned} x^{p^k - \phi(p^k)} h(x^{-1}) + h(x) &= \prod_{d|p^k, d \neq p^k} Q_d(x) = Q_1(x) Q_p(x) \cdots Q_{p^{k-1}}(x) \\ &= x^{p^{k-1}} + 1. \end{aligned} \tag{5.14}$$

From $t(x) = Q_{p^k}(x)h(x)$, we have $\deg(h(x)) \leq p^k - 1 - \phi(p^k) = p^{k-1} - 1$ because $\deg(t(x)) \leq p^k - 1$, where $\phi(p^k) = p^k \left(1 - \frac{1}{p}\right) = p^k - p^{k-1}$. Furthermore, $h(0) = 1$ from $t(0) = Q_{p^k}(0) = 1$. Thus, $h(x)$ can be written as $h(x) = 1 + \sum_{i=1}^{p^{k-1}-1} h_i x^i$ where $h_i \in \mathbb{F}_2$. Hence,

$$\begin{aligned} x^{p^k - \phi(p^k)} h(x^{-1}) &= x^{p^{k-1}} \left(\sum_{i=1}^{p^{k-1}-1} h_i x^{-i} + 1 \right) \\ &= \sum_{i=1}^{p^{k-1}-1} h_i x^{p^{k-1} - i} + x^{p^{k-1}} = \sum_{i=1}^{p^{k-1}-1} h_{p^{k-1}-i} x^i + x^{p^{k-1}} \end{aligned} \tag{5.15}$$

where the last equality is from a change of variables from $p^{k-1} - i$ to i. Applying this to (5.14), the requirement of coefficients h_i's is obtained, i.e.,

$$\sum_{i=1}^{\frac{p^{k-1}-1}{2}} (h_i + h_{p^{k-1}-i})(x^i + x^{p^{k-1}-i}) = 0$$

or equivalently,

$$h_i = h_{p^{k-1}-i}, \quad 1 \le i \le \frac{p^{k-1}-1}{2}.$$

From Definition 1, $h(x)$ is circular symmetric.

(b) By definitions,

$$t(x) = Q_{p^k}(x)h(x) = Q_{p^k}(x) + Q_{p^k}(x) \sum_{i=1}^{p^{k-1}-1} h_i x^i.$$

From $Q_{p^k}(x) = Q_p(x^{p^{k-1}}) = \sum_{j=0}^{p-1} x^{p^{k-1}j}$,

$$t(x) = Q_{p^k}(x) + \sum_{j=0}^{p-1} \sum_{i=1}^{p^{k-1}-1} h_i x^{i+p^{k-1}j} = Q_{p^k}(x) + A(x)$$

where $A(x)$ is the double summation of the second term. The exponent of a monomial term in $A(x)$ has a form $i+p^{k-1}j$ where $1 \le i \le p^{k-1}-1$ and $0 \le j \le p-1$, while the exponent of a monomial term of $Q_{p^k}(x)$ has a form $p^{k-1}u$ where $0 \le u \le p-1$. Thus, all monomial terms in $A(x)$ are distinct from the monomials in $Q_{p^k}(x)$. Therefore, all terms of $Q_{p^k}(x)$ survive in $t(x)$. □

Similar to the case of $n = 2^v p$, we need to investigate $\bar{c_k}(x)$, $c(x)$ reduced modulo $x^{p^k}+1$, $1 \le k \le r$ for $n = 2^v p^r$ with $r \ge 2$. In order to do so, the following two lemmas on $\bar{c_k}(x)$ are needed.

Lemma 21 *Let $\bar{c_k}(x)$ be $c(x)$ reduced modulo $x^{p^k}+1$ for each k with $1 \le k \le r$, where $c(x)$ is given by (5.3) for $n = 2^v p^r$ with $v \ge 1$ and $r \ge 2$. Then,*

$$\begin{aligned} \bar{c_k}(x) \equiv c(x) \pmod{x^{p^k}+1} &= \sum_{i=1}^{p^k-1} w_{i,k}(x^i + x^{p^k-i}) + 1 \\ &= \sum_{i=1}^{\frac{p^k-1}{2}} (w_{i,k} + w_{p^k-i,k})(x^i + x^{p^k-i}) + 1 \end{aligned} \tag{5.16}$$

where $w_{i,k} = \sum_{l=0}^{\frac{n}{2p^k}-1} c_{i+lp^k}$. *Furthermore,* $\bar{c_k}(x)$ *is circular symmetric.*

Proof. Similar to (5.8), $c(x)$ can be rewritten as

$$c(x) = \sum_{i=1}^{p^k-1} c_i(x^i + x^{n-i}) + c_{p^k}(x^{p^k} + x^{n-p^k}) + \sum_{i=p^k+1}^{2p^k-1} c_i(x^i + x^{n-i}) + c_{2p^k}(x^{2p^k} + x^{n-2p^k})$$
$$+ \cdots + \sum_{i=(\frac{n}{2p^k}-1)p^k+1}^{\frac{n}{2}-1} c_i(x^i + x^{n-i}) + x^{\frac{n}{2}}.$$

With $w_{i,k} = \sum_{l=0}^{\frac{n}{2p^k}-1} c_{i+lp^k}$, therefore, $\bar{c_k}(x)$ is given by (5.16), where $w_{i,k} + w_{p^k-i,k}$ is a coefficient of both x^i and x^{p^k-i}. Thus, $\bar{c_k}(x)$ is circular symmetric. □

Lemma 22 *With the notation of Lemma 21,*

(*a*) $\gcd(c(x), Q_{p^k}(x)) \neq 1 \Longleftrightarrow \gcd(\bar{c_k}(x), Q_{p^k}(x)) \neq 1$.

(*b*) *Let* p *be odd prime with* $\mathrm{ord}_p(2) = p-1$ *or* $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ *where* s *is odd. Then,* $\gcd(c(x), Q_{p^k}(x)) \neq 1$ *if and only if* $\gcd(\bar{c_k}(x), Q_{p^k}(x)) = Q_{p^k}(x)$.

Proof. (*a*) From the definition of $\bar{c_k}(x)$,

$$c(x) = b(x)(x^{p^k} + 1) + \bar{c_k}(x) = b(x)Q_1(x)\prod_{i=1}^{k} Q_{p^i}(x) + \bar{c_k}(x)$$

where $b(x)$ is a quotient of $c(x)$ divided by $x^{p^k} + 1$. Hence, if $Q_{p^k}(x)$ shares a common factor with $\bar{c_k}(x)$, then it also shares the common factor with $c(x)$, and vice versa.

(*b*) If $\mathrm{ord}_p(2) = p-1$, then $Q_{p^k}(x) = Q_p(x^{p^{k-1}})$ is irreducible over $\mathbb{F}_2$ for a given k from Property 6. From the irreducibility of $Q_{p^k}(x)$ and Lemma 22-(*a*),

$$\begin{aligned} \gcd(c(x), Q_{p^k}(x)) \neq 1 &\Longleftrightarrow \gcd(\bar{c_k}(x), Q_{p^k}(x)) \neq 1 \\ &\Longleftrightarrow \gcd(\bar{c_k}(x), Q_{p^k}(x)) = Q_{p^k}(x). \end{aligned} \tag{5.17}$$

If $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ where s is odd, on the other hand, we have

$$Q_{p^k}(x) = Q_p(x^{p^{k-1}}) = G_1(x)G_2(x) \text{ for } G_i(x) = g_i(x^{p^{k-1}}),\ i = 1, 2$$

where $g_1(x)$ and $g_2(x)$ are the irreducible factors of $Q_p(x)$ such that $Q_p(x) = g_1(x)g_2(x)$ and $g_2(x) = x^s g_1(x^{-1})$ [18]. From Property 6, $G_1(x)$ and $G_2(x)$ are irreducible over $\mathbb{F}_2$. Also, they are reciprocal to each other, i.e., $G_2(x) = x^{sp^{k-1}} G_1(x^{-1})$. Let z be a root of $Q_{p^k}(x)$ for $k \geq 1$. Then, $z \neq 1$. If z is a root of $G_1(x)$, then z^{-1} is a root of $G_2(x)$, and vice versa. (Note that if z is a root of $G_1(x)$, then z^{-1} cannot be its root. If it were true, then this holds for any other roots of $G_1(x)$ and thus the number of valid roots of $G_1(x)$ should be even, which is impossible because sp^{k-1}, the degree of $G_1(x)$, is odd.) Since z is a root of $Q_{p^k}(x)$, we have $z^{p^k} = z^n = 1$. Thus, from (5.16),

$$\bar{c_k}(z) = \sum_{i=1}^{p^k-1} w_{i,k}(z^i + z^{p^k-i}) + 1 = \sum_{i=1}^{p^k-1} w_{i,k}(z^i + z^{-i}) + 1.$$

Hence, z^{-1} is a root of $\bar{c_k}(x)$ if z is its root. Therefore, if $\bar{c_k}(x)$ has the irreducible factor $G_1(x)$, then it simultaneously has the other irreducible factor $G_2(x)$, and vice versa. Similar to the case of $\mathrm{ord}_p(2) = p - 1$, we have (5.17) for a given k. □

In (5.16), denote

$$u_{i,k} = w_{i,k} + w_{p^k-i,k}, \quad 1 \leq i \leq p^k - 1 \tag{5.18}$$

for each k with $1 \leq k \leq r$. Let U_k be a $p \times p^{k-1}$ matrix in Figure 5.1 whose entries are given by $u_{i,k}$ where $u_{0,k} = 1$. Then,

$$U_k = \begin{bmatrix} \mathbf{A}_{0,k} & \mathbf{A}_{1,k} & \cdots & \mathbf{A}_{p^{k-1}-1,k} \end{bmatrix} \tag{5.19}$$

where $\mathbf{A}_{i,k}$ is the ith column vector of U_k. From (5.18), $u_{i,k} = u_{p^k-i,k}$. Using the matrix U_k, we give the construction of all bent functions represented by (5.5) for $n = 2^v p^r$ with $v \geq 1$ and $r \geq 2$.

$$U_k = \begin{bmatrix} u_{0,k} & u_{1,k} & \cdots & u_{\frac{p^{k-1}-1}{2},k} & \cdots & u_{p^{k-1}-1,k} \\ u_{p^{k-1},k} & u_{p^{k-1}+1,k} & \cdots & u_{p^{k-1}+\frac{p^{k-1}-1}{2},k} & \cdots & u_{2p^{k-1}-1,k} \\ u_{2p^{k-1},k} & u_{2p^{k-1}+1,k} & \cdots & u_{2p^{k-1}+\frac{p^{k-1}-1}{2},k} & \cdots & u_{3p^{k-1}-1,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{(\frac{p-1}{2})p^{k-1},k} & u_{(\frac{p-1}{2})p^{k-1}+1,k} & \cdots & u_{\frac{p^k-1}{2},k} & \cdots & u_{\frac{p^k-1}{2}+\frac{p^{k-1}-1}{2},k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{(p-1)p^{k-1},k} & u_{(p-1)p^{k-1}+1,k} & \cdots & u_{(p-1)p^{k-1}+\frac{p^{k-1}-1}{2},k} & \cdots & u_{p^k-1,k} \end{bmatrix}$$

Figure 5.1: Structure of U_k

Theorem 14 *With the above notation, let p be odd prime with $\mathrm{ord}_p(2) = p - 1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ where s is odd, and $n = 2^v p^r$ with $v \geq 1$ and $r \geq 2$. Then, $f(x)$ given by (5.5) is bent if and only if for each k with $1 \leq k \leq r$, there exists at least one i for $0 \leq i \leq \frac{p^{k-1}-1}{2}$ such that $\mathbf{A}_{i,k}$ given by (5.19), is not a constant vector. In other words, $\mathbf{A}_{0,k} \neq (1, 1, \cdots, 1)$ or $\mathbf{A}_{i,k} \neq (c, c, \cdots, c)$ where $c \in \{0, 1\}$ for at least one i with $1 \leq i \leq \frac{p^{k-1}-1}{2}$.*

Proof. Similar to the proof of Theorem 13, a condition on $\bar{c_k}(x)$ will be first derived for $f(x)$ given by (5.5) to be non-bent.

From Fact 5 and Corollary 3, $f(x)$ is bent if and only if $\gcd(c(x), x^n + 1) = \gcd(c(x), x^{p^r}+1) = 1$. Since $Q_{p^k}(x)$'s for $1 \leq k \leq r$ are all factors of $(x^{p^r}+1)/(x+1)$, we have $\gcd(c(x), x^n + 1) = 1$ if and only if $\gcd(c(x), Q_{p^k}(x)) = 1$ for every k. Therefore, $f(x)$ is non-bent if and only if there exists at least one k for $1 \leq k \leq r$ such that $\gcd(c(x), Q_{p^k}(x)) \neq 1$. From Lemma 22, it is clear that if $\mathrm{ord}_p(2) = p - 1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ with odd s, then $\gcd(c(x), Q_{p^k}(x)) \neq 1$ if and only if $\gcd(\bar{c_k}(x), Q_{p^k}(x)) = Q_{p^k}(x)$. Thus, $\bar{c_k}(x)$ can be represented by

$$\bar{c_k}(x) = Q_{p^k}(x)h(x) \tag{5.20}$$

where $h(x) = 1 + \sum_{i=1}^{p^{k-1}-1} h_i x^i,\ h_i \in \mathbb{F}_2$.

From Lemma 21, $\bar{c_k}(x)$ is circular symmetric with $\deg(\bar{c_k}(x)) \leq p^k - 1$. Also, $h(x)$ is circular symmetric with $\deg(h(x)) \leq p^{k-1} - 1$ from Lemma 20. Thus,

together with (5.20), $\bar{c_k}(x)$ can be written as follows.

$$\begin{aligned}\bar{c_k}(x) &= Q_{p^k}(x) + \sum_{j=0}^{p-1}\sum_{t=1}^{p^{k-1}-1} h_t x^{t+jp^{k-1}} \\ &= \sum_{j=1}^{\frac{p-1}{2}}(x^{jp^{k-1}} + x^{p^k - jp^{k-1}}) + 1 + \sum_{t=1}^{p^{k-1}-1} h_t \sum_{j=0}^{\frac{p-1}{2}}(x^{t+jp^{k-1}} + x^{p^k - t - jp^{k-1}}).\end{aligned} \tag{5.21}$$

From Lemma 21, on the other hand, we have

$$\begin{aligned}\bar{c_k}(x) &= \sum_{i=1}^{p^k-1} w_{i,k}(x^i + x^{p^k-i}) + 1 \\ &= \sum_{j=1}^{p-1} w_{jp^{k-1},k}(x^{jp^{k-1}} + x^{p^k - jp^{k-1}}) + 1 + \sum_{i=1, i\neq jp^{k-1}}^{p^k-1} w_{i,k}(x^i + x^{p^k-i}) \\ &= \sum_{j=1}^{\frac{p-1}{2}}(w_{jp^{k-1},k} + w_{p^k - jp^{k-1},k})(x^{jp^{k-1}} + x^{p^k - jp^{k-1}}) + 1 \\ &\quad + \sum_{t=1}^{p^{k-1}-1}\sum_{j=0}^{\frac{p-1}{2}}(w_{t+jp^{k-1},k} + w_{p^k - t - jp^{k-1},k})(x^{t+jp^{k-1}} + x^{p^k - t - jp^{k-1}}), \\ &\quad (i = t + jp^{k-1}) \\ &= \sum_{j=1}^{\frac{p-1}{2}} u_{jp^{k-1},k}(x^{jp^{k-1}} + x^{p^k - jp^{k-1}}) + 1 \\ &\quad + \sum_{t=1}^{p^{k-1}-1}\sum_{j=0}^{\frac{p-1}{2}} u_{t+jp^{k-1},k}(x^{t+jp^{k-1}} + x^{p^k - t - jp^{k-1}}).\end{aligned} \tag{5.22}$$

By comparing (5.21) and (5.22),

$$u_{t+jp^{k-1},k} = \begin{cases} 1, & t = 0 \text{ and } 1 \le j \le \frac{p-1}{2}, \\ h_t, & 1 \le t \le p^{k-1} - 1 \text{ and } 0 \le j \le \frac{p-1}{2}. \end{cases} \tag{5.23}$$

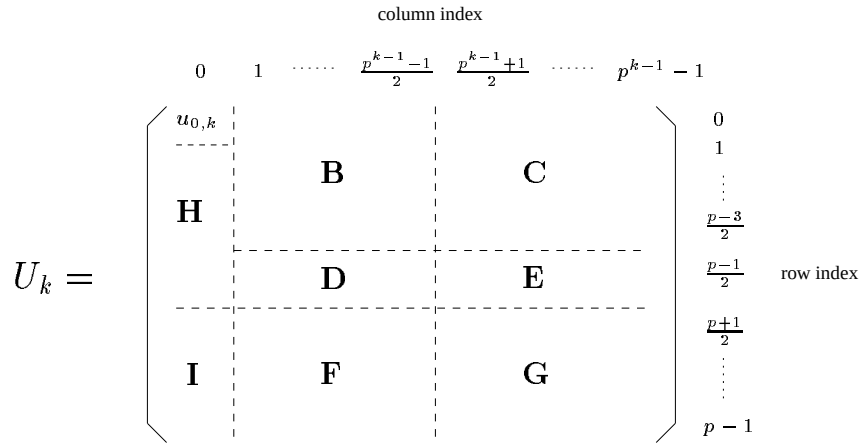


Figure 5.2: Submatrix structure of U_k

Thus, for non-bent cases, the entries of U_k, $u_{i,k}$'s with $1 \leq i \leq \frac{p^k-1}{2}$, are determined by (5.23). This is equivalently saying that each column of U_k is a constant vector if and only if $f(x)$ is non-bent. This completes the proof for Theorem 14. □

In the following, a submatrix structure of U_k is examined in detail in order to better understand Theorem 14. U_k consists of several submatrices in Figure 5.2, which are defined as follows. Each of $\mathbf{B}, \mathbf{C}, \mathbf{F}$, and $\mathbf{G}$ is an $l \times m$ matrix, where $l = \frac{p-1}{2}$ and $m = \frac{p^{k-1}-1}{2}$. Also, each of $\mathbf{D}$ and $\mathbf{E}$ is a $1 \times m$ matrix, and each of $\mathbf{H}$ and $\mathbf{I}$ is an $l \times 1$ matrix. From $u_{i,k} = u_{p^k-i,k}$ for $1 \leq i \leq \frac{p^k-1}{2}$, if $\mathbf{B} = [b_{i,j}]$, then $\mathbf{G} = [b_{l-i,m-j}]$. This relation is denoted by $\mathbf{B} \sim \mathbf{G}$. Similarly, we have $\mathbf{C} \sim \mathbf{F}$. Also, if $\mathbf{D} = [d_{i,j}]$, then $\mathbf{E} = [d_{i,m-j}]$, denoted by $\mathbf{D} \sim \mathbf{E}$. Thus, each element in $\mathbf{E}, \mathbf{F}$, and $\mathbf{G}$ is determined by each element in $\mathbf{B}, \mathbf{C}$, and $\mathbf{D}$, respectively. In other words, the column vectors $\mathbf{A}_{i,k}$ of U_k for $\frac{p^{k-1}+1}{2} \leq i \leq p^{k-1} - 1$ are determined by the column vectors $\mathbf{A}_{i,k}$ for $1 \leq i \leq \frac{p^{k-1}-1}{2}$.

In terms of the enumeration of quadratic bent functions for $n = 2^v p^r$, $r > 1$, Hu and Feng derived general results [42].

Example 14 *For $n = 18 = 2 \cdot 3^2$, $p = 3$ and $r = 2$. We consider the conditions for each $k, 1 \leq k \leq 2$ such that $f(x)$ defined by (5.5) is bent.*

For $k = 1$, we consider U_1, i.e.,

$$U_1 = \begin{bmatrix} u_{0,1} \\ u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} 1 \\ w_{1,1} + w_{2,1} \\ w_{2,1} + w_{1,1} \end{bmatrix} = \begin{bmatrix} 1 \\ c_1 + c_4 + c_7 + c_2 + c_5 + c_8 \\ c_2 + c_5 + c_8 + c_1 + c_4 + c_7 \end{bmatrix}$$

where a condition for $f(x)$ to be bent is

$$c_1 + c_4 + c_7 + c_2 + c_5 + c_8 = 0 \tag{5.24}$$

from Theorem 14. If $k = 2$, on the other hand,

$$U_2 = \begin{bmatrix} u_{0,2} & u_{1,2} & u_{2,2} \\ u_{3,2} & u_{4,2} & u_{5,2} \\ u_{6,2} & u_{7,2} & u_{8,2} \end{bmatrix} = \begin{bmatrix} 1 & w_{1,2} + w_{8,2} & w_{2,2} + w_{7,2} \\ w_{3,2} + w_{6,2} & w_{4,2} + w_{5,2} & w_{5,2} + w_{4,2} \\ w_{6,2} + w_{3,2} & w_{7,2} + w_{2,2} & w_{8,2} + w_{1,2} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & c_1 + c_8 & c_2 + c_7 \\ c_3 + c_6 & c_4 + c_5 & c_5 + c_4 \\ c_6 + c_3 & c_7 + c_2 & c_8 + c_1 \end{bmatrix}$$

where another condition for $f(x)$ to be bent is

$$c_3 + c_6 = 0 \text{ or the vector } (c_1 + c_8, c_4 + c_5, c_7 + c_2) \text{ is not constant.} \tag{5.25}$$

Both (5.24) and (5.25) must be achieved so that $f(x)$ is bent. Hence, $f(x)$ is bent at the following two exclusive cases.

(a) $c_1 + c_4 + c_7 + c_2 + c_5 + c_8 = 0$ *and* $c_3 + c_6 = 0 \Longrightarrow 64$ *cases, or*

(b) $(c_1 + c_8, c_4 + c_5, c_7 + c_2) = (0,1,1), (1,0,1), (1,1,0)$ *and* $c_3 + c_6 = 1 \Longrightarrow 48$ *cases.*

In (b), $c_3 + c_6 = 1$ is required to distinguish (b) from (a). With $(c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8) = (01110100)$, for example, $f(x) = Tr(x^5 + x^9 + x^{17} + x^{65}) + Tr_1^9(x^{513})$ is bent. Finally, there are in total $64 + 48 = 112$ bent functions of $f(x)$ whose c_i's satisfy either (a) or (b). It is also verified by computer experiments.

5.6 LFSR Implementation of Bent Sequences with A Quadratic Bent Function

This section introduces a family of bent sequences employing the quadratic bent function constructed in this chapter, and presents its linear feedback shift register (LFSR) implementation. The exact linear complexity of the bent sequences is also computed.

In terms of a polynomial form, a bent sequence [50] [71] of period $2^{2n}-1$ is represented by

$$f_\lambda(x) = f(Tr_n^{2n}(x)) + Tr_1^{2n}((\lambda+\sigma_0)x), \quad x \in \mathbb{F}_{2^{2n}}^* \tag{5.26}$$

where $\lambda \in \mathbb{F}_{2^n}$, $\sigma_0 \in \mathbb{F}_{2^{2n}} \setminus \mathbb{F}_{2^n}$, and f is a bent function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. A family of bent sequences constitutes a $(2^{2n}-1, 2^n, 1+2^n)$ signal set [33], where the maximum correlation achieves the equality of the Welch bound asymptotically. Therefore, bent sequences are a potential candidate for signature or scrambling sequences in future mobile CDMA communication systems.

By employing a quadratic bent function of a polynomial form for f, the bent sequence in (5.26) can be implemented by linear feedback shift registers (LFSRs). Let $h_i(x)$ be a single trace function with a coefficient c_i in a quadratic bent function $f(x)$ given by (5.2), i.e.,

$$h_i(x) = \begin{cases} Tr_1^n(x^{2^i+1}), & \text{if } 1 \le i < \frac{n}{2} \\ Tr_1^{n/2}(x^{2^{n/2}+1}), & \text{if } i = \frac{n}{2}. \end{cases}$$

Then, $f(x)$ is a combination of $h_i(x)$'s according to $c_i = 1$, and hence $f(Tr_n^{2n}(x))$

is implemented by a sum of $h_i(Tr_n^{2n}(x))$'s each of which is given by

$$\begin{aligned} h_i(Tr_n^{2n}(x)) &= Tr_1^u\left(\left(Tr_n^{2n}(x)\right)^{2^i+1}\right) \\ &= Tr_1^u\left(x^{2^i+1}+x^{2^i+2^n}+x^{2^{n+i}+1}+x^{2^{n+i}+2^n}\right) \\ &= \begin{cases} Tr_1^{2n}(x^{2^i+1}+x^{2^{n+i}+1}) & \text{if } 1\le i<\frac{n}{2} \\ Tr_1^{2n}(x^{2^{n/2}+1}) & \text{if } i=\frac{n}{2} \end{cases} \end{aligned} \tag{5.27}$$

where $u=n$ for $1\le i<\frac{n}{2}$, or $u=\frac{n}{2}$ for $i=\frac{n}{2}$. Since a trace function $Tr_1^{2n}(x)$ is implemented by a $2n$-stage LFSR, $h_i(Tr_n^{2n}(x))$ can be implemented by two $2n$-stage LFSRs for $1\le i<\frac{n}{2}$, or one $2n$-stage LFSR for $i=\frac{n}{2}$, where each LFSR has a different feedback polynomial. Consequently, $f(Tr_n^{2n}(x))$ can be implemented by a combination of such LFSRs, which allows the bent sequences employing a quadratic bent function to be implemented by LFSR structures.

The linear complexity of the bent sequences is computed as follows. If $f(x)$ given by (5.2) has l_c nonzero c_i's for $1\le i\le\frac{n}{2}-1$, or equivalently l_c nonzero trace terms from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$, then the linear complexity L of the bent sequences in (5.26) is given by

$$L=2\cdot 2n\cdot l_c+2n+2n=4n(l_c+1).$$

Thus, the maximum linear complexity is obtained by $L_{\max}=2n^2$ at $l_c=\frac{n}{2}-1$, where all c_i's are equal to 1 for $1\le i\le\frac{n}{2}-1$.

Example 15 *For $n=6$, consider a quadratic bent function $f(x)=Tr_1^6(x^3)+Tr_1^6(x^5)+Tr_1^3(x^9)$ constructed from Theorem 13. Then, the bent sequences employing $f(x)$ as a bent function are represented by*

$$\begin{aligned} f_\lambda(x) &= f(Tr_n^{12}(x))+Tr_1^{12}((\lambda+\sigma_0)x) \\ &= Tr_1^{12}(x^3)+Tr_1^{12}(x^{129})+Tr_1^{12}(x^5)+Tr_1^{12}(x^{257}) \\ &\quad +Tr_1^{12}(x^9)+Tr_1^{12}((\lambda+\sigma_0)x) \end{aligned} \tag{5.28}$$

where $\lambda\in\mathbb{F}_{2^6}$ and $\sigma_0\in\mathbb{F}_{2^{12}}\setminus\mathbb{F}_{2^6}$. In (5.28), if x runs through all elements of $\mathbb{F}_{2^{12}}^*$ generated by a primitive polynomial $x^{12}+x^6+x^4+x+1$, then $f_\lambda(x)$ generates a

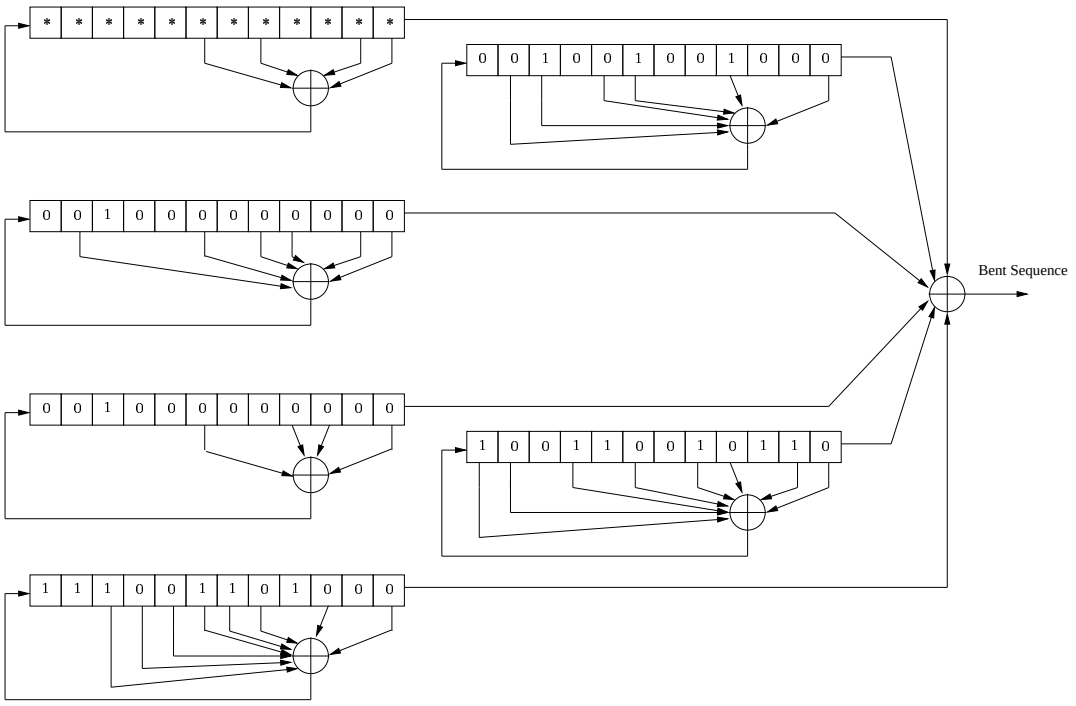


Figure 5.3: LFSR implementation of bent sequences of period $2^{12}-1$

bent sequence of period $2^{12}-1$ *for each* λ.

The bent sequences are implemented by Figure 5.3, where '$*$*'*s denote varying initial states corresponding to a specific user index* λ. *From Figure 5.3, it is clear that the linear complexity of the bent sequences is given by* $4n(l_c+1) = 4\times 6\times 3 = 72$.

5.7 Conclusion

This chapter has constructed all bent functions represented by a polynomial form (5.2) by giving necessary and sufficient conditions on c_i's for $n = 2^v p^r$ with $v \geq 1$ and $r \geq 1$, where p is odd prime with $\mathrm{ord}_p(2) = p-1$ or $\mathrm{ord}_p(2) = \frac{p-1}{2} = s$ with odd s. The enumeration for such bent functions has been also given for $n = 2^v p^r$ with $v \geq 1$ and $r = 1$. By employing the quadratic bent functions of a polynomial form, a family of bent sequences is implemented by LFSR structures.

Chapter 6

Concluding Remarks

In this thesis, periodic correlation of binary sequences has been mainly studied. Several interesting results of binary sequences and correlation have been presented; 3- and 5-valued crosscorrelation properties of binary sequences, new binary sequences with optimal autocorrelation, a new sequence family with low correlation and a large family size, and quadratic bent functions of a polynomial form. The contributions of this thesis and future works are summarized in this chapter.

6.1 Summary of Contributions

- *Crosscorrelation Properties:* The recently discovered binary sequences of period $2^n - 1$ with ideal two-level autocorrelation have been investigated in terms of their crosscorrelation for odd n. At most 5-valued crosscorrelaion of binary m-sequences is clarified, which is linked to the crosscorrelation of a pair of other binary sequences with ideal two-level autocorrelation. Through the Hadamard equivalence, 3- and 5-valued crosscorrelation properties of several pairs of the binary sequences are determined by either theoretical proofs or conjectures.

- *Binary Sequences with Optimal Autocorrelation Magnitude:* New binary sequences of period $N = 4(2^m - 1)$, $m = 2k$, $k > 1$ with three-valued out-

of-phase autocorrelation of $\{0, \pm 4\}$ have been found, which are theoretically optimal with respect to autocorrelation magnitude. It has been shown that the sequences provide large linear complexities by the implementation with linear feedback shift registers (LFSRs) and a simple logic. This sequence construction is the first application of the array structure of sequences [34] to the analysis and design of binary sequences with optimal autocorrelation.

- *Binary Sequence Family with Low Correlation and A Large Size:* A new binary sequence family with low correlation and a large family size has been constructed. With a variable integer ρ, the sequence family provides various sub-sequence families each of which has the maximum correlation and family size determined by ρ. The flexibility enables us to choose a sub-sequence family possessing the parameters suitable to specific practical demands. Sequences in the family also have large linear complexities favorable to the potential cryptographic applications. With given maximum correlation, the new sequence family of period $2^n - 1$ for odd n and $\rho \leq 2$ provides the largest family size of all known binary sequence families.

- *Quadratic Bent Functions of A Polynomial Form:* The construction of bent functions of a special quadratic polynomial form has been presented by the equivalent necessary and sufficient conditions. By employing the quadratic bent function of a polynomial form, one is able to implement the bent sequences with linear feedback shift registers (LFSRs). Furthermore, the exact linear complexity of the bent sequences can be also computed.

6.2 Future Works

Since sequences are the basic theoretical background of communication and cryptography, theoretical researches for them will crucially affect the performance of communication and cryptographic systems. In terms of the correlation of sequences, there are several interesting research topics for theory and applications of sequences.

6.2.1 Theoretical Researches

- *Correlation Properties of Binary Sequences:* For the recently discovered binary sequences of period $2^n - 1$ with ideal two-level autocorrelation, crosscorrelation properties have been investigated for odd n in this thesis. As an extension of the research, the crosscorrelation properties for even n are also worth studying. In the meantime, aperiodic autocorrelation properties of the sequences have not been clearly examined. Since the aperiodic correlation better reflects real situations of communication systems, it is also worth studying for the sequences. This mathematical research shall make theoretical contributions to binary sequences and correlation for digital communications.

- *Binary Sequence Family with Low Correlation:* A goal of sequence design for DS-CDMA systems is to construct a new binary sequence family where the maximum correlation of the sequences asymptotically achieves or approaches to the equalities of its theoretical lower bounds. In Figure 4.7, however, no known binary sequence families achieve the asymptotic bounds for $\log_L M > 1$, while there exists a gap between the maximum correlation and the bounds. Thus, it would be theoretically interesting to design a new binary sequence family with low correlation achieving the equalities of the bounds. Exponential sums [72] and the Weil-Carlitz-Uchiyama bound [40] can be revisited for this study. Also, the structure of Z_4-linear binary sequence family can be intensively examined for the improvement.

6.2.2 Practical Researches

- *Sequences for MC-CDMA:* Recently, there are active discussions on the beyond third or the fourth generation mobile communication systems. Several techniques are considered as a candidate for multiple access schemes; DS-CDMA, MC-CDMA, OFDMA, etc. Among them, MC-CDMA (multicarrier CDMA) [39] has much attention from researchers because of its robustness to multipath fading and the FFT-based efficient implementation. In order to identify users or channels in MC-CDMA, spreading sequences are also needed

as in DS-CDMA. In MC-CDMA, however, each chip of a spreading sequence is modulated by a complex valued signal with a different carrier. Therefore, the requirements for sequences are a bit different from those of DS-CDMA, which are described in [75].

1) *Crest factors*: The peak-to-average power ratio (PAPR) of spreading sequences is required to be as low as possible for avoiding the nonlinear distortion of transmitted signals.

2) *Dynamic power ranges*: The ratio of maximum to minimum values of complex signal envelopes should be also minimized to avoid the increase of a required linearity range of a power amplifier.

3) *Correlation*: Complex valued multicarrier sequences must have good correlation such that the corresponding multiple access interference (MAI) is minimized.

Therefore, efforts should be made to design the optimum or nearly optimum sequences achieving the above requirements for the application to MC-CDMA.

- *Sequences for ultra-wideband (UWB) radio:* Ultra-wideband (UWB) radio technology is anticipated to cause paradigm shifts in design and implementation of communication systems, and to play an important role in future wireless local area networks [82]. The UWB radio has many advantages over narrowband wireless technologies, including the coexistence with current radio services, the low cost design from simple transceiver architectures, and the high performance in multipath environments. UWB systems should employ spread spectrum techniques to protect themselves against the interference from other radio systems, and to give themselves multiple access capability. Also, UWB signals show the flat power spectral density so that they should meet the regulation for the satisfactory coexistence with other narrowband and wideband signals. For these purposes, direct-sequence (DS-SS) and time-hopped spread spectrum (TH-SS) techniques are considered in proposed UWB systems [81].

1) *Design of a sequence family for TH-SS UWB systems:* In a TH-SS, a sequence family should be designed such that the sequences in the family have low coincidence correlation and flat power spectral densities. In terms of the spectral flatness, the cyclic difference set sequences are known to be ideal [81]. Thus, we need to make efforts to the design of a new sequence family with low Hamming correlation and spectral flatness based on the cyclic difference sets.

2) *Design of a sequence family for DS-SS UWB systems:* In a DS-SS, the sequence design for UWB systems is quite similar to the one for DS-CDMA systems. In DS-SS UWB systems, the individual sequence is further required to have the optimal or nearly optimal power spectral density for the spectral flatness. Therefore, we need to evaluate the power spectral densities of known sequence families with low correlation. Then, efforts will be made to design a new sequence family to guarantee the spectral flatness and the low mutual interference for DS-SS UWB systems.

Whether a system employs a DS-SS or a TH-SS, it is a key to successful UWB systems that one is able to design a sequence family where each sequence has a flat power spectral density and low correlation [81].

Bibliography

[1] *Bluetooth Specification, ver. 1.1*, available at www.bluetooth.org/spec/.

[2] *3GPP TS 25.213: Technical Specification Group Radio Access Network; Spreading and modulation (FDD) (Release 1999), v.3.9.0*, Dec. 2003.

[3] *IEEE Standard 802.11, Information Technology - Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks - Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2003.

[4] *Physical Layer Standard for cdma2000 Spread Spectrum Systems, Ver. 1.0, Revision D*, Feb. 2004.

[5] M. Antweiler, "Cross-correlation of p-ary GMW sequences," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1253–1261, 1994.

[6] K. T. Arasu, C. Ding, T. Helleseth, P. V. Kumar, and H. Martinsen, "Almost difference sets and their sequences with optimal autocorrelation," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2934–2943, Nov. 2001.

[7] R. H. Barker, "Group synchronization of binary digital systems," in *Proceedings of the Second London Symposium on Information Theory*, London, 1953, pp. 273–287.

[8] L. D. Baumert, *Cyclic Difference Sets*, ser. Lecture Notes in Mathematics. Springer-Verlag, 1971.

[9] H. Beker and F. Piper, *Cipher Systems.* New York: John Wiley and Sons, 1982.

[10] E. R. Berlekamp, *Algebraic Coding Theory.* New York: McGraw-Hill, 1968.

[11] S. Boztas and P. V. Kumar, "Binary sequences with Gold-like correlation but larger linear span," *IEEE Trans. Inform. Theory*, vol. 40, no. 2, pp. 532–537, Mar. 1994.

[12] M. Briceno, I. Goldberg, and D. Wagner, "A pedagogical implementation of A5/1," May 1999, available at http://www.scard.org.

[13] A. Canteaut, C. Carlet, P. Charpin, and C. Fontaine, "On cryptographic properties of the cosets of $R(1,m)$," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1494–1513, 2001.

[14] A. Canteaut and P. Charpin, "Decomposing bent functions," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 2004–2019, 2003.

[15] C. Carlet, "A larger class of cryptographic boolean functions via a study of the Maiorana-McFarland construction," in *Advances in Cryptology - CRYPTO 2002, no. 2442 in Lecture Notes in Computer Science*, 2002, pp. 549–564.

[16] C. Carlet, P. Charpin, and V. A. Zinoviev, "Codes, bent functions and permutations suitable for DES-like cryptosystem," *Designs, Codes, and Cryptography*, vol. 15, pp. 125–156, 1998.

[17] A. Chang, P. Gaal, S. W. Golomb, G. Gong, T. Helleseth, and P. V. Kumar, "On a conjectured ideal autocorrelation sequence and a related triple-error correcting cyclic code," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 680–687, 2000.

[18] P. Charpin, E. Pasalic, and C. Tavernier, "On bent and semi-bent quadratic boolean functions," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4286–4298, Dec. 2005.

[19] D. C. Chu, "Polyphase codes with periodic correlation properties," *IEEE Trans. Inform. Theory*, pp. 531–532, July 1972.

[20] J. Davis, "Almost difference sets and reversible difference sets," *Arch. Math.*, vol. 59, pp. 595–602, 1992.

[21] J. F. Dillon, "Elementary Hadamard difference sets," Ph.D. dissertation, Univ. of Maryland, 1974.

[22] ——, "Multiplicative difference sets via additive characters," *Designs, Codes, and Cryptography*, pp. 225–236, 1999.

[23] J. F. Dillon and H. Dobbertin, "New cyclic difference sets with Singer parameters," *Finite Fields and Their Applications*, no. 10, pp. 342–389, 2004.

[24] C. Ding, "Binary cyclotomic generators," *Lecture Notes in Computer Science*, vol. 1008, pp. 29–60, 1995, in Fast Software Encryption.

[25] C. Ding, T. Helleseth, and K. Y. Lam, "Several classes of binary sequences with three-level autocorrelation," *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp. 2606–2612, Nov. 1999.

[26] C. Ding, T. Helleseth, and H. Martinsen, "New families of binary sequences with optimal three-valued autocorrelation," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 428–433, Jan. 2001.

[27] H. Dobbertin, G. Leander, A. Canteaut, C. Carlet, P. Felke, and P. Gaborit, "Construction of bent functions via Niho power functions," *Journal of Combinatorial Theory, Series A*, to appear, Available at www.sciencedirect.com.

[28] R. L. Frank and S. A. Zadoff, "Phase shift pulse codes with good periodic correlation functions," *IRE Trans. Inform. Theory*, pp. 381–382, Oct. 1962.

[29] R. A. Games, "Crosscorrelation of m-sequences and GMW-sequences with the same primitive polynomial," *Discrete Applied Mathematics*, vol. 12, pp. 139–146, 1985.

[30] M. J. E. Golay, "Complementary series," *IRE Trans. Inform. Theory*, pp. 82–87, Apr. 1961.

[31] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Inform. Theory*, vol. 14, pp. 154–156, Jan. 1968.

[32] S. W. Golomb, *Shift Register Sequences.* San Francisco: Holden-Day, Inc., 1967.

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

[34] G. Gong, "Theory and applications of q-ary interleaved sequences," *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 400–411, Mar. 1995.

[35] ——, "New designs for signal sets with low cross correlation, balance property, and large linear span: GF(p) case," *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2847–2867, Nov. 2002.

[36] G. Gong and S. W. Golomb, "The decimation-Hadamard transform of two-level autocorrelation sequences," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 853–865, Apr. 2004.

[37] B. Gordon, W. H. Mills, and L. R. Welch, "Some new difference sets," *Canadian J. Math.*, vol. 14, no. 4, pp. 614–625, 1962.

[38] M. Hall, "A survey of difference sets," *Proc. Amer. Math. Soc.*, vol. 7, pp. 975–986, 1956.

[39] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, vol. 35, pp. 126–133, Dec. 1997.

[40] T. Helleseth and P. V. Kumar, "Sequences with low correlation," in *Handbook of Coding Theory*, V. Pless and C. Huffman, Eds. Elsevier Science Publishers, 1998.

[41] D. Hertel, "Cross-correlation properties of perfect binary sequences," *Lecture Notes in Computer Science*, vol. 3486, pp. 208–219, 2005.

[42] H. Hu and D. Feng, *Private communications*, 2006.

[43] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number theory.* New York: Springer-Verlag, 1991, 2nd ed.

[44] J. Jedwab, "A survey of the merit factor problem for binary sequences," *Lecture Notes in Computer Science*, vol. 3486, pp. 30–55, 2005.

[45] D. Jungnickel and A. Pott, "Difference sets: An introduction," in *Difference Sets, Sequences and their Correlation Properties*, A. Pott, P. V. Kumar, T. Helleseth, and D. Jungnickel, Eds. NATO Science Series C, 1999, pp. 259–296.

[46] T. Kasami, "Weight enumerators for several classes of subcodes of the 2nd-order Reed-Muller codes," *Information and Control*, vol. 18, pp. 369–394, 1971.

[47] K. Khoo, G. Gong, and D. R. Stinson, "A new family of Gold-like sequences," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Lausanne, Switzerland, 2002, p. 181.

[48] ——, "A new characterization of semi-bent and bent functions on finite fields," *Designs, Codes, and Cryptography*, vol. 38, no. 2, p. 279.295, Feb. 2006.

[49] S. H. Kim and J. S. No, "New families of binary sequences with low correlation," *IEEE Trans. Inform. Theory*, vol. 49, no. 11, pp. 3059–3065, Nov. 2003.

[50] P. V. Kumar, "On bent sequences and generalized bent functions," Ph.D. dissertation, University of Southern California, Los Angeles, CA, 1983.

[51] P. V. Kumar, T. Helleseth, A. R. Calderbank, and A. R. Hammons, "Large families of quaternary sequences with low correlation," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 579–592, Mar. 1996.

[52] P. V. Kumar and R. A. Scholtz, "Bounds on the linear span of bent sequences," *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 854–862, Nov. 1983.

[53] A. Lempel, M. Cohn, and W. Eastman, "A class of balanced binary sequences with optimal autocorrelation property," *IEEE Trans. Inform. Theory*, no. 1, pp. 38–42, Jan. 1977.

[54] V. I. Levenshtein, "Bounds for codes as solutions of extremum problems for system of orthogonal polynomials," *AAECC-93, Lecture Notes in Computer Science*, vol. 673, pp. 25–42, 1983.

[55] ——, "Lower bounds on crosscorrelation of codes," in *Proceedings of IEEE International Symposium on Spread Spectrum Techniques and Applications*, 1996, pp. 657–661.

[56] R. Lidl and H. Niederreiter, *Finite Fields*. Reading, MA: Addison-Wesley, 1983, vol. 20, Encyclopedia of Mathematics and Its Applications.

[57] H. D. Lüke, "Sequences and arrays with perfect periodic correlation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 3, pp. 287–294, May 1988.

[58] H. D. Lüke, H. D. Schotten, and H. Hadinejad-Mahram, "Binary and quadriphase sequences with optimal autocorrelation properties: A survey," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 3271–3282, Dec. 2003.

[59] W. Ma, M. Lee, and F. Zhang, "A new class of bent functions," *IEICE Trans. Fundamentals*, no. 7, pp. 2039–2040, July 2005.

[60] F. J. MacWilliams and N. J. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[61] A. Maschietti, "Difference sets and hyperovals," *Designs, Codes and Cryptography*, vol. 14, pp. 89–98, 1998.

[62] J. L. Massey, "Shift-register synthesis and BCH decoding," *IEEE Trans. Inform. Theory*, no. 1, pp. 122–127, Jan. 1969.

[63] R. J. McEliece, *Finite Fields for Computer Scientists and Engineers.* Kluwer academic publishers, 1987.

[64] R. L. McFarland, "A new family of noncyclic difference sets," *Journal of Combinatorial Theory, Series A*, pp. 1–10, 1973.

[65] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography.* CRC Press, 2001.

[66] A. A. Nechaev, "Kerdock code in a cyclic form," *Discr. Math. Appl.*, vol. 1, pp. 365–384, 1991.

[67] Y. Niho, "Multi-valued cross-correlation functions between two maximal linear recursive sequences," Ph.D. dissertation, University of Southern California, Los Angeles, CA, 1972.

[68] J. S. No, H. Chung, H. Y. Song, K. Yang, J. D. Lee, and T. Helleseth, "New construction for binary sequences of period p^m-1 with optimal autocorrelation using $(z+1)^d+az^d+b$," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1638–1644, May 2001.

[69] J. S. No, H. Chung, and M. S. Yun, "Binary pseudorandom sequences of period 2^m-1 with ideal autocorrelation generated by the polynomial $z^d+(z+1)^d$," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1278–1282, May 1998.

[70] J. S. No, S. W. Golomb, G. Gong, H. K. Lee, and P. Gaal, "Binary pseudorandom sequences of period 2^m-1 with ideal autocorrelation," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 814–817, Mar. 1998.

[71] J. D. Olsen, R. A. Scholtz, and L. R. Welch, "Bent-function sequences," *IEEE Trans. Inform. Theory*, no. 6, pp. 858–864, Nov. 1982.

[72] K. G. Paterson, "Applications of exponential sums in communication theory," *Lecture Notes in Computer Science*, vol. 1746, pp. 1–24, 1999.

[73] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread-spectrum communications - A tutorial," *IEEE Trans. Commun*, no. 5, pp. 855–884, May 1982.

[74] B. M. Popovic, "Synthesis of power efficient multitone signals with flat amplitude spectrum," *IEEE Trans. Commun.*, vol. 39, no. 7, pp. 1031–1033, July 1991.

[75] ——, "Spreading sequences for multicarrier CDMA systems," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 918–926, June 1999.

[76] M. B. Pursely, "Performance evaluation for phase-coded spread-spectrum multiple-access communication - part I: System analysis," *IEEE Trans. Commun*, no. 8, pp. 795–799, Aug. 1977.

[77] R. Rivest, *The RC4 Encryption Algorithm.* RSA Data Security, Inc., Mar. 1992.

[78] O. S. Rothaus, "On bent functions," *Journal of Combinatorial Theory*, vol. 20, pp. 300–305, 1976.

[79] ——, "Modified gold codes," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 654–656, Mar. 1993.

[80] B. Schmidt, "Cyclotomic integers and finite geometry," *J. Amer. Math. Soc.*, vol. 12, pp. 929–952, 1999.

[81] R. A. Scholtz, P. V. Kumar, and C. Corrada-Bravo, "Some problems and results in ultra-wideband signal design," in *Proceedings of Sequences and Their Applications (SETA'01)*, Bergen, Norway, May 2001.

[82] R. A. Scholtz, D. M. Pozar, and W. Namgoong, "Ultra-Wideband radio," *EURASIP Journal on Applied Signal Processing*, vol. 3, pp. 252–272, 2005.

[83] J. Seberry and M. Yamada, "Hadamard matrices, sequences, and block designs," in *Contemporary Design Theory: A Collection of Surveys*, J. H. Dinitz and D. R. Stinson, Eds. John Wiley & Sons, Inc., 1992.

[84] A. G. Shanbhag, P. V. Kumar, and T. Helleseth, "Improved binary codes and sequence families from Z_4-linear codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 5, pp. 1582–1587, 1996.

[85] V. M. Sidelnikov, "Some k-valued pseudo-random sequences and nearly equidistant codes," *Probl. Inform. Transm.*, vol. 5, no. 1, pp. 12–16, 1969.

[86] ——, "On mutual correlation of sequences," *Soviet Math. Dokl*, vol. 12, pp. 197–201, 1971.

[87] M. K. Simon, J. Omura, R. Scholtz, and K. Levitt, *Spread Spectrum Communications Handbook (revised ed.).* McGraw-Hill, New York, 1994.

[88] R. J. Turyn and J. Storer, "On binary sequences," *Proc. Amer. Math. Soc.*, vol. 12, pp. 394–399, 1961.

[89] P. Udaya, "Polyphase and frequency hopping sequences obtained from finite rings," Ph.D. dissertation, Dept. Elec. Eng., Indian Inst. Technol., Kanpur, India, 1992.

[90] A. J. Viterbi, *CDMA - Principles of Spread Spectrum Communication.* Reading, MA: Addison-Wesley, 1995.

[91] L. R. Welch, "Lower bounds on the maximum cross correlation of the signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.

[92] N. Y. Yu and G. Gong, "Constructions of quadratic bent functions in polynomial forms," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3291–3299, July 2006.

[93] ——, "Crosscorrelation properties of binary sequences with ideal two-level autocorrelation," *Lecture Notes in Computer Science*, vol. 4086, pp. 104–118, 2006.

[94] ——, "Interleaved construction of binary sequences with optimal autocorrelation magnitude," in *Proceedings of IEEE Information Theory Workshop (ITW) 2006*, Chengdu, China, Oct. 2006, accepted.

[95] ——, "A new binary sequence family with low correlation and large size," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1624–1636, Apr. 2006.

[96] ——, "New binary sequences with optmal autocorrelation magnitude," *IEEE Trans. Inform. Theory*, 2006, submitted.

[97] ——, "On asymptotic optimality of binary sequence families," *The Centre for Applied Cryptographic Research (CACR) Technical Report CACR 2006-28*, Aug. 2006, available at http://www.cacr.math.uwaterloo.ca/techreports/2006/cacr2006-28.pdf.

[98] ——, "Quadratic bent functions of polynomial forms and their applications to bent sequences," in *Proceedings of the 23rd Biennial Symposium on Communications*, Kingston, Ontario, Canada, 2006, pp. 128–131.

[99] Y. Zhang, J. G. Lei, and S. P. Zhang, "A new family of almost difference sets and some necessary conditions," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 2052–2061, May 2006.

Appendix A

List of Binary Sequences of Period $2^n - 1$ with Ideal Two-level Autocorrelation ($n = 13, 15, 17, 19$)

Table A.1: Binary two-level autocorrelation sequences of period $2^{13}-1$: $\mathbb{F}_{2^{13}}$ is defined by $x^{13}+x^4+x^3+x+1$

Classes	Trace Representation	Number of Terms
m-sequence	1	1
KPF, B_6	1, 65, 127	3
KPF, B_4	1, 17, 273, 241, 497	5
KPF, B_3	1, 9, 73, 57, 585, 457, 569, 969, 911	9
KPF, B_5	1, 33, 265, 287, 297, 295, 997, 1189, 1181, 1253, 809, 745, 933, 1245, 869	15
KPF, B_2	1, 5, 21, 13, 85, 53, 77, 341, 213, 309, 333, 205, 1365, 853, 1237, 1333, 821, 1357, 845, 1229, 1877, 1703, 1693, 1653, 1639	25
WG	1, 9, 19, 5, 21, 11, 23, 3, 25, 13, 27, 7, 29, 15, 31, 265, 305, 133, 337, 267, 369, 67, 401, 269, 433, 135, 465, 271, 497	29
Glynn I	11, 19, 21, 33, 39, 41, 65, 81, 89, 101, 105, 133, 135, 137, 149, 155, 197, 269, 273, 327, 423, 465, 1173	23
Glynn II	65, 71, 83, 89, 95, 133, 139, 141, 145, 163, 165, 169, 177, 181, 187, 189, 267, 273, 279, 291, 327, 329, 339, 363, 369, 375, 377, 553, 559, 569, 571, 713, 761, 857, 937, 949, 1333	37

Table A.2: Binary two-level autocorrelation sequences of period $2^{15}-1$: $\mathbb{F}_{2^{15}}$ is defined by $x^{15}+x+1$

Classes	Trace Representation	Number of Terms
m-sequence	1	1
KPF, B_7	1, 129, 255	3
KPF, B_4	1, 17, 273, 241, 2185, 1929, 2191, 1937, 1935	9
KPF, B_2	1, 5, 21, 13, 85, 53, 77, 341, 213, 309, 333, 205, 1365, 853, 1237, 1333, 821, 1357, 845, 1229, 5461, 3413, 4949, 5333, 3285, 5429, 3381, 4917, 5453, 3405, 4941, 5325, 3277, 7509, 6823, 6813, 6773, 6567, 6613, 6605, 6557	41
Glynn I	11, 19, 21, 33, 39, 41, 65, 81, 89, 101, 105, 129, 133, 135, 137, 149, 155, 197, 269, 273, 321, 327, 417, 423, 449, 465, 541, 545, 553, 561, 779, 901, 1061, 1093, 1095, 1219, 1315, 1329, 1353, 1637, 1671, 1863, 2189, 2709, 2851	45
Glynn II	1, 5, 13, 17, 29, 37, 49, 61, 65, 77, 101, 113, 125, 133, 145, 157, 193, 205, 229, 241, 253, 261, 273, 285, 321, 333, 357, 369, 381, 389, 401, 413, 449, 461, 485, 497, 509, 523, 531, 535, 541, 547, 555, 561, 567, 579, 583, 589, 595, 613, 619, 631, 637, 651, 653, 657, 663, 675, 677, 689, 701, 711, 723, 747, 753, 759, 1169, 1265, 1329, 1425, 1521, 1809, 1905, 2187, 2217, 2223, 2231, 2247, 2265, 2271, 2283, 2857, 2905, 2917, 3429, 3667, 3691, 5307, 5331	89

Table A.3: Binary two-level autocorrelation sequences of period $2^{17}-1$: $\mathbb{F}_{2^{17}}$ is defined by $x^{17}+x^3+1$

Classes	Trace Representation	Number of Terms
m-sequence	1	1
KPF, B_8	1, 257, 511	3
KPF, B_6	1, 65, 2081, 2111, 2017	5
KPF, B_4	1, 17, 273, 241, 4369, 3857, 4337, 7953, 7711	9
KPF, B_7	1, 129, 1033, 1151, 1161, 9209, 1159, 9289, 9337, 7305, 7241, 9335, 9161, 9145, 7113	15
KPF, B_3	1, 9, 73, 57, 585, 569, 457, 4681, 4665, 4553, 3657, 3641, 18725, 14665, 14633, 14629, 14565, 18727, 14649, 14631, 14921, 14569, 14621, 14793, 14567	25
KPF, B_5	1, 33, 1057, 993, 4229, 3973, 4255, 4261, 4005, 5279, 4259, 4003, 5285, 5029, 20389 5283, 5027, 5221, 4965, 12925, 21141, 20117, 21149, 13477, 13469, 12965, 19861, 12957, 12949, 11925, 21147, 12941, 11917, 20133, 20125, 19877, 11941, 11933, 12917, 20123, 12909	41
KPF, B_2	1, 5, 21, 13, 85, 53, 77, 341, 213, 309, 333, 205, 1365, 853, 1237, 1333, 821, 1357, 845, 1229, 5461, 3413, 4949, 5333, 3285, 5429, 3381, 4917, 5453, 3405, 4941, 5325, 3277, 21845, 13653, 19797, 21333, 13141, 21717, 13525, 19669, 21813, 13621, 19765, 21301, 13109, 21837, 13645, 19789, 21325, 13133, 21709, 13517, 19661, 30037, 27303, 27293, 27253, 26279, 27093, 27047, 26269, 26453, 26421, 26445, 26229, 26215	67
WG	1041, 2083, 521, 2085, 1043, 2087, 261, 2089, 1045, 2091, 523, 2093, 1047, 2095, 131, 2097, 1049, 2099, 525, 2101, 1051, 2103, 263, 2105, 1053, 2107, 527, 2109, 1055, 2111, 97, 161, 225, 289, 353, 417, 481, 545, 609, 673, 737, 801, 865, 929, 993, 1057, 1121, 1185, 1249, 1313, 1377, 1441, 1505, 1569, 1633, 1697, 1761, 1825, 1889, 1953, 2017	61
Glynn I	1, 5, 13, 17, 29, 33, 45, 49, 61, 67, 71, 75, 77, 81, 85, 87, 145, 163, 177, 201, 225, 259, 265, 267, 271, 337, 385, 405, 433, 457, 529, 547, 553, 565, 571, 643, 651, 655, 661, 773, 835, 841, 931, 1033, 1039, 1041, 1047, 1129, 1177, 1419, 1549, 1611, 1617, 1623, 1713, 1803, 2081, 2085, 2091, 2179, 2185, 2191, 2309, 2357, 2371, 2381, 2575, 2765, 2849, 3101, 3139, 3337, 3343, 3721, 3727, 3907, 4229, 4233, 4281, 4433, 4677, 5769, 8777, 8793, 8873, 10797, 13465	87
Glynn II	257, 263, 275, 281, 287, 323, 329, 335, 353, 359, 371, 377, 383, 517, 523, 525, 529, 547, 549, 553, 561, 565, 571, 573, 577, 643, 645, 649, 657, 661, 667, 669, 673, 705, 709, 715, 717, 721, 739, 741, 745, 753, 757, 763, 765, 1035, 1041, 1047, 1059, 1089, 1095, 1107, 1131, 1137, 1143, 1155, 1287, 1289, 1299, 1313, 1323, 1329, 1335, 1337, 1347, 1419, 1425, 1431, 1433, 1443, 1473, 1479, 1481, 1491, 1505, 1515, 1521, 1527, 1529, 2081, 2089, 2095, 2105, 2107, 2119, 2185, 2191, 2201, 2203, 2209, 2215, 2249, 2251, 2257, 2263, 2273, 2281, 2287, 2297, 2299, 2577, 2673, 2825, 2849, 2865, 2873, 2961, 2969, 3017, 3041, 3057, 3065, 3353, 3401, 3425, 3449, 3601, 3625, 3633, 3637, 3721, 3729, 3733, 3745, 3781, 3793, 3817, 3825, 3829, 4233, 4239, 4377, 4383, 4395, 4405, 4425, 4431, 4491, 4501, 4521, 4527, 4549, 4551, 4569, 4575, 4587, 4597, 5173, 5269, 5317, 5365, 6693, 6885, 9005, 9101, 9197, 9581, 9769, 9817, 9913, 9925, 9961, 10587, 10635, 10683, 10695, 10707, 11443, 11467, 11491, 13005, 13767, 22899	173

Table A.4: Binary two-level autocorrelation sequences of period $2^{19}-1$: $\mathbb{F}_{2^{19}}$ is defined by $x^{19}+x^5+x+1$

Classes	Trace Representation	Number of Terms
m-sequence	1	1
KPF, B_9	1, 513, 1023	3
KPF, B_6	1, 65, 4161, 4033, 8129	5
KPF, B_5	1, 33, 1057, 993, 16913, 15889, 16927, 15905, 15903,	9
KPF, B_4	1, 17, 273, 241, 4369, 4337, 3857, 34953, 30865, 30857, 34959, 30863, 30993, 30841, 30961	15
KPF, B_3	1, 9, 73, 57, 585, 457, 569, 4681, 3657, 4553, 4665, 3641, 37449, 29257, 36425, 37321, 29129, 37433, 29241, 36409, 62025, 58511, 58489, 58313, 58255	25
KPF, B_8	1, 257, 2057, 2303, 2313, 2311, 18425, 18505, 14409, 14601, 18681, 18679, 18761, 18759, 14665, 18745, 18743, 63785, 57125, 63273, 75045, 59685, 58661, 74981, 74973, 75557, 59173, 75493, 51529, 51017, 51527, 47433, 46921, 59689, 56613, 59177, 59687, 75485, 55593, 56101, 55081	41
KPF, B_7	1, 129, 4129, 4223, 4257, 4255, 16261, 17029, 17021, 20357, 12449, 12193, 20645, 21117, 20389, 20643, 20861, 12453, 21115, 12197, 21157, 21149, 81557, 20901, 20893, 21155, 21147, 48789, 12965, 12957, 81301, 12709, 12701, 84629, 84597, 85653, 53925, 52901, 53909, 84589, 52885, 51861, 51829, 85397, 50837, 50805, 78501, 77477, 78493, 45733, 44709, 53907, 77221, 52883, 80533, 85621, 79509, 53877, 52853, 47765, 85365, 46741, 80277, 85619, 79253, 53875, 52851	67
KPF, B_2	1, 5, 21, 13, 85, 53, 77, 341, 213, 309, 333, 205, 1365, 853, 1237, 1333, 821, 1357, 845, 1229, 5461, 3413, 4949, 5333, 3285, 5429, 3381, 4917, 5453, 3405, 4941, 5325, 3277, 21845, 13653, 19797, 21333, 13141, 21717, 13525, 19669, 21813, 13621, 19765, 21301, 13109, 21837, 13645, 19789, 21325, 13133, 21709, 13517, 19661, 87381, 54613, 79189, 85333, 52565, 86869, 54101, 78677, 87253, 54485, 79061, 85205, 52437, 87349, 54581, 79157, 85301, 52533, 86837, 54069, 78645, 87373, 54605, 79181, 85325, 52557, 86861, 54093, 78669, 87245, 54477, 79053, 85197, 52429, 120149, 109223, 109213, 109173, 105127, 109013, 108199, 105117, 108373, 108341, 108189, 105077, 104871, 105813, 105685, 105781, 105805, 105063, 104917, 104909, 104861	109
WG	1, 33, 67, 17, 69, 35, 71, 9, 73, 37, 75, 19, 77, 39, 79, 5, 81, 41, 83, 21, 85, 43 87, 11, 89, 45, 91, 23, 93, 47, 95, 3, 97, 49, 99, 25, 101, 51, 103, 13, 105, 53, 107, 27, 109, 55, 111, 7, 113, 57, 115, 29, 117, 59, 119, 15, 121, 61, 123, 31, 125, 63, 127, 4129, 4289, 2065, 4417, 4131, 4545, 1033, 4673, 4133, 4801, 2067, 4929, 4135, 5057, 517, 5185, 4137, 5313, 2069, 5441, 4139, 5569, 1035, 5697, 4141, 5825, 2071, 5953, 4143, 6081, 259, 6209, 4145, 6337, 2073, 6465, 4147, 6593, 1037, 6721, 4149, 6849, 2075, 6977, 4151, 7105, 519, 7233, 4153, 7361, 2077, 7489, 4155, 7617, 1039, 7745, 4157, 7873, 2079, 8001, 4159, 8129	125

Table A.5: Binary two-level autocorrelation sequences of period $2^{19}-1$: $\mathbb{F}_{2^{19}}$ is defined by $x^{19}+x^5+x+1$ (*Cont'd*)

Classes	Trace Representation	Number of Terms
Glynn I	1, 5, 13, 17, 29, 33, 45, 49, 61, 67, 71, 75, 77, 81, 85, 87, 145, 163, 177, 201, 225, 259, 265, 267, 271, 337, 385, 405, 433, 457, 529, 547, 553, 565, 571, 643, 651, 655, 661, 773, 835, 841, 931, 1027, 1033, 1039, 1041, 1047, 1129, 1153, 1177, 1419, 1549, 1611, 1617, 1623, 1713, 1803, 1921, 2081, 2085, 2091, 2113, 2177, 2179, 2185, 2191, 2309, 2357, 2369, 2371, 2381, 2563, 2569, 2575, 2753, 2765, 2849, 2945, 3089, 3101, 3139, 3331, 3337, 3343, 3715, 3721, 3727, 3845, 3907, 4185, 4201, 4229, 4233, 4281, 4291, 4305, 4385, 4433, 4621, 5153, 5267, 5769, 5921, 6193, 6205, 6229, 6305, 7185, 7191, 7693, 8327, 8331, 8421, 8461, 8469, 8517, 8525, 8713, 8719, 9091, 9245, 9283, 9377, 9479, 9481, 9487, 9569, 9617, 9859, 10345, 10435, 10437, 10531, 10821, 11015, 11399, 11541, 12449, 12553, 12559, 12745, 12939, 13417, 13465, 13707, 14669, 14863, 15625, 15631, 16913, 16925, 16941, 17105, 17285, 17543, 17735, 19237, 21065, 21161, 21571, 22725, 23085, 23687, 25009, 26933, 42277	167
Glynn II	1, 5, 13, 17, 29, 37, 49, 61, 65, 77, 101, 113, 125, 133, 145, 157, 193, 205, 229, 241, 253, 257, 269, 293, 305, 317, 389, 401, 413, 449, 461, 485, 497, 509, 517, 529, 541, 577, 589, 613, 625, 637, 769, 781, 805, 817, 829, 901, 913, 925, 961, 973, 997, 1009, 1021, 1029, 1041, 1053, 1089, 1101, 1125, 1137, 1149, 1281, 1293, 1317, 1329, 1341, 1413, 1425, 1437, 1473, 1485, 1509, 1521, 1533, 1541, 1553, 1565, 1601, 1613, 1637, 1649, 1661, 1793, 1805, 1829, 1841, 1853, 1925, 1937, 1949, 1985, 1997, 2021, 2033, 2045, 2059, 2067, 2071, 2077, 2083, 2091, 2097, 2103, 2115, 2119, 2125, 2131, 2149, 2155, 2167, 2173, 2179, 2187, 2193, 2199, 2211, 2241, 2247, 2259, 2283, 2289, 2295, 2307, 2311, 2317, 2323, 2341, 2347, 2359, 2365, 2371, 2437, 2443, 2455, 2461, 2467, 2503, 2509, 2515, 2533, 2539, 2551, 2557, 2571, 2573, 2577, 2583, 2595, 2597, 2609, 2621, 2625, 2631, 2643, 2667, 2673, 2679, 2691, 2693, 2705, 2717, 2753, 2765, 2789, 2801, 2813, 2823, 2835, 2859, 2865, 2871, 2883, 2955, 2961, 2967, 2979, 3009, 3015, 3027, 3051, 3057, 3063, 4625, 4673, 4721, 4913, 5009, 5057, 5105, 5169, 5265, 5313, 5361, 5649, 5697, 5745, 5937, 6033, 6081, 6129, 7185, 7233, 7281, 7473, 7569, 7617, 7665, 8323, 8329, 8331, 8335, 8361, 8367, 8391, 8409, 8415, 8427, 8457, 8463, 8467, 8473, 8479, 8497, 8503, 8515, 8521, 8527, 8593, 8599, 8611, 8617, 8623, 8659, 8665, 8671, 8689, 8695, 8715, 8745, 8751, 8753, 8759, 8775, 8793, 8799, 8811, 8841, 8847, 8849, 8855, 8903, 8945, 8951, 8967, 8985, 8991, 9003, 9033, 9039, 9099, 9129, 9135, 9159, 9177, 9183, 9195, 10769, 10865, 11057, 11153, 11249, 11305, 11353, 11365, 11377, 11401, 11545, 11557, 11569, 11593, 11653, 11665, 11689, 11737, 11749, 11761, 12401, 12593, 12689, 12785, 13413, 13605, 13701, 13797, 13841, 13937, 14129, 14225, 14321, 14419, 14443, 14611, 14635, 14659, 14731, 14755, 14803, 14827, 14897, 14993, 15089, 17509, 17515, 17701, 17707, 17797, 17803, 17893, 17899, 20667, 20679, 20691, 20763, 20779, 20875, 20971, 21051, 21063, 21075, 21147, 21267, 21339, 21387, 21435, 21447, 21459, 23653, 23845, 24037, 25189, 25381, 25573, 25677, 25869, 26061, 26823, 26917, 27109, 27207, 27591, 35417, 35609, 35801, 35929, 35977, 36025, 36169, 36217, 36313, 37337, 38489, 38681, 38873, 39091, 39283, 39307, 39379, 39641, 45415, 45463, 45511, 45747, 45939, 45987, 50765, 51149, 51917, 52119, 60745, 60841, 76213	383