Sequence Design and Construction of Cryptographic Boolean Functions

by

Khoongming Khoo

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2013

©Khoongming Khoo 2013
I hereby declare that I am the sole author of this thesis.

I authorize the University of Waterloo to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the University of Waterloo to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.
The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.
Abstract

In this thesis, we study the application of well-known sequences from communications theory, in the construction of cryptographic Boolean functions. First, we explain the basic connection between binary sequences, polynomials over the finite field $GF(2^n)$, and Boolean functions. Second, we classify highly nonlinear quadratic polynomials which may have useful applications in the design and cryptanalysis of finite field based crypto-systems. Third, we construct Boolean functions to be used in stream cipher systems. They have good cryptographic properties like balance, resiliency, high nonlinearity and low additive auto-correlation for protection against various statistical attacks. Fourth, we explore using an S-box instead of a Boolean function in stream cipher systems for higher communication speed. However in that case, we require the S-box to have an additional requirement: low maximum correlation. We construct S-boxes with good cryptographic properties and having low maximum correlation which improves on currently known bounds. These Boolean function and S-box constructions are based on the m-sequences, GMW sequences and ideal 2-level auto-correlation sequences from communications theory. Finally, we consider efficient methods to compute the Hadamard transform of polynomials, which is useful in the design of Boolean functions from sequences. We also present some experimental results to demonstrate the application of these methods.
Acknowledgements

I would like to thank my supervisors Professor Guang Gong and Professor Douglas R. Stinson for their patient and caring guidance. They have been very generous with their time and effort to teach me about cryptography and the theory of sequences, and I am grateful to them for inspiring the many interesting ideas presented in this thesis. Their supervision made this research a very enjoyable experience.

Special thanks goes out to DSO National Laboratories of Singapore, who sponsored my PhD program. They have shown me great faith and generosity throughout my studies.

I would like to thank the many friends I met at Waterloo. To Charles Chung and Ada Chan who helped me settle down and feel welcomed when I first came to Canada. The friends I met while at Minota Hagey: Hermia Ip, John Tsang, Jennifer Lee, Richard Chen, Raymond Li, Lawrence Tang, Paul Goswami and Han Gang. Especially Hermia and Jennifer who shared their good food with my undercooked rice for many memorable dinners. The late-night gatherings, table-tennis and tennis games, and outings made school life really fun.

Thanks also to Charles Lam and Shaoquan Jiang for their intellectual interactions. To Shaoquan for letting me stay over at his place to attend school at Waterloo while I was in Toronto.

I would like to thank my family who showed care and concern for my well-being while in Canada with their phone-calls and well wishes.

Finally, I would like to thank my girlfriend Hermia who shared her life with me for the past three years. Her caring company made my stay in Canada the most memorable three years of my life, and her emotional support and delicious food sustained me throughout my program. She is definitely a most positive influence in my life. Last but certainly not least, I wish to thank God for his many blessings and the completion of this thesis.
To my family and girlfriend Hermia.
# Contents

1 Introduction

2 Relationships between Sequences, Polynomials and Boolean Functions
   2.1 Sequences and Polynomial Functions
   2.2 Polynomial and Boolean Functions
   2.3 Cryptographic Requirements of Boolean Functions
   2.4 Cryptographic Requirements of S-boxes

3 Optimal Quadratic Polynomials
   3.1 Characterization of Optimal Quadratic Polynomials
   3.2 Optimal Quadratic Polynomials for All Choices of Coefficients
      3.2.1 $p$ is odd prime and $ord_p(2) = p - 1$
      3.2.2 $p = 2s + 1$ is prime, $s$ is odd and $ord_p(2) = s$
   3.3 Characterizing Optimal Quadratic Polynomials
   3.4 Summary of Chapter 3

4 Additive Autocorrelation of Resilient Boolean Functions
   4.1 Cryptographic Properties of the Dual Function
      4.1.1 Correlation Immunity and Non-Affine Dual Function
7 Computational Results for Boolean Functions based on $GF(2^n)$

7.1 Efficient Search for Highly Nonlinear Resilient Boolean Functions

7.1.1 Example: Search for Resilient Preferred Functions in $GF(2^{11})$

7.1.2 Search Result

7.2 The Patterson-Wiedemann Functions

8 Conclusion

9 Future Work

Bibliography
Chapter 1

Introduction

The theory of sequences has a long and rich history, with wide ranging applications in communication systems, random number generation and cryptography.

One notable application of sequences in communication systems is the construction of code division multiple access (CDMA) systems [33]. In traditional communication systems, different users have to use different carrier frequencies to avoid interference of transmission signals. But in a CDMA system, they can all use the same frequency because each user can identify his message by computing the cross correlation between his signature sequence and the message’s carrier sequence. Only the message meant for him will have high cross-correlation while the message of other users have low cross-correlation. The family of sequences which are used in CDMA applications includes certain m-sequences [32] and the interleaved sequences [26].

In random number generation, linear and nonlinear feedback shift register sequences have been used to compute pseudorandom strings from a purely random seed. These pseudorandom keystreams can be used as part of a stream cipher for asynchronous encryption in cryptography or simply to create a statistically uniform sample for experimental testing. Golomb postulated four properties a pseudorandom sequence should possess to behave like a random sequence [23]. These are balance, equal distribution of runs, ideal 2-level autocorrelation (i.e. zero autocorrelation for non-zero shifts) and uniform \(k\)-tuple distribution. The m-sequences were shown to possess all four properties. On a related note, the property of ideal 2-level autocorrelation is particularly attractive because it can be used to construct sets of orthogonal sequences for engineering applications and
has a close connection with cyclic Hadamard difference sets in design theory [23]. Besides the m-sequences, the GMW sequences by Gordon, Mills and Welch [30] were shown to possess the ideal 2-level autocorrelation property. Recently, Dillon and Dobbertin proved the validity of all conjectured 2-level autocorrelation sequences using a technique based on Parseval’s equation [16, 17, 18].

The early applications of sequences in cryptography mainly involved the use of Linear Feedback Shift Registers in stream ciphers. Techniques from sequence analysis like Berlekamp-Massey Decoding and Discrete Fourier Transforms were used to determine the linear span and algebraic complexity of the keystream. In 1993, Nyberg pioneered the use of power functions for constructing cryptographic S-boxes to be used in block ciphers [60]. The S-boxes were highly nonlinear and resistant to differential attacks. These power functions are closely related to the m-sequences used in CDMA systems [32]. This led to increased attention in such m-sequences by the cryptographic community and led to a proof of the Welch’s conjecture (first stated in the early 70’s) by Canteaut, Dobbertin and Charpin in 1998 [5]. During the same period that m-sequences were introduced for S-box construction, Chang, Dai and Gong used the same m-sequences to construct highly nonlinear Boolean functions which satisfy the strict avalanche criteria [10]. Later in 2000, Gong and Youssef used other sequences like the Welch-Gong Transformation sequences to construct resilient Boolean functions with high nonlinearity [29]. At Eurocrypt 2001, the same authors applied ideas from interleaved sequences to construct hyper-bent functions which have large distances from all monomials [75]. Finally, we note that sequences can also be used to construct public key cryptosystems. In 1999, Gong and Harn proposed the GH public key cryptosystem which is based on the 3rd-order shift register sequences in a finite field $GF(q)$ [27]. At Crypto 2000, Lenstra and Verheul proposed the XTR public key cryptosystem [48] which can be viewed as a special case of the GH cryptosystem.

In this thesis, we continue the study of sequences in cryptography. Our main area of study will be in the construction of cryptographic Boolean functions. We extend some known constructions and also introduce new application of sequences. These are summarized in the following paragraphs.

There is a close relationship between sequences of period $2^n - 1$, polynomials over $GF(2^n)$, and Boolean functions with $n$ input bits. In Chapter 2, we explore the inter-connections between them and how they can be applied to cryptography. Some useful properties which protect
CHAPTER 1. INTRODUCTION

the Boolean functions and S-boxes against well-known cryptanalytic attacks are also introduced. These properties include balance, high nonlinearity, resiliency, low additive autocorrelation, low maximal differential and low maximal correlation.

Boolean functions with high nonlinearity are important in cryptography because they provide protection against linear cryptanalysis [53]. In general, it is not an easy problem to identify all Boolean functions with high nonlinearity. However, this problem has been completely solved for quadratic Boolean functions [54]. In Chapter 3, we attempt to solve a related problem in the finite field \( GF(2^n) \), i.e. to find all quadratic polynomials with optimal nonlinearity. Although we cannot classify all such polynomials, we manage to identify large classes of optimal quadratic polynomials based on linear combinations of the Gold function \( Tr(x^{2^i+1}) \) [22]. The motivation for our study is that finite fields have played an increasingly important role in cryptography over the years. They can be found in applications like the Digital Signature Standard, the elliptic curve cryptosystem and the GH, XTR cryptosystems. Therefore these quadratic polynomials with high nonlinearity may be useful in the construction and cryptanalysis of cryptographic systems based on finite fields.

Another important criteria for a Boolean function is resiliency, which ensures it cannot be approximated by linear functions with few variables. When used in stream cipher systems, a Boolean function is required to have high nonlinearity and resiliency for protection against correlation attacks [71]. We might also want the Boolean function to have low additive autocorrelation for protection against differential-like cryptanalysis [2]. In Chapter 4, we find Boolean functions which are resilient, have high nonlinearity and optimally low additive autocorrelation. To achieve this, we introduce a new notion called the dual function to analyze Boolean functions whose Hadamard transform only takes on 3 values. Then we use the dual function to deduce that some 2-level autocorrelation sequences, recently found by Dillon and Dobbertin [16, 17, 18], give rise to Boolean functions which achieve the above mentioned properties.

In Chapter 5, we construct highly nonlinear Boolean functions which are balanced or resilient. Our construction is based on the following idea. If we compose a balanced (or resilient) highly nonlinear Boolean function with a cascaded GMW sequence function [44], then the resulting function is also highly nonlinear and balanced (or resilient). Our main result in that chapter is the construction of balanced Boolean functions whose nonlinearity exceeds the quadratic bound \( 2^{n-1} - 2^{(n-1)/2} \). Our construction of such functions is based on a new approach using cascaded
GMW sequences; whereas previous approaches were to take the direct sum of two highly nonlinear Boolean functions, one of which is balanced.

In Chapter 6, we consider using an S-box as a filter function in a stream cipher system. One advantage of an S-box over a Boolean function in such a system is that more bits of the keystream are generated per clock cycle. But in exchange for a higher transmission rate, we need an additional security requirement: low maximal correlation [77]. This ensures that the cipher is not prone to a general correlation attack using nonlinear functions of output bits of the S-box. Based on the m-sequences and GMW sequences, we construct two classes of balanced S-boxes with high nonlinearity that improves the best known bound on maximum correlation [77] by a factor of $\sqrt{2}$. Thus, our S-boxes offer better protection against general correlation attacks.

The cryptographic properties of a Boolean function depend on its Hadamard transform. For example, a Boolean function has high nonlinearity if it has low Hadamard transform at all vectors. And it is resilient of order $k$ if its Hadamard transform at vectors with weight not greater than $k$ is zero. In Chapter 7, we study efficient computations for the Hadamard transform of polynomials over $GF(2^n)$. Then we use these fast transform methods to exhaustively search for certain polynomials over $GF(2^n)$ which correspond to cryptographically useful Boolean functions.
Chapter 2

 Relationships between Sequences, Polynomials and Boolean Functions

An overview of some basic concepts and definitions commonly used in later chapters are presented. First, we explain how binary sequences of period $2^n - 1$ can be viewed as polynomials over $GF(2^n)$. Thus, we will always present sequences in their polynomial form in later chapters. Then we show that polynomial functions on $GF(2^n)$ correspond naturally to Boolean functions with $n$ input bits. Therefore we can construct cryptographic Boolean functions through their inter-connections with sequences and polynomials on $GF(2^n)$. Finally we present some cryptographic requirements of Boolean functions and S-boxes for stream and block cipher applications.

Let $\mathbb{Z}_2[x]$ be the ring of all binary polynomials and let $f(x) \in \mathbb{Z}_2[x]$ be an irreducible polynomial of degree $n$. Then the quotient ring $\mathbb{Z}_2[x]/(f(x))$ is a finite field with $2^n$ elements. We denote this finite field by $GF(2^n)$. It can be proven that the multiplicative group $GF(2^n)^*$ is cyclic [35] and there exist polynomials $f(x)$ such that $x^i \mod f(x)$, $i = 0, \ldots, 2^n - 2$, generate all the elements of $GF(2^n)^*$. Such an $f(x)$ is called a primitive polynomial [23]. Therefore $GF(2^n)^*$ always contains a primitive element $\alpha$ and all other primitive elements of the group are given by $\alpha^i$, $\gcd(i, 2^n - 1) = 1$. 

CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

### Notation 1.

In this thesis, we denote a primitive element of $GF(2^n)$ by $\alpha$.

The trace function $Tr_m^n : GF(2^n) \rightarrow GF(2^m)$ is a linear function defined by:

$$Tr_m^n(x) = \sum_{i=0}^{n/m-1} x^{2^i m}.$$  

It satisfies the following relations:

$$Tr_m^n(x + y) = Tr_m^n(x) + Tr_m^n(y) \text{ for all } x, y \in GF(2^n),$$

$$Tr_m^n(ax) = aTr_m^n(x) \text{ for all } a \in GF(2^m), x \in GF(2^n).$$

The special case $m = 1$ is commonly used in the construction of binary sequences and Boolean functions for engineering and cryptographic applications. In that case, the trace function is:

$$Tr_1^n(x) = \sum_{i=0}^{n-1} x^{2^i}.$$  

We usually denote $Tr_1^n(x)$ by $Tr(x)$.

### 2.1 Sequences and Polynomial Functions

Our focus is mainly on binary sequences with period $2^n - 1$. We usually denote them by

$$a_0, a_1, \ldots, a_{2^n - 2}$$

where each $a_i = 0$ or $1$. They have many useful applications in engineering like the CDMA communications system.

A binary sequence with period $2^n - 1$ can be represented by a polynomial $f : GF(2^n) \rightarrow GF(2)$. We say $f(x)$ is the trace (or polynomial) representation of the binary sequence $a = \{a_i\}_{i=0}^{2^n-2}$ if $a_i = f(\alpha^i)$ for all $i$.

Let $f(x) = \sum_{i=0}^{2^n-2} A_i x^i$ be the polynomial representation of the binary sequence $a = \{a_k\}_{k=0}^{2^n-2}$ on $GF(2^n)$. Then the following formula [24] determines the coefficients of $f(x)$:

$$A_i = \sum_{k=0}^{2^n-2} a_k \alpha^{-ik}. \quad (2.1)$$
It can be proven [24] that
\[ A_{i}^{2^j} = A_{2^j i}, \quad j = 0, \ldots, n - 1. \] (2.2)

The cyclotomic coset of \( i \) is defined to be the set:
\[ C_i = \{ 2^j i \mod 2^n - 1, j = 0, 1, \ldots, n - 1 \}. \]

We call the smallest element of each cyclotomic coset a cyclotomic coset leader.

By equation (2.2), \( f(x) \) can be written in trace representation form given by:
\[ f(x) = \sum_i Tr_{n_i}(A_i x^i), \] (2.3)
where the sum is taken over all cyclotomic coset leaders \( i \) and \( n_i \) is the size of the cyclotomic coset \( C_i \). For more details, please refer to [24].

**Example 1.** Let us look at the binary sequence of period \( 2^4 - 1 = 15 \) given by:
\[ a = 010100111000001. \]

To obtain the trace representation of \( a \), we need to compute the coefficients \( A_i \) in equation (2.3) for the cyclotomic coset leaders \( i \). The set \( \mathbb{Z}_{15} \) is partitioned into the following cyclotomic cosets:
\[ C_0 = \{ 0 \}, C_1 = \{ 1, 2, 4, 8 \}, C_3 = \{ 3, 6, 12, 9 \}, C_5 = \{ 5, 10 \}, C_7 = \{ 7, 14, 13, 11 \}. \]

Thus the cyclotomic coset leaders are 0, 1, 3, 5, 7.

We present in Table 2.1 the elements of \( GF(2^4)^* \), which is needed in the computation of \( A_i \). The powers of the primitive element \( \alpha \) are computed using the relation \( \alpha^4 + \alpha + 1 = 0 \).

We note that \( a_k = 1 \) for \( k \in \{ 1, 3, 6, 7, 8, 14 \} \), otherwise \( a_k = 0 \). By equation (2.1), the
Table 2.1: Elements of $GF(2^4)^*$ generated by $\alpha^4 + \alpha + 1 = 0$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\alpha^i$</th>
<th>$i$</th>
<th>$\alpha^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>8</td>
<td>$\alpha^2 + 1$</td>
</tr>
<tr>
<td>1</td>
<td>$\alpha$</td>
<td>9</td>
<td>$\alpha^3 + \alpha$</td>
</tr>
<tr>
<td>2</td>
<td>$\alpha^2$</td>
<td>10</td>
<td>$\alpha^2 + \alpha + 1$</td>
</tr>
<tr>
<td>3</td>
<td>$\alpha^3$</td>
<td>11</td>
<td>$\alpha^3 + \alpha^2 + \alpha$</td>
</tr>
<tr>
<td>4</td>
<td>$\alpha + 1$</td>
<td>12</td>
<td>$\alpha^3 + \alpha^2 + \alpha + 1$</td>
</tr>
<tr>
<td>5</td>
<td>$\alpha^2 + \alpha$</td>
<td>13</td>
<td>$\alpha^3 + \alpha^2 + 1$</td>
</tr>
<tr>
<td>6</td>
<td>$\alpha^3 + \alpha^2$</td>
<td>14</td>
<td>$\alpha^3 + 1$</td>
</tr>
<tr>
<td>7</td>
<td>$\alpha^3 + \alpha + 1$</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

coefficients $A_i$ are:

$$ A_0 = \sum_{k=1}^{14} a_k = 0. $$

$$ A_1 = \sum_{k=1}^{14} a_k \alpha^{-k} = \alpha + \alpha^7 + \alpha^8 + \alpha^9 + \alpha^{12} + \alpha^{14} = 0. $$

$$ A_3 = \sum_{k=1}^{14} a_k \alpha^{-3k} = \alpha^3 + \alpha^9 = \alpha. $$

$$ A_5 = \sum_{k=1}^{14} a_k \alpha^{-5k} = 0. $$

$$ A_7 = \sum_{k=1}^{14} a_k \alpha^{-7k} = \alpha^3 + \alpha^4 + \alpha^7 + \alpha^8 + \alpha^9 + \alpha^{11} = 1. $$

Therefore, the trace representation of $a = 010100111000001$ is:

$$ f(x) = Tr_4(\alpha x^3 + x^7). $$

That is, $a_i = f(\alpha^i)$.

An m-sequence is a sequence whose trace representation is given by $Tr(x^r)$, for some $r$, $1 \leq r \leq 2^n - 2$. Let $m$ be the m-sequence represented by $Tr(x)$. The cross correlation of a binary
sequence \( a = \{a_k\}_{k=0}^{2^n-2} \) with \( m \) is:

\[
C_{a,m}(\tau) = \sum_{k=0}^{2^n-2} (-1)^{m_k+\tau+a_k}, \quad 0 \leq \tau \leq 2^n - 2.
\]

Let \( f : GF(2^n) \to GF(2) \) be the trace representation of the sequence \( a \), there is a related concept called the Hadamard transform of \( f \) which is defined by:

\[
\hat{f}(\lambda) = \sum_{x \in GF(2^n)} (-1)^{Tr(\lambda x)+f(x)}.
\]

(2.4)

We can deduce that

\[
\hat{f}(\lambda) = 1 + C_{a,m}(\tau)
\]

where \( m_k+\tau = Tr(\alpha^{k+\tau}), \lambda = \alpha^\tau \) and \( a_k = f(\alpha^k) \).

**Example 2.** Let us consider sequences of period \( 2^5 - 1 = 31 \) and the related finite field \( GF(2^5) \) generated by the relation \( \alpha^3 + \alpha^3 + 1 = 0 \) where \( \alpha \) is primitive. The \( m \)-sequence \( m \) represented by \( Tr_5^1(x) \) is given by:

\[
m = 100000101011100011110011001.
\]

Consider the sequence:

\( a = 0111111111010111100011000100100 \),

which is represented by the polynomial \( f(x) = Tr_5^1(x^3 + x^5) \).

We will compute the cross-correlation between \( a \) and \( m \) for no shift, i.e. \( \tau = 0 \). In the sequence

\[
\{a_i + m_i\}_{i=0}^{30} = 111011001100111110001101000000,
\]

there are sixteen 1’s and fifteen 0’s which implies

\[
C_{a,m}(0) = \sum_{i=0}^{30} (-1)^{a_i+m_i} = -1 \quad \text{and} \quad \hat{f}(1) = 1 + C_{a,m}(0) = 0.
\]

The cross-correlation between \( a \) and \( m \) for a shift of \( \tau = 1 \) is based on:

\[
\{a_i + m_{i+1}\}_{i=0}^{30} = 0111010100000110100000100001101.
\]

There are twelve 1’s and nineteen 0’s which implies

\[
C_{a,m}(1) = \sum_{i=0}^{30} (-1)^{a_i+m_{i+1}} = 7 \quad \text{and} \quad \hat{f}(\alpha) = 1 + C_{a,m}(1) = 8.
\]

In fact, it can be verified that \( C_{a,m}(\tau) \in \{-9,-1,7\} \) for all \( 0 \leq \tau \leq 30 \) and that \( \hat{f}(\lambda) \in \{0, \pm 8\} \) for all \( \lambda \in GF(2^5) \).
From the relationship between $\hat{f}(\lambda)$ and $C_{a,m}(\tau)$, we see that binary sequences with low cross-correlation from CDMA applications correspond to highly nonlinear polynomials that have large distances from all affine functions.

2.2 Polynomial and Boolean Functions

A Boolean function is a function from $GF(2)^n$ to $GF(2)$. They are widely used as cryptographic components in stream ciphers, block ciphers and hash functions, where they are required to satisfy various uniformity and nonlinearity criteria.

There is a natural correspondence between Boolean functions $g : GF(2)^n \rightarrow GF(2)$ and polynomial functions $f : GF(2^n) \rightarrow GF(2)$. Let $\{\alpha_{n-1}, \ldots, \alpha_0\}$ be a basis of $GF(2^n)$ and $g(x)$ be the Boolean function representation of $f(x)$, then this correspondence is given by:

$$g(x_{n-1}, \ldots, x_1, x_0) = f(x_{n-1}\alpha_{n-1} + \cdots + x_1\alpha_1 + x_0\alpha_0).$$  \hspace{1cm} (2.5)

Example 3. Let us look at the Boolean representation of the polynomial $Tr_3^1(x^3)$ on $GF(2^3) = \mathbb{Z}_2[x]/(x^3 + x + 1)$.

In Table 2.2, we look at the powers of the primitive element $\alpha$ which is generated from the relation $\alpha^3 = \alpha + 1$. From it, we can compute $Tr(x^3)$ for $x = \alpha^i$, $i = 0, 1, \ldots, 6$. For example, when $x = \alpha$,

$$Tr(\alpha^3) = \alpha^3 + \alpha^6 + \alpha^5 = (\alpha + 1) + (\alpha^2 + 1) + (\alpha^2 + \alpha + 1) = 1.$$

Using the basis $\{\alpha^2, \alpha, 1\}$, the Boolean representation of $f(x) := Tr(x^3)$ is given by:

$$g(x_2, x_1, x_0) = f(x_2\alpha^2 + x_1\alpha + x_0).$$

The truth (output) table of $g : GF(2)^3 \rightarrow GF(2)$ is given in Table 2.3.

The algebraic expression of $g(x)$ is the quadratic function:

$$g(x_2, x_1, x_0) = x_2x_1 + x_2 + x_1 + x_0.$$

The Hadamard Transform of a Boolean function $f : GF(2)^n \rightarrow GF(2)$ is:

$$\hat{f}(w) = \sum_{x \in GF(2)^n} (-1)^{w \cdot x + f(x)}. \hspace{1cm} (2.6)$$
Table 2.2: The function $Tr(x^3)$ on $GF(2^3)$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$x = \alpha^i$</th>
<th>$Tr(x^3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>$\alpha$</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$\alpha^2$</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>$\alpha + 1$</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>$\alpha^2 + \alpha$</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>$\alpha^2 + \alpha + 1$</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>$\alpha^2 + 1$</td>
<td>0</td>
</tr>
</tbody>
</table>

Remark: Note that the sequence represented by $Tr(x^3)$ is 1110100. In general, a more efficient way to generate the sequence of a polynomial $\sum_i Tr(x^i)$ is to first compute the m-sequence represented by $Tr(x)$ using the recurrence relation of the corresponding LFSR. Each sequence of $Tr(x^i)$ is easily computed by decimating the sequence of $Tr(x)$. Then, these decimated sequences are XORed to give the sequence of $\sum_i Tr(x^i)$.

Table 2.3: Truth table of $g(x_2, x_1, x_0)$

<table>
<thead>
<tr>
<th>$x$</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g(x)$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Remark: The truth table of a Boolean function $g$ is the output values listed according to lexicographical ordering of the input. For example, a 4-bit Boolean function $g$ has truth table: 
g(0000), g(0001), g(0010), g(0011), \ldots, g(1111).
CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Table 2.4: Hadamard transform of \( f(x) = \text{Tr}_3^1(x^3) \)

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>0</th>
<th>1</th>
<th>( \alpha )</th>
<th>( \alpha^2 )</th>
<th>( \alpha^3 )</th>
<th>( \alpha^4 )</th>
<th>( \alpha^5 )</th>
<th>( \alpha^6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{f}(\lambda) )</td>
<td>0</td>
<td>-4</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Equations (2.4) and (2.6) are equivalent definitions of the Hadamard transform over different algebraic structures. This is because the set of linear polynomials \( \{ \text{Tr}(\lambda x) | \lambda \in GF(2^n) \} \) is equivalent to the set of linear Boolean functions \( \{ w \cdot x | w \in GF(2^n) \} \) by the following argument:

Let

\[
l(x_{n-1}, \ldots, x_1, x_0) = w \cdot x = w_{n-1}x_{n-1} + \cdots + w_1x_1 + w_0x_0.
\]

The field element corresponding to the input vector \((x_{n-1}, \ldots, x_1, x_0)\), for the basis \(\{ \alpha_{n-1}, \ldots, \alpha_1, \alpha_0 \}\) of \(GF(2^n)\), is given by:

\[
x = x_{n-1}\alpha_{n-1} + \cdots + x_1\alpha_1 + x_0\alpha_0 \in GF(2^n).
\]

Under this correspondence, we see that \(l(x_{n-1}, \ldots, x_1, x_0)\) is the Boolean representation of the polynomial:

\[
\text{Tr}(\lambda x) = \text{Tr}(\lambda(x_{n-1}\alpha_{n-1} + \cdots + x_1\alpha_1 + x_0\alpha_0))
\]

\[
= x_{n-1}\text{Tr}(\lambda\alpha_{n-1}) + \cdots + x_1\text{Tr}(\lambda\alpha_1) + x_0\text{Tr}(\lambda\alpha_0),
\]

where

\[
w_i = \text{Tr}(\lambda\alpha_i) \text{ for } 0 \leq i \leq n - 1.
\]

**Example 4.** From example 3, the polynomial \( f(x) = \text{Tr}_3^1(x^3) \) on \( GF(2^3) = \mathbb{Z}_2[x]/(x^3 + x + 1) \) has Boolean representation \( g(x_2, x_1, x_0) = x_2x_1 + x_2 + x_1 + x_0 \). We compute the Hadamard transform of these two representations to demonstrate their equivalence.

The Hadamard transform of \( f(x) \) is computed in Table 2.4 where we make use of Table 2.2 for \( GF(2^3) \) computation. For example, the polynomials \( \text{Tr}(\alpha^3x) \) and \( \text{Tr}(x^3) \) represent the sequences:

\[
1011100 \text{ and } 1110100
\]

respectively. They agree in five places and disagree in two places. Thus their cross-correlation is 3 which means \( \hat{f}(\alpha^3) = 1 + 3 = 4 \).
Table 2.5: Hadamard transform of $g(x) = x_2x_1 + x_2 + x_1 + x_0$

<table>
<thead>
<tr>
<th>$w$</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{g}(w)$</td>
<td>0</td>
<td>−4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

The Hadamard transform of $g(x)$ is computed in Table 2.5. For example, the truth tables of $x_2 + x_0$ and $g(x_2, x_1, x_0)$ are

01011010 and 01101010 respectively. They agree in six places and disagree in two places. Therefore $\hat{g}(101) = 4$.

The fact that the Hadamard transform is 4 in both our computations is no coincidence. This is because the Boolean representation of $Tr(\alpha^3 x)$ is $x_2 + x_0$, by equation (2.7) using the basis $\{\alpha^2, \alpha, 1\}$ of $GF(2^3)$:

Coefficient of $x_2$: $w_2 = Tr(\alpha^3 \alpha^2) = Tr(\alpha^5) = 1.

Coefficient of $x_1$: $w_1 = Tr(\alpha^3 \alpha) = Tr(\alpha^4) = 0.

Coefficient of $x_0$: $w_0 = Tr(\alpha^3) = 1.$

And we can see that the Hadamard transforms of $f(x)$ and $g(x)$ take exactly the same values: $−4$ once, $4$ three times and $0$ four times.

From the inter-connections described above, we can construct cryptographic Boolean functions from well-known sequences with good correlation properties, through polynomials over $GF(2^n)$.

### 2.3 Cryptographic Requirements of Boolean Functions

A Boolean function $f : GF(2)^n \rightarrow GF(2)$ is balanced if $f(x)$ takes on the values 0’s and 1’s an equal number of time. This is equivalent to saying that $\hat{f}(0) = 0$.

We say the function $f : GF(2)^n \rightarrow GF(2)$ has 3-valued spectrum if its Hadamard Transform $\hat{f}(w)$ only takes on the values $0, \pm 2^i$ for some $i$. 
CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Definition 1. Let \( n \) be odd and \( f : \mathbb{GF}(2)^n \to \mathbb{GF}(2) \). If \( \hat{f}(w) \) only takes on the values 0, \( \pm 2^{(n+1)/2} \), then we say \( f \) is a preferred function.

Remark 1. It is desirable for a Boolean function to have low Hadamard transform for cryptographic applications. By Parseval’s equation:

\[
\sum_{w \in \mathbb{GF}(2)^n} \hat{f}(w)^2 = 2^{2n},
\]

we deduce that \( \max_{w \in \mathbb{GF}(2)^n} |\hat{f}(w)| \geq 2^{n/2} \).

When equality is achieved, we say that \( f \) is a bent function and \( \hat{f}(w) = \pm 2^{n/2} \) for all \( w \in \mathbb{GF}(2)^n \). When \( f \) has 3-valued spectrum 0, \( \pm 2^i \), \( f \) cannot be bent and therefore, \( i \geq (n + 1)/2 \).

Thus a preferred function has the lowest Hadamard transform among functions with 3-valued spectrum. That is the reason such functions are called preferred in [33].

The nonlinearity of a function \( f : \mathbb{GF}(2)^n \to \mathbb{GF}(2) \) is defined as

\[
N_f = \min_{\{a \text{ affine function}\}} \text{dist}(f(x), a(x)).
\]

where \( \text{dist}(f(x), g(x)) = |\{x|f(x) \neq g(x)\}| \) is the distance between \( f(x) \) and \( g(x) \), and an affine Boolean function \( a(x) \) is of the form \( w \cdot x + c \).

From the definition of nonlinearity, we can deduce that the approximation of \( f(x) \) by any linear function \( w \cdot x \) satisfies the following probability bound:

\[
\frac{N_f}{2^n} \leq \Pr(f(x) = w \cdot x) \leq 1 - \frac{N_f}{2^n} \text{ for all } w \in \mathbb{GF}(2)^n.
\]

The nonlinearity of \( f(x) \) is related to its Hadamard transform by

\[
N_f = 2^{n-1} - \frac{1}{2} \max_{w \in \mathbb{GF}(2)^n} |\hat{f}(w)|.
\]

Thus low Hadamard transform ensures high nonlinearity.

Example 5. Let us look at the Boolean function \( g(x_2, x_1, x_0) = x_2 x_1 + x_2 + x_1 + x_0 \) from Example 4. In each column of Table 2.6, we present the distance \( d \) of \( g(x) \) from the linear function \( w \cdot x \), and the distance \( 8 - d \) of \( g(x) \) from the affine function \( w \cdot x + 1 \).

The nonlinearity of \( g(x) \) is:

\[
N_g = \min_{\text{affine}} \text{dist}(g(x), a(x)) = 2.
\]
Table 2.6: Distance of \( g(x) = x_2x_1 + x_1 + x_0 \) from affine function \( w \cdot x + c \), \( c = 0, 1 \)

<table>
<thead>
<tr>
<th>( w )</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>4,4</td>
<td>6,2</td>
<td>4,4</td>
<td>2,6</td>
<td>4,4</td>
<td>2,6</td>
<td>4,4</td>
<td>2,6</td>
</tr>
</tbody>
</table>

Alternatively, we see from Example 4 that \( \hat{g}(w) = 0, \pm 4 \) for all \( w \in GF(2)^3 \). Therefore by equation (2.9),

\[
N_g = 2^{3-1} - \frac{1}{2} \max_w |\hat{g}(w)| = 4 - 1/2 \times 4 = 2.
\]

By equation (2.8), linear approximations of \( g(x) \) satisfy the following probability bound:

\[
\frac{1}{4} \leq \text{Pr}(g(x) = w \cdot x) \leq \frac{3}{4} \text{ for all } w \in GF(2)^3.
\]

A high nonlinearity is desirable as it ensures linear approximation of \( f \) is ineffective. This offers protection against linear cryptanalysis [53] and correlation attack [71].

Let \( n \) be the number of input variables of a Boolean function. The highest achievable nonlinearity is \( 2^{n-1} - 2^{n/2-1} \) and Boolean functions achieving such nonlinearity are called bent functions [64] (see also Remark 1). Bent functions only exist when \( n \) is even and their constructions can be found in [59, 64, 75]. In many applications, we require the Boolean function to be balanced. However, a bent function is not balanced and thus the nonlinearity of a balanced function is less than \( 2^{n-1} - 2^{n/2-1} \).

When \( n \) is even, a balanced function with nonlinearity exceeding \( 2^{n-1} - 2^{n/2} \) is considered high. Balanced Boolean functions with such nonlinearity were constructed by Seberry, Zhang and Zheng in [68] and Dobbertin in [19] for even \( n \geq 6 \).

When \( n \) is odd, it has been conjectured that the highest achievable nonlinearity for a Boolean function is \( 2^{n-1} - 2^{(n-1)/2} \). This conjecture was proven for \( n = 3, 5, 7 \) and it was believed to be true until the early 1980’s (see [62]). But in 1983, Patterson and Wiedemann disproved the conjecture by constructing two functions with 15 input bits and nonlinearity 16276 > \( 2^{14} - 2^7 \) [62, 63]. However, the Patterson-Wiedemann functions were not balanced which may not be suitable for cryptographic applications. Later, balanced functions with nonlinearity exceeding \( 2^{n-1} - 2^{(n-1)/2} \) were constructed by Seberry, Zhang, Zheng for odd \( n \geq 27 \) in [68] and by Sarkar, Maitra for odd \( n \geq 15 \) in [66].
Table 2.7: Recommended nonlinearity for balanced Boolean functions

<table>
<thead>
<tr>
<th>n-bit Balanced Function</th>
<th>Recommended Nonlinearity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 3, 5, 7$</td>
<td>$2^{n-1} - 2^{(n-1)/2}$</td>
<td>[62]</td>
</tr>
<tr>
<td>$n = 9, 11, 13^*$</td>
<td>$\geq 2^{n-1} - 2^{(n-1)/2}$</td>
<td>[62, 68]</td>
</tr>
<tr>
<td>$n \geq 15$ odd</td>
<td>$&gt; 2^{n-1} - 2^{(n-1)/2}$</td>
<td>[66, 68]</td>
</tr>
<tr>
<td>$n = 4$</td>
<td>$2^4 - 2^{4/2} = 4$</td>
<td>[68]</td>
</tr>
<tr>
<td>$n \geq 6$ even</td>
<td>$&gt; 2^{n-1} - 2^{n/2}$</td>
<td>[19, 68]</td>
</tr>
</tbody>
</table>

* Remark: It has not been proven that the maximal nonlinearity for a balanced function is $2^{n-1} - 2^{(n-1)/2}$ for $n = 9, 11, 13$. It is easy to construct balanced functions achieving such nonlinearity [68] but no construction for a balanced function with higher nonlinearity is known for these cases.

Here, we note that the numbers $2^{n-1} - 2^{(n-1)/2}$ and $2^{n-1} - 2^{n/2}$ are called the quadratic bound when $n$ is odd or even respectively. This is because they are the highest achievable nonlinearity for a quadratic Boolean function.

Based on the above discussion, we present Table 2.7 to summarize the recommended nonlinearity of balanced Boolean functions.

A Boolean function $f : GF(2)^n \rightarrow GF(2)$ is $k$-th order correlation immune, denoted $CI(k)$, if

$$\hat{f}(w) = 0 \text{ for all } 1 \leq wt(w) \leq k,$$

where $wt(w)$ is the number of ones in the binary representation of $w$. Correlation immunity ensures that $f$ cannot be approximated by linear functions with too few terms, which offers protection against correlation attacks [71]. Furthermore, if $f$ is balanced and $CI(k)$, we say $f$ is resilient of order $k$.

**Example 6.** We note that the 3-bit Boolean function $g(x_2, x_1, x_0) = x_2x_1 + x_2 + x_1 + x_0$ from Example 4 is balanced but not 1-resilient because $\hat{g}(001) = -4 \neq 0$.

An example of a resilient function is given by:

$$g(x_3, x_2, x_1, x_0) = x_0x_1 + x_2 + x_3.$$
Table 2.8: Hadamard transform of \( g(x_3, x_2, x_1, x_0) = x_0x_1 + x_2 + x_3 \)

<table>
<thead>
<tr>
<th>( x )</th>
<th>( g(x) )</th>
<th>( w )</th>
<th>( \hat{g}(w) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>0</td>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>0001</td>
<td>0</td>
<td>0001</td>
<td>0</td>
</tr>
<tr>
<td>0010</td>
<td>0</td>
<td>0010</td>
<td>0</td>
</tr>
<tr>
<td>0011</td>
<td>1</td>
<td>0011</td>
<td>0</td>
</tr>
<tr>
<td>0100</td>
<td>1</td>
<td>0100</td>
<td>0</td>
</tr>
<tr>
<td>0101</td>
<td>1</td>
<td>0101</td>
<td>0</td>
</tr>
<tr>
<td>0110</td>
<td>1</td>
<td>0110</td>
<td>0</td>
</tr>
<tr>
<td>0111</td>
<td>0</td>
<td>0111</td>
<td>0</td>
</tr>
<tr>
<td>1000</td>
<td>1</td>
<td>1000</td>
<td>0</td>
</tr>
<tr>
<td>1001</td>
<td>1</td>
<td>1001</td>
<td>0</td>
</tr>
<tr>
<td>1010</td>
<td>1</td>
<td>1010</td>
<td>0</td>
</tr>
<tr>
<td>1011</td>
<td>0</td>
<td>1011</td>
<td>0</td>
</tr>
<tr>
<td>1100</td>
<td>0</td>
<td>1100</td>
<td>8</td>
</tr>
<tr>
<td>1101</td>
<td>0</td>
<td>1101</td>
<td>8</td>
</tr>
<tr>
<td>1110</td>
<td>0</td>
<td>1110</td>
<td>8</td>
</tr>
<tr>
<td>1111</td>
<td>1</td>
<td>1111</td>
<td>−8</td>
</tr>
</tbody>
</table>

From Table 2.8, we see that \( g(x) \) is a 4-bit 1-resilient Boolean function because

\[
\hat{g}(w) = 0 \quad \text{for } w = 0000, 0001, 0010, 0100, 1000.
\]

Since \( \max_w |\hat{g}(w)| = 8 \), we deduce from equation (2.9) that the nonlinearity is

\[
N_g = 2^{4-1} - \frac{1}{2} \times 8 = 4.
\]

By equation (2.8), linear approximations of \( g(x) \) satisfy the following probability bound:

\[
\frac{1}{4} \leq \Pr(g(x) = w \cdot x) \leq \frac{3}{4} \quad \text{for all } w \in GF(2)^4.
\]

The correlation attack is a powerful attack against stream cipher systems. To protect against such an attack, we require a Boolean function to possess both high nonlinearity and resiliency.
Chapter 2. Relationships between Sequences, Polynomials and Boolean Functions

Until the late 1990’s, the resilient functions constructed in the literature had maximal nonlinearity $2^{n-1} - 2^{(n-1)/2}$ when $n$ is odd and $2^{n-1} - 2^{n/2}$ when $n$ is even [6, 69]. This led Pasalic and Johansson to conjecture that these are the maximal nonlinearity for resilient functions in 1999 [61]. But at Eurocrypt 2000, Sarkar and Maitra disproved their conjecture by constructing 1-resilient functions with higher nonlinearity for odd $n \geq 41$ and even $n \geq 10$. Their construction requires more input bits to achieve high nonlinearity if the order of resiliency is higher.

On the other hand, when $n$ is odd and small, a resilient function with nonlinearity $2^{n-1} - 2^{(n-1)/2}$ is still considered high [6]. When $n$ is even and small, it was proven in [61] that the quadratic bound is the maximal nonlinearity for resilient functions when $n = 4$ and 6. For $n = 8$, a 1-resilient function with nonlinearity $116 > 2^7 - 2^4 = 112$ was constructed by Maitra and Pasalic in 2002 [51].

At Crypto 2000, Sarkar and Maitra proved that $\hat{f}(w) \equiv 0 \pmod{2^{k+2}}$ for $k$-resilient functions [65]. This implies that $\hat{f}(w) \geq 2^{k+2}$ and thus the maximal achievable nonlinearity is $2^{n-1} - 2^{k+1}$ by equation (2.9). Resilient functions achieving this optimal nonlinearity are called saturated [65]. In the same paper, Sarkar and Maitra gave a construction for almost all saturated functions. Because $N_f \leq 2^{n-1} - 2^{n/2-1}$ for any Boolean function $f(x)$, their bound for optimal nonlinearity is meaningful only for

$$2^{n-1} - 2^{k+1} < 2^{n-1} - 2^{n/2-1} \implies k > n/2 - 2.$$

Based on the above discussion, we present Table 2.9 to summarize the recommended nonlinearity of resilient Boolean functions.

We will construct balanced and resilient Boolean functions, whose nonlinearity achieve the values of Tables 2.7 and 2.9, in Chapters 4 and 5 [28, 40].

Let $f : GF(2)^n \to GF(2)$. The additive autocorrelation at $a \in GF(2)^n$ is defined as

$$\Delta_f(a) = \sum_{x \in GF(2)^n} (-1)^{f(x) + f(x+a)}.$$

This value is also called the propagation of $f$ at $a$ in the literature.

Many useful properties of a Boolean function can be deduced by analysing the relationship between $\Delta_f(a)$ and $\hat{f}(\lambda)$. Some examples can be found in [4, 45, 57, 66, 68, 70, 82, 83].
CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Table 2.9: Recommended nonlinearity for resilient Boolean functions

<table>
<thead>
<tr>
<th>n-bit Resilient Function</th>
<th>Recommended Nonlinearity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 5, 7$</td>
<td>$2^{n-1} - 2^{(n-1)/2}$</td>
<td>[62, 69]</td>
</tr>
<tr>
<td>$n$ small, odd and $\geq 9$</td>
<td>$\geq 2^{n-1} - 2^{(n-1)/2}$</td>
<td>[6, 69]</td>
</tr>
<tr>
<td>1-resilient, $n \geq 41$ and odd</td>
<td>$\geq 2^{n-1} - 2^{(n-1)/2}$</td>
<td>[66]</td>
</tr>
<tr>
<td>$n = 4, 6$</td>
<td>$2^{n-1} - 2^{n/2}$</td>
<td>[61]</td>
</tr>
<tr>
<td>1-resilient, $n \geq 8$ and even</td>
<td>$&gt; 2^{n-1} - 2^{n/2}$</td>
<td>[51, 66]</td>
</tr>
<tr>
<td>$k$-resilient and $k &gt; n/2 - 2$</td>
<td>$2^{n-1} - 2^{k+1}$</td>
<td>[65]</td>
</tr>
</tbody>
</table>

Table 2.10: Additive autocorrelation values of $g(x) = x_2x_1 + x_2 + x_1 + x_0$

<table>
<thead>
<tr>
<th>$a$</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta g(a)$</td>
<td>8</td>
<td>-8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

We say $f$ satisfies the propagation criteria of order $k$, denoted $PC(k)$, if

$$\Delta f(a) = 0$$ for all $1 \leq \text{wt}(a) \leq k,$

i.e. changing any $k$ or fewer bits will change the output bit with probability $1/2$. Therefore, it is not possible to gather any information from $f$ by tweaking certain input bits.

If $\Delta f(a) = \pm 2^n$, then $a$ is called a linear structure of $f$. Non-zero linear structures are undesirable because they leak the maximum amount of information when certain input bits are toggled.

**Example 7.** Consider the Boolean function $g(x_2, x_1, x_0) = x_2x_1 + x_2 + x_1 + x_0$ from Example 4.

The additive autocorrelation values of $g(x)$ are listed in Table 2.10.

For example, the truth tables of $g(x)$ and $g(x + 010)$ are:

01101010 and 10011010

respectively. They agree and disagree in four places. Therefore $\Delta g(010) = 0$. 

CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Table 2.11: Output of \( f(x) \) and \( f(x + \alpha) \) for \( f(x) = \text{Tr}(x^3) \)

<table>
<thead>
<tr>
<th>( x )</th>
<th>0</th>
<th>1</th>
<th>( \alpha )</th>
<th>( \alpha^2 )</th>
<th>( \alpha^3 )</th>
<th>( \alpha^4 )</th>
<th>( \alpha^5 )</th>
<th>( \alpha^6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(x) )</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>( x + \alpha )</td>
<td>( \alpha )</td>
<td>( \alpha^3 )</td>
<td>0</td>
<td>( \alpha^4 )</td>
<td>1</td>
<td>( \alpha^2 )</td>
<td>( \alpha^6 )</td>
<td>( \alpha^5 )</td>
</tr>
<tr>
<td>( f(x + \alpha) )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

We see from Table 2.10 that \( g(x) \) does not satisfy the \( PC(k) \) condition for \( k \geq 1 \) because \( \Delta_g(001) = -8 \neq 0 \). But under a change of basis (i.e., non-singular linear transformation of input vectors), \( g(x) \) can be made to satisfy \( PC(2) \).

We also see that \( \Delta_g(001) = -2^3 \). Therefore, \( g(x) \) has a non-zero linear structure at \( 001 \) which is undesirable.

**Example 8.** Note that the definition of additive autocorrelation is the same for polynomials \( f : GF(2^n) \rightarrow GF(2) \). Let us look at the polynomial \( \text{Tr}_1(x^3) \) from example 4. We will compute \( \Delta_{f}(\alpha) \) in the finite field \( GF(2^3) \) defined by \( \alpha^3 = \alpha + 1 \).

The output table for \( \text{Tr}(x^3) \) and \( \text{Tr}((x + \alpha)^3) \) is given in Table 2.11. Since \( f(x) \) and \( f(x + \alpha) \) agree and disagree in four places, \( \Delta_{f}(\alpha) = 0 \).

**Definition 2.** The additive autocorrelation of \( f \) is \( \Delta_{f} := \max_{a \neq 0} |\Delta_{f}(a)| \).

This quantity was first introduced and studied in detail by Zhang and Zheng in [78]. They called it the maximum indicator of \( f \). For cryptographic applications, we require a Boolean function to have low additive autocorrelation so that toggling any number of input bits will complement the output bit with probability close to \( 1/2 \). Thus, very little information is leaked from attacks based on tweaking input bits.

One motivation for the concept of low additive autocorrelation is that the \( PC(k) \) condition is too strict. It forces the Boolean function to have rigid structure which in many cases, causes linear structures to occur [78]. Low additive autocorrelation is less restrictive and will avoid this problem.

Low additive autocorrelation is also related to a low maximal differential in S-boxes (see Section 2.4), which provides protection against differential cryptanalysis [2].
CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Remark 2. For a balanced function $f$, we want $\Delta_f$ to be low so that any change in the input bits will complement the output with probability close to $1/2$. It was conjectured by Zhang and Zheng in [78, Conjecture 1] that $\Delta_f \geq 2^{(n+1)/2}$ for a balanced function $f : GF(2)^n \rightarrow GF(2)$. Although this conjecture was later disproved by Clark et. al. when $n$ is even (see [13, Table 3]), it still holds when $n$ is odd. Thus, we make the following definition.

Definition 3. Let $n$ be odd and $f : GF(2)^n \rightarrow GF(2)$ be balanced. $f$ has optimal additive autocorrelation if $\Delta_f = 2^{(n+1)/2}$.

The algebraic expression of a Boolean function $f : GF(2^n) \rightarrow GF(2)$ or its algebraic normal form (ANF) is given by:

$$f(x_{n-1}, \ldots, x_1, x_0) = \sum_{a_{n-1}, \ldots, a_1, a_0 \in GF(2)} g(a_{n-1}, \ldots, a_1, a_0) x_{n-1}^{a_{n-1}} \cdots x_1^{a_1} x_0^{a_0}.$$  

The algebraic degree of $f(x)$, denoted by $\deg(f)$, is the degree of the highest term in the ANF. The coefficients $g(a_{n-1}, \ldots, a_1, a_0)$ of $f(x)$ in the ANF can be computed by the following formula:

$$g(a_{n-1}, \ldots, a_1, a_0) = \sum_{(b_{n-1}, \ldots, b_0) \subset (a_{n-1}, \ldots, a_0)} f(b_{n-1}, \ldots, b_1, b_0) \quad (2.10)$$

where by $(b_{n-1}, \ldots, b_0) \subset (a_{n-1}, \ldots, a_0)$, we mean that if $b_j = 1$ then $a_j = 1$. It enables us to easily compute the ANF of a Boolean function from its truth table. This is the formula we used to compute the ANF of the Boolean representation of $Tr_{1}^3(x^3)$ in Example 4.

The Hamming weight of an integer $a$, denoted by $wt(a)$, is the number of ones in the binary representation of $a$. When we have a polynomial function $f : GF(2^n) \rightarrow GF(2)$ defined by:

$$f(x) = \sum_i \beta_i x^{s_i}, \beta_i \in GF(2^n), \beta_i \neq 0,$$

it can be shown that the algebraic degree $\deg(f)$ of the corresponding Boolean function is given by the maximum weight of the exponents $\max_i, wt(s_i)$. We want it to be high so that algebraic analysis is complex.

The linear span is given by the number of monomials in the polynomial $f(x)$. We want it to be large to protect against an interpolation attack [36].

Example 9. Consider the polynomial:

$$Tr_{1}^3(\alpha^3 x) = \alpha^3 x + \alpha^6 x^2 + \alpha^5 x^4.$$
from Example 4. The linear span is 3 because it is a sum of three monomials. It corresponds to a linear Boolean function because all the exponents of $x$: 1, 2 and 4 have Hamming weight 1. This is confirmed by the fact that its Boolean representation is $x_2 + x_0$ from Example 4.

Next, consider the polynomial:

$$Tr_1^3(x^3) = x^3 + x^6 + x^5.$$ 

from Example 4. The linear span is 3 because it is a sum of three monomials. It corresponds to a quadratic Boolean function because all the exponents of $x$: 3, 6 and 5 have Hamming weight 2. This is confirmed by the fact that its Boolean representation is $x_2x_1 + x_2 + x_1 + x_0$ from Example 4.

### 2.4 Cryptographic Requirements of S-boxes

An S-box is a function from $GF(2)^n$ to $GF(2)^m$, $m \geq 2$. In cryptographic applications, S-boxes are also called vectorial Boolean functions. They are used in block ciphers to provide confusion, e.g., in a Substitution Permutation Network-based block cipher like the Advanced Encryption Standard (AES) or a Feistel cipher like the Data Encryption standard (DES). A good description and security analysis of these ciphers can be found in [73]. In a stream cipher, they are used as a combiner or filter function for higher throughput [39, 77].

There is a natural correspondence between Boolean S-boxes $F : GF(2)^n \to GF(2)^m$ and polynomial functions from $GF(2^n)$ to $GF(2^m)$. Let $\{\alpha_{n-1}, \ldots, \alpha_0\}$ be a basis for $GF(2^n)$ and $\{\beta_{m-1}, \ldots, \beta_0\}$ be a basis for $GF(2^m)$. The correspondence is given by

- **Input vector:** $(x_{n-1}, \ldots, x_0) \leftrightarrow \alpha_{n-1}x_{n-1} + \ldots + \alpha_0x_0 \in GF(2^n)$
- **Output vector:** $(y_{m-1}, \ldots, y_0) \leftrightarrow \beta_{m-1}y_{m-1} + \ldots + \beta_0y_0 \in GF(2^m)$

For convenience, we call a function $F : GF(2)^n \to GF(2)^m$ an $n$-by-$m$ S-box.

An $n$-by-$m$ S-box, $n \geq m$, is *balanced* if every output occurs $2^{n-m}$ times.

Let $F(x) = (f_{m-1}(x), \ldots, f_0(x))$ be an $n$-by-$m$ S-box where each $f_i(x)$ represents the Boolean function of the $i$-th output bit. Then the *algebraic degree* of $F(x)$ is defined by:

$$deg(F) = \min_i deg(f_i).$$
Let $F : GF(2^n) \rightarrow GF(2^m)$, $F(x) = \sum \beta_i x^{s_i}$, be the polynomial representation of an $n$-by-$m$ S-box. Then the algebraic degree is the maximum weight of the exponents, i.e., $\text{deg}(F) = \max_i s_i$.

The linear span is the number of monomials in the expression of $F(x)$.

The nonlinearity $N_F$ of an S-box is given by:

$$N_F = \min_{b \neq 0} N_{b \cdot F}$$

where $b \cdot F(x)$ is a Boolean function formed from a linear combination of output bits. Like the case of Boolean functions, high nonlinearity for S-boxes provides protection against linear cryptanalysis in block ciphers [53] and correlation attacks in stream ciphers [71].

**Example 10.** Let us consider the 3-by-3 bijective S-box:

$$F(x_2, x_1, x_0) = (f_2(x_2, x_1, x_0), f_1(x_2, x_1, x_0), f_0(x_2, x_1, x_0))$$

corresponding to the permutation $x^3$ on $GF(2^3)$. The values of $x^3$ and the truth table of the corresponding S-box are listed in Table 2.12. The basis of $GF(2^3)$ for converting both the input and output to Boolean form is $\{\alpha^2, \alpha, 1\}$, i.e.

3-bit vector: $(x_2, x_1, x_0) \leftrightarrow \alpha^2 x_2 + \alpha x_1 + x_0 \in GF(2^3)$.

Because the weight of the exponent of $x^3$ is 2, $F(x)$ is a quadratic S-box. This is verified by

<table>
<thead>
<tr>
<th>$i$</th>
<th>$x = \alpha^i$</th>
<th>$x^3$</th>
<th>$(x_2, x_1, x_0)$</th>
<th>$(f_2, f_1, f_0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>0</td>
<td>0</td>
<td>000</td>
<td>000</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>001</td>
<td>001</td>
</tr>
<tr>
<td>1</td>
<td>$\alpha$</td>
<td>$\alpha + 1$</td>
<td>010</td>
<td>011</td>
</tr>
<tr>
<td>2</td>
<td>$\alpha^2$</td>
<td>$\alpha^2 + 1$</td>
<td>011</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>$\alpha + 1$</td>
<td>$\alpha^2$</td>
<td>100</td>
<td>101</td>
</tr>
<tr>
<td>4</td>
<td>$\alpha^2 + \alpha$</td>
<td>$\alpha^2 + \alpha + 1$</td>
<td>101</td>
<td>110</td>
</tr>
<tr>
<td>5</td>
<td>$\alpha^2 + \alpha + 1$</td>
<td>$\alpha$</td>
<td>110</td>
<td>111</td>
</tr>
<tr>
<td>6</td>
<td>$\alpha^2 + 1$</td>
<td>$\alpha^2 + \alpha$</td>
<td>111</td>
<td>010</td>
</tr>
</tbody>
</table>
equation (2.10), where we deduce that the algebraic normal form of $F(x)$ is:

$$F(x_2, x_1, x_0) = (x_0 x_1 + x_2, x_0 x_2 + x_0 x_1 + x_1, x_1 x_2 + x_0 + x_1 + x_2).$$

The linear span of $x^3$ is 1 which is not desirable. In fact, block ciphers which use monomials for S-boxes has been susceptible to interpolation and algebraic attacks. For example, there is a summary of such attacks on the S-box $x^{-1}$ of the AES in a recent survey article by Courtois [14].

We can verify that for all $(v_2, v_1, v_0) \in GF(2)^3 - (0,0,0)$, the nonlinearity of:

$$v \cdot F(x_2, x_1, x_0) = v_2 (x_0 x_1 + x_2) + v_1 (x_0 x_2 + x_0 x_1 + x_1) + v_0 (x_1 x_2 + x_0 + x_1 + x_2)$$

is 2. Therefore the nonlinearity of $F(x)$ is $N_F = 2$. By equation (2.8), linear approximations of $F(x)$ satisfy the following probability bound:

$$\frac{1}{4} \leq Pr(v \cdot F(x) = w \cdot x) \leq \frac{3}{4} \text{ for all } v \in GF(2)^3 - \{0\}, w \in GF(2)^3.$$

We can deduce from the nonlinearity bound for Boolean functions that the highest nonlinearity an S-box can achieve is $2^n - 1 - 2^{n/2} - 1$. Such S-boxes are called bent S-boxes, they only exists when the number of input bits $n$ is even and the number of output bits satisfies $m \leq n/2$ [59]. Some constructions can be found in [59, 67, 75].

There is also a general bound for the nonlinearity of $n$-by-$m$ S-boxes due to Chabaud and Vaudenay [9, Theorem 4]:

$$N_F \leq 2^{n-1} - \frac{1}{2} \sqrt{3 \times 2^n - 2 - \frac{2(2^n - 1)(2^{n-1} - 1)}{2^m - 1}}. \quad (2.11)$$

One drawback is that inequality (2.11) only gives meaningful bounds when $n \leq m$, i.e. the number of output bits is at least the number of input bits. When equality is achieved, $F$ is called an almost bent S-box. In that case, the Hadamard transform $\hat{b} \cdot \hat{F}(w)$ where $w \in GF(2)^n, b \in GF(2)^m$, only takes on three values 0, ±Λ for some Λ ∈ $\mathbb{R}$.

We can deduce from inequality (2.11) that when $m = n$ is odd, the maximal achievable nonlinearity is the quadratic bound: $2^{n-1} - 2^{(n-1)/2}$. We can easily construct almost bent permutations achieving such nonlinearity from the m-sequences [5, 60].

Although we noted in the previous section that there exist many constructions for Boolean functions ($m = 1$) having nonlinearity $> 2^{n-1} - 2^{(n-1)/2}$ when $n$ is odd [62, 63, 66, 68], to the best
CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Table 2.13: Recommended nonlinearity for cryptographic S-boxes

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Recommended Nonlinearity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$ even, $m \leq n/2$</td>
<td>$2^{n-1} - 2^{n/2-1}$</td>
<td>[59, 67, 75]</td>
</tr>
<tr>
<td>$n \leq m$</td>
<td>$2^{n-1} - 1/2 \sqrt{3 \times 2^n - 2 - 2(2^{n-1})(2^{n-1}-1)/2^{m-1}}$</td>
<td>[9]</td>
</tr>
<tr>
<td>$n$ odd and $m = n$</td>
<td>$2^{n-1} - 2^{(n-1)/2}$</td>
<td>[5, 9]</td>
</tr>
<tr>
<td>$n$ odd and $2 \leq m \leq n-1$</td>
<td>$\geq 2^{n-1} - 2^{(n-1)/2}$</td>
<td>-</td>
</tr>
</tbody>
</table>

of our knowledge, there is no known construction for S-boxes ($m \geq 2$) whose nonlinearity exceeds the quadratic bound when $n$ is odd. Therefore we consider a nonlinearity $\geq 2^{n-1} - 2^{(n-1)/2}$ to be high for an S-box.

Based on the above discussion, we present Table 2.13 to summarize the recommended nonlinearity of cryptographic S-boxes functions. In Table 2.13, $n$ is the number of input bits while $m$ is the number of output bits.

Let $F : GF(2)^n \to GF(2)^m$. The maximum differential $\Delta_F$ is defined by:

$$\Delta_F = \max_{a \in GF(2)^n \setminus \{0\}, b \in GF(2)^m} \{|x| F(x) + F(x + a) = b\}.$$  

This implies that:

$$0 \leq Pr_{x \in GF(2)^n} (F(x) + F(x + a) = b) \leq \Delta_F / 2^n \text{ for all } a \neq 0.$$  

We want $\Delta_F$ to be as small as possible for protection against differential cryptanalysis [2]. Note that the maximal differential is closely related to the additive autocorrelation of Boolean functions. When an S-box $F(x)$ has low maximal differential, all Boolean functions formed from linear combination of output bits have low additive autocorrelation.

The smallest possible value for $\Delta_F$ is $2^{n-m}$ and it is achieved only when $n$ is even [59]. S-boxes whose $\Delta_F$ attains this lower bound are called perfect nonlinear. It can be proven that the bent S-boxes are equivalent to the perfect nonlinear S-boxes [59].

When $n$ is odd, we want $\Delta_F$ to be as close to $2^{n-m}$ as possible. When $m = n$ is odd, the smallest possible value for $\Delta_F$ is 2 and S-boxes which achieves this bound are called almost perfect nonlinear (APN) [60]. It is proven in [9, Theorem 4] that almost bent S-boxes are APN.
Table 2.14: Difference distribution table of S-box corresponding to $x^3$

<table>
<thead>
<tr>
<th>$\Delta F(a) \setminus a$</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>001</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>010</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>011</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>101</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>110</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>111</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Therefore, both bent and almost bent S-boxes provide protection against linear and differential cryptanalysis of block ciphers. This makes them particularly attractive to be used in applications.

**Example 11.** Consider the S-box corresponding to the permutation $x^3$ on $GF(2^3)$ in Example 10. Its algebraic normal form is:

$$F(x_2, x_1, x_0) = (x_0x_1 + x_2, x_0x_2 + x_0x_1 + x_1, x_1x_2 + x_0 + x_1 + x_2),$$

and its truth table can be found in Table 2.12. The difference distribution table is listed in Table 2.14 where $\Delta F(a)$ represents $F(x) + F(x + a)$ over all $x \in GF(2)^3$.

For example, for $a = 111$, the computation of $\Delta F(111)$ is:

1. Input pair: (000, 111), (111, 000). Output difference: 010 twice.
2. Input pair: (001, 110), (110, 001). Output difference: 110 twice.
4. Input pair: (011, 100), (100, 011). Output difference: 001 twice.

We see that $\Delta F = 2$, i.e. $F(x)$ is APN, which implies:

$$0 \leq Pr_{x \in GF(2)^3}(F(x) + F(x + a) = b) \leq 1/4$$ for all $a \neq 000$. 
CHAPTER 2. RELATIONSHIPS BETWEEN SEQUENCES, POLYNOMIALS AND BOOLEAN FUNCTIONS

Table 2.15: \( k \) such that \( x^k \) is an almost bent permutation, \( n \) odd

<table>
<thead>
<tr>
<th>Exponent ( k )</th>
<th>Condition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^r + 1 ) (Gold)</td>
<td>( \gcd(r, n) = 1 )</td>
<td>[22]</td>
</tr>
<tr>
<td>( 2^{2r} - 2^r + 1 ) (Kasami)</td>
<td>( \gcd(r, n) = 1 )</td>
<td>[38]</td>
</tr>
<tr>
<td>( 2^{(n-1)/2} + 3 ) (Welch)</td>
<td>none</td>
<td>[5]</td>
</tr>
<tr>
<td>( 2^{2r} + 2^r - 1 ) (Niho)</td>
<td>( 4r + 1 \equiv 0 \pmod{n} )</td>
<td>[58]</td>
</tr>
</tbody>
</table>

The Niho exponent is conjectured. These permutations correspond to the m-sequences \( Tr(x^k) \) which has been used in CDMA applications since the 1970’s.

Alternatively, we see from Example 10 that \( F(x) \) is almost bent because \( N_F = 2 \) which implies that \( F(x) \) is APN from [9, Theorem 4].

In Table 2.15, we list the known values of \( k \) for which \( x^k \) is an almost bent permutation [5, 22, 38, 58]. They will be used to construct cryptographic S-boxes in Chapters 5 and 6 [39, 40].

At Crypto 2000, Zhang and Chan introduced the maximum correlation [77] of an \( n \)-by-\( m \) S-box at the point \( w \in GF(2)^n \), defined by:

\[
C_F(w) = \max_g [\text{Prob}(g(F(x)) = w \cdot x) - \text{Prob}(g(F(x)) \neq w \cdot x)].
\] (2.12)

Here, the maximum is taken over all \( g : GF(2)^m \rightarrow GF(2) \). Low maximal correlation provides protection against linear approximation attacks on any nonlinear function of output bits in stream cipher systems. This is a natural generalization of high nonlinearity which provides protection against linear approximation attacks on linear combinations of output bits.

**Remark 3.** We do not consider \( w = 0 \) in equation (2.12) because \( C_F(0) = 1 \) for all \( F \) (by letting \( g(z) = 0 \)); instead, we require that \( F \) be balanced.

**Proposition 1.** (Zhang-Chan [77, Theorem 4]) Let \( F : GF(2)^n \rightarrow GF(2)^m \). Then the maximum correlation of \( F \) at \( w \neq 0 \) satisfies

\[
C_F(w) \leq 2^{m/2 - n \max_v} \sum_x (-1)^{w \cdot F(x) + w \cdot x},
\] (2.13)
Proposition 1 implies that a high nonlinearity guarantees low maximum correlation. In Chapter 6, we will construct highly nonlinear balanced S-boxes whose maximum correlation improves the bound in equation (2.13) [39].

Definition 4. The maximum correlation of $F$ is defined as $C_F = \max_{w \neq 0} C_F(w)$.

If an S-box has low maximum correlation, then it is secure against generalized linear approximation attacks in stream ciphers. From the definition of $C_F(w)$ and $C_F$, we deduce that:

$$\frac{1}{2} - \frac{C_F}{2} \leq \Pr(g(F(x)) = w \cdot x) \leq \frac{1}{2} + \frac{C_F}{2},$$

for all $w \in GF(2)^n$ and $g : GF(2)^m \rightarrow GF(2)$.

By noticing that $C_F(w) = 2^{-n} \max_g \sum_x (-1)^{g(F(x)) + w \cdot x}$, we get the following corollary of Proposition 1.

Corollary 1. Let $F : GF(2)^n \rightarrow GF(2)^m$. Then the maximum correlation of $F$ satisfies

$$2^{-n/2} \leq 2^{-n} \max_{w \neq 0, v \neq 0} |\hat{v} \cdot F(w)| \leq C_F \leq 2^{m/2-n} \max_{w \neq 0, v \neq 0} |\hat{v} \cdot F(w)|.$$

Some examples of S-boxes with high nonlinearity and low maximum correlation can be found in Chapters 5 and 6.
Chapter 3

Optimal Quadratic Polynomials

Boolean functions with high nonlinearity have important applications in cryptography. This is because linear approximation of these functions are difficult, thus making certain cryptographic attacks like linear cryptanalysis infeasible. It is well-known that when the number of input bits $n$ is odd, the highest achievable nonlinearity for a quadratic Boolean function $f : GF(2)^n \rightarrow GF(2)$ defined by,

$$f(x_0, \ldots, x_{n-1}) = \sum_{i<j} c_{ij} x_i x_j + \sum_i c_i x_i + c,$$

where $c_{ij}, c_i, c \in \{0, 1\}$, is $2^{n-1} - 2^{(n-1)/2}$. Thus, this number is commonly called the quadratic bound. The quadratic Boolean functions which achieve this nonlinearity have been well studied and a complete classification can be found in [54, Chapter 15].

We note that recently there has been much usage of finite fields in cryptography. For example, the S-box for the Advanced Encryption Standard uses the inverse function over $GF(2^n)$ to provide confusion [73, Chapter 3]. Implementation of elliptic curve cryptosystems over $GF(2^n)$ provides efficiency and security [55]. The Digital Signature Standard signs messages using exponentiation in finite fields [73, Chapter 7].

In view of the importance of finite fields in cryptography, and the natural correspondence between a polynomial in $GF(2^n)$ and a Boolean function (see equation (2.5)), it will be interesting to find the quadratic polynomials in $GF(2^n)$ whose nonlinearity achieve the quadratic bound. We call such polynomials the optimal quadratic polynomials.
The quadratic polynomials studied in this chapter are motivated by the Gold sequences and the Gold-like sequences of Boztas-Kumar, which are represented by the polynomials $T r(x^{2^i+1})$ and $\sum_{i=1}^{(n-1)/2} T r(x^{2^i+1})$ respectively [3, 22]. They are XORed with shifts of the m-sequence $T r(x)$ to form sequences with low cross-correlation for CDMA applications. From this, we can deduce that both the Gold function and the Boztas-Kumar function are optimal quadratic polynomials on $GF(2^n)$, $n$ odd. Therefore, it might be interesting to look at whether the related function $\sum_{i=1}^{(n-1)/2} c_i T r(x^{2^i+1})$, $c_i \in \{0,1\}$ is optimal. By combining techniques from linear algebra, coding theory and number theory, we can determine if this quadratic polynomial is optimal by an efficient GCD computation.

Using the tools we develop, we find two new classes of primes $p$ for which the polynomial functions $\sum_{i=1}^{\frac{p-1}{2}} c_i T r(x^{2^i+1})$, $c_i \in \{0,1\}$, $x \in GF(2^p)$ are optimal for all choices of $c_i$. One class consists of odd primes $p$ for which 2 is a primitive element modulo $p$; while the other includes the Sophie-Germain primes of the form $p = 2q + 1$, $p, q$ prime. Assuming the Riemann Hypothesis, the first class was proven to be infinite while it is a well-known conjecture that the second class of primes is infinite, too. Furthermore, we prove that these primes are the only odd integers $n$ with this property.

Finally, we find large classes of optimal quadratic functions which can be found with minimal or no computations. These include the sum of Gold functions based on arithmetic progression and the sum of all but one Gold function $T r(x^{2^i+1})$. Our result can be applied to all finite fields $GF(2^n)$ where $n$ is odd.

This chapter is organized as follows. In Section 3.1, we characterize the optimal quadratic polynomials by an efficient GCD computation. In Section 3.2, we find finite fields $GF(2^n)$ for which all linear combination of Gold functions yield optimal quadratic polynomials. In Section 3.3, we identify the optimal quadratic polynomials which can be found with minimal or no computations.

### 3.1 Characterization of Optimal Quadratic Polynomials

**Definition 5.** Let $n$ be odd. The function $f : GF(2^n) \rightarrow GF(2)$ defined by

$$f(x) = \sum_{i=1}^{(n-1)/2} c_i T r(x^{2^i+1}), c_i \in \{0,1\},$$
is called an optimal quadratic polynomial\(^1\) if its Hadamard transform takes the three values 0, ±2\(^{n+1}/2\).

By equation (2.5), optimal quadratic polynomials correspond to quadratic Boolean functions with maximal nonlinearity 2\(^n−1−2^{(n−1)/2}\).

In this section, we prove that for \(n\) odd, we can determine very efficiently if the polynomial \(f(x) = \sum_{i=1}^{n−1} c_i Tr(x^{2^{i+1}})\) is optimal using elementary algebraic techniques.

**Lemma 1.** For \(n\) odd, let \(f(x) = \sum_{i=1}^{n−1} c_i Tr(x^{2^{i+1}})\), \(c_i \in \{0, 1\}\), \(x \in GF(2^n)\). Then \(f(x)\) is optimal if and only if the cyclic matrix

\[
L = \begin{pmatrix}
c_0 & c_1 & c_2 & c_3 & \ldots & c_{n-1} \\
c_{n-1} & c_0 & c_1 & c_2 & \ldots & c_{n-2} \\
c_{n-2} & c_{n-1} & c_0 & c_1 & \ldots & c_{n-3} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
c_1 & c_2 & c_3 & c_4 & \ldots & c_0
\end{pmatrix}
\]  

(3.1)

has rank \(n−1\), where \(c_0 = 0\) and \(c_{n-i} = c_i\) for \(i = 1, \ldots, \frac{n-1}{2}\).

**Proof.** We use the Welch squaring method:

\[
\hat{f}(\lambda)^2 = \sum_{x,y} (-1)^{Tr(\lambda x)+f(x)}(-1)^{Tr(\lambda y)+f(y)} = \sum_{w,x} (-1)^{Tr(\lambda x)+f(x)+Tr(\lambda(x+w))+f(x+w)} \\
\text{(where } y = x + w \text{)} \\
= \sum_{w} (-1)^{Tr(\lambda w)+f(w)} \sum_{x} (-1)^{\Phi(x,w)}
\]

\(^1\)Such a polynomial is quadratic because when we convert it to a Boolean function using a basis of \(GF(2^n)\), it is a quadratic Boolean function. It is optimal because \(2^{(n+1)/2}\) is the lower bound for \(\max_{\lambda} |\hat{f}(\lambda)|\) when \(f\) is quadratic.
where $\Phi(x, w) = f(w) + f(x) + f(x+w)$. We simplify $\Phi$ as follows:

$$\Phi(x, w) = \sum_{i=1}^{n-1} c_i \left[ \text{Tr}(x^{2^i+1}) + \text{Tr}(w^{2^i+1}) + \text{Tr}((x+w)^{2^i+1}) \right]$$

$$= \sum_{i=1}^{n-1} c_i \text{Tr}(x^{2^i}w + w^{2^i}x)$$

$$= \sum_{i=1}^{n-1} c_i \text{Tr}(x^{2n-i} + w^{2^i})$$

$$= \text{Tr}(xL(w)),$$

where $L(w) = \sum_{i=1}^{n-1} c_i (w^{2^i} + w^{2n-i}).$

Note that $L$ is a linear function, and under a normal basis $\{\alpha, \alpha^2, \alpha^4, \ldots, \alpha^{2^n-1}\}$ of $GF(2^n)$, the matrix representation of $L$ is given by matrix (3.1). Also,

$$\sum_x (-1)^{\Phi(x, w)} = \sum_x (-1)^{\text{Tr}(xL(w))} = 2^n$$

if and only if $L(w) = 0$, otherwise the sum is 0. Therefore

$$\hat{f}(\lambda)^2 = 2^n \sum_{w \in \ker(L)} (-1)^{\text{Tr}(\lambda w) + f(w)}.$$ (3.3)

Let $\dim(\ker(L)) = k$. By the definition of $\Phi$, $\text{Tr}(\lambda w) + f(w)$ is a linear function on $\ker(L)$. Therefore

$$\sum_{w \in \ker(L)} (-1)^{\text{Tr}(\lambda w) + f(w)} = 2^k \text{ or } 0,$$ (3.4)

depending on whether the exponent is the zero function ($2^{n-k}$ cases) or a non-zero linear function ($2^n - 2^{n-k}$ cases). That means that $\hat{f}$ is 0, $\pm 2^{n-k}$ with frequency $2^n - 2^{n-k}, 2^{n-k}$ if and only if $\dim(\ker(L)) = k$. As a special case, $f$ is optimal if and only if $\dim(\ker(L)) = 1$, i.e., $\text{rank}(L) = n-1$.

Remark 4. The bilinear form $\Phi(x, w)$, defined in the above proof, corresponds to the symplectic of $f(x)$ under the boolean representation. For results on quadratic Boolean functions, see [54, Chapter 15].

Note that the rows of matrix (3.1) span a cyclic code $C$ generated by the vector $(c_0, c_1, \ldots, c_{n-1})$. In the study of cyclic codes, it is useful to represent the vectors of $C$ by polynomials modulo $x^n+1$, \]
where

\[ C = \text{span}\{c(x), xc(x), \ldots, x^{n-1}c(x)\} \] (3.5)

and \( c(x) = \sum_{i=1}^{n-1} c_i(x^i + x^{n-i}) \). Then we have the following well-known useful facts about the cyclic code \( C \) (e.g., see [34]):

1. There exists a unique monic polynomial \( g(x) \), called the generator polynomial, such that \( g(x)|v(x) \) for all \( v(x) \in C \).
2. \( \text{rank}(L) = \text{dim}(C) = n - \deg(g(x)) \).
3. \( g(x) = \gcd(c(x), x^n + 1) \).

Thus, if we can show that \( g(x) = \gcd(c(x), x^n + 1) = x + 1 \), then we will have proven that \( \text{rank}(L) = \text{dim}(C) = n - 1 \), the condition required in Lemma 1. We summarize our discussion in the following theorem:

**Theorem 1.** For \( n \) odd, let \( f(x) = \sum_{i=1}^{n-1} c_i \text{Tr}(x^{2^i+1}) \), \( c_i \in \{0, 1\}, \ x \in GF(2^n) \). Then \( f \) is optimal if and only if \( g(x) = x + 1 \) where \( g(x) = \gcd(c(x), x^n + 1) \) and \( c(x) = \sum_{i=1}^{n-1} c_i(x^i + x^{n-i}) \).

To demonstrate the application of Theorem 1, we can deduce that:

1. \( f(x) = \text{Tr}(x^3 + x^5) \) is optimal over \( GF(2^{155}) \) because \( \gcd(x+x^2+x^{153}+x^{154}, x^{155}+1) = x+1 \).
2. \( f(x) = \text{Tr}(x^3 + x^{17}) \) is not optimal over \( GF(2^{155}) \) because \( \gcd(x+x^4+x^{151}+x^{154}, x^{155}+1) = x^5 + 1 \).

The above computations can be done efficiently using the mathematical package MAPLE, while it would take a long time if we were to verify directly that \( f(x) \) is optimal over a large field like \( GF(2^{155}) \).

Using Theorem 1, we give a short proof of two known results regarding optimal polynomials. Let \( n \) be odd.

1. The function \( f(x) = \text{Tr}(x^{2^i+1}) \) is optimal if and only if \( \gcd(i, n) = 1 \) [22]. This is because the generator polynomial \( g(x) = \gcd(x^i+x^{n-i}, x^n+1) = \gcd(x^{n-2i}+1, x^n+1) = x^{\gcd(n,n-2i)}+1 \).
   So \( g(x) = x + 1 \) if and only if \( \gcd(n, i) = 1 \).
2. The function \( f(x) = \sum_{i=1}^{n-1} Tr(x^{2^i+1}) \) is optimal [3]. This is because the associated cyclic code \( C \) is generated by the polynomial \( c(x) = \sum_{i=1}^{n-1} x^i \) and \( 1+x = (xc(x) \text{ mod } x^n+1)+c(x) \) is a vector in \( C \). So \( \gcd(c(x), x^n+1) = x+1 \).

From Theorem 1, we see that it is useful to analyze the generator polynomial \( g(x) \) of the associated cyclic code. Let us state a useful lemma regarding \( g(x) \).

**Lemma 2.** Let \( n \) be odd and \( g(x) \) be the generator polynomial of the cyclic code \( C \) generated by \( c(x) = \sum_{i=1}^{n-1} c_i(x^i + x^{n-i}) \). Then \( g(x) \) has the form
\[
g(x) = (x+1)h(x), \quad \text{where } \deg(h(x)) \text{ is even.} \tag{3.6}
\]

**Proof.** If \( \beta \) is a root of \( g(x) = \gcd(c(x), x^n+1) \), then it is a root of both \( x^n+1 \) and \( c(x) \). \( \beta^n = 1 \) implies \( (\beta^{-1})^i = \beta^{n-i} \) and thus \( c(\beta^{-1}) = c(\beta) = 0 \). Also, \( (\beta^{-1})^n = (\beta^n)^{-1} = 1 \). Therefore, \( \beta^{-1} \) is a root of \( c(x) \) and \( x^n+1 \), and thus of \( g(x) \). So the roots of \( g(x) \) come in pairs \( \beta, \beta^{-1} \). If \( \beta \neq 1 \), then the assumption that \( n \) is odd and \( \beta^n = 1 \) imply \( \text{ord}(\beta) \neq 2 \) which means \( \beta \neq \beta^{-1} \). Thus \( \beta = 1 \) gives the factor \( x+1 \), while the other pairs of roots give an even degree polynomial, \( h(x) \). \( \square \)

### 3.2 Optimal Quadratic Polynomials for All Choices of Coefficients

For what \( n \) is \( f(x) = \sum_{i=1}^{n-1} c_i Tr(x^{2^i+1}) \), \( c_i \in \{0,1\} \), \( x \in GF(2^n) \) optimal for all choices of coefficients? By Gold’s result [22], we know that \( Tr(x^{2^i+1}) \) is not optimal when \( \gcd(i,n) \neq 1 \). So we have the following lemma.

**Lemma 3.** Let \( n \) be odd. If \( f(x) = \sum_{i=1}^{n-1} c_i Tr(x^{2^i+1}) \), \( c_i \in \{0,1\} \), \( x \in GF(2^n) \) is optimal for all choices of coefficients, then \( n \) is prime.

Since the generator polynomial \( g(x) \) is closely related to the factorization of the polynomial \( x^n+1 \), let us examine this factorization more clearly in Lemma 4. We need the following related definition:

**Definition 6.** Let \( p \) be prime. We denote the multiplicative order of a modulo \( p \) by \( \text{ord}_p(a) \).
Lemma 4. Let $p$ be an odd prime. Over $\mathbb{Z}_2[x]$, the factorization of $x^p + 1$ into irreducible factors is of the form

$$x^p + 1 = (x + 1)h_1(x)h_2(x)\ldots h_t(x)$$

(3.7)

where $t = \frac{p-1}{\text{ord}_p(2)}$ and each $h_i(x)$ is an irreducible polynomial of degree $\text{ord}_p(2)$.

Proof. Let $\zeta_p$ be a primitive $p$th-root of unity and let $H = \{2^s \mod p, s \in \mathbb{Z}\}$ be the subgroup of $\mathbb{Z}_p^*$ generated by 2. The irreducible factors of $x^p + 1$ are given by $\prod_{i \in kH}(x - \zeta_i^p)$ where $k \in \mathbb{Z}_p$.

Except for $k = 0$, $kH$ is a coset of $H$ in $\mathbb{Z}_p^*$. Thus, except for $k = 0$, which corresponds to the linear factor $x + 1$, the degree of each nonlinear irreducible factor is $|kH| = |H| = \text{ord}_p(2)$, and the number of nonlinear factors is $t = \frac{p-1}{\text{ord}_p(2)}$. \qed

If we want $f(x)$ to be optimal for all choices of coefficients, we might choose a prime $p$ for which $x^p + 1$ has as few irreducible factors as possible, for then it would be “harder” for $g(x) = \gcd(v(x), x^p + 1)$ to contain any factor other than $x + 1$. We look at the cases where $x^p + 1$ has two and three factors in the following sections.

3.2.1 $p$ is odd prime and $\text{ord}_p(2) = p - 1$

This is the case where $x^p + 1$ has two irreducible factors.

Theorem 2. Let $f(x) = \sum_{i=1}^{p-1} c_i Tr(x^{2^i} + 1)$, $c_i \in \{0, 1\}$, $x \in GF(2^p)$ where $p$ is a prime such that $\text{ord}_p(2) = p - 1$. Then $f$ is optimal.

Proof. By Lemma 4, $x^p + 1$ has factorization

$$x^p + 1 = (x + 1)(1 + x + x^2 + \ldots + x^{p-1}).$$

(3.8)

Because the generator polynomial $g(x)$ divides $x^p + 1$ properly, it must be $x + 1$ or $1 + x + \ldots + x^{p-1}$. By Lemma 2, $(x + 1)|g(x)$, which implies $g(x) = x + 1$, and hence we are done by Theorem 1. \qed

In view of the above theorem, it is natural to ask about the existence and distribution of primes $p$ for which $\text{ord}_p(2) = p - 1$. The first ten such primes are as follows:

$$3, 5, 11, 13, 19, 29, 37, 53, 59, 61.$$  

(3.9)

A conjecture of Artin states that there exists an infinite number of such primes; this conjecture was proven, assuming that the Riemann Hypothesis holds [72].
3.2.2 \( p = 2s + 1 \) is prime, \( s \) is odd and \( \text{ord}_p(2) = s \)

This is the case where \( x^p + 1 \) has three irreducible factors of odd degree.

**Theorem 3.** Let \( f(x) = \sum_{i=1}^{\frac{p-1}{2}} c_i \text{Tr}(x^{2^i}+1), \ c_i \in \{0,1\}, \ x \in GF(2^p) \) where \( p = 2s + 1 \) is a prime such that \( s \) is odd and \( \text{ord}_p(2) = s \). Then \( f \) is optimal.

**Proof.** By Lemma 4, the factorization of \( x^p + 1 \) is

\[
x^p + 1 = (x + 1)h_1(x)h_2(x), \tag{3.10}
\]

where \( h_i(x) \) has degree \( s \), \( i = 1, 2 \). The generator polynomial is a proper divisor of \( x^p + 1 \) and has \( x + 1 \) as a factor by Lemma 2. So \( g(x) = x + 1 \) or \( (x + 1)h_i(x) \), \( i = 1, 2 \). But Lemma 2 says that \( g(x) \) is a product of \( x + 1 \) and an even degree polynomial. Therefore \( g(x) = x + 1 \) and we are done by Theorem 1. \( \square \)

The first ten primes of the form of Theorem 3 are as follows:

\[
7, 23, 47, 71, 103, 167, 191, 199, 239, \tag{3.11}
\]

and computer simulations suggest that there are an infinite number of such primes.

As a corollary of Theorems 2 and 3, we prove a similar statement for the Sophie Germain primes, \( p = 2q + 1, \ p, q \) prime.

**Corollary 2.** Let \( f(x) = \sum_{i=1}^{\frac{p-1}{2}} c_i \text{Tr}(x^{2^i}+1), \ c_i \in \{0,1\}, \ x \in GF(2^p) \) where \( p = 2q + 1, \ p, q \) prime. Then \( f \) is optimal.

**Proof.** If \( \text{ord}_p(2) = p - 1 \), then we are done by Theorem 2. If not, then \( \text{ord}_p(2) \) is a proper divisor of \( p - 1 = 2q \), which implies \( \text{ord}_p(2) = 2 \) or \( q \). But \( \text{ord}_p(2) \neq 2 \) because \( p \geq 5 \), so \( \text{ord}_p(2) = q \), which is an odd prime, and we are done by Theorem 3. \( \square \)

The first ten Sophie-Germain primes are as follows:

\[
5, 7, 11, 23, 47, 59, 83, 107, 167, 179. \tag{3.12}
\]

As shown in the proof of Corollary 2, it can be verified that the Sophie Germain primes are primes of the form mentioned in Theorem 2 and 3. The Sophie Germain primes are well studied in number theory and it is conjectured that there are an infinite number of such primes.
CHAPTER 3. OPTIMAL QUADRATIC POLYNOMIALS

What about primes which are not of the form considered in Theorems 2 and 3? We show that non-optimal polynomials exists in these cases.

**Theorem 4.** The only odd integers \( n \) for which \( f(x) = \sum_{i=1}^{n-1} c_i \text{Tr}(x^{2^i+1}) \), \( c_i \in \{0, 1\} \), \( x \in GF(2^n) \) is optimal for all choices of coefficients are the primes mentioned in Theorems 2 and 3.

**Proof.** Let \( p \) be a prime not of the form considered in Theorems 2 and 3. We will show there exists a polynomial \( c(x) \) such that

1. \( \text{deg}(c(x)) < p \).
2. \( c(x) \) is a \( p \)-symmetric polynomial, i.e. \( c(x) = x^p c(1/x) \). Then \( c(x) = \sum_{i=1}^{p-1} c_i x^i \) where \( c_i = c_{p-i} \).
3. \( \text{gcd}(c(x), x^p + 1) \neq x + 1 \).

Then the polynomial function \( f(x) = \sum_{i=1}^{(p-1)/2} c_i \text{Tr}(x^{2^i+1}) \) formed from the coefficients \( c_i \) of \( c(x) \) is not optimal.

Let \( \zeta_p \) be a primitive \( p \)th-root of unity. Form the polynomial \( u(x) \) whose roots are \( \zeta_p^i, i \in \{\pm 2^s \mod p, s \in \mathbb{Z}\} \). Then \( r := \text{deg}(u) \) is even because \( \zeta_p^{-1} \neq \zeta_p \) is a root of \( u(x) \) whenever \( \zeta_p \) is a root.

\( u(x) \) is \( r \)-symmetric, i.e., \( u(x) = x^r u(1/x) \). This can be shown by the following computation:

Let \( a_1, a_2, \ldots, a_r \) be the roots of \( u \). Then \( 1/a_1, \ldots, 1/a_r \) are also the roots of \( u \). Thus

\[
(x + a_1) \cdots (x + a_r) = \left( x + \frac{1}{a_1} \right) \cdots \left( x + \frac{1}{a_r} \right) = x^r \prod_{\gamma \in C_1} \left( a_1 x + 1 \right) \cdots \left( a_r x + 1 \right) = x^r \prod_{\gamma \in C_1} \left( \frac{a_1 + 1}{x} \cdots \left( \frac{a_r + 1}{x} \right) \right)
\]

because \( a_1 a_2 \cdots a_r = 1 \). The top line of the above computation is \( u(x) \) while the last line is \( x^r u(1/x) \). We see that \( (x + 1)u(x) \) is \( r + 1 \)-symmetric by a similar reason.

We claim that \( (x + 1)u(x) \) has degree less than \( p \). Let

\[
C_1 = \{ \zeta_p^{2^k}, k = 0, 1, \ldots \} \text{ and } m_1(x) = \prod_{\gamma \in C_1} (x - \gamma),
\]
where $m_1(x)$ is the minimal polynomial of $\zeta_p$. Let

$$C_{-1} = \{\zeta_p^{-2^k}, k = 0, 1, \ldots\}$$

and $m_{-1}(x) = \prod_{\gamma \in C_{-1}} (x - \gamma)$,

where $m_{-1}(x)$ is the minimal polynomial of $\zeta_p^{-1}$. We split the proof of our claim into two cases:

Case $C_1 = C_{-1}$. Then $u(x) = m_1(x) = m_{-1}(x)$, $\deg(u) = |C_1| = \text{ord}_p(2) < p - 1$ because $p$ is not from Theorem 2.

Case $C_1 \neq C_{-1}$. Then $C_1, C_{-1}$ are disjoint and $s := \text{ord}_p(2)$ cannot be even. This is because otherwise $2^s = 1$ means $2^{s/2} = -1$ and this implies $C_1$ contains $\zeta_p^{-1}$. Thus $s$ must be odd and $s < (p - 1)/2$ because $p$ is not from Theorem 3. We have $u(x) = m_1(x)m_{-1}(x)$ and $\deg(m_1(x)) = \deg(m_{-1}(x)) = s$ implies $\deg(u(x)) = 2s < p - 1$.

Then $c(x) = x^{(p-(r+1)/2)}(x+1)u(x)$ is a polynomial of degree

$$\frac{p-(r+1)}{2} + 1 + r = \frac{p+r+1}{2} < p \quad (3.13)$$

because $1 + r = \deg((x+1)u(x)) < p$. We also have

$$x^p c(1/x) = x^{(p+r+1)/2}(1+1/x)u(1/x) = x^{(p+r+1)/2}x^{-(r+1)}(1+x)u(x) = c(x).$$

Thus $\deg(c(x)) < p$ and $c(x)$ is a $p$-symmetric polynomial. The corresponding function $f(x)$ over $GF(2^p)$ will not be optimal by Theorem 1, because $\gcd(c(x), x^p + 1) = (x+1)u(x) \neq x+1$. □

**Example 12.** Consider the prime $p = 17$ which is not of the form considered in Theorem 2 and 3. Let $\beta$ be a primitive element of the finite field $GF(2^8)$. Then $\zeta_{17} = \beta^{15}$ is a primitive 17th-root of unity because $255 = 15 \times 17$. It can be verified that $C_1 = C_{-1} = \{\zeta_{17}, \zeta_{17}^2, \ldots, \zeta_{17}^2\}$,

$$u(x) = (x + \zeta_{17})(x + \zeta_{17}^2)\ldots(x + \zeta_{17}^2) = x^8 + x^7 + x^6 + x^5 + x^2 + x + 1,$$

and $r = \deg(u) = 8$. Then, following the procedure in the proof of Theorem 4, we get

$$c(x) = x^{17-(8+1)/2}(1+x)u(x) = x^{13} + x^{10} + x^9 + x^8 + x^7 + x^4$$

where $\gcd(c(x), x^{17} + 1) = (x+1)u(x)$. Thus $f(x) = Tr(x^{17} + x^{129} + x^{257})$, $x \in GF(2^{17})$, is not optimal.
3.3 Characterizing Optimal Quadratic Polynomials

In this section, we identify certain classes of optimal quadratic polynomials which can be found with no or minimal computations.

**Theorem 5.** Let $n$ be odd. Then

1. $f(x) = Tr(x^{2^i+1} + x^{2^j+1}), \ x \in GF(2^n)$ is optimal if and only if $\gcd(i + j, n) = 1 = \gcd(i - j, n)$.

2. Consider the function

$$f(x) = Tr(x^{2^a + d + 1} + \ldots + x^{2^a + (r-1)d + 1} + x^{2^a + rd + 1})$$

over $GF(2^n)$, $n$ odd, where

$$a, a + d, \ldots, a + (r - 1)d, a + rd$$

forms an arithmetic progression with common difference $d$. Then $f$ is optimal if $\gcd(2^a + rd, n) = 1 = \gcd((r + 1)d, n)$. Moreover, the converse is true if $\gcd(d, n) = 1$.

**Proof.** We analyze the two cases by applying Theorem 1.

1. Without loss of generality, assume $i > j$. The corresponding polynomial $c(x)$ is

$$c(x) = x^i + x^j + x^{n-j} + x^{n-i}$$

$$= (x^i + x^j)(1 + x^{n-i+j})$$

$$= x^j(1 + x^{i-j})(1 + x^{n-i+j}).$$

Thus $\gcd(c(x), x^n + 1) = x + 1$ if and only if

$$\gcd(i - j, n) = 1 = \gcd(i + j, n).$$

2. The corresponding polynomial $c(x)$ is

$$c(x) = x^a + \ldots + x^{a+rd} + x^{n-a-rd} + \ldots + x^{n-a}$$

$$= (1 + x^{n-(2a+rd)})(x^a + \ldots + x^{a+rd})$$

$$= (1 + x^{n-(2a+rd)})x^a \left( \frac{1 + x^{(r+1)d}}{1 + x^d} \right).$$
The gcd of the numerator with $x^n + 1$ is $x + 1$ if
\[
gcd(2a + rd, n) = 1 = gcd((r + 1)d, n).
\]

Moreover, we see that the converse is true if
\[
gcd(1 + x^d, 1 + x^n) = 1 + x, \text{ i.e. } gcd(d, n) = 1.
\]

Next, we present a result on optimal quadratic polynomials with all but one trace term. The result states that for two-thirds of all integers $n$, the sum of all but one $Tr(x^{2^i+1})$ term is always optimal on $GF(2^n)$.

**Theorem 6.** Consider the quadratic polynomial
\[
f(x) = Tr(x^3 + \ldots + x^{2^{i-1}+1} + x^{2^{i+1}+1} + \ldots + x^{2^{(n-1)/2}+1})
\]
on $GF(2^n), n$ odd, i.e., $f$ is a sum of all but one trace term $Tr(x^{2^i+1})$. Then

1. If 3 does not divide $n$, $f$ is always optimal.
2. If $n = 3^k$ for some $k$, $f$ is never optimal.
3. If $n = 3^k m$, $m > 1$ and 3 does not divide $m$, $f$ is optimal if and only if $i$ is a multiple of $3^k$.

**Proof.** We analyse the corresponding polynomial $c(x)$:
\[
c(x) = x + \ldots + x^{i-1} + x^{i+1} + \ldots + x^{n-(i+1)} + x^{n-(i-1)} + \ldots + x^{n-1}
\]
The relation between $x^n + 1$ and $c(x)$ is
\[
x^n + 1 = (x + 1)(1 + x + \ldots + x^{n-1})
\]
\[
= (x + 1)(c(x) + 1 + x^i + x^{n-i}).
\]
Thus we have
\[
x^n + 1 + (x + 1)c(x) = (x + 1)(1 + x^i + x^{n-i})
\]
which enables us to simplify $g(x) := \gcd(c(x), x^n + 1)$ as
\[
g(x) = \gcd(c(x), (x + 1)(1 + x^i + x^{n-i})) = (x + 1) \gcd(c(x), 1 + x^i + x^{n-i}).
\]
The last equality is true because $c(1) = 0$ and $1 + 1^i + 1^{n-i} \neq 0$. By Theorem 1, we deduce that $f$ is not optimal if and only if
\[
g(x) \neq x + 1 \iff \exists a \neq 1 \text{ such that } g(a) = 0
\]
\[
\iff \exists a \text{ such that } 1 + a^i + a^{n-i} = 0 \text{ and } c(a) = 0
\]
\[
\iff \exists a \text{ such that } 1 + a^i + a^{n-i} = 0 \text{ and } a^n = 1
\]
\[
\iff \exists a \text{ such that } a^i + a^{2i} + 1 = 0 \text{ and } a^n = 1
\]
\[
\iff \exists a \text{ such that } a^{3i} = 1, a^i \neq 1 \text{ and } a^n = 1.
\]
Using this condition, we prove parts 1, 2 and 3.

1. $3$ does not divide $n$. Suppose there exists an $i$ for which $f$ is not optimal. This means that $g(x)$ has a root $a \neq 1$. By the conditions we derived, $a$ satisfies $a^i \neq 1$ and $a^{3i} = 1$, which implies $\text{ord}(a) = 3$ or $3i$. Therefore $3$ divides $\text{ord}(a)$ and $a^n = 1$ implies $\text{ord}(a)$ divides $n$ which means that $3|n$. However, this is a contradiction to the hypothesis that $3$ does not divide $n$.

2. $n = 3^k$. Since $n$ is odd, there exists a primitive $n$th root of 1 in some field of characteristic 2, say $w$ (see [35]). We can write $i$ as $i = 3^rd$, where $r < k$ and $3$ does not divide $d$. Let $a = w^{3^kr-1}$, then
\[
a^{3i} = w^{3^kd} = 1, \quad a^i = w^{3^{k-1}d} \neq 1 \text{ and } a^n = 1.
\]
Thus $f$ is never optimal.

3. $n = 3^k m$, $3$ does not divide $m$. Let $i = 3^kd$, where $3$ does not divide $d$ and suppose $f$ is not optimal, i.e. there exists an $a$ such that
\[
a^{3i} = 1, \quad a^i \neq 1 \text{ and } a^n = 1.
\]
$a^n = 1$ implies $\text{ord}(a)|3^k m$ and $3$ does not divide $m$. So the maximum exponent of $3$ in $\text{ord}(a)$ is $3^k$. Now we have
\[
a^{3i} = 1 \implies a^{3k+1}d = 1 \implies a^{3^kd} = 1 \implies a^i = 1,
\]
CHAPTER 3. OPTIMAL QUADRATIC POLYNOMIALS

42

a contradiction! Thus, $f$ is optimal for $i = 3^k d$.

Let $i = 3^r d$, where $r < k$ and 3 does not divide $d$. Since $n$ is odd, there exists a primitive $n$th root of 1, say $w$ (see [35]). Let $a = w^{a^k - r - 1 m}$; then we can verify that

$$a^{3^i} = 1, \quad a^l \neq 1 \quad \text{and} \quad a^n = 1,$$

as we have done in part (b) to show that $f$ is not optimal.

Example 13. Let $p$ be a prime. By Theorem 4, not all linear combinations of Gold functions $\sum_i c_i Tr(x^{2^i+1})$ give optimal quadratic polynomials if $p$ is not a prime from Theorem 2 or 3, e.g. $p = 17$. But by Gold’s result [22], Theorem 5 and Theorem 6, all quadratic polynomials given by sum of one or two Gold functions:

$$f(x) = Tr(x^{2^i+1}), \quad f(x) = Tr(x^{2^i+1} + x^{2^j+1}),$$

the polynomials based on arithmetic progression:

$$f(x) = Tr(x^{2^i+1} + x^{2^s + d+1} + \ldots + x^{2^s + (r-1)d+1} + x^{2^s + rd+1}),$$

and the sum of all but one Gold function:

$$f(x) = Tr(x^3 + \ldots + x^{2^n-1+1} + x^{2^n+1} + \ldots + x^{2^{(n-1)/2}+1}),$$

are optimal on $GF(2^p)$. Therefore, we have a large number of optimal quadratic polynomials without any computation required for verification.

3.4 Summary of Chapter 3

By generalizing the Gold and Boztas-Kumar polynomials from the theory of CDMA communications [3, 22], we have found a new class of quadratic polynomials on $GF(2^n)$ with optimal nonlinearity for cryptographic applications. These polynomials are easily identified by an efficient polynomial GCD computation. We also identified large classes of optimal quadratic polynomials which require little or no computation.
Chapter 4

Additive Autocorrelation of Resilient Boolean Functions

Resiliency and high nonlinearity are two of the most important properties required of Boolean functions when they are used as combiners in stream cipher systems. Resiliency ensures the cipher is not prone to correlation attack [71] while high nonlinearity offers protection against linear approximation attack [53]. Another criteria, studied in many recent papers, is low additive autocorrelation [74, 78, 79, 81]. This ensures that the output of the function is complemented with a probability close to 1/2 when any fixed number of input bits are complemented. As a result, the cipher is not prone to differential-like cryptanalysis [2]. This is a more practical condition than the propagation criterion of order $k$ which, in the case of high nonlinearity, may cause linear structures to occur (as pointed out in [78]).

In this chapter, we study the autocorrelation and resiliency properties of Boolean functions with 3-valued spectrum. This is an important class of functions with many applications in cryptography, for example see [4, 7, 9, 65, 82]. Based on functions with 3-valued spectrum, we derive several new constructions of resilient Boolean functions with high nonlinearity and optimally low additive autocorrelation. Then we show that the functions obtained from our constructions have better additive autocorrelation than known highly nonlinear resilient functions. Our findings are summarized in the following paragraphs.
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

First, we introduce a new notion called the \textit{dual function}, which is defined as the characteristic function of the Hadamard transform. This notion turns out to be a very useful tool for studying functions with 3-valued spectrum. We prove that the additive autocorrelation of such functions is proportional to the Hadamard transform of their duals. Therefore, we show that propagation criteria of order \( k \) and correlation immunity of order \( k \) are dual concepts. Moreover, low additive autocorrelation and high nonlinearity are also dual concepts. Furthermore, we show that a function with 3-valued spectrum is correlation immune of order 1 if and only if its dual function is not affine.

Second, we look at functions which have the lowest 3-valued spectrum \( 0, \pm 2^{(n+1)/2} \). We prove that, if a balanced preferred function \( f(x) \) has a dual function with this property, then \( f(x) \) has several optimal cryptographic properties. The function will be 1-resilient, which offers protection against correlation attack [71]. The nonlinearity is \( 2^{n-1} - 2^{(n-1)/2} \), which is considered high among resilient functions, according to [6]. Sarkar and Maitra constructed resilient functions with higher nonlinearity but their construction only works for odd \( n \geq 41 \) [66] while our construction works for all odd \( n \geq 5 \). The functions have optimal additive autocorrelation, which achieves a lower bound conjectured by Zhang and Zheng [78]. Therefore, our functions have useful cryptographic properties when used as Boolean functions in cipher systems.

We present some balanced functions which can be used in the construction of certain ideal 2-level autocorrelation sequences, for which the duals are non-affine or preferred. They include the Kasami functions, the Dillon-Dobbertin functions, the Segre hyperoval functions and the Welch-Gong Transformation functions [16, 17, 18, 29] which all give rise to Boolean functions that achieve the desirable cryptographic properties mentioned above. Moreover, some of these functions have high algebraic degree (for algebraic complexity) and large linear span (which offers protection against an interpolation attack [36]).

Third, we compute the additive autocorrelation of some known resilient functions with high nonlinearity. We show that our constructed functions have better additive autocorrelation than resilient preferred functions based on the Maiorana-McFarland construction [6]. The Maiorana-McFarland construction is based on the concatenation of linear functions. This may be a weakness as the function becomes linear when certain input bits are fixed. We show with some examples that our construction can avoid this weakness.

We also investigate an important class of functions with 3-valued spectrum: the saturated
functions constructed in [65] (which are resilient functions optimizing the Siegenthaler and Sarkar-Maitra inequalities). We compute a lower bound for the additive autocorrelation which improves on a bound given in [74]. We show that they have very high additive autocorrelation close to $2^n$. Moreover an $n$-bit saturated function becomes linear when we fix just very few bits ($\log_2(n)$ or fewer bits). Thus, although a saturated function satisfies some very strong cryptographic properties, it may lead to a rigid structure which causes other weaknesses to occur. These potential weaknesses have to be considered before we deploy them in applications.

This chapter is organized as follows. In Section 4.1, we derive a useful relation between a function with 3-valued spectrum and its dual. In Section 4.1.1, it is shown that a function with 3-valued spectrum is correlation immune if and only if its dual is non-affine. In Section 4.1.2, we prove that a balanced preferred function has optimal additive autocorrelation if its dual is preferred. In Section 4.2, we construct four classes of Boolean functions, from certain ideal 2-level autocorrelation sequences, with optimal cryptographic properties. In Section 4.3.1, our constructions were shown to have better additive autocorrelation than a well-known class of nonlinear resilient functions, the Maiorana-McFarland class. In Section 4.3.2, the saturated functions were shown to have certain weaknesses that should be considered before deploying them.

### 4.1 Cryptographic Properties of the Dual Function

We give the definition of the dual function of a Boolean function as follows.

**Definition 7.** Let $f : GF(2)^n \to GF(2)$. Its dual function $\sigma_f : GF(2)^n \to GF(2)$ is defined as

$$\sigma_f(w) = \begin{cases} 
0 & \text{if } \hat{f}(w) = 0 \\
1 & \text{if } \hat{f}(w) \neq 0.
\end{cases}$$

**Remark 5.** From Parseval’s equation, $\sum_w \hat{f}(w)^2 = 2^{2n}$, we see that $\sigma_f(w)$ has weight $2^{2(n-i)}$ if $f$ has 3-valued spectrum $0, \pm 2^i$. When $f$ is a preferred function, its dual function is balanced.

Next, we show that the Hadamard transform of the dual function is proportional to the additive autocorrelation of a function with 3-valued spectrum. It can be applied to derive useful cryptographic properties in the next few subsections.
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

Lemma 5. If \( f : GF(2)^n \rightarrow GF(2) \) is a Boolean function with 3-valued spectrum \( 0, \pm 2^i \), then for all \( a \neq 0 \)

\[
\Delta_f(a) = -2^{2i-(n+1)} \hat{\sigma}_f(a).
\]

Proof. We make use of the well known Wiener-Khintchine theorem (e.g. see [4, Lemma 1]):

\[
\hat{f}(w)^2 = \sum_a \Delta_f(a)(-1)^a\cdot w.
\]

By applying the inverse Hadamard transform\(^1\), we have the equivalent formula:

\[
\Delta_f(a) = \frac{1}{2^n} \sum_w \hat{f}(w)^2(-1)^a\cdot w. \tag{4.1}
\]

By definition, \( \hat{f}(w)^2 = 2^{2i} \sigma_f(w) \). Substituting this into equation (4.1), we get

\[
\Delta_f(a) = 2^{2i-n} \sum_w \sigma_f(w)(-1)^a\cdot w.
\]

By noting that \( 2\sigma_f(w) = 1 - (-1)^\sigma_f(w) \), we get

\[
\Delta_f(a) = 2^{2i-(n+1)} \left( \sum_w (-1)^a\cdot w - \sum_w (-1)^\sigma_f(w)+a\cdot w \right).
\]

This is \(-2^{2i-(n+1)} \hat{\sigma}_f(a)\) when \( a \neq 0 \) and \( 2^n \) when \( a = 0 \).

4.1.1 Correlation Immunity and Non-Affine Dual Function

In this section, we prove that there exists a basis for which the Boolean representation of a polynomial function with 3-valued spectrum is correlation immune if and only if the dual function is non-affine.

We also show that correlation immunity of order \( k \) and propagation criteria of order \( k \) can be viewed as dual concepts for Boolean functions with 3-valued spectrum in the following proposition.

Proposition 2. Let \( f : GF(2)^n \rightarrow GF(2) \) be a Boolean function with 3-valued spectrum \( 0, \pm 2^i \). Then \( f \) is \( PC(k) \) if and only if \( \sigma_f \) is \( CI(k) \).

Proof. \( f \) is \( PC(k) \iff \Delta_f(a) = 0 \) for \( 1 \leq wt(a) \leq k \iff \hat{\sigma}_f(a) = 0 \) for \( 1 \leq wt(a) \leq k \) (by Lemma 5) \iff \( \sigma_f \) is \( CI(k) \).

\(^1\)The inverse Hadamard transform is given by the formula: \( F(w) = \sum_x f(x)(-1)^{w\cdot x} \implies f(x) = \frac{1}{2^n} \sum_w F(w)(-1)^{w\cdot x} \).
We need the following result from [4] for proving correlation immunity of Boolean functions.

**Proposition 3.** (Canteaut, Carlet, Charpin and Fontaine [4, Theorem 7]) Let \( f : GF(2^n) \rightarrow GF(2) \) be a polynomial function with 3-valued spectrum 0, ±2\(^i\). Then there exists a basis of \( GF(2^n) \) such that the Boolean representation of \( f \) is CI(1) if and only if \( f \) is not PC\((n - 1)\) under any basis representation.

**Remark 6.** We stated Proposition 3 in a modified form (from the original in [4]) so that it applies to polynomial functions. From the proof of Proposition 3, we see that the set \( \{ \lambda \in GF(2^n) | \hat{f}(\lambda) = 0 \} \) contains \( n \) linearly independent vectors. Based on these \( n \) vectors, Gong and Youssef [29] gave an algorithm to find a basis of \( GF(2^n) \) such that the Boolean form of \( f \) is 1-resilient. This is presented as Algorithm 1 in this section.

We also note that Proposition 3 was proven in another form for preferred functions (also called plateaued functions of order \( n - 1 \)) by Zheng and Zhang [82, Theorem 2].

The following theorem is a corollary of Proposition 2 and 3. It is well suited for studying polynomial functions with 3-valued spectrum whose duals are known. Some applications can be found in Section 4.2.

**Theorem 7.** Let \( f : GF(2^n) \rightarrow GF(2) \) be a polynomial function with 3-valued spectrum 0, ±2\(^i\). Then there exists a basis of \( GF(2^n) \) such that the Boolean representation of \( f(x) \) is CI(1) if and only if \( \sigma_f \) is not affine.

**Proof.** \( f \) is CI(1) in some basis \( \iff \) \( f \) is not PC\((n - 1)\) in any basis (by Proposition 3) \( \iff \) \( \sigma_f \) is not CI\((n - 1)\) in any basis (by Proposition 2) \( \iff \) \( \sigma_f \) is not affine.

The last equivalence is because of Siegenthaler’s inequality [71]:

\[
k + \deg(f) \leq n
\]

for an \( n \)-bit Boolean function which is CI\((k)\). Thus \( \sigma_f \) is CI\((n - 1)\) if and only if \( \deg(\sigma_f) \leq 1 \). But \( \sigma_f \) cannot be a constant function because \( f \) is 3-valued. Thus \( \sigma_f \) must be affine.

For the sake of completeness, we list in Algorithm 1 a method to find a 1-resilient Boolean representation for a polynomial function \( f : GF(2^n) \rightarrow GF(2) \) when the set \( \{ \lambda \in GF(2^n) | \hat{f}(\lambda) = 0 \} \) contains \( n \) linearly independent vectors.
Algorithm 1. ([29, Algorithm 1])

1. Input: A polynomial function \( f : \mathbb{GF}(2^n) \rightarrow \mathbb{GF}(2) \) and \( n \) linearly independent vectors \( \lambda_1, \ldots, \lambda_n \in \mathbb{GF}(2^n) \) such that \( \hat{f}(\lambda_i) = 0 \) for all \( i \).

   Output: A 1-resilient Boolean representation of \( f(x) \) given by \( g : \mathbb{GF}(2^n) \rightarrow \mathbb{GF}(2) \).

2. Let \( \alpha \) be a generator for \( \mathbb{GF}(2^n) \). Form the binary matrix \( A \) whose \( i \)th row is:

   \[
   (\text{Tr}_1^n(\lambda_i), \text{Tr}_1^n(\alpha \lambda_i), \ldots, \text{Tr}_1^n(\alpha^{n-1} \lambda_i)).
   \]

3. Let \( x \in \mathbb{GF}(2^n) \) be represented by

   \[
   x = x_0 + x_1 \alpha + \cdots + x_{n-1} \alpha^{n-1}.
   \] (4.2)

   Define \( g : \mathbb{GF}(2^n) \rightarrow \mathbb{GF}(2) \) by

   \[
   g((x_0, x_1, \ldots, x_{n-1}) A^t) = f(x)
   \] (4.3)

   where \( A^t \) is the transpose of \( A \). Then \( g \) is a Boolean representation of \( f \) which is 1-resilient.

Some examples illustrating the application of Algorithm 1 can be found in Section 4.2.1.

4.1.2 Computing Additive Autocorrelation from the Dual Function

In this section, we derive formulas to compute the additive autocorrelation of functions with 3-valued spectrum.

**Proposition 4.** Suppose \( f : \mathbb{GF}(2^n) \rightarrow \mathbb{GF}(2) \) has 3-valued spectrum \( 0, \pm2^i \). Then \( \Delta_f \leq 2^{2i-1} - 2^{2i-n} N_{\sigma_f} \).

**Proof.** By Lemma 5, we have \( \hat{\sigma_f}(a) = -2^{n+1-2i} \Delta_f(a) \) for \( a \neq 0 \). By substituting this in the formula \( N_{\sigma_f} = 2^{n-1} - 1/2 \max_a |\hat{\sigma_f}(a)| \), we have

\[
\Delta_f = \max_{a \neq 0} |\Delta_f(a)| = 2^{2i-(n+1)} \max_{a \neq 0} |\hat{\sigma_f}(a)| \leq 2^{2i-(n+1)} \max_{a \neq 0} |\hat{\sigma_f}(a)| = 2^{2i-1} - 2^{2i-n} N_{\sigma_f}.
\]

\( \square \)
Thus, if $N_{\sigma_f}$ is high, then $\Delta_f$ is low.

The concept of propagation criteria of order $k$ was well studied in the literature but it has its shortcomings; see Section 2.3. Most Boolean functions with high nonlinearity that satisfy $PC(k)$ have linear structures or high $\Delta_f$ as pointed out by Zhang and Zheng in [78], which is not desirable. Instead, we require that balanced functions have low additive autocorrelation, i.e. toggling any number of input bits will result in the output being complemented with probability close to $1/2$. The next theorem shows when low additive autocorrelation can be achieved. Some applications can be found in Section 4.2.

**Theorem 8.** If $f : GF(2)^n \rightarrow GF(2)$ is a balanced preferred function and $\sigma_f$ is preferred, then $f$ has optimal additive autocorrelation, i.e. $\Delta_f = 2^{(n+1)/2}$. Moreover, $\Delta_f(a) = 0$ for $2^{n-1} - 1$ a’s.

**Proof.** By Lemma 5, $\hat{\sigma}_f(a) = -\Delta_f(a)$ for $a \neq 0$ when $f$ is preferred. $\sigma_f$ is preferred means $\hat{\sigma}_f(a) = 0, \pm 2^{(n+1)/2}$ which implies $\Delta_f(a) = 0, \pm 2^{(n+1)/2}$ for all $a \neq 0$. Thus, $\Delta_f = 2^{(n+1)/2}$.

Let $v$ be the number of elements $a$ such that $\hat{\sigma}_f(a) = 0$. By Parseval’s equation and the fact that $\sigma_f$ is preferred, we have

$$\sum_a \hat{\sigma}_f(a)^2 = 2^{2n} \implies (2^n - v)2^{n+1} = 2^{2n} \implies v = 2^{n-1}.$$ 

From remark 5, $\hat{\sigma}_f(0) = 0$ because $\sigma_f$ is balanced (note that $\Delta_f(0) = 2^n$). Therefore $\hat{\sigma}_f(a) = 0$ for $2^{n-1} - 1$ non-zero a’s. By Lemma 5, $\Delta_f(a) = 0$ for $2^{n-1} - 1$ elements a’s.

4.1.3 Nonlinearity of Functions with 3-valued Spectrum

From equation 2.9 in Chapter 2, the following proposition follows easily.

**Proposition 5.** A function $f : GF(2)^n \rightarrow GF(2)$ with 3-valued spectrum $0, \pm 2^i$ has nonlinearity $2^{n-1} - 2^{i-1}$. By remark 1, a preferred function has the highest nonlinearity $2^{n-1} - 2^{(n-1)/2}$ among functions with 3-valued spectrum.

**Remark 7.** In Section 4.2 and 4.3.1, we will concentrate on resilient preferred functions. When $n$ is odd and small, their nonlinearity $2^{n-1} - 2^{(n-1)/2}$ is considered high among resilient functions according to Carlet [6]. Sarkar and Maitra constructed resilient functions with nonlinearity $> 2^{n-1} - 2^{(n-1)/2}$ [66, Theorem 6]. But their construction only works when $n \geq 41$ for 1-resilient
4.1.4 Algebraic Degree of Functions with 3-valued Spectrum

The following upper bound on algebraic degree, based on the divisibility of the Hadamard transform, is known, e.g. see [84].

Lemma 6. If \( f : GF(2)^n \rightarrow GF(2) \) satisfies \( \hat{f}(\lambda) \equiv 0 \pmod{2^i} \) for all \( \lambda \), then \( \deg(f) \leq n - i + 1 \).

The following proposition follows easily.

Proposition 6. A function with 3-valued spectrum \( 0, \pm 2^i \) satisfies

\[
\deg(f) \leq n - i + 1.
\]

By remark 1, the preferred functions have maximal algebraic degree \((n + 1)/2\) among functions with 3-valued spectrum.

For example, the maximal algebraic degree \((n + 1)/2\) is achieved by the preferred Kasami function \( f(x) = Tr(x^{2^k - 2^{k+1}}) \) when \( k = (n - 1)/2 \) [38].

4.2 Construction of Resilient Highly Nonlinear Boolean Functions with Optimal Additive Autocorrelation

Our main result in this section is to construct four classes of Boolean functions with desirable cryptographic properties, from functions used in the construction of ideal 2-level autocorrelation sequences. This is achieved by applying the results we have derived in Section 4.1, which are summarized in Theorems 9 and 10.

Theorem 9. If \( f : GF(2^n) \rightarrow GF(2) \) is a balanced preferred function such that its dual \( \sigma_f \) is non-affine, then

1. \( f \) is resilient of order 1 for some basis conversion.
2. \( f \) has high nonlinearity \( 2^{n-1} - 2^{(n-1)/2} \).

Proof. This is a direct consequence of Theorem 7 and Proposition 5.

Theorem 10. If \( f : GF(2^n) \to GF(2) \) is a balanced preferred function such that its dual \( \sigma_f \) is preferred, then

1. \( f \) is resilient of order 1 for some basis conversion.
2. \( f \) has high nonlinearity \( 2^{n-1} - 2^{(n-1)/2} \).
3. \( f \) has optimal additive autocorrelation, i.e. \( \Delta_f = 2^{(n+1)/2} \) and \( \Delta_f(a) = 0 \) for \( 2^{n-1} - 1 \) a’s.

Proof. This is a direct consequence of Theorem 7, Proposition 5 and Theorem 8.

In the following discussion, we present some balanced preferred functions for which their duals are non-affine or preferred. By Theorem 9 and 10, they have desirable properties as Boolean functions for cryptographic applications. Moreover, some of them have high algebraic degree and large linear span for added algebraic complexity.

Our first construction is based on a class of Kasami functions whose Hadamard transform distribution is found by Dillon.

Lemma 7. (Kasami-Dillon [16, 38]) Let \( n \) be odd, \( \gcd(n, 3) = 1 \) and \( f : GF(2^n) \to GF(2) \) be defined by \( f(x) = Tr(x^d) \) where \( d = 2^{2k} - 2^k + 1 \), \( 3k \equiv 1 \pmod{n} \). Then \( f \) is preferred and satisfies

\[
\hat{f}(\lambda) = \begin{cases} 
0 & \text{if } Tr(\lambda^{2^{k+1}}) = 0 \\
\pm 2^{(n+1)/2} & \text{if } Tr(\lambda^{2^{k+1}}) = 1.
\end{cases}
\]

Theorem 11. The Kasami function \( f(x) \) in Lemma 7 is 1-resilient, \( N_f = 2^{n-1} - 2^{(n-1)/2} \) and \( \Delta_f = 2^{(n+1)/2} \). Moreover, the algebraic degree is \( \deg(f) = \left\lceil n/3 \right\rceil + 1 \).

Proof. By Lemma 7, \( f \) is balanced because \( \hat{f}(0) = 0 \). Also by Lemma 7, \( \sigma_f(x) = Tr(x^{2^k+1}) \) which is preferred by [22]. Therefore, we can apply Theorem 10 because \( f \) is balanced and both functions \( f, \sigma_f \) are preferred.

For any \( k \), the degree of \( f \) is \( wt(2^{2k} - 2^k + 1) = k+1 \) for \( 1 \leq k \leq (n-1)/2 \) and \( wt(2^{2k} - 2^k + 1) = (n-k) + 1 \) when \( (n-1)/2 \leq k \leq n-1 \) [38]. When \( 3k \equiv 1 \pmod{n} \), \( k \equiv \pm \left\lfloor n/3 \right\rfloor \pmod{n} \). Therefore \( \deg(f) = \left\lceil n/3 \right\rceil + 1 \) in this case.
Our next construction is based on a class of functions from the construction of cyclic Hadamard
difference sets by Dillon and Dobbertin [17]. Denote by \( \text{Im}(f) \) the image set of the function \( f(x) \).

**Lemma 8.** (Dillon-Dobbertin [17]) Let \( n \) be odd. Define \( f : \mathbb{GF}(2^n) \to \mathbb{GF}(2) \) by

\[
f(x) = \begin{cases} 
0 & \text{if } x^{2^{k+1}} \in \text{Im}(\Delta_k) \\
1 & \text{if } x^{2^{k+1}} \notin \text{Im}(\Delta_k),
\end{cases}
\]

where \( \Delta_k(x) = (x+1)^d + x^d + 1 \), \( d = 2^{2k} - 2^k + 1 \) and \( \gcd(k,n) = 1 \). Then \( f(x) \) is preferred and satisfies

\[
\hat{f}(\lambda) = \begin{cases} 
0 & \text{if } \text{Tr}(\lambda^{x^{2^{k+1}}}) = 0 \\
\pm 2^{(n+1)/2} & \text{if } \text{Tr}(\lambda^{x^{2^{k+1}}}) = 1.
\end{cases}
\]

**Theorem 12.** Let \( f(x) \) be the Dillon-Dobbertin function in Lemma 8

1. \( f(x) \) is 1-resilient and \( N_f = 2^{n-1} - 2^{(n-1)/2} \).
2. Furthermore, if \( k = 3 \) or \( 3k \equiv 1 \pmod{n} \), then \( f(x) \) is 1-resilient, \( N_f = 2^{n-1} - 2^{(n-1)/2} \) and \( \Delta_f = 2^{(n+1)/2} \).
3. If \( 3k \equiv 1 \pmod{n} \), then the linear span of \( f \) is \( 5n \) where

\[
f(x) = \text{Tr}(x^3 + x^{2^{k+1}} + x^{2^{2k+1}+1} + x^{2^{2k+2}+2^{k+1}+1} + x^{2^{2k+1}+2^{k+1}+3}). \tag{4.4}
\]

The algebraic degree satisfies \( \deg(f) = 4 \) for \( n \neq 5 \). \( \deg(f) = 3 \) for \( n = 5 \).

**Proof.**

1. By Lemma 8, \( f \) is balanced because \( \hat{f}(0) = 0 \). Also by Lemma 8, \( f \) is preferred and the dual function is \( \sigma_f(x) = \text{Tr}(x^{(2^{k+1})/3}) \) which is not affine. Therefore, we can apply Theorem 9.

2. If \( k = 3 \), then \( (2^k + 1)/3 = 3 \) and \( \sigma_f(x) = \text{Tr}(x^3) \) which is preferred by [22].

If \( 3k \equiv 1 \pmod{n} \), then

\[
2^{3k} + 1 \equiv 2 + 1 \equiv 3 \pmod{2^n - 1}.
\]

By using the above identity and the equation \( a^3 + 1 = (a + 1)(a^2 - a + 1) \) with \( a = 2^k \), we have

\[
(2^k + 1)/3 \equiv (2^k + 1)/(2^{3k} + 1) \equiv 1/(2^{2k} - 2^k + 1) \equiv d^{-1} \pmod{2^n - 1}.
\]
where \( d = 2^{2k} - 2^k + 1 \) as in Lemma 8. Therefore the dual function is \( \sigma_f(x) = Tr(x^{d-1}) \) which is preferred by the following argument:

We define \( g(x) := Tr(x^d) \) which is preferred from [38]. This implies \( \sigma_f(x) \) is preferred from the following computation:

\[
\tilde{\sigma}_f(\lambda) = \sum_x (-1)^{Tr(x^{d-1}) + Tr(\lambda x)} = \sum_y (-1)^{Tr(\lambda^{-d-1} y) + Tr(y^d)} = \tilde{g}(\lambda^{-d-1}),
\]

where we let \( x = \lambda^{-1} y^d \). Therefore, we can apply Theorem 10 because \( f \) is balanced and both functions \( f, \sigma_f \) are preferred.

3. \( f(x) \) is a \( 2^k + 1 \)-decimation of the characteristic function \( b_k(x) \) of the Hadamard difference set \( B_k \) in [17, 18], i.e. \( f(x) = b_k(x^{2^k+1}) \). When \( 3k \equiv 1 \) (mod \( n \)), the trace representation of \( b_k(x) \) in [17, 18] is

\[
b_k(x) = Tr(x^{2^{2k}+2^k+1} + x^{2^{2k}+2^k} + x^{2^{2k}-2^k+1} + x^{2^k+1} + x).
\]

The exponents of \( f(x) \) in Equation 4.4 are obtained by multiplying all the exponents of \( b_k(x) \) by \( 2^k + 1 \), and noting that \( 3k \equiv 1 \) (mod \( n \)) implies \( 2^{3k} \equiv 2 \) modulo \( 2^n - 1 \). When we expand the 5 trace terms of \( f(x) \), we get \( 5n \) monomials, i.e. \( LS(f) = 5n \).

The maximum weight of the exponents of \( f(x) \) in Equation 4.4 is 4 which comes from the exponent \( 2^{2k+1} + 2^k + 3 \) for \( n > 5 \). Thus, \( deg(f) = 4 \). For \( n = 5 \), we have \( k = 2 \) and the exponent \( 2^{2k+1} + 2^k + 3 \) modulo 31 is 12 which has weight 2 instead of 4. Thus the maximum weight of the exponents of \( f(x) \) is 3 (from the exponent \( 2^4 + 2^3 + 1 = 25 \)) which implies \( deg(f) = 3 \).

\( \square \)

Our final construction is based on the hyperoval functions in [16].

**Lemma 9.** (Hyperoval [16]) Let \( n \) be odd. Define \( f : GF(2^n) \rightarrow GF(2) \) by

\[
f(x) = \begin{cases} 
0 & \text{if } x \in Im(D_k) \\
1 & \text{if } x \notin Im(D_k)
\end{cases}
\]

where \( D_k(x) = x + x^k \) is a 2-to-1 map on \( GF(2^n) \). Then \( f(x) \) satisfies

\[
\hat{f}(\lambda) = \hat{g}(\lambda^{\frac{2^n-1}{n}})
\]

where \( g(x) = Tr(x^k) \).
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

The known values of \(k\) for which \(x + x^k\) is a 2-to-1 map are \(k = 2\) (Singer), \(k = 6\) (Segre) and two cases due to Glynn. We will consider the Segre case \(k = 6\).

**Theorem 13.** Let \(f(x)\) be the Segre hyperoval function for \(k = 6\) in Lemma 9. Then \(f(x)\) is 1-resilient and \(N_f = 2^{n-1} - 2^{(n-1)/2}\).

**Proof.** We see from Lemma 9 that when \(k = 6\), \(f\) is balanced because \(\hat{f}(0) = \hat{g}(0) = 0\) where \(g(x) = Tr(x^6) = Tr(x^3)\). \(f\) is preferred because \(g\) is preferred [22] and \(\hat{f}(\lambda) = \hat{g}(\lambda^{5/6})\) by Lemma 9. By the distribution of \(\hat{g}\) from [22] and Lemma 9,

\[
\hat{f}(\lambda) = \hat{g}(\mu) = \begin{cases} 
0 & \text{if } Tr(\mu) = 0 \\
\pm 2^{(n+1)/2} & \text{if } Tr(\mu) = 1 
\end{cases}
\]

where \(\mu = \lambda^{5/6}\). Therefore, \(\sigma_f(x) = Tr(x^{5/6})\) which is not affine. Thus, we can apply Theorem 9 because \(f\) is a balanced preferred function and \(\sigma_f\) is not affine.

The Welch-Gong Transformation functions also correspond to optimal Boolean functions with low additive autocorrelation, as indicated in the following remark.

**Remark 8.** In [29], the Welch-Gong Transformation function was shown to have good cryptographic properties: 1-resiliency, high nonlinearity \(2^{n-1} - 2^{(n-1)/2}\), high algebraic degree \(\deg(f) = [n/3] + 1\) and large linear span \(LS(f) = n(2^{[n/3]} - 3)\). A description of the function can be found in [29, Section 2].

Here, we remark that the Welch-Gong function has the additional property of optimal additive autocorrelation. This is because the Welch-Gong Transformation function \(f(x)\) is a balanced preferred function and its dual function is \(\sigma_f(x) = Tr(x^{d^{-1}})\), \(d = 2^{2k} - 2^k + 1\) where \(3k \equiv 1 \pmod{n}\) [29, Lemma 2]. By the same reason as in Theorem 12 part (2), we can apply Theorem 10 because \(f\) is balanced and both \(f\) and \(\sigma_f\) are preferred.

In [16, 17], the ideal 2-level autocorrelation sequences were proven via Parseval’s relation, by showing that the associated function is preferred and the dual is an m-sequence (which is usually non-affine or preferred). This is a powerful and useful technique, which will be applied to find more 2-level autocorrelation sequences for applications in communication systems. By Theorems 9 and 10, they give rise to Boolean functions with high nonlinearity, resiliency and low additive autocorrelation. We present Table 4.1 to summarize the results we have obtained.
### Table 4.1: Cryptographic properties of preferred functions

<table>
<thead>
<tr>
<th></th>
<th>Kasami* (Theorem 11)</th>
<th>Dillon-Dobbertin* (Theorem 12)</th>
<th>Segre hyperoval* (Theorem 13)</th>
<th>Welch-Gong* (Remark 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Balance</strong></td>
<td>Balance</td>
<td>Balance</td>
<td>Balance</td>
<td>Balance</td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td><strong>Nonlinearity</strong></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td><strong>Resiliency</strong></td>
<td>Resiliency</td>
<td>Resiliency</td>
<td>Resiliency</td>
<td>Resiliency</td>
</tr>
<tr>
<td><strong>Optimal</strong></td>
<td>Optimal</td>
<td>-</td>
<td>Optimal</td>
<td>Additive</td>
</tr>
<tr>
<td><strong>Additive</strong></td>
<td>Additive</td>
<td>-</td>
<td>Additive</td>
<td>Additive</td>
</tr>
<tr>
<td><strong>Autocorrelation</strong></td>
<td>Additive Autocorrelation</td>
<td>-</td>
<td>Autocorrelation</td>
<td></td>
</tr>
<tr>
<td>For odd $n$, $3 \n$</td>
<td>For odd $n \geq 5$</td>
<td>For odd $n \geq 5$</td>
<td>For odd $n$, $3 \n$</td>
<td></td>
</tr>
<tr>
<td>$\deg(f) = \lceil n/3 \rceil + 1$</td>
<td>$\deg(f) = 3, 4$</td>
<td>$\deg(f) = 3, 4$</td>
<td>$\deg(f) = \lceil n/3 \rceil + 1$</td>
<td></td>
</tr>
<tr>
<td>when $k = 3^{-1}$</td>
<td>$-\n$</td>
<td>$-\n$</td>
<td>$LS(f) = n(2^{\lceil n/3 \rceil} - 3)$</td>
<td></td>
</tr>
</tbody>
</table>

**Remark:** All the functions listed in Table 4.1 satisfy 2-level (multiplicative) autocorrelation, i.e. $C_f(\lambda) = \sum_{x \in GF(2^n)} (-1)^{f(x) + f(\lambda x)} = 0$ for all $\lambda \neq 1$ [16, 17, 18].
4.2.1 Examples

We present some examples to demonstrate our construction for Boolean functions with desirable properties.

**Example 14.** Let $GF(2^5)^*\rightarrow GF(2^5)$ be generated by the element $\alpha$ satisfying $\alpha^5 + \alpha^3 + 1 = 0$. Define $f : GF(2^5) \rightarrow GF(2)$ by $f(x) = Tr_1^2(x^{13})$, which is the Kasami function of Theorem 11 with $3k \equiv 1$ (mod 5), i.e. $k = 2$. By applying Algorithm 1 with the five linearly independent vectors $(\alpha^i | i = 3, 6, 7, 11, 12)$ satisfying $f(\alpha^i) = 0$, a 1-resilient Boolean form of $f$ is given by

$$g(x_4, x_3, x_2, x_1, x_0) = x_4x_3x_2 + x_4x_3x_0 + x_4x_2x_0 + x_4x_1x_0 + x_3x_2x_1 + x_3x_2x_0 + x_3x_1x_0 + x_3x_2x_1 + x_2x_0 + x_3x_1 + x_2 + x_0.$$ 

$g$ is a Boolean function with five input bits, is 1-resilient and has algebraic degree 3, which is highest among all preferred functions by Proposition 6. The nonlinearity is $2^4 - 2^2 = 12$ which is optimal among 5-bit Boolean functions, see [62]. The additive autocorrelation is optimal, given by $\Delta_f = 2^{(5+1)/2} = 8$.

**Example 15.** Let $GF(2^7)^*\rightarrow GF(2^7)$ be generated by the element $\alpha$ satisfying $\alpha^7 + \alpha + 1 = 0$. Define $f : GF(2^7) \rightarrow GF(2)$ by $f(x) = Tr_1^7(x^3 + x^5 + x^9 + x^{19} + x^{29})$, which is the Dillon-Dobbertin function of Theorem 12 with $3k \equiv 1$ (mod 7), i.e. $k = 5$ (we have reduced the exponents of $f$ to their cyclotomic coset leaders). By applying Algorithm 1 with the seven linearly independent vectors $(\alpha^i | i = 1, 2, 3, 4, 5, 6, 13)$ satisfying $f(\alpha^i) = 0$, a 1-resilient Boolean form of $f$ is given by

$$g(x_6, x_5, x_4, x_3, x_2, x_1, x_0) = x_6x_3x_1x_0 + x_5x_4x_1x_0 + x_5x_3x_2x_0 + x_5x_3x_2x_0 + x_4x_2x_1 + x_4x_2x_1 + x_4x_1x_0 + x_3x_2x_0 + x_3x_1x_0 + x_2x_0 + x_5x_4 + x_5x_3 + x_5x_2 + x_5x_0 + x_4x_3 + x_4x_2 + x_4x_0 + x_3x_1 + x_3x_0 + x_2x_1 + x_2x_0 + x_1x_0 + x_6 + x_4 + x_2 + x_1 + x_0.$$ 

$g$ is a Boolean function with seven input bits, is 1-resilient and has algebraic degree 4, which is highest among all preferred functions by Proposition 6. The nonlinearity is $2^6 - 2^3 = 56$ which is optimal among 7-bit Boolean functions, see [62]. The additive autocorrelation is optimal, given by $\Delta_f = 2^{(7+1)/2} = 16$. 
4.3 Additive Autocorrelation of Known Resilient Functions with High Nonlinearity

4.3.1 Comparison with Known Resilient Preferred Functions

In the known literature, there are many constructions for resilient preferred functions. Some examples include [4, 6, 29, 69]. One common construction is the Maiorana-McFarland class; see [6] for a summary. We present in Proposition 7 a general construction for Maiorana-McFarland resilient preferred functions, which can be deduced from a description given in [6, page 555]. Then, we show that their additive autocorrelation is not as good as for our resilient preferred functions, which are listed in Table 4.1.

**Proposition 7.** (Carlet [6]) Let \( n \) be odd and \( f : \mathbb{GF}(2)^n \rightarrow \mathbb{GF}(2) \) be defined by

\[
f(x, y) = x \cdot \phi(y) + g(y), \quad x \in \mathbb{GF}(2)^{(n+1)/2}, y \in \mathbb{GF}(2)^{(n-1)/2}
\]

where \( g : \mathbb{GF}(2)^{(n-1)/2} \rightarrow \mathbb{GF}(2) \) is any Boolean function, and \( \phi : \mathbb{GF}(2)^{(n-1)/2} \rightarrow \mathbb{GF}(2)^{(n+1)/2} \) is an injection such that \( wt(\phi(y)) \geq k + 1 \) and

\[
|\{x \in \mathbb{GF}(2)^{(n+1)/2} | wt(x) \geq k + 1\}| \geq 2^{(n-1)/2}.
\]

Then \( f \) is a \( k \)-resilient function with nonlinearity \( 2^n - 2^{(n-1)/2} \).

The above construction is quite useful. For example, when \( n \equiv 1 \pmod{4} \), we can construct \((n - 1)/4\)-resilient functions having nonlinearity \( 2^n - 2^{(n-1)/2} \) [6, page 555].

We will show that the function \( f(x, y) \) in Proposition 7 is preferred, and deduce its dual function and additive autocorrelation in Theorem 14. Note that the lower bound for additive autocorrelation of resilient functions in [74, Theorem 4] and [81, Theorem 2] are sharp only for high order of resiliency. Here, we consider functions which may have a low order of resiliency, thus our computation using the dual function can produce a sharper bound. We define the characteristic function \( \chi_A(x) \) on a set \( A \) to be: \( \chi_A(x) = 1 \) if \( x \in A \) and 0 otherwise. We also denote the image set of \( \phi \) by \( \text{Im}(\phi) \).

**Theorem 14.** Let \( f(x, y) \) be a function as defined in Proposition 7. Then it is preferred with dual function

\[
\sigma_f(x, y) = \chi_{\text{Im}(\phi)}(x).
\]
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

The additive autocorrelation of \( f \) satisfies

\[
\Delta f \geq 2^{(n-1)/2} \sqrt{(2^{n+1}/(2^{(n+1)/2} - 1))},
\]

where the lower bound is approximately \( 2^{3n/4} > 2^{(n+1)/2} \).

Proof. From [6], the Hadamard transform of \( f \) is:

\[
\hat{f}(a, b) = \sum_{y \in GF(2)^{(n-1)/2}} (-1)^{g(y) + b \cdot y} \sum_{x \in GF(2)^{(n+1)/2}} (-1)^{x(a + \phi(y))}
\]

\[
= 2^{(n+1)/2} \sum_{y \in \phi^{-1}(a)} (-1)^{g(y) + b \cdot y}
\]

\[
= \begin{cases} 
0, & \text{if } a \notin Im(\phi), \\
\pm 2^{(n+1)/2}, & \text{if } a \in Im(\phi),
\end{cases}
\]

because \( \phi \) is injective. Therefore, \( f \) is preferred and \( \sigma_f(x, y) = \chi_{Im(\phi)}(x) \). To find the additive autocorrelation of \( f \), we can compute \( \widehat{\sigma_f} \):

\[
\widehat{\sigma_f}(a, b) = \sum_{x \in GF(2)^{(n+1)/2}, y \in GF(2)^{(n-1)/2}} (-1)^{\chi_{Im(\phi)}(x) + a \cdot x + b \cdot y}
\]

\[
= \sum_{y \in GF(2)^{(n-1)/2}} (-1)^{b \cdot y} \sum_{x \in GF(2)^{(n+1)/2}} (-1)^{\chi_{Im(\phi)}(x) + a \cdot x}
\]

\[
= \begin{cases} 
0, & \text{if } b \neq 0, \\
2^{(n-1)/2} \chi_{Im(\phi)}(a), & \text{if } b = 0.
\end{cases}
\]

By Lemma 5, the additive autocorrelation

\[
\Delta f = \max_{(a, b) \neq (0, 0)} |\widehat{\sigma_f}(a, b)| = 2^{(n-1)/2} \max_{a \neq 0} |\chi_{Im(\phi)}(a)|.
\]

Note that \( \chi_{Im(\phi)}(0) = 0 \) because \( \chi_{Im(\phi)} \) is balanced. Therefore, by Parseval’s equation: \( \sum_{a \neq 0} \chi_{Im(\phi)}(a)^2 = 2^{n+1} \). Thus \( \max_{a \neq 0} |\chi_{Im(\phi)}(a)| \geq \sqrt{2^{n+1}/(2^{(n+1)/2} - 1)} \) and

\[
\Delta f = \max_{(a, b) \neq (0, 0)} |\widehat{\sigma_f}(a, b)| \geq 2^{(n-1)/2} \sqrt{2^{n+1}/(2^{(n+1)/2} - 1)}.
\]

This lower bound is approximately \( 2^{n/2} \times \sqrt{2^{n/2}} = 2^{3n/4} \).

Remark 9. We note that for \( n \) odd, Canteaut et. al. [4, Corollary 3] constructed 1-resilient preferred functions that can achieve all algebraic degrees between 2 and \( (n + 1)/2 \) by restricting a Maiorana-McFarland bent function to a hyperplane. The dual function and additive autocorrelation can be computed by a method similar to the proof of Theorem 14.
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

By Theorem 14, we see that the Maiorana-McFarland construction of resilient preferred functions have higher (worse) additive autocorrelation than our constructed functions from Table 4.1. Note that the Maiorana-McFarland function in equation (4.5) becomes linear when we fix \((n-1)/2\) input bits \(y\). This may be a weakness. We demonstrate with two examples that our construction can avoid this weakness:

Example 16. The construction of Proposition 7 for a 5-bit 1-resilient preferred function is

\[
f(x, y) = x \cdot \phi(y) + g(y), \ x \in GF(2)^3, y \in GF(2)^2.
\]

where \(\phi : GF(2)^2 \rightarrow GF(2)^3\) maps to the vectors \(\{011, 101, 110, 111\}\).

\(f\) becomes linear when we fix two bits \(y = (y_0, y_1)\). In comparison, our 1-resilient preferred function of Example 14 is not linear when we fix any two bits. By Theorem 14, \(\Delta_f \geq \lceil 2^2 \times \sqrt{2^6/(2^3-1)} \rceil = 13\) while the additive autocorrelation of our function in Example 14 is 8.

Example 17. The construction of Proposition 7 for a 7-bit 1-resilient preferred function is

\[
f(x, y) = x \cdot \phi(y) + g(y), \ x \in GF(2)^4, y \in GF(2)^3.
\]

where \(\phi : GF(2)^3 \rightarrow GF(2)^4\) is an injection such that \(wt(\phi(y)) \geq 2\).

\(f\) becomes linear when we fix three bits \(y = (y_0, y_1, y_2)\). In comparison, our 1-resilient preferred function in Example 15 is not linear when we fix any three bits. By Theorem 14, we deduce that \(\Delta_f \geq \lceil 2^3 \times \sqrt{2^8/(2^3-1)} \rceil = 34\) while the additive autocorrelation of our function in Example 15 is 16.

4.3.2 Potential Weaknesses of Saturated Functions

In this section, we investigate an important class of Boolean functions with 3-valued spectrum. They are the saturated functions introduced by Sarkar and Maitra [65]. We will give a description of the saturated functions and their strengths. Then we point out some potential weaknesses that need to be taken into consideration before deploying them.

If an \(n\)-bit Boolean function \(f\) has algebraic degree \(d\), then the maximal order of resiliency it can achieve is \(k = n - 1 - d\) by Siegenthaler’s inequality [71]. If the order of resiliency is \(k\), then the maximal nonlinearity it can achieve is \(2^{n-1} - 2^{k+1}\) by the Sarkar-Maitra inequality [65,
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

Theorem 2]. When both these conditions are achieved, we say \( f \) is a saturated function and it necessarily has 3-valued spectrum \( 0, \pm 2^{k+2} \) [65]. For fixed \( d \geq 2 \), Sarkar and Maitra constructed an infinite class of saturated functions having algebraic degree \( d \) in [65]. They denoted their class of functions by \( SS(d - 2) \). We will present their construction using the Maiorana-McFarland functions.

**Proposition 8.** ([65, Theorem 5] Sarkar and Maitra) Construction for functions in \( SS(d - 2) \).

Fix \( d \geq 2 \), let \( n = r + s + t \) where \( r = 2^{d-1} - 1 \), \( s = d - 1 \) and \( t \geq 0 \). Define \( f : GF(2)^n \rightarrow GF(2) \) by

\[
f(x, y, z) = x \cdot \phi(y) + g(y) + z_0 + \cdots + z_{t-1}, \ x \in GF(2)^r, y \in GF(2)^s, z \in GF(2)^t
\]

(4.6)

where \( g : GF(2)^s \rightarrow GF(2) \) is any function and \( \phi : GF(2)^t \rightarrow GF(2)^r \) is an injection such that \( wt(\phi(y)) \geq r - 1 \) for all \( y \in GF(2)^s \). Then \( \text{deg}(f) = d \), and \( f \) is \( k \)-resilient having nonlinearity \( 2^{n-1} - 2^{k+1} \) where \( k = n - 1 - d \). \( f \) necessarily has 3-valued spectrum \( 0, \pm 2^{(k+2)} \).

The Boolean functions in Proposition 8 achieve two very desirable properties, maximal resiliency and nonlinearity, for applications in stream cipher systems. However, we will see in Theorem 15 that they have certain weaknesses that have to be considered in applications.

A good lower bound for the additive autocorrelation of a general \( k \)-resilient function\(^2\) is given by Tarannikov et. al. in [74, Theorem 4]. In Theorem 15, we give a better (sharper) bound for the autocorrelation of the saturated functions in Proposition 8.

**Theorem 15.** Let \( d \geq 2 \) and let \( f(x, y, z) \) be a function in \( SS(d - 2) \) as defined in Proposition 8.

1. When \( t = 0 \), \( \Delta_f \geq (1 - \frac{2}{n})2^n \). This bound is sharper than a general bound obtained by [74, Theorem 4]. Thus \( \Delta_f \) has an asymptotic linear structure: \( \Delta_f/2^n \rightarrow 1 \) as \( n \rightarrow \infty \).

2. When \( t \geq 1 \), \( \Delta_f = 2^n \) and the function has linear structures at \((0, 0, a)\) for all \( a \in GF(2)^t \).

Furthermore, when we fix \( \log_2(n) \) or fewer input bits, \( f \) becomes linear.

\(^2\)We note that such a lower bound was first given by Zheng and Zhang in [81, Theorem 2]. Later this bound was improved by Tarannikov et. al. in [74, Theorem 4] for functions with order of resiliency \( \geq (n - 3)/2 \).
CHAPTER 4. ADDITIVE AUTOCORRELATION OF RESILIENT BOOLEAN FUNCTIONS

Proof. 1. When \( t = 0 \), we do not have any \( z \)-terms in equation (4.6):

\[
\Delta_f(a, b) = \sum_{x \in GF(2)^s, y \in GF(2)^s} (-1)^{x \cdot \phi(y) + g(y) + (x + a) \cdot \phi(y + b) + g(y + b)}
\]

\[
= \sum_{y \in GF(2)^s} (-1)^{y \cdot \phi(y + b) + g(y + b)} \sum_{x \in GF(2)^s} (-1)^{x \cdot \phi(y) + \phi(y + b)}
\]

\[
= \begin{cases} 
0, & \text{if } b \neq 0, \\
2^r \sum_{y \in GF(2)^s} (-1)^{a \cdot \phi(y)}, & \text{if } b = 0.
\end{cases}
\]

where we use the fact that \( \phi \) is injective, so \( \phi(y) + \phi(y + b) = 0 \) if and only if \( b = 0 \).

Note that \( |\{ x \in GF(2)^r | wt(x) \geq r - 1 \}| = r + 1 = 2^s = |Im(\phi)| \). Therefore we necessarily have

\[
Im(\phi) = \{ x \in GF(2)^r | wt(x) \geq r - 1 \}
= \{ 011 \ldots 11, 101 \ldots 11, 110 \ldots 11, \ldots, 111 \ldots 10, 111 \ldots 11 \},
\]
and \( \{ 100 \ldots 0 \cdot \phi(y) | y \in GF(2)^s \} = \{ 0, 1, 1, \ldots, 1, 1 \} \). This implies

\[
|\Delta_f(100 \ldots 0, 000 \ldots 0)| = 2^r \left| \sum_{y \in GF(2)^s} (-1)^{100 \ldots 0 \cdot \phi(y)} \right|
= 2^r (2^s - 2) = 2^n - 2^{r+1} = 2^n - 2^{n-s+1}.
\]

Therefore, \( \Delta_f \geq 2^n - 2^{n-s+1} \geq 2^n - 2^{n - \log_2(n)+1} = (1 - 2/n)2^n \). The last inequality is true because \( n = r + s = 2^s - 1 + s \implies 2^s \leq n \implies s \leq \log_2(n) \).

By [74, Theorem 4], \( \Delta_f \geq \left( \frac{2k-n+3}{n+1} \right)2^n = \left( 1 - \frac{2(s+1)}{n+1} \right)2^n \) where \( k = r - 2 \) is the order of resiliency of \( f \). It is easy to see by elementary algebra that their lower bound \( \left( 1 - \frac{2(s+1)}{n+1} \right)2^n \) is not as sharp as our bound \( (1 - \frac{2}{n})2^n \).

2. When \( t \geq 1 \), \( \Delta_f(0, 0, a) = 2^{s+r} \sum_z (-1)^{11 \ldots 1z+11 \ldots 1(z+a)} = \pm 2^{s+r+t} = \pm 2^n \).

It is easy to see that equation (4.6) becomes linear when we fix \( y \) which has \( s \leq \log_2(n) \) bits.

\( \square \)

Remark 10. We have used the direct approach for computing additive autocorrelation in Theorem 15 because it gives sharper bounds than by using dual functions.

We have tried this approach on the Maiorana-McFarland preferred functions in Theorem 14. However, the computation of \( \sum_y (-1)^{a \cdot \phi(y)} \) is more complex. Usually, there are several ways to
choose $\text{Im}(\phi)$ in Proposition 7 and it is more difficult to repeat our computation in Theorem 15. For example, in Example 17 where $|\{x \in GF(2)^4 | \text{wt}(x) \geq 2\}| = 11$, there are $\binom{11}{8} = 165$ choices for the set $\text{Im}(\phi)$. Thus we used the dual function approach.

The saturated functions constructed by Proposition 8 are ‘nearly linear’ because they become linear when we fix very few ($\leq \log_2(n)$) bits. Moreover, the function used in the basic construction (i.e., when $t = 0$) has asymptotic linear structures. And all the functions extended from the basic function by concatenation (i.e., those with $t \geq 1$) have linear structures.

We see that, although the resilient functions in Proposition 8 optimize Siegenthaler [71] and Sarkar-Maitra [65, Theorem 2] inequality, these strong conditions may cause the function to have a linear-like structure. These potential weaknesses have to be considered before using them in applications.

### 4.4 Summary of Chapter 4

The dual function was introduced as a useful concept in the study of functions with 3-valued spectrum. A function with 3-valued spectrum is correlation immune if and only if its dual function is non-affine. Moreover, cryptographic properties like propagation criterion and correlation immunity, high nonlinearity and low additive autocorrelation are dual concepts.

From a recent technique by Dillon and Dobbertin for constructing ideal 2-level (multiplicative) autocorrelation sequences, we obtain preferred functions for which the dual functions are non-affine or preferred. They correspond to Boolean functions with high nonlinearity, resiliency and optimal additive autocorrelation, suitable for cryptographic applications. From [23], the ideal 2-level autocorrelation sequences also correspond to cyclic Hadamard difference sets. Thus we have a useful link between the construction of cyclic difference sets in design theory, 2-level autocorrelation sequences in engineering and the construction of cryptographically useful Boolean functions.

Finally, we showed that our functions have better (lower) additive autocorrelation than highly nonlinear resilient Boolean functions based on the Maiorana-McFarland and Sarkar-Maitra constructions.
“GMW sequences” is an abbreviation for a class of binary sequences first discovered by Gordon, Mills and Welch in 1962 [30]. This was later generalized to cascaded GMW sequences by Klapper, Chan and Goresky where the GMW sequence construction is recursively applied to obtain new sequences with good auto-correlation and cross-correlation properties [44]. The binary cascaded GMW sequences were further generalized by Gong to the q-ary case in [25].

In this chapter, we generalize the cascaded GMW construction in [44] and study three applications in the construction of cryptographic Boolean functions. These are summarized in the following paragraphs.

First, we consider the problem of constructing highly nonlinear resilient functions. We look at preferred functions, i.e. functions whose Hadamard transform only takes on the values 0, ±2^{(n+1)/2}, n odd. Based on a result of Zhang and Zheng [82], we deduce an efficient test for determining when a preferred function is 1-resilient. From any one such function, we can obtain an infinite number of 1-resilient preferred functions by applying the generalized cascaded GMW sequence construction. These functions have nonlinearity 2^{n-1} − 2^{(n-1)/2} (n odd), which is considered
high among resilient functions when $n < 41$ (see Chapter 2, Table 2.9). Moreover, the 3-valued Hadamard transform of our functions limit the number of parity check equations that can be used for the soft output joint attack [49]. By taking the direct sum of our construction for the odd case with the highly nonlinear Patterson-Wiedemann function [62, 63], we construct 1-resilient functions with nonlinearity $> 2^{n-1} - 2^{n/2}$ for even number of input bits $n$.

Second, we consider the problem of constructing balanced function with nonlinearity exceeding $2^{n-1} - 2^{(n-1)/2}$ when $n$ is odd (see Chapter 2, Table 2.7). Previous approaches have been to take the direct sum of two highly nonlinear functions, one of which is balanced [66, 68]. Our approach is a new one based on recursive composition of a highly nonlinear balanced function with quadratic functions. By applying our construction to the highly nonlinear balanced Boolean functions of [66, 68], we obtain new balanced Boolean functions with high nonlinearity $> 2^{n-1} - 2^{(n-1)/2}$, large linear span and high algebraic degree.

Finally, we consider the problem of constructing balanced vectorial Boolean functions to be used as nonlinear combiners in stream ciphers. Such functions will have higher throughput than 1-bit output functions. However, they need to have high nonlinearity to protect against linear approximation attacks [53, 71], and low maximum correlation to protect against approximation by nonlinear functions of output bits [77]. Our construction yields balanced vectorial Boolean functions with nonlinearity $2^{n-1} - 2^{(n-1)/2}$ and low maximum correlation.

This chapter is organised as follows. In Section 5.1, we derive a useful result on cascaded GMW functions, which is basic to all the construction methods in this chapter. In Section 5.2, we construct highly nonlinear resilient Boolean functions. In Section 5.3, we construct highly nonlinear balanced Boolean function. In Section 5.4, we construct nonlinear balanced vectorial Boolean functions with low maximum correlation.

### 5.1 Generalization of Cascaded GMW Functions

We derive a useful result on cascaded functions that will be applied to construct cryptographic Boolean functions in Sections 5.2, 5.3 and 5.4. We change the notation for a subfield trace function slightly for easier representation of cascaded functions.

**Notation 2.** Let $q = 2^n$. In this Chapter, the trace function from $GF(q^m)$ to the subfield $GF(q)$
Define the **imbalance** of \( f : \text{GF}(2^n) \rightarrow \text{GF}(2) \) as:

\[
I(f) = \sum_{x \in \text{GF}(2^n)} (-1)^{f(x)},
\]

and the correlation between polynomial functions \( f(x) \) and \( g(x) \) at \( \lambda \in \text{GF}(2^n) \) as:

\[
C_{f,g}(\lambda) = \sum_{x \in \text{GF}(2^n)} (-1)^{f(\lambda x) + g(x)}.
\]

**Lemma 10.** (Klapper, Chan, Goresky [44, Theorem 3]) Let \( q = 2^n \), and consider \( g, h : \text{GF}(q) \rightarrow \text{GF}(2) \). The correlation between the pair of functions

\[
g(\text{Tr}_{q}^m(x)) \quad \text{and} \quad h(\text{Tr}_{q}^m(x^{q^i+1}))
\]

at \( \Lambda \in \text{GF}(q^m)^* = \text{GF}(q^m) \setminus \{0\} \) takes on the values:

1. \( q^{m-2}I(g)I(h) \),

2. \( (q^{m-2} - q^{-(m-d)/2-2})I(g)I(h) \pm q^{m-(m-d)/2-1}C_{g(x), h(x^2)}(\lambda), \lambda \in \text{GF}(q)^* \),

where \( d = \gcd(i, m) \).

From Lemma 10, we derive the following result that is basic to all our construction methods.

**Theorem 16.** Let \( q = 2^n \), \( n, n_j \) be odd for \( j = 1, \ldots, l \) and let \( f : \text{GF}(q) \rightarrow \text{GF}(2) \). Define recursively the functions

\[
\begin{align*}
f_0(x) &= f(x) \\
f_j(x) &= f_{j-1}(\text{Tr}_{q_j}^{q_{j-1}}(x^{k_j})), j = 1, \ldots, l
\end{align*}
\]

where \( q_0 = q, q_j = q_{j-1}^{n_j}, k_j = q_{j-1}^{r_j} + 1 \) and \( \gcd(r_j, n_j) = 1 \). Then \( \hat{f}_i(\Lambda), \Lambda \in \text{GF}(q_l)^* = \text{GF}(q_l) \setminus \{0\} \) takes on the values \( \pm q^{n_1 n_2 \cdots n_{l-1} - \frac{1}{2}} \hat{f}(\lambda), \lambda \in \text{GF}(q_l)^* \).

**Proof.** We proceed by induction on \( j \). To prove the base case \( j = 1 \), we apply Lemma 10 by letting \( g(x) = \text{Tr}_2^q(x) \) and \( h(x) = f(x) \) to find the correlation between the functions

\[
g(\text{Tr}_{q_0}^q(x)) = \text{Tr}_2^q(x) \quad \text{and} \quad h(\text{Tr}_{q_0}^q(x^{k_1})) = f_1(x).
\]
CHAPTER 5. NEW CONSTRUCTIONS FOR CRYPTOGRAPHIC BOOLEAN FUNCTIONS FROM CASCADED GMW SEQUENCES

Note that $I(g) = I(\text{Tr}^q_2) = 0$, $C_{g(x), h(x^2)}(\lambda) = \hat{f}(\lambda^2)$ by the properties of the trace function; and $\hat{f}_1(\Lambda)$, $\Lambda \in GF(q_1)^*$ is the correlation between the functions in equation (5.1). By Lemma 10, it takes on the values:

$$0, \pm q^{n_1 - (n_1 - 1)/2 - 1} \hat{f}(\lambda^2) = \pm q^{(n_1 - 1)/2} \hat{f}(\lambda), \lambda \in GF(q)^*.$$

We can substitute $\hat{f}(\lambda^2)$ with $\hat{f}(\lambda)$ since $\lambda \mapsto \lambda^2$ is a permutation on $GF(q)^*$. Therefore, the base case is true.

Suppose that case $j-1$ is true, i.e. $\hat{f}_{j-1}(\Lambda)$, $\Lambda \in GF(q_{j-1})^*$ takes on the values $0, \pm q^{(n_1 \cdots n_{j-1} - 1)/2} \hat{f}(\lambda)$, $\lambda \in GF(q)^*$. To prove the $j$th case, we apply Lemma 10 by letting $g(x) = \text{Tr}^{q_{j-1}}_2(x)$ and $h(x) = f_{j-1}(x)$ to find the correlation between the functions

$$g(\text{Tr}^{q_{j-1}}_{q_j}(x)) = \text{Tr}^{q_{j}}_2(x) \text{ and } h(\text{Tr}^{q_{j-1}}_{q_j}(x^k)) = f_j(x).$$

(5.2)

Similar to the proof of the base case, we deduce from Lemma 10 that $\hat{f}_j(\Lambda)$, $\Lambda \in GF(q_j)^*$ takes on the values:

$$0, \pm q^{n_j - (n_j - 1)/2 - 1} q^{(n_1 \cdots n_{j-1} - 1)/2} \hat{f}(\lambda) = \pm q^{(n_1 \cdots n_j - 1)/2} \hat{f}(\lambda), \lambda \in GF(q)^*.$$

Therefore, the statement is true for all $j = 1, \ldots, l$ by induction.

Remark 11. In [44], Klapper, Chan and Goresky proved that if $f(x) = \text{Tr}(x^{2i+1})$ where $\gcd(i, n) = 1$, then $f_1(x)$ in Theorem 16 is preferred. They called their construction a cascaded GMW sequence. Here we generalize it so that it applies to any function $f(x)$.

Note that Theorem 16 only holds for $\hat{f}_1(\Lambda)$ where $\Lambda \neq 0$. The imbalance of $f_1(x)$, which is $\hat{f}_1(0)$, is given by $q^{n_1 n_2 \cdots n_l - 1} \hat{f}(0)$.

### 5.2 New Construction for Resilient Functions

#### 5.2.1 Construction for Odd Number of Input Bits

In this section, we explore constructions for resilient preferred functions. Such functions have nonlinearity $2^{n-1} - 2^{(n-1)/2}$ which is considered high among resilient functions according to Carlet
[6]. Sarkar and Maitra constructed 1-resilient functions with nonlinearity \( > 2^{n-1} - 2^{(n-1)/2} \) for odd \( n \geq 41 \) [66, Theorem 6]. However, their construction [66, Corollary 2], as well as the many Maiorana-McFarland type constructions in the existing literature (see [6]), correspond to concatenation of linear functions. This may be a weakness as the functions become linear when certain input bits are fixed [6]. Our construction will avoid this weakness. Moreover, our functions are 3-valued which makes the soft output joint attack less efficient [49, Corollary 1].

First, we derive an efficient test for 1-resilient preferred functions based on a result of Zheng and Zhang [82]. Then, we prove that by applying the geometric sequence construction of Klapper, Chan and Goresky [44] on certain resilient preferred functions, we can obtain an infinite number of resilient highly nonlinear functions from them.

**Lemma 11.** (Zheng and Zhang [82, Theorem 2]) Let \( n \) be odd and \( f : GF(2^n) \rightarrow GF(2) \) be a balanced preferred function. If \( f \) does not have a non-zero linear structure, then the Boolean representation of \( f \) is 1-resilient for some basis of \( GF(2^n) \).

**Remark 12.** We stated Lemma 11 in a modified form (from the original in [82]) so that it applies to polynomial functions. Lemma 11 has also been proven in a more general form in [4, Theorem 7]. From the proof [82] of Lemma 11, we see that the set \( \{ \lambda | \hat{f}(\lambda) = 0 \} \) contains \( n \) linearly independent vectors. Based on these \( n \) vectors, Gong and Youssef gave an algorithm to find a basis of \( GF(2^n) \) such that the Boolean representation of \( f \) is 1-resilient [29]; see Algorithm 1 of Chapter 4.

The following equation is well-known, e.g. see [4, 82]:

\[
\frac{1}{2^n} \sum_{\lambda \in GF(2^n)} \hat{f}(\lambda)^4 = \sum_{a \in GF(2^n)} \Delta_f(a)^2. \tag{5.3}
\]

It can be used to derive Corollary 3 which is a more applicable form of Lemma 11.

**Corollary 3.** Let \( n \) be odd and \( f : GF(2^n) \rightarrow GF(2) \) be a balanced preferred function. If there exists \( a \in GF(2^n) \) such that \( \Delta_f(a) \neq 0 \) or \( \pm 2^{n/2} \), then \( N_f = 2^{n-1} - 2^{(n-1)/2} \) and the Boolean representation of \( f \) is 1-resilient for some basis of \( GF(2^n) \).

**Proof.** By equation (2.9) of Chapter 2, a preferred function has nonlinearity \( 2^{n-1} - 2^{(n-1)/2} \). From equation (5.3), the preferred function \( f \) satisfies

\[
1/2^n \times 2^{n-1} \times 2^{2n+2} = \Delta_f(0)^2 + \sum_{x \neq 0} \Delta_f(x)^2.
\]
Since $\Delta_f(0) = 2^n$, we deduce that $\sum_{x \neq 0} \Delta_f(x)^2 = 2^{2n}$. When $f$ has a non-zero linear structure $a$, it must be the case that $|\Delta_f(a)| = 2^n$ and $\Delta_f(x) = 0$ for $x \neq 0, a$. Therefore, if $|\Delta_f(a)| \neq 0$ or $2^n$ for some $a \in GF(2^n)$, then $f$ has no non-zero linear structure and $f$ is 1-resilient by Lemma 11.

Remark 13. To test if a preferred function $f$ is resilient by Lemma 11, we have to check $\Delta_f(a) \neq \pm 2^n$ for all $a$. By using Corollary 3, we have a more efficient way to test if $f$ is resilient since we have to check very few $\Delta_f(a)$ as we shall see later on.

Remark 14. Let $f$ be a balanced preferred function formed from

$$h(x_1, \ldots, x_n) = g(x_1, \ldots, x_{n-1}) + x_n, \quad f(x) = h(xA),$$

where $g$ is bent and $A$ is an invertible $n \times n$ matrix. Then $\beta = (00\ldots01)A^{-1}$ satisfies $\Delta_f(\beta) = -2^n$. This implies $\Delta_f(a) = 0$ for all $a \neq 0, \beta$ by equation (5.3). Therefore, Corollary 3 fails for these functions, including all quadratic preferred functions.

We demonstrate some applications with the following two examples. Recall that the cyclotomic coset leaders modulo $2^n - 1$ are the smallest elements of the sets $\{2^i s \mod 2^n - 1, i = 0 \ldots n-1\}$. Since $Tr(x^{2^i}) = Tr(x^i)$, we just need to look at the cyclotomic coset leaders in the exponents of $Tr(x^i)$.

Example 18. Let $n = 5$. We exhaustively search for balanced preferred functions $f : GF(2^5) \to GF(2)$ of the form $f(x) = \sum_{i \in I} Tr(x^i)$, where $I$ are the cyclotomic coset leaders $\{1, 3, 5, 7, 11\}$. We obtain six non-quadratic preferred functions (all cubic):

$$Tr(x^7), Tr(x^{11}), Tr(x + x^3 + x^{11}),$$

$$Tr(x + x^5 + x^7), Tr(x + x^7 + x^{11}), Tr(x + x^3 + x^5 + x^7 + x^{11}).$$

Since $\Delta_f(1) = 8$ which is $\neq 0, \pm 2^5$ for all these functions, their Boolean representations are all 1-resilient for some basis of $GF(2^5)$. Such a basis can be found by applying Algorithm 1 of Chapter 4.

Example 19. Let $n = 11$. We exhaustively search for balanced preferred functions $f : GF(2^{11}) \to GF(2)$ of the form $f(x) = \sum_{i \in I} Tr(x^i)$, where $I$ is the set of cyclotomic coset leaders modulo 2047, from a restricted search space of size 78387228. We obtain 661 non-quadratic preferred functions, see Chapter 7.
By applying Corollary 3 to these functions, we verified that all these functions are 1-resilient in some basis representation. We just list the functions which achieves the maximal algebraic degree \((n + 1)/2 = 6\) (by Proposition 6 of Chapter 4) in our list:

\[
\begin{align*}
\Tr(x^{71} + x^{309} + x^{359}), & \quad \Tr(x + x^{43} + x^{171} + x^{683}) \\
\Tr(x + x^{13} + x^{411} + x^{413} + x^{423}), & \quad \Tr(x + x^{25} + x^{71} + x^{309} + x^{359}) \\
\Tr(x + x^{57} + x^{231} + x^{343} + x^{683}), & \quad \Tr(x + x^{63} + x^{95} + x^{189} + x^{315}) \\
\Tr(x + x^{143} + x^{151} + x^{365} + x^{429}). & \quad \\
\end{align*}
\]

These seven functions satisfy \(\Delta_f(1) \neq 0, \pm 2^{11}\). Therefore their Boolean representation are 1-resilient in some basis which can be found by applying Algorithm 1 of Chapter 4.

**Remark 15.** In both examples, all the balanced non-quadratic preferred functions we found by exhaustive search are 1-resilient for some basis representation. This seems to suggest that if we randomly choose a balanced non-quadratic preferred function, there is a high chance it is 1-resilient. Moreover, all the functions in our test satisfy \(\Delta_f(a) \neq 0, \pm 2^n\), see Chapter 7. So we can confirm \(\Delta_f(a) \neq 0, \pm 2^n\) at the beginning (if we are testing \(a = \alpha^i\) from \(i = 0, 1, \ldots\) onwards where \(\alpha\) generates \(GF(2^n)\)). This seems to suggest Corollary 3 gives a fast test.

Next, we derive a new construction for 1-resilient preferred functions. Precisely, we show that by composing a 1-resilient preferred function from Corollary 3 with a quadratic function, we can obtain infinitely more functions of the same kind.

**Theorem 17.** Let \(n\) be odd, \(q = 2^n\) and \(f : GF(q) \rightarrow GF(2)\) be a balanced preferred function. If there exists \(a \in GF(q)\) such that \(\Delta_f(a) \neq 0, \pm q\), then \(g : GF(2^N) \rightarrow GF(2), N = mn\), defined by

\[
g(x) = f(a \Tr_q(x^{q^m}(q^i+1))), \quad m \text{ odd, } \gcd(i,m) = 1
\]

is 1-resilient for some basis of \(GF(2^N)\) and preferred which implies \(N_g = 2^{N-1} - 2^{(N-1)/2}\). Moreover, \(\deg(g) = 2\deg(f)\).

**Proof.** First, we prove that \(g(x)\) is preferred. Let us consider the related function:

\[
h(x) = f(\Tr_q(x^{q^m}(q^i+1))).
\]

which is balanced because it is a composition of balanced functions. Therefore \(h(0) = 0\). We apply Theorem 16 with \(l = 1\) to see that \(\hat{h}(\lambda), \lambda \neq 0\), takes on the values:

\[
0, \pm q^{(m-1)/2}2^{(n+1)/2} = \pm 2^{(N+1)/2}, \text{ where } N = mn.
\]
Thus \( h(x) \) is preferred. Because the two functions \( g(x) \) and \( h(x) \) are related by a non-singular linear transform of input \( x \):

\[
g(x) = f(a\text{Tr}_{q^m}(x^{q^i+1})) = f(T_{q^m}(ax^{q^i+1})) = h(ax^{q^i+1}),
\]

we deduce that \( g(x) \) is also preferred. This implies \( N_g = 2^{N-1} - 2^{(N-1)/2} \) by equation (2.9).

Second, we prove that \( g \) is 1-resilient. Let \( y = \text{Tr}_{q^m}(x^{q^i+1}) \). Then

\[
g(x + 1) = f(a\text{Tr}_{q^m}((x+1)^{q^i+1}))
= f(a\text{Tr}_{q^m}(x^{q^i+1}) + \text{Tr}_{q^m}(x^{q^i}) + \text{Tr}_{q^m}(x) + 1))
= f(a(y+1)), \quad \text{because } \text{Tr}_{q^m}(x^{q^i}) = \text{Tr}_{q^m}(x).
\]

From the above identity and the fact that for each \( y \in GF(q) \), there are \( q^{m-1} \) elements \( x \in GF(q^m) \) such that \( y = \text{Tr}_{q^m}(x^{q^i+1}) \), we deduce that:

\[
\Delta_g(1) = \sum_{x \in GF(q^m)} (-1)^{g(x)+g(x+1)}
= q^{m-1} \sum_{y \in GF(q)} (-1)^{f(ay)+f(a(y+a)}
= q^{m-1} \Delta_f(a) \neq 0, \pm q^m, \quad \text{because } \Delta_f(a) \neq 0, \pm q.
\]

By Corollary 3, the Boolean representation of \( g \) is 1-resilient in some basis of \( GF(2^{mn}) \).

Third, the algebraic degree is \( 2\deg(f) \) from Theorem 18 part 3.

By applying Theorem 17 to the resilient functions constructed in Chapter 4, we obtain the following useful corollary.

**Corollary 4.** Let \( q = 2^n \) and let \( f : GF(q) \to GF(2) \) be a preferred function whose dual function \( \sigma_f : GF(q) \to GF(2) \) is non-affine (see Table 4.1).

Then \( g : GF(2^N) \to GF(2), \ N = mn, \) defined by

\[
g(x) = f(a\text{Tr}_{q^m}(x^{q^i+1})), \quad m \text{ odd, } \gcd(i,m) = 1
\]

is 1-resilient for some basis of \( GF(2^N) \) and preferred which implies \( N_g = 2^{N-1} - 2^{(N-1)/2} \). Moreover, \( \deg(g) = 2\deg(f) \).
Proof. Since \( \sigma_f(x) \) is not affine, we have \( \exists a \neq 0 \) such that \( \hat{\sigma}_f(a) \neq 0, \pm 2^n \) (because of Parsevals equation: \( \sum a \hat{\sigma}_f(a)^2 = 2^{2n} \)).

Hence, \( \exists a \neq 0 \) such that \( \Delta_f(a) \neq 0, \pm 2^n \) by Lemma 5 of Chapter 4. Therefore we can apply Theorem 17 on \( f(x) \).

Useful properties of our construction:

1. New construction for resilient function with nonlinearity \( 2^{N-1} - 2^{(N-1)/2}, N \text{ odd} \), for protection against linear and correlation attacks [71, 53]. Moreover, it is not based on concatenation of linear functions. Therefore it avoids a potential weakness of being linear when certain input bits are fixed.

2. Our function has 3-valued Hadamard transform which prevents the use of all even parity check equations in the soft output joint attack [49, Corollary 1]. This will limit the efficiency of that attack.

Example 20. Let \( f : GF(2^5) \rightarrow GF(2) \) be any of the preferred functions in Example 18. Suppose we take the function:

\[
f(x) = Tr_2^{2^5}(x + x^5 + x^7), \quad \deg(f) = 3.
\]

We apply Theorem 17 with \( n = 5, m = 3 \) and \( i = 1 \). Since \( \Delta_f(1) \neq 0, \pm 2^5 \), we can let \( a = 1 \) and define \( g : GF(2^{15}) \rightarrow GF(2) \) by:

\[
g(x) = f(Tr_2^{2^{15}}(x^{33})).
\]

Then the Boolean representation of \( g(x) \) has 15 input bits, algebraic degree = 6, is 1-resilient for some basis of \( GF(2^{15}) \) and has nonlinearity \( 2^{14} - 2^7 = 16256 \). This can be found by applying Algorithm 1 of Chapter 4.

5.2.2 Construction for Even Number of Input Bits

Pasalic and Johansson conjectured that the highest nonlinearity for 1-resilient functions with even number of input bits \( n \) is \( 2^{n-1} - 2^{n/2} \). We will disprove this conjecture by constructing functions with higher nonlinearity. This has also been done by Sarkar and Maitra in [66, 52]. We will make use of the Patterson Wiedemann functions with 15 input bits and nonlinearity 16276 from [62, 63]. The following theorem based on the direct sum construction is easy to prove, e.g. see [52].
**Proposition 9.** Let $f : \text{GF}(2)^n \rightarrow \text{GF}(2)$, $n$ odd, be a $k$-resilient function with nonlinearity $2^{n-1} - 2^{(n-1)/2}$ and $PW : \text{GF}(2)^{15} \rightarrow \text{GF}(2)$ be the 15-bit Patterson Wiedemann function from [62, 63]. Then

$$g(x, y) := f(x) + PW(y), x \in \text{GF}(2)^n, y \in \text{GF}(2)^{15}$$

has $m = n + 15$ (even) number of input bits, is $k$-resilient and has nonlinearity

$$N_g = 2^{m-1} - 27/32 \times 2^{m/2} > 2^{m-1} - 2^{m/2}.$$  

Moreover, $\deg(g) = \max(\deg(f), \deg(PW))$.

Useful properties of our construction:

1. $k$-resilient function with high nonlinearity $> 2^{m-1} - 2^{m/2}$ (which disproves the conjecture of [61]) for protection against linear and correlation attacks [71, 53]. Note that Sarkar and Maitra also disproved the conjecture of [61] by concatenation involving linear functions [66, Theorems 7, 8 and 9]

2. If we use a 1-resilient function $f(x)$ by exhaustive search in Examples 18, 19 or Theorem 17, then both $f(x)$ and the Patterson-Wiedemann function $PW(y)$ are constructed from finite fields which are not concatenation of linear functions. Therefore their direct sum $g(x, y)$ is not one too.

**Example 21.** We apply Proposition 9 by considering the function:

$$f(x) = Tr_{2^5}^{2^2}(x + x^5 + x^7), \deg(f) = 3,$$

from Example 18 and letting $PW(y)$ be the 15-bit Patterson-Wiedemann function with algebraic degree 9 from [62, 63]. We take the direct sum of the 1-resilient Boolean representation of $f(x)$ and $PW(y)$:

$$g(x, y) = f(x) + PW(y), x \in \text{GF}(2)^5, y \in \text{GF}(2)^{15},$$

which has 20 input bits, is 1-resilient, has nonlinearity $2^{19} - 2^{10} + 160$ and algebraic degree 9.

### 5.3 Highly Nonlinear Balanced Boolean Functions

In this section, we construct new classes of balanced Boolean functions with high nonlinearity $> 2^{n-1} - 2^{(n-1)/2}$, $n$ odd, high algebraic degree and large linear span. This is achieved by
applying Theorem 16 on the highly nonlinear balanced functions of Sarkar and Maitra [66] and Seberry, Zhang and Zheng [68].

**Theorem 18.** Let \( q = 2^n \) and \( n, n_j \) be odd for \( j = 1, \ldots, l \). Let \( f : GF(q) \to GF(2) \) be a balanced function with nonlinearity \( > 2^{n-1} - 2^{(n-1)/2} \) (such as the functions from [66, 68]). Define recursively the functions

\[
\begin{align*}
f_0(x) &= f(x) \\
f_j(x) &= f_{j-1}(Tr_{q_j}^{q_{j-1}}(x^{k_j})), j = 1, \ldots, l,
\end{align*}
\]

where \( q_0 = q, q_j = q_{j-1}^{n_j}, k_j = q_{j-1}^{r_j} + 1 \) and \( \gcd(r_j, n_j) = 1 \). Then \( f_l \) corresponds to a Boolean function with \( N = n_1 \cdots n_l \) input bits and

1. \( f_l \) is balanced having high nonlinearity

\[
2^{N-1} - 2^{N-1} \max_{\lambda} |\hat{f}(\lambda)| > 2^{N-1} - 2^{(N-1)/2}.
\]

2. If the polynomial expression for \( f \) is \( \sum_i \beta_i x^{t_i} \), then the linear span of \( f_l \) satisfies

\[
LS(f_l) = \sum_i (n_1 n_2^2 \cdots n_l^{2^{l-1}})^{wt(t_i)}.
\]

Thus \( LS(f_l) \geq (n_1 n_2^2 \cdots n_l^{2^{l-1}})^{\deg(f)} \).

3. The algebraic degree of \( f_l \) is \( 2^l \deg(f) \).

**Proof.** 1. \( f \) is balanced implies that \( f_l \) is balanced since it is a composition of balanced functions. Therefore, \( \hat{f_l}(0) = 0 \). We apply Theorem 16 to see that \( \hat{f_l}(\Lambda), \Lambda \in GF(2^N)^* \) takes on the values:

\[
0, \pm(2^n)^{(n_1 \cdots n_l-1)/2} \hat{f}(\lambda) = 2^{(N-n)/2} \hat{f}(\lambda), \lambda \in GF(q)^*.
\]

By equation (2.9),

\[
N_{f_l} = 2^{N-1} - 2^{\frac{N-1}{2}} \max_{\lambda} |\hat{f}(\lambda)|.
\]

Since \( f \) has nonlinearity \( > 2^{n-1} - 2^{(n-1)/2} \), \( \max_{\lambda} |\hat{f}(\lambda)| < 2^{(n+1)/2} \). This implies

\[
N_{f_l} > 2^{N-1} - 2^{\frac{N-1}{2}} 2^{\frac{n+1}{2}} = 2^{N-1} - 2^{(N-1)/2}.
\]
2. Note that $f_i(x)$ can be written as

$$f_i(x) = \sum_i \beta_i g(x)^{t_i}, \beta_i \neq 0, \beta_i \in GF(2^n)$$

where

$$g(x) = Tr_{q_1}(Tr_{q_2} \ldots Tr_{q_{l-1}}(x^{k_1}) \ldots)^{k_2})^{k_1},$$

is a cascaded GMW function from $GF(2^N)$ to $GF(2^n)$ with known linear span $n_1 n_2 \cdots n_{l-1}$ [44]. Applying Lemma 2-(i) in [25] to each term $g(x)^{t_i}$, we deduce that for each $i \neq j$, the monomial terms in $g(x)^{t_i}$ and $g(x)^{t_j}$ are distinct. According to Theorem 1 and Lemma 2-(ii) in [25], the number of distinct monomials of $g(x)^{t_i}$ is $(n_1 n_2 \cdots n_{l-1})^{wt(t_i)}$. Thus, equation (5.4) is established.

The inequality $LS(f_i) \geq (n_1 n_2 \cdots n_{l-1})^{deg(f)}$ follows from equation (5.4) because there is at least a monomial in the expression of $f$ that has exponent $deg(f)$.

3. According to the proof of Lemma 2 in [25], we can deduce that when we recursively expand each trace term of $f_i$ using the relation

$$\left(\sum_i x^{s_i}\right)^{2^n} = \prod_j \sum_i x^{2^{n_j}s_i},$$

there is no cancellation among the resulting sum of monomials. Therefore the maximal exponent is $2^l \max_i wt(t_i) = 2^l deg(f)$, which is the algebraic degree of $f_i$.

The following proposition will ensure our functions have high algebraic degree.

**Proposition 10.** (Sarkar and Maitra [66, Proposition 2 and 3]) Let $f : GF(2)^N \rightarrow GF(2)$ be a balanced Boolean function. If we change up to two bits in the Boolean truth table of $f$ such that the top half and bottom half truth tables have odd weight, then the new function $f$ will be balanced, $deg(f) = N - 1$ and the nonlinearity will decrease by at most 2.

Useful properties of our construction:

1. Theorem 18 gives a new construction for balanced function $f_i$ with high nonlinearity $> 2^{N-1} - 2^{(N-1)/2}$ not based on taking direct sum.
2. The function can achieve high algebraic degree and large linear span.

For the sake of completeness, we will outline some known constructions for balanced highly non-linear functions used in Theorem 18.

1. Seberry, Zhang and Zheng [68, Theorem 2]: They constructed a balanced function \( g : GF(2)^{2k} \rightarrow GF(2), \ k \geq 7 \), with high nonlinearity by modifying a bent function. Then they take the direct sum of \( g(x) \) with the 15-bit Patterson Wiedemann function \( PW(y) \) having nonlinearity \( 2^{14} - 2^7 \) \([62, 63]\). The resulting balanced function

\[
h(x, y) = g(x) + PW(y), \ x \in GF(2)^{2k}, y \in GF(2)^{15},
\]

has nonlinearity \( > 2^{n-1} - 2^{(n-1)/2} \) where \( n = 2k + 15 \). Their construction works for odd \( n \geq 29 \). (Later, Dobbertin gave a more general method to modify bent functions into balanced functions \( g \) in [19]).

2. Sarkar and Maitra [66, Theorem 13]: They modify the Patterson Wiedemann function \( PW_r \) of \([62, 63]\) for \( r = 15, 17, 19, 21 \) number of input bits to obtain balanced functions \( f_r \) with nonlinearities

\[
2^{14} - 2^7 + 6, \ 2^{16} - 2^8 + 18, \ 2^{18} - 2^9 + 46, \ 2^{20} - 2^{10} + 104,
\]

respectively. Then they take the direct sum of \( f_r \) with a bent function \( g : GF(2)^{2k} \rightarrow GF(2) \). The resulting balanced function

\[
h(x, y) = g(x) + f_r(y), \ x \in GF(2)^{2k}, y \in GF(2)^{r},
\]

has nonlinearity \( > 2^{n-1} - 2^{(n-1)/2} \) where \( n = 2k + r \). Their construction works for odd \( n \geq 15 \).

**Example 22.** We apply Theorem 18 by letting \( f(x) \) be the 15-bit balanced function with nonlinearity \( 2^{14} - 2^7 + 6 = 16262 \) from [66], \( l = 1, n_1 = 3 \) and \( r_1 = 1 \). Then

\[
f_1(x) = f(T_2^{245}(x^{2^{15}+1}))
\]

is balanced, has 45 input bits and nonlinearity

\[
2^{44} - 2^{(45-15)/2-1} \max_{\lambda} |\hat{f}(\lambda)| = 2^{44} - 2^{22} + 196608,
\]
where \( \max_{\lambda} |\hat{f}(\lambda)| = 244 \). As in Proposition 10, we can change at most 2 bits in the Boolean truth table of \( f_1 \) to get a balanced function with nonlinearity \( \geq 2^{44} - 2^{22} + 196606 \) and algebraic degree 44.

To ensure a large linear span, we could also apply Proposition 10 to the 15-bit function \( f \) to get a balanced function with nonlinearity \( \geq 2^{14} - 2^7 + 4 \) and algebraic degree 14. Then \( f_1 \) as defined above will have nonlinearity \( \geq 2^{44} - 2^{22} + 131072 \), algebraic degree \( 2 \times 14 = 28 \) and linear span \( \geq n_1^{\deg(f)} = 3^{14} = 4782969 \) by Theorem 18. The large linear span of \( f_1 \) provides protection against interpolation attack [36].

### 5.4 Highly Nonlinear Balanced Vectorial Functions with Low Maximum Correlation

We construct a new class of balanced vectorial Boolean functions with \( n \) (odd) input bits that have nonlinearity \( 2^{n-1} - 2^{(n-1)/2} \) through geometric sequences. They have low maximum correlation bounds by Proposition 1.

\[
C_F(w) \leq 2^{m/2-n(2^{(n+1)/2})} = 2^{(m-n+1)/2} \text{ for } w \neq 0.
\]  
(5.5)

In stream cipher applications, this ensures that linear approximation and approximation by nonlinear functions of output bits are difficult. We also point out that the maximum correlation can be further reduced by a factor of \( \sqrt{2} \).

Our S-box construction makes use of the almost bent permutations \( x^k \) on \( GF(2^n) \). They were introduced in Chapter 2 where a summary of such functions can be found in Table 2.15. From the theory of CDMA sequences, it is well known that \( Tr^{q_j}(x^k) \) is a preferred function if \( x^k \) is from Table 2.15 [32].

**Theorem 19.** Let \( q = 2^m \) and \( m, n_j \) be odd for \( j = 1 \cdots l \). Let \( x^k, x \in GF(q) \), be an almost bent permutation (see Table 2.15). Define inductively the functions

\[
F_0(x) = x^k \\
F_j(x) = F_{j-1}(Tr_{q_j}^{q_{j-1}}(x^{k_j})), \quad j = 1, \ldots, l
\]

where \( q_0 = q, q_j = q_j^{n_{j-1}}, k_j = q_j^{r_{j-1}} + 1 \) and \( \gcd(r_j, n_j) = 1 \). Let \( N = m n_1 n_2 \cdots n_l \) and \( F(x) = F_l(x) \) where \( F : GF(2^N) \to GF(2^m) \). Then
CHAPTER 5. NEW CONSTRUCTIONS FOR CRYPTOGRAPHIC BOOLEAN FUNCTIONS FROM CASCADED GMW SEQUENCES

1. F is balanced.

2. The nonlinearity is $N_F = 2^{N-1} - 2^{(N-1)/2}$.

3. The maximum correlation satisfies $C_F < 2^{(m-N+1)/2}$.

4. The linear span is $(n_1 n_2^2 \cdots n_l^{2i-1})^{wt(k)}$ and the algebraic degree is $\text{deg}(F) = 2^l wt(k)$.

Proof. 1. F is balanced because it is a composition of balanced functions.

2. Let $f(x) = Tr^2_q(F(x))$. Then $f$ is preferred by Theorem 16 and the fact that it is balanced. We will prove the nonlinearity of $F$ by showing that all linear combinations of output bits given by $Tr^2_q(bF(x))$, $b \in GF(q)$ are preferred. To achieve this, we need the following identity which can be proven by induction on $j$ using the properties of the trace function:

$$bF_j(x) = F_j(b^{(k_1 \cdots k_l)^{-1}_j}x), \text{ for all } b \in GF(q), j = 1, \ldots, l.$$ 

From this, we deduce that $bF(x) = F(b'x)$ where $b' = b^{(k_1 \cdots k_l)^{-1}}$. And the Hadamard transform of $Tr^2_q(bF(x))$ is given by:

$$\sum_x (-1)^{Tr(bF(x)) + Tr(\lambda x)} = \sum_x (-1)^{Tr(F(b'x)) + Tr(\lambda x)}, \text{ where } b' = b^{(k_1 \cdots k_l)^{-1}}$$

$$= \sum_y (-1)^{Tr(F(y)) + Tr((b')^{-1}y)}, \text{ where } y = b'x$$

$$= \hat{f}((b')^{-1} \lambda) = 0, \pm 2^{(N+1)/2}.$$ 

Thus all linear combinations of output bits correspond to Boolean functions with nonlinearity $2^{N-1} - 2^{(N-1)/2}$ by equation (2.9). Therefore $F$ has the same nonlinearity.

3. The maximum correlation bound is a direct application of Proposition 1, since $\hat{w} \cdot F(w)$ takes the values $0, \pm 2^{(N+1)/2}$. The inequality is strict because we can deduce from equation (2.12) that:

$$2^N C_F(w) = \max_g \sum_x (-1)^{g(F(x)) + w \cdot x} \text{ for all } w \neq 0.$$ 

This value is an integer while $2^N \times$ upper bound $= 2^{(m+N+1)/2}$ is not.

4. The algebraic degree and linear span can be proven in a similar way to Theorem 18 parts 2 and 3.
CHAPTER 5. NEW CONSTRUCTIONS FOR CRYPTOGRAPHIC BOOLEAN FUNCTIONS FROM CASCADED GMW SEQUENCES

Remark 16. The upper bound for maximum correlation in Theorem 19 holds for all S-boxes with nonlinearity $2^{n-1} - 2^{(n-1)/2}$. For the case $l = 1$, we can prove $C_F(w) \leq 2^{(m-N)/2}$, i.e. the upper bound for maximum correlation can be reduced by a factor of $\sqrt{2}$ [39, Theorem 2]; see also Chapter 6.

Example 23. We illustrate the application of Theorem 19 with parameters $l = 1, m = 5, n_1 = 3, k = 13, r_1 = 1$:

$$F(x) = \text{Tr}_{2^5}^{2^5} (x^{33})^{13}.$$  

$F$ corresponds to a balanced vectorial Boolean function with 15 input bits and 5 output bits. The function $\text{Tr}_{2^5}^{2^5} (x^{13})$ is preferred as it corresponds to the Kasami exponent in Table 2.15 for $r = 2$. By Theorem 19, the nonlinearity of $F$ is $2^{14} - 2^7 = 16256$. By Remark 16, the maximum correlation satisfies $C_F \leq 2^{(5-15)/2} = 0.03125$. The algebraic degree of $F$ is $2^{wt(k)} = 6$ and its linear span is $\nu_1^{wt(k)} = 3^3 = 27$. From this information, we deduce the following probability bounds for linear approximation of $v \cdot F(x)$ and $g(F(x))$:

$$0.496 \leq \Pr(v \cdot F(x) = w \cdot x) \leq 0.504 \text{ for all } w \in GF(2)^{15}, v \in GF(2)^5,$$

$$0.484 \leq \Pr(g(F(x)) = w \cdot x) \leq 0.516 \text{ for all } w \in GF(2)^{15}, g : GF(2)^5 \rightarrow GF(2).$$

5.5 Summary of Chapter 5

We have applied the theory of geometric sequences, due to Klapper, Chan and Goresky [44], to construct $n$-variable resilient Boolean functions with nonlinearity $2^{n-1} - 2^{(n-1)/2}$, $n$ odd. Moreover, the Hadamard transforms of these functions are 3-valued, which provides protection against the soft output joint attack [49]. They can be extended to construct highly nonlinear resilient functions that disprove Pasalic and Johansson’s conjecture for $n$ even. These functions do not have a weakness shared by Boolean functions formed from concatenating linear functions. We also applied geometric sequences to give a new construction for balanced Boolean functions having high nonlinearity $> 2^{n-1} - 2^{(n-1)/2}$, $n$ odd. This approach is different from previous constructions, which were based on direct sums of highly nonlinear Boolean functions. Finally, we constructed balanced vectorial Boolean functions with high nonlinearity and low maximum correlation. They can be used as combiners for stream ciphers with high throughput.
Chapter 6

Highly Nonlinear S-boxes with Reduced Bound on Maximum Correlation

In this chapter, we consider S-boxes with $n$ (odd) input bits and $m \geq 2$ output bits as combiners in stream cipher systems. S-boxes deployed in such ways can be cryptanalysed using correlation and linear approximation attacks [53, 71]. One way to defend against such attacks is for the S-boxes to have high nonlinearity. When the combining function is a Boolean function, i.e. $m = 1$, there are many constructions for functions having good cryptographic properties and high nonlinearity exceeding the quadratic bound $2^{n-1} - 2^{(n-1)/2}$ [62, 63, 66, 68]. On the other hand, when there is more than one output bit, the stream cipher will have higher throughput but lower nonlinearity. For example, when the input and output of the combining function have the same size, i.e., $m = n$, the highest achievable nonlinearity is $2^{n-1} - 2^{(n-1)/2}$ [9]. When the number of output bits of an S-box satisfies $2 \leq m \leq n - 1$, a nonlinearity of $2^{n-1} - 2^{(n-1)/2}$ is considered high (see Table 2.13).

We construct two classes of balanced S-boxes with high nonlinearity $2^{n-1} - 2^{(n-1)/2}$ for protection against correlation and linear approximation attacks [53, 71]. However, having a high nonlinearity may not be sufficient for security. In [77], Zhang and Chan considered a more general
correlation attack by using a nonlinear function of output bits. In that case, we require the maximum correlation coefficients to be low for protection against their attack. They proved an upper bound for maximum correlation that is low for functions with high nonlinearity. We improve their result for our S-boxes by reducing their upper bound by a factor of $\sqrt{2}$. Therefore our S-boxes are more secure against their general correlation attack.

Besides having high nonlinearity and reduced upper bound on maximum correlation, our S-boxes have other good cryptographic properties. The first class of S-boxes have maximum differential at most twice that of a perfect nonlinear S-box. Therefore they have low maximum differential to make differential cryptanalysis difficult [2]. The second class of S-boxes are based on GMW sequences [44] whose algebraic structure allows us to prove reduced maximum correlation for a larger class of functions. They also have larger linear span which offers more protection against interpolation attacks [36]. Moreover, both classes of S-boxes can achieve high algebraic degree.

This chapter is organized as follows. In Section 6.1, we will construct 2 classes of balanced highly nonlinear S-boxes with a $\sqrt{2}$ reduction in the upper bound on maximum correlation. The first class is constructed in Section 6.1.1 and the construction is based on power functions (equivalent to $m$-sequences used in CDMA applications). The reduced bound on maximum correlation is proven for S-boxes with Gold and Kasami exponents while the other cases are conjectured based on empirical evidence. Moreover, the S-boxes are also proven to have low maximum differential. The second class of S-boxes is constructed in Section 6.1.2 and the construction is based on applying ideas from GMW sequences [44] on power functions. The S-boxes are more abundant and have larger linear span than the first class.

### 6.1 S-box as Combining Function in Stream Ciphers

We consider the case when there is an odd number of input bits $n$ and construct two classes of $n$-by-$m$ S-boxes with nonlinearity $2^{n-1} - 2^{(n-1)/2}$. These S-boxes have high nonlinearity for protection against correlation and linear approximation attacks [53, 71]. Against the more general correlation attacks of [77], they give low bound for maximum correlation bounds by Corollary 1 of Chapter 2:

$$C_F \leq 2^{m/2 - n} \max_{v \neq 0, \lambda \neq 0} |v \cdot \hat{F}(\lambda)| = 2^{(m-n+1)/2},$$

(6.1)
because \( \max_{v \neq 0} |v \cdot \tilde{F}(\lambda)| = 2^{(n+1)/2} \). We will prove that the maximum correlation is \( \sqrt{2} \) times smaller than the upper bound of equation (6.1) for both of our S-box constructions. This implies that approximation by nonlinear function of the output bits is closer to \( \frac{1}{2} \) for protection against general correlation attack because:

\[
\frac{1}{2} - \frac{C_F}{2} \leq \Pr(g(F(x)) = w \cdot x) \leq \frac{1}{2} + \frac{C_F}{2}.
\]

### 6.1.1 S-box Construction I: m-Sequence Construction

We will first derive the nonlinearity, maximum differential and algebraic degree of our S-boxes, then we will prove the maximum correlation.

From Chapter 2, the power permutation \( F(x) = x^k \), \( \gcd(k, 2^n - 1) = 1 \), on \( GF(2^n) \) is almost bent if \( N_F = 2^{n-1} - 2^{(n-1)/2} \).

The following proposition on the cryptographic properties of almost bent power permutations is well known, e.g., see [5, Theorem 1] and [9, Theorem 4].

**Proposition 11.** Let \( n \) be odd and let \( F(x) = x^k \) be an almost bent permutation on \( GF(2^n) \). Then the Hadamard transform \( v \cdot \tilde{F}(w) \) takes on the values \( 0, \pm 2^{(n+1)/2} \) for all \( v \in GF(2^m) \). Moreover, \( F(x) \) is also an APN permutation, i.e., \( \Delta_F = 2 \).

The known almost bent permutations of the form \( x^k \) on \( GF(2^n) \) can be found in Table 2.15 of Chapter 2. It was conjectured in [5] that Table 2.15 contains all possible almost bent power permutations.

In the next theorem, we construct an S-box with odd number of input bits that has good cryptographic properties to be used as combiners in stream ciphers.

**Theorem 20.** Let \( n \) be odd and let \( m \) divide \( n \). Consider the function \( F : GF(2^n) \to GF(2^m) \) defined by

\[
F(x) = \text{Tr}_m^n(x^k)
\]

where \( x^k \) is an almost bent permutation on \( GF(2^n) \), i.e. \( k \) can be found in Table 2.15. Then

1. The nonlinearity satisfies \( N_F = 2^{n-1} - 2^{(n-1)/2} \).
2. When \( k = 2^r + 1 \) where \( \gcd(r, n) = 1 \) or \( k = 2^{2r} - 2^r + 1 \) where \( 3r \equiv 1 \mod n \), the maximum
correlation satisfies \( C_F \leq 2^{(m-n)/2} \). This is \( \sqrt{2} \) times smaller than the bound in equation
6.1.

**Proof.** Any linear combination of output bits is given by \( F_b(x) = \text{Tr}^m_1(bF(x)), b \in GF(2^m) \).
Because \( b^{2^m} = b \), we can simplify \( F_b(x) \) as
\[
\text{Tr}^n_1(b \text{Tr}^m_1(x^k)) = \text{Tr}^n_1(\text{Tr}^m_1(bx^k)) = \text{Tr}^n_1(bx^k)
\]
whose Hadamard transform takes on the values \( 0, \pm 2^{(n+1)/2} \) for all \( b \) because \( x^k \) is almost bent.
Thus the nonlinearity of \( F \) is \( 2^{n-1} - 2^{(n-1)/2} \) by applying equation (2.9).

The proof of the maximum correlation bound requires Proposition 12, Lemma 12 and Lemma
13. After presenting these results, we will present the proof that \( C_F \leq 2^{(m-n)/2} \).

The inequality in Proposition 12 is used in the proof of [77, Theorem 4].

**Proposition 12.** (Zhang and Chan [77]) The maximum correlation satisfies
\[
C_F(w)^2 \leq 2^{-2n} \sum_{v \in GF(2^n)} \hat{v} \cdot \hat{F}(w)^2.
\]

By Proposition 12, we can look at the distribution of \( \hat{v} \cdot \hat{F}(w) \) to analyse \( C_F(w) \). To do this,
we need Lemma 12 on the Hadamard Transform of Gold and Kasami functions.

**Lemma 12.** Let \( n \) be odd and \( f(x) = \text{Tr}(x^k) \) on \( GF(2^n) \).

1. (Gold [22]) Let \( k = 2^r + 1 \) where \( \gcd(r, n) = 1 \). Then \( \hat{f}(\lambda) = 0 \iff \text{Tr}(\lambda) = 0 \).

2. (Dillon [16, Theorem 7]) Let \( k = 2^{2r} - 2^r + 1 \) where \( 3r \equiv 1 \mod n \). Then \( \hat{f}(\lambda) = 0 \iff \text{Tr}(\lambda^{2^{2r} + 1}) = 0 \).

We also need the following lemma on the trace function of subfield elements.

**Lemma 13.** Let \( n \) be odd and \( GF(2^m) \) be a subfield of \( GF(2^n) \), i.e. \( m|n \). If \( \gcd(u, 2^{m} - 1) = 1 \),
then among the elements \( b \in GF(2^m) \), there are at least \( 2^{m-1} \) elements such that \( \text{Tr}^n_1(\lambda b^u) = 0 \)
for any fixed \( \lambda \in GF(2^n) \).
CHAPTER 6. HIGHLY NONLINEAR S-BOXES WITH REDUCED BOUND ON MAXIMUM CORRELATION

Proof. Note that
\[ \text{Tr}^n_1(\lambda b^u) = \text{Tr}^m_1(b^u \text{Tr}^n_m(\lambda)) \]
because \( b^u \in GF(2^m) \). Since \( \text{Tr}^n_m(\lambda) \) is fixed, \( \text{Tr}^m_1(\cdot) \) is balanced and \( b^u \) is a permutation on \( GF(2^m) \). Since \( \text{Tr}^n_m(\lambda) \) is fixed, \( \text{Tr}^m_1(\cdot) \) is balanced and \( b^u \) is a permutation on \( GF(2^m) \), half of \( \text{Tr}^m_1(b^u \text{Tr}^n_m(\lambda)) \) are 0’s when we vary \( b \) if \( \text{Tr}^n_m(\lambda) \neq 0 \). \( \text{Tr}^m_1(b^u \text{Tr}^n_m(\lambda)) \) is 0 for all \( b \in GF(2^m) \) if \( \text{Tr}^n_m(\lambda) = 0 \).

Using Proposition 12, Lemma 12 and Lemma 13, we can prove the maximum correlation of an S-box based on the Gold and certain Kasami exponents is \( \sqrt{2} \) times smaller than the bound given by Proposition 1.

Proof of Theorem 20 part 2: Let a linear combination of output bits be \( F_b(x) = \text{Tr}^n_m(bF(x)) = \text{Tr}^m_1(bx^k) \). We see that
\[
\hat{F}_b(\lambda) = \sum_{x \in GF(2^n)} (-1)^{\text{Tr}(bx^k) + \text{Tr}(\lambda x)}
\]
\[
= \sum_{y \in GF(2^n)} (-1)^{\text{Tr}(y^k) + \text{Tr}(\lambda b^{-k-1} y)}, \quad y = b^{-k-1} x
\]
\[
= \hat{f}(\lambda b^{-k-1}) \quad \text{where} \quad f(x) := \text{Tr}^n_1(x^k), \lambda \neq 0.
\]

Fix \( \lambda \neq 0 \). By Lemma 12,
\[
\hat{f}(\lambda b^{-k-1}) = 0 \iff \text{Tr}(\lambda b^{-k-1}) = 0 \quad \text{for} \quad k = 2^r + 1
\]
\[
\hat{f}(\lambda b^{-k-1}) = 0 \iff \text{Tr}(\lambda 2^{r+1} b^{-k-1}(2^{r+1}) = 0 \quad \text{for} \quad k = 2^{2r} - 2^r + 1.
\]

By Lemma 13, at least \( 2^{m-1} \) elements \( b \in GF(2^m) \) satisfy this condition. Therefore, \( \hat{F}_b(\lambda) = 0 \) for at least \( 2^{m-1} \) elements \( b \). For the other \( \leq 2^{m-1} \) elements \( b \), \( \hat{F}_b(\lambda) = \pm 2^{(n+1)/2} \) because the permutation \( x^k \) is almost bent. By Proposition 12,
\[
C_F(\lambda)^2 \leq 2^{-2n} \sum_{b \in GF(2^m)} \hat{F}_b(\lambda)^2 \leq 2^{-2n}(2^{m-1} 2^{n+1}) = 2^{m-n},
\]
for all \( \lambda \neq 0 \).

Remark 17. Besides high nonlinearity and reduced maximum correlation, the S-box constructed in Theorem 20 has other good cryptographic properties:

1. \( F \) is balanced.
\textbf{Proof.} $x^k$ is a permutation and the trace function $Tr^n_m(\cdot)$ is a balanced function. Therefore, $F$ is balanced because it is a composition of balanced functions. 

2. The maximum differential satisfies the inequality $2^{n-m} < \Delta_F \leq 2^{n-m+1}$.

\textit{Proof.} The lower bound $2^{n-m}$ for maximum differential is only achieved when $n$ is even, therefore $\Delta_F > 2^{n-m}$.

We count the number $\Delta_F(a,b)$ of $x$ such that

$$F(x) + F(x + a) = b$$

$$\implies \quad Tr^n_m(x^k + (x + a)^k) = b$$

for all $a \neq 0, b$. The equation $Tr^n_m(y) = b$ has $2^{n-m}$ solutions $y_1, \ldots, y_{2^{n-m}}$ in $GF(2^n)$. Each equation

$$x^k + (x + a)^k = y_i$$

has at most 2 solutions because $x^k$ is APN by Proposition 11. Therefore $\Delta_F(a,b) \leq 2 \times 2^{n-m} = 2^{n-m+1}$ which implies $\Delta_F \leq 2^{n-m+1}$.

3. The algebraic degree satisfies $\text{deg}(F) = \text{wt}(k)$ and this can be as high as $\frac{n+1}{2}$.

\textit{Proof.} When we write $Tr^n_m(x^k)$ as a polynomial equation, the weight of all the exponents is $\text{wt}(k)$. Therefore the algebraic degree, which is the maximum weight of the exponents, is also $\text{wt}(k)$.

Consider the Kasami functions in Table 2.15. The Kasami exponent $k = 2^{2r} - 2^r + 1$ has Hamming weight $r + 1$. When $k$ is reduced modulo $2^n - 1$, the maximum weight occurs at $r = \frac{n-1}{2}$. The corresponding algebraic degree is $r + 1 = \frac{n+1}{2}$.

\textbf{Example 24.} Let $n = 9$, $m = 3$ and $k = 3, 5, 17$ in Theorem 20. Therefore we are looking at the 9-by-3 S-boxes defined by:

$$F(x) = Tr^9_3(x^k), \quad k = 3, 5, 17.$$ 

These are balanced quadratic S-boxes having nonlinearity $2^8 - 2^4 = 240$ and maximum correlation $2^{(3-9)/2} = 0.125$. This implies the following bounds on linear approximation of $v \cdot F(x)$ and
CHAPTER 6. HIGHLY NONLINEAR S-BOXES WITH REDUCED BOUND ON MAXIMUM CORRELATION

\[ g(F(x)) : \]

\[ 0.469 \leq \Pr(v \cdot F(x) = w \cdot x) \leq 0.531 \text{ for all } w \in GF(2)^9, v \in GF(2)^3, \]

\[ 0.438 \leq \Pr(g(F(x)) = w \cdot x) \leq 0.563 \text{ for all } w \in GF(2)^9, g : GF(2)^3 \rightarrow GF(2). \]

We note that the Kasami exponent from Theorem 20 part 5 cannot be used here because we need the condition \(3r \equiv 1 \mod 9\) which is impossible.

Moreover, the S-boxes also satisfies \(\Delta_F \leq 2^{9-3+1} = 128\) which implies differential approximation satisfies:

\[ \Pr_x(F(x) + F(x + a) = b) \leq \frac{1}{4} \text{ for all } a \neq 0. \]

We performed some computer simulations and found out that \(C_F\) is exactly \(2^{m-n}\) for the functions in Theorem 20. This has been verified for the field \(GF(2^9)\) over the subfield \(GF(2^3)\) and the field \(GF(2^{15})\) over the subfields \(GF(2^5), GF(2^3)\). Thus we make the following conjecture.

**Conjecture 1.** Let \(n\) be odd and \(m\) divide \(n\). Consider the function \(F : GF(2^n) \rightarrow GF(2^m)\) defined by

\[ F(x) = Tr_m^n(x^k) \]

where \(k\) are the exponents listed in Table 2.15 of Chapter 2. Then the maximum correlation \(C_F\) is \(2^{(m-n)/2}\).

### 6.1.2 S-box Construction II: GMW Sequence Construction

In Theorem 19 of Chapter 5, we constructed \(N\)-by-\(m\) S-boxes with nonlinearity \(2^{N-1} - 2^{(N-1)/2}\) through cascaded GMW functions. By Corollary 1 of Chapter 2, they have low maximum correlation \(2^{(m-N+1)/2}\). In this section, we prove that when there is only one level of cascade in the GMW construction, the maximum correlation can be reduced by \(\sqrt{2}\).

**Theorem 21.** Let \(n\) be odd and \(F : GF(2^n) \rightarrow GF(2^m)\) be defined by

\[ F(x) = Tr_m^n(x^{2^m+1})^k, \]

where \(\gcd(i, n/m) = 1, \gcd(k, 2^m - 1) = 1\). Let \(x^k\) be almost bent on \(GF(2^m)\), i.e. \(k\) can be found in Table 2.15 of Chapter 2 where we consider \(m\) instead of \(n\). Then
CHAPTER 6. HIGHLY NONLINEAR S-BOXES WITH REDUCED BOUND ON MAXIMUM CORRELATION

1. The nonlinearity satisfies \( N_F = 2^{n-1} - 2^{(n-1)/2} \).

2. The maximum correlation satisfies \( C_F \leq 2^{(m-n)/2} \), which is \( \sqrt{2} \) times smaller than the bound in equation 6.1.

Proof. The nonlinearity is a direct consequence of Theorem 19 of Chapter 5.

The proof of the maximum correlation bound requires Lemma 14, 15 and 16. After presenting these results, we will present the proof that \( C_F \leq 2^{(m-n)/2} \).

Remark 18. We note that both Theorem 20 and 21 use the functions from Table 2.15 of Chapter 2. But while reduced maximum correlation can only be proven for the Gold and certain Kasami exponents in Theorem 20, it can be shown to hold for all the exponents of Table 2.15 in Theorem 21.

Lemma 14. Let \( n \) be odd and let \( m|n \). If \( g(x) = \text{Tr}_n^1(x^{2^{mi}+1}) \) where \( \gcd(i,n/m) = 1 \), then

\[
\hat{g}(\lambda)^2 = 2^n \sum_{z \in GF(2^m)} (-1)^g(z) + \text{Tr}_n^1(\lambda z).
\]

Proof.

\[
\hat{g}(\lambda)^2 = \sum_{x \in GF(2^n)} (-1)^g(x) + \text{Tr}_n^1(\lambda x) \sum_{y \in GF(2^n)} (-1)^g(y) + \text{Tr}_n^1(\lambda y)
\]

\[
= \sum_{z \in GF(2^n)} (-1)^g(z) + \text{Tr}_n^1(\lambda z) \sum_{x \in GF(2^n)} (-1)^g(x) + g(z) + g(x+z).
\]

where \( y = x + z \). We can simplify:

\[
g(x) + g(z) + g(x+z) = \text{Tr}_n^1(xz^{2^{mi}} + x^{2^{mi}}z)
\]

\[
= \text{Tr}_n^1(xL(z)),
\]

where \( L(z) = z^{2^{mi}} + z^{2^{n-mi}} \) is a linear function on \( GF(2^n) \).

Since \( \sum_z (-1)^{\text{Tr}_n^1(xL(z))} = 2^n \) if \( z \in \ker(L) \) and 0 otherwise, we see that

\[
\hat{g}(\lambda)^2 = 2^n \sum_{z \in \ker(L)} (-1)^g(z) + \text{Tr}_n^1(\lambda z).
\]

To complete the proof, we prove \( \ker(L) = GF(2^m) \) as follows. The elements of \( z \in \ker(L) \) are those \( z \in GF(2^n) \) that satisfy \( z^{2^{mi}} = z^{2^{n-mi}} \). This is the same as saying \( z^{2^{mi}-1} = 1 \) when \( z \neq 0 \).
But all non-zero \( z \in GF(2^n) \) satisfy \( z^{2^n-1} = 1 \), therefore \( z \in ker(L) \) if and only if
\[
z^{\gcd(2^n-1,2^{2mi}-1)} = 1 \iff z^{\gcd(n,2^{2mi})-1} = 1 \iff z^{2^m-1} = 1
\]
when \( z \neq 0 \), because \( \gcd(n,2mi) = m \). This means \( z \in GF(2^m) \).

**Lemma 15.** Let \( n \) be odd and let \( G(x) = Tr^n_m(x^{2^{mi}+1}) \) where \( \gcd(i,n/m) = 1 \). Then \( C_G \leq 2^{(m-n)/2} \).

**Proof.** Let \( G_u(x) = Tr^m(uG(x)), u \in GF(2^m) \), be a linear combination of output bits and let \( g(x) = Tr^n_m(x^{2^{mi}+1}) \). By Proposition 12, we have
\[
C_G(\lambda)^2 \leq 2^{-2n} \sum_{u \in GF(2^m)} \hat{G}_u(\lambda)^2
= 2^{-2n} \sum_{v \in GF(2^m)} \hat{g}(v\lambda)^2 \text{ where } v = u^{-(2^{mi}+1)-1}
= 2^{-2n} \sum_{v \in GF(2^m)} 2^n \sum_{z \in GF(2^m)} (-1)^{g(z)+Tr^n_m(v\lambda z)}, \text{ by Lemma 14}
= 2^{-n} \sum_{z \in GF(2^m)} (-1)^{g(z)} \sum_{v \in GF(2^m)} (-1)^{Tr^n_m(vzTr^n_m(\lambda))} \text{ because } v, z \in GF(2^m)
= 2^{m-n} \sum_{z \in GF(2^m), z Tr^n_m(\lambda) = 0} (-1)^{g(z)}.
\]
Suppose \( Tr^n_m(\lambda) \neq 0 \). Then \( z \) has to be 0 and the above inequality is
\[
C_G(\lambda)^2 \leq 2^{m-n}(-1)^{g(0)} = 2^{m-n}.
\]
Suppose \( Tr^n_m(\lambda) = 0 \). Then \( z \) runs through all elements in \( GF(2^m) \) and the above inequality is
\[
C_G(\lambda)^2 \leq 2^{m-n} \sum_{z \in GF(2^m)} (-1)^{g(z)} = 0.
\]
The last equality is true because for \( z \in GF(2^m) \), \( g(z) = Tr^n_1(z^{2^{mi}+1}) = Tr^n_1(z) \) since \( z^{2^m} = z \implies z^{2^{mi}+1} = z^2 \). Thus the above sum is 0 because the trace function is balanced on the subfield \( GF(2^m) \).

We will need Lemma 16 which gives an efficient way to compute the maximum correlation. It is similar to Theorem 1 of [77].

**Lemma 16.** For all \( \lambda \), \( C_F(\lambda) = 2^{-n} \sum_{z \in GF(2^m)} \left| \sum_{x \in F(x) = z} (-1)^{Tr(\lambda x)} \right| \).
CHAPTER 6. HIGHLY NONLINEAR S-BOXES WITH REDUCED BOUND ON MAXIMUM CORRELATION

Proof.

$$2^n C_F(\lambda) = \max_g \sum_{x \in GF(2^n)} (-1)^{g(F(x)) + Tr(\lambda x)}$$

$$= \max_g \sum_{z \in GF(2^m)} \sum_{\{x \in GF(2^n) : F(x) = z\}} (-1)^{g(z) + Tr(\lambda x)}$$

$$= \sum_{z \in GF(2^m)} \max_{g(z) = 0,1} \sum_{\{x \in GF(2^n) : F(x) = z\}} (-1)^{Tr(\lambda x)}$$

$$= \sum_{z \in GF(2^m)} \left| \sum_{\{x \in GF(2^n) : F(x) = z\}} (-1)^{Tr(\lambda x)} \right| .$$

\[ \square \]

Remark 19. If we were to compute \( C_F(w) \) from the definition, then we would have:

$$C_F(w) = \max_g [\text{Prob}(g(F(x)) = w \cdot x) - \text{Prob}(g(F(x)) \neq w \cdot x)]$$

$$\implies C_F(w) = 2^{-n} \max_g \sum_{x \in GF(2^n)} (-1)^{g(F(x)) + w \cdot x}.$$

We need to compute a sum of size \( 2^n \) for \( 2^{2m} \) \( m \)-bit Boolean functions \( g(x) \). However, by Lemma 16:

$$C_F(w) = 2^{-n} \sum_{z \in GF(2^m)} \left| \sum_{\{x \in GF(2^n) : F(x) = z\}} (-1)^{w \cdot x} \right|,$$

so we need just compute a sum of size \( 2^n \). Thus we have a \( 2^{2m} \) reduction in the time complexity when computing \( C_F \). This is how we confirmed Conjecture 1 for the field \( GF(2^9) \) over the subfield \( GF(2^3) \) and the field \( GF(2^{15}) \) over the subfields \( GF(2^5), GF(2^3) \).

Proof of Theorem 21 part 2: For \( \lambda \neq 0 \), we have by Lemma 16

$$C_F(\lambda) = \sum_{z \in GF(2^m)} \left| \sum_{F(x) = z} (-1)^{Tr(\lambda x)} \right|$$

$$= \sum_{z \in GF(2^m)} \left| \sum_{G(x)^k = z} (-1)^{Tr(\lambda x)} \right| \text{ where } G(x) = Tr_m(x^{2^m+1})$$

$$= \sum_{z' \in GF(2^m)} \left| \sum_{G(x) = z'} (-1)^{Tr(\lambda x)} \right| \text{ where } z' = z^{k^{-1}}$$

$$= C_G(\lambda).$$
Here, $z'$ will run through all elements of $GF(2^m)$ because $z^{k-1}$ is a permutation on $GF(2^m)$. By Lemma 15, $C_G(\lambda) \leq 2^{(m-n)/2}$.

Remark 20. Besides high nonlinearity and reduced maximum correlation, the S-box constructed in Theorem 21 has the following cryptographic properties which can be deduced from Theorem 19 of Chapter 5:

1. $F$ is balanced.
2. The algebraic degree satisfies $\deg(F) = 2^{\wt(k)}$.
3. The linear span is $(n/m)^{wt(k)}$.

Example 25. We apply Theorem 21 with $n = 9$, $m = 3$, $i = 1, 2$ and $k = 3$ (corresponding to the Gold exponent from Table 2.15 from Chapter 2). This corresponds to the 9-by-3 S-boxes:

$$F(x) = \text{Tr}_3^9(x^9)^3 \quad \text{and} \quad F(x) = \text{Tr}_3^9(x^{65})^3.$$ 

Both are balanced S-boxes having algebraic degree 4, nonlinearity $2^8 - 2^4 = 240$ and maximum correlation $2^{(3-9)/2} = 0.125$. This implies the following probability bounds on linear approximation of $v \cdot F(x)$ and $g(F(x))$:

$$0.469 \leq \Pr(v \cdot F(x) = w \cdot x) \leq 0.531 \text{ for all } w \in GF(2)^9, v \in GF(2)^3,$$

$$0.438 \leq \Pr(g(F(x)) = w \cdot x) \leq 0.563 \text{ for all } w \in GF(2)^9, g : GF(2)^3 \rightarrow GF(2).$$

The S-boxes from Theorem 20 and Theorem 21 have identical probability bounds against linear approximations. But for the case $n = 9$, Theorem 21 gives higher algebraic degree. Moreover, the linear span of the S-boxes in this example is $(\frac{9}{3})^2 = 9$ while that from Example 24 is 3.

6.2 Summary of Chapter 6

We considered S-boxes when used as combiners in stream ciphers. We constructed two classes of S-boxes with high nonlinearity for protection against linear correlation attacks [53, 71] and reduced the bound for maximum correlation by $\sqrt{2}$ for better protection against nonlinear correlation attacks [77]. Our S-boxes also possess other good properties like low maximum differential and
high algebraic degree. The proofs for the maximum correlation of our S-boxes linked several well-established results in sequence analysis, thus providing another example of the interconnection between the theory of sequences and cryptography.
Chapter 7

Computational Results for Boolean Functions based on $GF(2^n)$

In this chapter, we present some computational results for Boolean functions constructed from polynomials in $GF(2^n)$. First, we review some efficient methods to compute the Hadamard transform of a polynomial in $GF(2^n)$. These methods\(^1\) have useful applications in speeding up search algorithms for cryptographic Boolean functions based on polynomials in $GF(2^n)$.

Two applications of these fast transform methods are studied. First, we look at an efficient search algorithm for resilient and highly nonlinear Boolean functions from preferred functions in $GF(2^n)$. An application of this algorithm in the field $GF(2^{11})$ is presented. Second, we verify the search for the Patterson-Wiedemann (PW) functions performed by Gangopadhyay et. al. in [21], and compute the trace representations of these PW functions. From the trace representation, we discuss some cryptographic properties of these functions and their implications.

A binary sequence $\{a_i\}_{i=0}^{2^n-2}$ is called coset-constant if $a_i = a_{2i}$ for all $i$. It is an interleaved

---

\(^1\)For Boolean functions $f : GF(2^n) \rightarrow GF(2)$, we can speed up Hadamard transform computation with the fast fourier transform (FFT). However, they cannot be applied to Hadamard transform computations for polynomials in $GF(2^n)$.
sequence if for some \( p | (2^n - 1) \), \( a_i = a_{i+p} \) for all \( i \). Usually, we take \( p = 2^m - 1 \), \( m | n \), so that we can analyze the structure of the corresponding m-sequences for cross-correlation computation, e.g., see [26, 75, 76].

The following proposition shows that we can speed up the Hadamard transform of polynomials based on the coset-constant and interleaved sequences.

**Proposition 13.** Let \( f : GF(2^n) \rightarrow GF(2) \) and \( a_i = f(\alpha^i) \) where \( \alpha \) generates \( GF(2^n) \).

1. If \( f(x) = f(x^2) \) for all \( x \in GF(2^n) \), i.e. \( a_i = a_{2i} \), \( i = 0, \ldots, 2^n - 2 \), then \( \hat{f}(\lambda) = \hat{f}(\lambda^2) \).
2. Let \( p | (2^n - 1) \). If \( f(x) = f(\alpha^px) \) for all \( x \in GF(2^n) \), i.e. \( a_i = a_{i+p} \), \( i = 0, \ldots, 2^n - 2 \), then \( \hat{f}(\lambda) = \hat{f}(\alpha^p\lambda) \).

Proposition 13 is well-known in sequence analysis in the context of cross-correlation with m-sequences. It can be proven using elementary properties of finite fields.

### 7.1 Efficient Search for Highly Nonlinear Resilient Boolean Functions

In this section, we present an efficient search algorithm for highly nonlinear resilient Boolean functions from preferred functions in \( GF(2^n) \). Our algorithm is based on Corollary 3 of Chapter 5, which states that a balanced preferred function is 1-resilient if its additive autocorrelation does not take the extreme values 0, \( \pm 2^n \). First, we search for preferred functions among functions of the form:

\[
f(x) = \sum_{j=1}^{s} c_j Tr(x^{t_j}), \quad x \in GF(2^n),
\]

where \( t_1, \ldots, t_s \) are all cyclotomic coset leaders modulo \( 2^n - 1 \) and \( c_j \in \{0, 1\} \).

Because \( f(x) = f(x^2) \) in equation (7.1), Proposition 13 implies that \( \hat{f}(\lambda) = \hat{f}(\lambda^2) \). Therefore to test if \( f(x) \) is preferred, we just need to test whether \( \hat{f}(\alpha^i) = 0, \pm 2^{(n+1)/2} \) for all exponents \( i \) which are cyclotomic coset leaders modulo \( 2^n - 1 \).

Moreover, Proposition 6 of Chapter 4 states that \( deg(f) \leq (n + 1)/2 \) if \( f(x) \) is preferred. Therefore we can restrict the exponents of \( f(x) \) to \( wt(t_j) \leq (n+1)/2 \) when we search for preferred functions because \( deg(f) = \max_j wt(t_j) \).
Algorithm 2 is presented to find resilient preferred functions.

Algorithm 2. Repeat for all vectors \((c_1, \ldots, c_s) \in GF(2)^s\):

1. Let \(f(x) = \sum_{j=1}^{s} c_j \text{Tr}(x^{t_j}), x \in GF(2^n), \) where \(t_1, \ldots, t_s\) are all the cyclotomic coset leaders modulo \(2^n - 1\) satisfying \(\text{wt}(t_j) \leq (n + 1)/2.\)

2. If \(f(x)\) is quadratic or unbalanced, go to step 1.

3. For each cyclotomic coset leader \(i:\)

   Compute \(\hat{f}(\alpha^i).\) Go to step 1 once \(\hat{f}(\alpha^i) \neq 0, \pm 2^{(n+1)/2}.\)

4. Test if \(\Delta_f(\alpha^i) \neq 0, \pm 2^n\) for \(i = 0, 1, 2, \ldots\) Stop once such an \(i\) is found and output a 1-resilient Boolean form of \(f(x)\) with nonlinearity \(2^{n-1} - 2^{(n-1)/2}\) using Algorithm 1 of Chapter 4. Go to step 1.

A balanced quadratic preferred function is equivalent to a concatenation of quadratic bent functions [54]. Therefore it contains a linear structure and cannot be resilient by Lemma 11 of Chapter 5. Thus we exclude quadratic functions in step 2 of Algorithm 2.

The efficiency of Algorithm 2 is based on the following observations:

1. There are approximately \(2^n/n\) cyclotomic coset leaders modulo \(2^n - 1.\) In step 1, we are only looking at coset leaders \(t_j\) with Hamming weight \(\leq (n + 1)/2,\) where there are approximately \(2^{n-1}/n\) of them. So the search space is reduced from the order of \(2^{2n}/n\) to about \(2^{2n-1}/n.\)

2. In step 3, we are only testing \(\hat{f}(\lambda)\) for \(\lambda = \alpha^i, i\) a cyclotomic coset leader, instead of all \(\lambda \in GF(2^n).\) The number of Hadamard transforms to be tested is reduced from \(2^n\) to approximately \(2^n/n,\) a decrease by a factor of \(n.\)

3. In step 4, we observe for \(n = 5, 7, 9, 11\) that if \(f\) is resilient, the condition \(\Delta_f(\alpha^i) \neq 0, \pm 2^n\) is usually satisfied in the beginning when \(i = 0.\) So we probably need just one verification out of \(2^n\) possible iterations.

7.1.1 Example: Search for Resilient Preferred Functions in \(GF(2^{11})\)

As an example, we apply Algorithm 2 to find resilient preferred functions in \(GF(2^{11}).\) There are 187 cyclotomic coset leaders modulo \(2^{11} - 1 = 2047\) of which 136 have weight \(\leq 6.\) The
CHAPTER 7. COMPUTATIONAL RESULTS FOR BOOLEAN FUNCTIONS BASED ON $GF(2^N)$

Table 7.1: Weight distribution of cyclotomic cosets mod 2047 with weight $\leq 6$

<table>
<thead>
<tr>
<th>weight</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td># of coset leaders</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>15</td>
<td>30</td>
<td>42</td>
<td>42</td>
</tr>
</tbody>
</table>

weight distribution is given in Table 7.1. The search space for $f(x) = \sum_{j=1}^{136} c_j Tr(x^{t_j})$, $wt(t_j) \leq 6$ and $c_j \in \{0,1\}$, grows exponentially as the number of non-zero terms increases. Because this is intended to be a small example to demonstrate Algorithm 2, we consider the following restriction on the search space.

We restrict $f$ to a sum of at most seven trace terms. When $f$ is a sum of up to four trace terms, we will do an exhaustive search through all $t_j$. When $f$ is a sum of five trace terms, we restrict $weight(t_j) \leq 5$. When $f$ is a sum of six trace terms, we restrict $weight(t_i) \leq 4$. When $f$ is a sum of seven trace terms, we restrict $weight(t_i) \leq 3$.

Modification: We exclude the values $t_j = 0$ and $t_j = 1$. This is because $Tr(1)$ is constant and the Hadamard transform of $Tr(x) + f(x)$ takes on the same values as that of $f(x)$.

It can be observed from experimental data that if we allow unbalanced preferred functions to be tested in Algorithm 2, then they are always a sum of an even number of trace terms. Adding the function $Tr(x)$ to it makes it balanced. For example,

$$f(x) = Tr(x^3 + x^{13} + x^{69} + x^{83})$$

is an unbalanced preferred function while

$$f(x) = Tr(x + x^3 + x^{13} + x^{69} + x^{83})$$

is a balanced preferred function. Therefore we allow unbalanced functions to be tested in Algorithm 2 and we add the function $Tr(x)$ to it when necessary.

Based on these restrictions, the search space consists of

$$\sum_{i=1}^{4} \binom{134}{i} + \binom{92}{5} + \binom{50}{6} + \binom{20}{7} = 78387228 \approx 2^{26}$$

functions.
CHAPTER 7. COMPUTATIONAL RESULTS FOR BOOLEAN FUNCTIONS BASED ON $GF(2^N)$

Table 7.2: Algebraic degree = 4 (Exhaustive search for up to 6 non-trivial trace terms in $GF(2^{11})$)

<table>
<thead>
<tr>
<th>Exponents of Preferred Functions</th>
<th>Additive Autocorrelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 3, 13, 69, 83</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 25, 45, 85, 165</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>3, 5, 17, 73, 141</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>5, 25, 45, 85, 165</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 3, 7, 17, 19, 53, 67</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 5, 7, 25, 37, 43, 137</td>
<td>$\Delta_f(1) = -288$</td>
</tr>
<tr>
<td>1, 5, 9, 11, 19, 37, 57</td>
<td>$\Delta_f(1) = -288$</td>
</tr>
<tr>
<td>1, 5, 27, 45, 49, 81, 293</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 11, 13, 25, 45, 85, 165</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 11, 49, 73, 99, 163, 165</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
</tbody>
</table>

7.1.2 Search Result

We obtain 661 balanced non-quadratic preferred functions out of the 78387228 functions searched. It can be verified that $\Delta_f(1) \neq 0, \pm 2^{11}$ for all these functions. Therefore, they all have high nonlinearity $2^{10} - 2^5 = 992$ and a Boolean representation which is 1-resilient in some basis of $GF(2^{11})$. This demonstrates two things:

1. A randomly chosen balanced non-quadratic preferred function is probably 1-resilient in some basis representation\(^2\).

2. Step 4 of Algorithm 2 is fast, needing just one verification out of $2^{11}$ possible iterations.

Of the 661 preferred functions, 638 are cubic. There are too many cubic functions to be listed conveniently and their low degree is not as cryptographically interesting. Therefore, we just list the 23 non-cubic preferred functions in Tables 7.2, 7.3 and 7.4.

\(^2\)There exist many examples of balanced non-quadratic preferred functions which are not resilient, based on concatenation of bent functions.
CHAPTER 7. COMPUTATIONAL RESULTS FOR BOOLEAN FUNCTIONS BASED ON $GF(2^N)$

Table 7.3: Algebraic degree = 5 (Exhaustive search for up to 5 non-trivial trace terms in $GF(2^{11})$)

<table>
<thead>
<tr>
<th>Exponents of Preferred Functions</th>
<th>Additive Autocorrelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 31, 301</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 35, 79</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>31, 113, 301</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 5, 9, 17, 143</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 17, 121, 137, 143</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>11, 17, 31, 33, 301</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
</tbody>
</table>

Table 7.4: Algebraic degree = 6 (Exhaustive search for up to 4 non-trivial trace terms in $GF(2^{11})$)

<table>
<thead>
<tr>
<th>Exponents of Preferred Functions</th>
<th>Additive Autocorrelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>71, 309, 359</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 3, 43, 171, 683</td>
<td>$\Delta_f(1) = -24$</td>
</tr>
<tr>
<td>1, 13, 411, 413, 423</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 25, 71, 309, 359</td>
<td>$\Delta_f(1) = 64$</td>
</tr>
<tr>
<td>1, 57, 231, 343, 683</td>
<td>$\Delta_f(1) = -24$</td>
</tr>
<tr>
<td>1, 63, 95, 189, 315</td>
<td>$\Delta_f(1) = -200$</td>
</tr>
<tr>
<td>1, 143, 151, 365, 429</td>
<td>$\Delta_f(1) = 240$</td>
</tr>
</tbody>
</table>
CHAPTER 7. COMPUTATIONAL RESULTS FOR BOOLEAN FUNCTIONS BASED ON $GF(2^N)$

7.2 The Patterson-Wiedemann Functions

Patterson and Wiedemann introduced a novel search technique for Boolean functions with high nonlinearity, exceeding the quadratic bound $2^{n-1} - 2^{(n-1)/2}$ for $n$ odd, in [62, 63]. Their construction was based on the group action of $GF(2^3)^*$ and $GF(2^5)^*$ on the projective plane $PG(2, 2^5)$. This is equivalent to the interleaved sequence construction which has also been used to construct highly nonlinear Boolean functions by Dillon, Gong and Youssef, and Gangolyay and Maitra in [15, 75, 76, 20].

From [62, 63], we can describe the search space of the Patterson-Wiedemann (PW) functions as follows:

$$f(x) = \chi_A(x), \ x \in GF(2^{15}), \quad (7.2)$$

where

\[ A = \{ \alpha^{2^j(i+151k)} | i \in I, 0 \leq j < 15, 0 \leq j < 7 \times 31 \}, \]

\[ I = \{0\} \cup S \text{ and} \]

\[ S \subset \{1, 3, 5, 7, 11, 15, 17, 23, 35, 37\} \text{ and } |S| = 5. \]

Thus the weight of $f(x)$ is 16492. From equation (7.2), we deduce that:

$$f(x) = f(x^2) \text{ and } f(x) = f(\alpha^{151}x).$$

By Proposition 13, we have the relation:

$$\hat{f}(\lambda) = \hat{f}(\lambda^2) \text{ and } \hat{f}(\lambda) = \hat{f}(\alpha^{151}\lambda).$$

Therefore instead of computing $\hat{f}(\lambda)$ for all $\lambda \in GF(2^{15})$ to determine the nonlinearity of $f(x)$, we just need to compute $\hat{f}(\alpha^i)$ where $i$ is one of eleven cyclotomic coset leaders modulo 151:

$$C_{151} = \{0, 1, 3, 5, 7, 11, 15, 17, 23, 35, 37\}.$$

Because $2^{15} - 1 = 151 \times 217$, we see that among the Hadamard transforms $\hat{f}(\lambda)$ for $\lambda \in GF(2^{15})$: $\hat{f}(0)$ corresponds to the weight of $f(x)$, there are 217 repetitions of $\hat{f}(\alpha^0) = \hat{f}(1)$ and $217 \times 15 = 3255$ repetitions of $\hat{f}(\alpha^i)$ for the ten non-zero cyclotomic cosets $i$.

By repeating the search of Patterson-Wiedemann among the functions described by equation (7.2), we obtain the two PW functions originally found in [62, 63] with nonlinearity 16276 >
CHAPTER 7. COMPUTATIONAL RESULTS FOR BOOLEAN FUNCTIONS BASED ON GF(2^N)

Table 7.5: Weight distribution and Hadamard transform of PW1, PW2

<table>
<thead>
<tr>
<th>wt(f(x) + Tr(λx))</th>
<th>f(λ)</th>
<th>Frequency</th>
<th>Coset leaders λ</th>
</tr>
</thead>
<tbody>
<tr>
<td>16492</td>
<td>−216</td>
<td>13021</td>
<td>4L, 1Z</td>
</tr>
<tr>
<td>16428</td>
<td>−88</td>
<td>217</td>
<td>1S</td>
</tr>
<tr>
<td>16364</td>
<td>40</td>
<td>3255</td>
<td>1L</td>
</tr>
<tr>
<td>16300</td>
<td>168</td>
<td>16275</td>
<td>5L</td>
</tr>
</tbody>
</table>

Nonlinearity of PW1, PW2 is $2^{14} - \frac{1}{2} \times 216 = 16276$ by equation (2.9) of Chapter 2.

$2^{14} - 2^7 = 16256$. Moreover, there are two more PW functions with nonlinearity $16268 > 16256$. They are described by equation (7.2) with the set $I$ as follows:

- $PW1: \quad I = \{0, 1, 3, 7, 17, 35\}$ and $N_f = 16276$,
- $PW2: \quad I = \{0, 1, 7, 11, 15, 17\}$ and $N_f = 16276$,
- $PW3: \quad I = \{0, 3, 5, 23, 35, 37\}$ and $N_f = 16268$,
- $PW4: \quad I = \{0, 5, 11, 15, 23, 37\}$ and $N_f = 16268$.

These four PW functions were also found by Gangopadhyay et. al. in [21]. They provided the Hadamard transform and weight distribution of PW1 and PW2 according to the eleven cyclotomic coset leaders modulo 151, which we list in Table 7.5. Here, we also present the weight distribution of PW3 and PW4 in Table 7.6. Following the notation of [21], the zero coset with 217 elements is denoted by $S$, the ten non-zero cosets with 3255 elements are denoted by $L$ and the zero element of $GF(2^{15})$ is denoted by $Z$.

It is interesting to know the trace representation of these functions, as it reveals the linear span and algebraic degree of the functions. As noted in Chapter 2, a high algebraic degree and large linear span provides algebraic complexity and protection against interpolation attacks. Because the PW functions have period 151, their trace representation is of the form:

$$f(x) = \sum_{i \in C_{151}} c_i Tr(x^{217i}), c_i \in \{0, 1\},$$

where $C_{151}$ are the cyclotomic coset leaders modulo 151 and the exponents are multiples of
Table 7.6: Weight distribution and Hadamard transform of \( PW_3, PW_4 \)

<table>
<thead>
<tr>
<th>( wt(f(x) + Tr(\lambda x)) )</th>
<th>( \hat{f}(\lambda) )</th>
<th>Frequency</th>
<th>Coset leaders ( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>16492</td>
<td>-216</td>
<td>3256</td>
<td>1L, 1Z</td>
</tr>
<tr>
<td>16460</td>
<td>-152</td>
<td>13020</td>
<td>4L</td>
</tr>
<tr>
<td>16428</td>
<td>-88</td>
<td>217</td>
<td>1S</td>
</tr>
<tr>
<td>16396</td>
<td>-24</td>
<td>3255</td>
<td>1L</td>
</tr>
<tr>
<td>16300</td>
<td>168</td>
<td>3255</td>
<td>1L</td>
</tr>
<tr>
<td>16268</td>
<td>232</td>
<td>9765</td>
<td>3L</td>
</tr>
</tbody>
</table>

Nonlinearity of \( PW_3, PW_4 \) is \( 2^{14} - \frac{1}{2} \times 232 = 16268 \) by equation (2.9) of Chapter 2.

217 = \( \frac{2^{15} - 1}{15} \). By equation (2.1) of Chapter 2, the trace representation of the PW functions are:

\[
\begin{align*}
PW_1(x) &= Tr(x^{217} + x^{651} + x^{1085} + x^{2387} + x^{3255}), \\
PW_2(x) &= Tr(x^{651} + x^{1519} + x^{7595}), \\
PW_3(x) &= Tr(x^{217} + x^{1085} + x^{2387} + x^{3255} + x^{3689} + x^{4991} + x^{8029}), \\
PW_4(x) &= Tr(x^{1519} + x^{3689} + x^{4991} + x^{7595} + x^{8029}).
\end{align*}
\]

Using the fact that the algebraic degree is the maximum weight of the exponent and that each trace term contributes a linear span of 15, we summarize the cryptographic properties of the PW functions in Table 7.7.

Thus we see that, although \( PW_3 \) and \( PW_4 \) have lower nonlinearity than \( PW_1 \) and \( PW_2 \), they generally have higher algebraic degree and larger linear span. Moreover, they can be used in the same way as \( PW_1 \) and \( PW_2 \) are used in [40, 66, 68] to construct balanced/resilient Boolean functions with high nonlinearity. This allows for a larger class of cryptographic functions to be constructed.

Another useful measure of the effectiveness of a Boolean function, introduced in [75], is its *generalized nonlinearity*, denoted by \( NLG \) and defined as:

\[
NLG(f) = \min_{\gcd(c, 2^n - 1) = 1, \lambda \in GF(2^n)} dist(f(x), Tr(\lambda x^c)).
\]

A high \( NLG \) ensures that attacks based on approximation by bijective monomials are difficult. A
related concept is the extended Hadamard transform, denoted by EHT and defined as:

\[
\hat{f}(\lambda, c) = \sum_{x \in \mathbb{GF}(2^n)} (-1)^{f(x) + \text{Tr}(\lambda x^c)}, \gcd(c, 2^n - 1) = 1.
\]

It can be deduced that:

\[
\text{NLG}(f) = 2^{n-1} - \frac{1}{2} \max_{\gcd(c, 2^n - 1) = 1, \lambda \in \mathbb{GF}(2^n)} |\hat{f}(\lambda, c)|
\]

(7.3)

Youssef and Gong constructed functions with large distance from all bijective monomials via interleaved sequences in [75, 76]. Since the Patterson-Wiedemann functions are constructed from interleaved sequences, it might be possible that they have high NLG.

The structure of the PW functions allows for efficient computation of their NLG by Proposition 14. This result can be viewed as a generalization of Proposition 13.

**Proposition 14.** Let \( f : \mathbb{GF}(2^n) \rightarrow \mathbb{GF}(2) \) and denote \( a_i = f(\alpha^i) \), where \( \alpha \) generates \( \mathbb{GF}(2^n) \).

1. If \( f(x) = f(x^2) \) for all \( x \in \mathbb{GF}(2^n) \), i.e., \( a_i = a_{2i}, i = 0, \ldots, 2^n - 2 \), then \( \hat{f}(\lambda, c) = \hat{f}(\lambda^2, c) \) and \( \hat{f}(\lambda, c) = \hat{f}(\lambda, 2c) \).

2. Let \( p | (2^n - 1) \). If \( f(x) = f(\alpha^p x) \) for all \( x \in \mathbb{GF}(2^n) \), i.e., \( a_i = a_{i+p}, i = 0, \ldots, 2^n - 2 \), then \( \hat{f}(\lambda, c) = \hat{f}(\alpha^p \lambda, c) \) and \( \hat{f}(\lambda, c) = \hat{f}(\lambda, c + p) \).

Proposition 14 is well-known in sequence analysis in the context of cross-correlation with m-sequences. It can be proven using elementary properties of finite fields, e.g., see [21, Theorem 2].

From Proposition 14, we just need to compute \( \hat{f}(\alpha^i, c) \) for cyclotomic coset leaders \( i \) and \( c \) modulo 151, \( c \neq 0 \). That will yield \( 11 \times 10 = 110 \) Hadamard transform computations instead of
Table 7.8: Extended Hadamard Transform Values and Related Weights of the Functions
PW1, PW2, PW3, PW4

<table>
<thead>
<tr>
<th>$\text{wt}(f(x) + Tr(\lambda x^c))$</th>
<th>$\hat{f}(\lambda, c)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>16908</td>
<td>-1048</td>
</tr>
<tr>
<td>16588</td>
<td>-408</td>
</tr>
<tr>
<td>16556</td>
<td>-344</td>
</tr>
<tr>
<td>16524</td>
<td>-280</td>
</tr>
<tr>
<td>16492</td>
<td>-216</td>
</tr>
<tr>
<td>16460</td>
<td>-152</td>
</tr>
<tr>
<td>16428</td>
<td>-88</td>
</tr>
<tr>
<td>16396</td>
<td>-24</td>
</tr>
<tr>
<td>16364</td>
<td>40</td>
</tr>
<tr>
<td>16332</td>
<td>104</td>
</tr>
<tr>
<td>16300</td>
<td>168</td>
</tr>
<tr>
<td>16268</td>
<td>232</td>
</tr>
<tr>
<td>16236</td>
<td>296</td>
</tr>
<tr>
<td>16204</td>
<td>360</td>
</tr>
<tr>
<td>16172</td>
<td>424</td>
</tr>
<tr>
<td>15984</td>
<td>800</td>
</tr>
</tbody>
</table>

NLG of PW1, PW2, PW3, PW4 is $2^{14} - \frac{1}{2} \times 1048 = 15860$ by equation (7.3).

The original $2^{15} \times 1800 = 81920000$ Hadamard transform computations ($2^{15}$ for $\lambda \in GF(2^{15})$ and 1800 for cyclotomic coset leaders $c$ modulo $2^{15} - 1$, $\text{gcd}(c, 2^{15} - 1) = 1$). By applying it to the four PW functions, we see that they all have NLG = 15860 which is higher than the NLG of 15856 for the construction presented in [76]. This fact has also been verified by Gangopadyay et. al. in [21].

Here, we note that, besides having the same generalized nonlinearity, the extended Hadamard transform of the four PW functions also takes on the same values, though with different frequencies. These values are presented in Table 7.8.
Chapter 8

Conclusion

The Gold sequences, Gold-like sequences, m-sequences, GMW sequences, ideal 2-level (multiplicative) autocorrelation sequences and interleaved sequences are well-known sequences which are widely used in CDMA communications. In this thesis, we have studied how they can be applied to construct Boolean functions and S-boxes having good cryptographic properties like balance, resiliency, high nonlinearity, low additive autocorrelation, large linear span, high algebraic degree and low maximum correlation for protection against various statistical attacks. These applications are summarized in Table 8.1.
### Table 8.1: Cryptographic Application of Sequences

<table>
<thead>
<tr>
<th>Cryptographic Functions</th>
<th>Sequences</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal Quadratic Polynomials</td>
<td>Gold and Gold-like Sequences</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Highly Nonlinear Resilient Functions with Low Additive Autocorrelation</td>
<td>Ideal 2-level (Multiplicative) Autocorrelation Sequences</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>Highly Nonlinear Resilient Functions. Balanced Functions, Nonlinearity exceeds Quadratic Bound</td>
<td>Cascaded GMW Sequences</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>Highly Nonlinear S-boxes with Reduced Maximum Correlation</td>
<td>m-Sequences and GMW Sequences</td>
<td>Chapter 6</td>
</tr>
<tr>
<td>Patterson-Wiedemann Functions</td>
<td>Interleaved Coset-constant Sequences</td>
<td>Chapter 7</td>
</tr>
</tbody>
</table>
Chapter 9

Future Work

Listed below are some open problems and future research directions related to the topics presented in this thesis.

1. Let \( n \) be odd. In Chapter 3, it was shown that the quadratic function

\[
    f(x) = \frac{(n-1)/2}{\sum_{i=1}^{(n-1)/2} c_i Tr(x^{2^i+1})}, \quad x \in GF(2^n), c_i \in \{0, 1\},
\]

is optimal, i.e., has Hadamard transform \( 0, \pm 2^{(n+1)/2} \), if and only if \( \gcd(c(x), x^n+1) = x+1 \)

where \( c(x) = \sum_{i=1}^{(n-1)/2} c_i (x^i + x^{n-i}) \).

As an extension, we may try to find a simple characterization of when the more general quadratic function

\[
    f(x) = \frac{(n-1)/2}{\sum_{i=1}^{(n-1)/2} Tr(\beta_i x^{2^i+1})}, \quad x, \beta_i \in GF(2^n),
\]

is optimal.

Other directions to extend the work of Chapter 3 includes:

(a) Let \( n \) is even. We may try to characterize when the quadratic function

\[
    f(x) = \sum_{i \in I} c_i Tr(x^{2^i+1}), \quad x \in GF(2^n), c_i \in \{0, 1\},
\]

is bent, i.e. has Hadamard transform \( \pm 2^{n/2} \).
(b) Let $p$ be an odd prime. We may try to find when quadratic function

$$f(x) = \sum_{i \in I} c_i Tr(x^{p^i+1}), \ x \in GF(p^n), c_i \in GF(p),$$

has high nonlinearity. Some work has been done in this direction [42, 43].

2. Throughout the thesis, we have constructed highly nonlinear 1-resilient Boolean functions based on polynomials $f : GF(2^n) \to GF(2)$. We proved that they are 1-resilient by showing that the set

$$\{ \lambda \in GF(2^n) | \hat{f}(\lambda) = 0 \}$$

contains $n$ linearly independent vectors. Then by applying Algorithm 1 of Chapter 4, a 1-resilient Boolean representation can be found.

Most $GF(2^n)$-based constructions for highly nonlinear $k$-resilient Boolean functions satisfy $k \leq 1$ [28, 29, 40]. There are some $GF(2^n)$-based constructions for $k$-resilient function, $k > 1$, by Cheon and Chee [11, 12] based on algebraic curves, but the nonlinearity is not optimal.

It will be interesting to look for new $GF(2^n)$-based constructions of highly nonlinear $k$-resilient Boolean functions for $k > 1$.

3. Most known constructions for highly nonlinear resilient S-boxes are based on the Maiorana-McFarland constructions [31, 37, 46, 50]. There are constructions for resilient S-boxes based on finite fields [11, 12], but their nonlinearity bound (computed by applying Hasse-Weil bound on algebraic curves) are not high enough. It will be interesting to look for other constructions of resilient S-boxes based on finite fields which gives high nonlinearity.

4. There are many existing constructions [11, 12, 31, 37, 46, 50, 80] for resilient and highly nonlinear S-boxes. But they were not designed to have low maximum correlation (which is a relatively new concept introduced in 2000 [77]). It will be useful to compute the maximum correlation of these S-boxes to give a more complete measure of their security.

5. Most constructions for S-boxes with low maximum correlation have other good cryptographic properties like balance and high nonlinearity [8, 39, 77]. Since S-boxes with low maximum correlation are used as combiners in stream ciphers, we also require resiliency for protection against correlation attack. Therefore, it will be useful to construct S-boxes with high nonlinearity and low maximum correlation which are also resilient.
6. Extend the Patterson-Wiedemann search for highly nonlinear functions in Chapter 7 from $GF(2^{15})$ to $GF(2^{21})$. The search space can be further reduced by applying the method of Gangopadhyay et. al. [21, Section 2.1] for $n = 21$. 
Bibliography


