CONSTRUCTIONS OF RESILIENT BOOLEAN FUNCTIONS
WITH MAXIMUM NONLINEARITY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

M. ÖZGÜR ŞAHİN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

AUGUST 2005

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science

Prof. Dr. İsmet ERKMEN
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science

Assoc. Prof. Dr. Melek D. YÜCEL
Supervisor

Examining Committee Members

Prof. Dr. Yalçın Tanık (chairman) (METU, EEE)

Assoc. Prof. Dr. Melek D. Yücel (METU, EEE)

Prof. Dr. Kemal Leblebicioğlu (METU, EEE)

Assoc. Prof. Dr. Ferruh Özbudak (METU, MATH)

Assoc. Prof. Dr. Ali Doğanaksoy (METU, MATH)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : M. Özgür ŞAHİN

Signature :

ABSTRACT

CONSTRUCTIONS OF RESILIENT BOOLEAN FUNCTIONS WITH MAXIMUM NONLINEARITY

ŞAHİN, M. Özgür

M.S., Department of Electrical and Electronics Engineering

Supervisor: Assoc. Prof. Dr. Melek D. YÜCEL

August 2005, 59 pages

In this thesis, we work on the upper bound for nonlinearity of *t*-resilient Boolean functions given by Sarkar and Maitra, which is based on divisibility properties of spectral weights of resilient functions and study construction methods that achieve the upper bound.

One of the construction methods, introduced by Maity and Johansson, starts with a bent function and complements some values of its truth table corresponding to a previously chosen set of inputs, S, which satisfies three criteria. In this thesis, we show that a fourth criterion is needed for *t*-resiliency of the resulting function, and prove that three criteria of Maity and Johansson do not guarantee resiliency.

We also work on other constructions, one by Sarkar and Maitra, which uses a Maiorana-McFarland like technique to satisfy the upper bound and the other by Tarannikov, which satisfies the nonlinearity bound using a technique with low computational complexity. However, these methods have tendency to maximize the order of resiliency for a given number of variables, therefore one cannot construct functions for all possible resiliency values given the number of variables, using this method.

We further go into details and compute the auto-correlation functions of the constructed Boolean functions to find the absolute indicator and sum-of-squared-errors for each of them. We also provide a comparison of Boolean functions constructed by other techniques given in the literature, together with the ones studied in this thesis.

Keywords. Boolean function, nonlinearity, resiliency, correlation immunity, auto-correalation.

ÖZ

EN DOĞRUSAL OLMAYAN ESNEK BOOLE İŞLEVLERİNİN YAPIMI

ŞAHİN, M. Özgür

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Doç. Dr. Melek D. YÜCEL

Ağustos 2005, 59 sayfa

Bu tezde, Sarkar ve Maitra tarafından verilen ve esnek işlevlerin açılım değerlerinin bölünebilirlik özelliklerine dayanan, *t*-esnek Boole işlevlerinin doğrusal olmama üst sınırını ve bu üst sınıra ulaşan yapım yöntemlerini inceledik.

Anlatılan yapım yöntemlerinden biri, daha önce Maity ve Johansson tarafından önerilen bir yöntemdir. Yöntemlerinde, bükük bir işlevle başlayıp bu işlevin önceden belirlenmiş ve üç kısıtı sağlayan bir girdi kümesi, S'de bulunan girdilere karşılık gelen çıktıları evirmektedirler. Bu tezde, Maity ve Johansson'un kısıtlarının esnekliği kesinleştirmediğini ve dördüncü bir kısıtın gerekli olduğunu kanıtlıyoruz.

Ayrıca, doğrusal olmama üst sınırını sağlayan Sarkar ve Maitra'nın Maiorana-McFarland benzeri yöntemiyle Tarannikov'un hesap karmaşıklığı düşük yöntemi olmak üzere diğer yöntemleri inceliyoruz. Bu teknikler belirli bir değişken sayısı için esnekliği maksimuma çıkarma eğilimlerinden dolayı doğrusal olmama üst sınırının tüm noktaları için kullanılamamaktadır.

Ayrıca, yapılan Boole işlevlerinin kendi kendine bağlantı işlevlerini bulup mutlak gösterge ve kare-hatalar-toplamı değerlerini hesaplıyoruz. Bu tezde yapılan ve diğer yöntemlerle yapılmış Boole işlevlerinin bir karşılaştırmasını da yapıyoruz.

Anahtar Kelimeler. Boole işlevleri, doğrusal olmama, esneklik, bağlantı bağışıklığı, kendi kendine bağlantı.

To my beloved who has passed away

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Assoc. Prof. Dr. Melek D. YÜCEL for her guidance, advice, criticism, encouragements and insight throughout the research.

I am also grateful to Aselsan Inc. and my friends there for their tolerance and help in the completion of this thesis.

I would also like to give my special appreciation and gratitude to my family for their encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Boolean functions are important tools in cryptosystems, especially used for stream ciphers as combining functions in the construction of key stream generators based on Linear Feedback Shift Registers (LFSRs). Since security of a cryptosystem depends on its ability to guard against known attacks, Boolean functions – as a part of the system – should possess certain properties, such as high nonlinearity, high order of resiliency, high algebraic degree and low auto-correlation properties. As the complexity of the cryptosystem is proportional to the number of input variables to the Boolean function there should be an optimization between the number of input variables and maximization of the above parameters. Siegenthaler has shown that for an n-variable, unbalanced function of correlation immunity t and algebraic degree d, $t + d \leq n$ holds [Sieg1984]. If the function is balanced, then $t + d \leq n–1$.

Sarkar and Maitra [SarMai2000-1] have found an upper bound for the nonlinearity of t-resilient Boolean functions and inspired a number of researchers ([SarMai2000-1], [SarMai2000-2], [FedTar2001], [JohPas2003], [Tara2001], [PasJohMS2001], [Mait2000]) for discovering functions satisfying that upper bound, which is $nl_f \leq 2^{n-1} - 2^{t+1}$ for $t > n/2 - 2$ and $nl_f \leq 2^{n-1} - 2^{n/2-1} - 2^{t+1}$ for $t \leq n/2 - 2$. They have based their results to the fact that, for a t-resilient Boolean function f, $W_f(\omega) \equiv 0 \bmod 2^{t+2}$ for all $\omega \in GF(2)^n$, where $W_f(\omega)$ is the Walsh-Hadamard transform of f [SarMai2000-1]. This result certainly limits the possible values of the Walsh-Hadamard transform and therefore sets a limit to the nonlinearity of the function as described in Section 2.1.

In this thesis, we work on the nonlinearity bound given by Sarkar and Maitra [SarMai2000-1], and study construction methods that satisfy the upper bound on nonlinearity. Chapter 2 is devoted to introduce the notation and the preliminary concepts for comprehending the rest of the document and also includes the discussion on the nonlinearity bound of Sarkar and Maitra.

The construction method described in Chapter 3 is a previously introduced method by Maity and Johansson [MaiJoh2002] to construct *t*-resilient Boolean functions of nonlinearity $nl_{f'} = 2^{n-1} - 2^{n/2-1} - 2^{t+1}$, for $t \leq n/2$ - 2. In their method, Maity and Johansson start with a bent function and complement some values of its truth table corresponding to a previously chosen set of input vectors, S, which satisfies three criteria [MaiJoh2002]. In Theorem 3.2, we show that a fourth criterion is needed to prove that the resulting function is *t*-resilient and its nonlinearity satisfies $nl_{f'} \geq 2^{n-1} - 2^{n/2-1} - 2^{t+1}$. Combining this inequality with the upper bound of Sarkar and Maitra [SarMai2000-1], the construction yields a nonlinearity of $nl_{f'} = 2^{n-1} - 2^{n/2-1} - 2^{t+1}$, for $t \leq n/2$ - 2. We also show that, if $t > n/2$ - 2, the same construction yields a nonlinearity of $nl_{f'}$, where $2^{n-1} - 2^{n/2-1} - 2^{t+1} < nl_{f'} \leq 2^{n-1} - 2^{t+1}$. Without the introduction of the fourth criterion, one cannot guarantee the resiliency of the function *f*.

In Chapter 4, we study other construction methods, one by Sarkar and Maitra [SarMai2000-1], which uses a Maiorana-McFarland like technique to satisfy the upper bound and the other by Tarannikov [Tara2000], which satisfies the nonlinearity bound with a technique having low computational complexity. The construction method given by Sarkar and Maitra uses concatenation of functions with less number of variables and produces functions with high order of resiliency [SarMai2000-1]. A very important property of this construction technique is its being the first method to satisfy certain points of the nonlinearity bound, for example the function with nonlinearity 112, order of resiliency 3 and number of variables 8, given in Section 4.1. However, this method has tendency to maximize the order of resiliency for a given number of variables, therefore one cannot construct functions for all points of the upper bound for nonlinearity using this method, for example like the one we have constructed in Chapter 3, with nonlinearity 116, order of resiliency 1 and number of variables 8. This is because functions constructed by the method of Sarkar and Maitra must have the properties (n, t, d, z) obeying the rule $(3+2i+j, i+j, 2+i, 2^{2+2i+j}-2^{1+i+j})$ where i and j are natural numbers [SarMai2000-1].

The construction method given by Tarannikov – just like the one by Sarkar and Maitra – focuses only on functions with high order of resiliency, i.e. $t \geq \frac{2n-7}{3}$ [Tara2000]. Therefore, the same weakness of inability to construct Boolean functions of low order of resiliency is present for the method of Tarannikov. This is a major weakness because the construction of functions having small order of resiliency to achieve higher algebraic degree and better auto-correlation characteristics is usually desirable. On the other hand, the construction technique given by Tarannikov has a computational complexity linear on the number of variables, n, which is very preferable compared to the other methods having computational complexity 2^n [Tara2000].

In Chapter 5, we go more into details of the functions constructed in Chapters 3 and 4. We compute the auto-correlation functions of the constructed Boolean functions and find the absolute indicator and sum-of-squared-errors for each of them. We also provide a comparison of Boolean functions constructed by other techniques together with the ones studied in Chapters 3 and 4 in terms of nonlinearity, order of resiliency, the absolute indicator and sum-of-squared-errors.

CHAPTER 2

PRELIMINARIES

Let n be any positive integer. An n-variable Boolean function f: $GF(2)^n \rightarrow GF(2)$ maps each possible combination of n-bit variables to a single bit. Boolean functions play a major role in cryptosystems.

A Boolean function f can be represented as a polynomial over $GF(2)$:

$$f(x_1, x_2, ..., x_n) = a_0 \oplus \sum_{1 \le i \le n} a_i x_i \oplus \sum_{1 \le i < j \le n} a_{ij} x_i x_j \oplus ... \oplus a_{12...n} x_1 x_2 ... x_n$$

where a_0, a_{ij},..., $a_{\mathrm{12...n}} \in GF(2)$ and the multiplication and addition operations are in $GF(2)$. This representation of a Boolean function is called the Algebraic Normal Form (ANF) and the degree of this polynomial gives the algebraic degree of the function. Another representation of a Boolean function is its truth table, which lists the function output for all possible inputs. We denote the truth table of f with the 2^n dimensional vector f_t:

$$f_t = \{f(0), f(1), ..., f(2^n - 1)\}$$

There are a few properties of Boolean functions that are important in cryptographic applications. We now discuss those properties.

Definition 2.1. The *weight* of a Boolean function f, $wt(f)$, is the number of ones in its truth table.

Definition 2.2. An n-variable Boolean function f is said to be *balanced*, if the number of 0’s is equal to the number of 1’s in its truth table, i.e., $wt(f) = 2^{n-1}$.

Definition 2.3. A Boolean $g(x)$ is called an *affine* function of $x = (x_1, x_2,..., x_n) \in GF(2)^n$, if its degree is at most one, in other words, if its algebraic normal form is:

$$g(x_1, x_2,..., x_n) = a_0 \oplus a_1 x_1 \oplus a_2 x_2 \oplus ... \oplus a_n x_n = \omega \cdot x \oplus a_0$$

where $a_{0,} a_1,..., a_n \in GF(2)^n$, $\omega = (a_1,..., a_n) \in GF(2)^n$, and $\oplus$, $\cdot$ respectively denote addition and inner product operations in $GF(2)$. $g(x)$ is called linear if $a_0 = 0$.

Definition 2.4. The *Walsh-Hadamard transform* of a Boolean function f is defined as:

$$W_f(\omega) = \sum_{x \in GF(2)^n} (-1)^{f(x) \oplus x \cdot \omega}$$

For non-Boolean functions the Walsh Hadamard Transform is defined as:

$$W_r(\omega) = \sum_{x \in GF(2)^n} r(x)(-1)^{x \cdot \omega} \text{ [Yüce2001]}$$

and the inverse transform is:

$$W_{W_r(\omega)}{}^{-1}(x) = r(x) = 2^{-n} \sum_{\omega \in GF(2)^n} W_r(\omega)(-1)^{x \cdot \omega}$$

Definition 2.5. The *Hamming distance* between two functions $f(x)$ and $g(x)$ is defined as the number of inputs where the outputs differ, that is:

$$d_H(f, g) = \left|\{x \mid f(x) \neq g(x), x \in GF(2)^n\}\right|$$

Definition 2.6. The *nonlinearity* of a Boolean function is defined as its minimum Hamming distance to the set of affine functions

$$nl_f = \min_{g \in A_n} d_H(f, g)$$

where A_n is the set of n-variable affine Boolean functions. The Hamming distance between a Boolean function f and an affine function $g(x) = x \cdot \omega \oplus c$ can be calculated with the Walsh-Hadamard transform as

$$d_H(f,g) = 2^{n-1} - \frac{(-1)^c W_f(\omega)}{2}$$

Therefore, the nonlinearity of *f* can be obtained from the Walsh-Hadamard transform as

$$nl_f = 2^{n-1} - \frac{1}{2}\max_{\omega}\left|W_f(\omega)\right|$$

This measure of nonlinearity is important for linear cryptanalysis [Heys2000].

Carlet and Ding [CarDin2004] gives a more general definition for nonlinearity and investigates perfect nonlinear functions.

A very well known theorem says that for any Boolean function *f* the sum of squared values of the Walsh-Hadamard transform $W_f(\omega)$ is constant and equal to 2^{2n}

$$\sum_{\omega \in GF(2)^n} W_f^2(\omega) = 2^{2n}$$

Definition 2.7. For a perfectly nonlinear Boolean function, also called bent [Roth1976], the squared spectrum is flat and

$$W_f^2(\omega) = 2^n \text{, for all } \omega \in GF(2)^n.$$

Definition 2.8. Xiao and Massey [GouMas1988] provided a spectral characterization of correlation immune functions. Here we state that as a definition of *correlation immunity*:

A Boolean function *f* is *t*-th order correlation immune if and only if

$$W_f(\omega) = 0 \text{, for all } \omega \in GF(2)^n;\ 1 \leq wt(\omega) \leq t$$

If further *f* is balanced, $W_f(0) = 0$ and therefore

$$W_f(\omega) = 0 \text{, for all } \omega \in GF(2)^n;\ 0 \leq wt(\omega) \leq t$$

Balanced *t*-th order correlation immune functions are called *t*-resilient functions. From this point on, we will consider a balanced Boolean function as 0-resilient (this convention is accepted in [CamCarCS1991], [SarMai1999], [PasJoh1999], [Tara2000]) and an arbitrary Boolean function as (–1)-resilient.

The importance of correlation immunity is that, it provides zero mutual information between any *t*-tuple selected from the input vector and the output of the Boolean function i.e.,

$$I(x_{i_1},...,x_{i_t};Y)=0\text{, for } 1 \le i_1 < \ldots < i_t \le n$$

We define the mutual information between the two random variables *X* and *Y* as:

$$I(X;Y)=H(X)-H(X\,|\,Y)$$

where *H*(*X*) denotes the entropy and H(*X* |*Y*) denotes the conditional entropy, which are defined by:

$$H(X)=\sum_x P\{X=x\}\log\frac{1}{P\{X=x\}},$$

$$H(X\,|\,Y)=\sum_{x,y} P\{X=x,Y=y\}\log\frac{1}{P\{X=x\,|\,Y=y\}}$$

Therefore, having $I(X;Y)=0$ requires:

$H(X)=H(X\,|\,Y)$, hence $P\{X=x\}=P\{X=x\,|\,Y=y\}$

Theorem 2.1. Let *f* be an *n*-variable Boolean function and $\omega \in GF(2)^n$. *f* is *t*-resilient if and only if $P\{f(x)=\omega\bullet x\}=\frac{1}{2}$ for all ω of weight $0 \le wt(\omega) \le t$.

Proof.

We start the proof by finding the Hamming distance between the functions *f*(*x*) and $\omega \bullet x$, d_H(*f*(*x*), $\omega \bullet x$) in terms of the Walsh-Hadamard transform of *f*(*x*).

$$d_H(f(x), \omega \bullet x) = \left| \left\{ x \mid f(x) \neq \omega \bullet x, x \in GF(2)^n \right\} \right|$$

$$W_f(\omega) = \sum_{x \in GF(2)^n} (-1)^{f(x) \oplus x \cdot \omega}$$

$$= \left| \left\{ x \mid f(x) = \omega \bullet x, x \in GF(2)^n \right\} \right| - \left| \left\{ x \mid f(x) \neq \omega \bullet x, x \in GF(2)^n \right\} \right|$$

$$= 2^n - 2.\ d_H(f(x),\ \omega \bullet x)$$

Therefore; $d_H(f(x), \omega \bullet x) = 2^{n-1} - \dfrac{W_f(\omega)}{2}$

Since the functions *f*(*x*) and $\omega \bullet x$ differ in d_H(*f*(*x*), $\omega \bullet x$) number of places out of 2^n,

$$P\{f(x) \neq \omega \bullet x\} = \frac{d_H(f(x), \omega \bullet x)}{2^n} = \frac{1}{2} - \frac{W_f(\omega)}{2^{n+1}}$$

Hence,

$$P\{f(x) = \omega \bullet x\} = 1 - P\{f(x) \neq \omega \bullet x\} = \frac{1}{2} + \frac{W_f(\omega)}{2^{n+1}}$$

Using the definition of *t*-resiliency in the equation above, we arrive at the conclusion that *f* is *t*-resilient if and only if $P\{f(x) = \omega \bullet x\} = \dfrac{1}{2}$ for all ω of weight 0≤*wt*(ω)≤*t*.■

Theorem 2.2. Let *f* be an *n*-variable Boolean function and $\omega \in$ *GF*(2)^n. If the mutual information $I(f(x); \omega \bullet x) = 0$ for all ω of weight 0 ≤ *wt*(ω) ≤ *t*, then *f* is *t*-resilient.

Proof.

To check for resiliency by Theorem 2.1, we should compute the probability

$$P\{f(x)=\omega\bullet x\}=P\{f(x)=0,\ \omega\bullet x=0\}+P\{f(x)=1,\ \omega\bullet x=1\}$$

$$=P\{f(x)=0\big|\ \omega\bullet x=0\}P\{\omega\bullet x=0\}+P\{f(x)=1\big|\ \omega\bullet x=1\}P\{\omega\bullet x=1\}$$

for all ω of weight $0\le wt(\omega)\le t$.

In the equation above, we can substitute the followings:

1) $P\{f(x)\,|\,\omega\bullet x\}=P\{f(x)\}$. Because, by hypothesis the mutual information $I(f(x);\omega\bullet x)$ is equal to zero, which in turn requires $H(f(x))=H(f(x)\,|\,\omega\bullet x)$ for all ω of weight $0\le wt(\omega)\le t$.

2) $P\{\omega\bullet x=0\}=P\{\omega\bullet x=1\}=\frac{1}{2}$, since linear functions are balanced for all $\omega\neq 0$.

3) $P\{f(x)=0\}=p$ (note that $p=\frac{1}{2}$ if $f(x)$ is balanced).

Hence, for all ω of weight $0\le wt(\omega)\le t$,

$$P\{f(x)=\omega\bullet x\}=P\{f(x)=0\}P\{\omega\bullet x=0\}+P\{f(x)=1\}P\{\omega\bullet x=1\}$$

$$=p\frac{1}{2}+(1-p)\frac{1}{2}=\frac{1}{2}$$

Note that, for $\omega=0$, $P\{f(x)=\omega\bullet x\}=p$, which equals $\frac{1}{2}$ only if $f(x)$ is balanced, which proves *t*-resiliency of $f(x)$ by Theorem 2.1. ■

Note that a Boolean function satisfying the above property is uncorrelated with the linear combination of its inputs determined by the vector ω. If $0\le wt(\omega)\le t$, then the function is uncorrelated with all linear combinations of its inputs up to *t* variables.

Definition 2.9. The *algebraic degree* of a function f, denoted by deg(f), is defined as the number of variables in the longest term of f in its algebraic normal form. The *algebraic degree of variable* x_i in f, deg(f, x_i), is the number of variables in the longest term of f that contains x_i. If deg(f, x_i) = 0, then f does not depend on x_i. If deg(f, x_i) = 1, then f depends on x_i linearly. If deg(f, x_i) ≥ 2, we say that f depends on x_i nonlinearly.

Definition 2.10. The *autocorrelation function* of a Boolean function is given by

$$r_f(d) = \sum_{x \in GF(2)^n} (-1)^{f(x)} (-1)^{f(x \oplus d)} .$$

The maximum absolute value that we denote by $ac_f = \max_{d \neq 0 \in GF(2)^n} | r_f(d) |$ is also known as the absolute indicator [ZhaZhe1995].

Another measure related to the autocorrelation function is commonly called the sum-of-squares indicator [ZhaZhe1995], given by the sum $\sum_{d \in GF(2)^n} (r_f(d))^2$. We prefer to use the sum-of-squared-errors (SSE_f), $\sum_{d \neq 0 \in GF(2)^n} (r_f(d))^2$, instead of the sum-of-squares indicator, since SSE_f is proportional to the sum of squared spectrum deviations [Yüce2001] from that of the bent functions, that is

$$\sum_{d \neq 0 \in GF(2)^n} (r_f(d))^2 = 2^{-n} \sum_{\omega \in GF(2)^n} [(W_f(\omega))^2 - 2^n]^2 .$$

If f is affine, this sum of squared autocorrelation errors, i.e., the autocorrelation deviations from the autocorrelation of bent functions, is maximum and equal to $2^{3n} - 2^{2n}$. Hence, dividing the above equality by $2^{3n} - 2^{2n}$, one obtains the useful measure of mean squared error (MSE_f), which takes rational values in the interval [0,1]. The mean squared error percentage $100MSE_f$ of the Boolean function f shows the percentage of total squared deviations of its autocorrelation function $r_f(d)$ and squared spectrum $W_f^2(\omega)$ respectively, from the autocorrelation and the squared spectrum of bent functions [Yüce2001].

Another important properties of Boolean functions is that, for any t-resilient Boolean function f, $W_f(\omega) \equiv 0 \bmod 2^{t+2}$, for all $\omega \in GF(2)^n$ [SarMai2000-1]. We will make use of this property in Section 2.1.

2.1 The Nonlinearity Bound

In Chapter 3 and Chapter 4 we study methods to provide resilient Boolean functions of maximum nonlinearity. Sarkar and Maitra [SarMai2000-1] prove upper bounds of nonlinearity for n-variable, t-resilient Boolean functions, stated as the following:

1. If n is even and $t > n/2 - 2$, then $nl_f \le 2^{n-1} - 2^{t+1}$.
2. If n is even and $t \le n/2 - 2$, then $nl_f \le 2^{n-1} - 2^{n/2-1} - 2^{t+1}$.
3. If n is odd and $2^{t+1} > 2^{n-1} - nlmax(n)$, then $nl_f \le 2^{n-1} - 2^{t+1}$.
4. If n is odd and $2^{t+1} \le 2^{n-1} - nlmax(n)$, then nl_f is the highest multiple of 2^{t+1} which is less than or equal to $2^{n-1} - nlmax(n)$.

The proof is based on the congruency $W_f(\omega) \equiv 0 \bmod 2^{t+2}$ for any t-resilient Boolean function f [SarMai2000-1]. Having $nl_f = 2^{n-1} - 2^{n/2-1} - 2^{t+1}$ implies that there exist some ω_o such that $W_f(\omega_o) = 2^{n/2} + 2^{t+2}$. On the other hand, since $W_f(\omega_o) \equiv 0 \bmod 2^{t+2}$, the integer 2^{t+2} divides $2^{n/2} + 2^{t+2}$ and therefore it also divides $2^{n/2}$, which obviously requires $t \le n/2 - 2$. This result shows that for a t-resilient Boolean function with $t > n/2 - 2$, it is not possible to have a nonlinearity of $2^{n-1} - 2^{n/2-1} - 2^{t+1}$. We will make use of this result in Theorem 3.2 of Chapter 3, to evaluate the nonlinearity of the constructed Boolean function.

2.1.1 The Nonlinearity Bound Figures

In this section, we draw the figures of the mentioned upper bounds of nonlinearity for some specific values of n between 8-16. For $t > n/2 - 2$ we provide sketches of both $nl_f = 2^{n-1} - 2^{n/2-1} - 2^{t+1}$ and $nl_f = 2^{n-1} - 2^{t+1}$. The curve shows the upper bound of Sarkar and Maitra [SarMai2000-1], where the dotted part is plotted to show what would happen if the curve $nl_f = 2^{n-1} - 2^{n/2-1} - 2^{t+1}$ would continue for $t > n/2 - 2$ instead of $nl_f = 2^{n-1} - 2^{t+1}$. Note that, the difference between these curves is $2^{n/2-1}$ for all t.

The figures show the values of the maximum nonlinearity provided by Sarkar and Maitra [SarMai2000-1] for values of t less than n – 2. This is because, the value of maximum nonlinearity, $2^{n-1} - 2^{t+1}$ for $t > n/2 - 2$, is not positive for $t \geq n - 2$.

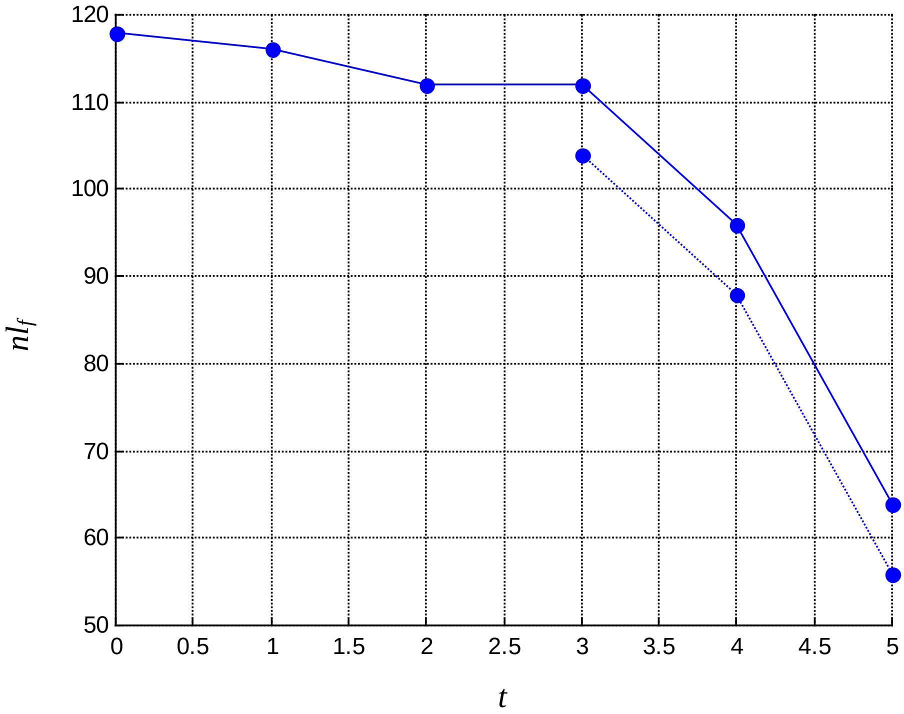

Figure 2-1. Nonlinearity Bounds For n = 8

Figure 2-1 shows the upper bound on nonlinearity for an 8-variable function. In Chapter 3 we provide an example function of order of resiliency 1 and nonlinearity 116, satisfying the upper bound of Sarkar and Maitra [SarMai2000-1]. In Chapter 4 we give another example function of resiliency 3 and nonlinearity 112, satisfying the upper bound at another point.

The following figures are sketched in order to give a general notion for the upper bound of nonlinearity for resilient Boolean functions. The existence of functions satisfying below figures is known and shown in Table 2.1 for (*n*, *t*) pairs ([SebZhaZ1994], [SarMai2000-1], [SarMai2000-2]), but we will not give examples of all points of the upper bound in this work.

Table 2-1. Known Functions Satisfying The Upper Bound

(*n*, *t*)	1	2	3	4	5	6	7	8
5	12	8	0					
6	24	24	16	0				
7	56	56*	48	32	0			
8	116	112	112	96	64	0		
9	244	240	240*	224	192	128	0	
10	492*	480	480	480*	448	384	256	0
Note: Existence of the functions marked with “*” is not known yet.								

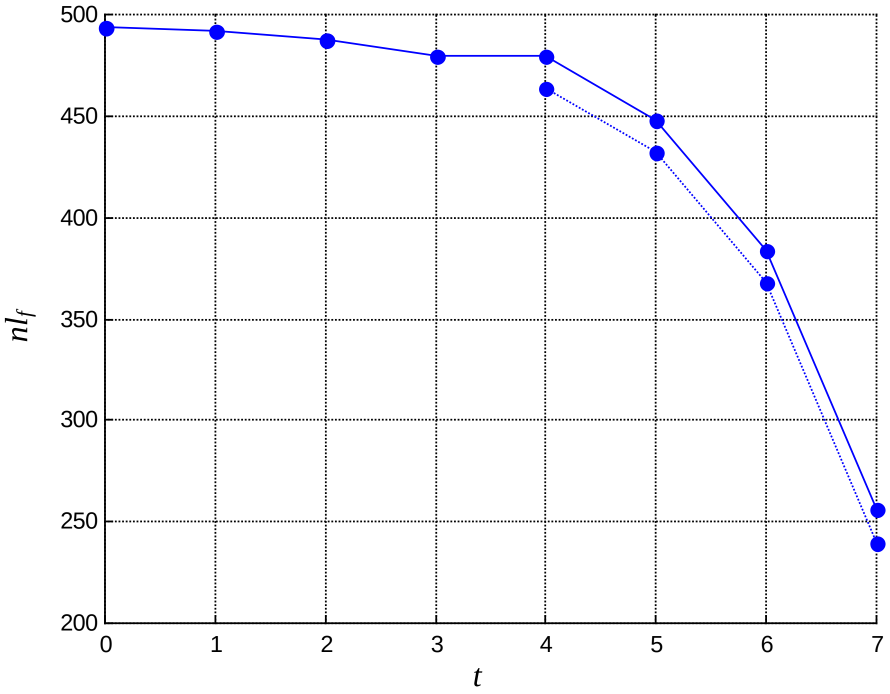

Figure 2-2. Nonlinearity Bounds For n = 10

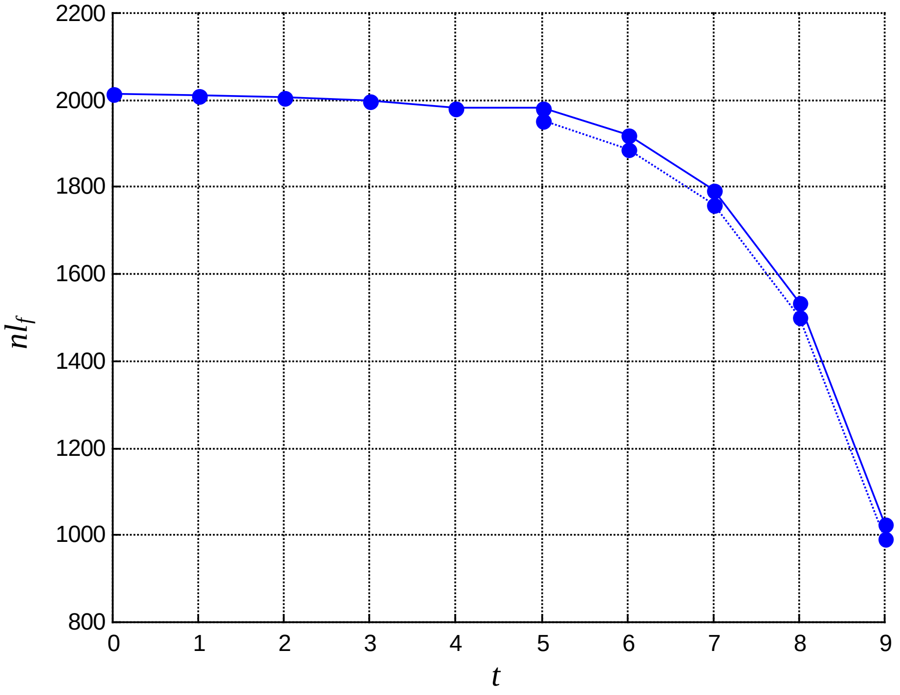

Figure 2-3. Nonlinearity Bounds For n = 12

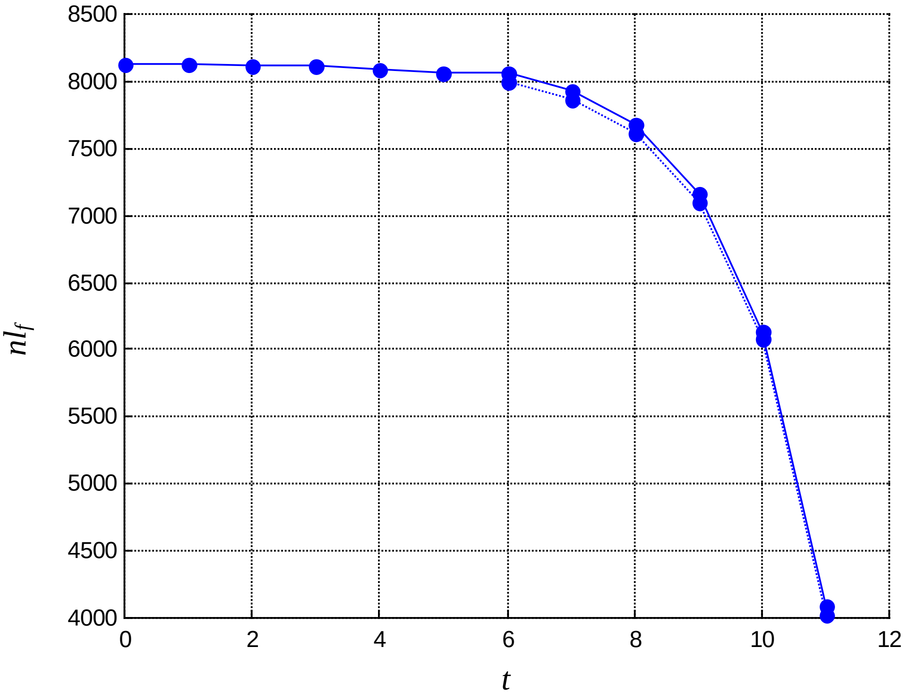

Figure 2-4. Nonlinearity Bounds For *n* = 14

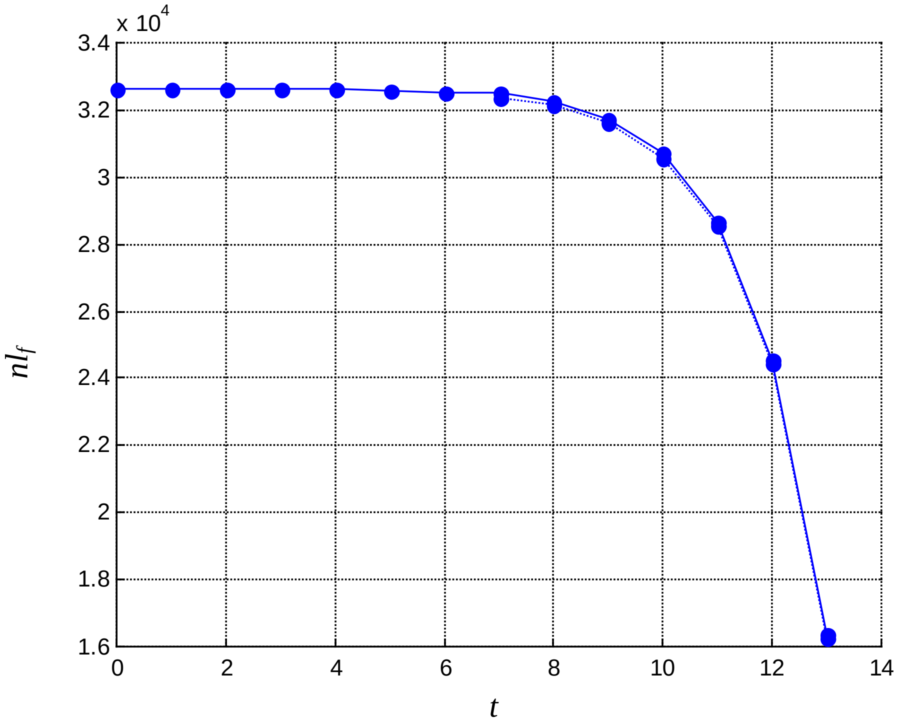

Figure 2-5. Nonlinearity Bounds For *n* = 16

CHAPTER 3

THE CONSTRUCTION METHOD

In this chapter, we investigate the construction method of Maity and Johansson [MaiJoh2002]. First we provide a proof for the nonlinearity of a 1-resilient function produced by this method and produce an example function using this method. Note that we use an additional criterion, which was missing in the paper of Maity and Johansson [MaiJoh2002], in the proof. After that we generalize the method to *t*-resiliency with Theorem 3.2.

3.1 Construction of 1-Resilient Boolean Functions

Lemma 3.1 for constructing 1-resilient Boolean functions where *n* is even and larger than 5, can be considered as a preliminary for the proof of Theorem 3.2 given for *t*-resilient functions. The construction method is very similar to the one proposed by Maity and Johansson [MaiJoh2002]. However, instead of criteria (i) and (ii) mentioned below, they give the following condition:

$$\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = +2^{n/2-1} \text{ for all } \omega \text{ such that and } 0 \leq wt(\omega) \leq 1, \tag{1}$$

The proof given below shows that (1) is not sufficient for 1-resiliency, unless (i) and (ii) of Lemma 3.1 are satisfied.

Lemma 3.1. Let *f* be an *n*-variable (*n* even and *n* > 5) bent function, and let $S \subseteq \{0,1\}^n$, such that

i. $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = +2^{n/2-1}$, for all ω such that $W_f(\omega) = +2^{n/2}$ and $0 \leq wt(\omega) \leq 1$,

ii. $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega}=-2^{n/2-1}$, for all ω such that $W_f(\omega) = -2^{n/2}$ and $0 \le wt(\omega) \le 1$,

iii. $-2^2 \le \sum_{x\in S}(-1)^{f\oplus x\cdot\omega} \le +2^{n/2}+2^2$, for all ω such that $W_f\ (\omega) = +2^{n/2}$ and $2 \le wt(\omega) \le n$,

iv. $-\left(2^{n/2}+2^2\right) \le \sum_{x\in S}(-1)^{f\oplus x\cdot\omega} \le +2^2$, for all ω such that $W_f\ (\omega) = -2^{n/2}$ and $2 \le wt(\omega) \le n$,

Then

$$f'(x)=\begin{cases} f(x)\oplus 1 & \text{if } x\in \mathrm{S} \\ f(x) & \text{otherwise}\end{cases}$$

is an n-variable 1-resilient function with nonlinearity $2^{n-1}-2^{n/2-1}-2^2$.

Proof.

Define $\overline{S}=\{0,1\}^n - \mathrm{S}$.

Assume $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega}=+2^{n/2-1}$, for all ω such that $W_f(\omega) = +2^{n/2}$ and $0 \le wt(\omega) \le 1$.

Since $f'(x)=f(x)\oplus 1$ for $x\in$ S,

$$\sum_{x\in S}(-1)^{f'\oplus x\cdot\omega}=-2^{n/2-1}.$$

Note that

$$\sum_{x\in S}(-1)^{f\oplus x\cdot\omega}+\sum_{x\in\overline{S}}(-1)^{f\oplus x\cdot\omega} = \sum_{x\in\{0,1\}^n}(-1)^{f\oplus x\cdot\omega} = W_f(\omega) \tag{2}$$

therefore,

$$\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = +2^{n/2-1}.$$

Since $f'(x)=f(x)$ for $x\in\overline{S}$,

$$\sum_{x\in\overline{S}}(-1)^{f'\oplus x\cdot\omega} = +2^{n/2-1}.$$

Therefore, $W_{f'}(\omega)=\sum_{x\in S}(-1)^{f'\oplus x\cdot\omega}+\sum_{x\in\overline{S}}(-1)^{f'\oplus x\cdot\omega}=0$ follows and shows the resiliency of *f'*, provided that $W_f(\omega) = +2^{n/2}$.

Assuming $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = -2^{n/2-1}$, for all ω such that $W_f(\omega) = -2^{n/2}$ and $0 \le wt(\omega) \le 1$, similarly completes the proof of the 1-resiliency condition for *f'*.

Now we have left to prove the nonlinearity. Assume $-2^2 \le \sum_{x\in S}(-1)^{f\oplus x\cdot\omega} \le +2^{n/2}+2^2$ for all ω such that $W_f(\omega) = +2^{n/2}$ and $2 \le wt(\omega) \le n$.

Again using the definition of *f'* we write the following,

$$-2^{n/2}-2^2 \le \sum_{x\in S}(-1)^{f'\oplus x\cdot\omega} \le +2^2. \tag{3}$$

Using (2) together with the definition of *f'* we find,

$$-2^2 \leq \sum_{x \in S} (-1)^{f' \oplus x \cdot \omega} \leq +2^{n/2} + 2^2 . \qquad (4)$$

Combining (3) and (4) gives us the following limit for the Walsh-Hadamard transform of *f'*,

$$-2^{n/2} - 2^3 \leq W_{f'}(\omega) \leq +2^{n/2} + 2^3$$

The condition $-\left(2^{n/2} + 2^2\right) \leq \sum_{x \in S} (-1)^{f \oplus x \cdot \omega} \leq +2^2$, for all ω such that $W_f(\omega) = -2^{n/2}$ and $2 \leq wt(\omega) \leq n$ gives the same interval for the Walsh-Hadamard transform of *f'* and therefore shows a bound for nonlinearity of *f'*:

$$nl_{f'} \geq 2^{n-1} - 2^{n/2 - 1} - 2^2$$

Combining this lower bound with the upper bound of [SarMai2000-1] ($nl_{f'} \leq 2^{n-1} - 2^{n/2-1} - 2^2$ for a 1-resilient function), we conclude

$$nl_{f'} = 2^{n-1} - 2^{n/2-1} - 2^2 \text{ for } n \geq 6 \qquad ■$$

The above equality does not hold for n ≤ 6, because then $2^{n/2} + 2^3$ would not be a multiple of 2^3 (remember the discussion made in Chapter 2).

The following 1-resilient function, f_{C1}, of nonlinearity 116, was produced using this method.

f_{C1}={E880D555B33366668F0F5A5A3C3C696980FF55AA33CC66990FF05AA53CC36997}

For finding an initial bent function, we make use of Theorem 3.1, which constructs the bent function by concatenating affine functions.

Theorem 3.1. Let l_i be independent affine functions for $0 \le i \le {}^{n}\!/_{2} - 1$. Let f_{bent} be concatenation of the ${}^{n}\!/_{2}$ affine functions, l_i's. Then f_{bent} is an n-variable bent function.

Proof.

Recall the well-known fact that the Walsh-Hadamard transform of an ${}^{n}\!/_{2}$-variable affine function is an impulse of magnitude $2^{n/2}$. Concatenating the ${}^{n}\!/_{2}$-variable functions to form an n-variable function makes any jth element of the truth table of the constructed n-variable function to be the sum of kth elements of all the concatenated functions' spectrums with (+) or (–) polarity. Since all the concatenated functions are independent affine functions, only – and exactly – one of those functions have a $\pm\ 2^{n/2}$ term in the kth position of its spectrum. Therefore, every term in the spectrum of the constructed function, f_{bent}, is $\pm\ 2^{n/2}$, which means the function f_{bent} is bent. ■

We construct the initial bent function, f_{C1b}, using this idea as:

f_{C1b}={000055AA0F0F5AA5333366993C3C6996555500FF5A5A0FF0666633CC69693CC3}

and select the set S by inspection as:

S = {(0,0,0,0,0,0,0,0), (0,0,0,0,0,0,0,1), (0,0,0,0,0,0,1,0), (0,0,0,0,0,1,0,0), (0,0,0,0,1,0,0,0), (0,0,0,1,0,0,0,0), (0,0,1,0,0,0,0,0), (0,1,0,0,0,0,0,0), (1,0,0,0,0,0,0,0), (1,1,1,1,1,1,1,1)}

Note that this function has sum of squares error of 32640 and autocorrelation value of 32.

Selection of Elements of S

For constructing the set S, we first selected a set S$_1$ = {(0,0,0,0,0,0,0,1), (0,0,0,0,0,0,1,0), (0,0,0,0,0,1,0,0), (0,0,0,0,1,0,0,0), (0,0,0,1,0,0,0,0), (0,0,1,0,0,0,0,0), (0,1,0,0,0,0,0,0), (1,0,0,0,0,0,0,0)}, which consists of vectors of weight 1. Since we are constructing a 1-resilient function, ω with $wt(\omega) \leq 1$ are considered for assumptions (i) and (ii) of Theorem 3.1, this selection of inputs makes the inner product $x \bullet \omega = 1$ for $x = \omega$ and $x \bullet \omega = 0$ for $x \neq \omega$, and therefore $\sum_{x \in S_1} (-1)^{f \oplus x \cdot \omega} = 6$ assuming $f(x) = 0$ for $x \in$ S$_1$. This assumption is reasonable also for we want to make the resultant function balanced and the starting bent function has 16 ($2^{n/2}$) more zeros than ones and therefore 8 zeros should be complemented for balancedness. In order to satisfy the assumptions (i) and (ii) of Theorem 3.1 and 1-resiliency, we must have $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = 8$, hence we need to include another set S$_2$ in S such that, S = S$_1$ $\cup$ S$_2$ and $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = 8$ for all ω, where $wt(\omega) \leq 1$. We select S$_2$ = {(0,0,0,0,0,0,0,0), (1,1,1,1,1,1,1,1)} and satisfy the above condition for f(0,0,0,0,0,0,0,0) = 0 and f(1,1,1,1,1,1,1,1) = 1.

One thing left for finalizing the set S is to guarantee the assumptions made above on the output of the starting bent function at the input values selected from S. Recall that we construct the initial bent function by concatenation of linear functions. Hence, satisfying the above assumptions is easy by organizing the order of linear functions used in the concatenation.

For large values of n and t, a systematic method for obtaining the set S is not present, however a search algorithm can be used despite its high computational complexity.

3.2 Construction of *t*-Resilient Boolean Functions

We now generalize Lemma 3.1 to the construction of t-resilient functions. Note that, the construction method is very similar to the one proposed by Maity and Johansson

[MaiJoh2002] except for our additional constraint (ii) as in Lemma 3.1. Here we also evaluate the nonlinearity of the resultant function, where $t > {}^{n}\!/_{2}$ - 2.

Theorem 3.2. Let f be an n-variable (n even) bent function, and let $S \subseteq \{0,1\}^n$, such that

$\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = +2^{n/2-1}$, for all ω such that $W_f(\omega) = +2^{n/2}$ and $0 \le wt(\omega) \le t$,

$\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = -2^{n/2-1}$, for all ω such that $W_f(\omega) = -2^{n/2}$ and $0 \le wt(\omega) \le t$,

$-2^{t+1} \le \sum_{x\in S}(-1)^{f\oplus x\cdot\omega} \le +2^{n/2} + 2^{t+1}$, for all ω such that $W_f(\omega) = +2^{n/2}$ and $t+1 \le wt(\omega) \le n$,

$-\left(2^{n/2} + 2^{t+1}\right) \le \sum_{x\in S}(-1)^{f\oplus x\cdot\omega} \le +2^{t+1}$, for all ω such that $W_f(\omega) = -2^{n/2}$ and $t+1 \le wt(\omega) \le n$,

Then

$$f'(x) = \begin{cases} f(x)\oplus 1 & \text{if } x \in S \\ f(x) & \text{otherwise} \end{cases}$$

is an n-variable t-resilient function with nonlinearity $2^{n-1} - 2^{n/2-1} - 2^{t+1}$ if $t \le n/2 - 2$. Moreover, if $t > n/2 - 2$ then f' has nonlinearity larger than $2^{n-1} - 2^{n/2-1} - 2^{t+1}$, i.e.,

$$2^{n-1} - 2^{n/2-1} - 2^{t+1} < nl_{f'} \le 2^{n-1} - 2^{t+1}.$$

Proof.

Define $\overline{S}$ ={0,1}^{n} – S.

Assume $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega}=+2^{n/2-1}$ for all ω such that $W_f(\omega)$ = $+\,2^{n/2}$ and 0 ≤ $wt(\omega)$ ≤ t.

Since $f'(x)=f(x)\oplus 1$ for $x\in$ S,

$$\sum_{x\in S}(-1)^{f'\oplus x\cdot\omega}=-2^{n/2-1}.$$

Note that

$$\sum_{x\in S}(-1)^{f\oplus x\cdot\omega}+\sum_{x\in\overline{S}}(-1)^{f\oplus x\cdot\omega}=\sum_{x\in\{0,1\}^n}(-1)^{f\oplus x\cdot\omega}=W_f(\omega). \tag{5}$$

therefore,

$$\sum_{x\in\overline{S}}(-1)^{f\oplus x\cdot\omega}=+2^{n/2-1}.$$

Since $f'(x)=f(x)$ for $x\in\overline{S}$,

$$\sum_{x\in\overline{S}}(-1)^{f'\oplus x\cdot\omega}=+2^{n/2-1}.$$

Therefore, $W_{f'}(\omega)=\sum_{x\in S}(-1)^{f'\oplus x\cdot\omega}+\sum_{x\in\overline{S}}(-1)^{f'\oplus x\cdot\omega}=0$ follows and shows the resiliency of *f'*, provided that $W_f(\omega)$ = $+\,2^{n/2}$.

Assuming $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega}=-2^{n/2-1}$ for all ω such that $W_f(\omega)$ = $-\,2^{n/2}$ and 0 ≤ $wt(\omega)$ ≤ t, similarly completes the proof of the t-resiliency condition for *f'*.

Now we have left to prove the nonlinearity. Assume $-2^{t+1} \leq \sum_{x \in S} (-1)^{f \oplus x \cdot \omega} \leq +2^{n/2} + 2^{t+1}$ for all ω such that $W_f(\omega) = +2^{n/2}$ and $t+1 \leq wt(\omega) \leq n$.

Again using the definition of *f'* we write the following,

$$-2^{n/2} - 2^{t+1} \leq \sum_{x \in S} (-1)^{f' \oplus x \cdot \omega} \leq +2^{t+1}. \qquad (6)$$

Using (5) together with the definition of *f'* we find,

$$-2^{t+1} \leq \sum_{x \in \overline{S}} (-1)^{f' \oplus x \cdot \omega} \leq +2^{n/2} + 2^{t+1}. \qquad (7)$$

Combining (6) and (7) gives us the following limit for the Walsh-Hadamard transform of *f'*,

$$-2^{n/2} - 2^{t+2} \leq W_{f'}(\omega) \leq +2^{n/2} + 2^{t+2}$$

The condition $-\left(2^{n/2} + 2^{t+1}\right) \leq \sum_{x \in S} (-1)^{f \oplus x \cdot \omega} \leq +2^{t+1}$ for all ω such that $W_f(\omega) = -2^{n/2}$ and $t+1 \leq wt(\omega) \leq n$ gives the same interval for the Walsh transform of *f'* and therefore shows a bound for nonlinearity of *f'*:

$$nl_{f'} \geq 2^{n-1} - 2^{n/2-1} - 2^{t+1}$$

Combining this lower bound with the upper bound of [SarMai2000-1] ($nl_{f'} \leq 2^{n-1} - 2^{n/2-1} - 2^{t+1}$ for a *t*-resilient function where $t \leq n/2 - 2$), we conclude

$$nl_{f'} = 2^{n-1} - 2^{n/2-1} - 2^{t+1}, \text{ for } t \leq n/2 - 2$$

Moreover, if $t > {}^{n}\!/_{2} - 2$ then $nl_{f'} \leq 2^{n-1} - 2^{t+1}$ [SarMai2000-1] and $nl_{f'}$ = $2^{n-1} - 2^{n/2-1} - 2^{t+1}$ is not possible because of the discussion made in Chapter 2, therefore

$$2^{n-1} - 2^{n/2-1} - 2^{t+1} < nl_{f'} \leq 2^{n-1} - 2^{t+1}, \text{ for } t > {}^{n}\!/_{2} - 2$$

Also note that, since $t > {}^{n}\!/_{2} - 2$, the term $2^{n/2-1}$ above is smaller than 2^{t+1} and therefore the first meaningful term (remember that $W_f(\omega_o) \equiv 0 \bmod 2^{t+2}$ for a t-resilient Boolean function) larger than $2^{n-1} - 2^{n/2-1} - 2^{t+1}$ is $2^{n-1} - 2^{t+1}$. Hence,

$$nl_{f'} = 2^{n-1} - 2^{t+1}, \text{ for } t > {}^{n}\!/_{2} - 2 \qquad \blacksquare$$

Remember the first two assumptions of the construction method. We assume $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = +2^{n/2-1}$ for all ω such that $W_f\ (\omega) = +2^{n/2}$, $0 \leq wt(\omega) \leq t$ and $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = -2^{n/2-1}$ for all ω such that $W_f\,(\omega) = -2^{n/2}$, $0 \leq wt(\omega) \leq t$. In their work, Maity and Johansson make the assumption $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = 2^{n/2-1}$ for all ω such that $0 \leq wt(\omega) \leq t$. Here, we will show why this assumption is not sufficient for constructing resilient Boolean functions.

Let us assume $\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = +2^{n/2-1}$ for some ω such that $W_f\ (\omega) = -2^{n/2}$ and $0 \leq wt(\omega) \leq t$. Since $f'(x) = f(x) \oplus 1$ for $x \in$ S,

$$\sum_{x \in S} (-1)^{f' \oplus x \cdot \omega} = -2^{n/2-1}.$$

Making use of (5) and $f'(x) = f(x)$ for $x \in \overline{S}$ we arrive

$$\sum_{x\in\bar{S}}(-1)^{f\oplus x\cdot\omega} = \sum_{x\in\bar{S}}(-1)^{f'\oplus x\cdot\omega} = -3\cdot 2^{n/2-1} \text{ since } W_f(\omega) = -2^{n/2} \text{ for } \omega.$$

Therefore, $W_{f'}(\omega) = \sum_{x\in S}(-1)^{f'\oplus x\cdot\omega} + \sum_{x\in\bar{S}}(-1)^{f'\oplus x\cdot\omega} = -2^{n/2+1} \neq 0$ for some ω such that $0 \le wt(\omega) \le t$ and *f'* is not resilient.

The above discussion shows that assuming $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = 2^{n/2-1}$ for all ω such that $0 \le wt(\omega) \le t$ alone is not sufficient for resiliency of the constructed function. This assumption also misses the occasion for constructing resilient Boolean functions where we have some ω such that $0 \le wt(\omega) \le t$ where $\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = -2^{n/2-1}$. We have proved in Theorem 3.2 that it is possible to construct resilient functions in this case.

CHAPTER 4

OTHER CONSTRUCTIONS SATISFYING THE UPPER BOUND

In this chapter, we study some more methods satisfying the upper bound for nonlinearity given in Section 2.1. Eventhough the methods provided in this chapter satisfy the nonlinearity bound, they have their weaknesses as well. Here we comment on the weaknesses of these construction methods and investigate the autocorrelation properties of the constructed functions, which has not attracted enough attention in the literature.

4.1 Construction by Sarkar and Maitra

In this section, we investigate the method of Sarkar and Maitra [SarMai2000-1], which focuses on achieving the best possible trade-off among the cryptographic parameters: number of variables, order of resiliency, nonlinearity and algebraic degree. The authors show that functions achieving the best possible trade-off can be constructed by the Maiorana-McFarland like technique [SarMai2000-1].

We claim that, the weakness of the Maiorana-McFarland like technique is that it assigns the highest priority to resiliency. On the other hand, the algebraic degree or autocorrelation characteristics, for example, may have greater importance in the application. Moreover, some of the functions satisfying the previously discussed nonlinearity bound cannot be constructed with this technique. For example, an 8-variable function with order of resiliency 1 and nonlinearity 116, to which we gave an example in the previous chapter, cannot be constructed by the Maiorana-McFarland like technique. Furthermore, the intention to maximize the order of resiliency distorts the autocorrelation characteristics of the function, as we will see.

4.1.1 Preliminary Concepts

Before giving the description of the construction technique, let us first give the preliminary concepts for comprehension.

Maiorana-McFarland Like Construction Technique

The Maiorana-McFarland like technique is one of the construction techniques used to construct resilient Boolean functions and has been investigated previously in the literature ([CamCarCS1991], [Carl1997], [CheLeeLS1996], [SarMai2000-3], [SebZhaZ1994]). This technique of construction basically uses the following idea. Let π be a map from $\{0,1\}^r$ to $\{0,1\}^k$, where for any $x \in \{0,1\}^r$, $wt(\pi(x)) \geq t + 1$. Let f be a Boolean function from $\{0,1\}^{r+k}$ to $\{0,1\}$, such that, $f(x, y) = y \bullet \pi(x) \oplus g(x)$, where $x \in \{0,1\}^r$, $y \in \{0,1\}^k$ and $y \bullet \pi(x)$ is the inner product of y and $\pi(x)$. Then f is t-resilient.

f can be interpreted as a concatenation of 2^r affine functions $l_0, l_1, \dots, l_{2^r-1}$ from the set of k-variable affine functions, where $\left|\{x \mid l_i(x) \text{ is nondegenerate}\}\right| \geq t + 1$ for $0 \leq i \leq 2^r - 1$. The construction technique of this chapter is based on this idea.

Let us define an (n, t, d, z) function as an n-variable, t-resilient Boolean function of degree d with nonlinearity x. Also note that, given an n-variable function, there may be more than one possible values of order of resiliency t, such that the upper bound on the nonlinearity is the same.

Definition 4.1. An (n, t, d, z) function is said to be a *saturated maximum degree function* if:

1. z is the upper bound on nonlinearity for n-variable, t-resilient Boolean functions.

2. t is the maximum possible value for order of resiliency for given number of variables n, and upper bound on nonlinearity.

3. d satisfies Siegenthaler's inequality with equality, i.e, $d = n - t - 1$.

4. The spectrum of the function is three valued.

The definition of a saturated maximum degree function is important because, one can generate a sequence of Boolean functions using the technique described in Section 4.1.2, each of which is a saturated maximum degree function. This idea is based on the fact that if an (n, t, $n - t - 1$, z) function f is a saturated function, then so is an ($n + 1$, $t + 1$, $n - t - 1$, $2z$) function g [SarMai2000-1].

4.1.2 The Construction Method

In this section we will construct an (8, 3, 4, 112) function using the technique of Sarkar and Maitra.

For a Boolean function f let us define $NZ(W_f) = \{\omega \,|\, W_f(\omega) \neq 0\}$, where W_f is the Walsh-Hadamard transform of f.

Lemma 4.1. Let f_1, f_2 be two (7, 3, -, 48) functions such that $NZ(W_{f_1}) \cap NZ(W_{f_2}) = \varnothing$. Then the function f, $f = (1 \oplus x_8) f_1 \oplus x_8 f_2$, is an (8, 3, -, 112) function [SarMai2000-1].

One can use construction or search techniques to find (7, 3, -, 48) satisfying the above criteria. The authors use concatenation of smaller functions to get f_1, f_2 and construct an (8, 3, 4, 112) function out of them.

An example function, constructed by this method is given below. The given function is an (8, 3, 4, 112) function:

f_{C2}={077CE5A2F8831A5DF8831A5D077CE5A26996699669699696669999665AA5A55A}

4.1.3 Further Discussions

Note that the example function given has sum of squares error of 196608 and autocorrelation value of 128. These values are very large compared to the function of the previous chapter. This is because, the order of resiliency is higher, which leads to worse spectral characteristics. In fact, a function constructed using this Maiorana-McFarland like technique is to be a $(3+2i+j, i+j, 2+i, 2^{2+2i+j}-2^{1+i+j})$ function, where *i* and *j* are larger than or equal to zero [SarMai2000-1]. That is why this technique cannot construct an 8-variable function with order of resiliency 1 and nonlinearity 116.

4.2 Construction by Tarannikov

In this section, we investigate the method of Tarannikov [Tara2000], which is introduced in "On Resilient Boolean Functions with Maximal Possible Nonlinearity". In his paper Tarannikov proves that the nonlinearity of an *n*-variable ($n \geq 4$) *t*-resilient Boolean function does not exceed $2^{n-1}-2^{t+1}$, which is the upper bound mentioned in Section 2.1 for $t > n/2 - 2$. In Section 2.1 we also provide the bound given by Sarkar and Maitra [SarMai2000-1], $nl_f \leq 2^{n-1} - 2^{n/2-1} - 2^{t+1}$, for small *t*, i.e., $t \leq n/2 - 2$. The construction method explained in this section focuses only on functions with high order of resiliency, $t \geq \frac{2n-7}{3}$. In the rest of this section we study the method of Tarannikov and investigate the strengths and weaknesses of this method, in terms of cryptographically important properties other than nonlinearity.

4.2.1 Preliminary Concepts

In equivalent non-probabilistic formulation, a Boolean function *f* is called correlation immune of order *t* if $wt(f') = wt(f)/2^t$ for any its subfunction *f'* of $n - t$ variables [Tara2000], which is obtained by substituting *t* variables of *f* with some constants. Therefore, a Boolean function *f* is called *t*-resilient if $wt(f') = 2^{n-t-1}$ for any its subfunction *f'* of $n - t$ variables.

Definition 4.2. A Boolean function $f(x_1,\ldots,x_n)$ depends on a pair of its variables (x_i, x_j) quasilinearly if $f(\mathbf{x_1}) \neq f(\mathbf{x_2})$ for any two vectors $\mathbf{x_1}$ and $\mathbf{x_2}$ of length *n* that differ only in *i*th and *j*th components. A pair (x_i, x_j) in this case is called a *pair of* *quasilinear variables* in *f*. Note that, this definition yields to a *linear variable* in *f* if only one variable, x_i, is concerned.

Lemma 4.1. Let $f(x_1,\ldots,x_n)$ be a Boolean function represented in the form

$$f(x_1,...,x_n) = \sum_{\sigma_1,...,\sigma_n} (x_1 \oplus \sigma_1)...(x_l \oplus \sigma_l) f(\sigma_1 \oplus 1, \ldots, \sigma_l \oplus 1, x_{l+1}, \ldots, x_n).$$

Suppose that all 2^l subfunctions $f(\sigma_1 \oplus 1, \ldots, \sigma_l \oplus 1, x_{l+1}, \ldots, x_n)$ are *t*-resilient. Then the function *f* is also *t*-resilient [Tara2000].

Lemma 4.2. Let $f(x_1,\ldots,x_n)$ be a Boolean function. If *f* depends on some variable x_i linearly, then *f* is balanced [Tara2000].

Corollary 4.1. Let $f(x_1,\ldots,x_n)$ be a Boolean function. If *f* depends on some variables $x_1, \ldots , x_s$ linearly, then *f* is (*s*-1)-resilient [Tara2000].

This result is used for constructing resilient Boolean functions in Section 4.2.2 “The Construction Method”. Note that this result also emphasizes the fact that decreasing the nonlinearity of a function *f* may increase its degree of resiliency, which means, once more, resiliency and nonlinearity take counter parts for a cryptographically strong Boolean function.

Lemma 4.3. Let $f(x_1,\ldots,x_n, x_{n+1}) = g(x_1,\ldots,x_n) \oplus cx_{n+1}$ where $c \in \{0,1\}$. Then $nl_f = 2\, nl_g$ [Tara2000].

4.2.2 The Construction Method

The method of construction, which is explained in this section, is based on the following lemma, Lemma 4.4.

Lemma 4.4. Let *n* be a positive integer. Let $f_1(x_1,\ldots,x_n)$ and $f_2(x_1,\ldots,x_n)$ be *t*-resilient Boolean functions on $GF(2)^n$ such that $nl_{f_1} \geq N_0$, $nl_{f_2} \geq N_0$. Moreover, there exist two variables x_i and x_j such that f_1 depends on the variables x_i and x_j linearly, and f_2 depends on a pair the variables (x_i, x_j) quasilinearly. Then the function

$$f(x_1,\ldots,x_n,x_{n+1}) = (x_{n+1} \oplus 1) f_1(x_1,\ldots,x_n) \oplus x_{n+1} f_2(x_1,\ldots,x_n)$$

is a *t*-resilient Boolean function on $GF(2)^{n+1}$ with nonlinearity $nl_f \geq 2^{n-1} + N_0$ and the function

$$f(x_1,\ldots,x_n,x_{n+1},x_{n+2}) = (x_{n+1} \oplus x_{n+2} \oplus 1) f_1(x_1,\ldots,x_n) \oplus (x_{n+1} \oplus x_{n+2}) f_2(x_1,\ldots,x_n) \oplus x_{n+1}$$

is a (*t*+1)-resilient Boolean function on $GF(2)^{n+2}$ with nonlinearity $nl_f \geq 2^n + 2N_0$ and depends on a pair of the variables (x_{n+1}, x_{n+2}) quasilinearly.

The following function is derived using the method of Tarannikov [Tara2000]. The derivation starts from a 2-variable function and reaches to the final 8-variable Boolean function iteratively. The algebraic normal form of this example function is given by

$$\begin{aligned} f &= (x_8 \oplus 1)(x_1x_2x_7 \oplus x_1x_3x_7 \oplus x_1x_4x_7 \oplus x_1x_4 \oplus x_3x_7 \oplus x_4x_7 \oplus x_2 \oplus x_3 \oplus x_5 \oplus x_6) \\ &\oplus x_8(x_1x_2x_5 \oplus x_1x_2x_6 \oplus x_1x_3x_5 \oplus x_1x_3x_6 \oplus x_1x_4x_5 \oplus x_1x_4x_6 \oplus x_1x_4 \oplus x_3x_5 \oplus x_3x_6 \\ &\oplus x_4x_5 \oplus x_4x_6 \oplus x_2 \oplus x_3 \oplus x_5 \oplus x_7) \end{aligned}$$

The above function *f* is a 3-resilient Boolean function with nonlinearity 112, which satisfies the upper bound for nonlinearity given in Section 2.1.

4.2.3 Further Discussions

The construction technique, described above, produces resilient Boolean functions with maximum possible nonlinearity for the given order of resiliency where $t \geq \frac{2n-7}{3}$. Moreover, the computational complexity of the algorithm being used is linear on *n*, which is superior to the other algorithms being considered. However this technique guarantees that the constructed function depends on some variables linearly, which means that those variables decreases the overall nonlinearity in addition to weakening the function by directly introducing linear variables to the function.

Considering the result give in Lemma 4.3 one can see that addition of a linear variable to a function doubles the functions nonlinearity. However, this is a little increase in nonlinearity remembering the fact that the upper bound on nonlinearity increases exponentially with the number of variables. Considering the condition c = 0 makes it easier to realize this result.

In Section 2.1, we have provided an upper bound on nonlinearity of resilient Boolean functions, which lies on two different lines for $t > n/2 - 2$ and $t \leq n/2 - 2$. The nonlinearity bound given by Tarannikov [Tara2000] is proved to be valid for $t \geq \frac{2n-7}{3}$. Note that, $n/2 - 2 < \frac{2n-7}{3}$ for $n > 2$, which means that the nonlinearity bound provided by Tarannikov is already included in the upper bound given in Section 2.1.

CHAPTER 5

ADDITIONAL RESULTS

In this chapter we give and study the results that we provide in this work. Specifically, we investigate the auto-correlation characteristics of the functions constructed in Chapters 3 and 4, which is not considered in any of the works trying to satisfy the upper bound on nonlinearity. Those works mentioned in Chapters 3 and 4 mainly consider a trade-off between nonlinearity and degree of resiliency. They focus on constructing Boolean functions of maximum nonlinearity for a given degree of resiliency or vice versa. Despite the fact that, the auto-correlation properties of Boolean functions are cryptographically important, they are not investigated in literature.

Recall from Definition 2.10 that, the auto-correlation function of a Boolean function gives the relation between the input difference and the output. In other words, the auto-correlation function gives the difference between the number of outputs that are same for a specific input difference vector, *d*, and the number of outputs that are different for that input difference vector, in the function's truth table. Therefore, the weakness in the auto-correlation properties of a function can be used for differential cryptanalysis of the function [Heys2000]. Since it is cryptographically important, here we study the auto-correlation properties of the functions introduced in Chapters 3 and 4.

Theorem 5.1. Let $r_f(d) = \sum_{x \in \mathrm{GF}(2)^n} (-1)^{f(x)} (-1)^{f(x \oplus d)}$ be the auto-correlation function of *f*. Then $r_f(d) = W_{W_f^2(\omega)}^{-1}(d)$.

Proof.

We start the proof from the definition of auto-correlation function. Then we take the Walsh-Hadamard transform of the auto-correlation function.

$$r_f(d) = \sum_x (-1)^{f(x)} (-1)^{f(x \oplus d)}$$

$$R(\omega) = W_{r_f(d)}(\omega) = \sum_d r_f(d)(-1)^{\omega \bullet d}$$

$$R(\omega) = \sum_d \sum_x (-1)^{f(x)} (-1)^{f(x \oplus d)} (-1)^{\omega \bullet d}$$

$$R(\omega) = \sum_x (-1)^{f(x)} \sum_d (-1)^{f(x \oplus d)} (-1)^{\omega \bullet d}$$

Let $a \in GF(2)^n$ and substitute $a = x \oplus d$.

$$R(\omega) = \sum_x (-1)^{f(x)} (-1)^{\omega \bullet x} \sum_a (-1)^{f(a)} (-1)^{\omega \bullet a}$$

Hence, we reach the important result

$$W_{r_f(d)}(\omega) = R(\omega) = W_f{}^2(\omega)$$

where, $r_f(d) = W_{W_f^2(\omega)}{}^{-1}(d)$ follows. ■

Definition 5.1. Let f be an n-variable Boolean function. We define the *sequence vector* or the *polar form* of the function f as follows:

$$f_s = \left\{(-1)^{f(0)}, (-1)^{f(1)}, \ldots, (-1)^{f(2^n - 1)}\right\}$$

Definition 5.2. We define the 2^n x 2^n *Hadamard matrix* as,

$$H_1 = [1] \text{ and } H_{2^n \times 2^n} = \begin{bmatrix} H_{2^{n-1} \times 2^{n-1}} & H_{2^{n-1} \times 2^{n-1}} \\ H_{2^{n-1} \times 2^{n-1}} & -H_{2^{n-1} \times 2^{n-1}} \end{bmatrix}$$

Hadamard matrix is important for us because it offers easy computation of the Walsh-Hadamard transform of a Boolean function.

Theorem 5.2. One can form the 2^n dimensional spectral vector of *f* $W_f = \{W_f(0), W_f(1), \ldots, W_f(2^n - 1)\}$, which is the output vector of the Walsh - Hadamard transform of the Boolean function *f*, by multiplying the sequence vector f_s by the 2^n x 2^n Hadamard matrix.

$$W_f = f_s \times H_{2^n \times 2^n}$$

Proof.

The proof is based on the properties of Hadamard matrix, and is given by Piotr Porwik [Porw2003]. ■

Gupta and Sarkar have introduced an alternative computation method for computing Walsh-Hadamard values of a Boolean function from its algebraic normal form [GupSar2003]. Although their method is good for calculating the Walsh-Hadamard transform at a specific point, the algorithm does not find all values of the spectrum, which is necessary for finding the nonlinearity of the function. Moreover, we do not investigate functions of a large number of variables and only the function constructed by Tarannikov is in its algebraic normal form. Therefore we do not use their method.

5.1 The Absolute Indicator and Sum-of-Squared-Errors

In Definition 2.10 we give the definition for the auto-correlation function of a Boolean function, $r_f(d)$, and define the absolute indicator, as $ac_f = \max_{d \neq 0 \in \mathrm{GF}(2)^n} | r_f(d) |$. The absolute indicator is a measure of strength of the function against differential attacks, where the attack uses the high auto-correlation value at a single point. Therefore, we need a low valued absolute indicator for functions to resist this kind of attacks.

For finding the absolute indicator, we make use of Theorem 5.1 and Theorem 5.2. We use Theorem 5.2 to find the spectrum of the Boolean function, whereas, Theorem 5.1 is used as an easy computation method to calculate its auto-correlation function.

The absolute indicator is not the only measure related to the auto-correlation function. As the high auto-correlation at one point cause cryptographic weakness, the distribution of high auto-correlation to more than one point is also a weakness. The sum-of-squares indicator suggested by Zhang et al [ZhaZhe1995], is a measure of strength against attacks, using the distributed high auto-correlation. Note that $r_f(0)$ for any n-variable Boolean function f is constant and equal to 2^n. Therefore the $r_f^{\ 2}(0)$ term in the sum-of-squares indicator does not carry information about the function. Moreover, the sum-of squared-errors (SSE_f) defined as,

$$SSE_f = \sum_{d \neq 0 \in \ GF(2)^n} (r_f(d))^2$$

is proportional to the sum of squared deviations from that of the bent functions [Yüce2001].

5.1.1 The AI and SSE of the Function Constructed in Chapter 3

Recall from Section 3.1 that the truth table of the constructed function was:

f_{C1}={E880D555B33366668F0F5A5A3C3C696980FF55AA33CC66990FF05AA53 CC36997}.

We use Theorem 5.2 and calculate the Walsh-Hadamard transform of the function f_{C1} as:

$$W_{f_{C1}} = f_{C1_s} \times H_{2^8 \times 2^8}$$

$W_{f_{C1}} =$

{0 0 0 8 0 8 8 8 0 8 8 8 8 8 8 16 0 8 8 8 8 8 8 16 -24 -24 -24 -16 -24 -16 -16 -16 0 8 8 8 -24 -24 -24 -16 8 8 8 16 -24 -16 -16 -16 8 8 8 16 -24 -16 -16 -16 -24 -16 -16 -16

16 16 16 24 0 8 -24 -24 8 8 -24 -16 8 8 -24 -16 8 16 -16 -16 8 8 -24 -16 8 16 -16 -16 -24 -16 16 16 -16 -16 16 24 8 8 -24 -16 -24 -16 16 16 8 16 -16 -16 -16 -16 16 24 8 16 -16 -16 -16 -16 16 24 -16 -16 16 24 16 24 -8 -8 0 -24 8 -24 8 -24 8 -16 -24 8 -24 16 -24 16 -16 16 8 -24 8 -16 8 -16 16 -16 8 -16 16 -16 16 -16 16 -8 8 -24 8 -16 -24 16 -16 16 -24 16 -16 16 16 -16 16 -8 8 -16 16 -16 -16 16 -16 24 16 -16 16 -8 -16 24 -8 24 8 -24 -24 16 8 -16 -16 16 -24 16 16 -16 -16 16 16 -8 8 -16 -16 16 16 -16 -16 24 16 -16 -16 24 16 -8 -8 24 8 -16 -16 16 -16 16 16 -8 -16 16 16 -8 16 -8 -8 24 16 -16 -16 24 -16 24 24 -8 16 -8 -8 24 -8 24 24 0}

One can see that $\max|W_{f_{C1}}(\omega)|$ = 24 and therefore the nonlinearity, $nl_{f_{C1}} = 2^{8-1} - \frac{1}{2}\max_{\omega}|W_{f_{C1}}(\omega)|$ = 116.

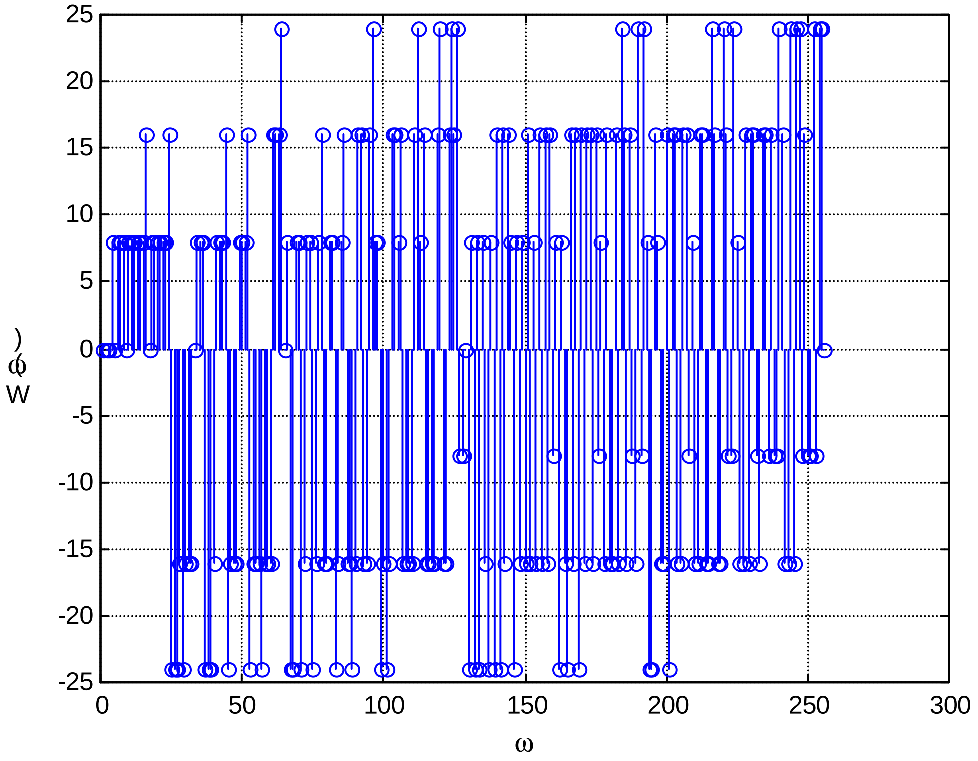

Figure 5-1. The Walsh-Hadamard Transform of f_{C1}

Note that, $W_{f_{C1}}(0) = W_{f_{C1}}(2^n) = 0$ for $0 \le n \le 7$, which implies that the function is 1-resilient. Also note that, $W_{f_{C1}}(\omega)$ is nonzero for all other values of *ω*, so that the overall spectrum of the function has a small maximum absolute value in order to satisfy high nonlinearity.

Now we calculate the autocorrelation function $r_{f_{C1}}(d)$ using Theorem 5.1:

$$r_{f_{C1}}(d) = W_{W^2_{f_{C1}}(\omega)}{}^{-1}(d)$$

$$r_{f_{C1}} =$$

{256 -16 -16 0 -16 0 -16 -8 -16 -16 -16 -24 -16 -24 -8 -8 -16 16 0 -8 0 -8 -8 -8 16 -8 8 -8 8 -8 -8 8 -16 0 -16 -8 32 8 8 -8 -16 -24 -8 -8 8 8 8 8 0 -8 -8 -8 8 8 -8 24 8 -8 -8 8 8 -8 8 -8 -16 0 32 8 -16 -8 8 -8 -16 -24 8 8 -8 -8 8 8 0 -8 8 8 -8 -8 -8 24 8 -8 8 -8 -8 8 8 -8 -16 -8 8 -8 8 -8 -8 8 -8 -8 8 8 8 8 8 8 -8 -8 -8 24 -8 24 8 -24 -8 8 8 -8 8 -8 -24 16 -16 16 -16 8 -16 8 -8 -8 -16 24 -8 8 -8 8 -8 8 -16 -8 -24 -8 -24 -8 -8 8 24 24 8 8 8 8 8 8 -16 8 -8 -8 8 8 8 8 -8 8 -8 8 8 -8 -8 8 -24 -8 -8 8 8 -8 8 -8 8 8 8 8 -8 -8 8 0 -16 8 8 8 -8 -8 8 8 -8 8 8 -8 -8 8 -8 8 -24 -8 8 -8 -8 8 8 -8 8 8 -8 -8 8 8 8 0 -8 -8 8 8 8 8 8 -24 -8 8 -8 8 -8 8 8 -16 -8 8 8 -8 8 -8 8 16 8 8 8 0 8 0 -16 -16}

Let us sketch the auto correlation function as:

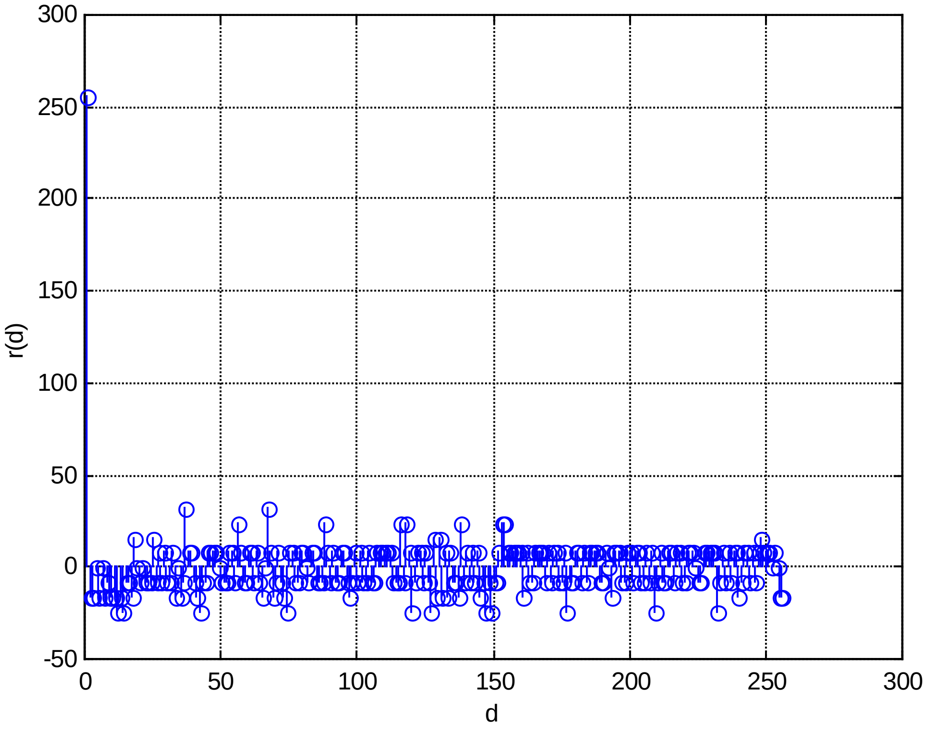

Figure 5-2. The auto-correlation function of f_{C1}

The vector form and the sketch of the auto-correlation function show us that the absolute indicator is:

$$ac_{f_{C1}} = 32.$$

We calculate the sum-of-squared errors using its definition and the auto-correlation vector to get the result:

$$SSE_{f_{C1}} = 32640$$

5.1.2 The AI and SSE of the Function by Sarkar and Maitra

Recall from Section 4.1 that the truth table of the constructed function was:

f_{C2}={077CE5A2F8831A5DF8831A5D077CE5A2699669966696996966669999665AA5A55A}.

We use Theorem 5.2 and calculate the Walsh-Hadamard transform of the function f_{C2} as:

$$W_{f_{C2}} = f_{C2_s} \times H_{2^8 \times 2^8}$$

$$W_{f_{C2}} =$$

{0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 32 0 0 0 0 0 0 0 32 0 0 0 32 0 32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 32 0 0 0 0 0 0 0 -32 0 0 0 32 0 -32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 32 0 0 0 0 0 0 0 32 0 0 0 -32 0 -32 0 0 0 0 0 32 0 -32 0 0 0 32 0 0 32 -32 -32 32 0 32 0 0 32 32 32 -32 32 -32 32 -32 32 32 -32 -32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -32 0 0 0 0 0 0 0 -32 0 0 0 -32 0 -32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -32 0 0 0 0 0 0 0 32 0 0 0 -32 0 32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -32 0 0 0 0 0 0 0 -32 0 0 0 32 0 32 0 0 0 0 0 32 0 -32 0 0 0 32 0 0 32 -32 -32 -32 0 32 0 0 32 32 32 32 32 -32 32 32 32 -32 -32 -32}

One can see that $\max\left|W_{f_{C2}}(\omega)\right|$ = 32 and therefore the nonlinearity, $nl_{f_{C2}} = 2^{8-1} - \frac{1}{2}\max_{\omega}\left|W_{f_{C2}}(\omega)\right|$ = 112.

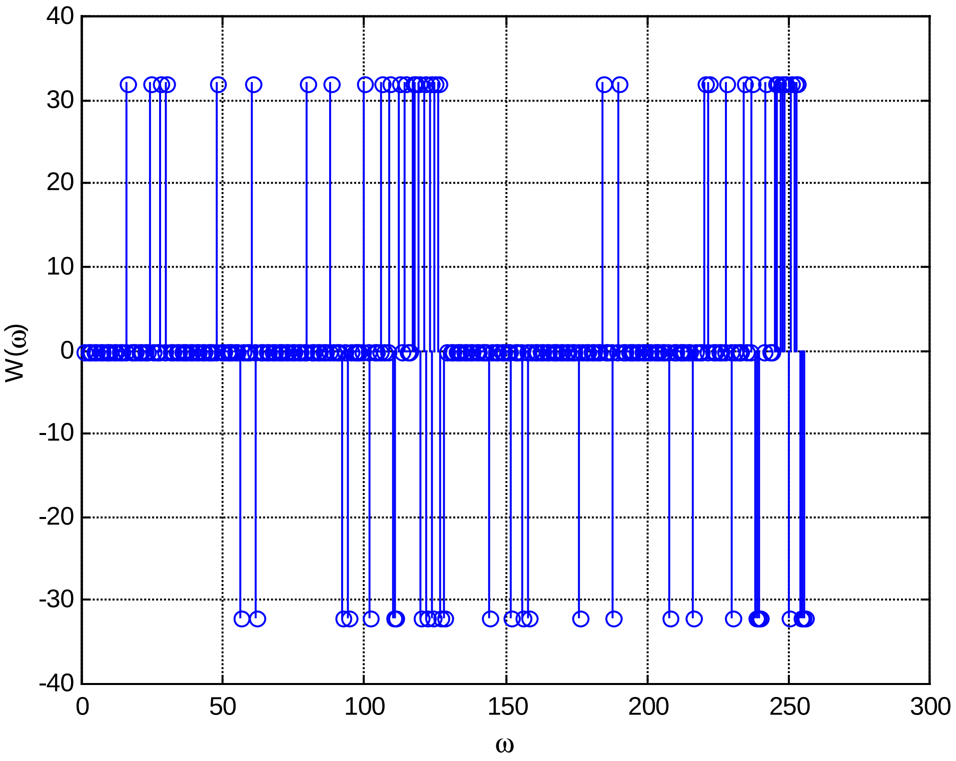

Figure 5-3. The Walsh-Hadamard Transform of f_{C2}

One can see that $\max\left|W_{f_{C2}}(\omega)\right|$ = 32 and therefore the nonlinearity, $nl_{f_{C2}} = 2^{8-1} - \frac{1}{2}\max_{\omega}\left|W_{f_{C2}}(\omega)\right|$ = 112. Remember that 112 is the best achievable nonlinearity for an 8-variable, 3-resilient Boolean function, and $W_f(\omega) \equiv 0 \bmod 2^{t+2}$ for any t-resilient Boolean function [SarMai2000-1]. Therefore, 32 is the minimum and maximum possible absolute value of the spectrum to achieve both 112 nonlinearity and 3-resiliency. Note that $W_{f_{C2}}(\omega)$ is a three-valued function, where the absolute value of all nonzero elements is 32, which inherently implies a good spectral distribution. Also note that, since the spectrum is three valued, the number of nonzero terms is constant, remembering

$$\sum_{\omega \in GF(2)^n} W_f^2(\omega) = 2^{2n}$$

from Chapter 2.

Now we calculate the autocorrelation function $r_{f_{C2}}(d)$ using Theorem 5.1:

$$r_{f_{C2}}(d) = W_{W_{f_{C2}}^2(\omega)}^{-1}(d)$$

$r_{f_{C2}} =$

{256 -128 -32 32 -96 32 0 0 -96 32 0 0 0 0 32 -32 -96 32 0 0 0 0 32 -32 0 0 32 -32 96 -32 -128 128 -128 0 -32 32 32 32 0 0 32 32 0 0 0 0 32 -32 32 32 0 0 0 0 32 -32 0 0 32 -32 -32 -32 0 0 -128 0 -32 32 32 32 0 0 32 32 0 0 0 0 32 -32 32 32 0 0 0 0 32 -32 0 0 32 -32 -32 -32 0 0 128 0 32 -32 -32 -32 0 0 -32 -32 0 0 0 0 -32 32 -32 -32 0 0 0 0 -32 32 0 0 -32 32 32 32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0}

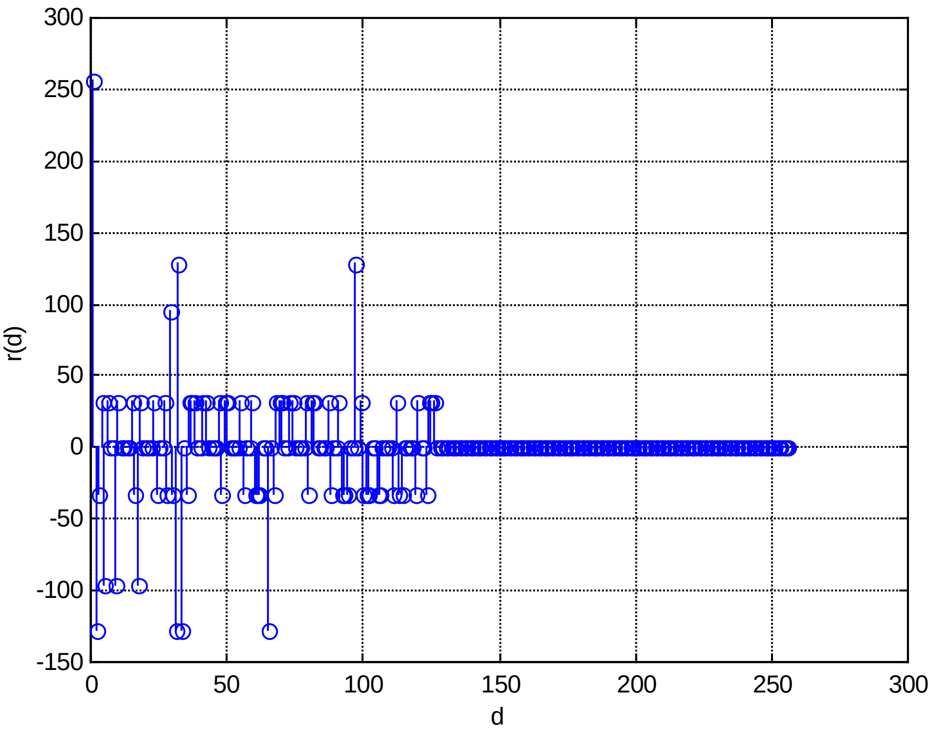

Figure 5-4. The auto-correlation function of f_{C2}

There are two basic differences between the auto-correlation functions of the two functions, f_{C1} and f_{C2}. The auto-correlation function of f_{C1} has small values for nonzero terms and distributed over the entire d axis, which shows a slight shift from the bent function. However, f_{C2} has all its nonzero places in the first 126 terms of the vector form of the auto-correlation function and with higher values. Furthermore, we calculate the SSE for this function as:

$$SSE_{f_{C2}} = 196608$$

which is a worse value compared to f_{C1} and further from a bent function.

The vector form and the sketch of the auto-correlation function show us that the absolute indicator is:

$$ac_{f_{C2}} = 128.$$

5.1.3 The AI and SSE of the Function by Tarannikov

In Section 4.2.2, we give the function f_{C3} as:

$$f_{C3}(x_1, x_2, x_3, ..., x_8) = (x_8 \oplus 1)(x_1x_2x_7 \oplus x_1x_3x_7 \oplus x_1x_4x_7 \oplus x_1x_4 \oplus x_3x_7 \oplus x_4x_7 \oplus x_2 \oplus x_3 \oplus x_5 \oplus x_6) \oplus x_8(x_1x_2x_5 \oplus x_1x_2x_6 \oplus x_1x_3x_5 \oplus x_1x_3x_6 \oplus x_1x_4x_5 \oplus x_1x_4x_6 \oplus x_1x_4 \oplus x_3x_5 \oplus x_3x_6 \oplus x_4x_5 \oplus x_4x_6 \oplus x_2 \oplus x_3 \oplus x_5 \oplus x_7)$$

From here we generate the truth table of the function, in order to use in the computation of its cryptographic values.

f_{C3}={1BE43C96C369E41BE41BC3693C961BE41BE4E41BE41B1BE4C3693C963 C96C369}.

We use Theorem 5.2 and calculate the Walsh-Hadamard transform of the function f_{C3} as:

$$W_{f_{C3}} = f_{C3_s} \times H_{2^8 \times 2^8}$$

$W_{f_{C3}} =$

{0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 -32 32 0 0 32 -32 32 32 -32 -32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 -32 32 0 0 32 -32 32 32 -32 -32 0 0 0 0 0 0 32 -32 0 0 32 -32 32 32 32 32 0 0 0
0 0 0 32 -32 0 0 32 -32 32 32 32 32 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 32 -32 0 0 -32 32 -32 -32 32 32
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -32 32 0 0 32 -32 32 32 -32 -32 0 0 0 0 0 0
32 -32 0 0 32 -32 32 32 32 32 0 0 0 0 0 0 -32 32 0 0 -32 32 -32 -32 -32 -32}

One can see that $\max|W_{f_{C1}}(\omega)|$ = 32 and therefore the nonlinearity, $nl_{f_{C1}} = 2^{8-1} - \frac{1}{2}\max_{\omega}|W_{f_{C1}}(\omega)|$ = 112.

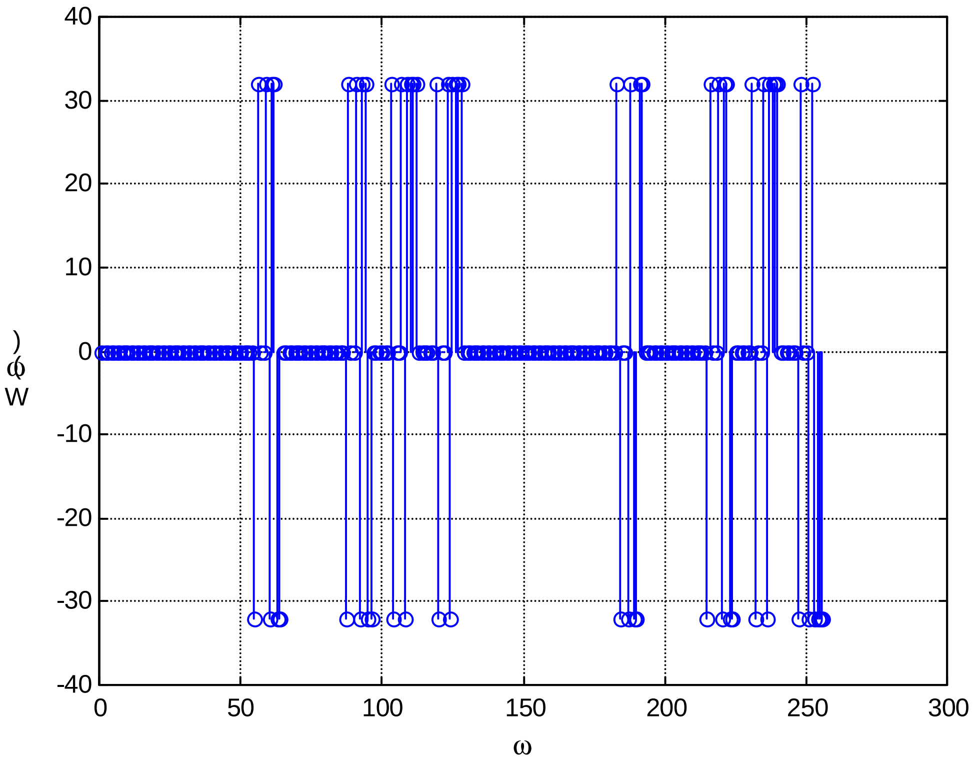

Figure 5-5. The Walsh-Hadamard Transform of f_{C3}

As in the case of f_{C2}, the function f_{C3} is also an 8-variable, 3-resilient Boolean function with nonlinearity 112 and its Walsh-Hadamard transform is a three-valued function, where the absolute value of all nonzero elements is 32. The difference of the two spectra is that, the spectrum of f_{C2} is more spread than that of f_{C3}. Eventhough this difference in the spectrum, the nonlinearity, absolute indicator and sum-of-squared-errors – as we will see in Section 5.2 – values of both functions are the same.

Now we calculate the autocorrelation function $r_{f_{C3}}(d)$ using Theorem 5.1:

$$r_{f_{C3}}(d) = W_{W^2_{f_{C3}}(\omega)}^{-1}(d)$$

$$r_{f_{C3}} =$$

{256 0 -128 0 -128 0 0 0 -128 0 0 0 0 0 128 0 -128 0 64 0 64 0 0 0 64 0 0 0 0 0 -64 0
-128 0 64 0 64 0 0 0 64 0 0 0 0 0 -64 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -128 0 64 0 64
0 0 0 64 0 0 0 0 0 -64 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
128 0 -64 0 -64 0 0 0 -64 0 0 0 0 0 64 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0}

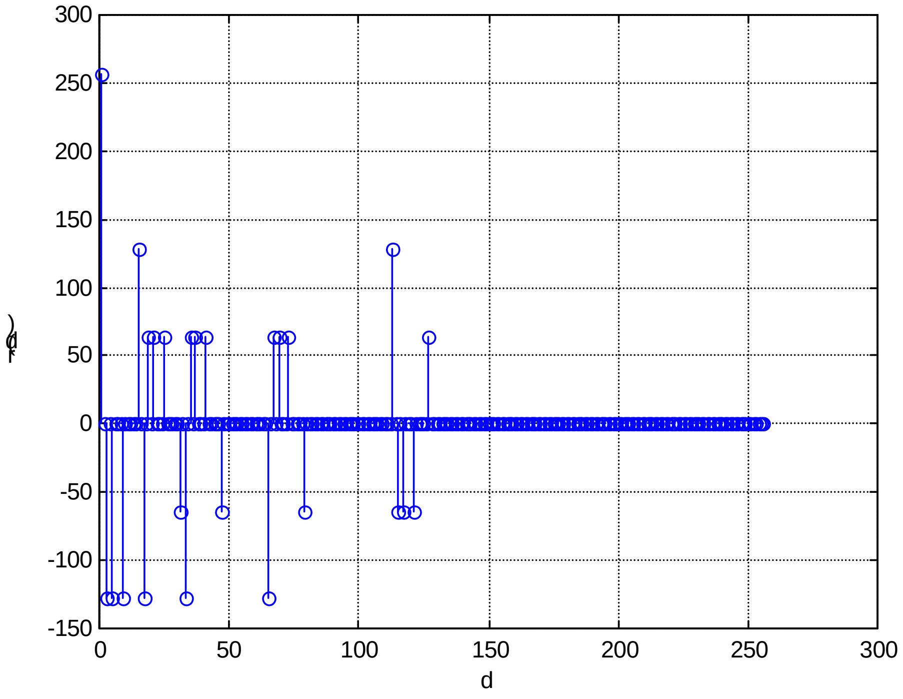

Figure 5-6. The auto-correlation function of f_{C3}

As can be seen in Figure 5-6, the auto-correlation characteristics of f_{C3} is very similar to that of f_{C2} and the SSE is the same for both functions.

$$SSE_{f_{C3}} = 196608$$

The vector form and the sketch of the auto-correlation function show us that the absolute indicator is:

$$ac_{f_{C3}} = 128.$$

5.2 Comparison of the Constructed Functions

In this section, we make a comparison of the functions constructed in the previous chapters of this work, together with some other functions provided by some previous works [KavYüc2003]. Table 5-1 gives the comparison of functions in terms of nonlinearity, order of resiliency, absolute indicator and sum-of-squared-errors.

Table 5-1. Comparison of 8-Variable Functions

Function f	nl_f	**Order of Resiliency,** t	$\lvert r_f(d)\rvert_{\max_{d\neq 0}}$ (ac_f)	$\sum_{all\ d\neq 0} r_f^2(d)$ (SSE_f)	$\frac{\sum_{all\ d\neq 0} r_f^2(d)}{167116.8}$ $(100MSE_f)$
Affine	0	-1 or 0	256	16711680	100 %
Function constructed in Chapter 3	116	1	32	32640	0.1953125%
Function by Sarkar and Maitra [SarMai2000-1]	112	3	128	196608	1.176471 %

Table 5-1 Comparison of 8-Variable Functions (Continued)

Function f	nl_f	**Order of Resiliency,** t	$\lvert r_f(d)\rvert_{\max_{d\neq 0}}$ **(ac_f)**	$\sum_{all\ d\neq 0} r_f^2(d)$ **(SSE_f)**	$\frac{\sum_{all\ d\neq 0} r_f^2(d)}{167116.8}$ **(100MSE_f)**
Function by Tarannikov [Tara2000]	112	3	128	196608	1.176471 %
Stanica, Sung (see [KavYüc2003])	112		256	196608	1.176471 %
Cauteaut et al (see [KavYüc2003])	112		256	172032	1.029412 %
Maitra (see [KavYüc2003])	116		128	55296	0.330882 %
Kavut, Yücel [KavYüc2003]	114	0	16	23424	0.140165 %
Kavut, Yücel [KavYüc2003]	116	0	24	21120	0.126378 %
Bent	120	-1	0	0	0 %

CHAPTER 6

CONCLUSIONS

In this thesis work, we have studied the upper bound on nonlinearity of a resilient Boolean function, provided by Sarkar and Maitra [SarMai2000-1]. In Section 2.1, we investigated the given upper bound, given to be $nl_f \leq 2^{n-1} - 2^{t+1}$ for $t > n/2 - 2$ and $nl_f \leq 2^{n-1} - 2^{n/2-1} - 2^{t+1}$ for $t \leq n/2 - 2$, and based on the fact that for any given *t*-resilient Boolean function *f*, $W_f(\omega) \equiv 0 \bmod 2^{t+2}$ for all $\omega \in GF(2)^n$. We then commented on the meaning of this upper bound in terms of cryptographic properties and sketch it for different number of variables for easy interpretation.

In Chapter 3 we worked on a construction method, which was previously introduced by Maity and Johansson [MaiJoh2002]. We pointed out that, in their work Maity and Johansson omitted a criterion on the set, S, they use in the construction. Moreover, we give a complete description of the construction method – together with the criterion missed by Maity and Johansson – and prove that any *t*-resilient Boolean function constructed by this technique satisfies the upper bound on nonlinearity, studied in Section 2.1. We also proved that, if $t > n/2$ - 2, the construction technique yields a nonlinearity of $nl_{f'}$, where $2^{n-1} - 2^{n/2-1} - 2^{t+1} < nl_{f'} \leq 2^{n-1} - 2^{t+1}$. We show that without the introduction of the fourth criterion, one cannot guarantee the resiliency of the function *f*.

The construction method of Chapter 3 is a method, which starts with a bent function and modifies it at a set of pre-selected points. The strength of this method is that, the selection of the set, S, depends on the required order of resiliency and therefore any order of resiliency can be obtained using this method. On the other hand, obtaining the set, S, has no obvious algorithm and is a major weakness of this method (see Appendix A).

We have also studied two more construction techniques achieving the nonlinearity bound of Sarkar and Maitra, in Chapter 4. One of these methods was by Sarkar and Maitra, and uses a Maiorana-McFarland like technique to construct a Boolean function having the maximum order of resiliency for given number of variables. We have pointed that the intention to maximize order of resiliency distorts other cryptographic parameters – like the autocorrelation characteristics – of the function, as we have investigated in Chapter 5. Also it is impossible to construct functions having small order of resiliency – like the one constructed in Chapter 3 – using this method.

The other construction method studied in Chapter 4 is by Tarannikov and emphasizes the computational complexity. The author starts with functions of less number of variables and adds new variables linearly to the initial function to have resilient functions of large number of variables. This method again produces functions with high order of resiliency, namely functions for which $t \geq \frac{2n-7}{3}$. The most important property of this construction technique is its advantageous computational complexity, although it cannot produce functions with small order of resiliency.

In the final chapter of the thesis, we investigated the functions that have been constructed using the methods introduced in Chapters 3 and 4, in terms of their auto-correlation characteristics, pointing the fact that although auto-correlation is an important parameter against differential attacks, it has not attracted attention in the literature for resilient functions. We have computed the spectrum, and auto-correlation function of each constructed function and calculated the absolute indicator and sum-of-squared-errors for each of them.

REFERENCES

[CamCarCS1991] P. Camion, C. Carlet, P. Charpin and N. Sendrier, On correlation-immune functions, In Advances in Cryptology – Crypto’91 Lecture Notes in Computer Science, V. 576, pages 86-100, 1991.

[CarDin2004] C. Carlet and C. Ding, Highly nonlinear mappings, Journal of Complexity 20(2-3), pages 205-244, 2004.

[Carl1997] C. Carlet, More correlation immune and resilient functions over Galois fields and Galois rings, In Advances in Cryptology – EUROCRYPT’97, pages 422-433, 1997.

[CheLeeLS1996] S. Chee, S. Lee, D. Lee and S. H. Sung, On the correlation immune functions and their nonlinearity, In Advances in Cryptology, Asiacrypt 96, Lecture Notes in Computer Science, V. 1163, pages 232-243, 1996.

[FedTar2001] M. Fedorova and Y. V. Tarannikov, On the constructing of highly nonlinear resilient Boolean functions by means of special matrices, In Progress in Cryptology – INDOCRYPT 2001, Lecture Notes in Computer Science, V. 2247, pages 254-266, 2001.

[GouMas1988] X. Guo-Zhen and J. Massey, A spectral characterization of correlation immune combining functions, IEEE Transactions on Information Theory, 34(3), pages 569-571, 1988.

[GupSar2003] K. C. Gupta and P. Sarkar, Computing Walsh Transform from the Algebraic Normal Form of a Boolean Function, http://eprint.iacr.org/2003/040, 2003.

[Heys2000] H. M. Heys, A tutorial on linear and differential cryptanalysis, http://www.engr.mun.ca/~howard/PAPERS/ldc_tutorial.pdf, 2000.

[JohPas2003] T. Johansson and E. Pasalic, A construction of resilient functions with high nonlinearity, IEEE Transactions on Information Theory, V. 49, pages 494-501, 2003.

[KavYüc2003] S. Kavut and M. D. Yücel, Improved Cost Function in the Design of Boolean Functions Satisfying Multiple Criteria, In Progress in Cryptology – INDOCRYPT 2003, pages 121-134, 2003.

[MaiJoh2002] S. Maity and T. Johansson, Construction of cryptographically important Boolean functions, In Progress in Cryptology – INDOCRYPT 2002, pages 234-245, 2002.

[Mait2000] S. Maitra, Correlation Immune Boolean Functions with Very High Nonlinearity, http://eprint.iacr.org/2000/054/, 2000.

[PasJoh1999] E. Pasalic and T. Johansson, Further results on relation between nonlinearity and resiliency for Boolean functions, IMA Conference on Cryptography and Coding, Lecture Notes in Computer Science, V. 1746, pages 35-44, 1999.

[PasJohMS2001] E. Pasalic, T. Johansson, S. Maitra, and P. Sarkar, New constructions of resilient and correlation immune Boolean functions achieving upper bounds on nonlinearity, In Workshop on Coding and Cryptography, Electronic Notes in Discrete Mathematics, 2001.

[Porw2003] P. Porwik, The Spectral Test of the Boolean Function Linearity, International Journal of Applied Mathematics and Computer Science, V. 13, No. 4, pages 567-575, 2003.

[Roth1976] O. S. Rothaus, On Bent Functions, Journal of Combinatorial Theory, Series A, V. 20, pages 300-305, 1976.

[SarMai1999] P. Sarkar and S. Maitra, Construction of nonlinear resilient Boolean functions, Technical Report No. ASD/99/30, Indian Statistical Institute, 1999.

[SarMai2000-1] P. Sarkar and S. Maitra, Nonlinearity bounds and constructions of resilient Boolean functions, Advances in Cryptology - CRYPTO 2000, Lecture Notes in Computer Science, V. 1880, pages 515-532, 2000.

[SarMai2000-2] P. Sarkar and S. Maitra, Highly nonlinear balanced Boolean functions with important cryptographic properties. Advances in Cryptology - EUROCRYPT2000, Lecture Notes in Computer Science, V. 1807, pages 485-506, 2000.

[SarMai2000-3] P. Sarkar and S. Maitra, Construction of nonlinear Boolean functions with important cryptographic properties, In Advances in Cryptology – EUROCRYPT 2000, Lecture Notes in Computer Science, V. 1807, pages 491-512, 2000.

[SebZhaZ1994] J. Seberry, X. M. Zhang and Y. Zheng, On constructions and nonlinearity of correlation immune Boolean functions, In Advances in Cryptology – CRYPTO'93, pages 181-199, 1994.

[Sieg1984] T. Siegenthaler, Correlation-immunity of nonlinear combining functions for cryptographic applications, IEEE Transactions on Information Theory, IT-30 (5), pages 776-780, 1984.

[Tara2000] Y. Tarannikov, On Resilient Boolean Functions with Maximal Possible Nonlinearity, In Progress in Cryptology - INDOCRYPT 2000, pages 19-30, 2000.

[Tara2001] Y. Tarannikov. New constructions of resilient Boolean functions with maximal nonlinearity. FSE 2001, pages 66-77, 2001.

[Yüce2001] M. D. Yücel, Alternative Nonlinearity Criteria for Boolean Functions. Electrical and Electronics Engineering Departmental Memorandum, No.2001-1, Middle East Technical University (METU), 2001.

[ZhaZhe1995] X. M. Zhang and Y. Zheng, GAC–the criterion for global avalanche characteristics of cryptographic functions, Journal for Universal Computer Science, 1(5), pages 316-333, 1995.

APPENDIX A

SEARCH FOR THE SET S

Although we have selected elements of the set S, used in Chapter 3, by inspection, there is not a certain way to select elements of S. Therefore, we have tried searching for a suitable set S satisfying the criteria listed in Chapter 3.

Note that since the initial function is bent, there is a difference of $2^{n/2}$ between the number of ones and zeros and therefore the set S should at least contain $2^{n/2-1}$ elements to satisfy balancedness. Additional elements should have outputs with same number of ones and zeros, which implies the set S has even number of elements. Moreover, using the set S or its complement $\overline{\mathrm{S}}$ gives the same function except an inversion. Therefore, choosing S or $\overline{\mathrm{S}}$ is the same in terms of nonlinearity, resiliency and auto-correlation, which sets an upper limit to the number of elements of S to be 2^{n-1}.

A.1 Search Starting with an Initial Set

In this algorithm we first select an initial set S_1 and then search for a second set S_2 in order to form $S = S_1 \cup S_2$. We first select S_1 to be {1, 2, 4, 8, 16, 32, 64, 128}, where the numbers represent the decimal equivalent of the input vectors. Note that this is the same set selected in Chapter 3.

Now we run an algorithm shown in Figure A-1 to form an S_2 such that $S = S_1 \cup S_2$ satisfies our criteria. Our approach is to use a pair of inputs, with one element being the complement of the other. Since the number of elements to be searched is small (128 for *n* = 8) this is a deterministic search.

In order to organize the search in a systematic manner, in other words, to finish the search after the first 128 tries, we take the input vectors to the algorithm in the order of gray code, excluding the vectors {(0,0,0,0,0,0,0,1), (0,0,0,0,0,0,1,0), (0,0,0,0,0,1,0,0), (0,0,0,0,1,0,0,0), (0,0,0,1,0,0,0,0), (0,0,1,0,0,0,0,0), (0,1,0,0,0,0,0,0), (1,0,0,0,0,0,0,0)} since they are already included in S_1.

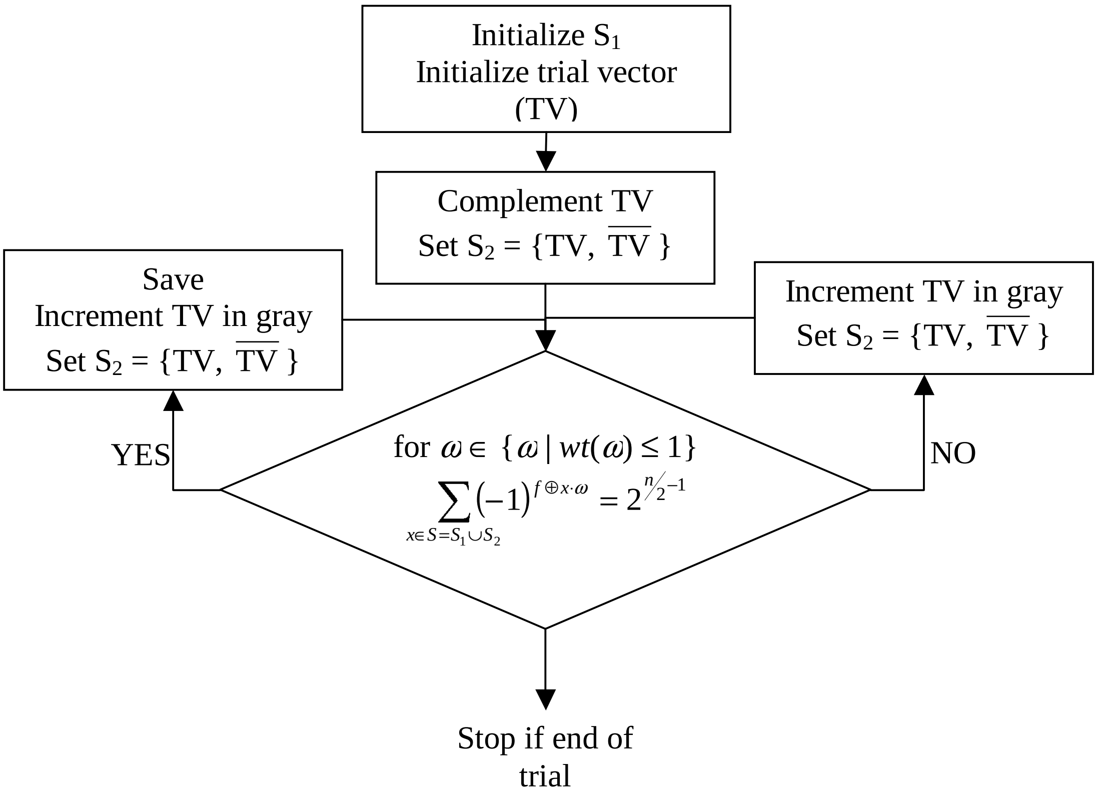
Initialize S_1
Initialize trial vector (TV)
Complement TV
Set $S_2 = \{TV, \overline{TV}\}$
Save
Increment TV in gray
Set $S_2 = \{TV, \overline{TV}\}$
Increment TV in gray
Set $S_2 = \{TV, \overline{TV}\}$
YES
NO
for $\omega \in \{\omega \mid wt(\omega) \le 1\}$
$\sum_{x \in S = S_1 \cup S_2} (-1)^{f \oplus x \cdot \omega} = 2^{n/2 - 1}$
Stop if end of trial

Figure A-1. Search Algorithm Starting with an Initial Set

This search finds only S_2 = {0, 255}, which is the one used in Chapter 3. There is no other S to be constructed with this search.

Instead of S_1 above, if another set $S_1{}'$ = {127, 191, 223, 239, 247, 251, 253, 254} (the complemented versions of vectors of S_1) is selected initially, the search does not give any result satisfying our criteria.

A.2 Random Search

Another algorithm used to find a suitable S set is randomized search for all elements of the set. Here we have used two basic approaches:

1. Set number of variables of S and randomly select elements.

2. Start with a random set of size $2^{n/2-1}$, randomly select new elements, incrementing the number of elements after a certain number of trials.

By randomized search we have tried over 820 million sets for number of elements 8, 10, 12 and up to 32 elements with the incremental search and not been successful in finding any other S set.

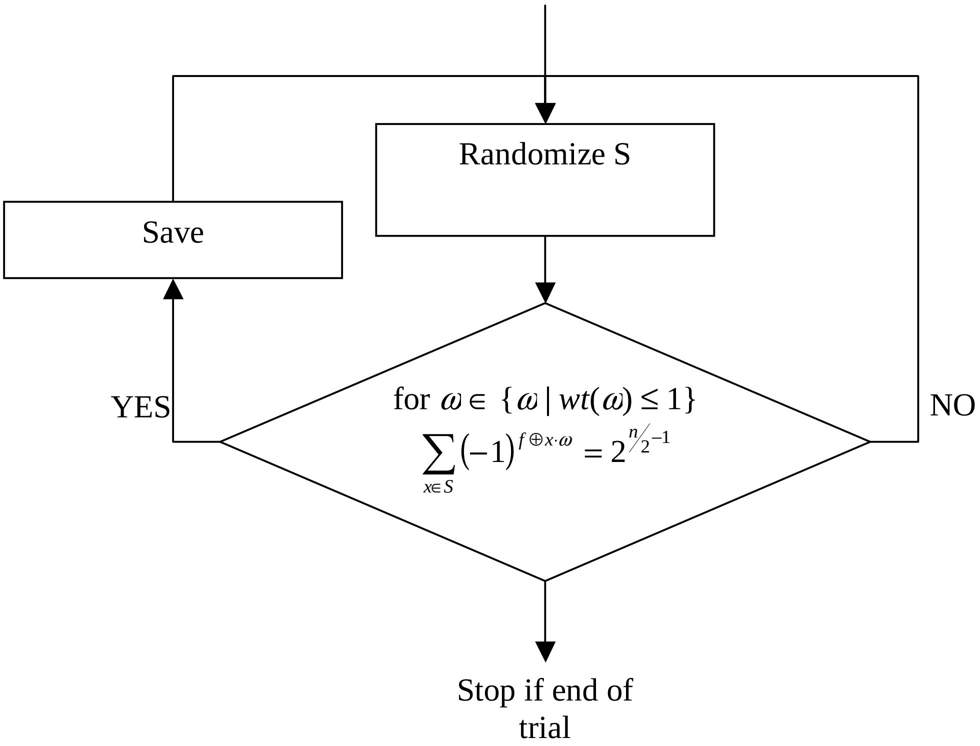
Randomize S
Save
YES
for $\omega \in \{\omega \mid wt(\omega) \le 1\}$
$\sum_{x\in S}(-1)^{f\oplus x\cdot\omega} = 2^{n/2-1}$
NO
Stop if end of trial

Figure A-2. Random Search

Figure A-2 and Figure A-3 show the two algorithms that apply randomized search. Modifying the number of variables of S enlarges the search space, but also slows down the computation.

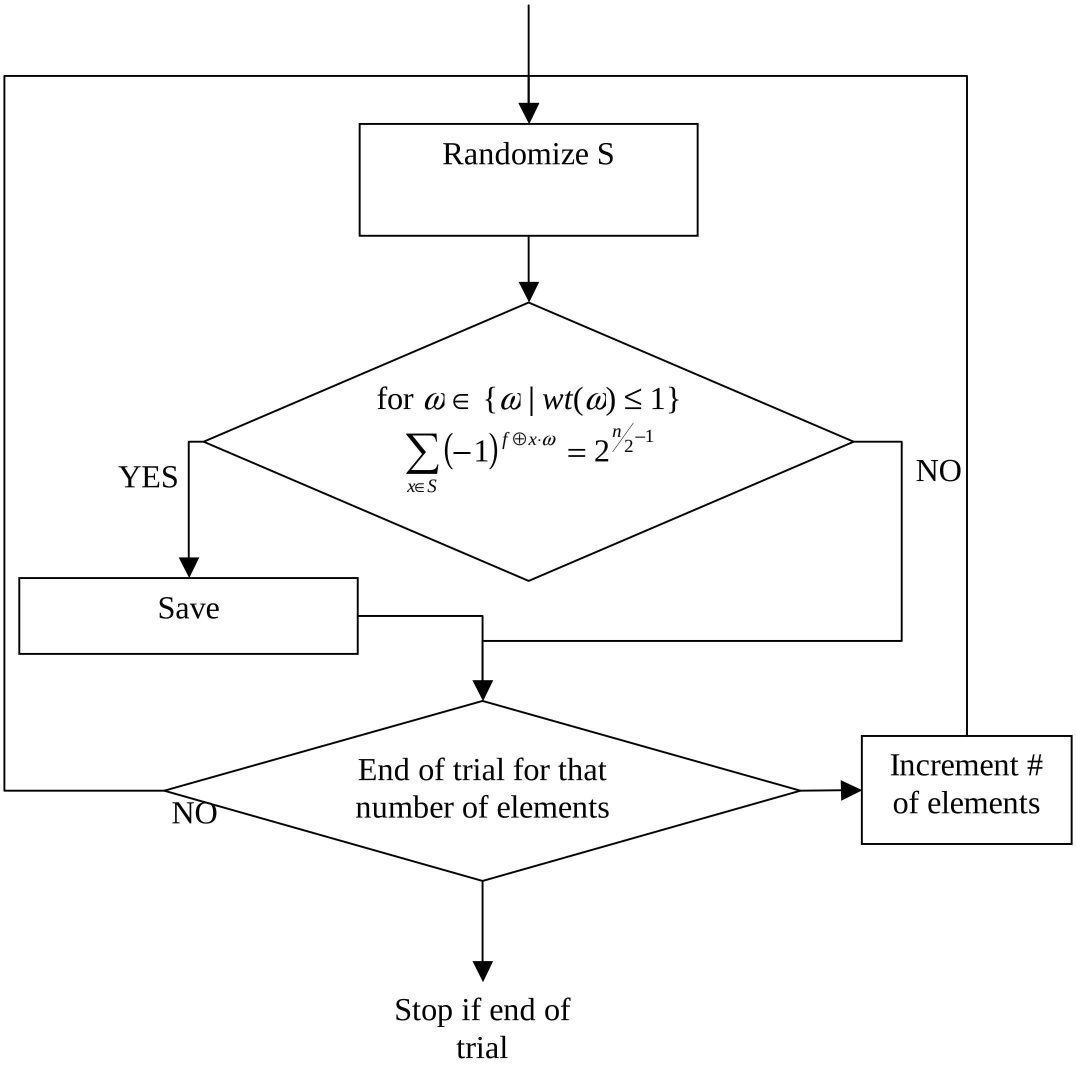
Randomize S
for $\omega \in \{\omega \mid wt(\omega) \leq 1\}$
$\sum_{x \in S} (-1)^{f \oplus x \cdot \omega} = 2^{n/2 - 1}$
YES
NO
Save
End of trial for that number of elements
NO
Increment # of elements
Stop if end of trial

Figure A-3. Random Search with Variable Number of Elements

A.3 Further Discussions

The two basic search methods we have investigated, the one with an initial set and randomized search, does not give many sets that satisfy the required criteria.

The only set obtained is the one obtained by the search starting with an initial set and is same as the one used in Chapter 3. Although this method decreases the number of elements to be searched by selecting some elements initially, still there is no obvious way to form the initial set. However, forming S_1 from the vector of weight one, for example, as done in this work may be a wise guess. The advantage here is the possibility to perform a deterministic search since the search space is small.

The random search on the other hand is performed in a very large search space. For example, there are 2.79 x 10^{17} possible sets for n = 8 and |S| = 10. Moreover, the number of elements of S may have many other values rather than 10 for n = 8 and this number increases exponentially with increasing n.

We have tried 820 million different sets with number of variables 8, 10 and 12, which takes days to complete such a search on a computer with 2800 GHz processor. Note that this number of trials, even though it is large and requires a very long time, is very small compared to the number of possible sets.