NONLINEARITY PRESERVING POST-TRANSFORMATIONS

İSA SERTKAYA

JUNE 2004

NONLINEARITY PRESERVING POST-TRANSFORMATIONS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

İSA SERTKAYA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF CRYPTOGRAPHY

JUNE 2004

Approval of the Graduate School of Applied Mathematics

Prof. Dr. Aydın AYTUNA
Director

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

Prof. Dr. Ersan AKYILDIZ
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. Ali DOĞANAKSOY
Supervisor

Examining Committee Members

Prof. Dr. İsmail Ş. GÜLOĞLU _____

Prof. Dr. Ersan AKYILDIZ _____

Assoc. Prof. Dr. Ali DOĞANAKSOY _____

Assist. Prof. Dr. Ali A. SELÇUK _____

Dr. Emrah ÇAKÇAK _____

ABSTRACT

NONLINEARITY PRESERVING POST-TRANSFORMATIONS

SERTKAYA, İsa
M.Sc., Department of Cryptography
Supervisor: Assoc. Prof. Dr. Ali DOĞANAKSOY

June 2004, 61 pages

Boolean functions are accepted to be cryptographically strong if they satisfy some common pre-determined criteria. It is expected that any design criteria should remain invariant under a large group of transformations due to the theory of similarity of secrecy systems proposed by Shannon. One of the most important design criteria for cryptographically strong Boolean functions is the nonlinearity criterion. Meier and Staffelbach studied nonlinearity preserving transformations, by considering the invertible transformations acting on the arguments of Boolean functions, namely the pre-transformations. In this thesis, first, the results obtained by Meier and Staffelbach are presented. Then, the invertible transformations acting on the truth tables of Boolean functions, namely the post-transformations, are studied in order to determine whether they keep the nonlinearity criterion invariant. The equivalent counterparts of Meier and Staffelbachs results are obtained in terms of the post-transformations. In addition, the existence of nonlinearity preserving post-transformations, which are not equivalent to pre-transformations, is proved. The necessary and sufficient conditions for an affine post-transformation to preserve nonlinearity are proposed and proved. Moreover, the sufficient conditions for an non-affine post-transformation to keep nonlinearity invariant are proposed. Furthermore, it is proved that the smart hill climbing method, which is introduced to improve nonlinearity of Boolean func-

tions by Millan et. al., is equivalent to applying a post-transformation to a single Boolean function. Finally, the necessary and sufficient condition for an affine pre-transformation to preserve the strict avalanche criterion is proposed and proved.

Keywords: Cryptography, Boolean functions, Balancedness, Nonlinearity, Strict Avalanche Criterion, Propagation Criterion, Correlation Immunity, Hadamard matrices, Sylvester-Hadamard matrices, Walsh-Hadamard transformation, Pre-transformations, Post-transformations, Smart Hill Climbing Method

ÖZ

NONLİNEERİTEYİ KORUYAN ARD DÖNÜŞÜMLER

SERTKAYA, İsa
Yüksek Lisans, Kriptografi Bölümü
Tez Yöneticisi: Doç. Dr. Ali DOĞANAKSOY

Haziran 2004, 61 sayfa

Boole fonksiyonları önceden belirlenmiş bir takım kriterleri sağladığında kriptografik açıdan sağlam kabul edilir. Shannon'ın gizlilik sistemlerinin benzerliği teorisi gereği, her tasarım kriterinin büyük bir dönüşüm grubu altında sabit kalması beklenir. Kriptografik açıdan sağlam Boole fonksiyonları için en önemli tasarım kriterlerinden birisi nonlineerite kriteridir. Meier ve Staffelbach, Boole fonksiyonlarının argümanları üzerinde tanımlı tersinir dönüşümleri, yani "ön dönüşümleri", göz önünde bulundurarak nonlineeriteyi koruyan dönüşümleri incelemişlerdir. Bu tezde, önce, Meier ve Staffelbach'ın elde ettikleri neticeler takdim olunmuştur. Müteâkiben, Boole fonksiyonlarının doğruluk tabloları üzerinde tanımlı tersinir dönüşümler, nam-ı diğer "ard dönüşümler", nonlineeriteyi sabit bırakmaları açısından çalışılmıştır. Meier ve Staffelbach'ın tarafından verilen sonuçların ard dönüşümler cinsinden mütekabil karşılıkları elde edilmiştir. Buna ek olarak, herhangi bir ön dönüşüme eşdeğer olmayan ve nonlineeriteyi koruyan ard dönüşümlerin varlığı da ispatlanmıştır. Afin bir ard dönüşümün nonlineeriteyi koruması için gerek ve yeter şartlar önerilmiş ve ispatlanmıştır. Dahası, afin olmayan ard dönüşümlerin nonlineeriteyi koruması için bazı yeter şartlar da önerilmiştir. Üstelik, Millan ve arkadaşları tarafından Boole fonksiyonlarının nonlineeritesini yükseltmek amacıyla tanımladıkları "Smart Hill Climbing" metodunun, bir Boole fonksiyonuna bir ard dönüşüm uygulamaktan ibaret olduğu gösterilmiştir. Son olarak da, afin ön dönüşümlerin keskin çığ kriterini

sabit bırakması için gerek ve yeter şart önerilmiş ve ispatlanmıştır.

Anahtar Kelimeler: Kriptografi, Boole fonksiyonları, Dengelilik, Nonlineerite, Keskin çığ kriteri, Yayılma kriteri, Korelasyon muafiyeti, Hadamard matrisleri, Sylvester-Hadamard matrisleri, Walsh-Hadamard dönüşümü, Ön dönüşümler, Ard dönüşümler, Smart Hill Climbing Metodu

ACKNOWLEDGMENTS

My first, and the most earnest, acknowledgment must go to my supervisor Assoc. Prof. Dr. Ali Doğanaksoy not only for patiently guiding, motivating, and encouraging me throughout this study, but also for being instrumental in ensuring my academic, professional and moral wellbeing ever since. In every sense, none of this work would have been possible without him.

I owe a huge debt of gratitude to Prof. Dr. İsmail Güloğlu whom I have had the honor to talk many times, for his excellent suggestions and comments on this thesis, especially for the Remark 3.2.11 in Chapter 3.

I also wish to express my deep gratitude to Kudret Özkal, for her patience and loving encouragement, especially for being with me all the way.

I also would like to thank to all people in Boolean functions studying group at IAM, especially to Dr. Muhiddin Uğuz for his valuable comments and suggestions.

It is a pleasure to thank also Serhat Sağdıçoğlu and Koray Karabina for helping me while preparing this manuscript and their comments on the manuscript.

Special thanks must go to my colleagues at UEKAE for creating an enjoyable and stimulating atmosphere. In particular, I want to acknowledge my managers Önder Yetiş and Alparslan Babaoğlu for their patience and support.

Last but certainly not the least, warm thanks go to my family for their incredible love, help, and extraordinary patience, especially to my parents for their never-ending support and belief in what I do, not only during the time spent on this thesis, but also throughout the many years of education which preceded it. They deserve far more credit than I can ever give them.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Cryptographic mappings, in particular iterated block ciphers, are often called "strong" when they satisfy certain cryptographic criteria. Two main concepts, *confusion* and *diffusion* were suggested by Shannon in [22]. The principle of confusion can be thought as designing a cryptosystem so as the dependence of the key on the plaintext and the ciphertext should be complex enough to make cryptanalysis unsuccessful. The principle of diffusion can be stated as designing the cryptosystem so that there should be no statistical dependence between the simple structures in the plaintext and the simple structures in the ciphertext, in other words, statistical properties of the plaintext should be dissipated into each component of the ciphertext. In order to respect these principles, several criteria are proposed to be satisfied by cryptosystems.

H. Feistel, in [5], proposed that a cryptosystem should possess *avalanche effect*. That is, when a single input bit is complemented, half of the output bits should change. In 1979, Kam and Davida ([12]) pointed out that a cryptosystem should be *complete*, that is, each output bit should depend on each input bit. Later, in 1985, Webster and Tavares in [28], combined these two criteria into one. Namely, a cryptographic transformation is said to satisfy *strict avalanche criterion* if each output bit changes with the probability one half, whenever a single input bit is complemented.

Due to the computational infeasibility of testing these criteria, in the late 80's, they were redefined for the core components of cryptosystems, particularly for *Boolean functions* which are in fact transformations that map binary n-tuples to 0 or 1.

In 1988, R. Forré ([6]), showed the way of checking the strict avalanche criterion by using Walsh transform of Boolean functions, and introduced a generalization for strict avalanche criterion. This generalization is stated as follows. A Boolean function is said to satisfy *strict avalanche criterion of order* m, if it satisfies strict avalanche criterion of order $m-1$ and if any function obtained from the Boolean function by fixing its m input bits constant, satisfies strict avalanche criterion as well. In this generalization, the strict avalanche criterion is equivalent to the strict avalanche criterion of order 0. On the other hand, Preneel et. al., proposed another generalization in [19] namely, the propagation criterion of degree k. A Boolean function satisfies the *propagation criterion of degree* k, if the outputs of the Boolean function changes with a probability one half whenever at most k bits of the input bits are complemented. In this point of view, strict avalanche criterion coincides with propagation criterion of degree 1.

Another criterion, that a cryptographically good Boolean function should possess, is correlation immunity, which was introduced by Siegenthaler ([26]). A Boolean function, whose output distribution probability is unchanged when any m input bits are kept constant, is called *m-th order correlation immune*.

The criteria mentioned up to now are proposed as a way of satisfying diffusion. However, the concept of confusion is more important, since otherwise the cryptosystem can be easily broken. *Nonlinearity* which was introduced by Pieprzyk and Finkelstein in [18], is the most important criterion for any cryptosystem, and

the purpose of nonlinearity criterion is to achive confusion. The most widely used nonlinearity criterion is defined as the minimum distance of the Boolean function to the affine Boolean functions.

In [22], Shannon proposed similarity of secrecy systems. Two cryptosystems are said to be similar if by applying a transformation to one of them, one gets the other. Then, a cryptosystem is weak if it can be mapped by a simple transformation to a weak cryptosystem. Therefore, only proposing the design criteria is not enough to test a cryptosystem's strength. The proposed criteria should remain invariant under simple transformations. Meier and Staffelbach in [15] investigated under which transformation these criteria preserved by considering the invertible transformations acting on the inputs of Boolean functions, which we call *pre-transformations*.

In this thesis, after presenting results obtained in [15], by considering the invertible transformations acting on the truth tables of Boolean functions, which we call *post-transformations*, we analyze these criteria, especially the nonlinearity criterion. We wish to note that, the transformations we consider belong only to a small group of post-transformations. We obtain counterparts of the previous results in terms of the post-transformations. We propose and prove the necessary and sufficient conditions for an affine post-transformations to preserve nonlinearity. Moreover, we give some remarks on non-affine post-transformations to keep nonlinearity invariant. Particularly, we present sufficient conditions for an non-affine post-transformation to preserve nonlinearity. Furthermore, we prove that the hill climbing method, which was introduced to find highly nonlinear Boolean functions in [16], can be viewed as applying a post-transformation to a single Boolean function. Finally, we present some remarks for the other design criteria.

The rest of the thesis is organized as follows

In Chapter 2, we establish the notations that we use, and recall the properties of Boolean functions. Then, we state the characteristics of Hadamard matrices, and their automorphism groups. Later, we present the Walsh-Hadamard transform of Boolean functions, its properties and relations with Sylvester-Hadamard matrices. By stating the common design criteria mentioned above and revising useful theorems to compute them, we end this chapter.

We devote the Chapter 3 to the analysis of the pre-transformations and the post-transformations. We first recall the results proposed in [15], then with the help of post-transformations we obtain those results again and classify which post-transformations correspond linear and affine pre-transformations. We propose and prove the necessary and sufficient conditions on affine post-transformations for preserving nonlinearity criterion. We give sufficient conditions for non-affine post-transformations to keep nonlinearity invariant.

In Chapter 4, we shall present an application of post-transformation, that is smart hill climbing, which was proposed by Millan et. al. in [16].

We end the thesis in Chapter 5, by presenting a summary of the thesis, and giving some remarks on the other design criteria and discussing some possible future work.

CHAPTER 2

PRELIMINARIES

In this chapter, we are going to state the definitions and the notations that we use in the following chapters. Since we do not give detailed explanation for the most of the definitions and notations, the reader may refer to [21] for further information.

2.1 Boolean Functions

Let $\mathcal{V}_n$ be the vector space which is composed of all n-tuples of elements from $GF(2)$. An element $\alpha_k = (a_1, a_2, \ldots, a_n)$ in $\mathcal{V}_n$, can be represented by the integer $k = \sum_{i=1}^{n} a_i 2^{n-i}$. In this representation, $\mathcal{V}_n$ assumes a natural ordering called **lexicographic ordering**. For $\alpha, \beta \in \mathcal{V}_n$, the sum $\alpha \oplus \beta \in \mathcal{V}_n$ is obtained by adding corresponding components of α and β modulo 2. The standard basis of $\mathcal{V}_n$ is denoted by $\{e_1, e_2, \ldots, e_n\}$, where e_i stands for the vector having all zero's except a 1 at the i-th position.

The **Hamming weight** of an element $\alpha \in \mathcal{V}_n$ is the number of components that are equal to 1 and is denoted by $w(\alpha)$. The **Hamming distance** between two elements $\alpha, \beta \in \mathcal{V}_n$ is the number of components in which they differ and is denoted by $d(\alpha, \beta)$. Obviously, $d(\alpha, \beta)$ is the Hamming weight of $\alpha \oplus \beta$. From now on by "the weight" and "the distance" we shall mean the Hamming weight and

the Hamming distance, respectively. Obviously, if the weight of each $\alpha, \beta \in \mathcal{V}_n$ is even, then $d(\alpha, \beta)$ is also even.

Let $\alpha = (a_1, a_2, \ldots, a_n)$, $\beta = (b_1, b_2, \ldots, b_n) \in \mathcal{V}_n$. The **standard inner product** on $\mathcal{V}_n$ is denoted by $\langle , \rangle$ which is defined as

$$\langle \alpha, \beta \rangle = \sum_{i=1}^{n} a_i b_i. \tag{2.1.1}$$

A **Boolean function** is a map, with domain $\mathcal{V}_n$, that takes values from $GF(2)$. The set of all Boolean functions is denoted by $\mathcal{F}_n$. It is easy to show that $\mathcal{F}_n$ is isomorphic to $\mathcal{V}_{2^n}$, and hence $|\mathcal{F}_n| = 2^{2^n}$. From now on, unless otherwise stated explicitly, by "a function" we mean a Boolean function in $\mathcal{F}_n$

Any function $f \in \mathcal{F}_n$ can be uniquely represented by different forms:

- The ordered tuple,

  $$(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$$

  is called the **truth table** of f and denoted by T_f.

- In some cases, instead of the truth table of f, it may be more convenient to use the real valued function, namely the **sign** function $\hat{f}$, of f. It is defined as $\hat{f}(\alpha) = (-1)^{f(\alpha)} = 1 - 2f(\alpha)$ for all $\alpha \in \mathcal{V}_n$. The truth table of the sign function $\hat{f}$ is called the **sequence** of f and is denoted by ζ_f. That is

  $$\zeta_f = ((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-1)^{f(\alpha_{2^n-1})}).$$

- The multivariate polynomial in $GF(2)[x_1, x_2, \ldots, x_n]$ of the form

$$f(x_1, x_2, \ldots, x_n) = \sum_{\alpha=(a_1,a_2,\ldots,a_n)\in\mathcal{V}_n} \lambda_\alpha x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$$

where $\lambda_\alpha \in GF(2)$. By considering the above statement as a function over $GF(2)$, one gets:

$$f(x_1, x_2, \ldots, x_n) = c_0 \oplus c_1 x_1 \oplus \cdots \oplus c_n x_n \oplus c_{12} x_1 x_2 \oplus \cdots \oplus c_{12\cdots n} x_1 x_2 \cdots x_n$$

where $c_0, c_1, \ldots, c_{12\cdots n} \in GF(2)$. This polynomial representation is called the **algebraic normal form** of f. The degree of the algebraic normal form of f is called the **degree** of f, and is denoted by $deg(f)$. In the literature, $deg(f)$ is sometimes called as the "nonlinearity order" of f.

For any $f \in \mathcal{F}_n$, by the weight of f, we mean the weight of its truth table T_f on $\mathcal{V}_{2^n}$. If the weight of a function is 2^{n-1}, i.e. the numbers of 0's and 1's are equal, then the function is called a **balanced** function. We denote the set of all balanced functions by $\mathcal{B}_n$. Obviously, $|\mathcal{B}_n| = \binom{2^n}{2^{n-1}}$.

The sum $f \oplus g$ of functions $f, g \in \mathcal{F}_n$, is defined by setting $(f \oplus g)(\alpha) = f(\alpha) \oplus g(\alpha)$ for all $\alpha \in \mathcal{V}_n$. Let $f, g \in \mathcal{F}_n$. Then by the distance between f and g, we mean the distance between T_f and T_g on $\mathcal{V}_{2^n}$. Thus, $d(f, g) = w(f \oplus g)$. Now, we can state the following important lemma which will be used in Section 2.4 while defining the nonlinearity criterion of a Boolean function.

Lemma 2.1.1. ([23]) *For any $f, g \in \mathcal{F}_n$, $d(f, g) = 2^{n-1} - \frac{1}{2} \langle \zeta_f, \zeta_g \rangle$.*

Two constant functions, whose weights are equal to 0 and 2^n will be denoted by 0_n and 1_n, respectively. For any function $f \in \mathcal{F}_n$, the **complement function** $\bar{f}$ is defined to be $\bar{f} = f \oplus 1_n$. Trivially, it follows that $w(\bar{f}) = 2^n - w(f)$.

Consequently, f is balanced if and only if $\bar{f}$ is balanced.

For a function $f \in \mathcal{F}_n$, the **support** of f is defined to be the set $\{\alpha \in V_n | f(\alpha) = 1\}$ and is denoted by $Supp(f)$. It is obvious that $|Supp(f)| = w(f)$ and that $Supp(f) \cap Supp(\bar{f}) = \emptyset$.

For a function $f \in \mathcal{F}_n$, if $f(\alpha \oplus \beta) = f(\alpha) \oplus f(\beta)$ holds for all $\alpha, \beta \in \mathcal{V}_n$ then f is called a **linear** function and such a function is of the form $f(x) = a_1x_1 \oplus a_2x_2 \oplus \cdots \oplus a_nx_n$, $a_i \in GF(2)$. The set of all linear functions is denoted by $\mathcal{L}_n$.

A function $f \in \mathcal{F}_n$ that is of the form $f(x) = a_0 \oplus a_1x_1 \oplus a_2x_2 \oplus \cdots \oplus a_nx_n$, $a_i \in GF(2)$, is called an **affine** function, and the class of all affine functions is denoted by $\mathcal{A}_n$.

Some properties of linear and affine functions are given below. These results are not novel but for the sake of completeness, we include some of the proofs.

It is easy to show that $\mathcal{L}_n \subset \mathcal{A}_n$ and $2|\mathcal{L}_n| = |\mathcal{A}_n| = 2^{n+1}$.

Theorem 2.1.2. ([*21*]) *For any linear function $f \in \mathcal{L}_n$, there exists a unique vector $\alpha \in \mathcal{V}_n$ such that $f(\beta) = \langle \alpha, \beta \rangle$ for all $\beta \in \mathcal{V}_n$.*

Hereinafter, for any $\alpha = (a_1, a_2, \ldots, a_n) \in \mathcal{V}_n$, by f_α, we denote the linear function $f_\alpha : \beta \to \langle \alpha, \beta \rangle$, for all $\beta \in \mathcal{V}_n$, and ℓ_k will stand for the sequence of $\hat{f}_{\alpha_k}$, where $k = \sum_{i=1}^{n} a_i 2^{n-i}$.

Theorem 2.1.3. ([*21*]) *Every non-constant affine function is balanced.*

Theorem 2.1.4. ([*21*]) *Let $f_{\alpha_i}, f_{\alpha_j} \in \mathcal{L}_n$, with $0 \le i \le j \le 2^n - 1$, we have the following:*

$$d(f_{\alpha_i}, f_{\alpha_j}) = \begin{cases} 0 & if \quad \alpha_i = \alpha_j, \\ 2^{n-1} & otherwise. \end{cases} \tag{2.1.2}$$

Proof: Consider the sequences ℓ_i, ℓ_j of $f_{\alpha_i}, f_{\alpha_j}$,respectively. Then, we have the following:

$$\begin{aligned}
\langle \ell_i, \ell_j \rangle &= \sum_{x \in \mathcal{V}_n} (-1)^{\langle \alpha_i, x \rangle \oplus \langle \alpha_j, x \rangle} \\
&= \sum_{x \in \mathcal{V}_n} (-1)^{\langle \alpha_i \oplus \alpha_j, x \rangle} \\
&= \begin{cases} 2^n & if \quad \alpha_i = \alpha_j \\ 0 & if \quad \alpha_i \neq \alpha_j \end{cases} \\
&= 2^n \delta(\alpha_i + \alpha_j). \qquad (2.1.3)
\end{aligned}$$

where $\delta(\omega)$ is the **Kronecker delta function** which is equal to 1 if ω is the zero vector and 0 otherwise.

Moreover, from Lemma 2.1.1, we know that $d(f_{\alpha_i}, f_{\alpha_j}) = 2^{n-1} - \frac{1}{2} \langle \ell_i, \ell_j \rangle$. Therefore, for $\alpha_i = \alpha_j$,

$$\begin{aligned}
d(f_{\alpha_i}, f_{\alpha_j}) &= 2^{n-1} - \frac{1}{2} \langle \ell_i, \ell_i \rangle \\
&= 2^{n-1} - \frac{1}{2} \cdot 2^n \\
&= 0.
\end{aligned}$$

Similarly, for $\alpha_i \neq \alpha_j$,

$$\begin{aligned}
d(f_{\alpha_i}, f_{\alpha_j}) &= 2^{n-1} - \frac{1}{2} \langle \ell_i, \ell_j \rangle \\
&= 2^{n-1} - \frac{1}{2} \cdot 0 \\
&= 2^{n-1}
\end{aligned}$$

which completes the proof. □

The set $\{\ell_0, \ell_1, \ldots, \ell_{2^n-1}\}$ of sequences of all linear functions, forms an orthogonal basis for $\mathbb{R}^{2^n}$ over the set of real numbers $\mathbb{R}$ with respect to the standard inner product on $\mathbb{R}^{2^n}$. It follows that the sign function of any function can be uniquely written as a linear combination of $\ell_0, \ell_1, \ldots, \ell_{2^n-1}$ ([21]).

2.2 Sylvester-Hadamard Matrices

Definition 2.2.1. ([9]) An $n \times n$ matrix H with all entries 1 or -1 is called a **Hadamard** matrix if $H \cdot H^t = nI_n$, where H^t is the transpose of H and I_n is the $n \times n$ identity matrix.

Jacques Hadamard introduced these matrices in 1893 as solutions to an extremal problem in analysis. An enormous literature concerned with both the applications and combinatorial aspects of Hadamard matrices now exists. However, we only focus on a particular class of the Hadamard matrices, known as "Sylvester-Hadamard matrices" or "Sylvester matrices".

Definition 2.2.2. ([14]) If $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is $p \times q$ matrix, then the **Kronecker product** of A and B is the $mp \times nq$ matrix given as follows:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

where $a_{ij}B$ denotes the $p \times q$ matrix obtained by multiplying each entry of B with a_{ij}.

For extra information about the Kronecker product, the reader may refer to [14] . In order to be complete, we state some useful theorems without including proofs about Hadamard matrices. From now on, by "a matrix of order n" we mean an $n \times n$ matrix.

Theorem 2.2.3. ([*11*]) *If a Hadamard matrix of order n exists, then* $n = 1, 2$ *or* $n \equiv 0(mod\ 4)$.

There are several operations on Hadamard matrices which preserves the Hadamard property:

1. permuting rows and multiplying some rows by -1,
2. permuting columns and multiplying some columns by -1,
3. transposition.

Remark 2.2.4. *Let* $\mathcal{S}_n$ *be the group consisting of all permutation matrices of order n, and let* $\mathcal{D}_n$ *be the group consisting of all diagonal matrices with all diagonal entries equal to* ± 1. *Then* $\mathcal{S}_n^{\pm}$, *the group generated by* $\mathcal{S}_n$ *and* $\mathcal{D}_n$, *is a semidirect product of* $\mathcal{S}_n$ *and* $\mathcal{D}_n$. *The elements of* $\mathcal{S}_n^{\pm}$ *are called* ***monomial matrices***. *Note that, to apply an operation of type* (1) *to a matrix amounts multiplying that matrix by a monomial matrix from the left. Similarly, an operation of type* (2) *corresponds to multiplying by a monomial matrix from the right.*

Definition 2.2.5. ([10]) Two Hadamard matrices H_1 and H_2 of order n, are **equivalent** if one can be obtained from the other by operations of types (1)

and/or (2), stated above. That is to say, H_1 and H_2 are equivalent if $H_2 = P^{-1}H_1Q$, for some monomial matrices P and Q in $\mathcal{S}_n^{\pm}$.

Definition 2.2.6. ([10]) The **automorphism group** of a Hadamard matrix H of order n, is the group consisting of all pairs (P, Q) of monomial matrices, satisfying $P^{-1}HQ = H$, where the group operation is

$$(P_1, Q_1) \circ (P_2, Q_2) = (P_1P_2, Q_1Q_2).$$

In 1979, J. S. Leon ([13]) , proposed an algorithm for computing the automorphism group of a Hadamard matrix. We denote the automorphism group of a Hadamard matrix H by $Aut(H)$, that is

$$Aut(H) = \{(P, Q) \in \mathcal{S}_n^{\pm} | P^{-1}HQ = H\}. \tag{2.2.4}$$

An Hadamard matrix whose first row and first column entries equal to $+1$ is referred as **normalized**. It is clear that any Hadamard matrix can be normalized by applying the operations mentioned above.

Corollary 2.2.7. ([*11*]) *If H is a normalized Hadamard matrix of order $4n$, then every row (column) except the first has $2n$ entries equal to -1, $2n$ entries equal to $+1$. Furthermore, n -1's in any row (column) overlap with n -1's in any other row (column).*

Proof: Follows from Theorem 2.2.3. □

Theorem 2.2.8. ([*9*]) *If H is a Hadamard matrix of order n, then*

$$det(H) = \pm n^{n/2}.$$

Theorem 2.2.9. ([9]) *Let H_1 and H_2 be Hadamard matrices of orders n_1 and n_2, respectively. Then, $H_1 \otimes H_2$ is a Hadamard matrix of order $n_1 n_2$.*

Theorem 2.2.10. ([27]) *There is a Hadamard matrix of order 2^n for all non-negative integers n.*

The simplest construction of new Hadamard matrices is by means of the Kronecker product. The "Sylvester-Hadamard" matrix of order 2^n, denoted by H_n, is the iterated Kronecker product of n copies of the Hadamard matrix $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ of order 2. This result follows from Theorem 2.2.9. The recursive relation of Sylvester-Hadamard matrices can be stated as follows:

$$H_0 = [1]\,,\ H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.2.5}$$

and

$$H_n = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes H_{n-1} = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}. \tag{2.2.6}$$

In the following lemma, it is shown that Sylvester-Hadamard matrices can be written as a product of matrices. This property of Sylvester-Hadamard matrices will be used in the following section.

Lemma 2.2.11. ([14]) *Let H_n be the Sylvester-Hadamard matrix of order 2^n. H_n can be written as the product of n matrices such that $H_n = M_n^{(1)} M_n^{(2)} \cdots M_n^{(n)}$, where $M_n^{(i)} = I_{2^{n-i}} \otimes H_1 \otimes I_{2^{i-1}}$ and I_{2^i} is the identity matrix of order 2^i.*

Lemma 2.2.12. ([23]) *Let* $H_n = \begin{bmatrix} \ell_0 \\ \ell_1 \\ \vdots \\ \ell_{2^n-1} \end{bmatrix}$ *be the Sylvester-Hadamard matrix of order* 2^n *for* $n \geq 0$*, where* ℓ_i *denotes the* i*-th row of* H_n*. Then,* ℓ_i *is the sequence of linear function* f_{α_i} *for any* $i = 0, 1, \ldots, 2^n - 1$*, where* $\alpha_i \in \mathcal{V}_n$*.*

Corollary 2.2.13. *If the rows and the columns of a Sylvester - Hadamard matrix* H_n *are indexed by all elements in* $\mathcal{V}_n$ *with respect to lexicographic ordering, then the entry in row* $\alpha = (a_1, a_2, \ldots, a_n)$ *and column* $\beta = (b_1, b_2, \ldots, b_n)$ *is equal to* $(-1)^{\langle \alpha, \beta \rangle}$*, for all* α, β *in* $\mathcal{V}_n$*, where* $\langle \alpha, \beta \rangle$ *is the inner product defined on* $\mathcal{V}_n$*.*

Since H_n is a symmetric matrix, the lemma above is also true for the columns of H_n. In addition, all rows of H_n comprise the sequence of all linear functions in $\mathcal{L}_n$. Consequently, the set containing the columns of H_n and $-H_n$ together consists of all elements of $\mathcal{A}_n$. Since, for any $f \in \mathcal{F}_n$ and its sequence ζ_f, we have the following:

$$H_n \zeta_f^t = \begin{bmatrix} \langle \zeta_f, \ell_0 \rangle \\ \langle \zeta_f, \ell_1 \rangle \\ \vdots \\ \langle \zeta_f, \ell_{2^n-1} \rangle \end{bmatrix} \tag{2.2.7}$$

From which, it follows that, if we want to compute the distance of a function to each of the linear functions, it is enough to compute $\frac{1}{2}\left(2^n 1_n - \zeta_f H_n\right)$, where 1_n is the vector in $\mathcal{V}_{2^n}$ with all components equal to 1. In particular, the transformation $f \to \zeta_f H_n$ is another representation of $f \in \mathcal{F}_n$ with the inner product values of sequence of f and ℓ_i where $i \in \{0, 1, \ldots, 2^n - 1\}$, as defined in equation (2.1.1). This transformation is, in fact, called **Walsh-Hadamard** transformation and

will be mentioned in the following section.

2.3 Walsh-Hadamard Transform

Recall that, in the Section 2.1, we mentioned the representations of a Boolean function. In addition to those representations, a function f can also be represented by its **Walsh-Hadamard** transform which contains lots of information about $f \in \mathcal{F}_n$. In this section, we will give a brief information about Walsh - Hadamard transform and its properties. For more information on Walsh-Hadamard transform, reader may refer to [1]. Hereinafter, by "Walsh transform" and "inverse Walsh transform", we will mean the Walsh - Hadamard transform and the inverse Walsh - Hadamard transform, respectively.

Definition 2.3.1. ([6]) The **Walsh transform** of a real-valued function f with $f : \mathcal{V}_n \to \mathbb{R}$, is defined as:

$$W_f(\omega) = \sum_{\alpha \in \mathcal{V}_n} f(\alpha)(-1)^{\langle \alpha, \omega \rangle}, \tag{2.3.8}$$

where $w \in \mathcal{V}_n$.

The function $f(\alpha)$ can be recovered by the **inverse Walsh transform** by:

$$f(\alpha) = 2^{-n} \sum_{\omega \in \mathcal{V}_n} W_f(\omega)(-1)^{\langle \alpha, \omega \rangle}. \tag{2.3.9}$$

Observe that, the Walsh transform and its inverse are both valid only for real valued functions. Therefore, for a Boolean function f, while computing its Walsh transform, the sum and values of inner product are treated as integer. We denote the Walsh transform of f and $\hat{f}$ by W_f and $W_{\hat{f}}$, respectively. The relation

between W_f and $W_{\hat{f}}$ is stated in the following lemma. Similarly, we will denote the inverse Walsh transform by W_f^{-1} and $W_{\hat{f}}^{-1}$.

Lemma 2.3.2. ([6]) *If $\hat{f}$ is the sign function of f, then*

$$W_{\hat{f}}(\omega) = -2W_f(\omega) + 2^n\delta(\omega), \tag{2.3.10}$$

which is equivalent to

$$W_f(\omega) = 2^{n-1}\delta(\omega) - \frac{1}{2}W_{\hat{f}}(\omega), \tag{2.3.11}$$

where $\delta(\omega)$ is the Kronecker delta function.

Nevertheless, in the literature, both $W_f(\omega)$ and $W_{\hat{f}}(\omega)$ are used interchangeably. However, from now on unless otherwise stated explicitly, for a Boolean function f, by "the Walsh transform" of f, we shall mean the Walsh transform of its sign function, and denote it by $W_{\hat{f}}$.

The ordered tuple, $(W_{\hat{f}}(\alpha_0), W_{\hat{f}}(\alpha_1), \ldots, W_{\hat{f}}(\alpha_{2^n-1}))$, i.e. the truth table of the Walsh transform of f is called the **Walsh spectrum** of f and is denoted by $T_{W_{\hat{f}}}$. The absolute maximum value in the Walsh spectrum is called **spectral amplitude** and we shall denote by $\left\|T_{W_{\hat{f}}}\right\|$. Note that, by "Walsh spectrum" we do not mean the **Walsh-Hadamard energy spectrum** of f which is defined as $\left(W_{\hat{f}}(\omega)\right)^2$. By $T_{(W_{\hat{f}})^2}$, we shall mean the ordered sequence

$$((W_{\hat{f}}(\alpha_0))^2, (W_{\hat{f}}(\alpha_1))^2, \ldots, (W_{\hat{f}}(\alpha_{2^n-1})^2).$$

The following theorem gives the necessary and sufficient condition for a Walsh transform belongs to a sign function of a Boolean function:

Theorem 2.3.3. ([6]) *$g : \mathcal{V}_n \longrightarrow \mathbb{R}$ is the Walsh transform of a sign function $\hat{f}$ of a Boolean function $f \in \mathcal{F}_n$ if and only if the following holds for all λ in $\mathcal{V}_n$:*

$$\sum_{\omega \in \mathcal{V}_n} g(\omega)g(\omega+\lambda) = 2^n\delta(\lambda) = \begin{cases} 2^n & for\ \lambda = \alpha_0 \\ 0 & otherwise \end{cases} \tag{2.3.12}$$

Proof: Let $\hat{f} : \mathcal{V}_n \longrightarrow \mathbb{R}$ be defined as

$$\hat{f}(\alpha) = \sum_{\omega \in \mathcal{V}_n} g(\omega)(-1)^{\langle \alpha, \omega \rangle},$$

then $W_{\hat{f}} = g$. Now, $\hat{f}$ is the sign function of a Boolean function if and only if $(\hat{f}(\alpha))^2 = 1$ for all $\alpha \in \mathcal{V}_n$. For all $\alpha \in \mathcal{V}_n$ we have

$$\begin{aligned} (\hat{f}(\alpha))^2 &= \frac{1}{2^{2n}} \sum_{\beta \in \mathcal{V}_n} g(\beta)(-1)^{\langle \beta, \alpha \rangle} \sum_{\gamma \in \mathcal{V}_n} g(\gamma)(-1)^{\langle \gamma, \alpha \rangle} \\ &= \frac{1}{2^{2n}} \sum_{\beta, \gamma \in \mathcal{V}_n} g(\beta)g(\gamma)(-1)^{\langle \beta \oplus \gamma, \alpha \rangle} \\ &= \frac{1}{2^{2n}} \sum_{\lambda \in \mathcal{V}_n} \underbrace{\left[\sum_{\beta \in \mathcal{V}_n} g(\beta)g(\beta \oplus \lambda) \right]}_{h(\lambda)} (-1)^{\langle \lambda, \alpha \rangle} \end{aligned} \tag{2.3.13}$$

and hence

$$\begin{aligned} \sum_{\alpha \in \mathcal{V}_n} (\hat{f}(\alpha))^2 (-1)^{\langle \alpha, \mu \rangle} &= \frac{1}{2^{2n}} \sum_{\alpha \in \mathcal{V}_n} \sum_{\lambda \in \mathcal{V}_n} h(\lambda)(-1)^{\langle \mu \oplus \lambda, \alpha \rangle} \\ &= \frac{1}{2^{2n}} \sum_{\lambda \in \mathcal{V}_n} h(\lambda) \sum_{\alpha \in \mathcal{V}_n} h(\lambda)(-1)^{\langle \mu \oplus \lambda, \alpha \rangle} \\ &= \frac{1}{2^n} h(\mu). \end{aligned} \tag{2.3.14}$$

Thus, if $(\hat{f}(\alpha))^2 = 1$ for all $\alpha \in \mathcal{V}_n$ then $h(\mu) = 2^{2n}\delta(\mu)$ for all $\mu \in \mathcal{V}_n$.

Conversely, if $h(\mu) = 2^{2n}\delta(\mu)$ for all $\mu \in \mathcal{V}_n$ then

$$(\hat{f}(\alpha))^2 = \frac{1}{2^{2n}} \sum_{\lambda \in \mathcal{V}_n} 2^{2n}\delta(\lambda)(-1)^{\langle \lambda, \alpha \rangle} = 1$$

for all $\alpha \in \mathcal{V}_n$. □

If we re-compute the equation (2.3.12) with $\lambda = \alpha_0$, we get the following well-known equation:

Corollary 2.3.4. ([*14*]) *Parseval's Equation*

$$\sum_{\omega \in \mathcal{V}_n} \left(W_{\hat{f}}(\omega) \right)^2 = 2^{2n}. \tag{2.3.15}$$

It is obvious that, for any $f \in \mathcal{F}_n$,

$$\begin{aligned} W_{\hat{f}}(\alpha_i) &= \sum_{\beta \in \mathcal{V}_n} (-1)^{f(\beta)} (-1)^{\langle \beta, \alpha_i \rangle} \\ &= \sum_{\beta \in \mathcal{V}_n} (-1)^{f(\beta)} (-1)^{f_{\alpha_i}(\beta)} \\ &= \langle \zeta_f, \ell_i \rangle . \end{aligned} \tag{2.3.16}$$

Thus, $W_{\hat{f}}(\alpha_i)$ is in fact, nothing but the difference between the number of 0's and the number of 1's in $T_{(f \oplus f_{\alpha_i})}$. Then, it is easy to see that:

$$d(f, f_{\alpha_i}) = \frac{1}{2}(2^n - W_{\hat{f}}(\alpha_i)) \tag{2.3.17}$$

From the equation above, it follows that:

$$T_{W_{\hat{f}}} = \zeta_f H_n, \tag{2.3.18}$$

as it is mentioned in the previous section.

The complexity of computing the Walsh spectrum of a function f is 2^{2n}. Beauchamp, in [1], by defining the **butterfly algorithm**, for the **Fast Walsh Transform** proposed by Brown in [2], reduced this complexity to $n2^n$. This algorithm follows from the lemma (2.2.11). For more information, the reader may refer to [21].

Recall that a function $f \in \mathcal{F}_n$ is called balanced if the number of zeros and the number ones in its truth table are equal. In order to check the balancedness of a function f, it is enough to look at $W_{\hat{f}}(\alpha_0)$. A function f is balanced if and only if $W_{\hat{f}}(\alpha_0) = 0$.

We will end this section with some properties of the Walsh transform:

Proposition 2.3.5. *([14]) For a function $f \in \mathcal{F}_n$, the Walsh transform of its complement function $\bar{f}$ at $\omega \in \mathcal{V}_n$ is equal to $-W_{\hat{f}}(\omega)$.*

Proof: Since $\bar{f}(\alpha) = f(\alpha) \oplus 1$, for all α in $\mathcal{V}_n$, we have the following:

$$\begin{aligned} W_{\hat{\bar{f}}}(\omega) &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{\bar{f}(\alpha) \oplus \langle \alpha, \omega \rangle} \\ &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{(f(\alpha) \oplus 1) \oplus \langle \alpha, \omega \rangle} \\ &= -\sum_{\alpha \in \mathcal{V}_n} (-1)^{f(\alpha) \oplus \langle \alpha, \omega \rangle} \\ &= -W_{\hat{f}}(\omega). \end{aligned} \tag{2.3.19}$$

□

Note that, $T_{W_{\hat{f}}}$ is composed of the distance of f to the functions in $\mathcal{A}_n \setminus \mathcal{L}_n$.

Proposition 2.3.6. *([21]) For a function $f \in \mathcal{F}_n$, the Walsh transform of the function $g = f \oplus f_\beta$ at $\omega \in \mathcal{V}_n$ is equal to $W_{\hat{f}}(\omega \oplus \beta)$, where f_β is the function*

defined by $f_\beta(\alpha) = \langle \alpha, \beta \rangle$ *for all* $\alpha \in \mathcal{V}_n$.

Proof:

$$\begin{aligned} W_{\hat{g}}(\omega) &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{g(\alpha) \oplus \langle \alpha, \omega \rangle} \\ &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{(f(x) \oplus f_\beta(x)) \oplus \langle \alpha, \omega \rangle} \\ &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(x) \oplus \langle \alpha, \beta \rangle \oplus \langle \alpha, \omega \rangle} \\ &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(x) \oplus \langle \alpha, \beta \oplus \omega \rangle} \\ &= W_{\hat{f}}(\omega \oplus \beta). \end{aligned} \tag{2.3.20}$$

□

From the above theorem, it follows that adding the truth table of a linear function to the truth table of another function f results in a dyadic shift on $T_{W_{\hat{f}}}$, which is determined by the linear function added.

Theorem 2.3.7. *For any* $f \in \mathcal{F}_n$,

$$W_{\hat{f}}(\alpha_i) \equiv \begin{cases} 0 \text{ (modulo 4)}, & \text{if } w(f) \text{ is even} \\ 2 \text{ (modulo 4)}, & \text{if } w(f) \text{ is odd} \end{cases} \tag{2.3.21}$$

for all $\alpha_i \in \mathcal{V}_n$.

Proof: Consider the function $0_n \in \mathcal{F}_n$. It is trivial that the Walsh spectrum of 0_n is $\{2^n, 0, 0, \ldots, 0\}$. We know that, any function $f \in \mathcal{F}_n$ can be obtained by complementing a 0 to 1 in truth table of 0_n iteratively. Then the number of these steps is equal to the $w(f)$. By Theorem 2.3.8, after each step, $W_{\hat{f}}(\alpha_i)$ increases or decreases by $+2$ for all $\alpha_i \in \mathcal{V}_n$. Therefore, if the number of steps is

even, then $W_{\hat{f}}(\alpha_i) \equiv 0$ (modulo 4). Conversely, if number of steps is odd, then $W_{\hat{f}}(\alpha_i) \equiv 2$ (modulo 4), which completes the proof. □

Theorem 2.3.8. *([17]) Let f and g be functions in $\mathcal{F}_n$ with the following property:*

$$g(\alpha) = \begin{cases} f(\alpha_i) \oplus 1 & if\ \alpha = \alpha_i \\ f(\alpha) & otherwise. \end{cases} \tag{2.3.22}$$

Then, Walsh transform of g is equal to:

$$W_{\hat{g}}(\omega) = W_{\hat{f}}(\omega) - 2(-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle}. \tag{2.3.23}$$

Proof:

$$\begin{aligned} W_{\hat{g}}(\omega) &= \sum_{\alpha\in\mathcal{V}_n} (-1)^{g(\alpha)\oplus\langle\alpha,\omega\rangle} \\ &= (-1)^{(f(\alpha_i)\oplus 1)\oplus\langle\alpha_i,\omega\rangle} + \sum_{\substack{\alpha\in\mathcal{V}_n \\ \alpha\neq\alpha_i}} (-1)^{f(\alpha)\oplus\langle\alpha,\omega\rangle} \\ &= -2(-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle} + \sum_{\alpha\in\mathcal{V}_n} (-1)^{f(\alpha)\oplus\langle\alpha,\omega\rangle} \\ &= W_{\hat{f}}(\omega) - 2(-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle}. \end{aligned} \tag{2.3.24}$$

□

Theorem 2.3.9. *([17]) Let $f, g \in \mathcal{F}_n$ be functions with the following property:*

$$g(\alpha) = \begin{cases} f(\alpha_i) \oplus 1 & if\ \alpha = \alpha_i \\ f(\alpha_j) \oplus 1 & if\ \alpha = \alpha_j \\ f(\alpha) & otherwise. \end{cases} \tag{2.3.25}$$

Then, Walsh transform of g is equal to:

$$W_{\hat{g}}(\omega) = W_{\hat{f}}(\omega) - 2((-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle} + (-1)^{f(\alpha_j)\oplus\langle\alpha_j,\omega\rangle}). \tag{2.3.26}$$

Proof:

$$\begin{aligned} W_{\hat{g}}(\omega) &= \sum_{\alpha\in\mathcal{V}_n} (-1)^{g(\alpha)\oplus\langle\alpha,\omega\rangle} \\ &= (-1)^{(f(\alpha_i)\oplus 1)\oplus\langle\alpha_i,\omega\rangle} + (-1)^{(f(\alpha_j)\oplus 1)\oplus\langle\alpha_j,\omega\rangle} + \sum_{\substack{\alpha\in\mathcal{V}_n \\ \alpha\neq\alpha_i,\alpha_j}} (-1)^{f(\alpha)\oplus\langle\alpha,\omega\rangle} \\ &= -2(-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle} - 2(-1)^{f(\alpha_j)\oplus\langle\alpha_j,\omega\rangle} + \sum_{\alpha\in\mathcal{V}_n} (-1)^{f(\alpha)\oplus\langle\alpha,\omega\rangle} \\ &= W_{\hat{f}}(\omega) - 2((-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle} + (-1)^{f(\alpha_j)\oplus\langle\alpha_j,\omega\rangle}). \end{aligned} \tag{2.3.27}$$

□

Theorem 2.3.10. *([21]) Let $f, g \in \mathcal{F}_n$. Let $h \in \mathcal{F}_n$ be the function defined as,*

$$h(\alpha) = (f \oplus g)(\alpha) = f(\alpha) \oplus g(\alpha) \text{ for all } \alpha \in \mathcal{V}_n.$$

Then Walsh transform of h is equal to

$$W_{\hat{h}}(\omega) = \frac{1}{2^n} \sum_{\alpha\in\mathcal{V}_n} W_{\hat{f}}(\omega \oplus \alpha) W_{\hat{g}}(\alpha), \tag{2.3.28}$$

for all $\omega \in \mathcal{V}_n$.

2.4 Design Criteria

Design and evaluation of Boolean functions used in cryptographic applications require some pre-defined design criteria. Although, the choice of the criteria depends on the cryptographic system in use, some are quite common, such as balancedness, strict avalanche criterion, nonlinearity, high algebraic degree, correlation immunity, propagation criterion. A cryptographically strong Boolean function ought to satisfy these criteria in order to resist well-known cryptanalytic attacks.

Recall that, we have defined that a function $f \in \mathcal{F}_n$ is said to be balanced if the number of 0's and 1's are equal in T_f, or equivalently, $W_{\hat{f}}(\alpha_0) = 0$. In order to avoid statistical dependence between input and output, which can be used in some cryptanalytic attacks, Boolean functions should be balanced.

H. Feistel, in [5], proposed that a cryptosystem should possess **avalanche effect**, that is when a single input bit complemented, half of the output bits change. In 1979, Kam and Davida ([12]) pointed out that a cryptosystem must be **complete**, that is, each output bit must depend on each input bit. Later, Webster and Tavares, in [28], combined these two criteria into one as **strict avalanche criterion**. A cryptographic transformation is said to satisfy strict avalanche criterion if each output bit changes with the probability one half whenever a single input bit is changed. Although the strict avalanche criterion is defined for a cryptosystem, it can be redefined for Boolean functions as follows,

Definition 2.4.1. ([6]) A function $f \in \mathcal{F}_n$ satisfies strict avalanche criterion if

$$\sum_{\alpha \in \mathcal{V}_n} f(\alpha) \oplus f(\alpha \oplus e_i) = 2^{n-1}, \text{ for all } i \in \{1, 2, \ldots, n\}. \tag{2.4.29}$$

or equivalently,

$$\sum_{\alpha \in \mathcal{V}_n} \hat{f}(\alpha)\hat{f}(\alpha \oplus e_i) = 0, \text{ for all } i \in \{1, 2, \ldots, n\}. \tag{2.4.30}$$

For any $f \in \mathcal{F}_n$, the function $\alpha \longmapsto f(\alpha) \oplus f(\alpha \oplus \beta)$ for all $\alpha \in \mathcal{V}_n$ is called the **difference function** of f corresponding to $\beta \in \mathcal{V}_n$, and denoted by f^β. The **auto-correlation function** of f with shift β is defined as

$$\Delta_f(\beta) = \sum_{\alpha \in \mathcal{V}_n} \hat{f}(\alpha)\hat{f}(\alpha \oplus \beta) = \langle \zeta_f, \zeta_f(\beta) \rangle, \tag{2.4.31}$$

where $\zeta_f(\beta)$ hereinafter denotes the sequence of the function $\alpha \longmapsto f(\alpha \oplus \beta)$ for all $\alpha \in \mathcal{V}_n$. It is easy to show that Walsh transform of the auto-correlation function of a function f is $(W_{\hat{f}}(\omega))^2$ for all $\omega \in \mathcal{V}_n$, ([19]).

Then the following theorem is immediate.

Theorem 2.4.2. *([19]) Any function $f \in \mathcal{F}_n$ satisfies the strict avalanche criterion if and only if*

$$\Delta_f(\beta) = 0, \text{ for all } \beta \in \mathcal{V}_n \text{ with } w(\beta) = 1. \tag{2.4.32}$$

Definition 2.4.3. ([3],[4]) A function $f \in \mathcal{F}_n$ is said to have a **linear structure** if there exists a vector $\beta \in \mathcal{V}_n$ such that f^β takes the same value for all $\alpha \in \mathcal{V}_n$.

The most crucial design criterion for a cryptographically strong Boolean function is the **nonlinearity** criterion, since cryptosystems with linear Boolean functions can be easily broken. The notion of nonlinearity, firstly introduced by Pieprzyk and Finkelstein in [18], defined as the minimum of the distances of a

function $f \in \mathcal{F}_n$ to the affine functions, and is denoted by N_f. In particular, N_f is

$$N_f = \min_{g \in \mathcal{A}_n} d(f, g). \tag{2.4.33}$$

Obviously, for any function $f \in \mathcal{F}_n$, $N_f = 0$ if and only if f is an affine function.

Beside the nonlinearity definition above, there are two more nonlinearity notion defined in the literature, ([15]). One is the distance to the set of functions having linear structure, and the other is high algebraic degree which is $deg(f)$. However, the one expressed above is the most widely accepted.

From Lemma 2.1.1, we get the following useful theorem,

Theorem 2.4.4. *([24])For any $f \in \mathcal{F}_n$,*

$$N_f = 2^{n-1} - \frac{1}{2} \max_{i=0,1,\ldots,2^n-1} \{|\langle \zeta_f, \ell_i \rangle|\} \tag{2.4.34}$$

By using the equation 2.3.17, we can restate the above theorem as follows,

$$\begin{aligned} N_f &= 2^{n-1} - \frac{1}{2} \max_{\omega \in \mathcal{V}_n} \left\{ \left| W_{\hat{f}}(\omega) \right| \right\} && (2.4.35) \\ &= 2^{n-1} - \frac{1}{2} \left\| T_{W_{\hat{f}}} \right\|. && (2.4.36) \end{aligned}$$

In the next chapter, the nonlinearity criterion will be investigated in details.

Another criterion, which a cryptographically good Boolean function should satisfy is **correlation immunity**, firstly introduced by Siegenthaler ([26]). A Boolean function, whose output distribution probability is unchanged when any m of input bits are kept constant, is called **m-th order correlation immune** where $m \in \{1, 2, \ldots, n\}$. Furthermore, if a balanced boolean function is m-th order correlation immune, f is then said to be **m-resilient**.

Correlation immunity and resiliency of a function f can be characterized through the Walsh transform of f as follows,

Theorem 2.4.5. *([8]) Any $f \in \mathcal{F}_n$ is m-th order correlation immune where $m \in \{1, 2, \ldots, n\}$ if and only if $W_{\hat{f}}$ satisfies*

$$W_{\hat{f}}(\omega) = 0, \text{ for all } \omega \in \mathcal{V}_n \text{ with } 1 \leq w(\omega) \leq m. \tag{2.4.37}$$

There are trade-offs between n, $deg(f)$, N_f and the resiliency order of a function $f \in \mathcal{F}_n$. Siegenthaler in [26], states that for any m-resilient function f, where $0 \leq m < n-1$, we have the relation $deg(f) \leq n - m - 1$ and that any $(n-1)$-resilient function is affine. Similarly, for any m-th order correlation-immune function f, $deg(f)$ is at most $n - m$. Obviously, due to the Parseval's equation, while increasing the order of correlation immunity, $\left\| T_{W_{\hat{f}}} \right\|$ will also increase, which means N_f will decrease.

After the concept of strict avalanche criterion (SAC) introduced, two different generalizations were proposed. One of them was introduced by Forré as follows,

Definition 2.4.6. ([6]) Any $f \in \mathcal{F}_n$ is said to satisfy **strict avalanche criterion of order m** if

- f fulfills SAC of order $m - 1$
- any function obtained from f by fixing m of its input bits constant satisfies SAC as well.

From which it follows that the strict avalanche criterion corresponds to strict avalanche criterion of order 0.

The other was proposed by Preneel et. al. which is stated as follows,

Definition 2.4.7. ([19]) Any $f \in \mathcal{F}_n$ satisfies the **propagation criterion of degree k** ($PC(k)$) where $k \in \{1, 2, \ldots, n\}$, if T_f changes with a probability one half whenever i with $1 \leq i \leq k$ bits of the input are complemented. That is,

$$\Delta_f(\beta) = 0 \text{ for all } \beta \in \mathcal{V}_n \text{ with } 1 \leq w(\beta) \leq k. \tag{2.4.38}$$

Note that SAC is equivalent to the propagation criterion of degree 1.

All the design criteria defined up to now but the nonlinearity criterion measure the statistical dependence between the input and output bits. In fact, they measure the diffusion property mentioned in the previous chapter. On the other hand, the nonlinearity criterion measures how far are the Boolean functions, used in cryptosystem, from the affine functions, in particular, the confusion property mentioned in the previous chapter. In most of the cryptosystem, S-boxes, that are mostly constructed by composing the Boolean functions, are the only nonlinear parts. Due to the weaknesses of linear cryptosystems, nonlinearity criterion is the core property of Boolean functions. Meier and Staffelbach in [15] introduced the functions with highest nonlinearity, which are called **perfect nonlinear**, and proved that the class of these functions coincides with the class of **bent functions**. Bent functions are first defined by Rothaus in [20]. We will end this chapter with some properties of bent functions.

Definition 2.4.8. ([15]) Any function $f \in \mathcal{F}_n$ is perfect nonlinear if it satisfies $PC(n)$.

Definition 2.4.9. ([20]) Any function $f \in \mathcal{F}_n$ is called a bent function if

$$2^{-\frac{n}{2}} \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(\alpha) \oplus \langle \alpha, \omega \rangle} = \pm 1, \tag{2.4.39}$$

or equivalently,

$$W_{\hat{f}}(\omega) = \pm 2^{\frac{n}{2}}, \tag{2.4.40}$$

for all $\omega \in \mathcal{V}_n$.

From the theorem above it follows that bent functions are not balanced.

Theorem 2.4.10. *([20]) Bent functions only exist for even number n of variables, and the degree of their algebraic normal form is always bounded above by* $\frac{n}{2}$.

Theorem 2.4.11. *([15])The class of perfect nonlinear functions is the class of functions with maximum distance to both the affine functions and the functions having linear structure.*

CHAPTER 3

Nonlinearity Criteria

In the Section 2.4, we mentioned some common design criteria for cryptosystems. Due to the trade-offs in between these criteria, it is not possible to construct a function which is highly nonlinear, balanced, correlation immune of higher order, and satisfying propagation criterion of higher degree at the same time. Many attempts were made to construct Boolean functions satisfying all or some of those criteria. One can foresee that by analyzing the transformations under which, those criteria are invariant, new construction methods for Boolean functions with desired properties may raise. On the other hand, any design criteria should remain invariant under a large group of transformations, because of the theory of similarity of secrecy systems proposed by Shannon ([22]). In this chapter, first we present the bijective transformations acting on input variables of Boolean functions that preserve nonlinearity, which were initially analyzed in [15]. Next, we analyze counterparts of these transformations via the bijective transformations acting on the truth table of Boolean functions. Finally, we investigate the bijective transformations acting on the truth table of Boolean functions which are not equivalent to the bijective transformations acting on the input variables of Boolean functions.

3.1 Pre-Transformations

Let $\Omega(n)$ denote the group of all bijective transformations from $\mathcal{V}_n$ to $\mathcal{V}_n$ so that any $\varphi \in \Omega(n)$ can be written as

$$\varphi(x_1, x_2, \ldots, x_n) = (\varphi_1(x_1, x_2, \ldots, x_n), \varphi_2(x_1, x_2, \ldots, x_n), \ldots, \varphi_n(x_1, x_2, \ldots, x_n)),$$

where each φ_j $(j = 1, 2, \ldots, n)$ is a Boolean function. Let $\mathcal{A}(n)$ denote the subgroup of $\Omega(n)$ consisting of all affine transformations. Note that any element $\varphi \in \mathcal{A}(n)$ can be described as

$$\varphi(\alpha) = \alpha A \oplus \beta \tag{3.1.1}$$

where A is a nonsingular square matrix of order n with entries from $GF(2)$, and $\alpha, \beta \in \mathcal{V}_n$. Obviously, if β is the zero vector then φ turns out to be a linear transformation. The group of linear transformations is denoted by $\mathcal{L}(n)$. From now on, unless otherwise stated explicitly, by a **pre-transformation**, we shall mean a bijective transformation on the input variables of a Boolean function, i.e. an element of $\Omega(n)$.

We denote the composition of a function $f \in \mathcal{F}_n$ and a pre-transformation φ by $(f \circ \varphi)(\alpha) = f(\varphi(\alpha))$ for all $\alpha \in \mathcal{V}_n$. Since φ is a bijection, it maps the ordered tuple $\{\alpha_0, \alpha_1, \ldots, \alpha_{2^n-1}\}$ to the ordered tuple $\{\alpha_{i_0}, \alpha_{i_1}, \ldots, \alpha_{i_{2^n-1}}\}$, where the latter is a permutation of the former.

It is clear that applying a pre-transformation φ to a function f, yields only a permutation of the components of T_f, hence any pre-transformation preserves $w(f)$. Particularly, the function obtained by applying a pre-transformation to a

balanced Boolean function is also balanced.

Lemma 3.1.1. *([15]) For any $f \in \mathcal{F}_n$,*

$$d(f, g) = d(f \circ \varphi, g \circ \varphi) \tag{3.1.2}$$

where $g \in \mathcal{F}_n$ and $\varphi \in \mathcal{A}(n)$.

Proof: Since φ induces a permutation on $\{\alpha_0, \alpha_1, \ldots, \alpha_{2^n-1}\}$, in particular say $\left\{\alpha_{i_0}, \alpha_{i_1}, \ldots, \alpha_{i_{2^n-1}}\right\}$, $f(\alpha_j) = g(\alpha_j)$ if and only if $f(\alpha_{i_j}) = g(\alpha_{i_j})$, from which we get the result. □

We denote the set of pre-transformations preserving the nonlinearity for all $f \in \mathcal{F}_n$ with $\mathcal{Q}(N)$. It can be described as,

$$\mathcal{Q}(N) = \{\varphi \in \Omega(n) \mid N_f = N_{f\circ\varphi}, \text{ for all } f \in \mathcal{F}_n\}. \tag{3.1.3}$$

Theorem 3.1.2. *([15]) The group of pre-transformations keeping nonlinearity invariant for all functions is the group of affine pre-transformations. Namely,*

$$\mathcal{Q}(N) = \mathcal{A}(n). \tag{3.1.4}$$

Proof: From Lemma 3.1.1, it follows that $\mathcal{A}(n) \subseteq \mathcal{Q}(N)$. Now suppose that there exists $\varphi \in \Omega(n) \setminus \mathcal{A}(n)$ such that $\varphi \in \mathcal{Q}(N)$. Then, there exists at least one non-affine component of $\varphi(x_1, x_2, \ldots, x_n)$. Without loss of generality say the first component, namely $\varphi_1(x_1, x_2, \ldots, x_n)$. Then, for the projection function $f(x_1, x_2, \ldots, x_n) = x_1$, $f \circ \varphi$ is not an affine function, therefore $N_f < N_{f\circ\varphi}$ which contradicts with the assumption that $\varphi \in \mathcal{Q}(N)$. Thus, $\mathcal{Q}(N) = \mathcal{A}(n)$. □

We can proceed in the same way and obtain similar results when the minimum distance of a function $f \in \mathcal{F}_n$ to the functions having a linear structure is preferred as nonlinearity measure as in [15]. We denote the set of functions having a linear structure by $\mathcal{LS}_n$, and the minimum distance of $f \in \mathcal{F}_n$ to the functions in $\mathcal{LS}_n$ by $\acute{N}_f$. Then,

$$\mathcal{Q}(\acute{N}) = \left\{ \varphi \in \Omega(n) \mid \acute{N}_f = \acute{N}_{f \circ \varphi}, \text{ for all } f \in \mathcal{F}_n \right\}. \tag{3.1.5}$$

Thus, the exact statement of the analogous theorem is as follows:

Corollary 3.1.3. *([15]) The group of affine pre-transformations is a subset of the group of pre-transformations preserving the minimum distance of Boolean functions to the functions in $\mathcal{LS}_n$. That is,*

$$\mathcal{A}(n) \subset \mathcal{Q}(\acute{N}). \tag{3.1.6}$$

Proof: Let $f \in \mathcal{LS}_n$ with a nonzero vector $\beta \in \mathcal{V}_n$ as its linear structure, that is $f^\beta(\alpha) = f(\alpha) \oplus f(\alpha \oplus \beta) = c$, where $c \in GF(2)$, for all $\alpha \in \mathcal{V}_n$ and let $\varphi \in \mathcal{A}(n)$ be an affine pre-transformation. Then,

$$\begin{aligned} f(\varphi(\alpha)) \oplus f(\varphi(\alpha \oplus \varphi^{-1}(\beta))) &= f(\varphi(\alpha)) \oplus f(\varphi(\alpha) \oplus \varphi(\varphi^{-1}(\beta))) \\ &= f(\varphi(\alpha)) \oplus f(\varphi(\alpha) \oplus \beta) \\ &= c, \end{aligned}$$

from which it follows that $\varphi^{-1}(\beta)$ is a linear structure of $f \circ \varphi$. Since $d(f, f^\beta) = d(f \circ \varphi, f^\beta \circ \varphi)$, it follows that $\mathcal{A}(n) \subset \mathcal{Q}(\acute{N})$. □

Similarly, when the nonlinear order, in other words $deg(f)$ is considered as the

nonlinearity measure, by applying the same process one can obtain that $deg(f)$ for all $f \in \mathcal{F}_n$, remains invariant only under the affine transformations, in particular, $\mathcal{A}(n)$. We denote the set of transformations that preserve $deg(f)$ for all $f \in \mathcal{F}_n$ by $\mathcal{Q}(deg)$,

$$\mathcal{Q}(deg) = \{\varphi \in \Omega(n) \mid deg(f) = deg(f \circ \varphi), \text{ for all } f \in \mathcal{F}_n\}. \tag{3.1.7}$$

Theorem 3.1.4. *([15]) The group of pre-transformations that preserves the non-linear order is the group of affine pre-transformations. Namely,* $\mathcal{Q}(deg) = \mathcal{A}(n)$.

Proof: Let $f \in \mathcal{F}_n$ be a function with $deg(f)$ and $\varphi \in \mathcal{A}(n)$ be an affine pre-transformation. Then, obviously

$$deg(f \circ \varphi) \leq deg(f),$$

since the highest degree term of f may disappear in $f \circ \varphi$. On the other hand, this is also true for φ^{-1}, then

$$deg(f) = deg((f \circ \varphi) \circ \varphi^{-1}) \leq deg(f \circ \varphi).$$

Then we get $deg(f) = deg(f \circ \varphi)$, and therefore $\mathcal{A}(n) \subseteq \mathcal{Q}(deg)$.

Conversely, assume that, there exists a $\varphi \in \Omega(n) \backslash \mathcal{A}(n)$, such that $\varphi \in \mathcal{Q}(deg)$, then at least one component is not affine. Without loss of generality say the first, $\varphi_1(x_1, x_2, \ldots, x_n)$. Now, consider the function $f(x_1, x_2, \ldots, x_n) = x_1 \in \mathcal{F}_n$. Then one gets $f \circ \varphi = \varphi_1$ with $deg(f \circ \varphi) > 1$ while $deg(f) = 1$, which contradicts with the assumption that $\varphi \in \mathcal{Q}(deg)$. Thus, $\mathcal{Q}(deg) \subseteq \mathcal{A}(n)$.

From which it follows that $\mathcal{Q}(deg) = \mathcal{A}(n)$. □

3.2 Post-Transformations

Let $\Theta(n)$ be the group of all invertible transformations over V_{2^n}. Any transformation $\psi \in \Theta(n)$, can be written as follows,

$$\psi(x_1, x_2, \ldots, x_{2^n}) = (\psi_0(x_1, x_2, \ldots, x_{2^n}), \ldots, \psi_{2^n-1}(x_1, x_2, \ldots, x_{2^n})), \quad (3.2.8)$$

where each $\psi_i(x_1, x_2, \ldots, x_{2^n})$ is a Boolean function with 2^n variables. From now on, we call a bijective transformation acting on the truth tables of Boolean functions, i.e. a transformation in $\Theta(n)$, as a **post-transformation**. We denote the application of a post-transformation $\psi \in \Theta(n)$ to a function $f \in \mathcal{F}_n$ by $\psi * f = \psi(T_f)$. Note that $(\psi * f)(\alpha_i) = \psi_i(T_f)$, for all $i \in \{0, 1, \ldots, 2^n - 1\}$. Thus, $\psi * f$ represents the truth table of the resulting function, and $(\psi * f)(\alpha_i)$ represents the i-th $(i = 0, 1, \ldots, 2^n - 1)$ component of the truth table, namely the image of the resulting function at $\alpha_i \in \mathcal{V}_n$.

It is obvious that, for any $f \in \mathcal{F}_n$, in order to keep N_f invariant, it is enough for a post-transformation $\psi \in \Theta(n)$, to preserve $\left\|T_{W_{\hat{f}}}\right\|$. From Theorem 2.4.4, it follows that $N_f = N_{\psi * f}$ if and only if $\left\|T_{W_{\hat{f}}}\right\| = \left\|T_{W_{\widehat{\psi * f}}}\right\|$. By $\mathcal{P}(N)$, we denote the set of post-transformations, which preserve N_f for all $f \in \mathcal{F}_n$, that is,

$$\mathcal{P}(N) = \{\psi \in \Theta(n) \mid N_f = N_{\psi * f}, \text{ for all } f \in \mathcal{F}_n\} \quad (3.2.9)$$

Consider a bijective transformation $\varphi \in \Omega(n)$. Naturally, φ induces a permutation on the set $\{\alpha_0, \alpha_1, \ldots, \alpha_{2^n-1}\}$ and consequently on the truth table of any

function $f \in \mathcal{F}_n$. That is, φ can be regarded as a permutation matrix $P \in \mathcal{S}_{2^n}$,

$$T_{f\circ\varphi} = T_f P \tag{3.2.10}$$

or equivalently,

$$\zeta_{f\circ\varphi} = \zeta_f P. \tag{3.2.11}$$

Also note that any permutation is induced by a unique pre-transformation. Thus, it follows that $\Omega(n) \simeq \mathcal{S}_{2^n} \subset \Theta(n)$. From now on, we identify the permutations in $\Omega(n) \subset \Theta(n)$ with the matrix $P_\varphi \in \mathcal{S}_{2^n}$, $\varphi \in \Omega(n)$. Thus, $\Omega(n)$ corresponds only to a small subset of $\Theta(n)$.

Theorem 3.2.1. *A permutation $P_\varphi \in \mathcal{S}_{2^n}$ is in $\mathcal{P}(N)$ if and only if $\varphi \in \mathcal{Q}(N)$, or equivalently $\varphi \in \mathcal{A}(n)$.*

Proof:

($\Longleftarrow$) Suppose $\varphi \in \mathcal{A}(n)$. Equivalently, $\varphi \in \mathcal{Q}(N)$. Then, the permutation matrix $P_\varphi \in \mathcal{S}_{2^n}$ corresponding to φ preserves $\left\|T_{W_{\hat{f}}}\right\|$ and by Theorem 2.4.4, $P_\varphi \in \mathcal{P}(N)$.

($\Longrightarrow$) Conversely, suppose that $P_\varphi \in \mathcal{S}_{2^n}$ is in $\mathcal{P}(N)$. Then,

$$\left\|H_n P_\varphi \zeta_f^t\right\| = \left\|H_n \zeta_{f\circ\varphi}^t\right\| = \left\|H_n \zeta_f^t\right\|,$$

for all $f \in \mathcal{F}_n$. Now, consider an affine function $g \in \mathcal{A}_n$. Thus, we have

$$\left\|H_n \zeta_g^t\right\| = \left\|H_n \zeta_{g\circ\varphi}^t\right\| = 2^n,$$

by the assumption $P_\varphi \in \mathcal{P}(N)$. Therefore, it follows that $g \circ \varphi \in \mathcal{A}_n$, and

consequently, $\varphi \in \mathcal{A}(n)$. $\square$

Lemma 3.2.2. *A post-transformation $P_\varphi \in \mathcal{S}_{2^n}$ preserves nonlinearity for all $f \in \mathcal{F}_n$, i.e. $P_\varphi \in \mathcal{P}(N)$ if and only if, applying P_φ to T_f, results in a signed permutation of $T_{W_{\hat{f}}}$ for all $f \in \mathcal{F}_n$.*

Proof:

($\Longleftarrow$) Suppose, application of $P_\varphi \in \mathcal{S}_{2^n}$ to T_f results in a signed permutation of $T_{W_{\hat{f}}}$ for all $f \in \mathcal{F}_n$. Then, obviously,

$$\left\| T_{W_{\hat{f}}} \right\| = \left\| T_{W_{\widehat{\psi \circ f}}} \right\|.$$

By Theorem 2.4.4, we conclude that $P_\varphi \in \mathcal{S}_{2^n}$ preserves nonlinearity for all $f \in \mathcal{F}_n$, i.e. $P_\varphi \in \mathcal{P}(N)$.

($\Longrightarrow$) Conversely, suppose that a post-transformation $P_\varphi \in \mathcal{S}_{2^n}$ preserves nonlinearity for all $f \in \mathcal{F}_n$, i.e. $P_\varphi \in \mathcal{P}(N)$. Consequently, by Theorem 3.2.1, $\varphi \in \mathcal{A}(n)$ so that $\varphi(\alpha) = \alpha A \oplus \beta$ with $\alpha, \beta \in \mathcal{V}_n$ and A being a non-singular matrix of order n. Then, for all $f \in \mathcal{F}_n$ we have,

$$\begin{aligned}
W_{\widehat{f \circ \varphi}}(w) &= \sum_{\alpha \in \mathcal{V}_n} (-1)^{(f \circ \varphi)(\alpha) \oplus \langle \omega, \alpha \rangle} \\
&= \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(\alpha) \oplus \left\langle \omega, \varphi^{-1}(\alpha) \right\rangle} \qquad (3.2.12) \\
&= \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(\alpha) \oplus \left\langle \omega, A^{-1}\alpha \oplus A^{-1}\beta \right\rangle} \\
&= \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(\alpha) \oplus (\left\langle \omega, A^{-1}\alpha \right\rangle \oplus \left\langle \omega, A^{-1}\beta \right\rangle)} \\
&= (-1)^{\left\langle \omega, A^{-1}\beta \right\rangle} \sum_{\alpha \in \mathcal{V}_n} (-1)^{f(\alpha) \oplus \left\langle \omega (A^{-1})^t, \alpha \right\rangle}, \qquad (3.2.13)
\end{aligned}$$

for all $\omega \in \mathcal{V}_n$. Equation (3.2.12) holds since φ is a bijection and equation (3.2.13)

holds since $\varphi \in \mathcal{A}(n)$. Without loss of generality, let $W_{\widehat{f\circ\varphi}}(\alpha_i) = \left\|T_{W_{\widehat{f\circ\varphi}}}\right\|$ and $W_{\hat{f}}(\alpha_j) = \left\|T_{W_{\hat{f}}}\right\|$ for some $i, j \in \{0, 1, \ldots, 2^n - 1\}$. Then, since $\varphi \in \mathcal{A}(n)$ and $\left\|T_{W_{\widehat{f\circ\varphi}}}\right\| = \left\|T_{W_{\hat{f}}}\right\|$, by letting $\alpha_j = (\varphi^t)^{-1}(\alpha_i)$, we see that $\varphi \in \mathcal{A}(n)$ results in a signed permutation of $T_{W_{\hat{f}}}$. Thus, we can conclude that $P_\varphi \in \mathcal{S}_{2^n}$ is in $\mathcal{P}(N)$ if and only if applying P_φ to T_f results in a signed permutation in $W_{\hat{f}}$ for all $f \in \mathcal{F}_n$. □

Theorem 3.2.3. *For any permutation matrix $P_\varphi \in \mathcal{S}_{2^n}$, P_φ is in $\mathcal{P}(N)$ if and only if there exists a monomial matrix $Q \in \mathcal{S}_{2^n}^{\pm}$ such that $QH_n = H_nP_\varphi$.*

Proof: From Lemma 3.2.2, we know that the sufficient condition for $P_\varphi \in \mathcal{S}_{2^n}$ to be an element of $\mathcal{P}(N)$ is to permute the Walsh spectrum $T_{W_{\hat{f}}}$ for all $f \in \mathcal{F}_n$ up to possible changes in the signs of components, which is in fact, equivalent to the existence of a monomial matrix $Q \in \mathcal{S}_{2^n}^{\pm}$ such that

$$QT_{W_{\hat{f}}}^t = T_{W_{\widehat{f\circ\varphi}}}^t. \tag{3.2.14}$$

Since $T_{W_{\hat{f}}} = \zeta_f H_n$, the above equation can be re-written as follows

$$\begin{aligned} QH_n\zeta_f^t &= H_n\zeta_{f\circ\varphi}^t \\ &= H_nP_\varphi\zeta_f^t. \end{aligned} \tag{3.2.15}$$

From which it follows that $QH_n = H_nP_\varphi$. The converse can also be shown in a similar way. □

In fact, an equivalent statement for the above theorem can be given as follows.

Corollary 3.2.4. *For any permutation matrix $P_\varphi \in \mathcal{S}_{2^n}$, $P_\varphi \in \mathcal{P}(N)$ if and only if there exists a monomial matrix $Q \in \mathcal{S}_{2^n}^{\pm}$ such that $(Q^{-1}, P_\varphi) \in Aut(H_n)$.*

Proof: Let $P_\varphi \in \mathcal{P}(N)$ be a permutation matrix. Then, Theorem 3.2.3, says that there exists a monomial matrix $Q \in \mathcal{S}_{2^n}^{\pm}$ such that $QH_n = H_nP_\varphi$. Then,

$$H_n = Q^{-1}H_nP_\varphi. \tag{3.2.16}$$

Recall from Remark 2.2.4 that $Aut(H_n) = \{(Q, P) \in \mathcal{S}_{2^n}^{\pm} \mid H_n = Q^{-1}H_nP\}$. Therefore,

$$(Q^{-1}, P_\varphi) \in Aut(H_n). \tag{3.2.17}$$

It is easy to show that the converse also holds. □

Now, we state a useful theorem that classifies the permutation matrices $P_\varphi \in \mathcal{P}(N)$ as being in $\mathcal{L}(n)$ or $\mathcal{A}(n) \setminus \mathcal{L}(n)$.

Theorem 3.2.5. *Given a permutation matrix $P_\varphi \in \mathcal{P}(N)$, φ is in $\mathcal{L}(n)$ if and only if $Q \in \mathcal{S}_{2^n}$, where $(Q^{-1}, P_\varphi) \in Aut(H_n)$.*

Proof: Since $\varphi \in \mathcal{L}_n$, φ can be written as $\varphi(\alpha) = \alpha A$, where A is a non-singular matrix of order n and $\alpha \in \mathcal{V}_n$. Then, by recomputing the equation (3.2.13) by setting β to be the zero vector in $\mathcal{V}_n$, we get

$$W_{\widehat{f\circ\varphi}}(w) = \sum_{\alpha\in\mathcal{V}_n} (-1)^{f(\alpha)\oplus\left\langle\omega(A^{-1})^t,\alpha\right\rangle},$$

for all $\omega \in \mathcal{V}_n$. It follows that, applying P_φ to T_f, corresponds to a permutation on $T_{W_{\hat{f}}}$ but does not change the values. Therefore, the matrix $Q \in \mathcal{S}_{2^n}^{\pm}$, satisfying the equation (3.2.15), has no entry equal to -1, then it follows that $Q \in \mathcal{S}_{2^n}$. Similarly, suppose there exists $Q \in \mathcal{S}_{2^n}$, where $(Q^{-1}, P_\varphi) \in Aut(H_n)$. Then, P_φ corresponds to a permutation on $T_{W_{\hat{f}}}$ which does not change the values. This holds only when the β is the zero vector in the equation (3.2.13). Thus, we

conclude that φ is a linear pre-transformation. $\square$

Up to now, we recalled Meier and Staffelbach's results. One of them is that the group of nonlinearity preserving pre-transformations is the group of affine pre-transformations. Next, we proposed the permutations on the truth tables of Boolean functions which are equivalent to pre-transformations. For the elements in $\mathcal{S}_{2^n}$, we proved the necessary and sufficient conditions to preserve nonlinearity. Moreover, we give the classification for permutations whether they are equivalent to a linear pre-transformation or not. Now, we investigate the post-transformations that are not equivalent to a pre-transformation.

Lemma 3.2.6. *Fix $g \in \mathcal{F}_n$, and define a post-transformation $\psi \in \Theta(n)$ by setting $\psi * f = \psi(T_f) = T_f \oplus T_g$, for all $f \in \mathcal{F}_n$. Then, $\psi \in \mathcal{P}(N)$ if and only if $g \in \mathcal{A}_n$.*

Proof:

($\Longleftarrow$) Suppose $\psi \in \Theta(n)$ such that $\psi(T_f) = T_f \oplus T_g$ where $g \in \mathcal{A}_n$, and let $T_h = \psi * f = T_f \oplus T_g$. Then, from Theorem 2.3.10, we know that, Walsh transform of the function h is

$$W_{\hat{h}}(\omega) = \frac{1}{2^n} \sum_{\alpha \in \mathcal{V}_n} W_{\hat{f}}(\omega \oplus \alpha) W_{\hat{g}}(\alpha),$$

for all $\omega \in \mathcal{V}_n$. Since $g \in \mathcal{A}_n$, $T_{W_{\hat{g}}} = (0, 0, \ldots, \pm 2^n, 0, \ldots, 0)$. That is, Walsh spectrum of g has only one non-zero value, which is equal to $\pm 2^n$. Then, the above equation can be simplified as,

$$W_{\hat{h}}(\omega) = \pm W_{\hat{f}}(\omega \oplus \alpha_k)$$

for all $\omega \in \mathcal{V}_n$, where $k \in \{0, 1, \ldots, 2^n - 1\}$. Thus, ψ preserves $\left\| T_{W_{\hat{f}}} \right\|$ for all

$f \in \mathcal{F}_n$. Therefore, it follows that $\psi \in \mathcal{P}(N)$.

($\Longrightarrow$) Conversely, suppose $\psi \in \Theta(n)$ such that $\psi(T_f) = T_f \oplus T_g$ where $f, g \in \mathcal{F}_n$ and $\psi \in \mathcal{P}(N)$. Assume that $g \notin \mathcal{A}_n$. Consider a function $f \in \mathcal{A}_n$. Then, obviously, the function $h \in \mathcal{F}_n$, with $T_h = \psi * f = T_f \oplus T_g$ is not an affine function. Thus, $N_h > N_f$, which contradicts with the assumption $\psi \in \mathcal{P}(N)$. Thus, the assertion follows. □

The above lemma, proves that there exist post-transformations which are not in $\mathcal{S}_{2^n}$, that keep N_f invariant for all $f \in \mathcal{F}_n$, that is,

$$\mathcal{P}(N) \setminus \{P_\varphi \in \mathcal{S}_{2^n} \mid \varphi \in \mathcal{A}(n)\} \neq \emptyset. \tag{3.2.18}$$

On the other hand, note that for some functions $f \in \mathcal{F}_n$, by satisfying some necessary conditions for the ψ considered above, N_f can be increased. In fact, iterative application of such ψ's, is nothing but the smart hill climbing method which is introduced in [16]. In the next chapter, we revise smart hill climbing as an application of such post-transformations.

Lemma 3.2.7. *Let $A \in GL(2^n, GF(2))$ be fixed. Define $\psi \in \Theta(n)$ such that $\psi(T_f) = (A{T_f}^t)^t$ for all $f \in \mathcal{F}_n$. If $\psi \in \mathcal{P}(N)$, then $A = (a_{i,j})$, where $i, j \in \{1, 2, \ldots, 2^n\}$, satisfies the following:*

(*i.*) *Any column of A is the truth table of some function whose nonlinearity is equal to one.*

(*ii.*) *The vector obtained by x-or of any two columns of A is the truth table of some function whose nonlinearity is equal to two.*

Proof: Let $\psi \in \Theta(n)$ be defined as above, and suppose that $\psi \in \mathcal{P}(N)$. Then,

($i.$) Consider the function $f \in \mathcal{F}_n$ such that $T_f = (1, 0, 0, \ldots, 0)$. Since $\psi \in \mathcal{P}(N)$, ψ preserves N_f. Then, the function $h \in \mathcal{F}_n$, such that $T_h = \psi * f = (a_{1,1}, a_{2,1}, \ldots, a_{2^n,1})$, i.e. T_h is equal to the first column of A. Since, $\psi \in \mathcal{P}(N)$, we also have $N_h = N_f = 1$. Thus, first column of A is the truth table of some function whose nonlinearity is equal to one. On the other hand, this must hold for all $f \in \mathcal{F}_n$ with $w(f) = 1$, therefore, it follows that this is true for any column of A. This proves the first condition.

($ii.$) Similarly, consider the function $f \in \mathcal{F}_n$ such that $T_f = (1, 1, 0, \ldots, 0)$. Due to the assumption that $\psi \in \mathcal{P}(N)$, the function $h \in \mathcal{F}_n$ obtained by $\psi(T_f) = ((a_{1,1} \oplus a_{1,2}), (a_{2,1} \oplus a_{2,2}), \ldots, (a_{2^n,1} \oplus a_{2^n,2}))$ has also the same nonlinearity with f. In fact, T_h is equal to the x-or of the first two columns of A. Since, $N_h = N_f = 2$, we see that the x-or of the first two columns of A satisfies the proposed condition. Furthermore, this holds for all functions $f \in \mathcal{F}_n$ with $w(f) = 2$, which completes the proof.

□

Lemma 3.2.8. *Any matrix $A \in GL(2^n, GF(2))$ satisfies the two conditions in Lemma 3.2.7 if and only if $A = B \oplus P_\varphi$, where $P_\varphi \in \mathcal{S}_{2^n}$ and B is a matrix of order 2^n over $GF(2)$ whose columns are the truth table of affine functions, not necessarily distinct.*

Proof:

($\Longleftarrow$) Let $A \in GL(2^n, GF(2))$ be such that $A = B \oplus P_\varphi$, where $P_\varphi \in \mathcal{S}_{2^n}$ and B is a matrix of order 2^n over $GF(2)$ whose columns are the truth tables of affine functions, not necessarily distinct. That is, $B = [C_1\ C_2\ \ldots\ C_{2^n}\ ]$, where C_i's are the truth tables of some affine functions. Since, P_φ is a matrix whose columns

have only one non-zero entry, it follows that the distance of the columns of A to their nearest affine function is one. Thus, the columns of A are truth tables of some functions whose nonlinearities are one. Similarly, since rows of P_φ have only one non-zero entry, the entries of columns in which they differ from their nearest affine functions can not be in the same row. Then, the vector obtained by x-oring any two columns of A is truth table of some function whose nonlinearity is equal to two. Therefore, A satisfies the two conditions in Lemma 3.2.7.

($\Longrightarrow$) Conversely, suppose $A \in GL(2^n, GF(2))$ satisfies the two conditions in Lemma 3.2.7. Then, each column of A is truth table of some function which differ from truth table of some affine function in only one entry, that are not in the same row. Thus, without loss of generality, assume those entries are in diagonal of A. Then, we can write $A = B \oplus I_{2^n}$, where I_{2^n} is the identity matrix of order 2^n and B is the matrix whose columns are truth table of some affine functions, from which the assertion follows. □

Lemma 3.2.9. *Given $A \in GL(2^n, GF(2))$ satisfying the conditions in Lemma 3.2.7, define the linear post-transformation $\psi \in \Theta(n)$ so that $\psi(T_f) = (AT_f{}^t)^t$ for all $f \in \mathcal{F}_n$. Then $\psi \in \mathcal{P}(N)$ if and only if the corresponding $\varphi \in \mathcal{A}(n)$.*

Proof: By Lemma 3.2.8, we know that $A \in GL(2^n, GF(2))$ satisfies the conditions in Lemma 3.2.7 if and only if it is of the form $B \oplus P_\varphi$, where B and P_φ satisfies the conditions mentioned in Lemma 3.2.8. Let $h \in \mathcal{F}_n$ be such that $T_h = \psi * f = (AT_f^t)^t$ for all $f \in \mathcal{F}_n$. Then, we have $T_h = (BT_f^t \oplus P_\varphi T_f^t)^t$. Since, each column of B is the truth table of an affine function, the function $g \in \mathcal{F}_n$ with $T_g = (BT_f^t)^t$ is again an affine function whose truth table is equal to the x-or of the columns of B corresponding to nonzero components in T_f. Let $d \in \mathcal{F}_n$ be such that $T_d = (P_\varphi T_f^t)^t$. Thus, we have $h(\alpha) = (d \oplus g)(\alpha)$ for all $\alpha \in \mathcal{V}_n$, then

by Theorem 2.3.10, we get,

$$W_{\hat{h}}(\omega) = \frac{1}{2^n} \sum_{\alpha \in \mathcal{V}_n} W_{\hat{d}}(\omega \oplus \alpha) W_{\hat{g}}(\alpha), \tag{3.2.19}$$

for all $\omega \in \mathcal{V}_n$. Since $g \in \mathcal{A}_n$, $T_{W_{\hat{g}}}$ has only one non-zero component that is equal to $\pm 2^n$, without loss of generality say at $\omega = \alpha_1$. Then, (3.2.19) can be simplified into

$$W_{\hat{h}}(\omega) = \pm W_{\hat{d}}(\omega \oplus \alpha_1),$$

for all $\omega \in \mathcal{V}_n$. Then, $N_h = N_f$ if and only if $N_d = N_f$. On the other hand, since $P_\varphi \in \mathcal{S}_{2^n}$, we know that, the function d is obtained from f by applying the pre-transformation φ, i.e. $d = f \circ \varphi$. From Theorem 3.1.2, we know that $\varphi \in \mathcal{Q}(N)$ if and only if $\varphi \in \mathcal{A}_n$. Thus, $N_h = N_f$ if and only if $\varphi \in \mathcal{A}_n$, which completes the proof. □

Theorem 3.2.10. *Let $\psi \in \Theta(n)$ be an affine post-transformation so that for all $f \in \mathcal{F}_n$,*

$$\psi(T_f) = (A{T_f}^t \oplus T_g^t)^t,$$

where $g \in \mathcal{F}_n$ and $A \in GL(2^n, GF(2))$ are fixed. Then $\psi \in \mathcal{P}(N)$ if and only if $g \in \mathcal{A}_n$ and $A = B \oplus P_\varphi$, where $P_\varphi \in \mathcal{S}_{2^n}$ with $\varphi \in \mathcal{A}(n)$, and B is the matrix of order 2^n over $GF(2)$ whose columns are the truth table of affine functions, not necessarily distinct.

Proof:

($\Longleftarrow$) Obviously, from Lemma 3.2.6, 3.2.7, 3.2.8, and 3.2.9, it follows that if $g \in \mathcal{A}_n$ and $A = B \oplus P_\varphi$, where $P_\varphi \in \mathcal{S}_{2^n}$ with $\varphi \in \mathcal{A}(n)$, and B is the matrix of order 2^n over $GF(2)$, whose columns are the truth table of affine functions,

which are not necessarily distinct, and $\psi(T_f) = (AT_f{}^t \oplus T_g^t)^t$, then $\psi \in \mathcal{P}(N)$.

($\Longrightarrow$) Conversely, for fixed $g \in \mathcal{F}_n$ and $A \in GL(2^n, GF(2))$, suppose an affine post-transformation $\psi \in \Theta(n)$ so that $\psi(T_f) = (AT_f{}^t \oplus T_g^t)4$, preserves N_f for all $f \in \mathcal{F}_n$, that is $\psi \in \mathcal{P}(N)$. Now, consider the affine function 0_n, that is $T_{0_n} = (0, 0, \ldots, 0)$. Then, by the assumption, the function $h \in \mathcal{F}_n$ such that $T_h = \psi(T_{0_n})$ has the same nonlinearity with 0_n. In other words, $h \in \mathcal{A}_n$. Since, $T_h = \psi(T_{0_n}) = T_g$, it follows that $g \in \mathcal{A}_n$. Since ψ is an affine post-transformation, we can write $\psi * f = \psi_2 * (\psi_1 * f)$, where $\psi_1 * f = (AT_f^t)^t$ and $\psi_2 * f = T_f \oplus T_g$. We have shown that, ψ_2 preserves N_f, for all $f \in \mathcal{F}_n$, since $g \in \mathcal{A}_n$. Then, $\psi \in \mathcal{P}(N)$ if and only if ψ_1 preserves N_f, for all $f \in \mathcal{F}_n$, for which the sufficient and necessary conditions are shown in Lemma 3.2.9. Therefore, the assertion follows. □

Remark 3.2.11. *([7]) Lemma 3.2.9 states that the nonlinearity preserving linear post-transformations are of the form $((B \oplus P_\varphi)T_f^t)^t$ where $P_\varphi \in \mathcal{S}_{2^n}$ and $\varphi \in \mathcal{A}(n)$, and B is a matrix of order 2^n over $GF(2)$ whose columns are truth tables of some affine functions, not necessarily distinct. Let ψ be a nonlinearity preserving linear post-transformation. Trivially, for all $f \in \mathcal{F}_n$ we have $\psi * f = ((B \oplus P_\varphi)T_f^t)^t$ where B and P_φ are as mentioned above. Furthermore, we can express ψ as $\psi = \psi_1 \oplus \psi_2$ where $\psi_1 * f = (BT_f^t)^t$ and $\psi_2 * f = (P_\varphi T_f^t)^t$. Obviously, image of ψ_1 is a subset of affine functions, i.e. $Im(\psi_1) \subseteq \mathcal{A}_n$. So, we may replace ψ_1 with a nonlinear map defined over V_{2^n}, say $\tilde{\psi}_1$, with $Im(\tilde{\psi}_1) \subseteq \mathcal{A}_n$, which is not necessarily invertible. Then, the resulting transformation $\tilde{\psi} = \tilde{\psi}_1 \oplus \psi_2$ is not an affine transformation. However, it is easy to show that $\tilde{\psi}$ keeps nonlinearity invariant. Therefore, if there exists a post-transformation $\tilde{\psi}$ such that $\tilde{\psi} = \tilde{\psi}_1 \oplus \psi_2$ where $Im(\tilde{\psi}_1) \subseteq \mathcal{A}_n$ and $\psi_2 * f = (P_\varphi T_f^t)^t$ with $P_\varphi \in \mathcal{S}_{2^n}$ and $\varphi \in \mathcal{A}(n)$, then $\tilde{\psi}$ is a nonlinear*

post-transformation which preserves nonlinearity.

Let ψ be a post-transformation. By definition ψ can be written as follows,

$$\psi(x_1, x_2, \ldots, x_{2^n}) = (\psi_0(x_1, x_2, \ldots, x_{2^n}), \ldots, \psi_{2^n-1}(x_1, x_2, \ldots, x_{2^n})),$$

where each $\psi_i(x_1, x_2, \ldots, x_{2^n})$ is a Boolean function with 2^n variables. Then ψ_i's can be represented by their algebraic normal form uniquely, i.e.,

$$\psi_i(x_1, x_2, \cdots, x_{2^n}) = c_0^{(i)} \oplus c_1^{(i)} x_1 \oplus \ldots \oplus c_{12\cdots 2^n}^{(i)} x_1 x_2 \cdots x_{2^n}.$$

So we have,

$$\psi : T_f \longmapsto \begin{pmatrix} \psi_0(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1})) \\ \psi_1(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1})) \\ \vdots \\ \psi_{2^n-1}(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1})) \end{pmatrix}^t.$$

By representing each ψ_i with its algebraic normal form, the above equation can be written as follows,

$$\psi : T_f \longmapsto \begin{pmatrix} c_0^{(0)} \oplus c_1^{(0)} f(\alpha_0) \oplus \ldots \oplus c_{12\cdots 2^n}^{(0)} f(\alpha_0) f(\alpha_1) \cdots f(\alpha_{2^n-1}) \\ c_0^{(1)} \oplus c_1^{(1)} f(\alpha_0) \oplus \ldots \oplus c_{12\cdots 2^n}^{(1)} f(\alpha_0) f(\alpha_1) \cdots f(\alpha_{2^n-1}) \\ c_0^{(2^n-1)} \oplus c_1^{(2^n-1)} f(\alpha_0) \oplus \ldots \oplus c_{12\cdots 2^n}^{(2^n-1)} f(\alpha_0) f(\alpha_1) \cdots f(\alpha_{2^n-1}) \end{pmatrix}^t$$

Then we get,

$$\psi : T_f \longmapsto \left(\underbrace{\begin{bmatrix} c_0^{(0)} \\ c_0^{(1)} \\ \vdots \\ c_0^{(2^n-1)} \end{bmatrix}}_{\lambda_0} \oplus \underbrace{\begin{bmatrix} c_1^{(0)} \\ c_1^{(1)} \\ \vdots \\ c_1^{(2^n-1)} \end{bmatrix}}_{\lambda_1} f(\alpha_0) \oplus \cdots \underbrace{\begin{bmatrix} c_{2^n}^{(0)} \\ c_{2^n}^{(1)} \\ \vdots \\ c_{2^n}^{(2^n-1)} \end{bmatrix}}_{\lambda_{2^n}} f(\alpha_{2^n-1}) \oplus \right.$$

$$\left. \underbrace{\begin{bmatrix} c_{12}^{(0)} \\ c_{12}^{(1)} \\ \vdots \\ c_{12}^{(2^n-1)} \end{bmatrix}}_{\lambda_{12}} f(\alpha_0)f(\alpha_1) \oplus \cdots \oplus \underbrace{\begin{bmatrix} c_{12\cdots 2^n}^{(0)} \\ c_{12\cdots 2^n}^{(1)} \\ \vdots \\ c_{12\cdots 2^n}^{(2^n-1)} \end{bmatrix}}_{\lambda_{12\cdots 2^n}} f(\alpha_0)\cdots f(\alpha_{2^n-1}) \right)^t,$$

or equivalently,

$$\psi : T_f \longmapsto \lambda_0 \oplus AT_f^t \oplus \lambda_{12} f(\alpha_0)f(\alpha_1) \oplus \cdots \oplus \lambda_{12\cdots 2^n} f(\alpha_0)f(\alpha_1)\cdots f(\alpha_{2^n-1}), \tag{3.2.20}$$

where A is the matrix of the form $A = [\lambda_0\ \lambda_1\ \ldots\ \lambda_{2^n}]$.

For a $\psi \in \Theta(n)$, if $\lambda_i = [0\ 0\ \ldots\ 0]^t$ for all $i \in \{12, 13, \ldots, 12\cdots 2^n\}$, then ψ is an affine post-transformation. Otherwise, if there exists a $\lambda_i \neq [0\ 0\ \ldots\ 0]^t$ for some $i \in \{12, 13, \ldots, 12\cdots 2^n\}$, then ψ is a nonlinear post-transformation. By Theorem 3.2.10 we know the necessary and sufficient conditions for an affine post-transformations to preserve the nonlinearity.

Remark 3.2.12. *Let ψ be a post-transformation satisfying the following conditions:*

(*i.*) λ_0 *is the truth table of some affine Boolean function,*

(*ii.*) *the matrix A satisfies the conditions mentioned in Lemma 3.2.9,*

(*iii.*) λ_i *is the truth table of some affine Boolean function for all* $i \in \{12, 13, \ldots, 12 \cdots 2^n\}$.

Unless $\lambda_i = [0\ 0\ \ldots\ 0]^t$ *for all* $i \in \{12, 13, \ldots, 12 \cdots 2^n\}$, ψ *is a non-affine post-transformation. Moreover, applying* ψ *to a function is nothing but permuting the function's truth table by an affine pre-transformation and x-oring it with the truth table of an affine function which is determined by the function itself. Therefore* ψ *keeps nonlinearity invariant.*

So far, we proved that there exists nonlinearity preserving post transformations that are not equivalent to the pre-transformations. Particularly, we gave the exact set of nonlinearity preserving affine post-transformations. Moreover, we showed that there exists a class of nonlinearity preserving non-affine post-transformations.

CHAPTER 4

An Application of Post-Transformations

In this chapter, we revise the smart hill climbing method, which is proposed in [16]. It is used to find highly nonlinear Boolean functions, by using post transformations.

4.1 Smart Hill Climbing Method

Let $\psi \in \Theta(n)$ such that $\psi * f = \psi(T_f) = T_f \oplus T_g$, for a fixed $g \in \mathcal{F}_n$ and for all $f \in \mathcal{F}_n$. In the Section 3.2, we showed that if $g \in \mathcal{A}_n$, then $\psi \in \mathcal{P}(N)$. Consider the function $g \in \mathcal{F}_n$ with $w(g) = 1$, then applying ψ to a given function $f \in \mathcal{F}_n$, can be expressed as follow,

$$h(\alpha) \;=\; \begin{cases} f(\alpha_i) \oplus 1 & if\ \alpha = \alpha_i \\ f(\alpha) & otherwise, \end{cases}$$

where $T_h = \psi * f$ and non-zero component of T_g is $g(\alpha_i)$. Then, by Theorem 2.3.8, Walsh transform of h is

$$W_{\hat{h}}(\omega) = W_{\hat{f}}(\omega) - 2(-1)^{f(\alpha_i) \oplus \langle \alpha_i, \omega \rangle}, \tag{4.1.1}$$

for all $\omega \in \mathcal{V}_n$. In fact,

$$2(-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle} = \begin{cases} -2 & if\ f(\alpha_i) \neq \langle\alpha_i,\omega\rangle\,, \\ 2 & if\ f(\alpha_i) = \langle\alpha_i,\omega\rangle\,. \end{cases} \tag{4.1.2}$$

Then, $W_{\hat{h}}(\omega) = W_{\hat{f}}(\omega) \pm 2$ for all $\omega \in \mathcal{V}_n$. In order to satisfy $N_h > N_f$, we must have

$$\left\|T_{W_{\hat{h}}}\right\| < \left\|T_{W_{\hat{f}}}\right\|. \tag{4.1.3}$$

By constructing the following sets,

$$\begin{aligned} W_1^+ &= \left\{\omega \in \mathcal{V}_n \mid W_{\hat{f}}(\omega) = \left\|T_{W_{\hat{f}}}\right\|\right\}, \\ W_1^- &= \left\{\omega \in \mathcal{V}_n \mid W_{\hat{f}}(\omega) = -\left\|T_{W_{\hat{f}}}\right\|\right\}, \end{aligned} \tag{4.1.4}$$

one can easily check whether the equation (4.1.3) is satisfied or not, without re-computing the Walsh transform of h. For a fixed $\alpha \in \mathcal{V}_n$, if we have $f(\alpha) = \langle\alpha,\omega\rangle$ for all $\omega \in W_1^+$ and $f(\alpha) \neq \langle\alpha,\omega\rangle$ for all $\omega \in W_1^-$, then obviously, it follows that the function obtained by complementing the value of f at α, has higher nonlinearity than f. Thus, we proved the following theorem.

Theorem 4.1.1. *([16]) Given a function $f \in \mathcal{F}_n$, the function $h \in \mathcal{F}_n$ obtained by applying a post-transformation $\psi \in \Theta(n)$ such that $\psi * f = T_f \oplus T_g$, where fixed $g \in \mathcal{F}_n$ and $w(g) = 1$, has higher nonlinearity if and only if*

(i.) $f(\alpha_i) = \langle\alpha_i,\omega\rangle$ for all $\omega \in W_1^+$,

(ii.) $f(\alpha_i) \neq \langle\alpha_i,\omega\rangle$ for all $\omega \in W_1^-$,

where $g(\alpha_i)$ is the nonzero component of T_g.

Remark 4.1.2. *Note that, in [16], in order to improve nonlinearity by complementing one component in the truth table of a function $f \in \mathcal{F}_n$, the components of Walsh spectrum which are equal to $\pm\left(\left\|T_{W_{\hat{f}}}\right\| - 2\right)$ are also considered. However, from the Theorem 2.3.7, we know that there are no such values. Therefore, the sets constructed as W_2^+, W_2^-, in Definition 2 in [16], are empty for all $f \in \mathcal{F}_n$.*

Throughout the process above, by considering the complemented component's value, we may obtain a balanced function. In order to use in a cryptosystem, the function $f \in \mathcal{F}_n$ should be balanced. However, complementing only one component in the truth table of a Boolean function does not preserve its weight and therefore balancedness. Instead, complementing two components of distinct value of a balanced function's truth table, preserves the weight.

Similarly, let $\psi \in \Theta(n)$ such that $\psi * f = \psi(T_f) = T_f \oplus T_g$, for a fixed $g \in \mathcal{F}_n$ with $w(g) = 2$ and for all $f \in \mathcal{F}_n$. Given a function $f \in \mathcal{F}_n$, applying ψ to f is in fact, nothing but the following:

$$h(\alpha) \;=\; \begin{cases} f(\alpha_i) \oplus 1 & if\ \alpha = \alpha_i, \\ f(\alpha_j) \oplus 1 & if\ \alpha = \alpha_j, \\ f(\alpha) & otherwise, \end{cases}$$

where $T_h = \psi * f$ and non-zero components in T_g are $g(\alpha_i)$ and $g(\alpha_j)$. Then, by Theorem 2.3.9, Walsh transform of h is

$$W_{\hat{h}}(\omega) = W_{\hat{f}}(\omega) - 2((-1)^{f(\alpha_i)\oplus\langle\alpha_i,\omega\rangle} + (-1)^{f(\alpha_j)\oplus\langle\alpha_j,\omega\rangle}), \tag{4.1.5}$$

for all $\omega \in \mathcal{V}_n$.

In addition to W_1^+ and W_1^- defined in (4.1.4), construct the following sets,

$$\begin{aligned} W_2^+ &= \left\{\omega \in \mathcal{V}_n \mid W_{\hat{f}}(\omega) = \left\|T_{W_{\hat{f}}}\right\| - 4\right\}, \\ W_2^- &= \left\{\omega \in \mathcal{V}_n \mid W_{\hat{f}}(\omega) = -(\left\|T_{W_{\hat{f}}}\right\| - 4)\right\}. \end{aligned} \tag{4.1.6}$$

Theorem 4.1.3. *([16]) Given a function $f \in \mathcal{B}_n$, compute the sets W_1^+, W_1^-, $W_1 = W_1^+ \cup W_1^-, W_2^+, W_2^-$. The function $h \in \mathcal{B}_n$ obtained by applying a post-transformation $\psi \in \Theta(n)$ such that $\psi * f = T_f \oplus T_g$, where fixed $g \in \mathcal{F}_n$ and $w(g) = 2$, has higher nonlinearity if and only if the following hold,*

$(i.)$ $f(\alpha_i) \neq f(\alpha_j)$,

$(ii.)$ $\langle \alpha_i, \omega \rangle \neq \langle \alpha_j, \omega \rangle$ *for all* $\omega \in W_1$,

$(iii.)$ $f(\alpha) = \langle \alpha, \omega \rangle$ *where* $\alpha = \alpha_i, \alpha_j$, *for all* $\omega \in W_1^+$,

$(iv.)$ $f(\alpha) \neq \langle \alpha, \omega \rangle$ *where* $\alpha = \alpha_i, \alpha_j$, *for all* $\omega \in W_1^-$,

$(v.)$ *for all* $\omega \in W_2^+$, *if* $\langle \alpha_i, \omega \rangle \neq \langle \alpha_j, \omega \rangle$, *then* $f(\alpha) = \langle \alpha, \omega \rangle$, *where* $\alpha = \alpha_i, \alpha_j$,

$(vi.)$ *for all* $\omega \in W_2^-$, *if* $\langle \alpha_i, \omega \rangle \neq \langle \alpha_j, \omega \rangle$, *then* $f(\alpha) \neq \langle \alpha, \omega \rangle$, *where* $\alpha = \alpha_i, \alpha_j$,

where $g(\alpha_i)$ and $g(\alpha_j)$ are the nonzero components of T_g.

Proof: Condition $(i.)$ is the necessary condition to retain balancedness. Then, we have,

$$W_{\hat{h}}(\omega) - W_{\hat{f}}(\omega) = \begin{cases} -4 & if\ f(\alpha_i) \neq \langle \alpha_i, \omega \rangle \text{ and } f(\alpha_j) \neq \langle \alpha_j, \omega \rangle, \\ 4 & if\ f(\alpha_i) = \langle \alpha_i, \omega \rangle \text{ and } f(\alpha_j) = \langle \alpha_j, \omega \rangle, \\ 0 & otherwise, \end{cases} \tag{4.1.7}$$

for all $\omega \in \mathcal{V}_n$. Since, in order to decrease $\left\|T_{W_{\hat{f}}}\right\|$, trivially, we need to have $W_{\hat{h}}(\omega) - W_{\hat{f}}(\omega) = -4$, (or 4) for all $\omega \in W_1^+$ (W_1^-, respectively) which proves

the conditions $(ii.)$, $(iii.)$ and $(iv.)$.

Furthermore, for all $\omega \in W_2^+$, $W_{\hat{h}}(\omega) - W_{\hat{f}}(\omega) \neq 4$ and for all $\omega \in W_2^-$, $W_{\hat{h}}(\omega) - W_{\hat{f}}(\omega) \neq -4$ must hold, again in order to decrease $\left\| T_{W_{\hat{f}}} \right\|$. Obviously, by checking all the possibilities in order to satisfy the statements above, we prove the conditions $(v.)$ and $(vi.)$.

□

Millan et. al. proposed algorithms in [16], that iteratively search for candidates satisfying the Theorem 4.1.1 and 4.1.3. Their algorithms are as fast as randomly generating functions and computing their nonlinearity. For example, for balanced functions with $n = 8$, mostly the algorithm results in functions with nonlinearity equal to 110. Note that, for $n = 8$, bent Boolean functions have nonlinearity equal to $2^{n-1} - \frac{1}{2}2^{\frac{n}{2}} = 2^7 - 2^3 = 120$, and it is still not known whether there exists a balanced function whose nonliearity is equal to 118. Later, in [17], the authors modified this algorithm so that beside nonlinearity criterion, it also searches for improvement on some other criteria as well.

CHAPTER 5

Conclusion

5.1 Remarks on The Other Design Criteria

Meier and Staffelbach in [15] also classified the pre-transformations keeping the order of correlation immunity invariant. This classification can be stated as follows,

Theorem 5.1.1. *([15])The subgroup of $\Omega(n)$ whose elements keep the order of correlation immunity invariant for all functions in $\mathcal{F}_n$ is the group of permutations (i.e. $\mathcal{S}_n$) and complementation of input variables of the functions.*

Recall that, Theorem 2.4.2 states that any function $f \in \mathcal{F}_n$ satisfies the strict avalanche criterion if and only if $\Delta_f(\beta) = 0$, for all $\beta \in \mathcal{V}_n$ with $w(\beta) = 1$. We know that the Walsh transform of the auto-correlation function of f with shift β, at ω is equal to $(W_{\hat{f}}(w))^2$, ([19]). Let $h : \mathbb{R}^{2^n} \to \mathbb{R}^{2^n}$ be defined as $h : (x_1, x_2, \ldots, x_{2^n}) \longmapsto ((x_1)^2, (x_2)^2, \ldots, (x_{2^n})^2)$. Thus, we have,

$$W_{\Delta_f}(\omega) = (W_{\hat{f}}(\omega))^2, \text{ for all } \omega \in \mathcal{V}_n$$

then,

$$(\Delta_f(\alpha_0), \Delta_f(\alpha_1), \ldots, \Delta_f(\alpha_{2^n-1}))^t = H_n^{-1} h(H_n \zeta_f^t), \tag{5.1.1}$$

where H_n^{-1} is the inverse of the Sylvester-Hadamard matrix of order 2^n.

Let $\varphi \in \Omega(n)$ be an affine pre-transformation, and $P_\varphi \in \mathcal{S}_{2^n}$ be the corresponding post-transformation. Then by Theorem 3.2.3, we know that φ is an affine pre-transformation if there exists $Q \in \mathcal{S}_{2^n}^{\pm}$ such that $H_n P_\varphi = Q H_n$. Since, the transformation h defined above commutes with any permutation matrix, and $P_\varphi H_n^{-1} = H_n^{-1} Q$, the auto-correlation function of $g \in \mathcal{F}_n$, such that $T_g = T_{f \circ \varphi}$ can be written as follows,

$$\begin{aligned}
(\Delta_g(\alpha_0), \Delta_g(\alpha_1), \ldots, \Delta_g(\alpha_{2^n-1}))^t &= (\Delta_{f\circ\varphi}(\alpha_0), \Delta_{f\circ\varphi}(\alpha_1), \ldots, \Delta_{f\circ\varphi}(\alpha_{2^n-1}))^t \\
&= H_n^{-1} h(H_n \zeta_{f\circ\varphi}^t) \\
&= H_n^{-1} h(H_n P_\varphi \zeta_f^t) \\
&= H_n^{-1} h(Q H_n \zeta_f^t) \\
&= H_n^{-1} Q h(H_n \zeta_f^t) \\
&= P_\varphi H_n^{-1} h(H_n \zeta_f^t) \\
&= P_\varphi(\Delta_f(\alpha_0), \Delta_f(\alpha_1), \ldots, \Delta_f(\alpha_{2^n-1}))^t
\end{aligned}$$

Thus, it follows that, φ induces the same permutation on the truth table and on the auto-correlation function's truth table for all function in $\mathcal{F}_n$. Then, the following is immediate,

Lemma 5.1.2. *An affine pre-transformation $\varphi \in \Omega(n)$ preserves strict avalanche property if the restriction of φ onto $\mathcal{B}$ is a bijection where*

$$\mathcal{B} = \{\alpha \in \mathcal{V}_n \mid w(\alpha) = 1\}.$$

Moreover, in the above theorem, by setting $\mathcal{B}$ such that $\mathcal{B} = \{\alpha \in \mathcal{V}_n \mid 1 \leq$

$w(\alpha) \leq k\}$, we get the following,

Lemma 5.1.3. *An affine pre-transformation $\varphi \in \Omega(n)$ preserves the order of propagation criterion if the restriction of φ onto $\mathcal{B}$ is a bijection where*

$$\mathcal{B} = \{\alpha \in \mathcal{V}_n \mid 1 \leq w(\alpha) \leq k\}.$$

Remark 5.1.4. *Let f be a Boolean function with $\Delta_f(\beta_i) = 0$ where $\beta_i \in \acute{\mathcal{B}} = \{\beta_i \in \mathcal{V}_n | i = 1, 2, \ldots, n\}$. If we have $\mathcal{B} = \acute{\mathcal{B}}$ where $\mathcal{B} = \{\alpha \in \mathcal{V}_n \mid w(\alpha) = 1\}$, then trivially f satisfies srict avalanche criterion. On the other hand, if $\mathcal{B} \neq \acute{\mathcal{B}}$ holds and the matrix A of order n whose rows are equal to β_i where $i = 1, 2, \ldots, n$, is invertible, then the function f can be transformed to a function g satisfying strict avalanche criterion by applying a pre-transformation φ that maps α to αA for all $\alpha \in \mathcal{V}_n$. In fact, this was first proposed and proved by Seberry et. al. ([25]), and recalled in the following theorem.*

Theorem 5.1.5. *([25]) Let f be a Boolean function and $A \in GL(n, GF(2))$. Suppose that $\Delta_f(\beta_i) = 0$, for each row β_i of A where $i = 1, 2, \ldots, n$. In other words, f satisfies propagation criterion with respect to all rows of A. Then $f \circ \varphi$ satisfies the strict avalanche criterion where φ is a pre-transformation such that $\varphi(\alpha) = \alpha A$ for all $\alpha \in \mathcal{V}_n$.*

5.2 Conclusion

In this thesis, after reviewing the basic concepts of Boolean functions and Walsh-Hadamard transform, we recalled some common design criteria, particularly, balancedness, strict avalanche criterion, correlation immunity, nonlinearity

and propagation criterion. Due to the similarity of secrecy systems proposed in [22], it is expected that these criteria should be preserved under a large group of transformations. Meier and Staffelbach in [15], studied under which pre-transformations, nonlinearity , algebraic degree and correlation immunity independently remain invariant.

First, we presented the results obtained in [15], in particular, nonlinearity criterion remains invariant under the group of affine pre-transformations. We explored the equivalent counterpart of these pre-transformations, that is to say, the permutations,which are induced by the pre-transformations, on the truth tables of Boolean functions. We proposed and proved that for any $P_\varphi \in \mathcal{S}_{2^n}$ preserves nonlinearity for all functions in $\mathcal{F}_n$ if there exists a monomial matrix $Q \in \mathcal{S}_{2^n}^{\pm}$ such that $(Q^{-1}, P) \in Aut(H_n)$. Furthermore, we showed that for any permutation $P_\varphi \in \mathcal{P}(N)$, $\varphi \in \mathcal{L}(n)$ if and only if the monomial matrix Q such that $(Q^{-1}, P) \in Aut(H_n)$ is in $\mathcal{S}_{2^n}$.

Next, we revised that a post-transformation $\psi \in \Theta(n)$ of the form $\psi * f = \psi(T_f) = T_f \oplus T_g$ for all $f \in \mathcal{F}_n$, where $g \in \mathcal{F}_n$ fixed, preserves nonlinearity if and only if $g \in \mathcal{A}_n$. Then, for the linear post-transformations, we proposed and proved the necessary and sufficient conditions to keep nonlinearity invariant. Namely, a linear post-transformation $\psi \in \Theta(n)$ so that $\psi * f = (AT_f^t)^t$, preserves nonlinearity if and only if the matrix $A \in GL(2^n, GF(2))$ is of the form $A = B \oplus P_\varphi$, where each column of B is the truth table of some affine function, which are not necessarily distinct, and $P_\varphi \in \mathcal{S}_{2^n}$ such that $\varphi \in \mathcal{A}(n)$. As a consequence of these results, we proved that an affine post-transformation $\psi \in \Theta(n)$ so that $\psi * f = (AT_f^t \oplus T_g^t)^t$, preserves nonlinearity if and only if $A \in GL(2^n, GF(2))$ is of the form $A = B \oplus P_\varphi$ as mentioned above, and $g \in \mathcal{A}_n$. Therefore, we showed

that besides the affine pre-transformations, there exists post-transformations that keeps nonlinearity invariant. Moreover, we proved that there exists a class of non-affine post-transformations that keep nonlinearity invariant.

Later, we presented the equivalence of smart hill climbing method proposed to find highly non-linear Boolean functions, by some post- transformations. In addition to this, we propose and prove some propositions on the other design criteria as well.

As a future work, the non-linear post-transformations should be examined to determine the exact set of post-transformations that keeps nonlinearity invariant. In addition, the post-transformations may also be investigated for the other design criteria. One can foresee that by analyzing the post-transformations keeping the design criteria invariant, new construction methods for Boolean functions satisfying optimized design criteria may raise. Another direction will be to investigate the effects of both pre-transformations and post-transformations on S-boxes that are constructed by Boolean functions.

REFERENCES

[1] Beauchamp K., *Applications of Walsh and related functions with an introduction to sequency functions*, Microelectronics and Signal Processing, Academic Press, London, New York, Tokyo (1984).

[2] Brown R., *A recursive algorithm for sequency-ordered Fast Walsh Transforms*, IEEE Transactions on Computers, C-26(8): 819-822 (1977).

[3] Chaum D. and Evertse J.H., *Cryptanalysis of DES with a reduced number of rounds*, Advances in Cryptology - CRYPTO'85 ed. H.C. Williams (Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, New York) 218: 192-211 (1985).

[4] Evertse J.H., *Linear structures in block ciphers*, Advances in Cryptology - EUROCRYPT'87 (Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York) 304: 249-266 (1988).

[5] Feistel H., *Cryptography and computer privacy*, Scienctific American, 228(5): 15-23 (1973).

[6] Forré R., *The strict avalanche criterion: Spectral properties of Boolean functions and extended definition*, Advances in Cryptology - CRYPTO'88 (Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, New York) 403: 450-468 (1988).

[7] Güloğlu İsmail Ş, personal communication (2004).

[8] Guo-Zhen X., and Massey J.L., *A spectral characterization of correlation-immune combining functions*, IEEE Transactions on Information Theory, Vol.34, No. 3: 569-571 (1988).

[9] Hadamard J., *Resolution d'une question relative aux determinant*, Bull. des. Sci. Math. 17: 240-246 (1893).

[10] Hall Jr. M., *Note on the Mathieu group* M_{12}, Arch. Math. 13: 334-340 (1962).

[11] Hall Jr. M., *Combinatorial theory*, Blaisdell, Walthem, Mass (1967).

[12] Kam J.B. and Davida G.I., *Structured design of substitution permutation encryption networks.*, IEEE Transactions on Computers, C-28(10):747-753 (1979).

[13] Leon J.S., *An algorithm for computing Authomorphism group of Hadamaard matrices*, Journal of Combinatorial Theory, Series A, Vol.27: 289-306 (1979).

[14] MacWilliams F.J., Sloane N.J.A., *The theory of error-correcting codes*, Amsterdam, New York, Oxford:North-Holland (1978).

[15] Meier W. and Staffelbach O., *Nonlinearity Criteria for cryptographic functions*, Advances in Cryptology - EUROCRYPT'89 (Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, New York 1990) 434: 549-562 (1989).

[16] Millan W., Clark A., and Dawson E., *Smart hill climbing finds better Boolean functions*, In Workshop on Selected Areas in Cryptology, Workshop Record: 50-63 (1997).

[17] Millan W., Clark A., and Dawson E., *Boolean function design using hill climbing methods*, In 4-th Australasian Conference on Information, Security and Privacy, (Lecture Notes on Computer Science, Springer Verlag, Berlin, Heidelberg, New York) Vol. 1587: 1-11 (1999).

[18] Pieprzyk J. and Finkelstein G., *Towards effective nonlinear cryptosystem design*, IEE Proceedings, Vol.135, Pt. E, No. 6: 325-335 (1988).

[19] Preneel B., Leekwijck W.V., Linden L.V., Govaerts R., and Vandewalle J., *Propagation characteristics of Boolean functions*, Advances in Cryptology - EUROCRYPT'90 (Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, New York 1991) 437: 155-165 (1990).

[20] Rothaus O.S., *On "bent" functions*, Journal of Combinatorial Theory, Ser. A, 20: 300-305 (1976).

[21] Sağdıçoğlu S., *Cryptological viewpoint of Boolean functions*, M. Sc. Thesis, The Department of Mathematics, Middle East Technical University, Ankara, Turkey, (2003).

[22] Shannon C.E., *Communication Theory of secrecy systems*, Bell System Technical Journal, Vol. 28: 656-715 (1949).

[23] Seberry J. and Zhang X. M., *Highly nonlinear 0-1 balanced Boolean functions satisfying strict avalanche criterion*, Advances in Cryptology - AUSCRYPT-92 (Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York) 718: 145-155 (1993).

[24] Seberry J., Zhang X. M. and Zheng Y., *Nonlinearly balanced Boolean functions and their propagation characteristics*, Advances in Cryptology -

CRYPTO-93 (Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York) 773: 49-60 (1994).

[25] Seberry J., Zhang X. M. and Zheng Y., *Improving the strict avalanche characteristics of cryptographic functions*, Information Processing Letters, 50: 37-41 (1994).

[26] Siegenthaler T., *Correlation-immunity of nonlinear combining functions for cryptographic applications*, IEEE Transactions on Information Theory, IT-30, No. 5: 776-779 (1984).

[27] Sylvester J.J., *Thoughts on the inverse orthogonal matrices, simultaneous sign succesions, and tesselated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers*, Phil. Mag. 34: 461-475 (1867).

[28] Webster A.F. and Tavares S.E., *On the design of S-boxes*, Advances in Cryptology - CRYPTO'85 ed. H.C. Williams (Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, New York) 218: 523-524 (1986).