Designs, Codes and Cryptography, 19, 45–63 (2000)
© 2000 Kluwer Academic Publishers, Boston. Manufactured in The Netherlands.

Relating Differential Distribution Tables to Other Properties of of Substitution Boxes

XIAN-MO ZHANG xianmo@cs.uow.edu.au
School of Information Technology and Computer Science, University of Wollongong, Wollongong, NSW 2522, Australia

YULIANG ZHENG yuliang.zheng@monash.edu.au
School of Computing and Information Technology, Monash University, Frankston, Melbourne, VIC 3199, Australia

HIDEKI IMAI imai@iis.u-tokyo.ac.jp
The Third Department, Institute of Industrial Science, University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8558, Japan

Communicated by: P. C. van Oorschot

Received July 22, 1997; Revised September 29, 1998; Accepted October 22, 1998

Abstract. Due to the success of differential and linear attacks on a large number of encryption algorithms, it is important to investigate relationships among various cryptographic, including differential and linear, characteristics of an S-box (substitution box). After discussing a precise relationship among three tables, namely the difference, auto-correlation and correlation immunity distribution tables, of an S-box, we develop a number of results on various properties of S-boxes. More specifically, we show (1) close connections among three indicators of S-boxes, (2) a tight lower bound on the sum of elements in the leftmost column of its differential distribution table, (3) a non-trivial and tight lower bound on the differential uniformity of an S-box, and (4) two upper bounds on the nonlinearity of S-boxes (one for a general, not necessarily regular, S-box and the other for a regular S-box).

Keywords: Boolean Functions, cryptography, differential attack, linear attack, nonlinearity, S-boxes

1. Introduction

This paper deals with $n \times m$ S-boxes with $n > m$. Success of the notable differential cryptanalysis on various block ciphers [3, 4] has motivated researchers to investigate properties of the difference distribution tables of S-boxes. A core topic in this endeavor is to discover relationships between differential distribution tables and other properties of S-boxes. In this paper we first introduce two additional tables associated with an S-box, these being the auto-correlation and correlation immunity distribution tables. Then we establish a precise relationship among the three tables of an S-box (i.e., the difference, auto-correlation and correlation immunity distribution tables). With this relationship as a basis, we show that an S-box is regular (or balanced) if and only if the sum of the values in the leftmost column of its difference distribution table is 2^{2n-m}. In a sense, this result complements a well-known fact about the regularity of an S-box which states that an S-box is regular if and only if the non-zero linear combinations of its component functions are all balanced.

The next issue addressed in this paper is on the lower bound on the differential uniformity of an S-box which is defined as the largest non-zero value in the differential distribution table of the S-box, not taking into account the first entry in the top row. For an $n \times m$ S-box, it is easy to see that its differential uniformity is at least 2^{n-m}. As another contribution of this paper, we will show a new tight lower bound that improves the "trivial" bound of 2^{n-m}.

The final issue addressed in this work relates more specifically the nonlinearity of an S-box to its difference distribution table. In particular, we give two upper bounds on the nonlinearity of the S-box, one for the case when the S-box is an arbitrary mapping and the other when it is regular. These two bounds are expressed in terms of three parameters: the number of input bits, the number of output bits and the number of non-zero entries in the entire difference distribution table or in the leftmost column of the difference distribution table of the S-box, respectively. We also compare the second new upper bound with previous works in the same area.

The remainder of this paper is organized as follows: Section 2 introduces formal notations and definitions used in this paper. The difference, auto-correlation and correlation immunity distribution tables of an S-box are defined in Section 3 where a precise relationship among the three tables is also established. An interesting connection between the regularity of an S-box and columns of its differential distribution table is presented in Section 4. A tight lower bound on the differential uniformity of an S-box is presented in Section 5, and then two upper bounds on the nonlinearity of an S-box and its difference distribution table are proved in Section 6. Section 7 closes the paper with some concluding remarks.

2. Basic Notations and Definitions

This section is intended as a summary of the minimum amount of mathematical knowledge required in rigorously treating issues on S-boxes to be discussed in this paper.

The vector space of n tuples of elements from $GF(2)$ is denoted by V_n. These vectors, in ascending lexicographic order, are denoted by $\alpha_0, \alpha_1, \ldots, \alpha_{2^n-1}$. As vectors in V_n and integers in $[0, 2^n - 1]$ have a natural one-to-one correspondence, it allows us to switch from a vector in V_n to its corresponding integer in $[0, 2^n - 1]$, and vice versa.

Let f be a function from V_n to $GF(2)$ (or simply, a function on V_n). The *sequence* of f is defined as $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-1)^{f(\alpha_{2^n-1})})$, while the *truth table* of f is defined as $(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$. f is said to be *balanced* if its truth table assumes an equal number of zeros and ones. We call $h(x) = a_1x_1 \oplus \cdots \oplus a_nx_n \oplus c$ an *affine function*, where $x = (x_1, \ldots, x_n)$ and $a_j, c \in GF(2)$. In particular, h will be called a *linear function* if $c = 0$. The sequence of an affine (linear) function will be called an *affine (linear) sequence*.

The *Hamming weight* of a vector v, denoted by $W(v)$, is the number of ones in v. Let f and g be functions on V_n. Then $d(f, g) = \sum_{f(x)\neq g(x)} 1$, where the addition is over the reals, is called the *Hamming distance* between f and g. Let $\varphi_0, \ldots, \varphi_{2^{n+1}-1}$ be the affine functions on V_n. Then $N_f = \min_{i=0,\ldots,2^{n+1}-1} d(f, \varphi_i)$ is called the *nonlinearity* of f. It is well-known that the nonlinearity of f on V_n satisfies $N_f \leq 2^{n-1} - 2^{\frac{1}{2}n-1}$. The equality holds if and only if f is bent (see P. 426 of [12]).

Given two sequences $a = (a_1, \ldots, a_m)$ and $b = (b_1, \ldots, b_m)$, their component-wise product is denoted by $a * b$, while the scalar product (sum of component-wise products) is denoted by $\langle a, b\rangle$.

Definition. Let f be a function on V_n. For a vector $\alpha \in V_n$, denote by $\xi(\alpha)$ the sequence of $f(x \oplus \alpha)$. Thus $\xi(0)$ is the sequence of f itself and $\xi(0) * \xi(\alpha)$ is the sequence of $f(x) \oplus f(x \oplus \alpha)$. Define the *auto-correlation* of f with a shift α by

$$\Delta(\alpha) = \langle \xi(0), \xi(\alpha) \rangle.$$

The *Sylvester-Hadamard matrix (or Walsh-Hadamard matrix)* of order 2^n, denoted by H_n, is generated by the recursive relation

$$H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, \ n = 1, 2, \ldots, \ H_0 = 1.$$

Each row (column) of H_n is a linear sequence of length 2^n.

The following two formulas are well known to researchers (for a proof see for instance [14, 23]).

Let ξ be the sequence of a function f on V_n. Then the nonlinearity of f, N_f can be calculated by

$$N_f = 2^{n-1} - \frac{1}{2} \max\{|\langle \xi, \ell_i \rangle|, 0 \le i \le 2^n - 1\} \tag{1}$$

where ℓ_i is the ith row of H_n, $i = 0, 1, \ldots, 2^n - 1$, and

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1}))H_n = (\langle \xi, \ell_0 \rangle^2, \langle \xi, \ell_1 \rangle^2, \ldots, \langle \xi, \ell_{2^n-1} \rangle^2) \tag{2}$$

where α_i is the binary representation of an integer i and ℓ_i is the ith row of H_n, $i = 0, 1, \ldots, 2^n - 1$.

An $n \times m$ S-box or substitution box is a mapping from V_n to V_m, i.e., $F = (f_1, \ldots, f_m)$, where n and m are integers with $n \ge m \ge 1$ and each component function f_j is a function on V_n. In this paper, we use the terms of mapping and S-box interchangeably.

As can be seen from the design of many practical block ciphers, researchers are mainly concerned with *regular* S-boxes only. A mapping $F = (f_1, \ldots, f_m)$ is said to be regular if $F(x)$ runs through each vector in V_m 2^{n-m} times while x runs through V_n once.

The following lemma states a useful result on the regularity of an S-box. This result has appeared in many different forms in the literature. Our description can be viewed as the binary version of Corollary 7.39 of [11].

LEMMA 1 *Let $F = (f_1, \ldots, f_m)$ be a mapping from V_n to V_m, where n and m are integers with $n \ge m \ge 1$ and each $f_j(x)$ is a function on V_n. Then F is regular if and only if every non-zero linear combination of $f_1, \ldots, f_m$ is balanced.*

The concept of nonlinearity can be extended to the case of an S-box [16].

Definition. The standard definition of the *nonlinearity* of $F = (f_1, \ldots, f_m)$ is

$$N_F = min_g \left\{ N_g \mid g = \bigoplus_{j=1}^{m} c_j f_j, \ c_j \in GF(2), g \neq 0 \right\}.$$

Now we consider an S-box in terms of its usefulness in designing a block cipher secure against differential cryptanalysis [3, 4]. The essence of a differential attack is to exploit particular entries in the difference distribution tables of S-boxes employed by a block cipher. The difference distribution table of an $n \times m$ S-box is a $2^n \times 2^m$ matrix. The rows of the matrix, indexed by the vectors in V_n, represent the changes in the inputs, while the columns, indexed by the vectors in V_m, represent the change in the output of the S-box. An entry in the table indexed by (α, β) indicates the number of input vectors which, when changed by α (in the sense of bit-wise XOR), result in a change in the output by β (also in the sense of bit-wise XOR). It should be pointed out that while in this paper the notation of difference is restricted to XOR differences, in general other differences are also of interest, such as those based on modular addition and multiplication.

Note that an entry in a difference distribution table can only take an even value, the sum of the values in a row is always 2^n, and the top row is always $(2^n, 0, \dots, 0)$. As entries with higher values in the table are particularly useful to differential cryptanalysis, a desirable condition for an S-box not to be exploited in differential cryptanalysis would be that it does not have large values in its differential distribution table (not taking into account the leftmost entry in the top row).

In measuring the strength of an S-box (in terms of the security of a block cipher that employs the S-box) against differential attacks, a useful indicator commonly used is *differential uniformity* which is defined as follows [17].

Definition. Let F be an $n \times m$ S-box, where $n \geq m$. Let δ be the largest value in the differential distribution table of the S-box (not taking into account the leftmost entry in the top row), namely,

$$\delta = \max_{\alpha \in V_n, \alpha \neq 0} \max_{\beta \in V_m} \#\{x \mid F(x) \oplus F(x \oplus \alpha) = \beta\}$$

Then F is said to be *differentially δ-uniform*, and accordingly, δ is called the differential uniformity of F.

An important ingredient in designing cryptographic Boolean functions is bent functions. Below is the formal definition of bent functions.

Definition. Let f be a function on V_n and ξ denote the sequence of f. Then f is called a *bent* function if $|\langle \xi, \ell_i \rangle| = 2^{\frac{n}{2}}$, $i = 0, 1, \dots, 2^n - 1$, where ℓ_i denotes the ith row of H_n.

Bent functions can be characterized in various ways [2, 8, 20, 23, 26]. A characterization of particular interest can be found in [8, 20] which states that bent functions on V_n exist only when n is even, and that they achieve the highest possible nonlinearity on V_n, namely, $2^{n-1} - 2^{\frac{n}{2}-1}$.

3. Relationships among Three Tables

Now we introduce three more notations, $k_j(\alpha)$, $\Delta_j(\alpha)$ and η_j, associated with an S-box $F = (f_1, \dots, f_m)$.

Definition. Let $F = (f_1, \ldots, f_m)$ be an $n \times m$ S-box, $\alpha \in V_n$, $j = 0, 1, \ldots, 2^m - 1$ and $\beta_j = (b_1, \ldots, b_m)$ be the vector in V_m that corresponds to the binary representation of j. In addition, set $g_j = \bigoplus_{u=1}^{m} b_u f_u$ be the jth linear combination of the component functions of F. Then we define

1. $k_j(\alpha)$ as the number of times $F(x) \oplus F(x \oplus \alpha)$ equals $\beta_j \in V_m$ while x runs through V_n once,
2. $\Delta_j(\alpha)$ as the auto-correlation of g_j with a shift α,
3. η_j as the sequence of g_j.

Since both η_0 and ℓ_0 are the all-one sequence of length 2^n and ℓ_j is $(1, -1)$ balanced for $j > 0$, we have

$$\langle \eta_0, \ell_0 \rangle = 2^n, \;\; \langle \eta_0, \ell_j \rangle = 0, \;\; j = 1, \ldots, 2^n - 1. \tag{3}$$

From the definition of $k_j(\alpha_i)$, one can see that the sum of the entries in each row of K is 2^n, and that the first row has the form of $(2^n, 0, \ldots, 0)$. Namely,

$$\sum_{j=0}^{2^m-1} k_j(\alpha_i) = 2^n, \;\; i = 0, 1, \ldots, 2^n - 1, \tag{4}$$

and

$$k_0(\alpha_0) = 2^n, \;\; k_j(\alpha_0) = 0, \;\; j = 1, \ldots, 2^m - 1. \tag{5}$$

Using the three notations introduced above, we formally define three tables/matrices related to $F = (f_1, \ldots, f_m)$.

Definition. For an S-box $F = (f_1, \ldots, f_m)$, set

$$K = \begin{bmatrix} k_0(\alpha_0) & k_1(\alpha_0) & \ldots & k_{2^m-1}(\alpha_0) \\ k_0(\alpha_1) & k_1(\alpha_1) & \ldots & k_{2^m-1}(\alpha_1) \\ & & \vdots & \\ k_0(\alpha_{2^n-1}) & k_1(\alpha_{2^n-1}) & \ldots & k_{2^m-1}(\alpha_{2^n-1}) \end{bmatrix}$$

$$D = \begin{bmatrix} \Delta_0(\alpha_0) & \Delta_1(\alpha_0) & \ldots & \Delta_{2^m-1}(\alpha_0) \\ \Delta_0(\alpha_1) & \Delta_1(\alpha_1) & \ldots & \Delta_{2^m-1}(\alpha_1) \\ & & \vdots & \\ \Delta_0(\alpha_{2^n-1}) & \Delta_1(\alpha_{2^n-1}) & \ldots & \Delta_{2^m-1}(\alpha_{2^n-1}) \end{bmatrix}$$

and

$$P = \begin{bmatrix} \langle \eta_0, \ell_0 \rangle^2 & \langle \eta_1, \ell_0 \rangle^2 & \cdots & \langle \eta_{2^m-1}, \ell_0 \rangle^2 \\ \langle \eta_0, \ell_1 \rangle^2 & \langle \eta_1, \ell_1 \rangle^2 & \cdots & \langle \eta_{2^m-1}, \ell_1 \rangle^2 \\ & & \vdots & \\ \langle \eta_0, \ell_{2^n-1} \rangle^2 & \langle \eta_1, \ell_{2^n-1} \rangle^2 & \cdots & \langle \eta_{2^m-1}, \ell_{2^n-1} \rangle^2 \end{bmatrix}$$

where ℓ_i is the ith row of H_n, $i = 0, 1, \ldots, 2^n - 1$. The three tables (or matrices) K, D and P share the same size of $2^n \times 2^m$. Clearly K is the difference distribution table of F that has already been (informally) introduced in Section 2. The other two tables, D and P, are called *auto-correlation distribution table and correlation immunity distribution table* of the S-box F, respectively.

In designing a strong S-box, many cryptographic criteria should be examined not only against component functions, but also against their linear combinations. Such criteria include those related to nonlinearity, propagation characteristics [19] and difference distribution tables. The matrix K characterizes the differential characteristics of an S-box. The matrix D indicates the auto-correlation of all linear combinations of the component functions. While the matrix P represents the inner product between the sequence of each linear combination of the component functions and each linear sequence. P is helpful in studying the correlation immunity, as well as the nonlinearity, of each linear combination of the component functions (see [22]).

The following lemma shows an intimate relationship between the three tables K, D and P defined above. The lemma can be easily shown to be correct by the use of a connection between the Hamming distance between rows and the distribution of ones in the columns in a (0, 1) matrix. For completeness, a full proof for the lemma is provided in the appendix. It turns out that the lemma is very useful in examining cryptographic properties of an S-box, and it will be used in proving many of the main results in this paper.

LEMMA 2 *Let $F = (f_1, \ldots, f_m)$ be a mapping from V_n to V_m, where n and m are integers with $n \geq m \geq 1$ and each f_j is a function on V_n. Set $g_j = \bigoplus_{u=1}^{m} c_u f_u$ where $(c_1, \ldots, c_m)$ is the binary representation of an integer j, $j = 0, 1, \ldots, 2^m - 1$. Then*

(i) $(k_0(\alpha_i), k_1(\alpha_i), \ldots, k_{2^m-1}(\alpha_i))H_m = (\Delta_0(\alpha_i), \Delta_1(\alpha_i), \ldots, \Delta_{2^m-1}(\alpha_i))$, *where α_i is the binary representation of an integer i,*

(ii) $D = KH_m$,

(iii) $P = H_n D$,

(iv) $P = H_n K H_m$.

When $n = m$ and F is regular, a similar relation between matrices K and P has been derived in [7]. As permutations, a special type of S-boxes, are used in many cryptographic algorithms, it is of interest is to look into how the three tables of a permutation are connected to the three corresponding tables of the inverse of the permutation. The following result is easy to verify.

COROLLARY 1 *Let F be a permutation on V_n and F^{-1} denote the inverse of F. Let $K = (k_i(\alpha_j))$, $D = (\Delta_i(\alpha_j))$ and $P = (\langle \eta_i, \ell_j \rangle)$ be the difference distribution, auto-*

correlation distribution and correlation immunity distribution tables of F. *Similarly, let* $K^* = (k_i^*(\alpha_j))$, $D^* = (\Delta_i^*(\alpha_j))$ *and* $P^* = (\langle \eta_i^*, \ell_j \rangle)$ *be the difference distribution, auto-correlation distribution and correlation immunity distribution tables of* F^{-1}. *Then*

(i) $K^* = K^T$,

(ii) $P^* = P^T$,

(iii) $D^* = H_n^{-1} D^T H_n$.

4. Regularity of S-boxes and Difference Distribution Tables

Using Lemma 2, we now show that the regularity of an S-box can be characterized by its difference distribution table. This characterization nicely complements Lemma 1 which is stated in terms of the balance of non-zero linear combinations of component functions of an S-box.

COROLLARY 2 *Let* $F = (f_1, \ldots, f_m)$ *be a mapping from* V_n *to* V_m, *where* n *and* m *are integers with* $n \geq m \geq 1$ *and each* f_j *is a function on* V_n. *Then* F *is regular if and only if the sum of the entries in each column in the difference distribution table is* 2^{2n-m}, *i.e.,* $\sum_{\alpha \in V_n} k_i(\alpha) = 2^{2n-m}$, $i = 0, 1, \ldots, 2^m - 1$.

Proof. Compare the first rows in both sides of the formula in Part (iv) of Lemma 2,

$$\left(\sum_{\alpha \in V_n} k_0(\alpha), \sum_{\alpha \in V_n} k_1(\alpha), \ldots, \sum_{\alpha \in V_n} k_{2^m-1}(\alpha) \right) H_m$$
$$= (\langle \eta_0, \ell_0 \rangle^2, \langle \eta_1, \ell_0 \rangle^2, \ldots, \langle \eta_{2^m-1}, \ell_0 \rangle^2). \tag{6}$$

Obviously, if $\sum_{\alpha \in V_n} k_i(\alpha) = 2^{2n-m}$, $i = 0, 1, \ldots, 2^m - 1$ then $\langle \eta_1, \ell_0 \rangle^2 = \cdots = \langle \eta_{2^m-1}, \ell_0 \rangle^2 = 0$. Note that ℓ_0 is the all-one sequence of length 2^n. Hence $g_1, \ldots, g_{2^m-1}$ are balanced, where $g_1, \ldots, g_{2^m-1}$ are defined in Lemma 2. By Lemma 1, F is regular.

Conversely, suppose F is regular. By Lemma 1, $g_1, \ldots, g_{2^m-1}$ are balanced. Hence $\langle \eta_1, \ell_0 \rangle^2 = \cdots = \langle \eta_{2^m-1}, \ell_0 \rangle^2 = 0$. Note that $\langle \eta_0, \ell_0 \rangle^2 = 2^{2n}$. Rewrite (6) as

$$2^m \left(\sum_{\alpha \in V_n} k_0(\alpha), \sum_{\alpha \in V_n} k_1(\alpha), \ldots, \sum_{\alpha \in V_n} k_{2^m-1}(\alpha) \right) = (2^{2n}, 0, \ldots, 0) H_m.$$

This proves that $\sum_{\alpha \in V_n} k_i(\alpha) = 2^{2n-m}$, $i = 0, 1, \ldots, 2^m - 1$. ∎

Corollary 2 has also been obtained independently by Tapia-Recillas, Daltabuit and Vega [25].

The following corollary shows the uniqueness of the leftmost column of the difference distribution table of a regular mapping.

THEOREM 1 *Let $F = (f_1, \ldots, f_m)$ be a mapping from V_n to V_m, where n and m are integers with $n \geq m \geq 1$ and each f_j is a function on V_n. Then*

(i) $\sum_{\alpha \in V_n} k_0(\alpha) \geq 2^{2n-m}$,

(ii) the equality in (i) holds if and only if F is regular.

Proof. (i) Right-multiplying both sides of the equality in Part (iv) of Lemma 2 by e^T where, e denotes the all-one sequence of length 2^m. Hence we have

$$H_n \begin{bmatrix} k_0(\alpha_0) & k_1(\alpha_0) & \ldots & k_{2^m-1}(\alpha_0) \\ k_0(\alpha_1) & k_1(\alpha_1) & \ldots & k_{2^m-1}(\alpha_1) \\ & \vdots & & \\ k_0(\alpha_{2^n-1}) & k_1(\alpha_{2^n-1}) & \ldots & k_{2^m-1}(\alpha_{2^n-1}) \end{bmatrix} \begin{bmatrix} 2^m \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_0 \rangle^2 \\ \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2 \\ \vdots \\ \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2 \end{bmatrix}$$

and hence

$$2^m H_n \begin{bmatrix} k_0(\alpha_0) \\ k_0(\alpha_1) \\ \vdots \\ k_0(\alpha_{2^n-1}) \end{bmatrix} = \begin{bmatrix} \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_0 \rangle^2 \\ \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2 \\ \vdots \\ \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2 \end{bmatrix}. \tag{7}$$

Comparing the top element of the vector on the two sides of equality (7), the following is obtained

$$2^m \sum_{i=0}^{2^n-1} k_0(\alpha_i) = \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_0 \rangle^2. \tag{8}$$

Recall (3), $\langle \eta_0, \ell_0 \rangle^2 = 2^{2n}$. From (8), we have proved Part (i) of the theorem.

(ii) Suppose $\sum_{\alpha \in V_n} k_0(\alpha) = 2^{2n-m}$, then from (8), $\langle \eta_1, \ell_0 \rangle^2 = \cdots = \langle \eta_{2^m-1}, \ell_0 \rangle^2 = 0$. Note that ℓ_0 is the all-one sequence of length 2^n. Hence $g_1, \ldots, g_{2^m-1}$ are balanced, where $g_1, \ldots, g_{2^m-1}$ are defined in Lemma 2. By Lemma 1, F is regular.

Conversely, if F is regular, then by Corollary 2, $\sum_{\alpha \in V_n} k_0(\alpha) = 2^{2n-m}$. The proof of the theorem is completed. ■

5. A Lower Bound on Differential Uniformity

We turn our attention back to the differential uniformity, denoted by δ, of an $n \times m$ S-box. Recall that δ is defined as the largest value in the differential distribution table of the S-box (not taking into account the leftmost entry in the top row), namely,

$$\delta = \max_{\alpha \in V_n, \alpha \neq 0} \max_{\beta \in V_m} \#\{x \mid F(x) \oplus F(x \oplus \alpha) = \beta\}$$

(See Definition 2). As discussed earlier, δ is bounded by $2^{n-m} \leq \delta \leq 2^n$, and generally speaking S-boxes with a smaller δ are desirable in designing a block cipher secure against differential attacks. This motivates us to improve the "trivial" lower bound 2^{n-m} on the differential uniformity δ.

The following lemma will be used in our discussions. It is identical to Lemma 2 of [27].

LEMMA 3 *Let real valued sequences* $a_0, \ldots, a_{2^n-1}$ *and* $b_0, \ldots, b_{2^n-1}$ *satisfy*

$$(a_0, \ldots, a_{2^n-1})H_n = (b_0, \ldots, b_{2^n-1}).$$

For any integer p *and* q, $p + q = n$, $1 \le p, q \le n-1$, *set* $\sigma_j = \sum_{s=0}^{2^q-1} b_{j2^q+s}$, *where* $j = 0, 1, \ldots, 2^p - 1$. *Then*

$$2^q(a_0, a_{2^q}, a_{2\cdot 2^q}, \ldots, a_{(2^p-1)2^q})H_p = (\sigma_0, \sigma_1, \ldots, \sigma_{2^p-1}). \tag{9}$$

Now we prove another main result of this paper.

THEOREM 2 *Let* $F = (f_1, \ldots, f_m)$ *be an* $n \times m$ *S-box, where* n *and* m *are integers with* $n \ge m \ge 1$ *and each* f_j *is a function on* V_n. *Set* $g_j = \bigoplus_{u=1}^m c_u f_u$ *where* $(c_1, \ldots, c_m)$ *is the binary representation of an integer* j, $j = 0, 1, \ldots, 2^m - 1$. *Denote by* $\Delta_j(\alpha)$ *the auto-correlation of* g_j *with a shift* α, *and set* $\Delta_M = \max_{\alpha \in V_n, \alpha \ne 0} \max_{j=1,\ldots,2^m-1}\{|\Delta_j(\alpha)|\}$. *Then the differential uniformity* δ *of* F *is bounded from below by* $2^{n-m} + 2^{-m}\Delta_M$, *namely,* $\delta \ge 2^{n-m} + 2^{-m}\Delta_M$.

Proof. Let $\Delta_{j'}(\alpha_{i'}) = \Delta_M$. By Part (i) of Lemma 2, we have

$$2^{-m}(\Delta_0(\alpha_{i'}), \Delta_1(\alpha_{i'}), \ldots, \Delta_{2^m-1}(\alpha_{i'}))H_m = (k_0(\alpha_i'), k_1(\alpha_i'), \ldots, k_{2^m-1}(\alpha_i')) \tag{10}$$

Applying Lemma 3 to (10), we get

$$2^{m-1}2^{-m}(\Delta_0(\alpha_{i'}), \Delta_{2^{m-1}}(\alpha_{i'}))H_1 = (\sigma_0, \sigma_1)$$

where $\sigma_j = \sum_{s=0}^{2^{m-1}-1} k_{j2^{m-1}+s}$, $j = 0, 1$. Hence

$$2^{-1}(\Delta_0(\alpha_{i'}) + \Delta_{2^{m-1}}(\alpha_{i'})) = \sigma_0$$

and

$$2^{-1}(\Delta_0(\alpha_{i'}) - \Delta_{2^{m-1}}(\alpha_{i'})) = \sigma_1$$

Thus there is a $j_0 2^q + s_0$ for $0 \le s_0 \le 2^{m-1} - 1$ and $j_0 = 0$ or 1, such that

$$k_{j_0 2^q + s_0} \ge 2^{-m}(\Delta_0(\alpha_{i'}) + \Delta_{2^{m-1}}(\alpha_{i'})).$$

Recall that $\Delta_0(\alpha) = 2^n$ for all $\alpha \in V_n$. So we have

$$k_{j_0 2^q + s_0} \ge 2^{-m}(2^n + \Delta_{2^{m-1}}(\alpha_{i'})).$$

According to Section 5.3 of [21], the differential uniformity of F is invariant under a nonsingular linear transformation on the variables of F. Thus by choosing an appropriate nonsingular linear transformation on the variables of F, we have

$$k_{j_0 2^q + s_0} \geq 2^{n-m} + 2^{-m} \Delta_M$$

and hence $\delta \geq 2^{n-m} + 2^{-m} \Delta_M$. ■

Examining the new lower bound of $2^{n-m} + 2^{-m} \Delta_M$ on the differential uniformity δ, where Δ_M is the largest value among all $|\Delta_j(\alpha)|$ with $j = 1, \ldots, 2^m - 1$, $\alpha \in V_n$ and $\alpha \neq 0$, a natural question would be how large and small Δ_M can be and what could be its typical value.

First of all, by the definition of Δ_M, we have $0 \leq \Delta_M \leq 2^n$. When $\Delta_M = 0$, every non-zero linear combination of the components of F must be a bent function. And the converse is also true: if every non-zero linear combination of the components of F is a bent function, then we must have $\Delta_M = 0$. Note that in this case we have $\delta = 2^{n-m}$, which indicates that the bound in Theorem 2 is tight. Also note that such S-boxes do exist [1, 15], although they are not regular.

On the other hand, if $\Delta_M = 2^n$, then there must exist a non-zero vector α such that it is a linear structure of a non-zero linear combination, say g_j, of the component functions of F, i.e., $g_j(x) \oplus g_j(x \oplus \alpha)$ is a constant. Similarly, the converse is also true.

For other S-boxes, namely those whose non-zero linear combinations of component functions are not all bent, and do not have non-zero linear structures, their Δ_M will be a value between 0 and 2^n. Although it is not quite clear as to what would be the typical value of Δ_M for such S-boxes, from the bound $\delta \geq 2^{n-m} + 2^{-m} \Delta_M$, at least one thing can be said: if an S-box is designed to resist against differential attacks, then its differential uniformity must be small, and hence its Δ_M must be small too; conversely, if an S-box has a small Δ_M, we would expect that it could have a small differential uniformity too.

6. Upper Bounds on Nonlinearity of S-boxes

After the discovery of differential attacks in [4], an equally notable cryptanalysis method, the linear cryptanalytic attack, was subsequently introduced in [13]. Identifying relationships between these two types of attacks has been an interesting research area, both from the view point of cryptanalysis and the design of secure ciphers. We will first show a tight upper bound on the nonlinearity of a general S-box. This will be followed by another upper bound on the nonlinearity of a regular S-box. The usefulness of such an explicit relationship is obvious: the nonlinearity of an S-box represents a key indicator for the strength of a block cipher that employs the S-box. We also compare our result on the relationship with a related theorem in [6].

In studying an $n \times m$ S-box, a possible approach would be to use the two parameters n and m alone in determining information on the S-box. Success of this approach, however, seems to have been limited to the case of $m \geq n - 1$ with which an upper bound on nonlinearity has been obtained in [6] (but see discussions in the closing paragraph of this section.)

Another approach that can be used to obtain far more detailed information on an S-box is to take into account all the $k_j(\alpha)$, $\Delta_j(\alpha)$, or $\langle \eta_j, \ell_i \rangle^2$, for $j = 0, 1, \ldots, 2^m - 1$,

$i = 0, 1, \ldots, 2^n - 1$ and $\alpha \in V_n$ (see Definition 3). A potential problem with this approach is that it would be impractical to apply it to an S-box with relatively large n and/or m. In what follows, we adopt a different approach that employs more parameters other than n and m, and hence can be viewed as a compromise between the above two approaches. More specifically, we prove two theorems that relate the nonlinearity of an $n \times m$ S-box to three parameters, namely n, m and the number of non-zero entries in its difference distribution table K.

6.1. General Case

Here we consider $n \times m$ S-box that is not necessarily regular. In addition, the restriction of $n \geq m$ is not imposed on the S-box. We first introduce Hölder's Inequality which can be found in [9].

LEMMA 4 *Let $c_j \geq 0$ and $d_j \geq 0$ be real numbers, where $j = 1, \ldots, s$, and let p and q satisfy $\frac{1}{p} + \frac{1}{q} = 1$ and $p > 1$. Then*

$$\left(\sum_{j=1}^{s} c_j^p\right)^{1/p} \left(\sum_{j=1}^{s} d_j^q\right)^{1/q} \geq \sum_{j=1}^{s} c_j d_j$$

where the quality holds if and only if $c_j = \nu d_j$, $j = 1, \ldots, s$ for a constant $\nu \geq 0$.

When c_j, d_j, p and q satisfy the condition that $c_j \geq 0$, $d_j = \begin{cases} 1 & \text{if } c_j = 1 \\ 0 & \text{if } c_j = 0 \end{cases}$, and $p = q = \frac{1}{2}$, Hölder's Inequality gives

$$\sum_{j=1}^{s} c_j^2 \geq s^{-1} \left(\sum_{j=1}^{s} c_j\right)^2 \tag{11}$$

where the quality holds if and only if $c_1, \ldots, c_s$ are all identical. The inequality (11) will be used in the proof of the following two theorems regarding the upper bound on the nonlinearity of an S-box.

THEOREM 3 *Let F be an $n \times m$ S-box (F is not necessarily regular, and the restriction of $n \geq m$ is not imposed on it). Denote by T_{nz} the total number of all non-zero entries, except for $k_0(\alpha_0)$, in the difference distribution table K of the S-box (see Definition 3). Then*

(i) the nonlinearity of F satisfies

$$N_F \leq 2^{n-1} - \frac{1}{2}\left(\frac{2^{2n+m} - 2^{3n} + T_{nz}^{-1} 2^{2n+m}(2^n - 1)^2}{2^m - 1}\right)^{\frac{1}{4}},$$

(ii) the equality in (i) holds if and only if every non-zero linear combination of the component functions of F is a bent function.

Proof. We first prove Part (i) of the theorem. Using Part (iv) of Lemma 2, we have

$$P^T P = H_m K^T H_n^T H_n K H_m = 2^n H_m K^T K H_m = 2^{n+m} H_m^{-1} K^T K H_m.$$

Note that the sum of entries on the diagonal of $P^T P$ is equal to the sum of entries on the diagonal of $2^{n+m} K^T K$. Hence

$$\sum_{j=0}^{2^m-1} \sum_{i=0}^{2^n-1} \langle \eta_j, \ell_i \rangle^4 = 2^{n+m} \sum_{j=0}^{2^m-1} \sum_{i=0}^{2^n-1} k_j^2(\alpha_i).$$

From (3), (4) and (5) in Section 3, we have

$$2^{4n} + \sum_{j=1}^{2^m-1} \sum_{i=0}^{2^n-1} \langle \eta_j, \ell_i \rangle^4 = 2^{n+m} \left(2^{2n} + \sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j^2(\alpha_i) \right).$$

Now combining (4) with (11), a special form of Hölder's Inequality, we have

$$\sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j^2(\alpha_i) \geq T_{nz}^{-1} \left(\sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j(\alpha_i) \right)^2 = T_{nz}^{-1} 2^{2n} (2^n - 1)^2. \tag{12}$$

Hence there is a certain j_0, $1 \leq j_0 \leq 2^m - 1$, and a certain i_0, $0 \leq i_0 \leq 2^n - 1$, such that

$$\langle \eta_{j_0}, \ell_{i_0} \rangle^4 \geq \frac{2^{n+m}(2^{2n} + T_{nz}^{-1} 2^{2n}(2^n-1)^2) - 2^{4n}}{(2^m-1)2^n}$$

which implies

$$|\langle \eta_{j_0}, \ell_{i_0} \rangle| \geq \left(\frac{2^{2n+m} - 2^{3n} + T_{nz}^{-1} 2^{2n+m}(2^n-1)^2}{2^m - 1} \right)^{\frac{1}{4}}.$$

Now applying (1) we obtain Part (i) of the theorem.

Note that since $T_{nz} \leq 2^m(2^n - 1)$, we have

$$T_{nz}^{-1} 2^{2n+m} (2^n - 1)^2 \geq 2^{2n}(2^n - 1).$$

That is, the expression under the fourth root is always positive.

Now we prove Part (ii). First assume that the equality in Part (i) holds. From the definition of N_F, as well as (1), we have

$$|\langle \eta_j, \ell_i \rangle| \leq \left(\frac{2^{2n+m} - 2^{3n} + T_{nz}^{-1} 2^{2n+m}(2^n-1)^2}{2^m - 1} \right)^{\frac{1}{4}} \tag{13}$$

for all $j = 1, \ldots, 2^m - 1$ and $i = 0, 1, \ldots, 2^n - 1$. Returning to the proof of Part (i), we can see that (13) implies that the equality on the left hand side of (12) must hold. Namely,

$$\sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j^2(\alpha_i) = T_{nz}^{-1} \left(\sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j(\alpha_i) \right)^2.$$

Again using (11), the special form of Hölder's Inequality, there exists a constant k such that $k_j(\alpha_i) = k$, for all $j = 0, 1, \ldots, 2^m - 1$ and $i = 1, \ldots, 2^n - 1$. From (4), the constant k must satisfy the condition of $k = 2^{n-m}$. Note also that in this case, $T_{nz} = 2^m(2^n - 1)$. Thus due to Theorem 3.1 in [15], we conclude that every non-zero linear combination of the component functions of F is a bent function. A consequence of this conclusion is that in this case, n must be even and $m \leq \frac{1}{2}n$ [15].

Conversely, assume that every non-zero linear combination of the component functions of F is a bent function. Once again employing Theorem 3.1 in [15], we have $k_j(\alpha_i) = 2^{n-m}$ for $j = 0, 1, \ldots, 2^m - 1$ and $i = 1, \ldots, 2^n - 1$. In this case, the total number of non-zero entries in the table K is $T_{nz} = 2^m(2^n - 1)$. Now the inequality in Part (i) of the theorem becomes

$$N_F \leq 2^{n-1} - 2^{\frac{n}{2}-1}. \tag{14}$$

On the other hand, since every non-zero linear combination of the component functions of F is a bent function, the equality in (14) must hold. That is, the equality in Part (i) of the theorem holds. ■

We note that for a permutation on V_n, results obtained in [18] imply that the expected value of T_{nz} approaches $(1 - e^{-\frac{1}{2}})(2^n - 1)^2$, when n is large, where $e = 2.718\ldots$. By using Theorem 3, the expected value of N_F for a permutation satisfies

$$N_F \leq 2^{n-1} - \frac{2^{\frac{3}{4}n-1}}{\sqrt[4]{(1 - e^{-\frac{1}{2}})(2^n - 1)}}$$

Before moving on to the next topic on regular S-boxes, we would like to stress that Theorem 3 shows a tight upper bound on the nonlinearity of a general S-box which does not have to be regular. We also note that an S-box that achieves the upper bound in theorem has a flat difference distribution table and hence is weak against differential cryptanalysis.

6.2. For a Regular S-box

As we mentioned earlier, most encryption algorithms employ regular S-boxes. Hence such S-boxes play a more important role than does an irregular one. Our research results to be described below show that the nonlinearity of a regular $n \times m$ S-box can be determined by n, m and a third parameter that counts only the number of non-zero entries in the leftmost column of the difference distribution table of the S-box.

We begin with examining partitions of the leftmost column of a difference distribution table.

LEMMA 5 *Let F be a mapping from V_n to V_m and K is the difference distribution table of F. Then the leftmost column of K is determined by a 2^m-partition of V_n, say $V_n = \Omega_0 \cup \cdots \cup \Omega_{2^m-1}$, that satisfies the condition that $\Omega_j \cap \Omega_i = \emptyset$ for all $j \neq i$.*

Proof. For each $\beta \in V_m$, define $\Omega_\beta = \{\alpha \in V_n \mid F(\alpha) = \beta\}$. Note that we use an integer in $[0, \dots, 2^m - 1]$ and a vector in V_m interchangeably. Clearly

$$V_n = \cup_{\beta \in V_m} \Omega_\beta \tag{15}$$

and $\Omega_{\beta'} \cap \Omega_{\beta''} = \emptyset$ if $\beta' \neq \beta''$. Note that $F(x) \oplus F(x \oplus \alpha) = 0$ if and only if both x and $x \oplus \alpha$ belong to the same class, say Ω_β.

Now we modify the mapping F into F' by applying an arbitrary permutation on V_m to the output of F. Clearly the partition in (15) remains unchanged, and $F'(x) \oplus F'(x \oplus \alpha) = 0$ if and only if both x and $x \oplus \alpha$ belong to the same class in (15). This proves that the leftmost columns of the difference distribution tables of F and F' are the same. ■

Armed with Lemma 5, we are ready to prove the following.

THEOREM 4 *Let F be a regular $n \times m$ S-box (For such an S-box $n \geq m$ is necessary). Denote by t_{nz} the total number of non-zero entries (except for $k_0(\alpha_0)$) in the leftmost column of the difference distribution table K of F. Then the nonlinearity of F satisfies*

$$N_F \leq 2^{n-1} - \frac{1}{2}\left(\frac{2^{3n+2m} - 2^{4n} + t_{nz}^{-1} \cdot 2^{3n+2m}(2^{n-m} - 1)^2}{(2^n - 1)(2^m - 1)^2}\right)^{\frac{1}{4}}.$$

Proof. Left-multiplying the transposes of the two sides in (7), we have

$$\left(\sum_{j=0}^{2^m-1} \langle \eta_j, \ell_0 \rangle^2\right)^2 + \left(\sum_{j=0}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2\right)^2 + \cdots + \left(\sum_{j=0}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2\right)^2$$
$$= 2^{2m+n} \sum_{i=0}^{2^n-1} k_0^2(\alpha_i) \tag{16}$$

Since both η_0 and ℓ_0 are an all-one sequence, we have $\langle \eta_0, \ell_0 \rangle = 2^n$. Recall that F is regular. By Lemma 1, each non-zero linear combination of the component functions of F is balanced. Thus for $j = 1, \dots, 2^m - 1$, η_j is $(1, -1)$ balanced and we have $\langle \eta_j, \ell_0 \rangle = 0$. Also recall the definition in (3) and the fact that ℓ_j is $(1, -1)$ balanced for $j > 0$, we can see that $\langle \eta_0, \ell_j \rangle = 0$ for $j = 1, \dots, 2^n - 1$.

Note that $k_0(\alpha_0) = 2^n$. So (16) can be transformed to

$$\left(\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2\right)^2 + \cdots + \left(\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2\right)^2$$
$$= 2^{2m+3n} - 2^{4n} + 2^{2m+n} \sum_{i=1}^{2^n-1} k_0^2(\alpha_i) \tag{17}$$

By using (11)

$$\sum_{i=1}^{2^n-1} k_0^2(\alpha_i) \geq t_{nz}^{-1} \left(\sum_{i=1}^{2^n-1} k_0(\alpha_i)\right)^2.$$

Note that F is regular and $k_0(\alpha_0) = 2^k$. By using Corollary 1, $\sum_{i=1}^{2^n-1} k_0(\alpha_i) \geq 2^{2n-m} - 2^n$. Hence

$$\sum_{i=1}^{2^n-1} k_0^2(\alpha_i) \geq t_{nz}^{-1} \cdot (2^{2n-m} - 2^m)^2.$$

Thus there is an i_0, $1 \leq i_0 \leq 2^n - 1$, such that

$$\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_{i_0} \rangle^2 \geq \left(\frac{2^{3n+2m} - 2^{4n} + t_{nz}^{-1} \cdot 2^n (2^{2n} - 2^{n+m})^2}{2^n - 1} \right)^{\frac{1}{2}}.$$

Since $t_{nz} \leq 2^n - 1$, it is easy to verify that the expression under the square root is always positive. Furthermore there is a j_0, $1 \leq j_0 \leq 2^m - 1$, such that

$$|\langle \eta_{j_0}, \ell_{i_0} \rangle| \geq \left(\frac{2^{3n+2m} - 2^{4n} + t_{nz}^{-1} \cdot 2^n (2^{2n} - 2^{n+m})^2}{(2^n - 1)(2^m - 1)^2} \right)^{\frac{1}{4}}.$$

Now the theorem follows immediately from (1). ■

For a permutation F on V_n, (F must be regular), again from results obtained in [18], we know that the expected value of t_{nz} approaches $(1 - e^{-\frac{1}{2}})(2^n - 1)$, while n is large enough, where $e = 2.718\ldots$. This, together with Theorem 4, shows that the expected value of N_F for regular S-boxes is bounded from above by $2^{n-1} - \frac{2^{n-1}}{\sqrt{2^n-1}}$. Namely,

$$N_F \leq 2^{n-1} - \frac{2^{n-1}}{\sqrt{2^n - 1}}.$$

6.3. Remarks on the Two Upper Bounds

Comparing Theorem 3 with Theorem 4, we note that while the former deals with a general S-box which is not necessarily regular, the latter is strictly on a regular S-box. Therefore the condition that $n \geq m$ is required only in Theorem 4. In addition to n and m, both theorems employ a third parameter in upper bounding the nonlinearity of an S-box. The third parameter T_{nz} used in Theorem 3 is the total number of non-zero entries in the *entire* difference distribution table of the S-box (not taking into account the first entry in the leftmost column). In contrast, the third parameter t_{nz} used in Theorem 4 is the total number of non-zero entries in the *leftmost column* in the difference distribution table of the S-box (again not taking into account the first entry in the column).

Another difference between Theorems 3 and 4 is that while the bound in the former is tight, it is unclear whether the same can be said with the latter. This is, however, not surprising, given that identifying the exact upper bound on the nonlinearity of a balanced function is one of the outstanding open problems in the study of nonlinear Boolean functions.

A direct consequence of Theorem 3 is that with any $n \times m$ S-box with $n > m$, be it regular or irregular, the larger the number of non-zero entries in the difference distribution table, the larger the upper bound on the nonlinearity of the S-box. To interpret the theorem in a different way, if one wishes to design an S-box that is resistant against linear attacks, namely highly nonlinear, then one should make sure that a large portion of entries in the difference distribution table of the S-box is non-zero. Interestingly, as a larger T_{nz} also means a wider spread of non-zero entries across the entire difference distribution table, such an S-box can potentially have a higher resilience against differential attacks.

What Theorem 4 implies is that for a regular S-box, t_{nz}, the number of non-zero entries in the leftmost column of its difference distribution table, effects the resistance against linear and differential attacks in a way similar to that of T_{nz}. Thus, in designing a regular S-box, one prefers both a large t_{nz} and a large T_{nz}. It should be pointed out, however, that other factors should be taken into account too. Examples of such factors include successful attacks that exploit non-zero entries in the leftmost column of a difference distribution table [4, 5, 21], and high order differential attacks recently developed in [10].

Before closing this section, we note that a paper by Chabaud and Vaudenay [6] is a prior work most relevant to this research. A main result in [6] is their Theorem 4 which is equivalent to stating that for every mapping from V_n to V_m, say F, the nonlinearity of F, N_F, satisfies

$$N_F \leq 2^{n-1} - \frac{1}{2}\left(3 \cdot 2^n - 2 - \frac{2(2^n-1)(2^{n-1}-1)}{2^m - 1}\right)^{\frac{1}{2}}. \tag{18}$$

Examining the part under the square root in the expression, one can see that it is negative if $m \leq n-2$. Therefore, (18) is applicable only to $n \times m$ S-boxes with $m \geq n-1$.

7. Concluding Remarks

We have introduced three tables associated with an S-box, and based on a relationship among the three tables, we have established a number of results ranging from regularity, nonexistence of certain quadratic S-boxes, to a tight lower bound on the differential uniformity and two upper bounds on the nonlinearity of an S-box.

In light of recent progress in interpolation and high order differential cryptanalysis [10, 24], a natural topic that deserves immediate attention is to research into high order differential distribution tables of S-boxes, together with connections to other cryptographic properties of S-boxes.

Acknowledgments

The first author was supported by a Queen Elizabeth II Research Fellowship (227 23 1002). Part of the second author's work was completed while on sabbatical at the University of Tokyo.

Appendix: The Proof of Lemma 2

There are close relationships between the Hamming distance between rows and the distribution of ones in the columns in a (0, 1) matrix. Such relationships have been very useful in constructing linear error correcting codes. In this appendix we review some of the relationships from the view point of Hadamard transforms. Once the relationships are clear, the proof of Lemma 2 becomes straightforward.

Let $t \geq s$, and A be an $s \times t$ (0, 1) matrix with rank s. Set

$$A = \begin{bmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_{s-1} \end{bmatrix} = (a_{ij}) = [\chi_0, \chi_1, \ldots, \chi_{t-1}], \tag{19}$$

where $\xi_i \in V_t$ is the ith row vector and $\chi_j \in V_s$ is the jth column vector of A.

We are concerned with all the linear combinations of $\xi_0, \xi_1, \ldots, \xi_{s-1}$, denoted by $\eta_0, \eta_1, \ldots, \eta_{2^s-1}$, where $\eta_j = \bigoplus_{u=0}^{s-1} c_u \xi_u$, $(c_0, c_1, \ldots, c_{s-1})$ is the binary representation of an integer j, $j = 0, 1 \ldots, 2^s - 1$. Now set

$$B = \begin{bmatrix} \eta_0 \\ \eta_1 \\ \vdots \\ \eta_{2^s-1} \end{bmatrix} = (b_{ij}) = [\gamma_0, \gamma_1, \ldots, \gamma_{t-1}], \tag{20}$$

where B is a (0, 1) matrix of order $2^s \times t$ and $\gamma_j \in V_{2^s}$ is the jth column vector of B. Replace every 0 entry in B with 1, and every 1 entry in B with -1. Then denote by B^* the new $(1, -1)$ matrix of order $2^s \times t$. Write

$$B^* = (b^*_{ij}) = \begin{bmatrix} R_0 \\ R_1 \\ \vdots \\ R_{2^s-1} \end{bmatrix} = [h_0, h_1, \ldots, h_{t-1}], \tag{21}$$

where R_i is the ith row vector and h_j is the jth column vector of B^*. One can verify that each h_j is a linear sequence of length 2^s.

Let B^* be the matrix defined in (21), $e_0, e_1, \ldots, e_{2^s-1}$ be the row vectors, from the top to the bottom, of H_s. Assume that e_j appears k_j times in the columns of B^*. We now prove

$$e_i B^* B^{*T} e_j^T = \begin{cases} k_j 2^{2s} & \text{if } e_i = e_j \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Write $e_i B^* = (c^*_0, \ldots, c^*_{t-1})$ where

$$c^*_u = \begin{cases} 2^s & \text{if } e_i^T = h_u \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

for all $u = 0, \ldots, t-1$. Similarly, write $e_j B^* = (d_0^*, \ldots, d_{t-1}^*)$, where

$$d_u^* = \begin{cases} 2^s & \text{if } e_j^T = h_u \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

for all $u = 0, \ldots, t-1$.

If $e_i = e_j$, then $e_i B^* B^{*T} e_j^T = \sum_{u=0}^{t-1} c_u^* c_u^* = k_j 2^{2s}$. On the other hand, if $e_i \neq e_j$, then by (23) and (24), $c_u^* \neq 0$ implies $d_u^* = 0$, which results in $e_i B^* B^{*T} e_j^T = \sum_{u=0}^{t-1} c_u^* d_u^* = 0$. This proves (22).

As the Sylvester-Hadamard matrix H_m is symmetric, (22) can be equivalently stated as:

$$H_s B^* B^{*T} H_s = 2^{2s} \operatorname{diag}(k_0, k_1, \ldots, k_{2^s-1}). \tag{25}$$

Let R_j be a row of B^* defined in (21) and k_j the number of times a row vector e_j in H_s appears in the columns of B^*. From (25) we have $B^* B^{*T} = H_s \operatorname{diag}(k_0, k_1, \ldots, k_{2^s-1}) H_s$. Comparing the first rows in the two sides of the equation, we have

$$(\langle R_0, R_0\rangle, \langle R_0, R_1\rangle, \ldots, \langle R_0, R_{2^s-1}\rangle) = (k_0, k_1, \ldots, k_{2^s-1}) H_s. \tag{26}$$

Now we are in a position to prove Lemma 2. Consider an $s \times t$ matrix A defined in (19) with $s = m$ and $t = n$. Let a row ξ_i in (19) be the truth table of $f_i(x) \oplus f_i(x \oplus \alpha)$, $i = 0, 1, \ldots, m-1$. Correspondingly, η_i in (20) denotes the truth table of $g_i(x) \oplus g_i(x \oplus \alpha)$, and R_i in (21) denotes the sequence of $g_i(x) \oplus g_i(x \oplus \alpha)$, $i = 0, 1, \ldots, 2^m - 1$.

As g_0 is the zero function, R_0 is the all-one sequence. Hence $\langle R_0, R_i\rangle$ is equal to the sum of the components in R_i. That is, $\langle R_0, R_i\rangle = \Delta_i(\alpha)$. Hence Part (i) of Lemma 2 follows from (26).

For $\alpha = \alpha_0, \alpha_1, \ldots, \alpha_{2^n-1}$, Part (i) of Lemma 2 gives 2^n equations. These equations can be written as Part (ii) of the lemma. Part (iii) of the lemma follows from (2). And finally Parts (ii) and (iii) of the lemma together give Part (iv) of the lemma.

References

1. C. M. Adams, On immunity against Biham and Shamir's "differential cryptanalysis, *Information Processing Letters*, Vol. 41 (1992) pp. 77–80.
2. C. M. Adams and S. E. Tavares. Generating and counting binary bent sequences, *IEEE Transactions on Information Theory*, Vol. IT-36 No. 5 (1990) pp. 1170–1173.
3. E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems, *Journal of Cryptology*, Vol. 4, No. 1 (1991) pp. 3–72.
4. E. Biham and A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, New York, Heidelberg, Tokyo (1993).
5. L. Brown, M. Kwan, J. Pieprzyk, and J. Seberry. Improving resistance to differential cryptanalysis and the redesign of LOKI. In *Advances in Cryptology—ASIACRYPT'91*, volume 739, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1993) pp. 36–50.
6. Florent Chabaud and Serge Vaudenay. Links between differential and linear cryptanalysis. In *Advances in Cryptology—EUROCRYPT'94*, volume 950, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1995) pp. 256–265.
7. J. Daemen, R. Govaerts, and J. Vandewalle. Correlation matrices. In *Fast Software Encryption*, volume 1008, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1994) pp. 275–285.
8. J. F. Dillon. A survey of bent functions. *The NSA Technical Journal*, (1972) pp. 191–215. (unclassified).

9. Friedhelm Erwe. *Differential And Integral Calculus*. Oliver And Boyd Ltd, Edinburgh And London (1967).
10. T. Jakobsen and L. Knudsen. The interpolation attack on block ciphers. In *Fast Software Encryption*, Lecture Notes in Computer Science, Springer-Verlag, Berlin, New York, Tokyo (1997).
11. Rudoff Lidl and Harald Niederreiter. *Finite Fields, Encyclopedia of Mathematics and Its Applications*. Cambridge University Press (1983).
12. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam, New York, Oxford (1978).
13. M. Matsui. Linear cryptanalysis method for DES cipher. In *Advances in Cryptology—EUROCRYPT'93*, volume 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1994) pp. 386–397.
14. W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. In *Advances in Cryptology—EUROCRYPT'89*, volume 434, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1990), pp. 549–562.
15. K. Nyberg. Perfect nonlinear S-boxes. In *Advances in Cryptology—EUROCRYPT'91*, volume 547, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1991) pp. 378–386.
16. K. Nyberg. On the construction of highly nonlinear permutations. In *Advances in Cryptology—EUROCRYPT'92*, volume 658, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1993) pp. 92–98.
17. K. Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptology—EUROCRYPT'93*, volume 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1994) pp. 55–65.
18. L. J. O'Connor. On the distribution of characteristics in bijective mappings. In *Advances in Cryptology—EUROCRYPT'93*, volume 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1994) pp. 360–370.
19. B. Preneel, W. V. Leekwijck, L. V. Linden, R. Govaerts, and J. Vandewalle. Propagation characteristics of boolean functions. In *Advances in Cryptology—EUROCRYPT'90*, volume 437, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1991) pp. 155–165.
20. O. S. Rothaus. On "bent" functions. *Journal of Combinatorial Theory*, Ser. A, Vol. 20 (1976) pp. 300–305.
21. J. Seberry, X. M. Zhang, and Y. Zheng. Systematic generation of cryptographically robust S-boxes. In *Proceedings of the first ACM Conference on Computer and Communications Security*, The Association for Computing Machinery, New York (1993) pp. 172–182.
22. J. Seberry, X. M. Zhang, and Y. Zheng, On constructions and nonlinearity of correlation immune functions. In *Advances in Cryptology—EUROCRYPT'93*, volume 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1994) pp. 181–199.
23. J. Seberry, X. M. Zhang, and Y. Zheng, Nonlinearity and propagation characteristics of balanced boolean functions. *Information and Computation*, Vol. 119, No. 1 (1995) pp. 1–13.
24. T. Shimoyama, S. Moriai, and T. Kaneko, Cryptanalysis of the cipher KN, May 1997, presented at the rump session of Eurocrypt'97.
25. H. Tapia-Recillas, E. Daltabuit, and G. Vega, Some results on regular mappings, preprint, 1996.
26. R. Yarlagadda and J. E. Hershey, Analysis and synthesis of bent sequences. *IEE Proceedings (Part E)*, Vol. 136 (1989) pp. 112–123.
27. X. M. Zhang and Y. Zheng, Auto-correlations and new bounds on the nonlinearity of boolean functions. In *Advances in Cryptology—EUROCRYPT'96*, volume 1070, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York (1996) pp. 294–306.