Non-existence of Certain Quadratic S-boxes and Two Bounds on Nonlinear Characteristics of General S-boxes

Xian-Mo Zhang
Department of Computer Science
The University of Wollongong
Wollongong, NSW 2522, AUSTRALIA
E-mail: xianmo@cs.uow.edu.au

Yuliang Zheng
School of Computing and Information Technology
Monash University
Melbourne, VIC 3199, AUSTRALIA
E-mail: yuliang@mars.fcit.monash.edu.au

Hideki Imai
Institute of Industrial Science
The University of Tokyo
7-22-1 Roppongi, Minato-ku, Tokyo 106, JAPAN
Email: imai@iis.u-tokyo.ac.jp

October 2, 1997

Abstract

Due to the success of differential and linear attacks on a large number of encryption algorithms, it is important to investigate relationships among the 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. These results include: (1) an interesting equivalence relationship between a regular (balanced) S-box and a tight lower bound on the sum of elements in the leftmost column of its differential distribution table, (2) a proof for the nonexistence of *quadratic* S-boxes with a uniformly half-occupied difference distribution table for the case of $n \geqq 2m-1$. This serves as a piece of evidence that further supports an important and unproven conjecture, namely, for all $n > m$, there exist no $n \times m$ S-boxes with a uniformly half-occupied difference distribution table. Prior to this work, the best known result that supports the conjecture is that there exist no *quadratic* S-boxes with a uniformly half-occupied difference distribution table *if n or m is even*, (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).

1 Introduction

This paper deals with $n \times m$ S-boxes with $n > m$. Success of the notable differential cryptanalysis on various block ciphers [4, 5] has motivated researchers to investigate properties of the difference distribution

tables of S-boxes. A core topic in the endeavor is to find out 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 different 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.

Our next concern is on the differential uniformity of an S-box. In order to resist against differential cryptanalysis, researchers started to search for S-boxes whose difference distribution tables are relatively flat. As S-boxes with a *completely flat* difference distribution table have been known to be weak in resisting against differential attacks, naturally one of the research focuses has been on designing S-boxes with a uniformly half-occupied difference distribution table (UHODDT), i.e., S-boxes whose differential distribution tables contain an equal number of zero and identical non-zero entries in each of their rows (not taking into account the top row). Previous works directly or indirectly related to this line of research include, but not limited to, [1, 3, 15, 16, 17, 18, 19].

Defying efforts by a number of researchers, no $n \times m$ S-box with a UHODDT has emerged. This has led to a conjecture which states that

for all $n > m$, there exists no $n \times m$ S-box with a UHODDT.

Some progress in proving the conjecture was made in [29] where it was shown that when n or m is even, there exists no *quadratic* $n \times m$ S-box with a UHODDT (see Theorem 1 of [29]). This paper reports further progress in proving the conjecture. In particular, we show that *when* $n \geqq 2m-1$, there exists no *quadratic* $n \times m$ S-box with a UHODDT. We hope that this new piece of evidence can be of some contribution to the eventual success in proving the conjecture.

The next issue addressed in this paper is on the lower bound of differential uniformity. The differential uniformity of an S-box 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 considerably 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, it shows 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 nonzero 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. This is followed by Section 5 where it is proved that for $n \geqq 2m-1$, there exists no quadratic $n \times m$ S-box with a UHODDT. A tight lower bound on the differential uniformity of an S-box is presented in Section 6, and then two upper bounds on the nonlinearity of an S-box and its difference distribution table are proved in Section 7. Section 8 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 alphabetical order, are denoted by α_0, α_1, ..., α_{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 $N_f \leqq 2^{n-1} - 2^{\frac{1}{2}n-1}$. The equality holds if and only if f is bent (see P. 426 of [13]).

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 1 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 the researchers.

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

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

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

Let ξ be the sequence of a function f on V_n. Then

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

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 $n \geqq m \geqq 1$ 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 [12].

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 $n \geqq m \geqq 1$ 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.

Definition 2 The standard definition of the *nonlinearity* of $F = (f_1, \ldots, f_m)$ is

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

Now we consider an S-box in terms of its usefulness in designing a block cipher secure against differential cryptanalysis [4, 5]. The essence of a differential attack is that it exploits 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).

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 $(2^n, 0, \ldots, 0)$. As entries with higher values in the table are particularly useful to differential cryptanalysis, a necessary condition for an S-box to be immune to differential cryptanalysis is 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* whose formal definition follows [17].

Definition 3 *Let F be an $n \times m$ S-box, where $n \geqq m$. 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,*

$$\delta = \max_{\alpha \in V_n, \alpha \neq 0} \max_{\beta \in V_s} |\{x | F(x) \oplus F(x \oplus \alpha) = \beta\}|$$

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 whose formal definition follows.

Definition 4 *Let f be a function on V_n and ξ denote the sequence of f. f is called a* bent *function if*

$$|\langle \xi, \ell_i \rangle| = 2^{\frac{n}{2}},$$

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

Bent functions can be characterized in various ways [2, 9, 21, 24, 27]. A characterization of particular interest can be found in [9, 21] 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 $F = (f_1, \ldots, f_m)$.

Definition 5 *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 $F(x) \oplus F(x \oplus \alpha)$ runs through $\beta_j \in V_m$ 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.*

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

Definition 6 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 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 are called *auto-correlation distribution table and correlation immunity distribution table* of the S-box F, respectively.

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

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

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 [20] 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 [23]).

To explore relationships between the difference distribution table and other cryptographic characteristics of an S-box, first we examine a relationship between the difference distribution table and the auto-correlations of the component functions of the S-box.

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 $n \geqq m \geqq 1$ and each $f_j(x)$ 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,

(i) $D = KH_m$,

(ii) $P = H_n D$,

(iii) $P = H_n K H_m$.

Permutations are a special type of S-boxes that are used in many cryptographic algorithms. Of particular 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 a regular S-box can be completely 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 $n \geqq m \geqq 1$ and each f_j is a function on V_n. Then F is regular if and only if the sum of a 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,

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

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

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

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 [26].

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 $n \geqq m \geqq 1$ and each f_j is a function on V_n. Then*

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

(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

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

Compare the two sides of equality (7), obtaining

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

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 Nonexistence of Certain Quadratic S-boxes

Recall that the differential uniformity δ of an S-box is defined as the largest value in the differential distribution table of the S-box (not taking into account the top row). Clearly δ is constrained by $2^{n-m} \leqq \delta \leqq 2^n$. Extensive research has been carried out to construct differentially δ-uniform S-boxes with low δ [1, 3, 15, 16, 17, 18, 19]. Some constructions, in particular those based on permutation polynomials on finite fields, are simple and elegant. However, caution must be taken with Definition 3. In particular, it should be noted that low differential uniformity (a small δ) is only a *necessary*, but not a *sufficient* condition for immunity to differential attacks. This is shown by the fact that S-boxes constructed in [1, 15], which have a flat difference distribution table, are extremely weak to differential attacks, despite the fact that they achieve the lowest possible differential uniformity $\delta = 2^{n-m}$ [5, 6, 22].

We are particularly interested in $n \times m$ S-boxes that have the following property: for each nonzero vector $\alpha \in V_n$, $F(x) \oplus F(x \oplus \alpha)$ runs through 2^{m-t}, $1 \leqq t \leqq m$, of the vectors in V_m, each 2^{n-m+t} times, but not through the other $2^m - 2^{m-t}$ vectors in V_n. With each row in the difference distribution table of such an S-box, 2^{m-t} of its entries contain a value 2^{n-m+t} while the remaining entries contain a value zero. For simplicity, we say such a difference distribution table to be *uniformly* 2^{m-t}*-occupied*.

For n odd, $n = m$ (i.e., permutation S-boxes) and $t = 1$, there has been a large body of research (see for instance [3, 15, 16, 17, 18, 19]). One of the properties of these permutations is that their differential distribution tables are all 2-uniform, namely, half of the entries in a row contain a value zero while the other half contain a value 2. For this reason, it is believed that these permutations achieve the highest possible robustness against the differential attack.

As an extension of the above observation to a $n \times m$ S-box with $n \geqq m$, one would expect that the S-box would be highly useful in resisting differential attacks if its difference distribution table is *uniformly* 2^{m-1}*-occupied*, i.e., each row in the difference distribution table contains an equal number of zero and non-zero entries with all the non-zero values being identical to 2^{n-m+1}. For simplicity, we say that such an $n \times m$ S-box has a *uniformly half-occupied difference distribution table (UHODDT)*.

Intuitively, an $n \times m$ S-box with a UHODDT is expected to be useful as it seems to sit nicely in the middle of two undesirable extremes: S-boxes whose differential distribution tables contain too few non-zero entries and S-boxes whose differential distribution tables contain too many non-zero entries. At one extreme, the differential distribution tables contain high-valued entries which may be exploited by differential attacks, while at the other extreme, the differential distribution tables may be so close to a flat one that the S-box is again exploitable by differential attacks.

As we mentioned earlier, despite efforts by a number of researchers around the world, we have not witnessed the appearance of an $n \times m$ S-box with a UHODDT, except for the case of $n = m$ with n odd. This has led us to a conjecture:

Conjecture 1 *For all $n > m$, there exists no $n \times m$ S-box with a UHODDT.*

The first major step towards proving the conjecture was made in Theorem 1 of [29] for a special class of S-boxes called *quadratic S-boxes* whose algebraic degrees are two. In particular, it has been proved in [29] that *for $n \geqq 4$, there exists no quadratic $n \times m$ S-box with a UHODDT if n or m is even.*

There are a few directions one can follow to improve the result in [29]. These directions may include (1) proving the conjecture for higher-degree (say cubic) S-boxes, (2) proving the conjecture for quadratic S-boxes, but with different parameters. In what follows we report our progress in the second direction.

Theorem 2 *There exists no quadratic $n \times m$ S-box with a UHODDT when $n \geqq 2m - 1$.*

Proof. Assume for contradiction that there exists a quadratic $n \times m$ S-box with a UHODDT, say $F = (f_1, \ldots, f_m)$, for $n \geqq 2m - 1$. Write all the nonzero linear combination of $f_1, \ldots, f_m$ as $g_1, \ldots, g_{2^m-1}$. From the proof of Theorem 1 of [29], each nonzero vector in V_n is a linear structure of a unique g_j, i.e., there is a unique g_j such that $g_j(x) \oplus g_j(x \oplus \alpha)$ is a constant. It is easy to verify that for each $j = 1, \ldots, 2^{m-1}$, the nonzero linear structures of g_j, together with the zero vector, form a t_j-dimensional subspace of V_n for an integer t_j. We denote the subspace by W_j.

Note that

$$V_n = W_1 \cup \cdots \cup W_{2^m-1} \tag{9}$$

where

$$W_j \cap W_i = \{0\} \text{ if } j \neq i. \tag{10}$$

Thus $2^{t_1} + \cdots + 2^{t_{2^m-1}} = 2^n + 2^m - 2$ and thus there is a j_0, $1 \leqq j_0 \leqq 2^m - 1$, such that

$$2^{t_{j_0}} \geqq \frac{2^n + 2^m - 2}{2^m - 1} \geqq 2^{n-m} + 1$$

From this it follows that

$$2^{t_{j_0}} \geqq 2^{n-m+1}.$$

Now consider W_{j_0}. From linear algebra, V_n can be expressed as a partition

$$V_n = U_0 \cup U_1 \cup \cdots \cup U_{2^{n-t_{j_0}}} \tag{11}$$

satisfying

(i) $U_0 = W_{j_0}$,

(ii) $|U_j| = 2^{t_{J_0}}$,

(iii) $U_j \cap U_i = \phi$ where ϕ denotes the empty set,

(iv) two vectors α', α'' belong the same class U_j (also called a coset) if and only if $\alpha' \oplus \alpha'' \in U_0$.

Now we focus on U_1. Since $U_1 \cap U_0 = \phi$, from (9), we have

$$U_1 \subseteq (W_1 \cup \cdots \cup W_{j_0-1} \cup W_{j_0+1} \cup \cdots \cup W_{2^m-1}). \tag{12}$$

Note that $|U_1| = 2^{t_{j_0}} \geqq 2^{n-m+1}$. By the assumption, we have $n - m + 1 \geqq m$. Thus $|U_1| > 2^m - 1$. (12) implies that there is i_0, $i_0 \in \{1, \cdots, j_0 - 1, j_0 + 1, \cdots, 2^m - 1\}$, such that $|W_{i_0} \cap U_1| \geqq 2$. Let $\alpha', \alpha'' \in W_{i_0} \cap U_1$.

Since $\alpha', \alpha'' \in U_1$, from the above property (iv), we have $\alpha' \oplus \alpha'' \in U_0 = W_{j_0}$. On the other hand, since $\alpha', \alpha'' \in W_{i_0}$ and W_{i_0} is a subspace, we must have $\alpha' \oplus \alpha'' \in W_{i_0}$. This proves that

$$\alpha' \oplus \alpha'' \in W_{i_0} \cap W_{j_0}. \tag{13}$$

Since $i_0 \in \{1, \cdots, j_0 - 1, j_0 + 1, \cdots, 2^m - 1\}$, we have $i_0 \neq j_0$. This contradicts (10). □

We note that both Theorem 2 in this paper and Theorem 1 in [29] can be extended to S-boxes with partially bent component functions introduced in [7].

6 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,

$$\delta = \max_{\alpha \in V_n, \alpha \neq 0} \max_{\beta \in V_s} |\{x | F(x) \oplus F(x \oplus \alpha) = \beta\}|$$

(See Definition 3). As discussed earlier, δ is bounded by $2^{n-m} \leqq \delta \leqq 2^n$, 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 [28].

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, $p + q = n$, $1 \leqq p, q \leqq n - 1$, set $\sigma_j = \sum_{s=0}^{2^q-1} b_{j2^q+s}$, where $j = 0, 1, \ldots, 2^p - 1$. Then

$$2^q (a_0, a_{2^q}, a_{2 \cdot 2^q}, \ldots, a_{(2^p-1)2^q}) H_p = (\sigma_0, \sigma_1, \ldots, \sigma_{2^p-1}). \tag{14}$$

Now we prove another main result of this paper.

Theorem 3 *Let $F = (f_1, \ldots, f_m)$ be an $n \times m$ S-box, where n and m are integers with $n \geqq m \geqq 1$ and each $f_j(x)$ 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 $\Delta_M = \max\{|\Delta_j(\alpha)| \mid j = 1, \ldots, 2^m - 1, \alpha \in V_n, \alpha \neq 0\}$. Then we have*

$$\delta \geqq 2^{n-m} + 2^{-m} \Delta_M.$$

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

$$2^{-m} (\Delta_0(\alpha_{i'}), \Delta_1(\alpha_{i'}), \ldots, \Delta_{2^m-1}(\alpha_{i'})) H_m = (k_0(\alpha'_i), k_1(\alpha'_i), \ldots, k_{2^m-1}(\alpha'_i)) \tag{15}$$

Applying Lemma 3 to (15), 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 $0 \leqq s_0 \leqq 2^{m-1} - 1$ and $j_0 = 0$ or 1, such that

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

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

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

According to Section 5.3 of [22], 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

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

and hence

$$\delta \geqq 2^{n-m} + 2^{-m}\Delta_M.$$

□

When $\Delta_M = 0$, every nonzero linear combination of the components of F is a bent function. (Such S-boxes do exist [1, 15], but are not regular.) In this case we have $\delta = 2^{n-m}$. This indicates that the bound in Theorem 3 is tight.

7 Upper Bounds on Nonlinearity of S-boxes

After the discovery of differential attacks in [5], an equally notable cryptanalysis method, the linear cryptanalytic attack, was subsequently introduced in [14]. 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 [8].

In studying a $n \times m$ S-box, the two parameters n and m alone are not adequate in finding out detailed information on the S-box, except that when $m \geqq n - 1$, an upper bound on nonlinearity was obtained in [8] (but see discussions in the closing paragraph of this section.)

On the other hand, it will be too complex 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 5). The two theorems to be proved in this section can be viewed as a compromise between the two approaches. These two theorems relate the nonlinearity of an $n \times m$ S-box to three parameters, namely n, m and the number of nonzero entries in its difference distribution table K.

7.1 General Case

Here we consider $n \times m$ S-box that is not necessarily regular. In addition, the restriction of $n \geqq m$ is not imposed on the S-box. We first introduce Hölder's Inequality which can be found in [10].

Lemma 4 *Let $c_j \geqq 0$ and $d_j \geqq 0$ 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*

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

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

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

$$\sum_{j=1}^{s} c_j^2 \geqq s^{-1} (\sum_{j=1}^{s} c_j)^2 \tag{16}$$

where the quality holds if and only if $c_1, \ldots, c_s$ are all identical. The inequality (16) will be used in the proof of the following two theorems regarding the upper bound on the nonlinearity of an S-box.

Theorem 4 *Let F be an $n \times m$ S-box (F is not necessarily regular, and the restriction of $n \geqq m$ is not imposed on it). Denote by T_{nz} the total number of all nonzero entries, except for $k_0(\alpha_0)$, in the difference distribution table K of the S-box (see Definition 6). Then*

(i) the nonlinearity of F satisfies

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

(ii) the equality in (i) holds if and only if every nonzero 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

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

Now combining (4) with (16), a special form of Hölder's Inequality, we have

$$\sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j^2(\alpha_i) \geqq T_{nz}^{-1} (\sum_{j=0}^{2^m-1} \sum_{i=1}^{2^n-1} k_j(\alpha_i))^2 = T_{nz}^{-1} 2^{2n} (2^n - 1)^2. \tag{17}$$

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

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

which implies

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

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

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

$$|\langle \eta_j, \ell_i \rangle| \leqq (\frac{2^{2n+m} - 2^{3n} + T_{nz}^{-1} 2^{2n+m}(2^n - 1)^2}{2^m - 1})^{\frac{1}{4}} \tag{18}$$

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 (18) implies that the equality on the left hand side of (17) must hold. Namely,

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

Again using (16), 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 nonzero 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 $m \leqq \frac{1}{2}n$ [15].

Conversely, assume that every nonzero 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 nonzero entries in the table K is $T_{nz} = 2^m(2^n - 1)$. Now the inequality in Part (i) of the theorem becomes

$$N_F \leqq 2^{n-1} - 2^{\frac{n}{2}-1}. \tag{19}$$

On the other hand, since every nonzero linear combination of the component functions of F is a bent function, the equality in (19) must hold i.e. the equality in Part (i) of the theorem holds. This completes the proof of Part (ii). □

Before moving on to the next topic on regular S-boxes, we would like to stress that Theorem 4 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.

7.2 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 a non-regular 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 nonzero 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 $\Omega_j \cap \Omega_i = \phi$ for all $j \neq i$.*

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

$$V_n = \cup_{\beta \in V_m} \Omega_\beta \tag{20}$$

and $\Omega_{\beta'} \cap \Omega_{\beta''} = \phi$ 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 (20) 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 (20). 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 5 *Let F be a regular $n \times m$ S-box (For such an S-box $n \geqq m$ is necessary). Denote by t_{nz} the total number of nonzero 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*

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

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

$$\left(\sum_{j=0}^{2^m-1} \langle \eta_j, \ell_0 \rangle^2\right)^2 + \left(\sum_{j=0}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2\right)^2 + \cdots + \left(\sum_{j=0}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2\right)^2 = 2^{2m+n} \sum_{i=0}^{2^n-1} k_0^2(\alpha_i) \tag{21}$$

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 nonzero linear combination of the component functions of F is balanced. Thus for $j = 1, \ldots, 2^m - 1$, 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 $j = 1, \ldots, 2^n - 1$.

Note that $k_0(\alpha_0) = 2^n$. So (21) can be specialized as

$$\left(\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2\right)^2 + \cdots + \left(\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2\right)^2 = 2^{2m+3n} - 2^{4n} + 2^{2m+n} \sum_{i=1}^{2^n-1} k_0^2(\alpha_i) \tag{22}$$

By using (16)

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

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

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

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

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

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

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

Now the theorem follows immediately from (1). □

Comparing Theorem 4 with Theorem 5, 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 $n \geqq m$ is required only in Theorem 5. 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 4 is the total number of nonzero 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 5 is the total number of nonzero 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 4 and 5 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.

Before closing this section, we note that a paper by Chabaud and Vaudenay [8] is a prior work most relevant to this research. A main result in [8] 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

$$N_F \leqq 2^{n-1} - \frac{1}{2}(3 \cdot 2^n - 2 - \frac{2(2^n - 1)(2^{n-1} - 1)}{2^m - 1})^{\frac{1}{2}}. \tag{23}$$

Examining the part under the square root in the expression, one can see that it is negative if $m \leqq n - 2$. Therefore, (23) is applicable only to $n \times m$ S-boxes with $m \geqq n - 1$.

8 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 tight upper bounds on the nonlinearity of an S-box.

The technique used in proving the nonexistence result is essentially similar to that used in [29]. This technique, however, seems to have its limitation in that it may not be applicable to proving the nonexistence of higher-degree S-boxes.

In light of recent progress in interpolation [11] and high order differential cryptanalysis [25], 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.

Acknowledgment

The first author was supported by a Queen Elizabeth II Research Fellowship (223 23 1001). Part of the second author's work was completed while on sabbatical at the University of Tokyo.

References

[1] Adams, C. M. On immunity against Biham and Shamir's "differential cryptanalysis". *Information Processing Letters 41* (1992), 77–80.

[2] Adams, C. M., and Tavares, S. E. Generating and counting binary bent sequences. *IEEE Transactions on Information Theory IT-36 No. 5* (1990), 1170–1173.

[3] Beth, T., and Ding, C. On permutations against differential cryptanalysis. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 65–76.

[4] Biham, E., and Shamir, A. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology Vol. 4, No. 1* (1991), 3–72.

[5] Biham, E., and Shamir, A. *Differential Cryptanalysis of the Data Encryption Standard.* Springer-Verlag, New York, Heidelberg, Tokyo, 1993.

[6] Brown, L., Kwan, M., Pieprzyk, J., and Seberry, J. Improving resistance to differential cryptanalysis and the redesign of LOKI. In *Advances in Cryptology - ASIACRYPT'91* (1993), vol. 739, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 36–50.

[7] Carlet, C. Partially-bent functions. *Designs, Codes and Cryptography 3* (1993), 135–145.

[8] Chabaud, F., and Vaudenay, S. Links between differential and linear cryptanlysis. In *Advances in Cryptology - EUROCRYPT'94* (1995), vol. 950, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 256–265.

[9] Dillon, J. F. A survey of bent functions. *The NSA Technical Journal* (1972), 191–215. (unclassified).

[10] Erwe, F. *Differential And Integral Calculus.* Oliver And Boyd Ltd, Edinburgh And London, 1967.

[11] Jakobsen, T., and Knudsen, L. The interpolation attack on block ciphers. In *Fast Software Encryption* (Berlin, New York, Tokyo, 1997), Lecture Notes in Computer Science, Springer-Verlag.

[12] Lidl, R., and Niederreiter, H. *Finite Fields, Encyclopedia of Mathematics and Its Applications.* Cambridge University Press, 1983.

[13] MacWilliams, F. J., and Sloane, N. J. A. *The Theory of Error-Correcting Codes.* North-Holland, Amsterdam, New York, Oxford, 1978.

[14] Matsui, M. Linear cryptanalysis method for DES cipher. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 386–397.

[15] Nyberg, K. Perfect nonlinear S-boxes. In *Advances in Cryptology - EUROCRYPT'91* (1991), vol. 547, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 378–386.

[16] Nyberg, K. On the construction of highly nonlinear permutations. In *Advances in Cryptology - EUROCRYPT'92* (1993), vol. 658, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 92–98.

[17] Nyberg, K. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 55–65.

[18] Nyberg, K., and Knudsen, L. R. Provable security against differential cryptanalysis. In *Advances in Cryptology - CRYPTO'92* (1993), vol. 740, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 566–574.

[19] Pieprzyk, J. Bent permutations. In *Proceeding of the International Conference on Finite Fields, Coding Theory, and Advances in Communications and Computing* (Las Vegas, 1991).

[20] Preneel, B., Leekwijck, W. V., Linden, L. V., Govaerts, R., and Vandewalle, J. Propagation characteristics of boolean functions. In *Advances in Cryptology - EUROCRYPT'90* (1991), vol. 437, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 155–165.

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

[22] Seberry, J., Zhang, X. M., and Zheng, Y. Systematic generation of cryptographically robust S-boxes. In *Proceedings of the first ACM Conference on Computer and Communications Security* (1993), The Association for Computing Machinery, New York, pp. 172 – 182.

[23] Seberry, J., Zhang, X. M., and Zheng, Y. On constructions and nonlinearity of correlation immune functions. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 181–199.

[24] Seberry, J., Zhang, X. M., and Zheng, Y. Nonlinearity and propagation characteristics of balanced boolean functions. *Information and Computation 119*, 1 (1995), 1–13.

[25] Shimoyama, T., Moriai, S., and Kaneko, T. Cryptanalysis of the cipher KN, May 1997. (presented at the rump session of Eurocrypt'97).

[26] Tapia-Recillas, H., Daltabuit, E., and Vega, G. Some results on regular mappings, 1996. (preprint).

[27] Yarlagadda, R., and Hershey, J. E. Analysis and synthesis of bent sequences. *IEE Proceedings (Part E) 136* (1989), 112–123.

[28] Zhang, X. M., and Zheng, Y. Auto-correlations and new bounds on the nonlinearity of boolean functions. In *Advances in Cryptology - EUROCRYPT'96* (1996), vol. 1070, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 294–306.

[29] Zhang, X. M., and Zheng, Y. On the difficulty of constructing cryptographically strong substitution boxes. *Journal of Universal Computer Science 2*, 3 (1996), 147–162. (available at `http://hgiicm.tu-graz.ac.at/`).

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

$$A = \begin{bmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_{s-1} \end{bmatrix} = (a_{ij}) = [\chi_0, \chi_1, \cdots, \chi_{t-1}], \tag{24}$$

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

$$B = \begin{bmatrix} \eta_0 \\ \eta_1 \\ \vdots \\ \eta_{2^s-1} \end{bmatrix} = (b_{ij}) = [\gamma_0, \gamma_1, \cdots, \gamma_{t-1}], \tag{25}$$

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

$$B^* = (b^*_{ij}) = \begin{bmatrix} R_0 \\ R_1 \\ \vdots \\ R_{2^s-1} \end{bmatrix} = [h_0, h_1, \cdots, h_{t-1}], \tag{26}$$

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 (26), $e_0, e_1, \ldots, e_{2^s-1}$ 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

$$e_i B^* B^{*T} e_j^T = \begin{cases} k_j 2^{2s} & \text{if } e_i = e_j \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

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

$$c^*_u = \begin{cases} 2^s & \text{if } e_i^T = h_u \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

for all $u = 0, \ldots, t-1$. Similarly, write $e_j B^* = (d^*_0, \ldots, d^*_{t-1})$, where

$$d^*_u = \begin{cases} 2^s & \text{if } e_j^T = h_u \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

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 (28) and (29), $c_u^* \neq 0$ 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 (27).

As the Sylvester-Hadamard matrix H_m is symmetric, (27) can be equivalently stated as:

$$H_s B^* B^{*T} H_s = 2^{2s} \operatorname{diag}(k_0, k_1, \ldots, k_{2^s-1}). \tag{30}$$

Let R_j be a row of B^* defined in (26) and k_j the number of times a row vector e_j in H_s appears in the columns of B^*. From (30) 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

$$(\langle R_0, R_0 \rangle, \langle R_0, R_1 \rangle, \ldots, \langle R_0, R_{2^s-1} \rangle) = (k_0, k_1, \ldots, k_{2^s-1}) H_s. \tag{31}$$

Now we are in a position to prove Lemma 2. Consider an $s \times t$ matrix A defined in (24) with $s = m$ and $t = n$. Let a row ξ_i in (24) be the truth table of $f_i(x) \oplus f_i(x \oplus \alpha)$, $i = 0, 1, \ldots, m-1$. Correspondingly, η_i in (25) denotes the truth table of $g_i(x) \oplus g_i(x \oplus \alpha)$, and R_i in (26) 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 (31).

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.