Relationships Among Nonlinearity Criteria

Jennifer Seberry

Xian-Mo Zhang

Yuliang Zheng

Department of Computer Science
The University of Wollongong
Wollongong, NSW 2522, AUSTRALIA
E-mail: {jennie,xianmo,yuliang}@cs.uow.edu.au

Abstract

An important question in designing cryptographic functions including substitution boxes (S-boxes) is the relationships among the various nonlinearity criteria each of which indicates the strength or weakness of a cryptographic function against a particular type of cryptanalytic attacks. In this paper we reveal, for the first time, interesting connections among the strict avalanche characteristics, differential characteristics, linear structures and nonlinearity of quadratic S-boxes. In addition, we show that our proof techniques allow us to treat in a unified fashion all quadratic permutations (namely, quadratic S-boxes that form permutations), regardless of the underlying construction methods. This greatly simplifies the proofs for a number of known results on the nonlinearity characteristics of quadratic permutation. As a by-product, we solve an open problem regarding the existence of differentially 2-uniform quadratic permutations on an even dimensional vector space. Another contribution of this paper is the identification of an error in a paper presented by Beth and Ding at EUROCRYPT'93.

1 Nonlinearity Criteria

This section introduces basic notions and definitions of several nonlinearity criteria for cryptographic functions.

Denote by V_n the vector space of n tuples of elements from $GF(2)$. Let $\alpha = (a_1, \ldots, a_n)$ and $\beta = (b_1, \ldots, b_n)$ be two vectors in V_n. The scalar product of α and β, denoted by $\langle \alpha, \beta \rangle$, is defined by $\langle \alpha, \beta \rangle = a_1 b_1 \oplus \cdots \oplus a_n b_n$, where multiplication and addition are over $GF(2)$. In this paper we consider functions from V_n to $GF(2)$ (or simply functions on V_n). We are particularly interested in functions whose algebraic degrees are 2, also called quadratic functions.

Let f be a function on V_n. The $(1,-1)$-sequence defined by $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-1)^{f(\alpha_{2^n-1})})$ is called the *sequence* of f, and the $(0,1)$-sequence defined by $(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$ is called the *truth table* of f, where $\alpha_0 = (0, \ldots, 0, 0)$, $\alpha_1 = (0, \ldots, 0, 1)$, $\ldots$, $\alpha_{2^{n-1}-1} = (1, \ldots, 1, 1)$. f is said *balanced* if its truth table has 2^{n-1} zeros (ones).

An *affine* function f on V_n is a function that takes the form of $f = a_1 x_1 \oplus \cdots \oplus a_n x_n \oplus c$, where $a_j, c \in GF(2), j = 1, 2, \ldots, n$. Furthermore f is called a *linear* function if $c = 0$. The sequence of an affine (or linear) function is called an *affine (or linear) sequence*.

The *Hamming weight* of a vector $\alpha \in V_n$, denoted by $W(\alpha)$, is the number of ones in the vector.

Now we introduce bent functions, an important combinatorial concept introduced by Rothaus in the mid 1960's (although his pioneering work was not published until some ten years later [15].)

Definition 1 *A function f on V_n is said to be bent if*

$$2^{-\frac{n}{2}} \sum_{x \in V_n} (-1)^{f(x) \oplus \langle \beta, x \rangle} = \pm 1$$

for every $\beta \in V_n$. Here $x = (x_1, \ldots, x_n)$ and $f(x) \oplus \langle \beta, x \rangle$ is considered as a real valued function.

From the definition, it can be seen that bent functions on V_n exist only when n is even. Another fact is that bent functions are not balanced, hence not directly applicable in most computer and communications security practices. Dillon presented a nice exposition of bent functions in [7]. In particular, he showed that bent functions can be characterized in various ways:

Lemma 1 *The following statements are equivalent:*

(i) f is bent.

(ii) $\langle \xi, \ell \rangle = \pm 2^{\frac{1}{2}n}$ for any affine sequence ℓ of length 2^n, where ξ is the sequence of f.

(iii) $f(x) \oplus f(x \oplus \alpha)$ is balanced for any non-zero vector $\alpha \in V_n$, where $x = (x_1, \ldots, x_n)$.

The strict avalanche criterion (SAC) was first introduced by Webster and Tavares [21, 22] when studying the design of cryptographically strong S-boxes.

Definition 2 *A function f on V_n is said to satisfy the strict avalanche criterion (SAC) if $f(x) \oplus f(x \oplus \alpha)$ is balanced for all $\alpha \in V_n$ with $W(\alpha) = 1$, where $x = (x_1, \ldots, x_n)$.*

It is widely accepted that the component functions of an S-box employed by a modern block cipher should all satisfy the SAC. A general technique for constructing SAC-fulfilling cryptographic functions can be found in [17].

While the SAC measures the avalanche characteristics of a function, the linear structure is a concept that in a sense is complementary to the former, namely, the linear structure indicates the smoothness of a function.

Definition 3 *Let f be a function on V_n. A vector $\alpha \in V_n$ is called a* linear structure *of f if $f(x) \oplus f(x \oplus \alpha)$ is a constant.*

By definition, the zero vector in V_n is a linear structure of all functions on V_n. As was pointed out in [11], the linear structures of a function f form a linear subspace of V_n. The dimension of the subspace is called the *linearity dimension* of f. Clearly, the linearity dimension of a function on V_n is bounded from the above by n, with the affine functions achieving the maximum dimension n. It is bounded from the below by 1 when n is even and by 2 when n is odd. The lower bound 1 is achieved only by bent functions, while 2 can be achieved by such functions as obtained by concatenating two bent functions (see [18, 20]).

In mathematical terms, an $n \times s$ S-box (i.e., with n input bits and s output bits), can be described as a mapping from V_n to V_s ($n \geqq s$). To avoid trivial statistical attacks, an S-box F should be *regular*, namely, $F(x)$ should run through all the vectors in V_s each 2^{n-s} times while x runs through V_n once. Note that an $n \times n$ S-box is a permutation on V_n and always regular.

Regularity of an $n \times s$ S-box F can be characterized by the balance of nonzero linear combinations of its component functions. It has been known that when $n = s$, F is regular if and only if all nonzero linear combinations of the component functions are balanced. A proof can be found in Remark 5.8 of [7]. The characterization can be extended to the case when $n > s$.

Theorem 1 *Let $F = (f_1, \ldots, f_s)$, where f_i is a function on V_n, $n \geqq s$. Then F is a regular mapping from V_n to V_s if and only if all nonzero linear combinations of $f_1, \ldots, f_n$ are balanced.*

A proof for the theorem is given in Appendix A. It seems to the authors that the proof for the case of $n = s$ as described in [7] can not be directly adapted to the general case of $n > s$, and hence the extension presented here is not trivial.

The next criterion is the nonlinearity that indicates the Hamming distance between a function and all the affine functions.

Definition 4 *Given two functions f and g on V_n, the* Hamming distance *between them, denoted by $d(f, g)$, is defined as the Hamming weight of the truth table of the function $f(x) \oplus g(x)$, where $x = (x_1, \ldots, x_n)$. The* nonlinearity *of f, denoted by N_f, is the minimal Hamming distance between f and all affine functions on V_n, i.e., $N_f = \min_{i=1,2,\ldots,2^{n+1}} d(f, \varphi_i)$ where φ_1, φ_2, ..., $\varphi_{2^{n+1}}$ denote the affine functions on V_n.*

The above definition can be extended to the case of mappings by defining the nonlinearity of a mapping from V_n to V_s as the minimum among the nonlinearities of nonzero linear combinations of the component functions.

The nonlinearity of a function f on V_n has been known to be bounded from the above by $2^{n-1} - 2^{\frac{1}{2}n-1}$. When n is even, the upper bound is achieved by bent functions. Constructions for highly nonlinear *balanced* functions can be found in [18, 20].

Nonlinearity has been considered to be an important criterion. Recent advances in *Linear cryptanalysis* put forward by Matsui [9] have made it explicit that nonlinearity is not just important, but essential to DES-like block encryption algorithms. Linear cryptanalysis exploits the low nonlinearity of S-boxes employed by a block cipher, and it has been successfully applied in attacking FEAL and DES. In [16], it has been shown that to immunize an S-box against linear cryptanalysis, it suffices for the Hamming distance between each nonzero linear combination of the component functions and each affine function not to deviate too far from 2^{n-1}, namely, *an S-box is immune to linear cryptanalysis if the nonlinearity of each nonzero linear combination of its component functions is high.*

Finally we consider a nonlinearity criterion that measures the strength of an S-box against differential cryptanalysis [3, 4]. 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 $n \times s$ S-box is a $2^n \times 2^s$ matrix. The rows of the matrix, indexed by the vectors in V_n, represent the change in the input, while the columns, indexed by the vectors in V_s, 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 first 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 counting the first entry in the first row).

Definition 5 *Let F be an $n \times s$ S-box, where $n \geqq s$. Let δ be the largest value in differential distribution table of the S-box (not counting the first entry in the first 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.*

Extensive research has been conducted in constructing differentially δ-uniform S-boxes with a low δ [10, 1, 11, 13, 12, 2]. Some constructions, in particular those based on permutation polynomials on finite fields, are simple and elegant. However, as pointed in [4, 5, 16], cautions must be taken with Definition 5. In particular, it should be noted that low differential uniformity (a small δ) is only a *necessary*, but not *sufficient* condition for immunity to differential attacks. A more complete measurement is the *robustness* introduced in [16]. The reader is directed to that paper for a comprehensive treatment of this subject.

We have discussed various cryptographic properties including the algebraic degree, the SAC, the linear structure, the regularity, the nonlinearity and the differential uniformity. As is stated in the following lemmas, some properties are invariant under a nonsingular linear transformation.

Lemma 2 *Let f be a function on V_n, A be a nonsingular matrix of order n over $GF(2)$, and let $g(x) = f(xA)$. Then f and g have the same algebraic degree, nonlinearity and linearity dimension.*

Lemma 3 *Let F be a mapping from V_n to V_s, where $n \geqq s$, A be a nonsingular matrix of order n over $GF(2)$, and B be a nonsingular matrix of order s over $GF(2)$. Let $G(x) = F(xA)$ and $H(x) = F(x)B$, where $x = (x_1, \ldots, x_n)$. Note that A is applied to the input, while B to the output of F. Then F, G and H all have the same regularity and differential uniformity.*

A proof for Lemma 3 can be found in Section 5.3 of [16].

2 Cryptographic Properties of Quadratic S-boxes

In this section we first prove a lower bound on the nonlinearity of S-boxes whose component functions are all quadratic (or simply, quadratic S-boxes). Then we reveal interesting relationships among the difference distribution table, linear structures and SAC of regular quadratic S-boxes.

2.1 Nonlinearity of Quadratic S-boxes

Consider a quadratic function f on V_n. Then $f(x) \oplus f(x \oplus \alpha)$ is affine, where $x = (x_1, \ldots, x_n)$ and $\alpha \in V_n$. Assume that f does not have nonzero linear structures. Then for any nonzero $\alpha \in V_n$, $f(x) \oplus f(x \oplus \alpha)$ is a nonzero affine function, hence balanced. By Part (iii) of Lemma 1, f is bent. Thus we have proved:

Lemma 4 *If a quadratic function f on V_n has no nonzero linear structures, then f is bent and n is even.*

The following lemma is a useful tool in calculating the nonlinearity of functions obtained via Kronecker product.

Lemma 5 *Let $g(x, y) = f_1(x) \oplus f_2(y)$, where $x = (x_1, \ldots, x_{n_1})$, $y = (y_1, \ldots, y_{n_2})$, f_1 is a function on V_{n_1} and f_2 is a function on V_{n_1}. Let d_1 and d_2 denote the nonlinearities of f_1 and f_2 respectively. Then the nonlinearity of g satisfies*

$$N_g \geqq d_1 2^{n_2} + d_2 2^{n_1} - 2 d_1 d_2.$$

In addition, we have $N_g \geqq d_1 2^{n_2}$ and $N_g \geqq d_2 2^{n_1}$.

Proof. The first half of the lemma can be found in Lemma 8 of [19]. The second half is true due to the fact that $d_1 \leqq 2^{n_1 - 1}$ and $d_2 \leqq 2^{n_2 - 1}$ (see also Section 3 of [18]). □

We now prove a lower bound on the nonlinearity of any quadratic $n \times s$ S-box ($n \geqq s$).

Lemma 6 *Let $F = (f_1, \ldots, f_s)$ be a quadratic $n \times s$ S-box. Also let $g(x) = \sum_{j=1}^{n} c_j f_j(x)$ be a nonzero linear combination of $f_1, \ldots, f_n$, and ℓ be the linearity dimension of g. Then*

(i) $n - \ell$ is even, and

(ii) the nonlinearity of g satisfies $N_g \geqq 2^{n-1} - 2^{\frac{1}{2}(n+\ell)-1}$.

Proof. (i) Recall that $\ell \leqq n$. If g is affine, then $\ell = n$, hence $n - \ell$ is even. Now suppose that g is not affine, i.e., $\ell < n$. Let $\{\beta_1, \ldots, \beta_\ell\}$ be a basis of the subspace consisting of the linear structures of g. $\{\beta_1, \ldots, \beta_\ell\}$ can be extended to $\{\beta_1, \ldots, \beta_\ell, \beta_{\ell+1}, \ldots, \beta_n\}$ such that the latter is a basis of V_n. Now let B be a nonsingular matrix with β_i as its ith row, and let $g^*(x) = g(xB)$. By Lemma 2, g^* and g have the same linearity dimension. Thus the question is transformed into the discussion of g^*.

Let e_j be the vector in V_n whose jth coordinate is one and others are zero. Then we have $e_j B = \beta_j$, and $g^*(e_j) = g(\beta_j)$, $j = 1, \ldots, n$. Thus $\{e_1, \ldots, e_\ell\}$ is a basis of the subspace consisting of the linear structures of g^*. As g^* is quadratic, it can be written as

$$g^*(x) = q(y) \oplus p(z) \oplus \sum_{j=1}^{\ell} r_j(z) x_j,$$

where $x = (x_1, \ldots, x_n)$, $y = (x_1, \ldots, x_\ell)$, and $z = (x_{\ell+1}, \ldots, x_n)$. In addition, each e_j can be written as $e_j = (\mu_j, 0)$, where $\mu_j \in V_\ell$ and $0 \in V_{n-\ell}$. Since each e_j is a linear structure of g^*, $g^*(x) \oplus g^*(x \oplus e_j) = q(y) \oplus q(y \oplus \mu_j) \oplus r_j(z)$ is a constant. Thus both $q(y) \oplus q(y \oplus \mu_j)$ and $r_j(z)$ are constants. This allows us to rewrite g^* as

$$\begin{aligned} g^*(x) &= q(y) \oplus p(z) \oplus \sum_{j=1}^{\ell} a_j x_j \\ &= q(y) \oplus \sum_{j=1}^{\ell} a_j x_j \oplus p(z) \\ &= h(y) \oplus p(z) \end{aligned}$$

where $a_j = r_j$ is a constant and $h(y) = q(y) \oplus \sum_{j=1}^{\ell} a_j x_j$.

Since linear structures form a subspace, $\{\mu_1, \ldots, \mu_\ell\}$ is a basis of V_ℓ and $q(y) \oplus q(y \oplus \mu_j)$ is a constant for each μ_j, $q(y) \oplus q(y \oplus \nu)$ must be a constant for all $\nu \in V_\ell$. In other words, q must be an affine function on V_ℓ. Thus h is also an affine function on V_ℓ. As the linearity dimension of g is ℓ, and h is an affine function on V_ℓ, p, a function on $V_{n-\ell}$, possesses no nonzero linear structures. By Lemma 4, p is a bent function on $V_{n-\ell}$ and hence $n - \ell$ is even.

(ii) This part is obviously true if g is affine. Now suppose that g is not affine. In this case, it has the same nonlinearity as that of g^*. As the function p is a bent function on V_ℓ, its nonlinearity satisfies $N_p = 2^{n-\ell-1} - 2^{\frac{1}{2}(n-\ell)-1}$. By Lemma 5, the nonlinearity of g^*, and hence of g, satisfies $N_g \geqq 2^\ell N_p = 2^{n-1} - 2^{\frac{1}{2}(n+\ell)-1}$. This completes the proof. □

2.2 Difference Distribution Table vs Linear Structure

First we show an interesting result stating that the number representing the differential uniformity of a quadratic S-box must be a power of 2.

Theorem 2 *Let δ be the differential uniformity of a regular quadratic $n \times s$ S-box. Then $\delta = 2^d$ for some $n - s + 1 \leqq d \leqq n$.*

Proof. Let $F = (f_1, \ldots, f_s)$. Let α be a nonzero vector in V_n. Then

$$F(x) \oplus F(x \oplus \alpha) = (f_1(x) \oplus f_1(x \oplus \alpha), \ldots, f_s(x) \oplus f_s(x \oplus \alpha)).$$

As f_1 is quadratic, $f_i(x) \oplus f_i(x \oplus \alpha)$ is affine, hence $F(x) \oplus F(x \oplus \alpha) = xD \oplus c$, where D is an $n \times s$ matrix over $GF(2)$ and c is a vector in V_s.

Assume that the rank of D is r with $0 \leqq r \leqq s$. Then $xD \oplus c$ runs through 2^r vectors in V_s, each 2^{n-r} times, while x runs through V_n, where n, s and r satisfy $n - s \leqq n - r \leqq n$. Thus the differential uniformity of F takes the form of 2^d, $n - s \leqq d \leqq n$.

We now prove $n - s + 1 \leqq d$. Assume $d = n - s$ holds. Since $\delta = 2^{n-s}$, for any nonzero $\alpha \in V_n$, $F(x) \oplus F(x \oplus \alpha)$ runs through at least $2^n/2^{n-s} = 2^s$ vectors in V_s while x runs through V_n once. On the other hand, $F(x) \oplus F(x \oplus \alpha)$, as a mapping from V_n to V_s, runs through at most 2^s vectors in V_s while x runs through V_n once. This proves that $F(x) \oplus F(x \oplus \alpha)$ runs through exactly 2^s vectors in V_s while x runs through V_n once. Since $F(x) \oplus F(x \oplus \alpha)$ is an affine transformation, it runs through 2^s vectors in V_s each $2^n/2^s = 2^{n-s}$ times while x runs through V_n once. In other words, $F(x) \oplus F(x \oplus \alpha)$ is regular. Note that α is an arbitrary nonzero vector in V_s. By Theorem 3.1 of [10], any nonzero linear combination of the components of $F(x)$ is a bent function on V_n. Since $F(x)$ is regular, any nonzero linear combination of the components of $F(x)$ is balanced (see Theorem 1). Since any bent function is not balanced (see [15]), the assumption of $n - s = d$ cannot hold. □

Theorem 3 *Let $F = (f_1, \ldots, f_s)$ be a differentially δ-uniform regular quadratic $n \times s$ S-box, where $\delta = 2^{n-s+t}$ for some $1 \leqq t \leqq s$ (see Theorem 2). Then*

(i) any nonzero vector $\alpha \in V_n$ is a linear structure of m nonzero linear combinations of $f_1, \ldots, f_s$, where m satisfies $1 \leqq m \leqq 2^t - 1$;

(ii) any nonzero nonzero linear combination of $f_1, \ldots, f_s$ has at least one linear structure $\alpha \in V_n$.

Proof. (i) Fix an arbitrary nonzero vector $\alpha \in V_n$. Note that $\delta > 2^{n-s}$. Then $F(x) \oplus F(x \oplus \alpha)$ is not regular. By Theorem 1 there exists a nonzero linear combination of $f_1, \ldots, f_s$, say $g = \sum_{j=1}^{n} c_j f_j$, such that $g(x) \oplus g(x \oplus \alpha)$ is not balanced. As $f_1, \ldots, f_s$ are all quadratic, g is quadratic or affine. Thus $g(x) \oplus g(x \oplus \alpha)$ must be a constant.

Now we proceed to proving that there exist at most $2^t - 1$ such combinations g in (i). First we note that there are $2^s - 1$ nonzero linear combinations of $f_1, \ldots, f_s$, denoted by $g_1, \ldots, g_{2^s-1}$, and $2^n - 1$ nonzero vectors in V_n, denoted by $\alpha_1, \ldots, \alpha_{2^n-1}$. Now suppose that there exist 2^t nonzero linear combinations $g_1, \ldots, g_{2^t}$, such that α is a linear structure of each g_j. Write $g_j(x) \oplus g_j(x \oplus \alpha) = a_j$, where a_j is constant, $j = 1, \ldots, 2^t$. Let $\Omega = \{g_1, \ldots, g_{2^t}\}$. We are interested in the rank of Ω, namely the maximum number of functions in Ω that are linearly independent. Recall that t linearly independent functions can generate only $2^t - 1$ distinct nonzero combinations. As Ω contains 2^t nonzero functions, its rank is at least $t+1$. Without loss of generality, suppose that $g_1, \ldots, g_{t+1}$ are linearly independent. Then there exist additional $s - t - 1$ nonzero linear combinations of $f_1, \ldots, f_s$, denoted by $h_{t+2}, \ldots, h_s$, such that $g_1, \ldots, g_{t+1}, h_{t+2}, \ldots, h_s$ are all linearly independent. Let G be an $n \times s$ mapping defined by $G = (g_1, \ldots, g_{t+1}, h_{t+2}, \ldots, h_s)$. Then G can be expressed as $G(x) = F(x)B$ for a nonsingular matrix B of order s over $GF(2)$.

By Lemma 3, G is also a differentially δ-uniform $n \times s$ S-box. Since $\delta = 2^{n-s+t}$ $(1 \leqq t \leqq s)$, $G(x) \oplus G(x \oplus \alpha)$ runs through at least $2^n/2^{n-s+t} = 2^{s-t}$ vectors. On the other hand,

$$G(x) \oplus G(x \oplus \alpha) = (a_1, \ldots, a_{t+1}, h_{t+2}(x) \oplus h_{t+2}(x \oplus \alpha), \ldots, h_s(x) \oplus h_s(x \oplus \alpha))$$

where a_1, ..., a_{t+1} are all constants. This indicates that $G(x) \oplus G(x \oplus \alpha)$ runs through at most 2^{s-t-1} vectors in V_s. This is a contradiction. Thus Part (i) is true.

(ii) Let $g = \sum_{j=1}^{s} c_j f_j$, where $(c_1, \ldots, c_s)$ is a nonzero vector in V_s. Assume that g has no nonzero linear structures. Then by Lemma 4, g is a bent function. This contradicts the fact that F is regular and that the nonzero linear combinations of its component functions are all balanced and have linear structures. This proves Part (ii). □

2.3 Difference Distribution Table vs SAC

Theorem 4 *Let $F = (f_1, \ldots, f_s)$ be a differentially δ-uniform regular quadratic $n \times s$ S-box, where $\delta = 2^{n-s+t}$, $1 \leqq t \leqq s$ (see Theorem 2) and $s \leqq 2^{s-t-2}$. Then there exists a nonsingular matrix of order n over $GF(2)$, say A, and a nonsingular matrix of order s over $GF(2)$, say B, such that $\Psi(x) = F(xA)B = (f_1(xA), \ldots, f_s(xA))B = (\psi_1(x), \ldots, \psi_s(x))$ is also a differentially δ-uniform regular quadratic $n \times s$ S-box whose component functions all satisfy the SAC.*

Proof. Again denote by g_1, ..., g_{2^s-1} the $2^s - 1$ nonzero linear combinations of $f_1, \ldots, f_s$, and by α_1, ..., α_{2^n-1} the $2^n - 1$ nonzero vectors in V_s. We construct a bipartite graph Γ whose vertices are g_1, ..., g_{2^s-1} and α_1, ..., α_{2^n-1}. An edge exists between g_i and α_j if and only if α_j is a linear structure of g_i. By Theorem 3, there exist at most $2^t - 1$ edges associated with each α. Thus there exist at most $(2^t-1)\cdot(2^n-1)$ edges in the graph Γ.

Denote by t_j the number of linear structures of g_j, $j = 1, \ldots, 2^s - 1$. Without loss of generality suppose that $t_1 \leqq t_2 \leqq \cdots \leqq t_{2^s-1}$. It can be seen that $t_j < 2^{n-s+t+1}$, $j = 1, \ldots, 2^{s-1}$. The reason is as follows. Suppose that it is not the case. Then we have $t_1 + \cdots + t_{2^s-1} \geqq 2^{s-1} \cdot 2^{n-s+t+1} = 2^{n+t} > (2^t - 1)\cdot(2^n - 1)$. This contradicts the fact that Γ has at most $2^{t-1} \cdot (2^n - 1)$ edges.

Now set $\Delta = \{g_1, \ldots, g_{2^{s-1}+1}\}$. As the rank of Δ is s, we can choose s functions from Δ, say $g_{j_1}, \ldots, g_{j_s}$, such that they are all linearly independent. Since $s \leqq 2^{s-t-2}$, we have $t_{j_1} + \cdots + t_{j_s} < s \cdot 2^{n-s+t+1} \leqq 2^{n-1}$. By Theorem 2 of [17], there exists a nonsingular matrix A of order n over $GF(2)$, such that all component functions of $(g_{j_1}(xA), \ldots, g_{j_s}(xA))$ satisfy the SAC. Furthermore, as each g_j is a nonzero linear combination of f_1, ..., f_s, there is a nonsingular matrix B of order s over $GF(2)$ such that $(g_{j_1}(x), \ldots, g_{j_s}(x)) = (f_1(x), \ldots, f_s(x))B$. Accordingly, by Lemma 3,

$$\Psi(x) = F(xA)B = (f_1(xA), \ldots, f_s(xA))B = (\psi_1(x), \ldots, \psi_s(x))$$

is a differentially δ-uniform regular quadratic $n \times s$ S-box, where each component function ψ_j satisfies the SAC. □

aaaaaa By (i) of Theorem 3, there exist at most $2^t - 1$ nonzero linear functions, without loss of generality say $U = \{g_1, \ldots, g_{2^t}\}$.

3 A Unified Treatment of Quadratic Permutations

This section is concerned with differentially 2-uniform quadratic $n \times n$ S-boxes. Such an S-box F has the following property: for any nonzero vector $\alpha \in V_n$, $F(x) \oplus F(x \oplus \alpha)$ runs through 2^{n-1} vectors in V_n, each twice, but not through the other 2^{n-1} vectors, while x runs through V_n.

Differentially 2-uniform quadratic $n \times n$ S-boxes have been extensively studied in the past years [14, 13, 6, 2, 12] and hence deserve special attention. Such S-boxes appear in various forms and researchers have employed different techniques, some of which are rather sophisticated, to prove their nonlinearity. By refining our proof techniques described in Section 2, we will show in this section that all differentially

2-uniform quadratic permutations, no matter how they are constructed, have the same nonlinearity and can be transformed into SAC-fulfilling S-boxes. This greatly simplifies the proof for a number of known results and could be a powerful tool in designing cryptographically strong block ciphers.

3.1 Linear Structure and Nonlinearity

Theorem 5 *Let $F = (f_1, \ldots, f_n)$ be a differentially 2-uniform quadratic permutation on V_n as described at the beginning of the section. Then there is a one-to-one correspondence between the nonzero vectors in V_n and the nonzero linear combinations of $f_1, \ldots, f_n$, namely,*

(i) each nonzero vector in V_n is the linear structure of a unique nonzero linear combination of $f_1, \ldots, f_n$,

(ii) each nonzero nonzero linear combination of $f_1, \ldots, f_n$ has a unique nonzero vector in V_n as its linear structure.

Proof. (i) follows from the first part of Theorem 3 (by letting $s = n$ and $t = 1$), while (ii) follows from (i) and (ii) of Theorem 3. □

Theorem 6 *Let $F = (f_1, \ldots, f_n)$ is a differentially 2-uniform quadratic permutation on V_n. Then*

(i) n is odd,

(ii) for any nonzero linear combination of $f_1, \ldots, f_n$, say $g = \sum_{j=1}^{n} c_j f_j$, the nonlinearity of g satisfies $N_g \geqq 2^{n-1} - 2^{\frac{1}{2}(n-1)}$.

Proof. (i) Let g be a nonzero linear combination of the n component functions. By Lemma 5, there is a unique nonzero vector $\alpha \in V_n$ such that $g(x) \oplus g(x \oplus \alpha)$ is a constant. Without loss of generality, we can suppose that $\alpha = e$, where $e = (0, \ldots 0, 1)$. On the other hand, g can be written as

$$g(x) = p(x_1, \ldots, x_{n-1})x_n \oplus q(x_1, \ldots, x_{n-1}).$$

Thus $g(x) \oplus g(x \oplus e) = p(x_1, \ldots, x_{n-1}) = a$ is a constant and

$$g(x) = ax_n \oplus q(x_1, \ldots, x_{n-1}).$$

Write $x = (x_1, \ldots, x_n)$, $y = (x_1, \ldots, x_{n-1})$. Let $\beta = (a_1, \ldots, a_{n-1}, 0)$ be any nonzero vector in V_n thus $\gamma = (a_1, \ldots, a_{n-1})$ is a nonzero vector in V_{n-1}. Due to the uniqueness of the vector e, $g(x) \oplus g(x \oplus \beta) = q(y) \oplus q(y \oplus \gamma)$ is a non-constant affine function and must be balanced. This proves that $q(y) \oplus q(y \oplus \gamma)$ is balanced for any nonzero vector $\gamma \in V_{n-1}$. Hence q does not have nonzero linear structures. By Lemma 4, q is bent and $n - 1$ is even. Thus n is odd.

(ii) From the proof of (i), we know that p is a constant a. Hence g can be expressed as $g(x) = ax_n \oplus q(x_1, \ldots, x_{n-1}) = (1 \oplus x_n)q(x_1, \ldots, x_{n-1}) \oplus x_n(a \oplus q(x_1, \ldots, x_{n-1}))$. Let ξ be the sequence of q. By (ii) of Lemma 1 in Section 1, $\langle \xi, \ell \rangle = \pm 2^{\frac{1}{2}(n-1)}$ for any affine sequence ℓ of length 2^{n-1}. By Lemma 7 of [18], $N_g \geqq 2^{n-1} - \frac{1}{2}(2^{\frac{1}{2}(n-1)} + 2^{\frac{1}{2}(n-1)}) = 2^{n-1} - 2^{\frac{1}{2}(n-1)}$. □

Theorem 6 indicates that differentially 2-uniform quadratic permutations are highly nonlinear and hence are immune to linear cryptanalysis.

Restating the part (i) of Theorem 6, we have:

Corollary 1 *There exists no differentially 2-uniform quadratic permutation on an even dimensional vector space.*

This gives a negative answer to an open problem regarding the existence of differentially 2-uniform quadratic permutations on an even dimensional vector space.

Now it is a right place to point out an error in [2]. Corollary 2 of [2] states that the permutation defined by a polynomial $P(x) = x^{2^\ell(2^k+1)}$ is a differentially 2-uniform quadratic permutation, where $x \in GF(2^n)$, ℓ, k and n are positive integers, and $\gcd(2^k + 1, 2^n - 1) = \gcd(k, n) = 1$. Beth and Ding claim that their corollary indicates the existence of differentially 2-uniform quadratic permutations on V_n, n even. This seemingly contradicts the non-existence result shown in our Corollary 1. However, one can see that when n is even, k must be odd in order for $\gcd(k, n) = 1$ to stand. On the other hand, if n is even and k is odd, then $\gcd(2^k + 1, 2^n - 1)$ has 3 as a factor. Thus $\gcd(2^k + 1, 2^n - 1) = \gcd(k, n) = 1$ can not stand for n even. In other words, Beth and Ding's corollary does not imply the existence of differentially 2-uniform quadratic permutations on V_n, n even.

3.2 SAC

Theorem 7 *Let $F = (f_1, \ldots, f_n)$ ($n \geqq 3$) be a differentially 2-uniform quadratic permutation. Then there exists a nonsingular matrix A of order n over $GF(2)$ such that $\Psi(x) = F(xA) = (f_1(xA), \ldots, f_n(xA)) = (\psi_1(x), \ldots, \psi_n(x))$ is also differentially 2-uniform, and each component function ψ_j satisfies the SAC.*

Proof. Let Φ denote the set of vectors γ such that $f_j \oplus f_j(x \oplus \gamma)$ is not balanced for some $1 \leqq j \leqq n$. By Lemma 5, we have $|\Phi| = n$. Since $|\Phi| < 2^{n-1}$ for all $n \geqq 3$, by Theorem 2 of [17], there exists a nonsingular matrix A of order n over $GF(2)$ that transforms F into a SAC-fulfilling S-box. □

4 Conclusion

We have proved that for quadratic S-boxes, there are close relationships among differential uniformity, linear structures, nonlinearity and the SAC. We have shown that by using our proof techniques, all differentially 2-uniform quadratic permutations can be treated in a unified fashion. In particular, general results regarding nonlinearity characteristics of these permutations are derived, regardless of the actual methods for constructing the permutations.

A future research direction is to extend the results to the more general case where component functions of an S-box can have an algebraic degree larger than 2. Another direction is to enlarge the scope of nonlinearity criteria so that it includes other cryptographic properties such as algebraic degree, propagation characteristics, and correlation immunity.

References

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

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

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

[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, pages 36–50. Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[6] J. Detombe and S. Tavares. Constructing large cryptographically strong S-boxes. In *Advances in Cryptology - AUSCRYPT'92*, volume 718, Lecture Notes in Computer Science, pages 165–181. Springer-Verlag, Berlin, Heidelberg, New York, 1993.

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

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

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

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

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

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

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

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

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

[16] 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*, pages 172 – 182. The Association for Computing Machinery, New York, 1993.

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

[18] J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearly balanced boolean functions and their propagation characteristics. In *Advances in Cryptology - CRYPTO'93*, volume 773, Lecture Notes in Computer Science, pages 49–60. Springer-Verlag, Berlin, Heidelberg, New York, 1994.

[19] 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, pages 181–199. Springer-Verlag, Berlin, Heidelberg, New York, 1994.

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

[21] A. F. Webster. Plaintext/ciphertext bit dependencies in cryptographic system. Master's Thesis, Department of Electrical Engineering, Queen's University, Ontario, 1985.

[22] A. F. Webster and S. E. Tavares. On the design of S-boxes. In *Advances in Cryptology - CRYPTO'85*, volume 219, Lecture Notes in Computer Science, pages 523–534. Springer-Verlag, Berlin, Heidelberg, New York, 1986.

Appendix

A Proof for Theorem 1

First we have

Lemma 7 *Let $L_i = (h_{i1}, \ldots, h_{i2^s})$ be the sequence of a linear function on V_s, where $i = 1, \ldots, 2^n$ $(n \geqq s)$. Set*

$$M = [L_1^T, \ldots, L_{2^n}^T].$$

If the rows of M are mutually orthogonal then each linear sequence of length 2^s appears as 2^{n-s} columns of M.

Proof. Let $\eta = (a_1, \ldots, a_{2^s})$ be a (1, -1) sequences of length 2^s. Since $\langle \eta, L_i \rangle = \sum_{p=1}^{2^s} a_p h_{ip}$, we have

$$\langle \eta, L_i \rangle^2 = 2^s + 2 \sum_{p<q} a_p a_q h_{ip} h_{iq}$$

and

$$\sum_{i=1}^{2^n} \langle \eta, L_i \rangle^2 = 2^{n+s} + 2 \sum_{i=1}^{2^n} \sum_{p<q} a_p a_q h_{ip} h_{iq} = 2^{n+s} + 2 \sum_{p<q} \sum_{i=1}^{2^n} a_p a_q h_{ip} h_{iq}.$$

Since rows of M are mutually orthogonal, we have $\sum_{j=1}^{2^n} h_{ip} h_{iq} = 0$ $(p \neq q)$ and hence

$$\sum_{j=1}^{2^n} \langle \eta, L_i \rangle^2 = 2^{n+s}. \tag{1}$$

Now suppose that L, an arbitrary linear sequence of length 2^s, appears as k columns of M. By noting

$$\langle L, L_i \rangle = \begin{cases} 2^s & \text{if } L = L_i \\ 0 & \text{otherwise} \end{cases}$$

we have

$$\sum_{j=1}^{2^n} \langle L, L_i \rangle^2 = k \cdot 2^{2s}. \tag{2}$$

Compare (1) and (2) we have

$$k \cdot 2^{2s} = 2^{n+s}$$

and hence $k = 2^{n-s}$. □

Note that (2) can be viewed as a generalization of Parseval's equation (Page 416, [8]). The following is the proof for Theorem 1.

Proof. (**for Theorem 1**) Suppose that F is a regular S-box, namely, $F(x)$ runs through each vector in V_s 2^{n-s} times while x runs through V_n, where $x = (x_1, \ldots, x_n)$. Then the truth table of each component function f_i must contain an equal number of ones and zeros, i.e., f_i is balanced.

Now we show that any nonzero linear combination, $f(x) = \sum_{j=1}^{s} c_j f_j(x)$, of the s component functions is also balanced. Recall that for any nonsingular matrix A of order s, $(f_1(x), \ldots, f_s(x))$ is regular if and only if $(f_1(x), \ldots, f_s(x))A$ is (see Lemma 3). Now suppose that the first column of A is $(c_1, \ldots, c_s)^T$. Let

$G(x) = (g_1(x), \ldots, g_s(x)) = (f_1(x), \ldots, f_s(x))A$. Then G is also regular, and hence its first component function $g_1(x) = f(x) = \sum_{j=1}^{s} c_j f_j(x)$ is balanced. This proves one direction of the theorem.

We now prove the other direction. Suppose that all nonzero linear combinations of the component functions are balanced. Let

$$\xi_i = (c_{i1}, \ldots, c_{i2^n})$$

be the truth table of f_i, $i = 1, \ldots, s$. From the s truth tables, we construct 2^n linear functions on V_s as follows:

$$\varphi_j(y) = c_{1j}y_1 \oplus c_{2j}y_2 \oplus \cdots \oplus c_{sj}y_s \tag{3}$$

where $y = (y_1, \ldots, y_s)$ and $j = 1, \ldots 2^n$.

Let

$$\eta_j = (b_{j1}, \ldots, b_{j2^s})$$

be the truth table of φ_j. Set

$$N = [\eta_1^T, \ldots, \eta_{2^n}^T].$$

Note that N is a $2^s \times 2^n$ matrix whose elements come from $GF(2)$.

N is constructed in such a way that its rows consist of precisely the 2^s different linear combinations of $\xi_1, \ldots, \xi_s$. To prove this is true, we take a close look at the rows of N. Let $\gamma_i = (b_{1i}, b_{2i}, \ldots, b_{2^n i})$ be the ith row of N, $0 \leqq i \leqq 2^s - 1$. Since $b_{ji} = \varphi_j(\alpha_i)$, where α_i is the vector in V_s corresponding to the integer i, we have $\gamma_i = (\varphi_1(\alpha_i), \varphi_2(\alpha_i), \ldots, \varphi_{2^n}(\alpha_i))$. Write $\alpha_i = (a_{i1}, \ldots a_{i2^s})$. Then

$$\begin{aligned}
\gamma_i &= (\sum_{j=1}^{s} c_{j1}a_{ij}, \sum_{j=1}^{s} c_{j2}a_{ij}, \ldots, \sum_{j=1}^{s} c_{j2^n}a_{ij}) \\
&= \sum_{j=1}^{s} a_{ij}(c_{j1}, c_{j2}, \ldots, c_{j2^n}) \\
&= \sum_{j=1}^{s} a_{ij}\xi_j.
\end{aligned}$$

This proves that γ_i, the ith row of N, is indeed a linear combination of $\xi_1, \ldots, \xi_s$. On the other hand, since any nonzero linear combination of $\xi_1, \ldots, \xi_s$ is balanced, $\xi_1, \ldots, \xi_s$ are linearly independent. Thus $\gamma_i \neq \gamma_j$ for any $i \neq j$. This proves our claim that the rows of N consist of precisely the 2^s different linear combinations of $\xi_1, \ldots, \xi_s$.

Now let M be an matrix obtained from N by substituting 0 with $+1$ and 1 with -1. Note that the sum of two different rows of N is a nonzero linear combination of $\xi_1, \ldots, \xi_s$ and hence balanced. This implies that the rows of M is mutually orthogonal. By Lemma 7 each linear sequence of length 2^s appears as 2^{n-s} columns of M. This in turn implies that the truth table of a linear function on V_s appears as 2^{n-s} columns of N, i.e. any linear function φ on V_s appears 2^{n-s} times in the set $\{\varphi_1, \ldots \varphi_{2^n}\}$, where φ_j is defined in (3). As there is a one to one correspondence between linear functions on V_s and vectors in V_s, we conclude that $F(x) = (f_1(x), \ldots, f_s(x))$ runs through each vector in V_s 2^{n-s} times while x runs through V_n. □