Relationships Among Nonlinearity Criteria (Extended Abstract)

Jennifer Seberry, Xian-Mo Zhang and Yuliang Zheng

Department of Computer Science, University of Wollongong
Wollongong, NSW 2522, Australia
{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, regardless of the underlying construction methods. This greatly simplifies the proofs for a number of known results on nonlinearity characteristics of quadratic permutations. As a by-product, we obtain a negative answer to an open problem regarding the existence of differentially 2-uniform quadratic permutations on an even dimensional vector space.

1 Nonlinearity Criteria

We first introduce 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. These functions take the form of $a_{00} \oplus \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j$, where a_{ij} is an element from $GF(2)$, while x_i is a variable in $GF(2)$.

Let f be a function on V_n. The $(1, -1)$-sequence defined by $((-1)^{f(\alpha_0)}$, $(-1)^{f(\alpha_1)}$, ..., $(-1)^{f(\alpha_{2^n-1})})$ is called the *sequence* of f, and the $(0, 1)$-sequence defined by $(f(\alpha_0)$, $f(\alpha_1)$, ..., $f(\alpha_{2^n-1}))$ is called the *truth table* of f, where $\alpha_0 = (0, \ldots, 0, 0)$, $\alpha_1 = (0, \ldots, 0, 1)$, ..., $\alpha_{2^n-1} = (1, \ldots, 1, 1)$. f is said to be *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 [18].)

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 2. *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 [24, 25] when studying the design of cryptographically strong substitution boxes (S-boxes).

Definition 3. 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 [22].

While the SAC measures the avalanche characteristics of a function, the linear structure is a concept that in a sense complements the former, namely, it indicates the straightness of a function.

Definition 4. 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.

Evertse apparently was the first person who studied implications of linear structures (in a sense broader than ours) on the security of encryption algorithms [8]. By definition, the zero vector in V_n is a linear structure of all functions on V_n. It is not hard to see that 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 0 when n is even and by 1 when n is odd. The lower bound 0 is achieved only by bent functions that have the zero vector as their only linear structure, while 1 can be achieved by functions that have only two linear structures (one is the zero vector and the other is a nonzero vector). Examples of the latter are those obtained by concatenating two bent functions (see [19, 23]).

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 ($n \geq s$). To avoid trivial statistical attacks, an S-box F should be *regular*, namely, $F(x)$ should run through all 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 5. *Let $F = (f_1, \ldots, f_s)$, where f_i is a function on V_n, $n \geq s$. 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 will be given in the full version. 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 6. 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, $\ldots$, $\varphi_{2^{n+1}}$ 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 [19, 23].

Nonlinearity has been considered to be an important criterion. Recent advances in *Linear cryptanalysis* put forward by Matsui [10, 11] have further 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 [21], 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 7. Let F be an $n \times s$ S-box, where $n \geq s$. 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.

Obviously the differential uniformity δ of an $n \times s$ S-box is constrained by $2^{n-s} \leq \delta \leq 2^n$. Extensive research has been carried out in constructing differentially δ-uniform S-boxes with a low δ [13, 1, 14, 16, 15, 2]. Some constructions, in particular those based on permutation polynomials on finite fields, are simple and elegant. However, caution must be taken with Definition 7. 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 [13, 1] are extremely weak to differential attacks, despite that they achieve the lowest possible differential uniformity $\delta = 2^{n-s}$ [4, 5, 21]. A more complete measurement is the *robustness* introduced in [21]. The reader is directed to that paper for a comprehensive treatment of this subject.

Note that an $n \times s$ S-box achieves the lowest possible differential uniformity $\delta = 2^{n-s}$ if and only if it has a *flat* difference distribution table. As has been noticed by many researchers (see for instance Page 62 of [4]), a flat difference distribution table is not associated with a regular S-box. This result, together with a formal proof, is now given explicitly.

Lemma 8. *The differential uniformity of a regular $n \times s$ S-box is larger than 2^{n-s}.*

Proof. Let F is a regular $n \times s$ S-box. By Theorem 5, nonzero linear combinations of the component functions of F are all balanced. Assume for contradiction that for each nonzero $\alpha \in V_n$, $F(x) \oplus F(x \oplus \alpha)$ is regular, namely it runs through all vectors in V_s, each 2^{n-s} times, while x runs through V_n once. Recall that Theorem 3.1 of [13] states that $F(x) \oplus F(x \oplus \alpha)$ is regular if and only if each nonzero linear combination of the component functions of F is a bent function. Thus the assumption contradicts the fact that each nonzero linear combination of the component functions of F is balanced.

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 9. *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.*

The next lemma was pointed out in Section 5.3 of [21]. It was also noticed by Beth and Ding in [2]. The lemma is followed by a short formal proof for the sake of completeness.

Lemma 10. *Let F be a mapping from V_n to V_s, where $n \geq s$, 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.*

Proof. Let β be a vector in V_s. Since $F(x) = G(xA^{-1})$, $F(x) = \beta$ if and only if $G(xA^{-1}) = \beta$. This implies that, while x runs through V_n, $F(x)$ and $G(x)$ run through β the same number of times.

Now consider $H(x) = F(x)B$. Clearly $F(x) = \beta$ if and only if $H(x) = F(x)B = \beta B$. As B is nonsingular, $F(x)$ runs through β exactly the same number of times as that $H(x)$ runs through βB, while x runs through V_n.

2 Cryptographic Properties of Quadratic S-boxes

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

2.1 Linear Structure vs Nonlinearity

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 2, f is bent. Thus we have:

Lemma 11. *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 12. *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_2}. Let d_1 and d_2 denote the nonlinearities of f_1 and f_2 respectively. Then the nonlinearity of g satisfies*

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

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

Proof. The first half of the lemma can be found in Lemma 8 of [20]. The second half is true due to the fact that $d_1 \leq 2^{n_1 - 1}$ and $d_2 \leq 2^{n_2 - 1}$ (see also Section 3 of [19]).

We now examine how the nonlinearity of a function on V_n relates to the linearity dimension of the function.

Let g be a (not necessarily quadratic) function on V_n, $\{\beta_1, \ldots, \beta_\ell\}$ 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 9, g^* and g have the same linearity dimension, algebraic degree and nonlinearity. Thus the question is transformed into the discussion of g^*.

Let e_i be the vector in V_n whose ith coordinate is one and others are zero. Then we have $e_j B = \beta_j$, and $g^*(e_i) = g(\beta_i)$, $i = 1, \ldots, n$. Thus $\{e_1, \ldots, e_\ell\}$ is a basis of the subspace consisting of the linear structures of g^*. Write g^* as

$$g^*(x) = q(y) \oplus \sum_j [m_j(y) r_j(z)] \tag{1}$$

where $x = (x_1, \ldots, x_n)$, $y = (x_1, \ldots, x_\ell)$, $z = (x_{\ell+1}, \ldots, x_n)$, $m_j \neq 0$, the algebraic degree of each r_j is at least 1 and $r_j \neq r_i$ for $j \neq i$. Also write e_i as

$e_i = (\mu_i, 0)$, where $\mu_j \in V_\ell$ and $0 \in V_{n-\ell}$. As e_i is a linear structure of g^*, the following difference

$$g^*(x) \oplus g^*(x \oplus e_i) = q(y) \oplus q(y \oplus \mu_i) \oplus \sum_j [(m_j(y) \oplus m_j(y \oplus \mu_i)) r_j(z)]$$

is a constant. This implies that $q(y) \oplus q(y \oplus \mu_i)$ is a constant (i.e. μ_i is a linear structure of $q(y)$) and each $m_j(y) \oplus m_j(y \oplus \mu_j) = 0$ (i.e. $m_j = 1$). Thus (1) can be rewritten as

$$g^*(x) = q(y) \oplus r(z). \tag{2}$$

Since all vectors in V_ℓ are linear structures of q, q is an affine function on V_ℓ. As the linearity dimension of g^* is also ℓ, r must be a function on $V_{n-\ell}$ that does *not* have nonzero linear structures. By Lemmas 9 and 12, we have $N_g = N_{g^*} = 2^\ell N_r$. This is precisely what Proposition 3 of [14] states.

As a special case, suppose that g in the above discussions is quadratic. Then the function r in (2) is a quadratic function on $V_{n-\ell}$ with no nonzero linear structures. By Lemma 11, r is a bent function on $V_{n-\ell}$ whose nonlinearity is $N_r = 2^{n-\ell-1} - 2^{\frac{1}{2}(n-\ell)-1}$. Thus we have:

Theorem 13. *Let g be a function on V_n whose algebraic degree is at most 2. Denote by ℓ the linearity dimension of g. Then*

(i) $n - \ell$ is even, and
(ii) the nonlinearity of g satisfies $N_g = 2^{n-1} - 2^{\frac{1}{2}(n+\ell)-1}$.

The lower bound on nonlinearity in Theorem 13 can be straightforwardly translated into that for quadratic (not necessarily regular) $n \times s$ S-boxes ($n \geq s$).

Now we take a closer look at the nonlinearity of a quadratic function g on V_n. As g is nonlinear, we have $\ell < n$, where ℓ is the linearity dimension of g. In addition since g is quadratic, by (i) of Theorem 13, $n - \ell$ is even. Thus we have $\ell \leq n-2$, and $N_g \geq 2^{n-1} - 2^{\frac{1}{2}(n+\ell)-1} \geq 2^{n-2}$. This proves the following:

Corollary 14. *The nonlinearity of a quadratic function on V_n is at least 2^{n-2}.*

Corollary 14 is a bit surprising in the sense that it indicates that all quadratic functions are fairly nonlinear, and there is no quadratic function whose nonlinearity is between 0 and 2^{n-2} (exclusive).

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 15. *Let δ be the differential uniformity of a quadratic $n \times s$ S-box. Then $\delta = 2^d$ for some $n - s \leq d \leq n$. Furthermore, if the S-box is regular, then we have $\delta = 2^d$ for some $n - s + 1 \leq d \leq n$.*

Let $F = (f_1, \ldots, f_s)$ be a regular quadratic $n \times s$ S-box, and let g be a nonlinear combination of the component functions of F. Then it can be shown that g has at least one nonzero linear structure. To prove the claim, we assume that g has no nonzero linear structures. Then by Lemma 11, 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 quadratic or affine functions and hence have linear structures.

Next we show that the differential uniformity of an S-box is closely related to the number of linear structures of an nonzero linear combinations of the component functions of the S-box.

Theorem 16. *Let $F = (f_1, \ldots, f_s)$ be a regular quadratic $n \times s$ S-box. Then the differential uniformity of F satisfies $\delta \le 2^{n-s+t}$, where $1 \le t \le s$ (see also Theorem 15), if and only if any nonzero vector $\alpha \in V_n$ is a linear structure of at most $2^t - 1$ nonzero linear combinations of $f_1, \ldots, f_s$.*

Theorem 16 indicates that with an S-box with a smaller δ, i.e., a smaller t, the nonzero linear combinations of its component functions have less linear structures. This coincides with our intuition that the nonlinearity of an S-box grows with the strength of its immunity to differential attacks.

2.3 Difference Distribution Table vs SAC

Armed with Theorem 16, we further reveal that differential uniformity is tightly associated with the strict avalanche characteristics.

Theorem 17. *Let $F = (f_1, \ldots, f_s)$ be a differentially δ-uniform regular quadratic $n \times s$ S-box, where $\delta = 2^{n-s+t}$, $1 \le t \le s$ (see also Theorem 15). If t and s satisfy $s \le 2^{s-t-2}$, 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 $g_1, \ldots, g_{2^s-1}$ the $2^s - 1$ nonzero linear combinations of $f_1, \ldots, f_s$, and by $\alpha_1, \ldots, \alpha_{2^n-1}$ the $2^n - 1$ nonzero vectors in V_s. We construct a bipartite graph Γ with $g_1, \ldots, g_{2^s-1}$ on one side and $\alpha_1, \ldots, \alpha_{2^n-1}$ on the other side. An edge exists between g_i and α_j if and only if α_j is a linear structure of g_i. By Theorem 16, 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 $t_1 \le t_2 \le \cdots \le t_{2^s-1}$. 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 $t_1 + \cdots + t_{2^s-1} \ge 2^{s-1} \cdot 2^{n-s+t+1} = 2^{n+t} > (2^t - 1) \cdot (2^n - 1)$. This contradicts the fact that Γ has at most $2^{t-1} \cdot (2^n - 1)$ edges.

Now set $\Omega = \{g_1, \ldots, g_{2^{s-1}+1}\}$. As the rank of Ω is s, we can choose s functions from Ω, say $g_{j_1}, \ldots, g_{j_s}$, such that they are all linearly independent.

Since $s \leq 2^{s-t-2}$, we have $t_{j_1} + \cdots + t_{j_s} < s \cdot 2^{n-s+t+1} \leq 2^{n-1}$. By Theorem 2 of [22], 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 $f_1, \ldots, f_s$, 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 10,

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

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

In Theorem 17, when the differential uniformity $\delta = 2^{n-s+t}$ is small, the parameter t is also small, and the condition $s \leq 2^{s-t-2}$ is readily satisfied. Equivalently we can say that S-boxes strong against differential attacks are also SAC-fulfilling, subject to a nonsingular linear transformation. Again, this coincides with our intuition.

3 A Unified Treatment of Quadratic Permutations

This section is concerned with differentially 2-uniform quadratic $n \times n$ S-boxes. Since such an S-box F is a permutation, $F(x) \oplus F(x \oplus \alpha)$ takes a vector two times or does not take it, while x runs through V_n once. 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.

Although there are many question marks regarding the applicability of differentially 2-uniform quadratic $n \times n$ S-boxes in computer security practices, primarily due to their low algebraic degree, these S-boxes have received extensive research in the past years [17, 16, 6, 2, 15] and hence deserve our special attention. These S-boxes appear in various forms and researchers have employed different techniques, some of which are rather sophisticated, to prove their nonlinearity characteristics. 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 proofs for a number of known results.

Theorem 18. *Let $F = (f_1, \ldots, f_n)$ be a quadratic permutation on V_n. Then the following statements are equivalent:*

(i) for any nonzero linear combination of $f_1, \ldots, f_n$, say $g(x) = \sum_{j=1}^{n} c_j f_j(x)$, its nonlinearity satisfies $N_g = 2^{n-1} - 2^{\frac{1}{2}(n-1)}$.
(ii) any nonzero linear combination of $f_1, \ldots, f_n$, say $g(x) = \sum_{j=1}^{n} c_j f_j(x)$, has a unique nonzero linear structure.
(iii) each nonzero vector in V_n is the linear structure of a unique nonzero linear combination of $f_1, \ldots, f_n$.

(iv) F is differentially 2-uniform, i.e. for each nonzero vector $\alpha \in V_n$, $F(x) \oplus F(x \oplus \alpha)$ runs through half of the vectors in V_n while x runs through V_n.
(v) every nonzero linear combination of the component functions, say g, can be expressed as $g(x) = xCx^T$, where $x = (x_1, \dots, x_n)$, C is a matrix over $GF(2)$ and the rank of $C \oplus C^T$ is $n-1$.

Proof. The equivalence of (i) and (ii): By (ii) of Theorem 13, a quadratic function has a nonlinearity $2^{n-1} - 2^{\frac{1}{2}(n-1)}$ if and only if its linearity dimension is 1.

The equivalence of (ii) and (iii): Let $\alpha_1, \dots, \alpha_{2^n-1}$ be the $2^n - 1$ nonzero vectors in V_n and $g_1, \dots, g_{2^n-1}$ be the $2^n - 1$ nonzero linear combinations of $f_1, \dots, f_n$. Similarly to the proof of Theorem 17, we construct a bipartite graph Γ with $\alpha_1, \dots, \alpha_{2^n-1}$ on one side and $g_1, \dots, g_{2^n-1}$ on the other side. A link exists between α_i and g_j if and only if α_i is a linear structure of g_j. Since each g_j is balanced, it must not be a bent function. By Lemma 11, each g_j has at least one nonzero linear structure. ¿From the construction of Γ, we can see that each g_j has an edge associated with it. On the other hand, for any nonzero vector, say α, $F(x) \oplus F(x \oplus \alpha)$ does not run through the vector zero, as $F(x)$ is a permutation on V_n. By Theorem 5, there exists a nonzero linear combination of the component functions of $F(x) \oplus F(x \oplus \alpha)$, say

$$\sum_{j=1}^{n} c_j [f_j(x) \oplus f_j(x \oplus \alpha)], \tag{3}$$

that is not balanced. Since f_j is quadratic, (3) is affine. Thus (3) must be a constant. Write $g_\alpha(x) = \sum_{j=1}^{n} c_j f_j(x)$. Then α is a nonzero linear structure of g_α. Thus each α has at least one edge associated with it. In summary, each g_j has at least one edge associated with it, and so does each α_j. As both sides of the bipartite graph have the same number of edges, (ii) and (iii) must stand in parallel.

The equivalence of (iii) and (iv): First we note that the differential uniformity of a permutation is at least 2. Let $s = n$ and $t = 1$. Then By Theorem 16, F is differential 2-uniform if and only if each nonzero vector in V_n is the linear structure of at most one nonzero linear combination of $f_1, \dots, f_n$. In the proof of the equivalence of (ii) and (iii), it is has been shown that each nonzero vector in V_n is a linear structure of at least one nonzero linear combination of the component functions. Thus F is differential 2-uniform if and only if each nonzero vector in V_n is the linear structure of a unique nonzero linear combination of the component functions.

The equivalence of (iv) and (v): Note that for any quadratic function g on V_n, there exists an $n \times n$ matrix C on $GF(2)$ such that $g(x) = xCx^T$. In [16], where the statement (v) is called the property (P), Nyberg and Knudsen proved that (v) implies (iv). We now show that the opposite is also true. Suppose that F is a differentially 2-uniform permutation on V_n. Let g be a nonzero linear combination of the component functions, and let C a matrix such that $g(x) = xCx^T$. By (ii), we have $\ell = 1$, where ℓ is the linearity dimension of g. By Proposition 4 of [14], the linearity dimension of g and the rank of $C \oplus C^T$ satisfy the following relation: $\ell = n - \text{rank}(C \oplus C^T)$. Hence we have $\text{rank}(C \oplus C^T) = n-1$, namely (iv) implies (v). This proves the equivalence of (iv) and (v).

An important corollary of Theorem 18 is:

Corollary 19. *There exists no differentially 2-uniform quadratic permutation on an even dimensional vector space.*

Proof. Let $F(x) = (f_1, \ldots, f_n)$ be a differentially 2-uniform quadratic permutation on V_n. By (ii) of Theorem 18, each component function f_i has a unique nonzero linear structure. Hence the linearity dimension of f_i is 1, and the corollary follows immediately from Part (i) of Theorem 13.

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 19. 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.

The following result has been pointed out by these authors in [22]. It is included here, together with its proof, for the sake of completeness.

Theorem 20. *Let $F = (f_1, \ldots, f_n)$ ($n \geq 3$) 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. When $n \geq 7$, it directly follows from Theorem 17. The proof described below applies to all $n \geq 3$.

Let Φ denote the set of vectors γ such that $f_j \oplus f_j(x \oplus \gamma)$ is not balanced for some $1 \leq j \leq n$. By (ii) and (iii) of Theorem 18, we have $|\Phi| = n$. Since $|\Phi| < 2^{n-1}$ for all $n \geq 3$, by Theorem 2 of [22], 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 underlying 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 examined so that it includes other cryptographic properties such as algebraic degree, propagation characteristics, and correlation immunity.

Acknowledgments

The first author was supported in part by the Australian Research Council under the reference numbers A49130102, A9030136, A49131885 and A49232172, the second author by A49130102, and the third author by A49232172.

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*. Springer-Verlag, Berlin, Heidelberg, New York, 1993. to appear.
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, Berlin, Heidelberg, London, Paris, Tokyo, Hong Kong, Barcelona, Budapest, 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*. Springer-Verlag, Berlin, Heidelberg, New York, 1991. to appear.
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. J.-H. Evertse. Linear structures in blockciphers. In *Advances in Cryptology - EUROCRYPT'87*, volume 304, Lecture Notes in Computer Science, pages 249–266. Springer-Verlag, Berlin, Heidelberg, New York, 1988.
9. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. New York: North-Holland, 1977.
10. M. Matsui. Linear cryptanalysis method for DES cipher. In *Advances in Cryptology - EUROCRYPT'93*. Springer-Verlag, Berlin, Heidelberg, New York, 1993. to appear.
11. M. Matsui. Linear cryptanalysis method for DES cipher (II). In *Proceedings of 1994 Symposium on Cryptography and Information Security*, Japan, 1994.

12. W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. In *Advances in Cryptology - EUROCRYPT'89*, volume 434, Lecture Notes in Computer Science, pages 549–562. Springer-Verlag, Berlin, Heidelberg, New York, 1990.
13. 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.
14. K. Nyberg. On the construction of highly nonlinear permutations. In *Advances in Cryptology - EUROCRYPT'92*. Springer-Verlag, Berlin, Heidelberg, New York, 1992. to appear.
15. K. Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT'93*. Springer-Verlag, Berlin, Heidelberg, New York, 1993. to appear.
16. K. Nyberg and L. R. Knudsen. Provable security against differential cryptanalysis. In *Advances in Cryptology - CRYPTO'92*, volume Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1992. to appear.
17. J. Pieprzyk. Bent permutations. In *Proceeding of the International Conference on Finite Fields, Coding Theory, and Advances in Communications and Computing*, Las Vegas, 1991.
18. O. S. Rothaus. On "bent" functions. *Journal of Combinatorial Theory*, Ser. A, 20:300–305, 1976.
19. J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearly balanced boolean functions and their propagation characteristics. In *Advances in Cryptology - CRYPTO'93*. Springer-Verlag, Berlin, Heidelberg, New York, 1993. to appear.
20. J. Seberry, X. M. Zhang, and Y. Zheng. On constructions and nonlinearity of correlation immune functions. In *Advances in Cryptology - EUROCRYPT'93*. Springer-Verlag, Berlin, Heidelberg, New York, 1993. to appear.
21. J. Seberry, X. M. Zhang, and Y. Zheng. Systematic generation of cryptographically robust S-boxes. In *Proceedings of the first ACM Conference on Computer and Commnications Security*, pages 172 – 182. The Association for Computing Machinery, New York, 1993.
22. J. Seberry, X. M. Zhang, and Y. Zheng. Improving the strict avalanche characteristics of cryptographic functions. *Information Processing Letters*, 50:37–41, 1994.
23. J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearity and propagation characteristics of balanced boolean functions. To appear in *Information and Computation*, 1994.
24. A. F. Webster. Plaintext/ciphertext bit dependencies in cryptographic system. Master's Thesis, Department of Electrical Engineering, Queen's University, Ontario, Cannada, 1985.
25. A. F. Webster and S. E. Tavares. On the designs 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.

This article was processed using the LaTeX macro package with LLNCS style