The Relationship Between Propagation Characteristics and Nonlinearity of Cryptographic Functions

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

The connections among the various nonlinearity criteria is currently an important topic in the area of designing and analyzing cryptographic functions. In this paper we show a quantitative relationship between propagation characteristics and nonlinearity, two critical indicators of the cryptographic strength of a Boolean function. We also present a tight lower bound on the nonlinearity of a cryptographic function that has propagation characteristics.

Key Words

Cryptography, Boolean functions, Encryption functions, Nonlinearity, Propagation Characteristics, SAC, S-boxes.

1 Introduction

Data Encryption Standard or DES is a cryptographic algorithm most widely used by industrial, financial and commercial sectors all over the world [23]. DES is also the root of many other data encryption algorithms proposed in the past decade, including LOKI [3], FEAL [12] and IDEA [9, 8, 7]. A core component of these encryption algorithms is so-called S-boxes or substitution boxes, each essentially a tuple of nonlinear Boolean functions. In most cases, these boxes are the only nonlinear component in an underlying encryption algorithm. The same can be said with one-way hashing algorithms which are commonly employed in the process of signing and authenticating electronic messages [27, 16, 13]. These all indicate the vital importance of the design and analysis of nonlinear cryptographic Boolean functions.

Encryption and authentication require cryptographic (Boolean) functions with a number of critical properties that distinguish them from linear (or affine) functions. Among the properties are high nonlinearity, high degree of propagation, few linear structures, high algebraic degree etc. These properties are often called *nonlinearity criteria*. An important topic is to investigate relationships among the various nonlinearity criteria. Progress in this direction has been made in [21], where connections have been revealed among the strict avalanche characteristic, differential characteristics, linear structures and nonlinearity, of *quadratic* functions.

In this paper we carry on the investigation initiated in [21] and bring together nonlinearity and propagation characteristic of a Boolean function (quadratic or non-quadratic). These two cryptographic criteria are seemly quite separate, in the sense that the former indicates the minimum distance between a Boolean function and all the affine functions whereas the latter forecasts the avalanche behavior of the function when some input bits to the function are complemented.

In particular we show that if f, a function on V_n, satisfies the propagation criterion with respect to all but a subset $\Re$ of V_n, then the nonlinearity of f satisfies $N_f \geqq 2^{n-1} - 2^{n-\frac{1}{2}\rho-1}$, where ρ is the maximum dimension a linear subspace contained in $\{0\} \cup (V_n - \Re)$ can achieve.

We also show that 2^{n-2} is the tight lower bound on the nonlinearity of f if f satisfies the propagation criterion with respect to at least one vector in V_n. As an immediate consequence, the nonlinearity of a function that fulfills the SAC or strict avalanche criterion is at least 2^{n-2}.

Two techniques are employed in the proofs of our main results. The first technique is in regard to the structure of $\Re$, the set of vectors where the function f does not satisfy the propagation criterion. By considering a linear subspace with the maximum dimension contained in $\{0\} \cup (V_n - \Re)$, together with its complementary subspace, we will be able to identify how the vectors in $\Re$ are distributed. The second technique is based on a novel idea of refining Parseval's equation, a well-known relationship in the theory of orthogonal transforms. A combination of these two techniques together with some careful analyses proves to be a powerful tool in examining the relationship among nonlinearity criteria.

The organization of the rest of the paper is as follows: Section 2 introduces basic notations and conventions, while Section 3 presents background information on the Walsh-Hadamard transform. The distribution of vectors where the propagation criterion is not satisfied is discussed in Section 4. This result is employed in Section 5 where a quantitative relationship between nonlinearity and propagation characteristics is derived. This relationship is further developed in Section 6 to identify a tight lower bound on nonlinearity of functions with propagation characteristics. The paper is closed by some concluding remarks in Section 7.

2 Basic Definitions

We consider Boolean functions from V_n to $GF(2)$ (or simply functions on V_n), V_n is the vector space of n tuples of elements from $GF(2)$. The *truth table* of a function f on V_n is a $(0,1)$-sequence defined by $(f(\alpha_0),\ f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$, and the *sequence* of f is a $(1,-1)$-sequence defined by $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-$ where $\alpha_0 = (0, \ldots, 0, 0)$, $\alpha_1 = (0, \ldots, 0, 1)$, $\ldots$, $\alpha_{2^{n-1}-1} = (1, \ldots, 1, 1)$. The *matrix* of f is a $(1,-1)$-matrix of order 2^n defined by $M = ((-1)^{f(\alpha_i \oplus \alpha_j)})$. f is said to be *balanced* if its truth table contains an equal number of ones and zeros.

An *affine* function f on V_n is a function that takes the form of $f(x_1, \ldots, x_n) = a_1x_1 \oplus \cdots \oplus a_nx_n \oplus c$, where $a_j, c \in GF(2)$, $j = 1, 2, \ldots, n$. Furthermore f is called a *linear* function if $c = 0$.

Definition 1 *The* Hamming weight *of a* $(0,1)$*-sequence s, denoted by $W(s)$, is the number of ones in the sequence. Given two functions f and g on V_n, the* Hamming distance $d(f,g)$ *between them is defined as the Hamming weight of the truth table of $f(x) \oplus g(x)$, 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, $\ldots$, $\varphi_{2^{n+1}}$ are all the affine functions on V_n.*

Note that the maximum nonlinearity of functions on V_n coincides with the covering radius of the first order binary Reed-Muller code $RM(1,n)$ of length 2^n, which is bounded from above by $2^{n-1} - 2^{\frac{1}{2}n-1}$ (see for instance [4]). Hence $N_f \leqq 2^{n-1} - 2^{\frac{1}{2}n-1}$ for any function on V_n. Next we introduce the definition of propagation criterion.

Definition 2 *Let f be a function on V_n. We say that f satisfies*

1. *the* propagation criterion with respect to α *if $f(x) \oplus f(x \oplus \alpha)$ is a balanced function, where $x = (x_1, \ldots, x_n)$ and α is a vector in V_n.*

2. *the* propagation criterion of degree k *if it satisfies the propagation criterion with respect to all $\alpha \in V_n$ with $1 \leqq W(\alpha) \leqq k$.*

$f(x) \oplus f(x \oplus \alpha)$ is also called the directional derivative of f in the direction α. The above definition for propagation criterion is from [15]. Further work on the topic can be found in [14]. Note that the strict avalanche criterion (SAC) introduced by Webster and Tavares [24, 25] is equivalent to the propagation criterion of degree 1 and that the perfect nonlinearity studied by Meier and Staffelbach [11] is equivalent to the propagation criterion of degree n where n is the number of the coordinates of the function.

While the propagation characteristic measures the avalanche effect 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 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. 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. We note that it was Evertse who first introduced the notion of linear structure (in a sense broader than ours) and studied its implication on the security of encryption algorithms [6].

A $(1, -1)$-matrix H of order m is called a *Hadamard* matrix if $HH^t = mI_m$, where H^t is the transpose of H and I_m is the identity matrix of order m. A Sylvester-Hadamard matrix of order 2^n, denoted by H_n, is generated by the following recursive relation

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

Let ℓ_i, $0 \leqq i \leqq 2^n - 1$, be the i row of H_n. By Lemma 2 of [20], ℓ_i is the sequence of a linear function $\varphi_i(x)$ defined by the scalar product $\varphi_i(x) = \langle \alpha_i, x \rangle$, where α_i is the ith vector in V_n according to the ascending alphabetical order.

Definition 4 *Let f be a function on V_n. The Walsh-Hadamard transform of f is defined as*

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

where $\alpha = (a_1, \ldots, a_n) \in V_n$, $x = (x_1, \ldots, x_n)$, $\langle \alpha, x \rangle$ is the scalar product of α and x, namely, $\langle \alpha, x \rangle = \bigoplus_{i=1}^n a_i x_i$, and $f(x) \oplus \langle \alpha, x \rangle$ is regarded as a real-valued function.

The Walsh-Hadamard transform, also called the discrete Fourier transform, has numerous applications in areas ranging from physical science to communications engineering. It appears in several slightly different forms [17, 10, 5]. The above definition follows the line in [17]. It can be equivalently written as

$$(\hat{f}(\alpha_0), \hat{f}(\alpha_1), \ldots, \hat{f}(\alpha_{2^n-1})) = 2^{-\frac{n}{2}} \xi H_n$$

where α_i is the ith vector in V_n according to the ascending order, ξ is the sequence of f and H_n is the Sylvester-Hadamard matrix of order 2^n.

Definition 5 *A function f on V_n is called a* bent *function if its Walsh-Hadamard transform satisfies*

$$\hat{f}(\alpha) = \pm 1$$

for all $\alpha \in V_n$.

Bent functions can be characterized in various ways [1, 5, 20, 26]. In particular the following four 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 satisfies the propagation criterion with respect to all non-zero vectors in V_n.

(iv) M, the matrix of f, is a Hadamard matrix.

Bent functions on V_n exist only when n is even [17]. Another important property of bent functions is that they achieve the highest possible nonlinearity $2^{n-1} - 2^{\frac{1}{2}n-1}$.

3 More on Walsh-Hadamard transform and Nonlinearity

As the Walsh-Hadamard transform plays a key role in the proofs of main results to be described in the following sections, this section provides some background knowledge on the transform. More information regarding the transform can be found in [10, 5]. In addition, Beauchamp's book [2] is a good source of information on other related orthogonal transforms with their applications.

Given two sequences $a = (a_1, \ldots, a_m)$ and $b = (b_1, \ldots, b_m)$, their component-wise product is defined by $a * b = (a_1 b_1, \ldots, a_m b_m)$. 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 $f(x) \oplus f(x \oplus \alpha)$.

Set

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

the scalar product of $\xi(0)$ and $\xi(\alpha)$. $\Delta(\alpha)$ is also called the auto-correlation of f with a shift α. Obviously, $\Delta(\alpha) = 0$ if and only if $f(x) \oplus f(x \oplus \alpha)$ is balanced, i.e., f satisfies the propagation criterion with respect to α. On the other hand, if $|\Delta(\alpha)| = 2^n$, then $f(x) \oplus f(x \oplus \alpha)$ is a constant and hence α is a linear structure of f.

Let $M = ((-1)^{f(\alpha_i \oplus \alpha_j)})$ be the matrix of f and ξ be the sequence of f. Due to a very pretty result by R. L. McFarland (see Theorem 3.3 of [5]), M can be decomposed into

$$M = 2^{-n} H_n \,\mathrm{diag}(\langle \xi, \ell_0 \rangle, \cdots, \langle \xi, \ell_{2^n-1} \rangle) H_n \tag{2}$$

where ℓ_i is the ith row of H_n, a Sylvester-Hadamard matrix of order 2^n.

Clearly

$$MM^T = 2^{-n} H_n \,\mathrm{diag}(\langle \xi, \ell_0 \rangle^2, \cdots, \langle \xi, \ell_{2^n-1} \rangle^2) H_n. \tag{3}$$

On the other hand, we always have

$$MM^T = (\Delta(\alpha_i \oplus \alpha_j)),$$

where $i, j = 0, 1, \ldots, 2^n - 1$.

Compare the two sides of (3), we have

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

Equivalently we write

$$(\Delta(\alpha_0), \Delta(\alpha_1), \dots, \Delta(\alpha_{2^n-1}))H_n = (\langle \xi, \ell_0 \rangle^2, \dots, \langle \xi, \ell_{2^n-1} \rangle^2). \tag{4}$$

In engineering, (4) is better known as (a special form of) the Wiener-Khintchine Theorem [2]. A closely related result is Parseval's equation (Corollary 3, p. 416 of [10])

$$\sum_{j=0}^{2^n-1} \langle \xi, \ell_j \rangle^2 = 2^{2n}$$

which also holds for any function f on V_n.

Let S be a set of vectors in V_n. The *rank* of S is the maximum number of linearly independent vectors in S. Note that when S forms a linear subspace of V_n, its rank coincides with its dimension.

The distance between two functions f_1 and f_2 on V_n can be expressed as $d(f_1, f_2) = 2^{n-1} - \frac{1}{2}\langle \xi_1, \xi_2 \rangle$, where ξ_1 and ξ_2 are the sequences of f_1 and f_2 respectively. (For a proof see for instance Lemma 6 of [20].) Immediately we have:

Lemma 1 *The nonlinearity of a function f on V_n can be calculated by*

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

where ξ is the sequence of f and ℓ_0, ..., ℓ_{2^n-1} are the rows of H_n, namely, the sequences of the linear functions on V_n.

The next lemma regarding splitting the power of 2 can be found in [21]

Lemma 2 *Let $n \geqq 2$ be a positive integer and $p^2 + q^2 = 2^n$ where both $p \geqq 0$ and $q \geqq 0$ are integers. Then $p = 2^{\frac{1}{2}n}$ and $q = 0$ when n is even, and $p = q = 2^{\frac{1}{2}(n-1)}$ when n is odd.*

In the next section we examine the distribution of the vectors in $\Re$.

4 Distribution of $\Re$

Let f be a function on V_n. Assume that f satisfies the propagation criterion with respect to all but a subset $\Re$ of V_n. Note that $\Re$ always contains the zero vector 0. Write $\Re = \{0, \gamma_1, \dots, \gamma_s\}$. Thus $|\Re| = s+1$.

Set $\Re^c = V_n - \Re$. Then f satisfies the propagation criterion with respect to all vectors in $\Re^c$.

Consider the set of vectors $\{0\} \cup \Re^c$. Then $\{0\}$ is a linear subspace contained in $\{0\} \cup \Re^c$. When $|\{0\} \cup \Re^c| > 1$, $\{0, \gamma\}$ is a linear subspace for any nonzero vector in $\Re^c$. We are particularly interested in linear subspaces with the maximum dimension contained in $\{0\} \cup \Re^c$. For convenience, denote by ρ the maximum dimension and by W a linear subspace in $\{0\} \cup \Re^c$ that achieves the maximum dimension.

Obviously, f is bent if and only if $\rho = n$, and f does not satisfy the propagation criterion with respect to any vector if and only if $\rho = 0$. The case when $1 \leqq \rho \leqq n - 1$ is especially interesting.

Now let U be a complementary subspace of W, namely $U \oplus W = V_n$. Then each vector $\gamma \in V_n$ can be uniquely expressed as $\gamma = \alpha \oplus \beta$, where $\alpha \in W$ and $\beta \in U$. As the dimension of W is ρ, the dimension of U is equal to $n - \rho$. Write $U = \{0, \beta_1, \dots, \beta_{2^{n-\rho}-1}\}$.

Proposition 1 $\Re \cap W = \{0\}$ *and* $\Re \cap (W \oplus \beta_j) \neq \phi$, *where* $W \oplus \beta_j = \{\alpha \oplus \beta_j | \alpha \in W\}$, $j = 1, \dots, 2^{n-\rho} - 1$.

Proof. $\Re \cap W = \{0\}$ follows from the fact that W is a subspace of $\{0\} \cup \Re^c$. Next we consider $\Re \cap (W \oplus \beta_j)$. Clearly,

$$V_n = W \cup (W \oplus \beta_1) \cup \cdots \cup (W \oplus \beta_{2^{n-\rho}-1}).$$

In addition,

$$W \cap (W \oplus \beta_j) = \phi$$

for $j = 1, \ldots, 2^{n-\rho} - 1$, and

$$(W \oplus \beta_j) \cap (W \oplus \beta_i) = \phi$$

for any $j \neq i$. Assume for contradiction that $\Re \cap (W \oplus \beta_{j_0}) = \phi$ for some j_0, $1 \leqq j_0 \leqq 2^{n-\rho} - 1$. Then we have $W \oplus \beta_{j_0} \subseteq \Re^c$. In this case $W \cup (W \oplus \beta_{j_0})$ must form a subspace of V_n. This contradicts the definition that W is a linear subspace with the maximum dimension in $\{0\} \cup \Re^c$. This completes the proof. □

The next corollary follows directly from the above proposition.

Corollary 1 *The size of* $\Re$ *satisfies* $|\Re| \geqq 2^{n-\rho}$ *and hence the rank of* $\Re$ *is at least* $n - \rho$*, where* ρ *is the maximum dimension a linear subspace in* $\{0\} \cup \Re^c$ *can achieve.*

5 Relating Nonlinearity to Propagation Characteristics

We proceed to the discussion of the nonlinearity of f. The main difficulty lies in finding a good approximation of $\langle \xi, \ell_i \rangle$ for each $i = 0, \ldots, 2^n - 1$, where ξ is the sequence of f and ξ_i is a row of H_n.

First we assume that

$$W = \{\gamma | \gamma = (a_1, \ldots, a_\rho, 0, \ldots, 0), a_i \in GF(2)\} \quad (5)$$

$$U = \{\gamma | \gamma = (0, \ldots, 0, a_{\rho+1}, \ldots, a_n), a_i \in GF(2)\} \quad (6)$$

where W is a linear subspace in $\{0\} \cup \Re^c$ that achieves the maximum dimension ρ and U is a complementary subspace of W. The more general case where (5) or (6) is not satisfied can be dealt with after employing a nonsingular transform on the input of f. This will be discussed in the later part of this section.

Recall that $\Re = \{0, \gamma_1, \ldots, \gamma_s\}$ and $\Delta(\alpha) = \langle \xi(0), \xi(\alpha) \rangle$, where $\xi(\alpha)$ is the sequence of $f(x \oplus \alpha)$. Since $\Delta(\gamma) \neq 0$ for each $\gamma \in \Re$ while $\Delta(\gamma) = 0$ for each $\gamma \in \Re^c = V_n - \Re$, (4) is specialized as

$$(\Delta(0), \Delta(\gamma_1), \ldots, \Delta(\gamma_s))Q = (\langle \xi, \ell_0 \rangle^2, \ldots, \langle \xi, \ell_{2^n-1} \rangle^2). \quad (7)$$

where ξ is the sequence of f, ℓ_i is the ith row of H_n and Q comprises the 0th, γ_1th, ..., γ_sth rows of H_n. Note that Q is an $(s+1) \times 2^n$ matrix.

Let ℓ be the γth row of H_n, where $\gamma \in \Re$. Note that γ can be uniquely expressed as $\gamma = \alpha \oplus \beta$, where $\alpha \in W$ and $\beta \in U$. Let ℓ' be the αth row of H_ρ and ℓ'' be the βth row of $H_{n-\rho}$. As $H_n = H_\rho \times H_{n-\rho}$, ℓ can be represented by $\ell = \ell' \times \ell''$, where $\times$ denotes the Kronecker product.

From the construction of $H_{n-\rho}$, we can see that the βth row of $H_{n-\rho}$ is an all-one sequence of length $2^{n-\rho}$ if $\beta = 0$, and a balanced $(1, -1)$-sequence of length $2^{n-\rho}$ if $\beta \neq 0$.

Recall that $\Re \cap W = \{0\}$ (see also Proposition 1). There are two cases associated with $\gamma = \alpha \oplus \beta \in \Re$: $\gamma = 0$ and $\gamma \neq 0$. In the first case, $\ell = \ell' \times \ell''$ is the all-one sequence of length 2^n, while in the second case, we have $\beta \neq 0$ which implies that ℓ'' is a balanced $(1, -1)$-sequence of length $2^{n-\rho}$ and hence $\ell = \ell' \times \ell''$ is a concatenation of 2^ρ balanced $(1, -1)$-sequences of length $2^{n-\rho}$.

Therefore we can write $Q = (Q_0, Q_1, \ldots, Q_{2^\rho-1})$, where each Q_i is a $(1, -1)$-matrix of order $(s+1) \times 2^{n-\rho}$. It is important to note that the top row of each Q_i is the all-one sequence, while the rest are balanced $(1, -1)$-sequences of length $2^{n-\rho}$.

With Q_0, we have

$$(\Delta(0), \Delta(\gamma_1), \ldots, \Delta(\gamma_s))Q_0 = (\langle\xi, \ell_0\rangle^2, \ldots, \langle\xi, \ell_{2^{n-\rho}-1}\rangle^2).$$

Let σ_0 be the all-one sequence of length $2^{n-\rho}$. Then

$$(\Delta(0), \Delta(\gamma_1), \ldots, \Delta(\gamma_s))Q_0\sigma_0^T = (\langle\xi, \ell_0\rangle^2, \ldots, \langle\xi, \ell_{2^{n-\rho}-1}\rangle^2)\sigma_0^T.$$

This causes

$$(\Delta(0), \Delta(\gamma_1), \ldots, \Delta(\gamma_s)) \begin{bmatrix} 2^{n-\rho} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{j=0}^{2^{n-\rho}-1} \langle\xi, \ell_j\rangle^2$$

and

$$\sum_{j=0}^{2^{n-\rho}-1} \langle\xi, \ell_j\rangle^2 = 2^{n-\rho}\Delta(0) = 2^{n-\rho+n} = 2^{2n-\rho}.$$

Similarly, with Q_i, $i = 1, \ldots, 2^\rho - 1$, we have

$$\sum_{j=0}^{2^{n-\rho}-1} \langle\xi, \ell_{j+i2^{n-\rho}}\rangle^2 = 2^{2n-\rho}.$$

Thus we have the following result:

Lemma 3 *Assume that f, a function on V_n, satisfy the propagation criterion with respect to all but a subset $\Re$ of vectors in V_n. Set $\Re^c = V_n - \Re$ and let W be a linear subspace with the maximum dimension ρ, in $\{0\} \cup \Re^c$, and U be a complementary subspace of W. Assume that W and U satisfy (5) and (6) respectively. Then*

$$\sum_{j=0}^{2^{n-\rho}-1} \langle\xi, \ell_{j+i2^{n-\rho}}\rangle^2 = 2^{2n-\rho}$$

for all $i = 0, 1, \ldots, 2^\rho - 1$, where ξ is the sequence of f and each ℓ_k is a row of H_n.

Lemma 3 can be viewed as a refinement of Parseval's equation $\sum_{j=0}^{2^n-1}\langle\xi, \ell_j\rangle^2 = 2^{2n}$. It implies that $|\langle\xi, \ell_j\rangle| \leqq 2^{n-\frac{1}{2}\rho}$ for all $j = 0, \ldots, 2^n - 1$. Therefore by Lemma 1 we have $N_f \geqq 2^{n-1} - 2^{n-\frac{1}{2}\rho-1}$.

So far we have assumed that W and U satisfy (5) and (6) respectively. When it is not the case, we can always find a nonsingular $n \times n$ matrix A whose entries are from $GF(2)$ such that the subspaces W' and U' associated with $f'(x) = f(xA)$ have the required forms. f' and f have the same algebraic degree and nonlinearity (see Lemma 10 of [18]). This shows that the following theorem is true.

Theorem 1 *For any function on V_n, the nonlinearity of f satisfies $N_f \geqq 2^{n-1} - 2^{n-\frac{1}{2}\rho-1}$, where ρ is the maximum dimension of the linear subspaces in $\{0\} \cup \Re^c$.*

Theorem 1 indicates that the nonlinearity of a function is determined by the maximum dimension that a linear subspaces in $\{0\} \cup \Re^c$ can achieve, but not by the size of $\Re^c$.

In [22], we have proved that $N_f \geqq 2^{n-1} - 2^{\frac{1}{2}(n+t)-1}$, where t is the rank of $\Re$. By Corollary 1, we have $t \geqq n - \rho$. This implies that $2^{n-1} - 2^{n-\frac{1}{2}\rho-1} \geqq 2^{n-1} - 2^{\frac{1}{2}(n+t)-1}$. Thus Theorem 1 is an improvement to the result in [22]. This improvement can be demonstrated by a concrete example. In [22] a function f_5 on

V_5 is constructed that satisfies the propagation criterion with respect to all but the following fives vectors in V_5:

$$\Re = \{(0,0,0,0,0),(0,0,0,0,1),(0,0,0,1,0),(0,0,1,0,0),(0,0,1,1,1)\}.$$

The rank t of $\Re$ is equal to 3. By using the result of [22], $N_{f_5} \geqq 2^{5-1} - 2^{\frac{1}{2}(5+3)-1} = 2^4 - 2^3 = 8$. On the other hand, we can set $W = \{(a_1,a_2,a_3,a_4,a_5)|a_i \in GF(2), a_1 \oplus a_2 \oplus a_3 = 0\}$. W is a four-dimensional subspace in $\{0\} \cup \Re^c$. Using Theorem 1 with $\rho = 4$, we have $N_{f_5} \geqq 2^{5-1} - 2^{5-\frac{1}{2}\rho-1} = 2^4 - 2^2 = 12 > 8$. According to [4], 12 is the maximum nonlinearity a function on V_5 can achieve.

6 A Tight Lower Bound on Nonlinearity of Functions with Propagation Characteristics

By Theorem 1, $N_f \geqq 2^{n-1} - 2^{n-\frac{3}{2}}$ if f, a function on V_n, satisfies the propagation criterion with respect to some vectors. This section shows that this lower bound can be significantly improved. Indeed we prove that $N_f \geqq 2^{n-2}$ and also show that it is tight.

Theorem 2 *If f, a function on V_n, satisfies the propagation criterion with respect to one or more vectors in V_n, then the nonlinearity of f satisfies $N_f \geqq 2^{n-2}$.*

Proof. As in the previous sections, we denote by $\Re$ the set of vectors in V_n with respect to which the propagation criterion is not satisfied by f. We also let $\Re^c = V_n - \Re$, and W be a linear subspace in $\{0\} \cup \Re^c$ that achieves the maximum dimension ρ.

By Theorem 1, the theorem is trivially true when $\rho > 1$. Next we consider the case when $\rho = 1$. We prove this part by further refining the Parseval's equation.

As in the proof of Lemma 3, without loss of generality, we can assume that

$$\begin{aligned} W &= \{\gamma|\gamma = (a_1, 0, \ldots, 0), a_1 \in GF(2)\} && (8)\\ U &= \{\gamma|\gamma = (0, a_2, \ldots, a_n), a_i \in GF(2)\} && (9) \end{aligned}$$

Similarly to Lemma 3, we have

$$\sum_{j=0}^{2^{n-1}-1} \langle \xi, \ell_{j+i2^{2^{n-1}}} \rangle^2 = 2^{2n-1},\ i = 0, 1, \qquad (10)$$

where ξ is the sequence of f and ℓ_k is a row of H_n.

Compare the first row of (2), we have

$$(a_0, a_1, \ldots, a_{2^n-1}) = 2^{-n}(\langle \xi, \ell_0 \rangle, \cdots, \langle \xi, \ell_{2^n-1} \rangle)H_n$$

or equivalently,

$$2^n(a_0, a_1, \ldots, a_{2^n-1}) = (\langle \xi, \ell_0 \rangle, \cdots, \langle \xi, \ell_{2^n-1} \rangle)H_n \qquad (11)$$

where each $a_j = \pm 1$ and $(a_0, a_1, \ldots, a_{2^n-1})$ is the first row of the matrix M described in (2).

Rewrite ℓ_i, the ith row of H_n, as $\ell(\alpha_i)$, where α_i is the binary representation of an integer i in the ascending alphabetical order. Set

$$N = (\langle \xi, \ell(\alpha_i \oplus \alpha_j) \rangle), 0 \leqq i, j \leqq 2^n - 1.$$

N is a symmetric matrix of order 2^n with integer entries. In [17], Rothaus has shown that $NN = NN^T = 2^{2n}I_{2^n}$. We can split N into four submatrices of equal size, namely

$$N = \begin{bmatrix} N_1 & N_2 \\ N_2 & N_1 \end{bmatrix}$$

where each N_j is a matrix of order 2^{n-1}. As $NN = 2^{2n}I_{2^n}$, we have $N_1N_2 = 0$.

Let $(c(\alpha_0), c(\alpha_1), \ldots, c(\alpha_{2^{n-1}-1}))$ be an arbitrary linear sequence of length 2^{n-1}. Then

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

is a linear sequence of length 2^n, and hence a row of H_n. Thus from (11), we have

$$\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle + \sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j \oplus 2^{n-1})\rangle = \pm 2^n.$$

Hence

$$(\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle + \sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j \oplus \alpha_{2^{n-1}})\rangle)^2 = 2^{2n}. \tag{12}$$

Rewrite the left hand side of (12) as

$$\begin{aligned} & (\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle)^2 + (\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j \oplus \alpha_{2^{n-1}})\rangle)^2 \\ + \quad & 2(\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle)(\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j \oplus \alpha_{2^{n-1}})\rangle) \end{aligned}$$

where

$$\begin{aligned} & (\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle)(\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j \oplus \alpha_{2^{n-1}})\rangle) \\ = \quad & \sum_{t=0}^{2^{n-1}-1} \sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle c(\alpha_j \oplus \alpha_t)\langle \xi, \ell(\alpha_j \oplus \alpha_t \oplus \alpha_{2^{n-1}})\rangle. \end{aligned} \tag{13}$$

As $(c(\alpha_0), c(\alpha_1), \ldots, c(\alpha_{2^{n-1}-1}))$ is a linear sequence, $c(\alpha_j)c(\alpha_j \oplus \alpha_t) = c(\alpha_t)$. Hence (13) can be written as

$$\sum_{t=0}^{2^{n-1}-1} c(\alpha_t) \sum_{j=0}^{2^{n-1}-1} \langle \xi, \ell(\alpha_j)\rangle\langle \xi, \ell(\alpha_j \oplus \alpha_t \oplus \alpha_{2^{n-1}})\rangle.$$

Since $N_1N_2 = 0$,

$$\sum_{j=0}^{2^{n-1}-1} \langle \xi, \ell(\alpha_j)\rangle\langle \xi, \ell(\alpha_j \oplus \alpha_t \oplus \alpha_{2^{n-1}})\rangle = 0.$$

This proves that (13) is equal to zero and hence

$$(\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle)^2 + (\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j \oplus \alpha_{2^{n-1}})\rangle)^2 = 2^{2n}.$$

By Lemma 2,

$$\sum_{j=0}^{2^{n-1}-1} c(\alpha_j)\langle \xi, \ell(\alpha_j)\rangle = 0 \text{ or } \pm 2^n. \tag{14}$$

Since $(c(\alpha_0), c(\alpha_1), \ldots, c(\alpha_{2^{n-1}-1}))$ is an arbitrary linear sequence of length 2^{n-1} and each linear sequence of length 2^{n-1} is a column of H_{n-1}, from (14) we have

$$(\langle \xi, \ell_0\rangle, \ldots, \langle \xi, \ell_{2^n-1}\rangle)H_{n-1} = 2^n(b_0, \ldots, b_{2^{n-1}-1}) \tag{15}$$

where $b_j = 0$ or ± 1. Therefore

$$(\langle \xi, \ell_0\rangle, \ldots, \langle \xi, \ell_{2^n-1}\rangle)2^{\frac{1}{2}(n-1)}H_{n-1} = 2^{\frac{1}{2}(n+1)}(b_0, \ldots, b_{2^{n-1}-1}).$$

Recall that a matrix A of order s is said to be orthogonal if $AA^T = I_s$. It is easy to verify that $2^{\frac{1}{2}(n-1)}H_{n-1}$ is an orthogonal matrix. Thus

$$\sum_{j=0}^{2^n-1} \langle \xi, \ell_{\alpha_j}\rangle^2 = 2^{n+1}\sum_{j=0}^{2^{n-1}-1} b_j^2.$$

On the other hand, by (10) we have

$$\sum_{j=0}^{2^n-1} \langle \xi, \ell_{\alpha_j}\rangle^2 = 2^{2n-1}.$$

Hence

$$\sum_{j=0}^{2^{n-1}-1} b_j^2 = \sum_{j=0}^{2^{n-1}-1} |b_j| = 2^{n-2}.$$

Now let $\sigma(\alpha_i)$ denote the ith row of H_{n-1}, where $\alpha_i \in V_{n-1}$ is the binary representation of i, $i = 0, 1, \ldots, 2^{n-1}-1$. From (15),

$$(\langle \xi, \ell_0\rangle, \cdots, \langle \xi, \ell_{2^n-1}\rangle)H_{n-1}\sigma(\alpha_i)^T = 2^n(b_0, \ldots, b_{2^{n-1}-1})\sigma(\alpha_i)^T. \tag{16}$$

Note that

$$\langle \sigma(\alpha_i), \sigma(\alpha_j)\rangle = \begin{cases} 2^{n-1} & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

Thus

$$H_{n-1}\sigma(\alpha_i)^T = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 2^{n-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{17}$$

where 2^{n-1} is on the ith position of the column vector.

Write $\sigma(\alpha_i) = (d_0, d_1, \ldots, d_{2^{n-1}-1})$. Then

$$(b_0, \ldots, b_{2^{n-1}-1})\sigma(\alpha_i)^T = \sum_{j=0}^{2^{n-1}-1} d_j b_j.$$

As $d_j = \pm 1$, we have

$$|\sum_{j=0}^{2^{n-1}-1} d_j b_j| \leqq \sum_{j=0}^{2^{n-1}-1} |b_j| = 2^{n-2}. \quad (18)$$

From (16), (17) and (18)

$$2^{n-1}|\langle \xi, \ell_i \rangle| \leqq 2^n \sum_{j=0}^{2^{n-1}-1} |b_j| = 2^{2n-2}$$

and hence

$$|\langle \xi, \ell_i \rangle| \leqq 2^{n-1}$$

where i is an arbitrary integer in $[0, \ldots, 2^{n-1}-1]$. Similarly,

$$|\langle \xi, \ell_i \rangle| \leqq 2^{n-1}$$

holds for all $i = 2^{n-1}, 2^{n-1}+1, \ldots, 2^n - 1$. By Lemma 1, the nonlinearity of f satisfies

$$N_f \geqq 2^{n-1} - 2^{n-2} = 2^{n-2}.$$

This completes the proof. □

As an immediate consequence, we have

Corollary 2 *Let f be a function on V_n. Then the following statements hold:*

1. *if the nonlinearity of f satisfies $N_f < 2^{n-2}$, then f does not satisfy the propagation criterion with respect to any vector in V_n.*
2. *if f satisfies the SAC, then the nonlinearity of f satisfies $N_f \geqq 2^{n-2}$.*

Finally we show that the lower bound 2^{n-2} is tight. We achieve the goal by demonstrating a function on V_n whose nonlinearity is equal to 2^{n-2}. Let $g(x_1, x_2) = x_1 x_2$ be a function on V_2. Then the nonlinearity of g is $N_g = 1$. Now let $f(x_1, \ldots, x_n) = x_1 x_2$ be a function on V_n. Then the nonlinearity of f is $N_f = 2^{n-2} N_g = 2^{n-2}$ (see for instance Lemma 8 of [19]). f satisfies the propagation criterion with respect to all vectors in V_n whose first two bits are nonzero, which count for three quarters of the vectors in V_n. It is not hard to verify that

$$\{(0,0,0,\ldots,0), (1,0,0,\ldots,0), (0,1,0,\ldots,0), (1,1,0,\ldots,0)\}$$

is the linear subspace that achieves the maximum dimension $\rho = 2$.

Thus we have a result described as follows:

Lemma 4 *The lower bound 2^{n-2} as stated in Theorem 2 is tight.*

7 Conclusion

We have shown quantitative relationships between nonlinearity, propagation characteristics and the SAC. A tight lower bound on the nonlinearity of a function with propagation characteristics is also presented.

This research has also introduced a number of interesting problems yet to be resolved. One of the problems is regarding the size and distribution of $\Re^c$, the set of vectors where the propagation criterion is satisfied by a function on V_n. For all the functions we know of, $\Re^c$ is either an empty set or a set with at least 2^{n-1} vectors. We believe that any further understanding of this problem will contribute to the research into the design and analysis of cryptographically strong nonlinear functions.

Acknowledgments

The first author was supported in part by ARC A49130102, A49131885 and A49232172, and ATERB C010/058, the second author by ARC A49130102 and ATERB C010/058, and the third author by ARC A49232172 and ATERB N069/412. All authors were supported by a University of Wollongong Research Program grant.

References

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

[2] K. G. Beauchamp. *Applications of Walsh and Related Functions with an Introduction to Sequency Functions.* Microelectronics and Signal Processing. Academic Press, London, New York, Tokyo, 1984.

[3] L. Brown, M. Kwan, J. Pieprzyk, and J. Seberry. Improving resistance to differential cryptanalysis and the redesign of LOKI. In H. Imai, R. Rivest, and T. Matsumoto, editors, *Advances in Cryptology - ASIACRYPT'91*, volume 739 of *Lecture Notes in Computer Science*, pages 36–50, Berlin, New York, Tokyo, 1993. Springer-Verlag.

[4] G. D. Cohen, M. G. Karpovsky, Jr. H. F. Mattson, and J. R. Schatz. Covering radius — survey and recent results. *IEEE Transactions on Information Theory*, IT-31(3):328–343, 1985.

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

[6] J.-H. Evertse. Linear structures in blockciphers. In *Advances in Cryptology - EUROCRYPT'87*, volume 304 of *Lecture Notes in Computer Science*, pages 249–266. Springer-Verlag, Berlin, Heidelberg, New York, 1988.

[7] X. Lai. *On the Design and Security of Block Ciphers.* ETH Series in Information Processing. Hartung-Gorre Verlag Konstanz, Zürich, 1992.

[8] X. Lai and J. L. Massey ans S. Murphy. Markov ciphers and differential cryptanalysis. In D. W. Davies, editor, *Advances in Cryptology - EUROCRYPT'91*, volume 547 of *Lecture Notes in Computer Science*, pages 17–38, Berlin, New York, Tokyo, 1991. Springer-Verlag.

[9] X. Lai and J. L. Massey. A proposal for a new block encryption standard. In I. B. Damgård, editor, *Advances in Cryptology - EUROCRYPT'90*, volume 473 of *Lecture Notes in Computer Science*, pages 389–404, Berlin, New York, Tokyo, 1991. Springer-Verlag.

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

[11] W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. In *Advances in Cryptology - EUROCRYPT'89*, volume 434 of *Lecture Notes in Computer Science*, pages 549–562. Springer-Verlag, Berlin, Heidelberg, New York, 1990.

[12] S. Miyaguchi. The FEAL cipher family. In *Advances in Cryptology - CRYPTO'90*, volume 537 of *Lecture Notes in Computer Science*, pages 627–638, Berlin, New York, Tokyo, 1991. Springer-Verlag.

[13] National Institute of Standards and Technology. Secure hash standard. Federal Information Processing Standards Publication FIPS PUB 180-1, U.S. Department of Commerce, April 1995.

[14] B. Preneel, R. Govaerts, and J. Vandewalle. Boolean functions satisfying higher order propagation criteria. In *Advances in Cryptology - EUROCRYPT'91*, volume 547 of *Lecture Notes in Computer Science*, pages 141–152. Springer-Verlag, Berlin, Heidelberg, New York, 1991.

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

[16] R. Rivest. The MD5 message digest algorithm. Request for Comments RFC 1321, IETF, 1992.

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

[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 of *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 of *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] J. Seberry, X. M. Zhang, and Y. Zheng. Relationships among nonlinearity criteria. In *Advances in Cryptology - EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 376–388. Springer-Verlag, Berlin, Heidelberg, New York, 1995.

[22] J. Seberry, X. M. Zhang, and Y. Zheng. The structure of cryptographic function with strong avalanche characteristics. In *Advances in Cryptology - ASIACRYPT'94*, volume 917 of *Lecture Notes in Computer Science*, pages 119–132. Springer-Verlag, Berlin, Heidelberg, New York, 1995.

[23] National Bureau Standards. Data encryption standard. *Federal Information Processing Standards Publication FIPS PUB 46, U.S. Department of Commerce*, 1977.

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

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

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

[27] Y. Zheng, J. Pieprzyk, and J. Seberry. HAVAL - a one-way hashing algorithm with varialbe length of output. In J. Seberry and Y. Zheng, editors, *Advances in Cryptology - AUSCRYPT'92*, volume 718 of *Lecture Notes in Computer Science*, pages 83–104, Berlin, New York, Tokyo, 1993. Springer-Verlag.