Improved Upper Bound on the Nonlinearity of High Order Correlation Immune Functions

Yuliang Zheng[1] and Xian-Mo Zhang[2]

[1] Monash University, Frankston, Melbourne, VIC 3199, Australia
yuliang.zheng@monash.edu.au, http://www.netcomp.monash.edu.au/links/
[2] The University of Wollongong, Wollongong, NSW 2522, Australia
xianmo@cs.uow.edu.au

Abstract. It has recently been shown that when $m > \frac{1}{2}n - 1$, the nonlinearity N_f of an mth-order correlation immune function f with n variables satisfies the condition of $N_f \leq 2^{n-1} - 2^m$, and that when $m > \frac{1}{2}n - 2$ and f is a balanced function, the nonlinearity satisfies $N_f \leq 2^{n-1} - 2^{m+1}$. In this work we prove that the general inequality, namely $N_f \leq 2^{n-1} - 2^m$, can be improved to $N_f \leq 2^{n-1} - 2^{m+1}$ for $m \geq 0.6n - 0.4$, regardless of the balance of the function. We also show that correlation immune functions achieving the maximum nonlinearity for these functions have close relationships with plateaued functions. The latter have a number of cryptographically desirable properties.

Key words: Correlation Immune Functions, Nonlinearity, Resilient Functions, Plateaued Functions, Stream Ciphers

1 Introduction

Correlation immunity has long been recognized as one of the critical indicators of nonlinear combining functions of shift registers in stream generators (see [12]). A high correlation immunity is generally a very desirable property, in view of various successful correlation attacks against a number of stream ciphers (see for instance [6]).

Another class of cryptanalytic attacks against stream ciphers, called best approximation attacks, were advocated in [4]. Success of these attacks in breaking a stream cipher is made possible by exploiting the low nonlinearity of functions employed by the cipher, and it highlights the significance of nonlinearity in the analysis and design of encryption algorithms.

Recently Sarkar and Maitra [10] have proved that when $m > \frac{1}{2}n - 1$, the nonlinearity N_f of an mth-order correlation immune function f with n variables satisfies the condition of $N_f \leq 2^{n-1} - 2^m$. In addition they have shown that if f is balanced and $m > \frac{1}{2}n - 2$, then the condition becomes $N_f \leq 2^{n-1} - 2^{m+1}$. (See also Section 8 for independent efforts by researchers other than Sarkar and Maitra.)

In this work we focus our attention on the case of $m \geq 0.6n - 0.4$. We show that for such m and n, the nonlinearity of an mth-order correlation immune

D.R. Stinson and S. Tavares (Eds.): SAC 2000, LNCS 2012, pp. 262–274, 2001.

function f with n variables must satisfy the condition of $N_f \leq 2^{n-1} - 2^{m+1}$, regardless of the balance of the function. This represents an improvement on the upper bound of $N_f \leq 2^{n-1} - 2^m$.

Plateaued functions are a new class of functions recently introduced in [16]. These functions have a number of properties that are deemed desirable in cryptography. We show that, interestingly, a correlation immune function with the maximum nonlinearity achievable by such a function can be identified with a plateaued function. This provides a new avenue for the analysis and design of cryptographically useful correlation immune functions.

The remaining part of this paper is organized as follows: Section 2 introduces basic definitions on Boolean functions, and Section 3 summarizes some of the important cryptographic criteria for Boolean functions. This will be followed by Section 4 where relevant properties of plateaued functions are discussed. Some useful results on correlation immune functions are introduced in Section 5. These results will then be used in Section 6 where our improved upper bound on the nonlinearity of correlation immune functions is proved. In the same section some relationships between correlation immune functions and plateaued functions are also examined. In Section 7, the new upper bound is demonstrated to be tight for balanced correlation immune functions. Finally the paper is closed by Section 8 where possible directions for future research are pointed out.

2 Boolean Functions

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

Given two sequences $\tilde{a} = (a_1, \cdots, a_m)$ and $\tilde{b} = (b_1, \cdots, b_m)$, their *component-wise product* is defined by $\tilde{a} * \tilde{b} = (a_1 b_1, \cdots, a_m b_m)$. In particular, if $m = 2^n$ and $\tilde{a}$, $\tilde{b}$ are the sequences of functions f and g on V_n respectively, then $\tilde{a} * \tilde{b}$ is the sequence of $f \oplus g$ where $\oplus$ denotes the addition in $GF(2)$.

Let $\tilde{a} = (a_1, \cdots, a_m)$ and $\tilde{b} = (b_1, \cdots, b_m)$ be two sequences or vectors, the *scalar product* of $\tilde{a}$ and $\tilde{b}$, denoted by $\langle \tilde{a}, \tilde{b} \rangle$, is defined as the sum of the component-wise multiplications. In particular, when $\tilde{a}$ and $\tilde{b}$ are from V_m, $\langle \tilde{a}, \tilde{b} \rangle = a_1 b_1 \oplus \cdots \oplus a_m b_m$, where the addition and multiplication are over $GF(2)$, and when $\tilde{a}$ and $\tilde{b}$ are $(1,-1)$-sequences, $\langle \tilde{a}, \tilde{b} \rangle = \sum_{i=1}^{m} a_i b_i$, where the addition and multiplication are over the reals.

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

A $(1, -1)$-matrix N of order n is called a *Hadamard* matrix if $NN^T = nI_n$, where N^T is the transpose of N and I_n is the identity matrix of order n. A Sylvester-Hadamard matrix of order 2^n, denoted by H_n, is generated by the following recursive relation

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

Obviously H_n is symmetric. Let ℓ_i, $0 \le i \le 2^n - 1$, be the i row of H_n. It is known that 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.

The *Hamming weight* of a $(0, 1)$-sequence ξ, denoted by $HW(\xi)$, 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 $x = (x_1, \ldots, x_n)$.

3 Cryptographic Criteria of Boolean Functions

The following criteria for cryptographic Boolean functions are often considered: balance, nonlinearity, propagation criterion, correlation immunity, algebraic degree and non-zero linear structures. In this paper we focus mainly on nonlinearity and correlation immunity.

The so called Parseval's equation (Page 416 [7]) is a useful tool in this work: Let f be a function on V_n and ξ denote the sequence of f. Then $\sum_{i=0}^{2^n-1} \langle \xi, \ell_i \rangle^2 = 2^{2n}$ where ℓ_i is the ith row of H_n, $i = 0, 1, \ldots, 2^n - 1$.

The *nonlinearity* of a function f on V_n, denoted by N_f, is the minimal Hamming distance between f and all affine functions on V_n, i.e.,
$N_f = \min_{i=1,2,\ldots,2^{n+1}} d(f, \psi_i)$ where ψ_1, ψ_2, …, $\psi_{2^{n+1}}$ are all the affine functions on V_n. High nonlinearity can be used to resist a linear attack. The following characterization of nonlinearity will be useful (for a proof see for instance [8]).

Lemma 1. *The nonlinearity of f on V_n can be expressed by*

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

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

From Lemma 1 and Parseval's equation, it is easy to verify that $N_f \le 2^{n-1} - 2^{\frac{1}{2}n-1}$ for any function f on V_n. If $N_f = 2^{n-1} - 2^{\frac{1}{2}n-1}$, then f is called a *bent function* [9]. It is known that a bent function on V_n exists only when n is even.

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 $\Delta_f(\alpha) = \langle \xi(0), \xi(\alpha) \rangle$, the scalar product of $\xi(0)$ and $\xi(\alpha)$. $\Delta(\alpha)$ is called the auto-correlation of f with a shift α. We omit the subscript of

$\Delta_f(\alpha)$ if no confusion occurs. 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. In the case that f does not satisfy the propagation criterion with respect to a vector α, it may be desirable for $f(x) \oplus f(x \oplus \alpha)$ to be almost balanced. That is, one may require $|\Delta_f(\alpha)|$ to be a small value.

The concept of correlation immune functions was introduced by Siegenthaler [12]. Xiao and Massey gave an equivalent definition [2,5]: A function f on V_n is called a *mth-order correlation immune function* if

$$\sum_{x \in V_n} f(x)(-1)^{\langle \beta, x \rangle} = 0$$

for all $\beta \in V_n$ with $1 \le HW(\beta) \le m$, where in the the sum, $f(x)$ and $\langle \beta, x \rangle$ are regarded as real-valued functions. From the first equality in Section 4.2 of [2], a correlation immune function can also be equivalently restated as follows: Let f be a function on V_n and let ξ be its sequence. Then f is called a *mth-order correlation immune function* if $\langle \xi, \ell \rangle = 0$ for every ℓ, where ℓ is the sequence of a linear function $\varphi(x) = \langle \alpha, x \rangle$ on V_n constrained by $1 \le HW(\alpha) \le m$. In fact, $\langle \xi, \ell_i \rangle = 0$, where ℓ_i is the ith row of H_n, if and only if $f(x) \oplus \langle \alpha_i, x \rangle$ is balanced, where α_i is the binary representation of an integer i, $0 \le i \le 2^n - 1$. Correlation immune functions are used in the design of running-key generators in stream ciphers to resist a correlation attack and the design of hash functions. Relevant discussions on correlation immune functions, more generally on resilient functions, can be found in [15].

Let f be a function on V_n and ξ denote the sequence of f. We introduce two new notations:

1. Set $\Im_f = \{i \mid \langle \xi, \ell_i \rangle \neq 0,\ 0 \le i \le 2^n - 1\}$ where ℓ_i is the ith row of H_n,
2. set $\Im_f^* = \{\alpha_i \mid \langle \xi, \ell_{\alpha_i} \rangle \neq 0,\ 0 \le i \le 2^n - 1\}$ where α_i is the binary representation of an integer i, $0 \le i \le 2^n - 1$ and ℓ_{α_i} is identified with ℓ_i.

$\Im_f^*$ is essentially the same as $\Im_f$ with the only difference being that its elements are represented by a binary vector in V_n. We will simply write $\Im_f$ as $\Im$ and $\Im_f^*$ as $\Im^*$ when no confusion arises. It is easy to verify that $\#\Im_f$ and $\#\Im_f^*$ are invariant under any nonsingular linear transformation on the variables of the function f. $\#\Im_f$ ($\#\Im_f^*$) together with the distribution of $\Im_f$ ($\Im_f^*$) determines the correlation immunity and other cryptographic properties of a function.

4 An Overview of Plateaued Functions

The concept of plateaued functions was introduced in [16].

Definition 1. *Let f be a function on V_n and ξ denote the sequence of f. If there exists an even number r, $0 \le r \le n$, such that $\#\Im = 2^r$ and each $\langle \xi, \ell_j \rangle^2$ takes the value of 2^{2n-r} or 0 only, where ℓ_j denotes the jth row of H_n, $j = 0, 1, \ldots, 2^n - 1$, then f is called a* rth-order plateaued function *on V_n. f is also simply called a* plateaued function *on V_n if we ignore the particular order r.*

Due to Parseval's equation, the condition that $\#\Im = 2^r$ can be obtained from the condition that "each $\langle \xi, \ell_j \rangle^2$ takes the value of 2^{2n-r} or 0 only, where ℓ_j denotes the jth row of H_n, $j = 0, 1, \ldots, 2^n - 1$". For the sake of convenience, however, we have mentioned both conditions in the definition of plateaued functions.

Some facts about plateaued functions follow: (1) if f is a rth-order plateaued function, then r must be even, (2) f is an nth-order plateaued function if and only if f is bent, (3) f is a 0th-order plateaued function if and only if f is affine.

All the following results can be found in [16].

Theorem 1. *Let f be a function on V_n and ξ denote the sequence of f. Set $p_M = \max\{|\langle \xi, \ell_j \rangle|, j = 0, 1, \ldots, 2^n - 1\}$, where ℓ_j is the jth row of H_n. Then the following statements are equivalent: (i) f is a plateaued function on V_n, (ii) $\sum_{j=0}^{2^n-1} \Delta^2(\alpha_j) = \frac{2^{3n}}{\#\Im}$, (iii) the nonlinearity N_f of f satisfies $N_f = 2^{n-1} - \frac{2^{n-1}}{\sqrt{\#\Im}}$, (iv) $p_M \sqrt{\#\Im} = 2^n$, (v) $N_f = 2^{n-1} - 2^{-\frac{n}{2}-1} \sqrt{\sum_{j=0}^{2^n-1} \Delta^2(\alpha_j)}$.*

Theorem 2. *Let f be a function on V_n and ξ denote the sequence of f. Then*

$$\sum_{j=0}^{2^n-1} \Delta^2(\alpha_j) \geq \frac{2^{3n}}{\#\Im}$$

where the equality holds if and only if f is a plateaued function.

Theorem 3. *Let f be a function on V_n and ξ denote the sequence of f. Then the nonlinearity N_f of f satisfies $N_f \leq 2^{n-1} - \frac{2^{n-1}}{\sqrt{\#\Im}}$, where the equality holds if and only if f is a plateaued function.*

Theorem 4. *Let f be a function on V_n and ξ denote the sequence of f. Then the nonlinearity N_f of f satisfies*

$$N_f \leq 2^{n-1} - 2^{-\frac{n}{2}-1} \sqrt{\sum_{j=0}^{2^n-1} \Delta^2(\alpha_j)}$$

where the equality holds if and only if f is a plateaued function on V_n.

Proposition 1. *Let f be a rth-order plateaued function on V_n. Then the nonlinearity N_f of f satisfies $N_f = 2^{n-1} - 2^{n-\frac{r}{2}-1}$.*

5 Some Useful Results on Correlation Immune Functions

Consider a function f on V_n. Denote by $\xi = (a_0, a_1, \ldots, a_{2^n-1})$, where $a_j = \pm 1$, the sequence of f. Obviously

$$(a_0, a_1, \ldots, a_{2^n-1}) H_n = (\langle \xi, \ell_0 \rangle, \langle \xi, \ell_1 \rangle, \ldots, \langle \xi, \ell_{2^n-1} \rangle) \tag{1}$$

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

Let p be an integer with $1 \le p \le n-1$. Rewrite (1) as

$$(a_0, a_1, \ldots, a_{2^n-1})(H_p \times H_{n-p}) = (\langle \xi, \ell_0 \rangle, \langle \xi, \ell_1 \rangle, \ldots, \langle \xi, \ell_{2^n-1} \rangle) \tag{2}$$

where $\times$ is the *Kronecker* Product [14].

Let e_i denote the ith row of H_{n-p}, $i = 0, 1, \ldots, 2^{n-p} - 1$. For any fixed j with $0 \le j \le 2^{n-p}-1$, comparing the jth, the $(j+2^{n-p})$th, ..., the $(j+(2^p-1)2^{n-p})$th terms in the two sides of (2), we have

$$(a_0, a_1, \ldots, a_{2^n-1})(H_p \times e_j^T) = (\langle \xi, \ell_j \rangle, \langle \xi, \ell_{j+2^{n-p}} \rangle, \ldots, \langle \xi, \ell_{j+(2^p-1)2^{n-p}} \rangle)$$

Write $\xi = (\xi_0, \xi_1, \ldots, \xi_{2^p-1})$ where each ξ_i is of length 2^{n-p}. Then we have

$$(\langle \xi_0, e_j \rangle, \langle \xi_1, e_j \rangle, \ldots, \langle \xi_{2^p-1}, e_j \rangle) H_p = (\langle \xi, \ell_j \rangle, \langle \xi, \ell_{j+2^{n-p}} \rangle, \ldots, \langle \xi, \ell_{j+(2^p-1)2^{n-p}} \rangle)$$

Hence

$$\begin{aligned} &2^p(\langle \xi_0, e_j \rangle, \langle \xi_1, e_j \rangle, \ldots, \langle \xi_{2^p-1}, e_j \rangle) \\ &= (\langle \xi, \ell_j \rangle, \langle \xi, \ell_{j+2^{n-p}} \rangle, \ldots, \langle \xi, \ell_{j+(2^p-1)2^{n-p}} \rangle) H_p \end{aligned} \tag{3}$$

Based on these discussions, we have the following lemma.

Lemma 2. *Let f be an mth-order correlation immune function on V_n, where $m \le n-2$, and ξ be the sequence of f. Then $\langle \xi, \ell_{2^{m+1}-1} \rangle \equiv 0 \pmod{2^{m+2}}$ if and only if $\langle \xi, \ell_0 \rangle \equiv 0 \pmod{2^{m+2}}$ where ℓ_0 is the top row of H_n.*

Proof. Set $W = \{\alpha_0, \alpha_{2^{n-m-1}}, \alpha_{2 \cdot 2^{n-m-1}}, \ldots, \alpha_{(2^{m+1}-1)2^{n-m-1}}\}$, where each α_j is the binary representation of an integer j. Note that W is an $(m+1)$-dimensional linear subspace of V_n.

Write $\xi = (\xi_0, \xi_1, \xi_2, \ldots . \xi_{2^{m+1}-1})$, where each ξ_i is of length 2^{n-m-1}. Let $p = m+1$ and $j = 0$ in (3), we have

$$\begin{aligned} &2^{m+1}(\langle \xi_0, e_0 \rangle, \langle \xi_1, e_0 \rangle, \ldots, \langle \xi_{2^{m+1}-1}, e_0 \rangle) \\ &= (\langle \xi, \ell_0 \rangle, \langle \xi, \ell_{2^{n-m-1}} \rangle, \langle \xi, \ell_{2 \cdot 2^{n-m-1}} \rangle, \ldots, \langle \xi, \ell_{(2^{m+1}-1)2^{n-m-1}} \rangle) H_{m+1} \end{aligned} \tag{4}$$

where e_0 denotes the 0th row of H_{n-m-1}, i.e., the all-one sequence of length 2^{n-m-1}.

As $HW(\alpha_{j \cdot 2^{n-m-1}}) \le m$, we have $\langle \xi, \ell_{j \cdot 2^{n-m-1}} \rangle = 0$, where $j = 1, \ldots, 2^{m+1} - 2$. Therefore (4) can be rewritten as

$$\begin{aligned} &2^{m+1}(\langle \xi_0, e_0 \rangle, \langle \xi_1, e_0 \rangle, \ldots, \langle \xi_{2^{m+1}-1}, e_0 \rangle) \\ &= (\langle \xi, \ell_0 \rangle, 0, \ldots, 0, \langle \xi, \ell_{(2^{m+1}-1)2^{n-m-1}} \rangle) H_{m+1} \end{aligned} \tag{5}$$

Comparing the rightmost term in the two sides of (5), we have

$$2^{m+1} \langle \xi_{2^{m+1}-1}, e_0 \rangle = \langle \xi, \ell_0 \rangle - \langle \xi, \ell_{(2^{m+1}-1)2^{n-m-1}} \rangle \tag{6}$$

Note that the length of $\xi_{2^{m+1}-1}$ and e_0 is even. Hence $\langle \xi_{2^{m+1}-1}, e_0 \rangle$ must be even. From this it follows that $2^{m+1} \langle \xi_{2^{m+1}-1}, e_0 \rangle \equiv 0 \pmod{2^{m+2}}$. Finally, by considering (6), we have proved that $\langle \xi, \ell_{(2^{m+1}-1)2^{n-m-1}} \rangle \equiv 0 \pmod{2^{m+2}}$ if and only if $\langle \xi, \ell_0 \rangle \equiv 0 \pmod{2^{m+2}}$. □

By choosing a different W in the proof of Lemma 2, we can prove the following lemma in a similar way.

Lemma 3. *Let f be an mth-order correlation immune function on V_n, where $m \leq n-2$, and ξ be the sequence of f. Let j_0 be an integer satisfying $0 < j_0 \leq 2^n - 1$ and $HW(\alpha_{j_0}) = m+1$, where α_{j_0} is the binary representation of the integer j_0. Then $\langle \xi, \ell_{j_0} \rangle \equiv 0 \pmod{2^{m+2}}$ if and only if $\langle \xi, \ell_0 \rangle \equiv 0 \pmod{2^{m+2}}$.*

Lemma 3 allows us to claim

Lemma 4. *Let f be an mth-order correlation immune function on V_n, where $m \leq n-2$, and ξ be the sequence of f. Let j_0 be an integer satisfying $0 < j_0 \leq 2^n - 1$ and $HW(\alpha_{j_0}) = m+1$. If $\langle \xi, \ell_{j_0} \rangle \equiv 0 \pmod{2^{m+2}}$ then $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$ for any integer j satisfying $HW(\alpha_j) = m+1$, where α_j is the binary representation of j.*

The condition of $HW(\alpha_j) = m+1$ in the lemma above can be removed, as is shown below.

Lemma 5. *Let f be an mth-order correlation immune function on V_n, where $m \leq n-2$, and ξ be the sequence of f. Let j_0 be an integer satisfying $0 < j_0 \leq 2^n - 1$ and $HW(\alpha_{j_0}) = m+1$, where α_{j_0} is the binary representation of j_0. If $\langle \xi, \ell_{j_0} \rangle \equiv 0 \pmod{2^{m+2}}$, then $\langle \xi, \ell_i \rangle \equiv 0 \pmod{2^{m+2}}$ for any row ℓ_i of H_n.*

Proof. We use induction on $HW(\alpha_j)$ to prove that $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$, where α_j is the binary representation of the subscript j of ℓ_j.

For $0 < HW(\alpha_j) \leq m$, since f is an mth-order correlation immune function, we have $\langle \xi, \ell_j \rangle = 0$. On the other hand, from Lemma 4, we have $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$, where ℓ_j is any row of H_n satisfying $HW(\alpha_j) = m+1$, and α_j is the binary representation of j. Due to Lemma 3, we also have $\langle \xi, \ell_0 \rangle \equiv 0 \pmod{2^{m+2}}$. Hence we have proved $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$, when $HW(\alpha_j) \leq m+1$.

Now assume that $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$, when $m+1 \leq HW(\alpha_j) \leq k \leq n-1$. Consider the case of $HW(\alpha_j) = k+1$. Obviously, W can be rewritten as $W = \{\alpha_0, \alpha_{2^{n-k-1}}, \alpha_{2 \cdot 2^{n-k-1}}, \ldots, \alpha_{(2^{k+1}-1)2^{n-k-1}}\}$, where each α_j is the binary representation of an integer j. One can see that W is a $(k+1)$-dimensional linear subspace.

Let $\xi = (\xi_0, \xi_1, \xi_2, \ldots . \xi_{2^{k+1}-1})$, where each ξ_i is of length 2^{n-k-1}. Furthermore, let $p = k+1$ and $j = 0$ in (3). Then we have

$$\begin{aligned} & 2^{k+1}(\langle \xi_0, e_0 \rangle, \langle \xi_1, e_0 \rangle, \ldots, \langle \xi_{2^{k+1}-1}, e_0 \rangle) \\ & = (\langle \xi, \ell_0 \rangle, \langle \xi, \ell_{2^{n-k-1}} \rangle, \langle \xi, \ell_{2 \cdot 2^{n-k-1}} \rangle, \ldots, \langle \xi, \ell_{(2^{k+1}-1)2^{n-k-1}} \rangle) H_{k+1} \end{aligned} \quad (7)$$

where e_0 denotes the 0th row of H_{n-k-1}, i.e., the all-one sequence of length 2^{n-k-1}.

By the assumption, we should have $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$ where $j = i \cdot 2^{n-k-1}$, $i = 0, 1, \ldots, 2^{k+1} - 2$. Note that $k \geq m+1$. From (7), we have $\langle \xi, \ell_{(2^{k+1}-1)2^{n-k-1}} \rangle \equiv 0 \pmod{2^{m+2}}$. Furthermore, note that HW

$(\alpha_{(2^{k+1}-1)2^{n-k-1}}) = k+1$. Taking into account Lemma 4, we can conclude that $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{m+2}}$, for $HW(\alpha_j) = k+1$, where α_j is the binary representation of j. This completes the proof. □

In the following section, we will use these results to improve the upper bond on the nonlinearity of correlation immune functions.

6 Improving Upper Bounds on Nonlinearity

The following lemma will be used in proving Theorem 5.

Lemma 6. *Let f be an mth-order correlation immune function on V_n, where $\frac{1}{2}n - 1 < m \leq n-2$, and ξ denotes the sequence of f. If $\binom{n}{m+1} > 2^{2n-2m-2}$, then there must be an integer j_0, $0 \leq j_0 \leq 2^n - 1$, such that $HW(\alpha_{j_0}) = m+1$ and $\langle \xi, \ell_{j_0} \rangle = 0$, where α_{j_0} is the binary representation of integer j_0.*

Proof. Since f is an mth-order correlation immune function on V_n, due to Theorem 3 of [10], we have $\langle \xi, \ell \rangle \equiv 0 \pmod{2^{m+1}}$, where ℓ is any row of H_n. Hence $\langle \xi, \ell \rangle \neq 0$ implies that $|\langle \xi, \ell \rangle| \geq 2^{m+1}$. Using Parseval's equation (Page 416 [7]), we have $\#\Im \leq 2^{2n-2m-2}$.

Note that the number of vectors α in V_n, satisfying $HW(\alpha) = m+1$, is equal to $\binom{n}{m+1} > 2^{2n-2m-2}$. Hence there must be a vector α_{j_0} such that $HW(\alpha_{j_0}) = m+1$ and $\alpha_{j_0} \notin \Im^*$. As a result, we have $\langle \xi, \ell_{j_0} \rangle = 0$, where α_{j_0} is the binary representation of j_0. □

Theorem 5. *Let f be an mth-order correlation immune function on V_n, where $\frac{1}{2}n - 1 < m \leq n-2$. If $\binom{n}{m+1} > 2^{2n-2m-2}$, then $N_f \leq 2^{n-1} - 2^{m+1}$, where the equality holds if and only if f is is a $2(n-m-2)$th-order plateaued function.*

Proof. By Lemma 6, there must be a vector α_{j_0} such that $HW(\alpha_{j_0}) = m+1$ and $\langle \xi, \ell_{j_0} \rangle = 0$. Now using Lemma 5, we have

$$\langle \xi, \ell \rangle \equiv 0 \pmod{2^{m+2}} \tag{8}$$

where ℓ is any row of H_n. Lemma 1 implies that $N_f \leq 2^{n-1} - 2^{m+1}$.

Assume that $N_f = 2^{n-1} - 2^{m+1}$. From Lemma 1, we have

$$\max\{|\langle \xi, \ell_i \rangle|, 0 \leq i \leq 2^n - 1\} = 2^{m+2} \tag{9}$$

Combining (8) and (9), we can conclude that $\langle \xi, \ell \rangle = 2^{m+2}$ if $\langle \xi, \ell \rangle \neq 0$. This proves that f is a $2(n-m-2)$th-order plateaued function.

Conversely, if f is a $2(n-m-2)$th-order plateaued function, due to Proposition 1, we must have $N_f = 2^{n-1} - 2^{m+1}$. □

Let n and m be two integers with $n > m > 0$. We claim that that the following inequality holds:

$$\binom{n}{m+1} > (\frac{n+m+2}{n-m})^{n-m-1} \tag{10}$$

To prove the claim, we set $\rho(i) = \frac{(n-i)(m+2+i)}{(n-m-1-i)(1+i)}$, where $0 \le i \le \frac{1}{2}(n-m-2)$. Since $\binom{n}{m+1} = \binom{n}{n-m-1}$, it is easy to verify that

$$\binom{n}{m+1} = \begin{cases} \rho(0)\rho(1)\cdots\rho(\frac{1}{2}(n-m-2)-1)(\frac{n+m+2}{n-m}), & \text{if } n-m \text{ is even} \\ \rho(0)\rho(1)\cdots\rho(\frac{1}{2}(n-m-3)), & \text{if } n-m \text{ is odd} \end{cases} \tag{11}$$

In addition, one can also verify that ρ satisfies the condition of $\rho(i) < \rho(i-1)$. Hence

$$\binom{n}{m+1} > \begin{cases} (\rho(\frac{1}{2}(n-m-2))^{\frac{1}{2}(n-m-2)}(\frac{n+m+2}{n-m}), & \text{if } n-m \text{ is even} \\ (\rho(\frac{1}{2}(n-m-3))^{\frac{1}{2}(n-m-1)}, & \text{if } n-m \text{ is odd} \end{cases} \tag{12}$$

There exist two cases to be considered: $n-m$ is even and $n-m$ is odd.

In the former case, we note that $\rho(\frac{1}{2}(n-m-2)) = (\frac{n+m+2}{n-m})^2$. Due to (12), we obtain $\binom{n}{m+1} > (\frac{n+m+2}{n-m})^{n-m-1}$.

In the latter case, as $\rho(\frac{1}{2}(n-m-3)) = \frac{(n+m+3)(n+m+1)}{(n-m+1)(n-m-1)} > (\frac{n+m+2}{n-m})^2$, taking into account (12), we have $\binom{n}{m+1} > (\frac{n+m+2}{n-m})^{n-m-1}$. Thus the inequality in (10) is indeed true.

Theorem 6. *Let f be an mth-order correlation immune function on V_n. If m and n satisfy the condition of $0.6n - 0.4 \le m \le n-2$, then $N_f \le 2^{n-1} - 2^{m+1}$, where the equality holds if and only if f is also a $2(n-m-2)$th-order plateaued function.*

Proof. One can verify that

$$\frac{n+\lambda_1+2}{n-\lambda_1} > \frac{n+\lambda_2+2}{n-\lambda_2}$$

for $n > \lambda_1 > \lambda_2 > 0$, where λ_1 and λ_2 are not necessarily integers. Since $m \ge 0.6n - 0.4$, we have

$$(\frac{n+m+2}{n-m})^{n-m-1} \ge (\frac{n+0.6n-0.4+2}{n-(0.6n-0.4)})^{n-m-1} = 2^{2n-2m-2}.$$

By using (10), we can conclude that $\binom{n}{m+1} > 2^{2n-2m-2}$. Taking into account Theorem 5, we know that the theorem is indeed true. □

Part (i) of Theorem 4 in [10] states that the nonlinearity N_f of an mth-order correlation immune function f on V_n satisfies $N_f \leq 2^{n-1} - 2^m$, when $m > \frac{1}{2}n - 1$. Our Theorem 6 represents an improvement on the result in [10], especially for the case of $m \geq 0.6n - 0.4$.

As a consequence of Theorem 5 or Theorem 6, a correlation immune function that achieves the maximum nonlinearity for such a function, also satisfies all the properties of plateaued function, as discussed in Section 4. As a result, by taking into account Theorem 2, we have

Corollary 1. *Let f be an mth-order correlation immune function on V_n. If $0.6n - 0.4 \leq m \leq n - 2$, then $N_f \leq 2^{n-1} - 2^{m+1}$, where the equality holds if and only if f is also a $2(n-m-2)$th-order plateaued function or the equality in Theorem 2 holds, i.e., $\sum_{j=0}^{2^n-1} \Delta^2(\alpha_j) = 2^{n+2m+4}$.*

An (n, m, t)-*resilient function* is an n-input m-output function or mapping F with the property that it runs through every possible output m-tuple an equal number of times when t arbitrary inputs are fixed and the remaining $n-t$ inputs runs through all the 2^{n-t} input tuples once. The concept was introduced by Chor *et al* in [3] and independently, by Bennett *et al* in [1]. Comparing the definition of resilient functions with that of correlation immune functions, one can see that an $(n, 1, t)$-resilient function coincides with a balanced tth-order correlation immune function on V_n. In this context, Theorem 1 of [15] is of special interest to practitioners alike, as it shows that each non-zero linear combination of the component functions of an (n, m, t)-resilient function is also a balanced tth-order correlation immune function on V_n, giving rise to $2^m - 1$ distinct, balanced tth-order correlation immune functions in total.

To close this section, we point out a result which follows from Theorem 2 of [10] and Theorem 2 in this paper.

Corollary 2. *Let f be an $(n, 1, m)$-resilient function, where $\frac{1}{2}n - 2 < m \leq n - 3$. Then the nonlinearity N_f of f satisfies $N_f \leq 2^{n-1} - 2^{m+1}$, where the equality holds if and only if f is also a $2(n - m - 2)$th-order plateaued function or the equality in Theorem 2 holds, i.e., $\sum_{j=0}^{2^n-1} \Delta^2(\alpha_j) = 2^{n+2m+4}$.*

7 Tightness of the Upper Bound

As Theorem 6 represents an improved upper bound on the nonlinearity of all the correlation immune functions including both balanced and unbalanced ones, we are further interested in the question as to whether the upper bound is tight or not. It turns out that the question can be answered in an affirmative way for balanced correlation immune functions. The approach we take is to actually demonstrate the existence of mth-order correlation immune, balanced functions on V_n, whose nonlinearity N_f satisfies $N_f = 2^{n-1} - 2^{m+1}$.

We note that [11] is the earliest paper to study the nonlinearity of correlation immune functions. Of particular importance are Theorems 9 and 14 in [11] which happen to be also relevant to the current work. Theorem 9 of

[11] proved the equivalence of two different methods for constructing correlation immune functions, while Theorem 4 in the same paper showed how to obtain highly nonlinear correlation immune functions. Let integers n and m satisfy $m+2 \geq 2^{n-m-2}$ and $n \geq 16$. For such n and m, there exist 2^{n-m-2} non-zero vectors in V_{m+2}, say $\gamma_0, \gamma_1, \ldots, \gamma_{2^{n-m-2}-1}$, such that $HW(\gamma_j) \geq m+1$, where $j = 0, 1, \ldots, 2^{n-m-2}-1$. Define a mapping P from V_{n-m-2} to V_{m+2} such that $P(V_{n-m-2}) = \{\gamma_0, \gamma_1, \ldots, \gamma_{2^{n-m-2}-1}\}$, where $P(V_{n-m-2}) = \{P(\delta) | \delta \in V_{n-m-2}\}$. Based on P, we construct a function f on V_n by $f(x) = f(y,z) = P(y)z^T$ where $x = (y,z)$, $y \in V_{n-m-2}$ and $z \in V_{m+2}$. By using Theorems 9 and 14 of [11], modifying the relevant parameters accordingly, and fixing t to 1, we can construct an $(n, 1, m)$-resilient (balanced) function f whose nonlinearity N_f reaches the upper bound of $2^{n-1} - 2^{m+1}$.

As a concrete example, let $n = 9$ and $m = 5$. Then $m + 2 \geq 2^{n-m-2}$. Set $\gamma_0 = (1,1,1,1,1,1,1)$, $\gamma_1 = (1,1,1,1,1,1,0)$, $\gamma_2 = (1,1,1,1,1,0,1)$ and $\gamma_3 = (1,1,1,1,0,1,1)$. Then each $\gamma_j \in V_7$ and $HW(\gamma_j) \geq 6$. Define a mapping P from V_2 to V_7 such that $P(0,0) = \gamma_0$, $P(0,1) = \gamma_1$ $P(1,0) = \gamma_2$ and $P(1,1) = \gamma_3$. Based on P, we construct a function f on V_9 by $f(x) = f(y,z) = P(y)z^T$ where $x = (y,z)$, $y \in V_2$ and $z \in V_7$. Theorems 9 and 14 in [11] tell us that f is a 5th-order correlation immune function on V_9, and the nonlinearity N_f of f achieves $N_f = 2^8 - 2^6 = 192$, the highest possible value for such a function. Since each γ_j is non-zero, f is balanced. One can verify that the function f takes the form of

$$\begin{aligned} f(y,z) = {} & y_1 z_6 \oplus y_2 z_7 \oplus y_1 y_2 (z_5 \oplus z_6 \oplus z_7) \\ & \oplus z_1 \oplus z_2 \oplus z_3 \oplus z_4 \oplus z_5 \oplus z_6 \oplus z_7 \end{aligned}$$

where $y = (y_1, y_2)$ and $z = (z_1, z_2, z_3, z_4, z_5, z_6, z_7)$.

The above discussions indicate that the upper bound $(2^{n-1} - 2^{m+1})$ is indeed tight for balanced correlation immune functions. While we have not been able to identify whether the bound is also tight for unbalanced correlation immune functions, its implication would be marginal, due to the fact that unbalanced correlation immune functions have found little use in practice.

To close this section, let us note that in [13], an unbalanced 3rd-order correlation immune function on V_6 whose nonlinearity achieves $2^5 - 2^3 = 24$ is constructed. This particular function does not contradict Theorem 5 or Theorem 6, as the specific parameters $n = 6$ and $m = 3$ satisfy neither $\binom{n}{m+1} > 2^{2n-2m-2}$ nor $m \geq 0.6n - 0.4$.

8 Concluding Remarks

Three separate research groups, Sarkar and Maitra, Tarannikov [13], and Zheng and Zhang, have apparently considered the same question on the upper bound on nonlinearity of correlation immune functions, independently of one another. All three groups submitted their research results to CRYPTO2000, although

only Sarkar and Maitra's got accepted. Our current paper contains essentially the same research results included in our CRYPTO2000 submission, minus those that happened to overlap results in Sarkar and Maitra's CRYPTO2000 paper.

Theorem 6 leaves open as to whether the condition of $0.6n - 0.4 \leq m \leq n - 2$ can be relaxed to $\frac{1}{2}n - 1 < m \leq n - 2$ where $n > 6$. We have recently successfully solved this problem [17].

It would also be interesting, albeit purely from a theoretical point of view, to examine whether the bound $N_f = 2^{n-1} - 2^{m+1}$, where $m \geq 0.6n - 0.4$, is also tight for unbalanced mth-order correlation immune functions, and if it is, how to construct such functions.

Acknowledgment

The second author was supported by a Queen Elizabeth II Fellowship (227 23 1002). Both authors would like to thank Yuriy Tarannikov for pointing out an error in an earlier version, and anonymous referees for SAC2000 for their comments that have helped further improve both the upper bound and the presentation of this paper.

References

1. C. H. Bennett, G. Brassard, and J. M. Robert. Privacy amplification by public discussion. *SIAM J. Computing*, 17:210–229, 1988.
2. P. Camion, C. Carlet, P. Charpin, and N. Sendrier. On correlation-immune functions. In *Advances in Cryptology - CRYPTO'91*, volume 576 of *Lecture Notes in Computer Science*, pages 87–100. Springer-Verlag, Berlin, Heidelberg, New York, 1991.
3. Benny Chor, Oded Goldreich, Johan Håstad, Joel Friedman, Steven Rudich, and Roman Smolensky. The bit extraction problem or t-resilient functions. *IEEE Symposium on Foundations of Computer Science*, 26:396–407, 1985.
4. C. Ding, G. Xiao, and W. Shan. *The Stability Theory of Stream Ciphers*, volume 561 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, Heidelberg, New York, 1991.
5. Xiao Guo-Zhen and J. L. Massey. A spectral characterization of correlation-immune combining functions. *IEEE Transactions on Information Theory*, 34(3):569–571, 1988.
6. M. Hermelin and K. Nyberg. Correlation properties of the bluetooth combiner generator. In *The 2nd International Conference on Information Security and Cryptology (ICISC'99), Seoul, Korea*, volume 1787 of *Lecture Notes in Computer Science*, pages 17–29. Springer-Verlag, Berlin, Heidelberg, New York, 1999.
7. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam, New York, Oxford, 1978.
8. 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.

9. O. S. Rothaus. On "bent" functions. *Journal of Combinatorial Theory*, Ser. A, 20:300–305, 1976.
10. P. Sarkar and S. Maitra. Nonlinearity bounds and constructions of resilient boolean functions. In *Advances in Cryptology - CRYPTO2000*, volume 1880 of *Lecture Notes in Computer Science*, pages 515–532. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
11. 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.
12. T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, IT-30 No. 5:776–779, 1984.
13. Yuriy Tarannikov. On resilient boolean functions with maximal possible nonlinearity. (`http://eprint.iacr.org/2000/005/`), 2000.
14. R. Yarlagadda and J. E. Hershey. Analysis and synthesis of bent sequences. *IEE Proceedings (Part E)*, 136:112–123, 1989.
15. X. M. Zhang and Y. Zheng. Cryptographically resilient functions. *IEEE Transactions on Information Theory*, 43(5):1740–1747, 1997.
16. Y. Zheng and X. M. Zhang. Plateaued functions. In *Advances in Cryptology - ICICS'99*, volume 1726 of *Lecture Notes in Computer Science*, pages 284–300. Springer-Verlag, Berlin, Heidelberg, New York, 1999.
17. Y. Zheng and X. M. Zhang. Two new results on correlation immune functions, August 2000. (submitted for publication).