

# Duality of Boolean Functions and Its Cryptographic Significance

Xian-Mo Zhang<sup>1</sup> and Yuliang Zheng<sup>2</sup> and Hideki Imai<sup>3</sup>

<sup>1</sup> The University of Wollongong, Wollongong, NSW 2522, Australia  
xianmo@cs.uow.edu.au

<sup>2</sup> Monash University, Frankston, Melbourne, VIC 3199, Australia  
yzheng@fcit.monash.edu.au, <http://www-pscit.fcit.monash.edu.au/~yuliang/>

<sup>3</sup> The University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106, JAPAN  
imai@iis.u-tokyo.ac.jp

**Abstract.** Recent advances in interpolation and high order differential cryptanalysis have highlighted the cryptographic significance of Boolean functions with a high algebraic degree. However, compared with other nonlinearity criteria such propagation, resiliency, differential and linear characteristics, apparently little progress has been made in relation to algebraic degree in the context of cryptography. The aim of this work is to research into relationships between algebraic degree and other nonlinearity criteria. Making use of duality properties of Boolean functions, we have obtained several results that are related to lower bounds on nonlinearity, as well as on the number of terms, of Boolean functions. We hope that these results would stimulate the research community's interest in further exploring this important area.

## 1 Introduction

The algebraic degree has long been believed by many designers of block ciphers and one-way hash functions to be an important nonlinearity indicator for the cryptographic strength of Boolean functions. Recent progress in interpolation cryptanalysis [1] and high order differential cryptanalysis [5] can be viewed as a proof for the correctness of the belief. Of particular interest is the work of [5] in which the authors showed how to break in less than 20 milli-seconds a block cipher that employs low algebraic degree (quadratic) Boolean functions as its S-boxes and is provably secure against linear and (the first order) differential attacks.

Investigation into the algebraic degree of Boolean functions has been a difficult topic. This is supported by the fact that, while the past few years have seen much progress in relation to other nonlinearity criteria such as propagation, differential profile, nonlinear profile, resiliency, correlation-immunity, local and global avalanche characteristics, little progress has been made in designing Boolean functions that have a high algebraic degree and also satisfy other important nonlinearity criteria.

In this paper we tackle algebraic degree, together with nonlinearity, propagation characteristics, correlation immunity and the number of terms in a Boolean

function by exploring the duality property of a Boolean function. Main contributions of this work are to show (1) two lower bounds, one on the nonlinearity and the other on the number of terms of a Boolean functions, and (2) a connection between the algebraic degree of a Boolean function and its Walsh-Hadamard transform.

## 2 Basic Definitions

We consider 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  $(f(\alpha_0), f(\alpha_1), \dots, f(\alpha_{2^n-1}))$ , and the *sequence* of  $f$  is a  $(1, -1)$ -sequence defined by  $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \dots, (-1)^{f(\alpha_{2^n-1})})$ , where  $\alpha_0 = (0, \dots, 0, 0)$ ,  $\alpha_1 = (0, \dots, 0, 1)$ ,  $\dots$ ,  $\alpha_{2^n-1} = (1, \dots, 1, 1)$ . 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  $GF(2)$ .  $f$  is said to be *balanced* if its truth table contains an equal number of ones and zeros.

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

Let  $\tilde{a} = (a_1, \dots, a_m)$  and  $\tilde{b} = (b_1, \dots, b_m)$  be two vectors (or sequences), 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  $V_m$ ,  $\langle \tilde{a}, \tilde{b} \rangle = a_1 b_1 \oplus \dots \oplus a_m b_m$ , 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  $f(x_1, \dots, x_n) = a_1 x_1 \oplus \dots \oplus a_n x_n \oplus c$ , where  $a_j, c \in GF(2)$ ,  $j = 1, 2, \dots, n$ . Furthermore  $f$  is called a *linear* function if  $c = 0$ .

**Definition 1.** The *Hamming weight* of a  $(0, 1)$ -sequence  $\xi$  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  $x = (x_1, \dots, x_n)$ . The *nonlinearity* of  $f$ , denoted by  $N_f$ , is the minimal Hamming distance between  $f$  and all affine functions on  $V_n$ , i.e.,  $N_f = \min_{i=1, 2, \dots, 2^{n+1}} d(f, \varphi_i)$  where  $\varphi_1, \varphi_2, \dots, \varphi_{2^{n+1}}$  are all the affine functions on  $V_n$ .

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

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

Let  $\ell_i$ ,  $0 \leq i \leq 2^n - 1$ , be the  $i$  row of  $H_n$ . By Lemma 2 of [4],  $\ell_i$  is the sequence of a linear function  $\varphi_i(x)$  defined by the scalar product  $\varphi_i(x) = \langle \alpha_i, x \rangle$ , where  $\alpha_i$  is the  $i$ th vector in  $V_n$  according to the ascending alphabetical order.

**Definition 2.** 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  $\alpha$ .

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  $\alpha$ . On the other hand, if  $|\Delta(\alpha)| = 2^n$ , then  $f(x) \oplus f(x \oplus \alpha)$  is a constant and hence  $\alpha$  is a linear structure of  $f$ .

A function  $f$  on  $GF(2)$  can be uniquely represented by a polynomial on  $GF(2)$  whose degree is at most  $n$ . Namely,

$$f(x_1, \dots, x_n) = \bigoplus_{\alpha \in V_n} g(a_1, \dots, a_n) x_1^{a_1} \cdots x_n^{a_n} \quad (1)$$

where  $\alpha = (a_1, \dots, a_n)$ , and  $g$  is also a function on  $V_n$ . Each  $x_1^{a_1} \cdots x_n^{a_n}$  is called a term (in the polynomial representation) of  $f$ .

The algebraic degree, or simply degree, of  $f$ , denoted by  $deg(f)$ , is defined as the number of variables in the longest term of  $f$ , i.e.,

$$deg(f) = \max\{W(a_1, \dots, a_n) \mid g(a_1, \dots, a_n) = 1\}.$$

**Definition 3.** Let  $f$  be a function on  $V_n$  and  $U$  be  $s$ -dimensional subspace of  $V_n$ . The restriction of  $f$  to  $U$ , denoted by  $f_U$ , is a function on  $U$ , defined by the following rule

$$f_U(\alpha) = f(\alpha) \text{ for every } \alpha \in U.$$

**Notation 1** Let  $W$  be a subspace of  $V_n$ . Denote the dimension of  $W$  by  $dim(W)$ .

**Notation 2**  $(b_1, \dots, b_n) \preceq (a_1, \dots, a_n)$  means that  $(b_1, \dots, b_n)$  is covered by  $(a_1, \dots, a_n)$ , namely if  $b_j = 1$  then  $a_j = 1$ . In addition,  $(b_1, \dots, b_n) \prec (a_1, \dots, a_n)$  means that  $(b_1, \dots, b_n)$  is properly covered by  $(a_1, \dots, a_n)$ , namely  $(b_1, \dots, b_n) \preceq (a_1, \dots, a_n)$  and  $(b_1, \dots, b_n) \neq (a_1, \dots, a_n)$ .

### 3 Duality of Boolean Functions

The dual of a Boolean function  $f$  is a function  $g$  that is uniquely determined by the coefficients of the terms of  $f$ . The main purpose of this section is to provide the minimum amount of knowledge on duality that is required in the rest part of this paper. A proof for the following result is provided, as we feel that understanding the proof would be helpful in studying other issues that are more directly related to cryptography.

**Theorem 4.** Let  $f$  be a function on  $V_n$ . Let  $\alpha, \beta \in V_n$   $\alpha = (1, \dots, 1, 0, \dots, 0)$  where only the first  $s$  components are one, and  $\beta = (0, \dots, 0, 1, \dots, 1, 0, \dots, 0)$  where only the  $(s+1)$ th,  $\dots$ , the  $(s+t)$ th components are one. Then the number of the terms among  $x_1 \cdots x_s$ ,  $x_1 \cdots x_s x_{s+1}$ ,  $\dots$ ,  $x_1 \cdots x_s x_{s+1} \cdots x_{s+t}$  that appear in the polynomial representation of  $f$ , is even if  $\bigoplus_{\gamma \preceq \alpha} f(\gamma \oplus \beta) = 0$ , and this number is odd if  $\bigoplus_{\gamma \preceq \alpha} f(\gamma \oplus \beta) = 1$ .

*Proof.* Consider a term

$$\chi(x) = x_{j_1} \cdots x_{j_{s'}} x_{i_1} \cdots x_{i_t} \quad (2)$$

in  $f$ , where  $x = (x_1, \dots, x_n)$ ,  $1 \leq j_1 \leq \dots \leq j_{s'} \leq s$  and  $s+1 \leq i_1 \leq \dots \leq i_{s'+t} \leq s+t$ . For  $s' < s$ , there are an even number of vectors  $\gamma$  in  $V_n$  such that  $\gamma \prec \alpha$  and  $\chi(\gamma \oplus \beta) = 1$ . Hence

$$\bigoplus_{\gamma \preceq \alpha} \chi(\gamma \oplus \beta) = 0. \quad (3)$$

For  $s' = s$ , there is only one vector in  $V_n$ ,  $\gamma = \alpha$ , such that  $\chi(\gamma \oplus \beta) = 1$ . Hence

$$\bigoplus_{\gamma \preceq \alpha} \chi(\gamma \oplus \beta) = 1. \quad (4)$$

Now consider a term

$$\omega(x) = x_{j_1} \cdots x_{j_k} \quad (5)$$

in  $f$ , where  $x = (x_1, \dots, x_n)$ ,  $1 \leq j_1 \leq \dots \leq j_k$ , and  $j_k > s+t$ . From (5) with  $j_k > s+t$ , and the structures of  $\alpha$  and  $\beta$ ,

$$\omega(\gamma \oplus \beta) = 0 \quad (6)$$

for each  $\gamma \preceq \alpha$ . Denote the set of terms given in (2) by  $\Gamma_1$  if  $s' < s$ , and by  $\Gamma_2$  if  $s' = s$ . And denote the set of terms given in (5) by  $\Omega$ . Then we can write  $f$  as

$$f = \bigoplus_{\chi \in \Gamma_1} \chi \oplus \bigoplus_{\chi \in \Gamma_2} \chi \oplus \bigoplus_{\omega \in \Omega} \omega.$$

From (3), (4) and (6),

$$\bigoplus_{\gamma \preceq \alpha} f(\gamma \oplus \beta) = \bigoplus_{\gamma \preceq \alpha} \bigoplus_{\chi \in \Gamma_2} \chi(\gamma \oplus \beta). \quad (7)$$

$\bigoplus_{\gamma \preceq \alpha} f(\gamma \oplus \beta) = 0$  implies that  $|\Gamma_2|$  is even, while  $\bigoplus_{\gamma \preceq \alpha} f(\gamma \oplus \beta) = 1$  implies that  $|\Gamma_2|$  is odd. This completes the proof.  $\square$

Set  $\beta = 0$  in Theorem 4 and reorder the variables, we obtain a result well known to coding theorists (see p.372 of [3]):

**Corollary 5.** Let  $f$  be a function on  $V_n$  and  $\alpha = (a_1, \dots, a_n)$ , a vector in  $V_n$ . Then the term  $x_1^{a_1} \cdots x_n^{a_n}$  appears in  $f$  if and only if  $\bigoplus_{\gamma \preceq \alpha} f(\gamma) = 1$ .

With the above two results, it is not hard to verify the correctness of the following theorem:

**Theorem 6.** *Let  $f$  and  $g$  be function on  $V_n$ . Then the following four statements are equivalent*

- (i)  $f(\alpha) = \bigoplus_{\beta \preceq \alpha} g(\beta)$  for every vector  $\alpha \in V_n$ .
- (ii)  $g(\alpha) = \bigoplus_{\beta \preceq \alpha} f(\beta)$  for every vector  $\alpha \in V_n$ .
- (iii)  $f(x_1, \dots, x_n) = \bigoplus_{\alpha \in V_n} g(a_1, \dots, a_n) x_1^{a_1} \cdots x_n^{a_n}$  where  $\alpha = (a_1, \dots, a_n)$ .
- (iv)  $g(x_1, \dots, x_n) = \bigoplus_{\alpha \in V_n} f(a_1, \dots, a_n) x_1^{a_1} \cdots x_n^{a_n}$  where  $\alpha = (a_1, \dots, a_n)$ .

## 4 Polynomial Representation and Nonlinearity

### 4.1 Restriction to Cosets

Let  $f$  be a function on  $V_n$  and  $U$  be an  $s$ -dimensional subspace of  $V_n$ . Then  $V_n$  is the union of  $2^{n-s}$  disjoint  $2^s$ -subsets

$$V_n = \Pi_0 \cup \Pi_1 \cup \cdots \cup \Pi_{2^{n-s}-1} \quad (8)$$

where

- (i)  $\Pi_0 = U$ ,
- (ii) for any  $\alpha, \beta \in V_n$ ,  $\alpha, \beta$  belong to the same class, say  $\Pi_j$ , if and only if  $\alpha \oplus \beta \in \Pi_0 = U$ . From (i) and (ii), it follows that
- (iii)  $\Pi_j \cap \Pi_i = \phi$  for  $j \neq i$ , where  $\phi$  denotes the empty set.

As each  $\Pi_j$  can be expressed as  $\Pi_j = \beta_j \oplus U$  for a  $\beta_j \in V_n$ , where  $\beta_j \oplus U = \{\beta_j \oplus \alpha \mid \alpha \in U\}$ , the definition of restriction (Definition 3) can be extended to each coset  $\Pi_j$ .

**Definition 7.** *Let  $f$  be a function on  $V_n$  and  $U$  be an  $s$ -dimensional subspace of  $V_n$ . The restriction of  $f$  to a coset  $\Pi_j = \beta_j \oplus U$ ,  $j = 0, 1, \dots, 2^{n-s} - 1$ , denoted by  $f_{\Pi_j}$ , is a function on  $U$ , and it is defined by  $f_{\Pi_j}(\alpha) = f(\beta_j \oplus \alpha)$  for every  $\alpha \in U$ .*

### 4.2 Maximal Odd Weighting Subspaces

**Definition 8.** *Let  $f$  be a function on  $V_n$ . A subspace  $U$  of  $V_n$  is called a maximal odd weighting subspace of  $f$  if the Hamming weight of  $f_U$  is odd and the Hamming weight of  $f_{U'}$ , where  $U'$  is any subspace with  $U' \supset U$  (i.e.  $U$  is a proper subset of  $U'$ ), is even.*

A maximal odd weighting subspace of a function is not necessarily a subspace with the maximum dimension, even if the Hamming weight of the restrictions of  $f$  to the subspace is odd. This is best explained with the following example.

*Example 1.* let  $f(x_1, x_2, x_3, x_4) = x_1x_2x_3 \oplus x_1x_2x_4 \oplus x_3x_4 \oplus x_3$  be a function on  $V_4$ , whose truth table is 001000100010000. The eight vectors (0000), (0001), (0100), (0101), (1000), (1001), (1100) and (1101) form a 3-dimensional subspace, say  $W$ , such that the Hamming weight of  $f_W$  is one (odd). By a direct verification, 3 is the maximum dimension of the subspaces, the Hamming weight of the restrictions of  $f$  to these subspaces is odd. However, the four vectors (0000), (0001), (0010) and (0011) form a 2-dimensional subspace, say  $U$ , such that the Hamming weight of  $f_W$  is one (odd). There are four 3-dimensional subspaces containing  $U$ :

$$U' = \{(0000), (0001), (0010), (0011), (0100), (0101), (0110), (0111)\}$$

$$U'' = \{(0000), (0001), (0010), (0011), (1000), (1001), (1010), (1011)\}$$

$$U''' = \{(0000), (0001), (0010), (0011), (1000), (1001), (1010), (1011)\}$$

We note that the Hamming weights of  $f_{U'}$ ,  $f_{U''}$  and  $f_{U'''}$  are all two (even). We also note that the 4-dimensional subspace containing  $U$  is  $V_4$  itself and the Hamming weight of  $f$  is four (even). Hence both  $W$  and  $U$  are a maximal odd weighting subspace of  $f$ .

As will be shown in the forthcoming sections, the concept of maximal odd weighting subspace of a function plays an important role, primarily due to the fact that the dimension of a subspace is relevant to the structure of the function. In particular, we will show in the next section a connection between the dimension of a maximal odd weighting subspace of a function and the lower bound on nonlinearity of a function.

### 4.3 A Lower Bound on Nonlinearity

**Definition 9.** Let  $f$  be a function on  $V_n$ ,  $x_{j_1} \cdots x_{j_t}$  and  $x_{i_1} \cdots x_{i_s}$  be two terms in the polynomial representation of function  $f$ .  $x_{j_1} \cdots x_{j_t}$  is said to be covered by  $x_{i_1} \cdots x_{i_s}$  if  $\{j_1, \dots, j_t\}$  is a subset of  $\{i_1, \dots, i_s\}$ , and  $x_{j_1} \cdots x_{j_t}$  is said to be properly covered by  $x_{i_1} \cdots x_{i_s}$  if  $\{j_1, \dots, j_t\}$  is a proper subset of  $\{i_1, \dots, i_s\}$ .

**Theorem 10.** Let  $f$  be a function on  $V_n$  and  $U$  be a maximal odd weighting subspace of  $f$ . If  $\dim(U) = s$  then the Hamming weight of  $f$  is at least  $2^{n-s}$ .

*Proof.* Let  $U$  be a subspace defined in (8). And let  $N_j = |\{\alpha | \alpha \in \Pi_j, f(\alpha) = 1\}|$ , where  $\Pi_j$  is defined in (8),  $j = 0, 1, \dots, 2^s - 1$ . Since  $\Pi_0 = U$ ,  $N_0$  is odd. Note that  $\Pi_0 \cup \Pi_j$  is a  $(s+1)$ -dimensional subspace of  $V_n$ ,  $j = 1, \dots, 2^{n-s} - 1$ .

Since  $\Pi_0 = U$  is a maximal odd weighting subspace of  $f$ , Hamming weight of the restriction of  $f$  to  $\Pi_0 \cup \Pi_j$  is even. In other words,  $N_0 + N_j$  is even. This proves that each  $N_j$  is odd,  $j = 1, \dots, 2^{n-s} - 1$ . Hence  $N_0 + N_1 + \cdots + N_{2^{n-s}-1} \geq 2^{n-s}$ , namely, the Hamming weight of  $f$  is at least  $2^{n-s}$ .  $\square$

**Theorem 11.** Let  $f$  be a function on  $V_n$  and  $U$  be a maximal odd weighting subspace of  $f$ . Let  $\dim(U) = s$  ( $s \geq 2$ ) then the nonlinearity of  $f$ ,  $N_f$ , satisfies  $N_f \geq 2^{n-s}$ .

*Proof.* Let  $\varphi$  be an affine function on  $V_n$ . Since  $s \geq 2$  the Hamming weight of  $\varphi_U$  must be even. Hence the Hamming weight of  $\varphi_U$  must be even. Hence the Hamming weight of  $(f \oplus \varphi)_U$  must be odd. According to Lemma 10, the Hamming weight of  $f \oplus \varphi$  is at least  $2^{n-s}$ . As the Hamming weight of  $f \oplus \varphi$  determines  $d(f, \varphi)$ , the theorem is proved.  $\square$

**Theorem 12.** *Let  $t \geq 2$ . If  $x_{j_1} \cdots x_{j_t}$  is a term in a function  $f$  on  $V_n$  and it is not properly covered (see Definition 9) by any other terms in the same function, then the nonlinearity of  $f$ ,  $N_f$ , satisfies  $N_f \geq 2^{n-t}$ .*

*Proof.* Write  $\alpha = (a_1, \dots, a_n)$  where  $a_j = 1$  for  $j \in \{j_1, \dots, j_t\}$  and  $a_j = 0$  for  $j \notin \{j_1, \dots, j_t\}$ . Set

$$U = \{\gamma \mid \gamma \preceq \alpha\}.$$

Obviously  $U$  is a  $t$ -dimensional subspace of  $V_n$ . Since  $x_{j_1} \cdots x_{j_t}$  is a term in  $f$  on  $V_n$ , by using Corollary 5,  $\bigoplus_{\gamma \preceq \alpha} f(\gamma) = 1$  or  $\bigoplus_{\gamma \in U} f(\gamma) = 1$  i.e. the Hamming weight of  $f_U$  is odd.

We now prove that  $U$  is a maximal odd weighting subspace of  $f$ . Suppose  $U$  is not a maximal odd weighting subspace of  $f$ . Hence there is a  $s$ -dimensional subspace of  $V_n$ , say  $W$ , such that  $U$  is a proper subset of  $W$  i.e.  $s > t$  and the Hamming weight of  $f_W$  is odd i.e.  $\bigoplus_{\gamma \in W} f(\gamma) = 1$ . Since  $U$  is a proper subspace of  $W$ , by using linear algebra,  $W$  can be expressed as a union of  $2^{s-t}$  disjoint  $2^t$ -subsets

$$W = U \cup (\beta_1 \oplus U) \cup \cdots \cup (\beta_{2^{s-t}-1} \oplus U) \quad (9)$$

where each  $\beta \preceq \bar{\alpha}$ , where  $\bar{\alpha} \oplus \alpha = (1, \dots, 1)$ . Since both the Hamming weights of  $f_U$  and  $f_W$  are odd, there is a coset, say  $\beta_k \oplus U$ ,  $1 \leq k \leq 2^{s-t} - 1$ , such that the Hamming weight of  $f_{\beta_k \oplus U}$  is even or  $\bigoplus_{\gamma \in U} f(\beta_k \oplus \gamma) = 0$  i.e.

$$\bigoplus_{\gamma \preceq \alpha} f(\beta_k \oplus \gamma) = 0. \quad (10)$$

Applying Theorem 4 to (10), there are even number of terms covering  $x_{j_1} \cdots x_{j_t}$ . Since the term  $x_{j_1} \cdots x_{j_t}$  itself appears in  $f$ , there is another term properly covering  $x_{j_1} \cdots x_{j_t}$ . This contradicts the condition in the theorems, that the term  $x_{j_1} \cdots x_{j_t}$  is not properly covered by any other terms in  $f$ . The contradiction proves that  $U$  is a maximal odd weighting subspace of  $f$ . By using Theorem 11, the proof is completed.  $\square$

*Example 2.* Let

$$f(x_1, \dots, x_{10}) = x_1 x_2 x_3 x_4 x_5 x_6 x_7 \oplus x_3 x_4 x_5 x_6 x_7 x_8 x_9 \oplus x_7 x_8 x_9 x_{10} \oplus x_4 x_6 x_8 x_{10} \oplus x_1 x_5 x_9 \oplus x_2 x_4 \oplus x_6$$

be a function on  $V_{10}$ . term  $x_1 x_5 x_9$  is not properly covered by any other terms in  $f$ . By using Corollary 12, the nonlinearity of  $f$ ,  $N_f$ , satisfies  $N_f \geq 2^{10-3} = 2^7$ .

*Example 3.* Let

$$f(x_1, \dots, x_{10}) = x_1 x_2 x_3 x_4 x_5 x_6 x_7 \oplus x_3 x_4 x_5 x_6 x_7 x_8 x_9 \oplus x_7 x_8 x_9 x_{10} \oplus x_4 x_6 x_8 x_{10} \oplus x_1 x_3 x_5 \oplus x_2 x_8 \oplus x_1 \oplus x_2$$

be a function on  $V_{10}$ . The term  $x_2 x_8$  is not properly covered by any other terms in  $f$ . By using Corollary 12, the nonlinearity of  $f$ ,  $N_f$ , satisfies  $N_f \geq 2^{10-2} = 2^8$ .

We note that the lower bound in Theorem 11 is tight:

**Corollary 13.** *For any  $n$  and any  $s$ ,  $2 \leq s \leq n$ , there are a function on  $V_n$ , say  $f$ , and a  $s$ -dimensional subspace, say  $U$ ,  $U$  be a maximal odd weighting subspace of  $f$  and the nonlinearity of  $f$ ,  $N_f$ , satisfies  $N_f = 2^{n-s}$ .*

*Proof.* We prove the corollary by an example. Let  $g$  be a function on  $V_s$ , defined as  $g(\beta) = 1$  if and only if  $\beta = 0$ . Set  $f(z, y) = g(y)$ , a function on  $V_n$ , where  $z \in V_{n-s}$  and  $y \in V_s$ . Since the Hamming weight of  $f$  is  $2^{n-s}$  ( $s \geq 2$ ),  $d(f, h) \geq 2^{n-s}$  where  $h$  is any affine function on  $V_n$  and the equality holds if  $h$  is the zero function on  $V_n$ . Hence the nonlinearity of  $f$ ,  $N_f$ , satisfies  $N_f = 2^{n-s}$ . On the other hand, set

$$U = \{(0, \dots, 0, b_1, \dots, b_s) | b_j \in GF(2)\}$$

where the number of zeros is  $n-s$ . It is easy to verify that  $s$ -dimensional subspace  $U$  is a maximal odd weighting subspace of  $f$ .  $\square$

Finally we note that for  $s = 2$ , the value of  $2^{n-s}$  in Theorem 11 is very close to  $2^{n-1} - 2^{\frac{1}{2}n-1}$ , the upper bound on the nonlinearity of functions on  $V_n$  [4]. However Theorem 11 cannot be further improved by extending  $s$  to  $s = 1$ , as the condition of  $s \geq 2$  in the proof of the theorem cannot be removed. For example, let  $f$  be a function on  $V_n$ , whose truth table is given as follows

$$0110011010011001.$$

It is easy to verify that (0000), (0001) form a maximal 1-dimensional subspace, denoted by  $U$ . Theorem 11 is not applicable due to the fact that  $\dim(U) = 1$ . In fact,  $f$  is a linear function, hence its nonlinearity is 0.

Nevertheless, Theorem 10 can be applied, which gives us  $\geq 2^{4-1} = 8$  as the Hamming weight of  $f$ .

#### 4.4 A Lower Bound on the Number of Terms

**Theorem 14.** *Let  $f$  be a function on  $V_n$  such that  $f(\alpha) = 1$  for a vector  $\alpha \in V_n$ , and  $f(\beta) = 0$  for every vector  $\beta$  with  $\alpha \prec \beta$  where  $\prec$  is defined as in Notation 2. Then  $f$  has at least  $2^{n-t}$  terms where  $t$  denotes the Hamming weight of  $\alpha$ .*

*Proof.* We first give Theorem 10 an equivalent statement, that we call Theorem 10', as follows

**Theorem 10'** Let  $f$  be a function on  $V_n$  and  $g$  be defined in (1). Let  $g(\alpha) = 1$  for a vector  $\alpha \in V_n$ , and  $g(\beta) = 0$  for every vector  $\beta$  with  $\alpha \prec \beta$  where  $\prec$  is defined as in Notation 2. Then the Hamming weight of  $f$  is at least  $2^{n-t}$ .

The equivalence between (iii) and (iv) in Theorem 6 allows us to interchange  $f$  and  $g$  in Theorem 10'. Thus we have

**Theorem 10''** Let  $f$  be a function on  $V_n$  and  $g$  be defined in (1). Let  $f(\alpha) = 1$  for a vector  $\alpha \in V_n$ , and  $f(\beta) = 0$  for every vector  $\beta$  with  $\alpha \prec \beta$  where  $\prec$  is defined as in Notation 2. Then the Hamming weight of  $g$  is at least  $2^{n-t}$ .

This completes the proof.  $\square$

**Corollary 15.** *Let  $f$  be a function on  $V_n$  such that  $f(\alpha) = 0$  for a vector  $\alpha \in V_n$ , and  $f(\beta) = 1$  for every vector  $\beta$  with  $\alpha \prec \beta$  where  $\prec$  is defined as in Notation 2. then  $f$  has at least*

- (i)  $2^{n-s} - 1$  terms if  $f(0) = 0$ ,
- (ii)  $2^{n-s} + 1$  terms if  $f(0) = 1$ ,

where  $s$  denotes the Hamming weight of  $\alpha$ .

*Proof.* Set  $f' = 1 \oplus f$ . Hence  $f'(\alpha) = 1$  and  $f'(\beta) = 0$  for every  $\beta \in V_n$ . By using Theorem 14,  $f'$  has at least  $2^{n-s}$  terms and hence  $f$  has at least  $2^{n-s} - 1$  terms. This proves (i) of the corollary.

In the above the proof, we have already proved that  $f'$  has at least  $2^{n-s}$  terms. Suppose  $f(0) = 1$ . Note that  $f'(0) = 0$ . Hence  $f$  has at least  $2^{n-s} + 1$  terms.  $\square$

*Example 4.* Let  $f$  be a function on  $V_6$ , whose truth table is given as follows

1000110111110010001101001100100001111100011001101001011010001010

Note that the value of  $f(001011)$  is one, while the values of  $f(001111)$ ,  $f(011011)$ ,  $f(011111)$ ,  $f(101011)$ ,  $f(101111)$ ,  $f(111011)$  and  $f(111111)$  are all zero. Applying Theorem 14 to the vector  $(001011)$ , we conclude that  $f$  has at least  $2^{6-3} = 8$  terms.

*Example 5.* Let  $f$  be a function on  $V_6$ , whose truth table is given as follows

1000110111110011001101011101100101111101011101111001011110011010

Note that  $f(000011)$  assumes the value zero, while  $f(000111)$ ,  $f(001011)$ ,  $f(001111)$ ,  $f(010011)$ ,  $f(010111)$ ,  $f(011011)$ ,  $f(011111)$ ,  $f(100011)$ ,  $f(100111)$ ,  $f(101011)$ ,  $f(101111)$ ,  $f(110011)$ ,  $f(110111)$ ,  $f(111011)$  and  $f(111111)$  all assume the value one. Applying (ii) of Corollary 15 to the vector  $(000011)$ , one can see that  $f$  has at least  $2^{6-2} + 1 = 17$  terms.

The lower bounds on the number of terms given by Theorem 14 and Corollary 15 are tight, due to Corollary 13 and Theorem 6.

## 5 Relating Algebraic Degree to Other Criteria

Note that the algebraic degree of any function, say  $f$ , on  $V_n$  is invariant under a non-singular linear transformation on the variables, and for any vector  $\alpha \in V_n$ , the subset  $W = \{\beta | \beta \preceq \alpha\}$  is a  $s$ -dimensional subspace, where  $s$  denotes the Hamming weight of  $\alpha$ . Using Theorem 6 it is not difficult to prove

**Theorem 16.** *Let  $f$  be a function on  $V_n$  ( $n \geq 2$ ). Then*

$$\deg(f) = \max\{\dim(U) \mid U \text{ is a subspace and Hamming weight of } f|_U \text{ is odd}\}.$$

The following lemma is called ‘‘Poisson Summation’’ whose proof can be found in [2].

**Lemma 17.** *Let real valued sequences  $a_0, \dots, a_{2^n-1}$  and  $b_0, \dots, b_{2^n-1}$  satisfy*

$$(a_0, \dots, a_{2^n-1})H_n = (b_0, \dots, b_{2^n-1}).$$

*Then for any  $p$ -dimensional subspace  $1 \leq p \leq n-1$ , say  $W$ ,*

$$\sum_{\alpha \in W} a_\alpha = 2^{p-n} \sum_{\alpha \in W^\perp} b_\alpha$$

*where  $W^\perp = \{\beta | \beta \in V_n, \langle \beta, \alpha \rangle = 0, \text{ for each } \alpha \in W\}$ .*

The next theorem shows a relationship between algebraic degree and Walsh-Hadamard transforms of a function.

**Theorem 18.** *Let  $f$  be a function on  $V_n$  ( $n \geq 2$ ),  $\xi$  be the sequence of  $f$ , and  $p$  is an integer,  $2 \leq p \leq n$ . If  $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{n-p+2}}$ , where  $\ell_j$  is the  $j$ th row (column) of  $H_n$ ,  $j = 0, 1, \dots, 2^n - 1$ , then  $\deg(f) \leq p - 1$ .*

*Proof.* Let  $\xi = (a_0, a_1, \dots, a_{2^n-1})$ . Note that

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

Then from Lemma 17

$$\sum_{\alpha \in W} a_\alpha = 2^{p-n} \sum_{\alpha \in W^\perp} \langle \xi, \ell_\alpha \rangle \tag{11}$$

holds for each  $p$ -dimensional subspace  $W$  of  $V_n$ , where  $W^\perp = \{\alpha | \alpha \in V_n, \langle \alpha, \beta \rangle = 0, \text{ for each } \beta \in W\}$  and  $a_\alpha = a_j$  if  $\alpha$  is the binary representation of integer  $j$ . From (11) and the condition that  $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{n-p+2}}$ ,  $j = 0, 1, \dots, 2^n - 1$ , we have  $\sum_{\alpha \in W} a_\alpha \equiv 0 \pmod{4}$ . Note that  $\xi = (a_0, a_1, \dots, a_{2^n-1})$  is the sequence of  $f$ . It is easy to verify that  $\sum_{\alpha \in W} a_\alpha \equiv 0 \pmod{4}$  if and only if the Hamming weight of  $f|_W$  is even.

Since  $W$  is an arbitrary  $p$ -dimensional subspace, using Theorem 16, the Hamming weight of the restriction of  $f$  to any  $q$ -dimensional subspace is even,  $q = p, p+1, \dots, n$ . So from Theorem 16, we have  $\deg(f) \leq p - 1$ .  $\square$

**Corollary 19.** *Let  $f$  be a function on  $V_n$  ( $n \geq 2$ ) and  $\xi$  be the sequence of  $f$ , and  $p$  is an integer,  $2 \leq p \leq n$ . If  $\Delta(\alpha) \equiv 0 \pmod{2^p}$ , for each  $\alpha \in V_n$ , then  $\deg(f) \leq n + 1 - \frac{1}{2}p$  for  $p$  even, and  $\deg(f) \leq n + 1 - \frac{1}{2}(p + 1)$  for  $p$  odd.*

*Proof.* From [6]

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

where  $\ell_j$  is the  $j$ th row (column) of  $H_n$ . Since  $\Delta(\alpha) \equiv 0 \pmod{2^p}$  for each  $\alpha \in V_n$ , we have  $\langle \xi, \ell_j \rangle^2 \equiv 0 \pmod{2^p}$  for  $j = 0, 1, \dots, 2^n - 1$ . Hence  $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{\frac{1}{2}p}}$  if  $p$  is even, and  $\langle \xi, \ell_j \rangle \equiv 0 \pmod{2^{\frac{1}{2}(p+1)}}$  if  $p$  is odd. Now the corollary follows from Theorem 18.  $\square$

We note that in Theorem 18,  $\langle \xi, \ell_j \rangle$  is closely related to nonlinearity [4], and in Corollary 19,  $\Delta(\alpha)$  is related to propagation characteristics [6].

## References

1. T. Jakobsen and L. Knudsen. The interpolation attack on block ciphers. In *Fast Software Encryption*, Lecture Notes in Computer Science, Berlin, New York, Tokyo, 1997. Springer-Verlag.
2. R. J. Lechner. *Harmonic Analysis of Switching Functions*. in Recent Developments in Switching Theory, cited by Amar Mukhopadhyay. Academic Press, New York, 1971.
3. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam, New York, Oxford, 1978.
4. J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearity and propagation characteristics of balanced boolean functions. *Information and Computation*, 119(1):1–13, 1995.
5. T. Shimoyama, S. Moriai, and T. Kaneko. Cryptanalysis of the cipher KN, May 1997. (presented at the rump session of Eurocrypt'97).
6. X. M. Zhang and Y. Zheng. Characterizing the structures of cryptographic functions satisfying the propagation criterion for almost all vectors. *Design, Codes and Cryptography*, 7(1/2):111–134, 1996. special issue dedicated to Gus Simmons.