The lower bound of second-order nonlinearity of a class of Boolean functions*

Luozhong GONG† and Shangzhao LI††, Nonmembers

SUMMARY The $r$-th nonlinearity of Boolean functions is an important cryptographic criterion associated with higher order linearity attacks on stream and block ciphers. In this paper, we tighten the lower bound of the second-order nonlinearity of a class of Boolean function over finite field $F_{2^n}$, $f_2(x) = Tr(\lambda x^d)$, where $\lambda \in F_{2^n}^*$, $d = 2 r + 2^r+1$ and $n = 7r$. This bound is much better than the lower bound of Iwata-Kurosawa.

key words: Boolean function; Higher-order nonlinearity; Higher-order derivative

1. Introduction

To resist the many kinds of crypt analysis, Boolean functions used in stream ciphers should have many good cryptographic properties: high algebraic degree, balancedness, high algebraic immunity and high nonlinearity etc. Now, many classes of Boolean functions with some good cryptographic properties have been constructed. In [1], [7], [10], [12], [14], [19], [26], many classes of Boolean functions achieving optimum algebraic immunity have been introduced. The Carlet-Feng functions have optimum algebraic degree, optimum algebraic immunity and higher nonlinearity [10], but it is not enough to resistance to fast correlation attack. In [25], the Tu-Ding functions are another class Boolean functions with optimum algebraic degree, optimum algebraic immunity and a provable good nonlinearity. However, they are also weak against fast algebraic attacks.

A characteristic of Boolean functions, called their nonlinearity profile, plays an important role with respect to the linear approximation attack of the cryptosystems in which they are involved. For every non-negative integer $r \leq n$, we denote the $nl_r(f)$ the minimum distance of $f$ and all functions of algebraic degrees at most $r$. The nonlinearity profile of a function $f$ is the sequence of those values $nl_r(f)$ for $r(1 \leq r \leq n - 1)$. In the case $r = 1$, we simply write $nl(f)$. Clearly, it is the minimum Hamming distance between the function $f$ and all affine functions over $F_{2^n}$, called the nonlinearity of $f$. Attributed to the $nl(f)$’s relation with Walsh transform, most research work so far has been theoretically and practically focused on $nl(f)$ [see [2], [8]]. However, computing the $nl_r(f)$ of a given function with algebraic degree strictly greater than $r$ is a hard task for $r > 1$, and, so far, few academic result has been achieved. Even proving lower bounds on the $nl_2(f)$ of functions is also a quite difficult task. In recently, Wang et al. [20] give an upper bound on the second-order of the hidden weighted bit function and Carlet (see [4]) introduce a new method for lower bounding the nonlinearity of a given function, which tell us how to derive a lower bound on the $r$-th order nonlinearity of a function $f$ from a lower bound on the $(r-1)$-th nonlinearity of at least one of the derivatives of $f$. Using this approach, G. Sun and C. Wu in [16], S. Gangopadhyay et al. in [22] and L. Gong and G. Fan in [18] recently also obtained the lower bounds of the second-order nonlinearity of several classes of Boolean functions.

Let $f(x) = x^{2 r + 2^r + 1}$ be a function defined on $F_{2^n}$, then $f$ has a low differential uniformity of four and higher $nl(tr(bf))$. So, it is an interesting problem whether its second-order nonlinearity is also high so that it can withstand the second-order affine approximation attack. When $n = 3r, 4r, 6r$, the lower bound of $nl(tr(bf))$ has been obtained [see [11], [16], [17]]. The present paper is engaged in deducing the lower bound of the second-order of nonlinearity of the above function with $n = 7r$.

2. Notation and Preliminaries

Let $F_2 = \{0, 1\}$ be the prime field of characteristic 2, $F_2^n$ be an $n$-dimensional vector space over $F_2$. Any mapping from $F_2^n$ to $F_2$ is called a Boolean function on $n$-variables. They play the core role in cryptography and error-correction coding. We denote by $B_n$ the set of the Boolean functions on $n$-vavriables. Any Boolean function is defined as

$$f(x_1, x_2, ..., x_n) = \bigoplus_{a=(a_1, a_2, ..., a_n) \in F_2^n} \mu_a(\prod_{i=1}^{n} x_i^{a_i}),$$

where $\mu_a \in F_2$ for all $a \in F_2^n$, which is called it’s algebraic normal form (ANF). Define $wt(a)$ the numbers of nozero components of vector $a$. The maximum value of

---

†The author is with the Faculty of the School of Mathematics, Changsha Normal University

†The author is with the Faculty of School of Mathematics and Statistics, Changshu Institute of Technology

*The paper was supported by the NNSFC (No.12071484, 11701046).
where $a_0,a_{2^n-1} \in F_2$, and $a_{2i(\text{mod}\ 2^n-1)} = a_i^2 \in F_2, 1 \leq i \leq 2^n - 2$. So the algebraic degree of the Boolean function

$$\text{deg}(f) = \max|a_j|, 0 \leq j \leq 2^n - 1,$$

where, given the 2-adic expansion $j = j_0 + j_1 2 + \cdots + j_{n-1} 2^{n-1}$, $j_i \in F_2, 0 \leq s \leq n-1$ and $a_2(j)$ denotes the number of all nonzero $j_i, 0 \leq s \leq n-1$. A Boolean function is affine if it has algebraic degree at most 1. The set of all affine functions is denoted by $A_n$.

Let $m|n, E = F_{2^m}$ and $L = F_{2^n}$. The function

$$tr_{L/E}(x) = \sum_{i=0}^{n-1} x^{2^{mi}}$$

called a trace function from $L$ to $E$. If $m = 1$, namely $E = F_2$, we denote $tr_{L/E}$ simply by $tr$ which is called the absolute trace function. The trace function has the following properties [21]:

(i) $tr_{L/E}(ax + by) = a tr_{L/E}(x) + b tr_{L/E}(y)$ for all $x, y \in L$ and $a, b \in E$.

(ii) $tr_{L/E}(x^q) = tr_{L/E}(x)$ for all $x \in L$ and $q = 2^m$.

(iii) Let $K$ be a finite field, $F$ be a finite extension of $K$, and $E$ be a finite extension of $F$, that is $K \subset F \subset E$. Then $tr_{E/K} = tr_{F/K}(tr_{E/F}(x))$ for all $x \in E$.

**Definition 2.1:** Let $f \in B_n$ and $a \in L = F_{2^n}$, we called

$$W_f(a) = \sum_{x \in L} (-1)^{f(x)} \chi(ax), a \in L,$$

the Walsh transform of $f$, where $\chi(x) = (-1)^{\text{tr}(x)}$ is the canonical additive character on $L$. The set $\{W_f(a) | a \in F_{2^n}\}$ is said to be the Walsh spectrum of $f$.

It is trivial to deduce that the relation between the nonlinearity and the Walsh spectrum is

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{a \in F_{2^n}} |W_f(a)|.$$

By Parseval’s equality, $\sum_{a \in F_{2^n}} W_f(a)^2 = 2^{2n}$, we have $nl(f) \leq 2^{n-1} - 2^{\frac{n-1}{2}}$. When $nl(f) = 2^{n-1} - 2^{\frac{n-1}{2}}$, $f$ is called a bent function. Obviously, it is possible for a bent function to exist when $n$ is even. Since the nonlinearity of bent functions reaches the maximum value, it can withstand the linear attack (to be more precise, linear approximation or affine approximation attack) to the most extent ([15]), and can also well withstand the correlation attack ([2], [9]).

**Definition 2.2:** We call the Boolean function $D_a f(x) = f(x) + f(x + a)$ for any $x \in F_{2^n}$ as the derivative of $f \in B_n$ with respect to $a \in F_{2^n}$, which is denoted by $D_a f$. Let $V$ be a $k$ dimensional subspace of $F_{2^n}$ generated by $\alpha_1, \alpha_2, \cdots, \alpha_k$, the $k$-th order derivative of $f \in B_n$ is defined by

$$D_V f(x) = D_{\alpha_1} f(x) \cdots D_{\alpha_k} f(x) = \sum_{u \in F_{2^n}} f(x + \sum_{i=1}^{k} u_i \alpha_i).$$

for any $x \in F_{2^n}$, which $u = \sum_{i=1}^{k} u_i \alpha_i$.

It is to be noted that when $\alpha_1, \alpha_2, \cdots, \alpha_k$, are not linearly independent, then $D_{\alpha_1} \cdots D_{\alpha_k} f$ is zero; otherwise, the set $\{x + \sum_{i=1}^{k} u_i \alpha_i | u \in F_{2^n}\}$ is a $k$-dimensional flat. Also, the $k$-th order derivative of $f$ depends only on the choice of the $k$ dimensional subspace $V$ and is independent of the choice of the basis of $V$. On the Galois field $F_{2^n}$, a cyclotomic coset $C_s$ is defined by $C_s = \{s, 2s, 2^2 s, \cdots, 2^{n-1} s\}$, where $n_s$ is the smallest positive integer such that $s \equiv 2^{n_s} s$ (mod $2^n - 1$). The subscript $s$ is chosen as the smallest integer in $C_s$, and $s$ is called the coset leader of $C_s$.

**Definition 2.3:** Let $q$ be a power of 2 and $V$ be an $n$-dimensional vector space over $F_q$. A map $Q : V \rightarrow F_q$ is called a quadratic form on $V$ if

(a) $Q(cx) = c^2 Q(x)$ for any $c \in F_q$ and $x \in V$;

(b) $B(x,y) := Q(x+y) + Q(x) + Q(y)$ is bilinear on $V$.

The kernel $K$ of a bilinear form $Q$ is the subspace of $V$ defined by $K = \{x \in V | B(x,y) = 0, \forall y \in V\}$.

The following lemmas are obtained from the definitions.

**Lemma 2.4:** ([3]) Let $V$ be a vector space over a field $F_q$ of characteristic 2 and $Q : V \rightarrow F_q$ be a quadratic form. Then the dimension of $V$ and the dimension of the kernel of $Q$ have the same parity

**Lemma 2.5:** ([3]) If $f : F_{2^n} \rightarrow F_2$ is a quadratic Boolean function, then the Walsh spectrum of $f$ depends only on the dimension $k$ of the kernel of $f$. More precisely, the Walsh spectrum of $f$ is:

<table>
<thead>
<tr>
<th>$f$</th>
<th>Number of $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$2^n - 2^{n-k}$</td>
</tr>
<tr>
<td>$2^{n-k}$</td>
<td>$2^{n-\frac{k-1}{2}} + (-1)^{f(0)} 2^{\frac{n-k-2}{2}}$</td>
</tr>
<tr>
<td>$-2^{n-k}$</td>
<td>$2^{n-\frac{k-1}{2}} - (-1)^{f(0)} 2^{\frac{n-k-2}{2}}$</td>
</tr>
</tbody>
</table>

**Lemma 2.6:** ([3]) Let $f$ be any quadratic Boolean function. The kernel of $f$ is the subspace of those $b$ such that the derivative $D_b f$ is constant. That is

$$E_f = \{b \in F_{2^n} | D_b f = \text{constant}\}$$
3. Main results

Lemma 3.1: ([4]) Let $f$ be any $n$-variable function and $r$ be a positive integer smaller than $n$. Then we have

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2} \sum_{a \in \mathbb{F}_2^n} nl_{r-1}(D_{af}). \quad (2)$$

Lemma 3.2: Let $f_3(x) = Tr(\lambda x^p)$ with $p = 2r + 2' + 1, \lambda \in \mathbb{F}_2^r$ and $n = 7r$. Then the dimension of the kernel of bilinear form associated to $D_a(f_3(x))$ is either 3 or 5r.

Proof: The derivative of $f_3(x) = Tr(\lambda x^p)$ with respect to $a \in \mathbb{F}_2^n$ is

$$D_a(f_3(x)) = f_3(x) + f_3(x+a)$$

$$= Tr(\lambda x^{2r+2' + 1}) + Tr(\lambda (x + a)^{2r+2' + 1})$$

$$= Tr(\lambda x^{2r+2' + 1} + \lambda x^{2r+2' + 1} + \lambda (x + a)^{2r+2' + 1} + \lambda (x + a)^{2r+2' + 1})$$

Since the Walsh spectrum is affine invariant, the Walsh spectrum of the function $D_a(f_3(x))$ is equal to the one of the function $G(x) = Tr(\lambda x^{2r+2' + 1} + \lambda x^{2r+2' + 1})$.

Notice that $2^r + 1$ and $2^{2r} + 1$ are not in the same cyclotomic coset, so $G(x) \neq 0$ and $G(x)$ is a quadratic Boolean function. By Lemma 2.5, the Walsh spectrum of $G(x)$ only depends on the dimension $k$ of the kernel of $G(x)$. By lemma 2.6, the kernel of $G(x)$ is the subspace of those $b$ such that the derivative of $B_b(G(x))$ is constant. Since

$$D_b(G(x)) = G(x) + G(x + b)$$

$$= Tr(\lambda (x^{2r+2' + 1} + \lambda x^{2r+2' + 1}))$$

$$= Tr(\lambda ((ab^{2r} + a^{2r} b)x^{2r} + (ab^{2r} + a^{2r} b)x^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r}))$$

$$= Tr(\lambda(ab^{2r} + a^{2r} b)x^{2r} + (ab^{2r} + a^{2r} b)x^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r}))$$

Clearly, $D_b(G(x))$ is constant if and only if

$$(a^{2r} b^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r}) + (a^{2r} b^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r}) = 0$$

That is

$$(a^{2r} + a^{2r})b^{2r} + a^{2r} b^{2r} + a^{2r} b^{2r} + (a^{2r} + a^{2r})b^{2r} = 0$$

Raising $2^{2r}$-th power to the both sides of above equation gives the following equation

$$(a^{2r} + a^{2r})b^{2r} + a^{2r} b^{2r} + ab^{2r} + (a^{2r} + a^{2r})b^{2r} = 0 \quad (4)$$

If $a \in \mathbb{F}_2^n$, Eq.(4) is equivalent to the equation $a^{2r} b^{2r} + ab^{2r} = 0$. This follows $b \in \mathbb{F}_2^{2r}$, and so $k = 3r$. Hence, we only consider the case when $a \notin \mathbb{F}_2^n$.

In this case, Eq.(4) is a $2'$-polynomial. Write $P(b) := (a^{2r} + a^{2r})b^{2r} + a^{2r} b^{2r} + ab^{2r} + (a^{2r} + a^{2r})b^{2r}$.

We are all know, the dimension of the kernel of $P(b)$ is $k + 1 = 0, 1, 2, 3, 4, 5$. Because $a \notin \mathbb{F}_2^n$, $l \neq 0, 1, 2, 3, 4, 5$.

Now consider the quadratic form from $F_q^2 = F_{q^2}(q = 2')$

$$Q(x) = Tr_{E/L}(\lambda (a x^{2r+2' + 1} + a^{2r} x^{2r+2' + 1} + a^{2r} x^{2r+2' + 1}))$$

where $L = F_{q^2}$ and $E = F_q$.

In fact, the kernel of $Q(x)$ is the set of those $b$ such that $B(x, b) = 0$ for all $x$, where

$$B(x, b) = Q(x) + Q(b) + Q(x + b) = Tr_{E/L}(x(P(b))^{2r}).$$

Since

$$B(x, b) = Tr_{E/L}(\lambda (a x^{2r+2' + 1} + a^{2r} x^{2r+2' + 1} + a^{2r} x^{2r+2' + 1}))$$

Therefore the set of roots of $P(b)$ is also the kernel of $Q(x)$. By lemma 2.4, the kernel of $Q(x)$ must have the same parity 7, so it is odd. Hence the dimension of the kernel of $Q(x)$ is 3 or 5, which implies the one of roots space of $P(b)$ is 3 or 5r, that is the dimension of the kernel of the bilinear form associated to $D_a(f_3(x))$ is either 3 or 5r.

Theorem 3.3: Let $f_3(x) = Tr(\lambda x^p)$ with $p = 2^{2r} + 2' + 1, \lambda \in \mathbb{F}_2^r$ and $n = 7r$. Then

$$nl_2(f_3(x)) \geq 2^{7r-1} - 2^{7r-1} \sqrt{2^{2r} + 2}.$$ 

Proof: From Lemma 3.2, the dimension of the kernel of the bilinear form associated to $D_a(f_3(x))$ is either 3 or 5r. Thus the Walsh transform of $D_a(f_3(x))$ at any point $a$ is $|W_{D_a(f_3(x))}| = 2^{\frac{7r}{2}}$ or $2^{\frac{3r-1}{2}}$. And then by

$$nl_2(f_3(x)) \geq 2^{7r-1} - 2^{7r-1} \sqrt{2^{2r} + 2}.$$
Eq.(1), we get
\[ nl(D_a(f_X(x))) = 2^{n-1} - \frac{1}{2} \sum_{a \in F_{2^r}} nl(D_a) \leq 2^{n-1} - 2^{5r-1}, \]
if \( a \in F_{2^r} \).
\[ nl(D_a(f_X(x))) \geq 2^{n-1} - \frac{1}{2} \sum_{a \in F_{2^r}} nl(D_a) \geq 2^{n-1} - 2^{5r-1}, \]
if \( a \notin F_{2^r} \). Therefore,
\[ \sum_{a \in F_{2^n}} nl(D_a) = \sum_{a \in F_{2^r}} nl(D_a) + \sum_{a \notin F_{2^r}} nl(D_a) \geq 2^{14r-1} + 2^{10r-1} - 2^{5r-1} \]
By lemma 3.1, we have
\[ nl(f) \geq 2^{n-1} - \frac{1}{2} \sqrt{2^{2n-2} - \sum_{a \in F_{2^n}} nl(D_a)} \]
\[ \geq 2^{7r-1} - \frac{1}{2} \sqrt{2^{13r} + 2^{7r}} \]
\[ \geq 2^{7r-1} - 2^{r-1} \sqrt{2^{7r} + 2^{r}}. \]

4. Conclusion remarks

By studying the lower bound of the nonlinearity of the derivatives of the functions, the present paper obtains the lower bound of the second-order nonlinearity of a class of Boolean functions. Results show that the second-order nonlinearity of the class of Boolean functions with high nonlinearity is also high. We compare our lower bound obtained in Theorem 3.3 with the lower bound obtained by Iwata-Kurosawa [24] in the following table. It is seen from the following table that our lower bound is much better than the lower bound of Iwata-Kurosawa. In this case, the lower bounds cannot be obtained by the relation between algebraic immunity and the r-th order nonlinearity as studied in [5], [6].

<table>
<thead>
<tr>
<th>r</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bound obtained in Theorem 3.3</td>
<td>4088</td>
<td>677803</td>
<td>1.0066304</td>
<td>1.92414</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \times 10^8 )</td>
<td>( \times 10^{12} )</td>
</tr>
<tr>
<td>Iwata-Kurosawa bound</td>
<td>3072</td>
<td>393216</td>
<td>5.0332</td>
<td>8.42633</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \times 10^7 )</td>
<td>( \times 10^{11} )</td>
</tr>
</tbody>
</table>

5. Acknowledgment

The authors are thankful to the anonymous reviewers whose comments have improved the technical as well as editorial quality of the paper.

References


Luozhong Gong received the B.S. and M.S. degrees in Applied Mathematics from School of Mathematics and statistics, Central South University of China in 2005 and 2010, respectively. During 2005-2016, he stayed in School of science, Hunan University of Science and Engineering of China to study Cryptology and Group Theory. He is now with Changsha Normal University.

Shangzhao Li received the B.S. and M.S. degrees in Applied Mathematics from Central South University of China and Suzhou University in 2008 and 2014, respectively. During 2008 to now, he stayed in Changshu Institute of Technology of China to study Group Theory.