Balanced Boolean Functions with Optimum Algebraic Immunity and High Nonlinearity

Xiangyong Zeng [*], Claude Carlet [†], Jinyong Shan and Lei Hu [‡]

October 19, 2010

Abstract: In this paper, three constructions of balanced Boolean functions with optimum algebraic immunity are proposed. The cryptographical properties such as algebraic degree and nonlinearity of the constructed functions are also analyzed.

Keywords: Boolean function, stream cipher, balancedness, algebraic degree, algebraic immunity, nonlinearity, fast algebraic attack

1 Introduction

Boolean functions used in stream ciphers should have good cryptographical properties such as balancedness, high algebraic degree, high algebraic immunity, high nonlinearity and good immunity to fast algebraic attacks [4]. These properties are required to resist many kinds of known attacks [4, 16]. The concept of algebraic immunity was proposed very recently [13, 21] and there are several constructions of Boolean functions with optimum algebraic immunity [14, 8, 3, 15, 18, 19, 5, 10]. However, most of the constructed functions can not be proven to have a high nonlinearity.

In 2008, the second author and Feng proposed an infinite class of balanced functions with optimum algebraic immunity as well as a high nonlinearity [9]. It is also checked that at least for small values of the number of variables, the functions of this class have much higher nonlinearity

[*]X. Zeng and J. Shan are with the Faculty of Mathematics and Computer Science, Hubei University, Wuhan 430062, Hubei, China. Email: xiangyongzeng@yahoo.com.cn

[†]C. Carlet is with the LAGA, Universities of Paris 8 and Paris 13 and CNRS; address: University of Paris 8, 2 rue de la liberté, 93526 Saint-Denis, Cedex, France. Email: claude.carlet@inria.fr

[‡]L. Hu is the State Key Laboratory of Information Security, Graduate University of the Chinese Academy of Sciences, Beijing 100049, China. Email: hu@is.ac.cn.

than the previously found classes and also have a good behavior against fast algebraic attacks [1, 12, 17]. To further improve the nonlinearity of balanced Boolean functions with optimum algebraic immunity, Tu and Deng proposed a conjecture about binary strings [26]. Under the assumption that this conjecture is true, a class of balanced Boolean functions in even number of variables was proven to have optimum algebraic immunity and a very good nonlinearity [26]. Based on the same assumption, another class of balanced Boolean functions in even number of variables was also presented and they have optimum algebraic immunity and a higher nonlinearity [25]. Unfortunately, these two classes of functions are constructed from Bent functions and have small distances to Bent functions which also leads to their bad resistance to fast algebraic attacks [6, 27]. In the more recent paper [28] two constructions of balanced Boolean functions with optimum algebraic immunity and high nonlinearity were introduced. The first one was proven in [7] to be the same as that of [9], and the functions in the second class are in odd number of variables and they can have optimal algebraic immunity under some condition which is not investigated further in the paper.

The purpose of this paper is to construct more balanced Boolean functions with optimum algebraic immunity as well as a high nonlinearity. For an integer $n \geq 5$, three constructions of balanced n-variable Boolean functions with optimum algebraic immunity are proposed. In the case of n being odd, a family $\mathcal{F}$ of balanced n-variable Boolean functions is investigated. The algebraic degree and nonlinearity of the constructed functions are analyzed. For some small values of the number of variables, we find some new Boolean functions which have the same algebraic degree and nonlinearity as the function in [9]. Further, some n-variable functions constructed in this paper have higher nonlinearity than the function in [9]. The behavior against fast algebraic attacks of some functions is also considered.

The remainder of this paper is organized as follows. In Section 2, we introduce some necessary notation and related results of Boolean functions. In Section 3, for odd n, two constructions of balanced n-variable functions with optimum algebraic immunity are proposed. In Section 4, for even n, one construction of balanced n-variable functions with optimum algebraic immunity is proposed. Section 5 concludes the study.

2 Preliminaries

For an integer n, let $\mathbb{F}_2^n$ be the n-dimensional vector space over $\mathbb{F}_2$, and $\mathbb{B}_n$ be the set of all n-variable Boolean functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. A basic representation for a Boolean function $f(x_1, \cdots, x_n)$ is given by its image vector (the last column in its truth table), namely the binary

string of length 2^n which lists all of its output values,

$$f = [f(0,0,\cdots,0), f(1,0,\cdots,0), f(0,1,\cdots,0), f(1,1,\cdots,0), \cdots, f(1,1,\cdots,1)]\,.$$

The *Hamming weight* of f, $\mathrm{wt}(f)$, is the Hamming weight of this string, or in other words, the size of the support set $\mathrm{supp}(f) = \{x \in \mathbb{F}_2^n \,|\, f(x) = 1\}$. We similarly define the zero set of f as $\mathrm{zero}(f) = \{x \in \mathbb{F}_2^n \,|\, f(x) = 0\}$. The *Hamming distance* $\mathrm{d_H}(f,g)$ between two Boolean functions f and g is the Hamming weight of their difference $f + g$ (by abuse of notation, we use $+$ to denote the addition on $\mathbb{F}_2$, i.e., the XOR). A Boolean function f is *balanced* if its image vector contains an equal number of ones and zeros, that is, if its Hamming weight equals to 2^{n-1}.

Any Boolean function has a unique representation as a multivariate polynomial over $\mathbb{F}_2$, called the *algebraic normal form* (ANF),

$$f(x_1,\cdots,x_n) = \sum_{q=(q_1,q_2,\cdots,q_n)\in\mathbb{F}_2^n} c(q) x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n},$$

where the coefficients $c(q)$'s are in $\mathbb{F}_2$. The *algebraic degree*, $\deg(f)$, is the number of variables in a highest order term with non zero coefficient. A Boolean function is affine if it is of algebraic degree at most 1. The set of all affine functions is denoted by $\mathbb{A}_n$.

Let α be a primitive element of the finite field $\mathbb{F}_{2^n}$. By identifying the finite field $\mathbb{F}_{2^n}$ with the vector space $\mathbb{F}_2^n$, a Boolean function $f(x)$ can be defined from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ as

$$\left[f(0), f(1), f(\alpha), \cdots, f(\alpha^{2^n-2})\right],$$

which is equivalent to the image vector. The Boolean function f over $\mathbb{F}_{2^n}$ has then another representation by a univariate polynomial

$$f(x) = \sum_{i=0}^{2^n-1} a_i x^i$$

where $a_0, a_{2^n-1} \in \mathbb{F}_2$, $a_i \in \mathbb{F}_{2^n}$ for $1 \le i < 2^n - 1$ such that $a_i = a_{2i \,(\mathrm{mod}\, 2^n-1)}$, and the addition is modulo 2. The algebraic degree $\deg(f)$ equals $\max\{\mathrm{wt}(i) \,|\, a_i \neq 0,\, 0 \le i < 2^n\}$ [9]. In this representation, the set $\mathbb{A}_n$ consists of all functions $\mathrm{Tr}(ax) + b$ where $a \in \mathbb{F}_{2^n}$, $b \in \mathbb{F}_2$ and $\mathrm{Tr}(\cdot)$ is the trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ defined by $\mathrm{Tr}(x) = x + x^2 + \cdots + x^{2^{n-1}}$.

Boolean functions used in a cryptographic system must have high nonlinearity to withstand linear and correlation attacks [16, 22]. The *nonlinearity* of an n-variable function f is its distance from the set of all n-variable affine functions, i.e.,

$$nl(f) = \min_{g\in\mathbb{A}_n}(\mathrm{d_H}(f,g)).$$

This parameter can be expressed by means of the Walsh transform. Let $x = (x_1, \cdots, x_n)$ and $\lambda = (\lambda_1, \cdots, \lambda_n)$ both belong to $\mathbb{F}_2^n$ and $\lambda \cdot x$ be an inner product in $\mathbb{F}_2^n$, e.g. $\lambda \cdot x = \lambda_1 x_1 + \cdots + \lambda_n x_n$. The *Walsh transform* of an n-variable Boolean function $f(x)$ is an integer valued function over $\mathbb{F}_2^n$ defined as

$$W_f(\lambda) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+\lambda \cdot x}.$$

Then the nonlinearity of f can be expressed as

$$nl(f) = 2^{n-1} - \tfrac{1}{2} \max_{\lambda \in \mathbb{F}_2^n} |W_f(\lambda)|.$$

A Boolean function f is balanced if and only if $W_f(0) = 0$. Any Boolean function should also have high algebraic degree to be cryptographically secure [16, 24].

For an n-variable Boolean function f, different scenarios related to low degree multiples of f have been studied in [13, 21]. This led to the following definition.

Definition 1: For $f \in \mathbb{B}_n$, define $AN(f) = \{g \in \mathbb{B}_n \mid f * g = 0\}$. Any function $g \in AN(f)$ is called an *annihilator* of f. The *algebraic immunity* of f is the minimum degree of all the nonzero annihilators of f and of all those of $f + 1$. We denote it by $AI(f)$.

To resist the standard algebraic attack, Boolean functions should have large algebraic immunity. In [13] it was proved that $AI(f) \leq \lceil \frac{n}{2} \rceil$ for any n-variable Boolean function f.

Lemma 1: ([11]) For odd n, if a balanced n-variable Boolean function f does not have any nonzero annihilator with algebraic degree $< \frac{n+1}{2}$, then $f + 1$ has no nonzero annihilator with algebraic degree $< \frac{n+1}{2}$. Consequently, $AI_n(f) = \frac{n+1}{2}$.

A high algebraic immunity is necessary but not sufficient condition for resistance against all kinds of algebraic attacks. If one can find g of low degree and $h \neq 0$ of reasonable degree such that $f * g = h$, then a fast algebraic attack (FFA) is feasible [1, 12, 17]. An n-variable function f can be considered as optimal with respect to FFAs if there do not exist two nonzero functions g and h such that $f * g = h$ and $\deg(g) + \deg(h) < n$ with $\deg(g) < n/2$.

For an integer s with $1 \leq s \leq 2^n - 2$, the cyclotomic coset containing s consists of

$$\left\{s, 2s, \cdots, 2^{n_s - 1} s\right\},$$

where n_s is the smallest positive integer such that $s \equiv 2^{n_s} s \pmod{2^n - 1}$, and we denote the cyclotomic coset by C_s. The smallest positive integer in the coset C_s is called the coset leader of C_s. Let $\Gamma(n)$ denote the set of all coset leaders modulo $2^n - 1$, and $m_i(x)$ denote the minimal polynomial of α^i over $\mathbb{F}_2$ for $1 \leq i \leq 2^n - 2$, where α is a primitive element of $\mathbb{F}_{2^n}$.

In the sequel, we recall some notation from [23]. The polynomial $R_d(x)$ is defined as the product of the minimal polynomials over $\mathbb{F}_2$ of all elements $\alpha^{-i} \in \mathbb{F}_{2^n}$, where $wt(i) = d$ and

$i \in \Gamma(n)$:

$$R_d(x) = \prod_{i \in \Gamma(n),\, wt(i)=d} m_{2^n-1-i}(x) = \prod_{wt(j)=n-d} (x - \alpha^j)$$

for $1 \leq d \leq n-1$, $R_n(x) = x+1$ and $R_0(x) = x$. For $0 \leq d_1 \leq d_2 \leq n$, let

$$R_{d_1,d_2}(x) = \prod_{i=d_1}^{d_2} R_i(x).$$

Then for $d_1 < d_2$ and $E = \sum_{i=d_1+1}^{d_2} \binom{n}{i}$, the polynomial R_{d_1+1,d_2} has the form

$$R_{d_1+1,d_2} = 1 + r_1 x + r_2 x^2 + \cdots + r_{E-1} x^{E-1} + x^E \tag{1}$$

where $\deg(R_{d_1+1,d_2}) = E$. Let us denote $D_1 = \sum_{i=0}^{d_1} \binom{n}{i}$ and $D_2 = \sum_{i=0}^{d_2} \binom{n}{i}$, and define the $D_1 \times D_2$ matrix $\mathbf{R}_{d_1+1,d_2}$ such that its r-th row consists of the coefficients of the polynomial

$$x^r R_{d_1+1,d_2}(x) = x^r \sum_{j=0}^{E} r_j x^j$$

for $0 \leq r < \sum_{i=0}^{d_1} \binom{n}{i}$, appended with zeros, i.e.,

$$\mathbf{R}_{d_1+1,d_2} = \begin{pmatrix} r_0 & r_1 & \cdots & r_E & 0 & \cdots & 0 \\ 0 & r_0 & \cdots & r_{E-1} & r_E & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & \vdots & \vdots & \cdots & r_E \end{pmatrix} \tag{2}$$

where $r_0 = r_E = 1$. Note that this is a generator matrix of a binary cyclic code since $R_{d_1,d_2}(x)$ divides $x^{2^n-1} - 1$. Here we will enumerate the columns of the matrix $\mathbf{R}_{d+1,n-1}$ from the 0-th column to the (2^n-2)-th column where $1 \leq d < n-1$. Let $\mathbf{R}^{1_f}_{d+1,n-1}$ be the sub-matrix of $\mathbf{R}_{d+1,n-1}$ such that the j-th column of $\mathbf{R}_{d+1,n-1}$ belongs to $\mathbf{R}^{1_f}_{d+1,n-1}$ if $\alpha^j \in \mathrm{supp}(f)$, where $0 \leq j \leq 2^n - 2$. Similarly, $\mathbf{R}^{0_f}_{d+1,n-1}$ is the sub-matrix of $\mathbf{R}_{d+1,n-1}$ that consists of the j-th column, such that $\alpha^j \in \mathrm{zero}(f)$, for all $0 \leq j \leq 2^n - 2$. Thus, the $2^n - 1$ columns of $\mathbf{R}_{d+1,n-1}$ are divided into two disjoint subsets.

For $1 \leq d_1 < d_2 \leq n-1$, let the matrix $\mathbf{R}_{d_1+1,d_2} = \left[\boldsymbol{\eta}_0, \boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_{D_2-1}\right]$, where the vector $\boldsymbol{\eta}_i$ with D_1 components denotes the i-th column of $\mathbf{R}_{d_1+1,d_2}$ for $0 \leq i < D_2$. For $0 \leq j < D_1$, let $\boldsymbol{\eta}_i^{(j)}$ denote the j-th component of $\boldsymbol{\eta}_i$. If there is a non-negative integer l_1 such that $\boldsymbol{\eta}_i^{(l_1)} = 1$

and $\boldsymbol{\eta}_i^{(j)} = 0$ for all $0 \le j < l_1$ (if there exists such integer j), we define $\mathcal{N}^0(\boldsymbol{\eta}_i) = l_1$. Similarly, we can define $\mathcal{N}_0(\boldsymbol{\eta}_i) = l_2$ if there exists a non-negative integer l_2 such that $\boldsymbol{\eta}_i^{(D_1-1-l_2)} = 1$ and $\boldsymbol{\eta}_i^{(D_1-j-1)} = 0$ for all $0 \le j < l_2$ (if there exists such integer j). In particular, $\mathcal{N}^0(\boldsymbol{\eta}_i) = 0$ if $\boldsymbol{\eta}_i^{(0)} = 1$, and $\mathcal{N}_0(\boldsymbol{\eta}_i) = 0$ if $\boldsymbol{\eta}_i^{(D_1-1)} = 1$. With this notation, we can classify the vectors $\boldsymbol{\eta}_i$ according to the values $\mathcal{N}^0(\boldsymbol{\eta}_i)$ and $\mathcal{N}_0(\boldsymbol{\eta}_i)$ for $0 \le i < D_2$. For each j with $0 \le j \le D_1 - 1$, two sets S^j and S_j are defined as

$$S^j = \left\{ i \,|\, \mathcal{N}^0(\boldsymbol{\eta}_i) = j,\ K_1 \le i \le K_2 \right\} \tag{3}$$

and

$$S_j = \left\{ i \,|\, \mathcal{N}_0(\boldsymbol{\eta}_i) = j,\ K_3 \le i \le K_4 \right\}, \tag{4}$$

where the integers K_1, K_2, K_3 and K_4 will be given in Sections 3 and 4. Let $J^0 = \left\{ j \,|\, S^j \neq \varnothing \right\}$, and J^1 be a subset of J^0. For each $j \in J^1$, take exactly an integer i^j from the set S^j and define the set

$$I^0 = \left\{ i^j \,|\, j \in J^1 \right\}. \tag{5}$$

Similarly, we can define $J_0 = \{ j \,|\, S_j \neq \varnothing \}$, and J_1 is a subset of J_0 such that $S_j \backslash I^0 \neq \varnothing$ for all $j \in J_1$. For each $j \in J_1$, take exactly an integer i_j from the set $S_j \backslash I^0$ and define the set

$$I_0 = \left\{ i_j \,|\, j \in J_1 \right\}. \tag{6}$$

Define the sets W_0, Y_0, W^0, and Y^0 as

$$\begin{array}{c} W_0 = \{ i \,|\, 0 \le i \le D_1 - 1 \},\ Y_0 = \{ D_1 - 1 - j \,|\, j \in J_1 \} \\ W^0 = \{ i \,|\, D_2 - D_1 \le i \le D_2 - 1 \},\ Y^0 = \left\{ D_2 - D_1 + j \,|\, j \in J^1 \right\}. \end{array} \tag{7}$$

Lemma 2: (Theorem 1, [23]) Let $f(x) = \sum_{i=0}^{2^n-1} f_i x^i$ be the univariate representation of a Boolean function f and let $F(x) = \sum_{i=0}^{2^n-2} f(\alpha^i) x^i$.

(i) For $1 \le d < n-1$, $\deg(f) = d$ if and only if

$$R_{d+1,n-1}(x) | F(x), \quad R_n(x) | F(x) + f(0), \quad \text{and} \quad R_d(x) \nmid F(x).$$

(ii) For $d = n-1$, $\deg(f) = d$ if and only if

$$R_n(x) | F(x) + f(0) \quad \text{and} \quad R_d(x) \nmid F(x).$$

Lemma 3: (Theorem 4, [23]) There is an annihilator $g \in AN(f)$ with $\deg(g) \le d < n$, if and only if $\delta_g(d) > 0$, where

$$\delta_g(d) = \sum_{i=0}^{d} \binom{n}{i} - \operatorname{rank}\left(\mathbf{R}_{d+1,n-1}^{1_f} \right).$$

Lemma 4: (Theorem 5, [23]) There is an annihilator $h \in AN(f+1)$ with $\deg(h) \leq d < n$, if and only if, $\delta_h(d) > 0$, where

$$\delta_h(d) = \sum_{i=0}^{d} \binom{n}{i} - \text{rank}\left(\left[\boldsymbol{\gamma}_{1_f}(d) \quad \mathbf{R}^{0_f}_{d+1,n-1}\right]\right),$$

where $\boldsymbol{\gamma}_{1_f}(d) = \mathbf{R}^{1_f}_{d+1,n-1} \cdot \mathbf{1}^T_{|\text{supp}(f)|}$, and $\mathbf{1}^T_{|\text{supp}(f)|}$ is the transpose of the all ones vector with length $|\text{supp}(f)|$.

3 Two constructions of n-variable Boolean functions for odd n

In this section, the integer n is always assumed to be odd. We also assume that $d_1 = \frac{n-1}{2}$ and $d_2 = n-1$ in (1). Thus, $D_1 = 2^{n-1}$, $D_2 = 2^n - 1$ and $E = 2^{n-1} - 1$. Let $K_1 = 0$, $K_2 = D_1 - 2$, $K_3 = D_1$ and $K_4 = D_2 - 1$. We propose two constructions of balanced n-variable Boolean functions with optimum algebraic immunity. These functions can have optimum algebraic degree and high nonlinearity.

Construction 1: Let $f \in \mathbb{B}_n$ such that

$$\text{supp}(f) = \left\{\alpha^i \,|\, i \in (W_0 \backslash Y_0) \bigcup I_0\right\}, \tag{8}$$

where W_0, Y_0 and I_0 are given by (6) and (7).

Similarly, we have the following construction.

Construction 2: Let $f \in \mathbb{B}_n$ such that

$$\text{supp}(f) = \left\{\alpha^i \,|\, i \in (W^0 \backslash Y^0) \bigcup I^0\right\}, \tag{9}$$

where W^0, Y^0 and I^0 are given by (5) and (7).

Applying Lemmas 1 and 3, the balancedness and algebraic immunity of functions in Constructions 1 and 2 can obtained as below.

Theorem 1: The Boolean functions defined in Constructions 1 and 2 are balanced and have algebraic immunity $\frac{n+1}{2}$.

Proof: From Constructions 1 and 2, a function f defined by (8) or (9) is balanced. In the sequel, we will prove that every function f in Construction 1 has algebraic immunity $\frac{n+1}{2}$. The case for Construction 2 can be similarly proven and its proof is omitted.

Let $g \in AN(f)$. According to Construction 1, we have $\mathbf{R}^{1_f}_{\frac{n+1}{2},n-1} = (\boldsymbol{\eta}_i)_{i \in (W_0 \backslash Y_0) \bigcup I_0}$. By swapping the $(D_1 - 1 - j)$-th and i_j-th columns of the matrix $\mathbf{R}_{\frac{n+1}{2},n-1}$ for all $j \in J_1$, we obtain

a matrix and its sub-matrix consisting of i-th column with $0 \le i < D_1$ is an upper triangular matrix since

$$\mathcal{N}_0(\boldsymbol{\eta}_{D_1-1-j}) = \mathcal{N}_0(\boldsymbol{\eta}_{i_j}) = j,$$

where $j \in J_1$. Moreover, every entry in the main diagonal of this upper triangular matrix is 1, then it is invertible. By (8), the above upper triangular matrix can be obtained from the matrix $(\boldsymbol{\eta}_i)_{i\in(W_0\setminus Y_0)\bigcup I_0}$ only by a series of elementary column transformations. This shows the matrix $\mathbf{R}^{1_f}_{\frac{n+1}{2},n-1}$ has full rank, i.e., $\delta_g(\frac{n-1}{2}) = 0$. Thus $\deg(g) \ge \frac{n+1}{2}$ by Lemma 3.

Therefore, the algebraic immunity of Construction 1 is $\frac{n+1}{2}$ since n is odd by Lemma 1. □

3.1 A family of Boolean functions given by Construction 1 or 2

Before proposing a family of Boolean functions, we give a relation between the coefficients r_1 and r_{E-1} of the polynomial $R_{\frac{n+1}{2},n-1}$ in (1) as below.

Lemma 5: $r_1 + r_{E-1} = 1$.

Proof: By Vieta's Theorem, we have that $\prod\limits_{1\le wt(i)\le\frac{n-1}{2}} \alpha^i = 1$, $r_{E-1} = \sum\limits_{1\le wt(i)\le\frac{n-1}{2}} \alpha^i$ and

$$r_1 = \sum_{1\le wt(i)\le\frac{n-1}{2}} \left(\prod_{j\ne i,\ 1\le wt(j)\le\frac{n-1}{2}} \alpha^j \right) = \sum_{1\le wt(i)\le\frac{n-1}{2}} \alpha^{-i} = \sum_{\frac{n+1}{2}\le wt(i)\le n-1} \alpha^i.$$

Thus,

$$r_{E-1} + r_1 = \sum_{1\le wt(i)\le\frac{n-1}{2}} \alpha^i + \sum_{\frac{n+1}{2}\le wt(i)\le n-1} \alpha^i = \sum_{i=1}^{2^n-2} \alpha^i = 1 + \sum_{i=0}^{2^n-2} \alpha^i = 1. \quad \square$$

Lemma 5 shows $\{r_1, r_{E-1}\} = \{0, 1\}$. Thus, there exists an integer k_1 such that $r_i = 0$ for all $1 \le i \le k_1$ and $r_{k_1+1} = 1$, if $r_1 = 0$. Otherwise, there is an integer k_2 such that $r_{E-i} = 0$ for all $1 \le i \le k_2$ and $r_{E-k_2-1} = 1$.

Define a set J as

$$J = \begin{cases} \{1, 2, \cdots, k_1\}, & \text{if } r_1 = 0; \\ \{1, 2, \cdots, k_2\}, & \text{if } r_{E-1} = 0. \end{cases} \tag{10}$$

By Lemma 5, the set J is non-empty.

Let I be a subset of J. If $r_{E-1} = 0$, define the sets S, U and V as

$$S = \left\{0, 1, 2, \cdots, 2^{n-1} - 1\right\}, \quad U = \left\{2^n - 2 - i \,|\, i \in I\right\}, \quad V = \left\{2^{n-1} - 1 - i \,|\, i \in I\right\}. \tag{11}$$

If $r_1 = 0$, the sets are defined as

$$S = \left\{2^{n-1} - 1, 2^{n-1}, \cdots, 2^n - 2\right\}, \quad U = \{i \,|\, i \in I\}, \quad V = \left\{2^{n-1} - 1 + i \,|\, i \in I\right\}. \tag{12}$$

With the above preparations, we can present a family of Boolean functions.

Family $\mathcal{F}$: The family $\mathcal{F}$ consists of all Boolean functions $f \in \mathbb{B}_n$ with the support set as

$$\mathrm{supp}(f) = \left\{\alpha^i \,|\, i \in (S \backslash V) \bigcup U\right\} \tag{13}$$

where the sets S, U and V are defined in (11) for $r_{E-1} = 0$ and (12) for $r_1 = 0$.

When $r_{E-1} = 0$, take $W_0 = S$, $Y_0 = V$ and $I_0 = U$ and then we have that the family $\mathcal{F}$ is contained in Construction 1. Similarly, we have that the family $\mathcal{F}$ is contained in Construction 2 when $r_1 = 0$. Thus, the functions in $\mathcal{F}$ are balanced and have optimum algebraic immunity.

It is proven in [7] that the first construction of [28] is exactly the same as the construction of [9]. Thus, in order to make a comparison of the above family and the known ones in [9, 28], it is sufficient to consider the construction in [9]. Let f_1 be the function with

$$\mathrm{supp}(f_1) = \left\{0, 1, \alpha, \cdots, \alpha^{2^{n-1}-2}\right\} \tag{14}$$

in [9]. Then $f_1(x) = f_\varnothing(\alpha^{2^{n-1}}x) + 1$ for $r_{E-1} = 0$ and $f_1(x) = f_\varnothing(x) + 1$ for $r_1 = 0$, where $f_\varnothing(x)$ is the function in $\mathcal{F}$ corresponding to $I = \varnothing$. Therefore, some cryptographical properties of the function $f_\varnothing$ have been studied in [9] and [28]. In the following subsections, we are interested in studying the cryptographical properties of the functions of $\mathcal{F}$ for $\varnothing \neq I \subseteq J$. The number of these functions is $2^{|J|} - 1$.

Applying Lemma 2, we can characterize the algebraic degree of these functions as follows.

Proposition 1: A Boolean function f in the family $\mathcal{F}$ has algebraic degree $n-1$ if and only if $\sum\limits_{x \in \mathrm{supp}(f)} x \neq 0$. In particular, if there is a non-empty proper subset of J given by (10), then there always exists a Boolean function f in the family $\mathcal{F}$ such that $\deg(f) = n - 1$.

Proof: Notice that $|\mathrm{supp}(f)| = 2^{n-1}$ is even and $R_n(x) = x+1$. Consequently, $F(1)+f(0) = \sum\limits_{i=0}^{2^n-2} f(\alpha^i) + f(0) = 0$ and then $R_n(x) \,|\, F(x) + f(0)$, where $F(x)$ is given in Lemma 2. Thus, $\deg(f) = n - 1$ if and only if $R_{n-1}(x) \nmid F(x)$, i.e., $F(\alpha) = \sum\limits_{x \in \mathrm{supp}(f)} x \neq 0$ by Lemma 2 (ii).

If f is a function in $\mathcal{F}$ with $\deg(f) < n - 1$, then $\sum\limits_{x \in \mathrm{supp}(f)} x = 0$. Let I be a non-empty proper subset of J. Take an element i in $J \setminus I$, and replace an arbitrary element i' in I with i, then a function f' can be constructed from (13) by taking a subset $I' = (\{i\} \bigcup I) \setminus \{i'\}$ of J in

(11) or (12). Then, we have

$$\begin{aligned}\sum_{x\in \mathrm{supp}(f')} x &= \sum_{x\in \mathrm{supp}(f)} x + \alpha^{2^{n-1}-1-i} + \alpha^{2^n-2-i} + \alpha^{2^{n-1}-1-i'} + \alpha^{2^n-2-i'} \\ &= (1+\alpha^{2^{n-1}-1})(\alpha^{2^{n-1}-1-i} + \alpha^{2^{n-1}-1-i'}) \neq 0\end{aligned}$$

for $r_{E-1} = 0$ and

$$\sum_{x\in \mathrm{supp}(f')} x = (1+\alpha^{2^{n-1}-1})(\alpha^i + \alpha^{i'}) \neq 0$$

for $r_1 = 0$.

Thus, $\deg(f') = n-1$. $\square$

By Proposition 1, there always exists a Boolean function in the family $\mathcal{F}$ such that $\deg(f) = n-1$, if $|J| > 1$.

3.2 The nonlinearity of Boolean functions in $\mathcal{F}$

In this subsection, we will give the lower bound on the nonlinearity of the functions in the family $\mathcal{F}$, based on the method developed in [9] and [28]. The following lemmas from calculus will be used to measure the nonlinearity.

Lemma 6: For $0 \le x \le \frac{1}{2}$, $\sin(\pi x) \ge 3x - 2x^2$.

Lemma 7: $1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2^{n-1}-1} < \frac{n-1}{2}\ln 2 + 1$.

Theorem 2: For odd $n \ge 5$, the nonlinearity of the functions f in the family $\mathcal{F}$ satisfies

$$nl(f) > 2^{n-1} - \left(\frac{\ln 2}{3}(n-1) + \frac{5}{6} + \frac{1}{3\sqrt{3}} + \frac{1}{6\sqrt{2}}\right) 2^{\frac{n}{2}} - 2|I| - 1$$

where the set I is a subset of J given by (10).

Proof: For $r_{E-1} = 0$ and $\lambda \in \mathbb{F}_{2^n}^*$, the Walsh transform of the Boolean function f satisfies

$$\begin{aligned}W_f(\lambda) &= \sum_{i=0}^{2^n-2} (-1)^{f(\alpha^i)+\mathrm{Tr}(\lambda\alpha^i)} + 1 \\ &= -\sum_{i\in (S\setminus V)\bigcup U} (-1)^{\mathrm{Tr}(\lambda\alpha^i)} + \sum_{i\notin (S\setminus V)\bigcup U} (-1)^{\mathrm{Tr}(\lambda\alpha^i)} + 1 \\ &= -2\sum_{i\in (S\setminus V)\bigcup U} (-1)^{\mathrm{Tr}(\lambda\alpha^i)}\end{aligned}$$

and then

$$|W_f(\lambda)| \le 2\left|\sum_{i=0}^{2^{n-1}-1} (-1)^{\mathrm{Tr}(\lambda\alpha^i)}\right| + 4|I|. \tag{15}$$

Let $\lambda = \alpha^t$ and $\xi = e^{\frac{2\pi\sqrt{-1}}{2^n-1}}$. Let $\psi(\alpha^j) = \xi^j$ with $0 \le j \le 2^n - 2$ be a multiplicative character of $\mathbb{F}_{2^n}^*$, $\chi_1(x) = (-1)^{\mathrm{Tr}(x)}$ be the canonical additive character of $\mathbb{F}_{2^n}$, and the Gaussian sum $G(\psi, \chi_1)$ be defined by

$$G(\psi, \chi_1) = \sum_{i=0}^{2^n-2} \psi(\alpha^i)\chi_1(\alpha^i). \tag{16}$$

Notice that for $0 \le i \le 2^n - 2$,

$$\chi_1(\alpha^i) = (-1)^{\mathrm{Tr}(\alpha^i)} = \frac{1}{2^n-1} \sum_{j=0}^{2^n-2} G(\overline{\psi}^j, \chi_1)\psi^j(\alpha^i) \tag{17}$$

where the bar denotes complex conjugation (see [20], p. 195). By (15) and (17), we have

$$|W_f(\lambda)| \le \frac{2}{2^n-1} \left| \sum_{j=1}^{2^n-2} G(\overline{\psi}^j, \chi_1)\xi^{jt} \frac{1-\xi^{j2^{n-1}}}{1-\xi^j} \right| + 4|I| + \frac{2^n}{2^n-1}. \tag{18}$$

Notice that $\left|G(\overline{\psi}^j, \chi_1)\right| = 2^{\frac{n}{2}}$ for all $1 \le j \le 2^n - 2$ [20], and

$$\left| \frac{1-\xi^{j2^{n-1}}}{1-\xi^j} \right| = \left| \frac{\xi^{-j2^{n-2}} - \xi^{j2^{n-2}}}{\xi^{-j/2} - \xi^{j/2}} \right| = \left| \frac{\sin \frac{j\pi 2^{n-1}}{2^n-1}}{\sin \frac{j\pi}{2^n-1}} \right|.$$

Consequently, by (18) we have

$$|W_f(\lambda)| \le \frac{2^{\frac{n}{2}+2}}{2^n-1} \left(\sum_{j=1}^{2^{n-2}} \frac{1}{\sin \frac{(2j-1)\pi}{2^n-1}} + \sum_{j=1}^{2^{n-2}-1} \frac{1}{2\cos \frac{j\pi}{2^n-1}} \right) + 4|I| + \frac{2^n}{2^n-1}. \tag{19}$$

By Lemmas 6 and 7, we have

$$\sum_{j=1}^{2^{n-2}} \frac{1}{\sin \frac{(2j-1)\pi}{2^n-1}} \le (2^n-1)\left(\frac{n-1}{6}\ln 2 + \frac{5}{12} \right). \tag{20}$$

For the other summation in (19), we have

$$\sum_{j=1}^{2^{n-2}-1} \frac{1}{2\cos \frac{j\pi}{2^n-1}} < \left(\frac{1}{6\sqrt{3}} + \frac{1}{12\sqrt{2}} \right)(2^n-1). \tag{21}$$

By (19), (20) and (21), we have

$$|W_f(\lambda)| \le 2^{\frac{n}{2}+2}\left(\frac{n-1}{6}\ln 2 + \frac{5}{12} + \frac{1}{6\sqrt{3}} + \frac{1}{12\sqrt{2}} \right) + 4|I| + \frac{2^n}{2^n-1}$$

for all $\lambda \in \mathbb{F}_{2^n}^*$. Notice that $W_f(0) = 0$. Therefore,

$$nl(f) > 2^{n-1} - \left(\frac{\ln 2}{3}(n-1) + \frac{5}{6} + \frac{1}{3\sqrt{3}} + \frac{1}{6\sqrt{2}} \right) 2^{\frac{n}{2}} - 2|I| - 1.$$

For $r_1 = 0$, the conclusion can be obtained by a same analysis and this finishes the proof. □

Since the function f in the family $\mathcal{F}$ and $f_\varnothing$ (introduced in Section 3.1) take different values only in $2|I|$ elements of $\mathbb{F}_{2^n}$, we can establish a rough relation between the nonlinearities $nl(f_1)$ and $nl(f)$ as below.

Proposition 2: For odd n and a function f in the family $\mathcal{F}$, $|nl(f) - nl(f_1)| \leq 2|I|$ where the set I is a subset of J given by (10).

By Theorem 2 and Proposition 2, we have the following corollary.

Corollary 1: For odd $n \geq 5$, the nonlinearity of the functions f in the family $\mathcal{F}$ satisfies

$$nl(f) \geq \max\left\{2^{n-1} - \left(\frac{\ln 2}{3}(n-1) + \frac{5}{6} + \frac{1}{3\sqrt{3}} + \frac{1}{6\sqrt{2}}\right)2^{\frac{n}{2}} - 2|I| - 1,\ nl(f_1) - 2|I|\right\}$$

where the set I is a subset of J given by (10).

Notice that a function f in Construction 1 and $f_\varnothing$ in the family $\mathcal{F}$ take different values exactly in $2|I_0|$ elements of $\mathbb{F}_{2^n}$. By a similar analysis as in Theorem 2, Proposition 2 and the fact $nl(f_\varnothing) = nl(f_1)$, we have a lower bound of the nonlinearity of f as

$$nl(f) \geq \max\left\{2^{n-1} - \left(\frac{\ln 2}{3}(n-1) + \frac{5}{6} + \frac{1}{3\sqrt{3}} + \frac{1}{6\sqrt{2}}\right)2^{\frac{n}{2}} - 2|I_0| - 1,\ nl(f_1) - 2|I_0|\right\}.$$

Similarly, for a function f in Construction 2, we have

$$nl(f) \geq \max\left\{2^{n-1} - \left(\frac{\ln 2}{3}(n-1) + \frac{5}{6} + \frac{1}{3\sqrt{3}} + \frac{1}{6\sqrt{2}}\right)2^{\frac{n}{2}} - 2|I^0| - 1,\ nl(f_1) - 2|I^0|\right\}.$$

4 The construction of n-variable Boolean functions for even n

In this section, the integer n is always assumed to be even. Here we also assume that $d_1 = \frac{n-2}{2}$ and $d_2 = n - 1$. Then $D_1 = \sum_{i=0}^{\frac{n-2}{2}} \binom{n}{i}$, $D_2 = 2^n - 1$ and $E = \sum_{i=1}^{\frac{n}{2}} \binom{n}{i}$. Let $K_1 = K_3 = D_1$ and $K_2 = K_4 = 2^n - 2 - D_1$.

With the above preparations, we propose the following construction for even n.

Construction 3: Let $f \in \mathbb{B}_n$. The support and zero sets of f satisfy that

$$\mathrm{supp}(f) \supset \left\{\alpha^i \,|\, i \in (W_0 \backslash Y_0) \bigcup I_0\right\} \quad \text{and} \quad \mathrm{zero}(f) \supset \left\{\alpha^i \,|\, i \in (W^0 \backslash Y^0) \bigcup I^0\right\}, \tag{22}$$

where W_0, Y_0, I_0, W^0, Y^0, and I^0 are given by (5), (6) and (7).

Applying Lemmas 3 and 4, we can have the following result by a similar analysis as in Theorem 1.

Theorem 3: The Boolean functions defined in Construction 3 have algebraic immunity $\frac{n}{2}$.

By choosing suitable support set satisfying (22) and having cardinality 2^{n-1}, the Boolean function f in Construction 3 is balanced.

The existence of functions with optimal algebraic degree in Construction 3 is considered in the following proposition. It can be similarly proven as Proposition 1.

Proposition 3: For the sets W_0, Y_0, I_0, W^0, Y^0 and I^0 defined by (5), (6) and (7), there always exists a balanced Boolean function in Construction 3 whose algebraic degree is $n-1$.

For a balanced function f in Construction 3, let

$$L = \left\{\alpha^i \,|\, 0 \le i < 2^{n-1}\right\} \backslash \mathrm{supp}(f). \tag{23}$$

A similar analysis as in Section 3.2 gives the following result.

Theorem 4: The nonlinearity of a balanced function f in Construction 3 satisfies

$$nl(f) \ge \max\left\{2^{n-1} - \left(\frac{\ln 2}{3}(n-1) + \frac{5}{6} + \frac{1}{3\sqrt{3}} + \frac{1}{6\sqrt{2}}\right) 2^{\frac{n}{2}} - 2|L| - 1,\ nl(f_1) - 2|L|\right\},$$

where L is the set defined by (23).

5 Conclusion

We proposed three constructions of balanced Boolean functions with optimum algebraic immunity. These constructions provide a class of Boolean functions with optimal algebraic degree and high nonlinearity. It is also checked that for $5 \le n \le 19$, some new n-variable Boolean functions constructed in this paper have the same algebraic degree and nonlinearity as the function f_1 in [9]. Further, we also found some balanced functions with optimal algebraic degree, optimum algebraic immunity and higher nonlinearity than f_1. Experiments show that some functions in the proposed class have a strong immunity against FAA's. More details will be given in a full version of this paper.

Acknowledgement
The authors are indebted to Simon Fischer for his evaluation of the resistance of the functions to fast algebraic attacks.

References

[1] F. Armknecht, "Improving fast algebraic attacks," in Fast Software Encryption 2004, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3017, pp. 65-82, 2004.

[2] F. Armknecht, C. Carlet, P. Gaborit, S. Künzli, W. Meier and O. Ruatta. "Efficient computation of algebraic immunity for algebraic and fast algebraic attacks," in Advanceds in Cryptology-EUROCRYPT 2006, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 4004, pp. 147-164, 2006.

[3] A. Braeken and B. Preneel, "On the algebraic immunity of symmetric Boolean functions," in Progress in Cryptology-INDOCRYPT 2004, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3797, pp. 35-48, 2005.

[4] C. Carlet, "Boolean Functions for Cryptography and Error Correcting Codes", Chapter of the monography "Boolean Models and Methods in Mathematics, Computer Science, and Engineering", Cambridge University Press (Peter Hammer and Yves Crama editors), pp. 257-397, 2010.

[5] C. Carlet, "A method of construction of balanced functions with optimum algebraic immunity," Cryptology ePrint Archive. http://eprint.iacr.org/2006/149.

[6] C. Carlet, "On a weakness of the Tu-Deng function and its repair," Cryptology ePrint Archive. http://eprint.iacr.org/2009/606.

[7] C. Carlet, "Comment on 'constructions of cryptographically significant Boolean functions using primitive polynomials'," preprint.

[8] C. Carlet, D.K. Dalai, K.C. Gupta and S. Maitra,"Algebraic immunity for cryptographically significant Boolean functions: analysis and construction," IEEE Trans. Inf. Theory, vol. 52, pp. 3105-3121, 2006.

[9] C. Carlet and K. Feng, "An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity," in Advanceds in Cryptology-ASIACRYPT 2008, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 5350, pp. 425-440, 2008.

[10] C. Carlet, X. Zeng, C. Li and L. Hu, "Further properties of several classes of Boolean functions with optimum algebraic immunity," Des. Codes Cryptogr. vol. 52, pp. 303-338, 2009.

[11] A. Cauteaut, "Open problems related to algebraic attacks on stream ciphers," in Proceedings of WCC 2005, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3969, pp. 120-134, 2006.

[12] N. Courtois, "Fast algebraic attacks on stream ciphers with linear feedback," in Advances in Cryptology-CRYPTO 2003, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 2729, pp. 176-194, 2003.

[13] N. Courtois and W. Meier, "Algebraic attacks on stream ciphers with linear feedback," in Advances in Cryptology-EUROCRYPT 2003, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 2656, pp. 345-359, 2003.

[14] D.K. Dalai, K.C. Gupta and S. Maitra, "Cryptographically significant Boolean functions: construction and analysis in terms of algebraic immunity," in Fast Software Encryption, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3557 pp. 98-111, 2005.

[15] D.K. Dalai, S. Maitra and S. Sarkar, "Basic theory in construction of Boolean functions with maximum possible annihilator immunity," Des. Codes Cryptogr, vol. 40, pp. 41-58, 2006.

[16] C. Ding, G. Xiao and W. Shan, The Stability Theory of Stream Ciphers, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 561, 1991.

[17] P. Hawkes and G. Rose, "Rewriting variables: the complexity of fast algebraic attacks on stream ciphers," in Advances in Cryptology-CRYPTO 2004, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3152, pp. 390-406, 2004.

[18] N. Li and W. Qi, "Construction and analysis of boolean functions of $2t + 1$ variables with maximum algebraic immunity," in Advanceds in Cryptology-ASIACRYPT 2006, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 4284, pp. 84-98, 2006.

[19] N. Li, L. Qu, W. Qi, G. Feng, C. Li and D. Xie, "On the construction of Boolean functions with optimal algebraic immunity," IEEE Trans. Inf. Theory, vol. 54, pp. 1330-1334, 2008.

[20] R. Lidl and H. Niederreiter, "Finite felds," in Encyclopedia of Mathematics and Its Applications. Reading, MA: Addison-Wesley, vol. 20, 1983.

[21] W. Meier, E. Pasalic and C. Carlet, "Algebraic attacks and decomposition of Boolean functions," in Advances in Cryptology-EUROCRYPT 2004, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 3027, pp. 474-491, 2004.

[22] W. Meier and O. Staffelbach, "Fast correlation attacks on stream ciphers," in Advances in Cryptology-EUROCRYPT 1988, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, vol. 330, pp. 301-314, 1988.

[23] P. Rizomiliotis, "On the resistance of Boolean functions against algebraic attacks using univariate polynomial representation," IEEE Trans. Inform. Theory, vol. 56, pp. 4014-4024, 2010.

[24] S. Rønjom and T. Helleseth, "A new attack on the filter generator," IEEE Trans. Inform. Theory, vol. 53, pp. 1752-1758, 2007.

[25] X. Tang, D. Tang, X. Zeng and L. Hu, "Balanced Boolean functions with (almost) optimal algebraic immunity and very high nonlinearity," Cryptology ePrint Archive. http://eprint.iacr.org/2010/443.

[26] Z. Tu and Y. Deng, "A conjecture on binary string and its applications on constructing boolean functions of optimal algebraic immunity," Designs, Codes Cryptogr., 2010. Online First Articles. DOI 10.1007/s10623-010-9413-9

[27] Q. Wang and T. Johansson, "A note on fast algebraic attacks and higher order nonlinearities," to appear in INSCRYPT 2010.

[28] Q. Wang, J. Peng, H. Kan and X. Xue, "Constructions of cryptographically signicant Boolean functions using primitive polynomials," IEEE Trans. Inform. Theory, vol. 56, pp. 3048-3053, 2010.