Highly Nonlinear Boolean Functions with Optimal Algebraic Immunity and Good Behavior Against Fast Algebraic Attacks

Deng Tang* Claude Carlet† Xiaohu Tang

Abstract

In this paper, we present a new combinatorial conjecture about binary strings. Based on the new conjecture, two classes of Boolean functions of $2k$ variables with optimal algebraic immunity are proposed, where $k \geq 2$. The first class contains unbalanced functions having high algebraic degree and nonlinearity. The functions in the second one are balanced and have maximal algebraic degree and high nonlinearity. It is checked that, at least for small numbers of variables, both classes of functions have a good behavior against fast algebraic attacks. Compared with the known Boolean functions resisting algebraic attacks and fast algebraic attacks, the two classes of functions possess the highest lower bounds on nonlinearity. These bounds are however not enough for ensuring a sufficient nonlinearity for allowing resistance to the fast correlation attack. Nevertheless, as for previously found functions with the same features, there is a gap between the bound that we can prove and the actual values computed for small numbers of variables. Moreover, these values are very good and much better than for the previously found functions having all the necessary features for being used in the filter model of pseudo-random generators.

Keywords: Boolean functions, balancedness, algebraic immunity, fast algebraic attack, algebraic degree, nonlinearity.

1 Introduction

Boolean functions are the building blocks of symmetric cryptographic systems. They are used for S-box design in block ciphers and utilized as nonlinear filters and combiners

*D. Tang and X.H. Tang are with the Provincial Key Lab of Information Coding and Transmission, Institute of Mobile Communications, Southwest Jiaotong University, Chengdu, 610031, China. Email: dengtanghome@qq.com, xhutang@ieee.org

†C. Carlet is with the LAGA, Universities of Paris 8 and Paris 13; CNRS, UMR 7539; Address: University of Paris 8, Department of Mathematics, 2 rue de la liberté, 93526 Saint-Denis cedex 02, France. Email: claude.carlet@inria.fr

in stream ciphers. To resist the known attacks on each cryptosystem, Boolean functions should satisfy various criteria simultaneously. In the framework of stream ciphers with filter nonlinear model, the following criteria for cryptographic Boolean functions are mandatory: balancedness, nonlinearity, algebraic degree and algebraic immunity, for resisting many kinds of known attacks [6, 13]. In the case of the combiner model, the additional condition of resiliency is necessary for allowing resistance to the Siegenthaler correlation attack (which does not work for the filter model); we shall not address this model in this paper.

In 2003, Courtois and Meier successfully proposed algebraic attacks on several stream ciphers [9]. As a response to the standard algebraic attack, the notion of algebraic immunity of a Boolean function f was introduced [26], defined as the minimum algebraic degree of the nonzero functions g such that $f * g = 0$ or $(f + 1) * g = 0$, where $*$ is the multiplication of functions inherited from multiplication in $\mathbb{F}_2$, the finite field with two elements. For resisting the standard algebraic attack, a Boolean function should have algebraic immunity as high as possible, that is, close to the maximum $\lceil \frac{n}{2} \rceil$, where n is the number of variables [9]. Later, the standard attack was further improved in [10], where the so-called fast algebraic attack (FAA) was introduced. The fast algebraic attacks is feasible if one can find nonzero function g of low algebraic degree and h of algebraic degree not much larger than $\frac{n}{2}$, such that $f * g = h$ [1, 10, 18]. Particularly, if f admits a low algebraic degree annihilator, then the algebraic attack using g is more efficient since it needs less data; note that it is a special case of the fast algebraic attacks. Summarizing, for resisting the algebraic attacks, a high algebraic immunity is now an absolutely necessary property for cryptographic Boolean functions, but it is not sufficient for resisting the fast algebraic attacks.

Up to now, several classes of Boolean functions achieving optimal algebraic immunity have been proposed [3, 8, 12, 20, 21]. However, the nonlinearitiy of most of these functions are often not exceeding $2^{n-1} - \binom{n-1}{\lfloor \frac{n}{2} \rfloor}$, which is almost the worst possible value according to Lobanov's bound [23]. Even when they do exceed it, they are not much larger than this value. Hence, they are insufficient for the resistance to fast correlation attacks [25] and there remains to see whether these functions behave well against fast algebraic attacks.

In 2008, the first author and Feng studied an infinite class of n-variable balanced Boolean functions with optimal algebraic immunity which had been introduced in [15]. They clarified the reasons why these functions have optimal algebraic immunity and they proved they have additionally maximal algebraic degree and nonlinearity larger than $2^{n-1} + \frac{2^{n/2+1}}{\pi} \ln\left(\frac{\pi}{4(2^n-1)}\right) - 1 \approx 2^{n-1} - \frac{2\ln 2}{\pi} n 2^{n/2}$ [4]. This showed that these functions have much better nonlinearity than previously found functions. Moreover, they have a quite good nonlinearity computed for small values of n. In addition, it was also checked that, at least

for small number of variables, the functions in this class have a good behavior against fast algebraic attacks. It is the first class of Boolean functions almost satisfying all the criteria and potentially satisfying them completely (as checked for small values of n). In this paper, this class is called the Carlet-Feng function for short.

In [33], the Carlet-Feng function was presented by Wang *et al* in another way (as shown in [7]) with a very slightly improved lower bound on the nonlinearity: $\max\{6\lfloor\frac{2^{n-1}}{2n}\rfloor - 2, 2^{n-1} - \big(\frac{\ln 2}{3}(n-1) + \frac{3}{2}\big)\,2^{\frac{n}{2}}\}$. A minor modification of this class was also introduced (without proof) with the same parameters [33]. Furthermore, the Carlet-Feng function was analyzed in [27], and the author proposed new balanced Boolean functions with optimal algebraic immunity and gave an efficient method evaluating their behavior against fast algebraic attacks. Later in [35], Zeng *et al* presented more balanced functions with almost the same cryptographic properties as the Carlet-Feng function by extending the analytical method presented in [27].

More recently, Tu and Deng [32] constructed a class of bent functions with optimal algebraic immunity, that they could modify into $2k$-variable balanced Boolean functions with maximal algebraic degree and nonlinearity larger than $2^{n-1} - 2^{k-1} - 2^{\frac{k}{2}}k \cdot \ln 2 - 1$, where $n = 2k$. Most notably, based on a combinatorial conjecture, they were able to show that their Boolean functions possess optimal algebraic immunity. By means of the same assumption, another class of balanced Boolean functions in even number of variables was also presented [31], which have optimal algebraic immunity, maximal algebraic degree, and very high nonlinearity as good as the best result of the known balanced Boolean functions. But these two classes of functions are vulnerable to fast algebraic attacks [5, 34].

In this paper, we propose two classes of functions of $n = 2k$ variables with very good cryptographic properties, where $k \geq 2$. The functions in the first class are unbalanced. They have Hamming weight $2^{n-1} - 2^{k-1}$, algebraic degree $n-2$, and have nonlinearity larger than $2^{n-1} - \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k - 1$. The functions in the second class are balanced, having algebraic degree $n-1$, which is the maximal algebraic degree of balanced functions of n variables, and nonlinearity larger than $2^{n-1} - \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k - 2^{\lfloor\frac{k}{2}\rfloor} - 1$. Based on a new conjecture, these two classes of functions have optimal algebraic immunity. Further, experiments show that both of them have a good behavior against fast algebraic attacks (as checked for the even number of variables ranging from 4 to 16). Besides, in the process of the nonlinearity of our functions, we also find a tighter lower bound on the nonlinearity of the Carlet-Feng function. Even compared with this new lower bound, our two classes of functions possess the largest lower bounds among those functions with similar features which have been already found. And what seems to be the most interesting with these functions is that the actual

values of the nonlinearity computed for small numbers of variables happen to be far the best among all functions having large algebraic immunity, large algebraic degree and potentially good resistance to fast algebraic attacks. We believe also interesting to observe that the Tu-Deng conjecture can be extended to other similar ones. This gives the idea that a more general conjecture could be stated and maybe more easily proved (we actually checked a more general conjecture, for small values of the number of variables).

The remainder of this paper is organized as follows. In Section 2, the notations and the necessary preliminaries required for the subsequent sections are reviewed. In Section 3, the new combinatorial conjecture about binary strings is presented. In Section 4, we propose the construction of unbalanced functions and give their cryptographic properties. In Section 4, an infinite class of balanced functions satisfying all the main cryptographic properties are proposed and analyzed. Finally, Section 5 concludes the paper.

2 Preliminaries

Let $\mathbb{F}_2^n$ be the vector space of n-tuples over the field $\mathbb{F}_2 = \{0, 1\}$ of two elements, and $\mathbb{F}_{2^n}$ be the finite field of order 2^n. For a vector $a = (a_1, \cdots, a_n) \in \mathbb{F}_2^n$, its *support* $\mathrm{Supp}(a)$ is the set $\{1 \leq i \leq n \,|\, a_i = 1\}$, and its *Hamming weight* $\mathrm{wt}(a)$ is defined as the cardinality of its support, i.e., $\mathrm{wt}(a) = |\mathrm{Supp}(a)|$.

We denote by $\mathcal{B}_n$ the set of all the Boolean functions of n variables. A *Boolean function* of n variables is a function from $\mathbb{F}_2^n$ into $\mathbb{F}_2$. The basic representation of a Boolean function $f(x_1, \cdots, x_n)$ is by its *truth table*, i.e.,

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

We say that a Boolean function f is *balanced* if its truth table contains an equal number of ones and zeros, that is, if its Hamming weight equals 2^{n-1}. The *Hamming weight* of f, $\mathrm{wt}(f)$, is defined as the Hamming weight of this string, or in other words, the size of the support $\mathrm{Supp}(f) = \{x \in \mathbb{F}_2^n \,|\, f(x) = 1\}$.

Furthermore, any Boolean function $f \in \mathcal{B}_n$ can be uniquely represented by a multivariate polynomial over $\mathbb{F}_2$, called the *algebraic normal form* (ANF), of the form:

$$f(x_1, \cdots, x_n) = \bigoplus_{u \in \mathbb{F}_2^n} a_u \Big(\prod_{j=1}^{n} x_j^{u_j} \Big),$$

where $a_u \in \mathbb{F}_2$ and $u = (u_1, \cdots, u_n)$. The *algebraic degree*, denoted by $\deg(f)$, is the maximal value of $\mathrm{wt}(u)$ such that $a_u \neq 0$. A Boolean function is an *affine function* if its algebraic degree is at most 1. The set of all affine functions is denoted by A_n. The

cryptographic Boolean functions can be attacked by the Berlekamp-Massey algorithm [24] and by the Rønjom-Helleseth attack [28] if the functions have low algebraic degrees. It should be noted that the maximum algebraic degree of a balanced Boolean functions of n variables is $n-1$.

Note that $\mathbb{F}_{2^n}$ is isomorphic to $\mathbb{F}_2^n$ through the choice of some basis of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$. In this paper, for convenience, we shall represent the truth table of Boolean functions as

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

where α is a primitive element of $\mathbb{F}_{2^n}$

The Boolean functions over $\mathbb{F}_{2^n}$ can also be uniquely expressed 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 $1 \leq i < 2^n - 1$ such that $a_i = a_{2i \pmod{2^n-1}}$, and the addition is modulo 2. In [4], it was shown that the algebraic degree $\deg(f)$ equals $\max\{\mathrm{wt}(\bar{i}) | a_i \neq 0, 0 \leq i < 2^n\}$, where $\bar{i}$ is the binary expansion of i.

Besides, when n is even, the Boolean function of n variables can be viewed over $\mathbb{F}_{2^{n/2}}^2$ and uniquely expressed by a *bivariate polynomial*

$$f(x, y) = \sum_{i,j=0}^{2^{n/2}-1} a_{i,j} x^i y^j$$

where $a_{i,j} \in \mathbb{F}_{2^{n/2}}$. The algebraic degree $\deg(f)$ equals $\max\{\mathrm{wt}(\bar{i}) + \mathrm{wt}(\bar{j}) | a_{i,j} \neq 0\}$.

In order to resist the Best Affine Approximation (BAA) [13] and the fast correlation attack [25], Boolean functions used in a cryptographic system must have high nonlinearity. The *nonlinearity* N_f of a Boolean function $f \in \mathcal{B}_n$ is defined as

$$N_f = \min_{g \in A_n} (d_H(f, g)),$$

where $d_H(f, g)$ is the Hamming distance between f and g, i.e., $d_H(f, g) = |\{x \in \mathbb{F}_2^n \,|\, f(x) \neq g(x)\}|$. In other words, the nonlinearity N_f is the minimum Hamming distance between f and all the affine functions.

The nonlinearity can also be expressed by means of the Walsh transform of f. Let $x = (x_1, x_2, \cdots, x_n)$ and $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ both belong to $\mathbb{F}_2^n$ and let $x \cdot \alpha$ be the usual inner product $x \cdot \alpha = x_1\alpha_1 \oplus x_2\alpha_2 \oplus \cdots \oplus x_n\alpha_n$, then the Walsh transform of $f \in \mathcal{B}_n$ at α is defined by

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

Over $\mathbb{F}_{2^n}$, the Walsh transform of the Boolean function f can be defined by

$$W_f(a) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x)+tr(ax)}$$

where $tr(x) = \sum_{i=0}^{n-1} x^{2^i}$ is the trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$.
Over $\mathbb{F}^2_{2^{n/2}}$, the Walsh transform can be defined by

$$W_f(a,b) = \sum_{x,y \in \mathbb{F}_{2^{n/2}}} (-1)^{f(x,y)+tr(ax+by)}$$

where $a, b \in \mathbb{F}_{2^{n/2}}$ and tr is the trace function from $\mathbb{F}_{2^{n/2}}$ to $\mathbb{F}_2$.

Then, the nonlinearity of a Boolean function $f \in \mathcal{B}_n$ can be computed as

$$N_f = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_2^n} |W_f(a)| \ \big(\text{or } 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_{2^n}} |W_f(a)| \text{ or } 2^{n-1} - \frac{1}{2} \max_{a,b \in \mathbb{F}_{2^{n/2}}} |W_f(a,b)|\big).$$

A new kind of attack, called algebraic attack, has been introduced recently [9]. Algebraic Attack is a new powerful tool to cryptanalyse many stream ciphers which were previously believed very secure. The basic idea of algebraic attacks is to represent a cryptographic system by a large system of multivariate algebraic equations, and by using a trick allowing solving it with more efficiency; then one can recover the secret key. The idea was originated from Shannon [29], but this improvement in the efficiency of the method is recent. The standard algebraic attack leads to the following definition.

Definition 1 ([26]). *Given two n-variable Boolean functions f and h, h is said to be an annihilator of f if $f * h = 0$. The algebraic immunity $AI(f)$ of Boolean function f is defined to be the minimum algebraic degree of nonzero Boolean function h such that h is an annihilator of f or $f + 1$.*

In this paper, a Boolean function f of n variables is said to have optimal algebraic immunity if it has maximal algebraic immunity $\lceil \frac{n}{2} \rceil$. A high algebraic immunity is necessary but is not a sufficient condition for the resistance against all kinds of algebraic attacks. If one can find nonzero functions g of low algebraic degree and h of algebraic degree significantly lower than n such that $f * g = h$, then a fast algebraic attacks is feasible [1, 10, 18]. For an n-variable function f and any pair of integers (e, d) such that $e + d \geq n$, there is a nonzero function g of degree at most e such that $f * g$ has degree at most d, see [10]. In this sense, f can be considered as having optimal behavior against fast algebraic attacks if there do not exist two nonzero functions g and h such that $f * g = h$ and $\deg(g) + \deg(h) < n$ with $1 \leq \deg(g) < \lceil \frac{n}{2} \rceil$ (indeed, if the algebraic degree of g is not smaller than $\lceil \frac{n}{2} \rceil$, then we know that h can be taken equal to 0 or to g itself and we are in the framework of algebraic attacks).

3 New combinatorial conjecture

Very recently, Tu and Deng [32] presented a combinatorial conjecture on binary strings for proving that the Boolean functions they introduced have optimal algebraic immunity. Recall that $\overline{x}$ is the binary expansion of the integer $0 \leq x < 2^k - 1$.

Tu-Deng's Conjecture ([32]): Let $k > 1$ be an integer. For any $0 < t < 2^k - 1$, define

$$Q_t = \left\{(a, b) \mid 0 \leq a, b < 2^k - 1, a + b \equiv t \,(\mathrm{mod} 2^k - 1), \mathrm{wt}(\overline{a}) + \mathrm{wt}(\overline{b}) \leq k - 1\right\}$$

then $|Q_t| \leq 2^{k-1}$.

Tu and Deng validated the conjecture by computer for $k \leq 29$. Towards the proof of this conjecture, some advances have been achieved [11, 16, 17], in which many cases for t are solved. However, the complete proof remains open.

In the present paper, in order to show that the functions given in the subsequent two sections have optimal algebraic immunity, we propose a new combinatorial conjecture about the binary strings.

New Conjecture: Let $k > 1$ be an integer. For any $0 < t < 2^k - 1$, define

$$C_t = \left\{(a, b) \mid 0 \leq a, b < 2^k - 1, a - b \equiv t \,(\mathrm{mod} 2^k - 1), \mathrm{wt}(\overline{a}) + \mathrm{wt}(\overline{b}) \leq k - 1\right\} \qquad (1)$$

then $|C_t| \leq 2^{k-1}$.

We have the following two Lemmas about the new conjecture.

Lemma 1. *For any integer $k > 1$. Let C_t be the set defined in (1). Then we have $|C_t| = |C_{2t \pmod{2^k-1}}|$ for every $0 < t < 2^k - 1$.*

***Proof*:** For any fixed t. We can see that $(a, b) \in C_t$ if and only if $(2a, 2b) \in C_{2t}$ (all integers are modulo $2^k - 1$). Note that $\mathrm{wt}(a) = \mathrm{wt}(2a)$ and $\mathrm{wt}(b) = \mathrm{wt}(2b)$. So we have $\mathrm{wt}(a) + \mathrm{wt}(b) = \mathrm{wt}(2a) + \mathrm{wt}(2b)$, which implies that $|C_t| = |C_{2t \pmod{2^k-1}}|$.

Lemma 2. *For any integer $k > 1$. Let C_t be the set defined in (1). Then we have $|C_t| = |C_{2^k-1-t}|$ for any $0 < t < 2^k - 1$.*

***Proof*:** For any $0 < t < 2^k - 1$. We can see that $(a, b) \in C_t$ if and only if $(b, a) \in C_{-t}$. Note that $-t = 2^k - 1 - t$ and $\mathrm{wt}(a) + \mathrm{wt}(b) = \mathrm{wt}(b) + \mathrm{wt}(a)$. So we have $|C_t| = |C_{2^k-1-t}|$.

Based on these two lemmas, we can reduce the search space for checking the new conjecture. By C program, we validated our conjecture for $2 \leq k \leq 29$, whose corresponding values are given in Table 1. We believe that our conjecture is true, although we can not prove it mathematically up to now. There seems to be a close connection between these two

conjectures, but our new conjecture is essentially not equivalent to Tu-Deng's conjecture since we computed $|Q_t|$ for $2 \leq k \leq 20$ and found that $\max\limits_{0<t<2^k-1} |C_t| < \max\limits_{0<t<2^k-1} |Q_t| = 2^{k-1}$ for $3 \leq k \leq 20$.

Table 1. The corresponding values of new conjecture for small k

k	2	3	4	5	6	7	8	9	10	11
$\max\limits_{0<t<2^k-1} \|C_t\|$	2	3	7	13	28	55	114	227	463	925
2^{k-1}	2	4	8	16	32	64	128	256	512	1024
$2^{k-1} - \max\limits_{0<t<2^k-1} \|C_t\|$	0	1	1	3	4	9	14	29	49	99
k	12	13	14	15	16	17	18	19	20	21
$\max\limits_{0<t<2^k-1} \|C_t\|$	1873	3745	7555	15109	30415	60829	122284	244567	491190	982379
2^{k-1}	2048	4096	8192	16384	32768	65536	131072	262144	524288	1048576
$2^{k-1} - \max\limits_{0<t<2^k-1} \|C_t\|$	175	351	637	1275	2353	4707	8788	17577	33098	66197
k	22	23	24	25	26	27	28	29		
$\max\limits_{0<t<2^k-1} \|C_t\|$	1971676	3943351	7910416	15820831	31724162	63448323	127187158	254374315		
2^{k-1}	2097152	4194304	8388608	16777216	33554432	67108864	134217728	268435456		
$2^{k-1} - \max\limits_{0<t<2^k-1} \|C_t\|$	125476	250953	478192	956385	1830270	3660541	7030570	14061141		

Remark 1. *We checked, for* $2 \leq k \leq 15$*, a more general conjecture in which* $a - b$ *is replaced by* $ua \pm b$ *where* u *is a positive integer such that* $\gcd(u, 2^k - 1) = 1$ *is true.*

4 Boolean functions with very good cryptographic properties

In this section, we present a class of Boolean functions with optimal algebraic immunity, good immunity to fast algebraic attacks, high algebraic degree and a nonlinearity provably larger than all previously introduced functions with the same features. Moreover, the nonlinearity that we exactly computed for small values of n is very good.

Construction 1: Let $n = 2k \geq 4$. Let α be the primitive root of the finite field $\mathbb{F}_{2^k}$. Set $\Delta_s = \{\alpha^s, \cdots, \alpha^{2^{k-1}+s-1}\}$ where $0 \leq s < 2^k - 1$ is an integer. Then we construct a function $f \in \mathcal{B}_n$ as follows:

$$f(x, y) = g(xy), \tag{2}$$

where g is defined on $\mathbb{F}_{2^k}$ with $\mathrm{Supp}(g) = \Delta_s$.

4.1 Algebraic immunity of the constructed functions

In this subsection, we will show that the Boolean functions generated by Construction 1 have optimal algebraic immunity under the assumption that our new conjecture is true. The proof is very similar to the proof for the Tu-Deng functions.

Theorem 1. *Let f be the n-variable Boolean function generated by Construction 1. If the new conjecture is correct, then f has the optimal algebraic immunity, i.e., $AI(f) = k$.*

***Proof*:** From Construction 1, we know that $\mathrm{Supp}(f) = \{(\gamma y^{2^k-2}, y) \mid y \in \mathbb{F}_{2^k}^*, \gamma \in \Delta_s\}$.

First, assume that $h(x,y) = \sum_{i=0}^{2^k-1}\sum_{j=0}^{2^k-1} h_{i,j}x^i y^j \in \mathcal{B}_n$ is an annihilator of f with $\deg(h) < k$, i.e.,

(1) $h(\gamma y^{2^k-2}, y) = 0$ for $\forall y \in \mathbb{F}_{2^k}^*, \gamma \in \Delta_s$;

(2) $h_{i,j} = 0$ if $\mathrm{wt}(\bar{i})+\mathrm{wt}(\bar{j}) \geq k$, which implies $h_{2^k-1,j} = h_{i,2^k-1} = 0$ for all $0 \leq i,j < 2^k-1$.

We have:

$$\begin{aligned} h(\gamma y^{2^k-2}, y) &= \sum_{i=0}^{2^k-2}\sum_{j=0}^{2^k-2} h_{i,j}\gamma^i y^{j-i} \\ &= \sum_{t=0}^{2^k-2} h_t(\gamma) y^t \end{aligned}$$

where

$$\begin{aligned} h_t(\gamma) &= \sum_{0\leq i,j\leq 2^k-2, j-i\equiv t \,(\mathrm{mod}\, 2^k-1)} h_{i,j}\gamma^i \\ &= \sum_{i=0}^{2^k-2-t} h_{i,t+i}\gamma^i + \sum_{i=2^k-1-t}^{2^k-2} h_{i,t+i-(2^k-1)}\gamma^i. \end{aligned} \tag{3}$$

For arbitrary $\gamma \in \Delta_s$, the condition $h(\gamma y^{2^k-2}, y) = 0$ for $y \in \mathbb{F}_{2^k}^*$, results in

$$h_t(\gamma) = 0, 0 \leq t \leq 2^k - 2.$$

Then, given $0 \leq t \leq 2^k-2$, the vector $(h_{0,t}, h_{1,t+1}, \cdots, h_{2^k-1-t,0}, h_{2^k-t,1}, \cdots, h_{2^k-2,t-1})$ is a BCH codeword of length 2^k-1 over $\mathbb{F}_{2^k}$, having the elements in Δ_s as zeros and the designed distance $2^{k-1}+1$. According to the BCH bound, the codeword has Hamming weight at least $2^{k-1}+1$ if it is nonzero. However, from the new conjecture, its Hamming weight should be no more than 2^{k-1}. Hence the codeword must be zero, i.e.,

$$h_{0,t} = h_{1,t+1} = \cdots = h_{2^k-1-t,0} = h_{2^k-t,1} = \cdots = h_{2^k-2,t-1} = 0$$

for any $0 \leq t \leq 2^k-2$. That is, $h = 0$.

Next, let us consider the case for $f+1$. Suppose that $h(x,y) = \sum_{i=0}^{2^k-1}\sum_{j=0}^{2^k-1} h_{i,j}x^i y^j \in \mathcal{B}_n$ is an annihilator of $f+1$ with $\deg(h) < k$. Similarly, for all $0 \leq t \leq 2^k-2$, we have

$$h_t(\gamma) = 0, \forall \gamma \in \mathbb{F}_{2^k}^* \setminus \Delta_s,$$

where h_t is defined by (3) and $h_{i,0} = 0$ for $0 \le i \le 2^k - 2$.
Then the vector $(h_{0,t}, h_{1,t+1}, \cdots, h_{2^k-1-t,0}, h_{2^k-t,1}, \cdots, h_{2^k-2,t-1})$ is also a BCH codeword of length $2^k - 1$ over $\mathbb{F}_{2^k}$, having the elements in $\mathbb{F}_{2^k}^* \setminus \Delta_s$ as zeros and the designed distance 2^{k-1}. By the BCH bound, if the codeword is nonzero, then it has Hamming weight at least 2^{k-1}. But according to the new conjecture, the number of pairs (a, b) such that $0 \le a, b < 2^k - 1$, $a - b \equiv t \pmod{2^k - 1}$ and $\mathrm{wt}(\overline{a}) + \mathrm{wt}(\overline{b}) \le k - 1$ is then at most 2^{k-1} and since we have that $h_{i,0} = 0$ for $0 \le i \le 2^k - 2$, a contradiction follows. So, we have $h = 0$.

From what has been discussed above, both f and $f+1$ have no nonzero annihilators with algebraic degrees less than k. Then, $AI(f) = k$. It means that the constructed functions have optimal algebraic immunity.

□

4.2 The immunity to fast algebraic attacks

In this subsection, we analyze the immunity to fast algebraic attacks of the constructed functions in Construction 1 in small number of variables.

For $n = 2k$, let α be the default primitive root of $\mathbb{F}_{2^k}$ in Magma system and $g \in \mathcal{B}_k$ be the balanced function with $\mathrm{Supp}(g) = \{1, \alpha, ..., \alpha^{2^{k-1}-1}\}$ in Construction 1. By a Magma program, we can easily get the truth table of the n-variable Boolean function f given in Construction 1. Let $g_1, h_1 \in \mathcal{B}_n$ be two functions with $1 \le \deg(g_1) = e < k$ and $\deg(h_1) = d$ such that $f * g_1 = h_1$. We aim to find out the minimal sum $e + d$ to exploit its immunity against the fast algebraic attack for some small number of variables. Clearly, h_1 is an annihilator of $f + 1$, which implies $d \ge k$. Hence, we investigate all the combinations of e and d with $1 \le e < k$ and $k \le d$. Using the Algorithm 2 in [2], we have the following properties of f.

- For $n = 8, 12, 14, 16$, we only found pairs (e, d) exist for equations $e + d \ge n - 2$ by exhaustive check, where $1 \le e < n/2$.
- For $n = 4, 6, 10$, pairs (e, d) such that $e + d \le n - 1$ were never observed by exhaustive check, where $1 \le e < n/2$.

Recall that for any function of n variables always exists the pairs $e + d = n$, where $1 \le e < \lceil n/2 \rceil$. Therefore, f has the optimal immunity to fast algebraic attacks for $n = 4, 6, 10$, and the nearly optimal immunity to fast algebraic attacks for $n = 8, 12, 14, 16$. The above examples show that this class of functions, at least for small values of the number of variables, have a good behavior against fast algebraic attacks, even if not always optimal.

4.3 Polynomial representation and algebraic degree

We give now the univariate representation of the constructed functions and deduce their algebraic degree.

Theorem 2. *Let f be the n-variable Boolean function defined in Construction 1. Then its univariate representation equals*

$$f(x,y) = \sum_{i=1}^{2^k-2} \alpha^{-is}(1+\alpha^{-i})^{2^{k-1}-1}(xy)^i$$

Therefore, f has algebraic degree $n-2$.

***Proof*:** Define $g(y) = \sum_{i=0}^{2^k-1} g_i y^i$ to be the univariate representation of g with $\mathrm{Supp}(g) = \Delta_s$. It follows from the Fourier transform that

$$\begin{aligned} g_i &= \sum_{j=0}^{2^k-1} g(\alpha^j)\alpha^{-ij} = \sum_{j=s}^{2^{k-1}+s-1} \alpha^{-ij} = \alpha^{-is}\sum_{j=0}^{2^{k-1}-1} \alpha^{-ij} \\ &= \begin{cases} 0, & \text{if } i = 0 \text{ or } 2^k-1 \\ \alpha^{-is}\frac{1+\alpha^{-i2^{k-1}}}{1+\alpha^{-i}} = \alpha^{-is}(1+\alpha^{-i})^{2^{k-1}-1}, & \text{if } 1 \le i \le 2^k-2 \end{cases}. \end{aligned}$$

Then we have $g(y) = \sum_{i=1}^{2^k-2} \alpha^{-is}(1+\alpha^{-i})^{2^{k-1}-1} y^i$ and $\deg(g) = k-1$ due to $g_{2^k-1} = 0$ and $g_{2^k-2} \neq 0$. Immediately, by the definition of $f(x,y) = g(xy)$ in (2) we obtain

$$f(x,y) = g(xy) = \sum_{i=1}^{2^k-2} \alpha^{-is}(1+\alpha^{-i})^{2^{k-1}-1}(xy)^i$$

and therefore $\deg(f) = 2(k-1) = n-2$.

□

Remark 2. *Theorem 2 can also be deduced from the univariate representation of g given in [4].*

4.4 Nonlinearity

Before obtaining a lower bound on the nonlinearity of f, we need a few preliminary results. The following series is well known.

Lemma 3 ([30])**.** $\frac{1}{\sin(x)} = \frac{1}{x} + \frac{x}{6} + \frac{7x^3}{360} + \frac{31x^5}{15120} + \cdots + \frac{2(2^{2n-1}-1)B_n}{(2n)!}x^{2n-1} + \cdots = \frac{1}{x} + \sum_{n=1}^{\infty} M_n$, *where* $M_n = \frac{2(2^{2n-1}-1)B_n}{(2n)!}x^{2n-1}$ *and* $B_n = \frac{(2n)!}{\pi^{2n}2^{2n-1}}\sum_{m=1}^{\infty}\frac{1}{m^{2n}}$ *is the Bernoulli's number.*

We deduce a corollary and prove a lemma.

Corollary 1. *For every* $0 < x < \frac{\pi}{2}$,

$$\frac{1}{\sin(x)} < \frac{1}{x} + \frac{x}{4}.$$

***Proof*:** For $n \geq 2$ and $0 < x < \frac{\pi}{2}$,

$$\begin{aligned}
\frac{M_{n+1}}{M_n} &= \frac{(2^{2n+1}-1)\sum_{m=1}^{\infty}\frac{1}{m^{2n+2}}}{4\pi^2(2^{2n-1}-1)\sum_{m=1}^{\infty}\frac{1}{m^{2n}}}x^2 \\
&\leq \frac{(2^{2n+1}-1)\sum_{m=1}^{\infty}\frac{1}{m^{2n+2}}}{16(2^{2n-1}-1)\sum_{m=1}^{\infty}\frac{1}{m^{2n}}} \\
&\leq \frac{(2^{2n+1}-1)}{16(2^{2n-1}-1)} \\
&= \frac{1}{4} + \frac{3}{4(2^{2n+1}-4)} \\
&< \frac{1}{3}
\end{aligned}$$

Additionally, $M_2/M_1 < 0.288 < 1/3$. So we have $M_{n+1}/M_n < 1/3$ for any integer $n \geq 1$, which leads to

$$\frac{1}{\sin(x)} < \frac{1}{x} + \frac{x}{6}\sum_{m=0}^{\infty}\frac{1}{3^m} = \frac{1}{x} + \frac{x}{4}.$$

□

Lemma 4. *Let* α *be a primitive element of* $\mathbb{F}_{2^k}$ *and* $k \geq 2$ *a positive integer. Let* $\Delta_s = \{\alpha^s, \cdots, \alpha^{2^{k-1}+s-1}\}$ *where* $0 \leq s < 2^k - 1$ *is an integer. Define*

$$\Gamma_s = \sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(1/x+\gamma x)},$$

where $0 \leq s < 2^k - 1$. *Then,*

$$|\Gamma_s| < \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k.$$

***Proof*:** Let $\zeta = e^{\frac{2\pi\sqrt{-1}}{2^k-1}}$ be a primitive (2^k-1)-th root of 1 in the complex field $\mathbb{C}$, let χ be the multiplicative character of $\mathbb{F}_{2^k}^*$ defined by $\chi(\alpha^j) = \zeta^j$ $(0 \leq j \leq 2^k - 2)$ and let $\chi(0) = 0$. We define the Gauss sum:

$$G(\chi^\mu) = \sum_{x\in\mathbb{F}_{2^k}^*}\chi^\mu(x)(-1)^{tr(x)},\ 0 \leq \mu \leq 2^k - 2.$$

It is well-known that $G(\chi^0) = -1$ and $|G(\chi^\mu)| = 2^{\frac{k}{2}}$ for $1 \le \mu \le 2^k - 2$ [22]. By Fourier transform,

$$(-1)^{tr(\alpha^j)} = \frac{1}{2^k - 1} \sum_{\mu=0}^{2^k-2} G(\chi^\mu)\overline{\chi}^\mu(\alpha^j),\ 0 \le j \le 2^k - 2.$$

Denoting $q = 2^k$, we have then

$$\begin{aligned}
\Gamma_s &= \sum_{\gamma\in\Delta_s}\sum_{j=0}^{q-2}(-1)^{tr(\alpha^{-j})}(-1)^{tr(\gamma\alpha^j)} \\
&= \frac{1}{(q-1)^2}\sum_{i=s}^{\frac{q}{2}+s-1}\sum_{j=0}^{q-2}\sum_{\mu,\nu=0}^{q-2} G(\chi^\mu)G(\chi^\nu)\zeta^{\mu j-\nu(i+j)} \\
&= \frac{1}{(q-1)^2}\sum_{\mu,\nu=0}^{q-2} G(\chi^\mu)G(\chi^\nu)\Big(\sum_{i=s}^{\frac{q}{2}+s-1}\zeta^{-\nu i}\Big)\Big(\sum_{j=0}^{q-2}\zeta^{(\mu-\nu)j}\Big).
\end{aligned}$$

We can easily deduce that

$$\sum_{i=s}^{\frac{q}{2}+s-1}\zeta^{-\nu i} = \zeta^{-\nu s}\sum_{i=0}^{\frac{q}{2}-1}\zeta^{-\nu i} = \begin{cases} \frac{q}{2} & \text{if } \nu = 0 \\ \zeta^{-\nu s}\,\frac{\zeta^{-\nu\frac{q}{2}}-1}{\zeta^{-\nu}-1} & \text{if } \nu \ne 0 \end{cases}$$

and

$$\sum_{j=0}^{q-2}\zeta^{(\mu-\nu)j} = \begin{cases} q-1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \ne \nu \end{cases}.$$

Therefore,

$$\begin{aligned}
\Gamma_S &= \frac{q}{2(q-1)} + \frac{1}{q-1}\sum_{\nu=1}^{q-2} G^2(\chi^\nu)\zeta^{-\nu s}\frac{\zeta^{-\nu\frac{q}{2}}-1}{\zeta^{-\nu}-1} \\
&= \frac{q}{2(q-1)} + \frac{1}{q-1}\sum_{\nu=1}^{q-2} G^2(\chi^\nu)\zeta^{-\nu s}\frac{\zeta^{-\frac{\nu}{2}}-1}{(\zeta^{-\frac{\nu}{2}}-1)(\zeta^{-\frac{\nu}{2}}+1)} \\
&= \frac{q}{2(q-1)} + \frac{1}{q-1}\sum_{\nu=1}^{q-2} G^2(\chi^\nu)\zeta^{-\nu s}\frac{1}{\zeta^{-\frac{\nu}{2}}+1} \\
&= \frac{q}{2(q-1)} + \frac{1}{q-1}\sum_{\nu=1}^{q-2} G^2(\chi^\nu)\zeta^{-\nu s+\frac{\nu}{4}}\frac{1}{\zeta^{-\frac{\nu}{4}}+\zeta^{\frac{\nu}{4}}}.
\end{aligned}$$

So we have

$$\begin{aligned}
|\Gamma_S| &\leq \frac{q}{2(q-1)} + \frac{q}{2(q-1)} \sum_{\nu=1}^{q-2} \frac{1}{|\cos\left(\frac{\pi\nu}{2(q-1)}\right)|} \\
&= \frac{q}{2(q-1)} + \frac{q}{2(q-1)} \sum_{\nu=1}^{q-2} \frac{1}{\cos\left(\frac{\pi\nu}{2(q-1)}\right)} \\
&= \frac{q}{2(q-1)} + \frac{q}{2(q-1)} \sum_{\nu=1}^{q-2} \frac{1}{\sin\left(\frac{\pi}{2} - \frac{\pi\nu}{2(q-1)}\right)} \\
&= \frac{q}{2(q-1)} + \frac{q}{2(q-1)} \sum_{\nu=1}^{q-2} \frac{1}{\sin\left(\frac{\pi\nu}{2(q-1)}\right)}.
\end{aligned}$$

Applying Lemma 1, we get

$$\begin{aligned}
|\Gamma_S| &< \frac{q}{2(q-1)} + \frac{q}{2(q-1)} \sum_{\nu=1}^{q-2} \Big(\frac{2(q-1)}{\pi\nu} + \frac{\pi\nu}{8(q-1)}\Big) \\
&= \frac{q}{2(q-1)} + \frac{q(q-2)\pi}{32(q-1)} + \frac{q}{\pi} \sum_{\nu=1}^{q-2} \frac{1}{\nu} \\
&\leq \frac{q(q-2)\pi}{32(q-1)} + \frac{q}{\pi}\Big(1 + \int_1^{q-2} \frac{du}{u}\Big) + 1 \\
&\leq \frac{q(q-2)\pi}{32(q-1)} + \frac{q}{\pi}(1 + \ln q) + 1 \\
&< \big(\frac{\ln 2}{\pi}k + \frac{1}{\pi} + \frac{\pi}{32}\big)2^k + 1 \\
&< \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k + 1.
\end{aligned}$$

This completes the proof.

□

Now, we are ready to prove the lower bound on the nonlinearity of the functions generated by Construction 1.

Theorem 3. *For $n = 2k$. Let $f \in \mathcal{B}_n$ be the function given by Construction 1. Then we have*

$$N_f > 2^{n-1} - \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k - 1.$$

***Proof*:** Firstly, $W_f(0,0) = 2^k$ since $\mathrm{wt}(f) = (2^k - 1) \cdot 2^{k-1} = 2^{2k-1} - 2^{k-1}$. Secondly,

for any $(a,b) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k} \setminus \{(0,0)\}$, we have

$$\begin{aligned}
W_f(a,b) &= \sum_{x,y\in\mathbb{F}_{2^k}} (-1)^{f(x,y)+tr(ax+by)} \\
&= -2 \sum_{(x,y)\in \mathrm{Supp}(f)} (-1)^{tr(ax+by)} \\
&= -2 \sum_{x\in\mathbb{F}_{2^k}^*} (-1)^{tr(ax)} \sum_{\gamma\in\Delta_s} (-1)^{tr(b\gamma/x)} \\
&= -2 \sum_{\gamma\in\Delta_s} \sum_{x\in\mathbb{F}_{2^k}^*} (-1)^{tr(ax+b\gamma/x)} \\
&= \begin{cases} -2\sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(b\gamma/x)} & \text{if } a=0, b\in\mathbb{F}_{2^k}^* \\ -2\sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(ax)} & \text{if } b=0, a\in\mathbb{F}_{2^k}^* \\ -2\sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(1/x+ab\gamma x)} & \text{if } a\in\mathbb{F}_{2^k}^*, b\in\mathbb{F}_{2^k}^* \end{cases}
\end{aligned}$$

Recall that $1+\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(b\gamma/x)} = 1+\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(ax)} = 0$ if $b\gamma$ and a are nonzero. Then we have

$$W_f(a,b) = \begin{cases} 2^k & \text{if } a=0, b\in\mathbb{F}_{2^k}^* \\ 2^k & \text{if } b=0, a\in\mathbb{F}_{2^k}^* \\ -2\sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(1/x+ab\gamma x)} & \text{if } a\in\mathbb{F}_{2^k}^*, b\in\mathbb{F}_{2^k}^* \end{cases} \quad (4)$$

which implies

$$\max_{(a,b)\in\mathbb{F}_{2^k}\times\mathbb{F}_{2^k}} |W_f(a,b)| = \max\{2 \max_{0\le s<2^k-1} |\sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(1/x+\gamma x)}|,\ 2^k\}.$$

By Lemma 4, we have

$$\begin{aligned}
N_f &= 2^{n-1} - \frac{1}{2} \max_{(a,b)\in\mathbb{F}_{2^k}\times\mathbb{F}_{2^k}} |W_f(a,b)| \\
&> 2^{n-1} - \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k - 1.
\end{aligned}$$

This completes the proof. □

Further, we experiment the exact value of nonlinearity and show it is much better than what this lower bound gives. Let α be the default primitive root of $\mathbb{F}_{2^k}$ in Magma system. By Magma program, we can compute the values of $\sum_{\gamma\in\Delta_s}\sum_{x\in\mathbb{F}_{2^k}^*}(-1)^{tr(1/x+\gamma x)}$, $0 \le s < 2^k-1$, for some small values of k. Then we get their nonlinearities, which are given in Table 2 below, for even n ranging from 4 to 38. We can see that all the values of the nonlinearity are very close to $2^{n-1} - 2^{n/2}$.

Table 2. The values of the nonlinearity of f

n	4	6	8	10	12	14
$2^{n-1}-2^{n/2}$	4	24	112	480	1984	8064
N_f	6	24	112	480	1988	8036
n	16	18	20	22	24	26
$2^{n-1}-2^{n/2}$	32512	130560	523264	2095104	8384512	33546240
N_f	32520	130520	523164	2095012	8384528	33546056
n	28	30	32	34	36	38
$2^{n-1}-2^{n/2}$	134201344	536838144	2147418112	8589803520	34359476224	137438429184
N_f	134201532	536838180	2147416692	8589819224	34359469324	137438442132

5 Boolean functions with all main cryptographic properties

In this section, we slightly modify Construction 1 to get a class of $2k$-variables balanced Boolean functions with high nonlinearity and algebraic degree. The new functions have optimal algebraic immunity under the assumption that the new conjecture is true. It is shown that the functions also have a good behavior against fast algebraic attacks. It is a new class of Boolean functions almost satisfying all the criteria and potentially satisfying them completely (as checked for small numbers of n).

Construction 2: Let $n = 2k = 2^t m$ be an even integer no less than 4 such that $t \geq 1$ and $\gcd(m, 2) = 1$. Let α be the primitive root of the finite field $\mathbb{F}_{2^k}$. Set $\Delta_s = \{\alpha^s, \cdots, \alpha^{2^{k-1}+s-1}\}$ where $0 \leq s < 2^k - 1$ is an integer. We construct the function $F \in \mathcal{B}_n$ as follows

$$F(x,y) = \begin{cases} g(xy), & x \neq 0 \\ u(y), & x = 0 \end{cases} \tag{5}$$

where g is defined on $\mathbb{F}_{2^k}$ with $\mathrm{Supp}(g) = \Delta_s$ and $u(y)$ is a balanced Boolean function on $\mathbb{F}_{2^k}$ satisfying $u(0) = 0$, $\deg(u) = k-1$, and $\max_{a \in \mathbb{F}_2^k} |W_u(a)| \leq 2^{\frac{m+1}{2}}$ if $t = 1$ and $\max_{a \in \mathbb{F}_2^k} |W_u(a)| \leq \sum_{i=1}^{t-1} 2^{\frac{n}{2^{i+1}}} + 2^{\frac{m+1}{2}}$ if $t \geq 2$.

It should be noted that the required function u does exist, for instance it can be found in [31, 36]. Moreover, we can easily see that the function F is balanced since $\mathrm{wt}(F) = \mathrm{wt}(u) + 2^{k-1}(2^{k-1}-1) = 2^{k-1} + (2^{n-1} - 2^{k-1}) = 2^{n-1}$.

5.1 Algebraic degree and algebraic immunity

We shall see now that F has maximal algebraic degree for balanced function and optimal algebraic immunity by assuming that the new conjecture is correct.

Theorem 4. *Let F be the n-variable Boolean function generated by Construction 2. Then $deg(F) = n - 1$.*

***Proof*:** The constructed function F can be written as $F(x, y) = f(x, y)+U(x, y)$, where $f \in \mathcal{B}_{2k}$ is defined in Construction 1 and $U \in \mathcal{B}_{2k}$ equals $\delta_0(x)u(y)$ where δ_0 is the Kronecker symbol. We know that δ_0 has algebraic degree k and u has algebraic degree $k - 1$. Hence since $\delta_0(x)$ and $u(y)$ depend on independent variables, U has algebraic degree $n - 1$. Since f has algebraic degree $n - 2$, F has same algebraic degree as U. □

Theorem 5. *Let F be the n-variable Boolean function generated by Construction 2. If the new conjecture is true, then $AI(F) = n/2 = k$.*

***Proof*:** The proof is similar to that of Theorem 1. For completeness, we describe a sketch hereafter.

From Construction 2, we know that

$$\mathrm{Supp}(F) = \mathrm{Supp}(f) \bigcup (\{0\} \times \mathrm{Supp}(u)) .$$

Let $h(x, y) \in \mathcal{B}_n$ be an annihilator of F with $\deg(h) < k$. Straightforwardly, $F * h = 0 \Rightarrow f * h = 0$. Whereas by Theorem 1, f has no nonzero annihilators with algebraic degree less than k. Hence, F has no nonzero annihilators with algebraic degree less than k as well.

Next consider the annihilators of $F + 1$. Assume that $h(x, y) \in \mathcal{B}_n$ is an annihilator of $F + 1$ with $\deg(h) < k$. As the same as the case for $f + 1$ in the proof of Theorem 1, we still have

(1) $h(\gamma y^{2^k-2}, y) = 0$ for all $y \in \mathbb{F}_{2^k}^*, \gamma \in \mathbb{F}_{2^k}^* \setminus \Delta_s$;

(2) $h(x, 0) = 0, \forall x \in \mathbb{F}_{2^k}$,

by the definition of F in (5). Using the same argument, we get $h = 0$. In other words, $F+1$ has no nonzero annihilators with algebraic degree less than k too.

From the discussion above, we conclude that $AI(F) = k$. That is, the functions generated by Construction 2 have optimal algebraic immunity.

□

5.2 The immunity to fast algebraic attacks

In what follows, we consider the behavior against fast algebraic attacks of the functions generated in Construction 2 in small number of variables. For $n = 2k$, let α be the default primitive root of $\mathbb{F}_{2^k}$ in Magma system and $g \in \mathcal{B}_k$ be the balanced function with $\mathrm{Supp}(g) = \{1, \alpha, ..., \alpha^{2^{k-1}-1}\}$ in Construction 2. In our experiments, the following Boolean functions u are used, whose truth tables are given in hexadecimal format:

- u=6 for $n = 4$;
- u=E4 for $n = 6$;
- u=6536 for $n = 8$;
- u=25FB7054 for $n = 10$;
- u=5674C6B171A387E4 for $n = 12$;
- u=9696C3C3A5A5F0F06666333355556897 for $n = 14$;
- u=5C877C864CA7F350775076168CA78B717496B37076548BE1A37189E389A76536 for $n = 16$.

Accordingly, we can get the truth table of F.

Let $g_1, h_1 \in \mathcal{B}_n$ be two functions with $\deg(g_1) = e$ and $\deg(h_1) = d$ such that $F * g_1 = h_1$. Using the Algorithm 2 in [2], for even n ranging from 4 to 16, we only found pairs (e, d) such that $e + d \geq n - 1$ by exhaustive check, where $1 \leq e < n/2$. Therefore, our functions have a good behavior against the fast algebraic attacks, although there is no single function for even $4 \leq n \leq 16$ turned out to be optimal against fast algebraic attacks.

5.3 Nonlinearity

Theorem 6. *Let $n = 2k = 2^t m$ be an even integer no less than 4 such that $t \geq 1$ and $\gcd(m, 2) = 1$. Let $F \in \mathcal{B}_n$ be the function given by Construction 2. Then we have*

$$N_F > \begin{cases} 2^{n-1} - \left(\frac{\ln 2}{\pi}k + 0.42\right)2^k - 2^{\frac{k-1}{2}} - 1, & \text{if } t = 1 \\ 2^{n-1} - \left(\frac{\ln 2}{\pi}k + 0.42\right)2^k - \sum_{i=1}^{t-1} 2^{\frac{n}{2^{i+1}}-1} - 2^{\frac{m-1}{2}} - 1, & \text{if } t \geq 2. \end{cases}$$

***Proof*:** According to the definition of the Walsh transform, for any $(a,b) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$, we have (similarly to the computations made in [14])

$$\begin{aligned} W_F(a,b) &= \sum_{x\in\mathbb{F}_{2^k}} \sum_{y\in\mathbb{F}_{2^k}} (-1)^{F(x,y)+tr(ax+by)} \\ &= \sum_{y\in\mathbb{F}_{2^k}} (-1)^{u(y)+tr(by)} + \sum_{x\in\mathbb{F}_{2^k}^*} \sum_{y\in\mathbb{F}_{2^k}} (-1)^{F(x,y)+tr(ax+by)} \\ &= \sum_{y\in\mathbb{F}_{2^k}} (-1)^{u(y)+tr(by)} + \sum_{x\in\mathbb{F}_{2^k}^*} \sum_{y\in\mathbb{F}_{2^k}} (-1)^{f(x,y)+tr(ax+by)} \\ &= W_u(b) + W_f(a,b) - \sum_{y\in\mathbb{F}_{2^k}} (-1)^{tr(by)} \\ &= \begin{cases} 0, & \text{if } b=0 \\ W_u(b)+W_f(a,b), & \text{else} \end{cases} \end{aligned}$$

where in the last identity we use the facts that $W_u(0) = 0$ since u is balanced and $W_f(a,0) = 2^k$ from (4).

Consequently,

$$\max_{(a,b)\in\mathbb{F}_{2^k}\times\mathbb{F}_{2^k}} |W_F(a,b)| \le \max_{(a,b)\in\mathbb{F}_{2^k}\times\mathbb{F}_{2^k}^*} |W_f(a,b)| + \max_{b\in\mathbb{F}_{2^k}^*} |W_u(b)|$$

which results in

$$\begin{aligned} N_F &\ge \begin{cases} N_f - 2^{\frac{k-1}{2}}, & \text{if } t=1 \\ N_f - \sum_{i=1}^{t-1} 2^{\frac{n}{2^{i+1}}-1} - 2^{\frac{m-1}{2}}, & \text{if } t \ge 2. \end{cases} \\ &> \begin{cases} 2^{n-1} - \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k - 2^{\frac{k-1}{2}} - 1, & \text{if } t=1 \\ 2^{n-1} - \big(\frac{\ln 2}{\pi}k + 0.42\big)2^k - \sum_{i=1}^{t-1} 2^{\frac{n}{2^{i+1}}-1} - 2^{\frac{m-1}{2}} - 1, & \text{if } t \ge 2. \end{cases} \end{aligned}$$

□

From Theorem 6, we see that the nonlinearity of the function F given by Construction 2 decreases a little when compared with the function f in Construction 1. But in this way, we are able to obtain a class of balanced functions achieving (at least potentially) all the characteristics needed for being used in the filter model (maximal algebraic degree, optimal algebraic immunity, and good behavior against fast algebraic attacks and good nonlinearity as we shall see below). It is the first time that a function with all these features and such good nonlinearity is found.

5.4 Comparison with the known results

In this subsection, we compare the lower bounds on the nonlinearity of our result with the known Boolean functions resisting algebraic attacks and fast algebraic attacks. The

class of Carlet-Feng functions have optimal algebraic immunity, maximal algebraic degree, high nonlinearity, and a good behavior against the fast algebraic attacks. In [33], the same class was presented by Wang *et al* in another way (as shown in [7]) with a very slightly improved bound on the nonlinearity. Very recently, Zeng *et al* [35] presented more balanced functions by a modification of this class of functions.

There is a big gap between the lower bound on the nonlinearity of Carlet-Feng functions and their exact values. Herein, we intend to improve the lower bound of Carlet-Feng functions with the same method as the one used in Section 5.3. To do so, we need the following Lemma.

Lemma 5. *For every integer $n \geq 2$. Let $T = \sum_{\mu=1}^{2^n-2} \frac{|\sin(\frac{\pi\mu 2^{n-1}}{2^n-1})|}{\sin(\frac{\pi\mu}{2^n-1})}$, then we have $T \leq \frac{2(2^n-1)}{\pi}\Big(1+\frac{n}{2}\ln 2\Big) + \frac{\pi(2^{n-1}+1)}{8}$.*

***Proof*:** From the definition of T, we have

$$\begin{aligned}
T &= \sum_{\mu=1}^{2^n-2} \frac{|\sin(\frac{\pi\mu(2^{n-1}-\frac{1}{2})+\frac{1}{2}\pi\mu}{2^n-1})|}{\sin(\frac{\pi\mu}{2^n-1})} \\
&= \sum_{\mu=1}^{2^n-2} \frac{|\sin(\frac{\pi\mu}{2}+\frac{\pi\mu}{2(2^n-1)})|}{\sin(\frac{\pi\mu}{2^n-1})} \\
&= \sum_{\mu=1}^{2^{n-1}-1} \frac{\sin(\frac{\pi\mu}{2^n-1})}{\sin(\frac{2\pi\mu}{2^n-1})} + \sum_{\mu=0}^{2^{n-1}-2} \frac{\cos(\frac{\pi(2\mu+1)}{2(2^n-1)})}{\sin(\frac{\pi(2\mu+1)}{2^n-1})} \\
&= \frac{1}{2}\sum_{\mu=1}^{2^{n-1}-1} \frac{1}{\cos(\pi\cdot\frac{\mu}{2^n-1})} + \frac{1}{2}\sum_{\mu=0}^{2^{n-1}-2} \frac{1}{\sin(\pi\cdot\frac{2\mu+1}{2(2^n-1)})} \\
&= \frac{1}{2}\sum_{\mu=1}^{2^{n-1}-1} \frac{1}{\sin(\frac{\pi}{2}-\pi\cdot\frac{\mu}{2^n-1})} + \frac{1}{2}\sum_{\mu=0}^{2^{n-1}-2} \frac{1}{\sin(\pi\cdot\frac{2\mu+1}{2(2^n-1)})} \\
&= \sum_{\mu=0}^{2^{n-1}-2} \frac{1}{\sin(\pi\cdot\frac{2\mu+1}{2(2^n-1)})}
\end{aligned}$$

Applying Corollary 1, we get

$$\begin{aligned}
T &< \frac{2(2^n-1)}{\pi}\sum_{\mu=0}^{2^{n-1}-2}\frac{1}{2\mu+1} + \frac{\pi}{8(2^n-1)}\sum_{\mu=0}^{2^{n-1}-2}(2\mu+1) \\
&< \frac{2(2^n-1)}{\pi}\Big(1+\int_0^{2^{n-1}-2}\frac{1}{2\nu+1}d\nu\Big) + \frac{\pi(2^{n-1}-1)^2}{8(2^n-1)} \\
&< \frac{2(2^n-1)}{\pi}\Big(1+\frac{n}{2}\ln 2\Big) + \frac{\pi(2^{n-1}-1)^2}{8(2^n-1)}.
\end{aligned}$$

□

In fact, the estimation of T defined in Lemma 5 is very important for giving the lower bound on the nonlinearity of Carlet-Feng functions, which has been used in [4, 19, 33, 35]. By our new estimation in Lemma 5, we can improve its lower bound on the nonlinearity. We have the following Theorem.

Theorem 7. *For positive integer n. Let $f \in \mathcal{B}_n$ be the Carlet-Feng function with Supp$(f) = \{0, 1, \beta, \beta^2, \cdots, \beta^{2^{n-1}-2}\}$, where β is a primitive element of the field $\mathbb{F}_{2^n}$. Then we have $N_f > 2^{n-1} - \Big(\frac{n\ln 2}{2\pi} + 0.74\Big)2^{\frac{n}{2}} - 1$.*

***Proof*:** From the proof process of the nonlinearity in [4], we have

$$N_f \geq 2^{n-1} - \frac{2^{\frac{n}{2}}}{2^n-1}\sum_{\mu=1}^{2^n-2}\frac{|\sin(\frac{\pi\mu 2^{n-1}}{2^n-1})|}{\sin(\frac{\pi\mu}{2^n-1})} - \frac{2^n}{2(2^n-1)}$$

By Lemma 5, we can derive that

$$\begin{aligned} N_f &> 2^{n-1} - \Big(\frac{2+n\ln 2}{\pi} + \frac{\pi(2^{n-1}-1)^2}{8(2^n-1)^2}\Big)2^{\frac{n}{2}} - 1 \\ &> 2^{n-1} - \Big(\frac{2+n\ln 2}{\pi} + \frac{\pi(2^{n-1}-1)^2}{8(2(2^{n-1}-1))^2}\Big)2^{\frac{n}{2}} - 1 \\ &> 2^{n-1} - \Big(\frac{n\ln 2}{\pi} + \frac{2}{\pi} + \frac{\pi}{32}\Big)2^{\frac{n}{2}} - 1 \\ &> 2^{n-1} - \Big(\frac{n\ln 2}{\pi} + 0.74\Big)2^{\frac{n}{2}} - 1. \end{aligned}$$

□

Let us denote by N_{CF} the lower bound on the nonlinearity of the function given in [4] and N_F the lower bound given by Construction 2. We give in Table 3 below the concrete values of their lower bounds for integer ranging from 6 to 38. From which it is seen that our improved N_{CF} is better than previous estimations, and further N_F is even better than N_{CF}.

Table 3. Comparison of the known lower bounds of nonlinearity

n	N_{CF} in [4]	Improved N_{CF} in [19]	Improved N_{CF} in [33]	N_{CF} in Theorem 7	N_F in Theorem 6
6	10	12	10	14	20
8	70	79	78	86	102
10	366	396	397	416	458
12	1700	1780	1789	1830	1929
14	7382	7584	7615	7700	7931
16	30922	31409	31496	31673	32195
18	126927	128068	128292	128658	129823
20	515094	517704	518256	519010	521577
22	2076956	2082834	2084143	2085694	2091288
24	8344600	8357672	8360697	8363886	8376003
26	33459185	33487957	33494825	33501375	33527429
28	134012775	134075574	134090943	134104385	134160165
30	536432086	536568193	536602200	536629769	536748573
32	2146548157	2146841390	2146915941	2146972443	2147224628
34	8587947933	8588576434	8588738609	8588854346	8589387659
36	34355533697	34356874769	34357225267	34357462201	34358586905
38	137430081424	137432931707	137433684998	137434169787	137436534902

As we mentioned before that the exact value of nonlinearity is much better than what the lower bound gives. Finally, we compare the exact nonlinearity of the Boolean functions F generated by Construction 2 with the Carlet-Feng functions for some small numbers of variables. Let us denoted by $\mathcal{N}_{CF}$ and $\mathcal{N}_F$ the exact nonlinearity of the Carlet-Feng functions and Construction 2 respectively. For even n ranging from 4 to 26, their exact values are given in Table 4 (we also list the exact values of $\mathcal{N}_F$ for even n ranging from 28 to 38).

Table 4. The exact values of the nonlinearity of F

n	4	6	8	10	12	14
$2^{n-1}-2^{n/2}$	4	24	112	480	1984	8064
$\mathcal{N}_{CF}$	4	24	112	484	1970	8036
$\mathcal{N}_F$	4	22	108	476	1982	8028
n	16	18	20	22	24	26
$2^{n-1}-2^{n/2}$	32512	130560	523264	2095104	8384512	33546240
$\mathcal{N}_{CF}$	32530	130442	523154	2094972	8384536	33545716
$\mathcal{N}_F$	32508	130504	523144	2094980	8384490	33545992
n	28	30	32	34	36	38
$2^{n-1}-2^{n/2}$	134201344	536838144	2147418112	8589803520	34359476224	137438429184
$\mathcal{N}_F$	134201460	536838052	2147416552	8589818968	34359469052	137438441620

6 Conclusion

For an integer $k \geq 2$, a class of $2k$-variable unbalanced Boolean functions and a class of $2k$-variable balanced Boolean functions are proposed in this paper. The functions in the former class have high nonlinearity and good algebraic degree; The functions in the latter class are balanced have maximal algebraic and high nonlinearity. Based on a new conjecture, both of them have optimal algebraic immunity and good behavior against the fast algebraic attacks. We also further improved the lower bound on the nonlinearity of the Carlet-Feng functions. Even compared with this new lower bound, our two classes of functions possess the highest lower bounds. It is the first time that a class of functions is found with such good characteristics (at least for those which could be computed for small numbers of variables) for the filter model of pseudo-random generators.

Acknowledgement We wish to thank Simon Fischer for allowing us to use his program computing the resistance of functions to fast algebraic attacks.

References

[1] F. Armknecht, "Improving fast algebraic attacks," in *Fast Software Encryption 2004 (Lecture Notes in Computer Science)*, 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 *Advances in Cryptology-EUROCRYPT 2006 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 4004, pp. 147-164, 2006.

[3] 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, no. 7, pp. 3105-3121, 2006.

[4] 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 *Advances in Cryptology-ASIACRYPT 2008 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 5350, pp. 425-440, 2008.

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

[6] 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), pages 257-397, 2010.

[7] C. Carlet, "Comment on 'Constructions of Cryptographically Significant Boolean Functions Using Primitive Polynomials'," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4852-4853, 2011.

[8] 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.

[9] N. Courtois and W. Meier, "Algebraic attacks on stream ciphers with linear feedback," in *Advances in Cryptology - EUROCRYPTO 2003 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 2656, pp. 345-359, 2003.

[10] N. Courtois, "Fast algebraic attacks on stream ciphers with linear feedback", in *Advances in Cryptology - CRYPTO 2003 (Lecture Notes in Computer Science)*, Springer-Verlag, 2003, vol. 2729, pp. 176-194.

[11] T.W. Cusick, Y. Li, and P. Stănică, "On a combinatoric conjecture," Cryptology ePrint Archive, Report 2009/554, 2009. http: //eprint.iacr.org/2009/554.pdf.

[12] 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, no. 1, pp. 41-58, 2006.

[13] C. Ding, G. Xiao, and W. Shan, *The Stability Theory of Stream Ciphers*, Springer-Verlag, 1991.

[14] H. Dobbertin. Construction of bent functions and balanced Boolean functions with high nonlinearity. *Proceedings of Fast Software Encryption, Second International Workshop*, Lecture Notes in Computer Science 1008, pp. 61-74, 1995.

[15] K. Feng, Q. Liao, and J. Yang, "Maximal values of generalized algebraic immunity," *Designs, Codes and Cryptography* Vol. 50, No. 2, pp. 243 - 252, 2009 .

[16] J.P. Flori, H. Randriambololona, G. Cohen, and S. Mesnager, "On a conjecture about binary strings distribution," In*SETA 2010 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 6338, pp. 346-358, 2010.

[17] J.P. Flori and H. Randriambololona, "On the Number of Carries Occuring in an Addition mod 2^k-1," Cryptology ePrint Archive, Report 2011/245, 2011. http: //eprint.iacr.org/.

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

[19] R.M. Hakala and K. Nyberg, "On the Nonlinearity of the Discrete Logarithm in F_{2^n}," In *SETA 2010 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 6338, pp. 333-345, 2010.

[20] N. Li and W. Qi, "Construction and analysis of boolean functions of 2t+1 variables with maximum algebraic immunity," In *Advances in Cryptology - Asiacrypt 2006 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 4284, pp. 84-98, 2006.

[21] 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.

[22] R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, vol. 20, Addison-Wesley, Reading, Massachussetts (1983)

[23] M. Lobanov, "Tight bound between nonlinearity and algebraic immunity," Cryptology ePrint Archive, Report 2005/441, 2005. http: //eprint.iacr.org/2005/441.pdf

[24] J.L. Massey, "Shift-register analysis and BCH decoding," *IEEE Trans. Inf. Theory*, vol. 15, pp. 122-127, 1969.

[25] W. Meier and O. Staffelbach, " Fast correlation attacks on stream ciphers," in *Advances in Cryptology - EUROCRYPT 1988 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 330, pp. 301-314, 1988.

[26] W. Meier, E. Pasalic, and C. Carlet, "Algebraic attacks and decomposition of boolean functions," In *Advances in Cryptology - EUROCRYPT 2004 (Lecture Notes in Computer Science)*, vol. 3027, pp. 474-491, 2004.

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

[28] S. Rønjom and T. Helleseth, "A new attack on the filter generator," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1752-1758, 2007.

[29] C.E. Shannon, "Communication theory of secrecy systems," *Bell system technical journal*, 28, pp. 656-715, 1949.

[30] M.R. Spiegel, *Mathematical Handbook of Formulas and Tables*, McGraw-Hill, New York (1968).

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

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

[33] Q. Wang, J. Peng, H. Kan, and X. Xue, "Constructions of cryptographically significant Boolean functions using primitive polynomials," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 3048-3053, 2010.

[34] Q. Wang and T. Johansson, "A note on fast algebraic attacks and higher order nonlinearities," In *INSCRYPT 2010 (Lecture Notes in Computer Science)*, Springer-Verlag, vol. 6584, pp. 84-98, 2010.

[35] X. Zeng, C. Carlet, J. Shan, and L. Hu, "Balanced Boolean Functions with Optimum Algebraic Immunity and High Nonlinearity ," Cryptology ePrint Archive. Report 2010/534, 2010. http://eprint.iacr.org/2010/534

[36] X. Zeng and L. Hu, "Constructing boolean functions by modifying Maiorana-McFarland's superclass functions," *IEICE Trans. Fundamentals*, vol. 88-A, no. 1, pp. 59-66 (2005).