Almost Perfect Algebraic Immune Functions with Good Nonlinearity

Meicheng Liu and Dongdai Lin

State Key Laboratory of Information Security, Institute of Information Engineering,
Chinese Academy of Sciences, Beijing 100195, P. R. China

Abstract. In the last decade, algebraic and fast algebraic attacks are regarded as the most successful attacks on LFSR-based stream ciphers. Since the notion of algebraic immunity was introduced, the properties and constructions of Boolean functions with maximum algebraic immunity have been researched in a large number of papers. However, there are few results with respect to Boolean functions with provable good immunity against fast algebraic attacks. In previous literature, only Carlet-Feng function, which is affine equivalent to discrete logarithm function, was proven to be optimal against fast algebraic attacks as well as algebraic attacks.

In this paper, it is proven that a family of $2k$-variable Boolean functions, including the function recently constructed by Tang et al. [IEEE TIT 59(1): 653–664, 2013], are almost perfect algebraic immune for any integer $k \geq 3$. More exactly, they achieve optimal algebraic immunity and almost perfect immunity to fast algebraic attacks. The functions of such family are balanced and have optimal algebraic degree. A lower bound on their nonlinearity is obtained based on the work of Tang et al., which is better than that of Carlet-Feng function. It is also checked for $3 \leq k \leq 9$ that the exact nonlinearity of such functions is very good, which is slightly smaller than that of Carlet-Feng function, and some functions of this family even have a slightly larger nonlinearity than Tang et al.'s function. To sum up, among the known functions with provable good immunity against fast algebraic attacks, the functions of this family make a trade-off between the exact value and the lower bound of nonlinearity.

Keywords: Boolean functions, Fast algebraic attacks, Algebraic immunity, Perfect algebraic immune, Nonlinearity.

1 Introduction

Boolean functions are frequently used in the design of stream ciphers, block ciphers and hash functions. One of the most vital roles in cryptography of Boolean functions is to be used as filter and combination generators of stream ciphers based on linear feedback shift registers (LFSR). The study of the cryptographic criteria of Boolean functions is important because of the connections between known cryptanalytic attacks and these criteria.

In recent years, algebraic and fast algebraic attacks [1,7,6] have been regarded as the most successful attacks on LFSR-based stream ciphers. These attacks cleverly use over-defined systems of multi-variable nonlinear equations to recover the secret key. Algebraic attacks lower the degree of the equations by multiplying a nonzero function; fast algebraic attacks obtain equations of small degree by linear combination.

Thus the algebraic immunity ($\mathcal{AI}$), the minimum algebraic degree of annihilators of f or $f+1$, was introduced by W. Meier et al. [18] to measure the ability of Boolean functions to resist algebraic attacks. It was shown by N. Courtois and W. Meier [7] that maximum $\mathcal{AI}$ of n-variable Boolean functions is $\lceil \frac{n}{2} \rceil$. Constructions of Boolean functions with maximum $\mathcal{AI}$ were researched in a large number of papers, e.g., [9,15,14,4,21,23]. However, there are few results referring to constructions of Boolean functions with provable good immunity against fast algebraic attacks.

A preprocessing of fast algebraic attacks on LFSR-based stream ciphers, which use a Boolean function $f : GF(2)^n \to GF(2)$ as the filter or combination generator, is to find a function g of small degree such that the multiple gf has degree not too large. The resistance against fast algebraic attacks is not covered by algebraic immunity [8,2,16]. At Eurocrypt 2006, F. Armknecht et al. [2] introduced an effective algorithm for determining the immunity against fast algebraic attacks, and showed that a class of symmetric Boolean functions (the majority functions) have poor resistance against fast algebraic attacks despite their resistance against algebraic attacks. Later M. Liu et al. [16] stated that almost all the symmetric functions including these functions with good algebraic immunity behave badly against fast algebraic attacks.

In [6] N. Courtois proved that for any pair of positive integers (e, d) such that $e+d \geq n$, there is a nonzero function g of degree at most e such that gf has degree at most d. This result reveals an upper bound on maximum immunity to fast algebraic attacks. It implies that the function f has maximum possible resistance against fast algebraic attacks, if for any pair of positive integers (e, d) such that $e+d < n$ and $e < n/2$, there is no nonzero function g of degree at most e such that gf has degree at most d. Such functions are said to be perfect algebraic immune ($\mathcal{PAI}$) [17]. Note that one can use the fast general attack by splitting the function into two $f = h + l$ with l being the linear part of f [6]. In this case, e equals 1 (i.e. the degree of the linear function l) and d equals the degree of h (i.e. the degree of f), where g can be considered as the nonzero constant. Thus $\mathcal{PAI}$ functions have algebraic degree at least $n-1$. A $\mathcal{PAI}$ function also achieves maximum $\mathcal{AI}$. As a consequence, a $\mathcal{PAI}$ function has perfect immunity against classical and fast algebraic attacks. Besides, it is shown that a perfect algebraic immune function behaves good against probabilistic algebraic attacks as well [17]. Although preventing classical and fast algebraic attacks is not sufficient for resisting algebraic attacks on the augmented function [11], the resistance against these attacks depends on the update function and tap positions used in a stream cipher and in actual fact it is not a property of the Boolean function. In [17] M. Liu et al. proved that there are n-variable $\mathcal{PAI}$ functions if and only if $n = 2^s$ or $2^s + 1$. More precisely, there exist n-variable $\mathcal{PAI}$ functions with degree $n - 1$ (balanced functions) if and only if $n = 2^s + 1$; there exist n-variable $\mathcal{PAI}$ functions with degree n (unbalanced functions) if and only if $n = 2^s$.

Several classes of Boolean functions, e.g., [4,23,19,20], are observed through computer experiments to have good behavior against fast algebraic attacks, but in previous literature only Carlet-Feng function (see [10,4]), which is affine equivalent to discrete logarithm function [12], was proven in [17] to be optimal against fast algebraic attacks as well as algebraic attacks. The results of [17] imply that Carlet-Feng function is $\mathcal{PAI}$ for $n = 2^s + 1$ and is almost $\mathcal{PAI}$ for $n \neq 2^s + 1$.

In this paper, we investigate the cryptographic properties, especially in terms of immunity to fast algebraic attacks, for a large family of $2k$-variable functions which has a form as

$$F(x, y) = \phi(xy) + (x^{2^k-1} + 1)\psi(y) + (y^{2^k-1} + 1)\varphi(x),$$

where ϕ is a Carlet-Feng function from $\mathbb{F}_{2^k}$ into $\mathbb{F}_2$ and ψ and φ are Boolean functions from $\mathbb{F}_{2^k}$ into $\mathbb{F}_2$. The balanced function recently proposed by D. Tang et al. [20], which has a form as $\phi(xy)+(x^{2^k-1}+1)\psi(y)$, is contained in this class. Based on bivariate polynomial representation, it is proven that a Boolean function f admits no nonzero function g of degree at most e such that the product gf has degree at most d if and only if the matrix $B(f; e, d)$, whose elements are represented by the coefficients of the bivariate polynomial representation of the function f, has full column rank. After appropriate row transformations, the matrix $B(F; e, d)$ can be represented by

$$\begin{pmatrix} * \\ B^*(\phi(xy); e, d) \end{pmatrix},$$

where $B^*(\phi(xy); e, d)$ is a submatrix of $B(\phi(xy); e, d)$. After appropriate matrix transformations, the matrix $B(\phi(xy); e, d)$ can be transformed into a quasidiagonal matrix. Using the method treating Carlet-Feng functions in [17], it is shown that to ensure that the matrix $B^*(\phi(xy); e, d)$ has full column rank one only need to ensure the number of rows is greater than or equal to the number of columns of the submatrices. Based on the mentioned properties, we prove that the family of the functions F are almost $\mathcal{PAI}$, i.e., they achieve optimal algebraic immunity and almost perfect immunity against fast algebraic attacks. Since the function of Tang et al. falls into this family, it is also almost $\mathcal{PAI}$.

The functions of such family are balanced and have optimal algebraic degree. A lower bound on their nonlinearity is obtained by applying a similar method of [20]. This bound is better than that of Carlet-Feng function, and is slightly worse than that of Tang et al.'s function. It is also checked for $3 \leq k \leq 9$ that the functions of this family have very good nonlinearity, which is a little smaller than that of Carlet-Feng function, and the exact nonlinearity of some functions of this family is slightly larger than that of Tang et al.'s function. Among the known functions with provable good immunity against fast algebraic attacks, the functions of this family make a trade-off between the exact value and the lower bound of nonlinearity.

The remainder of this paper is organized as follows. In Section 2 some basic concepts and results are provided. Section 3 studies the cryptographic properties of the function F. The bivariate polynomial representation and algebraic degree are discussed in Section 3.1, the immunity to algebraic and fast algebraic attacks in Section 3.2, and the nonlinearity in Section 3.3. Section 4 concludes the paper.

2 Preliminary

Let $\mathbb{F}_2$ denote the binary field $GF(2)$ and $\mathbb{F}_2^n$ the n-dimensional vector space over $\mathbb{F}_2$. An n-variable Boolean function is a mapping from $\mathbb{F}_2^n$ into $\mathbb{F}_2$. Denote by $\mathbf{B}_n$ the set of all n-variable Boolean functions. An n-variable Boolean function f can be uniquely represented as its truth table, i.e., a binary string of length 2^n,

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

The support of f is given by $\text{supp}(f) = \{x \in \mathbb{F}_2^n \mid f(x) = 1\}$. The Hamming weight of f, denoted by $\text{wt}(f)$, is the number of ones in the truth table of f. An n-variable function f is said to be balanced if its truth table contains equal number of zeros and ones, that is, $\text{wt}(f) = 2^{n-1}$. The Hamming distance between n-variable functions f and g, denoted by $\text{d}(f,g)$, is the number of $x \in \mathbb{F}_2^n$ at which $f(x) \neq g(x)$. It is well known that $\text{d}(f,g) = \text{wt}(f+g)$.

An n-variable Boolean function f can also be uniquely represented as a multivariate polynomial over $\mathbb{F}_2$,

$$f(x_1,\cdots,x_n) = \sum_{c=(c_1,\cdots,c_n)\in\mathbb{F}_2^n} \lambda_c \prod_{i=1}^{n} x_i^{c_i},\ \lambda_c \in \mathbb{F}_2,$$

called the algebraic normal form (ANF). The algebraic degree of f, denoted by $\deg(f)$, is defined as $\max\{\text{wt}(c) \mid \lambda_c \neq 0\}$.

Let $\mathbb{F}_{2^n}$ denote the finite field $GF(2^n)$. The Boolean function f considered as a mapping from $\mathbb{F}_{2^n}$ into $\mathbb{F}_2$ can be uniquely represented as

$$f(x) = \sum_{i=0}^{2^n-1} a_i x^i,\ a_i \in \mathbb{F}_{2^n}, \tag{1}$$

where $f^2(x) \equiv f(x)(\bmod\, x^{2^n} - x)$. Expression (1) is called the univariate polynomial representation of the function f. It is well known that $f^2(x) \equiv f(x)(\bmod\, x^{2^n} - x)$ if and only if $a_0, a_{2^n-1} \in \mathbb{F}_2$ and for $1 \leq i \leq 2^n - 2$, $a_{2i \bmod (2^n-1)} = a_i^2$. The algebraic degree of the function f equals $\max\limits_{a_i\neq 0} \text{wt}(i)$, where $i = \sum_{k=1}^{n} i_k 2^{k-1}$ is considered as $(i_1, i_2, \cdots, i_n) \in \mathbb{F}_2^n$.

Let α be a primitive element of $\mathbb{F}_{2^n}$. The a_i's of Expression (1) are given by $a_0 = f(0)$, $a_{2^n-1} = f(0) + \sum_{j=0}^{2^n-2} f(\alpha^j)$ and

$$a_i = \sum_{j=0}^{2^n-2} f(\alpha^j)\alpha^{-ij},\ \text{for } 1 \leq i \leq 2^n - 2. \tag{2}$$

Let $n = n_1 + n_2$ $(n_1 \leq n_2)$ and denote by $\mathrm{lcm}(n_1, n_2)$ the least common multiple of positive integers n_1 and n_2. The Boolean function f considered as a mapping from $\mathbb{F}_{2^{n_1}} \times \mathbb{F}_{2^{n_2}}$ into $\mathbb{F}_2$ can be uniquely represented as

$$f(x, y) = \sum_{i=0}^{2^{n_1}-1} \sum_{j=0}^{2^{n_2}-1} a_{ij} x^i y^j, \ a_{ij} \in \mathbb{F}_{2^{\mathrm{lcm}(n_1,n_2)}}, \tag{3}$$

where $f^2(x, y) \equiv f(x, y)(\mathrm{mod}(x^{2^{n_1}} - x, y^{2^{n_2}} - y))$. Expression (3) is called the bivariate polynomial representation of the function f. We can see that $f^2(x, y) \equiv f(x, y)(\mathrm{mod}(x^{2^{n_1}} - x, y^{2^{n_2}} - y))$ if and only if $a_{2^{n_1}-1, 2^{n_2}-1} \in \mathbb{F}_2$ and for $0 \leq i \leq 2^{n_1} - 2$ and $0 \leq j \leq 2^{n_2} - 2$,

$$\begin{aligned} a_{2i,2j} &= a_{ij}^2, \\ a_{2^{n_1}-1,2j} &= a_{2^{n_1}-1,j}^2, \\ a_{2i,2^{n_2}-1} &= a_{i,2^{n_2}-1}^2, \end{aligned} \tag{4}$$

where $2i$ and $2j$ are considered as $2i \bmod (2^{n_1} - 1)$ and $2j \bmod (2^{n_2} - 1)$ respectively, which implies $a_{0,0}, a_{0,2^{n_2}-1}, a_{2^{n_1}-1,0} \in \mathbb{F}_2$. The algebraic degree of the function f equals $\max\limits_{a_{ij} \neq 0} \{\mathrm{wt}(i) + \mathrm{wt}(j)\}$.

In particular, for $n = 2k$, the Boolean function f considered as a mapping from $\mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$ into $\mathbb{F}_2$ can be uniquely represented as

$$f(x, y) = \sum_{i=0}^{2^k-1} \sum_{i=0}^{2^k-1} a_{ij} x^i y^j, \ a_{ij} \in \mathbb{F}_{2^k}, \tag{5}$$

where $f^2(x, y) \equiv f(x, y)(\mathrm{mod}(x^{2^k} - x, y^{2^k} - y))$.

Many properties of Boolean functions can be described by the Walsh spectra. For $x = (x_1, x_2, \cdots, x_n) \in \mathbb{F}_2^n$ and $w = (w_1, w_2, \cdots, w_n) \in \mathbb{F}_2^n$, let $w \cdot x = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \in \mathbb{F}_2$. The Walsh transform of the Boolean function f is an integer valued function over $\mathbb{F}_2^n$ which is defined as

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

The nonlinearity of f, defined as the minimum Hamming distance between f and the set of affine functions, can be given by

$$\mathcal{NL}(f) = 2^{n-1} - \frac{1}{2} \max_{w \in \mathbb{F}_2^n} |W_f(w)|.$$

A high nonlinearity is surely one of the most important cryptographic criteria.

The algebraic immunity of Boolean functions is defined as follows. Maximum algebraic immunity of n-variable Boolean functions is $\lceil \frac{n}{2} \rceil$ [7].

Definition 1 *[18] The algebraic immunity of a function $f \in \mathbf{B}_n$, denoted by $\mathcal{AI}(f)$, is defined as*

$$\mathcal{AI}(f) = \min\{\deg(g) \mid gf = 0 \textit{ or } g(f + 1) = 0, 0 \neq g \in \mathbf{B}_n\}.$$

If there is a nonzero Boolean function g with degree at most e such that the product gf has degree at most d, with e small and d not too large, then the Boolean function f is considered to be weak against fast algebraic attacks. The exact values of e and d for which a fast algebraic attack is feasible depend on several parameters, like the size of the memory and the key size of the stream cipher [6,13].

Theorem 1 *[17] Let $f \in \mathbf{B}_n$. If $\deg(f) < n$, then for $e < n/2$ such that $\binom{n-1}{e} \equiv 1(\mathrm{mod}\, 2)$, there exists a nonzero function g with degree at most e such that the product gf has degree at* *most $n-e-1$. Further, if $n \neq 2^s+1$ and $\deg(f) < n$, then there exist a positive integer $e < n/2$ and a nonzero function g with degree at most e such that the product gf has degree at most $n-e-1$.*

If $\deg(f) = n$, then for $e < n/2$ such that $\binom{n-1}{e} \equiv 0 (\mathrm{mod}\, 2)$, there exists a nonzero function g with degree at most e such that the product gf has degree at most $n-e-1$. Further, if $n \neq 2^s$ and $\deg(f) = n$, then there exist a positive integer $e < n/2$ and a nonzero function g with degree at most e such that the product gf has degree at most $n-e-1$.

The bounds of Theorem 1 can be achieved by Carlet-Feng function and modified Carlet-Feng function (see also [17]).

Definition 2 *Let f be an n-variable Boolean function. The function f is said to be almost perfect algebraic immune ($\mathcal{APAI}$) if for any positive integer $e < \frac{n-1}{2}$ the function f admits no nonzero function g of degree at most e such that gf has degree at most $n-e-2$.*

From the above definition, an $\mathcal{APAI}$ function has at least sub-optimal algebraic immunity (i.e. $\mathcal{AI} \geq \lceil \frac{n}{2} \rceil - 1$) for odd n and achieves optimal algebraic immunity for even n, since $\mathcal{AI}(f) > e$ if and only if there exists no nonzero function g of degree at most e such that gf has degree at most e.

2.1 Immunity of Boolean functions against fast algebraic attacks using bivariate polynomial representation

In this section we focus on the immunity of Boolean functions against fast algebraic attacks using bivariate polynomial representation.

For $0 \leq x, y \leq 2^n - 1$, we define $+_n$ and $-_n$ as

$$x +_n y = \begin{cases} 2^n - 1, & \text{if } x + y = 2^n - 1, \\ (x+y) \bmod (2^n - 1), & \text{otherwise,} \end{cases}$$

$$x -_n y = \begin{cases} 2^n - 1, & \text{if } x = 2^n - 1 \text{ and } y = 0, \\ (x-y) \bmod (2^n - 1), & \text{otherwise.} \end{cases}$$

Let

$$\mathcal{W}_e = \{(u,v) |\, \mathrm{wt}(u) + \mathrm{wt}(v) \leq e, 0 \leq u \leq 2^{n_1} - 1, 0 \leq v \leq 2^{n_2} - 1\},$$

$$\overline{\mathcal{W}}_d = \{(a,b) |\, \mathrm{wt}(a) + \mathrm{wt}(b) \geq d + 1, 0 \leq a \leq 2^{n_1} - 1, 0 \leq b \leq 2^{n_2} - 1\}.$$

For $(a,b) \in \mathcal{W}_{n_1+n_2}$ and $(u,v) \in \mathcal{W}_{n_1+n_2}$, $a \circ_{n_1} u$ and $b \circ_{n_2} v$ will be simply denoted by $a \circ u$ and $b \circ v$ respectively if there is no ambiguity, where "$\circ$" denotes the operations "$+$" and "$-$"; that is, the monomial $x^{a \circ t}$ and the monomial $y^{b \circ v}$ are considered as $x^{a \circ u} \bmod (x^{2^{n_1}} - x)$ and $y^{b \circ v} \bmod (y^{2^{n_2}} - y)$ respectively.

Let f, g, h be $(n_1 + n_2)$-variable functions and g be a function of algebraic degree at most e satisfying that $h = gf$ has algebraic degree at most d, where $n_1 \leq n_2$, $e < \frac{n_1+n_2}{2}$ and $e \leq d$. Let

$$f(x,y) = \sum_{i=0}^{2^{n_1}-1} \sum_{j=0}^{2^{n_2}-1} f_{ij} x^i y^j,\ f_{ij} \in \mathbb{F}_{2^{\mathrm{lcm}(n_1,n_2)}},$$

$$g(x,y) = \sum_{(i,j) \in \mathcal{W}_e} g_{ij} x^i y^j,\ g_{ij} \in \mathbb{F}_{2^{\mathrm{lcm}(n_1,n_2)}},$$

and

$$h(x,y) = \sum_{(i,j) \in \mathcal{W}_d} h_{ij} x^i y^j,\ h_{ij} \in \mathbb{F}_{2^{\mathrm{lcm}(n_1,n_2)}}$$

be the bivariate polynomial representations of f, g and h respectively. For $(a,b)\in\overline{\mathcal{W}}_d$, we have $h_{a,b}=0$ and thus

$$0 = h_{a,b} = \sum_{(u,v)\in\mathcal{W}_e} \lambda^f_{(a,b),(u,v)} g_{u,v}, \tag{6}$$

where $(a,b)\neq(u,v)$ (since $\mathcal{W}_e\cap\overline{\mathcal{W}}_d=\emptyset$ for $e\leq d$) and

$$\lambda^f_{(a,b),(u,v)} = \begin{cases} 0, & \text{if } a=0, u\neq 0 \text{ or } b=0, v\neq 0,\\ f_{0,b-v}+f_{2^{n_1}-1,b-v}, & \text{if } a=u\neq 0, b\neq 0, b\neq v,\\ f_{a-u,0}+f_{a-u,2^{n_2}-1}, & \text{if } a\neq 0, a\neq u, b=v\neq 0,\\ f_{a-u,b-v}, & \text{otherwise.} \end{cases} \tag{7}$$

The system of Equations (6) on $g_{u,v}$'s is homogeneous linear. Denote by $B(f;e,d)$ the coefficient matrix of the equations, that is,

$$B(f;e,d) = \left(\lambda^f_{(a,b),(u,v)}\right)_{\substack{(a,b)\in\overline{\mathcal{W}}_d\\ (u,v)\in\mathcal{W}_e}}. \tag{8}$$

The size of the matrix is $\sum_{i=d+1}^{n_1+n_2}\binom{n_1+n_2}{i}\times\sum_{i=0}^{e}\binom{n_1+n_2}{i}$.

Theorem 2 *Let n_1,n_2,e and d be positive integers such that $n_1\leq n_2$, $e<\frac{n_1+n_2}{2}$ and $e\leq d$. Let $f\in\mathbf{B}_{n_1+n_2}:\mathbb{F}_{2^{n_1}}\times\mathbb{F}_{2^{n_2}}\to\mathbb{F}_2$ and $B(f;e,d)$ be the matrix defined as (8). Then there exists no nonzero function g of degree at most e such that the product gf has degree at most d if and only if the matrix $B(f;e,d)$ has full column rank.*

Proof. If the matrix $B(f;e,d)$ has full column rank, i.e., the rank of $B(f;e,d)$ equals the number of $g_{u,v}$'s, then Equations (6) has no nonzero solution and thus f admits no nonzero function g of algebraic degree at most e such that $h=gf$ has algebraic degree at most d.

To prove the "only if" direction of the theorem, we need to show that if the matrix $B(f;e,d)$ has not full column rank, then there always exists a nonzero Boolean function satisfying Equations (6). If $g(x,y)=\sum_{(u,v)\in\mathcal{W}_e} g_{u,v}x^uy^v$ ($g_{u,v}\in\mathbb{F}_{2^{\mathrm{lcm}(n_1,n_2)}}$) satisfies (6), then

$$0 = h^2_{a,b} = \sum_{z\in\mathcal{W}_e}(\lambda^f_{(a,b),(u,v)})^2 g^2_{u,v} = \sum_{(u,v)\in\mathcal{W}_e}\lambda^f_{(2a,2b),(2u,2v)} g^2_{u,v},\ (a,b)\in\overline{\mathcal{W}}_d, \tag{9}$$

showing that $g^2(x,y)=\sum_{(u,v)\in\mathcal{W}_e} g^2_{u,v}x^{2u}y^{2v}$ satisfies (9) (noting that $f_{2i,2j}=f^2_{ij}$ and $\mathrm{wt}(2u)=\mathrm{wt}(u)$ and $\mathrm{wt}(2v)=\mathrm{wt}(v)$). Since (6) and (9) are actually the same equations, we can see that if $g(x,y)$ satisfies Equations (6) then $\mathrm{Tr}(g(x,y))$ satisfies Equations (6), where $\mathrm{Tr}(x)=x+x^2+\cdots+x^{2^{n-1}}$. Also it follows that if $g(x,y)$ satisfies Equations (6) then $\beta g(x,y)$ and $\mathrm{Tr}(\beta g(x,y))$ satisfy Equations (6) for any $\beta\in\mathbb{F}_{2^k}$. If $g(x,y)\neq 0$, then there is $c_x,c_y\in\mathbb{F}_{2^k}$ such that $g(c_x,c_y)=c\neq 0$, and there is $\beta\in\mathbb{F}_{2^k}$ such that $\mathrm{Tr}(\beta c)\neq 0$ and thus $\mathrm{Tr}(\beta g(x,y))\neq 0$. Now we can see that $\mathrm{Tr}(\beta g(x))$ is a nonzero Boolean function and satisfies (6). Hence, if $B(f;e,d)$ has not full column rank, then there exists a nonzero solution for (6) and therefore there exists a nonzero Boolean function satisfying (6).

Thus the theorem is obtained. □

The theorem shows that $\mathcal{AI}(f)>e$ if and only if the matrix $B(f;e,e)$ has full column rank.

3 The functions

Let k be a positive integer and α a primitive element of $\mathbb{F}_{2^k}$. Let ϕ be a univariate polynomial over $\mathbb{F}_{2^k}$ and

$$\phi(x)=\sum_{i=1}^{2^k-2}\frac{1}{1+\alpha^i}x^i. \tag{10}$$

Since $\phi^2 \equiv \phi(\bmod(x^{2^k} - x))$, ϕ is a Boolean function. From the above representation we can see that the algebraic degree of ϕ is equal to $k-1$. Applying $\sum_{x\in\mathbb{F}_{2^k}^*} x = 0$ gives $\phi(1) = \phi(\alpha) = 1$ and $\phi(x) + \phi(\alpha x) = 1$ for $x \notin \mathbb{F}_2$. Therefore, the support of ϕ is $\{1, \alpha, \alpha^3, \alpha^5, \cdots, \alpha^{2^k-3}\}$.

The function $\phi(\alpha x) + 1$ is equal to $\log_\alpha x$, where $\log_\alpha 0 = 1$, and the support of the function $\phi(\alpha^2 x^2) + 1$ is $\{0, 1, \alpha, \alpha^2, \cdots, \alpha^{2^{k-1}-2}\}$. Therefore, the function ϕ is affine equivalent to both discrete logarithm function and Carlet-Feng function.

In recent years, several constructions of Boolean functions with maximum algebraic immunity and good nonlinearity are proposed based on bivariate polynomial representation. The functions constructed by Z. Tu and Y. Deng [21] have the form $\phi(xy^{2^k-2}) + (x^{2^k-1} + 1)\psi(y)$ and the functions constructed by D. Tang et al. [20] have the form $\phi(xy) + (x^{2^k-1} + 1)\psi(y)$. Such functions have good nonlinearity and might have maximum algebraic immunity (depending on whether a binary conjecture is correct[1]). D. Tang et al.'s functions are observed through computer experiments to have good behavior against fast algebraic attacks, but no mathematical results are found in previous literature.

In this section, we study the $2k$-variable Boolean function

$$F(x, y) = \phi(xy) + (x^{2^k-1} + 1)\psi(y) + (y^{2^k-1} + 1)\varphi(x), \tag{11}$$

where ϕ is the function defined as (10), and ψ and φ are Boolean functions from $\mathbb{F}_{2^k}$ into $\mathbb{F}_2$ such that

$$\psi(0) = 0, \max\{\deg(\psi), \deg(\varphi)\} = k - 1 \text{ and } \operatorname{wt}(\psi) + \operatorname{wt}(\varphi) = 2^{k-1}. \tag{12}$$

Example 1 *Let $k \geq 2$ and $m \leq 2^{k-2}$ be positive integers. Let ψ be a k-variable function whose support is $\{\beta^l, \beta^{l+1}, \cdots, \beta^{l+2m-1}\}$ and φ be any k-variable function with Hamming weight of $2^{k-1} - 2m$, where β is a primitive element of $\mathbb{F}_{2^k}$. Then ψ and φ satisfy (12).*

Proof. We just need to show $\max\{\deg(\psi), \deg(\varphi)\} = k - 1$. Since ψ and φ have an even Hamming weight, we know $\max\{\deg(\psi), \deg(\varphi)\} \leq k - 1$. Let $\sum_{i=0}^{2^k-1} \psi_i x^i$ be the univariate polynomial representation of $\psi(x)$. By (2) we have $\psi_{2^n-2} = \sum_{j=0}^{2^k-2} f(\beta^j)\beta^j = \sum_{j=l}^{l+2m-1} \beta^j = \beta^l \frac{1+\beta^{2m}}{1+\beta} \neq 0$, so $\deg(\psi) = k - 1$. Therefore $\max\{\deg(\psi), \deg(\varphi)\} = k - 1$. □

Example 2 *Let $k \geq 3$ be an integer. Let ψ be a k-variable function whose support is $\{\beta^l, \beta^{l+1}, \cdots, \beta^{l+2^{k-2}-1}\}$ and φ be a k-variable function whose support is $\{\gamma^s, \gamma^{s+1}, \cdots, \gamma^{s+2^{k-2}-1}\}$, where β and γ are primitive elements of $\mathbb{F}_{2^k}$. Then ψ and φ satisfy (12).*

Example 3 *Let $k \geq 3$ be an even integer. Let ψ be a k-variable function whose support is $\{\beta^l, \beta^{l+1}, \cdots, \beta^{l+2^{\frac{k}{2}-1}-1}\}$ and φ be a k-variable Bent function, where β is a primitive element of $\mathbb{F}_{2^k}$. Then ψ and φ satisfy (12).*

3.1 Bivariate polynomial representation and algebraic degree

Hereinafter, denote $\phi_0 = \phi_{2^k-1} = 0$ and $\phi_i = \frac{1}{1+\alpha^i}$ for $1 \leq i \leq 2^k - 2$. Let $\sum_{i=0}^{2^k-1}\sum_{i=0}^{2^k-1} \Phi_{ij} x^i y^j$, $\Phi_{ij} \in \mathbb{F}_{2^k}$, be the bivariate polynomial representation of $\phi(xy)$. It is clear that

$$\Phi_{ij} = \begin{cases} \phi_i, & \text{if } 1 \leq i = j \leq 2^k - 2, \\ 0, & \text{otherwise}. \end{cases} \tag{13}$$

Let $\sum_{j=0}^{2^k-1} \psi_j y^j$ and $\sum_{i=0}^{2^k-1} \varphi_i x^i$ be the univariate polynomial representations of $\psi(y)$ and $\varphi(x)$ respectively, $\psi_j, \varphi_i \in \mathbb{F}_{2^k}$. It is clear that $\psi_0 = \psi(0) = 0$. Since $\max\{\deg(\psi), \deg(\varphi)\} =$

[1] The conjecture for D. Tang et al.'s functions was proven in [5].

$k-1$, we have $\psi_{2^k-1} = \varphi_{2^k-1} = 0$. Let $\sum_{i=0}^{2^k-1}\sum_{i=0}^{2^k-1} F_{ij}x^iy^j$ be the bivariate polynomial representation of $F(x, y)$. Then we have

$$F_{ij} = \begin{cases} \psi_j, & \text{if } i \in \{0, 2^k-1\} \text{ and } 1 \le j \le 2^k-2, \\ \varphi_i, & \text{if } 0 \le i \le 2^k-2 \text{ and } j \in \{0, 2^k-1\}, \\ \phi_i, & \text{if } 1 \le i = j \le 2^k-2, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

We can see that the algebraic degree of F is equal to $2k-1$ since $\max\{\deg(\psi), \deg(\varphi)\} = k-1$.

3.2 Immunity against algebraic and fast algebraic attacks

Before stating our main results, we give some useful notations and lemmas.

Hereinafter we consider $n_1 = n_2 = k$ and denote

$$\mathcal{W}_e = \{(u, v)|\, \mathrm{wt}(u) + \mathrm{wt}(v) \le e, 0 \le u, v \le 2^k-1\},$$

$$\overline{\mathcal{W}}_d = \{(a, b)|\, \mathrm{wt}(a) + \mathrm{wt}(b) \ge d+1, 0 \le a, b \le 2^k-1\},$$
$$\overline{\mathcal{W}}_d^* = \{(a, b) \in \overline{\mathcal{W}}_d | 1 \le a, b \le 2^k-2\}.$$

For $0 \le t \le 2^k-2$, let

$$\mathcal{W}_{e,t} = \{(u, v) \in \mathcal{W}_e | v-u \equiv t (\mathrm{mod}\, 2^k-1)\}, \tag{15}$$

$$\overline{\mathcal{W}}_{d,t} = \{(a, b) \in \overline{\mathcal{W}}_d | b-a \equiv t (\mathrm{mod}\, 2^k-1)\}. \tag{16}$$

Let

$$\overline{\mathcal{W}}_{d,0}^* = \overline{\mathcal{W}}_{d,0} \setminus \{(2^k-1, 0), (0, 2^k-1), (2^k-1, 2^k-1)\}, \tag{17}$$

and for $1 \le t \le 2^k-2$, let

$$\overline{\mathcal{W}}_{d,t}^* = \overline{\mathcal{W}}_{d,t} \setminus \{(0, t), (2^k-1-t, 0), (2^k-1, t), (2^k-1-t, 2^k-1)\}. \tag{18}$$

By (16), (17) and (18), it holds that

$$\overline{\mathcal{W}}_{d,t}^* = \overline{\mathcal{W}}_{d,t} \setminus \{(a, b)| a \in \{0, 2^k-1\} \text{ or } b \in \{0, 2^k-1\}\}$$

and thus $\overline{\mathcal{W}}_{d,t}^* \subset \overline{\mathcal{W}}_d^*$. In particular, if $d \ge k-1$, then $\overline{\mathcal{W}}_{d,t}^* = \overline{\mathcal{W}}_{d,t} \setminus \{(2^k-1, t), (2^k-1-t, 2^k-1)\}$ for $t \ne 0$; if $d \ge k$, then $\overline{\mathcal{W}}_{d,0}^* = \overline{\mathcal{W}}_{d,0} \setminus \{(2^k-1, 2^k-1)\}$.

Lemma 1 *Let $k \ge 3$ and $1 \le e \le k-1$. Then*
(1) $\#\overline{\mathcal{W}}_{2k-e-1,t} = \#\mathcal{W}_{e,t}$ for $0 \le t \le 2^k-2$.
(2) $\#\overline{\mathcal{W}}_{2k-e-2,t}^* \ge \#\mathcal{W}_{e,t}$ for $1 \le t \le 2^k-2$.

Proof. (1) Since $(a, b) \in \overline{\mathcal{W}}_{2k-e-1,t}$ if and only if $\mathrm{wt}(a)+\mathrm{wt}(b) \ge 2k-e$ and $b-a \equiv t (\mathrm{mod}\, 2^k-1)$, that is, $\mathrm{wt}(2^k-1-a) + \mathrm{wt}(2^k-1-b) \le e$ and $(2^k-1-a)-(2^k-1-b) \equiv t (\mathrm{mod}\, 2^k-1)$, it follows that $(a, b) \in \overline{\mathcal{W}}_{2k-e-1,t}$ if and only if $(2^k-1-b, 2^k-1-a) \in \mathcal{W}_{e,t}$. Therefore $\#\overline{\mathcal{W}}_{2k-e-1,t} = \#\mathcal{W}_{e,t}$.

(2) Before checking Lemma 1(2), we prove that the following statements are true for $k \le d \le 2k-1$ and $1 \le t \le 2^k-2$.

(2a) If $\mathrm{wt}(t) \ge d-k+2$, then $\#\overline{\mathcal{W}}_{d-1,t}^* - \#\overline{\mathcal{W}}_{d,t}^* \ge 2$; if $\mathrm{wt}(t) = d-k+1$, then $\#\overline{\mathcal{W}}_{d-1,t}^* - \#\overline{\mathcal{W}}_{d,t}^* \ge 1$.

(2b) If $\mathrm{wt}(t) \le 2k-d-2$, then $\#\overline{\mathcal{W}}_{d-1,t}^* - \#\overline{\mathcal{W}}_{d,t}^* \ge 2$; if $\mathrm{wt}(t) = 2k-d-1$, then $\#\overline{\mathcal{W}}_{d-1,t}^* - \#\overline{\mathcal{W}}_{d,t}^* \ge 1$.

First we prove (2a).

If $\mathrm{wt}(t) + k - d$ is even, then there are $\binom{\mathrm{wt}(t)}{(\mathrm{wt}(t)+k-d)/2}$ pairs of integers (t_a, t_b) such that $t_a + t_b = t$, $\mathrm{supp}(t_a) \subset \mathrm{supp}(t)$, $\mathrm{supp}(t_b) \subset \mathrm{supp}(t)$, $\mathrm{wt}(t_a) = (\mathrm{wt}(t) + k - d)/2$ and $\mathrm{wt}(t_b) = (\mathrm{wt}(t)+d-k)/2$. Let $(a, b) = (2^k - 1 - t_a, t_b)$. For $\mathrm{wt}(t) \geq d-k+1 \geq 1$, we know $\mathrm{wt}(t_a) \neq 0$ and $a \neq 2^k - 1$; noting that $\mathrm{wt}(b) = \mathrm{wt}(t_b) < k$, we have $b \neq 2^k - 1$. Then $(a, b) \notin \{(2^k - 1, t), (2^k - 1 - t, 2^k - 1)\}$. Since $b - a \equiv t_a + t_b = t (\mathrm{mod}\, 2^k - 1)$ and $\mathrm{wt}(a) + \mathrm{wt}(b) = k - \mathrm{wt}(t_a) + \mathrm{wt}(t_b) = k - (\mathrm{wt}(t) + k - d)/2 + (\mathrm{wt}(t) + d - k)/2 = d$, we know $(a, b) \in \overline{\mathcal{W}}^*_{d-1,t} \setminus \overline{\mathcal{W}}^*_{d,t}$ and therefore $\#\overline{\mathcal{W}}^*_{d-1,t} - \#\overline{\mathcal{W}}^*_{d,t} \geq \binom{\mathrm{wt}(t)}{(\mathrm{wt}(t)+k-d)/2} \geq 2$ when $\mathrm{wt}(t) \geq d - k + 2$.

If $\mathrm{wt}(t)+k-d$ is odd, then $\mathrm{wt}(t)+k-d-1$ is even and thus there are at least $\binom{\mathrm{wt}(t)-1}{(\mathrm{wt}(t)+k-d-1)/2}$ pairs of nonnegative integers (t_a, t_b) such that $t_a+t_b = t$, $\mathrm{supp}(t_a) \subset \mathrm{supp}(t)$, $\mathrm{supp}(t_b) \subset \mathrm{supp}(t)$, $\mathrm{wt}(t_a) = (\mathrm{wt}(t)+k-d-1)/2$, $\mathrm{wt}(t_b) = (\mathrm{wt}(t)+d+1-k)/2$ and $s+1 \in \mathrm{supp}(t_b)$, where s satisfies that $(s + 1) \bmod k \in \mathrm{supp}(t)$ and $s \notin \mathrm{supp}(t)$ (since $t \neq 2^k - 1$ we can always find such s). Let $(a, b) = (2^k-1-t_a-2^s, t_b-2^s)$. Since $\mathrm{supp}(t_b) \subset \mathrm{supp}(t)$, we know $s \notin \mathrm{supp}(t_a)$ and $s \notin \mathrm{supp}(t_b)$, and therefore $\mathrm{wt}(t_a + 2^s) = \mathrm{wt}(t_a) + 1$ and $\mathrm{wt}(t_b - 2^s) = \mathrm{wt}(t_b)$ (noting that $s + 1 \in \mathrm{supp}(t_b)$), which also shows that $a \neq 2^k-1$ and $b \neq 2^k-1$ and then $(a, b) \notin \{(2^k-1, t), (2^k-1-t, 2^k-1)\}$. Since $b-a \equiv t_a+t_b = t(\mathrm{mod}\, 2^k-1)$ and $\mathrm{wt}(a)+\mathrm{wt}(b) = k-\mathrm{wt}(t_a+2^s)+\mathrm{wt}(t_b-2^s) = k-\mathrm{wt}(t_a)-1 + \mathrm{wt}(t_b) = d$, we know $(a, b) \in \overline{\mathcal{W}}^*_{d-1,t} \setminus \overline{\mathcal{W}}^*_{d,t}$ and then $\#\overline{\mathcal{W}}^*_{d-1,t} - \#\overline{\mathcal{W}}^*_{d,t} \geq \binom{\mathrm{wt}(t)-1}{(\mathrm{wt}(t)+k-d-1)/2}$, which is greater than or equal to 2 when $\mathrm{wt}(t) \geq d-k+3$ and equal to 1 when $\mathrm{wt}(t) = d-k+1$.

Therefore (2a) has been proven. Then we check (2b). Since $(a, b) \in \overline{\mathcal{W}}_{d,t}$ if and only if $(b, a) \in \overline{\mathcal{W}}_{d,2^k-1-t}$, we have $\#\overline{\mathcal{W}}_{d,t} = \#\overline{\mathcal{W}}_{d,2^k-1-t}$, then (2b) is derived from (2a) by replacing t with $2^k - 1 - t$.

Now we prove Lemma 1(2).

By Lemma 1(1) we know $\#\overline{\mathcal{W}}_{2k-e-1,t} = \#\mathcal{W}_{e,t}$, then taking $d = 2k - e - 1$ in (2a) gives $\#\overline{\mathcal{W}}^*_{2k-e-2,t} \geq \#\overline{\mathcal{W}}^*_{2k-e-1,t}+2 \geq \#\mathcal{W}_{e,t}$ for $\mathrm{wt}(t) \geq k-e+1$; similarly, (2b) shows $\#\overline{\mathcal{W}}^*_{2k-e-2,t} \geq \#\mathcal{W}_{e,t}$ for $\mathrm{wt}(t) \leq e - 1$. Therefore we just need to check for $e \leq \mathrm{wt}(t) \leq k - e$ with $e \leq k/2$.

Denote $v_t = (2^k - 1, t)$, $v_{-t} = (2^k - 1 - t, 2^k - 1)$ and $\mathrm{wt}((a, b)) = \mathrm{wt}(a) + \mathrm{wt}(b)$. Then $\mathrm{wt}(v_t) = k + \mathrm{wt}(t)$ and $\mathrm{wt}(v_{-t}) = 2k - \mathrm{wt}(t)$.

For $e < k/2$, if $e < \mathrm{wt}(t) < k - e$, then $\mathrm{wt}(v_t) < 2k - e$ and $\mathrm{wt}(v_{-t}) < 2k - e$, and thus $v_t \notin \overline{\mathcal{W}}_{2k-e-1,t}$ and $v_{-t} \notin \overline{\mathcal{W}}_{2k-e-1,t}$, showing that $\#\overline{\mathcal{W}}^*_{2k-e-2,t} \geq \#\overline{\mathcal{W}}^*_{2k-e-1,t} = \#\overline{\mathcal{W}}_{2k-e-1,t} = \#\mathcal{W}_{e,t}$; if $\mathrm{wt}(t) = e$, then $\mathrm{wt}(v_t) = k+e < 2k-e$ and thus $v_t \notin \overline{\mathcal{W}}_{2k-e-1,t}$, and taking $d = 2k-e-1$ in (2b) gives $\#\overline{\mathcal{W}}^*_{2k-e-2,t} \geq \#\overline{\mathcal{W}}^*_{2k-e-1,t} + 1 \geq \#(\overline{\mathcal{W}}_{2k-e-1,t} \setminus \{v_t\}) = \#\overline{\mathcal{W}}_{2k-e-1,t} = \#\mathcal{W}_{e,t}$; if $\mathrm{wt}(t) = k-e$, then $\mathrm{wt}(v_{-t}) = k+e < 2k-e$ and thus $v_{-t} \notin \overline{\mathcal{W}}_{2k-e-1,t}$, and taking $d = 2k-e-1$ in (2a) gives $\#\overline{\mathcal{W}}^*_{2k-e-2,t} \geq \#\overline{\mathcal{W}}^*_{2k-e-1,t} + 1 \geq \#(\overline{\mathcal{W}}_{2k-e-1,t} \setminus \{v_{-t}\}) = \#\overline{\mathcal{W}}_{2k-e-1,t} = \#\mathcal{W}_{e,t}$.

For $e = k/2$ and $e \leq \mathrm{wt}(t) \leq k - e$ with k even, we have $\mathrm{wt}(t) = k/2$. Then there is s with $0 \leq s \leq k - 1$ such that $\mathrm{wt}(t - 2^s) = \mathrm{wt}(t) = k/2$ and there is s^* with $0 \leq s^* \leq k - 1$ such that $\mathrm{wt}(2^k-1-t-2^{s^*}) = \mathrm{wt}(2^k-1-t) = k/2$. We can check for $k \geq 4$ that $2^k-1-2^s \neq 2^k-1-t-2^{s^*}$, $(2^k-1-2^s, t-2^s) \in \overline{\mathcal{W}}^*_{3k/2-2,t} \setminus \overline{\mathcal{W}}^*_{3k/2-1,t}$ and $(2^k-1-t-2^{s^*}, 2^k-1-2^{s^*}) \in \overline{\mathcal{W}}^*_{3k/2-2,t} \setminus \overline{\mathcal{W}}^*_{3k/2-1,t}$, and therefore $\overline{\mathcal{W}}^*_{3k/2-2,t} \geq \overline{\mathcal{W}}^*_{3k/2-1,t} + 2 \geq \overline{\mathcal{W}}_{3k/2-1,t} = \overline{\mathcal{W}}_{k/2,t}$. □

Denote by $B^*(f; e, d)$ the matrix obtained by selecting rows $(a, b) \in \overline{\mathcal{W}}^*_d$ from $B(f; e, d)$, that is,

$$B^*(f; e, d) = \left(\lambda^f_{(a,b),(u,v)}\right)_{\substack{(a,b)\in\overline{\mathcal{W}}^*_d \\ (u,v)\in\mathcal{W}_e}}.$$

It is clear that $B^*(f; e, d)$ is a submatrix of $B(f; e, d)$.

Let $B(f; e, d; t)$ be the submatrix of $B(f; e, d)$ formed by selecting rows $(a, b) \in \overline{\mathcal{W}}_{d,t}$ and columns $(u, v) \in \mathcal{W}_{e,t}$, that is,

$$B(f; e, d; t) = \left(\lambda^f_{(a,b),(u,v)}\right)_{\substack{(a,b)\in\overline{\mathcal{W}}_{d,t} \\ (u,v)\in\mathcal{W}_{e,t}}}.$$

We can see that $B(f;e,d;t)$ is a $\#\overline{\mathcal{W}}_{d,t} \times \#\mathcal{W}_{e,t}$ matrix, where $\#$ denotes the number of elements in a set. The matrix $B(f;e,d;t)$ is conventionally considered as a full column rank matrix when $\#\mathcal{W}_{e,t} = 0$.

Let $B^*(f;e,d;t)$ be the matrix formed by removing rows $(0,t)$, $(2^k-1-t,0)$, $(2^k-1,t)$, $(2^k-1-t,2^k-1)$ and $(2^k-1,2^k-1)$, if any, from $B(f;e,d;t)$, that is,

$$B^*(f;e,d;t) = \left(\lambda^f_{(a,b),(u,v)}\right)_{\substack{(a,b)\in\overline{\mathcal{W}}^*_{d,t} \\ (u,v)\in\mathcal{W}_{e,t}}}.$$

It is clear that $B^*(f;e,d;t)$ is a submatrix of $B(f;e,d;t)$. Since $\overline{\mathcal{W}}^*_{d,t} \subset \overline{\mathcal{W}}^*_d$, we can see that $B^*(f;e,d;t)$ is also a submatrix of $B^*(f;e,d)$.

Denote

$$\overline{\mathcal{W}}^+_{d,0} = \overline{\mathcal{W}}_d \setminus \bigcup_{t\neq 0} \overline{\mathcal{W}}^*_{d,t}$$

and

$$B^+(f;e,d;0) = \left(\lambda^f_{(a,b),(u,v)}\right)_{\substack{(a,b)\in\overline{\mathcal{W}}^+_{d,0} \\ (u,v)\in\mathcal{W}_{e,0}}}.$$

Lemma 2 *[17] Let*

$$A = \left(\frac{1}{1+\beta_i\gamma_j}\right)_{m\times m}$$

be an $m \times m$ matrix with $\beta_i, \gamma_j \in \mathbb{F}^*_{2^k}$, $\beta_i\gamma_j \neq 1$, $1 \le i,j \le m$. If $\beta_i \neq \beta_j$ and $\gamma_i \neq \gamma_j$ for $i \neq j$, then $\det(A) \neq 0$.

Lemma 3 *Let $k \ge 3$, $1 \le e \le k-1 \le d \le 2k-e-2$. If $B^+(F;e,d;0)$ has full column rank, then $B(F;e,d)$ has full column rank.*

Proof. From (13) we know $\Phi_{ij} \neq 0$ only when $i = j$. Then from (7) we know $\lambda^{\phi(xy)}_{(a,b),(u,v)} \neq 0$ only when $b - v \equiv a - u (\operatorname{mod} 2^k - 1)$. In other words, $\lambda^{\phi(xy)}_{(a,b),(u,v)} \neq 0$ only when $b - a \equiv v - u \equiv t (\operatorname{mod} 2^k - 1)$, $0 \le t \le 2^k - 2$. Therefore, the matrix $B(\phi(xy);e,d)$ is a quasidiagonal matrix as

$$\begin{pmatrix} B(\phi(xy);e,d;0) & 0 & \cdots & 0 \\ 0 & B(\phi(xy);e,d;1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B(\phi(xy);e,d;2^k-2) \end{pmatrix}.$$

For $(u,v) \in \mathcal{W}_e$ and $(a,b) \in \overline{\mathcal{W}}^*_d$ with $a = u$, it holds that $b \neq 2^k - 1$ and $b - v \neq 2^k - 1$, so we have $\lambda^F_{(a,b),(u,v)} = \psi_{b-v} + \psi_{b-v} = 0$ by (7) and (14); and we also have $\lambda^{\phi(xy)}_{(a,b),(u,v)} = 0$ by (7) and (13). For $(u,v) \in \mathcal{W}_e$ and $(a,b) \in \overline{\mathcal{W}}^*_d$ with $b = v$, we similarly have $\lambda^F_{(a,b),(u,v)} = \lambda^{\phi(xy)}_{(a,b),(u,v)} = 0$. For $(u,v) \in \mathcal{W}_e$ and $(a,b) \in \overline{\mathcal{W}}^*_d$ with $a \neq u$ and $b \neq v$, it holds that $a - u \notin \{0, 2^k-1\}$ and $b - v \notin \{0, 2^k-1\}$, and therefore $\lambda^F_{(a,b),(u,v)} = \lambda^{\phi(xy)}_{(a,b),(u,v)}$ by (7), (14) and (13). Thus $B^*(F;e,d) = B^*(\phi(xy);e,d)$. Then, after appropriate matrix transformations, the matrix $B(F;e,d)$ can be represented as

$$\begin{pmatrix} * & * & \cdots & * \\ B^*(\phi(xy);e,d;0) & 0 & \cdots & 0 \\ 0 & B^*(\phi(xy);e,d;1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B^*(\phi(xy);e,d;2^k-2) \end{pmatrix}.$$

By the definition of $B^+(F;e,d;0)$, we can see that the above matrix is

$$\begin{pmatrix} B^+(F;e,d;0) & * & \cdots & * \\ 0 & B^*(\phi(xy);e,d;1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B^*(\phi(xy);e,d;2^k-2) \end{pmatrix}.$$

Thus, we just need to prove that all the matrices $B^*(\phi(xy);e,d;t)$ with $1 \le t \le 2^k-2$ have full column rank.

Next we show for $1 \le t \le 2^k-2$ the matrix $B^*(\phi(xy);e,d;t)$ has full column rank.

For $(a,b) \in \overline{\mathcal{W}}_{d,t}$ and $(u,v) \in \mathcal{W}_{e,t}$, when $a = u$, by (15) and (16) we have $b = 2^k-1$ and $v = 0$ (since $(a,b) \ne (u,v)$), and thus $a = u = (2^k-1-t) \bmod (2^k-1)$, which shows that: for $d+1-k \le \mathrm{wt}(2^k-1-t) \le e$, $a = u$ if and only if $(a,b) = (2^k-1-t, 2^k-1)$ and $(u,v) = (2^k-1-t, 0)$; for the other cases, there do not exist $(a,b) \in \overline{\mathcal{W}}_{d,t}$ and $(u,v) \in \mathcal{W}_{e,t}$ such that $a = u$. Similarly, one can obtain that: for $d+1-k \le \mathrm{wt}(t) \le e$, $b = v$ if and only if $(a,b) = (2^k-1, t)$ and $(u,v) = (0,t)$; for the other cases, there do not exist $(a,b) \in \overline{\mathcal{W}}_{d,t}$ and $(u,v) \in \mathcal{W}_{e,t}$ such that $b = v$.

Therefore, for $(a,b) \in \overline{\mathcal{W}}^*_{d,t}$ and $(u,v) \in \mathcal{W}_{e,t}$, we have $b-v \equiv a-u (\bmod\, 2^k-1)$, $a \notin \{0, 2^k-1\}$, $b \notin \{0, 2^k-1\}$, $a-u \notin \{0, 2^k-1\}$ and $b-v \notin \{0, 2^k-1\}$, then from (7) and (13) we obtain that

$$\lambda^{\phi(xy)}_{(a,b),(u,v)} = \Phi_{a-u,b-v} = \phi_{a-u}.$$

By Lemma 1(2), we have $\#\overline{\mathcal{W}}^*_{d,t} \ge \#\overline{\mathcal{W}}^*_{2k-e-2,t} \ge \#\mathcal{W}_{e,t}$ for $d \le 2k-e-2$. Let $\mathcal{U} = \{u \mid (u,v) \in \mathcal{W}_{e,t}\}$ and $\mathcal{A}$ be an arbitrary subset of $\{a \mid (a,b) \in \overline{\mathcal{W}}^*_{d,t}\}$ such that $\#\mathcal{A} = \#\mathcal{U}$. Let A be the matrix formed by selecting rows $\mathcal{A}$ from $B^*(\phi(xy);e,d;t)$, that is,

$$A = (\phi_{a-u})_{\substack{a\in\mathcal{A}\\ u\in\mathcal{U}}}.$$

For $a \in \mathcal{A}$ and $u \in \mathcal{U}$, we have $1 \le a -_k u \le 2^k-2$, and thus by (10),

$$\phi_{a-u} = \frac{1}{1+\alpha^a\alpha^{-u}}.$$

It is derived from Lemma 2 that $\det(A) \ne 0$. Hence the matrix $B^*(\phi(xy);e,d;t)$ has full column rank. □

Now we prove that the function F is $\mathcal{APAI}$.

Theorem 3 *Let $k \ge 3$. Then the $2k$-variable function F defined as (10) is $\mathcal{APAI}$. That is, for any positive integer e with $e < k$, there is no nonzero function $g \in \mathbf{B}_{2k}$ with* $\deg(g) \le e$ *such that* $\deg(gF) \le 2k-e-2$.

Proof. By Theorem 2 and Lemma 3 we just need to prove the matrix $B^+(F;e,d;0)$ has full column rank. Assume without loss of generality that the univariate polynomial representation of ψ has a monomial y^b with algebraic degree equal to $k-1$, that is, $\mathrm{wt}(b) = k-1$. Let $\psi_b \ne 0$ be the coefficient of y^b in the univariate polynomial representation of ψ, let $\sum_{i=0}^{2^k-1} \phi_i x^i$, $\phi_i \in \mathbb{F}_{2^k}$, be the univariate polynomial representation of ϕ, and let $\sum_{i=0}^{2^k-1}\sum_{i=0}^{2^k-1} F_{ij}x^iy^j$, $F_{ij} \in \mathbb{F}_{2^k}$, be the bivariate polynomial representation of $F(x,y)$. Since

$$\phi(xy) = \sum_{i=0}^{2^k-1} \phi_i x^i y^i$$

and

$$F(x,y) = \phi(xy) + (x^{2^k-1}+1)\psi(y) + (y^{2^k-1}+1)\varphi(x),$$

we have $F_{2^k-1,b}=\psi_b$ and $F_{2^k-1-j,b-j}=0$ for $1\le j\le 2^k-2$ and $j\ne b$ (since $2^k-1-j\ne b-_k j$, $2^k-1-j\notin\{2^k-1,0\}$ and $b-_k j\notin\{2^k-1,0\}$). By (15) we have $\mathcal{W}_{e,0}=\{(u,u)|\,\mathrm{wt}(u)\le\frac{e}{2}\}$. Thus for $(u,v)\in\mathcal{W}_{e,0}$, where $e<k$, we know $u=v$ and $\mathrm{wt}(v)<k/2\le k-1=\mathrm{wt}(b)$, where $k\ge 3$, and thus $u=v\ne 2^k-1$ and $u=v\ne b$. Therefore, for $(u,v)\in\mathcal{W}_{e,0}$, it follows from (7) that $\lambda^F_{(2^k-1,b),(u,v)}=F_{2^k-1-u,b-u}$ and thus, as mentioned above,

$$\lambda^F_{(2^k-1,b),(u,v)}=\begin{cases}\psi_b, & \text{if } (u,v)=(0,0),\\ 0, & \text{otherwise.}\end{cases}$$

Since $\mathrm{wt}(b)=k-1$, we know $(2^k-1,b)\in\overline{\mathcal{W}}_{2k-2}\subset\overline{\mathcal{W}}_d$ and thus $(2^k-1,b)\in\overline{\mathcal{W}}^+_{d,0}$, for $d=2k-e-2$. Since $\psi_b\ne 0$, from the definition of $B^+(F;e,d;0)$ it is sufficient to prove the matrix

$$B^*_*(F;e,d;0)=\left(\lambda^F_{(a,b),(u,v)}\right)_{\substack{(a,b)\in\overline{\mathcal{W}}^*_{d,0}\\(u,v)\in\mathcal{W}^*_{e,0}}}$$

has full column rank, where $\mathcal{W}^*_{e,0}=\mathcal{W}_{e,0}\backslash\{(0,0)\}$. By Lemma 1(1) we have $\#\overline{\mathcal{W}}_{2k-e-1,0}=\#\mathcal{W}_{e,0}$ and thus $\#\overline{\mathcal{W}}^*_{d,0}\ge\#\overline{\mathcal{W}}^*_{2k-e-1,0}=\#\mathcal{W}^*_{e,0}$. The same proof that $B^*(\phi(xy);e,d;t)$ has full column rank (Lemma 3) shows that $B^*_*(F;e,d;0)$ has full column rank. Hence we have proven that the matrix $B^+(F;e,d;0)$ has full column rank. □

Remark 1. The theorem shows that the function F achieves optimal algebraic immunity. The same proof of Theorem 3 gives that for $k=2^m t+1$ with $t>1$ odd, if $k-2^m-1\le\max\{\deg(\psi),\deg(\varphi)\}<k-1$, then the function F is also $\mathcal{APAI}$. In this case, however, the algebraic degree of F is equal to $2k-2$.

Remark 2. Since the balanced function f_2 proposed in [20] is a special case of functions defined as (10), it is also $\mathcal{APAI}$.

3.3 Nonlinearity

Lemma 4 *Let $k\ge 3$. Let ϕ be the k-variable function defined as (10) and Φ the $2k$-variable function $\phi(xy)$. Then*

$$\mathcal{NL}(\Phi)>2^{2k-1}-\frac{k\ln 2+\frac{25}{12}-\ln\frac{9\pi}{8}}{\pi}2^k-1.$$

Proof. For $x>0$, we have $\sin x>x-\frac{x^3}{6}$ by Taylor's theorem. Then, for $0<x<1$, it holds that

$$\frac{1}{x}-\frac{1}{\sin x}+\frac{x}{5}=\frac{\sin x-x+\frac{x^2}{5}\sin x}{x\sin x}>\frac{-\frac{x^3}{6}+\frac{x^2}{5}(x-\frac{x^3}{6})}{x\sin x}=\frac{\frac{x^2}{30}(1-x^2)}{\sin x}>0$$

and thus

$$\frac{1}{\sin x}<\frac{1}{x}+\frac{x}{5}. \tag{19}$$

Then, for $k\ge 3$, we have

$$\begin{aligned}\sum_{\mu=1}^{4}\frac{1}{\sin\frac{\mu\pi}{2(2^k-1)}}&<\sum_{\mu=1}^{4}\left(\frac{2(2^k-1)}{\mu\pi}+\frac{1}{5}\cdot\frac{\mu\pi}{2(2^k-1)}\right)\\&\le\frac{2(2^k-1)}{\pi}\sum_{\mu=1}^{4}\frac{1}{\mu}+\frac{\pi}{70}\sum_{\mu=1}^{4}\mu\\&<\frac{25(2^k-1)}{6\pi}+\frac{1}{2}.\end{aligned} \tag{20}$$

Since for $0 \le \theta < t$ and $t + \theta \le \pi$,

$$\frac{\theta}{\sin t} \le \int_{t-\frac{\theta}{2}}^{t+\frac{\theta}{2}} \frac{\mathrm{d}\,x}{x}, \tag{21}$$

taking $t = \frac{\mu\pi}{2(2^k-1)}$ and $\theta = \frac{\pi}{2(2^k-1)}$ gives

$$\begin{aligned}\sum_{\mu=5}^{2^k-2} \frac{\frac{\pi}{2(2^k-1)}}{\sin\frac{\mu\pi}{2(2^k-1)}} &\le \sum_{\mu=5}^{2^k-2} \int_{\frac{\mu\pi}{2(2^k-1)}-\frac{\pi}{4(2^k-1)}}^{\frac{\mu\pi}{2(2^k-1)}+\frac{\pi}{4(2^k-1)}} \frac{\mathrm{d}\,x}{\sin x} = \int_{\frac{9\pi}{4(2^k-1)}}^{\frac{\pi(2^k-\frac{3}{2})}{2(2^k-1)}} \frac{\mathrm{d}\,x}{\sin x}\\ &< \int_{\frac{9\pi}{4(2^k-1)}}^{\frac{\pi}{2}} \frac{\mathrm{d}\,x}{\sin x} = \Big[\ln(\tan\frac{x}{2})\Big]_{\frac{9\pi}{4(2^k-1)}}^{\frac{\pi}{2}} = -\ln(\tan\frac{9\pi}{8(2^k-1)})\\ &< -\ln(\frac{9\pi}{8(2^k-1)}) < k\ln 2 - \ln\frac{9\pi}{8}.\end{aligned} \tag{22}$$

The proofs of Theorem 3 and Lemma 1 of [20] show that

$$\mathcal{NL}(\varPhi) \ge 2^{2k-1} - \frac{2^k}{2(2^k-1)}\left(1 + \sum_{\mu=1}^{2^k-2} \frac{1}{\sin\frac{\mu\pi}{2(2^k-1)}}\right). \tag{23}$$

Hence, for $k \ge 3$, by (20) and (22) we can see that

$$\begin{aligned}\mathcal{NL}(\varPhi) &\ge 2^{2k-1} - \frac{2^k}{2(2^k-1)}\left(1 + \sum_{\mu=1}^{4} \frac{1}{\sin\frac{\mu\pi}{2(2^k-1)}} + \sum_{\mu=5}^{2^k-2} \frac{1}{\sin\frac{\mu\pi}{2(2^k-1)}}\right)\\ &> 2^{2k-1} - \frac{2^k}{2(2^k-1)}\left(\frac{3}{2} + \frac{25(2^k-1)}{6\pi} + \frac{2(2^k-1)}{\pi}(k\ln 2 - \ln\frac{9\pi}{8})\right)\\ &> 2^{2k-1} - \frac{2^k}{\pi}\left(k\ln 2 + \frac{25}{12} - \ln\frac{9\pi}{8}\right) - 1.\end{aligned}$$

This ends the proof of the lemma. □

In [20], the function $\varPhi(x, y) = \phi(xy)$ was proved to have nonlinearity more than

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

Lemma 4 shows that the nonlinearity of $\varPhi$ is larger than

$$2^{2k-1} - \frac{k\ln 2 + \frac{25}{12} - \ln\frac{9\pi}{8}}{\pi}2^k - 1 \approx 2^{2k-1} - (\frac{k\ln 2}{\pi} + 0.26)2^k,$$

which improved the previous result by a difference of about $0.16 \cdot 2^k$.

Theorem 4 *Let $k \ge 3$. Let F be the $2k$-variable function defined as (11). Then F is balanced and*

$$\mathcal{NL}(F) > 2^{2k-1} - \left(\frac{k\ln 2 + \frac{25}{12} - \ln\frac{9\pi}{8}}{\pi} + \frac{1}{2}\right)2^k - 1 \approx 2^{2k-1} - (\frac{k\ln 2}{\pi} + 0.76)2^k.$$

Proof. Let $\varPhi = \phi(xy)$ be the function of Lemma 4. Since $\mathrm{supp}(\varPhi) = \{(x, y) \mid xy \in \{1, \alpha, \alpha^3, \cdots, \alpha^{2^k-3}\}\}$ and $\psi(0) = 0$, from (11) we have

$$\mathrm{supp}(F) = \mathrm{supp}(\varPhi) \cup \{(0, y) \mid \psi(y) = 1\} \cup \{(x, 0) \mid \varphi(x) = 1\}.$$

It is clear that the three sets on the right side of the above equality are disjoint. Then we can see that

$$\mathrm{wt}(F) = \mathrm{wt}(\Phi) + \mathrm{wt}(\psi) + \mathrm{wt}(\varphi) = (2^k - 1)2^{k-1} + 2^{k-1} = 2^{2k-1},$$

and thus F is balanced.

Since $\mathrm{d}(\Phi, l) = \mathrm{wt}(\Phi + l) \le \mathrm{wt}(F + l) + \mathrm{wt}(F + \Phi) = \mathrm{d}(F, l) + 2^{k-1}$ for any $l \in \mathbf{B}_{2k}$, we have

$$\mathcal{NL}(F) \ge \mathcal{NL}(\Phi) - 2^{k-1}.$$

Then the theorem is derived from Lemma 4. □

To compared the function F with the function ϕ, we focus on the nonlinearity of ϕ. In [4], C. Carlet and K. Feng showed that the nonlinearity of ϕ is more than $2^{n-1} + \frac{2\ln\frac{\pi}{4(2^n-1)}}{\pi}2^{\frac{n}{2}} - 1 \approx 2^{n-1} - \frac{2n\ln 2}{\pi}2^{\frac{n}{2}}$. In [22], Q. Wang et al. proposed another form of the function ϕ and improved the lower bound on the nonlinearity: $\max\{6\lfloor\frac{2^{n-1}}{2n}\rfloor - 2, 2^{n-1} - (\frac{(n-1)\ln 2}{3} + \frac{3}{2})2^{\frac{n}{2}}\}$. At almost the same time, R. M. Hakala and K. Nyberg [12] also obtained a new lower bound $2^{n-1} - \frac{4\ln(2^n-1)+8}{\pi^2}2^{\frac{n}{2}}$ on the nonlinearity of ϕ, through analyzing the nonlinearity of the discrete logarithm in $\mathbb{F}_{2^n}$. Recently, D. Tang et al. [20] further improved the lower bound on the nonlinearity: $2^{n-1} - (\frac{n\ln 2}{\pi} + 0.74)2^{\frac{n}{2}} - 1$. Based on the results of [4] and [20], the following theorem gives our new bound: $2^{n-1} - \frac{n\ln 2 + \frac{8}{3} - \ln\pi}{\pi}2^{\frac{n}{2}} - 1 \approx 2^{n-1} - (\frac{n\ln 2}{\pi} + 0.48)2^{\frac{n}{2}}$. That is, Tang et al.'s lower bound on nonlinearity of Carlet-Feng function is improved by a difference of about $2^{\frac{n}{2}-2}$.

Theorem 5 *Let $n \ge 3$. Let ϕ be an n-variable defined as (10). Then*

$$\mathcal{NL}(\phi) > 2^{n-1} - \frac{n\ln 2 + \frac{8}{3} - \ln\pi}{\pi}2^{\frac{n}{2}} - 1.$$

Proof. By the proof of [4, Theorem 3], we know

$$\mathcal{NL}(\phi) \ge 2^{n-1} - \frac{2^{\frac{n}{2}}}{2^n - 1}\sum_{\mu=1}^{2^n-2}\frac{\left|\sin\frac{\pi\mu(2^{n-1}-1)}{2^n-1}\right|}{\sin\frac{\pi\mu}{2^n-1}} - \frac{2^n}{2(2^n-1)}.$$

The proof of [20, Lemma 3] shows that

$$\sum_{\mu=1}^{2^n-2}\frac{\left|\sin\frac{\pi\mu(2^{n-1}-1)}{2^n-1}\right|}{\sin\frac{\pi\mu}{2^n-1}} = \sum_{\mu=0}^{2^{n-1}-2}\frac{1}{\sin\frac{\pi(2\mu+1)}{2(2^n-1)}}.$$

By (19), for $n \ge 3$, we have

$$\frac{1}{\sin\frac{\pi}{2(2^n-1)}} + \frac{1}{\sin\frac{3\pi}{2(2^n-1)}} < \frac{2(2^n-1)}{\pi} + \frac{1}{5} + \frac{2(2^n-1)}{3\pi} + \frac{1}{5} = \frac{8(2^n-1)}{3\pi} + \frac{2}{5}. \tag{24}$$

Since for $0 \le \theta < t$ and $t + \theta \le \pi$,

$$\frac{\theta}{\sin t} \le \int_{t-\frac{\theta}{2}}^{t+\frac{\theta}{2}}\frac{\mathrm{d}\,x}{x}, \tag{25}$$

taking $t = \frac{\pi(2\mu+1)}{2(2^n-1)}$ and $\theta = \frac{\pi}{2^n-1}$ gives

$$\begin{aligned}\sum_{\mu=2}^{2^{n-1}-2}\frac{\frac{\pi}{2^n-1}}{\sin\frac{\pi(2\mu+1)}{2(2^n-1)}} &\le \sum_{\mu=2}^{2^{n-1}-2}\int_{\frac{\pi(2\mu+1)}{2(2^n-1)}-\frac{\pi}{2(2^n-1)}}^{\frac{\pi(2\mu+1)}{2(2^n-1)}+\frac{\pi}{2(2^n-1)}}\frac{\mathrm{d}\,x}{\sin x} = \sum_{\mu=2}^{2^{n-1}-2}\int_{\frac{\pi\mu}{2^n-1}}^{\frac{\pi(\mu+1)}{2^n-1}}\frac{\mathrm{d}\,x}{\sin x} = \int_{\frac{2\pi}{2^n-1}}^{\frac{\pi(2^{n-1}-1)}{2^n-1}}\frac{\mathrm{d}\,x}{\sin x}\\ &< \int_{\frac{2\pi}{2^n-1}}^{\frac{\pi}{2}}\frac{\mathrm{d}\,x}{\sin x} = \Big[\ln(\tan\frac{x}{2})\Big]_{\frac{2\pi}{2^n-1}}^{\frac{\pi}{2}} = -\ln(\tan\frac{\pi}{2^n-1})\end{aligned}$$

$$< -\ln(\frac{\pi}{2^n-1}) < n\ln 2 - \ln\pi.$$

Hence, for $n \geq 3$, we can obtain that

$$\begin{aligned}
\mathcal{NL}(\phi) &\geq 2^{n-1} - \frac{2^{\frac{n}{2}}}{2^n-1}\sum_{\mu=0}^{2^{n-1}-2}\frac{1}{\sin\frac{\pi(2\mu+1)}{2(2^n-1)}} - \frac{2^n}{2(2^n-1)} \\
&= 2^{n-1} - \frac{2^{\frac{n}{2}}}{2^n-1}\left(\frac{1}{\sin\frac{\pi}{2(2^n-1)}} + \frac{1}{\sin\frac{3\pi}{2(2^n-1)}} + \sum_{\mu=2}^{2^{n-1}-2}\frac{1}{\sin\frac{\pi(2\mu+1)}{2(2^n-1)}}\right) - \frac{2^n}{2(2^n-1)} \\
&> 2^{n-1} - \frac{2^{\frac{n}{2}}}{2^n-1}\left(\frac{8(2^n-1)}{3\pi} + \frac{2}{5} + \frac{2^n-1}{\pi}(n\ln 2 - \ln\pi)\right) - \frac{2^n}{2(2^n-1)} \\
&> 2^{n-1} - \frac{n\ln 2 + \frac{8}{3} - \ln\pi}{\pi}2^{\frac{n}{2}} - 1.
\end{aligned}$$

□

First we compare the lower bound on the nonlinearity of the function F defined as (11) with the function ϕ and the function f_2 constructed in [20]. Denote by N_ϕ, N_{f_2} and N_F respectively the lower bounds on the nonlinearity of ϕ, f_2 and F. We list in Table 1 these lower bounds for n from 6 to 18. From this table, one can see that N_F is better than N_ϕ and a little smaller than N_{f_2}. We should point out that the function f_2 is a special case of the functions defined as (11), and N_{f_2} can be slightly improved by using Lemma 4.

Table 1. Comparison of lower bounds on nonlinearity

n	N_ϕ in [20]	N_ϕ in Th.5	N_{f_2} in [20]	N_F in Th.4
6	15	17	21	20
7	38	41		
8	87	92	103	101
9	194	200		
10	417	425	459	452
11	880	892		
12	1831	1847	1930	1914
13	3769	3792		
14	7701	7734	7932	7896
15	15650	15697		
16	31674	31740	32196	32121
17	63910	64002		
18	128659	128790	129824	129665

Then we compare the exact nonlinearity of the function F with the functions ϕ, $\varPhi$ and f_2. Noting that the values of their nonlinearity are related to the primitive elements, we choose the primitive elements for the function ϕ and the function $\varPhi$ such that they achieve maximum, and give in Table 2 these values for n from 6 to 18. The primitive polynomials we choose are listed in Table 3.

To compute the nonlinearity of F, we test some of the functions in Example 2. In our experiment, we set $l = 0$ and $\beta = \gamma = \alpha^t$, and exhaust all of the possible functions, that is, any function, denoted by $F_{(\alpha^t,s)}$, having the form

$$F_{(\alpha^t,s)}(x,y) = \phi(xy) + (x^{2^k-1}+1)\psi(y) + (y^{2^k-1}+1)\varphi(x)$$

with

$$\mathrm{supp}(\psi) = \{1, \alpha^t, \cdots, (\alpha^t)^{2^{k-2}-1}\},\ 1 \leq t \leq 2^k-2,\ \gcd(2^k-1, t) = 1,$$

Table 2. Comparison of exact values of nonlinearity

n	ϕ	Φ	f_2 in [20]	$F_{(\alpha^2,1)}$	$F_{\max}$	$F_{\min}$
6	24	24	22	24	24	20
7	54					
8	112	112	108	112	112	108
9	236					
10	484	484	480	472	480	472
11	986					
12	1994	1988	1982	1982	1986	1972
13	4022					
14	8090	8072	8064	8060	8068	8048
15	16242					
16	32570	32520	32508	32504	32512	32480
17	65250					
18	130666	130632	130616	130602	130620	130580

Table 3. Primitive polynomials

n	ϕ	Φ
6	$1+x+x^6$	$1+x+x^3$
7	$1+x+x^7$	
8	$1+x^2+x^3+x^4+x^8$	$1+x+x^4$
9	$1+x^4+x^5+x^8+x^9$	
10	$1+x+x^2+x^3+x^5+x^6+x^{10}$	$1+x+x^2+x^4+x^5$
11	$1+x^2+x^3+x^4+x^6+x^8+x^9+x^{10}+x^{11}$	
12	$1+x+x^2+x^5+x^6+x^{10}+x^{12}$	$1+x+x^3+x^4+x^6$
13	$1+x^2+x^4+x^6+x^7+x^{11}+x^{13}$	
14	$1+x^3+x^8+x^9+x^{12}+x^{13}+x^{14}$	$1+x^2+x^5+x^6+x^7$
15	$1+x+x^2+x^3+x^5+x^8+x^{10}+x^{14}+x^{15}$	
16	$1+x+x^4+x^6+x^7+x^8+x^{12}+x^{14}+x^{16}$	$1+x^2+x^3+x^4+x^8$
17	$1+x^3+x^4+x^7+x^{11}+x^{16}+x^{17}$	
18	$1+x^2+x^4+x^5+x^6+x^7+x^9+x^{13}+x^{18}$	$1+x+x^4+x^5+x^6+x^8+x^9$

and

$$\mathrm{supp}(\varphi)=\{(\alpha^t)^s,(\alpha^t)^{s+1},\cdots,(\alpha^t)^{s+2^{k-2}-1}\},\ 0\le s\le 2^k-2.$$

The maximum and minimum values of the nonlinearity of these functions are listed in Table 2, for even $n=2k$ ranging from 6 to 18. We also list the values of nonlinearity for one of these functions, i.e. the function $F_{(\alpha^2,1)}$.

From Table 2, we have seen that the nonlinearity of F is very close to the nonlinearity of Φ and, for even n from 10 to 18, slightly smaller than that of ϕ, and that there always is F which have a slightly better nonlinearity than f_2. Here we should point out that sometimes the nonlinearity of F is even equal to that of Φ while any function with a form as $\phi(xy)+(x^{2^k-1}+1)\psi(y)$, e.g. f_2, always has a strictly smaller nonlinearity than Φ.

From the mentioned above, the function F has a good lower bound on nonlinearity and a very good nonlinearity, and provides a trade-off between the exact nonlinearity and the lower bound on nonlinearity.

4 Conclusion

In this paper, it was proven that a family of $2k$-variable balanced Boolean functions are almost perfect algebraic immune. The functions of this family also achieve almost all the other main cryptographic criteria, including balancedness, optimal algebraic degree and high nonlinearity.

The lower bound on nonlinearity of Carlet-Feng function was also slightly improved. Even compared with this new lower bound, the functions of that family have a better lower bound. The computer experiments for $3 \leq k \leq 9$ shows that the nonlinearity of such functions are very close to the maximum nonlinearity of Carlet-Feng function, and sometimes better than Tang et al.'s function.

Acknowledgment

The authors would like to thank Yin Zhang for his helpful discussions on the immunity to fast algebraic attacks, thank Shaoyu Du for her helpful discussions on the nonlinearity, and thank Tianze Wang for his careful checking of the results of this manuscript.

References

1. Armknecht, F.: Improving fast algebraic attacks. In: B.K. Roy, W. Meier (eds.) FSE 2004, *Lecture Notes in Computer Science*, vol. 3017, pp. 65–82. Springer (2004)
2. Armknecht, F., Carlet, C., Gaborit, P., Künzli, S., Meier, W., Ruatta, O.: Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. In: S. Vaudenay (ed.) EUROCRYPT 2006, *Lecture Notes in Computer Science*, vol. 4004, pp. 147–164. Springer (2006)
3. Carlet, C.: Boolean functions for cryptography and error correcting codes. In: Y. Crama, P.L. Hammer (eds.) Boolean Models and Methods in Mathematics, Computer Science, and Engineering, pp. 257–397. Cambridge University Press (2010)
4. Carlet, C., Feng, K.: An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity. In: J. Pieprzyk (ed.) ASIACRYPT 2008, *Lecture Notes in Computer Science*, vol. 5350, pp. 425–440. Springer (2008)
5. Cohen, G.D., Flori, J.P.: On a generalized combinatorial conjecture involving addition mod 2^{k}-1. IACR Cryptology ePrint Archive **2011**, 400 (2011)
6. Courtois, N.: Fast algebraic attacks on stream ciphers with linear feedback. In: D. Boneh (ed.) CRYPTO 2003, *Lecture Notes in Computer Science*, vol. 2729, pp. 176–194. Springer (2003)
7. Courtois, N., Meier, W.: Algebraic attacks on stream ciphers with linear feedback. In: E. Biham (ed.) EUROCRYPT 2003, *Lecture Notes in Computer Science*, vol. 2656, pp. 345–359. Springer (2003). `http://ntcourtois.free.fr/toyolili.pdf`
8. Courtois, N.T.: Cryptanalysis of sfinks. In: D. Won, S. Kim (eds.) Information Security and Cryptology - ICISC 2005, *Lecture Notes in Computer Science*, vol. 3935, pp. 261–269 (2006)
9. Dalai, D.K., Maitra, S., Sarkar, S.: Basic theory in construction of Boolean functions with maximum possible annihilator immunity. Designs Codes and Cryptography **40**(1), 41–58 (2006)
10. Feng, K., Liao, Q., Yang, J.: Maximal values of generalized algebraic immunity. Des. Codes Cryptography **50**(2), 243–252 (2009)
11. Fischer, S., Meier, W.: Algebraic immunity of S-boxes and augmented functions. In: A. Biryukov (ed.) FSE 2007, *Lecture Notes in Computer Science*, vol. 4593, pp. 366–381. Springer (2007)
12. Hakala, R.M., Nyberg, K.: On the nonlinearity of discrete logarithm in $\mathbb{F}_{2^n}$. In: C. Carlet, A. Pott (eds.) Sequences and Their Applications - SETA 2010, *Lecture Notes in Computer Science*, vol. 6338, pp. 333–345 (2010)
13. Hawkes, P., Rose, G.G.: Rewriting variables: The complexity of fast algebraic attacks on stream ciphers. In: M.K. Franklin (ed.) CRYPTO 2004, *Lecture Notes in Computer Science*, vol. 3152, pp. 390–406. Springer (2004)
14. Li, N., Qi, W.F.: Construction and analysis of Boolean functions of $2t{+}1$ variables with maximum algebraic immunity. In: X. Lai, K. Chen (eds.) ASIACRYPT 2006, *Lecture Notes in Computer Science*, vol. 4284, pp. 84–98. Springer (2006)
15. Li, N., Qu, L., Qi, W.F., Feng, G., Li, C., Xie, D.: On the construction of Boolean functions with optimal algebraic immunity. IEEE Transactions on Information Theory **54**(3), 1330–1334 (2008)
16. Liu, M., Lin, D., Pei, D.: Fast algebraic attacks and decomposition of symmetric Boolean functions. IEEE Transactions on Information Theory **57**(7), 4817–4821 (2011)
17. Liu, M., Zhang, Y., Lin, D.: Perfect algebraic immune functions. In: X. Wang, K. Sako (eds.) ASIACRYPT 2012, *Lecture Notes in Computer Science*, vol. 7658, pp. 172–189. Springer (2012). `http://eprint.iacr.org/2012/212/`
18. Meier, W., Pasalic, E., Carlet, C.: Algebraic attacks and decomposition of Boolean functions. In: C. Cachin, J. Camenisch (eds.) EUROCRYPT 2004, *Lecture Notes in Computer Science*, vol. 3027, pp. 474–491. Springer (2004)
19. Pasalic, E., Wei, Y.: On the construction of cryptographically significant Boolean functions using objects in projective geometry spaces. IEEE Transactions on Information Theory **58**(10), 6681–6693 (2012)

20. Tang, D., Carlet, C., Tang, X.: Highly nonlinear boolean functions with optimal algebraic immunity and good behavior against fast algebraic attacks. IEEE Transactions on Information Theory **59**(1), 653–664 (2013)
21. Tu, Z., Deng, Y.: A conjecture about binary strings and its applications on constructing Boolean functions with optimal algebraic immunity. Des. Codes Cryptography **60**(1), 1–14 (2011)
22. Wang, Q., Peng, J., Kan, H., Xue, X.: Constructions of cryptographically significant Boolean functions using primitive polynomials. IEEE Transactions on Information Theory **56**(6), 3048–3053 (2010)
23. Zeng, X., Carlet, C., Shan, J., Hu, L.: More balanced Boolean functions with optimal algebraic immunity and good nonlinearity and resistance to fast algebraic attacks. IEEE Transactions on Information Theory **57**(9), 6310–6320 (2011)