Perfect Algebraic Immune Functions ⋆

Meicheng Liu, Yin Zhang, and Dongdai Lin

SKLOIS, Institute of Information Engineering, CAS, Beijing 100195, P. R. China
meicheng.liu@gmail.com, zhangy@is.iscas.ac.cn, ddlin@iie.ac.cn

Abstract. A perfect algebraic immune function is a Boolean function with perfect immunity against algebraic and fast algebraic attacks. The main results are that for a perfect algebraic immune balanced function the number of input variables is one more than a power of two; for a perfect algebraic immune unbalanced function the number of input variables is a power of two. Also the Carlet-Feng functions on $2^s + 1$ variables and the modified Carlet-Feng functions on 2^s variables are shown to be perfect algebraic immune functions. Furthermore, it is shown that a perfect algebraic immune function behaves good against probabilistic algebraic attacks as well.

Keywords: Boolean functions, Algebraic immunity, Fast algebraic attacks, Probabilistic algebraic attacks

1 Introduction

The study of the cryptanalysis of the filter and combination generators of stream ciphers based on linear feedback shift registers (LFSRs) has resulted in a wealth of cryptographic criteria for Boolean functions, such as balancedness, high algebraic degree, high nonlinearity, high correlation immunity and so on. An overview of cryptographic criteria for Boolean functions with extensive bibliography is given in [5].

In recent years, algebraic and fast algebraic attacks [1,7,8] have been regarded as the most successful attacks on LFSR-based stream ciphers. These attacks cleverly use overdefined systems of multivariable nonlinear equations to recover the secret key. Algebraic attacks make use of the equations by multiplying a non-zero function of low degree, while fast algebraic attacks make use of the equations 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. [22] 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 $\lceil \frac{n}{2} \rceil$. The properties and constructions of Boolean functions with maximum $\mathcal{AI}$ are researched in a large number of papers, e.g., [10,17,18,6,27,28].

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 [8]. The resistance against fast algebraic attacks is not covered by algebraic immunity [9,2,19]. 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. [19] stated that almost all the symmetric functions including these functions with good algebraic immunity behavior badly against fast algebraic attacks. In [25] P. Rizomiliotis introduced a method to evaluate the behavior of Boolean functions against fast algebraic attacks using univariate polynomial representation. However, it is unclear what is maximum immunity to fast algebraic attacks.

In [8] N. Courtois proved that for any pair of positive integers (e, d) such that $e+d \geq n$, there is a non-zero 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

⋆ Supported by the National 973 Program of China under Grant 2011CB302400, the National Natural Science Foundation of China under Grants 10971246, 60970152, and 61173134, the Grand Project of Institute of Software of CAS under Grant YOCX285056 and the CAS Special Grant for Postgraduate Research, Innovation and Practice.

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 non-zero function g of degree at most e such that gf has degree at most d. Such functions are said to be perfect algebraic immune ($\mathcal{PAI}$). 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 [8]. In this case, e equals 1 and d equals the degree of the function f. 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. Although preventing classical and fast algebraic attacks is not sufficient for resisting algebraic attacks on the augmented function [14], 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.

It is an open question whether there are $\mathcal{PAI}$ functions for arbitrary number of input variables. This problem was also noticed in [6] at Asiacrypt 2008. It seems that $\mathcal{PAI}$ functions are quite rare. In [6] C. Carlet and K. Feng observed that the Carlet-Feng functions on 9 variables are $\mathcal{PAI}$. One can check that the Carlet-Feng functions on 5 variables are also $\mathcal{PAI}$ (see also [12]). However, no function is shown to be $\mathcal{PAI}$ for arbitrary number of variables. On the contrary, M. Liu et al. [19] proved that no symmetric functions are $\mathcal{PAI}$, and Y. Zhang et al. [29] proved that no rotation symmetric functions are $\mathcal{PAI}$ for even number (except a power of two) of variables.

In this paper, we study the upper bounds on the immunity to fast algebraic attacks, and solve the above question. The immunity against fast algebraic attacks is related to a matrix thanks to Theorem 1 of [2]. By a simple transformation on this matrix we obtain a symmetric matrix whose elements are the coefficients of the algebraic normal form of a given Boolean function. We improve the upper bounds on the immunity to fast algebraic attacks by proving that the symmetric matrix is singular in some cases. The results are that for an n-variable function, we have: (1) if n is a power of 2 then a $\mathcal{PAI}$ function has degree n; (2) if n is one more than a power of 2 then a $\mathcal{PAI}$ function has degree $n-1$ (which is also balanced); (3) otherwise, the function is not $\mathcal{PAI}$. We then prove that the Carlet-Feng functions, which have degree $n-1$, are $\mathcal{PAI}$ for n equal to one more than a power of 2, and are almost $\mathcal{PAI}$ for the other cases. Also we prove that the modified Carlet-Feng functions, which have degree n, are $\mathcal{PAI}$ for n equal to a power of 2, and are almost $\mathcal{PAI}$ for the other cases. The results show that our bounds on the immunity to fast algebraic attacks are tight, and that the Carlet-Feng functions are optimal against fast algebraic attacks as well as classical algebraic attacks. In contrast, P. Rizomiliotis [26] determined the immunity of the Carlet-Feng functions against fast algebraic attacks by computing the linear complexity of a sequence, which is infeasible for large n.

At Eurocrypt 2003, N. Courtois and W. Meier [7] described the probabilistic scenario of algebraic attacks as follows:

S4 There exists a non-zero function g of low degree such that gf can be approximated by a function of low degree with probability $1-\varepsilon$.

In [3], A. Braeken and B. Preneel generalized **S4** to the two scenarios:

S4a There exists a non-zero function g of low degree such that $gf=g$ on $\{x \mid f(x)=0\}$ with probability $1-\varepsilon$.

S4b There exists a non-zero function g of low degree such that $gf=0$ on $\{x \mid f(x)=1\}$ with probability $1-\varepsilon$.

The probability for the scenario **S4a** is equal to $p=1-\frac{\mathrm{d}(gf,g)}{2^n-\mathrm{wt}(f)}$, and equal to $p=1-\frac{\mathrm{d}(gf,0)}{\mathrm{wt}(f)}$ for the scenario **S4b**. Then $p_{\max}=1-\frac{\min\{\mathrm{d}(gf,g),\mathrm{d}(gf,0)\}}{2^{n-1}}$ for a balanced function.

At Crypto 2006, C. Carlet [4] proved that $\min\{\mathrm{d}(gf,g),\mathrm{d}(gf,0)\} \geq \sum_{i=0}^{\mathcal{AI}(f)-r-1}\binom{n-r}{i}$ holds for non-zero function g of degree at most r. This result gives an upper bound on the probability for applying probabilistic algebraic attacks. The details can also be found in [20].

In [24] E. Pasalic claimed that from time complexity point of view deterministic algebraic attacks are in general more efficient than probabilistic ones for practical sizes L (e.g. $L=256$) of LFSR in the context of certain LFSR-based stream ciphers under an assumption[1] that the minimum distance of the code derived by shortening Reed-Muller code (which depends on the filter function) meets the Gilbert-Varshamov (GV) bound. Nevertheless, one should still verify whether the structure of the function itself allows a low-degree approximation that is satisfied with high probability. In [20], M. Liu et al. gave two examples of filter functions for which probabilistic algebraic attacks outperform deterministic ones for practical sizes of the LFSR in the context of the nonlinear filter generator.

In this paper, based on Carlet's bound, we show that for a filter function with maximum $\mathcal{AI}$ probabilistic algebraic attacks are worse than exhaustive search in the context of the nonlinear filter generator if the length of the LFSR is greater than or equal to 46. This does not contradict the results of [20], since the filter functions shown in [20] do not have maximum $\mathcal{AI}$. Our work shows that a $\mathcal{PAI}$ function behaves good against probabilistic algebraic attacks since it has maximum $\mathcal{AI}$. Again, we do not consider probabilistic algebraic attacks on the augmented function here.

The remainder of this paper is organized as follows. In Section 2 some basic concepts are provided. Section 3 presents the improved upper bounds on the immunity of Boolean functions against fast algebraic attacks while Section 4 shows that the Carlet-Feng functions and their modifications achieve these bounds. Section 5 states that a $\mathcal{PAI}$ function has good immunity to probabilistic algebraic attacks. Section 6 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 $\mathrm{supp}(f) = \{x \in \mathbb{F}_2^n \mid f(x) = 1\}$. The Hamming weight of f, denoted by $\mathrm{wt}(f)$, 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, $\mathrm{wt}(f) = 2^{n-1}$. The Hamming distance between n-variable functions f and g, denoted by $\mathrm{d}(f,g)$, is the number of $x \in \mathbb{F}_2^n$ at which $f(x) \neq g(x)$. It is well known that $\mathrm{d}(f,g) = \mathrm{wt}(f+g)$.

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

$$f(x) = \sum_{c \in \mathbb{F}_2^n} a_c x^c,\ a_c \in \mathbb{F}_2,\ x^c = x_1^{c_1} x_2^{c_2} \cdots x_n^{c_n},\ c = (c_1, c_2, \cdots, c_n),$$

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

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 $\max\limits_{a_i \neq 0} \mathrm{wt}(i)$, where $i = \sum_{k=1}^{n} i_k 2^{k-1}$ is considered as $(i_1, i_2, \cdots, i_n) \in \mathbb{F}_2^n$.

[1] In some cases, this assumption (Eq. (3) in [24]) never holds. For example, there is no Boolean function such that the shortened Reed-Muller code of the second order achieves the GV bound. More precisely, the minimum distance of such code of r-th order is upper bounded by 2^{n-r-1} according to [3,20], and one can then check that the assumption never holds for the case $r = 2$.

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

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

For more details with regard to the representation of Boolean functions, we refer to [5].

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

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

The immunity of f against fast algebraic attacks is related to the degree e of a function g and the degree d of gf with $e \le d$. For an n-variable function f and any positive integer e with $e < n/2$, there is a non-zero function g of degree at most e such that gf has degree at most $n - e$ [8]. There are several notions about the immunity of Boolean functions against fast algebraic attacks in previous literatures, such as [15,23]. The perfect algebraic immune function we define below is actually a Boolean function which is algebraic attack resistant (see [23]) and has degree at least $n-1$. The latter is necessary for perfect algebraic immune function since for a function of degree less than $n-1$ the fast general attack uses $e = 1$ and $d = \deg(f) < n - 1 = n - e$.

Definition 2 *Let f be an n-variable Boolean function. The function f is said to be perfect algebraic immune ($\mathcal{PAI}$) if for any positive integers $e < n/2$, the product gf has degree at least $n - e$ for any non-zero function g of degree at most e.*

A $\mathcal{PAI}$ function also achieves maximum $\mathcal{AI}$. As a matter of fact, if a function does not achieves maximum $\mathcal{AI}$, then it admits a non-zero function g of degree less than $n/2$ such that $gf = 0$ or $gf = g$, which means that it is not $\mathcal{PAI}$. Therefore $\mathcal{PAI}$ functions are the class of Boolean functions perfectly resistant to classical and fast algebraic attacks.

3 The immunity of Boolean functions against fast algebraic attacks

In this section, we present the upper bounds on the immunity of Boolean functions against fast algebraic attacks. We first recall the previous results for determining the immunity against fast algebraic attacks, then state our bounds.

Denote by $\mathcal{W}_i$ the set $\{x \in \mathbb{F}_2^n \mid \mathrm{wt}(x) \le i\}$ in lexicographic order and by $\overline{\mathcal{W}}_i$ the set $\{x \in \mathbb{F}_2^n \mid \mathrm{wt}(x) \ge i+1\}$ in reverse lexicographic order. For $x \in \mathbb{F}_2^n$, let $\bar{x} = (x_1 + 1, \ldots, x_n + 1)$. If x is the j-th element in $\mathcal{W}_e$ and $\bar{x} \in \overline{\mathcal{W}}_d$, then $\bar{x}$ is the j-th element in $\overline{\mathcal{W}}_d$. Here are some additional notational conventions: for $y, z \in \mathbb{F}_2^n$, let $z \subset y$ be an abbreviation for $\mathrm{supp}(z) \subset \mathrm{supp}(y)$, where $\mathrm{supp}(x) = \{i \mid x_i = 1\}$, and let $y \cap z = (y_1 \wedge z_1, \ldots, y_n \wedge z_n)$, $y \cup z = (y_1 \vee z_1, \ldots, y_n \vee z_n)$, where $\wedge$ and $\vee$ are the AND and OR operations respectively. We can see that $z \subset y$ if and only if $y^z = y_1^{z_1} y_2^{z_2} \cdots y_n^{z_n} = 1$.

Let g be a function of algebraic degree at most e ($e < n/2$) such that $h = gf$ has algebraic degree at most d ($e \le d$). Let

$$f(x) = \sum_{c \in \mathbb{F}_2^n} f_c x^c, \ f_c \in \mathbb{F}_2,$$

$$g(x) = \sum_{z \in \mathcal{W}_e} g_z x^z, \ g_z \in \mathbb{F}_2,$$

and

$$h(x) = \sum_{y \in \mathcal{W}_d} h_y x^y, \ h_y \in \mathbb{F}_2$$

be the ANFs of f, g and h respectively. For $y \in \overline{\mathcal{W}}_d$, we have $h_y = 0$ and therefore

$$0 = h_y = \sum_{c \in \mathbb{F}_2^n} \sum_{\substack{c \cup z = y \\ z \in \mathcal{W}_e}} f_c g_z = \sum_{z \in \mathcal{W}_e} g_z \sum_{\substack{c \cup z = y \\ c \in \mathbb{F}_2^n}} f_c. \tag{3}$$

The above equations on g_z's are homogeneous linear. Denote by $V(f; e, d)$ the coefficient matrix of the equations, which is a $\sum_{i=d+1}^{n} \binom{n}{i} \times \sum_{i=0}^{e} \binom{n}{i}$ matrix with the ij-th element equal to

$$v_{yz} = \sum_{\substack{c \cup z = y \\ c \in \mathbb{F}_2^n}} f_c = \sum_{\substack{y \cap \bar{z} \subset c \subset y \\ z \subset y}} f_c = y^z \sum_{y \cap \bar{z} \subset c \subset y} f_c, \tag{4}$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$. Then f admits no non-zero function g of algebraic degree at most e such that $h = gf$ has algebraic degree at most d if and only if the rank of the matrix $V(f; e, d)$ equals the number of g_z's which is $\sum_{i=0}^{e} \binom{n}{i}$, i.e., $V(f; e, d)$ has full column rank (see also [2,12]).

Theorem 1 *[2,12] Let $f \in \mathbf{B}_n$. Then there exists no non-zero function g of degree at most e such that the product gf has degree at most d if and only if the matrix $V(f; e, d)$ has full column rank.*

Remark 1. The theorem shows that $\mathcal{AI}(f) > e$ if and only if the matrix $V(f; e, e)$ has full column rank (since $\mathcal{AI}(f) > e$ if and only if there exists no non-zero function g of degree at most e such that $h = gf$ has degree at most e). Then $\mathcal{AI}(f) = \lceil \frac{n}{2} \rceil$ if and only if the matrix $V(f; \lceil \frac{n}{2} \rceil - 1, \lceil \frac{n}{2} \rceil - 1)$ has full column rank.

Now we show that performing some column operations on the matrix $V(f; e, d)$ creates a matrix with f_c's as its elements.

Lemma 2 $\sum_{z^* \subset z} v_{yz^*} = f_{y \cap \bar{z}}$.

Proof. Note that $c \cup z = y$ if and only if $c \subset y, z \subset y$ and $y \subset c \cup z$, that is, $y^c = 1, y^z = 1$ and $(c \cup z)^y = 1$. By (4) we have

$$\begin{aligned}
\sum_{z^* \subset z} v_{yz^*} &= \sum_{z^* \subset z} \sum_{c \cup z^* = y} f_c \\
&= \sum_{z^* \subset z} \sum_{c \in \mathbb{F}_2^n} y^c y^{z^*} (c \cup z^*)^y f_c \\
&= \sum_{c \in \mathbb{F}_2^n} y^c f_c \sum_{z^* \subset z} y^{z^*} (c \cup z^*)^y \\
&= \sum_{c \subset y} f_c \sum_{\substack{z^* \subset y \cap z \\ y \subset c \cup z^*}} 1 \\
&= \sum_{c \subset y} f_c \sum_{y \cap \bar{c} \subset z^* \subset y \cap z} 1 \\
&= \sum_{c \subset y, y \cap \bar{c} = y \cap z} f_c \\
&= f_{y \cap \bar{z}}.
\end{aligned}$$

□

By Lemma 2 we know that the matrix $V(f;e,d)$ can be transformed into a matrix, denoted by $W(f;e,d)$, with the ij-th element equal to

$$w_{yz} = f_{y\cap\bar{z}},$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$. The ji-th element of $W(f;e,d)$ is equal to

$$w_{\bar{z}\bar{y}} = f_{\bar{z}\cap\bar{\bar{y}}} = f_{y\cap\bar{z}} = w_{yz},$$

since $\bar{z}$ is the j-th element in $\overline{\mathcal{W}}_d$ and $\bar{y}$ is the i-th element in $\mathcal{W}_e$ by the definition of $\overline{\mathcal{W}}_d$ and $\mathcal{W}_e$. Recall that $V(f;e,d)$ and $W(f;e,d)$ are $\sum_{i=d+1}^{n}\binom{n}{i}\times\sum_{i=0}^{e}\binom{n}{i}$ matrices. Therefore the matrix $W(f;e,n-e-1)$ is a symmetric $\sum_{i=0}^{e}\binom{n}{i}\times\sum_{i=0}^{e}\binom{n}{i}$ matrix, denoted by $W(f;e)$.

Theorem 3 *Let $f\in\mathbf{B}_n$ and $f(x)=\sum_{c\in\mathbb{F}_2^n} f_c x^c$. Then there exists no non-zero function g of degree at most e such that gf has degree at most d if and only if $W(f;e,d)$ has full column rank.*

Proof. Lemma 2 shows that $V(f;e,d)$ and $W(f;e,d)$ have the same rank. Then the theorem follows from Theorem 1. □

Remark 2. The theorem shows that $\mathcal{AI}(f)>e$ if and only if the matrix $W(f;e,e)$ has full column rank (since $\mathcal{AI}(f)>e$ if and only if there exists no non-zero function g of degree at most e such that $h=gf$ has degree at most e). Then $\mathcal{AI}(f)=\lceil\frac{n}{2}\rceil$ if and only if the matrix $W(f;\lceil\frac{n}{2}\rceil-1,\lceil\frac{n}{2}\rceil-1)$ has full column rank.

Next we concentrate on the upper bounds on the immunity of Boolean functions against fast algebraic attacks. As mentioned in Section 2, for an n-variable function f and any positive integer e with $e<n/2$, there is a non-zero function g of degree at most e such that gf has degree at most $n-e$. This can also be explained by Theorem 1 or Theorem 3: the matrices $V(f;e,n-e)$ and $W(f;e,n-e)$ have not full column rank since they are $\sum_{i=0}^{e-1}\binom{n}{i}\times\sum_{i=0}^{e}\binom{n}{i}$ matrices. From Theorem 3 the bounds on the immunity to fast algebraic attacks are related to the question whether the symmetric matrix $W(f;e)$ is invertible.

Before stating our main results, we list a useful lemma about the determinant of a symmetric matrix over a field with characteristic 2.

Lemma 4 *Let $A=(a_{ij})_{m\times m}$ be a symmetric $m\times m$ matrix over a field with characteristic 2, and $a_{ii}=a_{1i}^2$ for $2\le i\le m$, that is,*

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m}\\ a_{12} & a_{12}^2 & a_{23} & \cdots & a_{2m}\\ a_{13} & a_{23} & a_{13}^2 & \cdots & a_{3m}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{1m} & a_{2m} & a_{3m} & \cdots & a_{1m}^2 \end{pmatrix}. \tag{5}$$

If $a_{11}=(m+1)\bmod 2$, then $\det(A)=0$.

Proof. Let S_m be the symmetric group of degree m. Then

$$\begin{aligned}\det(A) &= \sum_{\sigma\in S_m}\prod_{i=1}^{m} a_{i,\sigma(i)}\\ &= \sum_{\sigma\in S_m,\sigma^2=1}\prod_{i=1}^{m} a_{i,\sigma(i)} + \sum_{\sigma\in S_m,\sigma^2\neq 1}\prod_{i=1}^{m} a_{i,\sigma(i)}\\ &\left(\text{since } \prod_{i=1}^{m} a_{i,\sigma(i)} = \prod_{i=1}^{m} a_{\sigma(i),i} = \prod_{i=1}^{m} a_{\sigma(i),\sigma^{-1}(\sigma(i))} = \prod_{i=1}^{m} a_{i,\sigma^{-1}(i)}\right)\end{aligned}$$

$$= \sum_{\sigma \in S_m, \sigma^2=1} \prod_{i=1}^{m} a_{i,\sigma(i)}.$$

If m is odd, then $a_{11} = 0$ and therefore

$$\begin{aligned}
\det(A) &= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1 \\ \sigma(1)=j}} a_{1j} \prod_{i=2}^{m} a_{i,\sigma(i)} \\
&= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1 \\ \sigma(1)=j}} a_{1j}^2 \prod_{\substack{2 \le i \le m \\ i \ne j}} a_{i,\sigma(i)} \\
&\text{(for odd } m \text{ and } \sigma^2 = 1\text{, there is } j' \text{ such that } j' \ne j \text{ and } \sigma(j') = j'\text{)} \\
&= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1 \\ \sigma(1)=j, \sigma(j')=j'}} a_{1j}^2 a_{1j'}^2 \prod_{\substack{2 \le i \le m \\ i \ne j, j'}} a_{i,\sigma(i)} \\
&\text{(there is unique } \sigma' \text{ such that } \sigma'(1) = j',\ \sigma'(j') = 1,\ \sigma'(j) = j,\ \sigma'(i) = \sigma(i) \text{ for } i \notin \{1, j, j'\}\text{)} \\
&= 0.
\end{aligned}$$

If m is even, then $a_{11} = 1$ and therefore

$$\begin{aligned}
\det(A) &= \sum_{\substack{\sigma^2=1 \\ \sigma(1)=1}} \prod_{i=2}^{m} a_{i,\sigma(i)} + \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1 \\ \sigma(1)=j}} a_{1j}^2 \prod_{\substack{2 \le i \le m \\ i \ne j}} a_{i,\sigma(i)} \\
&= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1 \\ \sigma(1)=1, \sigma(j)=j}} a_{1j}^2 \prod_{\substack{2 \le i \le m \\ i \ne j}} a_{i,\sigma(i)} + \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1 \\ \sigma(1)=j}} a_{1j}^2 \prod_{\substack{2 \le i \le m \\ i \ne j}} a_{i,\sigma(i)} \\
&= 0.
\end{aligned}$$

□

Remark 3. For the matrix A of Lemma 4 it holds that $\det(A) = \det(A^{(1,1)})$ if $a_{11} = m \bmod 2$, where $A^{(i,j)}$ is the $(m-1) \times (m-1)$ matrix that results from A by removing the i-th row and the j-th column.

Theorem 5 *Let $f \in \mathbf{B}_n$ and f_{2^n-1} be the coefficient of the monomial $x_1 x_2 \cdots x_n$ in the ANF of f. Let e be a positive integer less than $n/2$. If $f_{2^n-1} = \binom{n-1}{e} + 1 \bmod 2$, then there exists $g \ne 0$ with degree at most e such that gf has degree at most $n - e - 1$.*

Proof. According to Theorem 3 we need to prove that the square matrix $W(f; e)$ is singular when $f_{2^n-1} = \binom{n-1}{e} + 1 \bmod 2$. Let W_{ij} be the ij-th element of $W(f; e)$. Since $\mathbf{1} = (1, 1, \cdots, 1)$ and $\mathbf{0} = (0, 0, \cdots, 0)$ are the first elements in $\overline{\mathcal{W}}_{n-e-1}$ and $\mathcal{W}_e$ respectively, by the definition of $W(f; e)$ we have $W_{11} = w_{\mathbf{1},\mathbf{0}} = f_{2^n-1}$. Because $\sum_{i=0}^{e} \binom{n}{i} = \sum_{i=1}^{e} \binom{n-1}{i} + \sum_{i=1}^{e} \binom{n-1}{i-1} + 1 \equiv \binom{n-1}{e} (\bmod 2)$, we know $W_{11} = \sum_{i=0}^{e} \binom{n}{i} + 1 \bmod 2$ when $f_{2^n-1} = \binom{n-1}{e} + 1 \bmod 2$. As mentioned previously, $W(f; e)$ is a symmetric $\sum_{i=0}^{e} \binom{n}{i} \times \sum_{i=0}^{e} \binom{n}{i}$ matrix over $\mathbb{F}_2$. We wish to show that $W(f; e)$ has the form of (5). By the definition of $W(f; e)$ we have $W_{1i}^2 = W_{1i} = w_{\mathbf{1}z} = f_{\mathbf{1} \cap \bar{z}} = f_{\bar{z}} = f_{\bar{z} \cap \bar{z}} = w_{\bar{z}z} = W_{ii}$ where $\bar{z}$ is the i-th element in $\overline{\mathcal{W}}_{n-e-1}$ and z is the i-th element in $\mathcal{W}_e$. It follows from Lemma 4 that the matrix $W(f; e)$ is singular. □

Corollary 6 *Let n be an even number and $f \in \mathbf{B}_n$. If f is balanced, then there exists a non-zero function g with degree at most 1 such that the product gf has degree at most $n - 2$.*

Proof. If f is balanced, then $f_{2^n-1} = 0$. For even n, it holds that $\binom{n-1}{1} + 1 \equiv 0 (\bmod 2)$. Therefore the result follows from Theorem 5. □

From Corollary 6 it seems that for the number n of input variables, odd numbers are better than even ones from a cryptographic point of view (since cryptographic functions must be balanced).

Lucas' theorem states that for positive integers m and i, the following congruence relation holds:

$$\binom{m}{i} \equiv \prod_{k=1}^{s} \binom{m_k}{i_k} (\bmod 2),$$

where $m = \sum_{k=1}^{s} m_k 2^{k-1}$ and $i = \sum_{k=1}^{s} i_k 2^{k-1}$ are the binary expansion of m and i respectively. It means that $\binom{m}{i} \bmod 2 = 1$ if and only if $i \subset m$.

Note that $f_{2^n-1} = 1$ if and only if $\deg(f) = n$. Theorem 5 shows that for an n-variable function f of degree n and $e \not\subset n-1$, there is a non-zero function g of degree at most e such that gf has degree at most $n-e-1$, and that for an n-variable function f of degree less than n and $e \subset n-1$, there is a non-zero function g of degree at most e such that gf has degree at most $n-e-1$.

For the case $n-1 \notin \{2^s, 2^s-1\}$, there are integers e, e^* with $0 < e, e^* < n/2$ such that $e \subset n-1$ and $e^* \not\subset n-1$, and thus an n-variable function is not $\mathcal{PAI}$. This shows that for a $\mathcal{PAI}$ function the number n of input variables is 2^s+1 or 2^s. For $n = 2^s+1$ (resp. 2^s), it holds that $e \not\subset n-1$ (resp. $e \subset n-1$) for positive integer $e < n/2$, and thus an n-variable function with degree equal to n (resp. less than n) is not $\mathcal{PAI}$. Recall that a function on odd number of variables with maximum $\mathcal{AI}$ is always balanced [11]. For $n = 2^s+1$, a $\mathcal{PAI}$ function is balanced, since it has maximum $\mathcal{AI}$. For $n = 2^s$, a $\mathcal{PAI}$ function has degree n and is then unbalanced, since a function has an odd Hamming weight if and only if it has degree n. Consequently the following theorem is obtained.

Theorem 7 *Let $f \in \mathbf{B}_n$ be a perfect algebraic immune function. Then n is one more than or equal to a power of 2. Further, if f is balanced, then n is one more than a power of 2; if f is unbalanced, then n is a power of 2.*

4 The immunity of Boolean functions against fast algebraic attacks using univariate polynomial representation

In this section we focus on the immunity of Boolean functions against fast algebraic attacks using univariate polynomial representation and show that the bounds presented in Section 3 can be achieved.

Recall that $\mathcal{W}_e$ is the set $\{x \in \mathbb{F}_2^n | \operatorname{wt}(x) \le e\}$ in lexicographic order and $\overline{\mathcal{W}}_d$ is the set $\{x \in \mathbb{F}_2^n | \operatorname{wt}(x) \ge d+1\}$ in reverse lexicographic order. Hereinafter, an element $z = (z_1, z_2, \cdots, z_n)$ in $\mathcal{W}_e$ or $\overline{\mathcal{W}}_d$ is considered as an integer $z_1 + z_2 2 + \cdots + z_n 2^{n-1}$ from 0 to 2^n-1, and the operations "+" and "−" may be considered as addition and subtraction operations modulo 2^n-1 respectively if there is no ambiguity.

Let f, g and h be n-variable Boolean functions, and let g be a function of algebraic degree at most e ($e < n/2$) satisfying that $h = gf$ has algebraic degree at most d ($e \le d$). Let

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

$$g(x) = \sum_{z \in \mathcal{W}_e} g_z x^z, \ g_z \in \mathbb{F}_{2^n},$$

and

$$h(x) = \sum_{y \in \mathcal{W}_d} h_y x^y, \ h_y \in \mathbb{F}_{2^n}$$

be the univariate polynomial representations of f, g and h respectively. For $y \in \overline{\mathcal{W}}_d$, we have $h_y = 0$ and thus

$$0 = h_y = \sum_{\substack{k+z=y\\ z\in\mathcal{W}_e}} f_k g_z = \sum_{z\in\mathcal{W}_e} f_{y-z} g_z. \tag{6}$$

The above equations on g_z's are homogeneous linear. Denote by $U(f;e,d)$ the coefficient matrix of the equations, which is a $\sum_{i=d+1}^{n}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix with the ij-th element equal to

$$u_{yz} = f_{y-z} \tag{7}$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$. More precisely, for $(i,j) = (1,1)$ we have $(y,z) = (2^n-1, 0)$ and $u_{yz} = f_{2^n-1}$; for $(i,j) \neq (1,1)$ we have $y - z \notin \{0, 2^n-1\}$ and $u_{yz} = f_{(y-z) \bmod (2^n-1)}$ when $e \leq d$.

If the matrix $U(f;e,d)$ has full column rank, i.e., the rank of $U(f;e,d)$ equals the number of g_z's, then f admits no non-zero function g of algebraic degree at most e such that $h = gf$ has algebraic degree at most d.

If the matrix $U(f;e,d)$ has not full column rank, then there always exists a non-zero Boolean function $g(x)$ satisfying Equations (6). More precisely, if $g(x) = \sum_{z\in\mathcal{W}_e} g_z x^z$ ($g_z \in \mathbb{F}_{2^n}$) satisfies (6), then

$$0 = h_y^2 = \sum_{z\in\mathcal{W}_e} f_{y-z}^2 g_z^2 = \sum_{z\in\mathcal{W}_e} f_{2y-2z} g_z^2,\ y \in \overline{\mathcal{W}}_d, \tag{8}$$

where $f_{2(2^n-1)} = f_{2^n-1}$ and f_{2i} is considered as $f_{2i \bmod (2^n-1)}$ for $i \neq 2^n - 1$, showing that $g^2(x) = \sum_{z\in\mathcal{W}_e} g_z^2 x^{2z} \bmod (x^{2^n} - x)$ satisfies (8). Note that (6) and (8) are actually the same equations. It shows that if $g(x)$ satisfies Equations (6) then $\mathrm{Tr}(g(x))$ satisfies Equations (6), where $\mathrm{Tr}(x) = x + x^2 + \cdots + x^{2^{n-1}}$. Also it follows that if $g(x)$ satisfies Equations (6) then $\beta g(x)$ and $\mathrm{Tr}(\beta g(x))$ satisfy Equations (6) for any $\beta \in \mathbb{F}_{2^k}$. If $g(x) \neq 0$, then there is $c \in \mathbb{F}_{2^k}$ such that $g(c) \neq 0$, and there is $\beta \in \mathbb{F}_{2^k}$ such that $\mathrm{Tr}(\beta g(c)) \neq 0$ and thus $\mathrm{Tr}(\beta g(x)) \neq 0$. Now we can see that $\mathrm{Tr}(\beta g(x))$ is a non-zero Boolean function and satisfies (6). Hence if there is a non-zero solution for (6), then there always exists a non-zero Boolean function g satisfying (6).

Thus the following theorem is obtained.

Theorem 8 *Let $f \in \mathbf{B}_n$. Then there exists no non-zero function g of degree at most e such that the product gf has degree at most d if and only if the matrix $U(f;e,d)$ has full column rank.*

Remark 4. The theorem shows that $\mathcal{AI}(f) > e$ if and only if the matrix $U(f;e,e)$ has full column rank (since $\mathcal{AI}(f) > e$ if and only if there exists no non-zero function g of degree at most e such that $h = gf$ has degree at most e). Then $\mathcal{AI}(f) = \lceil \frac{n}{2} \rceil$ if and only if the matrix $U(f; \lceil \frac{n}{2} \rceil - 1, \lceil \frac{n}{2} \rceil - 1)$ has full column rank.

Remark 5. The matrix $U(f;e,n-e-1)$, denoted by $U(f;e)$, is symmetric since

$$u_{\bar{z}\bar{y}} = f_{\bar{z}-\bar{y}} = f_{(2^n-1-z)-(2^n-1-y)} = f_{y-z} = u_{yz}.$$

Further, we have

$$u_{y\bar{y}} = f_{y-\bar{y}} = f_{y-(2^n-1-y)} = f_{2y} = f_y^2 = u_{y,0}^2,$$

and therefore $U(f;e)$ has the form of (5). Hence Theorem 5 can also be derived from Theorem 8 and Lemma 4.

4.1 Carlet-Feng functions

The class of the Carlet-Feng functions were first presented in [13] and further studied by C. Carlet and K. Feng [6]. Such functions have maximum algebraic immunity and good nonlinearity. It was observed through computer experiments by Armknecht's algorithm [2] that the functions also have good behavior against fast algebraic attacks. In [26], P. Rizomiliotis determined the immunity of the Carlet-Feng functions against fast algebraic attacks by computing the linear complexity of a sequence, which is more efficient than Armknecht's algorithm but is not yet feasible for large n. In this section, we further discuss the immunity of the Carlet-Feng functions against fast algebraic attacks and prove that the functions achieve the bounds of Theorem 5.

Let n be an integer and α a primitive element of $\mathbb{F}_{2^n}$. Let $f \in \mathbf{B}_n$ and

$$\mathrm{supp}(f) = \{\alpha^l, \alpha^{l+1}, \alpha^{l+2}, \cdots, \alpha^{l+2^{n-1}-1}\}, 0 \le l \le 2^n - 2. \tag{9}$$

Then $\mathcal{AI}(f) = \lceil \frac{n}{2} \rceil$ according to [13,6]. As a matter of fact, the support of the function $f(\alpha^{l+2^{n-1}}x) + 1$ is $\{0, 1, \alpha, \cdots, \alpha^{2^{n-1}-2}\}$, which is the Carlet-Feng function. It means that these functions are affine equivalent.

A similar proof of [6, Theorem 2] applies to the following result. Here we give a proof for self-completeness.

Proposition 9 *Let $\sum_{i=0}^{2^n-1} f_i x^i (f_i \in \mathbb{F}_{2^n})$ be the univariate representation of the function f of (9). Then $f_0 = 0$, $f_{2^n-1} = 0$, and for $1 \le i \le 2^n - 2$,*

$$f_i = \frac{\alpha^{-il}}{1 + \alpha^{-i/2}}.$$

Hence the algebraic degree of f is equal to $n - 1$.

Proof. We have $f_0 = f(0) = 0$ and $f_{2^n-1} = 0$ since f has even Hamming weight and thus algebraic degree less than n. For $1 \le i \le 2^n - 2$, by Equality (2) we have

$$\begin{aligned} f_i &= \sum_{j=0}^{2^n-2} f(\alpha^j)\alpha^{-ij} = \sum_{j=l}^{l+2^{n-1}-1} \alpha^{-ij} = \alpha^{-il} \sum_{j=0}^{2^{n-1}-1} \alpha^{-ij} \\ &= \alpha^{-il} \frac{1 + \alpha^{-i2^{n-1}}}{1 + \alpha^{-i}} = \alpha^{-il} \frac{1 + \alpha^{-i/2}}{1 + \alpha^{-i}} = \frac{\alpha^{-il}}{1 + \alpha^{-i/2}}. \end{aligned}$$

We can see that $f_{2^n-2} \neq 0$ and therefore f has algebraic degree $n - 1$. $\square$

Remark 6. For the function f of (9), the ij-th element of the matrix $U(f; e, d)$ with $e \le d$ is equal to

$$u_{yz} = f_{y-z} = \frac{\alpha^{-yl}\alpha^{zl}}{1 + \alpha^{-y/2}\alpha^{z/2}}, \text{ for } (i, j) \neq (1, 1),$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$.

Lemma 10 *Let A be an $m \times m$ matrix with the ij-th element $a_{ij} = (1 + \beta_i\gamma_j)^{-1}$ in a field K of characteristic 2, $\beta_i, \gamma_j \in K$ and $\beta_i\gamma_j \neq 1$, $1 \le i, j \le m$. Then the determinant of A is equal to*

$$\prod_{1 \le i < j \le m} (\beta_i + \beta_j)(\gamma_i + \gamma_j) \prod_{1 \le i, j \le m} a_{ij}.$$

Furthermore, the determinant of A is non-zero if and only if $\beta_i \neq \beta_j$ and $\gamma_i \neq \gamma_j$ for $i \neq j$.

Proof. The second half part of this lemma is derived from the first half part. The proof of the first half part is given by induction on m. First we can check that the statement is certainly true for $m = 1$. Now we verify the induction step. Suppose that it holds for $m - 1$. Thus we suppose that

$$\det(A^{(1,1)}) = \prod_{2\le i<j\le m} (\beta_i + \beta_j)(\gamma_i + \gamma_j) \prod_{2\le i,j\le m} a_{ij},$$

where $A^{(i,j)}$ is the $(m-1) \times (m-1)$ matrix that results from A by removing the i-th row and the j-th column.

We wish to show that it also holds for m. Let $B = (b_{ij})_{m\times m}$ with $b_{1j} = a_{1j}$ and for $i > 1$,

$$\begin{aligned}
b_{ij} &= a_{ij} + a_{11}^{-1} a_{i1} a_{1j} \\
&= \frac{1}{1+\beta_i\gamma_j} + (\frac{1}{1+\beta_1\gamma_1})^{-1} \cdot \frac{1}{1+\beta_i\gamma_1} \cdot \frac{1}{1+\beta_1\gamma_j} \\
&= \frac{(1+\beta_i\gamma_1)(1+\beta_1\gamma_j) + (1+\beta_1\gamma_1)(1+\beta_i\gamma_j)}{(1+\beta_i\gamma_j)(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)} \\
&= \frac{\beta_i\gamma_1 + \beta_1\gamma_j + \beta_1\gamma_1 + \beta_i\gamma_j}{(1+\beta_i\gamma_j)(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)} \\
&= \frac{(\beta_1 + \beta_i)(\gamma_1 + \gamma_j)}{(1+\beta_i\gamma_j)(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)} \\
&= a_{ij} \cdot (\beta_1 + \beta_i) a_{i1} \cdot (\gamma_1 + \gamma_j) a_{1j}.
\end{aligned}$$

Let

$$P = \mathrm{diag}(1, (\beta_1 + \beta_2)a_{21}, \cdots, (\beta_1 + \beta_m)a_{m1})$$

and

$$Q = \mathrm{diag}(1, (\gamma_1 + \gamma_2)a_{12}, \cdots, (\gamma_1 + \gamma_2)a_{1m})$$

where $\mathrm{diag}(x_1, \cdots, x_m)$ denotes a diagonal matrix whose diagonal entries starting in the upper left corner are $x_1, \cdots, x_m$. Then

$$B = P \begin{pmatrix} a_{11} & * \\ 0 & A^{(1,1)} \end{pmatrix} Q.$$

Hence

$$\begin{aligned}
\det(A) &= \det(B) \\
&= \det(P) \cdot a_{11} \det(A^{(1,1)}) \cdot \det(Q) \\
&= \left(\prod_{i=2}^{m} (\beta_1 + \beta_i) a_{i1}\right) \cdot a_{11} \det(A^{(1,1)}) \cdot \left(\prod_{j=2}^{m} (\gamma_1 + \gamma_j) a_{1j}\right) \\
&= \prod_{1\le i<j\le m} (\beta_i + \beta_j)(\gamma_i + \gamma_j) \prod_{1\le i,j\le m} a_{ij}.
\end{aligned}$$

It has now been proved by mathematical induction that the first half part of this lemma holds for all positive integers m. □

Lemma 11 *Let $A = (a_{ij})_{m\times m}$ and $B = (b_{ij})_{m\times m}$ be $m \times m$ matrices with $a_{ij} = \beta_i\gamma_j b_{ij}$ and $\beta_i \neq 0, \gamma_j \neq 0$ for $1 \le i, j \le m$. Then $\det(A) \neq 0$ if and only if $\det(B) \neq 0$.*

Proof. Let $P = \mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_m)$ and $Q = \mathrm{diag}(\gamma_1, \gamma_2, \cdots, \gamma_m)$. Then $A = PBQ$ and hence $\det(A) = \det(B) \prod_{i=1}^{m} \beta_i\gamma_i$, which proves this lemma. □

Proposition 12 *Let e be a positive integer less than $n/2$ and f be the function of (9). Then $U(f;e)$ is invertible if $\binom{n-1}{e} \equiv 0 (\bmod 2)$, and $U(f;e,n-e-2)$ has full column rank if $\binom{n-1}{e} \equiv 1 (\bmod 2)$.*

Proof. Let $U = U(f;e)$ and U_{ij} be the ij-th element of U. We have $U_{11} = f_{2^n-1} = 0$. By Remark 5 we know that U is a symmetric matrix of order $\sum_{i=0}^{e}\binom{n}{i}$ in the form of (5). For the case $\binom{n-1}{e} \bmod 2 = 0$, we have $\sum_{i=0}^{e}\binom{n}{i} \bmod 2 = 0 = U_{11}$. By Remark 3 it holds that $\det(U) = \det(U^{(1,1)})$. Remark 6 shows that the ij-th element of $U^{(1,1)}$ is

$$U_{ij}^{(1,1)} = \frac{\alpha^{-yl}\alpha^{zl}}{1+\alpha^{-y/2}\alpha^{z/2}},$$

where y is the i-th element in $\overline{\mathcal{W}}_{n-e-1} \setminus \{2^n-1\}$ and z is the j-th element in $\mathcal{W}_e \setminus \{0\}$, since $e \le n-e-1$ for $e < n/2$. Let U^* be a $(\sum_{i=0}^{e}\binom{n}{i} - 1) \times (\sum_{i=0}^{e}\binom{n}{i} - 1)$ matrix with the ij-th element equal to

$$U_{ij}^{*} = \frac{1}{1+\alpha^{-y/2}\alpha^{z/2}}.$$

Since $\alpha^{-y/2} \ne \alpha^{-y'/2}$ for $y \ne y'$ $(y, y' \in \overline{\mathcal{W}}_{n-e-1} \setminus \{2^n-1\})$ and $\alpha^{z/2} \ne \alpha^{z'/2}$ for $z \ne z'$ $(z, z' \in \mathcal{W}_e \setminus \{0\})$, from Lemma 10 we have $\det(U^*) \ne 0$. Then by Lemma 11 it holds that $\det(U^{(1,1)}) \ne 0$. Hence, U is invertible.

For the case $\binom{n-1}{e} \bmod 2 = 1$, we consider the $\sum_{i=0}^{e+1}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix $U(f;e,n-e-2)$. For even n, we always have $e \le n-e-2$ for $e < n/2$. For odd n, we always have $e \le n-e-2$ for $e \le (n-3)/2$ and $\binom{n-1}{e} \bmod 2 = 0$ for $e = (n-1)/2$. Thus for $\binom{n-1}{e} \bmod 2 = 1$ and $e < n/2$, we always have $e \le n-e-2$. Let U^{**} be the $\sum_{i=0}^{e}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix that results from $U(f;e,n-e-2)$ by removing the first $\binom{n}{e+1}$ rows. A similar proof of $\det(U^{(1,1)}) \ne 0$ also applies to $\det(U^{**}) \ne 0$. Then $U(f;e,n-e-2)$ has full column rank. □

Theorem 13 *Let e be a positive integer less than $n/2$ and f be the function of (9). Then f admits no non-zero function g with degree at most e such that gf has degree at most $n-e-1$ if $\binom{n-1}{e} \equiv 0 (\bmod 2)$, and admits no non-zero function g with degree at most e such that gf has degree at most $n-e-2$ if $\binom{n-1}{e} \equiv 1 (\bmod 2)$.*

Proof. It is derived from Theorem 8 and Proposition 12. □

Corollary 14 *Let $n = 2^s + 1$ and $f \in \mathbf{B}_n$ be the function of (9). Then f is $\mathcal{PAI}$.*

Proof. It is obtained from Theorem 13 since $\binom{n-1}{e} = \binom{2^s}{e} \equiv 0 (\bmod 2)$ for $1 \le e < n/2$. □

Theorem 13 states that the Carlet-Feng functions achieve the bounds of Theorem 5 and thus the bounds of Theorem 5 are tight for the functions with algebraic degree less than n, while Corollary 14 states that the Carlet-Feng functions on 2^s+1 variables are $\mathcal{PAI}$.

Next we consider the Boolean functions with algebraic degree equal to n.

Let n be an integer and α a primitive element of $\mathbb{F}_{2^n}$. Let $f \in \mathbf{B}_n$ and

$$\mathrm{supp}(f) = \{0, \alpha^l, \alpha^{l+1}, \cdots, \alpha^{l+2^{n-1}-1}\}, 0 \le l \le 2^n - 2. \tag{10}$$

The function (10) is a function that results from the function (9) by flipping the output at $x = 0$.

A similar proof of Proposition 9 applies to the following result.

Proposition 15 *Let $\sum_{i=0}^{2^n-1} f_i x^i (f_i \in \mathbb{F}_{2^n})$ be the univariate representation of the function f of (10). Then $f_0 = 1$, $f_{2^n-1} = 1$, and for $1 \le i \le 2^n - 2$,*

$$f_i = \frac{\alpha^{-il}}{1+\alpha^{-i/2}}.$$

Hence the algebraic degree of f is equal to n.

A similar proof of Proposition 12 also applies to the following result.

Proposition 16 *Let e be a positive integer less than $(n-1)/2$ and f be the function of (10). Then $U(f;e)$ is invertible if $\binom{n-1}{e} \equiv 1 (\mathrm{mod}\, 2)$, and $U(f;e,n-e-2)$ has full column rank if $\binom{n-1}{e} \equiv 0 (\mathrm{mod}\, 2)$.*

Theorem 17 *Let e be a positive integer less than $(n-1)/2$ and f be the function of (10). Then f admits no non-zero function g with degree at most e such that gf has degree at most $n-e-1$ if $\binom{n-1}{e} \equiv 1 (\mathrm{mod}\, 2)$, and admits no non-zero function g with degree at most e such that gf has degree at most $n-e-2$ if $\binom{n-1}{e} \equiv 0 (\mathrm{mod}\, 2)$.*

Proof. It is confirmed by Theorem 8 and Proposition 16. □

Corollary 18 *Let $n = 2^s$ and $f \in \mathbf{B}_n$ be the function of (10). Then f is $\mathcal{PAI}$.*

Proof. It is obtained from Theorem 17 since $\binom{n-1}{e} = \binom{2^s-1}{e} \equiv 1 (\mathrm{mod}\, 2)$ for $1 \le e < n/2$. □

Theorem 17 states that the modified Carlet-Feng functions achieve the bounds of Theorem 5 and thus the bounds of Theorem 5 are tight for the functions with algebraic degree equal to n, while Corollary 18 states that the modified Carlet-Feng functions on 2^s variables are $\mathcal{PAI}$. Here we do not consider the case $e = (n-1)/2$ for odd n, since the algebraic immunity of an n-variable Boolean function with algebraic degree n is less than or equal to $(n-1)/2$ for odd n.

Consequently, as mentioned above, the bounds of Theorem 5 are tight and there exist $\mathcal{PAI}$ functions on 2^s and 2^s+1 variables. 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 $n = 2^s$.

5 The immunity of Boolean functions with maximum $\mathcal{AI}$ against probabilistic algebraic attacks

This section mainly focuses on the time complexities of probabilistic algebraic attacks on an LFSR-based nonlinear filter generator with the filter function achieving maximum $\mathcal{AI}$.

Let p be the probability for **S4a** or **S4b** (see Section 1). Then an overdetermined system of nonlinear equations with degree r is obtained where each equation holds with probability p. One can use the linearization algorithm to solve the system, where $R = \sum_{i=0}^{r} \binom{L}{i}$ equations are used and hold with probability p^R. Then the time complexity of probabilistic algebraic attacks is $p^{-R}R^w$, where $w \approx 2.807$ is the exponent of the Gaussian reduction.

In the affine case, probabilistic algebraic attacks are related to the (fast) correlation attacks [3], so we always consider the nonlinear case here. Recall that the maximum $\mathcal{AI}$ of an n-variable function is $\lceil \frac{n}{2} \rceil$. Then, for the case $r \ge \lceil \frac{n}{2} \rceil$, deterministic algebraic attacks can be used. Therefore hereinafter we always assume that $2 \le r \le \lceil \frac{n}{2} \rceil - 1$.

Let g be a non-zero function with degree at most r. For a balanced function we know that the maximum probability for applying **S4a** or **S4b** is

$$p_{\max} = 1 - \frac{\min\{\mathrm{d}(gf,g), \mathrm{d}(gf,0)\}}{2^{n-1}}.$$

According to [4, Proposition 5] we have

$$d_r = \min\{\mathrm{d}(gf,g), \mathrm{d}(gf,0)\} \ge \sum_{i=0}^{\mathcal{AI}(f)-r-1} \binom{n-r}{i}.$$

Since $\mathcal{AI}(f) \leq \lceil \frac{n}{2} \rceil$, we have $2\mathcal{AI}(f) - 2r - 1 < n - r$ and therefore for $r \leq \mathcal{AI}(f) - 1$,

$$\sum_{i=0}^{\mathcal{AI}(f)-r-1} \binom{n-r}{i} \geq \sum_{i=0}^{\mathcal{AI}(f)-r-1} \binom{2\mathcal{AI}(f) - 2r - 1}{i} = 2^{2\mathcal{AI}(f)-2r-2}.$$

Then for a function with maximum $\mathcal{AI}$ we have $d_r \geq 2^{n-2r-2}$ and therefore

$$p_{\max} = 1 - \frac{d_r}{2^{n-1}} \leq 1 - 2^{-2r-1}.$$

It is well known that the real function $1 - x - e^{-x}$ is decreasing when $x \geq 0$. Hence we have

$$p_{\max} \leq 1 - 2^{-2r-1} \leq e^{-2^{-2r-1}}$$

and the time complexity of probabilistic algebraic attacks

$$p^{-R} R^w \geq p_{\max}^{-R} \geq (e^{-2^{-2r-1}})^{-R} = e^{R/2^{2r+1}} \geq 2^{1.44R/2^{2r+1}} \geq 2^{1.44\binom{L}{r}/2^{2r+1}}.$$

For $r \leq L/5$, we have

$$\frac{1}{2^{2r+1}} \binom{L}{r} \geq \frac{1}{2^{2r-1}} \binom{L}{r-1},$$

and it then holds that

$$p^{-R} \geq 2^{1.44\binom{L}{r}/2^{2r+1}} \geq 2^{1.44\binom{L}{2}/2^5}. \tag{11}$$

Corollary 9 of [21, Page 310] states that for $0 < \mu < 1/2$,

$$\sum_{i=0}^{\mu L} \binom{L}{i} \geq \frac{2^{H_2(\mu)L}}{\sqrt{8L\mu(1-\mu)}},$$

where $H_2(\mu) = -\mu \log_2 \mu - (1-\mu)\log_2(1-\mu)$. For $L/5 < r < L/2$, it follows that

$$R^w \geq \left(\sum_{i=0}^{L/5} \binom{L}{i} \right)^{2.807} \geq \left(\frac{2^{H_2(1/5)L}}{\sqrt{32L/25}} \right)^{2.807} \geq \frac{2^{2.02L}}{1.42L^{1.41}}. \tag{12}$$

From (11) and (12) we can calculate that for $L \geq 46$,

$$p^{-R} R^w \geq 2^L.$$

Consequently, probabilistic algebraic attacks are worse than exhaustive key search in the context of their application to the nonlinear filter generator if the filter function achieves maximum $\mathcal{AI}$ and the size L of the LFSR is greater than or equal to 46. Since a $\mathcal{PAI}$ function has maximum $\mathcal{AI}$, the function also behaves good against probabilistic algebraic attacks.

As a matter of fact, a similar proof shows that for practical sizes L (e.g. $L = 256$) of the LFSR and reasonable number n of input variables, probabilistic algebraic attacks are worse than exhaustive key search if the $\mathcal{AI}$ of the filter function is a little smaller than the maximum value $\lceil \frac{n}{2} \rceil$.

6 Conclusion

In this paper, several open problems about the immunity of Boolean functions against algebraic attacks have been solved. We proved the maximum immunity to fast algebraic attacks, identified the immunity of the Carlet-Feng functions against fast algebraic attacks, and evaluated the resistance of Boolean functions with maximum algebraic immunity against probabilistic algebraic attacks. It seems that for a balanced function the optimal value of the number n of input variables is $2^s + 1$ in terms of immunity against fast algebraic attacks. The Carlet-Feng functions previously shown to have maximum algebraic immunity and good nonlinearity are proved to be optimal against fast algebraic attacks among the balanced functions. To the best of our knowledge this is the first time that a function is shown to have such cryptographic property.

Acknowledgement

The authors thank Tianze Wang for his careful reading of the manuscript, and thank Shaoyu Du, Lin Jiao, Yao Lu, Wenlun Pan, and Wenhao Wang for their participation in FAA seminar at SKLOIS in December 2011. The first author would like to thank Dingyi Pei for his enlightening conversations on the resistance of Boolean functions against algebraic attacks.

References

1. F. Armknecht. Improving fast algebraic attacks. In: B. Roy and W. Meier (eds.) FSE 2004. LNCS 3017, pp. 65–82. Berlin, Heidelberg: Springer, 2004.
2. F. Armknecht, C. Carlet, P. Gaborit, et al. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. In: S. Vaudenay (eds.) EUROCRYPT 2006. LNCS 4004, pp. 147–164. Berlin, Heidelberg: Springer, 2006.
3. A. Braeken and B. Preneel. Probabilistic algebraic attacks. In 10th IMA International Conference on Cryptography and Coding, 2005, LNCS 3796, pp. 290–303. Springer-Verlag, 2005.
4. C. Carlet. On the higher order nonlinearities of algebraic immune functions. CRYPTO 2006, LNCS 4117, 584–601. Berlin, Heidelberg: Springer, 2007.
5. C. Carlet. Boolean functions for cryptography and error correcting codes. In: Y. Crama, P. Hammer, eds. Boolean Methods and Models in Mathematics, Computer Science, and Engineering, pp. 257–397. Cambridge: Cambridge University Press, 2010.
6. 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: ASIACRYPT 2008, LNCS 5350, 425–440. Berlin, Heidelberg: Springer, 2008.
7. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. Advances in Cryptology-EUROCRYPT 2003, LNCS 2656, 345–359. Berlin, Heidelberg: Springer, 2003.
8. N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. Advances in Cryptology-CRYPTO 2003, LNCS 2729, 176–194. Berlin, Heidelberg: Springer, 2003.
9. N. Courtois. Cryptanalysis of Sfinks. ICISC 2005, LNCS 3935, 261–269. Berlin, Heidelberg: Springer, 2006.
10. D. K. Dalai, S. Maitra, and S. Sarkar. Basic theory in construction of Boolean functions with maximum possible annihilator immunity. Designs, Codes and Cryptography, vol. 40, no. 1, 41–58, 2006.
11. D. K. Dalai, K. C. Gupta, and S. Maitra. Results on algebraic immunity for cryptographically significant Boolean functions. INDOCRYPT 2004, LNCS 3348, 92–106. Berlin, Heidelberg: Springer, 2005.
12. Y. Du, F. Zhang, and M. Liu. On the resistance of Boolean functions against fast algebraic attacks. ICISC 2011, LNCS 7259, 261–274. Berlin, Heidelberg: Springer 2012.
13. 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.
14. S. Fischer and W. Meier. Algebraic immunity of S-boxes and augmented functions. In: Biryukov, A. (ed.) FSE 2007. LNCS 4593, pp. 366–381. Springer, 2007.
15. G. Gong. Sequences, DFT and resistance against fast algebraic attacks. SETA 2008, LNCS 5203, pp. 197–218, 2008.
16. P. Hawkes and G. Rose. Rewriting variables: the complexity of fast algebraic attacks on stream ciphers, in Crypto 2004, LNCS 3152, pp. 390–406. Springer, 2004.
17. N. Li, L. Qu, W. Qi, et al. On the construction of Boolean Functions with optimal algebraic immunity. IEEE Transactions on Information Theory, vol. 54, no. 3, 1330–1334, 2008.
18. N. Li and W. Qi. Construction and analysis of Boolean functions of $2t+1$ variables with maximum algebraic immunity. ASIACRYPT 2006, LNCS 4284, pp. 84–98. Berlin, Heidelberg: Springer, 2006.
19. M. Liu, D. Lin, and D. Pei. Fast algebraic attacks and decomposition of symmetric Boolean functions. IEEE Transactions on Information Theory, vol. 57, no. 7, pp. 4817–4821, 2011.
20. M. Liu, D. Lin, and D. Pei. Results on the immunity of Boolean functions against probabilistic algebraic attacks. U. Parampalli, P. Hawkes (Eds.): ACISP 2011, LNCS 6812, pp. 34–46. Berlin, Heidelberg: Springer, 2011.
21. F.J. MacWilliams and N.J.A. Sloane. The theory of error correcting codes. New York: North-Holland, 1977.
22. W. Meier, E. Pasalic, and C. Carlet. Algebraic attacks and decomposition of Boolean functions. Advances in Cryptology-EUROCRYPT 2004, LNCS 3027, 474–491. Berlin, Heidelberg: Springer, 2004.
23. E. Pasalic. Almost fully optimized infinite classes of Boolean functions resistant to (fast) algebraic cryptanalysis. ICISC 2008, LNCS 5461, 399–414. Berlin, Heidelberg: Springer, 2009.
24. E. Pasalic. Probabilistic versus deterministic algebraic cryptanalysis – a performance comparison. IEEE Transactions on Information Theory, vol. 55, no. 11, pp. 5233–5240, 2009.
25. P. Rizomiliotis. On the resistance of Boolean functions against algebraic attacks using univariate polynomial representation. IEEE Transactions on Information Theory, vol. 56, no. 8, pp. 4014–4024, 2010.
26. P. Rizomiliotis. On the security of the Feng-Liao-Yang Boolean functions with optimal algebraic immunity against fast algebraic attacks. Designs, Codes and Cryptography, vol. 57, no. 3, pp. 283-292, 2010.

27. Z. Tu and Y. Deng. A conjecture about binary strings and its applications on constructing Boolean functions with optimal algebraic immunity. Designs, Codes and Cryptography, vol. 60, no. 1, pp. 1–14, 2011.
28. X. Zeng, C. Carlet, J. Shan, and L. Hu. More balanced Boolean functions with optimal algebraic immunity and good nonlinearity and resistance to fast algebraic attacks. IEEE Transactions on Information Theory, vol. 57, no. 9, pp. 6310–6320, 2011.
29. Y. Zhang, M. Liu, and D. Lin. On the immunity of rotation symmetric Boolean functions against fast algebraic attacks. Cryptology ePrint Archive, Report 2012/111, http://eprint.iacr.org/