Resistance of S-boxes against Algebraic Attacks

Jung Hee Cheon[1] and Dong Hoon Lee[2]

[1] Department of Mathematics, Seoul National University
jhcheon@math.snu.ac.kr,
[2] National Security Research Institute (NSRI)
dlee@etri.re.kr

Abstract. We develop several tools to derive linear independent multivariate equations from algebraic S-boxes. By applying them to maximally nonlinear power functions with the inverse exponents, Gold exponents, or Kasami exponents, we estimate their resistance against algebraic attacks. As a result, we show that S-boxes with Gold exponents have very weak resistance and S-boxes with Kasami exponents have slightly better resistance against algebraic attacks than those with the inverse exponents.

Keywords: Algebraic Attack, S-boxes, Boolean Functions, Nonlinearity, Differential Uniformity

1 Introduction

Recently, Courtois and Pieprzyk proposed an algebraic attack for block ciphers [4]. Their attack on AES [11] exploits algebraic properties of S-boxes: If we can obtain many equations of small number of monomials from S-boxes, a block cipher with the S-boxes can be represented by many equations of small number of variables. By solving these multivariate equations by so called the *XSL* algorithm, we may find the key of the block cipher.

In the AES case, they introduce another viewpoint of the S-box as a quadratic equation $xy = 1$ in x and y rather than as a higher degree equation $y = 1/x$ in x, and obtain additional quadratic equations by multiplying appropriate monomials. More precisely, they obtain 23 quadratic equations with a total of 81 distinct terms from the S-box of AES and show that the equations are linearly independent by simulation.

In this paper, we give a theoretical approach to obtain *linearly independent* multivariate equations from algebraic S-boxes. Multivariate equations are said to be linearly independent if they are linearly independent when every distinct monomial is considered as a new variable. We develop three tools to prove *linear independence.* The first tool is that if a vector Boolean function is nonlinear, their component functions should be linearly independent as multivariate equations. We apply this to $n \times n$ S-boxes x^{2^k+1} and $n \times 2n$ S-boxes $(x^{2^k+1}, x^{2^{k+1}+1})$ over $\mathbb{F}_{2^n}$ which are known to be nonlinear when $\gcd(n, 2k) = 1$ and $|k - n/2| > 1$, respectively [5]. The second one is that if for a vector Boolean function $F(x, y) : \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ and $g : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, $F(x, g(x))$ has m linearly independent component functions, so has $F(x,y)$. The third one is that linear independence of multivariate functions is *invariant* under affine transformation of inputs and linear transformation of outputs.

By applying these tools, we can prove that $5n$ equations obtained from the inverse function $xy = 1$ in $\mathbb{F}_{2^n}$ (or its affine transformation) are linearly independent for any positive integer n. Further we apply them to estimate the resistance of power functions with well-known Gold exponents and Kasami exponents against algebraic attacks [7, 8]. Those S-boxes are the only power functions which are known to be maximally nonlinear (MN) and almost perfect nonlinear (APN) [6]. Note that 'MN' and 'APN' imply the best resistance against linear cryptanalysis and differential cryptanalysis, respectively [1, 2, 9]. Our analysis shows that the S-boxes with Gold exponents have very weak resistance and the S-boxes with Kasami exponents have better resistance against algebraic attacks while all of them have similar resistance against differential and linear cryptanalysis. It would be an interesting problem to apply algebraic attacks to the ciphers using Gold power functions as S-boxes such as MISTY [10] which is selected as standard block algorithms in NESSIE [12].

In Section 2, we introduce some preliminaries on nonlinearity, APN, and resistance against algebraic attacks. In Section 3, we propose some auxiliary lemmas used to show the linear independence of multivariate equations. In Section 4, we deal with the resistance of the above three families of S-boxes and compare them. We conclude in Section 5.

2 Preliminaries

In this section, we introduce the definitions of nonlinearity, APN, and resistance against algebraic attacks, and remind some useful results for algebraic S-boxes.

Definition 1. *A function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is called a* almost perfect nonlinear *(APN) if each equation*

$$F(x+a) - F(x) = b \quad \text{for } a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^n}$$

has at most two solutions $x \in \mathbb{F}_{2^n}$.

Note that APN functions have the best resistance against differential cryptanalysis. When n is odd, we have many classes of APN power functions. But when n is even, we have only two classes of APN power functions, that is, Gold exponents and Kasami exponents [7, 8, 6].

The *Hamming distance* between two Boolean functions $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ and $g : \mathbb{F}_{2^n} \to \mathbb{F}_2$ is the weight of $f + g$. The minimal distance between f and any affine function from $\mathbb{F}_{2^n}$ into $\mathbb{F}_2$ is the *nonlinearity* of f. Given a *vector Boolean function* $F = (f_1, \ldots, f_m) : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$, $b \cdot F$ denotes the Boolean function $b_1f_1 + b_2f_2 + \cdots + b_mf_m$ for each $b = (b_1, b_2, \cdots, b_m) \in \mathbb{F}_{2^m}$. Then the nonlinearity of F is defined as minimal nonlinearity of component functions as follows:

Definition 2. *The nonlinearity of F, $\mathcal{N}(F)$, is defined as*

$$\mathcal{N}(F) = \min_{b \in \mathbb{F}_{2^m}^*} \mathcal{N}(b \cdot F) = \min_{b \neq 0, \phi \in \mathcal{A}} wt(b \cdot F + \phi)$$

where $\mathcal{A}$ is the set of all affine functions over $\mathbb{F}_{2^n}$.

It is known that $\mathcal{N}(F) \leq 2^{n-1} - 2^{\frac{n-1}{2}}$. If n is odd, $\mathcal{N}(F)$ can be maximal, we call such functions *maximally nonlinear (MN)* functions. For even n, it is an open question to determine the maximal value. It is known that if n is odd and F is maximally nonlinear then F is almost perfect nonlinear [6].

Now we define the resistance against algebraic attacks as in [4].

Definition 3. *Given r equations of t monomials in $\mathbb{F}_2^n$, we define $\Gamma = ((t-r)/n)^{\lceil (t-r)/n \rceil}$ as the resistance of algebraic attacks (RAA).*

This quantity was introduced by Courtois and Pieprzyk [4]. They showed that the S-box of AES and the S-boxes of Serpent have $\Gamma \approx 2^{22.9}$ and $\Gamma \approx 2^{8.0}$, respectively. They claimed it can be a serious weakness of these ciphers and Γ should be greater than 2^{32} for secure ciphers.

Note that this measure is not an exact measure of XSL algorithm and an improvement of algorithm on solving multivariate equations may result in different measures. However, it is true that this quantity reflects a difficulty of solving multivariate equations in some sense. Thus we will use this quantity to measure the resistance of algebraic attacks in this paper.

3 Auxiliary Lemmas

Definition 4. *Given Boolean functions $f_1, \ldots, f_m$ from $\mathbb{F}_2^n$ to $\mathbb{F}_2$, they are said to be linearly independent over $\mathbb{F}_2$ if they are linearly independent as multivariate polynomials, or equivalently if $\sum_{i=1}^m a_i f_i(x) = 0$ for all $x \in \mathbb{F}_2^n$ with $a_1, \ldots, a_m \in \mathbb{F}_2$ implies $a_1 = \cdots = a_m = 0$.*

Lemma 1. *Consider two vector Boolean functions $F(x, y) : \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ and $g : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$. If $F(x, g(x))$ has m linearly independent component functions, so does $F(x, y)$ in $\mathbb{F}_2[x_1, \ldots, x_n, y_1, \ldots, y_n]$.*

Proof. Suppose that $F(x, y) = (f_1(x, y), \ldots, f_m(x, y))$ has m linearly dependent component functions, *i.e.* there are not-all-zero $a_1, \ldots, a_m \in \mathbb{F}_2$ such that $\sum_{i=1}^m a_i f_i(x, y) = 0$. Then we have $\sum_{i=1}^m a_i f_i(x, g(x)) = 0$, which implies that $f_i(x, g(x))$'s are linearly dependent. It contradicts that $F(x, g(x))$ has m linearly independent components. Therefore $F(x, y)$ should have m linearly independent component functions.

Lemma 2. *Any permutation $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ has n linearly independent component functions.*

Proof. Suppose that there exist not-all-zero $a_1, \ldots, a_n \in \mathbb{F}_2$ such that $\sum_{i=1}^{n} a_i f_i(x) = 0$ for $F = (f_1, \ldots, f_n)$. Then the image of F is a subset of the hyperplane given by $\sum_{i=1}^{n} a_i f_i(x) = 0$. Since the hyperplane has dimension less than n, F can not be a permutation. Therefore if F is a permutation, its n component functions should be linearly independent.

Lemma 3. *Consider a vector Boolean function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$. If the nonlinearity of F is non-zero, F has m linearly independent component functions.*

Proof. Suppose that there exist not-all-zero $a_1, \ldots, a_m \in \mathbb{F}_2$ such that $\sum_{i=1}^{m} a_i f_i(x) = 0$ for $F = (f_1, \ldots, f_m)$. If we take $b = (a_1, \ldots, a_m)$, we can see that $b \cdot F$ is a zero function and so has zero nonlinearity. Thus the nonlinearity of F, the minimum of nonlinearity of the component functions, is also zero. Therefore any nonlinear function should have m linearly independent component functions.

For the nonlinearity of S-boxes, we have the following results [5]:

$$\mathcal{N}(x^{2^k+1}) \geq 2^{n-1} - 2^{\frac{n+\gcd(n,2k)}{2}-1}, \tag{1}$$
$$\mathcal{N}(x^3, x^5, \cdots, x^{2k+1}) \geq 2^{n-1} - k \cdot 2^{\frac{n}{2}}. \tag{2}$$

By applying these results to Lemma 3, we obtain the following corollary:

Corollary 1. *Let k be a positive integer.*
(1) If n does not divide $2k$, x^{2^k+1} has n linearly independent component functions.
(2) If $k \leq 2^{n/2-1}$, $F = (x^3, x^5, \cdots, x^{2k+1}) : \mathbb{F}_{2^n} \to \mathbb{F}_{2^{kn}}$ have kn linearly independent component functions.

3.1 Invariants under Transformations

Now we show that linear independence is invariant under invertible transformations of inputs and invertible linear transformations of outputs.

Lemma 4. *Let $T : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an invertible transformation and $S : \mathbb{F}_2^m \to \mathbb{F}_2^m$ an invertible linear transformation. A vector Boolean function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ has m linearly independent component functions over $\mathbb{F}_2$ if and only if so does $S \circ F \circ T$.*

Proof. Since we consider invertible transformations T and S, we are enough to show that F has m linearly independent component functions when either $F \circ T$ or $S \circ F$ does.

Let $F(x) = (f_1(x), \ldots, f_m(x))$ for $x \in \mathbb{F}_2^n$. Assume $a_1, \ldots, a_m \in \mathbb{F}_2$ satisfies $\sum_{i=1}^{m} a_i f_i(x) = 0$ for all $x \in \mathbb{F}_2^n$. Since T is invertible, we have $\sum_{i=1}^{m} a_i f_i(Ty) = 0$ for all $y \in F_2^n$. Since $F \circ T$ has m linearly independent component functions, we have $a_1 = \cdots = a_m = 0$, which implies the independence of m component functions of F.

If we let $S^{-1} = (p_{ij})$ for p_{ij}'s $\in \mathbb{F}_2$ and $S \circ F = (g_1, \ldots, g_m)$, we have $f_i = \sum_{j=1}^{m} p_{ij} g_j$. If there are not-all-zero $a_1, \ldots, a_m \in \mathbb{F}_2$ satisfying $\sum_{i=1}^{m} a_i f_i(x) = 0$, we have

$$\sum_{i=1}^{m} \{\sum_{j=1}^{m} a_i p_{ij} g_j(x)\} = \sum_{j=1}^{m} \{\sum_{i=1}^{m} a_i p_{ij}\} g_j(x) = 0. \tag{3}$$

Since $g_1, \ldots, g_m$ are linearly independent, $\sum_{i=1}^{m} a_i p_{ij} = 0$ for all j. We can see $a_1 = \cdots = a_m = 0$ from the invertibility of $S^{-1} = (p_{ij})$. Hence m component functions of F should be linearly independent.

Remark that if S is an *affine* transformation, Lemma 4 does not hold. For example, $F : \mathbb{F}_2^2 \to \mathbb{F}_2^3 : (x_1, x_2) \mapsto (x_1 + 1, x_2 + 1, x_1 + x_2 + 1)$ has 3 linearly independent components, but after the affine transformation $S : \mathbb{F}_2^3 \to \mathbb{F}_2^3 : (x, y, z) \mapsto (x + 1, y + 1, z + 1)$ is taken to F, $S \circ F = (x_1, x_2, x_1 + x_2)$ is not linearly independent anymore. However, if we consider a constant term as one of variables, we can have this invariant property. That is, if $1, f_1, \ldots, f_m$ are linearly independent, $S(f_1, \ldots, f_m)$ are linearly independent. Also if all of f_i's do not have constant terms, independence property is preserved under an affine transformation S.

4 Independent Equations

From now on, we consider a polynomial over a finite field. If we fix a basis, this polynomial can be regarded as multivariate equations. Unless confused, we will consider a polynomial as multivariate equations without specifying a basis.

Because equations of higher degree than two do not help in the point of algebraic attacks to S-box, our purpose is to get linearly independent equations whose degree are at most two as many as possible. When we are given m quadratic equations from $F(x) = 0$, we can consider the following methods to get more quadratic equations:

1. Multiplication by linear or quadratic equations.
2. Composition with quadratic equations.

Note that composition of a monomial with affine equations gives only dependent equations and composition with equations of higher degree usually gives equations of higher degree.

The first case is restricted by the following lemma.

Lemma 5. *Suppose that $n > 2$ and $k \geq 1$. Assume that the Hamming weight of d is at most 2. The product x^m of two monomials x^{2^k+1} and x^d is linear or quadratic only in the following cases:*

1. *If $d = 1$, then $m = \begin{cases} 4 & \text{if } k = 1, \text{ (Linear)} \\ 2^k + 2 & \text{if } k \neq 1. \text{ (Quadratic)} \end{cases}$*
2. *If $d = 2^k$, then $m = 1 + 2^{k+1}$. (Quadratic)*

3. *If* $d = 3$*, then* $m = \begin{cases} 2^3 & \textit{if } k = 2, \textit{ (Linear)} \\ 2^k + 2^2 & \textit{if } k \neq 2. \textit{ (Quadratic)} \end{cases}$
4. *If* $d = 2^k + 1$*, then* $m = 2^{k+1} + 2$*. (Quadratic)*
5. *If* $d = 2^{k+1} + 2^k$*, then* $m = 2^{k+2} + 1$*. (Quadratic)*

Proof. It is sufficient to check the Hamming weight of $m = 2^k + 1 + d \mod (2^n - 1)$, since $x^{2^n - 1} = 1$. Assume that $w(d) = 1$, i.e $d = 2^l$ for some $l < n$. Then m becomes $1 + 2^k + 2^l < (2^n - 1)$. Unless two of $\{0, k, l\}$ are equal, x^m is cubic. This covers first two cases of the lemma.

Assume that $w(d) = 2$, i.e $d = 2^l + 2^s$ for some $l < s < n$. Then m becomes $1 + 2^k + 2^l + 2^s < (2^n - 1)$. If all of $\{0, k, l, s\}$ are distinct, then x^m is quartic. Hence at least two of them are equal, especially $l = 0$ or $l = k$ since $0 < k$. If $l = 0$ then s should be 1 or k (Case 3 and 4). If $l = k$ then s should be $k + 1$ (Case 5). This completes the proof.

4.1 Inverse Exponents

First we count the number of linearly independent equations from $xy - 1 = 0$. A composition of $xy - 1 = 0$ with any quadratic equation gives a equation of degree larger than two. In order to get another quadratic equations, we must multiply linear or quadratic equations:

1. The original equation: $F(x, y) = xy - 1$
2. Multiplied by x: $G_0(x, y) = x^2 y - x$
3. Multiplied by y: $H_0(x, y) = xy^2 - y$
4. Multiplied by x^3: $G_1(x, y) = x^4 y - x^3$
5. Multiplied by y^3: $H_1(x, y) = xy^4 - y^3$

First, we must show that each of equations has n linearly independent component functions. Using Lemma 1 and Lemma 2, we can easily see that $F(x, y)$ has n linearly independent component functions since $F(x, y) = xy - 1$ is permutation for any nonzero y. Each component of G_0 and H_0 has a unique variable x_i and y_i respectively, hence they are linearly independent. Both G_1 and H_1 have n linearly independent components by Lemma 1, Lemma 4, and Corollary 1 using the following equations:

$$G_1(x, ax^{2^n - 2}) = (a - 1)x^3$$
$$H_1(ay^{2^n - 2}, y) = (a - 1)y^3$$

since any non-zero $(a - 1)$ is an invertible linear transformation.

In order to show that all components produced by the above polynomials are linearly independent, it is better to look at the matrix form. Each row corresponds to the equations from $G = (G_0, G_1)$, $H = (H_0, H_1)$, and F.

$$\begin{pmatrix} M_1 & 0 & M_2 & 0 \\ 0 & M_3 & M_4 & 0 \\ 0 & 0 & M_5 & M_6 \end{pmatrix} \begin{pmatrix} x_i x_j \\ y_i y_j \\ x_i y_j \\ 1 \end{pmatrix} = 0,$$

where each M_i represents a nonzero matrix and each monomial in the column vector represents all monomials of similar forms (For example, x_ix_j represents all x_ix_j for $1 \le i, j \le n$.).

It is sufficient to show that the rank of the coefficient matrix is $5n$. If we consider the coefficient matrix as a 3×3 block matrix, we can see that the rank is the sum of the ranks of M_1, M_3, and $(M_5\ M_6)$. Since F has n linearly independent components, we know that the rank of $(M_5\ M_6)$ is n.

Lemma 6. *Each of the ranks of M_1 and M_3 is $2n$.*

Proof. We refine the monomials x_ix_j for $1 \le i, j \le n$ as x_i and x_ix_k for $1 \le i < k \le n$. Then $M_1(x_ix_j)$ is expressed as the following:

$$M_1(x_ix_j) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x_i \\ x_ix_k \end{pmatrix}.$$

Since $(A\ B)$ represents the term x in G_0, A is the identity matrix of size n and $B = 0$. Since $(C\ D)$ represents the term $-x^3$ in G_1, we can write $-x^3 = C(x_i) + D(x_ix_k)$. Since $C(x_i)$ is a linear function over $\mathbb{F}_2^n$, the nonlinearity of $D(x_ix_k)$ is equal to that of x^3. Therefore $D(x_ix_k)$ has n linearly independent components by Lemma 3, hence the rank of D is n. This implies that the rank of M_1 is $2n$.

We can show that the rank of M_3 is also $2n$ by the similar argument.

Now we are ready to measure the resistance of S-boxes with inverse exponents by Γ value. The type and the number of distinct monomials in the equations from F, G, and H is as the following table.

Table 1. The type and the number of distinct monomials

Eq.	Type	#
F	x_iy_j, 1	n^2+1
G_0	x_iy_j, x_i	n^2+n
H_0	x_iy_j, y_i	n^2+n
G_1	x_iy_j, x_ix_j, x_i	$\frac{3n(n+1)}{2}$
H_1	x_iy_j, y_iy_j, y_i	$\frac{3n(n+1)}{2}$

From Table 1, we have the following theorem.

Theorem 1. *Consider $xy = 1$ in $\mathbb{F}_{2^n}$. Let t be the number of monomials and r the number of linearly independent equations. Then we can have the following parameters (r, t, Γ) for $xy = 1$:*

1. $\left(n, n^2+1, \left(\frac{n^2-n+1}{n}\right)^{\lceil \frac{n^2-n+1}{n} \rceil}\right)$ *for F*

2. $\left(2n, n^2+n+1, \left(\frac{n^2-n+1}{n}\right)^{\lceil \frac{n^2-n+1}{n} \rceil}\right)$ *for* F *and* $\{G_0$ *or* $H_0\}$

3. $\left(3n, n^2+2n+1, \left(\frac{n^2-n+1}{n}\right)^{\lceil \frac{n^2-n+1}{n} \rceil}\right)$ *for* F*,* G_0*, and* H_0

4. $\left(4n, \frac{(3n+2)(n+1)}{2}, \left(\frac{3n^2-3n+2}{2n}\right)^{\lceil \frac{3n^2-3n+2}{2n} \rceil}\right)$ *for* F*,* G_0*,* H_0 *and* $\{G_1$ *or* $H_1\}$

5. $\left(5n, 2n^2+n+1, \left(\frac{2n^2-4n+1}{n}\right)^{\lceil \frac{2n^2-4n+1}{n} \rceil}\right)$ *for all 5 polynomials*

4.2 Gold Exponents

When $\gcd(k, n) = 1$, 2^k+1 is called a Gold exponent [7]. Note that any quadratic monomial can be changed into a monomial with a Gold exponent by an affine transformation. By multiplying monomials, we obtain

1. The original equation: $F_1(x, y) = x^{2^k+1} - y$
2. Multiplied by linear equations: $F_2(x, y) = x^{2^k+2} - xy$ and $F_3(x, y) = x^{2^{k+1}+1} - x^{2^k}y$
3. Multiplied by $x^{d_1}y^{d_2}$: $F_4(x, y) = x^4y - xy^2$ only for $k = 1$
4. Composition with x^d: $F_5(x, y) = x^9 - y^3$ only for $k = 1$.

Since the original equation consists of x^{2^k+1} and y, we should multiply monomials of type x^d or $x^{d_1}y^{d_2}$. In the first case, x^d should be linear so that we have $d = 1$ or $d = 2^k$ by Lemma 5. In the second case, $x^{2^k+1+d_1}$, y^{d_2}, x^{d_1}, and y^{1+d_2} should be linear so that $(d_1, d_2) = (1, 1)$.

For composition case, if d is 2^s, the product produces only dependent equations on the original equations. Thus the Hamming weight of d should be two. Then $m = (2^k + 1)(1 + 2^l) = 1 + 2^l + 2^k + 2^{k+l}$. Only when $l = k = 1$, x^m can be quadratic.

F_1 has n independent component functions since each component contains distinct y_i. We can see that $F_2(x, ax^{2^k+1}) = (1-a)x^{2^k+2}$ and $F_3(x, ax^{2^k+1}) = (1-a)x^{2^{k+1}+1}$. When $k = 1$, $F_2(x, ax^{2^k+1}) = (1-a)x^4$ and $F_3(x, ax^{2^k+1}) = (1-a)x^5$ are permutations unless $n \neq 2, 4$. Also each of $F_4(x, ax^3) = (1-a)x^7$ and $F_5(x, a^{1/3}x^3) = (1-a)x^9$ has n linearly independent components if $\gcd(n, 3) = 1$ and $n \neq 2, 4$ respectively. Thus F_4 and F_5 have by Lemma 1.

We show that all components produced by the above equations are linearly independent by the matrix argument similar to the inverse exponents case. At first, assume that $k = 1$. Each row corresponds to the equations from F_1, F_2, F_3, F_4 and F_5.

$$\begin{pmatrix} M_1 & M_2 & M_3 & 0 & 0 \\ M_4 & 0 & 0 & 0 & M_5 \\ M_6 & 0 & M_7 & 0 & M_8 \\ 0 & 0 & 0 & 0 & M_9 \\ M_{10} & M_{11} & M_{12} & M_{13} & 0 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ x_i x_k \\ y_i y_k \\ x_i y_j \end{pmatrix} = 0.$$

Each of M_2, M_4, M_7, M_9, and M_{13} represents $-y$, x^4, x^5, $x^4y - xy^2$, and y^3, respectively. Since all of them has n linearly independent component functions, each of the matrices has rank n. Further, if we consider the coefficient matrix by a 5×5 block matrix, we can easily convert it to a upper triangular matrix with diagonal M_2, M_4, M_7, M_9, and M_{13} by elementary row operations. Thus it has rank $5n$ and all components of the equations are linearly independent.

Next, assume that $k > 1$. Each row corresponds to the equations from F_1, F_2, and F_3.

$$\begin{pmatrix} M_1 & M_2 & M_3 & 0 \\ M_4 & 0 & M_5 & M_6 \\ M_7 & 0 & M_8 & M_9 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ x_i x_k \\ x_i y_j \end{pmatrix} = 0.$$

Since M_2 represents $-y$, it is invertible. Thus we are enough to show that all components of F_2 and F_3 are linearly independent. Let $F(x, y) = (F_2(x, y), F_3(x, y))$. We have $F(x, ax^{2^k+1}) = ((1-a)x^{2^k+2}, (1-a)x^{2^{k+1}+1})$. By Corollary 1 we can see $(x^{2^{k-1}+1}, x^{2^{k+1}+1})$ and $(x^{2^{n-k+1}+1}, x^{2^{n-k-1}+1})$ are nonlinear if $k < n/2 - 1$ and $k > n/2 + 1$, respectively. Note that both of them are affine transformations of $F(x, ax^{2^k+1})$. Thus unless $|k - n/2| \leq 1$, $F(x, y)$ has $2n$ linearly independent component functions.

Theorem 2. *Consider* $y = x^{2^k+1}$ *with* $\gcd(k, n) = 1$ *in* $\mathbb{F}_{2^n}$. *Let* t *be the number of monomials and* r *the number of linearly independent equations. Then we can have the following parameters* (r, t, Γ)*:*
(1) *If* $k = 1$*, we can obtain* 5 *linearly independent polynomials. Thus we get the followings:*

1. $\left(n, \frac{n(n+3)}{2}, \left(\frac{n+1}{2}\right)^{\lceil \frac{n+1}{2} \rceil}\right)$ *for* F_1
2. $\left(3n, \frac{n(3n+1)}{2}, \left(\frac{3n-5}{2}\right)^{\lceil \frac{3n-5}{2} \rceil}\right)$ *for* F_2*,* F_3*, and* F_4 *if* $n \neq 2, 4$ *and* $\gcd(n, 3) = 1$
3. $\left(4n, \frac{3n(n+1)}{2}, \left(\frac{3n-5}{2}\right)^{\lceil \frac{3n-5}{2} \rceil}\right)$ *for* F_1*,* F_2*,* F_3*, and* F_4 *if* $n \neq 2, 4$ *and* $\gcd(n, 3) = 1$
4. $\left(5n, n(2n+1), (2n-4)^{\lceil 2n-4 \rceil}\right)$ *for all polynomials if* $n \neq 2, 4$ *and* $\gcd(n, 3) = 1$.

(2) *Otherwise, we can obtain* 3 *linearly independent polynomials. Thus we get the followings:*

1. $\left(n, \frac{n(n+3)}{2}, \left(\frac{n+1}{2}\right)^{\lceil \frac{n+1}{2} \rceil}\right)$ *for* F_1
2. $\left(3n, \frac{3n(n+1)}{2}, \left(\frac{3n-3}{2}\right)^{\lceil \frac{3n-3}{2} \rceil}\right)$ *for* F_1*,* F_2*, and* F_3 *if* $|k - n/2| \leq 1$.

4.3 Kasami Exponents

When $\gcd(n, k) = 1$ and $k > 1$, $2^{2k} - 2^k + 1$ is called a Kasami exponent [8]. A Kasami exponent has the Hamming weight $k + 1$, but by applying composition by $2^k + 1$, we obtain a quadratic equation $F_1 : y^{2^k+1} - x^{2^{3k}+1}$.

By multiplying $x^{d_1}y^{d_2}$ to F_1, we have $x^{d_1+1+2^{3k}}y^{d_2} - x^{d_1}y^{d_2+1+2^k}$. Hence all x^{d_1}, y^{d_2}, $x^{d_1+1+2^{3k}}$, and $y^{d_2+1+2^k}$ should be linear monomials. It contradicts Lemma 5. Thus F_1 is the only quadratic equation we can obtain. F_1 has monomials of the type x_ix_j and y_iy_j. The number of monomials is n^2+n.

Theorem 3. [1] *Consider* $y = x^{2^{2k}-2^k+1}$ *with* $\gcd(k,n) = 1$ *in* $\mathbb{F}_{2^n}$. *We can obtain* n *linearly independent equations in* n^2+n *variables. Then RAA is* $\Gamma = n^n$.

4.4 Comparison

Exponent	Alg. Deg.	# of Eqns	# of Monomials	RAA Γ	When $n=8$
Inverse	$n-1$	$3n$	n^2+2n+1	$\left(\frac{n^2-n+1}{n}\right)^{\lceil \frac{n^2-n+1}{n} \rceil}$	$\Gamma = 2^{22.7}$
		$5n$	$2n^2+n+1$	$\left(\frac{2n^2-4n+1}{n}\right)^{\lceil \frac{2n^2-4+1}{n} \rceil}$	$\Gamma = 2^{46.8}$
Gold $(k=1)$	2	n	$\frac{n(n+3)}{2}$	$\left(\frac{n+1}{2}\right)^{\lceil \frac{n+1}{2} \rceil}$	$\Gamma = 2^{10.8}$
		$3n$	$\frac{n(3n+1)}{2}$	$\left(\frac{3n-5}{2}\right)^{\lceil \frac{3n-5}{2} \rceil}$	$\Gamma = 2^{32.5}$
		$4n$	$\frac{3n(n+1)}{2}$	$\left(\frac{3n-5}{2}\right)^{\lceil \frac{3n-5}{2} \rceil}$	$\Gamma = 2^{32.5}$
		$5n$	$n(2n+1)$	$(2n-4)^{\lceil 2n-4 \rceil}$	$\Gamma = 2^{43.0}$
Gold $(k>1)$ $\lvert k-n/2 \rvert > 1$	2	n	$\frac{n(n+3)}{2}$	$\left(\frac{n+1}{2}\right)^{\lceil \frac{n+1}{2} \rceil}$	$\Gamma = 2^{10.8}$
		$3n$	$\frac{3n(n+1)}{2}$	$\left(\frac{3n-3}{2}\right)^{\lceil \frac{3n-3}{2} \rceil}$	$\Gamma = 2^{37.3}$
Kasami	$k+1$	n	n^2+n	n^n	$\Gamma = 2^{24}$

Table 2. Comparison of RAA for Almost Perfect Nonlinear Functions.

Table 1 shows the comparison of the resistance of algebraic attacks. Surprisingly, more equations give larger RAA in each exponent. It is because RAA increases as $t-r$ increases and additional equations requires new variables more than new equations. From the table, we can see that the power functions with Kasami exponents have slightly better resistance against algebraic attacks, and the power functions with Gold exponents have very weak resistant against algebraic attacks.

5 Conclusion

In this paper, we developed several tools to prove linear independence of multivariate equations from algebraic S-boxes. By applying these tools to APN power

[1] By substituting $x = z^{2^k+1}$, we can obtain two independent quadratic equations $x = z^{2^k+1}$ and $y = z^{2^{3k}+1}$ with $n(n+5)/2$ variables, which reduces its RAA significantly. It will be introduced in the full version of this paper [3].

functions, we learned that a power function with a Gold exponent is very weak against algebraic attacks and a power function with a Kasami exponent has slightly stronger resistance against algebraic attacks. An open problem is to find S-boxes with $\Gamma > 2^{32}$ as indicated in [4]. Also, it is an interesting topic to apply algebraic attacks to block ciphers using a power function with a Gold exponent such as MISTY which is selected as standard block algorithms in NESSIE [12].

Acknowledgement

We are thankful to Hyun Soo Nam and Dae Sung Kwon for helpful discussions.

References

1. E. Biham and A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems," *Journal of Cryptology*, vol. 4, pp. 3–72, 1991.
2. E. Biham and A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.
3. J. Cheon and D. Lee, "Almost Perfect Nonlinear Power Functions and Algebraic Attacks," Manuscript, 2004
4. N. Courtois and J. Pieprzyk, "Cryptanalysis of Block Ciphers with Overdefined Systems of Equations," *Proc. of Asiacrypt 2002*, LNCS 2501, Springer-Verlag, pp. 267–287, 2002
5. J. Cheon, S. Chee and C. Park, "S-boxes with Controllable Nonlinearity," *Advances in Cryptology - Eurocrypt'99*, Springer-Verlag, pp. 286–294, 1999.
6. H. Dobbertin, "Almost Perfect Nonlinear Power Functions on $GF(2^n)$: The Welch Case," *IEEE Trans. Inform. Theory*, Vol. 45, No. 4, pp. 1271-1275, 1999.
7. R. Gold, "Maximal Recursive Sequences with 3-valued Recursive Cross-correlation Functions," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 154–156, 1968.
8. T. Kasami, "The Weight Enumerators for Several Classes of Subcodes of the Second Order Binary Reed-Muller Codes," *Infor. Contr.*, Vol. 18, pp. 369–394, 1971.
9. M. Matsui, "Linear Cryptanalysis Method for DES cipher," *Advances in Cryptology - Eurocrypt'93*, Springer-Verlag, pp. 386–397, 1993.
10. M. Matsui, "New Block Encryption Algorithm MISTY," *Proc. of FSE'97*, LNCS 1267, Springer-Verlag, pp. 54–68, 1997.
11. Advances Encryption Standards. `http://csrc.nist.gov/CryptoToolkit/aes/`
12. New European Schemes for Signatures, Integrity, and Encryption. `https://www.cosic.esat.kuleuven.ac.be/nessie/`