Few Properties of Coefficients in Bi-affine Equations for S-Boxes Based on Power Mappings

Technical Report No. ASU/2014/2

Dated : 14th February, 2014

Kishan Chand Gupta

Applied Statistics Unit

Indian Statistical Institute

203, B. T. Road, Kolkata 700108, INDIA.

kishan@isical.ac.in

Indranil Ghosh Ray

Applied Statistics Unit

Indian Statistical Institute

203, B. T. Road, Kolkata 700108, INDIA.

indranil_r@isical.ac.in

Few Properties of Coefficients in Bi-affine Equations for S-Boxes Based on Power Mappings

Kishan Chand Gupta and Indranil Ghosh Ray

Applied Statistics Unit, Indian Statistical Institute.
203, B. T. Road, Kolkata 700108, INDIA.
kishan@isical.ac.in, indranil_r@isical.ac.in

Abstract. S-boxes having large number of linearly independent multivariate bi-affine or quadratic equations may be susceptible to certain kinds of algebraic attack. In a 2011 IEEE-IT paper Kishan Chand Gupta et. al. provided a polynomial time algorithm to compute the maximal set of bi-affine and quadratic equations for S-boxes based on power mappings. In this paper, we study some structures and properties of coefficients of bi-affine equations.

Key words: Bi-affine equations, Quadratic equations, Power mapping, S-box, Algebraic attacks.

1 Introduction

Power mappings are of interest because unlike random permutations, they can be implemented in hardware without a lookup table. This facilitates compact and fast implementations of S-boxes in hardware. AES [9] uses S-boxes which are based on power mapping $x \mapsto x^{-1}$ over $\mathbb{F}_{2^8}$. Algebraic attacks are cryptanalytic attacks that reconstruct the secret key by solving the underlying equations. The idea behind the algebraic attacks is to express the cipher as a system of multivariate equations whose solution gives the secret key. The complexity of the attack depends on the number of such equations, their type (sparseness) and their algebraic degree. The first algebraic attack on a block cipher was discussed in [18]. For other developments in the area of algebraic attacks on block ciphers see [1, 2, 6, 7, 14, 15]. In [7], Courtois and Pieprzyk showed that AES [9] can be vulnerable by solving an overdefined system of algebraic equations. The authors presented an algorithm called Extended Sparse Linearization (XSL) to solve this multivariate equations.

Tools to calculate the number of linearly independent multivariate equations for S-boxes have been discussed in [3, 4, 7, 8, 16, 17]. The authors of [11] provided algorithms to compute the maximal set of actual bi-affine and quadratic equations in polynomial time. We study certain properties and patterns of such equations. We would like to mention here that tools developed in this paper are similar to [17] and many results of [17] are restated for easy reading.

In Section 2 we provide definitions and preliminaries. In Section 3 and Subsections therein, we study properties of coefficients of such equations. We conclude the paper in Section 4.

2 Definitions and Preliminaries

Let $\mathbb{F}_2 = \{0, 1\}$ be the finite field of two elements and $\mathbb{F}_{2^n}$ be the finite field of 2^n elements. We consider the domain of an n-variable Boolean function to be the vector space $(\mathbb{F}_2^n, +)$ over $\mathbb{F}_2$, where $+$ is used to denote the addition operator over both $\mathbb{F}_2$ and the vector space $\mathbb{F}_2^n = \{x_1, x_2, \cdots, x_n | x_i \in \mathbb{F}_2 = \{0, 1\}\}$ and n is a positive integer. Note, $x^{2^n} = x$ for all $x \in \mathbb{F}_{2^n}$ and $x_i x_i = x_i$ for all $x_i \in \mathbb{F}_2$. We will often denote a matrix by $((a_{i,j}))$, where $a_{i,j}$ is the (i, j)-th element of the matrix.

Any n variable Boolean function g: $\mathbb{F}_2^n \to \mathbb{F}_2$, can be uniquely represented as a multivariate polynomial over $\mathbb{F}_2$, called the *algebraic normal form*,

$$g(x_1, \ldots, x_n) = a_0 + \sum_{1 \le i \le n} a_i x_i + \sum_{1 \le i < j \le n} a_{i,j} x_i x_j + \ldots + a_{1,2,\ldots,n} x_1 x_2 \ldots x_n,$$

where the coefficients $a_0, a_i, a_{i,j}, \ldots, a_{1,2,\ldots,n} \in \mathbb{F}_2$. The degree of the Boolean function g, denoted by $\deg(g)$, is the same as the degree of the multivariate polynomial.

An (n, m) S-box (or vectorial function) is a map F: $\mathbb{F}_2^n \to \mathbb{F}_2^m$ and has component functions $f_1, \cdots, f_m$. We define the degree of an (n, m) S-box F to be the minimum of the degrees of all non zero linear combinations of its component functions.

An n variable affine function l is of the form $l(x_1, \ldots, x_n) = a_0 + \sum_{1 \le i \le n} a_i x_i$ where the coefficients $a_0, a_i \in \mathbb{F}_2$. If $a_0 = 0$, the function is called linear.

The Hamming weight of an integer i is the number of nonzero coefficients in the binary representation of i and is denoted by $\mathrm{H}(i)$. For example $\mathrm{H}(5) = 2$, $\mathrm{H}(8) = 1$.

$\mathbb{F}_{2^n}$ and $\mathbb{F}_2^n$ are isomorphic when both of them are regarded as vector space. The isomorphism is given by $x = (x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) \mapsto \{x_1, x_2 \cdots, x_n\}$, where $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis of $\mathbb{F}_{2^n}$. So a map F: $\mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is a (n, n) S-box.

Finite fields are polynomially complete i.e. any function F: $\mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ can be represented as a polynomial of degree $\leq 2^n - 1$ over $\mathbb{F}_{2^n}$. So any (n, n) S-box F can be represented as

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

where $a_i \in \mathbb{F}_{2^n}$ [12]. A function F: $\mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ such that $F(x) = x^a$, where $a \in \mathbb{Z}_{2^n-1}$ is called power function.

A cyclotomic coset C_s modulo $(2^n - 1)$ is defined as [13, page 104]

$$C_s = \{s, s \cdot 2, \cdots, s \cdot 2^{n_s-1}\}$$

where n_s is the smallest positive integer such that $s \equiv s2^{n_s} (\text{mod } 2^n - 1)$. The subscript s is chosen as the smallest integer in C_s, and s is called the coset leader of C_s. Note that n_s is the size of the coset C_s which will also be denoted by $|C_s|$. The set of all coset leaders modulo $(2^n - 1)$ is denoted by $\Upsilon(n)$. The computations in cosets are performed in $\mathbb{Z}_{2^n-1}$, the ring of integers modulo $(2^n - 1)$. For $n = 4$ the cyclotomic cosets modulo $2^4 - 1 = 15$ are: $C_0 = \{0\}, C_1 = \{1, 2, 4, 8\}, C_3 = \{3, 6, 12, 9\}, C_5 = \{5, 10\}, C_7 = \{7, 14, 13, 11\}$. Note $|C_5| = 2$, $|C_1| = 4$, and $\Upsilon(4) = \{0, 1, 3, 5, 7\}$.

A trace function Tr_m^n: $\mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$, is given by [12, page 51]

$$Tr_m^n(x) = \sum_{i=0}^{\frac{n}{m}-1} x^{2^{mi}} = x + x^{2^m} + x^{2^{2m}} + \ldots + x^{2^{n-m}}, \ x \in \mathbb{F}_{2^n}.$$

For $m = 1$, $Tr_1^n(x) = \sum_{i=0}^{n} x^{2^i} = x + x^2 + x^{2^2} + \ldots + x^{2^{n-1}}$. It is easy to check that for all $x, y \in \mathbb{F}_{2^n}$, $Tr_m^n(x+y) = Tr_m^n(x) + Tr_m^n(y)$ and $Tr_m^n(cx) = cTr_m^n(x)$, where $c \in \mathbb{F}_{2^m}$ and $x \in \mathbb{F}_{2^n}$. So the trace function Tr_m^n is a linear transformation from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$, when both $\mathbb{F}_{2^n}$ and $\mathbb{F}_{2^m}$ are viewed as vector spaces. Also note that $Tr_m^n(x) = Tr_m^n\left(x^{2^{ml}}\right)$ and $Tr_1^n(x) = Tr_1^n\left(x^{2^l}\right)$ for any non-negative integer l.

Any n variable Boolean function f: $\mathbb{F}_{2^n} \to \mathbb{F}_2$, can be uniquely represented as a sum of trace functions [10, page 178]

$$f(x) = \sum_{k \in \Upsilon(n)} Tr_1^{n_k}(A_k x^k) + A_{2^n-1} x^{2^n-1},\ A_k \in \mathbb{F}_{2^{n_k}},\ A_{2^n-1} \in \mathbb{F}_2,$$

where $\Upsilon(n)$ is the set consisting of all coset leaders modulo $2^n - 1$, n_k is the size of the coset C_k, and $Tr_1^{n_k}(x)$ is the trace function from $\mathbb{F}_{2^{n_k}} \to \mathbb{F}_2$. If $f(x)$ is balanced, we have [10]

$$f(x) = \sum_{k \in \Upsilon(n)} Tr_1^{n_k}(A_k x^k),\ A_k \in \mathbb{F}_{2^{n_k}},\ x \in \mathbb{F}_{2^n}. \tag{2}$$

The algebraic degree of f, denoted by $\deg(f)$, is given by the largest w such that $A_k \neq 0$ and $H(\mathrm{k})$=w. It is easy to check the following facts,

Fact: 1 *For $|C_k| = n_k$, $Tr_1^{n_k}(A_k x^k) = 0$ if and only if $A_k = 0$. If $A_k \neq 0$ then $Tr_1^{n_k}(A_k x^k)$ gives a Boolean function of degree H(k).*

Fact: 2 *If $|C_k| = m < n$, then $Tr_1^n(A_k x^{2^t k}) = Tr_1^m(A'_k x^{2^t k}) = Tr_1^m({A'_k}^{2^{m-t}} x^k)$, where $A'_k = Tr_m^n(A_k)$. There are 2^{n-m} elements of $\mathbb{F}_{2^n}$ for which Tr_m^n maps to 0 [13, page 116].*

Consider any arbitrary basis $\{\alpha_1, \ldots, \alpha_n\}$ of $\mathbb{F}_{2^n}$. A basis $\{\beta_1, \ldots, \beta_n\}$ is called dual basis of $\{\alpha_1, \ldots, \alpha_n\}$ if

$$Tr_1^n(\alpha_i \beta_j) = \begin{cases} 0 \text{ for } i \neq j \\ 1 \text{ for } i = j. \end{cases}$$

Fact: 3 *[13, page 118] Every basis of $\mathbb{F}_{2^n}$ has a dual basis.*

Let $\{\alpha_1, \ldots, \alpha_n\}$ be an arbitrary basis of $\mathbb{F}_{2^n}$. Then the inverse of the matrix

$$\begin{pmatrix} \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{2^{n-1}} \\ \alpha_2 & \alpha_2^2 & \ldots & \alpha_2^{2^{n-1}} \\ \ldots\ldots\ldots & & \ldots & \\ \alpha_n & \alpha_n^2 & \ldots & \alpha_n^{2^{n-1}} \end{pmatrix}$$

always exists and is of the form [13, page 117]

$$\begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_n \\ \beta_1^2 & \beta_2^2 & \dots & \beta_n^2 \\ \dots & \dots & \dots & \dots \\ \beta_1^{2^{n-1}} & \beta_2^{2^{n-1}} & \dots & \beta_n^{2^{n-1}} \end{pmatrix}.$$

The first row $\{\beta_1, \beta_2, \dots, \beta_n\}$ will be the dual basis of $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$.

Let $\{\alpha_1, ..., \alpha_n\}$ and $\{\beta_1, \dots, \beta_n\}$ be the dual basis of $\mathbb{F}_{2^n}$ and $x = x_1\alpha_1 + \dots + x_n\alpha_n$. Then $Tr_1^n(\beta_i x) = Tr_1^n(\beta_i(x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n)) = Tr_1^n(\sum_{k=1}^n x_k\beta_i\alpha_k) = \sum_{k=1}^n Tr_1^n(x_k\beta_i\alpha_k) = \sum_{k=1}^n x_k Tr_1^n(\beta_i\alpha_k) = x_i$.

We record this important result in the following fact.

Fact: 4 *Let $\{\alpha_1, ..., \alpha_n\}$ be any arbitrary basis of $\mathbb{F}_{2^n}$ and $\{\beta_1, \dots, \beta_n\}$ be it's dual basis and $x = x_1\alpha_1 + \dots + x_n\alpha_n$, then $x_i = Tr_1^n(\beta_i x)$.*

Let's fix any arbitrary basis $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ for $\mathbb{F}_{2^n}$. Then $\mathbb{F}_{2^n}$ and $\mathbb{F}_2^n$ are isomorphic and can be used interchangeably. Consider $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ to be an S-box based on a power mapping. Such S-boxes are classified according to the exponent a of the power mapping such that $y = F(x) = x^a$. *Bi-affine* equations are of the form

$$\sum_{i,j} a_{i,j}x_i y_j + \sum_j v_j y_j + \sum_i u_i x_i = 0, \quad a_{i,j}, v_j, u_i \in \mathbb{F}_2. \tag{3}$$

where $x = x_1\alpha_1 + \dots + x_n\alpha_n$ and $y = y_1\alpha_1 + \dots + y_n\alpha_n$.

Now we provide following Propositions which is mainly from [17]. For the sake of clarity, we also provide the proof.

Proposition 1. *[17, Proposition 1] Let $x = \sum_{i=1}^n x_i\alpha_i$, $y = \sum_{i=1}^n y_i\alpha_i$ and $a_{i,j} \in \mathbb{F}_2$. Then*

$$\sum_{i,j} a_{i,j}x_i y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} y), \text{ where } b_k = \sum_{i,j} a_{i,j}\beta_i^{2^k}\beta_j \in \mathbb{F}_{2^n}.$$

Proof. Let $\{\alpha_1, ..., \alpha_n\}$ and $\{\beta_1, \dots, \beta_n\}$ be the dual basis of $\mathbb{F}_{2^n}$ and $x = x_1\alpha_1 + \dots + x_n\alpha_n$ and $y = y_1\alpha_1 + \dots + y_n\alpha_n$. Now, from Fact 4, $x_i = Tr_1^n(\beta_i x)$ and $y_j = Tr_1^n(\beta_j y)$. Therefore $\sum_{i,j} a_{i,j}x_i y_j = \sum_{i,j} a_{i,j} Tr_1^n(\beta_i x) Tr_1^n(\beta_j y) = \sum_{i,j} a_{i,j} \sum_{k=0}^{n-1} Tr_1^n(\beta_i^{2^k}\beta_j x^{2^k} y)$
$= \sum_{k=0}^{n-1} Tr_1^n(\sum_{i,j} a_{i,j}\beta_i^{2^k}\beta_j x^{2^k} y) = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} y)$, where $b_k = \sum_{i,j} a_{i,j}\beta_i^{2^k}\beta_j$. □

Corollary 1. *[17, Corollary 1] Let $y = x^a$ be a power mapping. Then*

$$\sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(cx^a) \tag{4}$$

where

$$b_k = \sum_{i,j} a_{i,j} \beta_i^{2^k} \beta_j \in \mathbb{F}_{2^n}, \text{ and } c = \sum_j v_j \beta_j \in \mathbb{F}_{2^n}.$$

Proof. From Proposition 1, we have $\sum_{i,j} a_{i,j} x_i y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} y)$. We know, from Fact 1 that $y_j = Tr_1^n(\beta_j y)$. Thus we have $\sum_j v_j y_j = \sum_j v_j Tr_1^n(\beta_j y) = Tr_1^n(\sum_j v_j \beta_j y) = Tr_1^n(cy)$. where $c = \sum_j v_j \beta_j$. Putting $y = x^a$ we have $\sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} x^a) + Tr_1^n(cx^a) = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(cx^a)$. □

From Fact 2 and basic properties of trace function, we have the following very useful fact,

Fact: 5 *[17, Corollary 1]. Let $T_1 = \{a\} \cup \{2^k + a | k = 0, 1, \ldots, n-1\}$ and $S_1 = \{s_j | 0 \le j < K\}$ be the set consisting of different coset leaders modulo $2^n - 1$ of the elements in T_1. For $t \in S_1$, assume that there are v elements in T_1, say $J_t = \{i_1, \ldots, i_v\}$ which belong to C_t. Let $l = |C_t|$. Then*

$$\sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} x^a) + Tr_1^n(cx^a) = \sum_{t \in S_1} Tr_1^l(b'_t x^t) \tag{5}$$

where $b'_t = \sum_{i \in J_t} [\, Tr_l^n(b_i)]^{2^{l-u_i}}$. Here u_i is determined by $i \equiv t2^{u_i} \pmod{2^n - 1}$. Furthermore, the representation of Equation (5) is unique.

3 Properties of coefficients in Bi-affine equations

For a given power mapping $y = x^a$, the linearly independent bi-affine equations of the form $\sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j + \sum_i u_i x_i = 0, \quad a_{i,j}, v_j, u_i \in \mathbb{F}_2$. In characteristic 2, this is equivalent to $\sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j = \sum_i u_i x_i$, where $a_{i,j}, v_j, u_i \in \mathbb{F}_2$ From Corollary 1, we have $\sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(cx^a)$. Bi-affine equations exist if and only if

$$\sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(cx^a) = \sum_i u_i x_i \tag{6}$$

Therefore the number of bi-affine equations, of the form as given in Equation (3), is equal to the number of functions $\sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(cx^a)$ that are affine (i.e.,$\sum_i u_i x_i$). There will be three cases, where this will happen. Note that $\sum_i u_i x_i$ is a Boolean function which is either 0 or of degree 1.

Recall from Proposition 1 that for $k = 0, \ldots, n-1$, $b_k = \sum_{i,j} a_{i,j} \beta_i^{2^k} \beta_j$, which can be written as

$$\begin{pmatrix} \beta_1 & \beta_2 & \ldots & \beta_n \\ \beta_1^2 & \beta_2^2 & \ldots & \beta_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \beta_1^{2^{n-1}} & \beta_2^{2^{n-1}} & \ldots & \beta_n^{2^{n-1}} \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,n} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{pmatrix} \tag{7}$$

The Equation (7) is very important and we will derive various properties of $((a_{i,j}))$ using this equation. Hamming weight of $((a_{i,j}))$ matrix is number of nonzero entries in the matrix and will be denoted by $H|((a_{i,j}))|$

Let α be a primitive element of $\mathbb{F}_{2^n}$. The polynomial basis is $\{1, \alpha, \alpha^2, \ldots, \alpha^n\}$. Let us consider the (n, n) S-box based on the power mapping $y = x^a$ where bi-affine equations exist. Also assume that $H(2^k + a) = 1$ for some $k \in \{0, 1, \ldots, n-1\}$. We solve Equation (7) for $((a_{i,j}))$ and u_i's by using Algorithm 1 of [11] to get n linearly independent bi-affine equations, where we let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ to be the polynomial basis and b_k takes n values corresponding to each of these equations from polynomial basis. It is easy to check that $\alpha_i = \alpha^{i-1}$. We have the following properties of such $((a_{i,j}))$ matrices.

Lemma 1. *$((a_{i,j}))$ is a full rank matrix.*

Proof. It is easy to check that values at the right hand side of Equation (8) are linearly independent. Thus the rows of left hand side matrix i.e. $((a_{i,j}))$ will be linearly independent. Thus $((a_{i,j}))$ is a full rank matrix.

Lemma 2. *$n \leq H|((a_{i,j}))|$.*

Proof. That $H|((a_{i,j}))|$ attains the minimum weight of n is given by the fact that when $n = 4$, and primitive polynomial taken is $x^4 + x + 1$, one of the outputs of the Algorithm 1 of [11] contains $((a_{i,j}))$ having weight $n = 4$.

We are to show that $H|((a_{i,j}))|$ can not be reduced further. The Equation (7) can be written as

$$\begin{pmatrix} a_{1,1}\ a_{1,2}\ \dots\ a_{1,n} \\ a_{2,1}\ a_{2,2}\ \dots\ a_{2,n} \\ \dots\ \dots\ \dots\ \dots \\ \dots\ \dots\ \dots\ \dots \\ a_{n,1}\ a_{n,2}\ \dots\ a_{n,n} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1\ \alpha_1^2\ \dots\ \alpha_1^{2^{n-1}} \\ \alpha_2\ \alpha_2^2\ \dots\ \alpha_2^{2^{n-1}} \\ \dots\dots\dots\ \dots \\ \dots\dots\dots\ \dots \\ \alpha_n\ \alpha_n^2\ \dots\ \alpha_n^{2^{n-1}} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ b_k \\ \vdots \\ 0 \end{pmatrix} = b_k \begin{pmatrix} \alpha_1^{2^k} \\ \vdots \\ \alpha_l^{2^k} \\ \vdots \\ \alpha_n^{2^k} \end{pmatrix} \tag{8}$$

In this Equation (8), b_k is substituted with nonzero independent values for finding independent bi-affine equations. Since each entry of the right hand side matrix is nonzero, therefore no row of $((a_{i,j}))$ should be all-zero. Thus each row contains at least one entry as 1.

Remark 1. It might happen that dual basis of polynomial basis is a permutation, i.e. $\beta_i = \alpha_j$ for some i and j, $i, j \in \{1, 2, \dots, n\}$. Let us denote this permutation by π. So $\beta_i = \alpha_{\pi(i)}$. Also let $H(2^0 + a) = 1$. When we take $b_0 = 1$, the first value of polynomial basis, we observe from Equation (8) that $((a_{i,j}))$ is nothing but the permutation matrix corresponding to π having hamming weight n.

Lemma 3. $H|((a_{i,j}))| \leq n^2 - n + 1$.

Proof. That $H|((a_{i,j}))|$ attains the maximum weight of $n^2 - n + 1$ is given by the fact that when $n = 4$, and primitive polynomial taken is $x^4 + x + 1$, one of the outputs of the Algorithm 1 of [11] contains $((a_{i,j}))$ having weight $4^2 - 4 + 1 = 13$.

We are to show that $H|((a_{i,j}))|$ can not exceed $n^2 - n + 1$. Since from Lemma 1, $((a_{i,j}))$ is full rank, it can not have all-one i.e. $H|((a_{i,j}))| < n^2$. Let us consider some $((a_{i,j}))$ on $\mathbb{F}_2$, having weight $n^2 - n + 2$. number of ones. So the number of zeros are $n - 2$. It is easy to check that number of rows having all-one is at least 2, making $((a_{i,j}))$ a singular matrix which is a contradiction.

Lemma 4. $a_{i,j} = Tr_1^n(b_k \alpha_i^{2^k} \alpha_j)$.

Proof. From Equation (8), we have $a_{i,1}\beta_1 + a_{i,2}\beta_2 + \ldots + a_{i,n}\beta_n = b_k\alpha^{2^k}$. Multiplying by α_j and taking trace both sides, we have

$$Tr_1^n(\sum_{l=1}^{n} a_{i,l}\alpha_j\beta_l) = Tr_1^n(b_k\alpha_i^{2^k}\alpha_j) \implies \sum_{l=1}^{n} a_{i,l}Tr_1^n(\alpha_j\beta_l) = Tr_1^n(b_k\alpha_i^{2^k}\alpha_j)$$
$$\implies a_{i,j} = Tr_1^n(b_k\alpha_i^{2^k}\alpha_j)$$

Remark 2. When b_k is taking values from polynomial basis one by one giving n independent bi-affine equations, then some column of $((a_{i,j}))$ of one equation repeats in some other column position of $((a_{i,j}))$ corresponding to some other biaffine equation. To check this, let us fix $j = t$ and let $b_k = \alpha_l$. So we are constructing lth equation. From Lemma 3, $a_{i,t} = Tr_1^n(\alpha_l\alpha_i^{2^k}\alpha_t) = Tr_1^n(\alpha^{l-1}\alpha^{(i-1)2^k}\alpha^{t-1})$. So we have $a_{i,t} = Tr_1^n(\alpha^{i.2^k+l+t-2-2^k})$. When we compute $l+1$th equation, we have $b_k = \alpha_{l+1}$ and let us fix $j = t-1$. From the previous analogy, we have $a_{i,t-1} = Tr_1^n(\alpha^{i.2^k+l+1+t-1-2-2^k}) = Tr_1^n(\alpha^{i.2^k+l+t-2-2^k})$. From this result it is evident that t th column of coefficient matrix of lth equation is same as $t-1$ th column of coefficient matrix of $l+1$th equation.

Lemma 5. *When $k = 0$ i.e. $H(2^0 + a) = 1$, $((a_{i,j}))$ has got a pattern which is as follows: for any integer $m \in \{2, 3, \ldots, 2n\}$, the values $a_{i,j}$'s, where $i + j = m$ will be same.*

Proof. when $k = 0$, Equation (8) can be written as, $\sum_{l=1}^{n} a_{i,l}\beta_l = b_k\alpha_i$, for $i = 1, \ldots, n$. Now by multiplying $\sum_{l=1}^{n} a_{i,l}\beta_l = b_k\alpha_i$ by α_j we get $\sum_{l=1}^{n} a_{i,l}\alpha_j\beta_l = b_k\alpha_i\alpha_j$. Applying trace function both side of this equation, we get

$$Tr_1^n(\sum_{l=1}^{n} a_{i,l}\alpha_j\beta_l) = Tr_1^n(b_k\alpha_i\alpha_j) \implies \sum_{l=1}^{n} a_{i,l}Tr_1^n(\alpha_j\beta_l) = Tr_1^n(b_k\alpha_i\alpha_j)$$
$$\implies a_{i,j} = Tr_1^n(b_k\alpha_i\alpha_j) = Tr_1^n(b_k\alpha^{i+j-2}) = Tr_1^n(b_k\alpha^{m-2})$$

Thus for all i, j's such that $i + j = m$, $a_{i,j}$'s will be same.

Lemma 6. *When $k = 0$ i.e. $H(2^0 + a) = 1$, $((a_{i,j}))^t = ((a_{i,j}))$.*

Proof. From Lemma 5 this is easy to check that $a_{i,j} = a_{j,i}$ for all $i, j \in \{1, \ldots, n\}$. Hence the matrix $((a_{i,j}))$ is symmetric.

4 Conclusion

In this paper we explored few properties of coefficients of bi-affine equations on $\mathbb{F}_2$ that are obtained from the output of Algorithm 1 of [11]. We observed certain relations between various $a_{i,j}$s. Another interesting aspect of our work goes towards constructing good input candidates to make the XSL algorithm succeed. We studied the bounds of weights of coefficient matrices. It is believed that less is the number of monomials more is the chance of XSL algorithm to succeed. We have identified under what setting, Algorithm 1 of [11] can produce such equations with at least one lowest weight equation if they at all exist.

References

1. F. Armknecht, On the Existence of Low-degree Equations for Algebraic Attacks, *Cryptology ePrint Archive, Report 2004/185*, http://eprint.iacr.org/, 2004.
2. A. Biryukov and C. D. Canniere, Block Ciphers and Systems of Quadratic Equations, *Fast Software Encryption 2003*, LNCS 2887, pp. 274-289, Springer-Verlag, 2003.
3. J. Cheon and D. Lee, Resistance of S-Boxes Against Algebraic Attacks, *Fast Software Encryption 2004*, LNCS 3017, pp. 83-94, Springer-Verlag, 2004.
4. J. Cheon and D.H. Lee, Quadratic Equations from APN power functions. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E89-A(1), 19-27(2006)*
5. H. Cohen, A Course in Computational Algebraic Number Theory. *S*pringer
6. N. Courtois, Algebraic Attacks on Combiners with Memory and Several Outputs, *ICISC 2004*, LNCS 3506, pp. 3-20, Springer-Verlag, 2004.
7. N. Courtois and Pieprzyk J., Cryptanalysis of Block Ciphers with Overdefined Systems of Equations, *Advances in Cryptology - Asiacrypt 2002*, LNCS 2501. Springer-Verlag, 2002.
8. N. Courtois, B. Debraize and E. Garrido, On Exact Algebraic [Non]Immunity of S-boxes Based on Power Functions, *Cryptology ePrint Archive, Report 2005/203*, http://eprint.iacr.org/, 2005.
9. J. Daemen, and V.Rijmen, *The Design of Rijndael*, Springer-Verlag, 2002.
10. S. W. Golomb, and G. Gong, *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*, Cambridge University Press, ISBN 0521821045, 2005.
11. K.C. Gupta and I. Ghosh Ray, Finding Bi-affine and Quadratic Equations for S-Boxes Based on Power Mappings, preprint.
12. R. Lidl and H. Niederreiter, *Introduction to Finite Fields and their Applications.* Cambridge University Press, 1994.
13. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes.* North Holland, 1986.
14. S. Murphy and M. Robshaw, Essential Algebraic Structure within AES, *Advances in Cryptology - Crypto 2002*, LNCS 2442, pp.1-16, Springer-Verlag, 2002.
15. S. Murphy and M. Robshaw, Comments on the Security of the AES and the XSL Technique, *Electronic Letters*, Vol. 39, pp. 26-38, 2003.
16. Y. Nawaz, K.C. Gupta and G. Gong, Algebraic Immunity of S-Boxes Based on Power Mappings:Analysis and Construction, http://eprint.iacr.org/2006/322
17. Y. Nawaz, K.C. Gupta and G. Gong, Algebraic Immunity of S-Boxes Based on Power Mappings:Analysis and Construction, *IEEE Transactions on Information Theory*, Vol. 55, No. 9, pp. 4263-4273, 2009.
18. I. Schaumuller-Bichl, Cryptanalysis of the Data Encryption Standard by the Method of Formal Coding, *Advances in Cryptology - Eurocrypt 1982*, LNCS 149, pp.235-255, Springer-Verlag, 1983.