Algebraic Immunity of S-boxes Based on Power Mappings: Analysis and Construction

Yassir Nawaz, Kishan Chand Gupta, and Guang Gong, *Senior Member, IEEE*

***Abstract*—The algebraic immunity of an S-box depends on the number and type of linearly independent multivariate equations it satisfies. In this paper techniques are developed to find the number of linearly independent, multivariate, bi-affine and quadratic equations for S-boxes based on power mappings. These techniques can be used to prove the exact number of equations for any class of power mappings. Two algorithms to calculate the number of bi-affine and quadratic equations for any (n, n) S-box based on power mapping are also presented. The time complexity of both algorithms is only $O(n^2)$. To design algebraically immune S-boxes four new classes of S-boxes that guarantee zero bi-affine equations and one class of S-boxes that guarantees zero quadratic equations are presented. The algebraic immunity of power mappings based on Kasami, Niho, Dobbertin, Gold, Welch and Inverse exponents are discussed along with other cryptographic properties and several cryptographically strong S-boxes are identified. It is conjectured that a known Kasami like highly nonlinear power mapping is differentially 4-uniform. Finally an open problem to find an (n, n) bijective nonlinear S-box with more than $5n$ quadratic equations is solved.**

***Index Terms*—Algebraic immunity, multivariate equations, Bi-affine equations, Quadratic equations, Power mapping, S-box.**

I. Introduction

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 and their algebraic degree. The first algebraic attack on a block cipher was discussed in [19]. For recent developments in the area of algebraic attacks on block ciphers see [1], [2], [5]–[7], [16], [17]. In [7] Courtois and Pieprzyk showed that AES [9] can be attacked by solving an overdefined system of algebraic equations. This is possible because the only nonlinear component in AES, i.e., the S-box, can be expressed as an overdefined system of algebraic equations. The authors presented an algorithm called XSL to solve this system of multivariate equations and also introduced a parameter, Γ, for S-boxes where Γ is claimed to be an important parameter in measuring the complexity of the XSL algorithm. For a $GF(2^n) \rightarrow GF(2^n)$ S-box Γ was defined as $\Gamma = ((t-r)/n)^{\lceil (t-r)/n \rceil}$, where r is the number of equations and t is the number of monomials in these equations. A lower value of r means a higher value of Γ and therefore higher complexity of the algebraic attack. In [7] the authors showed that for AES S-box $\Gamma = 2^{22.9}$ and claimed that for secure ciphers Γ should be greater than 2^{32}. Although Γ was defined to be the measure of algebraic immunity of an S-box in [7], the relation between the number of equations and the complexity of solving that system of equations seems to be more complex. However it is generally believed that S-boxes with large number of quadratic equations have low algebraic immunity.

Inspired by [7] Cheon and Lee [4] developed tools to calculate the number of linearly independent multivariate equations for algebraic S-boxes. They used their results to estimate the algebraic immunity Γ of highly nonlinear and almost bent power S-boxes (based on Gold, Kasami and inverse exponents [10]). However Courtois et al. disputed their results in [8] by showing that in most cases the number of linearly independent multivariate equations calculated by them is incorrect. This was done by experimentally finding the total number of quadratic equations for Gold and Kasami power S-boxes.

In [8, Appendix A] Courtois et al. also used the polynomial representation of the algebraic S-boxes to prove that S-boxes based on inverse mapping in $GF(2^n)$ have $3n - 1$ bi-affine equations and $5n - 1$ quadratic equations. However they did not generalize their results to other power S-boxes and only provided experimental results for S-boxes based on Gold, Dobbertin, Niho, Welch and Kasami exponents. The largest S-box experimentally tested by them was $n = 17$. In this paper we expand and generalize the technique used in [8]. We use the polynomial representations to develop new techniques which can be used to prove the exact number of bi-affine and quadratic equations for any class of power mappings. From these techniques we also develop two very simple algorithms which, given an (n, n) S-box and the exponent of power mapping, calculate the exact number of bi-affine and quadratic equations respectively. The time complexity of both algorithms is $O(n^2)$ which is very small even for very large S-boxes (for example $n = 2^{20}$). The algorithms currently available in literature to find such equations have time complexities that are exponential in n and therefore impractical for $n > 25$. Note however that these algorithms find the actual equations whereas our algorithms only calculate the number of such equations. Towards designing S-boxes with highest algebraic immunity, we provide four classes of power S-boxes that guarantee zero bi-affine equations and one class

Yassir Nawaz is with Advanced Concepts & Technologies, Pitney Bowes Inc., USA, e-mail: yassirnawaz@gmail.com. This work was done while he was a doctoral candidate at the Departmentof Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada.

Kishan Chand Gupta is with Applied Statistics Unit, Indian Statistical Institute, 203 B T Road, Kolkata 700108, INDIA, e-mail: kishan_t@isical.ac.in. This work was done when he was a post doctoral fellow at Center for Applied Cryptographic Research at Universityof Waterloo, Waterloo, Canada.

Guang Gong is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada, e-mail: G.Gong@ece.uwaterloo.ca.

Manuscript received April **, 2006; revised March **, 2009.

that guarantees zero quadratic equations. We examine other cryptographic properties, i.e., nonlinearity, algebraic degree and uniform differential property of these S-boxes and list cryptographically strong power S-boxes for $n = 8$ and 10. We also provide a new conjecture regarding the uniform differential property of Kasami like power mappings. In addition to this we solve an open problem given in [7], [8] to find a bijective nonlinear (n, n) S-box with more than $5n$ quadratic equations.

II. Definitions and Preliminaries

Let $\mathbb{F}_2 = \{0, 1\}$ be the finite field of two 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. 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 a binary string, λ consecutive ones (1's) preceded by zero and followed by zero is called a run of ones of length λ. We consider the runs of ones to be cyclic. For example 1100011110011111 has two (not three) cyclic runs of ones.

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

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

where the coefficients $a_0, a_i, a_{i,j}, \ldots, a_{1,2,\ldots,n} \in \mathbb{F}_2$. An (n, m) S-box (or vectorial function) is a map F: $\mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ and has component functions $f_1, \cdots, f_m$.

The degree of the Boolean function g, denoted by $deg(g)$, is the same as the degree of the multivariate polynomial. 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 Walsh transform (WT) of an m-variable Boolean function g is an integer valued function $W_g : \{0, 1\}^m \rightarrow [-2^m, 2^m]$ defined by (see [15, page 414])

$$W_g(u) = \sum_{w \in \mathbb{F}_2^m} (-1)^{g(w) + \langle u, w \rangle}, u \in \mathbb{F}_2^m. \tag{1}$$

$\{W_g(u) | u \in \mathbb{F}_2^m\}$ is called the *spectrum* of g. Note that Walsh transform of $g(x)$ is actually the Fourier transform of $(-1)^{g(x)}$.

Nonlinearity of a Boolean function g measures the distance of the Boolean function from the set of all affine functions. The nonlinearity $nl(g)$ of an n-variable Boolean function g, can be written as

$$nl(g) = 2^{n-1} - \frac{1}{2} \max_{u \in \mathbb{F}_2^n} |W_g(u)|.$$

Nonlinearity of an (n, m) S-box is defined as the minimum nonlinearity of all the non zero linear combinations of its component functions.

Let $\mathbb{F}_{2^n}$ be the finite field with 2^n elements. Consider a mapping F: $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$. If n is odd F is called *almost bent* [10] if the Walsh spectra of all nonzero linear combinations of its component functions have precisely 3 values $\{0, \pm 2^{\frac{n+1}{2}}\}$. Note that the nonlinearity is $2^{n-1} - 2^{(n-1)/2}$ in this case. If n is even then it is conjectured [10] that maximum achievable nonlinearity by F is $2^{n-1} - 2^{\frac{n}{2}}$. If F achieves this nonlinearity then it is called *highly nonlinear*.

A mapping F: $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$ is called differentially k-uniform [18] if each equation

$$F(x + a) - F(x) = b, \text{ where } a \in \mathbb{F}_{2^n}, a \neq 0, b \in \mathbb{F}_{2^n},$$

has at most k solutions in $\mathbb{F}_{2^n}$. If $k = 2$, F is called *almost perfect nonlinar* (APN) function [10]. It is known that if n is odd and F is almost bent then F is APN [3].

A trace function Tr: $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^m}$, is given by [14, page 51]

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

A cyclotomic coset C_s modulo $(2^n - 1)$ is defined as [15, 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 $s \equiv s 2^{n_s} (\text{mod } 2^n - 1)$. The subscript s is chosen as the smallest integer in C_s, and s is called the coset leader of C_s. The computations of cosets are performed in $Z_{2^n - 1}$, the residue integer ring modulo $(2^n - 1)$. For example the cyclotomic cosets modulo 15 are: $C_0 = \{0\}, C_1 = \{1, 2, 4, 8\}, C_3 = \{3, 6, 12, 9\}, C_5 = \{5, 10\}, C_7 = \{7, 14, 13, 11\}$, where {0,1,3,5,7} are coset leaders modulo 15.

Any non-zero polynomial function f: $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_2$, can be represented as a sum of trace functions [12, page 178]:

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

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 $\mathbb{F}_{2^{n_k}} \rightarrow \mathbb{F}_2$. If $f(x)$ is balanced , we have [12]

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

Fact 1. *Assume that $f(x) = \sum_{k \in \Upsilon(n)} Tr_1^{n_k}(A_k x^k)$, and $g(x) = \sum_{k \in \Upsilon(n)} Tr_1^{n_k}(B_k x^k)$. Then $f(x) = g(x)$ if and only in $A_k = B_k$ for any $k \in \Upsilon(n)$.*

The algebraic degree of f, denoted by $deg(f)$, is given by the largest w such that $A_k \neq 0$ and $H(k) = w$. There is a natural correspondence between Boolean functions h and polynomial functions f [12, page 334]. Let $\{\alpha_0, \ldots, \alpha_{n-1}\}$ be a basis for $\mathbb{F}_{2^n}$, then this correspondence is given by:

$$h(x_0, \ldots, x_{n-1}) = f(x), x = \sum_{i=0}^{n-1} x_i \alpha_i, x_i \in \mathbb{F}_2.$$

A monomial or single trace term function f is a function that can be represented by a single trace term, $f(x) = Tr_1^n(\beta x^t)$ where $\beta \in \mathbb{F}_{2^n}$ and t is the coset leader of C_t.

III. Bi-affine and Quadratic Equations for Power Mappings

Let's fix any arbitrary basis $\{\alpha_0, \alpha_1, \cdots, \alpha_{n-1}\}$ 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 $y = F(x) = x^a$.

Proposition 1: Let $h(x, y) = \sum_{i,j} a_{i,j} x_i y_j$ be a bilinear Boolean function in $2n$ variables, where $x = \sum_{i=0}^{n-1} x_i \alpha_i$, $y = \sum_{i=0}^{n-1} y_i \alpha_i$ and $a_{i,j} \in \mathbb{F}_2$. Then $h(x, y)$ has a unique polynomial representation in $\mathbb{F}_{2^n}$ such that

$$h(x, y) = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} y), \; b_k \in \mathbb{F}_{2^n}.$$

Proof: Proof is given in the Appendix. ■

Lemma 1: (i) If $|C_r| = n$, $Tr_1^n(cx^{2^t r}) = Tr_1^n(c^{2^{n-t}} x^r)$. (ii) If $|C_r| = l < n$, then $Tr_1^n(cx^{2^t r}) = Tr_1^l(Tr_l^n(c)^{2^{l-t}} x^r)$ where $Tr_l^n(x)$ is the trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^l}$ and t is reduced by modulo l.

Corollary 1: Let $y = x^a$ and $h(x, y) = \sum_{i,j} a_{i,j} x_i y_j + \sum_j v_j y_j$ be a Boolean function in n variables where $a_{i,j}, b_j \in \mathbb{F}_2$. Then

(i) $h(x, y)$ has a polynomial representation in $\mathbb{F}_{2^n}$ such that

$$h(x, y) = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k + a}) + Tr_1^n(cx^a), \; b_k, c \in \mathbb{F}_{2^n}.$$

(ii) Let $T_1 = \{a, 2^k + a \,|\, k = 0, 1, \cdots, n-1\}$ and $S_1 = \{s_j \,|\, 0 \leq j < K\}$ 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 $J_t = \{i_1, \cdots, i_v\}$ which belong to C_t. Let $l = |C_t|$. Then

$$h(x, y) = \sum_{t \in S_1} Tr_1^l(b'_t x^t) \tag{3}$$

where

$$b'_t = \sum_{i \in J_t} [Tr_l^n(b_i)]^{2^{l-u_i}} \tag{4}$$

where u_i is determined by $i \equiv t2^{u_i} \pmod{2^n - 1}$. Furthermore, the representation in (3) is unique.

Proof: From Proposition 1 we have $\sum_{i,j} a_{i,j} x_i y_j = \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} y)$ and we know $\sum_j v_j y_j = Tr_1^n(\sum_j v_j \beta_j y) = Tr_1^n(cy)$. Putting $y = x^a$ we have the proof of the assertion (i). For the second assertion, the representation of (3) follows directly from the two results of Lemma 1. According to Fact 1, (3) is unique. ■

Proposition 2: Let $h(y) = \sum_{i \leq j} a_{i,j} y_i y_j$ be a Boolean function in n variables, where $a_{i,j} \in \mathbb{F}_2$. Then $h(y)$ has a polynomial representation in $\mathbb{F}_{2^n}$ such that:

- if n is even, $n = 2m$:

$$\begin{aligned} h(y) \;&=\; Tr_1^n(d_0 y) + \sum_{k=1}^{m-1} Tr_1^n(d_k y^{(2^k+1)}) + \\ &\quad Tr_1^m(d_m y^{(2^m+1)}), \; d_0, d_k \in \mathbb{F}_{2^n}, d_m \in \mathbb{F}_{2^m}. \end{aligned}$$

- if n is odd, $n = 2m + 1$:

$$h(y) = Tr_1^n(d_0 y) + \sum_{k=1}^{m} Tr_1^n(d_k y^{(2^k+1)}), \; d_0, d_k \in \mathbb{F}_{2^n}.$$

Proof: Proof is given in the Appendix. ■

Corollary 2: Let $y = x^a$ and $h(x, y) = \sum_{i,j} a_{i,j} x_i y_j + \sum_{i \leq j} e_{i,j} y_i y_j + \sum_i v_i y_i$ be a quadratic Boolean function in n variables where $a_{i,j}, e_{i,j}, v_j \in \mathbb{F}_2$. Then

(i) $h(x, y)$ has a polynomial representation in $\mathbb{F}_{2^n}$ such that:

- if n is even, $n = 2m$:

$$\begin{aligned} h(x, y) \;&=\; \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k + a}) + Tr_1^n(c' x^a) + \\ &\quad \sum_{k=1}^{m-1} Tr_1^n(d_k x^{(2^k+1)a}) + Tr_1^m(d_m x^{(2^m+1)a}), \\ &\quad b_k, d_k, c' \in \mathbb{F}_{2^n}, d_m \in \mathbb{F}_{2^m}. \end{aligned}$$

- if n is odd, $n = 2m + 1$:

$$\begin{aligned} h(x, y) \;&=\; \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k + a}) + Tr_1^n(c' x^a) + \\ &\quad \sum_{k=1}^{m} Tr_1^n(d_k x^{(2^k+1)a}), b_k, d_k, c' \in \mathbb{F}_{2^n}. \end{aligned}$$

(ii) For $n = 2m$, let

$$\begin{aligned} T_2 \;&=\; \{a, 2^k + a \,|\, 0 \leq k < n\} \cup \\ &\quad \{(2^m + 1)a, (2^k + 1)a \,|\, 1 \leq k < m\} \end{aligned}$$

and $S_2 = \{s_j \,|\, 0 \leq j < R\}$ be the set consisting of different coset leaders modulo $2^n - 1$ of the elements in T_2. For $t \in S_2$, assume that there are v elements in T_2, say $J_t = \{i_1, \cdots, i_v\}$ which belong to C_t. Let $l = |C_t|$. Then

$$h(x, y) = \sum_{t \in S_2} Tr_1^l(b'_t x^t) \tag{5}$$

where

$$b'_t = \sum_{i \in J_t} [Tr_l^n(b_i)]^{2^{l-u_i}} \tag{6}$$

where u_i is determined by $i \equiv t2^{u_i} \pmod{2^n - 1}$. Furthermore, the representation in (5) is unique. For n odd, we can obtain a similar unique representation as (5).

Proof: From Corollary 1 we have

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

From proposition 2, for even n, we have

$$\sum_{i\leq j} a_{i,j}y_iy_j = Tr_1^n(d_0x^a) + \sum_{k=1}^{m-1} Tr_1^n(d_k y^{(2^k+1)}) + Tr_1^m(d_m y^{(2^m+1)}).$$

Now we have

$$\sum_{i,j} a_{i,j}x_iy_j + \sum_j v_jy_j + \sum_{i\leq j} a_{i,j}y_iy_j = \sum_{k=0}^{n-1} Tr_1^n(b_kx^{2^k+a}) + Tr_1^n((c+d_0)x^a) + \sum_{k=1}^{m-1} Tr_1^n(d_kx^{(2^k+1)a}) + Tr_1^m(d_mx^{(2^m+1)a}),$$

where $c' = c+d_0$. The proof for the second assertion is similar to the proof given in Corollary 1. Also the proof for n odd is similar. ■

A. Bi-affine Equations

For a given power mapping $y = x^a$ we want to find the number of bi-affine equations of the form

$$\sum_{i,j} a_{i,j}x_iy_j + \sum_j v_jy_j + \sum_i u_ix_i + z = 0, \quad a_{i,j}, b_j, u_i, z \in \mathbb{F}_2. \tag{7}$$

Let $h(x,y)$ be a bilinear function as defined in Corollary 1. Then, a bi-affine equation exists if and only if

$$\begin{aligned} h(x,y) &= \sum_{k=0}^{n-1} Tr_1^n(b_kx^{2^k+a}) + Tr_1^n(cx^a) \qquad (8) \\ &= \sum_i u_ix_i + z. \end{aligned}$$

Therefore the number of bi-affine equations, of the form as given in (7), is equal to the number of functions $h(x,y)$ that are affine (i.e.,$\sum_i u_ix_i + z$).

$h(x,y)$ in (8) will be affine in the following cases:

1) If for any $0 \leq k \leq n-1$, $H(2^k+a) = 1$ then $Tr_1^n(b_kx^{2^k+a})$ will give 2^n affine functions (as $b_k \in \mathbb{F}_{2^n}$). Note only n of these functions are linearly independent. Similarly if $H(a) = 1$ we will get n linearly independent functions.
2) If for any $0 \leq k \leq n-1$, $H(2^k+a) > 1$ and $|C_{2^k+a}| = m' < n$, then $Tr_1^n(b_kx^{2^k+a})$ will give $2^{n-m'}$ affine functions. In this case $Tr_1^n(b_kx^{2^k+a}) = Tr_1^{m'}(b'_kx^{k'})$. Thus b'_k must be 0 for $Tr_1^n(b_kx^{2^k+a})$ to be an affine function. Since $b'_k \in \mathbb{F}_{2^{m'}}$ therefore $2^{n-m'}$ elements in $\mathbb{F}_{2^n}$ map to 0 in $\mathbb{F}_{2^{m'}}$. Only $n-m'$ of these functions are linearly independent. Same argument holds for $Tr_1^n(cx^a)$.
3) Let $\mathcal{A} = \{t = 2^k + a | 0 \leq k < n, H(t) > 1\}$. Now we define $\mathcal{S}$, the set of exponents such that
$$\mathcal{S} = \begin{cases} \mathcal{A} \cup \{a\}, & \text{if } H(a) > 1 \\ \mathcal{A}, & \text{otherwise.} \end{cases}$$
Note $\mathcal{S}$ can be a multiset. If any two exponents from $\mathcal{S}$ belong to the same coset, C_t, then they will give $2^{(2-1)|C_t|} = 2^{|C_t|}$ affine functions. For example if $2^{k_1}+a$ and $2^{k_2}+a$ ($k_1 \neq k_2$) belong to the same coset, C_t where $|C_t| = n$, then we can write $2^{k_1}+a = (2^{k_2}+a)2^l$ where $1 \leq l \leq n$. Now we have a term $Tr_1^n((b_{k_1} + (b_{k_2})^{2^l})x^{2^{k_1}+a})$ where $b_{k_1} + (b_{k_2})^{2^l}$ can be zero in 2^n ways. In general if there exist v exponents that belong to the same coset C_t, we will have $2^{(v-1)|C_t|}$ affine functions.
4) If $a \in C_{2^{n-1}-1}$ then $2^k + a = 2^n - 1$ for some $0 \leq k \leq n-1$. Then we have $Tr_1^n(b_kx^{2^n-1}) = Tr_1^n(b_k)$ which can be 0 in 2^{n-1} ways. Therefore we have $n-1$ linearly independent affine functions. Note $C_{2^{n-1}-1}$ is the only coset with Hamming weight $n-1$ and the inverse exponent belongs to this coset.

Based on the above 4 cases, we now give Algorithm 1 to compute the total number of linearly independent bi-affine equations for a given power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ ($n = 1$ case is trivial).

Remark 1: Cases 2 and 3 are used in Algorithm 1 to easily compute the number of bi-affine equations. The number of equations counted in Cases 2 and 3 is derived from the following result. From (4) in Corollary 1, $b'_t = 0$ gives bi-affine equations. We may write $b'_t = \sum_{j_t} B_i$ where $B_i \in \mathbb{F}_{2^l}$ is the ith trace term in (4). For $b'_t = 0$, we may write

$$B_{i_1} = B_{i_2} + \cdots + B_{i_v}. \tag{9}$$

For each $B_{i_j}, j = 1, \cdots, v$ there are 2^{n-l} choices of $b_{i_j} \in \mathbb{F}_{2^n}$. Thus for a fixed v-tuple $(B_{i_1}, B_{i_2}, \cdots, B_{i_v})$, the total number of the solutions for $b'_t = 0$ is given by $2^{v(n-l)}$. Since each B_{i_j}, $j = 2, \cdots, v$ can be any element in $\mathbb{F}_{2^l}$, there are $2^{(v-1)l}$ choices for $(v-1)$-tuple $(B_{i_2}, \cdots, B_{i_v})$. Thus the total number of the solutions for $b'_t = 0$ is given by $2^{v(n-l)} \cdot 2^{(v-1)l} = 2^{v(n-l)+(v-1)l} = 2^{vn-l}$. Therefore, for the coset leader t, there are $vn-l$ linear independent bi-affine equations.

Algorithm 1 Computing number of linearly independent bi-affine equations

```
Input $a, n > 1$.
Output Total number of independent bi-affine equations.
For given integers $a$ and $2^k + a, 0 \leq k \leq n-1$, compute their coset leaders modulo $2^n - 1$ in array $cst\_l$ and their coset sizes in array $cst\_s$;
sort array $cst\_l$ in ascending order and shuffle array $cst\_s$ accordingly;
$k \leftarrow 0, eqn \leftarrow 0$;
while $k \leq n$ do
    if($H(cst\_l[k] = 1)$): $eqn \leftarrow eqn + n$;
    else
        if($cst\_s[k] < n$): $eqn \leftarrow eqn + (n - cst\_s[k])$;
        if($cst\_l[k] = cst\_l[k+1]$):$eqn \leftarrow eqn + cst\_s[k]$;
    $k \leftarrow k + 1$;
```

Theorem 1: Let $y = x^a$ be a power mapping over $\mathbb{F}_{2^n}$. Then using Algorithm 1, the total number of linearly independent bi-affine equations can be computed in time $O(n^2)$.

Proof: In Algorithm 1, step 1 takes time $O(n^2)$. Step 2 takes time $O(nlog(n))$ and steps 4-9 take time $O(n)$. Therefore total time complexity is $O(n^2)$. Correctness of Algorithm 1 follows from the 4 cases discussed above. ■

B. Quadratic Equations

For a given power mapping $y = x^a$ we want to find the number of quadratic equations of the form

$$\sum_{i,j} a_{i,j} x_i y_j + \sum_{i \le j} b_{i,j} y_i y_j + \sum_i v_i y_i + \sum_{i \le j} c_{i,j} x_i x_j + \sum_i u_i x_i + z = 0, \quad a_{i,j}, b_{i,j}, c_{i,j}, v_i, u_i, z \in \mathbb{F}_2.$$

Let $h(x, y)$ be a quadratic function as defined in Corollary 2. Then, a quadratic equation exists if and only if

- for n even, $n = 2m$:

$$\begin{aligned} h(x,y) &= \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(c' x^a) + \\ &\quad \sum_{k=1}^{m-1} Tr_1^n(d_k x^{(2^k+1)a}) + Tr_1^m(d_m x^{(2^m+1)a}) \\ &= \sum_{i \le j} c_{i,j} x_i x_j + \sum_i u_i x_i + z. \end{aligned} \tag{10}$$

- for n odd, $n = 2m + 1$:

$$\begin{aligned} h(x,y) &= \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k+a}) + Tr_1^n(c' x^a) + \\ &\quad \sum_{k=1}^{m} Tr_1^n(d_k x^{(2^k+1)a}) \\ &= \sum_{i \le j} c_{i,j} x_i x_j + \sum_i u_i x_i + z. \end{aligned}$$

Therefore the number of quadratic equations, of the form as given in (10), is equal to the number of functions $h(x, y)$ that are quadratic (i.e.,$\sum_{i \le j} c_{i,j} x_i x_j + \sum_i u_i x_i + z$). Consider $n = 2m$ to be even then $h(x, y)$ in (10) will be quadratic in the following cases:

1) If for any $0 \le k \le n-1, 1 \le H(2^k + a) \le 2$ then $Tr_1^n(b_k x^{2^k+a})$ will give n linearly independent quadratic functions. If for any $1 \le k \le n-1, 1 \le H((2^k+1)a) \le 2$ then $Tr_1^n(d_k x^{(2^k+1)a})$ will give n linearly independent quadratic functions. If $1 \le H(a) \le 2$ then $Tr_1^n(c' x^a)$ will give n linearly independent quadratic functions. If $1 \le H((2^m+1)a) \le 2$ then $Tr_1^m(d_m x^{(2^m+1)a})$ will give m linearly independent quadratic functions (as $d_m \in \mathbb{F}_{2^m}$).
2) If for any $0 \le k \le n-1$, $H(2^k + a) > 2$ and $|C_{2^k+a}| = m_1 < n$, then $Tr_1^n(b_k x^{2^k+a})$ will give 2^{n-m_1} quadratic functions. If for any $1 \le k \le m-1$, $H((2^k+1)a) > 2$ and $|C_{(2^k+1)a}| = m_2 < n$, then $Tr_1^n(d_k x^{(2^k+1)a})$ will give 2^{n-m_2} quadratic functions. If $H(a) > 2$ and $|C_a| = m_3 < n$, then $Tr_1^n(c' x^a)$ will give 2^{n-m_3} quadratic functions. If $H((2^m+1)a) > 2$ and $|C_{(2^m+1)a}| = m_4 < m$, then $Tr_1^m(d_m x^{(2^m+1)a})$ will give 2^{m-m_4} quadratic functions.
3) Let
$$\mathcal{B} = \{t = 2^k + a | 0 \le k \le n-1, H(t) > 2\}$$
$$\cup \{u = (2^k+1)a | 1 \le k \le m-1, H(u) > 2\}.$$
Now we define the set of exponents $\mathcal{T}$ such that
$$\mathcal{T} = \begin{cases} \mathcal{B} \cup \{a\}, & \text{if } H(a) > 2 \\ \mathcal{B}, & \text{otherwise} \end{cases}$$
Note $\mathcal{T}$ can be a multiset. If any two exponents from $\mathcal{T}$ belong to the same coset, C_t, then they will give $2^{(2-1)|C_t|} = 2^{|C_t|}$ quadratic functions. In general if there exist v exponents that belong to the same coset, C_t, we will have $2^{(v-1)|C_t|}$ quadratic functions.
4) If $a \in C_{2^{n-1}-1}$ then $2^k + a = 2^n - 1$ for some $0 \le k \le n-1$. Therefore we have $n-1$ linearly independent quadratic functions.

Note that odd n can be handled in the similar manner and Algorithm 2 can be modified easily by merging the $(2^m + 1)a$ term in the first while loop. Based on the above 4 cases, we now give Algorithm 2 to compute the total number of linearly independent quadratic equations for a given power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ when n is even. Also note that Cases 2 and 3 above are derived from (6) in Corollary 2 using a similar argument as in Remark 1.

Algorithm 2 Computing number of linearly independent quadratic equations

```
Input $a, n = 2m$.
Output Total number of independent quadratic equations.
For given integers $a$ and $2^k + a, 0 \le k \le n-1$, and $(2^k+1)a, 1 \le k \le m-1$, compute their coset leaders modulo $2^n - 1$ in array cst_l and their coset sizes in array cst_s;
Compute coset leader of $(2^m+1)a$ in cstm_l and its coset size in cstm_s
sort array cst_l in ascending order and shuffle array cst_s accordingly;
$k \leftarrow 0, eqn \leftarrow 0$;
while $k \le n + m - 1$ do
    if($0 < H(cst\_l[k] \le 2)$): $eqn \leftarrow eqn + n$;
    else
        if($cst\_s[k] < n$)): $eqn \leftarrow eqn + (n - cst\_s[k])$;
        if($cst\_l[k] = cst\_l[k+1]$): $eqn \leftarrow eqn + cst\_s[k]$
    $k \leftarrow k + 1$;
if($0 < H(cstm\_l) \le 2$):$eqn \leftarrow eqn + m$;
else
    if($cstm\_s < m$):$eqn \leftarrow eqn + (m - cstm\_s)$;
    $k \leftarrow 0$;
    while $k \le n + m - 1$ do
        if($cstm\_l = cst\_l[k]$):$eqn \leftarrow eqn + cstm\_s$; break;
        $k \leftarrow k + 1$
```

From the above discussion we have the following theorem.

Theorem 2: Let $y = x^a$ be a power mapping over $\mathbb{F}_{2^n}$. Then using Algorithm 2, the total number of linearly independent quadratic equations can be computed in time $O(n^2)$.

IV. $\mathcal{AI}$ of Cryptographically Significant Power Mappings

S-boxes which satisfy zero bi-affine and/or quadratic equations provide optimal resistance against algebraic attacks and therefore are of great interest. In this section we provide several S-box constructions that satisfy this criterion. A cryptographically strong S-box must also have high nonlinearity, good uniform differential property, and high algebraic degree to resist linear, differential, and higher order differential attacks respectively. Unfortunately there are very few S-boxes that satisfy all the above requirements and as n increases, finding these S-boxes becomes extremely difficult. Therefore we identify classes of S-boxes which provably have all the above mentioned properties. We also provide experimental results for smaller values of n to identify good S-boxes and show various tradeoffs involved in the selection of an S-box.

A. Power Mappings with Zero Bi-affine Equations

It is easy to see that the probability of a randomly chosen S-box having zero bi-affine equations is high if n is large [7]. This holds true for S-boxes based on power mappings as well. For example for $n = 8$, 18.8 percent power mappings have zero bi-affine equations. For $n = 16$, 91.1 percent and for $n = 25$, 99.9 percent power mappings satisfy this condition. However note that several highly nonlinear S-boxes (for example inverse mapping) always have bi-affine equations and therefore it is important to calculate the exact number of bi-affine equations an S-box satisfies. If a power mapping is selected randomly then we can use Algorithm 1 in Section III-A to find the number of bi-affine equations it satisfies. Otherwise any one of the following S-box constructions can be used. Consider the power mapping $y = x^a$ over $\mathbb{F}_{2^n}$. From Section III-A, the 3 necessary and sufficient conditions for the power mapping to have zero bi-affine equations are:

1) $H(a) > 1$ and $H(2^i + a) > 1, 0 \leq i \leq n-1$.
2) $(2^i + a)2^l \not\equiv (2^i + a) \bmod 2^n - 1, 0 \leq i \leq n-1$ and $a2^l \not\equiv a \bmod 2^n - 1, l < n$.
3) $(2^i + a)2^l \not\equiv (2^j + a) \bmod 2^n - 1, i \neq j, 0 \leq i, j \leq n-1$ and $a2^l \not\equiv (2^k + a) \bmod 2^n - 1, 0 \leq k \leq n-1, l < n$.

Now we provide four (n, n) power S-box constructions which have zero bi-affine equations. The first construction, given in Theorem 3, consists of S-box based on Kasami-like exponents (See Section IV-C). A subclass of these exponenets, called Kasami exponents was studied by Dobbertin in [10] and mappings based on these exponents were shown to be APN.

Theorem 3: Let $n = 2m, n \geq 8$ and $a = 2^{m+1} + 2^{m-1} - 1$. Then power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ has no bi-affine equations.

Proof: Consider the binary representation of a and 2^i.

$$\begin{array}{rcl} a & = & \overbrace{00\cdots010}^{m}\overbrace{011\cdots11}^{m} \\ 2^i & = & \underbrace{00\cdots\cdots001}_{}\underbrace{0\cdots0}_{i} \end{array}$$

1) $H(a) > 1$ and from the binary representation of a it is obvious that $H(2^i + a) > 1, 0 \leq i \leq n-1$.
2) If $2^i + a, 0 \leq i \leq n-1 \in C_z$ where $|C_z| < n$ then $|C_z|$ divides n. This means that binary representation of $2^i + a$ must contain at least 2 runs of 1's of the same length and 2 runs of 0's of same length. We have 2 runs of 1's when $i = 0, m-1, m, m+1, m+2, 2m-1$. However in all these cases the lengths of the 2 run's of 1's is always different. For all the other cases we get 3 runs of 1's, however they will never be of the same length (we always have 1 run of length 1 and 1 run of length more than 1). Therefore $|C_z| = n$ and $(2^i + a)2^l \not\equiv (2^i + a), l < n$. Also it is evident from the binary representation of a that $a2^l \not\equiv (2^k + a), k < l$.
3) We know that if two integers b and c belong to the same coset then some cyclic shift of binary representation of b gives c. $H(2^i + a) = i + 2, 0 \leq i \leq m-1$, i.e., all belong to distinct cosets. $H(2^i + a) = m, i = m-2, m+1$ but the number of runs of 1's is different in the two cases. $H(2^i + a) = m+1, m \leq i \leq 2m-1, i \neq m+1$. Although the Hamming weight in all these cases is same but either the run's of 1's and 0's are of different length or they appear in different order. Therefore it is not possible to get one integer from the cyclic shift of the binary representation of another element. Therefore for each $0 \leq i \leq n-1$, $2^i + a$ belongs to a distinct coset. Also $H(a) = H(2^i + a)$ only when $i = m-2, m+1$, however in both cases the length of run's of 0's is different from a.

The 3 necessary and sufficient conditions for $y = x^a$ to have zero bi-affine equations (from Section IV-A) are satisfied. ■

Note that this mapping is highly nonlinear, has high algebraic degree and good uniform differential property (See section IV-C).

Theorem 4: Let $n \geq 8$ and $a = 1 + 2 + \sum_{i=3}^{r} 2^i, 3 \leq r \leq n-3$. Then power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ has zero bi-affine equations except when $n = 8, r = 5$.

Proof: The proof is given in the Appendix. ■

The proof technique for Theorems 5 and 6 is similar to that of Theorems 3 and 4. So we provide the theorems without proof.

Theorem 5: Let $n \geq 8$, and $a = 1 + \sum_{i=2}^{r} 2^i, 3 \leq r \leq n-3$. Then power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ has zero bi-affine equations except when n is even and $r = \frac{n}{2} - 1$.

Theorem 6: Let $n \geq 8$, and $a = 1 + 2 + 2^2 + \sum_{i=4}^{r} 2^i, 4 \leq r \leq n-3$. Then power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ has zero bi-affine equations except for the following: $(n = 8, r = 5)$, $(n = 9, r = 5, 6)$ and $(n = 10, r = 7)$.

Several mappings in the constructions given above have good nonlinearity and uniform differential property. For example consider the construction given in Theorem 4. For $n = 8$ power mapping $y = x^{11}$ has nonlinearity 96 and is differentially 10-uniform. For $n = 10$ power mapping $y = x^{11}$ has nonlinearity 480 and is differentially 10-uniform and for $n = 12$ power mapping $y = x^{27}$ has nonlinearity 472 and is differentially 6-uniform. Similarly for construction given in Theorem 5 the mapping $y = x^{13}$ in $\mathbb{F}_{2^{10}}$, has nonlinearity 480 and is differentially 4-uniform. For $n = 12$, mapping $y = x^{1021}$ has nonlinearity 1952 and is differentially 16-uniform.

B. Power Mappings with Zero Quadratic Equations

If a power mapping is selected randomly then we can use Algorithm 2 given in Section III-B to calculate the number of quadratic equations it satisfies. Otherwise we provide a construction below to obtain S-box with zero quadratic equations. Consider the power mapping $y = x^a$ on $\mathbb{F}_{2^n}$. We consider the case where $n = 2m$ is even. *The case where n is odd is similar.* From Section III-B, the 3 necessary and sufficient conditions for the power mapping to have zero quadratic equations are:

1)

$$\begin{aligned} H(a) &> 2, \text{ and } H(2^i + a) > 2, \text{ and} \\ H((2^k + 1)a) &> 2, 0 \le i \le n-1, 1 \le k \le m. \end{aligned}$$

2)

$$\begin{aligned} (2^i + a)2^l &\not\equiv (2^i + a) \bmod 2^n - 1, \\ & \quad 0 \le i \le n-1, l < n, \text{ and} \\ a2^l &\not\equiv a \bmod 2^n - 1, l < n, \text{ and} \\ (2^k + 1)a2^l &\not\equiv (2^k + 1)a, \bmod 2^n - 1, \\ & \quad 1 \le k \le m-1, l < n, \text{ and} \\ (2^m + 1)a2^{l_1} &\not\equiv (2^m + 1)a, \bmod 2^n - 1, l_1 < m. \end{aligned}$$

3) Let

$\mathcal{C} = \{2^i + a : 0 \le i \le n-1\} \cup \{(2^j + 1)a : 1 \le j \le m-1\} \cup \{a\} \cup \{(2^m + 1)a\}$.

$\mathcal{C}$ can be a multiset. Then any two elements in $\mathcal{C}$ belong to different cosets modulo $(2^n - 1)$.

Theorem 7: Let $n \ge 8$, and $a = 1 + 2 + \sum_{i=3}^{r} 2^i, 4 \le r \le n-3$. Then power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ has zero quadratic equations except for the following: $(n = 8, r = 5)$, $(n = 9)$, $(n = 10, r = 5)$ and $(n = 12, r = 4, 8)$.

Proof: Consider the binary representation of a.

$$a = \overbrace{00\cdots0}^{q}\overbrace{1\cdots11}^{l}011$$

First we consider the case $n = 2m, l < q$. Tables I and II show the run distribution (lengths of runs of 1's and 0's) of the binary representation of $(2^i + a), 0 \le i \le n-1$ and $(2^k + 1)a, 1 \le k \le m-1$ respectively. For example a run distribution $4, \underline{3}, 1, \underline{2}$ represents the binary form 0000111011 which is 59 in decimal representation. Note that underlined digits represent the length of run of 1's. For $k = l$ in Table II, there are three different cases for different values of l and for $k = l + 1$ there are 2 different cases for different values of l.

1) From the binary representation of a, Table I and Table II it is obvious that $H(a) > 3$, $H(2^i + a) > 2$, $H((2^k + 1)a) > 2$ and $H((2^m + 1)a) > 2, 0 \le i \le n-1, 1 \le k \le m-1$.
2) If $2^i + a, 0 \le i \le n-1 \in C_z$ where $|C_z| < n$ then $|C_z|$ divides n and binary representation of $2^i + a$ is periodic with period $|C_z|$. It is clear from Table I that the binary representation of $2^i + a$ is never periodic with period less than n. The same reasoning holds for $(2^k + 1)a, 1 \le k \le m-1$ (see Table II). Also from the binary representation of a it is obvious that $|C_a| = n$. Similarly it is easy to check that $(2^m + 1)a2^{l_1} \not\equiv (2^m + 1)a \bmod 2^n - 1, l_1 < m$.
3) If two integers b and c belong to the same coset then some cyclic shift of binary representation of b gives c. This means that both b and c must have same Hamming weight, the number of runs of 1's and 0's must be identical and they must appear in the same order (cyclic). From the binary representations of a and $(2^m + 1)a$, Tables I and II, it is clear that no two elements of the set $\mathcal{C}$ (see item 3 of Section IV-B) belong to the same coset.

TABLE I
RUN DISTRIBUTION IN THE BINARY REPRESENTATION OF $(2^i + a)$

i	run distr. in$(2^i + a)$	i	run distr. in$(2^i + a)$
0	q,$\underline{l+1}$,2	$l+2$	$q-1$,$\underline{1}$,1,$\underline{l-1}$,1,$\underline{2}$
1	q,$\underline{l+1}$,1,$\underline{1}$	$l+3$	$q-1$,$\underline{l+1}$,1,$\underline{2}$
2	q,$\underline{l+3}$	1+4	$q-2$,$\underline{1}$,1,$\underline{l}$,1,$\underline{2}$
3	$q-1$,$\underline{1}$,$l+1$,$\underline{2}$	$l+5$	$q-3$,$\underline{1}$,2,$\underline{l}$,1,$\underline{2}$
4	$q-1$,$\underline{1}$,$l-1$,$\underline{1}$,1,$\underline{2}$	$l+6$ ..	$q-4$,$\underline{1}$,3,$\underline{l}$,1,$\underline{2}$ ....
5 ..	$q-1$,$\underline{1}$,$l-2$,$\underline{2}$,1,$\underline{2}$	$n-1$	$\underline{3}$,$q-1$,$\underline{l}$,1($\underline{1}$,$q-1$,$\underline{l}$,1,$\underline{2}$)

TABLE II
RUN DISTRIBUTION IN THE BINARY REPRESENTATION OF $(2^k + 1)a$

k	run distr. in$(2^k + 1)a$	k	run distr. in$(2^k + 1)a$
1	$q-2$,$\underline{1}$,1,$\underline{l-1}$,3,$\underline{1}$	$l+1$	$q-k$,$\underline{l+2}$,1,$\underline{l-2}$,1, $\underline{2}$ $(l > 2)$ $q-3$,$\underline{4}$,2,$\underline{2}$ $(l = 2)$
2	$q-3$,$\underline{1}$,2,$\underline{l-2}$,2,$\underline{3}$	$l+2$	$q-k$,$\underline{l+1}$,2,$\underline{l-1}$,1,$\underline{2}$
3	$q-4$,$\underline{1}$,3,$\underline{l-3}$,1,$\underline{1}$,2,$\underline{2}$	$l+3$	$q-k$,$\underline{l}$,1,$\underline{l+2}$,1,$\underline{2}$
4	$q-5$,$\underline{1}$,4,$\underline{l-4}$,1,$\underline{1}$,1,$\underline{1}$,1,$\underline{2}$	$l+4$	$q-k$,$\underline{l}$,1,$\underline{2}$,1,$\underline{l}$,1,$\underline{2}$
5	$q-6$,$\underline{1}$,5,$\underline{l-5}$,1,$\underline{1}$,1,$\underline{2}$,1,$\underline{2}$	$l+5$	$q-k$,$\underline{l}$,1,$\underline{2}$,2,$\underline{l}$,1,$\underline{2}$
6 ..	$q-7$,$\underline{1}$,6,$\underline{l-6}$,1,$\underline{1}$,1,$\underline{3}$,1,$\underline{2}$	1+6 ..	$q-k$,$\underline{l}$,1,$\underline{2}$,3,$\underline{l}$,1,$\underline{2}$ ....
l	$q-l-1$,$\underline{1}$,$l+1$,$\underline{1}$,1,$\underline{l-3}$, 1,$\underline{2}$ $(l > 3)$ $q-3$,$\underline{1}$,4,$\underline{3}$ $(l = 2)$ $q-4$,$\underline{1}$,4,$\underline{1}$,2,$\underline{2}$ $(l = 3)$	$m-1$	$q-k$,$\underline{l}$,1,$\underline{2}$,$m-l-4$, $\underline{l}$,1,$\underline{2}$

For $n = 2m, l \ge q$, the run distribution of $(2^i + a)$ is the same as in Table I. If $k < q$ then the run distribution of $(2^k + 1)a$ is the same as given in Table II. Therefore for $k < q$ items 1, 2 and 3 hold with similar logic. For $k \ge q$, $(a2^k + a)$ may generate a carry which results in many possible run distributions depending on the actual value of a. Therefore representing the run distribution of $(a2^k + a)$ in a tabular form becomes cumbersome. It is tedious (but not very hard) to check that items 1, 2 and 3 hold for $k \ge q$ as well. Proof for odd n is very similar. ■

It can be proved easily that if n is even and $r = n - 3$ then power mapping $y = x^a$ in Theorem 7 is bijective if and only if $n \not\equiv 0 \bmod 18$. Also if $r = n - 5$ then the above mapping is bijective if and only if $n \not\equiv 0 \bmod 78$. So if $n < \text{lcm}(18, 78) = 234$, then by taking r to be either $n - 3$ or $n - 5$ we will always get a bijective S-box. Note that bijective mappings exist for many other values of r as well when r is odd. Similar conditions can be found for odd n easily. The power mapping in Theorem 7 is a subset of the power mapping given in Theorem 4 and several mappings with good nonlinearity and uniform differential property listed for Theorem 4 also belong to the class given in Theorem 7.

C. Cryptographically Strong Kasami Like Power Mappings

A Kasami exponent [13], $\acute{e}$, is defined as $\acute{e} = 2^{2s} - 2^s + 1$, gcd$(n, s)$=1 and $1 \leq s < \frac{n}{2}$. If we remove the condition gcd(n, s)=1 we can write $e = 2^{2s} - 2^s + 1$, $1 \leq s < \frac{n}{2}$. We will call e, the Kasami like exponent. Suppose $n = 2m$, and we take $s = m - 1$ then we have $e = 2^{2m-2} - 2^{m-1} + 1$ and $a = 2^{m+1}e = 2^{m+1} + 2^{m-1} - 1$. Note that a is the same exponent as in Theorem 3 which has zero bi-affine equations. Based on our extensive experimental results we provide the following conjecture.

Conjecture 1: Let $n = 2m$, m is odd, and $a = 2^{m+1} + 2^{m-1} - 1$. Then mapping $y = x^a$ is differentially 4-uniform. The validity of Conjecture 1 has been verified up to $n = 22$. Note that gcd$(m - 1, n) = 2$ and therefore the mapping $y = x^a$ is known to be highly nonlinear [11]. Also this mapping is always bijective. Now we provide the following theorem without proof. The proof technique is very similar to the one used in Theorem 7.

Theorem 8: Let $n = 2m, n \geq 8$ and $a = 2^{m+1} + 2^{m-1} - 1$. Then power mapping $y = x^a$ over $\mathbb{F}_{2^n}$ has $2n$ quadratic equations.

Note that this mapping has zero bi-affine equations (see Theorem 3), is highly nonlinear, has high algebraic degree and good uniform differential property. Therefore it is a suitable candidate for cryptographic applications.

D. $\mathcal{AI}$ of Some Well Known Power Mappings

Dobbertin investigated some well known power mappings in [10], i.e., Gold, Dobbertin, Niho, Welch, inverse and Kasami. The number of bi-affine and quadratic equations for some of these power mappings have been found experimentally in [8] (for $n \leq 17$). However using the proof techniques given in Sections IV-A and IV-B the exact number of bi-affine and quadratic equations for the above power mappings can be found out easily. For example Niho exponent e is defined as $e = 2^s + 2^{\frac{s}{2}} - 1$, $n = 2s + 1$ when s is even and $e = 2^s + 2^{\frac{3s+1}{2}} - 1$, $n = 2s + 1$ when s is odd. It can be easily proved that power mapping $y = x^e$ has zero bi-affine equations for $n \geq 7$ and n quadratic equations for $n > 7$. Since power mapping based on Niho exponent is almost bent and APN for odd n, it is a good choice for a cryptographically strong S-box in odd number of variables.

The power mapping with Dobbertin exponent although APN is not highly nonlinear and its algebraic degree is only $\frac{n}{5} + 3$. The algebraic degrees of power mappings with Gold and Welch exponents are 2 and 3 respectively. Therefore these mappings may not be of cryptographic interest for large values of n.

V. Experimental Results

In this section we give some cryptographically significant power mappings. Table III shows all bijective power mappings which have zero bi-affine equations for $n = 8$ and 10. The second column in the table lists the exponents a. Note that a is always the coset leader. The other exponents in the same coset, although not listed in the table, have same properties. The third, fourth and fifth columns show the algebraic degree, nonlinearity and differential uniform property of the mapping respectively. The AES S-box (inverse mapping) has degree 7, nonlinearity 112, and is differentially 4-uniform. From Table III it is clear that for $n = 8$, no bijective power mapping having zero bi-affine equations is as good as AES S-box. For $n = 10$ the optimal bijective power mapping with zero bilinear equations has exponent 79. This mapping has the same nonlinearity and differential uniform property as the inverse mapping, however there is a tradeoff between the degree and the number of bilinear equations. The inverse mapping has 29 bi-affine equations and its degree is 9 where as the mapping with exponent 79 has zero bilinear equations and its degree is 5. Note that exponent 79 for $n = 10$ is the same as defined in Theorems 3 and 8.

Table IV shows power mappings which have zero quadratic equations for $n = 8$ and 10. An * in the second column indicates that the mapping is non-bijective. From the table it is clear that for $n = 8$ and 10, there is no cryptographically good bijective power mapping with zero quadratic equations.

Table V shows highly nonlinear bijective power mappings for $n = 8$ and 10. Columns six and seven in the table give the number of bi-affine and quadratic equations respectively.

Tables III, IV and V are optimal in terms of bi-affine equations, quadratic equations and nonlinearity respectively. Our experiments show that there does not exist an S-box (for $n \leq 25$) which is optimal in terms of all the properties listed in Table V. Therefore we relax our criteria to obtain S-boxes that are optimal or close to optimal in terms of the desired properties. According to the relaxed criteria an S-box based on power mapping F must have the following properties:

- Bijective
- $2^{n-1} - 2^{\frac{n}{2}+1} \leq nl(F)$ for n even and $2^{n-1} - 2^{\frac{n+1}{2}} \leq nl(F)$ for n odd.
- At most n bi-affine equations
- At most differentially k-uniform where $k \leq 6$.
- $deg(F) \geq 3$.

A few S-boxes obtained according to the above criteria are listed in Table VI.

VI. Bijective Power Mappings with Maximum Number of Quadratic Equations

It was stated to be an open problem in [7], [8] to find a bijective nonlinear (n, n) S-box that would give strictly more than $5n$ linearly independent quadratic equations. We provide bijective nonlinear (n, n) S-boxes, where the number of linearly independent quadratic equations is much larger than $5n$. For example from Algorithm 2 it can be checked that for $n = 12$ exponent 683 gives 66 quadratic equations. Also for $n = 14, 16, 18, 20$ and 24 , exponents 2731, 21847, 43691, 174763 and 2796203 give 91, 120, 153, 190 and 276 quadratic equations respectively.

VII. Conclusions

In this paper we developed techniques to calculate the number of multivariate bi-affine and quadratic equations for S-boxes based on power mappings. We also provided two simple and efficient algorithms which are practical even for very large

TABLE III
POWER MAPPINGS WITH ZERO BI-AFFINE EQUATIONS

n	a	$deg(a)$	$nl(F)$	$k-uniform$
8	11	3	96	10
	23	4	96	16
	29	4	96	10
	61	5	96	16
10	13	3	480	4
	19	3	468	6
	23	4	472	6
	43	4	464	6
	47	5	448	34
	53	4	432	34
	59	5	464	6
	61	5	464	6
	71	4	464	6
	79	5	480	4
	89	4	472	6
	109	5	448	34
	119	6	464	6
	125	6	448	34
	151	5	464	6
	175	6	464	6
	191	7	464	6
	221	6	448	34
	245	6	464	6
	251	7	432	34

TABLE IV
POWER MAPPINGS WITH ZERO QUADRATIC EQUATIONS

n	a	$deg(a)$	$nl(F)$	$k-uniform$
8	27*	4	80	26
10	27*	4	472	6
	45*	4	432	6
	51*	4	432	8
	53	4	432	34
	75*	4	480	6
	87*	5	480	4
	105*	4	480	10
	111*	6	432	6
	117*	5	480	6
	123*	6	392	32
	183*	6	448	32
	237*	6	480	4
	251	7	432	34

TABLE V
HIGHLY NONLINEAR POWER MAPPINGS

n	a	$deg(a)$	$nl(F)$	$k-uniform$	bi-affine equations	quadratic equations
8	31	5	112	16	16	36
	91	5	112	16	16	36
	127	7	112	4	23	39
10	5	2	480	4	10	40
	13	3	480	4	0	20
	17	2	480	4	15	40
	25	3	480	8	5	10
	41	3	480	8	5	5
	49	3	480	8	5	15
	79	5	480	4	0	20
	107	5	480	8	5	15
	181	5	480	4	15	35
	205	5	480	4	10	40
	511	9	480	4	29	49

TABLE VI
CRYPTOGRAPHICALLY SIGNIFICANT POWER MAPPINGS

n	a	$deg(a)$	$nl(F)$	$k-uniform$	bi-affine equations	quadratic equations
7	29	4	56	2	0	21
8	29	4	96	10	0	24
9	27	4	240	2	0	9
	87	5	240	2	0	18
10	79	5	480	4	0	20
	223	7	472	4	5	5
11	157	5	976	6	0	0
	249	6	992	2	0	11
	367	7	960	4	0	0
12	731	7	1984	4	9	9
13	367	7	4032	2	0	13
	947	7	3968	4	0	0
	1691	7	4032	2	0	26
14	319	7	8064	4	0	28
	1883	8	8000	6	0	0
15	2033	8	16256	2	0	15

S-boxes. The S-box constructions given in Sections IV-A and IV-B have provable optimal $\mathcal{AI}$. We also gave experimental results for power S-boxes with high algebraic immunity, nonlinearity, algebraic degree and good uniform differential property. We identified two classes of Kasami like power mappings with good $\mathcal{AI}$. One class is known to be APN. The second class is known to be highly nonlinear and we conjectured that it is also differentially 4-uniform. We also identified bijective nonlinear S-boxes that have lowest $\mathcal{AI}$ and disproved a conjecture that AES S-box satisfies the largest number of quadratic equations.

APPENDIX A
PROOF OF PROPOSITION 1

Proof: Let $\{\alpha_0, ..., \alpha_{n-1}\}$ and $\{\beta_0, ..., \beta_{n-1}\}$ be the dual basis of $\mathbb{F}_{2^n}$ and $x = x_0\alpha_0 + \cdots + x_{n-1}\alpha_{n-1}$ and $y = y_0\alpha_0 + \cdots + y_{n-1}\alpha_{n-1}$. Now we can write $x_i = Tr_1^n(\beta_i x)$ and $y_j = Tr_1^n(\beta_j y)$.

Therefore

$$\begin{aligned} h(x,y) &= \sum_{i,j} a_{i,j} x_i y_j = \sum_{i,j} a_{i,j} Tr_1^n(\beta_i x) Tr_1^n(\beta_j y) \\ &= \sum_{i,j} a_{i,j} \sum_{k=0}^{n-1} Tr_1^n(\beta_i^{2^k} \beta_j x^{2^k} y) \\ &= \sum_{k=0}^{n-1} Tr_1^n(\sum_{i,j} a_{i,j} \beta_i^{2^k} \beta_j x^{2^k} y) \\ &= \sum_{k=0}^{n-1} Tr_1^n(b_k x^{2^k} y) \text{ where } b_k = \sum_{i,j} a_{i,j} \beta_i^{2^k} \beta_j \end{aligned}$$

■

APPENDIX B
PROOF OF PROPOSITION 2

Proof: Let $\{\alpha_0, ..., \alpha_{n-1}\}$ and $\{\beta_0, ..., \beta_{n-1}\}$ be the dual basis of $\mathbb{F}_{2^n}$ and $x = x_0\alpha_0 + \cdots + x_{n-1}\alpha_{n-1}$ and $y = y_0\alpha_0 + \cdots + y_{n-1}\alpha_{n-1}$. Now we can write $y_j = Tr_1^n(\beta_j y)$.

Therefore

$$\begin{aligned} h(y) &= \sum_{i \le j} a_{i,j} x_i y_j = \sum_{i \le j} a_{i,j} Tr_1^n(\beta_i y) Tr_1^n(\beta_j y) \\ &= \sum_{i \le j} a_{i,j} \sum_{k=0}^{n-1} Tr_1^n(\beta_i^{2^k} \beta_j y^{(2^k+1)}) \\ &= \sum_{k=0}^{n-1} Tr_1^n(\sum_{i \le j} a_{i,j} \beta_i^{2^k} \beta_j y^{(2^k+1)}) \end{aligned}$$

Now consider the case when n is even. Note that y and y^2 belong to the same coset. Also for $1 \le k < m$, $y^{(2^k+1)}$ and $y^{(2^{n-k}+1)}$ belong to the same coset. Note that $2^m + 1$ belongs to the coset of length m, i.e., $|C_{2^m+1}| = m$. Therefore

$$h(y) = Tr_1^n(d_0 y) + \sum_{k=1}^{m-1} Tr_1^n(d_k y^{(2^k+1)}) + Tr_1^m(d_m y^{(2^m+1)}),$$

where $d_0 = \sum_{i \le j} a_{i,j} (\beta_i \beta_j)^{2^{n-1}}$, $d_k = \sum_{i \le j} a_{i,j} (\beta_i^{2^k} \beta_j + \beta_i \beta_j^{2^k})$, $1 \le k < m$, and $d_m = \sum_{i \le j} a_{i,j} \beta_i^{2^m} \beta_j$. The proof for n odd is similar. ■

Appendix C
Proof of Theorem 4

Proof: Consider the binary representation of a

$$a \quad = \quad \overbrace{00}^{2}\,\overbrace{0 \cdots 0}\,\overbrace{1 \cdots 1011}^{r+1}$$

1) $H(a) > 1$ and from the binary representation of a it is obvious that $H(2^i + a) > 1, 0 \le i \le n - 1$.
2) If $2^i + a, 0 \le i \le n - 1 \in C_z$ where $|C_z| < n$ then $|C_z|$ divides n. This means that for even n binary representation of $2^i + a$ must contain at least 2 runs of 1's of the same length and 2 runs of 0's of same length. We get 2 runs of 1's when $i = 1, 3, 2^{r+1}, 2^{n-1}$. Except for the case where $n = 8, r = 5, i = 7$, $2^i + a$ can never fulfill this condition. We get 3 runs of 1's when $4 \le i \le n - 2, i \ne r + 1$. However we always have 1 run of 1's of length 1 and one run of 1's of length 2 so the above condition is not satisfied. More than 3 runs are impossible in the binary representation of $2^i + a$. For odd n we must have at least 3 runs of 1's of same length in the binary representation of $2^i + a$ which is impossible from the above reasoning. Therefore $|C_z| = n$ and $(2^i + a)2^l \not\equiv (2^i + a), l < n$. Also it is evident from the binary representation of a that $a2^l \not\equiv (2^k + a), k < l$.
3) We know that if two integers b and c belong to the same coset then some cyclic shift of binary representation of b gives c. The binary representation of $2^i + a$ has 1 run of 1's when $i = 0, 2$, however the length of the run is different for $i = 0$ and $i = 2$. We have 2 runs of 1's when $i = 1, 3, 2^{r+1}, 2^{n-1}$, however either the length of run's of 1's and 0's is different in each case or the runs of 1's and 0's appear in a different order. We have 3 runs of 1's when $4 \le i \le n - 2, i \ne r + 1$. For each $4 \le i \le k$, $2^i + a$ will have a different hamming weight. For each $k + 1 \le i \le n - 2$, $2^i + a$ has same hamming weight but either the run's of 1's and 0's are of different length or they appear in different order. Therefore for each $0 \le i \le n-1$, $2^i + a$ belongs to a distinct coset. Also $H(a) = H(2^i + a)$ only when $i = 1, k$, however in both cases the length of run's of 1's is different from a.

The 3 necessary and sufficient conditions for $y = x^a$ to have zero bi-affine equations (from Section IV-A) are satisfied. ■

Appendix D
Cryptographically Significant Power Mappings

Table VII lists all cryptographically significant S-boxes based on power mappings for $n = 7, 8, 9, 10, 12$, and 15. For $n = 11, 13$, and 14, the data has been omitted due to its large size and can be downloaded from G. Gong's website. Let F be a power mapping on which an S-box is based, then every S-box in the list satisfies the following criteria:

- Bijective
- $2^{n-1} - 2^{\frac{n}{2}+1} \le nl(F)$ for n even and $2^{n-1} - 2^{\frac{n+1}{2}} \le nl(F)$ for n odd.
- At most n bi-affine equations
- At most differentially k-uniform where $k \le 6$.
- $deg(F) \ge 3$.

Acknowledgment

This research was supported by NSERC (Natural Sciences and Engineering Research Council of Canada) Strategic Project Grant. The authors would like to thank Mr. NamYul Yu for many useful discussions during the course of this work.

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] F. Chabaud and S. Vaudenay, Links Between Differential and Linear Cryptanalysis,*Advances in Cryptology - Eurocrypt 1994*, LNCS 950, pp. 356-365, Springer-Verlag, 1995.

[4] J. Cheon and D. Lee, Resistance of S-Boxes Against Algebraic Attacks, *Fast Software Encryption 2004*, LNCS 3017, pp. 83-94, Springer-Verlag, 2004.

[5] N. Courtois, General Principles of Algebraic Attacks and New Design Criteria for Components of Symmetric Ciphers, *AES 4 Conference, Bonn*, LNCS 3373, pp. 67-83, Springer-Verlag, 2004.

[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, *Information Security and Privacy - ACISP 2006*, LNCS 4058, pp. 76-86, Springer-Verlag, 2006.

[9] J. Daemen, and V.Rijmen, *The Design of Rijndael*, Springer-Verlag, 2002.

[10] H. Dobbertin, Almost Perfect Nonlinear Power Functions on $GF(2^n)$: The Welch Case. *IEEE Transactions on Information Theory*, Vol. 45, No. 4, pp. 1271-1275, 1999.

[11] H. Dobbertin, One-to-One Highly Nonlinear Power Functions on $GF(2^n)$. *Applicable Algebra in Engineering, Communication and Computing*, Vol. 9, pp. 139-152, 1998.

[12] S. W. Golomb, and G. Gong, *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*, Cambridge University Press, ISBN 0521821045, 2005.

TABLE VII
CRYPTOGRAPHICALLY SIGNIFICANT POWER MAPPINGS

n	a	deg of a	$nl(F)$	diff-k unf	bi-aff eqns	quad eqns
7	11	3	56	2	0	21
	13	3	56	2	0	21
	23	4	56	2	0	21
	27	4	56	2	7	28
	29	4	56	2	0	21
8	11	3	96	10	0	24
	29	4	96	10	0	24
9	13	3	240	2	0	18
	19	3	240	2	0	9
	27	4	240	2	0	9
	45	4	232	4	0	9
	47	5	240	2	0	18
	59	5	240	2	0	18
	87	5	240	2	0	18
	103	5	240	2	9	36
	125	6	232	4	0	9
10	13	3	480	4	0	20
	19	3	464	6	0	10
	23	4	472	6	0	10
	43	4	464	6	0	10
	59	5	464	6	0	10
	61	5	464	6	0	10
	71	4	464	6	0	10
	79	5	480	4	0	20
	89	4	472	6	0	10
	91	5	464	6	5	5
	103	5	464	4	5	5
	115	5	464	6	5	15
	119	6	464	6	0	10
	149	4	464	4	5	5
	151	5	464	6	0	10
	167	5	456	6	5	5
	175	6	464	6	0	10
	191	7	464	6	0	10
	205	5	480	4	10	40
	215	6	464	6	5	5
	223	7	472	4	5	5
	235	6	464	6	5	5
	239	7	456	6	5	5
	245	6	464	6	0	10
	347	6	464	6	5	15
	367	7	472	4	5	5
	379	7	464	6	5	5
12	73	3	1984	4	9	9
	341	5	1952	6	10	10
	731	7	1984	4	9	9
	853	6	1952	6	10	10
15	13	3	16256	2	0	30
	73	3	16128	6	15	15
	131	3	16256	2	0	15
	241	5	16256	2	0	30
	383	8	16256	2	0	30
	521	3	16128	6	15	15
	1371	7	16256	2	0	15
	1935	8	16256	2	15	60
	2033	8	16256	2	0	15
	2523	8	16256	2	0	30
	3671	8	16256	2	0	30
	4717	7	16128	6	15	15
	4791	8	16256	2	0	30
	4815	8	16256	2	0	15
	4941	7	16128	6	15	15
	6555	8	16256	2	15	60

[13] 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.

[14] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and their Applications*. Cambridge University Press, 1994.

[15] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*. North Holland, 1986.

[16] S. Murphy and M. Robshaw, Essential Algebraic Structure within AES, *Advances in Cryptology - Crypto 2002*, LNCS 2442, pp.1-16, Springer-Verlag, 2002.

[17] S. Murphy and M. Robshaw, Comments on the Security of the AES and the XSL Technique, *Electronic Letters*, Vol. 39, pp. 26-38, 2003.

[18] K. Nyberg, Differentially Uniform Mappings for Cryptography, *Advances in Cryptology - Eurocrypt 1993*, LNCS 765, pp.55-64, Springer-Verlag, 1994.

[19] 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.