Two and Three Output Perfect Nonlinear and Maximally Nonlinear S-Boxes Satisfying Higher Order Strict Avalanche Criteria

Kishan Chand Gupta and Palash Sarkar
Cryptology Research Group
Applied Statistics Unit
Indian Statistical Institute
203, B.T. Road
Kolkata 700108, India
e-mail:{kishan_t,palash}@isical.ac.in

Abstract

We consider the problem of constructing perfect nonlinear S-boxes satisfying higher order strict avalanche criteria (SAC). Our first construction is an infinite family of 2-ouput perfect nonlinear functions satisfying higher order SAC. This construction is achieved using the theory of bilinear forms and symplectic matrices. Next we build on a known connection between 1-factorization of a complete graph and SAC to construct more examples of 2 and 3-output perfect nonlinear functions. In certain cases, the constructed S-boxes have optimal trade-off between the following parameters: numbers of input and output variables, nonlinearity and order of SAC. In the case the number of input variables is odd, we modify the construction for perfect nonlinear S-boxes to obtain a construction for maximally nonlinear S-boxes satisfying higher order SAC. Our constructions present the first examples of perfect nonlinear and maximally nonlinear multioutput S-boxes satisfying higher order SAC. Lastly, we present a simple method for improving the degree of the constructed functions with a small trade-off in nonlinearity and the SAC property. This yields functions which are well suited as building blocks for the design of block ciphers.
Keywords : S-box, SAC, bent function, bilinear form, symplectic matrix, nonlinearity, symmetric ciphers.

1 Introduction

An S-box is a mapping from $\{0,1\}^n$ to $\{0,1\}^m$. These are used in the design of symmetric ciphers. For example, the S-boxes used in DES have $n = 6$ and $m = 4$ and the S-box used in the design of AES has $n = m = 8$. In the case $m = 1$, an S-box is called a Boolean function. An S-box (or Boolean function) which is used in the design of a block (or stream) cipher must satisfy some desirable properties. These properties ensure resistance to certain classes of attacks.

Nonlinearity is one of the basic properties of an S-box. The nonlinearity of a Boolean function measures the distance of the function to the set of all affine functions. The nonlinearity of an S-box is a natural generalization of this notion. If the number of input variables is n, then the maximum nonlinearity achieveable is $2^{n-1} - 2^{(n-2)/2}$. If $m = 1$, functions achieving this value of nonlinearity are called bent [11] and for $m > 1$ they are called perfect nonlinear S-boxes [9]. It is known that such functions can exist only if n is even and $m \leq (n/2)$. Bent functions are important objects in coding theory and combinatorics. Also both bent functions and perfect nonlinear functions have studied quite extensively in the cryptology literature. In the case where n is odd and $m > 1$, the maximum achievable nonlinearity is $2^{n-1} - 2^{(n-1)/2}$. Functions achieving this nonlinearity are called maximally nonlinear.

Differential cryptanalysis is one of the basic attacks on block ciphers. This attack attempts to utilize the unevenness in the distribution of the output values of $g(x) \oplus g(x \oplus \alpha)$, where $g(x)$ is an S-box and α is a non zero bit string. A natural attempt to resist differential attacks is to design $g(x)$ so that $g(x) \oplus g(x \oplus \alpha)$ is a balanced function. This motivates the following property introduced in [10].

An S-box $f(x)$ is said to satisfy propagation characteristic of degree l and order k ($PC(l)$ of order k) if the following holds: Let $g(y)$ be a function obtained from $f(x)$ by fixing at most k inputs to constant values and let α be a non zero vector of maximum weight l. Then $g(y) \oplus g(y \oplus \alpha)$ is a balanced function.

If $k = 0$, then the function is simply said to satisfy PC(l). PC(l) of order k is a very general notion of resistance against differential cryptanalysis. This property has also been studied in [3, 4] and constructions of Boolean functions and S-boxes satisfying PC(l) of order k are known [7, 6, 12]. S-boxes satisfying PC(1) of order k are said to satisfy strict avalanche criteria of order k (SAC(k)). If $k = 0$, then the S-box is said to satisfy SAC. The notion of SAC was introduced in [13]. It is known [8] that any bent function or any perfect nonlinear S-box satisfies PC(n). It is also possible to construct bent functions satisfying SAC($n - 2$). However, for $m > 1$, construction of perfect nonlinear S-boxes satisfying SAC(k) for $k > 0$ has been a open problem.

In this paper, we (partially) solve this problem by providing constructions of perfect nonlinear S-boxes with $m = 2, 3$ and satisfying SAC(k) for $k \geq 1$. Our contributions are the following.

- Construction of an infinite family of 2-output perfect nonlinear S-boxes satisfying higher order SAC. More precisely, for each even $n \geq 6$, we construct a 2-output perfect nonlinear S-box satisfying SAC($(n/2) - 2$).

- In an earlier paper [7], a 1-factorization of the complete graph on n-vertices was used to construct S-boxes satisfying higher order SAC. However, the S-boxes constructed in [7] did not satisfy perfect nonlinearity. We make a more detailed analysis of the connection between 1-factorization and higher order SAC to construct 2 and 3 output *perfect nonlinear* S-boxes satisfying higher order SAC.

- In certain cases, the functions that we construct achieve the best possible trade-off among the following parameters: number of input variables, number of output variables, nonlinearity and order of SAC. Hence for such functions, it is not possible to improve any one parameter without changing some other parameter.

- For small n, our constructions provide S-boxes which cannot be obtained from the currently known constructions [7, 6, 12]. Some examples of such functions are the following.

  - 8-input, 2-output perfect nonlinear S-box satisfying SAC(2).
  - 8-input, 3-output perfect nonlinear S-box satisfying SAC(1).
  - 10-input, 3-output perfect nonlinear S-box satisfying SAC(3).

  The last example is also an example of an S-box achieving the best possible trade-off.

- Our constructions are based on bilinear forms and symplectic matrices used in the study of second order Reed-Solomon code. We show that if n is odd, then the construction for $(n+1)$ can be modified to obtain maximally nonlinear S-boxes satisfying higher order SAC.

- We provide a simple technique for improving the degree of an S-box with a small sacrifice in nonlinearity and the SAC property. This results in S-boxes which are well suited for use as building blocks in the design of block ciphers.

2 Preliminaries

Let $F_2 = GF(2)$. We consider the domain of a Boolean function to be the vector space $(F_2^n, \oplus)$ over F_2, where $\oplus$ is used to denote the addition operator over both F_2 and the vector space F_2^n. The inner product of two vectors $u, v \in F_2^n$ will be denoted by $\langle u, v\rangle$. The weight of an n-bit vector u is the number of ones in u and will be denoted by $\mathsf{wt}(u)$. The (Hamming) distance between two vectors $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n)$ is the number of places where they differ and is denoted by $d(x, y)$. The bitwise complement of a bit string x will be denoted by $\overline{x}$.

2.1 Boolean Functions

An n-variable Boolean function is a map $f : F_2^n \to F_2$. The weight of f, denoted by $\mathsf{wt}(f)$ is defined as $\mathsf{wt}(f) = |\{x : f(x) = 1\}|$. The function f is said to be balanced if $\mathsf{wt}(f) = 2^{n-1}$. The (Hamming) distance between two n-variable Boolean functions f and g is $d(f, g) = |\{x : f(x) \neq g(x)\}|$.

A parameter of fundamental importance in cryptography is the nonlinearity of a Boolean function. This quantity measures the distance of a Boolean function from the set of all affine functions. An n-variable affine function is of the form $l_{u,b}(x) = \langle u, x\rangle \oplus b$, where $u \in F_2^n$ and $b \in F_2$. Let A_n be the set of all n-variable affine functions. The nonlinearity $\mathsf{nl}(f)$ of an n-variable Boolean function is defined as $\mathsf{nl}(f) = \min_{l \in A_n} d(f, l)$. The maximum nonlinearity achievable by an n-variable Boolean function is $2^{n-1} - 2^{(n-2)/2}$. Functions achieving this value of nonlinearity are called bent and can exist only when n is even [11]. When n is odd, the maximum nonlinearity achievable by an n-variable Boolean function is not known. However, functions achieving a nonlinearity of $2^{n-1} - 2^{(n-1)/2}$ are easy to construct and are called maximally nonlinear [4].

An n-variable Boolean function f satisfies strict avalanche criteria (SAC) if $f(x) \oplus f(x \oplus \alpha)$ is balanced for any $\alpha \in F_2^n$ with $\mathsf{wt}(\alpha) = 1$ [13]. A function f satisfies SAC(k) if every subfunction obtained from $f(x_1, \cdots, x_n)$ by keeping at most k input bits constant satisfies SAC.

An n-variable Boolean function can be represented as a multivariate polynomial over F_2. The degree of this polynomial is called the degree of the function. Affine functions have degree one and functions of degree two are called quadratic.

2.2 S-Boxes

An (n, m) S-box (or vectorial function) is a map $f : \{0, 1\}^n \to \{0, 1\}^m$. Let $f : \{0, 1\}^n \to \{0, 1\}^m$ be an S-box and $g : \{0, 1\}^m \to \{0, 1\}$ be an m-variable Boolean function. The composition of g and f, denoted by $g \circ f$ is an n-variable Boolean function defined by $(g \circ f)(x) = g(f(x))$.

Let f be an (n, m) S-box. The nonlinearity of f is defined to be $\mathsf{nl}(f) = \min\{\mathsf{nl}(l \circ f) : l$ is a non-constant m-variable linear function$\}$. The maximum achievable nonlinearity of an n-variable function is $2^{n-1} - 2^{(n-2)/2}$ and S-boxes achieving this value of nonlinearity are called perfect nonlinear S-boxes. Such S-boxes exist only if n is even and $m \leq (n/2)$ [9]. For odd n and $m > 1$, the maximum possible nonlinearity achievable is $2^{n-1} - 2^{(n-1)/2}$ and S-boxes achieving this value of nonlinearity are called maximal nonlinear S-boxes.

We define the degree of an (n, m) S-box f to be the minimum of the degrees of $l \circ f$, where l ranges over all non constant m-variable linear functions. This definition is more meaningful to cryptography than the definition where the degree of an S-box is taken to be the maximum of the degrees of all the component functions. The later definition has been used in [2].

An (n, m) S-box f is said to be SAC(k), if $l \circ f$ is SAC(k) for every non-constant m-variable linear function l. By an (n, m, k) S-box we mean an (n, m) S-box which is SAC(k). We will be interested in (n, m, k) S-boxes with maximum possible nonlinearity. More specifically, we will be interested in (n, m, k)

perfect nonlinear S-boxes if n is even and in (n, m, k) maximally nonlinear S-boxes if n is odd. Such S-boxes have important applications in the design of secure block ciphers.

2.3 Binary Quadratic Form

An n-variable Boolean function g of degree ≤ 2 can be written as (see [8, page 434]) $g(x) = xQx^T \oplus Lx^T \oplus b$ where $Q = (q_{ij})$ is an upper triangular $n \times n$ binary matrix, $L = (l_1, \cdots, l_n)$ is a binary vector and b is 0 or 1. The expression xQx^T is called a quadratic form and Lx^T is called a linear form. Let $B = Q \oplus Q^T$. Then B is a binary symmetric matrix with zero diagonal. Such a matrix is called a symplectic matrix (see [8, page 435]). Thus from a quadratic Boolean function we can define a symplectic matrix. *Conversely, given a symplectic mmatrix B we can construct a quadratic Boolean function by reversing the above steps. We denote this Boolean function by f_B.*

It is known that the rank of a symplectic matrix is always even [8, page 436]. The nonlinearity of the Boolean function g is related to the rank of B by the following result [8, page 441].

Proposition 1 *Let g be a quadratic n-variable Boolean function and B be its associated symplectic form. Then the nonlinearity of g is equal to $2^{n-1} - 2^{n-h-1}$, where the rank of B is $2h$.*

Consequently, a quadratic Boolean function is bent if and only if the associated symplectic matrix is of full rank.

3 Basic Results

We will be interested in nonlinear quadratic functions satisfying higher order SAC. From Proposition 1, a convenient way to study the nonlinearity of quadratic functions is through the rank of the associated symplectic matrix. We now develop the basic relationships between the nonlinearity and SAC property of a quadratic S-box and the symplectic matrices associated with the component functions.

Proposition 2 *Let f be a quadratic Boolean function and B its associated symplectic matrix. Then f satisfies $SAC(k)$ if and only if for all $1 \leq i \leq n$, we have $\mathsf{wt}(r_i) \geq k+1$, where $r^{(i)}$ is the i^{th} row of B. (Since B is symmetric, a similar property holds for the columns of B.)*

Proof : Let $f(x) = xQx^T \oplus Lx^T \oplus b$. Let α be such that only the ith component of α is 1 and all other components are zero. Further, let the ith column of Q be $(a^{(i)})^T$ and the ith row of Q be $b^{(i)}$. Then $r^{(i)} = (a^{(i)})^T \oplus b^{(i)}$. We have

$$\begin{aligned}
f(x) \oplus f(x \oplus \alpha) &= xQx^T \oplus (x \oplus \alpha)Q(x \oplus \alpha)^T \oplus L\alpha^T \\
&= xQ\alpha^T \oplus \alpha Q x^T \oplus L\alpha^T \oplus \alpha Q \alpha^T \\
&= \langle b^{(i)} \oplus (a^{(i)})^T, x \rangle \oplus L\alpha^T \oplus \alpha Q \alpha^T \\
&= \langle r^{(i)}, x \rangle \oplus L\alpha^T \oplus \alpha Q \alpha^T
\end{aligned}$$

Note that $L\alpha^T \oplus \alpha Q \alpha^T$ is a constant. Now suppose $\mathsf{wt}(r^{(i)}) \geq k+1$. Let $g(x)$ be a function obtained by setting any k bits of $f(x) \oplus f(x \oplus \alpha)$ to constant values. Then $\langle r^{(i)}, x \rangle$ is a non constant linear function and hence $g(x)$ is balanced. Conversely, if $\mathsf{wt}(r^{(i)}) \leq k$, then we can set k variables to constant values in such a manner that $g(x)$ is a constant function. This proves the result. ∎

Let $f = (f_1, \cdots, f_m)$ be an (n, m) quadratic S-box. Then each of the component functions f_i is an n-variable quadratic Boolean function. For $1 \leq i \leq m$, let B_i be the symplectic matrix associated with the component function f_i. Clearly, any linear combination of symplectic matrices is also a symplectic matrix. We have the following extension of Proposition 2.

Lemma 3 *Let f be an (n, m) S-box with component functions f_i and associated symplectic forms B_i for $1 \le i \le m$. Then f satisfies SAC(k) if and only if the weight of every row in any non zero linear combination of the B_i's is at least $k+1$.*

A similar result for nonlinearity can be stated by extending Proposition 1.

Lemma 4 *Let f be an (n, m) S-box with component functions f_i and associated symplectic forms B_i for $1 \le i \le m$. The nonlinearity of f is $2^{n-1} - 2^{n-h-1}$, where $2h$ is the minimum of the ranks of any non zero linear combination of the B_i's. Consequently for even n, the S-box f is perfect nonlinear if and only if every non zero linear combination of the B_i's have full rank. Similarly, for odd n, the S-box f is maximally nonlinear if and only if every non zero linear combination of the B_i's have rank $(n-1)$.*

Lemmas 3 and 4 will be used in proving the correctness of our constructions in the next sections.

4 Construction of $(n, 2, \frac{n}{2} - 2)$ S-box

Our construction will be via symplectic matrices. Given any (n, r) quadratic S-box, it is clear from the above discussion that the symplectic matrices associated with the output component function uniquely defines the S-box. Thus to describe the construction, it is sufficient to define these symplectic matrices and use Lemmas 3 and 4 to prove the correctness of the construction.

In this section, we describe the construction of $(n, 2)$ S-boxes. Hence it is sufficient to define two symplectic matrices. We proceed to do this as follows. For each even $n \ge 6$, we define two sequences of $n \times n$ matrices and show that these matrices are the required symplectic matrices required in the construction. For the rest of this paper, we will use the following notation.

- For each $n \ge 1$, define v_n to be the alternating sequence of 0's and 1's starting with a 0. For example, $v_4 = 0101$ and $v_5 = 01010$. Define $w_n = 1\overline{v_{n-1}}$.
- For each even $n \ge 2$, define u_n as $u_n = \underbrace{1\ldots1}_{(n/2)}\underbrace{0\ldots0}_{(n/2)}$. For odd $n \ge 3$, define $x_n = 1\overline{u_{n-1}}$.

Define $M_4 = [0010, 0010, 1101, 0010]^T$ and $N_4 = [0101, 1011, 0101, 1110]^T$. Further, for even $n > 4$ define

$$
\begin{aligned}
M_n &= \begin{bmatrix} 0 & v_{n-2} & 0 \\ v_{n-2}^T & M_{n-2} & v_{n-2}^T \\ 0 & v_{n-2} & 0 \end{bmatrix}, \quad F_n = \begin{bmatrix} 0 & v_{n-2} & 0 \\ v_{n-2}^T & M_{n-2} & u_{n-2}^T \\ 0 & u_{n-2} & 0 \end{bmatrix}; \\
N_n &= \begin{bmatrix} 0 & \overline{v_{n-2}} & 1 \\ \overline{v_{n-2}^T} & N_{n-2} & \overline{v_{n-2}^T} \\ 1 & \overline{v_{n-2}} & 0 \end{bmatrix}, \quad G_n = \begin{bmatrix} 0 & \overline{v_{n-2}} & 1 \\ \overline{v_{n-2}^T} & N_{n-2} & \overline{u_{n-2}^T} \\ 1 & \overline{u_{n-2}} & 0 \end{bmatrix}.
\end{aligned} \tag{1}
$$

The following result is easy to prove by induction on even $n \ge 4$.

Lemma 5 *F_n, G_n and $F_n \oplus G_n$ are symplectic matrices, where F_n and G_n are defined by equation 1.*

The matrices F_n and G_n are our required symplectic matrices which define the two output component functions of the required $(n, 2)$ S-box. In particular, we have the following result.

Theorem 6 *Let $n \ge 6$ be an even integer. The S-box $f : F_2^n \to F_2$ defined by $f(x) = (f_{F_n}(x), f_{G_n}(x))$ is a perfect nonlinear S-box satisfying SAC($\frac{n}{2} - 2$).*

We now turn to the proof of correctness of Theorem 6. The proof is in two parts – in the first part we prove the statement about SAC and in the second part we prove the statement about nonlinearity.

Lemma 7 *The S-box f defined in Theorem 6 satisfy $SAC(\frac{n}{2}-2)$.*

Proof : Let r_j denote the j-th row of M_n. We make the following claim which can be routinely proved by induction on even $n \geq 4$.

$$\left.\begin{array}{lllll} \mathsf{wt}(r_j) & \geq & \frac{n}{2}-1 & \text{if } 1 \leq j \leq \frac{n}{2} & \text{and} \quad j \text{ is odd;} \\ \mathsf{wt}(r_j) & \geq & \frac{n}{2}-2 & \text{if } 1 \leq j \leq \frac{n}{2} & \text{and} \quad j \text{ is even;} \\ \mathsf{wt}(r_j) & \geq & \frac{n}{2} & \text{if } \frac{n}{2}+1 \leq j \leq n & \text{and} \quad j \text{ is odd;} \\ \mathsf{wt}(r_j) & \geq & \frac{n}{2}-1 & \text{if } \frac{n}{2}+1 \leq j \leq n & \text{and} \quad j \text{ is even.} \end{array}\right\} \tag{2}$$

We will use the notation r'_j for j-th row which is obtained by dropping first and last column of M_n. Let s_j denote the jth row of F_n. We now have several cases.

Case 1 : $1 \leq j \leq \frac{n}{2}$ and j odd: There are two subcases.

Subcase 1(a) : $j = 1$. In this case $\mathsf{wt}(s_j) = \mathsf{wt}(v_{n-2}) = \frac{n-2}{2} = \frac{n}{2} - 1$.

Subcase 1(b) : $j > 1$. In this case $\mathsf{wt}(s_j) = 1 + 1 + \mathsf{wt}(r'_j) \geq 2 + \frac{n-2}{2} - 2 = \frac{n}{2} - 1$.

Case 2 : $1 \leq j \leq \frac{n}{2}$ and j even: In this case $\mathsf{wt}(s_j) = 1 + \mathsf{wt}(r'_j) \geq 1 + \frac{n-2}{2} - 1 = \frac{n}{2} - 1$.

Case 3 : $\frac{n}{2} + 1 \leq j \leq n$ and j odd: In this case $\mathsf{wt}(s_j) = 1 + \mathsf{wt}(r'_j) \geq 1 + \frac{n-2}{2} - 1 = \frac{n}{2} - 1$.

Case 4 : $\frac{n}{2} + 1 \leq j \leq n$ and j even: There are two subcases.

Subcase 4(a) : $j < n$. In this case $\mathsf{wt}(s_j) = \mathsf{wt}(r'_j) \geq \frac{n-2}{2} = \frac{n}{2} - 1$.

Subcase 4(b) : $j = n$. In this case $\mathsf{wt}(s_j) = \mathsf{wt}(u_{n-2}) = \frac{n-2}{2} = \frac{n}{2} - 1$.

This proves that the weight of each row of F_n is at least $(n/2) - 1$ and hence the corresponding Boolean function satisfies SAC($(n/2) - 2$). By a similar argument the Boolean function associated with G_n also satisfies SAC($(n/2) - 2$). Also note

$$F_n \oplus G_n = \begin{bmatrix} 0 & J_{n-2} & 1 \\ J^T_{n-2} & M_{n-2} \oplus N_{n-2} & J^T_{n-2} \\ 1 & J_{n-2} & 0 \end{bmatrix},$$

where J_n is all 1 vector. From this it is simple to verify by induction that $F_n \oplus G_n$ satisfies SAC($\frac{n}{2} - 2$). Now using Lemma 3 we obtain the required result. ■

We next turn to the nonlinearity of the S-box defined in Theorem 6.

Lemma 8 *For even $n \geq 6$, the rank of F_n is n.*

Proof : First we prove that the rank of M_n is $n - 2$. It is easy to check that the rank of M_4 is 2. Assume that the rank of M_{n-2} is $n - 4$. It is clear that 1-st column and n-th column of M_n are identical. Likewise 1-st column and $n - 2$-th column of M_{n-2} are identical. Consider the matrix

$$M'_n = \begin{bmatrix} 0 & v_{n-2} \\ v^T_{n-2} & M_{n-2} \end{bmatrix}.$$

From the definition of v_n, we have that the first bit of v_{n-2} is 0 and $(n-2)$-th bit is 1. So v_{n-2} is linearly independent of rows of M_{n-2}. So rank of M'_n is at least $n - 4 + 1 = n - 3$. But M'_n is symplectic matrix and hence its rank must be even (see [8, page 436]). So the rank of M'_n (and hence M_n) is $n - 2$.

Now we turn to the rank of F_n. As M'_n has rank $n - 2$, the rank of F_n is at least $n - 2$. It is simple to verify by induction that $\frac{n}{2}$-th column and $(\frac{n}{2}+2)$-th column of M_n are identical. From definition, the $\frac{n}{2}$-th bit of $0u_{n-2}0$ is 1 and the $(\frac{n}{2} + 2)$-th bit is 0. Hence the last row $0u_{n-2}0$ of F_n is linearly independent of the previous $(n - 1)$ rows. Thus the rank of F_n is at least $n - 2 + 1 = n - 1$. But F_n is a binary symplectic matrix and hence its rank must be even. Hence the rank of F_n is n. ■

Lemma 9 *For even $n \geq 6$, the rank of $F_n \oplus G_n$ is n.*

Proof : Note $M_4 \oplus N_4 = [0111, 1001, 1000, 1100]^T$ and hence the rank of $M_4 \oplus N_4$ is 4. Assume that the rank of $M_{n-2} \oplus N_{n-2}$ is $n-2$. Note

$$F_n \oplus G_n = M_n \oplus N_n = \begin{bmatrix} 0 & J_{n-2} & 1 \\ J_{n-2}^T & M_{n-2} \oplus N_{n-2} & J_{n-2}^T \\ 1 & J_{n-2} & 0 \end{bmatrix},$$

where J_n is the all 1 vector. The row $1J_{n-2}0$ is linearly independent of rows of matrix $J_{n-2}^T(M_{n-2} \oplus N_{n-2})J_{n-2}^T$. So rank of

$$\begin{bmatrix} J_{n-2}^T & M_{n-2} \oplus N_{n-2} & J_{n-2}^T \\ 1 & J_{n-2} & 0 \end{bmatrix}$$

is at least $n-2+1 = n-1$ and hence the rank of $F_n \oplus G_n$ is at least $n-1$. Again since $F_n \oplus G_n$ is a symplectic matrix its rank must be even. Hence its rank is n. ■

We define $T_5 = [01010, 10101, 01011, 10101, 01110]^T$,

$$T_n = \begin{bmatrix} 0 & \overline{v_{n-2}} & 0 \\ \hline v_{n-2}^T & T_{n-2} & w_{n-2}^T \\ 0 & w_{n-2} & 0 \end{bmatrix} \text{ for odd } n > 5 \text{ and } H_n = \begin{bmatrix} T_{n-1} & x_{n-1}^T \\ x_{n-1} & 0 \end{bmatrix} \text{ for even } n \geq 6. \tag{3}$$

First we prove the following result.

Lemma 10 *$G_n = H_n$ for all even $n \geq 6$.*

Proof : We first prove the following statement by induction on n.

$$T_n = \begin{bmatrix} 0 & \overline{v_{n-1}} \\ \hline v_{n-1}^T & N_{n-1} \end{bmatrix} \text{ for odd } n \geq 5 \text{ and } N_n = \begin{bmatrix} T_{n-1} & w_{n-1}^T \\ w_{n-1} & 0 \end{bmatrix} \text{ for even } n \geq 6. \tag{4}$$

It is easy to verify that $T_5 = \begin{bmatrix} 0 & \overline{v_4} \\ \hline v_4^T & N_4 \end{bmatrix}$ and $N_6 = \begin{bmatrix} T_5 & w_5^T \\ w_5 & 0 \end{bmatrix}$. Assume that (4) holds for $(n-1)$. By definition and using $\overline{v_{n-1}} = \overline{v_{n-2}}0$ we have that for odd $n \geq 7$, $\begin{bmatrix} 0 & \overline{v_{n-1}} \\ \hline v_{n-1}^T & N_{n-1} \end{bmatrix} = \begin{bmatrix} 0 & \overline{v_{n-2}} & 0 \\ \hline v_{n-2}^T & T_{n-2} & w_{n-2}^T \\ 0 & w_{n-2} & 0 \end{bmatrix} = T_n$. Similarly, by definition and using $1\overline{v_{n-2}} = \overline{w_{n-1}}0$ we have that for even $n \geq 8$, $\begin{bmatrix} T_{n-1} & w_{n-1}^T \\ w_{n-1}^T & 0 \end{bmatrix} = \begin{bmatrix} 0 & \overline{v_{n-2}} & 1 \\ \hline v_{n-2}^T & N_{n-2} & v_{n-2}^T \\ 1 & \overline{v_{n-2}} & 0 \end{bmatrix} = N_n$. This completes the proof of (4). Now to prove $G_n = H_n$ it is sufficient to show $T_{n-1} = \begin{bmatrix} 0 & \overline{v_{n-2}} \\ \hline v_{n-2}^T & N_{n-2} \end{bmatrix}$ and $x_{n-1}0 = 1\overline{u_{n-2}}0$. The first statement follows from (4) and the second statement follows from the definition of x_n. ■

Lemma 11 *For odd $n \geq 5$, the following statements hold for $\underline{T_n}$.*
(1) The first column of T_{n-2} is v_{n-2}^T and the second column is $\overline{v_{n-2}^T}$; (3) The $\lfloor \frac{n}{2} \rfloor$-th column and $(\lfloor \frac{n}{2} \rfloor+2)$-th column of T_n are identical; (3) The rank of T_n is $(n-1)$.

Proof : All three statements are proved using induction on odd $n \geq 5$. We only describe the proof for the third statement. For $n=5$ it is easy to verify that the rank of T_5 is 4. Assume that the rank of T_{n-2} is

$n-3$. Consider the matrix $A_n = \begin{bmatrix} \overline{v_{n-2}^T} & T_{n-2} \\ 0 & w_{n-2} \end{bmatrix}$. By the first statement of the lemma, the first and third columns of the matrix $[\overline{v_{n-2}^T} \;\; T_{n-2}]$ are identical. At the same time the first and third bits of the vector $0w_{n-2}$ are 0 and 1 respectively. So the last row of A_n is linearly independent of other rows. Hence the rank of A_n is $n-3+1=n-2$. Consequently, T_n has rank at least $n-2$. Again since T_n is a symplectic matrix, its rank must be even and hence must be $n-1$. ■

Now we are in a position to prove that G_n is of full rank.

Lemma 12 *The rank of G_n is n.*

Proof : Consider $G_n = H_n = \begin{bmatrix} T_{n-1} & x_{n-1}^T \\ x_{n-1} & 0 \end{bmatrix}$. Since the rank of T_{n-1} is $(n-2)$ the rank of H_n is at least $n-2$. Again from Lemma 11, the $\lfloor\frac{n-1}{2}\rfloor$-th column and the $(\lfloor\frac{n-1}{2}\rfloor+2)$-th column of T_{n-1} are identical. But the $\lfloor\frac{n-1}{2}\rfloor$-th and the $(\lfloor\frac{n-1}{2}\rfloor+2)$-th bits of x_{n-1} are 0 and 1 respectively. Hence x_{n-1} is linearly independent of T_{n-1}. Thus the rank of G_n is at least $n-2+1=n-1$. Again since G_n is a symplectic matrix its rank must be even and hence its rank is n. ■

Thus we have the following result which completes the proof of Theorem 6.

Lemma 13 *The S-box f defined in Theorem 6 is a perfect nonlinear S-box.*

Proof : Using Lemmas 8, 9 and 12, we know that F_n, G_n and $F_n \oplus G_n$ have full rank. Hence the Boolean functions f_{F_n}, f_{G_n} and $f_{F_n} \oplus f_{G_n} = f_{F_n \oplus G_n}$ are bent. Thus the function f defined in Theorem 6 is a perfect nonlinear function. ■

5 Relation With One Factorization of a Complete Graph

Let K_n be the complete graph with n vertices. For even $n \geq 2$, it is well known that K_n can be decomposed into $(n-1)$ edge disjoint, one-factors [1]. One such decomposition of K_n is described as follows. For even n and $1 \leq i \leq n-1$, define

$$\mathcal{F}_i^n \;=\; \{(n,i)\} \cup \{((n-2-j+i) \bmod (n-1)+1, (i+j-1) \bmod (n-1)+1) : 1 \leq j \leq \frac{n}{2}-1\} \quad (5)$$

The collection $\mathcal{T}_n = \{\mathcal{F}_1^n, \ldots, \mathcal{F}_{n-1}^n\}$ is a one factorization of K_n where the vertices are labeled by the integers $1,\ldots,n$. When n is clear from the context we will write $\mathcal{F}_i$ instead of $\mathcal{F}_i^n$. The elements of $\mathcal{T}_8$ (i.e. a one factorization of K_8) are given below.

$$\begin{array}{rclrcl}
\mathcal{F}_1 &=& \{(8,1),(7,2),(6,3),(5,4)\} & \mathcal{F}_2 &=& \{(8,2),(1,3),(7,4),(6,5)\} \\
\mathcal{F}_3 &=& \{(8,3),(2,4),(1,5),(7,6)\} & \mathcal{F}_4 &=& \{(8,4),(3,5),(2,6),(1,7)\} \\
\mathcal{F}_5 &=& \{(8,5),(4,6),(3,7),(2,1)\} & \mathcal{F}_6 &=& \{(8,6),(5,7),(4,1),(3,2)\} \\
\mathcal{F}_7 &=& \{(8,7),(6,1),(5,2),(4,3)\} & & &
\end{array}$$

In [7], one factorization of K_n was used as a tool for construction of S-boxes satisfying SAC. We point out the connection of the construction of Section 4 to the one factorization of K_n. This connection will be developed in later sections to obtain other constructions of perfect nonlinear S-boxes satisfying higher order SAC.

Suppose $\mathcal{S} \subseteq \mathcal{T}_n$. We use $\mathcal{S}$ to define a symplectic matrix $B_{\mathcal{S}}$ in the following manner: For $1 \leq k,l \leq n$, the entry $B_{\mathcal{S}}[k,l]=1$ if and only if either (k,l) or (l,k) is in $\mathcal{F}_i^n$ for some $\mathcal{F}_i^n \in \mathcal{S}$.

Theorem 14 *Let $n \geq 4$ be an even integer, $\mathcal{S}_1 = \{\mathcal{F}_2, \ldots, \mathcal{F}_{\frac{n}{2}}\}$ and $\mathcal{S}_2 = \mathcal{T} \setminus \mathcal{S}_1$. Let $B_{\mathcal{S}_1}$ and $B_{\mathcal{S}_2}$ be the symplectic matrices associated with $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. Then*
1. F_n is obtained from $B_{\mathcal{S}_1}$ by changing the zeros in positions $(\frac{n}{2}+1, \frac{n}{2})$ and $(\frac{n}{2}, \frac{n}{2}+1)$ to ones.
2. G_n is obtained from $B_{\mathcal{S}_2}$ by changing the zeros in positions $(\frac{n}{2}+1, \frac{n}{2}+2)$ and $(\frac{n}{2}+2, \frac{n}{2}+1)$ to ones.

Theorem 14 shows the relationship between one factorization and two output S-boxes of Section 4. This can be generalized to more than two output S-boxes. In fact, the earlier work of [7] provides such a generalization. However, there is one major difficulty with the generalization. It becomes very difficult to ensure that the resulting S-box is a perfect nonlinear S-box. Thus while the generalization of [7] ensures the SAC property, it results in functions with quite weak nonlinearity. On the other hand, our motivation is to obtain *perfect nonlinear* S-boxes satisfying higher order SAC. The rest of the paper is devoted to identifying other perfect nonlinear S-boxes satisfying higher order SAC.

5.1 Improvements for Two Output S-Boxes

We know from [7] that for an $(n, 2, k)$-SAC function, $k \leq \lfloor \frac{2(n-1)}{3} \rfloor - 1$. Thus the construction in Section 4 is suboptimal with respect to the SAC property. (However, it is optimal with respect to nonlinearity).

Here we provide some examples of two output S-boxes with higher order SAC. All these examples were obtained using experimental trial and error method. The constructions are based on the relationship between the symplectic matrices and one factorization described above. These examples are summarized in Table 1. The interpretation of the entries in Table 1 is as follows. Each row describes a construction for the particular value of n. The second column describes two subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathcal{T}_n$. Let $B_{\mathcal{S}_1}$ and $B_{\mathcal{S}_2}$ be the symplectic matrices associated with these two sets. We set $B_1 = B_{\mathcal{S}_1}$ and B_2 is $B_{\mathcal{S}_2}$ with the following modification: If (k, l) is in the third column, then $B_{\mathcal{S}_2}[k, l]$ and $B_{\mathcal{S}_2}[l, k]$ are changed from 0 to 1. The desired S-box $f : F_2^n \to F_2^2$ is given by $f(x) = (f_{B_1}(x), f_{B_2}(x))$. *Each of these S-boxes is a perfect nonlinear S-box.* The fourth column provides the order of SAC that is achieved by the corresponding S-box. The fifth column provides the maximum order of SAC that can be achieved by an $(n, 2)$ S-box. *In the situation where this maximum is equal to the achieved order of SAC, the construction provides optimal trade-off between the following parameters : nonlinearity, order of SAC, number of input variables, number of output variables. None of these parameters can be improved without changing some other parameter.*

Table 1: Improved and Optimal Constructions of Two Output S-boxes.

n	Description	Modification	k	max k
8	$\mathcal{S}_1 = \{\mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_5, \mathcal{F}_7\}$	–	3	3
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_4, \mathcal{F}_5, \mathcal{F}_6\}$	(5,6)		
10	$\mathcal{S}_1 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_4, \mathcal{F}_7, \mathcal{F}_8\}$	–	4	5
	$\mathcal{S}_2 = \{\mathcal{F}_3, \mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_7, \mathcal{F}_8\}$	(6,9)		
12	$\mathcal{S}_1 = \{\mathcal{F}_1, \mathcal{F}_3, \mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_{11}\}$	–	6	6
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_4, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_9, \mathcal{F}_{10}\}$	(2,7)		
14	$\mathcal{S}_1 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_9, \mathcal{F}_{10}, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}\}$	–	7	7
	$\mathcal{S}_2 = \{\mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_9, \mathcal{F}_{10}, \mathcal{F}_{11}, \mathcal{F}_{12}\}$	(8,9)		
16	$\mathcal{S}_1 = \{\mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_9, \mathcal{F}_{15}\}$	–	8	9
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_9, \mathcal{F}_{10}, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}, \mathcal{F}_{14}\}$	(3,9)		

6 Construction of $(n, 3, k)$ S-boxes

We describe constructions of $(n, 3, k)$ perfect nonlinear S-boxes. These constructions were obtained by experimental trial and error methods. Some of the constructions seem to have a general pattern, though it has not been possible to prove a general result. There are several cases in the construction though the description of the constructions in all the cases is similar. We first identify three subsets $\mathcal{S}_1, \mathcal{S}_2$ and $\mathcal{S}_3$ of $\mathcal{T}_n$.

These three subsets define three symplectic matrices $B_{\mathcal{S}_1}, B_{\mathcal{S}_2}$ and $B_{\mathcal{S}_3}$. These matrices are then modified by changing a number of zeros to ones to obtain three other symplectic matrices B_1, B_2 and B_3. The positions where the changes are to be made are given by the third column. If (k,l) is in the third column, then $B_{\mathcal{S}_j}[k,l]$ and $B_{\mathcal{S}_j}[k,l]$ $(1 \leq j \leq 3)$ are changed from 0 to 1. The required $(n,3)$ S-box $f : F_2^n \to F_2^3$ is obtained from these three matrices in the following manner: $f(x) = (f_{B_1}(x), f_{B_2}(x), f_{B_3}(x))$. There are three cases.

1. Table 2 describes several cases of constructions for $n \equiv 0 \bmod 8$. For $n > 8$, there is a general heuristic which provides the required construction. For $n = 8$, a special construction is required.
2. Table 3 describes constructions for $n \equiv 4 \bmod 8$. These constructions have a general pattern.
3. Table 4 describes several constructions for $n \equiv 2 \bmod 4$. There does not appear to be any general pattern for these constructions.

The constructions for $n = 10, 22$ *provide optimal trade-off between the following parameters: numbers of input and output variables, nonlinearity and the order of SAC.* Further, for $n = 12, 16, 20$ and 24 the achieved value of k is only one less than the upper bound on k.

Table 2: Constructions for $n \equiv 0 \bmod 8$.

n	Description	Modification	k	$\max k$
8	$\mathcal{S}_1 = \{\mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_7\}$	(4,5)	1	2
	$\mathcal{S}_2 = \{\mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_5\}$	–		
	$\mathcal{S}_3 = \{\mathcal{F}_1, \mathcal{F}_3, \mathcal{F}_6\}$	(4,7)		
16,24,32	$\mathcal{S}_1 = \{\mathcal{F}_2, \mathcal{F}_3, \ldots, \mathcal{F}_{\frac{n}{2}-1}, \mathcal{F}_{n-1}\}$	$(\frac{n}{2}, \frac{n}{2}+1)$	$\frac{n}{2}-2$	min(
	$\mathcal{S}_2 = \{\mathcal{F}_{\frac{n}{4}+1}, \ldots, \mathcal{F}_{\frac{3n}{4}-1}\}$	$(\frac{n}{2}, \frac{n}{4}+1), (\frac{n}{2}, \frac{3n}{4}+1)$		$\lfloor\frac{4(n-1)}{7}\rfloor - 1$,
	$\mathcal{S}_3 = \{\mathcal{F}_1, \mathcal{F}_{\frac{n}{4}+1}, \ldots, \mathcal{F}_{\frac{n}{2}-1}, \mathcal{F}_{\frac{3n}{4}}, \ldots, \mathcal{F}_{n-2}\}$	$(\frac{n}{2}, \frac{3n}{4})$		$2\lfloor\frac{2n}{7}\rfloor - 1$)

Table 3: Constructions for $n \equiv 4 \bmod 8$.

n	Description	Modification	k	$\max k$
12,20,28	$\mathcal{S}_1 = \{\mathcal{F}_2, \mathcal{F}_3, \ldots, \mathcal{F}_{\frac{n}{2}-2}, \mathcal{F}_{n-1}\}$	$(\frac{n}{2}, \frac{n}{2}+1)$	$\frac{n}{2}-2$	min(
	$\mathcal{S}_2 = \{\mathcal{F}_{\frac{n}{4}+1}, \ldots, \mathcal{F}_{\frac{3n}{4}-1}\}$	$(\frac{n}{2}, \frac{n}{4}+2)$		$\lfloor\frac{4(n-1)}{7}\rfloor - 1$,
	$\mathcal{S}_3 = \{\mathcal{F}_1, \mathcal{F}_{\frac{n}{4}+1}, \ldots, \mathcal{F}_{\frac{n}{2}-1}, \mathcal{F}_{\frac{3n}{4}}, \ldots, \mathcal{F}_{n-2}\}$	$(\frac{n}{2}, \frac{3n}{4}+1)$		$2\lfloor\frac{2n}{7}\rfloor - 1$)

7 Maximally Nonlinear Functions

The constructions described so far hold when the number of input bits n is even. In the case n is odd, there does not exist any perfect nonlinear S-boxes. The best nonlinearity achieved by an (n,m) S-box with $m > 1$ is $2^{n-1} - 2^{(n-1)/2}$ and S-boxes achieving this value of nonlinearity are called maximally nonlinear. In this section, we describe a simple modification of the previously described constructions which provide maximally nonlinear S-boxes.

Theorem 15 *Let f be a $(2r, m, k)$ perfect nonlinear quadratic S-box where the symplectic matrices associated with the component functions are $B_1, \ldots, B_m$. For $1 \leq i \leq m$, let $B_i^{'}$ be obtained from B_i by deleting the first row and column. Then the S-box $f^{'} : F_2^{2r-1} \to F_2^m$ defined by $f^{'}(x) = (f_{B_1^{'}}(x), \ldots, f_{B_m^{'}}(x))$ is a $(2r-1, m, k-1)$ maximally nonlinear quadratic S-box.*

Table 4: Constructions for $n \equiv 2 \bmod 4$.

n	Description	Modification	k	$\max k$
10	$\mathcal{S}_1 = \{\mathcal{F}_3, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_9\}$	(6,9)	3	3
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_4, \mathcal{F}_7, \mathcal{F}_8\}$	–		
	$\mathcal{S}_3 = \{\mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_7, \mathcal{F}_8\}$	(5,6)		
14	$\mathcal{S}_1 = \{\mathcal{F}_1, \mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}\}$	(1,6)	4	6
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_6, \mathcal{F}_9, \mathcal{F}_{10}, \mathcal{F}_{11}, \mathcal{F}_{12}\}$	–		
	$\mathcal{S}_3 = \{\mathcal{F}_1, \mathcal{F}_5, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_{11}, \mathcal{F}_{12}\}$	(1,9)		
18	$\mathcal{S}_1 = \{\mathcal{F}_3, \mathcal{F}_9, \mathcal{F}_{10}, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}, \mathcal{F}_{14}, \mathcal{F}_{17}\}$	(5,10)	7	8
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_4, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}, \mathcal{F}_{14}, \mathcal{F}_{15}, \mathcal{F}_{16}\}$	–		
	$\mathcal{S}_3 = \{\mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_7, \mathcal{F}_8, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}, \mathcal{F}_{14}\}$	(9,11)		
22	$\mathcal{S}_1 = \{\mathcal{F}_1, \mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_5, \mathcal{F}_6, \mathcal{F}_{13}, \mathcal{F}_{14}, \mathcal{F}_{15}, \mathcal{F}_{16}, \mathcal{F}_{17}, \mathcal{F}_{21}\}$	(1,5)	9	9
	$\mathcal{S}_2 = \{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_8, \mathcal{F}_{13}, \mathcal{F}_{14}, \mathcal{F}_{15}, \mathcal{F}_{16}, \mathcal{F}_{17}, \mathcal{F}_{18}, \mathcal{F}_{19}, \mathcal{F}_{20}\}$	–		
	$\mathcal{S}_3 = \{\mathcal{F}_7, \mathcal{F}_9, \mathcal{F}_{10}, \mathcal{F}_{11}, \mathcal{F}_{12}, \mathcal{F}_{13}, \mathcal{F}_{14}, \mathcal{F}_{15}, \mathcal{F}_{16}, \mathcal{F}_{17}\}$	(1,16)		

Proof : There are two things to be proved – the nonlinearity and the order of SAC. Since f is a perfect nonlinear S-box, each nonzero linear combination of the B_i's have full rank (see Lemma 4). Dropping one row and one column decreases the rank by two. Hence the rank of any nonzero linear combination of the $B_i^{'}$'s is $2r-2$ and the nonlinearity of the corresponding Boolean function is $2^{2r-2} - 2^{r-1}$. Now using Lemma 4 we have $f^{'}$ to be a maximally nonlinear S-box.

Further, since f satisfies SAC(k), the number of ones in any nonzero linear combination of the B_i's is at least $k+1$. Dropping one row and one column decreases the number of ones in any row (or column) by at most one. Again using Lemma 3, it follows that the S-box $f^{'}$ satisfies SAC($k-1$). ■

8 Improving Algebraic degree

The constructions described in the previous sections provide quadratic functions. In this section, we describe a method of improving the degree of the constructed functions with a small trade-off in the nonlinearity and the SAC property. We first need to relax the notion of SAC. (See [5] for the notion of almost PC(l) of order k functions.)

Definition 16 *An n-variable Boolean function f is said to be (ϵ, k)-SAC if the following property holds: Let g be an $(n-i)$-variable Boolean function obtained from f by fixing at most $i \leq k$ input variables to constants. Then $\left|\frac{|g(x) \oplus g(x \oplus \alpha)|}{2^{n-i}} - \frac{1}{2}\right| \leq \epsilon$ for any α of weight 1. An (n,m) S-box is said to be (n,m,ϵ,k)-SAC if every nonzero linear combination of the component functions is an (ϵ,k)-SAC function.*

The next result shows how to convert an (n,m,k) S-box into an (n,m,ϵ,k) S-box for a small ϵ and with a small change in nonlinearity.

Theorem 17 *Let $f = (f_1, \ldots, f_m)$ be an (n,m,k) S-box where the degree of any f_i is less than $(n-1)$. Then it is possible to construct an (n,m,ϵ,k) S-box g with algebraic degree $n-1$, $\mathsf{nl}(g) \geq \mathsf{nl}(f) - (m+1)$ and $\epsilon = \frac{m+1}{2^{n-k-1}}$.*

Proof : Let the component functions of f be $f_1, \ldots, f_m$. We construct an (n,m) S-box g with component functions $g_1, g_2, \ldots, g_m$ in the following manner. For $1 \leq i \leq m$, define $g_i(x_1, \cdots, x_n) = f_i(x_1, \cdots, x_n) \oplus x_1 \ldots x_{i-1} x_{i+1} \ldots x_n$. By construction, the algebraic degree of any g_i is $n-1$. Further, the degree $(n-1)$

terms in the g_i's are distinct. Hence any nonzero linear combination of the g_i's also have degree $(n-1)$. Thus the degree of g is $(n-1)$.

We now prove the nonlinearity. The term $x_1 \ldots x_{i-1}x_{i+1} \ldots x_n$ which is XORed to f_i to obtain g_i changes exactly two output values of f_i. Thus $\mathsf{nl}(g_i) = \mathsf{nl}(f_i) - 2$. Further, the inputs for which the outputs are changed are the all one vector and the vector with a zero only in the ith position. Thus if h (resp. h') is a linear combination of i of the g_i's (resp. f_i's), then h and h' differ in at most $(i+1)$ positions. Since $1 \le i \le m$, we have $\mathsf{nl}(g) \ge \mathsf{nl}(f) - (m+1)$.

Now suppose that $h_1(x)$ (resp. $h_1'(x)$) is obtained from $h(x)$ (resp. $h'(x)$) by fixing at most j $(1 \le j \le k)$ input bits to constant values. Since $h(x)$ and $h'(x)$ differ in exactly $(i+1)$ positions, it follows that $h_1(x)$ and $h_1'(x)$ differ in at most $(i+1)$ positions. Further, since $h_1(x)$ and $h_1'(x)$ differ in at most $(i+1)$ positions, so does $h_1(x \oplus \alpha)$ and $h_1{}'(x \oplus \alpha)$. Let $\mu(x) = h(x) \oplus h(x \oplus \alpha)$ and $\mu' = h'(x) \oplus h'(x \oplus \alpha)$. Then it follows that $\mu(x)$ and $\mu'(x)$ differ in at most $2(i+1)$ positions. Since f satisfies SAC(k), it follows that $\mu'(x)$ is balanced and has weight 2^{n-j-1}. Also since $1 \le i \le m$ and $1 \le j \le k$, we obtain $\left|\frac{|\mu(x)|}{2^{n-j}} - \frac{1}{2}\right| = \left|\frac{|\mu(x)|}{2^{n-j}} - \frac{|\mu'(x)|}{2^{n-j}}\right| = \left|\frac{|\mu(x)|-|\mu'(x)|}{2^{n-j}}\right| \le \frac{2(i+1)}{2^{n-j}} \le \frac{m+1}{2^{n-k-1}}$. This completes the proof. ■

Table 5 provides some examples to illustrate Theorem 17. The interpretation of Table 5 is as follows.

Table 5: Values of k, ϵ and nonlinearity for 2 and 3 output S-boxes for different values of n (see Theorem 17).

n	degree	$m=2$	$m=3$
8	7	(3, 0.1875, 118)	(1, 0.0625, 116)
9	8	(3, 0.0938, 238)	(2, 0.0625, 236)
10	9	(4, 0.0938, 494)	(3, 0.0625, 492)
11	10	(5, 0.0938, 990)	(3, 0.0313, 988)
12	11	(6, 0.0938, 2014)	(4, 0.0313, 2012)

Each entry is of the form (k, ϵ, x), where k is the order of SAC, ϵ is defined in Theorem 17 and x is the nonlinearity of the modified function. (When m is even, the value of nonlinearity is one more than the lower bound given in Theorem 17.) Note that in each case the algebraic degree is $n-1$. The drop in nonlinearity is very small; for example for $n = 8$, the lower bound from Theorem 17 is 117 while the maximum possible nonlinearity is 120. Similarly, in each of the above cases, the value of ϵ is small. Hence the deviation from perfect nonlinearity and the (perfect) SAC property is small. On the other hand, the degree increases to the maximum possible. Thus such S-boxes are amply suited for use in the design of practical block cipher algorithms.

9 Conclusion

In this paper, we have considered the problem of constructing perfect nonlinear S-boxes satisfying higher order SAC. Previous work in this area [7] also provided constructions of S-boxes satisfying higher order SAC. However, the nonlinearity obtained was lower. To the best of our knowledge, we provide the first examples of S-boxes satisfying higher order SAC *and* perfect nonlinearity. Some of the constructed S-boxes also achieve optimal trade-off between the numbers of input and output variables, nonlinearity and the order of SAC. Our construction uses bilinear forms and symplectic matrices and yields quadratic functions. We show that the degree can be significantly improved by a small sacrifice in nonlinearity and the SAC property. This yields S-boxes which are well suited for use in design of block ciphers. Lastly, we would like to remark that more research is necessary to generalize our construction using symplectic matrices to more than 3 outputs and also to obtain direct constructions of higher degree S-boxes which satisfy higher order SAC and perfect nonlinearity.

References

[1] J. A. Bondy and U. S. R. Murthy. *Graph Theory with Applications.* London, Macmillan Press, 1977.

[2] A. Canteaut and M. Videau. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. *Advances in Cryptology – Eurocrypt 2002*, LNCS 2332, pages 518-533.

[3] C. Carlet. On cryptographic propagation criteria for Boolean functions. *Information and Computation*, 151:32–56, 1999.

[4] A. Canteaut, C. Carlet, P. Charpin and C. Fontaine. Propagation Characteristics and Correlation-Immunity of Highly Nonlinear Boolean Functions. In *Advances in Cryptology – Eurocrypt 2000*, pages 507–522, Lecture Notes in Computer Science, Springer-Verlag, 2000.

[5] K. Kurosawa. Almost security of cryptographic Boolean functions. Cryptology e-print archive, http://eprint.iacr.org/2003/075.

[6] K. Kurosawa and T. Satoh. Design of $SAC/PC(l)$ of order k Boolean functions and three other cryptographic criteria. In Advances in Cryptology - Eurocrypt'97, number 1233 in Lecture Notes in Computer Science Series, Springer-Verlag, 434-449, 1997.

[7] K. Kurosawa and T. Satoh. Generalization of Higher Order SAC to Vector Output Boolean Functions. In *Advances in Cryptology - Asiacrypt 1996*, Lecture Notes in Computer Science, Springer-Verlag, 1996.

[8] F. J. MacWillams and N. J. A. Sloane. *The Theory of Error Correcting Codes.* North Holland, 1977.

[9] K. Nyberg. Perfect Nonlinear S-boxes. In *Advances in Cryptology - EUROCRYPT 1991*, pages 378–386, Lecture Notes in Computer Science, Springer-Verlag, 1991.

[10] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts and J. Vandewalle. Propagation Characteristics of Boolean Functions. In *Advances in Cryptology - EUROCRYPT 1990*, pages 161–173, Lecture Notes in Computer Science, Springer-Verlag, 1991.

[11] O. S. Rothaus. On bent functions. *Journal of Combinatorial Theory, Series A*, 20:300–305, 1976.

[12] P. Sarkar and S. Maitra. Construction of Nonlinear Boolean Functions with Important Cryptographic Properties. *In Advances in Cryptology - EUROCRYPT 2000*, pages 485–506, Lecture Notes in Computer Science, Springer-Verlag, 2000.

[13] A. F. Webster and S. E. Tavares. On the Design of S-boxes. In *Advances in Cryptology - Crypto 1985*, pages 523–534, Lecture Notes in Computer Science, Springer-Verlag, 1986.