Construction of Perfect Nonlinear and Maximally Nonlinear Multi-Output Boolean Functions 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 multi-output Boolean functions satisfying higher order strict avalanche criteria (SAC). Our first construction is an infinite family of 2-output 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 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 have possible applications in the design of block ciphers.

Keywords: S-box, SAC, bent function, bilinear form, symplectic matrix, nonlinearity, symmetric ciphers.

1 Introduction

A Boolean function is a map from \(\{0,1\}^n\) to \(\{0,1\}\) and by a multi-output Boolean function we mean a map from \(\{0,1\}^n\) to \(\{0,1\}^m\). Multi-output Boolean functions are usually called S-boxes and are used as basic primitives for designing 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 \(n=m=8\). We next describe some properties of S-boxes which have been studied previously.

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. For even \(n\), functions achieving the maximum possible nonlinearity are called perfect nonlinear S-boxes [9]. If \(m=1\), such functions are called bent functions [11]. For odd \(n\) and \(m>1\), functions achieving the maximum possible nonlinearity are called maximally nonlinear functions.

The concept of propagation characteristic was introduced in the cryptology literature 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 $\alpha$ be a non zero vector of weight at most $l$. 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 $k$ functions have 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-Muller 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 have possible applications in the design of symmetric 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 $\text{wt}(u)$. 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
2.1 Boolean Functions

An $n$-variable Boolean function is a map $f : F_2^n \rightarrow F_2$. The weight of $f$, denoted by $\text{wt}(f)$ is defined as $\text{wt}(f) = \{ x : f(x) = 1 \}$. The function $f$ is said to be balanced if $\text{wt}(f) = 2^{n-1}$. 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 $nl(f)$ of an $n$-variable Boolean function is defined as $nl(f) = \min_{l \in A_n} d(f, l)$. 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 almost optimally 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 $\text{wt}(\alpha) = 1$ [13]. A function $f$ satisfies SAC($k$) if every subfunction obtained from $f(x_1, \ldots, x_n)$ 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 $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$. Let $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$ be an S-box and $g : \{0, 1\}^m \rightarrow \{0, 1\}$ 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 $nl(f) = \min\{nl(l \circ f) : l$ 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 $m = n$, 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. For odd $n$ and $1 < m < n$, the maximum possible achievable nonlinearity is an open problem. However, for odd $n$, $1 < m < n$, and quadratic functions the maximum possible achievable nonlinearity is $2^{n-1} - 2^{(n-1)/2}$. We will also call such functions to be maximally nonlinear.

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 $\leq 2$ 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, $L = (l_1, \ldots, l_n)$ 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 matrix $B$ 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 $\text{wt}(r^{(i)}) \geq k + 1$, 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 $f(x) = xQx^T \oplus La^T \oplus b$. Let $\alpha$ be such that only the $i^{th}$ component of $\alpha$ is 1 and all other components are zero. Further, let the $i^{th}$ column of $Q$ be $a^{(i)}$ and the $i^{th}$ row of $Q$ be $b^{(i)}$. Then $r^{(i)} = (a^{(i)})^T \oplus b^{(i)}$. We have

$$f(x) \oplus f(x \oplus \alpha) = xQx^T \oplus (x \oplus \alpha)Q(x \oplus \alpha)^T \oplus La^T \oplus \alpha Qa^T$$

$$= xQa^T \oplus \alpha Qx^T \oplus La^T \oplus \alpha Qa^T$$

$$= (b^{(i)} \oplus (a^{(i)})^T, x) \oplus La^T \oplus \alpha Qa^T$$

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

Let $f = (f_1, \ldots, f_m)$ 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 quadratic component functions $f_i$ and associated symplectic forms $B_i$ for $1 \leq i \leq m$. 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 quadratic component functions $f_i$ and associated symplectic forms $B_i$ for $1 \leq i \leq m$. 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 has 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 has rank $(n - 1)$. 

4
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 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 \(n \geq 6\), we define two sequences of \(n \times n\) matrices and show that these matrices are the symplectic matrices required in the construction. For the rest of this paper, we will use the following notation.

- For each \(n \geq 1\), define \(v_n\) to be a string of length \(n\) which is the alternating sequence of 0’s and 1’s starting with a 0. For example, \(v_4 = 0101\) and \(v_5 = 01010\). Define \(w_n = 1v_{n-1}\).
- For each even \(n \geq 2\), define \(u_n\) as \(u_n = \frac{1}{(n/2)!} 10\ldots 0\). For odd \(n \geq 3\), define \(x_n = 1u_{n-1}\).

Define \(M_4 = [0010, 0010, 1101, 0010]^T\) and \(N_4 = [0101, 1011, 0101, 1111]^T\). Further, for even \(n > 4\) define

\[
M_n = \begin{bmatrix}
0 & v_{n-2}^T & 0 \\
v_{n-2}^T & M_{n-2} & v_{n-2}^T \\
0 & v_{n-2} & 0
\end{bmatrix}, \quad F_n = \begin{bmatrix}
0 & v_{n-2}^T & 0 \\
v_{n-2}^T & M_{n-2} & u_{n-2}^T \\
0 & u_{n-2} & 0
\end{bmatrix};
\]

\[
N_n = \begin{bmatrix}
0 & v_{n-2}^T & 1 \\
v_{n-2}^T & N_{n-2} & v_{n-2}^T \\
1 & v_{n-2} & 0
\end{bmatrix}, \quad G_n = \begin{bmatrix}
0 & v_{n-2}^T & 1 \\
v_{n-2}^T & N_{n-2} & u_{n-2}^T \\
1 & u_{n-2} & 0
\end{bmatrix}.
\]

The following result is easy to prove by induction on even \(n \geq 6\).

**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 \(n \geq 6\) be an even integer. The S-box \(f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n\) 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 \(n \geq 4\).

\[
\begin{align*}
\text{wt}(r_j) &\geq \frac{n}{2} - 1 \quad \text{if } 1 \leq j \leq \frac{n}{2} \quad \text{and } \ j \text{ is odd}; \\
\text{wt}(r_j) &\geq \frac{n}{2} - 2 \quad \text{if } 1 \leq j \leq \frac{n}{2} \quad \text{and } \ j \text{ is even}; \\
\text{wt}(r_j) &\geq \frac{n}{2} \quad \text{if } \frac{n}{2} + 1 \leq j \leq n \quad \text{and } \ j \text{ is odd}; \\
\text{wt}(r_j) &\geq \frac{n}{2} - 1 \quad \text{if } \frac{n}{2} + 1 \leq j \leq n \quad \text{and } \ j \text{ is even}.
\end{align*}
\]

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 \(j\)th 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 \( wt(s_j) = wt(v_{n-2}) = \frac{n-2}{2} = \frac{n}{2} - 1 \).

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

Case 2: \( \frac{n}{2} \leq j \leq \frac{3n}{4} \) and \( j \) even: In this case \( wt(s_j) = 1 + wt(r'_j) \geq 1 + \frac{n-2}{2} - 1 = \frac{n}{2} - 1 \).

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

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 \( wt(s_j) = wt(r'_j) \geq \frac{n-2}{2} = \frac{n}{2} - 1 \).

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

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

\[
F_n \oplus G_n = \begin{bmatrix}
0 & J_{n-2} \\
J'_{n-2} & M_{n-2} \oplus N_{n-2} & 1 \\
1 & J_{n-2} & 0
\end{bmatrix},
\]

where \( J_n \) is all 1 vector. From this it is simple to verify 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

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

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 \( v_{n-20} \) is 1 and the \( (\frac{n}{2} + 2) \)-th bit is 0. Hence the last row \( 0v_{n-20} \) 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

\[
F_n \oplus G_n = M_n \oplus N_n = \begin{bmatrix}
0 & J_{n-2} & 1 \\
J'_{n-2} & M_{n-2} \oplus N_{n-2} & J_{n-2}^T \\
1 & J_{n-2} & 0
\end{bmatrix},
\]

where \( J_n \) is the all 1 vector. The row \( 1J_{n-2} \) 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 \( T_5 = \begin{bmatrix} 01010, 10101, 01011, 10101, 01110 \end{bmatrix}^T \),

\[
T_n = \begin{bmatrix} 0 & v_{n-2} & 0 \\ v_{n-2} & T_{n-2} & w_{n-2}^T \\ 0 & w_{n-2} & 0 \end{bmatrix}
\]

for odd \( n > 5 \) and \( H_n = \begin{bmatrix} T_{n-1} & x_{n-1}^T \\ x_{n-1} & 0 \end{bmatrix} \) for even \( n \geq 6 \). (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 \( n \).

\[
T_n = \begin{bmatrix} 0 & v_{n-1} \\ v_{n-1} & N_{n-1} \end{bmatrix}
\]

for odd \( n \geq 5 \) and \( N_n = \begin{bmatrix} T_{n-1} & w_{n-1}^T \\ w_{n-1} & 0 \end{bmatrix} \) for even \( n \geq 6 \). (4)

It is easy to verify that \( T_5 = \begin{bmatrix} 0 & v_{4} \\ v_{4} & N_4 \end{bmatrix} \) 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 \( v_{n-1} = v_{n-2}v_0 \) we have that for odd \( n \geq 7 \),

\[
\begin{bmatrix} 0 & v_{n-1} \\ v_{n-1} & N_{n-1} \end{bmatrix} = \begin{bmatrix} 0 & v_{n-2} & 0 \\ v_{n-2} & T_{n-2} & w_{n-2}^T \\ 0 & w_{n-2} & 0 \end{bmatrix}
\]

for even \( n \geq 8 \),

\[
\begin{bmatrix} T_{n-1} & w_{n-1}^T \\ w_{n-1} & 0 \end{bmatrix} = \begin{bmatrix} 0 & v_{n-2} & 1 \\ v_{n-2} & N_{n-2} & v_{n-2}^T \\ 1 & v_{n-2} & 0 \end{bmatrix}
\]

This completes the proof of (4). Now to prove \( G_n = H_n \) it is sufficient to show \( T_{n-1} = \begin{bmatrix} 0 & v_{n-2} \\ v_{n-2} & N_{n-2} \end{bmatrix} \) and \( x_{n-1} = 1v_{n-2} \). 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 \( T_n \).

(1) The first column of \( T_{n-2} \) is \( v_{n-2}^T \) and the second column is \( v_{n-2}^T \). (2) The \( \left[ \frac{n}{2} \right] \)-th column and \( \left( \left[ \frac{n}{2} \right] + 2 \right) \)-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 \( A_n = \begin{bmatrix} v_{n-2}^T & T_{n-2} \\ 0 & w_{n-2} \end{bmatrix} \). By the first statement of the lemma, the first and third columns of the matrix \( \begin{bmatrix} v_{n-2}^T & T_{n-2} \\ 0 & w_{n-2} \end{bmatrix} \) 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 \( \left[ \frac{n-1}{2} \right] \)-th column and the \( \left( \left[ \frac{n-1}{2} \right] + 2 \right) \)-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 $f_{F_n} \oplus f_{G_n}$ are bent. Thus the function $f$ defined in Theorem 6 is a perfect nonlinear function.

## 5 Relation With One Factorization of a Complete Graph

A one-factor of a graph $G$ is a one-regular spanning subgraph of $G$. A one-factorization of $G$ is a partition of the edges of $G$ into one-factors.

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

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

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.

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

In [7], a 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 $\left(\frac{n}{2} + 1, \frac{n}{2}\right)$ and $\left(\frac{n}{2}, \frac{n}{2} + 1\right)$ to ones.
2. $G_n$ is obtained from $B_{\mathcal{S}_2}$ by changing the zeros in positions $\left(\frac{n}{2} + 1, \frac{n}{2} + 2\right)$ and $\left(\frac{n}{2} + 2, \frac{n}{2} + 1\right)$ 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 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 \(S_1\) and \(S_2\) of \(\mathcal{T}_n\). Let \(B_{S_1}\) and \(B_{S_2}\) be the symplectic matrices associated with these two sets. We set \(B_1 = B_{S_1}\) and \(B_2 = B_{S_2}\) with the following modification: if \((k, l)\) is in the third column, then \(B_{S_1}[k, l]\) and \(B_{S_2}[k, l]\) are changed from 0 to 1. The desired S-box \(f : F^n_2 \rightarrow F^n_2\) 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 among 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.

<table>
<thead>
<tr>
<th>(n)</th>
<th>Description</th>
<th>Modification</th>
<th>(k)</th>
<th>(\text{max } k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>(S_1 = {F_1, F_3, F_4, F_5, F_7})</td>
<td>(-)</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(S_2 = {F_1, F_4, F_5, F_6})</td>
<td>( (5, 6))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>(S_1 = {F_3, F_6, F_7, F_8})</td>
<td>(-)</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(S_2 = {F_1, F_2, F_4, F_7, F_8})</td>
<td>((6, 9))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>(S_1 = {F_5, F_6, F_7, F_8, F_9, F_{11}})</td>
<td>(-)</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(S_2 = {F_1, F_2, F_4, F_7, F_8, F_9, F_{10}})</td>
<td>((2, 7))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>(S_1 = {F_5, F_6, F_7, F_9, F_{10}, F_{11}, F_{12}, F_{13}})</td>
<td>(-)</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>(S_2 = {F_5, F_6, F_7, F_8, F_9, F_{10}, F_{11}, F_{12}})</td>
<td>((8, 9))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>(S_1 = {F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9, F_{15}})</td>
<td>(-)</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>(S_2 = {F_1, F_7, F_8, F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{15}})</td>
<td>((3, 9))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 \(S_1, S_2\) and \(S_3\) of \(\mathcal{T}_n\). These three subsets define three symplectic matrices \(B_{S_1}, B_{S_2}\) and \(B_{S_3}\). 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 \(B_{S_j}[k, l]\) and \(B_{S_j}[k, l]\) \((1 \leq j \leq 3)\) are changed from 0 to 1. The required \((n, 3)\) S-box \(f : F^n_2 \rightarrow F^n_2\) 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 \(n \equiv 0 \mod 8\). 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 \( n \equiv 4 \mod 8 \). These constructions have a general pattern.
3. Table 4 describes several constructions for \( n \equiv 2 \mod 4 \). 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>
<thead>
<tr>
<th>( n )</th>
<th>Description</th>
<th>Modification</th>
<th>( k )</th>
<th>max ( k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>( S_1 = { F_2, F_3, F_7 } )</td>
<td>(4,5)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>( S_2 = { F_3, F_4, F_5 } )</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>( S_3 = { F_1, F_3, F_6 } )</td>
<td>(4,7)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16, 24, 32</td>
<td>( S_1 = { F_2, F_3, \ldots, F_{\frac{n-1}{2}}, F_{n-1} } )</td>
<td>( \left( \frac{n}{2}, \frac{n}{2} + 1 \right) )</td>
<td>( \frac{n}{2} - 2 )</td>
<td>( \min(\frac{4(n-1)}{7} - 1, 2\lfloor \frac{3n}{4} \rfloor - 1) )</td>
</tr>
<tr>
<td></td>
<td>( S_2 = { F_{\frac{n+1}{4}}, \ldots, F_{\frac{3n}{4} - 1} } )</td>
<td>( \left( \frac{n}{2}, \frac{n}{2} + 1, \frac{3n}{4} + 1 \right) )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>( S_3 = { F_1, F_{\frac{n+1}{4}}, \ldots, F_{\frac{3n}{4} - 1}, F_{\frac{3n}{4}}, \ldots, F_{n-2} } )</td>
<td>( \left( \frac{n}{2}, \frac{3n}{4} \right) )</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( n )</th>
<th>Description</th>
<th>Modification</th>
<th>( k )</th>
<th>max ( k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>12, 20, 28</td>
<td>( S_1 = { F_2, F_3, \ldots, F_{\frac{n-1}{2}}, F_{n-1} } )</td>
<td>( \left( \frac{n}{2}, \frac{n}{2} + 1 \right) )</td>
<td>( \frac{n}{2} - 2 )</td>
<td>( \min(\frac{4(n-1)}{7} - 1, 2\lfloor \frac{3n}{4} \rfloor - 1) )</td>
</tr>
<tr>
<td></td>
<td>( S_2 = { F_{\frac{n+1}{4}}, \ldots, F_{\frac{3n}{4} - 1} } )</td>
<td>( \left( \frac{n}{2}, \frac{n}{2} + 2 \right) )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>( S_3 = { F_1, F_{\frac{n+1}{4}}, \ldots, F_{\frac{3n}{4} - 1}, F_{\frac{3n}{4}}, \ldots, F_{n-2} } )</td>
<td>( \left( \frac{n}{2}, \frac{3n}{4} + 1 \right) )</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### 7 Maximally Nonlinear Functions

The constructions described so far hold when the number of input bits \( n \) is even. In case \( n \) is odd, there do not exist any perfect nonlinear S-boxes. The best nonlinearity achieved by an \((n, m)\) quadratic 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 \( f' : F_2^{2r-1} \rightarrow F_2^m \) 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.

**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 has full rank (see Lemma 4). Dropping one row and one column decreases the rank by two for symplectic matrices. 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 that \( f' \) is a maximally nonlinear S-box.
Table 4: Constructions for $n \equiv 2 \mod 4$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>Description</th>
<th>Modification</th>
<th>$k$</th>
<th>max $k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>$S_1 = {F_3, F_7, F_8, F_9}$</td>
<td>(6,9)</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>$S_2 = {F_1, F_2, F_4, F_7, F_8}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$S_3 = {F_5, F_6, F_7, F_8}$</td>
<td>(5,6)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>$S_1 = {F_{11}, F_{13}, F_{14}, F_{15}, F_{16}}$</td>
<td>(1,6)</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>$S_2 = {F_1, F_2, F_{10}, F_{11}, F_{12}}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$S_3 = {F_5, F_6, F_7, F_{11}, F_{12}}$</td>
<td>(1,9)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>$S_1 = {F_3, F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, F_{16}}$</td>
<td>(5,10)</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>$S_2 = {F_1, F_2, F_{13}, F_{15}, F_{16}}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$S_3 = {F_5, F_6, F_{11}, F_{12}, F_{13}, F_{14}}$</td>
<td>(9,11)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>$S_1 = {F_{11}, F_3, F_4, F_5, F_6, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}, F_{16}, F_{17}, F_{21}}$</td>
<td>(1,5)</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>$S_2 = {F_1, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}, F_{16}, F_{17}, F_{18}, F_{19}, F_{20}}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$S_3 = {F_7, F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}, F_{16}, F_{17}}$</td>
<td>(1,16)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

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

The next result shows how to convert an $(n, m, k)$ S-box into an $(n, m, \epsilon, k)$ S-box for a small $\epsilon$ 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 $(n, m, \epsilon, k)$ S-box $g$ with algebraic degree $n - 1$, $\epsilon = \frac{m + 1}{2^{m-n-k-1}}$ and $\text{nl}(g) \geq \text{nl}(f) - (m + 1)$ if $m$ is odd; $\text{nl}(g) \geq \text{nl}(f) - m$ if $m$ is even.

Proof: 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 $g_i(x_1, \ldots, x_n) = f_i(x_1, \ldots, x_n) \oplus x_1 \ldots x_{i-1}x_{i+1} \ldots x_n$. 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 has 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 $\text{nl}(g_i) = \text{nl}(f_i) - 2$. Further, the inputs for which the outputs are changed are the all one vector and the vector with a zero only in the $i$th 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 ≤ i ≤ m, we have nl(g) ≥ nl(f) − (m + 1) when m is odd. Since the nonlinearity of a function of n variables and degree < n is always even we have nl(g) ≥ nl(f) − m when m is even.

Now suppose that h_i(x) (resp. h'_i(x)) is obtained from h(x) (resp. h'(x)) by fixing at most j (1 ≤ j ≤ k) input bits to constant values. Since h(x) and h'(x) differ in exactly (i + 1) positions, it follows that h_i(x) and h'_i(x) differ in at most (i + 1) positions. Further, since h_i(x) and h'_i(x) differ in at most (i + 1) positions, so does h_i(x ⊕ α) and h'_i(x ⊕ α). Let μ(x) = h(x) ⊕ h(x ⊕ α) and μ' = h'(x) ⊕ h'(x ⊕ α). Then it follows that μ(x) and μ'(x) differ in at most 2(i + 1) positions. Since f satisfies SAC(k), it follows that μ'(x) is balanced and has weight 2^n−j−1. Also since 1 ≤ i ≤ m and 1 ≤ j ≤ k, we obtain

\[ \left| \frac{wt(\mu(x))}{2^n-j} - \frac{1}{2} \right| = \left| \frac{wt(\mu(x))}{2^n-j} - \frac{wt(\mu'(x))}{2^n-j} \right| = \left| \frac{wt(\mu(x)) - wt(\mu'(x))}{2^n-j} \right| \leq \frac{2(i+1)}{2^n-j} \leq \frac{m+1}{2^n-k-1}. \]

This completes the proof. ■

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

<table>
<thead>
<tr>
<th>n</th>
<th>degree</th>
<th>m = 2</th>
<th>m = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>7</td>
<td>(3, 0.1875, 118)</td>
<td>(1, 0.0625, 116)</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>(3, 0.0938, 238)</td>
<td>(2, 0.0625, 236)</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>(4, 0.0938, 494)</td>
<td>(3, 0.0625, 492)</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>(5, 0.0938, 990)</td>
<td>(3, 0.0313, 988)</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>(6, 0.0938, 2014)</td>
<td>(4, 0.0313, 2012)</td>
</tr>
</tbody>
</table>

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 have possible applications in the 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.

Acknowledgements: We wish to thank the reviewers for reading the paper and providing several suggestions.
References


