A GMM type construction for resilient S-boxes with higher-dimensional vectorial outputs and strictly almost optimal nonlinearity

WeiGuo Zhang and LuYang Li
ISN Laboratory, Xidian University, Xi'an 710071, China
e-mail: zwg@xidian.edu.cn

Enes Pasalic
University of Primorska, FAMNIT, Koper, Slovenia
e-mail: enes.pasalic6@gmail.com

Abstract

Resilient substitution boxes (S-boxes) with high nonlinearity are important cryptographic primitives in the design of certain encryption algorithms. There are several trade-offs between the most important cryptographic parameters and their simultaneous optimization is regarded as a difficult task. In this paper we provide a construction technique to obtain resilient S-boxes with so-called strictly almost optimal (SAO) nonlinearity for a larger number of output bits m than previously known. This is the first time that the nonlinearity bound $2^{n-1} - 2^{n/2}$ of resilient (n, m) S-boxes, where n and m denote the number of the input and output bits respectively, has been exceeded for $m > \lfloor \frac{n}{4} \rfloor$. Thus, resilient S-boxes with extremely high nonlinearity and a larger output space compared to other design methods have been obtained.

Keywords: Boolean functions, nonlinearity, resiliency, S-boxes, stream ciphers.

1 Introduction

The concept of resilient S-boxes was first introduced by Chor *et al,* [1] and independently by Bennett *et al.* [2]. Whereas in the block cipher design the resiliency is not considered as a relevant cryptographic criterion, in certain stream cipher encryption algorithms it is not only a desirable property but sometimes also necessary as well. On the other hand, the nonlinearity is a widely accepted cryptographic criterion which measures the Hamming distance of a given Boolean function to the set of affine functions. The same applies when S-boxes are considered since in this case a common approach is to consider the nonzero linear combinations of the output Boolean functions and to measure their distance to the affine functions. For instance, the best affine approximation attack [3] and linear approximation attack [4] both reflect the importance of designing highly nonlinear S-boxes. The construction of resilient S-boxes with high nonlinearity has been an important

challenge in cryptography since mid 1990s, and it was extensively studied in [6, 7, 8, 9, 10, 11, 12, 13].

An (n, m) S-box can be identified with a multiple-output (vectorial) Boolean function, thus with a mapping $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$. In addition, to specify the order of resiliency the notation (n, m, t) stands for a t-resilient (n, m) S-box. For even $n \geq 2m$, the (n, m) S-boxes achieving the maximum possible nonlinearity $2^{n-1} - 2^{n/2-1}$ are called perfect nonlinear S-boxes [5]. However, it is well-known that perfect nonlinear S-boxes can not be resilient. In [14], Zhang and Pasalic presented a novel approach of finding disjoint linear codes and obtained a large set of (n, m, t) S-boxes with strictly almost optimal (SAO) nonlinearity, where an (n, m) S-box is called SAO if its nonlinearity is strictly greater than $2^{n-1} - 2^{\lfloor n/2 \rfloor}$. For the first time the nonlinearity bound $2^{n-1} - 2^{n/2}$ has been exceeded by an (n, m, t) S-box in [14] but the number of output bits was bounded by $m \leq \lfloor \frac{n}{4} \rfloor$. In this paper, we present a new construction method for designing (n, m, t) S-boxes with SAO nonlinearity. whose number of output bits can exceed the value $\lfloor \frac{n}{4} \rfloor$. This implies the existence of resilient S-boxes with SAO nonlinearity with larger output space, thus potentially improving the throughput of encryption algorithms for certain stream cipher scheme that employ such S-boxes for achieving a proper nonlinear characteristic of the cipher.

The main idea behind the proposed method is to construct certain matrices whose m columns will define the m Boolean functions of a resilient (n, m) S-box. Similarly to the standard MM method, for any fixed column its entries are all different and each of them will specify a distinct linear function. To ensure that the order of resiliency is t, only those rows in the considered matrix are used for which any nonzero linear combination of the elements in the row is of Hamming weight at least $t + 1$. Otherwise, the rows not satisfying this condition are discarded and not used in the construction. The size of these row reduced matrices cannot be deduced theoretically and their number is established through computer simulations. Nevertheless, the simulations show that such an approach in many cases yields significant improvements in the number of outputs m for t-resilient S-boxes with SAO nonlinearity compared to the method in [14]. Notice that for the same purpose the method in [14] uses a set of disjoint linear $[u, m, t + 1]$ codes (these codes intersect in zero only) and therefore the request concerning the Hamming weight is automatically satisfied. Finally, we emphasize that our method can be employed for the cases when $m \leq \lfloor \frac{n}{4} \rfloor$, thus addressing the same range of main parameters as in [14], giving resilient S-boxes with slightly better nonlinearity (see Table 3).

The rest of this article is organized as follows. In Section 2 some basic definitions and notions are introduced. The construction method is described in Section 3 and the interesting parameters demonstrating the improvements upon other methods are given in a tabulated form. Some concluding remarks can be found in Section 4.

2 Preliminaries

The set of Boolean functions of n variables will be denoted by $\mathcal{B}_n$. A Boolean function $f \in \mathcal{B}_n$ maps the elements from the vector space $\mathbb{F}_2^n$ to $\mathbb{F}_2$, where $\mathbb{F}_2$ denotes the Galois field of two elements. The Galois field of order 2^n is denoted by $\mathbb{F}_{2^n}$. Thus, an element $X_n = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$ is mapped to $\mathbb{F}_2$ so that $f(X_n) \in \mathbb{F}_2$. To avoid confusion with the additions of integers in $\mathbb{R}$, denoted by $+$ and Σ_i, the additions over $\mathbb{F}_2$ are denoted by $\oplus$ and $\bigoplus_i$. Nevertheless, for simplicity, we denote by $+$ the addition of vectors of $\mathbb{F}_2^n$. A Boolean function $f(X_n)$ is commonly represented by its algebraic normal form (ANF):

$$f(X_n) = \bigoplus_{u \in \mathbb{F}_2^n} \lambda_u (\prod_{i=1}^{n} x_i^{u_i}) \tag{1}$$

where $\lambda_u \in \mathbb{F}_2,\ u = (u_1, \ldots, u_n)$. The algebraic degree of $f(X_n)$, denoted by $deg(f)$, is the maximal value of $wt(u)$ such that $\lambda_u \neq 0$, where $wt(u)$ denotes the Hamming weight of u. In particular, f is said to be affine if $deg(f) = 1$. An affine function with constant term equal to zero, thus $\lambda_0 = 0$, is called a linear function. Any linear function on $\mathbb{F}_2^n$ can be uniquely expressed using the standard inner (dot) product, denoted by "$\cdot$", as:

$$\omega \cdot X_n = \omega_1 x_1 \oplus \cdots \oplus \omega_n x_n,$$

where $\omega = (\omega_1, \ldots, \omega_n),\ X_n = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$. The Walsh spectral value of $f \in \mathcal{B}_n$ at any point $\omega \in \mathbb{F}_2^n$ is denoted by $W_f(\omega)$ and calculated as

$$W_f(\omega) = \sum_{X_n \in \mathbb{F}_2^n} (-1)^{f(X_n) \oplus \omega \cdot X_n}. \tag{2}$$

$f \in \mathcal{B}_n$ is said to be balanced if its output column in the truth table contains equal number of 0's and 1's (that is, $W_f(\mathbf{0}) = 0$), where the truth table is the evaluation of $f(X_n)$ when X_n goes through $\mathbb{F}_2^n$. In [15], a spectral characterization of resilient functions was given.

Lemma 1. *An n-variable Boolean function is t-resilient if and only if its Walsh transform satisfies*

$$W_f(\omega) = 0,\ \textit{for } 0 \leq wt(\omega) \leq t,\ \omega \in \mathbb{F}_2^n. \tag{3}$$

The nonlinearity of $f \in \mathcal{B}_n$, as the most important parameter, can be expressed in terms of the Walsh spectra as [16]

$$N_f = 2^{n-1} - \frac{1}{2} \cdot \max_{\omega \in \mathbb{F}_2^n} |W_f(\omega)|. \tag{4}$$

An (n, m) S-box can be viewed as a vectorial Boolean function $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ which in turn can be represented as a collection of m Boolean functions so that $F(X_n) = (f_1(X_n), \ldots f_m(X_n))$, where $f_1, \ldots, f_m \in \mathcal{B}_n$ are called component functions of F.

Definition 1. *The nonlinearity of an (n, m) S-box $F(X_n) = (f_1(X_n), \ldots, f_m(X_n))$, denoted by N_F, is defined as*

$$N_F = \min_{c \in \mathbb{F}_2^{m*}} N_{f_c} \tag{5}$$

where $f_c = \sum_{i=1}^{m} c_i f_i$, $c = (c_1, \ldots, c_m)$ and $\mathbb{F}_2^{m*} = \mathbb{F}_2^m \backslash \boldsymbol{0}$.

Definition 2. *An (n, m) S-box $F(X_n) = (f_1(X_n), \ldots f_m(X_n))$ is t-resilient if and only if all nonzero linear combinations of $f_1, \ldots, f_m$ are t-resilient functions.*

3 Main Construction

In this section, we first describe the construction details and give some explanations for greater readability when necessary.

Let u, m, t be integers with $2 \leq m < u$. Let α be a root of the primitive polynomial $p(x) = 1 + p_1 x + \cdots + p_{k-1}x^{k-1} + x^u \in \mathbb{F}_2[x]$ and $(1, \alpha, \alpha^2, \cdots, \alpha^{u-1})$ be a polynomial basis of $\mathbb{F}_{2^u}$. Define a bijection $\pi : \mathbb{F}_{2^u} \mapsto \mathbb{F}_2^u$ by

$$\pi(b_0 + b_1\alpha + \cdots + b_{u-1}\alpha^{u-1}) = (b_0, b_1, \cdots, b_{u-1}). \tag{6}$$

Consider the matrix $A^{(u)}$ of size $(2^u - 1) \times m$, defined by,

$$A^{(u)} = \begin{pmatrix} \pi(1) & \pi(\alpha) & \cdots & \pi(\alpha^{m-1}) \\ \pi(\alpha) & \pi(\alpha^2) & \cdots & \pi(\alpha^m) \\ \vdots & \vdots & \ddots & \vdots \\ \pi(\alpha^{2^u-2}) & \pi(1) & \cdots & \pi(\alpha^{m-2}) \end{pmatrix} = \begin{pmatrix} A_0^{(u)} \\ A_1^{(u)} \\ \vdots \\ A_{2^u-2}^{(u)} \end{pmatrix}. \tag{7}$$

It is not difficult to show that (cf. [11]), for any nonzero linear combination of columns of the matrix $A^{(u)}$ each nonzero vector of $\mathbb{F}_2^u$ appears exactly once.

To ensure the resiliency of order t, we delete some row vectors in $A^{(u)}$. More precisely, for any row vector $A_i^{(u)}$, $i = 0, 1, \cdots, 2^u - 2$, if there exists a vector $c \in \mathbb{F}_2^{m*}$ such that $wt(c \cdot A_i^{(u)}) \leq t$, then $A_i^{(u)}$ will be deleted from $A^{(u)}$. The remaining row vectors form a new matrix

$$\widetilde{A^{(u)}} = \begin{pmatrix} \pi(\alpha^{l_1}) & \pi(\alpha^{l_1+1}) & \cdots & \pi(\alpha^{l_1+m-1}) \\ \pi(\alpha^{l_2}) & \pi(\alpha^{l_2+1}) & \cdots & \pi(\alpha^{l_2+m-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \pi(\alpha^{l_{N(u,m,t)}}) & \pi(\alpha^{l_{N(u,m,t)}+1}) & \cdots & \pi(\alpha^{l_{N(u,m,t)}+m-1}) \end{pmatrix} = \begin{pmatrix} A_{l_1}^{(u)} \\ A_{l_2}^{(u)} \\ \vdots \\ A_{l_{N(u,m,t)}}^{(u)} \end{pmatrix} \tag{8}$$

of size $N(u, m, t) \times m$, where $0 \leq l_1 < l_2 < \cdots < l_{N(u,m,t)} \leq \alpha^{2^u-2}$ and

$$N(u, m, t) = \#\{A_i^{(u)} \mid wt(c \cdot A_i^{(u)}) > t \ \texttt{for any} \ c \in \mathbb{F}_2^{m*}, \ i = 0, 1, \cdots, 2^u - 2\}. \tag{9}$$

Table 1: $N(u, m, t)$

t	$m \setminus u$	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	2	113	238	492	1002	2024	4070	8164	16354	32737	65502	131036	262106	524248	1048534
	3	99	220	472	980	2002	4044	8136	16324	32707	65468	131000	262068	524208	1048492
	4	71	187	432	939	1957	3992	8080	16267	32647	65400	130931	261992	524128	1048411
	5	29	123	356	855	1867	3888	7972	16151	32527	65264	130791	261840	523968	1048247
	6	1	45	231	692	1687	3686	7753	15919	32287	64997	130511	261536	523648	1047919
	7	0	1	70	416	1350	3290	7315	15455	31802	64459	129951	260928	523014	1047263
1	8	-	0	1	115	791	2572	6466	14534	30847	63383	128831	259719	521741	1045951
	9	-	-	0	1	178	1438	4912	12760	28957	61239	126591	257295	519195	1043327
	10	-	-	-	0	1	307	2660	9555	25327	57028	122146	252456	514103	1038079
	11	-	-	-	-	0	1	487	4760	18842	49008	113437	242864	503958	1027593
	12	-	-	-	-	-	0	1	809	9327	35502	96951	224438	483872	1006716
	13	-	-	-	-	-	-	0	1	1364	16576	68432	189897	444768	965443
	14	-	-	-	-	-	-	-	0	1	2064	29618	131527	373576	885327
	2	79	189	429	922	1926	3948	8020	16185	32543	65276	130781	261818	523924	1048174
	3	40	133	356	831	1821	3808	7861	16002	32336	65032	130508	261508	523576	1047793
	4	2	51	219	651	1606	3526	7545	15637	31916	64535	129952	260873	522867	1047018
	5	0	0	63	362	1205	3016	6932	14914	31092	63556	128865	259627	521485	1045499
	6	0	0	0	67	587	2135	5741	13472	29425	61582	126667	257105	518682	1042428
	7	0	0	0	0	56	915	3759	10871	26231	57724	122338	252134	513121	1036329
2	8	-	0	0	0	0	74	1311	6590	20457	50374	113863	242313	502037	1024122
	9	-	-	0	0	0	0	53	1964	11606	37688	97826	223241	480169	999909
	10	-	-	-	0	0	0	0	59	2838	19258	70523	187967	438154	952192
	11	-	-	-	-	0	0	0	0	16	3574	33050	129858	361553	861031
	12	-	-	-	-	-	0	0	0	0	0	4929	55719	238452	696721
	13	-	-	-	-	-	-	0	0	0	0	0	6215	90566	439795
	14	-	-	-	-	-	-	-	0	0	0	0	0	7864	150083

Note that for any fixed $c \in \mathbb{F}_2^{m*}$,

$$\#\{A_i^{(u)} \mid wt(c \cdot A_i^{(u)}) \leq t,\ i = 0, 1, \cdots, 2^u - 2\} = \sum_{i=1}^{t} \binom{u}{i}. \tag{10}$$

We have

$$(2^u - 1) - (2^m - 1) - \sum_{i=1}^{t} \binom{u}{i} \geq N(u, m, t) \geq (2^u - 1) - (2^m - 1) \sum_{i=1}^{t} \binom{u}{i}. \tag{11}$$

By computer simulations, a detailed list of $N(u, m, t)$ is given in Table 1.

Remark 1. The main difference to the approach taken in [14] is that we cannot estimate the number $N_{u,m,t}$ theoretically but rather using computer simulations. On the other hand, the method in [14] uses a sophisticated way of finding disjoint linear $[u, m, t+1]$ codes where each code gives rise to $2^m - 1$ rows (only) satisfying the restrictions imposed by resiliency order t.

In what follows we specify our resilient S-box by using two matrices $\widetilde{A^{(n/2)}}$ and $\widetilde{A^{(k)}}$ for suitably chosen k which together provide enough rows to define an (n, m) S-box. The entries of $\widetilde{A^{(n/2)}}$ will correspond to linear functions in $n/2$ variables but since there are not sufficiently many rows from $\widetilde{A^{(n/2)}}$ the goal is to find the smallest integer k for which the construction is possible.

Let n be even, and k, m be two integers with $m < k < n/2$. Let $E_0 \subset \mathbb{F}_2^{n/2}$ with $\#E_0 = N_0$. Let $\overline{E_0} = \mathbb{F}_2^{n/2} \setminus E_0$, and $E_1 = \overline{E_0} \times \mathbb{F}_2^{n/2-k}$ with $\#E_1 = N_1$. Obviously,

$N_0 \cdot 2^{n/2} + N_1 \cdot 2^k = 2^n$. Let $E_0 = \{e_1, e_2, \cdots, e_{N_0}\}$ with

$$N_0 = N(n/2, m, t). \tag{12}$$

For $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, N(n/2, m, t)$, let

$$\varphi_i(e_j) = \pi(\alpha^{l_j+i-1}) \tag{13}$$

be an injective mapping from E_0 to $\mathbb{F}_2^{n/2}$. Then we have

$$\begin{pmatrix} \varphi_1(e_1) & \varphi_2(e_1) & \cdots & \varphi_m(e_1) \\ \varphi_1(e_2) & \varphi_2(e_2) & \cdots & \varphi_m(e_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(e_{N_0}) & \varphi_2(e_{N_0}) & \cdots & \varphi_m(e_{N_0}) \end{pmatrix} = \widetilde{A^{(n/2)}}. \tag{14}$$

Let $E_1 = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_{N_1}\}$. If $N_1 \leq N(k, m, t)$, we can build injective mappings ψ_i, $i = 1, 2, \ldots, m$, from E_1 to $\mathbb{F}_2^k$ such that

$$\begin{pmatrix} \psi_1(\epsilon_1) & \psi_2(\epsilon_1) & \cdots & \psi_m(\epsilon_1) \\ \psi_1(\epsilon_2) & \psi_2(\epsilon_2) & \cdots & \psi_m(\varepsilon_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\epsilon_{N_1}) & \psi_2(\varepsilon_{N_1}) & \cdots & \psi_m(\epsilon_{N_1}) \end{pmatrix} = \widetilde{A^{(k)}}^* = \begin{pmatrix} A_{l_1}^{(k)} \\ A_{l_2}^{(k)} \\ \vdots \\ A_{l_{N_1}}^{(k)} \end{pmatrix}, \tag{15}$$

where the row vectors of $\widetilde{A^{(k)}}^*$ consist of the upper N_1 rows of $\widetilde{A^{(k)}}$.

Let $X_n = (X'_{n/2}, X''_{n/2}) = (X'_{n-k}, X''_k) \in \mathbb{F}_2^n$, where $X'_{n/2}$, $X''_{n/2} \in \mathbb{F}_2^{n/2}$, $X'_{n-k} \in \mathbb{F}_2^{n-k}$ and $X''_k \in \mathbb{F}_2^k$. An (n, m) S-box is then defined as,

$$F(X_n) = (f_1(X_n), f_2(X_n), \cdots, f_m(X_n))$$

where for $i = 1, 2, \cdots, m$,

$$f_i(X_n) = \begin{cases} \varphi_i(X'_{n/2}) \cdot X''_{n/2}, & X'_{n/2} \in E_0 \\ \psi_i(X'_{n-k}) \cdot X''_k, & X'_{n-k} \in E_1. \end{cases} \tag{16}$$

Now we prove that the above approach indeed yields t-resilient S-boxes with SAO nonlinearity, which is summarized in Theorem 1 below. For any $\mathbf{0} \neq c = (c_1, \cdots, c_m) \in \mathbb{F}_2^m$, let $\varphi_c = c_1\varphi_1 + \cdots + c_m\varphi_m$. By (8),

$$\begin{pmatrix} \varphi_c(e_1) \\ \varphi_c(e_2) \\ \vdots \\ \varphi_c(e_{N_0}) \end{pmatrix} = \begin{pmatrix} \pi(\alpha^{l_1}(c_1 + c_2\alpha + \cdots + \alpha^{m-1})) \\ \pi(\alpha^{l_2}(c_1 + c_2\alpha + \cdots + \alpha^{m-1})) \\ \vdots \\ \pi(\alpha^{l_{N_0}}(c_1 + c_2\alpha + \cdots + \alpha^{m-1})) \end{pmatrix}. \tag{17}$$

Obviously, φ_c is bijective. Similarly, $\psi_c = c_1\psi_1 + \cdots + c_m\psi_m$ is bijective. Let $\alpha = (\beta', \beta'') = (\gamma', \gamma'') \in \mathbb{F}_2^n$, where $\beta', \beta'' \in \mathbb{F}_2^{n/2}$, $\gamma' \in \mathbb{F}_2^{n-k}$ and $\gamma'' \in \mathbb{F}_2^k$. Note that $\mathbb{F}_2^n = E_0 \times \mathbb{F}_2^{n/2} \cup E_1 \times \mathbb{F}_2^k$. Then

$$W_{f_c}(\alpha) = \sum_{X_n \in \mathbb{F}_2^n} (-1)^{f_c(X_n)+\alpha\cdot X_n} = I_1 + I_2 \tag{18}$$

where

$$I_1 = \sum_{X'_{n/2} \in E_0} (-1)^{\beta'\cdot X'_{n/2}} \sum_{X''_{n/2} \in \mathbb{F}_2^{n/2}} (-1)^{(\varphi_c(X'_{n/2})+\beta'')\cdot X''_{n/2}} \in \{0, \pm 2^{n/2}\} \tag{19}$$

and

$$I_2 = \sum_{X'_{n-k} \in E_1} (-1)^{\gamma'\cdot X'_{n-k}} \sum_{X''_k \in \mathbb{F}_2^k} (-1)^{(\psi_c(X'_{n-k})+\gamma'')\cdot X''_k} \in \{0, \pm 2^k\}. \tag{20}$$

Hence,

$$\max_{\alpha \in \mathbb{F}_2^n} |W_{f_c}(\alpha)| = 2^{n/2} + 2^k. \tag{21}$$

By (4),

$$N_{f_c} = 2^{n-1} - 2^{n/2-1} - 2^{k-1}, \tag{22}$$

which implies

$$N_F = 2^{n-1} - 2^{n/2-1} - 2^{k-1}. \tag{23}$$

Note that for $X'_k \in E_0$ (resp., $X'_{n-u} \in E_1$), we have $wt(\varphi_c(X'_k)) \geq t+1$ (resp. $wt(\psi_c(X'_{n-u})) \geq t+1$). When $0 \leq wt(\alpha) \leq t$, we have $wt(\beta'') \leq t$ and $wt(\gamma'') \leq t$. Clearly, $(\varphi_c(X'_k) + \beta'') \neq 0$, $(\psi_c(X'_u) + \gamma'') \neq 0$. Thus, $I_1 = I_2 = 0$. Then, we have $W_{f_c}(\alpha) = 0$. By (3), f_c is a t-resilient function, which implies F is a t-resilient S-box.

Thus we have proved the following result.

Theorem 1. *Let n be even, and k, m be two integers with $m < k < n/2$. If*

$$2^{n/2}N(n/2, m, t) + 2^k N(k, m, t) \geq 2^n, \tag{24}$$

then there exits an (n, m, t) S-box F with SAO nonlinearity

$$N_F = 2^{n-1} - 2^{n/2-1} - 2^{k-1}. \tag{25}$$

The condition $2^{n/2}N(n/2, m, t) + 2^k N(k, m, t) \geq 2^n$ simply means that the whole truth table of the component functions of the S-box (which is of length 2^n) can be covered by the linear functions derived from the matrices $\widetilde{A^{(n/2)}}$ and $\widetilde{A^{(k)}}$.

Example 1. *Let $n = 20$ and $m = 5$. Note that $N(10, 5, 1) = 855$ and $N(9, 5, 1) = 356$. We have $2^{10}N(10, 5, 1) + 2^9 N(9, 5, 1) \geq 2^{20}$, which implies that a $(20, 5, 1, 2^{19} - 2^9 - 2^8)$ S-box can be constructed.*

In Table 2, we list some SAO (n, m) S-boxes with $m \geq \lfloor n/4 \rfloor$ which were not known earlier.

Table 2: (n, m, t, N_F) S-boxes with $m \geq \lfloor n/4 \rfloor$

$(20, 5, 1, 2^{19} - 2^9 - 2^8)$	$(24, 6, 1, 2^{23} - 2^{11} - 2^{10})$
$(26, 7, 1, 2^{25} - 2^{12} - 2^{11})$	$(28, 8, 1, 2^{27} - 2^{13} - 2^{12})$
$(30, 9, 1, 2^{29} - 2^{14} - 2^{13})$	$(32, 8, 1, 2^{31} - 2^{15} - 2^{13})$
$(32, 10, 1, 2^{31} - 2^{15} - 2^{14})$	$(34, 9, 1, 2^{33} - 2^{16} - 2^{14})$
$(34, 11, 1, 2^{33} - 2^{16} - 2^{15})$	$(36, 10, 1, 2^{35} - 2^{17} - 2^{15})$
$(36, 12, 1, 2^{35} - 2^{17} - 2^{16})$	$(38, 11, 1, 2^{37} - 2^{18} - 2^{16})$
$(38, 13, 1, 2^{37} - 2^{18} - 2^{17})$	$(40, 10, 1, 2^{39} - 2^{19} - 2^{16})$
$(40, 12, 1, 2^{39} - 2^{19} - 2^{17})$	$(40, 14, 1, 2^{39} - 2^{19} - 2^{18})$
$(36, 9, 2, 2^{35} - 2^{17} - 2^{16})$	$(38, 10, 2, 2^{37} - 2^{18} - 2^{17})$

3.1 Extending the approach by using more matrices

In the previous construction, only two matrices $A^{(u)}$ for $u = n/2$ and $u = k$ were used. In the corollary below, we employ more matrices which may improve the nonlinearity while preserving the resiliency order. The main idea here is that in certain cases even though the parameters $n/2$ and k may be sufficient to provide enough rows to enable the construction, it might be the case that using the matrices $A^{(k_1)}, \ldots, A^{(k_s)}$, where $k > k_1 \cdots > k_s$, the construction is still possible but the nonlinearity is slightly better. It is worthy of noticing that using this approach one may also easily achieve better nonlineraty values by reducing the number of outputs m (most often it decreases by one).

Corollary 1. *Let n be even, and k, m be two integers with $n/2 > k_1 > k_2 > \cdots > k_s > m$ where $1 \leq s \leq n/2 - m - 1$. If*

$$2^{n/2} N(n/2, m, t) + \sum_{i=1}^{s} \left(2^{k_i} N(k_i, m, t) \right) \geq 2^n, \tag{26}$$

then there exits an (n, m, t) S-box F with SAO nonlinearity

$$N_F = 2^{n-1} - 2^{n/2-1} - \sum_{i=1}^{s} 2^{k_i - 1}. \tag{27}$$

Example 2. *Let $n = 22$ and $m = 6$. Noticing $2^{11} N(11, 6, 1) + 2^{10} N(10, 6, 1) + 2^9 N(9, 6, 1) \geq 2^{22}$, a $(22, 6, 1, 2^{21} - 2^{10} - 2^9 - 2^8)$ S-box can be obtained. Similarly, the S-boxes with parameters $(28, 7, 1, 2^{27} - 2^{13} - 2^{11} - 2^{10})$, $(30, 8, 1, 2^{29} - 2^{14} - 2^{12} - 2^{11})$, and $(40, 11, 2, 2^{39} - 2^{19} - 2^{18} - 2^{17})$ can also be constructed.*

Remark 2. *When $m < \lfloor n/4 \rfloor$, our construction still works. And we can obtain SAO S-boxes with better nonlinearity than those in [14], as illustrated in Table 3.*

Due to the size of these S-boxes we could not investigate their algebraic properties, in particular the algebraic immunity and their resistance to fast algebraic attacks remain unknown. We leave this issue as an open problem since to the best of our knowledge today's algorithms can only determine these values for $n \leq 16$.

Table 3: (n, m, t, N_F) S-boxes with better nonlinearity than [14]

Ours	[14]
$(18, 4, 1, 2^{17} - 2^8 - 2^7)$	$(18, 4, 1, 2^{17} - 2^9)$
$(22, 4, 1, 2^{21} - 2^{10} - 2^8)$	$(22, 4, 1, 2^{21} - 2^{10} - 2^9)$
$(30, 4, 1, 2^{29} - 2^{14} - 2^{10})$	$(30, 4, 1, 2^{29} - 2^{14} - 2^{11})$
$(28, 5, 1, 2^{27} - 2^{13} - 2^{10})$	$(28, 5, 1, 2^{27} - 2^{13} - 2^{11})$
$(36, 5, 1, 2^{35} - 2^{17} - 2^{12} - 2^{11})$	$(36, 5, 1, 2^{35} - 2^{17} - 2^{13})$
$(38, 5, 1, 2^{37} - 2^{18} - 2^{13})$	$(38, 5, 1, 2^{37} - 2^{18} - 2^{14})$
$(26, 6, 1, 2^{25} - 2^{12} - 2^{10} - 2^9)$	$(26, 6, 1, 2^{25} - 2^{12} - 2^{11})$
$(30, 6, 1, 2^{29} - 2^{14} - 2^{11} - 2^9)$	$(30, 6, 1, 2^{29} - 2^{14} - 2^{12})$
$(32, 6, 1, 2^{31} - 2^{15} - 2^{12})$	$(32, 6, 1, 2^{31} - 2^{15} - 2^{13})$
$(34, 6, 1, 2^{33} - 2^{16} - 2^{12} - 2^{11})$	$(34, 6, 1, 2^{33} - 2^{16} - 2^{14})$
$(30, 7, 1, 2^{29} - 2^{14} - 2^{12})$	$(30, 7, 1, 2^{29} - 2^{14} - 2^{13})$
$(32, 7, 1, 2^{31} - 2^{15} - 2^{12} - 2^{11})$	$(32, 7, 1, 2^{31} - 2^{15} - 2^{13})$
$(36, 7, 1, 2^{35} - 2^{17} - 2^{13} - 2^{12} - 2^{11})$	$(36, 7, 1, 2^{35} - 2^{17} - 2^{14})$
$(38, 7, 1, 2^{37} - 2^{18} - 2^{14})$	$(38, 7, 1, 2^{37} - 2^{18} - 2^{15})$
$(34, 8, 1, 2^{33} - 2^{16} - 2^{13} - 2^{12} - 2^{11})$	$(34, 8, 1, 2^{33} - 2^{16} - 2^{15})$
$(36, 8, 1, 2^{35} - 2^{17} - 2^{14})$	$(36, 8, 1, 2^{35} - 2^{17} - 2^{15})$
$(38, 8, 1, 2^{37} - 2^{18} - 2^{14} - 2^{13} - 2^{12})$	$(38, 8, 1, 2^{37} - 2^{18} - 2^{15})$
$(40, 8, 1, 2^{39} - 2^{19} - 2^{15})$	$(40, 8, 1, 2^{39} - 2^{19} - 2^{16})$
$(38, 9, 1, 2^{37} - 2^{18} - 2^{15})$	$(38, 9, 1, 2^{37} - 2^{18} - 2^{16})$
$(40, 9, 1, 2^{39} - 2^{19} - 2^{16})$	$(40, 9, 1, 2^{39} - 2^{19} - 2^{17})$
$(36, 2, 2, 2^{35} - 2^{17} - 2^{12} - 2^{11} - 2^{10})$	$(36, 2, 2, 2^{35} - 2^{17} - 2^{13})$

4 Conclusion

In this paper, we have presented a construction method for designing SAO resilient (n, m) S-boxes. In difference to the method in [14] our new construction can be used even for $m > \lfloor \frac{n}{4} \rfloor$ and consequently a large set of unknown functions with SAO nonlinearity can be generated. Furthermore, in Corollary 1 we give an improved method so that some higher nonlinearities can be get than those obtained by the original method.

References

[1] Chor, B., Goldreich, O., Hastad, J., Friedman, J., Rudich, S., Smolensky, R.: 'The bit extraction problem or t-resilient functions', *Proc. 26th Annu. Symp. Found. Comput. Sci.*, 1985, **26**, pp. 396–407

[2] Bennett, C.H., Brassard, G., Robert, J.M.: 'Privacy amplification by public discussion', *SIAM J. Comput.*, 1988, **7**, (2), pp. 210–229

[3] Ding, C., Xiao, G., Shan, W.: 'The stability theory of stream ciphers'. Berlin, Germany, 1991, (LNCS, 561)

[4] Matsui, M.: 'Linear cryptanalysis method for DES cipher'. Advances in Cryptography-EUROCRYPT 1993, 1994, (LNCS, **765**), pp. 386–397

[5] Nyberg, K.: 'Perfect nonlinear S-boxes'. Advances in Cryptography-EUROCRYPT 1991, 1991, (LNCS, **547**), pp. 378–386

[6] Chen, L., Fu, F.w.: 'On the construction of new resilient functions from old ones', *IEEE Trans. Inf. Theory*, 1999, **45**, (6), pp. 2077–2082

[7] Cheon, J.H.: 'Nonlinear vector resilient functions'. Advances in Cryptography-EUROCRYPT 2001, 2001, (LNCS, **2139**), pp. 458–469

[8] Gupta, K.C., Sarkar, P.: 'Improved construction of nonlinear resilient S-boxes', *IEEE Trans. Inf. Theory*, 2005, **51**, (1), pp. 339–348

[9] Johansson, T., Pasalic, E.: 'A construction of resilient functions with high nonlinearity', *IEEE Trans. Inf. Theory*, 2003, **49**, (2), pp. 494–501

[10] Kurosawa, K., Satoh, T., Yamamoto, K.: 'Highly nonlinear t-resilient functions', *J. Univ. Comput. Sci.*, 1997, **3**, (6), pp. 721–729

[11] Pasalic, E., Maitra, S.: 'Linear codes in generalized construction of resilient functions with very high nonlinearity', *IEEE Trans. Inf. Theory*, 2002, **48**, (8), pp. 2182–2191

[12] Zhang, X.M., Zheng, Y.: 'Cryptographically resilient functions', *IEEE Trans. Inf. Theory*, 1997, **43**, (5), pp. 1740–1747

[13] Zhang, W.G., Pasalic, E.: 'Highly nonlinear balanced S-boxes with good differential properties', *IEEE Trans. Inf. Theory*, 2014, **60**, (12), pp. 7970–7979

[14] Zhang, W.G., Pasalic, E.: 'Constructions of resilient S-boxes with SAO nonlinearity through disjoint linear codes', *IEEE Trans. Inf. Theory*, 2014, **60**, (3), pp. 1638–1651

[15] Xiao, G.Z., Massey, J.L.: 'A spectral characterization of correlationimmune combining functions', *IEEE Trans. Inf. Theory*, 1988, **34**, (3), pp. 569–571

[16] Meier, W., Staffelbach, O.: 'Nonlinearity criteria for cryptographic functions'. Advances in Cryptology-EUROCRYPT 1990, 1990, (LNCS, **434**), pp. 549–562.