Improved Construction of Nonlinear Resilient S-Boxes

Palash Sarkar and Kishan Chand Gupta
Indian Statistical Institute
203, B.T. Road, Calcutta 700 035, INDIA
E-mail: {Palash , kishan_t}@isical.ac.in

Abstract

We provide two new construction methods for nonlinear resilient functions. The first method is a simple modification of an elegant construction due to Zhang and Zheng and constructs n-input, m-output resilient S-boxes with degree $d > m$. We prove by an application of the Griesmer bound for linear error correcting codes that the modified Zhang-Zheng construction is superior to the previous method of Cheon in Crypto 2001. Our second construction uses a sharpened version of the Maiorana-McFarland technique to construct nonlinear resilient functions. The nonlinearity obtained by our second construction is better than previously known construction methods.

Keywords : S-box, Griesmer bound, Resiliency, nonlinearity, algebraic degree, stream cipher.

1 Introduction

An (n, m) S-box (or vectorial function) is a map $f : \{0,1\}^n \to \{0,1\}^m$. By an (n, m, t) S-box (or (n, m, t)-resilient function) we mean t-resilient (n, m) S-box. An $(n, 1, t)$-resilient S-box is a resilient Boolean function. The cryptographic properties (like resiliency, nonlinearity, algebraic degree) of Boolean functions necessary for stream cipher applications have already been extensively studied. The resiliency property of S-box was introduced by Chor et al [5] and Bennett et al [1]. However, to be used in stream ciphers several other properties of S-box like nonlinearity and algebraic degree are also very important. Stinson and Massey [18] considered nonlinear resilient functions but only to disprove a conjecture.

It was Zhang and Zheng [20] who first proposed a beautiful method of transforming a linear resilient S-box to construct a nonlinear resilient S-box with high nonlinearity and high algebraic degree keeping cryptography in mind. After that, serious efforts to construct nonlinear S-box with high nonlinearity and high algebraic degree has been made [8, 7, 12, 4](see Section 2.4).

The current state-of-art in resilient S-box design can be classified into the following two approaches.

1. Construction of (n, m, t)-resilient functions with very high nonlinearity.
2. Construction of (n, m, t)-resilient functions with degree $d > m$ and high nonlinearity.

The first problem has been studied in [20, 8, 7, 12]. The currently best known results are obtained using the construction described in [12], though in certain cases, for small number of variables, the search technique of [7] yields better results. The second problem has been less studied. To the best of our knowledge, the only known construction which provides functions of the second type is due to Cheon [4].

In this paper, we first prove that the correlation immunity of a resilient function is preserved under composition with an arbitrary Boolean function. This property is useful for possible application of resilient S-boxes in designing secure stream ciphers. Our main contribution consists of two different constructions for the above two classes of problems. In both cases our results provide significant improvement over all previous methods.

The construction for the second problem is a simple modification of the Zhang-Zheng method [20]. To get algebraic degree $d > m$, we start with an $[n, d+1, t+1]$ code. Then we apply Zhang-Zheng construction to obtain a nonlinear S-box. Finally we drop $d+1-m$ output columns to obtain an (n, m, t)-resilient S-box (see Section 4). This simple modification is powerful enough to improve upon the best known construction with algebraic degree greater than m [4]. This clearly indicates the power of the original Zhang-Zheng construction. Our contribution is to apply the Griesmer bound for linear error correcting codes to *prove* that the modified Zhang-Zheng construction is superior to the best known construction [4]. We know of no other work where such a provable comparison of construction has been presented.

The Maiorana-McFarland technique is a well known method to construct nonlinear resilient functions. The idea is to use affine functions on small number of variables to construct nonlinear resilient functions on larger number of variables. We provide a construction to generate functions of the first type using a sharpened version of the Maiorana-McFarland method. For Boolean functions, the Maiorana-McFarland technique to construct resilient functions was introduced by Camion et al [2]. Nonlinearity calculation for the construction was first performed by Seberry, Zhang and Zheng [16]. This technique was later sharpened by Chee et al [3] and Sarkar-Maitra [15]. For S-boxes this technique has been used by [7] and [12], though [7] uses essentially a heuristic search technique. Here we develop and sharpen the technique of affine function concatenation to construct nonlinear resilient S-boxes. This leads to significant improvement in nonlinearity over that obtained in [12]. Thus we obtain better results than [12] which provides currently best known nonlinearity results for most choices of input parameters n, m, t.

2 Preliminaries

This section has four parts. We cover preliminaries on Boolean functions and S-boxes in Sections 2.1 and 2.2 respectively. In Section 2.3, we mention the coding theory result that we require. In Section 2.4, we summarize the previous construction results.

2.1 Boolean Functions

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 $wt(u)$. 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 Walsh Transform of an m-variable Boolean function g is an integer valued function $W_g : \{0,1\}^m \rightarrow [-2^m, 2^m]$ defined by (see [9, page 414])

$$W_g(u) = \sum_{w \in F_2^m} (-1)^{g(w) \oplus \langle u,w\rangle}. \tag{1}$$

The Walsh Transform is called the spectrum of g. The inverse Walsh Transform is given by

$$(-1)^{g(u)} = \frac{1}{2^m} \sum_{w \in F_2^m} W_g(w)(-1)^{\langle u,w\rangle}. \tag{2}$$

An m-variable function is called *correlation immune* of order t (t-CI) if $W_g(u) = 0$ for all u with $1 \leq wt(u) \leq t$ [17, 19]. Further the function is balanced if and only if $W_g(0) = 0$. A balanced t-CI function is called *t-resilient*. For even n, an n-variable function f is called *bent* if $W_f(u) = \pm 2^{\frac{n}{2}}$, for all $u \in F_2^n$ (see [14]). This class of functions is important in both cryptography and coding theory.

A parameter of fundamental importance in cryptography is the non-linearity of a function (see [9]). This is defined to be the distance from the set of all affine functions. It is more convenient to define it in terms of the spectrum of a Boolean function. The non-linearity $nl(f)$ of an n-variable Boolean function f, is defined as

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{u \in F_2^n} |W_f(u)|.$$

For even n, bent functions achieve the maximum possible nonlinearity.

A Boolean function g can be uniquely represented by a multivariate polynomial over F_2. The degree of the polynomial is called the algebraic degree or simply the degree of g.

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 $(g \circ f)(x) = g(f(x))$. An (n, m) S-box f is said to be t-CI, if $g \circ f$ is t-CI for every non-constant m-variable linear function g (see [20]). Further, if f is balanced then f is called t-resilient. (f is said to be balanced if $g \circ f$ is balanced for every non-constant m-variable linear function g). By an (n, m, t) S-box we mean t-resilient (n, m) S-box. Let f be an (n, m) S-box. Then nonlinearity of f, denoted by $nl(f)$, is defined to be

$nl(f) = \min\{nl(g \circ f) : g$ is a non-constant m-variable linear function $\}$.

Similarly the algebraic degree of f, denoted by $deg(f)$, is defined to be

$deg(f) = \min\{deg(g \circ f) : g$ is a non-constant m-variable linear function $\}$.

We will be interested in (n, m) S-boxes with maximum possible nonlinearity. If $n = m$, the S-boxes achieving the maximum possible nonlinearity are called maximally nonlinear [6]. If n is odd, then maximally nonlinear S-boxes have nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$. For even n, it is possible to construct (n, m) S-boxes with nonlinearity $2^{n-1} - 2^{\frac{n}{2}}$, though it is an open question whether this value is the maximum possible.

An (n, m) S-box with nonlinearity $2^{n-1} - 2^{\frac{n}{2}-1}$ is called perfect nonlinear S-box. Nyberg [10] has shown that perfect nonlinear functions exist if and only if n is even and $n \geq 2m$. For odd $n \geq 2m$, it is possible to construct S-boxes with nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$.

If we fix an enumeration of the set $\{0,1\}^n$, then an (n, m) S-box f is uniquely defined by a $2^n \times m$ matrix M_f. Given a sequence of S-boxes $f_1, \cdots, f_k$; where f_i is an (n_i, m) S-box we define the *concatenation* of $f_1, \cdots, f_k$ to be the matrix

$$M = \begin{bmatrix} M_{f_1} \\ M_{f_2} \\ \vdots \\ M_{f_k} \end{bmatrix}.$$

If $n_1 + \cdots + n_k = 2^n$ for some n, then the matrix M uniquely defines an (n, m) S-box f. In this case we say f is the *concatenation* of $f_1, \cdots, f_k$.

2.3 Coding Theory Results

An $[n, k, d]$ binary linear code is a subset of F_2^n which is a vector space of dimension k over F_2 having minimum distance d. We here mention the Griesmer bound (see [9, page 546]). For an $[n, k, d]$ linear code let $N(k, d)$ = length of the shortest binary linear code of dimension k and minimum distance d.

The Griesmer bound states (see [9, page 547])

$$N(k, d) \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{2^i} \right\rceil. \tag{3}$$

We say that the parameters n, k, d satisfy the Griesmer bound with equality if $n = \sum_{i=0}^{k-1} \left\lceil \frac{d}{2^i} \right\rceil$. There is a general construction (see [9, page 550]) which gives large class of codes meeting the Griesmer bound with equality. Given d and k , define $s = \lceil \frac{d}{2^{k-1}} \rceil$ and $d = s2^{k-1} - \sum_{i=1}^{p} 2^{u_i - 1}$ where $k > u_1 > \cdots > u_p \geq 1$. Given d and k, there is an $[n = s(2^k - 1) - \sum_{i=1}^{p}(2^{u_i} - 1), k, d]$ code meeting the Griesmer bound with equality if $\sum_{i=1}^{min(s+1,p)} u_i \leq sk$ (see [9, page 552]). This condition is satisfied for most values of d and k.

2.4 Some Recent Constructions

Here we summarize the previous construction results.

1. Zhang and Zheng [20]: This is the first paper to provide an elegant general construction of nonlinear resilient S-boxes. The main result proved is the following [20, Corollary 6]. If there exists a linear (n, m, t)-resilient function, then there exists a nonlinear (n, m, t)-resilient function with algebraic degree $(m - 1)$ and nonlinearity $\geq (2^{n-1} - 2^{n-\frac{m}{2}})$.

2. Kurosawa, Satoh and Yamamoto [8, Theorem 18]: For any even l such that $l \geq 2m$, if there exists an $(n - l, m, t)$-resilient function , then there exists an (n, m, t)-resilient function, whose nonlinearity is at least $2^{n-1} - 2^{n-\frac{l}{2}-1}$.

3. Johansson and Pasalic [7]: They use a linear error correcting code to build a matrix A of small affine functions. Resiliency and nonlinearity is ensured by using non-intersecting codes along with the matrix A. The actual non-intersecting codes used were obtained by a heuristic search technique. It becomes difficult to carry out this search technique for $n > 12$.

4. Pasalic and Maitra [12]: They use the matrix A of the previous method (3) along with highly nonlinear functions for their construction. The nonlinearity obtained is higher than the previous methods, except in certain cases, where the search technique of (3) yields better results.

5. Cheon [4, Theorem 5]: Uses linearized polynomial to construct nonlinear resilient function. The nonlinearity calculation is based on Hasse-Weil bound for higher genus curves. The main result is the following. If there exists $[n, m, t]$ linear code then for any non-negative integer D there exists a $(n + D + 1, m, t - 1)$-resilient function with algebraic degree D and nonlinearity at least $(2^{n+D} - 2^n \lfloor \sqrt{2^{n+D+1}} \rfloor + 2^{n-1})$. *Till date, this is the only construction which provides (n, m, t) nonlinear resilient S-boxes with degree greater than m.*

3 A Composition Theorem for S-boxes

We consider the composition of an (n, m) S-box and an m-variable Boolean function. The following result describes the Walsh Transform of the composition.

Theorem 3.1 *Let $f : \{0,1\}^n \to \{0,1\}^m$ and $g : \{0,1\}^m \to \{0,1\}$. Then for any $w \in F_2^n$,*

$$W_{(g \circ f)}(w) = \frac{1}{2^m} \sum_{v \in F_2^m} W_g(v) W_{(l_v \circ f)}(w)$$

where $l_v = \langle v, x \rangle$ and $(l_v \circ f)(x) = \langle v, f(x) \rangle$.

Proof : By Equation 2 , we have $(-1)^{g(x)} = \frac{1}{2^m} \sum_{w \in F_2^m} W_g(w)(-1)^{<w,x>}$.

Hence , $(-1)^{(g \circ f)(x)} = (-1)^{g(f(x))} = \frac{1}{2^m} \sum_{v \in F_2^m} W_g(v)(-1)^{\langle v, f(x) \rangle} = \frac{1}{2^m} \sum_{v \in F_2^m} W_g(v)(-1)^{(l_v \circ f)(x)}$.

By Equation 1 , we have

$$W_{g\circ f}(w) = \sum_{x\in F_2^n} (-1)^{(g\circ f)(x)\oplus<w,x>} = \frac{1}{2^m} \sum_{x\in F_2^n} \sum_{v\in F_2^m} W_g(v)(-1)^{(l_v\circ f)(x)\oplus<w,x>}$$

$$= \frac{1}{2^m} \sum_{v\in F_2^m} W_g(v) \sum_{x\in F_2^n} (-1)^{(l_v\circ f)(x)\oplus<w,x>} = \frac{1}{2^m} \sum_{v\in F_2^m} W_g(v) W_{(l_v\circ f)}(w) \qquad \blacksquare$$

Corollary 3.1 *Let $f : \{0,1\}^n \to \{0,1\}^m$ be a balanced S-box. Let g be an m-variable Boolean function. Then $(g \circ f)$ is balanced if and only if g is balanced.*

Remark: It is possible for $(g \circ f)$ to be balanced even when either only f is unbalanced or both f and g are unbalanced. We present examples for these cases. Let $f : \{0,1\}^3 \to \{0,1\}^2$ be an unbalanced S-box and f_1, f_2 are component functions.

(a) Let $f_1(x_1,x_2,x_3) = x_1\oplus x_2\oplus x_1x_2\oplus x_1x_3\oplus x_1x_2x_3$ and $f_2(x_1,x_2,x_3) = x_2\oplus x_1x_2\oplus x_2x_3\oplus x_1x_3\oplus x_1x_2x_3$ and $g(x_1,x_2) = x_1\oplus x_2$. Here f is unbalanced but g is balanced. Observe $(g\circ f)(x_1,x_2,x_3) = f_1(x_1,x_2,x_3)\oplus f_2(x_1,x_2,x_3) = x_1 \oplus x_2x_3$ is balanced.

(b) Let $f_1(x_1,x_2,x_3) = x_3 \oplus x_1x_2 \oplus x_1x_2x_3$ and $f_2(x_1,x_2,x_3) = x_2 \oplus x_3 \oplus x_1x_2 \oplus x_2x_3 \oplus x_1x_2x_3$ and $g(x_1,x_2) = x_1x_2$. Here both f and g are unbalanced. Observe $(g\circ f)(x_1,x_2,x_3) = f_1(x_1,x_2,x_3)f_2(x_1,x_2,x_3)$ $= x_3$, which is balanced.

Theorem 3.1 and Corollary 3.1 provide the following theorem.

Theorem 3.2 *Let f be a t-resilient S-box and g be any arbitrary Boolean function then $(g \circ f)$ is t-CI. Further $(g \circ f)$ is t-resilient if and only if g is balanced.*

Theorem 3.2 shows that correlation immunity of an (n,m,t)-resilient S-box is preserved under composition with an *arbitrary* m-variable Boolean function. This is an important security property for the use of resilient S-boxes in stream cipher design.

4 Construction of (n,m,t)-Resilient S-box with Degree $> m$.

In this section we modify an elegant construction by Zhang and Zheng [20] to obtain high degree nonlinear resilient S-boxes. The following result is well known(see for example [20]).

Theorem 4.1 *Let C be a $[n,m,t+1]$ binary linear code. Then we can construct an (n,m,t)-resilient function.*

Modified Zhang-Zheng (MZZ) Construction .

1. Input: Number of output columns $= m$, degree $= d \geq m$ and resiliency $= t$.
2. Output: An $(n,m,t)-$resilient function with degree d and nonlinearity $2^{n-1} - 2^{n-\lceil\frac{d+1}{2}\rceil}$.

Procedure

1. Choose an $[n,d+1,t+1]$ code to obtain a linear $(n,d+1,t)$-resilient function f.
2. Define $g = G \circ f$, where $G : \{0,1\}^{d+1} \to \{0,1\}^{d+1}$ is a bijection and $deg(G) = d$, $nl(G) \geq 2^d - 2^{\lfloor\frac{d+1}{2}\rfloor}$ [11]. Then $nl(g) \geq 2^{n-d-1}(2^d - 2^{\lfloor\frac{d+1}{2}\rfloor}) = 2^{n-1} - 2^{n-\lceil\frac{d+1}{2}\rceil}$ and $deg(g) = d$ [20, Corollary 6] .
3. Drop $(d+1-m)$ columns from the output of g to obtain an(n,m,t)-resilient function with degree d and nonlinearity $2^{n-1} - 2^{n-\lceil\frac{d+1}{2}\rceil}$.

Remark: For Step 2 above, there are other bijections by which we get the same value of $nl(G)$ but $deg(G) = d$ is achieved only for inverse mapping [6].

The modification to the Zhang-Zheng construction is really simple. If we want degree d, then we start with an $[n, d+1, t+1]$ code. Then we apply the main step of Zhang-Zheng construction to obtain a nonlinear S-box. Finally we drop $d+1-m$ output columns to obtain an (n, m, t)-resilient S-box. Though simple, this modification is powerful enough to improve upon the best known construction with high algebraic degree [4]. This shows the power of the original Zhang-Zheng construction. Our contribution is to *prove* by an application of the Griesmer bound that the MZZ construction is superior to the best known construction [4, Cheon]. We know of no other work where such provable comparisons of construction has been presented.

Theorem 4.2 *Let n, m, d, t be such that the following two conditions hold.*
1. Either (a) $d < m$ or (b) $d \geq m \geq \log_2(t+1)$.
2. The parameters $n, d+1, t+1$ meet the Griesmer bound with equality.
Then it is not possible to construct an (n, m, t)-resilient function f with degree d using Cheon [4] method.

Proof : Recall the Cheon construction from Section 2.4. Given any $[N, M, T+1]$ and a non negative integer D, the Cheon construction produces an $(N+D+1, M, T)$-resilient function with degree D. Thus if f is obtained by the Cheon construction we must have $n = N+D+1$, $m = M$, $t = T$ and $d = D$. This means that an $[n-d-1, m, t+1]$ code will be required by the Cheon construction. Since the parameters $n, d+1, t+1$ satisfies Griesmar bound with equality we have $n = \sum_{i=0}^{d} \lceil \frac{t+1}{2^i} \rceil$.
Claim : If (a) $d < m$ or (b) $d \geq m \geq \log_2(t+1)$ then $n - d - 1 < \sum_{i=0}^{m-1} \lceil \frac{t+1}{2^i} \rceil$.
Proof of the claim: Since $n = \sum_{i=0}^{d} \lceil \frac{t+1}{2^i} \rceil$ we have that $n - d - 1 < \sum_{i=0}^{m-1} \lceil \frac{t+1}{2^i} \rceil$ if and only if $\sum_{i=0}^{d} \lceil \frac{t+1}{2^i} \rceil - d - 1 < \sum_{i=0}^{m-1} \lceil \frac{t+1}{2^i} \rceil$. If $d < m$, then the last mentioned condition is trivially true. So suppose $d \geq m \geq \log_2(t+1)$. Then the above inequality holds if and only if $\sum_{i=m}^{d} \lceil \frac{t+1}{2^i} \rceil < d+1$. Since $m \geq \log_2(t+1)$, $\sum_{i=m}^{d} \lceil \frac{t+1}{2^i} \rceil = d - m + 1 < d + 1$ for $m \geq 1$. This completes the proof of the claim.
Since $n - d - 1 < \sum_{i=0}^{m-1} \lceil \frac{t+1}{2^i} \rceil$, the parameters $n-d-1, m, t+1$ violate the Griesmer bound and hence an $[n-d-1, m, t+1]$ code do not exist. Thus Cheon method cannot be used to construct the function f. ■

The following result is a consequence of Theorem 4.2 and the MZZ construction.

Theorem 4.3 *Let n, m, d, t be such that the following two conditions hold.*
1. Either (a) $d < m$ or (b) $d \geq m \geq \log_2(t+1)$.
2. An $[n, d+1, t+1]$ code meeting the Griesmer bound with equality exist.
Then it is possible to construct an (n, m, t)-resilient function f with degree d by the MZZ method which cannot be constructed using Cheon [4] method.

Remark: As mentioned in [9, page 550] there is a large class of codes which meet the Griesmer bound. Further, the condition $d \geq m \geq \log_2(t+1)$ is quite weak. Hence there exists a large class of (n, m, t)-resilient functions which can be constructed using MZZ construction but cannot be constructed using Cheon [4] construction. See Section 6 for some concrete examples.

Nonlinearity in Cheon method is $(2^{N+D} - 2^N \lfloor \sqrt{2^{N+D+1}} \rfloor + 2^{n-1})$ (see item 5 of Section 2.4) which is positive if $D \geq N+1$ for $N \geq 2$. So for $D \leq N$ Cheon method do not provide any nonlinearity. Thus Cheon method may provide high algebraic degree but it does not provide good nonlinearity. In fact, in the next theorem we prove that nonlinearity obtained by MZZ method is larger than nonlinearity obtained by Cheon method.

Theorem 4.4 *Let f be an (n, m, t)-resilient function f of degree d and nonlinearity n_1 constructed by Cheon method. Suppose there exists a linear $[n, d+1, t+1]$ code. Then it is possible to construct an (n, m, t)-resilient function g with degree d and nonlinearity n_2 using MZZ method . Further $n_2 \geq n_1$.*

Proof : Since $[n, d+1, t+1]$ code exists, the MZZ construction can be applied to obtain an (n, m, t)-resilient function g with degree d and nonlinearity $nl(g) = n_2 = 2^{n-1} - 2^{n - \lceil \frac{d+1}{2} \rceil}$. It remains to show that

$n_2 \geq n_1$, which we show now. Recall $n_1 = 2^{n-1} - 2^{n-d-1}\lfloor\sqrt{2^n}\rfloor + 2^{n-d-2}$. Hence $n_2 - n_1 \geq -2^{n-\frac{d+1}{2}} + 2^{n-d-1}\lfloor\sqrt{2^n}\rfloor - 2^{n-d-2}$. Thus we have $n_2 \geq n_1$ if $-2^{\frac{-(d+1)}{2}} + 2^{-(d+1)}\lfloor\sqrt{2^n}\rfloor - 2^{-(d+2)} \geq 0$. The last condition holds if and only if $\lfloor\sqrt{2^n}\rfloor \geq 2^{d+1}\left(\frac{1}{2^{\frac{d+1}{2}}} + \frac{1}{2^{d+2}}\right)$. So $n_2 \geq n_1$ if $\sqrt{2^n} - 1 \geq 2^{\frac{d+1}{2}} + 2^{-1}$. i.e. if $2^{\frac{n}{2}} \geq 2^{\frac{d+1}{2}} + \frac{3}{2}$. Again the last condition hold for $1 \leq d \leq n-3$. Hence $n_2 \geq n_1$ for $1 \leq d \leq n-3$. The maximum possible degree of an S-box is $n-1$. For $d = n-1$ and $d = n-2$, Cheon construction requires $[0, m, t+1]$ and $[1, m, t+1]$ codes respectively. Clearly such code do not exist. Hence $n_2 \geq n_1$ holds for all d. ■

Lemma 4.1 *Let f be an (n, m, t)-resilient function f of degree $d \geq m$ constructed by Cheon method and $m \geq \log_2(t+1)$. Then the parameters $n, d+1, t+1$ satisfy the Griesmer bound with equality.*

Proof : Since f has been obtained from Cheon method, there exists an $[n-d-1, m, t+1]$ code. Hence the parameters $n-d-1, m$ and $t+1$ satisfy the Griesmar bound. Since $n-d-1, m$ and $t+1$ satisfy the Griesmar bound we have $n-d-1 \geq \sum_{i=0}^{m-1}\lceil\frac{t+1}{2^i}\rceil$. i.e. we have $n \geq d+1+\sum_{i=0}^{m-1}\lceil\frac{t+1}{2^i}\rceil$. As $m \geq \log_2(t+1)$ we have $\lceil\frac{t+1}{2^i}\rceil = 1$ for $i \geq m$. Hence $n \geq (d+1) - (d-m+1) + \sum_{i=m}^{d}\lceil\frac{t+1}{2^i}\rceil + \sum_{i=0}^{m-1}\lceil\frac{t+1}{2^i}\rceil$. This shows $n \geq m + \sum_{i=0}^{d}\lceil\frac{t+1}{2^i}\rceil$ and consequently $n \geq \sum_{i=0}^{d}\lceil\frac{t+1}{2^i}\rceil$. Thus the parameters $n, d+1, t+1$ satisfy the Griesmer bound with equality. ■

Remark: Since the parameters $n, d+1$ and $t+1$ satisfy the Griesmer bound, in most cases it is possible to obtain an $[n, d+1, t+1]$ code (see [9, page 550]) and apply Theorem 4.4. *In fact we do not know any case where a function can be constructed using the Cheon method but not by the MZZ method.* Theorems 4.3 and 4.4 *prove* the clear superiority of the MZZ method over the Cheon construction. Thus MZZ method is the currently known best method to construct $[n, m, t]$-resilient function with degree $d > m$.

5 A Construction to Obtain High Nonlinearity

In this section we concentrate on obtaining (n, m, t)-resilient S-boxes with high nonlinearity only. We present a construction method which improves the nonlinearity obtainable by the previously known methods. We start by mentioning the following result which is restatement of Lemma 7 in [7].

Theorem 5.1 *Let C be a $[u, m, t+1]$ code. Then it is possible to construct $(2^m - 1) \times m$ matrix D with entries from C, such that, $\{c_1 D_{i,1} \oplus \cdots \oplus c_m D_{i,m} : 1 \leq i \leq 2^m - 1\} = C \setminus \{(0, \cdots, 0)\}$ for each nonzero vector $(c_1, \cdots, c_m) \in F_2^m$.*

Let D be the matrix in Theorem 5.1. For $(1 \leq i \leq 2^m - 1)$ and $(1 \leq j \leq m)$ define a u-variable linear function $L_{i,j}(x_1, \cdots, x_u) \stackrel{\triangle}{=} \langle D_{i,j}, (x_1, \cdots, x_u)\rangle$. Given the code C we define a $(2^m - 1) \times m$ matrix $L(C)$ whose entries are u-variable linear functions by defining the i, j th entry of $L(C)$ to be $L_{i,j}(x_1, \cdots, x_u)$. We have the following result which follows directly from Theorem 5.1.

Proposition 5.1 *Let $c \in F_2^m$ be a nonzero row vector. Then all the entries of the column vector $L(C)c^T$ are distinct.*

For positive integers k, l with $k \leq l$, we define $L(C, k, l)$ to be the submatrix of $L(C)$ consisting of the rows k to l. Thus $L(C, 1, 2^m - 1) = L(C)$. Let $G(y_1, \cdots, y_p)$ be a (p, m) S-box whose component functions are $G_1, \cdots, G_m$. We define $G \oplus L(C, k, l)$ to be an $(l-k+1) \times m$ matrix whose i, j th entry is $G_j(y_1, \cdots, y_p) \oplus L_{k+i-1,j}(x_1, \cdots, x_u)$ for $1 \leq i \leq l-k+1$ and $1 \leq j \leq m$. If $l-k+1 = 2^r$ for some r then $G \oplus L(C, k, l)$ defines an S-box $F : \{0,1\}^{r+p+u} \rightarrow \{0,1\}^m$ in the following manner.

$$F_j(z_1, \cdots, z_r, y_1, \cdots, y_p, x_1, \cdots, x_u) \quad = \quad G_j(y_1, \cdots, y_p) \oplus L_{k+i-1,j}(x_1, \cdots, x_u) \qquad (4)$$

where $1 \leq j \leq m$, $1 \leq i \leq 2^r$, $F_1, \cdots, F_m$ are the component functions of F and $z_1 \cdots z_r$ is the binary representation of $i-1$. By $F = G \oplus L(C, k, l)$ we will mean the above representation of the S-box F. *Note* *that the function F is t-resilient, since each $L_{i,j}(x_1, \cdots, x_u)$ is non-degenerate on at least $(t+1)$ variables and hence t-resilient.*

In the matrix $M = G(y_1, \cdots, y_p) \oplus L(C, k, l)$ we say that the row $L_{i,*}$ of $L(C)$ is repeated 2^p times. Let $G(y_1, \cdots, y_p)$ and $H(y_1, \cdots, y_q)$ be (p, m) and (q, m) S-boxes respectively and $M_1 = G \oplus L(C, k, l)$, $M_2 = H \oplus L(C, k, l)$. Then we say that the row $L_{i,*}$ of $L(C)$, $(k \leq i \leq l)$ is repeated a total of $2^p + 2^q$ times in the matrix $[M_1\ M_2]^T$.

Proposition 5.1 has also been used by [12] in the construction of resilient S-boxes. However we improve upon the construction of [12] by utilizing the following two ideas.
1. We use all the $2^m - 1$ rows of the matrix $L(C)$. In contrast, [12] uses at most 2^{m-1} rows of $L(C)$.
2. We allow a row of $L(C)$ to be repeated 2^{r_1} or $2^{r_1} + 2^{r_2}$ or $2^{r_1} + 2^{r_2} + 2^{r_3}$ times as required. On the other hand, the number of times a row of $L(C)$ can be repeated in [12] is of the form 2^r.

It turns out that a proper utilization of the above two techniques result in significant improvement in nonlinearity. We will require (r, m) S-boxes with very high nonlinearity. For this we propose to use the best known results which we summarize in the following definition.

Definition 5.1 *Let G be an (r, m) S-box satisfying the following.*
1. If $r < m$, G is a constant S-Box.
2. If $m \leq r < 2m$, G is a maximally nonlinear S-Box [6].
3. If $r \geq 2m$ and r is even, G is a perfect nonlinear S-Box [11].
4. If $r \geq 2m$ and r is odd, G is concatenation of two perfect nonlinear S-Boxes(see Section 2.2).
Then we say that G is a PROPER S-box.

The following result summarizes the best known results on the nonlinearity of PROPER S-boxes.

Proposition 5.2 *Let G be an (r, m) PROPER S-box. Then*
1. If $r < m$, $nl(G) = 0$.
2. If $m \leq r < 2m$, then $nl(G) = 2^{r-1} - 2^{\frac{r-1}{2}}$ if r is odd and $nl(G) \geq 2^{r-1} - 2^{\frac{r}{2}}$ if r is even.
3. If $r \geq 2m$, then $nl(G) = 2^{r-1} - 2^{\frac{r}{2}-1}$ if r is even and $nl(G) = 2^{r-1} - 2^{\frac{r-1}{2}}$ if r is odd.

Now we are in a position to describe a new construction of resilient S-boxes. The construction has two parts. In Part-A, we compute the number of rows of $L(C)$ to be used and the number of times each row is to be repeated. The output of Part-A is a *list* of the form *list* $= \langle (n_1, R_1), (n_2, R_2), \cdots, (n_k, R_k) \rangle$ which signifies that n_i rows of $L(C)$ are to be repeated R_i times each. Part-A also computes a variable called effect which determines the nonlinearity of the S-box (see Theorem 5.2). In Part-B of the construction, we choose PROPER functions based on *list* and describe the actual construction of the S-box.

Construction-I.

1. Input: Positive integers (n, m) and t.
2. Output: A nonlinear (n, m, t)-resilient S-box F.

Part-A

1. Obtain minimum u such that $[u, m, t+1]$ code C exists.
2. Case: $n - u \leq 0$, then function cannot be constructed using this method. Hence stop.
3. Case: $n - u \geq 0$
   (a) $0 \leq n - u < m$; *list* $= \langle (2^{n-u}, 1) \rangle$ and effect $= 1$.
   (b) $m \leq n - u < 2m - 1$; *list* $= \langle (2^{m-1}, 2^{n-u-m+1}) \rangle$ and effect$= 2^{n-u-m+1}$.
   (c) $n - u = 2m - 1$; *list* $= \langle (2^{m-1}, 2^m) \rangle$ and effect$= 2^{\lfloor \frac{m}{2} \rfloor + 1}$.
   (d) $2m \leq n - u < 3m$.
      (i) $n - u = 2m + 2e$; m even; $0 \leq e < \frac{m}{2}$;

$list = \langle (1, 2^{m+2e+1}), (2^m - 2, 2^{m+2e}) \rangle$ and effect$= 2^{e+1+\frac{m}{2}}$.

(ii) $n-u = 2m + 2e + 1$; m even; $0 \le e \le \frac{m}{2} - 1$;

- $0 \le e \le \frac{m}{2} - 2$; $list = \langle (2, 2^{m+2e+1} + 2^{2e+1} + 2^{2e}), (2^m - 3, 2^{m+2e+1} + 2^{2e+1}) \rangle$ and effect$= 2^{2e+1} + 2^{2e} + 2^{e+1+\frac{m}{2}}$.
- $e = \frac{m}{2} - 1$; $list = \langle (2^{m-1}, 2^m) \rangle$ and effect$= 2^m$.

(iii) $n - u = 2m + 2e + 1$; m odd; $0 \le e \le \lfloor \frac{m}{2} \rfloor - 1$;
$list = \langle (1, 2^{m+2e+2}), (2^m - 2, 2^{m+2e+1}) \rangle$ and effect$= 2^{\frac{m+2e+3}{2}}$.

(iv) $n - u = 2m + 2e$; m odd; $0 \le e < \lfloor \frac{m}{2} \rfloor$;
$list = \langle (2^m - 2, 2^{m+2e} + 2^{2e+1}), (1, 2^{2e+2}) \rangle$ and effect$= 2^{2e+1} + 2^{\frac{m+2e+1}{2}}$.

(v) $n - u = 3m - 1$; m odd;
$list = \langle (2^{m-1}, 2^{2m}) \rangle$ and effect$= 2^m$.

(e) $n - u \ge 3m$.

(i) $n - u = 3m + 2e + 1$; $e \ge 0$;
$list = \langle (2^{m-1}, 2^{2m+2e+2}) \rangle$ and effect$= 2^{m+e+1}$.

(ii) $n - u = 3m + 2e$; (m even; $e \ge \frac{m}{2}$) or (m odd; $0 \le e < \lfloor \frac{m}{2} \rfloor$);
$list = \langle (2, 2^{2m+2e} + 2^{m+2e} + 2^{m+2e-1}), (2^m - 3, 2^{2m+2e} + 2^{m+2e}) \rangle$
and effect$= 2^{m+e} + 2^{e+1+\frac{m}{2}}$.

(iii) $n - u = 3m + 2e$; m even; $0 \le e < \frac{m}{2}$
$list = \langle (2^m - 2, 2^{2m+2e} + 2^{m+2e+1}), (1, 2^{m+2e+2}) \rangle$ and effect$= 2^{m+e} + 2^{e+1+\frac{m}{2}}$.

(iv) $n - u = 3m + 2e$; m odd; $e \ge \lfloor \frac{m}{2} \rfloor$
$list = \langle (2^m - 2, 2^{2m+2e} + 2^{m+2e+1}), (1, 2^{m+2e+2}) \rangle$ and effect$= 2^{m+e} + 2^{e+\frac{m+1}{2}}$.

Part-B

1. If $list = \langle (2^s, 2^r) \rangle$;
   - Obtain $L(C, 1, 2^s)$ from $L(C)$ by selecting first 2^s rows of $L(C)$.
   - Let G be an (r, m) PROPER S-box.
   - Define $F = G \oplus L(C, 1, 2^s)$.
   - This covers cases 3.(a),(b),(c),(d)(ii) second item, (d)(v) and e(i) of Part-A.
2. Case: 3(d)(i) of Part-A
   - Let G_1 and G_2 be $(m + 2e + 1, m)$ and $(m + 2e, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C, 1, 1)$, $F_2 = G_2 \oplus L(C, 2, 2^m - 1)$.
   - F is the concatenation of F_1 and F_2 .
3. Case: 3(d)(ii) first item of Part-A and $e = 0$
   - Let G_1 and G_2 be $(m + 1, m)$ and $(1, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C)$, $F_2 = G_2 \oplus L(C)$, $F_3 = L(C, 1, 2)$.
   - F is the concatenation of F_1, F_2 and F_3 .
4. Case: 3(d)(ii) first item of Part-A and $e \ne 0$
   - Let G_1, G_2 and G_3 be $(m + 2e + 1, m)$, $(2e + 1, m)$ and $(2e, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C)$, $F_2 = G_2 \oplus L(C)$, $F_3 = G_3 \oplus L(C, 1, 2)$.
   - F is the concatenation of F_1, F_2 and F_3 .
5. Case: 3(d)(iii) of Part-A
   - Let G_1 and G_2 be $(m + 2e + 2, m)$ and $(m + 2e + 1, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C, 1, 1)$, $F_2 = G_2 \oplus L(C, 2, 2^m - 1)$.
   - F is the concatenation of F_1 and F_2 .
6. Case: 3(d)(iv) of Part-A
   - Let G_1, G_2 and G_3 be $(m + 2e, m)$, $(2e + 2, m)$ and $(2e + 1, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C, 1, 2^m - 2)$, $F_2 = G_2 \oplus L(C, 2^m - 1, 2^m - 1)$, $F_3 = G_3 \oplus L(C, 1, 2^m - 2)$.

- F is the concatenation of F_1, F_2 and F_3 .

7. Case: 3(e)(ii) of Part-A
   - Let G_1, G_2 and G_3 be $(2m+2e, m)$, $(m+2e, m)$ and $(m+2e-1, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C)$, $F_2 = G_2 \oplus L(C)$, $F_3 = G_3 \oplus L(C, 1, 2)$.
   - F is the concatenation of F_1, F_2 and F_3 .
8. Case: 3(e)(iii) and 3(e)(iv) of Part-A
   - Let G_1, G_2 and G_3 be $(2m+2e, m)$, $(m+2e+2, m)$ and $(m+2e+1, m)$ PROPER S-boxes.
   - Define $F_1 = G_1 \oplus L(C, 1, 2^m - 2)$, $F_2 = G_2 \oplus L(C, 2^m - 1, 2^m - 1)$, $F_3 = G_3 \oplus L(C, 1, 2^m - 2)$.
   - F is the concatenation of F_1, F_2 and F_3 .

Theorem 5.2 *Construction-I provides a nonlinear (n, m, t)-resilient S-box with nonlinearity* $= (2^{n-1} - 2^{u-1} \times \mathsf{effect})$*, where* effect *is as computed in Part-A.*

Proof : There are several things to be proved.
(a) The output function F is an (n, m) S-box. *(b)* F is t-resilient. *(c)* $nl(f) = (2^{n-1} - 2^{u-1} \times \mathsf{effect})$.
Proof of (a) The output of Part-A is a *list* $= \langle (n_1, R_1), (n_2, R_2), \cdots, (n_k, R_k) \rangle$. Part-B ensures that for $1 \le i \le k$, n_i rows of $L(C)$ are repeated R_i times each. It is easy to verify that in each case of Part-A we have $\sum_{i=1}^{k} n_i R_i = 2^{n-u}$. Since each row $L_{i,*}$ of $L(C)$ defines a (u, m) S-box, ultimately F is an (n, m) S-box.
Proof of (b) Each row $L_{i,*}$ of $L(C)$ defines a t-resilient (u, m) S-box. F is formed by concatenating the rows of $L(C)$ one or more times. Hence F is t-resilient.
Proof of (c) The nonlinearity calculation is similar for all the cases. As an example, we perform the calculation for Case 3(e)(ii). In this case, Part-A computes *list* $= \langle (2, 2^{2m+2e} + 2^{m+2e} + 2^{m+2e-1}), (2^m - 3, 2^{2m+2e} + 2^{m+2e}) \rangle$. Let $R_1 = 2^{2m+2e} + 2^{m+2e} + 2^{m+2e-1}$ and $R_2 = 2^{2m+2e} + 2^{m+2e}$. Rows $L_{1,*}$ and $L_{2,*}$ of $L(C)$ are repeated R_1 times each and each of the rows $L_{3,*}$ to $L_{2^m-1,*}$ is repeated R_2 times each. Part-B uses three PROPER functions G_1, G_2 and G_3 to construct S-boxes F_1, F_2 and F_3 respectively. F is the concatenation of F_1, F_2 and F_3. We have to show that if ν is a non constant m-variable linear function and λ is an n-variable linear function, then $d(\nu \circ F, \lambda) \ge (2^{n-1} - 2^{u-1} \times \mathsf{effect})$. We write λ as $\lambda(y_1, \cdots, y_{n-u}, x_1, \cdots, x_u) = \lambda_1(y_1, \cdots, y_{n-u}) \oplus \lambda_2(x_1, \cdots, x_u)$. Let $\nu(z_1, \cdots, z_m) = \langle (c_1, \cdots, c_m), (z_1, \cdots, z_m) \rangle$ for some non-zero vector $c = (c_1, \cdots, c_m) \in F_2^m$. The Boolean function $\nu \circ F$ is a concatenation of Boolean functions $\nu \circ F_1$, $\nu \circ F_2$ and $\nu \circ F_3$. For $1 \le i \le 2$, $\nu \circ F_i = (\nu \circ G_i) \oplus (L(C)c^T)$ and $\nu \circ F_3 = (\nu \circ G_3) \oplus (L(C, 1, 2)c^T)$. Using Proposition 5.1, we know that all the entries of the column vector $L(C)c^T$ are distinct u-variable linear functions. Let $L(C)c^T = [\mu_1, \cdots, \mu_{2^m-1}]^T$. The function $\nu \circ F$ is a concatenation of the μ_i's and their complements. Further, μ_1 and μ_2 are repeated R_1 times and $\mu_3, \cdots, \mu_{2^m-1}$ are repeated R_2 times in the construction of $\nu \circ F$. If $\lambda \notin \{\mu_1, \cdots, \mu_{2^m-1}\}$ then $d(\lambda_2, \mu_i) = 2^{u-1}$ for each $1 \le i \le 2^m - 1$ and hence $d(\nu \circ F, \lambda) = 2^{n-u}(2^{u-1}) = 2^{n-1}$. Now suppose $\lambda_2 = \mu_i$ for some $i \in \{1, \cdots, 2^m - 1\}$. In this case $d(\nu \circ F, \lambda)$ will be less than 2^{n-1} and the actual value is determined by the repetition factors R_1 and R_2. There are two cases to consider.

Case 1: $\lambda_2 = \mu_1$ or μ_2. Without loss of generality we assume $\lambda_2 = \mu_1$, the other case being similar. Since $\lambda_2 = \mu_1$, we have $d(\lambda_2, \mu_i) = 2^{u-1}$ for $1 \le i \le 2^m - 1$. The function μ_2 is repeated R_1 times and each of the functions $\mu_3, \cdots, \mu_{2^m-1}$ is repeated R_2 times. Hence the total contribution of $\mu_2, \mu_3, \cdots, \mu_{2^m-1}$ to $d(\nu \circ F, \lambda)$ is $2^{u-1}(R_1 + (2^m - 3)R_2)$. We now have to compute the contribution of μ_1 to $d(\nu \circ F, \lambda)$. The function μ_1 is repeated in $\nu \circ F_i$ by XORing with $\nu \circ G_i$. Hence the contribution of μ_1 to $d(F, \lambda)$ is equal to $2^u(nl(\nu \circ G_1) + nl(\nu \circ G_2) + nl(\nu \circ G_3)) = 2^u(nl(G_1) + nl(G_2) + nl(G_3))$ since $nl(\nu \circ G_i) = nl(G_i)$. Each G_i is a PROPER function whose nonlinearity is given by 5.2.

Hence, $d(\nu \circ F, \lambda) = 2^{u-1}(R_1 + (2^m - 3)R_2 + 2(nl(G_1) + nl(G_2) + nl(G_3))$
$= 2^{u-1}(2^{n-u} - (R_1 - 2(nl(G_1) + nl(G_2) + nl(G_3)))) = 2^{n-1} - 2^{u-1}(R_1 - 2(nl(G_1) + nl(G_2) + nl(G_3)))$.
From the given conditions, it is easy to verify that $\mathsf{effect} = R_1 - 2(nl(G_1) + nl(G_2) + nl(G_3))$ and so $d(\nu \circ F, \lambda) = (2^{n-1} - 2^{u-1} \times \mathsf{effect})$.

Case 2: $\lambda_2 = \mu_i$ for some $i \in \{3, \cdots, 2^m - 1\}$. In this case we proceed as in the previous case to obtain $d(\nu \circ F, \lambda) = 2^{u-1}(2R_1 + (2^m - 4)R_2) + 2^u(nl(G_1) + nl(G_2)) = 2^{u-1}(2R_1 + (2^m - 4)R_2 + 2(nl(G_1) + nl(G_2))$ $= 2^{u-1}(2^{n-u} - R_2 + 2(nl(G_1) + nl(G_2)) = 2^{n-1} - 2^{u-1}(R_2 - 2(nl(G_1) + nl(G_2)))$ $> 2^{n-1} - 2^{u-1} \times$ effect, since effect $= R_1 - 2(nl(G_1) + nl(G_2) + nl(G_3)) > R_1 - 2(nl(G_1) + nl(G_2))$.

By *Case 1 and Case 2* above it follows that $nl(\nu \circ F) = 2^{n-1} - 2^{u-1} \times$ effect. Hence $nl(F) = 2^{n-1} - 2^{u-1} \times$ effect. ■

6 Results and Comparisons

Here we compare the construction methods described in this paper to the known construction methods.

6.1 Degree Comparison Based on MZZ Construction

We present examples to show the superiority of the MZZ method over the Cheon method. Cheon method cannot construct (n, m, t)-resilient function of degree $d \geq m \geq 2$ if the following two conditions hold.

(1)

t	1	2 to 3	4 to 7	8 to 15	16 to 31
m	$m \geq 1$	$m \geq 2$	$m \geq 3$	$m \geq 4$	$m \geq 5$

(2) The parameters $n, d+1, t+1$ satisfy Griesmer bound with equality.

We next present some examples of n, m, d and t satisfying condition (1) and (2) such that the MZZ method can be used to construct (n, m, t)-resilient function with degree d.

(a) $t = 1$, $2 \leq m \leq d$, $n = d + 2$. It is easy to check that a $[d+2, d+1, 2]$ code exists.

(b) $t = 2$, $2 \leq m \leq d$, $(n, d) = (6, 2), (7, 3), (8, 4), (9, 5), (10, 6), (11, 7)$. In each case an $[n, d+1, t+1]$ code exists.

(c) $t = 3$, $2 \leq m \leq d$, $(n, d) = (7, 2), (8, 3), (11, 6), (12, 7), (13, 8)$. In each case an $[n, d+1, t+1]$ code exists.

In (a) to (c) above an (n, m, t)-resilient function with degree d can be constructed using MZZ method, but cannot be constructed using Cheon method(see Theorem 4.3). Now we present some examples where both MZZ and Cheon method construct (n, m, t)-resilient function with degree d and compare their nonlinearity using Theorem 4.4. An (n, m, d, t) S-box is an (n, m, t)-resilient S-box with degree d.

Function	$(10, 3, 1, 5)$	$(18, 4, 2, 10)$	$(24, 5, 2, 15)$	$(24, 7, 3, 12)$	$(28, 6, 4, 14)$
Cheon [4, Theorem 5]	8	$2^{16} + 2^9$	$2^{23} - 2^{20} + 2^7$	2^{10}	2^{12}
MZZ	$2^9 - 2^7$	$2^{17} - 2^{12}$	$2^{23} - 2^{16}$	$2^{23} - 2^{17}$	$2^{27} - 2^{20}$

Table 1 : Comparison of nonlinearity obtained by MZZ Construction to that obtained by Cheon [4].

We see that in each case the nonlinearity obtained by the MZZ method is far superior to that obtained by the Cheon method.

6.2 Nonlinearity Comparison Based on Construction-I

We compare the nonlinearity obtained by Construction-I to the nonlinearity obtained in Theorem 4 of [12]. The nonlinearity obtained in [12] is better than the nonlinearity obtained by other methods. Hence we do not compare our method with the other methods. It is to be noted that in certain cases the search technique of [7] provides better nonlinearity than [12].

Our first observation is that *the nonlinearity obtained by Construction-I is at least as large the nonlinearity obtained in [12].* The intuitive reason is that we use all the rows of the matrix $L(C)$ and hence the repetition factor is less than that of [12]. The detailed verification of the superiority of Construction-I over [12] is straightforward but tedious. In the next table we summarize the cases under which Construction-I yields higher nonlinearity than [12].

Case	Nonlinearity of [12]	Construction-I nonlinearity
$2m \leq n-u < 3m-3$, π even	$2^{n-1} - 2^{(n+u-m+1)/2}$	$2^{n-1} - 2^{(n+u-m-1)/2} - 3 \times 2^{n-2m-2}$ **(1)** $2^{n-1} - 2^{(n+u-m-1)/2} - 2^{n-2m}$ **(2)**
$2m \leq n-u < 3m-3$, π odd	$2^{n-1} - 2^{(n+u-m+2)/2}$	$2^{n-1} - 2^{(n+u-m)/2}$ **(3)**
$n-u = 3m-3$	$2^{n-1} - 2^{(u+m-1)}$	$2^{n-1} - \frac{11}{16}2^{(u+m-1)}$ **(4)**
$n-u \geq 3m$, π odd	$2^{n-1} - 2^{(n+u-m)/2}$	$2^{n-1} - 2^{(n+u-m)/2}(\frac{1}{2} + \frac{1}{2^{m/2}})$ **(5)** $2^{n-1} - 2^{(n+u-m)/2}(\frac{1}{2} + \frac{1}{2^{((m+1)/2)}})$ **(6)**

Table 2 : Comparison of Construction-I nonlinearity with the nonlinearity of [12].

We list the different cases of Part-A corresponding to the different rows of the table.
(1) Case 3(d)(ii)first item; **(2)** Case 3(d)(iv); **(3)** Case 3(d)(i) and Case 3(d)(iii); **(4)** Case 3(d)(ii)first item; **(5)** Case 3(e)(iii), $m > 2$ and Case 3(e)(ii), $m > 2$; **(6)** Case 3(e)(iv), $m > 1$.

In Table 3 below and Tables 4 and 5 in the Appendix we provide some concrete examples of cases where the nonlinearity obtained by Construction-I is better than that obtained by [12]. Each entry of Tables 1 to 3 is of the form (a, b), where a is the nonlinearity obtained by [12] and b is the nonlinearity obtained by Construction-I.

Code	$n = 13$	$n = 14$	$n = 17$	$n = 19$
$[5, 4, 2]$	$(2^{12} - 2^8), (2^{12} - 2^7)$	$(2^{13} - 2^8), (2^{13} - \frac{11}{16}2^8)$	$(2^{16} - 2^9), (2^{16} - \frac{3}{4}2^9)$	$(2^{18} - 2^{10}), (2^{18} - \frac{3}{4}2^{10})$
Code	$n = 15$	$n = 16$	$n = 19$	$n = 21$
$[7, 4, 3]$	$(2^{14} - 2^{10}), (2^{14} - 2^9)$	$(2^{15} - 2^{10}), (2^{15} - \frac{11}{16}2^{10})$	$(2^{18} - 2^{11}), (2^{18} - \frac{3}{4}2^{11})$	$(2^{20} - 2^{12}), (2^{20} - \frac{3}{4}2^{12})$
Code	$n = 16$	$n = 17$	$n = 20$	$n = 22$
$[8, 4, 4]$	$(2^{15} - 2^{11}), (2^{15} - 2^{10})$	$(2^{16} - 2^{11}), (2^{16} - \frac{11}{16}2^{11})$	$(2^{19} - 2^{12}), (2^{19} - \frac{3}{4}2^{12})$	$(2^{21} - 2^{13}), (2^{21} - \frac{3}{4}2^{13})$

Table 3 : Comparison of Construction-I nonlinearity with [12] for $m = 4$ and resiliency $= 1, 2, 3$.

Nonlinearity of $(36, 8, t)$ resilient S-box has been used as very important examples in [8, 7, 12]. Now we compare our nonlinearity with those.

t	7	6	5	4	3	2	1
[8]	$2^{35} - 2^{27}$	$2^{35} - 2^{27}$	$2^{35} - 2^{26}$	$2^{35} - 2^{25}$	$2^{35} - 2^{24}$	$2^{35} - 2^{23}$	$2^{35} - 2^{22}$
[7]	$2^{35} - 2^{22}$	-	$2^{35} - 2^{23}$	$2^{35} - 2^{22}$	$2^{35} - 2^{22}$	$2^{35} - 2^{21}$	$2^{35} - 2^{21}$
[12]	$2^{35} - 2^{25}$	$2^{35} - 2^{24}$	$2^{35} - 2^{23}$	$2^{35} - 2^{23}$	$2^{35} - 2^{20}$	$2^{35} - 2^{20}$	$2^{35} - 2^{18}$
Ours	$2^{35} - 2^{24}$	$2^{35} - \frac{35}{64}2^{24}$	$2^{35} - \frac{19}{32}2^{23}$	$2^{35} - 2^{22}$	$2^{35} - 2^{20}$	$2^{35} - \frac{9}{16}2^{20}$	$2^{35} - 2^{18}$
Codes	$[20, 8, 8]$	$[19, 8, 7]$	$[17, 8, 6]$	$[16, 8, 5]$	$[13, 8, 4]$	$[12, 8, 3]$	$[9, 8, 2]$

Table 6 : Comparison of nonlinearity of $(36, 8, t)$-resilient S-boxes using different methods.

The results of [7] are not constructive. They show that resilient S-box with such parameter exist. *Note except for resiliency of order* 1 *and* 3 *our nonlinearity is better than nonlinearity of [12].* It should also be noted that in all the cases we provide construction with currently best known nonlinearity.

7 Conclusion

In this paper we consider the construction of nonlinear resilient S-boxes. We prove that the correlation immunity of a resilient S-box is preserved under composition with an arbitrary Boolean function. Our main contribution is to obtain two construction methods for nonlinear resilient S-boxes. The first construction is a simple modification of an elegant construction due to Zhang and Zheng [20]. This provides (n, m, t)-resilient S-boxes with degree $d > m$. We *prove* that the modified Zhang Zheng construction is superior to the only previously known construction [4] which provided degree $d > m$. Our second construction is based on concatenation of small affine function to build nonlinear resilient S-boxes. We sharpen the technique to construct (n, m, t)-resilient S-boxes with the currently best known nonlinearity.

References

[1] C. Bennett, G. Brassard and J. Robert. Privacy Amplification by Public Discussion. *SIAM Journal of Computing*, volume 17, pages 210–229, 1988.

[2] P. Camion, C. Carlet, P. Charpin and N. Sendrier . On correlation immune functions. *In Advances in Cryptology - CRYPT0 1991*, pages 86–100, Lecture Notes in Computer Science, Springer-Verlag, 1992.

[3] S. Chee, S. Lee, D. Lee and S. H. Sung . On the correlation immune functions and their nonlinearity. *In Advances in Cryptology - Asiacrypt 1996*, pages 232–243, Lecture Notes in Computer Science, Springer-Verlag, 1996.

[4] Jung Hee Cheon. Nonlinear Vector Resilient Functions. In *Advances in Cryptology - CRYPTO 2001*, pages 458–469, Lecture Notes in Computer Science, Springer-Verlag, 2001.

[5] B. Chor, O. Goldreich, J. Hastad, J. Friedman, S. Rudich and R. Smolensky. The Bit Extraction Problem or t-resilient Functions. *IEEE Symposium on Foundations of Computer Science*, volume 26, pages 396–407, 1985.

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

[7] T. Johansson and E. Pasalic. A construction of resilient functions with high nonlinearity. *International Symposium on Information Theory*, 2000.

[8] K. Kurosawa, T. Satoh and K. Yamamoto. Highly nonlinear t-resilient functions . Journal of Universal Computer Science, vol.3, no. 6, pp. 721-729, Springer Publishing Company, 1997.

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

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

[11] K. Nyberg. Differentially uniform mapping for cryptography. In *Advances in Cryptology - EUROCRYPT 1993*, pages 55–65, Lecture Notes in Computer Science, Springer-Verlag, 1994.

[12] E. Pasalic and S. Maitra. Linear Codes in Generalized Construction of Resilient Functions with Very High Nonlinearity. *Earlier version in SAC 2001. To appear in IEEE Transactions on Information Theory.*

[13] B. Preneel. Analysis and design of cryptographic hash functions, doctoral dissertation, K.U. Leuven, 1993.

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

[15] 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.

[16] J. Seberry, X.-M. Zhang and Y. Zheng . On construction and nonlinearity of correlation immune Boolean functions. *In Advances in Cryptology - EUROCRYPT 1993*, pages 181–199, Lecture Notes in Computer Science, Springer-Verlag, 1994.

[17] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, IT-30(5):776–780, September 1984.

[18] D. R. Stinson. An Infinite Class of Counterexamples to a Conjecture Concerning Nonlinear Resilient Functions. *Journal of Cryptology*, volume 8, pages 167–173, 1995.

[19] G. Xiao and J.L. Massey. A spectral characterization of correlation-immune combining functions. *IEEE Transactions on Information Theory*, pages 569–571, 1988.

[20] X.-M. Zhang and Y. Zheng, On Cryptographically Resilient Functions. *IEEE Transactions on Information Theory*, Vol 43 , No 5, pp. 1740-1747 , 1997.

A Further Examples

The linear codes used in Table 4 are $[6,5,2],[9,5,3]$ and $[10,5,4]$.

$n=16$	$n=17$	$n=18$	$n=21$
$(2^{15}-2^9),(2^{15}-\frac{5}{8}2^9)$	$(2^{16}-2^{10}),(2^{16}-2^9)$	$(2^{17}-2^{10}),(2^{17}-\frac{11}{16}2^{10})$	$(2^{20}-2^{11}),(2^{20}-\frac{5}{8}2^{11})$
$n=19$	$n=20$	$n=21$	$n=24$
$(2^{18}-2^{12}),(2^{18}-\frac{5}{8}2^{12})$	$(2^{19}-2^{13}),(2^{19}-2^{12})$	$(2^{20}-2^{13}),(2^{20}-\frac{11}{16}2^{13})$	$(2^{23}-2^{14}),(2^{23}-\frac{5}{8}2^{14})$
$n=18$	$n=19$	$n=20$	$n=25$
$(2^{17}-2^{11}),(2^{17}-\frac{5}{8}2^{11})$	$(2^{18}-2^{12}),(2^{18}-2^{11})$	$(2^{19}-2^{12}),(2^{19}-\frac{11}{16}2^{12})$	$(2^{24}-2^{15}),(2^{24}-\frac{5}{8}2^{15})$

Table 4: Comparison of Construction-I nonlinearity with [12] for $m=5$ and resiliency $=1,2,3$.

The linear codes used in Table 5 are $[7,6,2],[10,6,3]$ and $[10,6,4]$.

$n=19$	$n=20$	$n=21$	$n=22$
$(2^{18}-2^{11}),(2^{18}-2^{10})$	$(2^{19}-2^{11}),(2^{19}-\frac{19}{32}2^{11})$	$(2^{20}-2^{12}),(2^{20}-2^{11})$	$(2^{21}-2^{12}),(2^{21}-\frac{11}{16}2^{12})$
$n=22$	$n=23$	$n=24$	$n=25$
$(2^{21}-2^{14}),(2^{21}-2^{13})$	$(2^{22}-2^{14}),(2^{22}-\frac{19}{32}2^{14})$	$(2^{23}-2^{15}),(2^{23}-2^{14})$	$(2^{24}-2^{15}),(2^{24}-\frac{11}{16}2^{15})$
$n=22$	$n=23$	$n=24$	$n=25$
$(2^{21}-2^{14}),(2^{21}-2^{13})$	$(2^{22}-2^{14}),(2^{22}-\frac{19}{32}2^{14})$	$(2^{23}-2^{15}),(2^{23}-2^{14})$	$(2^{24}-2^{15}),(2^{24}-\frac{11}{16}2^{15})$

Table 5 : Comparison of Construction-I nonlinearity with [12] for $m=6$ and resiliency $=1,2,3$.