		$\{0, 38, 46, 125, 217, 280, 287\}$
6	11	$\{0, 1, 93, 132, 218, 286, 353\}$,
		$\{0, 5, 66, 121, 266, 270, 324\}$,
		$\{0, 10, 43, 162, 198, 300, 348\}$,
		$\{0, 6, 96, 146, 219, 302, 318\}$,
		$\{0, 23, 82, 100, 189, 297, 346\}$,
		$\{0, 32, 113, 158, 227, 321, 349\}$,
		$\{0, 19, 46, 147, 262, 271, 341\}$
		$\{0, 40, 47, 151, 223, 254, 328\}$,
		$\{0, 22, 60, 170, 231, 309, 329\}$,
		$\{0, 41, 98, 106, 242, 267, 351\}$,
		$\{0, 14, 56, 141, 185, 238, 350\}$,
		$\{0, 11, 35, 87, 255, 326, 343\}$,
		$\{0, 26, 63, 143, 205, 339, 342\}$,
		$\{0, 13, 15, 133, 263, 293, 314\}$
7	5	$\{0, 3, 52, 68, 94, 125, 168, 176\}$,
		$\{0, 5, 41, 75, 95, 139, 154, 158\}$,
		$\{0, 7, 13, 60, 93, 118, 164, 174\}$,
		$\{0, 9, 11, 39, 87, 142, 159, 171\}$,
		$\{0, 14, 37, 38, 99, 126, 144, 166\}$
7	6	$\{0, 12, 23, 86, 104, 156, 209, 213\}$,
		$\{0, 8, 10, 50, 93, 108, 181, 203\}$,
		$\{0, 17, 41, 47, 111, 143, 189, 202\}$,
		$\{0, 7, 55, 79, 120, 139, 176, 207\}$,
		$\{0, 14, 39, 90, 116, 151, 196, 205\}$,
		$\{0, 5, 6, 34, 135, 168, 184, 204\}$

Acknowledgment

The authors wish to thank the anonymous referees for their suggestions that helped improve the content and presentation of the correspondence. The computations were carried on the Beowulf cluster in the Math and CS Laboratory.

References

[1] Y. M. Chee and C. J. Colbourn, "Constructions for difference triangle sets," *IEEE Trans. Inf. Theory*, vol. 43, no. 4, pp. 1346–1349, Jul. 1997.

[2] C. J. Colbourn, "Difference triangle sets," in *The CRC Handbook of Combinatorial Designs*, C. J. Colbourn and J. Dintz, Eds. Boca Raton, FL: CRC, 1996, pp. 312–317.

[3] T. Kløve, "Bounds on the size of optimal difference sets," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 355–361, Mar. 1988.

[4] ——, "Bounds and construction for difference triangle sets," *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 879–886, Jul. 1989.

[5] A. Ling, "Difference triangle sets from affine planes," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2399–2401, Aug. 2002.

[6] R. Lorentzen and R. Nilsen, "Application of linear programming to the optimal difference triangle set problem," *IEEE Trans. Inf. Theory*, vol. 37, no. 5, pp. 1486–1488, Sep. 1991.

[7] J. B. Shearer, "Some new difference triangle sets," *J. Comb. Math. and Comb. Comput.*, vol. 27, pp. 65–76, 1998.

[8] ——, "Improved LP lower bounds for difference triangle sets," *Electron. J. Combin.*, vol. 6, no. 1, pp. R31–R31, 1999.

[9] ——, Difference Triangle Sets—Upper Bounds. [Online]. Available: http://www.research.ibm.com/people/s/shearer/dtsub.html#js3

[10] D. R. Stinson, "Hill-climbing algorithms for the construction of combinatorial designs," *Ann. Discr. Math.*, vol. 26, pp. 321–334, 1985.

Improved Construction of Nonlinear Resilient S-Boxes

Kishan Chand Gupta and Palash Sarkar

***Abstract*—We provide two new construction methods for nonlinear resilient functions. The first method is a simple modification of a 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.**

***Index Terms*—Algebraic degree, Griesmer bound, nonlinearity, resiliency, S-box, stream cipher.**

I. Introduction

An (n, m) S-box (or vectorial function) is a map $f : \{0,1\}^n \rightarrow \{0,1\}^m$. 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.* [7] and Bennett *et al.* [1]. However, to be used in stream ciphers, several other properties of the S-box, such as nonlinearity and algebraic degree, are also very important. Stinson and Massey [23] considered nonlinear resilient functions but only to disprove a conjecture.

Camion and Canteaut [2] described a general method of constructing a new resilient function by composing a resilient function and a bijection. A similar method for constructing resilient function from $\{0,1\}^n \rightarrow \{0,1\}^m$ was described by Zhang and Zheng [25]. After that, serious efforts to construct a nonlinear S-box with high nonlinearity and high algebraic degree has been made [13], [12], [17], [6] (see Section II-D).

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 [25], [13], [12], [17]. The currently best known results are obtained using the construction described in [17], though in certain cases, for a small number of variables, the search technique of [12] 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 [6].

In this correspondence, 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

Manuscript received January 6, 2004; revised September 8, 2004. The material of this correspondence was presented in part at the Conference on Advances in Cryptology–Asiacrypt 2002, Queenstown, New Zealand, December 2002.

The authors are with the Applied Statistics Unit, Indian Statistical Institute, Kolkata, Pin 700 108, India (e-mail: kishan_t@isical.ac.in; palash@isical.ac.in).

Communicated by T. Johansson, Associate Editor for Complexity and Cryptography.

Digital Object Identifier 10.1109/TIT.2004.839524

resilient S-boxes in designing secure stream ciphers. Our main contribution consists of two different constructions for the previously mentioned 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 [25]. To get algebraic degree $d > m$, we start with an $[u, d+1, t+1]$ code. Then we apply the 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 IV). This simple modification is powerful enough to improve upon the best known construction with algebraic degree greater than m [6]. 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 [6]. 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.* [3]. Nonlinearity calculation for the construction was first performed by Seberry, Zhang, and Zheng [21]. This technique was later sharpened by Chee *et al.* [5] and Sarkar–Maitra [20]. For S-boxes, this technique has been used by [12] and [17], though [12] 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 [17]. Thus, we obtain better results than [17] which currently provides the best known nonlinearity results for most choices of input parameters n, m, t.

In a recent work [10], the applicability of resilient S-boxes to stream cipher has been discussed. The work [10] also describes an efficient representation and software implementation method for resilient Maiorana–McFarland S-boxes. It is shown that such S-boxes can be implemented using very little memory and the output can be obtained using very few operations.

The correspondence is organized as follows. Section II provides basic definitions, notations, theory needed, and a quick review of recent construction. In Section III, we prove the composition theorem. Section IV provides a modified Zhang–Zheng construction and some theorems to prove its advantage over the Cheon construction. Section V provides some definitions and theory needed in that section. It also provides a construction by which we get an (n, m, t)-resilient S-box with nonlinearity greater than the nonlinearity obtained in [17] which has been known to be the best so far. In Section VI, we compare the modified Zhang–Zhang construction with the Cheon construction, and also compare Construction-I of Section V with the Pasalic and Maitra construction [17]. Section VII concludes this correspondence.

II. Preliminaries

This section consists of four parts. We cover preliminaries on Boolean functions and S-boxes in Sections II-A and B, respectively. In Section II-C, we mention the coding theory results that we require. In Section II-D, we summarize the previous construction results.

A. Boolean Functions

Let $\mathbb{F}_2 = \mathrm{GF}(2)$. We consider the domain of a Boolean function to be the vector space $(\mathbb{F}_2^n, \oplus)$ over $\mathbb{F}_2$, where $\oplus$ is used to denote the addition operator over both $\mathbb{F}_2$ and the vector space $\mathbb{F}_2^n$. The inner product of two vectors $u, v \in \mathbb{F}_2^n$ will be denoted by $\langle u, v\rangle$. The weight of an n-bit vector u is the number of ones in u and will be denoted by $\mathsf{wt}(u)$. The (Hamming) distance between two vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ 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 $W_g : \{0,1\}^m \to [-2^m, 2^m]$ defined by (see [14, p. 414])

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

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

$$(-1)^{g(u)} = \frac{1}{2^m} \sum_{w \in \mathbb{F}_2^m} W_g(w)(-1)^{\langle u, w\rangle}. \tag{2}$$

An m-variable function is called *correlation immune* of order t (t-CI) if $W_g(u) = 0$ for all u with $1 \le \mathsf{wt}(u) \le t$ [22], [24]. 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 $u \in \mathbb{F}_2^n$ (see [19]). This class of functions is important in both cryptography and coding theory.

A parameter of fundamental importance in cryptography is the nonlinearity of a function (see [14]). 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 nonlinearity $\mathsf{nl}(f)$ of an n-variable Boolean function f is defined as

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

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

A Boolean function g can be uniquely represented by a multivariate polynomial over $\mathbb{F}_2$. The degree of the polynomial is called the algebraic degree or simply the degree of g.

B. 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 nonconstant m-variable linear function g (see [25]). Further, if f is balanced then f is called t-resilient. (The function f is said to be balanced if $g \circ f$ is balanced for every nonconstant 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. The nonlinearity of f, denoted by $\mathsf{nl}(f)$, is defined to be

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

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

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

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 [9]. 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 a perfect nonlinear S-box. Nyberg [15] 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 $f_1,\dots,f_k$; where f_i is an (n_i,m) S-box, we define the *concatenation* of $f_1,\dots,f_k$ to be the matrix

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

If $2^{n_1} + \cdots + 2^{n_k} = 2^n$ 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 $f_1,\dots,f_k$.

C. Coding Theory Results

We will use some standard coding theory results and terminology all of which can be found in [14]. An $[n,k,d]$ binary linear code is a subset of $\mathbb{F}_2^n$ which is a vector space of dimension k over $\mathbb{F}_2$ having minimum distance d. We here mention the Griesmer bound (see [14, p. 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 [14, p. 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 [14, p. 550]) which gives a 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

$$\left[n = s(2^k - 1) - \sum_{i=1}^{p} (2^{u_i} - 1), k, d \right]$$

code meeting the Griesmer bound with equality if

$$\sum_{i=1}^{\min(s+1,p)} u_i \leq sk$$

(see [14, p. 552]). This condition is satisfied for most values of d and k.

D. Some Recent Constructions

Here we summarize the previous construction results.

1) Zhang and Zheng [25]: This is the paper to provide an elegant general construction of nonlinear resilient S-boxes. The same idea was also present in Camion and Canteaut [2]. The main result proved is as follows [25, 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 [13, 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 [12]: They use a linear error-correcting code to build a matrix A of small affine functions. Resiliency and nonlinearity is ensured by using nonintersecting codes along with the matrix A. The actual nonintersecting 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 [17]: Pasalic and Maitra use the matrix A of the 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 [6, Theorem 5]: Cheon 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 as follows. If there exists an $[n,m,t]$ linear code then for any nonnegative integer D there exists an $(n+D+1,m,t-1)$-resilient function with algebraic degree D and nonlinearity at least

$$\left(2^{n+D} - 2^n \left\lfloor \sqrt{2^{n+D+1}} \right\rfloor + 2^{n-1}\right).$$

To date, this is the only construction which provides (n,m,t) nonlinear resilient S-boxes with degree greater than m.

III. 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 1: Let $f:\{0,1\}^n \to \{0,1\}^m$ and $g:\{0,1\}^m \to \{0,1\}$. Then for any $w \in \mathbb{F}_2^n$

$$W_{(g\circ f)}(w) = \frac{1}{2^m} \sum_{v \in \mathbb{F}_2^m} W_g(v) W_{(l_v \circ f)}(w)$$

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

Proof: By (2), we have

$$(-1)^{g(x)} = \frac{1}{2^m} \sum_{w \in \mathbb{F}_2^m} W_g(w)(-1)^{\langle w,x \rangle}.$$

Hence,

$$\begin{aligned}(-1)^{(g\circ f)(x)} &= (-1)^{g(f(x))} \\ &= \frac{1}{2^m} \sum_{v \in \mathbb{F}_2^m} W_g(v)(-1)^{\langle v,f(x) \rangle} \\ &= \frac{1}{2^m} \sum_{v \in \mathbb{F}_2^m} W_g(v)(-1)^{(l_v\circ f)(x)}.\end{aligned}$$

By (1), we have

$$\begin{aligned}W_{g\circ f}(w) &= \sum_{x \in \mathbb{F}_2^n} (-1)^{(g\circ f)(x) \oplus \langle w,x \rangle} \\ &= \frac{1}{2^m} \sum_{x \in \mathbb{F}_2^n} \sum_{v \in \mathbb{F}_2^m} W_g(v) \\ &\quad \times (-1)^{(l_v\circ f)(x) \oplus \langle w,x \rangle}\end{aligned}$$

$$= \frac{1}{2^m}\sum_{v\in\mathbb{F}_2^m} W_g(v) \times \sum_{x\in\mathbb{F}_2^n}(-1)^{(l_v\circ f)(x)\oplus(w,x)} = \frac{1}{2^m}\sum_{v\in\mathbb{F}_2^m} W_g(v)W_{(l_v\circ f)}(w). \qquad \blacksquare$$

Corollary 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.

Proof: Since f is balanced, $W_{(l_v\circ f)}(w) = 0$ for all nonzero $v\in\mathbb{F}_2^m$. Thus,

$$W_{g\circ f}(0) = \frac{1}{2^m}W_g(0)2^m = W_g(0). \qquad \square$$

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 1 and Corollary 1 provide the following theorem.

Theorem 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 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.

IV. Construction of (n,m,t)-Resilient S-Box With Degree $> m$

In this section, we modify an elegant construction by Zhang and Zheng [25] to obtain high degree nonlinear resilient S-boxes. The Zhang–Zheng construction shows that highly nonlinear resilient functions can be constructed from linear resilient functions by applying highly nonlinear permutations in the transforming process. We take permutation to be an inverse function and then drop $(d+1-m)$ columns from the output. The following result is well known (see, for example, [25]).

Theorem 3: Let C be a $[n,m,t+1]$ binary linear code. Then we can construct an linear (n,m,t)-resilient function.

Modified Zhang–Zheng (MZZ) Construction

1. Input: Number of output columns $= m$, degree $= d\ge m$, 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 $\deg(G) = d$, $\mathsf{nl}(G) = 2^d - 2^{\lfloor\frac{d+1}{2}\rfloor}$. Then

$$\mathsf{nl}(g)\ge 2^{n-d-1}\left(2^d - 2^{\lfloor\frac{d+1}{2}\rfloor}\right) = 2^{n-1} - 2^{n-\lfloor\frac{d+1}{2}\rfloor}$$

and $\deg(g) = d$.

3. Drop $(d+1-m)$ columns from the output of g to obtain an (n,m,t)-resilient function with degree d and nonlinearity $2^{n-1}-2^{n-\lceil\frac{d+1}{2}\rceil}$.

In Step 2, we choose the function G to be the inverse function over GF(2^{d+1}) (with respect to a fixed irreducible polynomial). Then the nonlinearity of G is $2^d - 2^{\lfloor\frac{d+1}{2}\rfloor}$ and is given in [16]. There are other bijections by which we get the same value of $\mathsf{nl}(G)$ but $\deg(G) = d$ is achieved only for G obtained from the inverse map over GF(2^{d+1}) (see [4]). The fact that $g = G\circ f$ is t-resilient if f is t-resilient is given in a more general form in [2] and also appears in [25].

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 the 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 [6]. 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 of Cheon [6]. We know of no other work where such provable comparisons of construction has been presented.

Theorem 4: Let n,m,d,t be such that the following two conditions hold.

1) Either a) $d<m$ or b) $d\ge m\ge\log_2(t+1)$.
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's method [6].

Proof: Recall the Cheon construction from Section II-D. Given any $[N,M,T+1]$ and a nonnegative 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, $d+1$, $t+1$ satisfy the 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) $d\ge m\ge\log_2(t-1)$ then

$$n-d-1<\sum_{i=0}^{m-1}\left\lceil\frac{t-1}{2^i}\right\rceil.$$

Proof of the Claim: Since $n = \sum_{i=0}^{d}\lceil\frac{t+1}{2^i}\rceil$ we have that

$$n-d-1<\sum_{i=0}^{m-1}\left\lceil\frac{t+1}{2^i}\right\rceil$$

if and only if

$$\sum_{i=0}^{d}\left\lceil\frac{t+1}{2^i}\right\rceil - d - 1 < \sum_{i=0}^{m-1}\left\lceil\frac{t+1}{2^i}\right\rceil.$$

If $d<m$, then the last mentioned condition is trivially true. So suppose $d\ge m\ge\log_2(t+1)$. Then the above inequality holds if and only if

$$\sum_{i=m}^{d}\left\lceil\frac{t+1}{2^i}\right\rceil < d+1.$$

Since $m\ge\log_2(t+1)$,

$$\sum_{i=m}^{d}\left\lceil\frac{t+1}{2^i}\right\rceil = d-m+1<d+1, \qquad \text{for } m>1.$$

This completes the proof of the claim.

Since

$$n - d - 1 < \sum_{i=0}^{m-1} \left\lceil \frac{t+1}{2^i} \right\rceil$$

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, the Cheon method cannot be used to construct the function f. □

The following result is a consequence of Theorem 4 and the MZZ construction.

Theorem 5: 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's method [6].

Remark: As mentioned in [14, p. 550] there is a large class of codes which meet the Griesmer bound with equality. 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 the MZZ construction but cannot be constructed using the Cheon [6] construction. See Section VI for some concrete examples.

Nonlinearity in the Cheon method is

$$\left(2^{N+D} - 2^{N} \left\lfloor \sqrt{2^{N+D+1}} \right\rfloor + 2^{n-1}\right)$$

(see item 5 of Section II-D) which is positive if $D > N + 1$ for $N > 2$. So, for $D \leq N$, the Cheon method does not provide any nonlinearity. Thus, the 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 the MZZ method is larger than nonlinearity obtained by the Cheon method.

Theorem 6: Let f be an (n, m, t)-resilient function f of degree d and nonlinearity n_1 constructed by the 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 the MZZ method. Further, $n_2 \geq n_1$.

Proof: Since an $[n, d+1, t+1]$ code exists, the MZZ construction can be applied to obtain an (n, m, t)-resilient function g with degree d and nonlinearity $\mathrm{nl}(g) = n_2 = 2^{n-1} - 2^{n - \lceil \frac{d+1}{2} \rceil}$. It remains to show that $n_2 \geq n_1$, which we show now. Recall that

$$n_1 = 2^{n-1} - 2^{n-d-1} \lfloor \sqrt{2^n} \rfloor - 2^{n-d-2}.$$

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 holds 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 $d = n - 2$, the Cheon construction requires $[0, m, t + 1]$ and $[1, m, t + 1]$ codes, respectively. Clearly, such codes do not exist. Hence, $n_2 \geq n_1$ holds for all d. □

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

Proof: Since f has been obtained from the 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

$$n - d - 1 \geq \sum_{i=0}^{m-1} \left\lceil \frac{t+1}{2^i} \right\rceil$$

i.e, we have

$$n \geq d + 1 + \sum_{i=0}^{m-1} \left\lceil \frac{t-1}{2^i} \right\rceil.$$

As $m \geq \log_2(t + 1)$ we have $\left\lceil \frac{t-1}{2^i} \right\rceil = 1$ for $i \geq m$. Hence,

$$n \geq (d + 1) - (d - m + 1) + \sum_{i=m}^{d} \left\lceil \frac{t+1}{2^i} \right\rceil + \sum_{i=0}^{m-1} \left\lceil \frac{t-1}{2^i} \right\rceil.$$

This shows

$$n \geq m + \sum_{i=0}^{d} \left\lceil \frac{t-1}{2^i} \right\rceil$$

and consequently

$$n \geq \sum_{i=0}^{d} \left\lceil \frac{t+1}{2^i} \right\rceil.$$

Thus, the parameters n, $d + 1$, $t + 1$ satisfy the Griesmer bound. □

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 [14, p. 550]) and apply Theorem 6. In fact, we do not know of any case where a function can be constructed using the Cheon method but not by the MZZ method. Theorems 5 and 6 *prove* the clear advantage of the MZZ method over the Cheon construction. Thus, the MZZ method is the currently known best method to construct $[n, m, t]$-resilient function with degree $d > m$.

V. 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 [12].

Theorem 7: Let C be a $[n, m, t + 1]$ code. Then it is possible to construct $(2^m - 1) \times m$ matrix D with entries from C, such that

$$\{c_1 D_{i,1} \oplus \cdots \oplus c_m D_{i,m} : 1 \leq i \leq 2^m - 1\} = C \setminus \{(0, \ldots, 0)\}$$

for each nonzero vector $(c_1, \ldots, c_m) \in \mathbb{F}_2^m$.

Let D be the matrix in Theorem 7. For $(1 \leq i \leq 2^m - 1)$ and $(1 \leq j \leq m)$, define a n-variable linear function

$$L_{i,j}(x_1, \ldots, x_n) \triangleq \langle D_{i,j}, (x_1, \ldots, x_n) \rangle.$$

Given the code C, we define a $(2^m - 1) \times m$ matrix $L(C)$ whose entries are n-variable linear functions by defining the i, jth entry of $L(C)$ to be $L_{i,j}(x_1, \ldots, x_n)$. We have the following result which follows directly from Theorem 7.

Proposition 1: Let $c \in \mathbb{F}_2^u$ 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 $k < l$, 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 $G(y_1,\ldots,y_p)$ be a (p,m) S-box whose component functions are $G_1,\ldots,G_m$. We define $G \dotplus L(C,k,l)$ to be an $(l-k+1) \times m$ matrix whose i,jth entry is

$$G_j(y_1,\ldots,y_p) \oplus L_{k+i-1,j}(x_1,\ldots,x_u)$$

for $1 \leq i \leq l-k+1$ and $1 \leq j \leq m$. If $l-k+1 = 2^r$ for some r, then $G \dotplus L(C,k,l)$ defines an S-box $F : \{0,1\}^{r+p+u} \rightarrow \{0,1\}^m$ in the following manner:

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

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

In the matrix $M = G(y_1,\ldots,y_p) \oplus L(C,k,l)$ we say that the row $L_{i,*}$ of $L(C)$ is repeated 2^p times. Let $G(y_1,\ldots,y_p)$ and $H(y_1,\ldots,y_q)$ 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 $[M_1 \quad M_2]^T$.

Proposition 1 has also been used by [17] in the construction of resilient S-boxes. However, we improve upon the construction of [17] by utilizing the following two ideas.

1) We use all the $2^m - 1$ rows of the matrix $L(C)$. In contrast, [17] uses at most 2^{m-1} rows of $L(C)$.
2) We allow a row of $L(C)$ to be repeated 2^{i_1} or $2^{i_1} + 2^{i_2}$ or $2^{i_1} + 2^{i_2} + 2^{i_3}$ times as required. On the other hand, the number of times a row of $L(C)$ can be repeated in [17] is of the form 2^i.

It turns out that a proper utilization of the above two techniques results 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 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 [9].
3) If $r \geq 2m$ and r is even, G is a perfect nonlinear S-box [16].
4) If $r \geq 2m$ and r is odd, G is concatenation of two perfect nonlinear S-boxes (see Section II-B).

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 2: Let G be an (r,m) PROPER S-box. Then

1) If $r < m$, $\mathsf{nl}(G) = 0$.
2) If $m \leq r < 2m$, then $\mathsf{nl}(G) = 2^{r-1} - 2^{\frac{r-1}{2}}$ if r is odd and $\mathsf{nl}(G) \geq 2^{r-1} - 2^{\frac{r}{2}}$ if r is even.
3) If $r \geq 2m$, then $\mathsf{nl}(G) = 2^{r-1} - 2^{\frac{r}{2}-1}$ if r is even and $\mathsf{nl}(G) = 2^{r-1} - 2^{\frac{r-1}{2}}$ 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 $list = \langle(n_1,R_1),(n_2,R_2),\ldots,(n_k,R_k)\rangle$ 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 8). 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 $\mathsf{effect} = 1$.
   (b) $m \leq n - u < 2m - 1$; $list = \langle(2^{m-1},2^{n-u-m-1})\rangle$ and $\mathsf{effect} = 2^{n-u-m-1}$.
   (c) $n - u = 2m - 1$; $list = \langle(2^{m-1},2^m)\rangle$ and $\mathsf{effect} = 2^{\lfloor\frac{m}{2}\rfloor+1}$
   (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 $\mathsf{effect} = 2^{e+1+\frac{m}{2}}$.
      (ii) $n - u = 2m + 2e + 1$; m even; $0 \leq e \leq \frac{m}{2} - 1$;
      - $0 \leq e \leq \frac{m}{2} - 2$; $list = \langle(2,2^{m+2e+1}+2^{2e-1}+2^{2e}),(2^m-3,2^{m+2e-1}+2^{2e+1})\rangle$ and $\mathsf{effect} = 2^{2e-1}+2^{2e}+2^{e-1+\frac{m}{2}}$.
      - $e = \frac{m}{2} - 1$; $list = \langle(2^{m-1},2^{2m})\rangle$ and $\mathsf{effect} = 2^m$.
      (iii) $n - u = 2m + 2e + 1$; m odd; $0 \leq e \leq \lfloor\frac{m}{2}\rfloor - 1$; $list = \langle(1,2^{m+2e+2}),(2^m-2,2^{m+2e+1})\rangle$ and $\mathsf{effect} = 2^{\frac{m+1+2e}{2}}$.
      (iv) $n - u = 2m + 2e$; m odd; $0 \leq e < \lfloor\frac{m}{2}\rfloor$; $list = \langle(2^m-2,2^{m+2e}+2^{2e+1}),(1,2^{2e+2})\rangle$ and $\mathsf{effect} = 2^{2e-1}+2^{\frac{m+2e-1}{2}}$.
      (v) $n - u = 3m - 1$; m odd; $list = \langle(2^{m-1},2^{2m})\rangle$ and $\mathsf{effect} = 2^m$
   (e) $n - u \geq 3m$.
      (i) $n - u = 3m + 2e + 1$; $e \geq 0$; $list = \langle(2^{m-1},2^{2m+2e+2})\rangle$ and $\mathsf{effect} = 2^{m+e+1}$.
      (ii) $n - u = 3m + 2e$; (m even; $e \geq \frac{m}{2}$) or (m odd; $0 \leq e < \lfloor\frac{m}{2}\rfloor$); $list = \langle(2,2^{2m+2e}+2^{m+2e}-2^{m-2e-1}),(2^m-3,2^{2m+2e}+2^{m-2e})\rangle$ and $\mathsf{effect} = 2^{m+e}+2^{e-1+\frac{m}{2}}$.
      (iii) $n - u = 3m + 2e$; m even; $0 \leq e < \frac{m}{2}$ $list = \langle(2^m-2,2^{2m+2e}+2^{m+2e+1}),(1,2^{m+2e+2})\rangle$ and $\mathsf{effect} = 2^{m+e}+2^{e-1+\frac{m}{2}}$.
      (iv) $n - u = 3m + 2e$; m odd; $e \geq \lfloor\frac{m}{2}\rfloor$; $list = \langle(2^m-2,2^{2m+2e}+2^{m-2e+1}),(1,2^{m+2e-2})\rangle$ and $\mathsf{effect} = 2^{m+e}+2^{e+\frac{m-1}{2}}$.

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 $F = G \dotplus L(C,1,2^s)$.
   - 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 $e \neq 0$
   - 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 8: 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) $\mathsf{nl}(f) = (2^{n-1} - 2^{u-1} \times \mathsf{effect})$.

Proof of a): The output of Part A is a

$$list = ((n_1, R_1), (n_2, R_2), \ldots, (n_k, R_k)).$$

Part B ensures that for $1 \leq i \leq k$, 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

$$list = ((2, 2^{2m+2e} + 2^{m-2e} + 2^{m+2e-1}), (2^m - 3, 2^{2m+2e} + 2^{m-2e})).$$

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 nonconstant m-variable linear function and λ is an n-variable linear function, then $d(\nu \circ F, \lambda) \geq (2^{n-1} - 2^{u-1} \times \mathsf{effect})$. We write λ as

$$\lambda(y_1, \ldots, y_{n-u}, x_1, \ldots, x_u) = \lambda_1(y_1, \ldots, y_{n-u}) \oplus \lambda_2(x_1, \ldots, x_u).$$

Let $\nu(z_1, \ldots, z_m) = \langle (c_1, \ldots, c_m), (z_1, \ldots, z_m) \rangle$ for some nonzero vector $c = (c_1, \ldots, c_m) \in \mathbb{F}_2^m$. 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 $1 \leq i \leq 2$

$$\nu \circ F_i = (\nu \circ G_i) \oplus (L(C) c^T)$$

and

$$\nu \circ F_3 = (\nu \circ G_3) \oplus (L(C, 1, 2) c^T).$$

Using Proposition 1, we know that all the entries of the column vector $L(C) c^T$ are distinct u-variable linear functions. Let $L(C) c^T = [\mu_1, \ldots, \mu_{2^m-1}]^T$. 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 $\mu_3, \ldots, \mu_{2^m-1}$ are repeated R_2 times in the construction of $\nu \circ F$. If $\lambda \notin \{\mu_1, \ldots, \mu_{2^m-1}\}$ then $d(\lambda_2, \mu_i) = 2^{u-1}$ for each $1 \leq i \leq 2^m - 1$ and hence $d(\nu \circ F, \lambda) = 2^{n-u}(2^{u-1}) = 2^{n-1}$. Now suppose $\lambda_2 = \mu_i$ for some $i \in \{1, \ldots, 2^m - 1\}$. 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 $2 \leq i \leq 2^m - 1$. The function μ_2 is repeated R_1 times and each of the functions $\mu_3, \ldots, \mu_{2^m-1}$ is repeated R_2 times. So the total contribution of $\mu_2, \mu_3, \ldots, \mu_{2^m-1}$ 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

$$2^u(\mathsf{nl}(\nu \circ G_1) - \mathsf{nl}(\nu \circ G_2) + \mathsf{nl}(\nu \circ G_3)) \\ = 2^u(\mathsf{nl}(G_1) + \mathsf{nl}(G_2) + \mathsf{nl}(G_3))$$

since $\mathsf{nl}(\nu \circ G_i) = \mathsf{nl}(G_i)$. Each G_i is a PROPER function whose nonlinearity is given by Proposition 2.

Hence,

$$\begin{aligned} d(\nu \circ F, \lambda) &= 2^{u-1}\Big(R_1 + (2^m - 3)R_2 - 2(\mathsf{nl}(G_1) \\ &\qquad + \mathsf{nl}(G_2) + \mathsf{nl}(G_3))\Big) \\ &= 2^{u-1}(2^{n-u} - (R_1 - 2(\mathsf{nl}(G_1) \\ &\qquad + \mathsf{nl}(G_2) + \mathsf{nl}(G_3)))) \\ &= 2^{n-1} - 2^{u-1} \\ &\qquad \times (R_1 - 2(\mathsf{nl}(G_1) + \mathsf{nl}(G_2) + \mathsf{nl}(G_3))). \end{aligned}$$

From the given conditions, it is easy to verify that

$$\mathsf{effect} = R_1 - 2(\mathsf{nl}(G_1) + \mathsf{nl}(G_2) - \mathsf{nl}(G_3))$$

and so

$$d(\nu \circ F, \lambda) = (2^{n-1} - 2^{u-1} \times \mathsf{effect}).$$

TABLE I
COMPARISON OF NONLINEARITY OBTAINED BY MZZ CONSTRUCTION TO THAT OBTAINED BY CHEON [6]

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

TABLE II
COMPARISON OF CONSTRUCTION-I NONLINEARITY WITH THE NONLINEARITY OF [17]

Case	Nonlinearity of [17]	Construction-I nonlinearity	
$2m \le n-u < 3m-3$, π even	$2^{n-1}-2^{(n-u-m+1)/2}$	$2^{n-1}-2^{(n+u-m-1)/2}-3\times 2^{u-2m-2}$ $2^{n-1}-2^{(n-u-m-1)/2}-2^{u-2m}$	**(1)** **(2)**
$2m \le n-u < 3m-3$, π odd $n-u=3m-3$	$2^{n-1}-2^{(n-u-m+2)/2}$ $2^{n-1}-2^{(u+m-1)}$	$2^{n-1}-2^{(n-u-m)/2}$ $2^{n-1}-\frac{11}{13}2^{(u+m-1)}$	**(3)** **(4)**
$n-u > 3m$, π odd	$2^{n-1}-2^{(n+u-m)/2}$	$2^{n-1}-2^{(n+u-m)/2}(\frac{1}{2}+\frac{1}{2^{m/2}})$ $2^{n-1}-2^{(n-u-m)/2}(\frac{1}{2}+\frac{1}{2^{(m-1)/2}})$	**(5)** **(6)**

TABLE III
COMPARISON OF CONSTRUCTION-I NONLINEARITY WITH [17] FOR $m=4$ AND RESILIENCY $= \mathbf{1,2,3}$

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

Case 2: $\lambda_i = \mu_i$ for some $i \in \{3,\ldots,2^m-1\}$. In this case, we proceed as in the previous case to obtain

$$\begin{aligned} d(\nu\circ F,\lambda) &= 2^{u-2}(2R_1+(2^m-4)R_2)+2^u(\mathsf{nl}(G_1)+\mathsf{nl}(G_2)) \\ &= 2^{u-1}(2R_1+(2^m-4)R_2)+2(\mathsf{nl}(G_1)+\mathsf{nl}(G_2)) \\ &= 2^{u-1}\left(2^{n-u}-R_2+2(\mathsf{nl}(G_1)-\mathsf{nl}(G_2))\right) \\ &= 2^{n-1}-2^{u-1}(R_2-2(\mathsf{nl}(G_1)+\mathsf{nl}(G_2))) \\ &\geq 2^{n-1}-2^{u-1}\times \mathsf{effect} \end{aligned}$$

since

$\mathsf{effect}=R_1-2(\mathsf{nl}(G_1)+\mathsf{nl}(G_2)+\mathsf{nl}(G_3)) \geq R_2-2(\mathsf{nl}(G_1)+\mathsf{nl}(G_2))$.

By Cases 1 and 2, it follows that

$$\mathsf{nl}(\nu\circ F)=2^{n-1}-2^{u-1}\times \mathsf{effect}.$$

Hence, $\mathsf{nl}(F)=2^{n-1}-2^{u-1}\times \mathsf{effect}$. ∎

VI. RESULTS AND COMPARISONS

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

A. *Degree Comparison Based on MZZ Construction*

We present examples to show the advantage of the MZZ method over the Cheon method. The 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 an (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)–c), an (n,m,t)-resilient function with degree d can be constructed using the MZZ method, but cannot be constructed using the Cheon method (see Theorem 5). Now we present some examples where both the MZZ and Cheon methods construct (n,m,t)-resilient functions with degree d and compare their nonlinearity using Theorem 6. An (n,m,d,t) S-box is an (n,m,t)-resilient S-box with degree d.

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

B. *Nonlinearity Comparison Based on Construction-I*

We compare the nonlinearity obtained by Construction-I to the nonlinearity obtained in [17, Theorem 4]. The nonlinearity obtained in [17] 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 [12] provides better nonlinearity than [17].

Our first observation is that the nonlinearity obtained by Construction-I is at least as large the nonlinearity obtained in [17]. 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 [17]. The detailed verification of the superiority of Construction-I over [17] is straightforward but tedious. In Table II, we summarize the cases under which Construction-I yields higher nonlinearity than [17]. 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 Tables III–V, we provide some concrete examples of cases where the nonlinearity obtained by Construction-I is better than that obtained by [17]. Each entry of Tables III–V is of the form (a,b), where a is the nonlinearity obtained by [17] and b is the nonlinearity obtained by Construction-I.

TABLE IV
COMPARISON OF CONSTRUCTION-I NONLINEARITY WITH [17] FOR $m = 5$ AND RESILIENCY $= 1, 2, 3$

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

TABLE V
COMPARISON OF CONSTRUCTION-I NONLINEARITY WITH [17] FOR $m = 6$ AND RESILIENCY $= 1, 2, 3$

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

TABLE VI
COMPARISION OF CONSTRUCTION-I NONLINEARITY OF $(36, 8, t)$-RESILIENT S-BOXES USING DIFFERENT METHODS

t	7	6	5	4	3	2	1
[13]	$2^{35}-2^{27}$	$2^{35}-2^{27}$	$2^{35}-2^{26}$	$2^{35}-2^{25}$	$2^{35}-2^{24}$	$2^{35}-2^{23}$	$2^{35}-2^{22}$
[12]	$2^{35}-2^{24}$		$2^{35}-2^{23}$	$2^{35}-2^{22}$	$2^{35}-2^{22}$	$2^{35}-2^{21}$	$2^{35}-2^{20}$
[17]	$2^{35}-2^{26}$	$2^{35}-2^{24}$	$2^{35}-2^{23}$	$2^{35}-2^{22}$	$2^{35}-2^{20}$	$2^{35}-2^{20}$	$2^{35}-2^{18}$
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	$[20, 8, 8]$	$[10, 8, 7]$	$[17, 8, 6]$	$[16, 8, 5]$	$[13, 8, 4]$	$[12, 8, 3]$	$[9, 8, 2]$

The linear codes used in Table III are $[5, 4, 2]$, $[7, 4, 3]$, and $[8, 4, 4]$. The second, fourth, and sixth rows give the nonlinearity of (n, m, t)-resilient functions corresponding to the codes $[5, 4, 2]$, $[7, 4, 3]$, and $[8, 4, 4]$, respectively, for different values of n. The linear codes used in Table IV are $[6, 5, 2]$, $[9, 5, 3]$, and $[10, 5, 4]$.

The linear codes used in Table V are $[7, 6, 2]$, $[10, 6, 3]$, and $[10, 6, 4]$. Nonlinearity of $(36, 8, t)$ resilient S-box has been used as very important examples in [12], [13], [17]. Now we compare our nonlinearity with results in Table VI. The results of [12] are not constructive. They show that a resilient S-box with such parameter exists. Note that, except for resiliencies of order 1 and 3, our nonlinearity is better than nonlinearity of [17]. It should also be noted that in all the cases we provide construction with currently best known nonlinearity.

VII. CONCLUSION

In this correspondence, we considered the construction of nonlinear resilient S-boxes. We proved that the correlation immunity of a resilient S-box is preserved under composition with an arbitrary Boolean function. Our main contribution has been 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 [25]. This provides (n, m, t)-resilient S-boxes with degree $d > m$. We *prove* that the MZZ construction is superior to the only previously known construction [6] which provided degree $d > m$. Our second construction is based on concatenation of small affine functions to build nonlinear resilient S-boxes. We sharpen the technique to construct (n, m, t)-resilient S-boxes with the currently best known nonlinearity.

Algebraic attacks [8] are a new type of attack on stream ciphers. These attacks exploit the fact that even if a function may have high degree, it may have a low degree multiple. In this correspondence, we have not considered algebraic attacks. A possible future work is to identify the possible subclass of functions which can resist algebraic attacks.

REFERENCES

[1] C. Bennett, G. Brassard, and J. Robert, "Privacy amplification by public discussion," *SIAM J. Comput.*, vol. 17, pp. 210–229, 1988.

[2] P. Camion and A. Canteaut, "Construction of t-resilient functions over a finite alphabet," in *Advances in Cryptology—EUROCRYP 1996 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1996, pp. 283–293.

[3] P. Camion, C. Carlet, P. Charpin, and N. Sendrier, "On correlation immune functions," in *Advances in Cryptology—CRYPT0 1991 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1992, pp. 86–100.

[4] F. Chabaud and S. Vaudenay, "Links between differential and linear cryptoanalysis," in *Advances in Cryptology—EUROCRYPT 1994 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1995, pp. 356–365.

[5] S. Chee, S. Lee, D. Lee, and S. H. Sung, "On the correlation immune functions and their nonlinearity," in *Advances in Cryptology—Asiacrypt 1996 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1996, pp. 232–243.

[6] J. H. Cheon, "Nonlinear vector resilient functions," in *Advances in Cryptology—CRYPTO 2001 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2001, pp. 458–469.

[7] B. Chor, O. Goldreich, J. Hastad, J. Friedman, S. Rudich, and R. Smolensky, "The bit extraction problem or t-resilient functions," in *Proc. IEEE Symp. Foundations of Computer Science*, vol. 26, 1985, pp. 396–407.

[8] N. Courtois and W. Meier, "Algebraic attacks on stream ciphers with linear feedback," in *Advances in Cryptology—EUROCRYPT 2003 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2003, pp. 345–359.

[9] H. Dobbertin, "Almost perfect nonlinear power functions on GF (2^n): The welch case," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1271–1275, May 1999.

[10] K. C. Gupta and P. Sarkar, "Efficient representation and software implementation of resilient Maiorana-McFarland S-boxes," in *WISA 2004 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, to be published.

[11] ——, "Improved construction of nonlinear S-boxes," in *Advances in Cryptology—Asiacrypt 2002 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2002, pp. 466–483.

[12] T. Johansson and E. Pasalic, "A construction of resilient functions with high nonlinearity," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 494–501, Feb. 2003.

[13] K. Kurosawa, T. Satoh, and K. Yamamoto, "Highly nonlinear t-resilient functions," *J. Universal Comput. Sci.*, vol. 3, no. 6, pp. 721–729, 1997.

[14] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[15] K. Nyberg, "Perfect nonlinear S-boxes," in *Advances in Cryptology—EUROCRYPT 1991 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1991, pp. 378–386.

[16] ——, "Differentially uniform mapping for cryptography," in *Advances in Cryptology—EUROCRYPT 1993 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1994, pp. 55–65.

[17] E. Pasalic and S. Maitra, "Linear codes in generalized construction of resilient functions with very high nonlinearity," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2182–2191, Aug. 2002.

[18] B. Preneel, "Analysis and design of cryptographic hash functions," Ph.D. dissertation, K.U. Leuven, 1993.

[19] O. S. Rothaus, "On bent functions," *J. Comb. Theory*, ser. A, vol. 20, pp. 300–305, 1976.

[20] P. Sarkar and S. Maitra, "Construction of nonlinear boolean functions with important cryptographic properties," in *Advances in Cryptology—EUROCRYPT 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2000, pp. 485–506.

[21] J. Seberry, X.-M. Zhang, and Y. Zheng, "On construction and nonlinearity of correlation immune boolean functions," in *Advances in Cryptology—EUROCRYPT 1993 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1994, pp. 181–199.

[22] T. Siegenthaler, "Correlation-immunity of nonlinear combining functions for cryptographic applications," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 5, pp. 776–780, Sep. 1984.

[23] D. R. Stinson and J. L. Massey, "An infinite class of counterexamples to a conjecture concerning nonlinear resilient functions," *J. Cryptol.*, vol. 8, pp. 167–173, 1995.

[24] G. Xiao and J. L. Massey, "A spectral characterization of correlation-immune combining functions," *IEEE Trans. Inf. Theory*, pp. 569–571, 1988.

[25] X.-M. Zhang and Y. Zheng, "On cryptographically resilient functions," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1740–1747, May 1997.

The Maximum Squared Correlation, Sum Capacity, and Total Asymptotic Efficiency of Minimum Total-Squared-Correlation Binary Signature Sets

George N. Karystinos, *Member, IEEE,* and
Dimitris A. Pados, *Member, IEEE*

***Abstract*—The total squared correlation (TSC), maximum squared correlation (MSC), sum capacity ($C_{\rm sum}$), and total asymptotic efficiency (TAE) of underloaded signature sets, as well as the TSC and $C_{\rm sum}$ of overloaded signature sets are metrics that are optimized simultaneously over the real/complex field. In this present work, closed-form expressions are derived for the MSC, $C_{\rm sum}$, and TAE of minimum-TSC binary signature sets. The expressions disprove the general equivalence of these performance metrics over the binary field and establish conditions on the number of signatures and signature length under which simultaneous optimization can or cannot be possible. The sum-capacity loss of the recently designed minimum-TSC binary sets is found to be rather negligible in comparison with minimum-TSC real/complex-valued (Welch-bound-equality) sets.**

***Index Terms*—Binary sequences, code-division multiple access (CDMA), code division multiplexing, codes, signal design, spread-spectrum communications, Welch bound.**

I. Introduction and Background

In direct-sequence code-division-multiple-access (DS-CDMA) systems, individual user signals use distinct signatures (also known as spreading codes) to access a common, in time and frequency, communication channel. In conjunction with channel and receiver design specifics, the overall system performance is determined by the selection of the user signature set. Signature set metrics of interest include the total squared correlation (TSC) [1]–[6], maximum squared correlation (MSC) [1], [2], sum capacity $C_{\rm sum}$ [2], and total asymptotic efficiency (TAE) [7], [8]. We recall the definitions of these metrics below.

If

$$\boldsymbol{S} \triangleq [\boldsymbol{s}_1\ \boldsymbol{s}_2\ \ldots\ \boldsymbol{s}_K], \qquad \boldsymbol{s}_i \in \mathbb{C}^L,\ \|\boldsymbol{s}_i\| = 1,\ i = 1, 2, \ldots, K$$

is an $L \times K$ matrix that represents a set of K normalized (complex-valued in general) user signatures of length (spreading gain) L, then

i) the TSC of $\boldsymbol{S}$ is the sum of the squared magnitudes of all inner products between signatures

$$\text{TSC}(\boldsymbol{S}) \triangleq \sum_{i=1}^{K}\sum_{j=1}^{K} \boldsymbol{s}_i^H \boldsymbol{s}_j \Big|^2; \tag{1}$$

Manuscript received February 25, 2003; revised September 14, 2004. This work was supported in part by the National Science Foundation under Grant CCR-0219903 and the U.S. Air Force Office of Scientific Research under Grant FA9550-04-1-0256. The material in this correspondence was presented in part at the Conference on Information Sciences and Systems, Baltimore, MD, March 2003 and in part at IEEE Global Telecommunications Conference, Communications Theory Symposium, San Fransisco, CA, December 2003.

G. N. Karystinos was with the Department of Electrical Engineering, State University of New York at Buffalo, Amherst, NY 14260 USA. He is now with the Department of Electrical Engineering, Wright State University, Dayton, OH 45435 USA (e-mail: g.karystinos@wright.edu).

D. A. Pados is with the Department of Electrical Engineering, State University of New York at Buffalo, Amherst, NY 14260 USA (e-mail: pados@eng.buffalo.edu).

Communicated by D. N. C. Tse, Associate Editor for Communications.
Digital Object Identifier 10.1109/TIT.2004.839542