Highly Nonlinear Balanced S-Boxes With Good Differential Properties

WeiGuo Zhang (张卫国), *Senior Member, IEEE*, and Enes Pasalic

***Abstract*— Substitution boxes (S-boxes) play a central role in the modern design of iterative block ciphers. While in substitution-permutation networks the S-boxes are bijective, thus ensuring the invertibility of the encryption algorithm, the property of being bijective is not mandatory for Feistel kind of networks. In this paper, two methods of constructing highly nonlinear balanced S-boxes (whose nonlinearity $> 2^{n-1} - 2^{n/2}$ is better than the nonlinearity of the commonly used inverse S-box) with good algebraic and differential properties are given. The first method employs two vectorial Boolean functions from the Maiorana–McFarland class that need to fulfill certain conditions. In particular, these conditions are shown to be satisfied by maximum length sequences. The second method is based on a suitable modification of a certain class of vectorial bent functions. The differential properties of these boxes, measured as a deviation from an optimal uniform distribution, also appear to be better than those of the inverse S-box. Both methods are susceptible to further optimizations of the relevant cryptographic parameters due to the underlying design ideas.**

***Index Terms*— S-boxes, bent functions, differential properties, maximum-length sequences, substitution permutation networks, Feistel networks.**

I. Introduction

A FUNCTION $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, is called an (n, m) substitution box (S-box) and it represents a basic cryptographic primitive frequently used in the design of symmetric key algorithms. In iterative block ciphers, it is typically used to conceal the relationship between the key and the ciphertext, thus embedding the Shannon's concept of confusion [14]. Linear attacks [10] and differential attacks [3] are two main powerful attack on block ciphers, and the link between the two approaches was investigated in [5]. To ensure a good resistance to linear attacks, the nonlinearity of S-boxes must be as high as possible. To resist differential attacks, S-boxes with a low differential uniformity are desirable.

For even n, S-boxes achieving the maximum possible nonlinearity are called vectorial bent functions, and can only exist for $m \leq n/2$ [12]. Moreover, the output distributions of all derivatives of the vectorial bent functions are uniform but unfortunately vectorial bent functions are not balanced S-boxes. In the modern design of block ciphers two commonly adopted structures are substitution permutation networks (SPN) used in e.g. Advanced Encryption Standard (AES), and Feistel kind of networks used originally in Data Encryption Standard (DES). While the first approach requires the use of bijective S-boxes for making the encryption algorithm invertible, the latter approach is not restricted to bijective S-boxes only. For instance, a recent proposal PICARO [13] makes use of a non-bijective mapping in a Feistel type of network.

Nevertheless, while the use of bijective mappings is enforced by the structure of SPN the situation is different when Feistel kind of networks are considered. Though from the implementation point of view the use of symmetric S-boxes that substitute n-bit input blocks by n-bit output blocks might be more efficient, the use of smaller output blocks in Feistel kind of networks may improve the most important cryptographic properties by using such asymmetric S-boxes. Along these lines, the initiative towards improving the (non)linear and differential properties of the S-boxes used in DES was taken in [8]. This paper, therefore, focuses on the design of balanced S-boxes characterized by the property of having an exceptionally high resistance to linear and differential cryptanalysis. The resistance to linear cryptanalysis of the designed S-boxes is provably better than the resistance of some commonly used S-boxes such as the inverse S-box used in the AES, and much better than the nonlinearity of a non-injective mapping used in the PICARO block cipher. If the standard size of these boxes (being $(8, 8)$) is considered, the nonlinearity of the inverse S-box is 112 and the nonlinearity of the non-bijective mapping used in PICARO is only 94, whereas our both methods generate $(8, 4)$ S-boxes with nonlinearity 116.

Apart from having a higher resistance to linear cryptanalysis, our S-boxes also offer better resistance to differential cryptanalysis. We show that both our designs give a smaller standard deviation from the uniform distribution of differential values. In order to make a fair comparison between our asymmetric S-boxes and the symmetric inverse S-box, we consider a punctured version of the inverse S-box (obtained by deleting

Manuscript received June 14, 2014; revised September 18, 2014; accepted September 23, 2014. Date of publication September 30, 2014; date of current version November 18, 2014. This work was supported in part by the Open Foundation of State Key Laboratory of Networking and Switching Technology through the Beijing University of Posts and Telecommunications, Beijing, China, under Grant SKLNST-2013-1-07, in part by the 111 Project under Grant B08038, in part by the National Natural Science Foundation of China under Grant 11201359, Grant 61373008, and Grant 61003299, and in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2012JM8041.

W.-G. Zhang is with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China, and also with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: weiguozhang@vip.qq.com).

E. Pasalic is with FAMNIT, University of Primorska, Koper 6104, Slovenia (e-mail: enes.pasalic6@gmail.com).

Communicated by T. Helleseth, Associate Editor for Sequences.

Digital Object Identifier 10.1109/TIT.2014.2360880

$n/2$ outputs) and demonstrate that our differential values are more concentrated around the optimal value corresponding to a uniform distribution of differential values. The first design applies the idea of Maiorana-McFarland construction twice and the set of conditions for this method is shown to be satisfied through a suitable selection of maximum length sequences. The m component Boolean functions constructed in this way give rise to an (n, m) S-box so that these m functions, both individually and also their linear combinations, are characterized by exceptionally good nonlinear and differential properties.

The second design of balanced $(n, n/2)$ S-boxes gives some interesting theoretical results concerning the necessary and sufficient conditions of transforming vectorial bent functions into highly nonlinear balanced S-boxes. The algebraic structure of these S-boxes also allows a rather straightforward analysis of their differential properties, mainly due to "almost overall bentness" embedded by the design. Due to a small and controlled modification of a vectorial bent functions for the purpose of making these S-boxes balanced, only a small deviation from the uniform distribution is obtained and it is easily shown that neither the component functions nor their linear combinations admit linear structures and their Boolean derivatives are very close to balancedness.

The rest of this paper is organized as follows. Section II introduces some basic definitions. In Section III, we give a construction method for designing highly nonlinear balanced S-boxes based on the use of certain injective vectorial mappings which can be specified in terms of maximum length sequences. The other method to construct balanced S-boxes, based on a suitable modification of vectorial bent functions, is given in Section IV. In Section V, we compare the differential properties of our $(n, n/2)$ S-boxes to the inverse S-box obtained by puncturing $n/2$ output coordinates. Section VI concludes this paper.

II. Preliminaries

In this section we give some most important notions and definitions related to the cryptographic criteria for both Boolean and vectorial Boolean functions.

A. Boolean Functions

Let $\mathcal{B}_n$ denote the set of Boolean functions in n variables. A Boolean function $f(X_n) \in \mathcal{B}_n$ is a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$, where $X_n \in \mathbb{F}_2^n$ and $\mathbb{F}_2^n$ is the vector space of tuples of elements from $\mathbb{F}_2$. $f(X_n)$ is generally represented by its algebraic normal form (ANF):

$$f(X_n) = \bigoplus_{b \in \mathbb{F}_2^n} \lambda_b \left(\prod_{i=1}^{n} x_i^{b_i}\right) \tag{1}$$

where $\lambda_b \in \mathbb{F}_2, b = (b_1, \ldots, b_n)$. The algebraic degree of $f(X_n)$, denoted by $deg(f)$, is the maximal value of $wt(u)$ such that $\lambda_b \neq 0$, where $wt(b)$ denotes the Hamming weight of b. f is called an affine function when $deg(f) = 1$. An affine function with constant term equal to zero is called a linear function. Any linear function on $\mathbb{F}_2^n$ is denoted by:

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

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

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

Let $supp(f) = \{x \in \mathbb{F}_2^n \mid f(x) = 1\}$ denote the support of f. $f \in \mathcal{B}_n$ is said to be balanced if its output column in the truth table contains equal number of 0's and 1's, i.e., $\#supp(f) = 2^{n-1}$ or equivalently $W_f(0) = 0$.

Definition 1: The nonlinearity of a Boolean function $f \in \mathcal{B}_n$ is its distance to the set of all affine functions and is defined as

$$N_f = \min_{\rho \in A(n)} |\{X_n \in \mathbb{F}_2^n : f(X_n) \neq \rho(X_n)\}|$$

where $A(n)$ is the set of all affine functions on $\mathbb{F}_2^n$.

The nonlinearity of f can be obtained through the Walsh transform as follows [11]:

$$N_f = 2^{n-1} - \frac{1}{2}\mathcal{L}(f), \quad \text{where} \quad L(f) = \max_{\omega \in \mathbb{F}_2^n} |W_f(\omega)|. \tag{3}$$

The Parseval's equation [9] states that

$$\sum_{\omega \in \mathbb{F}_2^n} (W_f(\omega))^2 = 2^{2n} \tag{4}$$

and implies that

$$N_f \leq 2^{n-1} - 2^{n/2-1}.$$

The equality occurs if and only if $f \in \mathcal{B}_n$ are bent functions, thus n is even.

B. S-Boxes

An (n, m) S-box can be represented as a mapping $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ which in turn can be viewed as a collection of m Boolean functions so that $F(X_n) = (f_1(X_n), \ldots, f_m(X_n))$, where $f_1, \ldots, f_m \in \mathcal{B}_n$ are called component functions of F. In the whole paper, we identify the vector space $\mathbb{F}_2^n$ with the finite field $\mathbb{F}_{2^n}$ when needed, and such a function F is then expressed as a polynomial in $\mathbb{F}_{2^n}[x]$.

The *minimum degree* of F, denoted by $\deg(F)$ and which is simply called *degree* in this article, is defined as the minimum among the algebraic degrees of all nonzero linear combinations of the component functions of F, namely,

$$\deg(F) = \min_{c \in \mathbb{F}_2^{m*}} \deg\left(\bigoplus_{i=1}^{m} c_i f_i(X_n)\right) \tag{5}$$

where $c = (c_1, \ldots, c_m)$, $\mathbb{F}_2^{m*} = \mathbb{F}_2^m \backslash \{\mathbf{0}\}$.

Definition 2: The nonlinearity of an (n, m) S-box $F = (f_1, f_2, \ldots, f_m)$, denoted by N_F, is defined as

$$N_F = \min_{c \in \mathbb{F}_2^{m*}} N_{F_c} \tag{6}$$

where $F_c = \sum_{i=1}^{m} c_i f_i$. In a similar way using the extended Walsh transform defined as,

$$W_{F_c}(\omega) = \sum_{X_n \in \mathbb{F}_2^n} (-1)^{F_c(X_n) \oplus \omega \cdot X_n}, \tag{7}$$

the nonlinearity is given by

$$N_F = 2^{n-1} - \frac{1}{2} \max_{\substack{c\in\mathbb{F}_2^{m*} \\ \omega\in\mathbb{F}_2^n}} |W_{F_c}(\omega)|. \tag{8}$$

Definition 3: Let F be an (n, m) S-box. For any $a \in \mathbb{F}_2^{n*}$ and $b \in \mathbb{F}_2^m$, we denote

$$\delta_F(a, b) = \#\{x \in \mathbb{F}_2^n,\ F(X_n + a) + F(X_n) = b\} \tag{9}$$

where $\#S$ is the cardinality of any set S. We define

$$\delta(F) = \max_{a\in\mathbb{F}_2^{n*}, b\in\mathbb{F}_2^m} |\delta_F(a, b) - 2^{n-m}|. \tag{10}$$

For both symmetric and asymmetric S-boxes, the smaller $\delta(F)$ implies the better differential properties of F. This measure is however insufficient in capturing the deviation from the differential uniform distribution and a useful parameter is the standard deviation defined below.

Definition 4: Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be an (n, m) S-box. The standard deviation of a given differential distribution from the uniform distribution is defined as,

$$SD(F) = \left(\frac{\sum_{a\in\mathbb{F}_2^{n*}} \sum_{b\in\mathbb{F}_2^m} (\delta_F(a, b) - 2^{n-m})^2}{(2^n - 1)2^m} \right)^{1/2}. \tag{11}$$

Apparently, the smaller the value $SD(f)$ the closer is the distribution to the uniform.

Definition 5: An (n, m) S-box $F = (f_1, f_2, \ldots, f_m)$ is a balanced S-box if and only if for any $c = (c_1, \ldots, c_m) \in \mathbb{F}_2^{m*}$, $f_c = \sum_{i=1}^m c_i f_i$ is a balanced Boolean function.

III. Construction of Highly Nonlinear Balanced S-Boxes via ML-Sequences

In this section, we explore the possibilities of constructing highly nonlinear S-boxes with good differential properties by using certain ideas of the Maiorana-McFarland construction method. The idea of employing subfields $\mathbb{F}_{2^m}$ of $\mathbb{F}_{2^n}$, when $m \mid n$, to construct highly nonlinear S-boxes, $F : \mathbb{F}_n \to \mathbb{F}_m$, was originally considered in [7]. However, this method is conditioned on the existence of suitable subfields, thus without the possibility of applying it to a general case $m < n$. In what follows, we investigate the general case of constructing $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ by defining a suitable collection of m component functions separately, thus without any restrictions related to the factorization of n.

Construction 1: Let $\Phi = (\phi_1, \ldots, \phi_m)$ and $\Psi = (\psi_1, \ldots, \psi_m)$, where for each $1 \le i \le m$,

$$\phi_i : \mathbb{F}_2^{n-k} \to \mathbb{F}_2^k, \quad \psi_i : \mathbb{F}_2^{k-m} \to \mathbb{F}_2^m.$$

Let $G = (g_1, \ldots, g_m) : \mathbb{F}_2^{n-k} \to \mathbb{F}_2^m$ and $H = (h_1, \ldots, h_m) : \mathbb{F}_2^{k-m} \to \mathbb{F}_2^m$ be two vectorial functions. Define the vectorial Boolean function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ through its components as:

$$f_i(X_m, X'_{k-m}, X''_{n-k}) = \begin{cases} X_k \cdot \phi_i(X''_{n-k}) + g_i(X''_{n-k}), & \text{if } X''_{n-k} \neq 0, \\ X_m \cdot \psi_i(X'_{k-m}) + h_i(X'_{k-m}), & \text{if } X''_{n-k} = 0, \end{cases} \tag{12}$$

for $i = 1, \ldots, m$, where $X_m = (x_1, \ldots, x_m)$, $X'_{k-m} = (x_{m+1}, \ldots, x_k)$ and $X''_{n-k} = (x_{k+1}, \ldots, x_n)$.

Lemma 1: For any $c \in \mathbb{F}_2^m$ and $\lambda \in \mathbb{F}_2^n$ the extended Walsh spectra of the function $F_c = \sum_{i=1}^m c_i f_i$, where f_i are defined by means of Construction 1, satisfies,

$$W_F(c, \lambda) \le 2^m |\Psi_c^{-1}(\lambda_m)| + 2^k |\Phi_c^{-1}(\lambda_k)|, \tag{13}$$

where Ψ_c, Φ_c denotes the linear combinations of the component functions and $\lambda_k \in \mathbb{F}_2^k$, $\lambda_m \in \mathbb{F}_2^m$. Furthermore, $\deg(F) \le n - m + 1$.

Proof: We also use G_c and H_c to denote the linear combination of the component functions of G and H. Notice that $\Psi_c = c_1\psi_1 + \cdots + c_m\psi_m$ maps from $\mathbb{F}_2^{k-m}$ to $\mathbb{F}_2^m$ and similarly $\Phi_c : \mathbb{F}_2^{n-k} \to \mathbb{F}_2^k$. For any nonzero $c \in \mathbb{F}_2^m$ and $\lambda \in \mathbb{F}_2^n$ the extended Walsh transform is given by,

$$\begin{aligned} W_{F_c}(\lambda) &= \sum_{X_n\in\mathbb{F}_2^n} (-1)^{c\cdot F(X_n)+\lambda\cdot X_n} \\ &= \sum_{X''_{n-k}\in\mathbb{F}_2^{n-k}} \sum_{X_k\in\mathbb{F}_2^k} (-1)^{F_c(X_n)+\lambda\cdot X_n} \\ &= \sum_{X_k\in\mathbb{F}_2^k} (-1)^{X_m\cdot\Psi_c(X'_{k-m})+H_c(X'_{k-m})+\lambda_k\cdot X_k} \\ &\quad + \sum_{X''_{n-k}\neq 0} \sum_{X_k\in\mathbb{F}_2^k} (-1)^{X_k\cdot\Psi_c(X''_{n-k})+G_c(X''_{n-k})+\lambda_{n-k}\cdot X''_{n-k}} \\ &= 2^m \sum_{X'_{k-m}\in\Psi_c^{-1}(\lambda_m)} (-1)^{H_c(X'_{k-m})+\lambda_{k-m}\cdot X'_{k-m}} \\ &\quad + 2^k \sum_{X''_{n-k}\in\Phi_c^{-1}(\lambda_k)} (-1)^{G_c(X''_{n-k})+\lambda_{n-k}\cdot X''_{n-k}} \\ &\le 2^m |\Psi_c^{-1}(\lambda_m)| + 2^k |\Phi_c^{-1}(\lambda_k)|. \end{aligned}$$

For any fixed X'_{k-m} and X''_{n-k} of length $n-m$ the component functions f_i (and their linear combinations) are linear functions in the remaining variables. Thus, the degree of F is at most $n - m + 1$ which can be achieved by a careful choice of Ψ and Φ. ∎

Lemma 1 implies that in order to achieve the highest possible nonlinearity the functions Φ and Ψ should be injective. In addition, due to the nature of the construction the set of necessary conditions can be stated as follows:

Condition 1:

a) The input parameters satisfy: $n - k \le k$ and $k - m \le m$.
b) For any $c \in \mathbb{F}_2^{m*}$, $\Phi_c(X''_{n-k})$ and $\Psi_c(X'_{k-m})$ are injective.
c) $\Phi_c(X''_{n-k}) \neq \mathbf{0}_k$ for any $X''_{n-k} \neq \mathbf{0}$ and $\Psi_c(X'_{k-m}) \neq \mathbf{0}_m$ for any X'_{k-m}.

If Condition 1 is satisfied it follows from Lemma 1 that $N_F = 2^{n-1} - 2^{m-1} - 2^{k-1}$, since the equality in (13) essentially holds. We notice that there is no restriction on vectorial functions G and H, thus a large class of S-boxes can be constructed once Condition 1 is satisfied.

A. Finding Classes of Functions Satisfying Condition 1

In this section we address the problem of identifying the classes of functions $\{\Phi\}$ and $\{\Psi\}$ satisfying Condition 1 above. We show that such functions can be easily obtained by considering so called maximum length sequences

TABLE I
AN (8,3) S-BOX THROUGH TWO ML-SEQUENCES WITH PERIODS 15 AND 7

u	$f_1(x_1,\ldots,x_4,u)$	$f_2(x_1,\ldots,x_4,u)$	$f_3(x_1,\ldots,x_4,u)$
0000	x_2 $(x_4=0)$ x_1+x_3 $(x_4=1)$	x_1+x_3 $(x_4=0)$ x_2+x_3 $(x_4=1)$	x_2+x_3 $(x_4=0)$ $x_1+x_2+x_3$ $(x_4=1)$
0001	x_3	x_2	x_1+x_4
0010	x_2	x_1+x_4	x_3+x_4
0011	x_1+x_4	x_3+x_4	x_2+x_3
0100	x_3+x_4	x_2+x_3	$x_1+x_2+x_4$
0101	x_2+x_3	$x_1+x_2+x_4$	x_1+x_3
0110	$x_1+x_2+x_4$	x_1+x_3	x_2+x_4
0111	x_1+x_3	x_2+x_4	$x_1+x_3+x_4$
1000	x_2+x_4	$x_1+x_3+x_4$	$x_2+x_3+x_4$
1001	$x_1+x_3+x_4$	$x_2+x_3+x_4$	$x_1+x_2+x_3+x_4$
1010	$x_2+x_3+x_4$	$x_1+x_2+x_3+x_4$	$x_1+x_2+x_3$
1011	$x_1+x_2+x_3+x_4$	$x_1+x_2+x_3$	x_1+x_2
1100	$x_1+x_2+x_3$	x_1+x_2	x_1
1101	x_1+x_2	x_1	x_4
1110	x_1	x_4	x_3
1111	x_4	x_3	x_2

(ML-sequences), whereas the problem of finding other suitable classes is left open.

Binary ML-sequences have many different representations (polynomial, trace, interleaved representation etc.) though referring to the same property of having a maximum length period 2^k-1 if a linear recursion over $\mathbb{F}_2$ involving at most k terms of the initial sequence is used.

Lemma 2: Let α be a primitive element of $\mathbb{F}_{2^k}$. For any nonzero $\theta\in\mathbb{F}_{2^k}$, the sequence $s_0,s_1,s_2,\ldots$ defined by

$$s_t = Tr_1^k(\theta\alpha^t), \tag{14}$$

is an ML-sequence with period 2^k-1. Here, $Tr_1^k:\mathbb{F}_{2^k}\to\mathbb{F}_2$ denotes the absolute trace function defined as $Tr_1^k(x)=x+x^2+\ldots+x^{2^{k-1}}$. Moreover, among the 2^k-1 k-tuples of an ML-sequence below,

$$(s_t,s_{t+1},\ldots,s_{t+k-1}),\quad t=0,1,\ldots,2^k-2, \tag{15}$$

each nonzero binary vector of length k occurs once and only once.

Theorem 1: Let $n>k>m$ with $\lceil n/2\rceil\le k<2m$. Let α be a primitive element of $\mathbb{F}_{2^k}$, β be a primitive element of $\mathbb{F}_{2^m}$ and $X''_{n-k}\in\mathbb{F}_2^{n-k}$, $X'_{k-m}\in\mathbb{F}_2^{k-m}$. For $i=1,\ldots,m$, define $\phi_i:\mathbb{F}_2^{n-k}\to\mathbb{F}_2^k$ and $\psi_i:\mathbb{F}_2^{k-m}\to\mathbb{F}_2^m$ as follows:

$$\phi_i(X''_{n-k}) = (Tr_1^k(\alpha^{i+1+[X''_{n-k}]}),\ldots,Tr_1^k(\alpha^{i+k+[X''_{n-k}]})),$$
$$\psi_i(X'_{k-m}) = (Tr_1^m(\beta^{i+1+[X'_{k-m}]}),\ldots,Tr_1^m(\beta^{i+k+[X'_{k-m}]})),$$

where $[X''_{n-k}]$ and $[X'_{k-m}]$ is the decimal representation of X''_{n-k} and X'_{k-m}, respectively. Then, the functions ϕ_i and ψ_i satisfy Condition 1.

Proof: Due to similarity of the construction method we only show that $\Phi_c:\mathbb{F}_2^{n-k}\to\mathbb{F}_2^k$ is injective, and the same reasoning applies to Ψ_c. First notice that a) in Condition 1 is easily satisfied. For any nonzero vector $c=(c_1,\ldots,c_m)\in\mathbb{F}_2^m$, let $\Phi_c(X''_{n-k})=\sum_{i=1}^m c_i\phi_i(X''_{n-k})$. Then,

$$\Phi_c(X''_{n-k}) = \Big(\sum_{i=1}^m c_i Tr_1^k\big(\alpha^{i+1+[X''_{n-k}]}\big),\ldots,\sum_{i=1}^m c_i Tr_1^k\big(\alpha^{i+k+[X''_{n-k}]}\big)\Big) \tag{16}$$

$$= \Big(Tr_1^k\Big(\sum_{i=1}^m (c_i\alpha^i)\cdot\alpha^{1+[X''_{n-k}]}\Big),\ldots,Tr_1^k\Big(\sum_{i=1}^m (c_i\alpha^i)\cdot\alpha^{k+[X''_{n-k}]}\Big)\Big). \tag{17}$$

By Lemma 2, it is clear that $\Phi_c(X''_{n-k})\neq\mathbf{0}$ for any $X''_{n-k}\neq 0$, and furthermore each nonzero output k-tuple appears exactly once. Thus, Φ_c is injective. ∎

Remark 1: Notice that when ϕ_i and ψ_i are constructed by means of ML sequences, the functions f_i given by (12) in Construction 1 can be expressed as,

$$f_i(X_m,X'_{k-m},X''_{n-k}) = \begin{cases} \sum_{j=1}^k \big(Tr_1^k\big(\alpha^{i+j+[X''_{n-k}]}\big)\cdot x_j\big) + g_i(X''_{n-k}), & \text{if } X''_{n-k}\neq 0,\\ \sum_{j=1}^k \big(Tr_1^k\big(\beta^{i+j+[X'_{k-m}]}\big)\cdot x_j\big) + h_i(X'_{k-m}), & \text{if } X''_{n-k}=0,\end{cases}$$

where α and β are primitive elements of $\mathbb{F}_{2^k}$ and $\mathbb{F}_{2^m}$, respectively.

Corollary 1: Let $F:\mathbb{F}_2^n\to\mathbb{F}_2^m$ be as in Construction 1, with the choice of Φ and Ψ as in Theorem 1. Let n be even and $k=n/2$. Then F is a balanced S-box with nonlinearity $N_F=2^{n-1}-2^{n/2-1}-2^{m-1}$.

Example 1: Let α be a root of the primitive polynomial x^4+x+1. By Lemma 2, setting $\theta=1$ we obtain the ML-sequence $000100110101111\cdots$. Using Construction 1 and Theorem 1, with $g_i=h_i=0$ in Construction 1, we can construct an $(8,3)$ S-box, see Table I. The simulations show that $\delta(a,b)\in\{16,18,28,30,32,34,36,40,46\}$, where $a\in\mathbb{F}_{2^8}^*$ and $b\in\mathbb{F}_{2^3}$. The distribution table is given in Table II.

Open Problem 1: It is of interest to find other classes of functions Φ and Ψ satisfying Condition 1, which is left as an interesting open problem.

B. Constructing $(n,n/2)$ S-Boxes

Note that if the functions ϕ_i and ψ_i are designed by means of Theorem 1 then Construction 1 can not generate an $(n,n/2)$ S-box by using two ML-sequences with different periods. In what follows, we construct an $(n,n/2)$ S-box through an

TABLE II
THE DISTRIBUTION TABLE OF AN (8,3) S-BOX

$\delta_F(a,b)$	16	18	28	30	32	34	36	40	46
Number	7	8	128	608	675	440	128	14	32
Percent	0.34%	0.39%	6.27%	29.80%	33.09%	21.57%	6.27%	0.69%	1.57%

TABLE III
AN (8,4) S-BOX THROUGH AN ML-SEQUENCE WITH PERIOD 15

u	$f_1(x_1,\ldots,x_4,u)$	$f_2(x_1,\ldots,x_4,u)$	$f_3(x_1,\ldots,x_4,u)$	$f_4(x_1,\ldots,x_4,u)$
0000	$x_2{+}x_3{+}x_4{+}x_1x_4$ $+x_2x_4{+}x_3x_4{+}x_2x_3x_4$	$x_1x_2{+}x_3{+}x_4$ $+x_1x_3{+}x_1x_4{+}x_1x_3x_4$	$x_4{+}x_1x_2{+}x_1x_3$ $+x_2x_3{+}x_2x_4{+}x_1x_2x_4$	$x_1{+}x_2{+}x_3{+}x_4{+}x_1x_3$ $+x_2x_3{+}x_1x_2x_3{+}x_2x_3x_4$
0001	x_3	x_2	x_1+x_4	x_3+x_4
0010	x_2	x_1+x_4	x_3+x_4	x_2+x_3
0011	x_1+x_4	x_3+x_4	x_2+x_3	$x_1+x_2+x_4$
0100	x_3+x_4	x_2+x_3	$x_1+x_2+x_4$	x_1+x_3
0101	x_2+x_3	$x_1+x_2+x_4$	x_1+x_3	x_2+x_4
0110	$x_1+x_2+x_4$	x_1+x_3	x_2+x_4	$x_1+x_3+x_4$
0111	x_1+x_3	x_2+x_4	$x_1+x_3+x_4$	$x_2+x_3+x_4$
1000	x_2+x_4	$x_1+x_3+x_4$	$x_2+x_3+x_4$	$x_1+x_2+x_3+x_4$
1001	$x_1+x_3+x_4$	$x_2+x_3+x_4$	$x_1+x_2+x_3+x_4$	$x_1+x_2+x_3$
1010	$x_2+x_3+x_4$	$x_1+x_2+x_3+x_4$	$x_1+x_2+x_3$	x_1+x_2
1011	$x_1+x_2+x_3+x_4$	$x_1+x_2+x_3$	x_1+x_2	x_1
1100	$x_1+x_2+x_3$	x_1+x_2	x_1	x_4
1101	x_1+x_2	x_1	x_4	x_3
1110	x_1	x_4	x_3	x_2
1111	x_4	x_3	x_2	x_1+x_4

ML-sequence and a permuting $(n/2, n/2)$ S-box having a high nonlinearity value.

Construction 2: Let $n = 2r$ and α be a primitive element of $\mathbb{F}_{2^r}$. Let $X_n = (x_1,\ldots,x_n) \in \mathbb{F}_2^n$, $u = (u_{r+1},\ldots,u_n) \in \mathbb{F}_2^r$. Let $P(X_r) = (p_1, p_2, \ldots, p_r)$ be a permutation on $\mathbb{F}_2^r$ with $N_P = 2^{r-1} - 2^{\lfloor r/2 \rfloor}$, where $p_i \in \mathcal{B}_r$. For $i = 1, 2, \ldots, r$, let $f_i \in \mathcal{B}_n$ be defined by

$$f_i(X_n) = x_{r+1}^0 \cdots x_n^0 p_i(X_r) + \sum_{u \in \mathbb{F}_2^{r*}} x_{r+1}^{u_{r+1}} \cdots x_n^{u_n} \sum_{j=1}^{r} \left(Tr_1^r(\alpha^{i+j+[u]}) \cdot x_j \right), \quad (18)$$

where the notation $x_i^{u_i}$ means:

$$x_i^{u_i} = \begin{cases} 1, & \text{if } x_i = u_i \\ 0, & \text{if } x_i \neq u_i \end{cases} \quad (19)$$

Then an S-box, represented as F: $\mathbb{F}_2^n \to \mathbb{F}_2^r$, is defined as $F(X_n) = (f_1, f_2, \ldots, f_r)$.

Theorem 2: Let n be even and F: $\mathbb{F}_2^n \to \mathbb{F}_2^{n/2}$ be as in Construction 2. Then, F is a balanced S-box with nonlinearity

$$N_F = 2^{n-1} - 2^{n/2-1} - 2^{\lfloor n/4 \rfloor}. \quad (20)$$

The (minimum) degree of F is $n/2 + \deg(P)$, and it attains its maximum value $\deg(F) = n - 1$ when P is of degree $n/2 - 1$.

Proof: Let $n = 2r$ and denote $X_n = (X_r, X_r')$ and $\lambda_n = (\lambda_r, \lambda_r')$. In comparison to Construction 1, this method replaces the mappings $\psi_i : \mathbb{F}_2^{k-m} \to \mathbb{F}_2^m$ with a single nonlinear permutation P over $\mathbb{F}_2^m$. Denoting $P_c = \sum_{i=1}^m c_i p_i(X_r)$, the extended Walsh transform can be computed as,

$$\begin{aligned} W_{F_c}(\lambda) &= \sum_{X_r \in \mathbb{F}_2^r} \sum_{X_r' \in \mathbb{F}_2^r} (-1)^{F_c(X_n) + \lambda_r \cdot X_r + \lambda_r' \cdot X_r'} \\ &= \sum_{X_r \in \mathbb{F}_2^r; X_r' = 0} (-1)^{P_c(X_r) + \lambda_r \cdot X_r} \\ &\quad + \sum_{X_r' \in \mathbb{F}_2^{r*}} \sum_{X_r \in \mathbb{F}_2^r} (-1)^{F_c(X_n) + \lambda_r' \cdot X_r'} \\ &= 2^{\lfloor r/2 \rfloor + 1} + 2^r, \end{aligned}$$

where the latter term 2^r follows easily from the definition of f_i. Setting $r = n/2$, we obtain the result.

It is easily verified that the terms of degree larger than $r+1$ in $x_{r+1} \cdots x_n \sum_{i=1}^r c_i p_i(X_r)$ are not cancelled for any $c \in \mathbb{F}_2^{r*}$ (the existence of these terms follows from the assumption that P is a nonlinear permutation), thus $\deg(F_c) = n/2 + \deg(P)$. Then, $\deg(F) = n - 1$ if and only if $\deg(P) = n/2 - 1$. ■

The construction of $F : \mathbb{F}_2^8 \to \mathbb{F}_2^4$ is summarized in Table III. The simulations show that $\delta(a,b) \in \{0, 14, 16, 18, 20, 22\}$, where $a \in \mathbb{F}_{2^8}^*$ and $b \in \mathbb{F}_{2^4}$. The distribution of $\delta(a,b)$ is given in Table IV. The standard deviation corresponding to this differential table is easily computed using (11), which gives $SD(F) = 2.203$.

TABLE IV
THE DISTRIBUTION TABLE OF AN (8,4) S-BOX

$\delta_F(a,b)$	0	14	16	18	20	22
Number	15	1440	1400	1050	15	160
Percent	0.37%	35.29%	34.31%	25.74%	0.37%	3.92%

IV. BALANCED S-BOXES FROM VECTORIAL BENT FUNCTIONS

In this section we propose a method of constructing balanced S-boxes by a suitable modification of vectorial

bent functions. The method employs recently characterized vectorial bent functions that stem from the $\mathcal{PS}_{ap}$ class introduced by Dillon [6]. For the discussion that follows we first introduce some useful notation.

Let $\mathcal{U}$ be the cyclic group of (2^k+1)th roots of unity, and for $n=2k$, denote by $L=\mathbb{F}_{2^n}$ and $K=\mathbb{F}_{2^k}$. Notice that any element $x\in L^*$ can be uniquely represented as $x=u\gamma$, where $u\in\mathcal{U}$ and $\gamma\in K^*$, and furthermore $\cup_{u\in\mathcal{U}} uK^*=L^*$. The relative trace function $Tr_m^n:\mathbb{F}_{2^n}\to\mathbb{F}_{2^m}$, a mapping to a subfield $\mathbb{F}_{2^m}$, when $m\mid n$, is defined as

$$Tr_m^n(x)=x+x^{2^m}+x^{2^{2m}}+\cdots+x^{2^{(n/m-1)m}},\quad \text{for all } x\in\mathbb{F}_{2^n}.$$

In a recent article[1] [1] the following equivalent conditions were proved.

Theorem 3 ([1]): Let $n=2k$, and define $F(x)=Tr_k^n(P(x))$, where $P(x)=\sum_{i=1}^t a_i x^{i(2^k-1)}$ and $t\le 2^k$. Then the following conditions are equivalent:

a) *F is a vectorial bent function of dimension k.*
b) *$\sum_{u\in\mathcal{U}}(-1)^{Tr_1^k(\lambda F(u))}=1$, for all $\lambda\in K^*$.*
c) *There are two values $u\in\mathcal{U}$ such that $F(u)=0$, and furthermore if $F(u_0)=0$, then F is one-to-one and onto from $\mathcal{U}_0=\mathcal{U}\setminus u_0$ to K.*

This class of functions belongs to the partial spread ($\mathcal{PS}$) class of functions introduced by Dillon [6], and more precisely if F is bent then its component functions are in $\mathcal{PS}_{ap}$. The most fundamental property of any function in this class is that it is constant on the multiplicative cosets of the form uK^*. Indeed, for $x=u\gamma$, where $\gamma\in K^*$, we have,

$$\begin{aligned}F(u\gamma)&=Tr_k^n\Big(\sum_{i=1}^t a_i(u\gamma)^{i(2^k-1)}\Big)=Tr_k^n\Big(\sum_{i=1}^t a_i u^{i(2^k-1)}\Big)\\&=F(u).\end{aligned}$$

Theorem 4: The function F in Theorem 3 is never balanced for any choice of $a_i\in L$. On the other hand, the modification of F defined by,

$$\tilde{F}(x)=\begin{cases}F(x) & \text{if } x\in L\setminus K,\\ Q(x) & \text{if } x\in K,\end{cases}\tag{21}$$

where $Q(x)\in K[x]$ is a permutation over K, is balanced if and only if

$$\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))}=0,$$

for all $\lambda\in K^*$. That is, all nonzero linear combinations of the component functions of F are of weight 2^{n-1}.

Proof: The extended Walsh transform of F at σ is given by,

$$\begin{aligned}W_F(\lambda,\sigma)&=\sum_{x\in L}(-1)^{Tr_1^k(\lambda Tr_k^n(\sum_{i=1}^t a_i x^{r_i(2^k-1)}))+Tr_1^n(\sigma x)}\\&=1+\sum_{u\in\mathcal{U}}\sum_{z\in K^*}(-1)^{Tr_1^k(\lambda F(u))+Tr_1^n(\sigma uz)}\\&=1+\sum_{u\in\mathcal{U}}(-1)^{Tr_1^k(\lambda F(u))}\sum_{z\in K^*}(-1)^{Tr_1^n(\sigma uz)}\end{aligned}\tag{22}$$

[1]The result is given in a reduced form since there is one more equivalence originally in [1] which is not needed here.

Especially,

$$W_F(\lambda,0)=1+(2^k-1)\sum_{u\in\mathcal{U}}(-1)^{Tr_1^k(\lambda F(u))},$$

and F cannot be balanced. In particular, this also implies that $\sum_{u\in\mathcal{U}}(-1)^{Tr_1^k(\lambda F(u))}\neq 0$, since for any $\lambda\in K^*$ the Walsh coefficients of the Boolean function $Tr_1^k(\lambda F(x))$ are divisible by 2. Now for $\tilde{F}(x)$ we have,

$$\begin{aligned}W_{\tilde{F}}(\lambda,\sigma)&=\sum_{z\in K;u=1}(-1)^{Tr_1^k(\lambda\tilde{F}(z))+Tr_1^n(\sigma z)}\\&\quad+\sum_{u\in\mathcal{U}\setminus\{1\}}\sum_{z\in K^*}(-1)^{Tr_1^k(\lambda F(u))+Tr_1^n(\sigma uz)}\\&=\sum_{z\in K;u=1}(-1)^{Tr_1^k(\lambda Q(z))+Tr_1^n(\sigma z)}\\&\quad+\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))}\sum_{z\in K^*}(-1)^{Tr_1^n(\sigma uz)}\end{aligned}$$

In particular,

$$\begin{aligned}W_{\tilde{F}}(\lambda,0)&=\sum_{z\in K;u=1}(-1)^{Tr_1^k(\lambda Q(z))}\\&\quad+(2^k-1)\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))}\\&=(2^k-1)\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))},\end{aligned}$$

where we used the fact that $\sum_{z\in K}(-1)^{Tr_1^k(\lambda Q(z))}=0$ if and only if Q is a permutation over K. Thus, the linear combinations $Tr_1^k(\lambda\tilde{F}(x))$ of $\tilde{F}$ are balanced if and only if $\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))}=0$, for all $\lambda\in K^*$. ■

Remark 2: Notice that if F in Theorem 3 is bent, then all the linear combinations of the component functions of F have weight $2^{n-1}-2^{\frac{n}{2}-1}$. The modified function $\tilde{F}$ is obtained by replacing the all-zero function on the subfield K by a (nonlinear) permutation Q (whose any component function is of weight $2^{\frac{n}{2}-1}$).

Theorem 5: The condition that $\tilde{F}$ is balanced given by

$$\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))}=0,\quad \text{for all } \lambda\in K^*,\tag{23}$$

holds if and only if $F:\mathcal{U}\setminus\{1\}\to K$ is a bijection from $\mathcal{U}\setminus\{1\}$ to K.

Proof: The proof follows the same reasoning as in [1], but is given here for self-completeness. $G(x)\in K[x]$ is a permutation if and only if $\sum_{x\in K}(-1)^{Tr_1^k(\lambda G(x))}=0$, for any $\lambda\in K^*$. Since $\#\mathcal{U}\setminus\{1\}=\#K$, there is a bijection $\Psi:K\to\mathcal{U}\setminus\{1\}$. Thus, by setting $u=\Psi(x)$ for $u\in\mathcal{U}\setminus\{1\}$, $x\in K$, and letting $G:K\to K$, where $G=F\circ\Psi$, we have the condition

$$\begin{aligned}\sum_{u\in\mathcal{U}\setminus\{1\}}(-1)^{Tr_1^k(\lambda F(u))}&=\sum_{\Psi(x),x\in K}(-1)^{Tr_1^k(\lambda F(\Psi(x)))}\\&=\sum_{x\in K}(-1)^{Tr_1^k(\lambda G(x))}=0,\end{aligned}$$

where $F(\Psi(x)):K\to K$. This is satisfied if and only if $F(\Psi(x))$ permutes K. Since Ψ is a bijection and $F(\Psi(x))$ is a permutation, then $F:\mathcal{U}\setminus\{1\}\to K$ is also a bijection. ■

Theorem 6: Assume that $Q(x) \in K[x]$ is a permutation over K such that $N_Q = 2^{k-1} - \frac{1}{2}\Gamma$, that is, $\Gamma = \max_{\alpha \in K^*, \beta \in K} |W_Q(\alpha, \beta)|$. Let $F : L \to K$, defined in Theorem 4, be a vectorial bent function such that $F : \mathcal{U} \setminus \{1\} \to K$ is a bijection. Then, the nonlinearity of $\tilde{F} : L \to K$ is given by $N_{\tilde{F}} = 2^{n-1} - 2^{\frac{n}{2}-1} - \frac{1}{2}\Gamma$. Moreover, the degree of $\tilde{F}$ is $k + \deg(Q)$.

Proof: By Theorem 5, $\tilde{F}$ is balanced as $F : \mathcal{U} \setminus \{1\} \to K$ is a bijection, that is, $\sum_{u \in \mathcal{U} \setminus \{1\}} (-1)^{Tr_1^k(\lambda F(u))} = 0$. On the other hand, since F is bent $\sum_{u \in \mathcal{U}} (-1)^{Tr_1^k(\lambda F(u))} = 1$, for all $\lambda \in K^*$. This implies that F vanishes on K as we must have $(-1)^{Tr_1^k(\lambda F(1))} = 1$ for any $\lambda \in K^*$, that is, $F(z) = 0$ for $z \in K$. Notice that $\tilde{F}$ restricted to K equals to Q. Furthermore, $|W_F(\lambda, \sigma)| = 2^k$, for any $\lambda \in K^*$ and $\sigma \in L$. Now, we have

$$\begin{aligned} W_F(\lambda, \sigma) &= \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda F(z)) + Tr_1^n(\sigma z)} \\ &\quad + \sum_{u \in \mathcal{U} \setminus \{1\}} (-1)^{Tr_1^k(\lambda F(u))} \sum_{z \in K^*} (-1)^{Tr_1^n(\sigma u z)} \\ &= \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda F(z)) + Tr_1^n(\sigma z)} \\ &\quad + \sum_{u \in \mathcal{U} \setminus \{1\}} (-1)^{Tr_1^k(\lambda F(u))} \sum_{z \in K^*} (-1)^{Tr_1^n(\sigma u z)} \\ &\quad + \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda Q(z)) + Tr_1^n(\sigma z)} \\ &\quad - \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda Q(z)) + Tr_1^n(\sigma z)} \\ &= \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda F(z)) + Tr_1^n(\sigma z)} \\ &\quad + W_{\tilde{F}}(\lambda, \sigma) - \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda Q(z)) + Tr_1^n(\sigma z)}. \end{aligned}$$

Using the fact that $F(z) = 0$ on K, for $\sigma \neq 0$ the first sum in the last equality is 0, therefore

$$W_{\tilde{F}}(\lambda, \sigma) = W_F(\lambda, \sigma) + \sum_{z \in K; u=1} (-1)^{Tr_1^k(\lambda Q(z)) + Tr_1^n(\sigma z)},$$

so that $|W_{\tilde{F}}(\lambda, \sigma)| \leq 2^k + \Gamma$. Hence, $N_{\tilde{F}} = 2^{n-1} - 2^{\frac{n}{2}-1} - \frac{1}{2}\Gamma$. The result on degree follows easily from the construction. ■

Example 2: Let $n = 6$ and define $F(x) = Tr_3^6(\alpha^3 x^7 + \alpha^3 x^{21})$, where α is a primitive element in $L = \mathbb{F}_{2^6}$. The computation has been performed using the MAGMA software, where α is a root of the primitive polynomial $p(x) = x^6 + x^4 + x^3 + x + 1$. Then, F is of the form $Tr_k^n(\sum_{i=1}^t a_i x^{i(2^k-1)})$ and it can be checked that F is vectorial bent. The elements u of $\mathcal{U} = \{\beta^i : i = 0, 1, \ldots, 2^3\}$, where $\beta = \alpha^{2^k-1}$ and $\gamma = \alpha^{2^k+1}$ generate $\mathcal{U}$ and $K^* = \mathbb{F}_{2^3}^*$, respectively, are mapped by F as follows:

u	1	β^1	β^2	β^3	β^4	β^5	β^6	β^7	β^8
$F(u)$	0	γ	γ^6	0	γ^3	1	γ^5	γ^4	γ^2

Notice that $F(\mathcal{U} \setminus \{1\})$ permutes the subfield $\mathbb{F}_{2^3}$. Let $\tilde{F}$ be defined as in Theorem 4, and $Q(x) \in \mathbb{F}_{2^3}[x]$ be given as x^3. It is well-known that x^3 is a Gold like permutation over $\mathbb{F}_{2^k}$ for odd k, and its nonlinearity is $N_Q = 2^{k-1} - 2^{\frac{k-1}{2}}$. Therefore, in our case $\Gamma = 4$ and $N_{\tilde{F}} \geq 26$, which implies that $N_{\tilde{F}} = 26$ achieving the maximum possible nonlinearity of balanced Boolean functions. The differential properties of $\tilde{F}$ are summarized below, for instance the first column means there are 7 pairs (a, b) for which there are no solutions to $\tilde{F}(x) + \tilde{F}(x + a) = b$.

$\delta_{\tilde{F}}(a, b)$	0	6	8	10	12	14	16
Number	7	166	183	112	28	6	2

Example 3: For $n = 8$, define $F(x) = Tr_4^8(x^{15} + \alpha^{92} x^{45} + \alpha^{23} x^{60})$, which is a vectorial bent function. The computation has been performed using the MAGMA software, where α is a root of the primitive polynomial $p(x) = x^8 + x^4 + x^3 + x^2 + 1$. We select $Q(x) = x^{-1}$ and define $\tilde{F}(x) = Q(x)$ for $x \in \mathbb{F}_{2^4}$. Then, the maximal absolute value in the Walsh spectra of Q is $\Gamma = 2^{n/4+1} = 8$, and consequently $N_{\tilde{F}} \geq 116$. The actual nonlinearity is 116 which is the maximum known nonlinearity of balanced Boolean functions, and $\deg(\tilde{F}) = 7$. The differential properties of $\tilde{F}$ are summarized below. The standard deviation from the uniform case turns out to be $SD(\tilde{F}) = 2.203$.

Remark 3: Another approach, based on the use of a set of disjoint codes, has been proposed recently in [15, Corollary 3]. This method also first constructs a perfect nonlinear S-box, a vectorial bent function $F : \mathbb{F}_2^n \to \mathbb{F}_2^k$, and modify such an S-box by a suitable replacement of its output values. However, $F : \mathbb{F}_2^n \to \mathbb{F}_2^k$, $F = (f_1, \ldots, f_k)$, is designed via its component functions so that $supp(f_i) = \bigcup_{[j] \in supp(h_i)} C_j^*$, for $i = 1, \ldots, k$. Here, $\{C_0, C_1, \ldots, C_{2^{n/2}-1}, C_{2^k}\}$ is a set of disjoint linear codes ($C_j^* = C_j \setminus \{0\}$) and $H = (h_1, \ldots, h_k)$ is a permutation of $\mathbb{F}_2^k$. Using the trace representation it can be shown [2] that the cardinality of different vectorial bent functions of the form $Tr_k^n(\sum_{i=1}^t a_i x^{i(2^k-1)})$, where $a_i \in L$, equals to $(2^k + 1)!2^{k-1}$, whereas there are only $2^k!$ permutations H. It is not clear whether the trace class includes the class in [15] or these classes are possibly equal to each other.

A. Differential Properties of $\tilde{F}$

The differential properties of perfect nonlinear S-boxes $\mathbb{F}_q^r \to \mathbb{F}_q$, characterized by the property that for any fixed nonzero $a \in \mathbb{F}_q^n$ the difference $F(x+a) - F(x) = b$ has exactly q^{n-1} solutions in $\mathbb{F}_q^n$, were first analyzed by Nyberg in [12]. In the case of quadratic extensions, when $r = 2$ and $q = 2^k$, we can identify $\mathbb{F}_q$ with K and $\mathbb{F}_q \times \mathbb{F}_q \cong L$ in our notation. The difference $F(x + a) + F(x) = b$ has exactly 2^k solutions for each $a \in L^*$ and $b \in K$, thus $\delta_F(a, b) = 2^k$. Notice also, that the solutions come in pairs since if x_0 is a solution to $F(x+a) + F(x) = b$ so is $x_0 + a$. We now collect some useful observations related to the solution of $F(x + a) + F(x) = b$, for a vectorial bent function F in Theorem 3.

Lemma 3: Let $n = 2k$ and $F(x) = Tr_k^n(\sum_{i=1}^t a_i x^{i(2^k-1)})$ be vectorial bent. Furthermore, let x_0 denote an arbitrary solution to $F(x + a) + F(x) = b$ for a fixed $a \in L^*$ and $b \in K$. Then, (i)

a) *If $b = 0$ and $a \in K^*$, then any $x_0 \in K$ is a solution. If $b = 0$ and $a \in uK^*$, $u \neq 1$, then the set of solutions is given by $x_0 \in uK^* \setminus \{a\}$, where the remaining two solutions come from a different coset.*

b) *If $b \neq 0$ and $a \in uK^*$, then $x_0 \notin uK^*$ and the solutions comes from distinct 2^k cosets $u'K$, with $u' \neq u$.*

Proof: We frequently use the fact that $F(1) = 0$ and that $F(u\gamma) = F(u)$, for any $\gamma \in K^*$. Also, if $x_0 \in uK^*$ and $x_1 \in uK^*$, then $x_0 + x_1 \in uK^*$. This is because for $x_0 = u\gamma$, $x_1 = u\alpha$ we have $x_0 + x_1 = u(\gamma + \alpha) \in uK^*$. (i)

a) For $b = 0$, $x \in K$ and $a \in K^*$, we have $x + a \in K$ and therefore $F(x) + F(x + a) = F(1) + F(1) = 0$.
If $b = 0$ and $a \in uK^*$, $u \neq 1$, then for any $x_0 \in uK^*$, $x_0 \neq a$, we have $x_0 + a \in uK^*$. Thus, since $F(x_0) + F(x_0 + a) = 0$, $2^k - 2$ solutions $x_0 \in uK^* \setminus \{a\}$ come from uK^* and the remaining two from some other cosets.

b) For a fixed $b \neq 0$ and $a \in uK^*$ any solution $x_0 \notin uK^*$, as otherwise $x_0, x_0 + a \in uK^*$ and $F(x_0) + F(x_0 + a) = 0$. For this reason, x_0 and $x_0 + a$ do not belong to the same coset. Assume that two distinct solutions $x_0, x_1 \in u'K^*$ belong to the same coset, where $u' \neq u$ and $x_1 \neq x_0 + a$. Then,

$$F(x_0) + F(x_0 + a) = b$$
$$F(x_1) + F(x_1 + a) = b,$$

implies that $F(x_0 + a) = F(x_1 + a)$, since $F(x_0) = F(x_1) = F(u')$. Thus, $x_0 + a, x_1 + a \in u''K^*$, with $u'' \neq u'$. This is however not possible, since then (using that $x_0, x_1 u'K^*$) $u'\delta + u\gamma \in u''K^*$ and $u'\sigma + u\gamma \in u''K^*$ would imply that $u'(\delta + \sigma) \in u''K^*$, thus $u' = u''$ which contradicts the fact that x_0 and $x_0 + a$ do not belong to the same coset. ∎

Theorem 7: Let $\tilde{F}$ be defined as in Theorem 4. Then the differential values of $\tilde{F}$ satisfy the following,

$$\delta_{\tilde{F}}(a, 0) \in [2^k - 2, 2^{k+1}], \text{ for any } a \notin K^*,$$
$$\delta_{\tilde{F}}(a, 0) = 0, \text{ for any } a \in K^*,$$
$$\delta_{\tilde{F}}(a, b) \in [2^k - 2, 2^{k+1}], \quad \text{for any } a \notin K^*, b \neq 0,$$
$$\delta_{\tilde{F}}(a, b) = 2^k, \quad \text{for any } a \in K^*, b \neq 0.$$

Proof: We are concerned with the impact of changing the values of F on the subfield K, since $\tilde{F}(x) = Q(x)$ for $x \in K$ and $\tilde{F}(x) = F(x)$ elsewhere. We consider several cases depending on whether $a \in L^*$ and $b \in K$.

If $b = 0$ and $a \in K^*$, by Lemma 3, any $x_0 \in K$ is a solution to $F(x) + F(x + a) = 0$. Now, if $x \in K$ then also $x + a \in K$ and $\tilde{F}(x) + \tilde{F}(x + a) = 0$ equals to $Q(x) + Q(x + a) = 0$, which has no solution in K since Q is a permutation. If $x \in L \setminus K$, then $x_0 \in uK^*$, $u \neq 1$, and $x_0 + a$ must lie in the same coset uK^* which cannot be true. Thus, $\delta(a, 0) = 0$ for any $a \in K^*$.

If $b = 0$ and $a \in uK^*$, $u \neq 1$, by Lemma 3, the $2^k - 2$ solutions of $F(x) + F(x + a) = 0$ come from $uK^* \setminus \{a\}$. These are also the solutions of $\tilde{F}(x) + \tilde{F}(x + a) = 0$ since both $x_0, x_0 + a \notin K$. The solutions that might come from K would imply the analysis of $Q(x) + F(x + a) = 0$, which requires the specification of Q and F. However, when x goes through K then $x + a$ takes value from different cosets uK^*. In the worst case $Q(x) = F(x + a)$ and thus $\delta(a, 0) \in [2^k - 2, 2^{k+1}]$, for $a \notin K^*$.

If $b \neq 0$ and $a \in uK^*$, then by Lemma 3, the 2^k solutions comes from distinct 2^k cosets $u'K^*$, $u' \neq u$. Thus, for $a \in K^*$ the 2^k solutions to $F(x) + F(x + a) = b$ are also the solutions to $\tilde{F}(x) + \tilde{F}(x + a) = b$, whereas for $a \notin K^*$ we have $\delta(a, b) \in [2^k - 2, 2^{k+1}]$. ∎

TABLE V
THE DISTRIBUTION TABLE OF $\tilde{F}$ IN EXAMPLE 3

$\delta_{\tilde{F}}(a, b)$	0	14	16	18	20	22	24	26
Number	15	1417	1525	791	268	50	10	4

The differential properties of the component functions (or their linear combinations) are also easily derived from the properties of bent functions. Indeed, since all nonzero derivatives $f(x + a) + f(x)$ of a bent function f are balanced, then by extending the support of a vectorial bent function F (whose component functions and their linear combinations are again bent) the derivatives are still close to balancedness. More precisely, the derivatives of the linear combinations of the component functions, say $\tilde{F}_c$, of $\tilde{F}$ cannot be constant, and furthermore $2^{n-1} - 2^{n/2} \leq wt(\tilde{F}_c(x + a) + \tilde{F}_c(x)) \leq 2^{n-1} + 2^{n/2}$.

V. COMPARING THE DIFFERENTIAL PROPERTIES TO THE INVERSE S-BOX

Notice that perfectly nonlinear mappings (also known as planar mappings), which then cannot be bijective, only exist over the finite fields whose characteristic is different from two. These mapping are characterized by the property that $F(x + a) - F(x)$ is a permutation for any nonzero a, thus any differential of the form $F(x + a) - F(x) = b$ has one and only one solution. A mapping $F : GF(2)^n \to GF(2^m)$ has a uniform differential distribution if any differential $F(x + a) - F(x) = b$ has exactly 2^{n-m} solutions, which is only achieved by vectorial bent functions for suitable n and m. Nevertheless, it appears that there is no satisfactory criterion for comparing the differential properties of symmetric and non-symmetric S-boxes with respect to the differential cryptanalysis.

Even though the main purpose of this section is to provide a comparison of our $(n, n/2)$ S-boxes to the punctured inverse $(n, n/2)$ S-boxes (derived from the inverse mapping $G(x) = x^{-1}$ by deleting $n/2$ outputs) in terms of differential properties, our goal is also to initiate a development of a more general theoretical framework regarding this topic. That is, neglecting the issues related to hardware implementation, which might favour SPN based algorithms over Feistel networks, a rigorous mathematical analysis concerning the resistance to differential cryptanalysis for the two schemes seems to be needed. Note that both methods deviate from the optimal case: the deviation for SPN schemes is then measured as the divergence from APN functions, whereas for the Feistel networks that use (n, m) S-boxes the deviation amounts to measuring the distance to the optimal case of having 2^{n-m} solutions for any differential.

In what follows, we demonstrate that the differential properties of our S-boxes (for both our constructions) are better than those of the punctured inverse S-box, taking the standard value $n = 8$. The comparison is made in terms of the standard deviation of the corresponding differential tables. The punctured

TABLE VI
THE DIFFERENTIAL PROPERTY OF $G_1 = (g_1, g_2, g_3, g_4)$

$\delta_{G_1}(a,b)$	2	4	6	8	10	12	14	16	18	20	22	24	26	28
Number	1	4	30	117	263	488	749	806	699	495	283	103	39	3

TABLE VII
THE DIFFERENTIAL PROPERTY OF $G_2 = (g_3, g_4, g_5, g_6)$

$\delta_{G_2}(a,b)$	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
Number	1	6	28	105	271	524	715	775	737	507	264	108	31	7	1

TABLE VIII
THE DIFFERENTIAL PROPERTY OF $G_3 = (g_3, g_5, g_7, g_8)$

$\delta_{G_4}(a,b)$	2	4	6	8	10	12	14	16	18	20	22	24	26	28
Number	1	5	38	89	275	508	747	797	675	537	257	108	39	4

inverse S-box is obtained by simply deleting $n/2$ outputs if $G(x) = x^{-1}$ is represented as $G(x) = (g_1(x), \ldots, g_8(x))$, where $g_i(x)$ are its Boolean component functions. There are $\binom{8}{4} = 70$ possibilities of keeping four outputs for the inverse S-box $G : \mathbb{F}_{2^8} \to \mathbb{F}_{2^8}$, but due to space constraints we only give a comparison to three "typical" choices given in Table VI, Table VII and Table VIII.

Notice that the truth tables of the component functions for $G(x) = x^{-1}$ are obtained using the same primitive polynomial in MAGMA as in Example 3. Intuitively, due to a larger dispersion around the optimal value (which is 16 in this case), we expect that the standard deviation of the punctured inverse S-boxes is much larger than the one computed for the differential tables in Table IV or in Table V. Indeed, the standard deviation of the three punctured inverse S-boxes above attains the same value $SD(G_i) = 3.94$, for $i = 1, 2, 3$, which is much larger than the standard deviation of our S-boxes. Moreover, by testing the standard deviation of all 70 S-boxes, it turns out that the standard derivation remains the same regardless of the choice of punctured component functions.

Open Problem 2: Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a permutation and $\{F^{(m)}\} : \mathbb{F}_2^n \to F^m$ be a class of $\binom{n}{m}$ vectorial functions obtained by deleting $n - m$ component functions of F. What kind of permutations have the property that the standard deviation of the differential distribution of any function in $\{F^{(m)}\}$ attains the same value ? Notice that it is easy to find counterexamples of permutations not having this property.

In the absence of a fair comparison criterion related to the resistance to differential cryptanalysis of symmetric and asymmetric S-boxes, we once again mention the importance of establishing such criteria along with the development of a theoretical framework to compare SPN and Feistel networks adequately.

VI. CONCLUSIONS

This article provides two different methods of constructing non-symmetric S-boxes with exceptionally good resistance to linear and differential cryptanalysis. We strongly believe that there is a room for further improvements of both techniques towards achieving even better cryptographic properties. The increased complexity of implementation, requiring the doubling of the number of S-boxes if an (n, n) S-box is replaced by an $(n, n/2)$ S-box is motivated by a better resistance to linear cryptanalysis and possibly to differential cryptanalysis as well.

REFERENCES

[1] A. Muratović-Ribić, E. Pasalic, and S. Bajrić, "Vectorial bent functions from multiple terms trace functions," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1337–1347, Feb. 2014.

[2] A. Muratović-Ribić, E. Pasalic, and S. Ribić, "Vectorial hyperbent trace functions from the $\mathcal{PS}_{ap}$ class—Their exact number and specification," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4408–4413, Jul. 2014.

[3] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *J. Cryptol.*, vol. 4, no. 1, pp. 3–72, 1991.

[4] T. P. Berger, A. Canteaut, P. Charpin, and Y. Laigle-Chapuy, "On almost perfect nonlinear functions over $mmbF_2^n$," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4160–4170, Sep. 2006.

[5] F. Chabaud and S. Vaudenay, "Links between differential and linear cryptanalysis," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 950. Berlin, Germany: Springer-Verlag, 1995, pp. 356–365.

[6] J. Dillon, "Elementary Hadamard difference sets," Ph.D. dissertation, Dept. Comput. Sci., Univ. Maryland, College Park, MD, USA, 1974.

[7] K. Khoo, C.-W. Lim, and G. Gong, "Highly nonlinear balanced S-boxes with improved bound on unrestricted and generalized nonlinearity," *Appl. Algebra Eng., Commun. Comput.*, vol. 19, no. 4, pp. 323–338, 2008.

[8] G. Leander, C. Paar, A. Poschmann, and K. Schramm, "New lightweight DES variants," in *Fast Software Encryption* (Lecture Notes in Computer Science), vol. 4593. Berlin, Germany: Springer-Verlag, 2007, pp. 196–210.

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

[10] M. Matsui, "Linear cryptanalysis method for DES cipher," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 765. Berlin, Germany: Springer-Verlag, 1994, pp. 386–397.

[11] W. Meier and O. Staffelbach, "Nonlinearity criteria for cryptographic functions," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 434. Berlin, Germany: Springer-Verlag, 1990, pp. 549–562.

[12] K. Nyberg, "Perfect nonlinear S-boxes," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 547. Berlin, Germany: Springer-Verlag, 1991, pp. 378–386.

[13] G. Piret, T. Roche, and C. Carlet, "PICARO—A block cipher allowing efficient higher-order side-channel resistance," in *Applied Cryptography and Network Security* (Lecture Notes in Computer Science), vol. 7341. Berlin, Germany: Springer-Verlag, pp. 311–328, 2012.

[14] C. E. Shannon, "Communication theory of secrecy systems, *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, Oct. 1948.
[15] W.-G. Zhang and E. Pasalic, "Constructions of resilient S-boxes with strictly almost optimal nonlinearity through disjoint linear codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1638–1651, Mar. 2014.

Wei-Guo Zhang (M'10) received the Ph.D. degree in cryptography from Xidian University, Xi'an, China, in 2007. Since July 2007, he joined the State Key Laboratory of Integrated Services Networks at Xidian University, where he is currently a Professor. His research interests include symmetric cryptography, sequence design, and algebraic coding theory.

Enes Pasalic received the Ph.D. degree in cryptology from Lund University, Lund, Sweden, in 2003. His main research interest is in cryptology and in particular the design and analysis of symmetric encryption schemes. Since May 2003, he has been doing a postdoctoral research at INRIA (Versaille, France) crypto group, and later in 2005 at the Technical University of Denmark, Lyngby. He is currently with University of Primorska, FAMNIT and IAM, Koper, Slovenia.