ENHANCED CRITERIA ON DIFFERENTIAL UNIFORMITY AND NONLINEARITY OF CRYPTOGRAPHICALLY SIGNIFICANT FUNCTIONS

YIN TAN, GUANG GONG, AND BO ZHU

Abstract. The functions defined on finite fields with high nonlinearity are important primitives in cryptography. They are used as the substitution boxes in many block ciphers. To avoid the differential and linear attacks on the ciphers, the Sboxes must have low differential uniformity and high nonlinearity. In this paper, we generalize the notions of the differential uniformity and nonlinearity, which are called the *t-th differential uniformity* and the *diversity of the nonlinearity*, to measure the nonlinear property of the functions. We show that the Sboxes endorsed by `ZUC`, `SNOW 3G` and some lightweight block ciphers have poor performances under these new criteria. The properties and characterizations of these new notions are presented. Another contribution of this paper is to study the nonlinearity of the functions with the form $F = f \circ a$, where f is from $\mathbb{F}_2^k$ to $\mathbb{F}_2^n$ and a is a linear surjection from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$. The motivation of this study is that such a substitution-permutation composition structure is widely used in the design of modern ciphers, which is to bring the confusion and diffusion to the ciphers. We determine the nonlinearity of F for the linear function a with certain property. Using this result, we compute the diversity of the nonlinearity for F and f. It is found that the former value is greatly amplified, which weakens the ciphers against the linear attack.

1. Introduction

The functions defined on finite fields with high nonlinearity play an important role in cryptography. They are used as the Substitution boxes (Sbox) in many block ciphers. Actually in many cases these functions are the only nonlinear part of the ciphers. To avoid various attacks on the ciphers, Sboxes are required to satisfy certain properties, for instance having high algebraic degree [2, 12, 13, 26, 28], low differential uniformity [3], high nonlinearity [33], etc. Furthermore, Sboxes usually are permutations defined on the fields with even degree, i.e. $\mathbb{F}_{2^{2k}}$.

The extensive study on the cryptographically significant functions during the past decades shows that it is difficult to design a function attaining all cryptographic criteria, see for instance [9, 10] and the references therein. A widely adopted method to design a cryptographically strong cipher is to iterate the round function (usually defined as the composition of the carefully chosen Sboxes and linear functions) by several rounds, for instance the ciphers with the substitution-permutation network (SPN) structure. Interestingly, it was shown in some recent work [5–7, 28, 30] that the weakness of the round function may lead to serious security flaws of the ciphers for any number of rounds. For example, Leander et al. [30] demonstrated that, for the block cipher `PRINT`cipher [27], there exist many weak keys such that an attacker may obtain strongly biased linear approximations for any number of rounds. This weakness of the `PRINT`cipher is caused by the existence of the invariant subspace of the round function (defined in [30, page 209]). Another example of the influence of the Sboxes on the security of ciphers can be found in [5, 7, 28], where the authors showed that, if the Sbox is a quadratic function, the algebraic degree of the ciphers increases slower than expected during the iteration of the round

Date: April 15, 2015.

Key words and phrases. Substitution box, almost perfect nonlinear function, perfect nonlinear function, truncated differential attack.

function. These authors made use of this property to develop the zero-sum attack on the hash functions `Keccak` and `Luffa`.

It is the aim of this paper to develop more cryptographic criteria for functions defined on finite fields and study their properties.

As mentioned above, low differential uniformity and high nonlinearity are two important criteria for functions that are endorsed as Sboxes. In Section 3, we generalize the notion of differential uniformity, which are called the *t-th differential uniformity* (Definition 1). We show that some of the well-known Sboxes have very poor t-th differential uniformity, although the commonly defined differential uniformity is optimal. For example, for the Sbox used in `Present` [4], its 1-st differential uniformity is the same as a linear function. This observation has been used in [47] to give an attack on the round-reduced `Present`. We systematically study the property of the t-th differential uniformity for any t. The lower and upper bounds on the t-th differential uniformity for a function from $\mathbb{F}_2^n$ to itself (Theorem 1) is given. It is shown that the functions with the ideal t-th differential uniformity (attaining the lower bound) are actually the perfect nonlinear functions. A characterization of the t-th differential uniformity using the fourth-power Walsh transform is also provided (Theorem 2). Furthermore, for the inverse function from $\mathbb{F}_2^n$ to itself, we theoretically determine its 1-st and $(n-1)$-th differential uniformity. We should mention that the t-th differential uniformity was also used by Nyberg in [36] (called the *chopping of functions*) for the study of almost perfect nonlinear functions and bent functions.

Another contribution of this paper is to study the nonlinearity of the functions with the form $F = f \circ a$, where f is from $\mathbb{F}_2^k$ to $\mathbb{F}_2^n$ and a is a linear surjection from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$. The motivation of this study is that the substitution-permutation composition structure is widely used in many modern ciphers as it may bring the confusion and diffusion to the ciphers. The stream ciphers `SNOW 3G` [38] and `ZUC` [37] are two recent examples with this structure. In Section 4, we first propose the notion *diversity of the nonlinearity* (Definition 3) of a function F, which is adapted from the communication theory, to measure the nonlinearity of a function. It can be seen from the definition that the diversity of the nonlinearity for a linear function is $\frac{1}{\sqrt{2^k}}$. Hence, we will prefer the functions with small diversity of nonlinearity for the cryptographic applications. Usually the diversity of the nonlinearity for a function F is difficult to be determined. However, for the linear functions a with the property that the image array of a and $\mathrm{Tr_n^k(x)}$ is simple (Definition 5), we may determine the nonlinearity of F (Theorem 3) and then make use of it to compute the diversity of the nonlinearity of F. As a result, one may see that $\mathrm{Div}_F > \mathrm{Div}_f$, which shows the nonlinear property of F is poorer than f in some sense.

The rest of the paper is organized as follows. In Section 2, we give the necessary definitions and results which are used throughout the paper. The definition and properties of the t-th differential uniformity is given in Section 3. We study the diversity of the nonlinearity in Section 4. The nonlinearity of the functions with the form $F = f \circ a$ is also presented therein. Some concluding remarks are given in Section 5.

2. Preliminaries

In this section, we introduce the necessary definitions and results which will be used throughout the paper.

2.1. Vectorial and Boolean functions.

Given two positive integers n and m, a function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ is called an (n, m)-*function*. Particularly, when $m = 1$, F is called a *Boolean function* and we usually use the notation $\mathcal{BF}_n$ to denote the set of all Boolean functions with n variables. Clearly, a Boolean function may be regarded as a vector with elements on $\mathbb{F}_2$ of length 2^n by identifying $\mathbb{F}_{2^n}$ with a vector space $\mathbb{F}_2^n$ of dimension n over $\mathbb{F}_2$. In the following, we will switch between these two points of view without explanation if the context is clear. When $m \geq 2$, an (n, m)-function F is called *vectorial Boolean* and the

notation $\mathcal{VF}^{n,m}$ denotes the set of all (n,m)-functions. In particular, when $n=m$, we simply denote it by $\mathcal{VF}^n$. For a function $F \in \mathcal{VF}^n$ and a nonzero element $a \in \mathbb{F}_{2^n}$, the function $F_a(x) \triangleq \mathrm{Tr}\left(\mathrm{aF(x)}\right)$ is called a *component function* of F, where $\mathrm{Tr(x)} = \sum_{\mathrm{i}=0}^{\mathrm{n}-1} \mathrm{x}^{2^{\mathrm{i}}}$ denotes the absolute trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$. Finally, we should mention the well known fact that any (n,n)-function may be represented uniquely as a polynomial in $\mathbb{F}_{2^n}[x]$ mod $(x^{2^n}+x)$, and hence, in the following we express (n,n)-functions by their polynomial forms. An (n,n)-function F is called a *permutation polynomial* (PP) if it is a bijective mapping on $\mathbb{F}_{2^n}$. It is known that if F is a PP on $\mathbb{F}_{2^n}$, then $\deg F \leq n-1$ and we call F has the *optimal* algebraic degree if the aforementioned equality is attained. If $\deg F \leq 1$, then $F(x)$ is called an *affine function*. For an (n,n)-function F, we call the set $\mathcal{G}_F = \{(x, F(x)) : x \in \mathbb{F}_{2^n}\}$ the *graph* of F.

2.2. **Differential and Walsh spectrum.** Denote by $\mathbb{F}_{2^n}^*$ the set of all nonzero elements of $\mathbb{F}_{2^n}$. For an (n,n)-function F and any $(a,b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^n}$, define

$$\delta_F(a,b) = \sharp\{x : x \in \mathbb{F}_{2^n} | F(x+a) + F(x) = b\}.$$

Note that we denote the cardinality of S by $\sharp S$. The multiset $\{* \ \delta_F(a,b) : (a,b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^n} \ *\}$ is called the *differential spectrum* of F. The value

$$\Delta_F \triangleq \max_{(a,b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^n}} \delta_F(a,b)$$

is called the *differential uniformity* of F, or we call F a *differentially Δ_F-uniform* function. In particular, we call F *almost perfect nonlinear* (APN) if $\Delta_F = 2$. It is easy to see that APN functions achieve the lowest possible differential uniformity for functions defined on fields with even characteristic.

Another commonly used method to characterize the nonlinearity of F is as follows. For a function F from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$, the *Walsh (Fourier) transform* $\mathcal{W}_F : \mathbb{F}_{2^n} \times \mathbb{F}_{2^m} \to \mathbb{C}$ of F is defined by:

$$\mathcal{W}_F(a,b) := \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_1^{\mathrm{n}}(\mathrm{ax}) + \mathrm{Tr}_1^{\mathrm{m}}(\mathrm{bF(x)})}.$$

The multiset $W_F := \{* \ \mathcal{W}_F(a,b) : a \in \mathbb{F}_{2^n}, b \in \mathbb{F}_{2^m}^* \ *\}$ is called the *Walsh spectrum* of F. Some researchers call the multiset $\{* \ |x| : x \in W_F \ *\}$ the *extended Walsh spectrum* of F. The *nonlinearity* $\mathrm{NL}(F)$ of F is defined as

$$\mathrm{NL}(F) \triangleq 2^{n-1} - \frac{1}{2} \max_{x \in W_F} |x|.$$

When $n=m$, it is known that, if n is odd, the nonlinearity $\mathrm{NL}(F)$ is upper-bounded by $2^{n-1} - 2^{\frac{n-1}{2}}$; and if n is even, it is conjectured that $\mathrm{NL}(F)$ is upper-bounded by $2^{n-1} - 2^{\frac{n}{2}}$. When n is even, we call a function *known maximal nonlinear* if its nonlinearity attains the aforementioned bound.

Two (n,n)-functions F and G are called *extended affine* (EA) equivalence if there exist affine permutations $L, L' : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ and an affine function A such that $G = L' \circ F \circ L + A$. Two functions F and G are called *Carlet-Charpin-Zinoviev* (CCZ) equivalence if their graphs $\mathcal{G}_F = \{(x, F(x)) : x \in \mathbb{F}_{2^n}\}$ and $\mathcal{G}_G = \{(x, G(x)) : x \in \mathbb{F}_{2^n}\}$ are affine equivalent, that is, if there exists an affine automorphism L of $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ such that $L(\mathcal{G}_F) = \mathcal{G}_G$. It is well known that EA equivalence implies CCZ equivalence, but not vise versa. Moreover, both EA and CCZ equivalence preserve the differential spectrum and the extended Walsh spectrum, and EA equivalence also preserves the algebraic degree when $n \geq 2$.

2.3. **Group rings and characters.** Group rings and character theory of finite fields are useful tools to study functions defined on $\mathbb{F}_{p^n}$. We briefly review some definitions and results. For more details on the character theory and group rings, one may refer to [31] and [41] respectively.

In the following, we assume G is a finite Abelian group. The group algebra $\mathbb{C}[G]$ consists of all formal sums $\sum_{g\in G} a_g g, a_g \in \mathbb{C}$. We define the component-wise addition

$$\sum_{g\in G} a_g g + \sum_{g\in G} b_g g = \sum_{g\in G} (a_g + b_g) g,$$

and multiplication by

$$\sum_{g\in G} a_g g \cdot \sum_{g\in G} b_g g = \sum_{g\in G} \left(\sum_{h\in G} a_h b_{gh^{-1}} \right) g.$$

A subset S of G is identified with the group ring element $\sum_{s\in S} s$ in $\mathbb{C}[G]$, which is also denoted by S (by abuse of notation). For $A = \sum_{g\in G} a_g g$ in $\mathbb{C}[G]$ and t an integer, we define $A^{(t)} := \sum_{g\in G} a_g g^t$.

A character χ of a finite Abelian group G is a homomorphism from G to $\mathbb{C}^*$. A character χ is called *principal* if $\chi(c) = 1$ for all $c \in G$, otherwise it is called *non-principal*. All characters form a group which is denoted by $\widehat{G}$. This *character group* $\widehat{G}$ is isomorphic to G. We denote its unity by χ_0 and the unity of G by 1_G. The action of any character χ is extended to $\mathbb{C}[G]$ by $\chi(\sum_{g\in G} a_g g) = \sum_{g\in G} a_g \chi(g)$. For a group ring element $D = \sum_{g\in G} a_g g \in \mathbb{C}[G]$, the following equation

$$a_g = \frac{1}{|G|} \sum_{\chi\in\widehat{G}} \chi(D)\chi(g^{-1}). \tag{1}$$

is called the *Inversion formula*. It is easy to see that $DD^{(-1)} = \sum_{g,h\in G} a_g a_h g h^{-1}$. By letting $t = gh^{-1}$, we may rewrite it as

$$DD^{(-1)} = \sum_{t\in G} \left(\sum_{h\in G} a_{ht} a_h \right) t.$$

Now, applying the Inversion formula on $DD^{(-1)}$ and we compute the coefficient of t, we get

$$\sum_{h\in G} a_{ht} a_h = \frac{1}{|G|} \sum_{\chi\in\widehat{G}} \chi\left(DD^{(-1)}\right) \chi(t^{-1}) = \frac{1}{|G|} \sum_{\chi\in\widehat{G}} |\chi(D)|^2 \chi(t^{-1}),$$

where $|\cdot|$ denotes the magnitude of a complex number. As a special case, if $t = 1$, we have the *Parseval's equation*:

$$\sum_{h\in G} |a_h|^2 = \frac{1}{|G|} \sum_{\chi\in\widehat{G}} |\chi(D)|^2 . \tag{2}$$

3. The t-th Differential Uniformity

In this section, we introduce the notion of the t-th differential uniformity of a function F from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ and present its basic properties. The relationships among the t-th differential uniformity, (almost) perfect nonlinear functions and the truncated differential attack are also discussed. We start with the definition.

Definition 1. *Let F be a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ and t be an integer with $1 \le t \le m$. Clearly F can be written as the form $F = (F_1, \ldots, F_m)$, where F_i is a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. For any subset $T = \{i_1, i_2, \ldots, i_t\} \subseteq \{1, 2, \cdots, m\}$ with size t, define the function $F_T : \mathbb{F}_2^n \to \mathbb{F}_2^t$ by*

$$F_T(x) = (F_{i_1}(x), \cdots, F_{i_t}(x)) . \tag{3}$$

For any $(a,b) \in (\mathbb{F}_2^n)^* \times \mathbb{F}_2^t$, *define* $\delta_F^T(a,b)$ *by*

$$\delta_F^T(a,b) = \sharp\{x : x \in \mathbb{F}_{2^n} \mid F_T(x+a) + F_T(x) = b\}.$$

The value

$$\Delta_F^t = \max_{\substack{T \in \Omega_t \\ a \in \mathbb{F}_{2^n}^*, b \in \mathbb{F}_{2^t}}} \delta_F^T(a,b)$$

is called the t*-th differential uniformity of* F*, where* Ω_t *is the set of all subsets of* $\{1, \cdots, m\}$ *with size* t*. Furthermore, the value* $\mathcal{D}_F^t = \frac{\Delta_F^t}{2^n}$ *is called the normalized* t*-th differential uniformity.*

For an (n,n)-function, clearly the n-th differential uniformity is the commonly defined differential uniformity in Section 2.1. The notion of the t-th differential uniformity was firstly proposed in [36, Section 6]. It was called the *chopping* of the functions in [36]. We prefer the notion of the t-th differential uniformity for the convenience of discussions in the following sections. As we will see below, this notion is useful to identify the nonlinear property of a function since some functions which are endorsed as Sboxes have poor t-th differential uniformity when $t < n$, while their n-th differential uniformity is optimal.

3.1. **Properties of the t-th differential uniformity.** Throughout this section we focus on the (n,n)-functions. One may easily see that $\Delta_F^t = 2^n$ for any t with $1 \le t \le n$ if F is a linear function. Therefore, for the applications in cryptography, we should require a function F having the property that Δ_F^t is as small as possible. The following result gives the lower and upper bound for the t-th differential uniformity.

Theorem 1. *Let* F *be an* (n,n)*-function and* t *be an integer with* $1 \le t \le n$*. Then* $2^{n-t} \le \Delta_F^t \le 2^{n-t}\Delta_F^n$.

Proof. For any subset $T \in \Omega_t$, let $F_T : \mathbb{F}_2^n \to \mathbb{F}_2^t$ be the function defined in (3). Define the group ring elements $D_F = \sum_{x \in \mathbb{F}_2^n} (x, F(x)) \in \mathbb{Z}[\mathbb{F}_2^n \times \mathbb{F}_2^n]$ and $D_{F_T} = \sum_{x \in \mathbb{F}_2^n} (x, F_T(x)) \in \mathbb{Z}[\mathbb{F}_2^n \times \mathbb{F}_2^t]$. It is easy to see that

$$D_F D_F^{(-1)} = 2^n + \sum_{a,b' \in \mathbb{F}_2^n, a \neq 0} \delta_F(a,b')(a,b'), \tag{4}$$

$$D_{F_T} D_{F_T}^{(-1)} = 2^n + \sum_{a \in \mathbb{F}_2^n, b \in \mathbb{F}_2^t, a \neq 0} \delta_F^T(a,b)(a,b), \tag{5}$$

where $\delta_F(a,b')$ and $\delta_F^T(a,b)$ are defined in Section 2.2 and Definition 1, respectively. Indeed, one can see that

$$\begin{aligned} D_F D_F^{(-1)} &= \left(\sum_{x \in \mathbb{F}_2^n} (x, F(x))\right)\left(\sum_{x \in \mathbb{F}_2^n} (x, F(x))\right) \\ &= \sum_{x,y \in \mathbb{F}_2^n} (x+y, F(x)+F(y)) \\ &= \sum_{a,x \in \mathbb{F}_2^n} (a, F(x+a)+F(x)) \\ &= 2^n + \sum_{a,x \in \mathbb{F}_2^n, a \neq 0} (a, F(x+a)+F(x)) \\ &= 2^n + \sum_{a,b' \in \mathbb{F}_2^n, a \neq 0} \delta_F(a,b')(a,b'). \end{aligned}$$

The Eq. (5) can be proven similarly. Now, let us apply the principal character of the group $\mathbb{F}_2^n \times \mathbb{F}_2^t$ on both sides of Eq. (5), then we have

$$\sum_{(a,b) \in (\mathbb{F}_2^n)^* \times \mathbb{F}_2^t} \delta_F^T(a,b) = 2^n(2^n - 1).$$

Since the maximal value of ℓ non-negative numbers must be greater than or equal to the average of their sum, we get

$$\Delta_F^t \ge \max_{(a,b) \in (\mathbb{F}_2^n)^* \times \mathbb{F}_2^t} \delta_F^T(a,b) \ge \frac{2^n(2^n-1)}{(2^n-1)2^t} = 2^{n-t}.$$

On the other hand, let $\rho : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^n \times \mathbb{F}_2^t$ be the natural homomorphism. Applying ρ on D_F we get $\rho(D_F) = \sum_{x\in\mathbb{F}_2^n}(x, F_T(x)) = D_{F_T}$. Therefore $\rho(D_F D_F^{(-1)}) = D_{F_T} D_{F_T}^{(-1)}$ and then $2^n + \sum_{a,b'\in\mathbb{F}_2^n, a\neq 0} \delta_F(a,b')(a,\rho(b')) = 2^n + \sum_{a\in\mathbb{F}_2^n, b\in\mathbb{F}_2^t, a\neq 0} \delta_F^T(a,b)(a,b)$. By comparing the coefficient of $(a,b) \in \mathbb{F}_2^n \times \mathbb{F}_2^t$ on both sides of the above equation we get $\sum_{a,b'\in\mathbb{F}_2^n, a\neq 0, \rho(b')=b} \delta_F(a,b') = \delta_F^T(a,b)$. For any $b \in \mathbb{F}_2^t$, clearly there are 2^{n-t} elements $b' \in \mathbb{F}_2^n$ satisfies $\rho(b') = b$. Since by definition Δ_F^t is the maximal value of $\delta_F^T(a,b)$ for $(a,b) \in (\mathbb{F}_2^n)^* \times \mathbb{F}_2^t$ and each $\delta_F^T(a,b) \leq 2^{n-t}\Delta_F^n$, we get the upper bound of Δ_F^t stated in the theorem. □

We call a function F has the *ideal t-th order differential uniformity* if $\Delta_F^t = 2^{n-t}$. Notice that, for a collection of nonnegative numbers, if the maximal value equals to the average value of them, then all the numbers are equal to the average value. Therefore, for the functions with the ideal t-th order differential uniformity, $\delta_F^T(a,b) = 2^{n-t}$ or 0 for all $(a,b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^t}$. The functions with the ideal t-th differential uniformity have been studied in other contexts. For instance, in [34], [42], a function $G : \mathbb{F}_{2^n} \to \mathbb{F}_{2^t}$ is called *perfect nonlinear* if $G(x+a)+G(x) = b$ has at most 2^{n-t} solutions for all $(a,b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^t}$. Using our language, the function F has ideal t-th order differential uniformity if and only if the function F_T are all perfect nonlinear for any $T \in \Omega_t$. It is well-known (see for instance [34]) that perfect nonlinear function does not exist when $t > n/2$. Therefore, we obtain the following result.

Proposition 1. *Let F be a function on $\mathbb{F}_{2^n}$. For any integer t with $n \geq t > n/2$, the t-th differential uniformity cannot be ideal.*

Remark 1. *It is well-known that EA and CCZ equivalence (see definitions in [10]) preserve the n-th differential uniformity. However, when $t < n$, the t-th order differential uniformity is not invariant under these equivalences, while they are invariant under the permutation equivalence defined in [43].*

The following definition serves to be a measurement for the distance between the function with ideal t-th order differential uniformity and an arbitrary function.

Definition 2. *Let F be an (n,n)-function. For an integer t with $1 \leq t \leq n$, the differential distinguisher U_F^t is defined as*

$$U_F^t = \left| \frac{\Delta_F^t}{2^n} - \frac{2^{n-t}}{2^n} \right| = \left| \frac{\Delta_F^t}{2^n} - \frac{1}{2^t} \right|.$$

Remark 2. *The differential distinguisher U_F^t is close to zero if a function is indistinguishable from the one with the ideal t-th differential uniformity.*

Particularly, the Sbox $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ used in many block ciphers is of the form

$$S = (S_1, S_2, \ldots, S_m) : (\mathbb{F}_{2^{n/m}})^m \to (\mathbb{F}_{2^{n/m}})^m,$$

where S_i is a function from $\mathbb{F}_{2^{n/m}}$ to itself for $1 \leq i \leq m$. The following result gives the lower bound for the n-th differential uniformity of S.

Proposition 2. *Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a function defined by $S(x) = \big(S_1(x_1), S_2(x_2), \ldots, S_{n/m}(x_{n/m})\big)$, where $x_i \in \mathbb{F}_{2^m}$ and $x = (x_1, \ldots, x_{n/m})$. Then $\Delta_S^n \geq 2^{n/m-1} \max\limits_{1\leq i\leq n/m} \Delta_{S_i}^m$.*

Proof. W.l.o.g. we may assume $\Delta_{S_1}^m = \max\limits_{1\leq i\leq n/m} \Delta_{S_i}^m$. Let $a, b \in \mathbb{F}_{2^m}$ such that $\delta_{S_1}(a,b) = \Delta_{S_1}^m$. Choosing $A = (a, 0, \ldots, 0), B = (b, 0, \ldots, 0) \in \mathbb{F}_{2^m}^{n/m}$, it is not difficult to see $\delta_S(A,B) = 2^{n/m-1}\delta_{S_1}(a,b)$ and then the result follows. □

Another characterization of the t-th differential uniformity for the (n,n)-functions is to use the Walsh transform. Recall that we denote the character group of a group G by $\widetilde{G}$.

Theorem 2. *Let F be an (n,n)-function and $F_T:\mathbb{F}_2^n\to\mathbb{F}_2^t$ be the function defined above. Then we have*

$$\frac{1}{2^{n+t}}\sum_{(a,b)\in\mathbb{F}_2^n\times\mathbb{F}_2^t}|\mathcal{W}_{F_T}(a,b)|^4=2^{2n}+\sum_{(a,b)\in(\mathbb{F}_2^n)^*\times\mathbb{F}_2^t}\delta_F^T(a,b)^2.$$

Proof. Define the group ring element $D_{F_T}=\{(x,F_T(x)):x\in\mathbb{F}_2^n\}\in\mathbb{Z}[\mathbb{F}_2^n\times\mathbb{F}_2^t]$. By Eq. (5) and by the Parseval's equation (2), we get the following

$$2^{2n}+\sum_{\substack{(a,b)\in\mathbb{F}_2^n\times\mathbb{F}_2^t\\ a\neq 0}}\delta_F^T(a,b)^2=\frac{1}{2^{n+t}}\sum_{\chi\in\widetilde{\mathbb{F}_2^n\times\mathbb{F}_2^t}}\chi\left(D_{F_T}D_{F_T}^{(-1)}\right)^2=\frac{1}{2^{n+t}}\sum_{\chi\in\widetilde{\mathbb{F}_2^n\times\mathbb{F}_2^t}}\chi\left|D_{F_T}\right|^4.$$

Notice that any character $\chi\in\widetilde{\mathbb{F}_2^n\times\mathbb{F}_2^t}$ can be represented by $\chi=\chi_a\chi_b$ defined by $\chi((x,y))=\chi_a(x)\chi_a(y)=(-1)^{\mathrm{Tr_1^n(ax)+Tr_1^m(by)}}$. Therefore, it is not difficult to see that $\chi\left(D_{F_T}\right)=\mathcal{W}_{F_T}(a,b)$ for some $(a,b)\in\mathbb{F}_2^n\times\mathbb{F}_2^t$. We finish the proof. □

By Theorem 2, we give another lower bound for the t-th differential uniformity as following.

Corollary 1. *Let the notations be the same as Proposition 2. Denoting*

$$C=\frac{1}{2^{n+t}}\sum_{(a,b)\in\mathbb{F}_2^n\times\mathbb{F}_2^t}|\mathcal{W}_{F_T}(a,b)|^4-2^{2n}.$$

Then $\max\limits_{(a,b)\in(\mathbb{F}_2^n)^*\times\mathbb{F}_2^t}\delta_F^T(a,b)^2\geq\frac{1}{2^t(2^n-1)}C$.

3.2. **The t-th differential uniformity of well-known Sboxes.** For the design of cryptographic algorithms, there are a lot of study on the construction of functions with good nonlinearity. We will discuss the t-th differential uniformities for some of them.

3.2.1. *Golden Sboxes.* In [29], all 4-bit Sboxes are classified up to EA and CCZ equivalence. These Sboxes are further studied in [43] and four of them are called *Golden* Sboxes. The following table lists the t-th differential uniformity and the corresponding differential distinguisher (see Definition 2) of these four Golden Sboxes. One may refer to the Appendix for the computational results of all 4-bit Sboxes used in previously proposed ciphers.

TABLE 1. The t-th differential uniformities of Golden Sboxes

No.	$(\Delta_S^1,U_{S,1})$	$(\Delta_S^2,U_{S,2})$	$(\Delta_S^3,U_{S,3})$	$(\Delta_S^4,U_{S,4})$
1	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
2	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
3	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
4	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)

From Table 1 we see that the Golden Sboxes have poor t-th ($t<4$) differential uniformity, while the 4-th differential uniformity of them is optimal.

3.2.2. *Sboxes endorsed in ZUC, SNOW 3G and AES.* The Sboxes endorsed in ZUC [37], SNOW 3G [38] and AES [40] are as following:

(1). ZUC [37]: The Sbox used in ZUC is of the form $S=(S_0,S_1,S_0,S_1)$ from $\mathbb{F}_{2^{32}}$ to itself, where S_0,S_1 are from $\mathbb{F}_{2^8}$ to itself, S_1 is the inverse function x^{-1} and S_0 can be found in [37, Page 12].

(2). `SNOW 3G` [38]: The Sbox $S_Q : \mathbb{F}_{2^8} \to \mathbb{F}_{2^8}$ used in `SNOW 3G` is $x + x^9 + x^{13} + x^{15} + x^{33} + x^{41} + x^{45} + x^{47} + x^{49}$, where the field $\mathbb{F}_{2^8}$ is defined by $x^8 + x^6 + x^5 + x^3 + 1$.

(3). `AES` [40]: The Sbox used in `AES` is of the form $(I, \cdots, I) : \mathbb{F}_{2^k} \to \mathbb{F}_{2^k}$, where $k = 128, 196, 256$ and I is a function that is EA-equivalent to the inverse function x^{-1}. One may refer to [40] for the details.

We use the following table to give the t-th differential uniformity for the functions S_0, S_Q and I.

TABLE 2. Differential uniformities of I, S_0, S_Q

t	$(\Delta^t_{S_0}, U^t_{S_0})$	$(\Delta^t_{S_Q}, U^t_{S_Q})$	(Δ^t_I, U^t_I)
1	(192, 1/4)	(176, 3/16)	(144, 1/16)
2	(160, 3/8)	(108, 11/64)	(88, 3/32)
3	(128, 3/8)	(78, 23/128)	(52, 5/64)
4	(96, 5/16)	(48, 1/8)	(30, 7/128)
5	(64, 7/32)	(34, 13/128)	(18, 5/128)
6	(32, 7/64)	(24, 5/64)	(10, 3/128)
7	(16, 11/128)	(16, 11/128)	(6, 1/32)
8	(8, 7/256)	(8, 7/256)	(4, 3/256)

For an (n, n)-function F, denote Γ_F by the highest differential distinguisher U^t_F for $1 \le t \le n$. From Table 2 we see that $\Gamma_{S_0} = 3/8$, $\Gamma_{S_Q} = 3/16, \Gamma_I = 3/32$. Recall that we prefer a function with small differential distinguisher, therefore from this point of view the Sboxes used in `AES` is the best.

3.2.3. *More on the inverse function.* It is well-known that the inverse function x^{-1} defined on $\mathbb{F}_{2^n}$ is an APN permutation if n is odd; and is a differentially 4-uniform permutation when n is even. It is also proven that the nonlinearity of the inverse function attains the known upper bound. Therefore it is chosen as the Sbox in `AES` and other ciphers. In the following, we will determine the 1-st and $(n-1)$-th differential uniformity of it.

Proposition 3. *Let $Inv(x) = x^{-1}$ be the inverse function from $\mathbb{F}_{2^n}$ to itself with $n \ge 3$. Let $k_n = \max\{k \equiv 0 \mod 8 | -2^{n/2+1} - 3 < k < 2^{n/2+1} + 1\}$. Then the followings hold:*

(1) *the 1-st differential uniformity of $Inv(x)$ is*

$$\Delta^1_{Inv} = \begin{cases} 2^{n-1} + 2^{n/2} & \text{if } n \text{ is even,} \\ 2^{n-1} + \frac{1}{2}k_n & \text{if } n \text{ is odd.} \end{cases}$$

(2) *The $(n-1)$-th differential uniformity of $Inv(x)$ equals 6 when n is even; and equals 4 when n is odd.*

Proof. Let the inverse function $\mathrm{Inv}(x)$ be written as the form $\mathrm{Inv}(x) = (F_1, \ldots, F_n)$, where F_i is a function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ for $1 \le i \le n$. Assume that $F_i(x) = \mathrm{Tr}(\beta_i \mathrm{Inv(x)})$ for some $\beta_i \in \mathbb{F}_{2^n}$.

(1) For $b \in \mathbb{F}_2$, we need to determine the number of solutions of the equation $b = F_i(x+a) + F_i(x)$. Assuming that there are A_0 elements $x \in \mathbb{F}_{2^n}$ such that $F_i(x+a) + F_i(x) = 0$ and there are A_1 elements $x \in \mathbb{F}_{2^n}$ such that $F_i(x+a) + F_i(x) = 1$. Obviously we have $A_0 - A_1 = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{F_i(x+a)+F_i(x)} := \mathcal{F}(D_a F_i)$ and $A_0 + A_1 = 2^n$. Therefore, we get $A_0 = 2^{n-1} + \frac{1}{2}\mathcal{F}(D_a F_i)$ and $A_1 = 2^{n-1} - \frac{1}{2}\mathcal{F}(D_a F_i)$. By the definition of Δ^1_{Inv} we have $\Delta^1_F = 2^{n-1} + \frac{1}{2}\max_{a \in \mathbb{F}^*_{2^n}, 1 \le i \le n} |\mathcal{F}(D_a F_i)|$. By [11, Corollary 1], the values of $\mathcal{F}(D_a F_i)$ takes any value s divisible by 8 in the range $[-2^{n/2+1} - 3, 2^{n/2+1} + 1]$. When n is even, the maximal value of $\mathcal{F}(D_a F_i)$ can attain is clearly $2^{n/2+1}$; and when n is odd we denote this maximal value by k_n. The rest of the proof follows from the definition of Δ^1_{Inv}.

(2) By the definition of $\Delta_{\mathrm{Inv}}^{n-1}$ we have

$$\Delta_{\mathrm{Inv}}^{n-1} = \max_{\substack{T\in\Omega_{n-1}\\(a,b)\in(\mathbb{F}_2^n)^*\times\mathbb{F}_2^{n-1}}} \{\delta_{\mathrm{Inv}}^T(a,b)\},$$

where $\rho : \mathbb{F}_2^n \to \mathbb{F}_2^{n-1}$ is the natural homomorphism defined by

$$\rho(x_1,\ldots,x_n) = (x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

and $T = \{1,\ldots,n\}\setminus\{i\}$ for some integer i with $1 \le i \le n$. By the proof of Theorem 1 we have that $\delta_{\mathrm{Inv}}^T(a,b) = \sum_{b'\in\mathbb{F}_2^n,\rho(b')=b} \delta_{\mathrm{Inv}}(a,b')$. Note that there are two elements $b', b'' \in \mathbb{F}_2^n$ such that $\rho(b') = \rho(b'') = b$ and they are related by $b'' = b' + (0,\ldots,0,1,0,\ldots,0)$ (the element 1 appears in the i-th bit). Therefore $\delta_{\mathrm{Inv}}^T(a,b) = \delta_{\mathrm{Inv}}(a,b') + \delta_{\mathrm{Inv}}(a,b'')$.

When n is even, the inverse function $\mathrm{Inv}(x)$ is a differentially 4-uniform function. The equation $(x+a)^{-1} + x^{-1} = c$ has 4 solutions if and only if $ac = 1$; and it has 2 solutions if and only if $ac \neq 1$ and $\mathrm{Tr}(1/(\mathrm{ac})) = 0$ (here $c \neq 0$ since $\mathrm{Inv}(x)$ is a permutation polynomial of $\mathbb{F}_{2^n}$, and $(x+a)^{-1} + x^{-1} = c = 0$ will not have 2 solutions). Note that a non-zero element $c \in \mathbb{F}_2^n$ with the property that $\mathrm{Tr}(1/\mathrm{c}) = 0$ if and only if $c = c_1 + c_1^{-1}$ for some nonzero element $c_1 \in \mathbb{F}_2^n$. First we can see that $\delta_{\mathrm{Inv}}^T(a,b) = \delta_{\mathrm{Inv}}(a,b') + \delta_{\mathrm{Inv}}(a,b'') \le 4+2 = 6$. Now let us prove there exists a subset $T \in \Omega_{n-1}$ and $a \in \mathbb{F}_2^n, b \in \mathbb{F}_2^{n-1}$ such that $\delta_{\mathrm{Inv}}^T(a,b) = 6$, and then $\Delta_{\mathrm{Inv}}^{n-1} = 6$ will follow. For convenience, we denote the element $(0,\ldots,0,1,0,\ldots,0)$ by w_i, where 1 appears in the i-th bit. Given any nonzero element $\beta \in \mathbb{F}_2^n$, define

$$\begin{array}{rcl} b' & = & \left((\beta+\beta^{-1}+1)w_i^{-1}\right)^{-1}, \\ b'' & = & b' + w_i, \text{ and} \\ a & = & b'^{-1}. \end{array} \tag{6}$$

Clearly $b := \rho(b') = \rho(b'') \in \mathbb{F}_2^{n-1}$. In the following we show that $\delta_{\mathrm{Inv}}^T(a,b) = \delta_{\mathrm{Inv}}(a,b') + \delta_{\mathrm{Inv}}(a,b'') = 6$, where $T = \{1,\ldots,n\}\setminus\{i\}$. First, we have $\delta_{\mathrm{Inv}}(a,b') = \{x \in \mathbb{F}_2^n | (x+a)^{-1} + x^{-1} = b'\} = 4$ since $ab' = 1$. Second, $\delta_{\mathrm{Inv}}(a,b'') = 2$ because by the definitions in (6) we have $ab'' = a(b'+w_i) = ab' + aw_i = 1 + ((\beta+\beta^{-1}+1)w_i^{-1})w_i = \beta + \beta^{-1}$, which follows that $\mathrm{Tr}(1/(\mathrm{ab}'')) = 0$ and then the equation $(x+a)^{-1} + x^{-1} = b''$ has 2 solutions. Finally, by the discussions above we have $\delta_{\mathrm{Inv}}(a,b) = 6$ and then $\Delta_{\mathrm{Inv}}^{n-1} = 6$.

When n is odd, notice that the inverse function $\mathrm{Inv}(x)$ is an APN function. The rest of proof is similar to the case n is even, we omit the details here. □

3.3. **Application of the t-th differential uniformity on truncated differential attack.** In the following we discuss the relationship between truncated differential attack on the block ciphers with the SPN structure and the t-th differential uniformity. Given a block cipher $\mathcal{B}$ with the SPN structure, assume that $\mathcal{B}$ has block size n, plaintext space $\mathcal{M}$, key space $\mathcal{K}$ and ciphertext space $\mathcal{C}$. Let $\mathrm{Enc}_K(\cdot)$ be the encryption function of $\mathcal{B}$ for $K \in \mathcal{K}$. An idea behind the truncated differential attack is as following. For the Sbox $S : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ of $\mathcal{B}$, the attacker needs to find an input difference $\Delta_1 \in \mathbb{F}_{2^n}$ and a set of output differences with the form

$$X = \{\ (*,\cdots,i_1,*,\cdots,*,i_k,\cdots,*) \in \mathbb{F}_2^n\ \}, \tag{7}$$

where $i_i,\cdots,i_k \in \mathbb{F}_2$ are constants and $*$ can be either 0 or 1, such that the probability

$$\Pr_{m\in\mathcal{M}}\left(\mathrm{Enc}_K(m) + \mathrm{Enc}_K(m+\Delta_1) \in X\right) = p \gg 0, \tag{8}$$

for some subset N of $\mathcal{M}$. To recover the key K, the attacker first guesses a key K' and compute the probability (8) to see whether it is around p. If this is true, there is a large probability that K' is the right key; otherwise the attacker excludes K' as the correct key.

On the other hand, we may regard the encryption function $\mathrm{Enc}_K(\cdot)$ as a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$. Finding the two-tuple (Δ_1, X) defined above is equivalent to finding the tuple

$a \in \mathbb{F}_2^n, b \in \mathbb{F}_2^k$ such that the $\mathrm{Enc}_K(a,b)$ is large for the function $\mathrm{Enc}_K : \mathbb{F}_2^n \to \mathbb{F}_2^{n-k}$ defined by $\mathrm{Enc}_K(x) = (f_{i_1}(x), \cdots, f_{i_k}(x))$, where $i_1, \cdots, i_k$ are the ones in (7).

4. Cryptographic Properties of the Compositions of Vectorial Boolean Functions

As mentioned in Section 1, in the design of many block ciphers (e.g. those with SPN structure) and many stream ciphers (e.g. `SNOW 3G` [38], `ZUC` [37]), there is one structure consisting of the composition of a linear function $a : \mathbb{F}_{2^k} \to \mathbb{F}_{2^n}$ and a nonlinear function (Sbox) $f : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$. However, as we will show below, for some linear function a, the nonlinearity of $F = f \circ a$ is much smaller than the nonlinearity of f, which weakens the ciphers against the linear attack.

4.1. The diversity of the nonlinearity.

Recall that the nonlinearity of a vectorial Boolean and a Boolean function is given in Section 2.2. Let F be a vectorial Boolean function from $\mathbb{F}_{2^n}$ to itself. Then the average of the absolute Walsh coefficients of F is $\mathrm{Ave}_{\lambda,\eta\in\mathbb{F}_{2^n}}|\mathcal{W}_F(\lambda,\eta)| = \sqrt{2^n}$. This is because that, by the Parseval's equation (2) (see also [10, page 24]), we have $2^{3n} = \sum_{\lambda,\eta\in\mathbb{F}_{2^n}} \mathcal{W}_F^2(\lambda,\eta)$ and then $\mathrm{Ave}(\mathcal{W}_F^2(\lambda,\eta)) = 2^{3n}/2^{2n} = 2^n$ and hence $\mathrm{Ave}(|\mathcal{W}_F(\lambda,\eta)|) = \sqrt{2^n}$. We introduce the following notion, which is adapted from the communication theory.

Definition 3. *For an (n,n)-function F, the diversity of the nonlinearity of F is defined as the ratio of the average value of extended Walsh spectrum (the set of the absolute values of the Walsh coefficients) to the maximal value of the Walsh spectrum of F:*

$$\mathrm{Div}_F = \left| \frac{\sqrt{2^n}}{2^n} - \frac{\max\limits_{\lambda,\eta\in\mathbb{F}_{2^n},\eta\neq 0} \left\{ \left| \mathcal{W}_F(\lambda,\eta) \right| \right\}}{2^n} \right|.$$

Assume that $\max\limits_{\lambda,\eta\in\mathbb{F}_{2^n},\eta\neq 0} \{|\mathcal{W}_F(\lambda,\eta)|\} = \sqrt{2^n}c$, we may further write $\mathrm{Div}_F = \left|\frac{1-c}{\sqrt{2^n}}\right|$.

Remark 3. (1) *Note that $1 \le c \le \sqrt{2^n}$ (the left inequality holds since the maximal value of the extended Walsh spectrum is greater or equal than the average value; and the right inequality holds since the linear functions having the maximal extended Walsh value 2^n). Hence, the diversity of the nonlinearity of F is to converge to 0 if a vectorial Boolean function has a good nonlinearity (clearly $\mathrm{Div}_F = 1/\sqrt{2^n}$ if F is a linear function). Therefore, to obtain a nonlinear function, we need Div_F to be as small as possible.*

(2) *Since the nonlinearity of an (n,n)-function is upper bounded by $2^{n-1} - 2^{(n-1)/2}$ when n is odd; and is conjectured upper bounded by $2^{n-1} - 2^{n/2}$ when n is even, the nonlinearity diversity of the functions attain the upper bounds is $\frac{\sqrt{2}-1}{\sqrt{2^n}}$ (resp. $\mathrm{Div}_I(x) = \frac{1}{\sqrt{2^n}}$) when n is odd (resp. even).*

Definition 4. *Let F, G be two (n,m) functions. The* cross-correlation function *between F and G is defined as*

$$C_{F,G}(\lambda,u,v) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_1^m(\mathrm{uF(x)+vG(\lambda x)})},$$

where $\lambda \in \mathbb{F}_{2^n}$ and $u, v \in \mathbb{F}_{2^m}$. Obviously, if $m = 1$ and G is the trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$, the cross-correlation between F and G becomes the Walsh transform of F.

4.2. **Diversity of the nonlinearity for substitution-permutation composition.** Let a, b be functions from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^n}$. For a given $\lambda \in \mathbb{F}_{2^k}$, we define the following $2^k \times 2n$ array:

$$M_{a,b}(\lambda) = \begin{pmatrix} a(x_0) & b(\lambda x_0) \\ a(x_1) & b(\lambda x_1) \\ \vdots & \\ a(x_{2^k-1}) & b(\lambda x_{2^k-1}) \end{pmatrix} \tag{9}$$

which is called an *image array* of a and b at λ (note that we write $a(x_j), b(\lambda x_j)$ as row vectors with length n).

Definition 5 (Simple image array)**.** *For $k \geq 2n$, the image array of $a(x)$ and $b(x)$ is called simple if, for any $\lambda \in \mathbb{F}_{2^k}$, the followings satisfy:*

1. *If for any $\eta \in \mathbb{F}_{2^n}$, there exists $x \in \mathbb{F}_{2^k}$ such that $b(\lambda x) \neq \eta b(x)$, and then the image array of a and b at λ defined in (9) is a $(2^k, 2n)$ orthogonal array, i.e. each $2n$-bit vector in $\mathbb{F}_2^{2n}$ occurs exactly 2^{k-2n} times in the image array;*
2. *If there exists $\eta \in \mathbb{F}_{2^n}$ such that $b(\lambda x) = \eta b(x)$ for all $x \in \mathbb{F}_{2^k}$, and then there exists a permutation function $h : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ such that any pair $(h(x), \eta x) \in \mathbb{F}_{2^n}^2$ occurs 2^{k-n} times in the image array of a and b, which now becomes*

$$M_{a,b}(\lambda) = \begin{pmatrix} a(x_0) & \eta b(x_0) \\ a(x_1) & \eta b(x_1) \\ \vdots & \vdots \\ a(x_{2^k-1}) & \eta b(x_{2^k-1}) \end{pmatrix}. \tag{10}$$

We will show below that if a, b are special linear functions, their image array is simple.

Proposition 4. *Let k and n be two positive integers with $k \geq 2n$ and $n \mid k$. Let a, b be two linear surjections from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^n}$ defined by $a(x) = \mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\gamma_1 \mathrm{x})$ and $b(x) = \mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\gamma_2 \mathrm{x})$, where $\gamma_1, \gamma_2 \in \mathbb{F}_{2^k}$. Then the image array of a and b is simple.*

Proof. First note that $\mathbb{F}_{2^k}$ can be regarded as the vector space over $\mathbb{F}_{2^n}$ of dimension $\ell = k/n$. For each $y \in \mathbb{F}_{2^n}$, we define the subsets $H_y^a = \{x \in \mathbb{F}_{2^k} | a(x) = y\}$ and $H_y^b = \{x \in \mathbb{F}_{2^k} | b(x) = y\}$. Since a, b are linear surjections, it is easy to see that the sizes of H_y^a and H_y^b are $2^{k-n} = 2^{n(\ell-1)}$ for any $y \in \mathbb{F}_{2^n}$. Furthermore, it is well known that the sets H_0^a, H_0^b are subspaces over $\mathbb{F}_{2^n}$ with dimension $\ell - 1$; and H_y^a, H_y^b are affine hyperplanes over $\mathbb{F}_{2^n}$ (i.e. $H_y^a = \alpha + H_0^a$ and $H_y^b = \beta + H_0^b$ for some $\alpha, \beta \in \mathbb{F}_{2^k}$).

Given $\lambda \in \mathbb{F}_{2^k}$, first assume that λ satisfies Definition 5(1), we need to show that each vector in $\mathbb{F}_{2^n}^2$ appears in the image array 2^{k-2n} times. For each $(y_1, y_2) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$, by [24, Lemma 4.2], there are 2^{k-2n} elements $x \in \mathbb{F}_{2^k}$ such that $\left(\mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\gamma_1 \mathrm{x}), \mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\lambda\gamma_2 \mathrm{x})\right) = (y_1, y_2)$, which implies (9) is a $(2^k, 2n)$ orthogonal array with every element in $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ appearing 2^{k-2n} times. Secondly, if λ satisfies Definition 5(2), namely $b(\lambda x) = \eta b(x)$ for some $\eta \in \mathbb{F}_{2^k}$, by [24, Lemma 4.1], there are 2^{k-n} elements $x \in \mathbb{F}_{2^k}$ such that $\left(\mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\gamma_1 \mathrm{x}), \mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\lambda\gamma_2 \mathrm{x})\right) = (y_1, y_2)$, where $y_2 \in \mathbb{F}_{2^n}$ and $y_1 = \lambda^{-1} h(y_2)$ for some permutation h of $\mathbb{F}_{2^n}$. We complete the proof. □

Remark 4. *(1) Proposition 4 is not true if a, b are linear functions but not of the form $\mathrm{Tr}_{\mathrm{n}}^{\mathrm{k}}(\gamma \mathrm{x})$. A counter example is given below. Let w be a primitive element of the field $\mathbb{F}_{2^4}$. Define the functions $a, b : \mathbb{F}_{2^4} \to \mathbb{F}_{2^2}$ as*

$$\begin{aligned} a(x) &= w^{12}x^8 + w^2x^4 + w^7x^2 + x, \\ b(x) &= w^{10}x^8 + w^9x^4 + w^9x^2 + w^8x. \end{aligned}$$

It is easy to verify that $\lambda = w^{14}$ *satisfies Def. 5(1) but the image array is*

$$\begin{array}{cccc} (0,0)^2 & (0,1)^2 & (1,w)^2 & (1,w^4)^2 \\ (w,w)^2 & (w,w^4)^2 & (w^4,0)^2 & (w^4,1)^2, \end{array}$$

where x^y *denotes the element* x *appears* y *times. Nevertheless, it is well-known that all linear Boolean functions can be written as the form* $\mathrm{Tr}(\gamma \mathrm{x})$*. Therefore, the image arrays for any two linear surjective Boolean functions are simple.*

(2) The condition that $n \mid k$ *is necessary. One may verify the image array of the functions* $a, b : \mathbb{F}_2^5 \to \mathbb{F}_2^2$ *that are defined by* $a(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2)$ *and* $b(x_1, x_2, x_3, x_4, x_5) = (x_2, x_1)$ *is not simple.*

Now let f, g be two functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$ and define the functions F, G from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^m}$ by $F = f \circ a$ and $G = g \circ b$. If the image array of a, b is simple, the cross-correlation of F and G is determined below.

Proposition 5. *Let* a, b *be functions from* $\mathbb{F}_{2^k}$ *to* $\mathbb{F}_{2^n}$ *and* f, g *functions from* $\mathbb{F}_{2^n}$ *to* $\mathbb{F}_{2^m}$*, where* $k \geq 2n$*. Let* $F = f \circ a$ *and* $G = g \circ b$ *be functions from* $\mathbb{F}_{2^k}$ *to* $\mathbb{F}_{2^m}$*. If* $a(x)$ *and* $b(x)$ *have the simple image array and* a, b *are surjections, then the cross-correlation between* F *and* G *is given by*

$$C_{F,G}(\lambda, u, v) = \begin{cases} 2^{k-2n}\mathcal{W}_{uf}(0)\mathcal{W}_{vg}(0), & \text{if } \lambda \text{ satisfies Def. 5(1),} \\ 2^{k-n}C_{f\circ h,g}(\eta, u, v), & \text{if } \lambda \text{ satisfies Def. 5(2).} \end{cases}$$

Proof. First assume that λ satisfies Def. 5(1). Now

$$\begin{aligned} C_{F,G}(\lambda, u, v) &= \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathrm{Tr}_1^m(\mathrm{uF(x)+vG}(\lambda \mathrm{x}))} \\ &= \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathrm{Tr}_1^m(\mathrm{uf(a(x))+vg(b}(\lambda \mathrm{x})))} \\ &= 2^{k-2n} \sum_{y_1,y_2\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_1^m(\mathrm{uf}(y_1)+\mathrm{vg}(y_2))} \text{ (letting } y_1 = a(x), y_2 = b(\lambda x)) \\ &= 2^{k-2n} \sum_{y_1\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_1^m(\mathrm{uf}(y_1))} \sum_{y_2\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}_1^m(\mathrm{vg}(y_2))} \\ &= 2^{k-2n}\mathcal{W}_{uf}(0)\mathcal{W}_{vg}(0). \end{aligned}$$

The case where λ satisfies Def. 5(2) can be proven similarly and we omit it here. □

In particular, when the functions G, a, b are linear, the cross-correlation between F and G becomes the Walsh transform of F, and we have the following result.

Theorem 3. *Let* a *be a linear surjection from* $\mathbb{F}_{2^k}$ *to* $\mathbb{F}_{2^n}$*, where* $k \geq 2n$ *and* $n \mid k$*. Let* f *be a function from* $\mathbb{F}_{2^n}$ *to* $\mathbb{F}_{2^m}$ *with* $m \mid n$*. Define the function* $F = f \circ a$*. If the image array of* a *and* $b(x) = \mathrm{Tr}_n^k(\mathrm{x})$ *is simple, then*

$$\mathcal{W}_F(\alpha, \beta) = \begin{cases} 0, & \text{if } \forall \eta \in \mathbb{F}_{2^n}, \exists x \in \mathbb{F}_{2^k} \text{ s.t. } \mathrm{Tr}_n^k(\alpha \mathrm{x}) \neq \eta \mathrm{Tr}_n^k(\mathrm{x}). \\ 2^{k-n}\mathcal{W}_{f\circ h}(\eta, \beta), & \text{if } \exists \eta \in \mathbb{F}_{2^n} \text{ s.t. } \mathrm{Tr}_n^k(\alpha \mathrm{x}) = \eta \mathrm{Tr}_n^k(\mathrm{x}) \text{ for all } x \in \mathbb{F}_{2^n}, \end{cases}$$

where h *and* η *is defined in Definition 5 and* $\alpha \in \mathbb{F}_{2^k}, \beta \in \mathbb{F}_{2^m}^*$*. Furthermore, the nonlinearity of* F *is given by* $\mathrm{NL}(F) = 2^{k-n}\mathrm{NL}(f \circ h)$*. The diversity of nonlinearity of* F *is*

$$\mathrm{Div}_F = \left| \frac{1 - 2^{k/2-n} \max_{\alpha'\in\mathbb{F}_{2^n}, \beta'\in\mathbb{F}_{2^m}^*} \{|\mathcal{W}_{f\circ h}(\alpha', \beta')|\}}{\sqrt{2^k}} \right|.$$

Proof. For any $\alpha \in \mathbb{F}_{2^k}$ and $\beta \in \mathbb{F}_{2^m}^*$, we have that

$$\begin{aligned}\mathcal{W}_F(\alpha,\beta) &= \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathrm{Tr_1^m}(\beta \mathrm{F(x)})+\mathrm{Tr_1^k}(\alpha \mathrm{x})} \\ &= \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathrm{Tr_1^m}(\beta \mathrm{f(a(x))}+\mathrm{Tr_m^n}(\mathrm{Tr_n^k}(\alpha \mathrm{x})))}.\end{aligned}$$

Now define $b(x) = \mathrm{Tr_n^k(x)}$ from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^n}$ and $g(x) = \mathrm{Tr_m^n(x)}$ from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$. By the above equation, we have $\mathcal{W}_F(\alpha,\beta) = \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathrm{Tr_1^m}(\beta \mathrm{f(a(x))+g(b(\alpha x))})} = C_{f\circ a, g\circ b}(\alpha,\beta,1)$. According to Proposition 5 and the assumption that the image array of a, b is simple, we have the following two cases:

(i) if $\forall \eta \in \mathbb{F}_{2^n}, \exists x \in \mathbb{F}_{2^k}$ s.t. $b(\alpha x) = \mathrm{Tr_n^k}(\alpha \mathrm{x}) \neq \eta \mathrm{Tr_n^k(x)}$, then $\mathcal{W}_F(\alpha,\beta) = C_{f\circ a, g\circ b}(\alpha,\beta,1) = 2^{k-2n}\mathcal{W}_{\beta f}(0)\mathcal{W}_g(0)$. It is easy to see that $\mathcal{W}_g(0) = 0$ since g is linear. Hence $\mathcal{W}_F(\alpha,\beta) = 0$ in this case.

(ii) if $\exists \eta \in \mathbb{F}_{2^n}$ s.t. $b(\alpha x) = \mathrm{Tr_n^k}(\alpha \mathrm{x}) = \eta \mathrm{Tr_n^k(x)}$ for all $x \in \mathbb{F}_{2^n}$, then there exists a permutation h over $\mathbb{F}_{2^n}$ such that $(h(y), \eta y)$ appears exactly 2^{k-n} times in the image array of a, b. Then

$$\begin{aligned}\mathcal{W}_F(\alpha,\beta) &= C_{f\circ a, g\circ b}(\alpha,\beta,1) = 2^{k-n} C_{f\circ h, g}(\eta,\beta,1) \\ &= 2^{k-n} \sum_{y\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr_1^m}(\beta \mathrm{f(h(y))})+\mathrm{Tr_1^n}(\eta \mathrm{y})} \\ &= 2^{k-n}\mathcal{W}_{f\circ h}(\eta,\beta).\end{aligned}$$

Following the above obtained result, for the function $F = f \circ a : \mathbb{F}_{2^k} \to \mathbb{F}_{2^m}$, we have

$$\begin{aligned}\mathrm{NL}(F) &= 2^{k-1} - \tfrac{1}{2}\max_{\alpha\in\mathbb{F}_{2^k},\beta\in\mathbb{F}_{2^m},\beta\neq 0} |\mathcal{W}_F(\alpha,\beta)| \\ &= 2^{k-1} - \tfrac{1}{2}\max_{\eta\in\mathbb{F}_{2^n},\beta\in\mathbb{F}_{2^m},\beta\neq 0} (2^{k-n}|\mathcal{W}_{f\circ h}(\eta,\beta)|) \\ &= 2^{k-1} - 2^{k-n-1}\max_{\eta\in\mathbb{F}_{2^n},\beta\in\mathbb{F}_{2^m},\beta\neq 0} |\mathcal{W}_{f\circ h}(\eta,\beta)| \\ &= 2^{k-1} - 2^{k-n-1}(2^n - 2\mathrm{NL}(f\circ h)) \\ &= 2^{k-n}\mathrm{NL}(f\circ h),\end{aligned}$$

where the last second equation uses the fact $\mathrm{NL}(f\circ h) = 2^{n-1} - \frac{1}{2}\max_{\eta\in\mathbb{F}_{2^n},\beta\in\mathbb{F}_{2^m},\beta\neq 0} |\mathcal{W}_{f\circ h}(\eta,\beta)|$. For the value of Div_F, we have

$$\begin{aligned}\mathrm{Div}_F &= \left| \frac{\sqrt{2^k}}{2^k} - \frac{\max\limits_{\alpha\in\mathbb{F}_{2^k},\beta\in\mathbb{F}_{2^m}^*} \left\{|\mathcal{W}_F(\alpha,\beta)|\right\}}{2^k} \right| \\ &= \left| \frac{1}{\sqrt{2^k}} - \frac{\max\limits_{\eta\in\mathbb{F}_{2^n},\beta\in\mathbb{F}_{2^m}^*} \left\{2^{k-n}|\mathcal{W}_{f\circ h}(\eta,\beta)|\right\}}{2^k} \right|.\end{aligned}$$

By simplifying the above equation we may get the value of Div_F in the theorem. □

Remark 5. *For $k \geq 2n$, from $\mathcal{W}_f(\alpha,\beta) \geq \sqrt{2^n}$ we may see that $\mathrm{Div}_F \gg 0$, which is not the ideal case for a function with high nonlinearity. This observations possibly can be exploited to propose some distinguishing attacks on the ciphers with the structure of Sbox and linear function composition.*

Now we give the t-th differential uniformity of the function $F = f \circ a$.

Theorem 4. *Let the notations be the same as above. For each t with $1 \leq t \leq m$, the t-th differential uniformity of F and f is related by $\Delta_F^t = 2^{k-n}\Delta_f^t$.*

Proof. Recall that F is from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^m}$ and f is from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$ with $k \geq 2n, n \geq m$. We write F as the form $F = (F_i, \ldots, F_m)$, where $F_i : \mathbb{F}_{2^k} \to \mathbb{F}_2$ is the component function

of F. For each t with $1 \leq t \leq m$ and for each $(b_1, \cdots, b_t) \in \mathbb{F}_2^t$, we need to consider the maximal number of solutions of the equations

$$\begin{cases} F_{i_1}(x+s) + F_{i_1}(x) &= b_1, \\ & \cdots \\ F_{i_t}(x+s) + F_{i_t}(x) &= b_t, \end{cases} \tag{11}$$

where $T = \{i_1, \cdots, i_t\}$ runs through Ω_t (the set of all subsets of $\{1, \ldots, m\}$ with cardinality t). Choosing a set of basis $\{\beta_1, \cdots, \beta_m\}$ of $\mathbb{F}_2^m$ over $\mathbb{F}_2$ such that $F_i(x) = \mathrm{Tr_1^m}(\beta_{\mathrm{i}}\mathrm{F(x)})$ for $1 \leq i \leq m$. Now, we have $F_{i_j}(x+c) + F_{i_j}(x) = \mathrm{Tr_1^m}(\beta_{\mathrm{i_j}}\mathrm{F(x+c)}) + \mathrm{Tr_1^m}(\beta_{\mathrm{i_j}}\mathrm{F(x)}) = \mathrm{Tr_1^m}(\beta_{\mathrm{i_j}}(\mathrm{f(a(x+c))+f(a(x))})) = \mathrm{Tr_1^m}(\beta_{\mathrm{i_j}}(\mathrm{f(a(x)+a(c))+f(a(x))}))$. By denoting the function $\mathrm{Tr_1^m}(\beta_{\mathrm{i_j}}\mathrm{f})$ by f_{ij} and letting $y = a(x)$, the equation (11) becomes

$$\begin{cases} f_{i_1}(y + a(c)) + f_{i_1}(y) &= b_1, \\ & \cdots \\ f_{i_t}(y + a(c)) + f_{i_t}(y) &= b_t. \end{cases}$$

By noticing that f_{ij} is the component function of f, the above equation has at most Δ_f^t solutions y. Clearly the size of the set $\{x : x \in \mathbb{F}_{2^k} | a(x) = y\}$ is 2^{k-n} for any $y \in \mathbb{F}_{2^n}$ since a is a linear surjection, and each solution $y \in \mathbb{F}_{2^n}$ will give rise to 2^{k-n} solutions $x \in \mathbb{F}_{2^k}$. Therefore, the system of equations (11) has at most $\Delta_F^t = 2^{k-n}\Delta_f^t$ solutions. □

We conclude this section by computing the diversity of the nonlinearity and the t-th differential uniformity for Sboxes used in ZUC [37].

Example 1. *In the design of* ZUC*, the Sboxes S_0, S_1 are composed with two linear function $L_1, L_2 : \mathbb{F}_{2^{32}} \to \mathbb{F}_{2^8}$ (L_i defined in [37] is from $\mathbb{F}_{2^{32}}$ to itself. Since we study the composition function $S_j \circ L_i$, by abuse of notations, the L_i from $\mathbb{F}_{2^{32}}$ to $\mathbb{F}_{2^8}$ we mentioned is the function $(L_{i,1}, \ldots, L_{i,8})$, where $L_{i,j}$ is the j-th component function of L_i). Following the notations in Theorem 3, we have $k = 32$ and $n = m = 8$. By a computer, one may verify that the image array of L_i and $\mathrm{Tr_n^k(x)}$ is simple, and then by Theorems 3 and 4, the nonlinearity diversity and differential distinguisher of $S_0 \circ L_i$ are as follows:*

$$\begin{aligned} \mathrm{Div}_{S_0 \circ L_i} &= \left|\frac{1-2^8 \max \mathcal{W}_{S_0}}{2^{16}}\right| = \left|\frac{1-2^{14}}{2^{16}}\right| = \frac{1}{4} \gg 0, \\ U^3_{S_0 \circ L_i} &= U^3_{S_0} = \frac{3}{8} \gg 0. \end{aligned}$$

One may also compute that $\mathrm{Div}_{S_0} = 0.1875$. Then we can see that the nonlinearity diversity of $S_0 \circ L_i$ is larger than S_0. We do not know whether this observation can be exploited to propose an attack on ZUC *or not.*

5. Concluding remarks and future work

In this paper we propose new criteria, called the t-th differential uniformity and the diversity of the nonlinearity, on the nonlinear property for functions that are defined on finite fields. These notions are the generalizations of the differential uniformity and nonlinearity, which are the measurements for a cipher against differential and linear cryptanalysis. We show that, though the differential uniformity of an Sbox is optimal, its t-th differential uniformity can be very poor. The lower and upper bounds, properties and characterizations of the t-th differential uniformity are studied. We also study the nonlinearity of a $f : \mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$ and the nonlinearity of the function $F = f \circ a : \mathbb{F}_{2^k} \to \mathbb{F}_{2^m}$, where a is a linear surjection from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^n}$. We show that, for special a, the nonlinearity of F is greatly decreased, which weakens the ciphers against the linear attack.

Recently, for the cipher Present, by observing the 1-st differential uniformity of its Sbox is the same as a linear function, Tezcan succeeded in giving an attack on the round-reduced Present. As we provided in the paper, the Sboxes endorsed by SNOW 3G, ZUC and some lightweight block ciphers also have similar weakness, it will be very interesting to propose more attacks on these ciphers by exploiting these weakness.

Acknowledgement

We thank the anonymous reviewers for the valuable comments, which significantly improve the quality and presentation of this paper. We thank Cihangir Tezcan for sending us his paper [47].

References

[1] R. Anderson, E. Biham and L. Knudsen, `Serpent`: A proposal for the Advanced Encryption Standard, `http://www.cl.cam.ac.uk/~rja14/Papers/serpent.pdf`, (1999).
[2] F. Armknecht, Improving fast algebraic attacks, in 11th International Workshop on Fast Software Encryption, FSE 2004. Lecture Notes in Computer Science, vol. 3017, 65–82, (2004).
[3] E. Biham and A. Shamir, Differential cryptanalysis of DES-like cryptosystems, Journal of Cryptology: 4(1): 3–72, (1991).
[4] A. Bogdanov, L.R. Knudsen, G. Leander, C. Paar, A. Poschmann, M.J.B. Robshaw, Y. Seurin, and C. Vikkelsoe, `PRESENT`: An ultra-lightweight block cipher. In P. Paillier and I. Verbauwhede (Eds.): CHES 2007, LNCS 4727, 450–466, (2007).
[5] C. Boura, A. Canteaut, C. Cannire, Higher-Order Differential Properties of Keccak and Luffa, In A. Joux (Eds.): FSE 2011, LNCS 6733, 252–269, (2011).
[6] S. Bulygin, M. Walter, J. Buchmann, Full analysis of `PRINTcipher` with respect to invariant subspace attack: efficient key recovery and countermeasures, Designs, Codes and Cryptography 73(3), 997-1022, (2014).
[7] C. Boura, A. Canteaut, On the Influence of the Algebraic Degree of F^{-1} on the Algebraic Degree of $G \circ F$, IEEE Transactions on Information Theory 59(1): 691-702, (2013).
[8] C. De Cannire, H. Sato, and D. Watanabe: Hash Function `Luffa` - Specification Ver. 2.0.1, NIST SHA-3 Submission, Round 2 document, (2009).
[9] C. Carlet, Boolean functions for cryptography and error correcting codes, chapter of a monograph. In: Crama Y., Hammer P. (eds.) Boolean Models and Methods in Mathematics, Computer Science, and Engineering, pp. 257397. Cambridge University Press, Cambridge (2010). Preliminary version available at `http://www.math.univ-paris13.fr/~carlet/`.
[10] C. Carlet, Vectorial Boolean functions for cryptography, chapter of a monograph. In: Crama Y.,Hammer P. (eds.) Boolean Models and Methods in Mathematics, Computer Science, and Engineering, pp. 398 472. Cambridge University Press, Cambridge (2010). Preliminary version available at `http://www.math.univ-paris13.fr/~carlet/`.
[11] P. Charpin, T. Helleseth, V. Zinoviev, Progagation characteristics of $x \mapsto x^{-1}$ and Kloosterman sums, Finite Fields and Their Applications: 13, 366–381, (2007).
[12] N. Courtois, Fast algebraic attacks on stream ciphers with linear feedback. In Advances in Cryptology - CRYPTO 2003, volume 2729 of Lecture Notes in Computer Science, pages 177–194. Springer-Verlag, (2003).
[13] N. Courois and W. Meier, Algebraic attack on stream ciphers with linear feedback, in Advances in Cryptology - EUROCRYPTO 2003 Lecture Notes in Computer Science, Springer-Verlag, vol. 2656, 345-359, (2003).
[14] J. Daemen, M. Peeters, G. Van Assche, and V. Rijmen: Nessie Proposal: `NOEKEON`: NESSIE Proposal, 27 October 2000.
[15] J.F. Dillon, APN polynomials: an update, In Conference Finite Fields and Applications Fq9, Dublin, Ireland (2009).
[16] M. Duan, X. Lai, Improved zero-sum distinguisher for full round Keccak-f permutation, IACR ePrint Archive 2011: 23 (2011).
[17] V. Dolmatov, `GOST` 28147-89: Encryption, Decryption, and Message Authentication Code (MAC) Algorithms, Internet Engineering Task Force RFC 5830, March 2010.
[18] D. Engels, X. Fan, G. Gong, H. Hu, and E. M. Smith: `Hummingbird`: Ultra-Lightweight Cryptography for Resource-Constrained Devices, In R. Sion et al. (Eds.): FC 2010 Work- shops, LNCS 6054, pp. 3–8. Springer (2010).
[19] D. Engels, M.-J. O. Saarinen, P. Schweitzer, and E. M. Smith: The `Hummingbird-2` Lightweight Authenticated Encryption Algorithm, RFIDSec 2011, The 7th Workshop on RFID Security and Privacy: 2628, June 2011, Amherst, Massachusetts, USA (2011).
[20] X. Fan, K. Mandal, G. Gong, `WG-8`: A Lightweight Stream Cipher for Resource-Constrained Smart Devices, `http://cacr.uwaterloo.ca/techreports/2012/cacr2012-28.pdf`, (2012).
[21] S. Golomb and G. Gong, Signal design for good correlation, *Cambridge University Press*, (2005).
[22] G. Gong and S.W. Golomb, The Decimation-Hadamard transform of two-level autocorrelation sequences, IEEE Transaction on Information Theory: 48 (4), 853-865, (2002).

[23] J. Guo, T. Peyrin, A. Poschmann, and M. Robshaw, The LED Block Cipher, In B. Preneel and T. Takagi (Eds.): CHES 2011, 326-341, (2011).
[24] A. Klapper, A. H. Chan and M. Goresky, Cross-correlations of linearly and quadratically related geometric sequences and GMW sequences, Discrete Applied Mathematics: 46, 1–20, (1993).
[25] O. Kucuk, The Hash Function Hamsi, NIST SHA-3 Submission, Round 2 document, 14 September 2009.
[26] L. Knudsen, Truncated and Higher Order Differentials, FSE'94, 196211, (1995).
[27] L.R. Knudsen, G. Leander, A. Poschmann, and M.J.B. Robshaw, Printcipher: A block cipher for IC-Printing, In Stefan Mangard and Francois-Xavier Standaert, editors, Proc. CHES 2010, vol. 6225 of LNCS, 16–32, 2010.
[28] X. Lai, Higher order derivatives and differential cryptanalysis, in Proc. Symp. Commun., Coding Cryptography, 227-233, in honor of J. L. Massey on the occasion of his 60th birthday, Kluwer Academic Publishers, (1994).
[29] G. Leander, A. Poschmann, On the classification of 4 bit Sboxes, C. Carlet and B. Sunar (Eds.): WAIFI 2007, LNCS 4547, 159-176, (2007).
[30] G. Leander, M. A. Abdelraheem, H. Alkhzaimi and E. Zenner, A Cryptanalysis of PRINTcipher: The invariant subspace attack, In: Proceedings of Cypto 2011, LNCS Vol 6841, 206–221, (2011).
[31] R. Lidl and H. Niederreiter, Finite Fields, Cambridge University Press, (1997).
[32] Y. Y. Luo, Q. Chai, G. Gong and X. J. Lai, A Lightweight Stream Cipher WG-7 for RFID Encryption and Authentication, GLOBECOM, 1–6, (2010).
[33] M. Matsui, Linear cryptanalysis method for DES cipher, Advances in Cryptology - EUROCRYPT 1993.
[34] K. Nyberg, Perfect Nonlinear Sboxes. EUROCRYPT 1991: 378-386.
[35] K. Nyberg, L. R. Knudsen: Provable Security Against Differential Cryptanalysis. CRYPTO 1992: 566-574
[36] K. Nyberg, Sboxes and round functions with controlled linearity and differential uniformity, FSE 94, LNCS 1008, 111-130, (1995).
[37] NIST, Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3. Document 2: ZUC Specification, http://www.dacas.cn/thread.aspx/ID=2304.
[38] NIST, Specification of the 3GPP Confidentiality and Integrity Algorithms UEA2 & UIA2. Document 2: SNOW 3G Specification, www.gsma.com/technicalprojects/wp-content/uea2uia2.
[39] NIST: Data Encryption Standard, FIPS PUB 46. National Bureau of Standards, U.S. Department of Commerce, Washington D.C., 15 January 1977.
[40] NIST, The Advanced Encryption Standard, FIPS 197, csrc.nist.gov/publications/fips/fips197/fips-197.pdf.
[41] D.S. Passman, The Algebraic Structure of Group Rings, Wiley-Interscience, New York, (1977).
[42] A. Pott, Nonlinear functions in abelian groups and relative difference sets, Discrete Applied Mathematics: 138(1-2): 177–193, (2004).
[43] M. O. Saarinen, Cryptographic Analysis of All 4×4 Sboxes, In A. Miri and S. Vaudenay (Eds.) SAC 2011, LNCS 7118, 118-133, (2011).
[44] B. Schmidt, On (p^a, p^b, p^a, p^{a-b})-relative difference set, Journal of Algebraic Combinatorics: 6, 279–297, (1997).
[45] A. Sorkin, Lucifer: A cryptographic algorithm, Cryptologia, Vol. 8, No. 1, pp. 22-32. (1984).
[46] F.-X. Standaert, G. Piret, G. Rouvroy, J.-J. Quisquater, and J.-D. Legat, ICEBERG: An Involutional Cipher Efficient for Block Encryption in Reconfigurable Hardware, In B. Roy and W. Meier (Eds.): FSE 2004, LNCS 3017, pp. 279-299, Springer (2004).
[47] C. Tezcan, Improbable differential attacks on Present using undisturbed bits, Journal of Computational and applied mathematics 259(B), 503–511, (2014).
[48] J. V. Uspensky, Introduction to Mathematical Probability, New York McGraw-Hill, (1937).
[49] H. Wu: The Hash Function JH, NIST SHA-3 Submission, Round 3 document, 16 January, 2011.

Appendix

The t-th differential uniformity of the 4-bit Sboxes used in previous ciphers

Table 3. t-th Differential Uniformities of 4-bit Sboxes

Ref	$(D_S^1, U_{S,1})$	$(D_S^2, U_{S,2})$	$(D_S^3, U_{S,3})$	$(D_S^4, U_{S,4})$
Lucifer S_0 [45]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(6, 5/16)
Lucifer S_1 [45]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(6, 5/16)
Present [4]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Present^{-1} [4]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
JH S_0 [49]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
JH S_1 [49]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(4, 3/16)
ICEBERG0 [46]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
ICEBERG1 [46]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
LUFFA [8]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
NOEKEON [14]	(16, 1/2)	(16, 3/4)	(8, 3/8)	(4, 3/16)
HAMSI [25]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
HB1 S_0 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1 S_1 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1 S_2 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1 S_3 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1-1 S_0 [18]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
HB1-1 S_1 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1-1 S_2 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1-1 S_3 [18]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2 S_0 [19]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2 S_1 [19]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2 S_2 [19]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
HB2 S_3 [19]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S_0 [19]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S_1 [19]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S_2 [19]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S_3 [19]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
DES S_0-0 [39]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
DES S_0-1 [39]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(8, 7/16)
DES S_0-2 [39]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
DES S_0-3 [39]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S_1-0 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S_1-1 [39]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(8, 7/16)
DES S_1-2 [39]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S_1-3 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S_2-0 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S_2-1 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S_2-2 [39]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S_2-3 [39]	(16, 1/2)	(16, 3/4)	(8, 3/8)	(8, 7/16)
DES S_3-0 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S_3-1 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S_3-2 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S_3-3 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S_4-0 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)

DES S_4-1 [39]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(6, 5/16)
DES S_4-2 [39]	(16, 1/2)	(10, 3/8)	(8, 3/8)	(6, 5/16)
DES S_4-3 [39]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(6, 5/16)
DES S_5-0 [39]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
DES S_5-1 [39]	(12, 1/4)	(10, 3/8)	(10, 1/2)	(8, 7/16)
DES S_5-2 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S_5-3 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S_6-0 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S_6-1 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S_6-2 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S_6-3 [39]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S_7-0 [39]	(16, 1/2)	(12, 1/2)	(12, 5/8)	(6, 5/16)
DES S_7-1 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(10, 9/16)
DES S_7-2 [39]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S_7-3 [39]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(8, 7/16)
Serpent S_0 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
Serpent S_1 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent S_2 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
Serpent S_3 [1]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(8, 7/16)
Serpent S_4 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
Serpent S_5 [1]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
Serpent S_6 [1]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
Serpent S_7 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(8, 7/16)
Serpent-1 S_0 [1]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
Serpent-1 S_1 [1]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
Serpent-1 S_2 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S_3 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S_4 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S_5 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
Serpent-1 S_6 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S_7 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K_1 [17]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K_2 [17]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
GOST K_3 [17]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K_4 [17]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K_5 [17]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K_6 [17]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
GOST K_7 [17]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
GOST K_8 [17]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)

Department of Electrical and Computer Engineering, University of Waterloo, Canada
E-mail address, Yin Tan: yin.tan@uwaterloo.ca

Department of Electrical and Computer Engineering, University of Waterloo, Canada
E-mail address, Guang Gong: ggong@uwaterloo.ca

Department of Electrical and Computer Engineering, University of Waterloo, Canada
E-mail address, Bo Zhu: bo.zhu@uwaterloo.ca