Enhanced Criteria on Differential Uniformity and Nonlinearity of Cryptographically Significant Functions

Guang Gong, Yin Tan, and Bo Zhu

Department of Electrical and Computer Engineering, University of Waterloo
Waterloo, Ontario, Canada.
{ggong, y24tan, bo.zhu}@uwaterloo.ca

Abstract. The Substitution box (Sbox) plays an important role in the design of both block and stream ciphers as, in most cases, it is the only nonlinear layer of the cipher. There are many requirments of the choice of the Sbox to aviod various attacks. In this paper, we mathematically characterizes the truncated differential cryptanalysis invented by Kundsen, which is called the t-th differential uniformity in this paper. It is found that, although some S-boxes used in certain previously ciphers have optimal differential uniformity, they have very bad t-th differential uniformity for some integer t. This gives an entile possibility to improve the cryptanalysis of these ciphers. As an application of this concept, we propose a construction of a distinguisher of a block cipher by the bias rate of some differential sequences defined in this paper, which is obtained via the t-th differential uniformity. By this algorithm, we found that there are many keys in the recently proposed lightweight LED-64 are weak, namely they are distinguishable from a uniform cipher. Furthermore, we compute the t-th differential uniformity of all 4-bit S-boxes used in previous designed ciphers, and the result of the Sbox used in ZUC, SNOW 3G and AES. Particularly, it is found that the Sbox used in ZUC has very bad 3rd differential uniformity. The second contribution of this paper is to target the following scenario. For many ciphers, there exists a structure consisting of the composition of a linear function a from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^n}$ and an Sbox S defined on $\mathbb{F}_{2^n}$, which aims to bring the confusion and diffusion into the cipher. We define a new criteria, *diversity of nonlinearity*, to measure the nonlinearity of such a function. By using this tool, we show that the nonlinearity of $S \circ a$ is greatly decreased compared to S, which makes the cipher weaker against the linear attack. Finally, we present the relationship between the t-th differential uniformity and the nonlinearity. By this relationship, we give a fast method to tell whether a function has ideal t-th differential uniformity.

Keywords: Substitution box, block cipher, differential attack, linear attack, distinguish attack.

1 Introduction

The Substitution box, or Sbox in short, plays an important role in the design of block and stream ciphers as it is the only nonlinear part of the cipher in most cases. To avoid various attacks on the ciphers and for efficient software implementation, S-boxes are required to satisfy a lot of properties, for instance being a permutation defined on the fields with even degrees [31], with a high algebraic degree [9], a low differential uniformity [2] and a high nonlinearity [26], etc. However, it seems very difficult to find an Sbox to satisfy all the criteria. It is widely believed that, although the Sbox has certain weakness, we may still get cryptographically strong function by iterating the Sbox several rounds. Recently, some attacks have been discovered due to the weakness of the Sbox even the cipher iterates many rounds. For instance, if the Sbox is an almost bent function, the algebraic degree of the cipher increases slow and it yields the zero-sum attack ([5], [4], [12]). The aim of this paper is to discover more attacks by using certain weakness of the S-boxes.

The differential cryptanalysis proposed in [2] and the techniques derive from it are powerful tools to analyze block ciphers. To avoid these attacks, we usually require the Sbox of the cipher to have a low differential uniformity. For instance the inverse function, endorsed by AES, is a differentially 4 uniform permutation on $\mathbb{F}_{2^8}$. In this paper, we give a mathematical characterization of the truncated differential attack invented by Kundsen in [22], called the t-th differential uniformity (see the details in Section 3). We found that, although some S-boxes endorsed by certain ciphers have good differential uniformity, they have very bad t-th differential uniformity for some t. Actually, this method has also been used by Nyberg in [29] to study the properties of APN and bent functions. The relationship between the t-th differential uniformity and the (almost) perfect nonlinear functions, relative difference sets, and the basic properties of t-th differential uniformity can be found in Section 3.

Thanks to the classification of the 4-bit S-boxes in [23] and [34], we compute the t-th differential uniformities for all 4-bit S-boxes. The computational results reflect that many of them do not have good t-th differential uniformity for some integer $1 \leq t \leq 3$, while they all achieve the lowest differential uniformity 4. We give a list of the computational results for all 4-bit S-boxes used in previously proposed ciphers in Section 4. Except the S-boxes used in block ciphers, the t-th differential uniformities of the S-boxes endorsed by the recently standarized stream ciphers ZUC [40] and SNOW 3G [41] are also computed. It is found that the Sbox S_0 in ZUC has very poor 3-rd differential uniformity, which gives a possibility to improve the cryptanalysis on it.

Inspired by the bad t-th differential uniformity of some S-boxes, we propose a construction of a distinguisher on a general block cipher $\mathcal{B}$ with known S-boxes in Section 5. As an application, we analyze the random properties of the recently proposed lightweight cipher LED-64 ([19]), and found there exist many keys such that the cipher is distinguishable to a uniform cipher.

Another contribution of this paper is as follows. In the design of many stream ciphers, there is a structure consisting of the composition of a linear function $a : \mathbb{F}_{2^k} \rightarrow \mathbb{F}_{2^n}$ and the S-box $S : \mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^m}$. The purpose of this design is to bring the confusion and diffusion into the cipher. For any function $F : \mathbb{F}_{2^k} \rightarrow \mathbb{F}_{2^m}$, we define the *nonlinearity diversity* of F by (see details in Section

6.2)

$$Div_F = \left| \frac{\sqrt{2^k} - \max\{\mathcal{W}_F\}}{2^k} \right|.$$

One may see that, for a linear function, the nonlinearity diversity is $\frac{1}{\sqrt{2^k}}$ and hence to obtain a highly nonlinear function, we need the nonlinearity diversity to be as **small** as possible. In Section 6, we determine the nonlinearity diversity of the composition of an Sbox and a linear function mentioned above (Theorem 6). As a result, we see that the nonlinearity diversity of the S-box S is greatly amplified by its composition with the linear function f, which points out a drawback of such design. We need to mention that, a similar problem has also been considered by Nyberg in [29], however, in this paper, we first present a more general result on the crosscorrelation of two functions F and G (defined in Section 6) which is of independent interest, and then, as an application, the nonlinearity of the composition of an S-box and a linear function is obtained.

Finally, for a function F defined on $\mathbb{F}_{2^n}$, to the best of our knowledge, there is no fast method to determine its t-th differential uniformity. We present the relationship between the Walsh spectrum and t-th differential spectrum in Section 7. By the butterfly algorithm introduced in [7] and the relationship, we may efficiently determine a lower bound of the t-th differential uniformity. Particularly, if the function F is an APN function or differentially 4-uniform function, we give an explicit formula of its differential spectrum by means of its Walsh spectrum.

The rest of the paper is organized as follows. In Section 2, necessary definitions and results are given. The definitions of t-th differential uniformity and its relationship to truncated differential attack, include its basic properites are presented in Section 3. The discussions of the t-th differential uniformity of all 4-bit S-boxes and S-boxes used in ZUC, SNOW 3G, AES can be found in Section 4. In Section 5, we propose an algorithm to construct a distinguisher for a general block cipher by using the t-th differential uniformity. The test of the random properties of LED-64 is given therein. In Section 6, we show the influence of the composition of an Sbox $\mathbb{F}_{2^m} \rightarrow \mathbb{F}_{2^m}$ with a linear function $\mathbb{F}_{2^k} \rightarrow \mathbb{F}_{2^m}$ on the nonlinearity diversity. The relationship between t-th differential uniformity and nonlinearity is presented in Section 7.

2 Preliminaries

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

2.1 Group rings and characters.

We briefly introduce some algebraic tools which is used in the following. For the definitions and properties of the character theory and group ring theory of the abelian groups in this subsection, one may refer to [24] and [32], respectively. Notice that for a subset D of G, we may regard it as a group ring element $\sum_{d\in D} d \in \mathbb{C}[G]$. By abusing using the notations, we still denote it by D.

Let G be an abelian group and $\widetilde{G}$ be its character group. For a group ring elemnt $D = \sum_{g\in G} d_g g \in \mathbb{C}[G]$, the following equation

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

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

$$DD^{(-1)} = \sum_{t\in G} \left(\sum_{g\in G} d_g d_{g^{-1}t} \right) t.$$

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

$$\sum_{g\in G} d_g d_{g^{-1}t} = \frac{1}{|G|} \sum_{\chi\in\widehat{G}} \chi\left(DD^{(-1)}\right) \chi(t^{-1}) = \frac{1}{|G|} \sum_{\chi\in\widehat{G}} |\chi(D)|^2 \chi(t^{-1}).$$

As a special case, if $t = 1$, we have the *Parseval's equation*

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

2.2 Bias rate

Let $\mathbf{a}$ be a binary sequence of length N, we call the value $|N_0 - N_1|$ the *imbalance* of a, where N_0, N_1 are the number of 0 and 1 appeared in $\mathbf{a}$ respectively. Furthermore, we call the value $|N_0 - N_1|/N$ the *imbalance rate* of $\mathbf{a}$. The following result gives the expectation of the imbalance rate of a random sequence $\mathbf{a}$. We include a short proof for the convenience of the readers.

Theorem 1. *[38] Let $\mathbf{a}$ be a sequence generated by a random generator with length N, then the expectation of the imbalance of $\mathbf{a}$ is $\sqrt{N/2\pi}$, or equivalently, the expectation of the imbalance rate is $\sqrt{1/2\pi N}$.*

Proof. Let $\xi = |S_0 - S_1|$ be the imbalance of $\mathbf{a}$, where $S_0 = \sharp\{i \in [1..n] | a_i = 0\}$ and $S_1 = n - S_0$. Then ξ is a random variable with values in the range $0 \le \xi \le n$. In the following we only discuss the case n is an even integer. The arguments of the case n is odd is similar and we omit it here. Now assume $\xi = k$ and we get $k = |S_0 - S_1| = |2S_0 - n|$ and then $S_0 = \frac{n\pm k}{2}$. Clearly the equation is only meaningful when k is even. By the symmetry we only consider the case $S_0 = \frac{n+k}{2}$. Now we have

$$\Pr(\xi = k) = 2 \cdot \binom{n}{\frac{n+k}{2}} \cdot \frac{1}{2^n} = \frac{1}{2^{n-1}} \cdot \binom{n}{\frac{n+k}{2}} .$$

Therefore, the expectation of ξ is

$$E(\xi) = \sum_{k=0,k\ even}^{n} \Pr(\xi_k)k = \frac{1}{2^{n-1}} \sum_{k} \binom{n}{\frac{n+k}{2}} k.$$

By the fact that $\lim\limits_{n\to\infty} \frac{1}{2^{n-1}} \sum\limits_{k} \binom{n}{\frac{n+k}{2}} k = \sqrt{n/2\pi}$ (which can be shown by $\sum\limits_{k} \binom{n}{\frac{n+k}{2}} k = \frac{n}{2}\binom{n}{n/2}$ and the sterling formula $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$), we finish the proof. □

3 The t-th Differential Uniformity

In this section, we define the t-th differential uniformity of an S-box and give their basic properties. The relationship of the t-th differential uniformity and the truncated differential attack, (almost) perfect nonlinear functions is also discussed. We first give the definition of t-th differential uniformity.

Let F be a function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$ and t be an integer with $1 \leq t \leq n$. For any subset $T = \{i_1, i_2, \cdots, i_t\} \subseteq \{1, 2, \cdots, n\}$ with size t, define the function $F_T : \mathbb{F}_{2^n} \to \mathbb{F}_{2^t}$ by

$$F_T(x) = \left(f_{i_1}(x), \cdots, f_{i_t}(x)\right), \tag{3}$$

where $f_{i_j}(x)$ is the i_j-th component functions of F for $1 \leq j \leq t$. For any $(a, b) \in \mathbb{F}_{2^n}^* \times \mathbb{F}_{2^t}$, define $\delta_F^T(a, b)$ by

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

We call the value

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

the *t-th differential uniformity* of F, where Ω_t is the set of all subsets of $\{1, \cdots, n\}$ with size t. Note that the n-th differential uniformity is the commonly defined differential uniformity, see for instance [7]. Furthermore, the value

$$\mathcal{D}_F^t = \frac{\Delta_F^t}{2^n}$$

is called the *normalized* t-th differential uniformity.

We should mention that, in [29, Section 6], Nyberg was the first who studies the properties of APN and bent functions by t-th differential uniformity, which is called *chopping* of functions therein. It is also easy to see that, for functions defined on fields with even characteristic, the lowest possible n-th differential uniformity is 2, which are called *almost perfect nonlinear* (APN) functions in [28]. However, as we will see below, some functions used as S-boxes in certain block ciphers have bad t-th differential uniformity when $t < n$, while they have good n-th differential uniformity. It is entirely possible that such a property offers improvements to the cryptanalysis of the ciphers using these functions as S-boxes. We will illustrate this by an example in the next sections.

3.1 Properties of t-th differential uniformity

One may easily see that $\Delta_F^t = 2^n$ if F is a linear function. Therefore, to obtain a highly nonlinear function F, we should require Δ_F^t to be as **small** as possible. The following result gives a lower bound of the t-th differential uniformity.

Proposition 1. *Let F be a function defined on $\mathbb{F}_{2^n}$, for each integer t with $1 \leq t \leq n$, the t-th differential uniformity*

$$\Delta_F^t \geq 2^{n-t}.$$

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). Clearly, 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 their sum, we get

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

We finish the proof. □

We call the function F has the *ideal t-th order differential uniformity* if $\Delta_F^t = 2^{n-t}$. Note 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 function having 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 following definition serves to be a measurement for the distance between the function with ideal t-th order differential uniformity and an arbitrary function F:

Definition 1. *A differential distinguisher $U_{F,t}$ is defined as*

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

This value is close to zero if a function is indistinguishable from a function with ideal t-th order of differential uniformity. Thus, the **larger** U_F^t, the easier it is distinguishable from the one with ideal differential uniformity. The tuple $\left(t, \Delta_F^t, U_F^t, A\right)$ is called a t-differential, where A is the set of elements a of $\mathbb{F}_{2^n}$ such that $\delta_F^T(a,b) = \Delta_F^t$ for some element $b \in \mathbb{F}_2^t$ and some subset T.

We should mention that functions with the ideal t-th differential uniformity have been discussed before in other contexts. For instance, in [27], [33], 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 known in [27] that, when $p = 2$ and $t > n/2$, perfect nonlinear function does not exist. Therefore, we obtain the following result.

Theorem 2. *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, i.e. $\Delta_F^t > 2^{n-t}$.*

Remark 1. It is well-known that EA and CCZ equivalence (see definitions in [7]) 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 [34].

We conclude this section by the following result. In the design of many block ciphers, the Sbox $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is of the form

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

where S_i are functions from $\mathbb{F}_{2^{n/m}}$ to itself. The following result gives the lower bound of differential uniformity of S.

Proposition 2. *Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a function defined by*

$$S(x) = \left(S_1(x_1), S_2(x_2), \ldots, S_{n/m}(x_{n/m})\right),$$

where $x_i \in \mathbb{F}_{2^m}$ and $x = (x_1, \ldots, x_{n/m})$. Then

$$\Delta_S \geq 2^{n/m-1} \max_{1 \leq i \leq n/m} \Delta_{S_i}.$$

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

3.2 Relation to truncated differential attack

Before we begin to study the properties of the t-th differential uniformity of F, we first show that the t-th differential uniformity of F is a general mathematical characterization of the *truncated differential attack*.

Given a block cipher $\mathcal{B}$ with block size n, message space $\mathcal{M}$, key space $\mathcal{K}$. Let $\mathrm{Enc}_K(\cdot)$ be the encrypt function of $\mathcal{B}$. The main idea behind the *truncated differential attack* is to find an input difference $\Delta_1 \in \mathbb{F}_{2^n}$ and a set of output difference with the form

$$S = \{\ (*, \cdots, i_1, *, \cdots, *, i_k, \cdots, *) \in \mathbb{F}_2^n\ \}, \tag{4}$$

where $i_i, \cdots, i_k \in \mathbb{F}_2$ are constants, and $*$ denotes the corresponding bit which may be either 0 or 1. Assume the following probability

$$\mathrm{Pr}_{m\in N}\left(\mathrm{Enc}_K(m) + \mathrm{Enc}_K(m+\varDelta_1) \in S\right) = p \gg 0,\ N \subseteq \mathcal{M}. \tag{5}$$

The attacker then guess a key K', to see whether the probability in (5) is around p. If this is true, there is a large probability that K' is the right key; otherwise it is not. The tuple $(\varDelta_1, S)$ is called a *differential* of $\mathcal{B}$.

It is clear that, for any key $K \in \mathcal{K}$, we may regard the encryption function $\mathrm{Enc}_K(\cdot)$ as a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$, denoted by F. Finding the differential $(\varDelta_1, S)$ is equivalent to finding the tuple $a \in \mathbb{F}_2^n, b \in \mathbb{F}_2^k$ such that the $\delta_F^T(a,b)$ is large for the function $F_T : \mathbb{F}_2^n \to \mathbb{F}_2^{n-k}$ defined by $F_T(x) = (f_{i_1}(x), \cdots, f_{i_k}(x))$, where $i_1, \cdots, i_k$ are the ones in (4).

4 *t*-th Differential uniformity of well-known S-boxes

In this section, we study the t-differential uniformity of the S-boxes used in some previously proposed ciphers.

4.1 *t*-th differential uniformity of Golden S-boxes

In [23], all S-boxes defined on $\mathbb{F}_{2^4}$ are classified up to EA and CCZ equivalence. These S-boxes are further studied in [34] and four of them are called *golden* S-boxes. In the following table, we compute their t-th differential uniformity and the corresponding differential distinguishers. Please refer to the Appendix for the full computation results of the 4-bit S-boxes used in previously proposed ciphers.

Table 1. t-th Differential Uniformities of Golden S-boxes

No.	$(D_S^1, U_{S,1})$	$(D_S^2, U_{S,2})$	$(D_S^3, U_{S,3})$	$(D_S^4, U_{S,4})$	Comment
1	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)	Golden Sbox
2	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)	Golden Sbox
3	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)	Golden Sbox
4	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)	Golden Sbox

From Table 1, we may see that, while the Sbox has optimal n-th differential uniformity, it may have bad t-th differential uniformity. For instance, the first differential uniformity of the No. 1 function is 12, while its 4-th differential uniformity attains the optimal value 4.

4.2 t-th differential uniformity of S-boxes in AES, ZUC and SNOW 3G

We will discuss the S-boxes endorsed in AES [31], ZUC [40] and SNOW 3G [41]. The following provides a brief introduction to their S-boxes.

1. ZUC [40]: The S-box S used in ZUC of the form $S = (S_0, S_1, S_0, S_1) : \mathbb{F}_{2^{32}} \to \mathbb{F}_{2^{32}}$, where S_1 is the inverse function x^{-1} from $\mathbb{F}_{2^8}$ to $\mathbb{F}_{2^8}$, and $S_0 : \mathbb{F}_{2^8} \to \mathbb{F}_{2^8}$ can be found in [40, Page 12].
2. SNOW 3G [41]: The S-box $S_Q : \mathbb{F}_{2^8} \to \mathbb{F}_{2^8}$ 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 [31]: The S-box 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 EA equivalent to the inverse function x^{-1}, the definition may be found in [31].

The following table gives the t-th differential uniformity of the functions S_0, S_Q and I.

Table 2. Differential uniformities of I, S_0, S_Q

t	$(D^t_{S_0}, U^t_{S_0})$	$(D^t_{S_Q}, U^t_{S_Q})$	(D^t_I, U^t_I)	Ideal Values
1	(192, 1/4)	(176, 3/16)	(144, 1/16)	(128, 0)
2	(160, 3/8)	(108, 11/64)	(88, 3/32)	(64, 0)
3	(128, 3/8)	(78, 23/128)	(52, 5/64)	(32, 0)
4	(96, 5/16)	(48, 1/8)	(30, 7/128)	(16, 0)
5	(64, 7/32)	(34, 13/128)	(18, 5/128)	(8, 0)
6	(32, 7/64)	(24, 5/64)	(10, 3/128)	(4, 0)
7	(24, 11/128)	(24, 11/128)	(10, 1/32)	(2, 0)
8	(8, 7/256)	(8, 7/256)	(4, 3/256)	(1, 0)

By the above experimental results, we have the following result.

Theorem 3. *Denote Γ_F by the highest differential distinguisher of F. Then*

$$\Gamma_{S_0} = 3/8,\ \Gamma_{S_Q} = 3/16, \Gamma_I = 3/32.$$

We take a closer look at the Sbox S_0 used in ZUC. By Table 2, we may see that the worst case is $\Delta^2_{S_0} = \Delta^3_{S_0} = 3/8 \gg 0$. We investigate this case in details. For $t = 3$, let a 3-subset be $\{i_1, i_2, i_3\}$. Then the following system of equations

$$\begin{aligned} y_0 &= f_{i_1}(x_0, \cdots, x_7) + f_{i_1}(x_0 + a_0, \cdots, x_7 + a_7) \\ y_2 &= f_{i_2}(x_0, \cdots, x_7) + f_{i_2}(x_0 + a_0, \cdots, x_7 + a_7) \\ y_3 &= f_{i_3}(x_0, \cdots, x_7) + f_{i_3}(x_0 + a_0, \cdots, x_7 + a_7) \end{aligned}$$

has 128 solutions for the following two cases of 3-subsets and the vector $\mathbf{a}$, listed in Table 3.

Table 3. 3rd Differential Uniformity of S_0

(y_0, y_1, y_2)	(i_1, i_2, i_3)	$\mathbf{a} = (a_0, \cdots, a_7)$
011	(0, 3, 6)	00000101
100	(5, 6, 7)	00000110

4.3 The first differential uniformity of the inverse function.

In general, it is difficult to determine the t-th differential uniformity for an (n, n)-function. In this following, for the inverse function, we determine its first differential uniformity. We first give a result on the first differential uniformity for any function F.

Theorem 4. *Let F be a function on $\mathbb{F}_{2^n}$ and $\{\beta_1, \cdots, \beta_n\}$ is a basis of $\mathbb{F}_{2^n}$ on $\mathbb{F}_2$. Then the first differential uniformity*

$$\Delta_F^1 = 2^{n-1} + \frac{1}{2} \max_{\substack{a\in\mathbb{F}_{2^n}^* \\ 1\le i\le n}} |\mathcal{F}(D_a F_i)|,$$

where $\mathcal{F}(D_a F_i) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{F_i(x+a)+F_i(x)}$ and $F_i(x) = \mathrm{Tr}(\beta_i F(x))$ for $1 \le i \le n$.

Proof. Letting $b \in \mathbb{F}_2$, we need to determine the number of solutions of the equation $b = F_i(x + a) + F_i(x)$. Assuming there are A_0 elements x such that $F_i(x+a)+F_i(x) = 0$ and A_1 elements such that $F_i(x + a) + F_i(x) = 1$. Then 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 $A_1 = 2^{n-1} - \frac{1}{2}\mathcal{F}(D_a F_i)$. The rest of the proof is followed by the definition of Δ_F^1. □

By Theorem 4 and [8, Theorem 1], we may get the first differential uniformity of the inverse function. The proof is simple and we omit it here. For convenience, define the number $k_n = \max\{k \equiv 0 \mod 4 | k < 2^{n/2+1} + 1\}$.

Corollary 1. *Let $I : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be the inverse function. Then the first differential uniformity*

$$\Delta_I^1 = \begin{cases} 2^{n-1} + 2^{n/2} & \text{if } n \text{ is even,} \\ 2^{n-1} + \frac{1}{2}(k_n \pm 4) & \text{if } n \text{ is odd.} \end{cases}$$

5 Building Distinguishers from the t-th Differential Uniformity

In this Section, we first introduce the method of building distuiguishers of a block cipher by the bias rate of the differential sequences, which are determined by t-th differential uniformity of the

S-box. As an application, we compute the bias rate of the differential sequences of the lightweight cipher LED-64. The computation results imply that there are many keys of LED-64 may generate a differential sequence that has large bias rate.

5.1 Differential sequences of a block cipher

Let $\mathcal{B}$ be a block cipher with block size $m = ns$. Assume its S-box $S : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$ is of the form $S = (S_0, \cdots, S_{s-1})$, where each S_i is a permutation defined on $\mathbb{F}_{2^n}$ for $0 \leq i \leq s-1$. Randomly choose N pairs of inputs, say p_{i1}/p_{i2} for $1 \leq i \leq N$, of $\mathcal{B}$ with a fixed nonzero difference Δ_a, i.e. $p_{i1} \oplus p_{i2} = \Delta_a \in \mathbb{F}_2^m$, and let c_{i1}, c_{i2} be the ciphertexts of p_{i1}, p_{i2} for some master key K, and denoting the difference of each c_{i1}, c_{i2} by $\mathbf{b}_i$, namely $\mathbf{b}_i = c_{i1} \oplus c_{i2} \in \mathbb{F}_2^m$ for $1 \leq i \leq N$,

$$\begin{aligned}
\mathbf{b}_1 &= c_{11} \oplus c_{12} = (b_{11}, b_{12}, \ldots, b_{1m}), \\
\mathbf{b}_2 &= c_{21} \oplus c_{22} = (b_{21}, b_{22}, \ldots, b_{2m}), \\
&\cdots \\
\mathbf{b}_N &= c_{N1} \oplus c_{N2} = (b_{N1}, b_{N2}, \ldots, b_{Nm}).
\end{aligned}$$

Now, instead of considering $\mathbf{b}_i$ as in truncated differential attack, for $1 \leq j \leq m$, we consider the sequence $\mathbf{d}_j$ defined by

$$\mathbf{d}_j = (b_{1j}, b_{2j}, \cdots, b_{Nj}),$$

where b_{ji} is the i-th element of the sequence $\mathbf{b}_j$. The sequence $\mathbf{d}_j$ is called the j-th *differential sequence* with respect to Δ_a.

Remark 2. One may understand the j-th differential sequence $\mathbf{d}_j$ as the following way. We may write the m sequences $\mathbf{b}_i$ above as an $N \times m$ matrix $\mathcal{M}$, with each row is b_i. The j-th differential sequence is then the j-th column of $\mathcal{M}$.

By Result 1, if the cipher $\mathcal{B}$ has good random property, the imbalance rate of the differential sequence $\mathbf{d}_j$ defined above should approximate to $\sqrt{1/2\pi N}$. In the following, we give an algorithm to test the randomness of a block cipher $\mathcal{B}$, by measuring it j-th differential sequence $\mathbf{d}_j$ with respect to certain input difference Δ_a, which is chosen by the t-differential defined in Section 3. In the following algorithm, assume the specifications of $\mathcal{B}$ is public, and its S-box S is of the form $(S_0, \cdots, S_0)$. We randomly choose K keys, N pairs of inputs and request their ciphertexts.

Algorithm 1 *(Computing the bias rate of differential sequence)*

Input: K keys, N pairs of inputs p_{i1}/p_{i2} for $1 \leq i \leq N$;

```
  Output: The number of keys which may generate the $j$-th differential sequence with much larger
  bias rate than the ideal value;

Step 1 For the function $S_0$, compute its $i$-differential $(i, D^i_{S_0}, U^i_{S_0}, A_i)$ (see definition in Section 3.2)
       for $1 \leq i \leq n$. Let $t$ be the the integer $i$ with the highest value of $D^i_{S_0}$;
Step 2 Choose an element $a \in A_t$ and let $\Delta_a = (a, 0, \cdots, 0) \in \mathbb{F}^s_{2^n}$;
Step 3 $i \leftarrow 1$;
Step 4 While $(i \leq K$ and $A \neq \emptyset)$
         Randomly choose a key and $N$ pairs of plaintexts with difference $\Delta_a = (a, 0, \cdots, 0) \in \mathbb{F}^s_{2^n}$;
         Request the ciphertexts of the $N$ pairs of plaintexts by the key chosen above, compute the
         imbalance rate of the $j$-th differential sequence (defined in Section 5.1) for $1 \leq j \leq m$ and
         save it as a sequence $\boldsymbol{IR}_i$;
         $i \leftarrow i + 1$;
Step 5 $S \leftarrow \{\}$;
Step 6 For $(1 \leq i \leq K)$
         If $\min(\boldsymbol{IR}_i) \gg 1/\sqrt{2N\pi}$ Then $S \leftarrow S \cup \{\boldsymbol{IR}_i\}$;
Step 7 If $S \neq \emptyset$ Then return $S$;
       Else if $A \neq \emptyset$ go to Step 2;
       Else return FAIL.
```

Remark 3. (1) Algorithm 1 aims to compare the randomness of a block cipher to a uniform cipher. If there are too many keys which may generate a j-th differential with much larger bias rate, the j-th differential sequence may be used as a distinguisher of this cipher. However, we should mention that, this distinguisher **cannot** yield an attack in general as the bias rate varies with the choice of the key.

(2) Obviously, the success probability of Algorithm 1 is influenced by the design of the cipher. In general, the good choice of the S-box and design diffusion layer will decrease the success probability to find a distinguisher by Algorithm 1.

5.2 The bias rate of the first differential sequence of LED-64

As an application of Algorithm 1, we test the randomness of LED-64, which is a lightweight cipher presented at CHES 2011 in [19], see Fig. 1. Its Sbox $S = (S_0, \cdots, S_0) : \mathbb{F}_{2^{32}} \rightarrow \mathbb{F}_{2^{32}}$, where S_0 is a permutation on $\mathbb{F}_{2^4}$ defined by

$$[C, 5, 6, B, 9, 0, A, D, 3, E, F, 8, 4, 7, 1, 2]$$

and the field $\mathbb{F}_{2^4}$ is generate by the primitive polynomial $x^4 + x + 1$. Please refer to [19] for the detailed specifications.

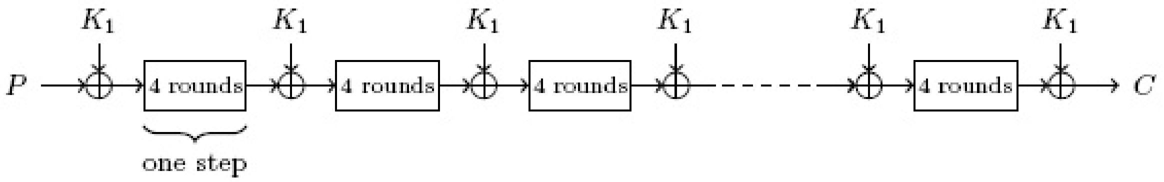

Fig. 1. Block diagram of LED-64 [19]

First, we compute the t-differential for the function S_0. For $t = 3, 4$, we do not list the set A_t but only give its size since we do not need them in the following computations.

Table 4. t-differential of the LED Sbox

t	$(D^t_{S_0}, U^t_{S_0})$	A_t
1	$(16, 1/2)$	$\{[1,0,0,0],[1,1,1,1]\})$
2	$(12, 1/2)$	$\{[0,0,0,1],[1,0,0,1]\})$
3	$(8, 3/8)$	$\|A\| = 4$
4	$(4, 3/16)$	$\|A\| = 14$

Now, letting the input difference of the cipher LED be

$$\Delta_a = [1, 0, 0, 0, 0, 0, 0, 0, \cdots, 0]$$

and we randomly pick up $N = 2^{15}$ pairs of plaintexts with the difference Δ_a. Repeating the experiment described in Algorithm 1 for $K = 4,000$ randomly chosen keys and compute the bias rate of the first differential sequence with respect to Δ_a. Note that, by Theorem 4, for $N = 2^{15}$, the ideal value of the bias rate is 0.002210. In the following table, we use BR and ID to denote the bias rate of the first differential sequence, and the ideal bias rate respectively. The value $|BR - ID|$ denotes the absolute value between the difference of actual bias rate and ideal value. For any two numbers a, b, the notation $[a, b)$ denotes the set of numbers c with $a \leq c < b$.

Table 5. The bias rate of the first differential sequence of LED-64 with respect to Δ_a

$\|BR - ID\|$	# of keys
$[0, 0.001)$	853
$[0.001, 0.01)$	2940
$[0.01, 0.021]$	207

From the above table, we may see that, for many keys, the difference of the bias rate of the first differential sequence is large, which shows that some keys are weak.

6 Cryptographic Properties of the Compositions of Vectorial Boolean Functions

In the design of many stream ciphers, for instance in SNOW 3G [41], ZUC [40] etc., there is one structure consisting of the composition of a linear function $\mathbb{F}_{2^k} \rightarrow \mathbb{F}_{2^n}$ and an S-box $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$. The purpose of this design is to bring the confusion and diffusion into the cipher. However, as we will see below, the nonlinearity of the composition will be greatly decreased, which makes the cipher weaker against linear attack. We first introduce some definitions.

Let f and g be two (n, m)-functions and $a(x), b(x)$ be two (k, n)-functions. Let $F = f \circ a$ and $G = g \circ b$ be the compositions. The following figure shows the relationship among these mappings.

$$F(x) = f \circ a(x) = f(a_0(x), \cdots, a_{n-1}(x)) \text{ and } G(x) = g \circ b(x) = g(b_0(x), \cdots, b_{n-1}(x)) \tag{6}$$

$$\begin{tikzcd}
\mathbb{F}_2^k \arrow[d, "a(x)"] & \\
\mathbb{F}_2^n \arrow[d, "f(x)"] & \Longrightarrow F = f \circ a : \mathbb{F}_2^k \to \mathbb{F}_2^m \\
\mathbb{F}_2^m &
\end{tikzcd}$$

$$\begin{tikzcd}
\mathbb{F}_2^k \arrow[d, "b(x)"] & \\
\mathbb{F}_2^n \arrow[d, "g(x)"] & \Longrightarrow G = g \circ b : \mathbb{F}_2^k \to \mathbb{F}_2^m \\
\mathbb{F}_2^m &
\end{tikzcd}$$

Fig. 2. Composition of linear function and S-box

Note that the nonlinearity of the composition of an S-box and a linear function is also considered in [29], however, in the following, we will first give a more general result on the crosscorrelation of F and G (defined below) for certain functions a, b, and then, as an application, the nonlinearity of the composition of an S-box and a linear function is obtained. In the following, we consider the nonlinearity and differential distinguisher (see Definition 1) of the functions F and G.

6.1 Nonlinearity and its diversity

A commonly used method to characterize the nonlinearity of an (n, n)-function F is as follows. The *Walsh (Fourier) transform* $F^{\mathcal{W}} : \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \to \mathbb{C}$ of F is defined by:

$$F^{\mathcal{W}}(\lambda, \eta) := \sum_{x \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(\lambda x + \eta F(x))}.$$

The multiset $\mathcal{W}_F := \{F^{\mathcal{W}}(\lambda, \eta) : \lambda \in \mathbb{F}_{2^n}, \eta \in \mathbb{F}_{2^n}^*\}$ is called the *Walsh spectrum* of F. The *nonlinearity* $\mathrm{NL}(F)$ of F is defined by

$$\mathrm{NL}(F) \triangleq 2^{n-1} - \frac{1}{2} \max_{x \in \mathcal{W}_F} |x|,$$

and the *normalized nonlinearity* of F is defined by

$$\rho_F = \mathrm{NL}_F / 2^n.$$

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 $2^{n-1} - 2^{\frac{n}{2}}$.

Similarly, for a Boolean function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, we may define its Walsh transform by

$$f^{\mathcal{W}}(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) + \mathrm{Tr}(\lambda x)}.$$

For simplicity, we denote the Walsh coefficients $F^{\mathcal{W}}(\lambda, \eta), f^{\mathcal{W}}(\lambda)$ by $\widehat{F}(\lambda, \eta), \widehat{f}(\lambda)$ respectively.

Note that for a random vectorial boolean function G, the average value of its Walsh spectrum of G is given by $\sqrt{2^n}$. Indeed, by the Parseval Equation (2), we have $2^{3n} = \sum_{\lambda, \eta \in \mathbb{F}_{2^n}} \widehat{G}^2(\lambda, \eta)$ and then $\mathrm{ave}(\widehat{G}^2(\lambda, \eta)) \geq 2^{3n}/2^{2n} = 2^n$ and hence $\mathrm{ave}(|\widehat{G}(\lambda, \eta)|) \geq 2^{n/2}$. We introduce the following concept, which is adapted from communication theory.

Definition 2. *For an (n, n)-function F, the diversity of nonlinearity of F is defined as the ratio of the average value of Walsh spectrum, $\sqrt{2^n}$, to the maximal value of the Walsh spectrum of F:*

$$Div_F = \left| \frac{\sqrt{2^n}}{2^n} - \frac{\max\{\mathcal{W}_F\}}{2^n} \right|.$$

We may further write it as

$$Div_F = \left| \frac{\sqrt{2^n} - \max\{\mathcal{W}_F\}}{2^n} \right| = \left| \frac{1 - c}{\sqrt{2^n}} \right|,$$

where $\max\{\mathcal{W}_F\} = \sqrt{2^n} c$ *and c is a constant.*

Note that $1 \leq c < \sqrt{2^n}$ as $\max\{\mathcal{W}_F\} < 2^n$. Hence, the diversity of nonlinearity of F is to converge to 0 if a vectorial boolean has a good nonlinearity. Specially, if F is a linear function, we can easily see that $\mathrm{Div}_F = \frac{1}{\sqrt{2^n}}$. Therefore, to obtain a nonlinear function, we need Div_F as **small** as possible. We assume that the functions considered in the paper have zero constant term.

The *cross-correlation functions* between two functions F and G are defined as

$$C_{F,G}(\lambda, \mathbf{u}, \mathbf{v}) = \sum_{x \in \mathbb{F}_{2^k}} (-1)^{\mathbf{u}\cdot F(x) + \mathbf{v}\cdot G(\lambda x)}, \lambda \in \mathbb{F}_{2^k}$$

where $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^m$. If $m = 1$ and G is linear, then the cross-correlation between F and G becomes the Walsh transform of F.

6.2 Nonlinearity diversity of the composition

Now we consider the nonlinearity diversity of the function $F = f \circ a$. For the later use, we introduce two definitions below. For simplicity, in the following, we denote $Q = 2^k$, $q = 2^n$, $\mathbb{F}_{2^k} = \{x_0, \cdots, x_{Q-1}\}$ and $\mathbb{F}_{2^n} = \{z_0, \cdots, z_{q-1}\}$.

For a given $\lambda \in \mathbb{F}_{2^k}$, we define the following $2^k \times 2n$ array:

$$M_{a,b}(\lambda) = \begin{pmatrix} a(x_0) & b(\lambda x_0) \\ a(x_1) & b(\lambda x_1) \\ \vdots & \\ a(x_{Q-1}) & b(\lambda x_{Q-1}) \end{pmatrix} \tag{7}$$

which is called an *image array* of a and b at λ.

Definition 3. *(Simple image set) For $k \geq 2n$, $a(x)$ and $b(x)$ are said to have simple image set if they satisfy the following conditions:*

Case 1: For $\lambda \in \mathbb{F}_Q$, if for any η, there exists x such that $b(\lambda x) \neq \eta b(x)$, then the image array of a and b at λ defined in (7) is a $(2^k, 2n)$ orthogonal array, i.e. each $2n$-bit vector occurs 2^{k-2n} times in the image array;

Case 2: For $\lambda \in \mathbb{F}_Q$, if there exists some $\eta \in \mathbb{F}_q$ such that $b(\lambda x) = \eta b(x)$ for all $x \in \mathbb{F}_{2^k}$, then there exists some permutation function $h : \mathbb{F}_q \to \mathbb{F}_q$ such that any pair $(h(y), \eta y) \in \mathbb{F}_Q^2$ occurs 2^{k-n} times in the image array of a and b, which now becomes

$$M_{a,b}(\lambda) = \begin{pmatrix} a(x_0) & \eta b(x_0) \\ a(x_1) & \eta b(x_1) \\ \vdots & \\ a(x_{Q-1}) & \eta b(x_{Q-1}) \end{pmatrix}. \tag{8}$$

Under this definition, we get a similar result as the one in [20].

Theorem 5. *With the above notations, if $a(x)$ and $b(x)$ have the simple image set and a, b are surjections, then the cross-correlation between F and G is given by*

$$C_{F,G}(\lambda, \mathbf{u}, \mathbf{v}) = \begin{cases} 2^{k-2n}\widehat{\mathbf{u}\cdot f}(0)\widehat{\mathbf{v}\cdot g}(0), \text{ if } \lambda \text{ belongs to Case 1,} \\ 2^{k-n}C_{(\mathbf{u}\cdot f)\circ h,(\mathbf{v}\cdot g)}(\eta), \text{ if } \lambda \text{ belongs to Case 2 and } h, \eta \text{ are in Definition 3.} \end{cases}$$

Proof. We first prove the case λ belongs to Case 1. Now

$$\begin{aligned} C_{F,G}(\lambda, \mathbf{u}, \mathbf{v}) &= \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathbf{u}\cdot F(x)+\mathbf{v}\cdot G(\lambda x)} = \sum_{x\in\mathbb{F}_{2^k}} (-1)^{\mathbf{u}\cdot f(a(x))+\mathbf{v}\cdot g(b(\lambda x))} \\ &= 2^{k-2n} \sum_{y_1,y_2\in\mathbb{F}_{2^n}} (-1)^{\mathbf{u}\cdot f(y_1)+\mathbf{v}\cdot g(y_2)}, \quad y_1 = a(x), y_2 = b(\lambda x) \\ &= 2^{k-2n} \sum_{y_1\in\mathbb{F}_{2^n}} (-1)^{\mathbf{u}\cdot f(y_1)} \sum_{y_2\in\mathbb{F}_{2^n}} (-1)^{\mathbf{v}\cdot g(y_2)} \\ &= 2^{k-2n}\widehat{\mathbf{u}\cdot f}(0)\widehat{\mathbf{v}\cdot g}(0). \end{aligned}$$

The case λ belongs to Case 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 6. *If the functions G, a, b are all linear and a, b are surjections, the following assertions hold.*

1. *The image array of a and b is a simple image array, and the Walsh transform of F is given by*

$$C_{F,F}(\lambda, \mathbf{u}, \mathbf{v}) = \begin{cases} 0, & \text{if } \lambda \text{ belongs to Case 1 in Definition 2,} \\ 2^{k-n}\hat{f}(\eta), & \text{if } \lambda \text{ belongs to Case 2 in Definition 2.} \end{cases}$$

2. *The nonlinearity of F is given by*

$$NL_F = 2^{k-1} - 2^{k-n-1}NL_f, \text{ or equivalently, } \max\{\mathcal{W}_F\} = 2^{k-n}\max\{\mathcal{W}_f\}.$$

3. *The diversity of nonlinearity is*

$$Div_F = \left| \frac{1 - 2^{k/2-n}\max\{\mathcal{W}_f\}}{\sqrt{2^k}} \right|.$$

4. *The t-th differential uniformity of F and f is related by $\Delta_F^t = 2^{k-n}\Delta_f^t$, and hence $U_F^t = U_f^t$.*

Proof. (1) Note that for any linear function L from $\mathbb{F}_{2^k}$ to $\mathbb{F}_{2^n}$, there exists an element $\gamma \in \mathbb{F}_{2^k}$ such that $L(x) = \mathrm{Tr}_n^k(\gamma x)$. For each $y \in \mathbb{F}_{2^n}$, the set $H_y = \{x \in \mathbb{F}_{2^k} | L(x) = y\}$ is an affine hyperplane, i.e. $H_y = a + \mathcal{L}$, where $a \in \mathbb{F}_{2^k}$ and $\mathcal{L}$ is a subspace of $\mathbb{F}_{2^k}$ over $\mathbb{F}_{2^n}$ with dimension $k/n - 1$.

Now assume that $a(x) = \mathrm{Tr}_n^k(\gamma_1 x)$ and $b(x) = \mathrm{Tr}_m^n(\gamma_2 x)$. If for any $\eta \in \mathbb{F}_q$, there exists x such that $b(\lambda x) \neq \eta b(x)$, we have

$$(a(x), b(\lambda x)) = \left(\mathrm{Tr}_n^k(\gamma_1 x), \mathrm{Tr}_n^k(\lambda\gamma_2 x)\right).$$

For each $(y_1, y_2) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$, by[20, Lemma 2], there are 2^{k-2n} elements $x \in \mathbb{F}_{2^k}$ such that $\big(\mathrm{Tr}_n^k(\gamma_1 x), \mathrm{Tr}_n^k(\lambda\gamma_2 x)\big) = (y_1, y_2)$, which implies (7) is a $(2^k, 2n)$ orthogonal array with every element in $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ appears 2^{k-2n} times. If $b(\lambda x) = \eta b(x)$ for some $\eta \in \mathbb{F}_q$, by [20, Lemma 1], there are 2^{k-n} elements $x \in \mathbb{F}_{2^k}$ such that $\big(\mathrm{Tr}_n^k(\gamma_1 x), \mathrm{Tr}_n^k(\lambda\gamma_2 x)\big) = (y_1, y_2)$, where $y_2 \in \mathbb{F}_{2^n}$ and $y_1 = \lambda^{-1}h(y_2)$ for some permutation h of $\mathbb{F}_q$. Therefore, we show that the image set of a, b is simple. The value of $CF, F(\lambda, \mathbf{u}, \mathbf{v})$. can be obtained by Theorem 5.

The results in (2) and (3) follows from the definition.

(4) For each t with $1 \le t \le n$, we need to consider the maximal number of solutions of the equation

$$\begin{cases} b_1 = & F_{i_1}(x+s) + F_{i_1}(x), \\ \quad\cdots & \\ b_t = & F_{i_t}(x+s) + F_{i_t}(x), \end{cases} \tag{9}$$

where $T = \{i_1, \cdots, i_t\}$ runs through Ω_t and $(b_1, \cdots, b_t)$ runs through $\mathbb{F}_2^t$. Recall that there exists a basis $\{\beta_1, \cdots, \beta_n\}$ such that $F_i = \mathrm{Tr}(\beta_i F(x))$ for $1 \le i \le n$. Now, we have $F_{i_j}(x+c) + F_{i_j}(x) = \mathrm{Tr}(\beta_{i_j} F(x+c)) + \mathrm{Tr}(\beta_{i_j} F(x)) = \mathrm{Tr}(\beta_{i_j}(f(a(x+c)) + f(a(x)))) = f_{i_j}(a(x)+a(c)) + f_{i_j}(a(x))$. Letting $y = a(x)$, the equation (9) becomes

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

The above equation has at most Δ_f^t solutions y. Note that $\sharp\{x : x \in \mathbb{F}_{2^k} | a(x) = y\} = 2^{k-n}$ for any $y \in \mathbb{F}_{2^n}$, then 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 equation (9) has at most $\Delta_F^t = 2^{k-n}\Delta_f^t$ solutions. □

Remark 4. For $k \ge 2n$, since $\max\{\mathcal{W}(f)\} \ge \sqrt{2^n}$, we see that $Div_F \gg 0$. This shows that the nonlinearity of the composition of a linear function and an Sbox is very poor, which makes the cipher weaker against linear attack.

6.3 Nonlinearity diversity and differential distinguisher of S_0, S_Q, I

The following table lists the nonlinearity diversity of the S-boxes S_0, S_Q, I.

Table 6. Nonlinearity diversity of S_0, S_Q, I

Function	$\max \mathcal{W}_F$	Div_F
S_0	64	0.1875
S_Q	64	0.1875
I	32	0.0625

It is well-known that when n is even (resp. odd), the nonlinearity of the inverse function $I(x) = x^{-1}$ attains the known maximal value $2^{n-1} - 2^{n/2}$ (resp. $2^{n-1} - 2^{(n-1)/2}$). Equivalently, $\max(\mathcal{W}_I) = 2^{n/2+1}$ (resp. $2^{(n+1)/2}$). Therefore, its nonlinearity diversity $Div_I(x) = \frac{1}{\sqrt{2^n}}$ (resp. $\frac{\sqrt{2}-1}{\sqrt{2^n}}$).

In the design of ZUC, the S-boxes S_0, S_1 are composed with two linear function $L_1, L_2 : \mathbb{F}_{2^{32}} \to \mathbb{F}_{2^8}$. One may refer to the figure at the beginning of Section 3 by letting $k = 32$ and $n = m = 8$. By Table 6 and Theorem 6, we may determine the maximal differential distinguisher and nonlinearity diversity of $S_0 \circ L_i$ for $i = 1, 2$.

Theorem 7. *For $i = 1, 2$,*

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

Proof. The result can be obtained from Theorem 6. □

Theorem 7 shows that $S_0 \circ L_i$ has both poor 3rd differential uniformity and nonlinearity diversity. Furthermore, by comparing $Div_{S_0 \circ L_i} = 0.25$ with $Div_{S_0} = 0.1875$ in Table 6 and Theorem 6, the nonlinearity diversity is amplified. Therefore, in the design of ciphers, the structure of the composition of the linear functions and S-boxes makes the cipher weaker against linear attack.

7 Relationship between t-order differential uniformity and nonlinearity

Finally, we present a relationship between the t-th differential uniformity and its nonlinearity by the Parseval relation. For the convenience of the readers, we include a short proof of the following result. For any integer t with $1 \le t \le n$ and any $T \in \Omega_t$, let the function F_T be the same as the one in Section 2. Recall that, for a group G, its character group is denoted by $\widetilde{G}$.

Theorem 8. *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*

$$\frac{1}{2^{n+t}} \sum_{(a,b)\in\mathbb{F}_{2^n}\times\mathbb{F}_{2^t}} \left|\widehat{F_T}(a,b)\right|^4 = 2^{2n} + \sum_{(a,b)\in\mathbb{F}^*_{2^n}\times\mathbb{F}_{2^t}} \delta^T_F(a,b)^2.$$

Proof. Denoting by $D_{F_T} = \{(x, F_T(x)) : x \in \mathbb{F}_{2^n}\}$. We may regard D_{F_T} as an element of the group ring $\mathbb{C}[G]$, where $G = \mathbb{F}_{2^n} \times \mathbb{F}_{2^t}$. By abusing the notations, we write $D_{F_T} = \sum\limits_{x\in\mathbb{F}_{2^n}} (x, F_T(x))$. Now we have

$$\begin{aligned} D := D_{F_T} D_{F_T}^{(-1)} &= \sum_{x,y\in\mathbb{F}_{2^n}} (x+y, F_T(x) + F_T(y)) \\ &= \sum_{a,x\in\mathbb{F}_{2^n}} (a, F_T(x+a) + F_T(x)) = \sum_{a,b\in\mathbb{F}_{2^n}} \delta(a,b)\,(a,b) \\ &= 2^n + \sum_{\substack{(a,b)\in\mathbb{F}_{2^n}\times\mathbb{F}_{2^t} \\ a\neq 0}} \delta(a,b)\,(a,b)\,. \end{aligned}$$

By the Parseval's equation (2), we get the following

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

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 $\chi((x,y)) = \chi_a(x)\chi_a(y) = (-1)^{\mathrm{Tr}_1^n(ax)+\mathrm{Tr}_1^m(by)}$. Therefore, it is not difficult to see that $\chi(D_{F_T}) = \widehat{F}(a,b)$ for some $(a,b)\in\mathbb{F}_{2^n}\times\mathbb{F}_{2^t}$. We finish the proof. □

To the best of our knowledge, there is no fast method to determine the differential uniformity of a function F. However, by the butterfly algorithm ([7]), the Walsh coefficients $\widehat{F_T}(a,b)$ can be determined efficiently. Therefore, by Theorem 8, we give a lower bound for the t-th differential uniformity. Denote by

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

Corollary 2. *Let the notations be the same as above.*

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

Furthermore, if the function $F:\mathbb{F}_{2^n}\to\mathbb{F}_{2^n}$ is a differentially 4-uniform function, i.e the value $\delta_F(a,b)\in\{0,2,4\}$ for any $(a,b)\in\mathbb{F}_{2^n}^*\times\mathbb{F}_{2^n}$, we may get the differential spectrum from its Walsh spectrum.

Theorem 9. *Let $n \geq 2$ be an integer and F be a differentially 4-uniform function on $\mathbb{F}_{2^n}$. Denoting by*

$$A = \sum_{(a,b)\in\mathbb{F}_{2^n}\times\mathbb{F}_{2^t}} \left| \widehat{F_T}(a,b) \right|^4.$$

Then the differential spectrum of F is

$$\begin{aligned}
&0 \text{ appears } 9\cdot 2^{2n-2} - 2^{n+1} - 5\cdot 2^{n-3} - \frac{3}{2^{n+3}}A,\\
&2 \text{ appears } 5\cdot 2^{2n-2} - 2^{n} - \frac{1}{2^{n+2}}A,\\
&4 \text{ appears } 2^{n-2} - 3\cdot 2^{2n-3} + \frac{1}{2^{n+3}}A.
\end{aligned}$$

Proof. Let $\mathcal{G}_F = \{(x,F(x)) : x\in\mathbb{F}_{2^n}\}$ be the graph of F. Since F is a differentially 4-uniform function, clearly we have

$$\mathcal{G}_F \cdot \mathcal{G}_F^{(-1)} = \sum_{\substack{a,b\in\mathbb{F}_{2^n},\\ a\neq 0}} \delta_F(a,b)(a,b) = 2^n + 2D_1 + 4D_2, \tag{10}$$

where D_1, D_2 are two subsets of $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$. Note that $|D_1|, |D_2|$ are the multiplicities of 2 and 4. By the Parseval's equation, we obtain the identity

$$2^{2n} + 4|D_1| + 16|D_2| = \frac{1}{2^n} \sum_{\chi \in \widehat{G}} \chi(\mathcal{G}_F)^4. \tag{11}$$

Applying the principal character on (10) we get

$$2^n + 2|D_1| + 4|D_2| = 2^{2n}. \tag{12}$$

Denoting $|D_1| = x_1, |D_2| = x_2$ and $A = \sum\limits_{\chi \in \widehat{G}} \chi(\mathcal{G}_F)^4$. Solving the equations (11) and (12) we have

$$\begin{aligned} x_1 &= 5 \cdot 2^{2n-2} - 2^n - \frac{1}{2^{n+2}} A, \\ x_2 &= 2^{n-2} - 3 \cdot 2^{2n-3} + \frac{1}{2^{n+3}} A. \end{aligned}$$

Now it is easy to the multiplicity of 0 is that $2^n(2^n-1)-x-y$, which is $9 \cdot 2^{2n-2} - 2^{n+1} - 5 \cdot 2^{n-3} - \frac{3}{2^{n+3}} A$. We finish the proof. □

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. E. Biham and A. Shamir, Differential cryptanalysis of DES-like cryptosystems, Journal of Cryptology: 4(1): 3–72, (1991).
3. 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. Pail- lier and I. Verbauwhede (Eds.): CHES 2007, LNCS 4727, 450–466, (2007).
4. 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).
5. 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).
6. C. De Cannire, H. Sato, and D. Watanabe: Hash Function Luffa - Specification Ver. 2.0.1, NIST SHA-3 Submission, Round 2 document, (2009).
7. C. Carlet, Vectorial Boolean Functions for Cryptography, Idem, 398–469, (2010).
8. 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).
9. 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 177C194. Springer-Verlag, (2003).
10. J. Daemen, M. Peeters, G. Van Assche, and V. Rijmen: Nessie Proposal: NOEKEON: NESSIE Proposal, 27 October 2000.
11. J.F. Dillon, APN polynomials: an update, In Conference Finite Fields and Applications Fq9, Dublin, Ireland (2009).

12. M. Duan, X. Lai, Improved zero-sum distinguisher for full round Keccak-f permutation, IACR ePrint Archive 2011: 23 (2011).
13. V. Dolmatov, GOST 28147-89: Encryption, Decryption, and Message Authentication Code (MAC) Algorithms, Internet Engineering Task Force RFC 5830, March 2010.
14. 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).
15. 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).
16. 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).
17. S. Golomb and G. Gong, Signal deisgn for good correlation, *Cambridge University Press*, (2005).
18. 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).
19. 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).
20. A. Klapper, A. H. Chan and M. Goresky, Cross-correlations of linearly and quadratically related geometric sequences and GMW sequences, Discrete Appllied Mathematics: 46, 1–20, (1993).
21. O. Kucuk: The Hash Function Hamsi, NIST SHA-3 Submission, Round 2 document, 14 September 2009.
22. L. Knudsen, Truncated and Higher Order Differentials, FSE'94, 196211, (1995).
23. G. Leander, A. Poschmann, On the classification of 4 bit S-boxes, C. Carlet and B. Sunar (Eds.): WAIFI 2007, LNCS 4547, 159-176, (2007).
24. R. Lidl and H. Niederreiter, Finite Fields, Cambridge University Press, (1997).
25. 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).
26. M. Matsui, Linear cryptanalysis method for DES cipher, Advances in Cryptology - EUROCRYPT 1993.
27. K. Nyberg, Perfect Nonlinear S-Boxes. EUROCRYPT 1991: 378-386.
28. K. Nyberg, L. R. Knudsen: Provable Security Against Differential Cryptanalysis. CRYPTO 1992: 566-574
29. K. Nyberg, S-boxes and round functions with controlled linearity and differential uniformity, FSE 94, LNCS 1008, 111-130, (1995).
30. NIST: Data Encryption Standard, FIPS PUB 46. National Bureau of Standards, U.S. Department of Commerce, Washington D.C., 15 January 1977.
31. NIST, The Advanced Encryption Standard, FIPS 197, "csrc.nist.gov/publications/fips/fips197/fips-197.pdf".
32. D.S. Passman, The Algebraic Structure of Group Rings, Wiley-Interscience, New York, (1977).
33. A. Pott, Nonlinear functions in abelian groups and relative difference sets, Discrete Applied Mathematics: 138(1-2): 177–193, (2004).
34. M. O. Saarinen, Cryptographic Analysis of All 4×4 S-Boxes, In A. Miri and S. Vaudenay (Eds.) SAC 2011, LNCS 7118, 118-133, (2011).
35. B. Schmidt, On (p^a, p^b, p^a, p^{a-b})-relative difference set, Journal of Algebraic Combinatorics: 6, 279–297, (1997).
36. A. Sorkin, Lucifer: A cryptographic algorithm, Cryptologia, Vol. 8, No. 1, pp. 22-32. (1984).
37. 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).

38. J. V. Uspensky, Introduction to Mathematical Probability, New York McGraw-Hill, (1937).

39. H. Wu: The Hash Function JH, NIST SHA-3 Submission, Round 3 document, 16 January, 2011.

40. 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".

41. NIST, Specification of the 3GPP Confidentiality and Integrity Algorithms UEA2 & UIA2. Document 2: SNOW 3G Specification, "www.gsma.com/technicalprojects/wp-content/uea2uia2".

Appendix

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

Table 7: t-th Differential Uniformities of 4-bit S-boxes

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 S0 [36]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(6, 5/16)
Lucifer S1 [36]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(6, 5/16)
Present [3]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Present^{-1} [3]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
JH S0 [39]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
JH S1 [39]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(4, 3/16)
ICEBERG0 [37]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
ICEBERG1 [37]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
LUFFA [6]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
NOEKEON [10]	(16, 1/2)	(16, 3/4)	(8, 3/8)	(4, 3/16)
HAMSI [21]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
HB1 S0 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1 S1 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1 S2 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1 S3 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1-1 S0 [14]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
HB1-1 S1 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1-1 S2 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB1-1 S3 [14]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2 S0 [15]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2 S1 [15]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2 S2 [15]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)

HB2 S3 [15]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S0 [15]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S1 [15]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S2 [15]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
HB2-1 S3 [15]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
DES S0-0 [30]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
DES S0-1 [30]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(8, 7/16)
DES S0-2 [30]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
DES S0-3 [30]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S1-0 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S1-1 [30]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(8, 7/16)
DES S1-2 [30]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S1-3 [30]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S2-0 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S2-1 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S2-2 [30]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
DES S2-3 [30]	(16, 1/2)	(16, 3/4)	(8, 3/8)	(8, 7/16)
DES S3-0 [30]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S3-1 [30]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S3-2 [30]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S3-3 [30]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S4-0 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S4-1 [30]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(6, 5/16)
DES S4-2 [30]	(16, 1/2)	(10, 3/8)	(8, 3/8)	(6, 5/16)
DES S4-3 [30]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(6, 5/16)
DES S5-0 [30]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
DES S5-1 [30]	(12, 1/4)	(10, 3/8)	(10, 1/2)	(8, 7/16)
DES S5-2 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S5-3 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S6-0 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S6-1 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(8, 7/16)
DES S6-2 [30]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
DES S6-3 [30]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)

DES S7-0 [30]	(16, 1/2)	(12, 1/2)	(12, 5/8)	(6, 5/16)
DES S7-1 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(10, 9/16)
DES S7-2 [30]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
DES S7-3 [30]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(8, 7/16)
Serpent S0 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
Serpent S1 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent S2 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
Serpent S3 [1]	(12, 1/4)	(12, 1/2)	(8, 3/8)	(8, 7/16)
Serpent S4 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(8, 7/16)
Serpent S5 [1]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
Serpent S6 [1]	(16, 1/2)	(12, 1/2)	(10, 1/2)	(6, 5/16)
Serpent S7 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(8, 7/16)
Serpent-1 S0 [1]	(16, 1/2)	(16, 3/4)	(12, 5/8)	(8, 7/16)
Serpent-1 S1 [1]	(12, 1/4)	(10, 3/8)	(6, 1/4)	(6, 5/16)
Serpent-1 S2 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S3 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S4 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S5 [1]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
Serpent-1 S6 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
Serpent-1 S7 [1]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K1 [13]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K2 [13]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
GOST K3 [13]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K4 [13]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K5 [13]	(16, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
GOST K6 [13]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
GOST K7 [13]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
GOST K8 [13]	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)

Table 8. t-th Differential Uniformities of All 4-bit S-boxes in [23]

No.	$(D_S^1, U_{S,1})$	$(D_S^2, U_{S,2})$	$(D_S^3, U_{S,3})$	$(D_S^4, U_{S,4})$
1	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
2	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
3	(**16**, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
4	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
5	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
6	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
7	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
8	(12, 1/4)	(12, 1/2)	(8, 3/8)	(4, 3/16)
9	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
10	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
11	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
12	(12, 1/4)	(10, 3/8)	(6, 1/4)	(4, 3/16)
13	(12, 1/4)	(10, 3/8)	(6, 1/4)	(4, 3/16)
14	(**16**, 1/2)	(12, 1/2)	(8, 3/8)	(4, 3/16)
15	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)
16	(12, 1/4)	(10, 3/8)	(8, 3/8)	(4, 3/16)