OPTIMAL S-BOXES AGAINST ALTERNATIVE OPERATIONS

MARCO CALDERINI, ROBERTO CIVINO, AND RICCARDO INVERNIZZI

Abstract. Civino et al. have characterised diffusion layers that expose an SPN to vulnerability from differential cryptanalysis when employing alternative operations coming from groups isomorphic to the translation group on the message space. In this study, we present a classification of diffusion layers that exhibit linearity in *parallel* alternative operations for ciphers with 4-bit s-boxes, enabling the possibility of an alternative differential attack simultaneously targeting all the s-boxes within the block. Furthermore, we investigate the differential behaviour with respect to alternative operations for all classes of optimal 4-bit s-boxes, as defined by Leander and Poschmann (2007). Our examination reveals that certain classes contain weak permutations w.r.t. alternative differential attacks, and we leverage these vulnerabilities to execute a series of experiments.

1. Introduction and preliminaries

Differential cryptanalysis, originally introduced by Biham and Shamir in the late 1980s [BS91] and subsequently generalised [Wag99, Knu95, BCJW02, BBS05], has become one of the cornerstones for evaluating the robustness of various symmetric primitives. The fundamental premise of differential cryptanalysis is that analysing the differences (differentials) between pairs of plaintexts and the corresponding ciphertexts can unveil undesired biases. While differentials can be calculated with respect to any difference operator, regardless of which operation is responsible for performing the sum with the round key during encryption, it is usual for the two operations to coincide. For this reason, classical differential cryptanalysis of a cipher in which the key is xor-ed to the state is typically performed by studying the distribution of xor-differentials, whose propagation is traditionally prevented by the combined action of the linear diffusion layer and the s-box layer. In particular, s-boxes are pivotal for ensuring the security of almost all contemporary block ciphers, serving as the primary non-linear component within the cipher, particularly in the case of SPNs. Equally relevant, the efficiency of a cipher is significantly influenced by the size of the s-boxes. In practical scenarios, s-boxes typically have a size of 4 or 8 bits, with 4 being the most popular choice for ciphers designed to operate on power-constrained devices [BAK98, BKL+07, SIH+11, BBI+15]. It is clear that the selection of appropriate s-boxes is critical to fortify the cipher against various types of attacks. In this sense, Leander and Poschmann have classified 4-bit s-boxes which are optimal w.r.t. standard criteria that guarantee poor propagation of xor-differentials [LP07].

A recent line of research is focused on the study of alternative difference operators for the differential cryptanalysis of xor-based ciphers [CBS19, CCS21, Teş22, CCI24]. These new operators are designed to induce a novel operation with respect to which differentials are computed. Within this approach, a large class of possible alternative operations has been studied, all of

2010 *Mathematics Subject Classification.* 20B35, 94A60, 68P25.

Key words and phrases. Differential cryptanalysis; alternative operations; 4-bit s-boxes.

M. Calderini and R. Civino are members of INdAM-GNSAGA (Italy). R. Civino is partially funded by the Centre of excellence ExEMERGE at University of L'Aquila. R. Invernizzi is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement ISOCRYPT - No. 101020788) and by CyberSecurity Research Flanders with reference number VR20192203.

which have in common that they are induced by a group of translations isomorphic to the group of translations acting on the message space by means of the xor addition with the key. In the context of an SPN, where the encrypted message is generated by iterating through a sequence of s-box layers, (xor)-linear diffusion, and xor-based key addition layers, altering the differential operator yields a dual impact. On one hand, it is highly probable that differentials traverse the s-box layer more effectively, given that its non-linearity is maximised with respect to xor. On the other hand, differentials do not deterministically propagate through the diffusion layer, as observed in classical scenarios. This pivotal limitation effectively restricts the success of the attack only to cases where the target layer is linear not only concerning xor but also with respect to the operation under consideration for computing differentials.

A first successful attempt based on the study of the alternative differential properties of a xor-based toy cipher of the SPN family has shown that it is possible to highlight a bias in the distribution of the differences calculated compared to an alternative operation which is instead not detectable by means of the standard xor-differential-based approach [CBS19]. The target cipher featured five 3-bit s-boxes and the operation used to perform the attack acted as the xor on the last four s-boxes, while on the first one matched with one of the alternative sums defined by Calderini et al. [CCS21], coming from another translation groups. The advantage of employing an alternative operation in this case was only derived from the benefit induced by a single s-box. In a more recent experimental approach [CCI24], we showed that better results in a similar context can be obtained using an *alternative parallel operation*, in which every s-box can be targeted. In this case, the diffusion layer of the cipher was determined through an algorithm, ensuring that it adheres to the constraint of linearity with respect to both xor and the target operation.

In this paper, we establish a general result that, in the context of an SPN with 4-bit s-boxes, characterises all xor-linear maps that are concurrently linear with respect to a parallel alternative operation (Sec. 2). This finding enables the execution of a differential attack wherein each s-box affected by a non-trivial differential contributes to the final differential probability with increased efficacy compared to the conventional xor differentials. Additionally, differentials propagate deterministically through the linear layer in this scenario. Moreover, we examine all possible alternative operations on 4 bits and investigate the differential properties of optimal 4-bit s-boxes, following the classification outlined by Leander and Poschmann (a comparable methodology, albeit in the context of modular addition, was recently employed by Zajac and Jókay [ZJ20]. Our analysis demonstrates that each class comprises potentially weak permutations (Sec. 3). When coupled with a diffusion layer as described earlier, these permutations have the potential to render the cipher susceptible to differential attacks with alternative operations. To substantiate our findings, we conclude the paper by presenting experimental results on a family of toy SPNs (Sec. 4).

1.1. **Notation.** Let V be an n-dimensional vector space over $\mathbb{F}_2$ which represents the message space. We write $V = V_1 \oplus V_2 \oplus \cdots \oplus V_b$, where each V_j is isomorphic to a vector space B such that $\dim(B) = s$ on which every s-box acts. Thefore we have $n = sb$. We denote by $\{e_i\}_{i=1}^n$ the canonical basis of V. If G is any finite group acting on V, for each $g \in G$ and $v \in V$ we denote the action of g on v as vg, i.e. we use postfix notation for every function evaluation. We denote by $\mathrm{Sym}(V)$ the symmetric group acting on V, i.e. the group of all permutation on the message space, by $\mathrm{GL}(V,+)$ the group of linear transformations, and by $\mathrm{AGL}(V,+)$ the group of affine permutations. The identity matrix of size l is denoted by $\mathbb{1}_l$ and the zero matrix of size $l \times h$ is denoted by $\mathbb{0}_{l,h}$, or simply $\mathbb{0}_l$ if $l = h$. We finally denote by T_+ the group of translations on V, i.e. $T_+ := \{\sigma_a \mid a \in V, x \mapsto x + a\} < \mathrm{Sym}(V)$. We remind that the translation σ_k acts on a

vector x in the same way the key-addition layer of an SPN acts xor-ing the round key k to the message x, i.e. $x\sigma_k = x + k$.

1.2. **Preliminaries on alternative operations.** An alternative operation on V can be defined given any 2-elementary abelian regular subgroup $T < \mathrm{AGL}(V, +)$, that we can write as $T = \{\tau_a \mid a \in V\}$, where τ_a is the unique element in T which maps 0 into a. Consequently, for all $a, b \in$V, we can define $a \circ b := a\tau_b$, resulting in $(V, \circ)$ forming an additive group. The operation $\circ$ induces a vector space structure on V, with the corresponding group of translation being $T_\circ = T$. Additionally, for each $a \in V$, there exists $M_a \in \mathrm{GL}(V, +)$ such that $\tau_a = M_a\sigma_a$, meaning that for every $x \in V$,

$$x \circ a = x\tau_a = xM_a + a.$$

It is also assumed throughout that $T_+ < \mathrm{AGL}(V, \circ)$, where $\mathrm{AGL}(V, \circ)$ is the normaliser in $\mathrm{Sym}(V)$ of $T_\circ$ (i.e., the group of affine permutations w.r.t. $\circ$). This crucial technical assumption renders the key-addition layer an affine operator concerning the new operation, enabling the prediction of how the key addition affects the differentials with a reasonable probability. Further details on this aspect, which may not be directly relevant to the scope of the current paper, can be found in Civino et al. [CBS19]. In this context, we define the *weak keys subspace* as

$$W_\circ := \{a \mid a \in V, \sigma_a = \tau_a\} = \{k \mid k \in V,\ \forall x \in V\ x \circ k = x + k\}.$$

$W_\circ$ is a vector subspace of both $(V, +)$ and $(V, \circ)$. It is known [CDVS06, CCS21] that $W_\circ$ is non empty and that

$$2 - (n \bmod 2) \leq \dim(W_\circ) \leq n - 2. \tag{1}$$

Moreover, up to conjugation we can always assume $W_\circ$ to be the span of the last d canonical vectors of V [CCS21]. This allows to represent the new sum in a canonical way [CCS21]: for each $a \in V$ there exists a matrix $E_a \in \mathbb{F}_2^{(n-d)\times d}$ such that

$$M_a = \begin{pmatrix} \mathbb{1}_{n-d} & E_a \\ \mathbb{0}_{d,n-d} & \mathbb{1}_d \end{pmatrix}. \tag{2}$$

Fixing such an operation as above is therefore equivalent to defining the matrices

$$M_{e_i} = \begin{pmatrix} \mathbb{1}_{n-d} & E_{e_i} \\ \mathbb{0}_{d,n-d} & \mathbb{1}_d \end{pmatrix} = \left(\begin{array}{c|c} & \mathbf{b}_{i,1} \\ \mathbb{1}_{n-d} & \vdots \\ & \mathbf{b}_{i,n-d} \\ \hline \mathbb{0}_{d,n-d} & \mathbb{1}_d \end{array}\right)$$

for $1 \leq i \leq n$, where $\mathbf{b}_{i,j} \in \mathbb{F}_2^d$. The assumptions on $T_\circ$ and on $W_\circ$ imply that $E_{e_i} = 0$ for $n - d + 1 \leq i \leq n$, $\mathbf{b}_{i,i} = \mathbf{0}$ and $\mathbf{b}_{i,j} = \mathbf{b}_{j,i}$. In conclusion, the following result characterises the criteria that the vectors $\mathbf{b}_{i,j}$ must adhere to in order to define an alternative operation as previously described.

Theorem 1.1 ([CBS19])**.** *Let $T_\circ < \mathrm{AGL}(V, +)$ be 2-elementary, abelian, and regular, and let $d \leq n - 2$. The operation $\circ$ induced by $T_\circ$ is such that $d = \dim(W_\circ)$, $T_+ < \mathrm{AGL}(V, \circ)$, and $W_\circ = \mathrm{Span}\{e_{n-d+1}, \ldots, e_n\}$ if and only if the matrix $\Theta_\circ \in (\mathbb{F}_{2^d})^{(n-d)\times(n-d)}$ defined as*

$$\Theta_\circ := \begin{pmatrix} \mathbf{b}_{1,1} & \mathbf{b}_{1,2} & \cdots & \mathbf{b}_{n-d,1} \\ \mathbf{b}_{2,1} & \mathbf{b}_{2,2} & \cdots & \mathbf{b}_{n-d,2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{b}_{n-d,1} & \mathbf{b}_{n-d,2} & \cdots & \mathbf{b}_{n-d,d} \end{pmatrix}$$

is zero-diagonal, symmetric and no $\mathbb{F}_2$-linear combination of its columns is the null vector. The matrix $\Theta_\circ$ is also called the defining matrix *for $\circ$.*

In the subsequent discussion, the term *alternative operation* refers to an additive law $\circ$ on V as defined above.

2. Parallel operations and their automorphism groups

Let $\circ$ be an alternative operation on the block-sized space V. As outlined in the introduction, if $\lambda \in \mathrm{GL}(V,+)$ represents a (xor)-linear diffusion layer, and $\Delta \in V$ is an input difference traversing λ, predicting the output difference with respect to $\circ$, i.e.,

$$x\lambda \circ (x \circ \Delta)\lambda,$$

becomes inherently challenging without additional assumptions on λ that ensure a sufficiently high predictive probability. For this reason, the examination of the following object becomes crucial: in cryptographic terms, it contains potential diffusion layers that allow differentials, whether computed with respect to xor or $\circ$, to propagate with a probability of 1.

Definition 2.1. Let $\circ$ be an alternative operation on V. Let us define

$$H_\circ := \{f \in \mathrm{GL}(V,+) \mid \forall a,b \in V : \ (a \circ b)f = af \circ bf\}$$

to be the subgroup of $\mathrm{GL}(V,+)$ of permutations that are linear w.r.t. the operation $\circ$. More precisely, denoting by $\mathrm{AGL}(V,\circ)$ the normaliser in $\mathrm{Sym}(V)$ of $T_\circ$ and by $\mathrm{GL}(V,\circ)$ the stabiliser of 0 in $\mathrm{AGL}(V,\circ)$, we have $H_\circ = \mathrm{GL}(V,+) \cap \mathrm{GL}(V,\circ)$.

The structure of the group $H_\circ$ in its most general case has not been understood yet. This work addresses this challenge in a specific scenario, guided by assumptions that are deemed reasonable within the context of differential cryptanalysis.

Assumption 1: $\circ$ is a parallel operation. While the operation $\circ$ could, in theory, be defined on the entire message space V, studying the differential properties of the s-box layer, considered as a function with 2^n inputs, is impractical for standard-size ciphers. For this reason, we focus on operations applied in a *parallel* way to each s-box-sized block, i.e., $\circ = (\circ_1, \circ_2, \ldots, \circ_b)$, where for each $1 \le j \le b$, $\circ_j$ is an operation on V_j. In this scenario, every operation is acting independently on the s-box space B, regardless of the others. This motivates the following definition.

Definition 2.2. Let $\circ$ be an alternative operation on V. We say that $\circ$ is *parallel* if for each $1 \le j \le b$ there exists an alternative operation $\circ_j$ on V_j such that for each $x, y \in V$ we have

$$x \circ y = \begin{pmatrix} x_1 \\ \vdots \\ x_b \end{pmatrix} \circ \begin{pmatrix} y_1 \\ \vdots \\ y_b \end{pmatrix} = \begin{pmatrix} x_1 \circ_1 y_1 \\ \vdots \\ x_b \circ_b y_b \end{pmatrix},$$

where $x = (x_1, x_2, \ldots, x_b), y = (y_1, y_2, \ldots, y_b)$ and each component belongs to the s-box-sized space, i.e., $x_j, y_j \in V_j \cong B$ for $1 \le j \le b$.

In the notation of Sec. 1.2, up to a block matrix conjugation, we can assume that every element $x \in V$ is associated to a translation $\tau_x = M_x \sigma_x$, with

$$M_x = \begin{pmatrix} M_{x_1}^{\circ_1} & \cdots & \mathbb{0} \\ \vdots & \ddots & \vdots \\ \mathbb{0} & \cdots & M_{x_b}^{\circ_b} \end{pmatrix}$$

where $M_{x_i}^{\circ_i}$ is the matrix associated to the translation τ_{x_i} with respect to the sum $\circ_i$, as defined in Eq. (2). Notice that it can be assumed, without loss of generality, that all the operations $\circ_j$ coincide.

Assumption 2: $\dim(W_{\circ_j}) = s-2$. According to Eq. (1), every operation $\circ_j$ defined at the s-box level must satisfy the bound $\dim(W_{\circ_j}) \leq s-2$, being $s = \dim(B)$. The situation where the (upper) bound is reached holds particular interest for several reasons, as elaborated further in Civino et al. [CBS19]. Notably,

- if the s-box size s is four, the case where $\dim(W_{\circ_j}) = 2$ is the sole possibility;
- the considered case stands today as the only one for which the structure of $H_{\circ_j}$ is well understood.

For the reader's convenience, we present the classification result for $H_{\circ_j}$ obtained by Civino et al. in the considered case. Additionally, it is worth recalling that, according to Theorem 1.1, any $\circ_j$ for which $\dim(W_{\circ_j}) = s-2$ is determined by a single non-null vector $\mathbf{b} \in (\mathbb{F}_2)^{s-2}$.

Theorem 2.3 ([CBS19])**.** *Let $\circ_j$ be an alternative operation such that $d = \dim(W_{\circ_j}) = s-2$ defined by a vector $\boldsymbol{b} \in (\mathbb{F}_2)^{s-2}$, and let $\lambda \in (\mathbb{F}_2)^{s\times s}$. The following are equivalent:*

- *$\lambda \in H_{\circ_j}$;*
- *there exist $A \in \mathrm{GL}((\mathbb{F}_2)^2, +)$, $D \in \mathrm{GL}((\mathbb{F}_2)^d, +)$, and $B \in (\mathbb{F}_2)^{2\times d}$ such that*

$$\lambda = \begin{pmatrix} A & B \\ \mathbb{0}_{d,2} & D \end{pmatrix}$$

and $\boldsymbol{b}D = \boldsymbol{b}$.

We are now prepared to present the first novel contribution of this work, wherein we characterise the group $H_\circ$ for a parallel operation $\circ = (\circ_1, \circ_2, \ldots, \circ_b)$ with components at the s-box level satisfying $\dim(W_{\circ_j}) = s-2$. For the sake of simplicity and without losing generaly, we assume that the b operations at the s-box level coincide.

Theorem 2.4. *Let $\circ = (\circ_1, \circ_2, \ldots, \circ_b)$ be a parallel alternative operation on V such that for each $1 \leq j \leq b$ $\circ_j$ is an alternative operation on V_j. Let us assume that every $\circ_j$ is such that $\dim(W_{\circ_j}) = s-2$ and it is defined by a vector $\boldsymbol{b} \in (\mathbb{F}_2)^{s-2}$. Let $\lambda \in (\mathbb{F}_2)^{n\times n}$. Then, $\lambda \in H_\circ$ if and only if it can be represented in the block form*

$$\lambda = \left(\begin{array}{cc|c|cc} A_{11} & B_{11} & \cdots & A_{1b} & B_{1b} \\ C_{11} & D_{11} & & C_{1b} & D_{1b} \\ \hline \multicolumn{2}{c|}{\vdots} & \ddots & \multicolumn{2}{c}{\vdots} \\ \hline A_{b1} & B_{b1} & \cdots & A_{bb} & B_{bb} \\ C_{b1} & D_{b1} & & C_{bb} & D_{bb} \end{array}\right),$$

where

1. *$A_{ij} \in (\mathbb{F}_2)^{2\times 2}$ such that for each row and each column of blocks there exists one and only one non-zero A_{ij}; moreover, all the non-zero A_{ij} are invertible;*
2. *$B_{ij} \in (\mathbb{F}_2)^{2\times(s-2)}$;*
3. *$C_{ij} = \mathbb{0}_{(s-2)\times 2}$;*
4. *$D_{ij} \in (\mathbb{F}_2)^{(s-2)\times(s-2)}$ such that if A_{ij} is zero, then $\boldsymbol{b}D_{ij} = \boldsymbol{0}$, and if A_{ij} is invertible, then $\boldsymbol{b}D_{ij} = \boldsymbol{b}$. Moreover, the matrix D defined by*

$$D := \begin{pmatrix} D_{11} & \cdots & D_{1b} \\ \vdots & \ddots & \vdots \\ D_{b1} & \cdots & D_{bb} \end{pmatrix}$$

is invertible.

Proof. The proof involves standard linear algebra techniques, but its extensive and laborious nature necessitates omission due to page limitations. □

TABLE 1. Optimal 4-bit permutations according to Leander and Poschmann

	0_x	1_x	2_x	3_x	4_x	5_x	6_x	7_x	8_x	9_x	A_x	B_x	C_x	D_x	E_x	F_x
G_0	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	B_x	C_x	9_x	3_x	E_x	A_x	5_x
G_1	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	B_x	E_x	3_x	5_x	9_x	A_x	12_x
G_2	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	B_x	E_x	3_x	A_x	C_x	5_x	9_x
G_3	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	C_x	5_x	3_x	A_x	E_x	B_x	9_x
G_4	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	C_x	9_x	B_x	A_x	E_x	5_x	3_x
G_5	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	C_x	B_x	9_x	A_x	E_x	3_x	5_x
G_6	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	C_x	B_x	9_x	A_x	E_x	5_x	3_x
G_7	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	C_x	E_x	B_x	A_x	9_x	3_x	5_x
G_8	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	9_x	5_x	A_x	B_x	3_x	12_x
G_9	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	B_x	3_x	5_x	9_x	A_x	12_x
G_{10}	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	B_x	5_x	A_x	9_x	3_x	12_x
G_{11}	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	B_x	A_x	5_x	9_x	C_x	3_x
G_{12}	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	B_x	A_x	9_x	3_x	C_x	5_x
G_{13}	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	C_x	9_x	5_x	B_x	A_x	3_x
G_{14}	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	C_x	B_x	3_x	9_x	5_x	10_x
G_{15}	0_x	1_x	2_x	D_x	4_x	7_x	F_x	6_x	8_x	E_x	C_x	B_x	9_x	3_x	A_x	5_x

3. Differential properties of optimal s-boxes

In this section we delve into the examination of the differential properties exhibited by all possible 4-bit permutations, with respect to all possible alternative operations defined as in Sec. 1.2. In particular, we set $s = 4$ and therefore consider $B = \mathbb{F}_2^4$. We begin by acknowledging that, despite the compact size of the space, the count of alternative operations on B is considerable:

Proposition 3.1 ([CCS21]). *There exist 105 different elementary abelian regular subgroups groups $T_\circ$ in $AGL(\mathbb{F}_2^4, +)$. Furthermore, each of them satisfies $T_+ < AGL(\mathbb{F}_2^4, \circ)$ and* $\dim W_\circ = s - 2 = 2$.

We recall that given a permutation $f \in \mathrm{Sym}(B)$ we can define

$$\delta_f(a, b) = \#\{x \in B \mid xf + (x + a)f = b\}.$$

The *differential uniformity* of f is defined as $\delta_f := \max_{a \neq 0} \delta_f(a, b)$ and it represent the primary metric to consider when assessing the resistance of an s-box to differential cryptanalysis [Nyb93].

Several cryptographic properties, including differential uniformity, are preserved under affine equivalence for vectorial Boolean functions. Two functions, denoted as f and g, are considered *affine equivalent* if there exist two affine permutations, α and β, in $\mathrm{AGL}(V, +)$ such that $g = \beta f \alpha$.

Leander and Poschmann [LP07] provided a comprehensive classification (up to affine equivalence) of permutations over $B = \mathbb{F}_2^4$. They identified 16 classes with *optimal* cryptographic properties. All 16 classes exhibit a classical differential uniformity equal to 4, which represents the best possible value for s-boxes in $\mathrm{Sym}(B)$. The representatives of the 16 classes are listed in Table 3, where each vector is interpreted as a binary number, most significant bit first.

3.1. Dealing with affine equivalence.

Our goal is to analyse the differential uniformity of each optimal s-box class, with respect to every alternative operation $\circ$ on B. The definitions given above can be generalised in the obvious way setting $\delta_f^\circ(a, b) = \#\{x \in B \mid xf \circ (x \circ a)f = b\}$ and calling $\circ$-*differential uniformity* of f the value $\delta_f^\circ := \max_{a \neq 0} \delta_f^\circ(a, b)$.

It is noteworthy that, unlike in the case of classic differential uniformity, the value of δ_f° is not invariant under affine equivalence. However, verifying the $\circ$-differential uniformity of $g_2 G_i g_1$ for any optimal class and every pair $g_1, g_2 \in \mathrm{AGL}(V,+)$ would be impractical. Therefore, a reduction in the number of permutations to be checked is necessary, and for this purpose, we make the following observations. First, similar to the classical case, the $\circ$-differential uniformity is preserved under affine transformations w.r.t. $\circ$.

Proposition 3.2. *Given $f \in \mathrm{Sym}(B)$ and $g_1, g_2 \in \mathrm{AGL}(B,\circ)$ we have*

$$\delta^\circ_{g_1 f g_2}(a,b) = \delta^\circ_f(g_2(a), g_1^{-1}(b)).$$

Moreover, Proposition 3.1 establishes that for any $\circ$ derived from a translation group in $\mathrm{AGL}(B,+)$, the $+$-translations are affine with respect to $\circ$. This initial observation allows us to narrow down the analysis to $g_2 G_i g_1$ with $g_1, g_2 \in \mathrm{GL}(B,+)$, which still remains impractical. Furthermore, considering that $H_\circ = \mathrm{GL}(B,+) \cap \mathrm{GL}(B,\circ)$, Proposition 3.2 establishes that left and right multiplication by elements in $H_\circ$ preserves both $\circ$ and $+$-differential uniformity. It is noteworthy that during this process, the rows of the matrix containing all the $\delta_f^\circ(a,b)$ (DDT°) are merely shuffled, thereby preserving the highest element of each row. Therefore, the following conclusion can be easily obtained.

Proposition 3.3. *Let $g_1, g_2 \in \mathrm{GL}(B,+)$ and $f \in \mathrm{Sym}(B)$. For any $g_1' \in g_1 H_\circ$ and $g_2' \in H_\circ g_2$ we have*

$$\delta^\circ_{g_2 f g_1} = \delta^\circ_{g_2' f g_1'}.$$

Proof. Take $h_1, h_2 \in H_\circ$ such that $g_1' = g_1 h_1$ and $g_2' = h_2 g_2$. Then,

$$x g_2' f g_1' \circ (x \circ a) g_2' f g_1' = x h_2 g_2 f g_1 \circ (x h_2 \circ a h_2 g_2 f g_1) h_1,$$

implying that $\delta^\circ_{g_2' f g_1'}(a,b) = \delta^\circ_{g_2 f g_1}(a h_2, b h_1^{-1})$. So, $\delta^\circ_{g_2' f g_1'} = \delta^\circ_{g_2 f g_1}$. □ □

The final proposition allows us to focus solely on g_1 and g_2 within the left and right cosets of $H_\circ$. These reductions facilitate the analysis of the potential $\circ$-differential uniformities attainable across all classes of optimal permutations for the 105 conceivable alternative sums defined over B. For each of the 105 alternative operations, we systematically explored each of the 16 classes, following the described procedure, and we recorded the $\circ$-differential uniformity for every candidate. To streamline the presentation, we calculated the average across the 105 operations and presented the consolidated results in Tab. 3.1.

In our examination, we observe that if, for a given operation $\circ$, certain elements within an affine equivalence class yield a $\circ$-differential uniformity δ, then this value δ is achieved by some element in the entire class for all alternative operations. Our analysis reveals that certain optimal functions may exhibit the highest differential uniformity (16) for alternative operations, specifically the classes G_0 (containing, e.g., the s-box S1 of Serpent [BAK98]), G_1 (containing, e.g., the s-box of Present [BKL+07]), G_2, and G_8. Conversely, the classes G_3, G_4, G_5, G_6, G_{11}, and G_{12} demonstrate more favorable behavior concerning alternative operations.

4. Experiments on a 16-bit block cipher with 4-bit s-boxes

In this concluding section, we aim to apply the results obtained above to a family of (toy) ciphers. These ciphers may exhibit security under classical differential cryptanalysis but reveal vulnerabilities to the alternative differential approach.

In our experiments, we set $V = \mathbb{F}_2^{16}$, $n = 4$, and $s = 4$, defining $\circ$ as the parallel sum by applying the alternative operation defined by the vector $\mathbf{b} = (0,1)$ to each 4-bit block. Moreover, all our ciphers will feature the 4-bit permutation $\gamma : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ defined by the sequence (0_x, E_x,

TABLE 2. Avg. number of functions with given ∘-differential uniformity

Class \ δ°	2	4	6	8	10	12	14	16
G_0	0	914	7842	3463	420	19	0	14
G_1	0	1019	10352	4226	560	0	0	18
G_2	0	1003	8604	3805	462	21	0	16
G_3	0	16733	117740	27639	1779	0	0	0
G_4	0	1101	9295	2715	179	0	0	0
G_5	0	2479	24135	5402	639	0	0	0
G_6	0	1632	10842	3071	218	0	0	0
G_7	0	1257	10679	2994	186	28	0	0
G_8	0	1691	12821	6113	583	93	0	24
G_9	0	1228	7734	2693	154	39	0	0
G_{10}	0	1228	8063	2763	166	41	0	0
G_{11}	0	1637	9940	2941	214	0	0	0
G_{12}	0	2541	16832	5308	352	0	0	0
G_{13}	0	1124	9520	2416	217	15	0	0
G_{14}	0	1207	7641	2584	160	51	0	0
G_{15}	0	1227	7776	2630	163	52	0	0

$\mathtt{B}_\mathtt{x}$, $\mathtt{1}_\mathtt{x}$, $\mathtt{7}_\mathtt{x}$, $\mathtt{C}_\mathtt{x}$, $\mathtt{9}_\mathtt{x}$, $\mathtt{6}_\mathtt{x}$, $\mathtt{D}_\mathtt{x}$, $\mathtt{3}_\mathtt{x}$, $\mathtt{4}_\mathtt{x}$, $\mathtt{F}_\mathtt{x}$, $\mathtt{2}_\mathtt{x}$, $\mathtt{8}_\mathtt{x}$, $\mathtt{A}_\mathtt{x}$, $\mathtt{5}_\mathtt{x}$) as its s-box. Precisely, four copies of γ will act on the 16-bit block. Notice that the s-box $\gamma \in$ belongs to G_0 and has $\delta_\gamma = 4$ and $\delta^\circ_\gamma = 16$.

In all the experiments described below, we consider the SPN whose i-th round is obtained by the composition of the parallel application of the s-box γ on every 4-bit block, a 'diffusion layer' λ sampled random from $H_\circ$, and the xor with the i-th random round key. We study the difference propagation in the cipher in a long-key scenario, i.e., the key-schedule selects a random long key $k \in \mathbb{F}_2^{16r}$ where r is the number of rounds. To avoid potential bias from a specific key choice, we conduct our experiments by averaging over 2^{15} random long-key generations. This approach gives us a reliable estimate of the expected differential probability for the best differentials in this cipher.

In 150 distinct executions, spanning a range of rounds from 3 to 10, we calculated the discrepancy between the most effective ∘-trail and +-trail. To manage computational resources, our focus was narrowed down to input differences with a Hamming weight of 1.

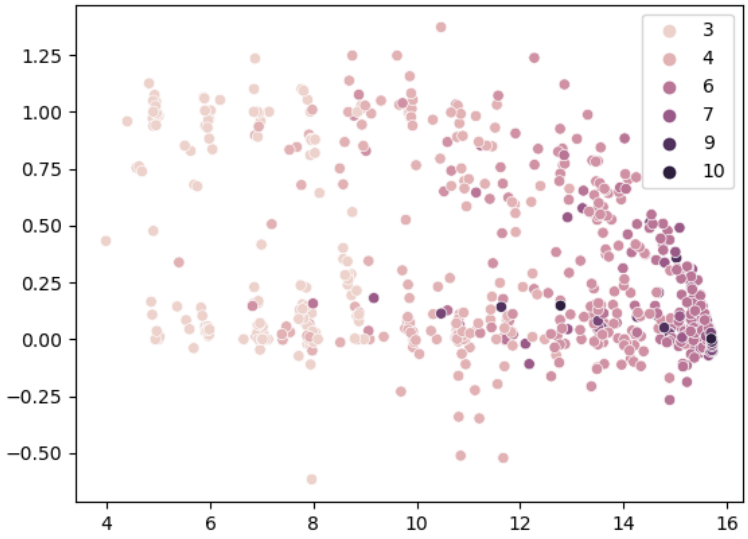
The results are depicted in Fig. 1, where each dot represents an individual simulation. The x axis corresponds to the negative logarithm of the probability of the best ∘ differential, while the y axis represents the difference between that value and the negative logarithm of the probability of the best + differential. Darker dots indicate a higher number of rounds, as explained in the legend. Notably, about half of the dots lie above zero, suggesting that the best ∘ differential consistently outperforms the best + differential until they become indistinguishable. Interestingly, this convergence often occurs when the ∘ probability is already very close to 16, providing potential candidates for our distinguisher attack.

REFERENCES

[BAK98] Eli Biham, Ross Anderson, and Lars Knudsen, *Serpent: A new block cipher proposal*, International workshop on fast software encryption, Springer, 1998, pp. 222–238.

[BBI+15] Subhadeep Banik, Andrey Bogdanov, Takanori Isobe, Kyoji Shibutani, Harunaga Hiwatari, Toru Akishita, and Francesco Regazzoni, *Midori: A block cipher for low energy*, Advances in Cryptology–ASIACRYPT 2015: 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29–December 3, 2015, Proceedings, Part II 21, Springer, 2015, pp. 411–436.

FIGURE 1. Comparison of + and ∘ trails for random mixing layers



[BBS05] Eli Biham, Alex Biryukov, and Adi Shamir, *Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials*, Journal of Cryptology **18** (2005), 291–311.

[BCJW02] Nikita Borisov, Monica Chew, Rob Johnson, and David Wagner, *Multiplicative differentials*, Fast Software Encryption: 9th International Workshop, FSE 2002 Leuven, Belgium, February 4–6, 2002 Revised Papers 9, Springer, 2002, pp. 17–33.

[BKL+07] Andrey Bogdanov, Lars R Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew JB Robshaw, Yannick Seurin, and Charlotte Vikkelsoe, *Present: An ultra-lightweight block cipher*, Cryptographic Hardware and Embedded Systems-CHES 2007: 9th International Workshop, Vienna, Austria, September 10-13, 2007. Proceedings 9, Springer, 2007, pp. 450–466.

[BS91] Eli Biham and Adi Shamir, *Differential cryptanalysis of DES-like cryptosystems*, Journal of CRYPTOLOGY **4** (1991), 3–72.

[CBS19] Roberto Civino, Céline Blondeau, and Massimiliano Sala, *Differential attacks: using alternative operations*, Designs, Codes and Cryptography **87** (2019), 225–247.

[CCI24] Marco Calderini, Roberto Civino, and Riccardo Invernizzi, *Differential experiments using parallel alternative operations*, Journal of Mathematical Cryptology **18** (2024), no. 1, 20230030.

[CCS21] Marco Calderini, Roberto Civino, and Massimiliano Sala, *On properties of translation groups in the affine general linear group with applications to cryptography*, Journal of Algebra **569** (2021), 658–680.

[CDVS06] A. Caranti, F. Dalla Volta, and M. Sala, *Abelian regular subgroups of the affine group and radical rings*, Publ. Math. Debrecen **69** (2006), no. 3, 297–308.

[Knu95] Lars R Knudsen, *Truncated and higher order differentials*, Fast Software Encryption: Second International Workshop Leuven, Belgium, December 14–16, 1994 Proceedings 2, Springer, 1995, pp. 196–211.

[LP07] Gregor Leander and Axel Poschmann, *On the classification of 4 bit s-boxes*, Arithmetic of Finite Fields: First International Workshop, WAIFI 2007, Madrid, Spain, June 21-22, 2007. Proceedings 1, Springer, 2007, pp. 159–176.

[Nyb93] Kaisa Nyberg, *Differentially uniform mappings for cryptography*, Workshop on the Theory and Application of of Cryptographic Techniques, Springer, 1993, pp. 55–64.

[SIH+11] Kyoji Shibutani, Takanori Isobe, Harunaga Hiwatari, Atsushi Mitsuda, Toru Akishita, and Taizo Shirai, *Piccolo: an ultra-lightweight blockcipher*, Cryptographic Hardware and Embedded Systems–CHES 2011: 13th International Workshop, Nara, Japan, September 28–October 1, 2011. Proceedings 13, Springer, 2011, pp. 342–357.

[Teş22] George Teşeleanu, *The security of quasigroups based substitution permutation networks*, International Conference on Information Technology and Communications Security, Springer, 2022, pp. 306–319.

[Wag99] David Wagner, *The boomerang attack*, International Workshop on Fast Software Encryption, Springer, 1999, pp. 156–170.

[ZJ20] Pavol Zajac and Matúš Jókay, *Cryptographic properties of small bijective s-boxes with respect to modular addition*, Cryptography and Communications **12** (2020), 947–963.

Department of Mathematics - University of Trento
Email address: marco.calderini@unitn.it

DISIM - University of L'Aquila
Email address: roberto.civino@univaq.it

COSIC - KU Leuven
Email address: riccardo.invernizzi@kuleuven.be