IACR Transactions on Symmetric Cryptology
ISSN 2519-173X, Vol. 2019, No. 1, pp. 170–191. DOI:10.13154/tosc.v2019.i1.170-191

A General Proof Framework for Recent AES Distinguishers

Christina Boura[1,2], Anne Canteaut[1] and Daniel Coggia[3,1]

[1] Inria, Paris, France
Anne.Canteaut@inria.fr,Daniel.Coggia@inria.fr

[2] University of Versailles Saint-Quentin-en-Yvelines (UVSQ), Versailles, France
Christina.Boura@uvsq.fr

[3] Direction Générale de l'Armement, Paris, France

Abstract. In this paper, a new framework is developed for proving and adapting the recently proposed *multiple-of-8* property and *mixture-differential* distinguishers. The above properties are formulated as immediate consequences of an equivalence relation on the input pairs, under which the difference at the output of the round function is invariant. This approach provides a further understanding of these newly developed distinguishers. For example, it clearly shows that the branch number of the linear layer does not influence the validity of the property, on the contrary of what was previously believed. We further provide an extension of the mixture-differential distinguishers and *multiple-of-8* property to any SPN and to a larger class of subspaces. These adapted properties can then be exhibited in a systematic way for other ciphers than the AES. We illustrate this with the examples of Midori, Klein, LED and Skinny.

Keywords: AES, Distinguisher, Subspace Trail Cryptanalysis

1 Introduction

The Advanced Encryption Standard (AES) is a block cipher designed by Daemen and Rijmen in 1997 and standardised by the NIST in 2001 [AES01]. It is since then the most used and the most analysed symmetric primitive worldwide. Since its submission to the NIST competition, many different cryptanalytic techniques, including among others, integral attacks [FKL+01], or sophisticated meet-in-the-middle attacks [DS08, DF14, DF16] have been developed and applied to the AES, exploiting different properties of the algorithm. Up to now, all the developed techniques, at least in the single-key model, have only managed to break reduced-round versions of the standard[1].

Most of the attacks against block ciphers are based on the existence of a distinguisher, that is a non-random property that permits to distinguish within reasonable time and by using reasonable data and memory resources, a reduced-round version of the cipher instantiated with a random secret key from a random permutation. Until recently, all known distinguishers of the AES in the single-key model could reach at most 4 rounds. However, since 2016, the first 5-round AES-distinguishers appeared [SLG+16, GRR17, RBH17, Gra18] and this topic has become again a subject of broad and current interest. The importance of these distinguishers is that they exhibit new, unexplored properties of the AES. They led to improved attacks on reduced-round versions of the cipher, like the attack on 5 rounds described in [BDK+18] based on the distinguisher exhibited in [Gra18].

[1] In 2011, biclique attacks were applied against all full-round versions of the AES [BKR11]. This technique permitted to reduce the exhaustive key-search by a few bits. However, we consider here that this kind of attack is a form of accelerated key-search.

Received: 2018-11-23, Accepted: 2019-01-23, Published: 2019-03-08

The main breakthrough in these attacks is the identification by Grassi, Rechberger and Rønjom [GRR17] of the following property of the AES round function $\mathcal{R}$. There exist two well-chosen linear subspaces V and W of $\mathbb{F}_2^{128}$ satisfying the following property: for any coset of V, $(c+V)$, the number of distinct pairs of elements x, x', $x \neq x'$ in $(c+V)$ such that $\mathcal{R}(x)$ and $\mathcal{R}(x')$ belong to the same coset of W is always divisible by 8. This behaviour, known as the *multiple-of-8* property, is then combined with two 2-round deterministic subspace trails to form a 5-round distinguisher. However, despite the theoretical interest of this distinguisher, because of the nature of the particular subspace V used, it could not be exploited directly for mounting a key-recovery attack. For this reason, Grassi presented in [Gra18] new 4-round distinguishers, that exploit a property appearing in the proof of the multiple-of-8 property but that is expressed in a more convenient way and facilitates key-recovery attacks. These new distinguishers were given the name of *mixture-differential distinguishers*.

The proof of the multiple-of-8 property given in [GRR17] is divided into many special cases and each case needs to be proved separately. An adaptation of the initial property, considering slightly more general input subspaces is also provided in the same paper and it requires a similar case-by-case proof. The disadvantage of these redundant proofs is that it is not clear from them what are the characteristics and the properties of the inner components of the AES that have an influence on the multiple-of-8 behaviour. Furthermore, the question whether this kind of property is proper to the AES or whether it can be adapted to other ciphers is unclear and the original proofs do not provide hints to answer this question. The same questions can be asked for the mixture-differential distinguishers.

Then, the aim of our paper is to provide a general formulation of the mixture-differential distinguisher and of the multiple-of-8 property which can be applied *in a systematic way* to any cipher following the SPN construction. It then avoids all these redundant proofs which were previously required for each new occurrence of these properties. Also, our result precisely identifies the conditions to be satisfied for the property to hold. Most notably, it shows that these distinguishers apply to a more general class of SPN than the ones mentioned in [GRR17, Gra18].

Our contributions. We show that the mixture-differential distinguishers, and by extension the multiple-of-8 property, revealed in [GRR17, Gra18] are direct consequences of the fact that the difference between two outputs of the AES round function, $\mathcal{R}(p^0) + \mathcal{R}(p^1)$, is invariant under an *equivalence relation between the plaintext pairs*. The definition of this equivalence relation leads to a simple and compact proof of the distinguishing properties exhibited in [GRR17, Gra18]. Also it clarifies which parts of the AES have an influence on the property. Most notably, the validity of the property *does not depend* on the branch number of `MixColumns`, on the contrary of what was previously believed in [GRR17, Gra18]. This disproves for instance the statement in the abstract of [GRR17], which points out that "this new structural property [...] is independent of the details of the `MixColumns` matrix (with the exception that the branch number must be maximal)". We show here that the property holds even if `MixColumns` has a lower branch number. Also, we describe the form of the linear subspaces that can be used in the distinguisher in place of the *mixed spaces* $\mathcal{M}_I$ in the original article. This permits to adapt the property and makes it directly applicable to other ciphers. As an illustration, we show that the same kind of property holds for the block ciphers Midori [BBI+15], Klein [GNL12], LED [GPPR11] and Skinny [BJK+16].

Organisation of the paper. The rest of the paper is organised as follows. A description of the AES and basic definitions on subspace trails are recalled in Section 2. Section 3 describes the multiple-of-8 property and the mixture-differential distinguishers presented in [GRR17, Gra18]. Section 4 then exhibits an equivalence relation between pairs of AES

states, which leads to a new proof of the above properties. It also discusses the influence of the branch number of `MixColumns` on these results. Then, an adaptation to any SPN and other linear subspaces is presented in Section 5. Finally, applications on different ciphers are provided in Section 6.

2 Preliminaries

We start by providing a brief description of the AES and introduce in parallel the notation that we will use. We work throughout this paper with finite fields of characteristic 2 that are fields containing 2^d elements, seen as extensions of $\mathbb{F}_2$.

2.1 Description of the AES

The Advanced Encryption Standard [AES01] is a *Substitution-Permutation network* operating on 128-bit plaintexts. The master-key size can be 128, 192 or 256 bits and the round-key size is 128 bits. The number of rounds N_r is respectively 10, 12 or 14. The internal state is represented as a 4×4 matrix over the field $\mathbb{F}_{2^8}$, called the state array. An AES round $\mathcal{R}$ is the composition $\mathcal{K} \circ \mathcal{L} \circ \mathcal{S}$ where:

- $\mathcal{S}$ is the `SubBytes` operation applying the same invertible S-box to each $\mathbb{F}_{2^8}$-entry of the state array.
- $\mathcal{L} = \mathsf{MC} \circ \mathsf{SR}$ is the linear layer. SR, is the `ShiftRows` operation consisting of a cyclic shift of each row to the left and MC, which stands for `MixColumns`, consists in the left multiplication of the state array by a 4×4 constant matrix over $\mathbb{F}_{2^8}$. This matrix, denoted M_{MC}, is defined by

$$M_{\mathsf{MC}} = \begin{pmatrix} 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 3 \\ 3 & 1 & 1 & 2 \end{pmatrix},$$

  where $\mathbb{F}_{2^8}$ is identified with $\mathbb{F}_2^8$ based on the irreducible polynomial $X^8 + X^4 + X^3 + X + 1$, and the elements of $\mathbb{F}_2^8$ are represented as integers in $\{0, \ldots, 2^8 - 1\}$.
- $\mathcal{K}$ is the `AddRoundKey` operation bitwise adding to the state array a 128-bit round-key derived from the master key.

An important notion for measuring the quality of the linear layer of a block cipher is the so-called *differential branch number*. This notion, that will play a role later in our analysis is introduced below.

Definition 1 (Differential branch number [Dae95])**.** Let $M : (\mathbb{F}_2^d)^N \to (\mathbb{F}_2^d)^N$ be an $\mathbb{F}_2$-linear function. The *differential branch number* of M with respect to $\mathbb{F}_2^d$ is

$$\min_{x \in (\mathbb{F}_2^d)^N \setminus \{0\}} (wt_d(x) + wt_d(Mx))$$

where $wt_d(x) = \#\{i \in \{0, \ldots, N-1\} \mid x_i \neq 0\}$ when $x = (x_0, \ldots, x_{N-1}) \in (\mathbb{F}_2^d)^N$.

Notably, AES `MixColumns` MC has branch number 5 with respect to $\mathbb{F}_2^8$.

2.2 Subspace trails for AES

We now recall the basic notation on subspace-trail cryptanalysis of AES, first introduced in [GRR16].

Let d be the degree of the extension over $\mathbb{F}_2$ on which the S-box operates and $\mathbb{K} = \mathbb{F}_{2^d}$. For the AES, $d = 8$, but we rather present all results for an arbitrary d since simulations are often performed on small scale variants of the AES with a 4-bit S-box (see e.g. Section 5.2 in [Gra18]). Let $\mathcal{M}_4(\mathbb{K})$ denote the set of all 4×4-matrices over $\mathbb{K}$ and let $(e_{i,j})_{i,j\in\{0,\dots,3\}}$ be the canonical basis of $\mathcal{M}_4(\mathbb{K})$:

$$e_{i,j} = \begin{array}{c} \\ \\ \leftarrow i \end{array} \begin{array}{c} \begin{array}{c} \\ \\ j \\ \downarrow \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \end{array}$$

In the following, $\mathrm{vect}_{\mathbb{K}}(v_0, \dots, v_{k-1})$ denotes the linear space formed by all linear combinations with coefficients in $\mathbb{K}$ of the vectors $v_0, \dots, v_{k-1} \in \mathcal{M}_4(\mathbb{K})$. As in [GRR16, GRR17], we define the following subspaces of $\mathcal{M}_4(\mathbb{K})$ for $i \in \{0, \dots, 3\}$, with indices computed modulo 4.

$$\begin{array}{lcrcl}
\text{The column spaces} & : & \mathcal{C}_i & = & \mathrm{vect}_{\mathbb{K}}(e_{0,i}, e_{1,i}, e_{2,i}, e_{3,i}), \\
\text{The diagonal spaces} & : & \mathcal{D}_i & = & \mathrm{vect}_{\mathbb{K}}(e_{0,i}, e_{1,i+1}, e_{2,i+2}, e_{3,i+3}) = \mathsf{SR}^{-1}(\mathcal{C}_i), \\
\text{The anti-diagonal spaces} & : & \mathcal{ID}_i & = & \mathrm{vect}_{\mathbb{K}}(e_{0,i}, e_{1,i-1}, e_{2,i-2}, e_{3,i-3}) = \mathsf{SR}(\mathcal{C}_i), \\
\text{The mixed spaces} & : & \mathcal{M}_i & = & \mathsf{MC}(\mathcal{ID}_i).
\end{array}$$

For example, if $x_0, x_1, x_2, x_3 \in \mathbb{K}$,

$$\begin{pmatrix} x_0 & 0 & 0 & 0 \\ x_1 & 0 & 0 & 0 \\ x_2 & 0 & 0 & 0 \\ x_3 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{C}_0, \quad \begin{pmatrix} x_0 & 0 & 0 & 0 \\ 0 & x_1 & 0 & 0 \\ 0 & 0 & x_2 & 0 \\ 0 & 0 & 0 & x_3 \end{pmatrix} \in \mathcal{D}_0,$$

$$\begin{pmatrix} x_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 \\ 0 & 0 & x_2 & 0 \\ 0 & x_3 & 0 & 0 \end{pmatrix} \in \mathcal{ID}_0, \quad \begin{pmatrix} 2\cdot x_0 & x_1 & x_2 & 3\cdot x_3 \\ x_0 & x_1 & 3\cdot x_2 & 2\cdot x_3 \\ x_0 & 3\cdot x_1 & 2\cdot x_2 & x_3 \\ 3\cdot x_0 & 2\cdot x_1 & x_2 & x_3 \end{pmatrix} \in \mathcal{M}_0.$$

If $I \subseteq \{0, 1, 2, 3\}$, we then define :

$$\mathcal{C}_I = \bigoplus_{i\in I} \mathcal{C}_i, \quad \mathcal{D}_I = \bigoplus_{i\in I} \mathcal{D}_i, \quad \mathcal{ID}_I = \bigoplus_{i\in I} \mathcal{ID}_i, \quad \mathcal{M}_I = \bigoplus_{i\in I} \mathcal{M}_i.$$

Now that we have linear subspaces of $\mathcal{M}_4(\mathbb{K})$, we define their *cosets* as affine subspaces of $\mathcal{M}_4(\mathbb{K})$. More precisely, a *coset* of the linear subspace $V \subseteq \mathcal{M}_4(\mathbb{K})$ is a set of the form $V + a = \{v + a \mid v \in V\}$ where $a \in \mathcal{M}_4(\mathbb{K})$. Moreover, we are going to pay attention to subspaces that satisfy a specific property defined below.

Definition 2 (Subspace trail [GRR16])**.** Let $\mathcal{F} : \mathbb{K}^N \to \mathbb{K}^N$ be any map. Two linear subspaces $U, V \subseteq \mathbb{K}^N$ form a (one-round) $\mathcal{F}$-subspace trail if

$$\forall a \in \mathbb{K}^N, \exists b \in \mathbb{K}^N : \mathcal{F}(U + a) \subseteq V + b, \tag{1}$$

which is denoted by $U \overset{\mathcal{F}}{\rightrightarrows} V$. The negation is denoted by $U \overset{\mathcal{F}}{\not\rightrightarrows} V$. An $(r+1)$-tuple of subspaces $(U_0, \dots, U_r)$ is called a subspace trail (over r rounds) if

$$\forall i \in \{0, \dots, r-1\},\ U_i \overset{\mathcal{F}}{\rightrightarrows} U_{i+1}.$$

For example, we have the trivial subspace trails $\{0\} \overset{\mathcal{F}}{\rightrightarrows} \{0\}$ and $U \overset{\mathcal{F}}{\rightrightarrows} \mathbb{K}^N$. In this paper, we only consider *exact* subspace trails, i.e., for which equality holds in (1).

3 Distinguishers based on subspace trails

We describe in this section three distinguishers based on subspace trail cryptanalysis in order to have a complete understanding of the context in which our contribution lies. The first one, describes the multiple-of-8 property presented by Grassi, Rechberger and Rønjom at Eurocrypt 2017 [GRR17], while the other two are based on the mixture differential property and were published recently by Grassi in [Gra18].

3.1 The distinguisher from [GRR17]

We first describe the distinguisher presented in [GRR17]. We begin with this easy to verify lemma describing a two-round subspace trail for the AES.

Lemma 1 ([GRR16]). *Let $I \subseteq \{0,1,2,3\}$, then $\mathcal{D}_I \overset{\mathcal{R}}{\rightrightarrows} \mathcal{C}_I \overset{\mathcal{R}}{\rightrightarrows} \mathcal{M}_I$.*

Now comes a more subtle lemma which is the keystone of Theorem 1. In the whole paper, as in [GRR17], we always consider *unordered pairs* of elements and denote them as pair sets, i.e. $\{a, b\}$.

Lemma 2 ([GRR17]). *Let $a \in \mathcal{M}_4(\mathbb{K})$, $i \in \{0,1,2,3\}$, $J \subseteq \{0,1,2,3\}$. We define*

$$n = \#\{\ \{p^0, p^1\} \text{ with } p^0, p^1 \in \mathcal{M}_i + a \mid \mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{D}_J\}.$$

Then $n \equiv 0 \mod 8$.

A proof of Lemma 2 is given in Section 6 of [GRR17] but we provide in Section 4 a much more compact proof of this same result. A direct consequence of Lemmas 1 and 2 is the following theorem.

Theorem 1 ([GRR17]). *Let $a \in \mathcal{M}_4(\mathbb{K})$, $i \in \{0,1,2,3\}$, $J \subseteq \{0,1,2,3\}$. We define*

$$n = \#\{\ \{p^0, p^1\} \text{ with } p^0, p^1 \in \mathcal{D}_i + a \mid \mathcal{R}^5(p^0) + \mathcal{R}^5(p^1) \in \mathcal{M}_J\}.$$

Then $n \equiv 0 \mod 8$.

Theorem 1 directly provides a distinguisher for five rounds of the AES independent of the secret key. Indeed, given an oracle simulating either five rounds of the AES, either a random permutation, one can compute the number n from Theorem 1 with only the 2^{32} plaintexts belonging to the same coset of $\mathcal{D}_i$. This distinguisher is fully described in [GRR17].

Exploiting the above distinguisher for mounting a key-recovery attack on more rounds revealed to be however a difficult task because of the form of the output subspace $\mathcal{M}_J$. Indeed, as $\mathcal{M}_J$ affects the whole AES-state, a key-recovery attack requires the guess of the entire subkey in the last round. For this reason, Grassi presented in [Gra18] new distinguishers that exploit similar properties but have a description that is more adapted to a key-recovery attack. The counterpart is that these distinguishers cover one round fewer than in [GRR17].

3.2 The two distinguishers from [Gra18]

In the following, we use the notation from [Gra18]: for a given basis $(g_0, \ldots, g_{k-1})$ and a given element $a \in \mathcal{M}_4(\mathbb{K})$, a vector p in the affine subspace $a + \mathrm{vect}_{\mathbb{K}}(g_0, \ldots, g_{k-1})$ defined by $p = a + \sum_{i=0}^{k-1} x_i g_i$ with $x_i \in \mathbb{K}$ is denoted by $p \equiv (x_0, \ldots, x_{k-1})$.

Theorem 2 ([Gra18]). *Let $a \in \mathcal{M}_4(\mathbb{K})$, $i \in \{0,1,2,3\}, J \subseteq \{0,1,2,3\}$ and let $p^0, p^1, q^0, q^1 \in \mathrm{vect}_{\mathbb{K}}(e_{0,i}, e_{1,i}) + a$. Suppose that*

1. $p^0 \equiv (x_0, x_1), \quad p^1 \equiv (y_0, y_1)$,
2. $q^0 \equiv (x_0, y_1), \quad q^1 \equiv (y_0, x_1)$.

Then

$$\mathcal{R}^4(p^0) + \mathcal{R}^4(p^1) \in \mathcal{M}_J \text{ if and only if } \mathcal{R}^4(q^0) + \mathcal{R}^4(q^1) \in \mathcal{M}_J.$$

We will provide an alternative proof of Theorem 2 in Section 5. The following variant of the distinguisher, involving another input subspace, is also exhibited in [Gra18].

Theorem 3 ([Gra18]). *Let $a \in \mathcal{M}_4(\mathbb{K})$, $i \in \{0,1,2,3\}$, $J \subseteq \{0,1,2,3\}$ and let $p^0, p^1 \in \mathcal{C}_i + a$. Suppose that $p^0 \equiv (x_0, x_1, x_2, x_3)$ and that $p^1 \equiv (y_0, y_1, y_2, y_3)$. Then,*

$$\mathcal{R}^4(p^0) + \mathcal{R}^4(p^1) \in \mathcal{M}_J \text{ if and only if } \mathcal{R}^4(q^0) + \mathcal{R}^4(q^1) \in \mathcal{M}_J,$$

for all sets of plaintexts $\{q^0, q^1\}$ with $q^0, q^1 \in \mathcal{C}_i + a$ of the following form:

1. $q^0 \equiv (y_0, x_1, x_2, x_3), \quad q^1 \equiv (x_0, y_1, y_2, y_3)$
2. $q^0 \equiv (x_0, y_1, x_2, x_3), \quad q^1 \equiv (y_0, x_1, y_2, y_3)$
3. $q^0 \equiv (x_0, x_1, y_2, x_3), \quad q^1 \equiv (y_0, y_1, x_2, y_3)$
4. $q^0 \equiv (x_0, x_1, x_2, y_3), \quad q^1 \equiv (y_0, y_1, y_2, x_3)$
5. $q^0 \equiv (x_0, x_1, y_2, y_3), \quad q^1 \equiv (y_0, y_1, x_2, x_3)$
6. $q^0 \equiv (x_0, y_1, x_2, y_3), \quad q^1 \equiv (y_0, x_1, y_2, x_3)$
7. $q^0 \equiv (x_0, y_1, y_2, x_3), \quad q^1 \equiv (y_0, x_1, x_2, y_3)$

We will provide a proof of Theorem 3 in Section 4.

4 A more concise and general proof

This section is dedicated to our proof of Lemma 2 from [GRR17] (Lemma 2 here). This new proof is a much more concise version of the case-by-case proof given in the original paper. To be more precise, instead of proving Lemma 2, we prove directly a more general variant. This generalisation is present in the original paper but its proof is only sketched in Appendix A of [GRR17]. Indeed, the proof framework of Lemma 2 in [GRR17] does not allow a compact proof of this generalisation.

Our approach for proving Lemma 2 can be divided into three steps:

- there exists an equivalence relation between pairs of elements in a certain subspace of $\mathcal{M}_4(\mathbb{K})$ (Definition 4);
- some function on those pairs derived from the round function is invariant under this equivalence relation (Theorem 4);
- the cardinality of the equivalence classes is always a multiple of 8 (Proposition 1).

In the following, we fix $a \in \mathcal{M}_4(\mathbb{K})$, $I \subseteq \{0,1,2,3\}$ and $J \subseteq \{0,1,2,3\}$. Here it might help to remind that

$$\mathcal{ID}_I = \mathrm{vect}_{\mathbb{K}}(e_{k,i-k} \mid k \in \{0,1,2,3\}, i \in I) = \mathrm{vect}_{\mathbb{K}}(e_{i-k,k} \mid k \in \{0,1,2,3\}, i \in I),$$
$$\mathcal{M}_I = \mathsf{MC}(\mathcal{ID}_I) = \mathrm{vect}_{\mathbb{K}}(\mathsf{MC}(e_{i-k,k}) \mid k \in \{0,1,2,3\}, i \in I).$$

4.1 An equivalence relation between pairs of states

Definition 3 (Information set)**.** Let $\{p^0, p^1\}$ be a set of elements in $\mathcal{M}_I + a$, written as

$$p^0 = \sum_{k=0}^{3} \sum_{i \in I} p^0_{i,k} \mathsf{MC}(e_{i-k,k}) + a \quad \text{and} \quad p^1 = \sum_{k=0}^{3} \sum_{i \in I} p^1_{i,k} \mathsf{MC}(e_{i-k,k}) + a$$

for some (uniquely defined) $p^0_{i,k}, p^1_{i,k} \in \mathbb{K}$, $i \in I, 0 \leq k \leq 3$. The *information set* K of $\{p^0, p^1\}$ is defined as

$$K = \{k \in \{0,1,2,3\} \mid \exists i \in I : p^0_{i,k} \neq p^1_{i,k}\}\ .$$

Definition 4 (Equivalence relation)**.** Let $P = \{p^0, p^1\}$ and $Q = \{q^0, q^1\}$ with $p^0, p^1, q^0, q^1 \in \mathcal{M}_I + a$. We say that $P \sim Q$ if:

- $\{p^0, p^1\}$ and $\{q^0, q^1\}$ have the same information set K.
- $\forall k \in K, \exists b \in \{0,1\} : \forall i \in I, q^0_{i,k} = p^b_{i,k}$ and $q^1_{i,k} = p^{1-b}_{i,k}$.

Clearly, $\sim$ is an equivalence relation on unordered pairs of $\mathcal{M}_I + a$.

Example 1. The following two sets $\{p_0, p_1\}$ and $\{q_0, q_1\}$, with $p_0, p_1, q_0, q_1 \in \mathcal{M}_0$ are equivalent.

$$\{p_0, p_1\} = \left\{ \begin{pmatrix} 2 \cdot x_0 & x_1 & z_2 & 3 \cdot z_3 \\ x_0 & x_1 & 3 \cdot z_2 & 2 \cdot z_3 \\ x_0 & 3 \cdot x_1 & 2 \cdot z_2 & z_3 \\ 3 \cdot x_0 & 2 \cdot x_1 & z_2 & z_3 \end{pmatrix}, \begin{pmatrix} 2 \cdot y_0 & y_1 & z_2 & 3 \cdot z_3 \\ y_0 & y_1 & 3 \cdot z_2 & 2 \cdot z_3 \\ y_0 & 3 \cdot y_1 & 2 \cdot z_2 & z_3 \\ 3 \cdot y_0 & 2 \cdot y_1 & z_2 & z_3 \end{pmatrix} \right\}$$

$$\sim$$

$$\{q_0, q_1\} = \left\{ \begin{pmatrix} 2 \cdot x_0 & y_1 & w_2 & 3 \cdot w_3 \\ x_0 & y_1 & 3 \cdot w_2 & 2 \cdot w_3 \\ x_0 & 3 \cdot y_1 & 2 \cdot w_2 & w_3 \\ 3 \cdot x_0 & 2 \cdot y_1 & w_2 & w_3 \end{pmatrix}, \begin{pmatrix} 2 \cdot y_0 & x_1 & w_2 & 3 \cdot w_3 \\ y_0 & x_1 & 3 \cdot w_2 & 2 \cdot w_3 \\ y_0 & 3 \cdot x_1 & 2 \cdot w_2 & w_3 \\ 3 \cdot y_0 & 2 \cdot x_1 & w_2 & w_3 \end{pmatrix} \right\}$$

The information set of these pairs has cardinality $|K| = 2$.

Now that we have the right definitions, we can state and prove the following key theorem that has also other applications than just proving Lemma 2. It is worth noticing that, in the original proof of Lemma 2, the authors split the proof procedure in several cases, each case corresponding to a different size of the information set. In our approach we assemble all cases together by using the above introduced equivalence relation.

Theorem 4. *For any $a \in \mathcal{M}_4(\mathbb{K})$, the function Δ operating on unordered pairs of elements in $\mathcal{M}_I + a$ and defined by*

$$\Delta : \quad \{p^0, p^1\} \quad \longmapsto \quad \mathcal{R}(p^0) + \mathcal{R}(p^1)$$

is constant over the equivalence classes for $\sim$.

Proof. Let $P = \{p^0, p^1\}$ and $Q = \{q^0, q^1\}$ with $p^0, p^1, q^0, q^1 \in \mathcal{M}_I + a$ such that $P \sim Q$. We write as in Definition 3

$$p^0 = \sum_{k=0}^{3} \sum_{i \in I} p^0_{i,k} \mathsf{MC}(e_{i-k,k}) + a \quad \text{and} \quad p^1 = \sum_{k=0}^{3} \sum_{i \in I} p^1_{i,k} \mathsf{MC}(e_{i-k,k}) + a.$$

We also write the `MixColumns` matrix $M_{\mathsf{MC}} = (m_{\ell,k})_{0 \le \ell,k \le 3}$. Hence

$$p^0 = \sum_{k,\ell} \sum_{i \in I} p^0_{i,k} m_{\ell,i-k} e_{\ell,k} + a = \sum_{k,\ell} \left(\sum_{i \in I} p^0_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) e_{\ell,k}.$$

Then $\mathcal{S}(p^0) = \sum_{k,\ell} \text{S-box} \left(\sum_{i \in I} p^0_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) e_{\ell,k}$ and

$$\mathcal{S}(p^0) + \mathcal{S}(p^1) = \sum_{k,\ell} \left[\text{S-box} \left(\sum_{i \in I} p^0_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) + \text{S-box} \left(\sum_{i \in I} p^1_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) \right] e_{\ell,k} \,. \tag{2}$$

It is now clear with Definition 4 and Equation (2) that $\mathcal{S}(p^0) + \mathcal{S}(p^1)$ and $\mathcal{S}(q^0) + \mathcal{S}(q^1)$ are equal in $\mathcal{M}_4(\mathbb{K})$. Indeed, with K the information set of P and Q,

$$\begin{aligned}
&\mathcal{S}(q^0) + \mathcal{S}(q^1) \\
&= \sum_{k,\ell \in \{0,1,2,3\}} \left[\text{S-box} \left(\sum_{i \in I} q^0_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) + \text{S-box} \left(\sum_{i \in I} q^1_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) \right] e_{\ell,k} \\
&= \sum_{\substack{\ell \in \{0,1,2,3\}, \\ k \in K}} \left[\text{S-box} \left(\sum_{i \in I} q^0_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) + \text{S-box} \left(\sum_{i \in I} q^1_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) \right] e_{\ell,k} \\
&= \sum_{\substack{\ell \in \{0,1,2,3\}, \\ k \in K}} \left[\text{S-box} \left(\sum_{i \in I} q^{b(k)}_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) + \text{S-box} \left(\sum_{i \in I} q^{1-b(k)}_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) \right] e_{\ell,k} \\
&= \sum_{\substack{\ell \in \{0,1,2,3\}, \\ k \in K}} \left[\text{S-box} \left(\sum_{i \in I} p^0_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) + \text{S-box} \left(\sum_{i \in I} p^1_{i,k} m_{\ell,i-k} + a_{\ell,k} \right) \right] e_{\ell,k} \\
&= \mathcal{S}(p^0) + \mathcal{S}(p^1) \,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Delta(P) &= \mathcal{R}(p^0) + \mathcal{R}(p^1) \\
&= \mathcal{K} \circ \mathcal{L} \circ \mathcal{S}(p^0) + \mathcal{K} \circ \mathcal{L} \circ \mathcal{S}(p^1) \\
&= \mathcal{L}(\mathcal{S}(p^0) + \mathcal{S}(p^1)) \quad \text{by characteristic 2 and linearity of } \mathcal{L}; \\
&= \mathcal{L}(\mathcal{S}(q^0) + \mathcal{S}(q^1)) \quad \text{by our previous observation;} \\
&= \Delta(Q).
\end{aligned}$$

□

We would like to adapt in Section 5 the previous theorem to any SPN structure and any linear space. For this, it is important to identify the key argument that makes the proof work. Looking at the proof carefully, it appears that it relies on the fact that the coordinates $p_{i,k}$ of the elements in $\mathcal{M}_I$ can be decomposed into disjoint sets, in such a

way that each individual S-box involves only one coordinate subset. Here, the four subsets correspond to the coordinates $p_{i,k}$ sharing the same index k. Indeed, the S-box at Row ℓ and Column k involves all $p_{i,k}, i \in I$. This particular structure then makes possible to exchange between the two pairs $\{p^0, p^1\}$ and $\{q^0, q^1\}$ the coordinates corresponding to one of the subsets. This is exactly the property that we will consider in the generalisation presented in Section 5.

4.2 The multiple-of-8 property

The multiple-of-8 property presented in Lemma 2 is then a direct consequence of the previous theorem. It is derived by combining the theorem with the following proposition, which computes the cardinality of the equivalence classes.

Proposition 1. *Let $\mathfrak{C}$ be an equivalence class with information set K. The cardinality of $\mathfrak{C}$ is*

$$|\mathfrak{C}| = 2^{|K|-1+d|I|(4-|K|)}.$$

It is always a multiple of 8.

Proof. Since for a given set $\{p^0, p^1\}$ with information set K, we have that $\forall k \notin K, \forall i \in I, p^0_{i,k} = p^1_{i,k}$, we have $(2^d)^{|I|\times(4-|K|)}$ choices for the shared coordinates in a pair of $\mathfrak{C}$. Those coordinates fixed, we have to make for all $k \in K$ the choice $b = 0$ or $b = 1$, i.e. $2^{|K|}$ choices. Since we are counting unordered pairs, we have $2^{|K|-1+d|I|(4-|K|)}$ elements in $\mathfrak{C}$. Obviously, the exponent $|K| - 1 + d|I|(4 - |K|)$ is minimal for $|K| = 4$. We deduce that

$$|K| - 1 + d|I|(4 - |K|) \geq 3 \,,$$

leading to $|\mathfrak{C}| \equiv 0 \mod 8$. □

By combining Proposition 1 and Theorem 4, we deduce the following corollary which generalises Lemma 2 in the sense that we are not restricted to the case $|I| = 1$ as in Lemma 2.

Corollary 1. *Let $n = \#\{\ \{p^0, p^1\}$ with $p^0, p^1 \in \mathcal{M}_I + a \mid \mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{D}_J\}$. Then, $n \equiv 0 \mod 8$.*

Proof. Let $\mathcal{P}^2(\mathcal{M}_I + a)$ denote the set of all unordered pairs of elements in $\mathcal{M}_I + a$, and $\mathcal{P}^2(\mathcal{M}_I + a)/\sim$ denote the set of all equivalence classes for $\sim$. Since the equivalence classes form a partition of $\mathcal{P}^2(\mathcal{M}_I + a)$, we have that

$$n = \left|\Delta^{-1}(\mathcal{D}_J)\right| = \sum_{\mathfrak{C} \in \mathcal{P}^2(\mathcal{M}_I+a)/\sim} \left|\Delta^{-1}(\mathcal{D}_J) \cap \mathfrak{C}\right| .$$

We know from Theorem 4 that Δ is constant on the equivalence classes, implying that $\left(\Delta^{-1}(\mathcal{D}_J) \cap \mathfrak{C}\right)$ equals either 0 or $|\mathfrak{C}|$. In other words, there exists a function δ from $\mathcal{P}^2(\mathcal{M}_I + a)/\sim$ into $\{0, 1\}$ such that

$$\left|\Delta^{-1}(\mathcal{D}_J) \cap \mathfrak{C}\right| = \delta(\mathfrak{C}) \times |\mathfrak{C}| \,.$$

It follows that

$$n = \sum_{\mathfrak{C} \in \mathcal{P}^2(\mathcal{M}_I+a)/\sim} \delta(\mathfrak{C}) \times |\mathfrak{C}| \equiv 0 \mod 8,$$

since, by Proposition 1, all equivalence classes have a cardinality divisible by 8. □

4.3 Influence of the branch number

In [GRR17], the proof of Lemma 2 uses the fact that the differential branch number of MC denoted by b is maximal and equal to 5 but as we have seen, the branch number b does not have any importance for this lemma. It affects the exact value of n, but not the fact that it is a multiple of 8.

Indeed, in the formula given in the proof of Corollary 1 for computing n, we can distinguish between the different equivalent classes according to the size of their information set:

$$n = \sum_{\mathfrak{C}\in\mathcal{P}^2(\mathcal{M}_I+a)/\sim} \delta(\mathfrak{C}) \times |\mathfrak{C}| = \sum_{h=0}^{4} \sum_{\mathfrak{C}:|K(\mathfrak{C})|=h} \delta(\mathfrak{C}) \times |\mathfrak{C}| \ .$$

Obviously, for any $\{p^0, p^1\}$ with information set K,

$$p^0 + p^1 = \sum_{k=0}^{3}\sum_{i\in I}(p^0_{i,k} + p^1_{i,k})\mathsf{MC}(e_{i-k,k}) = \sum_{k\in K}\sum_{i\in I}(p^0_{i,k} + p^1_{i,k})\mathsf{MC}(e_{i-k,k}) \in \mathcal{C}_K,$$

implying, by Lemma 1, that $(\mathcal{R}(p^0) + \mathcal{R}(p^1))$ belongs to $\mathcal{M}_K$.

However, it has been proved in [GRR16, Lemma 5] that, for any two sets $I, J \subseteq \{0, \ldots, 3\}$ such that $|I| + |J| < b$ where b is the branch number of MC, we have $\mathcal{D}_I \cap \mathcal{M}_J = \{0\}$. It then follows that, if K has size $h < b - |J|$, then $\mathcal{R}(p^0) + \mathcal{R}(p^1) \notin \mathcal{D}_J$ unless $p^0 = p^1$. We can then express the influence of the branch number on n with the formula

$$n = \sum_{h=b-|J|}^{4} \sum_{\mathfrak{C}:|K(\mathfrak{C})|=h} \delta(\mathfrak{C}) \times |\mathfrak{C}| \ ,$$

which does not affect the multiple-of-8 property.

4.4 An alternative proof of Theorem 3

The multiple-of-8 property is a consequence of Theorem 4, which states that the function Δ is constant over each equivalence class. However, this invariance can be directly used as a distinguishing property. This is actually what is done by the second *mixture-differential distinguisher* exhibited in [Gra18] and detailed in Theorem 3: Theorem 3 is nothing else than the combination of Theorem 4 with the subspace trail given by Lemma 1. Indeed, consider p^0 and p^1 in $\mathcal{C}_i + a$ with $p^0 \equiv (x_0, x_1, x_2, x_3)$ and $p^1 \equiv (y_0, y_1, y_2, y_3)$. Then, from Lemma 1, $\mathcal{R}(p^0)$ and $\mathcal{R}(p^1)$ belong to the same coset of $\mathcal{M}_i$ and their decompositions over the basis $(\mathsf{MC}(e_{i,0}), \mathsf{MC}(e_{i-1,1}), \mathsf{MC}(e_{i-2,2}), \mathsf{MC}(e_{i-3,3}))$ are given by

$$\begin{aligned} \mathcal{R}(p^0) &\equiv (\text{S-box}(x_0 + a_{0,i}), \ldots, \text{S-box}(x_3 + a_{3,i})) \\ \mathcal{R}(p^1) &\equiv (\text{S-box}(y_0 + a_{0,i}), \ldots, \text{S-box}(y_3 + a_{3,i})) \ . \end{aligned}$$

It follows that, if the two pairs $\{p^0, p^1\}$ and $\{q^0, q^1\}$ satisfy one of the relations given in Theorem 3, then $\{\mathcal{R}(p^0), \mathcal{R}(p^1)\}$ and $\{\mathcal{R}(q^0), \mathcal{R}(q^1)\}$ belong to the same equivalence class for $\sim$. We then deduce from Theorem 4 that Δ takes the same value on these two pairs, i.e.,

$$\mathcal{R}^2(p^0) + \mathcal{R}^2(p^1) = \mathcal{R}^2(q^0) + \mathcal{R}^2(q^1) \ .$$

Consequently,

$$\mathcal{R}^2(p^0) + \mathcal{R}^2(p^1) \in \mathcal{D}_J \text{ if and only if } \mathcal{R}^2(q^0) + \mathcal{R}^2(q^1) \in \mathcal{D}_J.$$

Moreover, we know from Lemma 1 that $\mathcal{D}_J \overset{\mathcal{R}}{\rightrightarrows} \mathcal{C}_J \overset{\mathcal{R}}{\rightrightarrows} \mathcal{M}_J$, implying that

$$\mathcal{R}^4(p^0) + \mathcal{R}^4(p^1) \in \mathcal{M}_J \text{ if and only if } \mathcal{R}^4(q^0) + \mathcal{R}^4(q^1) \in \mathcal{M}_J.$$

5 Adaptation to a general SPN construction

We provide in this section an extensive version of the equivalence relation and the multiple-of-8 property for a more general SPN cipher than the AES. A natural question for this extension is to find out what is the particular property of the subspaces $\mathcal{M}_I$ for Theorem 4 and Lemma 2 to work and whether these spaces could be replaced by others without altering the result. For a general SPN cipher, we analyse the form of such subspaces with respect to the non-linear layer of the cipher and provide the necessary conditions for their successful combination.

5.1 A more general setting for Theorem 4 and Lemma 2

Consider a general SPN cipher working on a state of N words, where the size of each word equals the cipher's S-box size. Suppose that the cipher is iterated for an arbitrary number of rounds and that the round-keys as well as the internal state are represented as word-vectors in $\mathbb{K}^N$ (and not as matrices). An SPN round $\mathcal{R}$ is the composition $\mathcal{K} \circ \mathcal{L} \circ \mathcal{S}$ where:

- $\mathcal{S}$ is the substitution layer applying an invertible S-box : $\mathbb{K} \to \mathbb{K}$ to each word of the internal state in a certain basis $\{f_0, \ldots, f_{N-1}\}$ of $\mathbb{K}^N$. It is important to notice that we define $\mathcal{S}$ and the basis together.
- $\mathcal{L}$ is the linear layer, a bijective $\mathbb{F}_2$-linear map of $\mathbb{K}^N$.
- $\mathcal{K}$ is the `AddRoundKey` operation adding to the internal state a round-key of the same size.

We now want to describe a more general subspace V of $\mathbb{K}^N$ that could play the role of $\mathcal{M}_I$ in an adaptation of Theorem 4 to the previously described SPN. As we have already noticed in the previous section, the proof of Theorem 4 relies on the fact that the coordinates of the elements in the input subspace V can be decomposed into several subsets in such a way that each S-box involves only one coordinate subset. This property is captured by the following definition: it requires the existence of a basis of V which can be decomposed over the original basis $\{f_0, \ldots, f_{N-1}\}$ by a block-diagonal matrix.

Definition 5. Let V be a subspace of $\mathbb{K}^N$. We say that V is *compatible with* $\mathcal{S}$ if there exists a basis of V whose elements written in the basis $\{f_0, \ldots, f_{N-1}\}$ form a block-diagonal matrix (with blocks having potentially different dimensions) for a certain order of the elements in both bases. We call such a basis of V a *compatibility basis*.

Given an arbitrary basis of a subspace V, it is quite easy to check whether V is compatible with $\mathcal{S}$ by computing the unique reduced echelon form of the corresponding matrix. If, for a given ordering of the rows, this matrix has a reduced echelon form which is block-diagonal, then V is compatible with $\mathcal{S}$ and the reduced echelon form provides a compatibility basis. Otherwise, V is not compatible with $\mathcal{S}$.

We provide now the necessary notation for describing a compatibility basis g of a compatible subspace V. For this notation to be as clear as possible, we first give a representation of g as a collection of column vectors written in the basis f.

$$
\begin{array}{c}
\left(\begin{array}{ccccccccc}
* & \cdots & * & & & & & & \\
\vdots & \lambda_{0,\ell,i} & \vdots & & 0 & & & 0 & \\
* & \cdots & * & & & & & & \\
 & & & * & \cdots & * & & & \\
 & 0 & & \vdots & \lambda_{k,\ell,i} & \vdots & & 0 & \\
 & & & * & \cdots & * & & & \\
 & & & & & & * & \cdots & * \\
 & 0 & & & 0 & & \vdots & \lambda_{h-1,\ell,i} & \vdots \\
 & & & & & & * & \cdots & * \\
 & 0 & & & 0 & & & 0 &
\end{array}\right) \\
\begin{array}{ccc}
\uparrow & \uparrow & \uparrow \\
g_{0,i} & g_{k,i} & g_{h-1,i} \\
i < i_0 & i < i_k & i < i_{h-1}
\end{array}
\end{array}
\qquad
\begin{array}{l}
\leftarrow \text{ coordinate on } f_{j_0+\ell} \\
\\
\leftarrow \quad j_k + \ell \\
\\
\leftarrow \quad j_{h-1} + \ell
\end{array}
$$

- We denote by h the number of blocks. The number of basis vectors in the k-th block, $0 \leq k < h$ will be denoted by i_k. It must obviously hold that $\sum_{k=0}^{h-1} i_k = \dim V$.
- The basis of V will be denoted by $(g_{k,i})_{k<h,i<i_k} \in (\mathbb{K}^N)^{\dim V}$, which means that $V = \mathrm{vect}_{\mathbb{K}}(g_{k,i} \mid k < h, i < i_k)$. The index k of a basis element $g_{k,i}$ stands for the block-number, while i represents the position of the vector inside the block k.
- There exist $(h+1)$ integers $j_0, \ldots, j_h$, with $j_0 = 0$ and $j_h \leq N$, such that for all vectors inside the k-th block, all coordinates outside the interval $\{j_k, \ldots, j_{k+1} - 1\}$ are zero.

Then, each basis vector $g_{k,i}$ can be written as

$$g_{k,i} = \sum_{\ell=0}^{j_{k+1}-j_k-1} \lambda_{k,\ell,i} f_{j_k+\ell} \tag{3}$$

for some $\lambda_{k,\ell,i} \in \mathbb{K}$ with $0 \leq k < h$ and $0 \leq i < i_k$.

Example 2. For the AES, we have $N = 16$ and the original basis $\{f_0, \ldots, f_{15}\}$ is formed by the vectors $e_{i,j}$ with $i, j \in \{0, \ldots, 3\}$. Let $I = \{t_0, \ldots, t_{r-1}\}$ be a subset of size r of $\{0, \ldots, 3\}$. Then, $\mathcal{M}_I$ is compatible with the AES S-box layer. A compatibility basis of $\mathcal{M}_I$ is given by

$$g_{k,i} = \mathsf{MC}(e_{t_i-k,k}) \text{ for } 0 \leq k < 4 \text{ and } 0 \leq i < r \ .$$

Indeed, when the elements of the original basis are ordered by $f_{4j+i} = e_{i,j}$ for $i, j \in \{0, \ldots, 3\}$, $(g_{0,0}, \ldots, g_{3,r-1})$ is obtained by multiplying $(f_0, \ldots, f_{15})$ by a matrix with $h = 4$ blocks, all of size r. This comes from the fact that, for $j_k = 4k$, we have

$$g_{k,i} = \mathsf{MC}(e_{t_i-k,k}) = \sum_{\ell=0}^{3} m_{\ell,t_i-k} e_{\ell,k} = \sum_{\ell=0}^{j_{k+1}-j_k-1} m_{\ell,t_i-k} f_{j_k+\ell} \ ,$$

where $m_{i,j}$ are the coefficients of the 4×4 matrix defining `MixColumns`. This is exactly the block-diagonal form described by (3).

For example, a basis of $\mathcal{M}_0$ can be written as follows, where . corresponds to 0.

$$\begin{pmatrix} 2 & . & . & . \\ 1 & . & . & . \\ 1 & . & . & . \\ 3 & . & . & . \\ . & 1 & . & . \\ . & 1 & . & . \\ . & 3 & . & . \\ . & 2 & . & . \\ . & . & 1 & . \\ . & . & 3 & . \\ . & . & 2 & . \\ . & . & 1 & . \\ . & . & . & 3 \\ . & . & . & 2 \\ . & . & . & 1 \\ . & . & . & 1 \end{pmatrix}$$

From now on, we fix $a \in \mathbb{K}^N$ and a subspace V compatible with $\mathcal{S}$ with compatibility basis g. The notion of information set and the equivalence relation between pairs of elements in $(a+V)$ now involve the decomposition of the elements in V over the basis $\{g_{k,i}\}$, where the coordinates corresponding to the same k are gathered together. The next definition adapts the notion of information set to this context.

Definition 6 (Information set)**.** Let $\{p^0, p^1\}$ be an unordered pair of elements from $V+a$, written as

$$p^0 = \sum_{k=0}^{h-1} \sum_{i=0}^{i_k-1} p^0_{i,k} g_{i,k} + a \quad \text{and} \quad p^1 = \sum_{k=0}^{h-1} \sum_{i=0}^{i_k-1} p^1_{i,k} g_{i,k} + a$$

for some (uniquely defined) $p^0_{i,k}, p^1_{i,k} \in \mathbb{K}$. The *information set* K of $\{p^0, p^1\}$ is defined as

$$K = \{k \in \{0, \ldots, h-1\} \mid \exists i < i_k : p^0_{i,k} \neq p^1_{i,k}\} .$$

Similarly, we define an equivalence relation between pairs of inputs by considering all pairs of elements in $(a+V)$ obtained by exchanging the sets of coordinates corresponding to the same k.

Definition 7 (Equivalence relation)**.** Let $P = \{p^0, p^1\}$ and $Q = \{q^0, q^1\}$ with $p^0, p^1, q^0, q^1 \in (V+a)$. We say that $P \sim Q$ if:

- $\{p^0, p^1\}$ and $\{q^0, q^1\}$ have the same information set K.
- $\forall k \in K, \exists b \in \{0,1\} : \forall i < i_k, q^0_{i,k} = p^b_{i,k}$ and $q^1_{i,k} = p^{1-b}_{i,k}$.

$\sim$ is an equivalence relation on the set of unordered pairs of elements in $(V+a)$.

The next theorem is an adaptation of Theorem 4 to any subspace V compatible with $\mathcal{S}$.

Theorem 5. *For any $a \in \mathbb{K}^N$, the function Δ operating on unordered pairs of elements in $(V+a)$ and defined by*

$$\Delta : \quad \{p^0, p^1\} \quad \longmapsto \quad \mathcal{R}(p^0) + \mathcal{R}(p^1)$$

is constant over the equivalence classes for $\sim$.

Proof. Let $P = \{p^0, p^1\}, Q = \{q^0, q^1\}$ such that $P \sim Q$. We have with the previously introduced notation

$$p^0 = \sum_{k=0}^{h-1} \sum_{i=0}^{i_k-1} p^0_{i,k} g_{i,k} + a = \sum_{k=0}^{h-1} \sum_{\ell=0}^{j_{k+1}-j_k-1} \left(\sum_{i=0}^{i_k-1} p^0_{i,k} \lambda_{i,k,\ell} + a_{j_k+\ell} \right) f_{j_k+\ell}$$

where the last equality is obtained by replacing each $g_{i,k}$ by its decomposition on the basis $(f_0, \ldots, f_{N-1})$ given by (3). Then

$$\mathcal{S}(p^0) = \sum_{k=0}^{h-1} \sum_{\ell=0}^{j_{k+1}-j_k-1} \text{S-box} \left(\sum_{i=0}^{i_k-1} p^0_{i,k} \lambda_{i,k,\ell} + a_{j_k+\ell} \right) f_{j_k+\ell}$$

and the difference $\mathcal{S}(p^0) + \mathcal{S}(p^1)$ can be written as:

$$\sum_{k,\ell} \left[\text{S-box} \left(\sum_{i=0}^{i_k-1} p^0_{i,k} \lambda_{i,k,\ell} + a_{j_k+\ell} \right) + \text{S-box} \left(\sum_{i=0}^{i_k-1} p^1_{i,k} \lambda_{i,k,\ell} + a_{j_k+\ell} \right) \right] f_{j_k+\ell} \tag{4}$$

It is now clear with Definition 7 and Equation (4) that $\mathcal{S}(p^0) + \mathcal{S}(p^1)$ and $\mathcal{S}(q^0) + \mathcal{S}(q^1)$ are equal in $\mathbb{K}^N$. Therefore,

$$\begin{aligned}
\Delta(P) &= \mathcal{R}(p^0) + \mathcal{R}(p^1) \\
&= \mathcal{K} \circ \mathcal{L} \circ \mathcal{S}(p^0) + \mathcal{K} \circ \mathcal{L} \circ \mathcal{S}(p^1) \\
&= \mathcal{L}(\mathcal{S}(p^0) + \mathcal{S}(p^1)) \quad \text{by characteristic 2 and linearity of } \mathcal{L} \text{ ;} \\
&= \mathcal{L}(\mathcal{S}(q^0) + \mathcal{S}(q^1)) \quad \text{by our previous observation ;} \\
&= \Delta(Q).
\end{aligned}$$

□

From this invariance theorem, we can derive, as in the case of the AES distinguisher, some information on the number of unordered pairs $\{p^0, p^1\}$ such that $\Delta(\{p^0, p^1\})$ belongs to a given set $\mathcal{E}$. This consequence is similar to the multiple-of-8 property, but the divisibility of this number depends on the structure of the subspace V we consider. This comes from the sizes of the equivalence classes, which are determined in the following proposition.

Proposition 2. *Let $\mathfrak{C}$ be an equivalence class with information set K. The cardinality of $\mathfrak{C}$ is*

$$|\mathfrak{C}| = 2^{|K|-1+d\sum_{k \notin K} i_k}.$$

It is always a multiple of 2^{h-1}.

Proof. We have $\prod_{k \notin K} (2^d)^{i_k}$ choices for the shared coordinates in a pair of $\mathfrak{C}$. Those coordinates fixed, we have to make for all $k \in K$ the choice $b = 0$ or $b = 1$, i.e. $2^{|K|}$ choices. Since we are counting unordered pairs, we have $2^{|K|-1+d\sum_{k \notin K} i_k}$ pairs in $\mathfrak{C}$. The exponent $\left(|K| - 1 + d\sum_{k \notin K} i_k\right)$ is minimal for $|K| = h$. Indeed,

$$\left(|K| - 1 + d \sum_{k \notin K} i_k \right) = \left(d \cdot \dim V - 1 - \sum_{k \in K} (d \cdot i_k - 1) \right)$$

which obviously decreases as K gets bigger. Hence $|\mathfrak{C}| \equiv 0 \mod 2^{h-1}$. □

We then deduce the following generalisation of the multiple-of-8 property.

Corollary 2. *Let $\mathcal{E}$ be any subset of $\mathbb{K}^N$ and*

$$n = \#\{\ \{p^0, p^1\}\ \textit{with}\ p^0, p^1 \in (V + a)\ |\ \mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{E}\}.$$

Then $n \equiv 0 \mod 2^{h-1}$.

The proof is the same as the proof of Corollary 1, but we detail it for the sake of completeness.

Proof. We denote by $\mathcal{P}^2(V + a)$ the set of all unordered pairs of elements in $V + a$ and by $\mathcal{P}^2(\mathcal{M}_I + a)/\sim$ the set of all equivalence classes for $\sim$. Since the equivalence classes form a partition of $\mathcal{P}^2(\mathcal{M}_I + a)$, we have that

$$n = \left|\Delta^{-1}(\mathcal{E})\right| = \sum_{\mathfrak{C} \in \mathcal{P}^2(\mathcal{M}_I + a)/\sim} \left|\Delta^{-1}(\mathcal{E}) \cap \mathfrak{C}\right|.$$

We know from Theorem 4 that Δ is constant on the equivalence classes, implying that

$$\left|\Delta^{-1}(\mathcal{E}) \cap \mathfrak{C}\right| = \delta(\mathfrak{C}) \times |\mathfrak{C}|$$

for some function δ from $\mathcal{P}^2(\mathcal{M}_I + a)/\sim$ into $\{0, 1\}$. It follows that

$$n = \sum_{\mathfrak{C} \in \mathcal{P}^2(\mathcal{M}_I + a)/\sim} \delta(\mathfrak{C}) \times |\mathfrak{C}| \equiv 0 \mod 2^{h-1},$$

since, by Proposition 2, all equivalence classes have a cardinality divisible by 2^{h-1}. □

5.2 A new proof of Theorem 2

As a first illustration, we now show that the mixture-differential distinguisher described by Theorem 2 and originally stated in [Gra18] can be seen as a direct application of Theorem 5. In this case, the compatible subspace V is $\mathrm{vect}_{\mathbb{K}}(\mathsf{MC}(e_{0,i}), \mathsf{MC}(e_{1,i})) = \mathcal{M}_i \cap \mathcal{C}_{0,1}$ with $i \in \{0, \ldots, 3\}$. The proof of Theorem 2 is then similar to the one presented in Section 4.4, but for a different subspace V. We fix $a \in \mathcal{M}_4(\mathbb{K})$ for the rest of this section.

Proof. We define V as the subspace $\mathrm{vect}_{\mathbb{K}}(\mathsf{MC}(e_{0,i}), \mathsf{MC}(e_{1,i}))$. A basis of this subspace is composed of the two column vectors $g_{0,0} = \mathsf{MC}(e_{0,i})$ and $g_{1,0} = \mathsf{MC}(e_{1,i})$, whose decomposition over the canonical basis is defined by:

$$\begin{pmatrix} 2 & . \\ 1 & . \\ 1 & . \\ 3 & . \\ . & 1 \\ . & 1 \\ . & 3 \\ . & 2 \\ . & . \\ . & . \\ . & . \\ . & . \\ . & . \\ . & . \\ . & . \\ . & . \end{pmatrix}$$

which justifies that V is compatible with the AES substitution layer. Let $p^0, p^1, q^0, q^1 \in \text{vect}_{\mathbb{K}}(e_{0,i}, e_{1,i}) + a$ such that

$$p^0 \equiv (x_0, x_1), \quad p^1 \equiv (y_0, y_1) \text{ and } q^0 \equiv (x_0, y_1), \quad q^1 \equiv (y_0, x_1) \ .$$

Then,

$$\mathcal{R}(p^0) \equiv (\text{S-box}(x_0{+}a_{0,i}), \text{S-box}(x_1{+}a_{1,i})) \text{ and } \mathcal{R}(p^1) \equiv (\text{S-box}(y_0{+}a_{0,i}), \text{S-box}(y_1{+}a_{1,i})) \ .$$

This exactly means that $\{\mathcal{R}(p^0), \mathcal{R}(p^1)\}$ and $\{\mathcal{R}(q^0), \mathcal{R}(q^1)\}$ are equivalent pairs of a coset of V. Theorem 5 then gives that $\mathcal{R}^2(p^0) + \mathcal{R}^2(p^1) = \mathcal{R}^2(q^0) + \mathcal{R}^2(q^1)$. Consequently,

$$\mathcal{R}^2(p^0) + \mathcal{R}^2(p^1) \in \mathcal{D}_J \text{ if and only if } \mathcal{R}^2(q^0) + \mathcal{R}^2(q^1) \in \mathcal{D}_J.$$

Since $\mathcal{D}_J \overset{2\mathcal{R}}{\rightrightarrows} \mathcal{M}_J$, this equivalently means that

$$\mathcal{R}^4(p^0) + \mathcal{R}^4(p^1) \in \mathcal{M}_J \text{ if and only if } \mathcal{R}^4(q^0) + \mathcal{R}^4(q^1) \in \mathcal{M}_J.$$

□

Obviously the same result also holds for any subspace V formed by the intersection between $\mathcal{M}_i$ and another subset (unless the matrix defining the corresponding basis has a single block).

6 Applications

In this section we provide some applications of Theorem 5 to SPN ciphers other than the AES. The goal is to show in practice that the mixture-differential distinguishers and the *multiple-of* property are not proper to the AES but that they hold for many other SPN constructions. Furthermore, as a result of the adaptation provided in the previous section, the application to other ciphers is almost straightforward. Therefore, we adapt Theorem 1 to the SPN ciphers LED [GPPR11], Midori [BBI+15], Klein [GNL12] and Skinny [BJK+16] and discuss why the result does not adapt to Crypton [Lim99] or Prince [BCG+12].

In practice, for finding a *multiple-of* property for some cipher, two conditions must be met. The first one is the existence of a subspace V compatible with the substitution layer for the *multiple-of* property through one round to hold. This condition is described through Corollary 2, which can be seen as a general replacement of Lemma 2. The second condition is the existence of exact subspace trails covering some rounds before and after the central round on which the *multiple-of* property is found. For this, we need a result equivalent to Lemma 1 for each analysed cipher. For searching for subspace trails that can replace the subspaces $\mathcal{D}_I, \mathcal{C}_I$ and $\mathcal{M}_I$ in Lemma 1, we use Algorithm 1 proposed by Leander, Tezcan and Wiemer in [LTW18].

Midori, Klein, Skinny and Crypton all follow our general SPN description and operate on a state composed of $N = 16$ words, where the size of each word (i.e. the dimension of the S-box alphabet) is $d \in \{4, 8\}$. LED and a part of Prince also follow this general description with $d = 4$. We study for each cipher the exact subspace trails given by Algorithm 1 in [LTW18] and the compatibility of the last subspace of the trail with the substitution layer in the sense of Definition 5.

6.1 The cases of AES, LED, Midori, Klein and Skinny

AES. First of all, a natural idea would be to use our adaptation with a longer subspace trail than the ones of Lemma 1 for the AES. However, as shown in [LTW18], the exact trails $\mathcal{D}_I \overset{\mathcal{R}}{\rightrightarrows} \mathcal{C}_I \overset{\mathcal{R}}{\rightrightarrows} M_I$ are the longest possible trails for this cipher, which means that Corollary 2 cannot give any improvement for the AES.

LED. LED is a lightweight block cipher proposed by Guo et al. [GPPR11]. Its round function has the same structure as the AES and exhibits the same two-round subspace trails $\mathcal{D}_I \overset{\mathcal{R}}{\rightrightarrows} \mathcal{C}_I \overset{\mathcal{R}}{\rightrightarrows} \mathcal{M}_I$, leading to a 5-round distinguisher.

Midori. Midori is a lightweight block cipher designed by Banik et al. [BBI+15], optimized with respect to the energy consumption. The round function $\mathcal{R}_{\mathsf{Mi}}$ is the composition $\mathcal{L}_{\mathsf{Mi}} \circ \mathcal{S}_{\mathsf{Mi}}$, where $\mathcal{L}_{\mathsf{Mi}} = \mathsf{MC}_{\mathsf{Mi}} \circ \mathsf{SC}$. $\mathsf{MC}_{\mathsf{Mi}}$ is the `MixColumns` operation, applying the binary involutive matrix

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

to the columns of the state. The branch number of the above matrix is 4. SC is the so-called `ShuffleCell` permutation that permutes the words of the state as follows:

$$(s_0, s_1, \ldots, s_{15}) \leftarrow (s_0, s_7, s_{14}, s_9, s_5, s_2, s_{11}, s_{12}, s_{15}, s_8, s_1, s_6, s_{10}, s_{13}, s_4, s_3)$$

where the words are numbered column-wise. As in Section 2, we define the following subspaces depending on the linear components of Midori's round-function. If $i \in \{0, 1, 2, 3\}$,

$$\begin{array}{rcl} \mathcal{C}_i & = & \mathrm{vect}_{\mathbb{K}}(e_{0,i}, e_{1,i}, e_{2,i}, e_{3,i}), \\ \mathcal{D}_i^{\mathsf{Mi}} & = & \mathsf{SC}^{-1}(\mathcal{C}_i), \\ \mathcal{M}_i^{\mathsf{Mi}} & = & \mathcal{L}_{\mathsf{Mi}}(\mathcal{C}_i). \end{array}$$

Applying Algorithm 1 in [LTW18] gives that the longest exact subspace trails are the two-round trails of the form $\mathcal{D}_I^{\mathsf{Mi}} \overset{\mathcal{R}_{\mathsf{Mi}}}{\rightrightarrows} \mathcal{C}_I \overset{\mathcal{R}_{\mathsf{Mi}}}{\rightrightarrows} \mathcal{M}_I^{\mathsf{Mi}}$. Besides, $\mathcal{M}_I^{\mathsf{Mi}}$ has a basis whose matrix is block-diagonal with four blocks. For example, $\mathcal{M}_0^{\mathsf{Mi}}$ has a basis whose representation as a collection of column vectors is given in Figure 1a. We then have by Corollary 2 that for all $I, J \subseteq \{0, 1, 2, 3\}, a \in (\mathbb{F}_{2^d})^{16}$,

$$\#\{\{p^0, p^1\} \text{ with } p^0, p^1 \in \mathcal{D}_I^{\mathsf{Mi}} + a \mid \mathcal{R}_{\mathsf{Mi}}^5(p^0) + \mathcal{R}_{\mathsf{Mi}}^5(p^1) \in \mathcal{M}_J^{\mathsf{Mi}}\} \equiv 0 \mod 8,$$

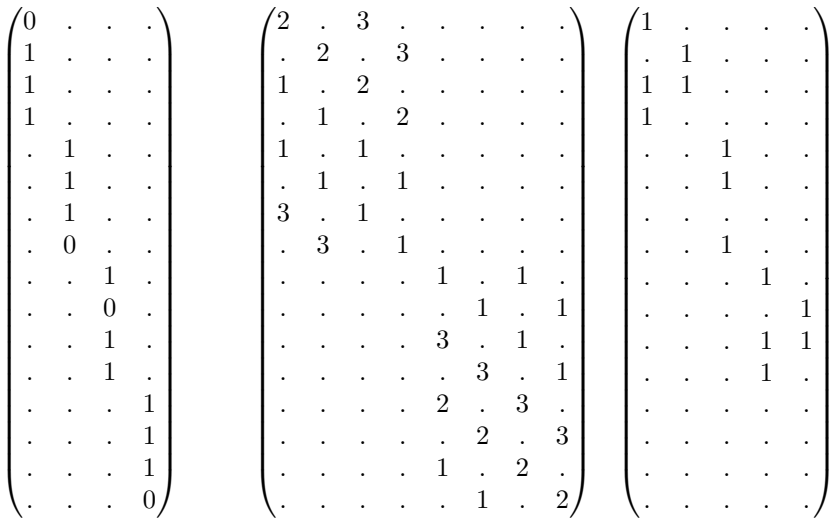
which gives a 5-round distinguisher similar to the AES-one. It is worth noticing that the property holds even if the branch number of the `MixColumns` operation in Midori is only 4.

Klein. Klein is a lightweight block cipher proposed in 2011 by Gong et al. [GNL12]. The round function of Klein $\mathcal{R}_{\mathsf{Kl}} = \mathcal{L}_{\mathsf{Kl}} \circ \mathcal{S}_{\mathsf{Kl}}$ can be seen as the application of a non-linear layer $\mathcal{S}_{\mathsf{Kl}}$ followed by a linear layer $\mathcal{L}_{\mathsf{Kl}}$. The linear layer $\mathcal{L}_{\mathsf{Kl}} = \mathsf{MN} \circ \mathsf{RN}$ is the composition of RN, standing for the `RotateNibbles` permutation rotating the state two bytes to the left and MN, standing for the `MixNibbles` permutation. This last operation applies the AES `MixColumns` transformation to each half of the state. We denote the canonical basis of $\mathbb{F}_{2^4}^{16}$ by $(f_i)_{0 \le i < 16}$ and then define the following subspaces for $i \in \{0, 1\}$:

$$\begin{array}{rcl} \mathcal{C}_i & = & \mathrm{vect}_{\mathbb{K}}(f_k \mid 0 \le k < 8), \\ \mathcal{D}_i^{\mathsf{Kl}} & = & \mathsf{RN}^{-1}(\mathcal{C}_i), \\ \mathcal{M}_i^{\mathsf{Kl}} & = & \mathcal{L}_{\mathsf{Kl}}(\mathcal{C}_i). \end{array}$$

Algorithm 1 in [LTW18] gives that the longest exact subspace trails are two-round trails of the form

$$\mathcal{D}_i^{\mathsf{Kl}} \overset{\mathcal{R}_{\mathsf{Kl}}}{\rightrightarrows} \mathcal{C}_i \overset{\mathcal{R}_{\mathsf{Kl}}}{\rightrightarrows} \mathcal{M}_i^{\mathsf{Kl}} .$$



(a) $\mathcal{M}_0^{\mathsf{Mi}}$ basis **(b)** $\mathcal{M}_0^{\mathsf{Kl}}$ basis **(c)** W_0 basis

Figure 2: Compatibility basis representations for Midori (left) and Klein (right)

Besides, $\mathcal{M}_i^{\mathsf{Kl}}$ has a basis whose matrix is block-diagonal with two blocks. For example, $\mathcal{M}_0^{\mathsf{Kl}}$ has a basis whose representation as a collection of column vectors is given in Figure 1b. We then have by Corollary 2 for all $i, j \in \{0,1\}, a \in (\mathbb{F}_{2^4})^{16}$,

$$\#\{\{p^0, p^1\} \text{ with } p^0, p^1 \in \mathcal{D}_i^{\mathsf{Kl}} + a \mid \mathcal{R}_{\mathsf{Kl}}^5(p^0) + \mathcal{R}_{\mathsf{Kl}}^5(p^1) \in \mathcal{M}_j^{\mathsf{Kl}}\} \equiv 0 \mod 2\,,$$

meaning that Klein has a *multiple-of-2* property for 5 rounds. Note that even if we get only a multiple of 2 in the case of Klein, as the pairs are not ordered, this can still be considered as a distinguishing property.

Skinny. Skinny is a family of tweakable lightweight block ciphers, designed in 2016 by Beierle et al. [BJK+16]. The round function follows a classical SPN construction $\mathcal{R}_{\mathsf{Sk}} = \mathcal{L}_{\mathsf{Sk}} \circ \mathcal{S}_{\mathsf{Sk}}$ where $\mathcal{L}_{\mathsf{Sk}} = \mathsf{MC}_{\mathsf{Sk}} \circ \mathsf{SR}_{\mathsf{Sk}}$. The operation $\mathsf{SR}_{\mathsf{Sk}}$ is similar to the AES `ShiftRows` operation, with the only difference that the shift is performed to the right. $\mathsf{MC}_{\mathsf{Sk}}$ is the `MixColumns` operation, where each column of the state is multiplied by the following binary matrix of branch number 2:

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Applying here again Algorithm 1 of [LTW18] gives as result that the longest exact subspace trails are two-round-long. There are 1294 such trails. A Gauss elimination on the last subspace of each trail gives that among these trails, 1282 end with a subspace compatible with the substitution layer. This allows to adapt Theorem 1 for 5-round Skinny, concluding that 5-round Skinny always has the *multiple-of-*2^{h-1} property. However, it is interesting to note here that depending on the trail, the value of h varies. More precisely, $h \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14\}$.

We now give an example of such a distinguisher for $h = 3$. For Skinny, the internal state is classically represented as an element of $\mathcal{M}_4(\mathbb{K})$. We first exhibit two 2-round

subspace trails, $U_i \overset{\mathcal{R}_{\mathsf{Sk}}}{\rightrightarrows} V_i \overset{\mathcal{R}_{\mathsf{Sk}}}{\rightrightarrows} W_i$ for $i \in \{0,1\}$ where

$$U_0 = \mathrm{vect}_{\mathbb{K}}(e_{1,1}, e_{1,2}, e_{1,3}, e_{3,1}, e_{3,3}),\ V_0 = \mathcal{L}_{\mathsf{Sk}}(U_0),\ W_0 = \mathcal{L}_{\mathsf{Sk}}(V_0)$$

and

$$U_1 = \mathrm{vect}_{\mathbb{K}}(e_{0,3}, e_{1,0}, e_{1,2}, e_{1,3}, e_{2,1}, e_{2,3}, e_{3,0}, e_{3,1}, e_{3,2}, e_{3,3}),\ V_1 = \mathcal{L}_{\mathsf{Sk}}(U_1),\ W_1 = \mathcal{L}_{\mathsf{Sk}}(V_1)\ .$$

Moreover, W_0 is compatible with $h = 3$. One of its compatibility basis is represented in Figure 1c. We then have that

$$\#\{\{p^0, p^1\} \text{ with } p^0, p^1 \in U_0 + a \mid \mathcal{R}^5_{\mathsf{Sk}}(p^0) + \mathcal{R}^5_{\mathsf{Sk}}(p^1) \in W_1\} \equiv 0 \mod 4\ .$$

Finally, we recall that this section only aims at giving examples of how Theorem 1 can be adapted to other SPN ciphers and does not claim new cryptanalytic results. In particular, this property on 5 rounds does not threaten the overall security of the cipher that is composed of 32 rounds.

6.2 The cases of Crypton and Prince

Crypton The block cipher Crypton [Lim98], designed by Lim in 1998 was among the candidates to the NIST AES competition. It has a structure very similar to the one of AES and this is why we considered it as a natural candidate for the *multiple-of* property. Indeed, the round function of Crypton is naturally decomposed as $\mathcal{R}_{\mathsf{Cr}} = \mathcal{L}_{\mathsf{Cr}} \circ \mathcal{S}_{\mathsf{Cr}}$, where $\mathcal{L}_{\mathsf{Cr}}$ is the composition of a byte transposition of columns into rows with respect to the anti-diagonal of the internal state and a permutation at the bit level applied column-wise. Algorithm 1 in [LTW18] gives two-round exact subspace trails. However, the problem here is that $\mathcal{L}_{\mathsf{Cr}}$ is only $\mathbb{F}_2$-linear and not $\mathbb{F}_{2^8}$-linear and this implies that the last subspaces in the subspace trails are not $\mathbb{F}_{2^8}$-vector subspaces and cannot verify the compatibility hypothesis of Corollary 2. As a consequence, Theorem 1 cannot be adapted to 5-round Crypton. However, Crypton has 1-round exact subspace trails that end with subspaces of the form $\mathrm{vect}_{\mathbb{K}}(e_{i,j} | i \in I, j \in \{0,1,2,3\})$. Those subspaces are obviously compatible with the substitution layer, and we have a 4-round version of Theorem 1 for Crypton: the first round is a 1-round exact subspace trail ending with a compatible subspace, the second round exploits Corollary 2 and the last two rounds are a 2-round exact subspace trail. Indeed, the trail used after Corollary 2 does not need to end with a compatible subspace.

Prince Prince [BCG+12] is a block cipher proposed by Borghoff et al. in 2012 with a specific structure named α-reflection. The main parts of this structure follow the SPN construction. As for Crypton, the round function is defined as $\mathcal{R}_{\mathsf{Pr}} = \mathcal{L}_{\mathsf{Pr}} \circ \mathcal{S}_{\mathsf{Pr}}$ with $\mathcal{L}_{\mathsf{Pr}}$ a $\mathbb{F}_2$-linear map which is not $\mathbb{F}_{2^4}$-linear. Again, as for Crypton, Prince exhibits two-round exact subspace trails but the last subspaces of those trails are not $\mathbb{F}_{2^4}$-linear subspaces. We then cannot mount a 5-round distinguisher by adapting Theorem 1. However, one-round exact subspace trails ending with compatible diagonal $\mathbb{F}_{2^4}$-linear subspaces allow to have a 4-round version of Theorem 1 for Prince.

7 Conclusion

We have presented a general result which allows cryptanalysts to search for mixture-differential distinguishers, or *multiple-of* properties, in a systematic way, for any SPN. This result then avoids the redundant proofs which were previously needed for each new occurrence of these distinguishing properties. Also, it highlights the properties of the ciphers which have to be taken into account for establishing the existence of such

distinguishers and it shows that mixture-differential distinguishers directly apply to a more general class of SPN than what was previously believed. As shown in the previous examples, all these distinguishing properties can be exhibited by combining our framework with the search for subspace trails, which is investigated in [LTW18]. Since our result, exploiting an appropriate equivalence relation, applies in many situations, it appears that the main limitation for finding efficient distinguishers is the existence of long subspace trails for the ciphers.

Acknowledgments

We would like to thank Hadi Soleimany and the anonymous reviewers for their helpful comments and suggestions.

References

[AES01] Advanced Encryption Standard (AES). National Institute of Standards and Technology (NIST), FIPS PUB 197, U.S. Department of Commerce, November 2001.

[BBI+15] Subhadeep Banik, Andrey Bogdanov, Takanori Isobe, Kyoji Shibutani, Harunaga Hiwatari, Toru Akishita, and Francesco Regazzoni. Midori: A block cipher for low energy. In Tetsu Iwata and Jung Hee Cheon, editors, *ASIACRYPT 2015, Part II*, volume 9453 of *LNCS*, pages 411–436. Springer, Heidelberg, November / December 2015.

[BCG+12] Julia Borghoff, Anne Canteaut, Tim Güneysu, Elif Bilge Kavun, Miroslav Knežević, Lars R. Knudsen, Gregor Leander, Ventzislav Nikov, Christof Paar, Christian Rechberger, Peter Rombouts, Søren S. Thomsen, and Tolga Yalçin. PRINCE - A low-latency block cipher for pervasive computing applications - extended abstract. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 208–225. Springer, Heidelberg, December 2012.

[BDK+18] Achiya Bar-On, Orr Dunkelman, Nathan Keller, Eyal Ronen, and Adi Shamir. Improved key recovery attacks on reduced-round AES with practical data and memory complexities. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO 2018, Part II*, volume 10992 of *LNCS*, pages 185–212. Springer, Heidelberg, August 2018.

[BJK+16] Christof Beierle, Jérémy Jean, Stefan Kölbl, Gregor Leander, Amir Moradi, Thomas Peyrin, Yu Sasaki, Pascal Sasdrich, and Siang Meng Sim. The SKINNY family of block ciphers and its low-latency variant MANTIS. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part II*, volume 9815 of *LNCS*, pages 123–153. Springer, Heidelberg, August 2016.

[BKR11] Andrey Bogdanov, Dmitry Khovratovich, and Christian Rechberger. Biclique cryptanalysis of the full AES. In Dong Hoon Lee and Xiaoyun Wang, editors, *ASIACRYPT 2011*, volume 7073 of *LNCS*, pages 344–371. Springer, Heidelberg, December 2011.

[Dae95] Joan Daemen. *Cipher and hash function design, strategies based on linear and differential cryptanalysis.* PhD thesis, K.U.Leuven, 1995.

[DF14] Patrick Derbez and Pierre-Alain Fouque. Exhausting Demirci-Selçuk meet-in-the-middle attacks against reduced-round AES. In Shiho Moriai, editor,

FSE 2013, volume 8424 of *LNCS*, pages 541–560. Springer, Heidelberg, March 2014.

[DF16] Patrick Derbez and Pierre-Alain Fouque. Automatic search of meet-in-the-middle and impossible differential attacks. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part II*, volume 9815 of *LNCS*, pages 157–184. Springer, Heidelberg, August 2016.

[DS08] Hüseyin Demirci and Ali Aydin Selçuk. A meet-in-the-middle attack on 8-round AES. In Kaisa Nyberg, editor, *FSE 2008*, volume 5086 of *LNCS*, pages 116–126. Springer, Heidelberg, February 2008.

[FKL+01] Niels Ferguson, John Kelsey, Stefan Lucks, Bruce Schneier, Michael Stay, David Wagner, and Doug Whiting. Improved cryptanalysis of Rijndael. In Bruce Schneier, editor, *FSE 2000*, volume 1978 of *LNCS*, pages 213–230. Springer, Heidelberg, April 2001.

[GNL12] Zheng Gong, Svetla Nikova, and Yee Wei Law. KLEIN: A new family of lightweight block ciphers. In Ari Juels and Christof Paar, editors, *RFIDSec 2011*, volume 7055 of *LNCS*, pages 1–18. Springer, Heidelberg, June 2012.

[GPPR11] Jian Guo, Thomas Peyrin, Axel Poschmann, and Matthew J. B. Robshaw. The LED block cipher. In Bart Preneel and Tsuyoshi Takagi, editors, *CHES 2011*, volume 6917 of *LNCS*, pages 326–341. Springer, Heidelberg, September / October 2011.

[Gra18] Lorenzo Grassi. Mixture differential cryptanalysis: a new approach to distinguishers and attacks on round-reduced AES. *IACR Trans. Symm. Cryptol.*, 2018(2):133–160, 2018.

[GRR16] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. Subspace trail cryptanalysis and its applications to AES. *IACR Trans. Symm. Cryptol.*, 2016(2):192–225, 2016. http://tosc.iacr.org/index.php/ToSC/article/view/571.

[GRR17] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. A new structural-differential property of 5-round AES. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *EUROCRYPT 2017, Part II*, volume 10211 of *LNCS*, pages 289–317. Springer, Heidelberg, April / May 2017.

[Lim98] Chae Hoon Lim. CRYPTON: a New 128-bit Block Cipher. AES Proposal, 1998.

[Lim99] Chae Hoon Lim. A revised version of Crypton - Crypton v1.0. In Lars R. Knudsen, editor, *FSE'99*, volume 1636 of *LNCS*, pages 31–45. Springer, Heidelberg, March 1999.

[LTW18] Gregor Leander, Cihangir Tezcan, and Friedrich Wiemer. Searching for subspace trails and truncated differentials. *IACR Trans. Symm. Cryptol.*, 2018(1):74–100, 2018.

[RBH17] Sondre Rønjom, Navid Ghaedi Bardeh, and Tor Helleseth. Yoyo tricks with AES. In Tsuyoshi Takagi and Thomas Peyrin, editors, *ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 217–243. Springer, Heidelberg, December 2017.

[SLG+16] Bing Sun, Meicheng Liu, Jian Guo, Longjiang Qu, and Vincent Rijmen. New insights on AES-like SPN ciphers. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 605–624. Springer, Heidelberg, August 2016.