On subspace trails cryptanalysis

Daniel Coggia

On subspace trails cryptanalysis

Daniel Coggia,
supervised by Anne Canteaut and Christina Boura
Inria Project Team SECRET

August 20, 2018

Summary

General context

Cryptography is about designing cryptosystems and cryptanalysis is about determining their security. Among well-known cryptosystems, the Advanced Encryption Standard (AES) chosen as a standard in 2001 [10] is probably the most widely used cryptographic primitive nowadays and determining its security remains a central problem in cryptanalysis. As cryptanalytic techniques still emerge in that goal, Grassi et al. presented at Eurocrypt 2017 [7] a new kind of distinguisher for a reduced version of the AES. A distinguisher is an algorithm able to distinguish between the cipher and a random oracle, which can be a first step in designing an attack against the cipher. This new kind of distinguisher is based on the notion of *subspace trails*.

The research problem

The proof of the existence of the distinguisher given by Grassi et al. in [7] is quite tedious and cumbersome while having a clear understanding of the distinguisher is crucial for determining its true limits. The research problem adressed during this internship is to fully understand this proof and to explore its possible extensions. All the details about the proof are presented in Section 1. Sections 2, 3 and 4 present the possible extensions.

My contribution

Always in a tight working relationship with Anne Canteaut and Christina Boura, I proved the properties behind Grassi's distinguisher in the most concise way and in the most general design I could. From the better understanding and intuition I got, I explored ways of improving the distinguisher on the AES and on another cipher with a similar design, Midori [2].

Arguments supporting its validity

The best evidence that the first part of my work (the rewriting of a proof in Section 1) is a good improvement is that I provide in this report a two-page-long proof general enough for embracing all the cases proved by Grassi et al. in ten pages of [7]. Moreover, my proof works without a rather

strong hypothesis used by Grassi et al. in [7]. The second part of my work, exploring possible improvements of the distinguisher, mostly led me to understanding the main papers on similar subjects and I have produced C programs running all the algorithms I present in this report.

Summary and future work

The distinguisher published by Grassi et al. is first presented and studied in Section 1. We will then see that this distinguisher is based on two properties verified by the AES.

1. The AES exhibits two-round subspace trails (defined in Definition 11).
2. The function $f(p^0, p^1) = \mathcal{R}(p^0) + \mathcal{R}(p^1)$ — where $\mathcal{R}$ is the round function — restricted to pairs of inputs (p^0, p^1) such that $p^0 + p^1$ belongs to a certain linear subspace is constant over well-defined equivalence classes.

Section 1 shows that the second property, unefficiently proved in [7], can be generalized to other ciphers. We give an example of this fact by applying the distinguisher to another cipher with similar design, Midori [2], in Section 3. However, Section 2 explains that for the AES, this generalisation is not helpful for improving the initial distinguisher presented by Grassi et al. in [7] because the subspace trails from the first property are optimal in some sense. Finally, Section 4 studies the influences of the components of an AES-like cipher on the existence of subspace trails.

The main problems left open by this report are whether it is possible to find general criteria on those components to avoid subspace trails and whether subspace trails can lead to other kinds of attacks, for example by finding a result better than the second property above. I will personnally continue to work on this subject as a PhD student under the supervision of Anne Canteaut and Christina Boura.

Contents

1 Starting point : Grassi's distinguisher

In [7], Grassi, Rechberger and Rønjom present a structural property of the 5-round AES leading to a distinguisher on the same number of rounds. The existence of this property strongly relies on Lemma 2 in [7] which is proved in a very ineffective way. The aim of this section is to give a better proof of this lemma in order to get a better understanding of the phenomenon that leads to the property.

Section 1.1 presents some important definitions and properties initially presented in [6] and [7]. Section 1.2 gives a better proof for Lemma 2 of [7] in the case of the AES. Finally, Section 1.4 proves a very similar result for a more general cipher.

1.1 Context

1.1.1 Description of the AES

The Advanced Encryption Standard [10] is a *Substitution-Permutation network* 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.
- $\mathcal{L} = MC \circ SR$ is the linear layer. SR is a cyclic shift of each row to the left and MC is the left multiplication of the state array by a constant 4×4 matrix over $\mathbb{F}_{2^8}$ denoted M_{MC} and defined by

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

- $\mathcal{K}$ is the *AddRoundKey* operation adding to the state array a 128-bit round-key derived from the master key.

1.1.2 First definitions

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, $d = 8$. Let $(e_{i,j})_{i,j \in [\![0,3]\!]}$ be the canonical basis of $\mathcal{M}_4(\mathbb{K})$.

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

As in [7], we define the following subspaces of $\mathcal{M}_4(\mathbb{K})$ for $i \in [\![0,3]\!]$, with indices computed modulo 4.

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

For example, if $x_1, x_2, x_3, x_4 \in \mathbb{K}$,

$$\begin{pmatrix} x_1 & 0 & 0 & 0 \\ x_2 & 0 & 0 & 0 \\ x_3 & 0 & 0 & 0 \\ x_4 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{C}_0, \quad \begin{pmatrix} x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & x_3 & 0 \\ 0 & 0 & 0 & x_4 \end{pmatrix} \in \mathcal{D}_0, \quad \begin{pmatrix} x_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 \\ 0 & 0 & x_3 & 0 \\ 0 & x_4 & 0 & 0 \end{pmatrix} \in \mathcal{ID}_0,$$

$$\begin{pmatrix} 2 \cdot x_1 & x_2 & x_3 & 3 \cdot x_4 \\ x_1 & x_2 & 3 \cdot x_3 & 2 \cdot x_4 \\ x_1 & 3 \cdot x_2 & 2 \cdot x_3 & x_4 \\ 3 \cdot x_1 & 2 \cdot x_2 & x_3 & x_4 \end{pmatrix} \in \mathcal{M}_0.$$

If $I \subseteq [\![0,3]\!]$, 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 vector 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 vector subspace $V \subseteq \mathcal{M}_4(\mathbb{K})$ is a set of the form $V + a = \{v + a \mid v \in V\}$ where $a \in \mathcal{M}_4(\mathbb{K})$.

Here comes the definition of the *differential branch number*. Understanding this definition will not be crucial for the rest of this section but we will make reference to it later.

Definition 1 (Differential branch number [4])**.** Let $N \in \mathbb{N}^\star$, $M : \mathbb{K}^N \to \mathbb{K}^N$ be a $\mathbb{F}_2$-linear function. The *differential branch number* of M over $\mathbb{K}^N$ with respect to $\mathbb{K}$ is

$$\min_{x \in \mathbb{K}^N \setminus \{0\}} (wt(x) + wt(Mx))$$

where $wt(x) = |\{i \in [\![0, N-1]\!] \mid x_i \neq 0\}|$ if $x = (x_0, \ldots, x_{N-1}) \in \mathbb{K}^N$.

For example, the AES MixColumns MC has branch number 5.

1.1.3 The five-round distinguisher

We now describe the distinguisher presented in [7] in order to have a complete understanding of the context in which lies our contribution. We begin with this easy to verify lemma.

Lemma 1 ([6])**.** *Let $I \subseteq [\![0,3]\!]$ and $a_1, a_2, a_3 \in \mathcal{M}_4(\mathbb{K})$. There exist $b_1, b_2, b_3 \in \mathcal{M}_4(\mathbb{K})$ such that*

- $\mathcal{R}(\mathcal{D}_I + a_1) = \mathcal{C}_I + b_1$*;*
- $\mathcal{R}(\mathcal{C}_I + a_2) = \mathcal{M}_I + b_2$*;*
- $\mathcal{R}^2(\mathcal{D}_I + a_3) = \mathcal{M}_I + b_3$*.*

Now comes a more subtle lemma which is the keystone of Theorem 1. For any set $\mathcal{E}$, we denote by $\mathcal{P}^2(\mathcal{E})$ the set of pairs of elements from $\mathcal{E}$, $\{\{a, b\} \mid a, b \in \mathcal{E}\}$.

Lemma 2 ([7])**.** *Let $a \in \mathcal{M}_4(\mathbb{K})$, $i \in [\![0,3]\!]$, $J \subseteq [\![0,3]\!]$. We define*

$$n = \#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{M}_i + a) \mid \mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{D}_J\}.$$

Then $n \equiv 0 \mod 8$.

Lemma 2 has a five-page-long proof in Section 6 of [7]. Section 1.2 gives an alternative proof of this result.

Theorem 1 ([7]). *Let $a \in \mathcal{M}_4(\mathbb{K})$, $i \in [\![0,3]\!]$, $J \subseteq [\![0,3]\!]$. We define*

$$n = \#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{D}_i + a) \mid \mathcal{R}^5(p^0) + \mathcal{R}^5(p^1) \in \mathcal{M}_J\}.$$

Then $n \equiv 0 \mod 8$.

Proof. We know by Lemma 1 that there exists $b \in \mathcal{M}_4(\mathbb{K})$ such that $\mathcal{R}^2$ is bijective rom $\mathcal{D}_i + a$ to $\mathcal{M}_i + b$. Moreover, Lemma 2 asserts that the number of pairs of state arrays from $\mathcal{M}_i + b$ whose respective images after one round belong to the same coset of D_J is a multiple of 8, i.e.

$$\#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{M}_i + b) \mid \mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{D}_J\} \equiv 0 \mod 8.$$

Hence with our first observation,

$$\#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{D}_i + a) \mid \mathcal{R}^3(p^0) + \mathcal{R}^3(p^1) \in \mathcal{D}_J\} \equiv 0 \mod 8.$$

Again by Lemma 1, we know that the images by $\mathcal{R}^2$ of two state arrays from the same coset of $\mathcal{D}_J$, for example $\mathcal{R}^3(p^0)$ and $\mathcal{R}^3(p^1)$ if $\{p^0, p^1\} \in \mathcal{P}^2(\mathcal{D}_i + a)$, always belong to the same coset of $\mathcal{M}_J$. Finally,

$$\#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{D}_i + a) \mid \mathcal{R}^5(p^0) + \mathcal{R}^5(p^1) \in \mathcal{M}_J\} \equiv 0 \mod 8.$$

□

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 [7] and we provide an implementation in [1].

1.2 A better proof

Now that the context of Lemma 2 in [7] (Lemma 2 here) is clear, we provide in this section a proof of this lemma which is no more than a much more concise version of the five-page-long proof given in [7]. This proof is in fact a formal proof of the generalisation of Lemma 2 in [7] only sketched in Appendix A in [7] through five other pages. Let us begin with two custom definitions.

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

$$\begin{aligned}\mathcal{ID}_I &= \mathrm{vect}_{\mathbb{K}}(e_{k,i-k} \mid k \in [\![0,3]\!], i \in I) = \mathrm{vect}_{\mathbb{K}}(e_{i-k,k} \mid k \in [\![0,3]\!], i \in I),\\ \mathcal{M}_I &= MC(\mathcal{ID}_I) = \mathrm{vect}_{\mathbb{K}}(MC(e_{i-k,k}) \mid k \in [\![0,3]\!], i \in I).\end{aligned}$$

Definition 2. Let $\{p^0, p^1\}$ be a pair of state arrays from $\mathcal{M}_I + a$. There exists a unique pair $x \in \mathbb{K}^{|I|\times 4}$ and $y \in \mathbb{K}^{|I|\times 4}$ such that

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

We then define the information set K of the pair $\{p^0, p^1\}$ as $\{k \in [\![0,3]\!] \mid \exists i \in I : x_{i,k} \neq y_{i,k}\}$.

Definition 3. Let $P = \{p^0, p^1\}$, $Q = \{q^0, q^1\} \in \mathcal{P}^2(\mathcal{M}_I + a)$. We say that $P \sim Q$ if :

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

$\sim$ is an equivalence relation on $\mathcal{P}^2(\mathcal{M}_I + a)$ and we denote

$$\Pi : \quad \mathcal{P}^2(\mathcal{M}_I + a) \quad \longrightarrow \quad \mathcal{P}^2(\mathcal{M}_I + a)/\sim$$

the canonical surjection.

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 pair $\{p^0, p^1\}$ 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 pairs and not tuples, we have $2^{|K|-1+d|I|(4-|K|)}$ pairs in $\mathfrak{C}$.

$|K| - 1 + d|I|(4 - |K|)$ is minimal for $|K| = 4$. Hence $\#\mathfrak{C} \equiv 0 \mod 8$. □

Now that we have the right definitions, we can state and prove a key lemma before proving a generalised version of Lemma 2.

Lemma 3. *The function*

$$\begin{array}{rccc} f : & \mathcal{P}^2(\mathcal{M}_I + a) & \longrightarrow & \mathcal{M}_4(\mathbb{K}) \\ & \{p^0, p^1\} & \longmapsto & \mathcal{R}(p^0) + \mathcal{R}(p^1) \end{array}$$

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

Proof. Let $P = \{p^0, p^1\}, Q = \{q^0, q^1\} \in \mathcal{P}^2(\mathcal{M}_I + a)$ such that $P \sim Q$. We write as in Definition 2

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

We also write the MixColumns matrix $M_{MC} = (m_{\ell,k})_{(\ell,k)\in[\![0,3]\!]^2}$. 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

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

It is now clear with Definition 3 and Equation (1) 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,

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

Therefore,

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

□

Now comes Lemma 4, which generalizes Lemma 2 in the sense that we are not restricted to the case $|I| = 1$ as in Lemma 2.

Lemma 4. *If* $n = \#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{M}_I + a) \mid \mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{D}_J\}$, *then* $n \equiv 0 \mod 8$.

Proof. The function f from Lemma 3 is constant on the equivalence classes of $\mathcal{P}^2(\mathcal{M}_I + a)$. This justifies the existence of a function $\tilde{f} : \mathcal{P}^2(\mathcal{M}_I + a)\big/\sim \to \mathcal{M}_4(\mathbb{K})$ such that $f = \tilde{f} \circ \Pi$. Since the equivalence classes form a partition of $\mathcal{P}^2(\mathcal{M}_I + a)$, we have that

$$\begin{aligned}
n &= \#f^{-1}(\mathcal{D}_J) \\
&= \sum_{\mathfrak{C}\in\mathcal{P}^2(\mathcal{M}_I+a)\big/\sim} \#(f^{-1}(\mathcal{D}_J) \cap \mathfrak{C}) \\
&= \sum_{\mathfrak{C}\in\mathcal{P}^2(\mathcal{M}_I+a)\big/\sim} 1_{\tilde{f}(\mathfrak{C})\in\mathcal{D}_J} \#\mathfrak{C} \\
&\equiv 0 \mod 8 \quad \text{since from Proposition 1 all equivalence classes have a cardinality divisible by 8.}
\end{aligned}$$

□

1.3 The influence of the branch number

In [7], the proof of Lemma 2 (also Lemma 2 here) uses the fact that the differential branch number of MC denoted by b is 5 but as we have seen, the branch number b does not seem to be of any importance for this lemma. However, it does have some influence and to show this, we begin with this proposition from [6].

Proposition 2 ([6])**.** *Let $I, J \subseteq [\![0,3]\!]$ and b be the differential branch number of MC. Then*

$$|I| + |J| < b \quad \Rightarrow \quad \mathcal{D}_I \cap \mathcal{M}_J = \{0\} \quad \textit{and} \quad \mathcal{ID}_I \cap \mathcal{M}_J = \{0\}.$$

Proof. Let $u \in \mathcal{D}_I \cap \mathcal{M}_J$. There exists $x \in \mathcal{M}_4(\mathbb{K})$ such that its nonzero entries are on the rows indexed by I and such that:

$$u = \sum_{k=0}^{3} \sum_{i \in I} x_{i,k} e_{k,i+k}.$$

There also exists $y \in \mathcal{M}_4(\mathbb{K})$ such that its nonzero entries are on the rows indexed by J and such that:

$$u = \sum_{\ell=0}^{3} \sum_{j \in J} y_{i,\ell} MC(e_{\ell,j-\ell}).$$

On column $c \in [\![0,3]\!]$ we get

$$\sum_{i \in I} x_{i,c-i} e_{c-i,c} + \sum_{j \in J} y_{j,j-c} MC(e_{j-c,c}) = 0.$$

We can write this equation as

$$\left(\begin{array}{c|c} M_{MC} & I_4 \end{array} \right) \times \begin{pmatrix} y_{c,0} \\ \vdots \\ y_{c+3,3} \\ x_{c,0} \\ \vdots \\ x_{c-3,3} \end{pmatrix} = 0$$

where $y_{c+\ell,\ell} = 0$ if $c + \ell \notin J$ and $x_{c-\ell,\ell} = 0$ if $c - \ell \notin I$. Hence, since $|I| + |J| < b$, the $\mathbb{K}$-vector on the right has weight at most $b - 1$. Moreover, $\left(\begin{array}{c|c} M_{MC} & I_4 \end{array} \right)$ is a parity-check matrix for the code $\mathcal{C}_{MC}^{\perp}$ whose minimal distance is b by definition of the differential branch number. Hence, $x = y = 0$, $u = 0$ and finally, $\mathcal{D}_I \cap \mathcal{M}_J = \{0\}$.

The proof for $\mathcal{ID}_I \cap \mathcal{M}_J = \{0\}$ is very similar.

□

Now, if we go back to the proof of Lemma 4, we have:

$$\begin{aligned} n &= \sum_{\mathfrak{C} \in \mathcal{P}^2(\mathcal{M}_I + a) \big/ \sim} 1_{\tilde{f}(\mathfrak{C}) \in \mathcal{D}_J} \# \mathfrak{C} \\ &= \sum_{h=0}^{4} \sum_{\mathfrak{C} : |K(\mathfrak{C})| = h} 1_{\tilde{f}(\mathfrak{C}) \in \mathcal{D}_J} \# \mathfrak{C} \end{aligned}$$

Let $\mathfrak{C}$ be such that its information set $K(\mathfrak{C})$ has size $h < b - 1 - |J|$. If $\{p^0, p^1\} \in \mathfrak{C}$, it is clear that

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

After one round, by Lemma 1, $\mathcal{R}(p^0) + \mathcal{R}(p^1) \in \mathcal{M}_{K(\mathfrak{C})}$. Since $p^0 \neq p^1$ and by Proposition 2 $\mathcal{M}_{K(\mathfrak{C})} \cap \mathcal{D}_J = \{0\}$ then $\mathcal{R}(p^0) + \mathcal{R}(p^1) \notin \mathcal{D}_J$ and we finally have that $1_{\tilde{f}(\mathfrak{C}) \in \mathcal{D}_J} = 0$.

We can then express the influence of the branch number on n with the formula

$$n = \sum_{h=b-|J|}^{4} \sum_{\mathfrak{C} : |K(\mathfrak{C})| = h} 1_{\tilde{f}(\mathfrak{C}) \in \mathcal{D}_J} \# \mathfrak{C}.$$

1.4 A more general proof

This section is very similar to Section 1.2. It aims to state a more general version of Lemma 2 for a more general SPN cipher than the AES. We fix $N \in \mathbb{N}^\star$. The proofs, very similar to the ones of Section 1.2 can be found in appendix A.

Our general SPN cipher is composed of an arbitrary number of rounds, the round keys and the internal states are represented as vectors in $\mathbb{K}^N$, not as matrices. A 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 : $\mathbb{K} \to \mathbb{K}$ to each coordinate of the internal state in a certain basis $(f_i)_{i \in [\![0,N-1]\!]}$ of $\mathbb{K}^N$. It is important to notice that we define $\mathcal{S}$ and $(f_i)_{i \in [\![0,N-1]\!]}$ together.
- $\mathcal{L}$ is the linear layer, a bijective $\mathbb{K}$-linear map of $\mathbb{K}^N$.
- $\mathcal{K}$ is the *AddRoundKey* operation adding to the internal state a round-key of same size.

Definition 4. Let V be a subspace of $\mathbb{K}^N$. We say that V is compatible with $\mathcal{S}$ if it has a basis such that its elements written in the basis $(f_i)_{i \in [\![0,N-1]\!]}$ of $\mathbb{K}^N$ form a block-diagonal matrix of any size blocks. More formally, V is compatible with $\mathcal{S}$ if there exists $h \in [\![1, N]\!]$ such that :

$$\begin{aligned}
&\exists (i_0, \ldots, i_{h-1}) \in \mathbb{N}^h : \sum_{k=0}^{h-1} i_k = \dim V, \\
&\exists (j_0, \ldots, j_h) \in \mathbb{N}^{h+1} : j_0 = 0 \text{ and } j_h \leq N, \\
&\exists (g_{k,i})_{k<h, i<i_k} \in (\mathbb{K}^N)^{\dim V} : V = \mathrm{vect}_{\mathbb{K}}(g_{k,i} \mid k < h, i < i_k), \\
&\text{and } \exists (\lambda_{k,\ell,i}) \in \mathbb{K}^{3N} : \forall k \in [\![0, h-1]\!], \forall i \in [\![0, i_k - 1]\!],\ g_{k,i} = \sum_{\ell=0}^{j_{k+1}-j_k-1} \lambda_{k,\ell,i} f_{j_k+\ell}.
\end{aligned}$$

We call such a basis of V a compatibility basis. Written as a collection of column vectors written in the basis f, the compatibility basis looks like this :

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

Example 1. If $N = 16, I \subseteq [\![0,3]\!]$, $\mathcal{M}_I$ is compatible with the AES S-box layer (denoted by $\mathcal{S}_{\text{AES}}$). Indeed, we consider the basis f of $\mathbb{K}^{16}$ defined by $f_{4j+i} = e_{i,j}$ then we take $h = 4, i_0 = \ldots = i_3 = |I|, \phi : [\![0, |I|-1]\!] \to I$ a bijection, $j_k = 4k$ for all $k \leq 4$ and for all $k \leq 4, i < |I|$, $g_{k,i} = MC(e_{\phi(i)-k,k})$. Hence

$$\mathcal{M}_I = \text{vect}_{\mathbb{K}}(g_{k,i}) \quad \text{and} \quad \forall k, i,\ g_{k,i} = \sum_{\ell=0}^{3} m_{\ell,\phi(i)-k} e_{\ell,k} = \sum_{\ell=0}^{j_{k+1}-j_k-1} m_{\ell,\phi(i)-k} f_{j_k+\ell}.$$

Example 2. If $I \subset [\![0, N-1]\!]$, $V = \text{vect}_{\mathbb{K}}(f_i \mid i \in I)$ is compatible with $\mathcal{S}$. Indeed, we have $h = |I|$ and $\forall k,\ i_k = 1, j_k = k$ and $g_{k,0} = f_{\phi(k)}$ where $\phi : [\![0, h-1]\!] \to [\![0, N-1]\!]$ is a bijection.

From now on, we fix $a \in \mathbb{K}^N$ and a subspace V compatible with $\mathcal{S}$ with compatibility basis g.

Definition 5. Let $\{p^0, p^1\}$ be a pair of states from $V + a$. There exists a unique pair $x \in \mathbb{K}^{\dim V}$ and $y \in \mathbb{K}^{\dim V}$ such that

$$p^0 = \sum_{k=0}^{h-1} \sum_{i \in I_k} x_{i,k} g_{i,k} + a \quad \text{and} \quad p^1 = \sum_{k=0}^{h-1} \sum_{i \in I_k} y_{i,k} g_{i,k} + a.$$

We define the information set K of the pair $\{p^0, p^1\}$ as $\{k \in [\![0, h-1]\!] \mid \exists i < i_k : x_{i,k} \neq y_{i,k}\}$.

Definition 6. Let $P = \{p^0, p^1\},\ Q = \{q^0, q^1\} \in \mathcal{P}^2(V + a)$. We say that $P \sim Q$ if :

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

$\sim$ is an equivalence relation on $\mathcal{P}^2(V + a)$ and we denote

$$\Pi : \ \mathcal{P}^2(V + a) \ \longrightarrow \ \mathcal{P}^2(V + a)/\sim$$

the canonical surjection

Proposition 3. *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}.

Lemma 5. *The function*

$$\begin{array}{rccc} f: & \mathcal{P}^2(V+a) & \longrightarrow & \mathbb{K}^N \\ & \{p^0,p^1\} & \longmapsto & \mathcal{R}(p^0)+\mathcal{R}(p^1) \end{array}$$

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

Now comes the most general result we managed to obtain on this subject.

Lemma 6. *Let $\mathcal{E}$ be any subset of $\mathbb{K}^N$. We define*

$$n = \#\{\ \{p^0,p^1\} \in \mathcal{P}^2(V+a) \mid \mathcal{R}(p^0)+\mathcal{R}(p^1) \in \mathcal{E}\}.$$

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

The initial distinguisher relies on the combination of two properties respectively given by Lemma 1 and Lemma 4. This subsection only generalises Lemma 4 and the aim of the next section will be to replace Lemma 1 to obtain a full generalisation of the distinguisher. In particular, it aims at finding a property similar to Lemma 1 that holds for strictly more rounds than two in the case of the AES.

2 Study for more rounds

Lemma 1 gives a property that links the subspaces $\mathcal{D}_I$ and $\mathcal{M}_I$ over two rounds, i.e. gives a two-round *subspace trail* (see Definition 11). This property and the one-round property of Lemma 4 allow to build a five-round distinguisher. If we could find a property similar to the one of Lemma 1 over strictly more than two rounds, it would directly yield a distinguisher over more rounds of the AES. However, Leander, Tezcan and Wiemer showed in [9] that we cannot find any better subspace trail than the one exhibited in [6].

Subsection 2.1 shows that the AES S-box has no non-linear structure, which is useful for Subsection 2.2 that shows like in [9] that the two-round subspace trails presented in [6] are optimal. This whole section sums up the parts of [9] that apply to the AES.

2.1 Linear structures in the AES S-box

In this section, we are going to focus on the notion of linear structures with basic definitions and properties. We fix m and n in $\mathbb{N}$.

Definition 7. Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $a \in \mathbb{F}_2^n$. The derivative of F in direction a is defined as

$$\begin{array}{rcccc} \Delta_a(F) & : & \mathbb{F}_2^n & \to & \mathbb{F}_2^m \\ & & x & \mapsto & F(x)+F(x+a). \end{array}$$

Definition 8 ([9])**.** Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. The set of linear structures of f is

$$\mathrm{LS}(f) = \{a \in \mathbb{F}_2^n \mid \exists c_a \in \mathbb{F}_2 : \forall x, \Delta_a(f)(x) = c_a\}.$$

Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$. The set of linear structures of F is

$$\mathrm{LS}(F) = \{(b, a) \in \mathbb{F}_2^m \times \mathbb{F}_2^n \mid \exists c_{b,a} \in \mathbb{F}_2 : \forall x, \langle b|\Delta_a(F)(x)\rangle = c_{b,a}\}.$$

For example, the set of trivial linear structures is $(\{0\} \times \mathbb{F}_2^n) \cup (\mathbb{F}_2^m \times \{0\}) \subset \mathrm{LS}(F)$. Now, given a function F, we want to find if it has any non-trivial linear structures. In order to do this, we will need the Walsh transform.

Definition 9. Let $\phi : \mathbb{F}_2^n \to \mathbb{Z}$. The Walsh transform of ϕ is the function

$$\begin{array}{rcl} \mathcal{W}(\phi) : \mathbb{F}_2^n & \to & \mathbb{Z} \\ a & \mapsto & \sum_{x \in \mathbb{F}_2^n} (-1)^{\langle a|x\rangle} \phi(x). \end{array}$$

The entire table of the Walsh transform can be efficiently computed from the entire table of ϕ in $O(n2^n)$ operations in $\mathbb{Z} \cap [-2^n, 2^n]$. To exhibit the link between the Walsh transform and the linear structures, we now define the autocorrelation of a Boolean function f and we give easy to verify properties.

Definition 10. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. The autocorrelation of f is the function

$$\begin{array}{rcl} \mathcal{A}(f) : \mathbb{F}_2^n & \to & \mathbb{Z} \\ a & \mapsto & \sum_{x \in \mathbb{F}_2^n} (-1)^{\Delta_a(f)(x)}. \end{array}$$

Proposition 4. *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. Then $\mathcal{W}(\mathcal{W}(x \mapsto (-1)^{f(x)})^2) = 2^n \mathcal{A}(f)$.*

Proposition 5. *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. Then f has no non-trivial linear structures if and only if $\forall a \in \mathbb{F}_2^n \backslash \{0\},\ |\mathcal{A}(f)(a)| < 2^n$.*

Proposition 6. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$. Then F has no non-trivial linear structures if and only if all its non-zero components have no non-trivial linear structures, i.e. for all $b \in \mathbb{F}_2^n \backslash \{0\}$, the Boolean function $x \mapsto \langle b|F(x)\rangle$ has no non-trivial linear structures.*

Finally, to check whether a function F has non-trivial linear structures, we can check for all its components whether they have non-trivial linear structures by computing their autocorrelation with two Walsh transforms and by comparing the autocorrelation values with 2^n. This yields a simple algorithm — Algorithm B — running in $O(n2^{n+m})$ operations in $\mathbb{Z} \cap [-2^{2n}, 2^{2n}]$. Running this algorithm on the AES S-box with $n = m = 8$ answers that the AES S-box has no non-trivial linear structures.

2.2 Searching for other subspace trails

Definition 11 (Subspace trail [9])**.** Let $\mathcal{F} : \mathbb{K}^N \to \mathbb{K}^N$ be any map. $\mathbb{F}_2$-linear subspaces $U, V \subseteq \mathbb{K}^N$ are called a (one-round) $\mathcal{F}$-subspace trail if

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

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

$$\forall i \in [\![0, r-1]\!],\ U_i \rightrightarrows_{\mathcal{F}} U_{i+1}.$$

For example, we have the trivial subspace trails $\{0\} \rightrightarrows_{\mathcal{F}} \{0\}$ and $U \rightrightarrows_{\mathcal{F}} \mathbb{K}^N$ and this trivial lemma.

Lemma 7 ([9])**.** *If* $U \rightrightarrows_{\mathcal{F}} V$ *then* $\forall u \in U,\ \mathrm{Im}(\Delta_u(\mathcal{F})) \subseteq V$.

Moreover, if $U \rightrightarrows_{\mathcal{F}} V$, then for all $U' \subseteq U$ and $V' \supseteq V$, $U' \rightrightarrows_{\mathcal{F}} V'$. Therefore, in order to search for subspace trails, it is sufficient to stick to *essential* subspace trails which correspond to optimal subspace trails with respect to inclusion.

Definition 12 (Essential subspace trail [9])**.** Let $U, V \subseteq \mathbb{K}^N$ be a $\mathcal{F}$-subspace trail. If

$$\begin{aligned} &\forall U' \supset U,\ U' \not\rightrightarrows_{\mathcal{F}} V \\ \text{and} \quad &\forall V' \subset V,\ U \not\rightrightarrows_{\mathcal{F}} V', \end{aligned}$$

then (U, V) is called an essential $\mathcal{F}$-subspace trail.

In the following, if $I \subseteq [\![0, N-1]\!]$, we will denote $\mathrm{vect}_{\mathbb{K}}(e_i \mid i \in I)$ as $\mathcal{E}_I$.

Lemma 8 (Prop. 2 in [9])**.** *Let* $U, V \subseteq \mathbb{K}^N$ *be an essential* $\mathcal{S}$*-subspace trail. If the* S-box *has no non-trivial linear structures, then*

$$\exists I \subseteq [\![0, N-1]\!] \ :\ U = V = \mathcal{E}_I.$$

Lemma 9. *Let* $U, V \subseteq \mathbb{K}^N$*. Then we have that:*

1. $U \rightrightarrows_{\mathcal{S}} V$ *if and only if* $U \rightrightarrows_{\mathcal{R}} \mathcal{L}(V)$*.*
2. $U \rightrightarrows_{\mathcal{S}} V$ *is essential if and only if* $U \rightrightarrows_{\mathcal{R}} \mathcal{L}(V)$ *is essential.*

Proof. First, we consider the trail $U \rightrightarrows_{\mathcal{S}} V$. Let $a \in \mathbb{K}^N$, there exists $b \in \mathbb{K}^N$ such that

$$\begin{aligned} & \mathcal{S}(U+a) && \subseteq V + b \\ \Rightarrow \quad & \mathcal{L} \circ \mathcal{S}(U+a) && \subseteq \mathcal{L}(V) + \mathcal{L}(b) \\ \Rightarrow \quad & \mathcal{R}(U+a) && \subseteq \mathcal{L}(V) + \mathcal{K} \circ \mathcal{L}(b), \end{aligned}$$

which shows that we consequently have $U \rightrightarrows_{\mathcal{R}} \mathcal{L}(V)$. Similarly, if $U \rightrightarrows_{\mathcal{R}} \mathcal{L}(V)$ we can show that $U \rightrightarrows_{\mathcal{S}} V$.

We now assume that $U \rightrightarrows_{\mathcal{S}} V$ is essential. If $U' \supseteq U$ and $U' \rightrightarrows_{\mathcal{R}} \mathcal{L}(V)$, the first point shows that we also have $U' \rightrightarrows_{\mathcal{S}} V$. Since $U \rightrightarrows_{\mathcal{S}} V$ is essential, we get $U = U'$. Moreover, if $V' \subseteq \mathcal{L}(V)$ is such that $U \rightrightarrows_{\mathcal{R}} V'$, we have that $U \rightrightarrows_{\mathcal{S}} \mathcal{L}^{-1}(V')$. Again, since $U \rightrightarrows_{\mathcal{S}} V$ is essential, we get $V = \mathcal{L}^{-1}(V')$ and $V' = \mathcal{L}(V)$. Finally, $U \rightrightarrows_{\mathcal{R}} \mathcal{L}(V)$ is essential.

The converse can be shown in the same way. □

Theorem 2. *Let* $U, V \subseteq \mathbb{K}^N$ *be an essential* $\mathcal{R}$*-subspace trail. If the* S-box *has no non-trivial linear structures, then*

$$\exists I \subseteq [\![0, N-1]\!] \ :\ U = \mathcal{E}_I \quad \textit{and} \quad \mathcal{L}(U) = V.$$

Proof. We have by Lemma 9 the trail $U \rightrightarrows_{\mathcal{S}} \mathcal{L}^{-1}(V)$ and the fact that this trail is essential. By Lemma 8, $\exists I \subseteq [\![0, N-1]\!] : U = \mathcal{L}^{-1}(V) = \mathrm{vect}_{\mathbb{K}}(e_i | i \in I)$. □

We now want to find two-round subspace trails. We then consider the trail $U \rightrightarrows_{\mathcal{R}} V \rightrightarrows_{\mathcal{R}} W$ decomposed in the essential subspace trails

$$U \rightrightarrows_{\mathcal{R}} V', \quad V' \subseteq V, \quad V \rightrightarrows_{\mathcal{R}} W$$

We know with Theorem 2 that there exist $I,\ J \subseteq [\![0, N-1]\!]$ such that

1. $U = \mathcal{E}_I$
2. $V = \mathcal{E}_J$
3. $V' = \mathcal{L}(U) \subseteq V$ and $W = \mathcal{L}(V)$.

The subspace trails we want to find must allow us to build a distinguisher and then the subspace W should not be the entire set $\mathbb{K}^N$. Equivalently, the subspace V should not be equal to $\mathbb{K}^N$, which adds a fourth constraint.

4. $V \neq \mathbb{K}^N$.

The exhaustive search over I checking whether the smallest J such that $\mathcal{L}(\mathcal{E}_I) \subseteq \mathcal{E}_J$ has cardinality N then provides a very simple algorithm — Algorithm B — to find interesting two round subspace trails. For the AES, this algorithm answers that the only possibilities for U are the diagonals defined in Section 1.1 and that the length cannot be strictly longer than two rounds. For other versions of Rijndael (the AES is a special instance of a more general cipher called Rijndael [5]), this algorithm also answers that it is not possible to find a trail strictly longer than two rounds.

3 Adapting to Midori

Midori is a recent SPN cipher designed by Banik et al. and presented at Asiacrypt 2015 [2]. It is designed to be lightweight in the sense that an encryption/decryption consumes much less energy when implemented in hardware than the AES.

3.1 Description of Midori

The notation here is the same as the one in Section 1.4. As for the AES, for Midori $N = 16$ and the internal state is represented as a state array. There are two different versions of Midori: Midori64 has $d = 4$ and hence $\mathbb{K} = \mathbb{F}_{2^4}$ and for Midori128, $d = 8$ like for the AES. The round function $\mathcal{R}$ for both versions is the composition of:

- $\mathcal{S}$, the parallel operation of the S-boxes. For Midori64, the S-box is always the same and given in Table 1 whereas for Midori128, the S-box depends on the row of the nibble.
- $\mathcal{L} = MC \circ SC$ where SC is the ShuffleCell operation, a permutation of the $\mathbb{K}$-entries of the state array given by Table 2 and $\forall i \in [\![0, N-1]\!], SC(s)_{\Pi(i)} = s_i$. MC is the left multiplication of the state array by a constant 4×4 binary matrix denoted by M_{MC}.

$$M_{MC} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

- $\mathcal{K}$, the classical *AddRoundKey* operation.

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
S-box(x)	c	a	d	3	e	b	f	7	8	9	1	5	0	2	4	6

Table 1: Midori64 S-box in hexadecimal form

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\Pi(i)$	0	7	14	9	5	2	11	12	15	8	1	6	10	13	4	3

Table 2: Midori ShuffleCell

3.2 Distinguisher on Midori

It is obvious that Midori fits into our general cipher description of Section 1.4. Moreover, as in Section 1.1, we define the following subspaces :

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

It is also obvious that for all $I \subseteq [\![0,3]\!]$, $\mathcal{M}'_I$ is compatible with $\mathcal{S}$ with $h = 4$ as in Example 1. We can then transpose Theorem 1 for Midori thanks to Lemma 4, yielding a distinguisher for five-round Midori64 and five-round Midori128.

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

$$n = \#\{\ \{p^0, p^1\} \in \mathcal{P}^2(\mathcal{D}'_i + a) \mid \mathcal{R}^5(p^0) + \mathcal{R}^5(p^1) \in \mathcal{M}'_J\}.$$

Then $n \equiv 0 \mod 8$.

As for the AES, we searched for other possible trails to mount a similar distinguisher on more rounds, but Algorithm B answers that the only interesting two-round trails are $\mathcal{D}'_I \xrightarrow{\mathcal{R}} \mathcal{C}_I \xrightarrow{\mathcal{R}} \mathcal{M}'_I$ for $I \subseteq [0,3]$ and that no non-trivial three-round trail with the first and last subspaces having the same dimension exists.

However, as shown in Section 2.2, Algorithm B performs an exhaustive search if the S-boxes have no non-trivial linear structures. This observation motivates a deeper study of subspace trails in Section 4.

4 A broader study of subspace trails

Until now, we have mostly been interested in understanding Grassi's distinguisher and we have restricted our view on subspace trails in that purpose. This section aims at giving more details on subspace trails, starting from the results of [9] before showing some interesting observations.

4.1 Linear structures and subspace trails

If an S-box has non-trivial linear structures, both Lemma 8 and Theorem 2 are no longer appliable. However, Leander et al. gave in [9] an algorithm that gives a bound on the length of the longest possible subspace trail for a given SPN. We will need the following definition.

Definition 13. Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $u \in \mathbb{F}_2^n$, then we define for $c \in \mathbb{F}_2$

$$L_u^c(F) = \{\alpha \in \mathbb{F}_2^m | \forall x \in \mathbb{F}_2^n, \langle \alpha \mid \Delta_u(F)(x)\rangle = c\}$$

and $L_u(F) = L_u^0(F) \cup L_u^1(F)$.

4.1.1 Length of the longest subspace trails

The idea of the algorithm is to avoid performing an exhaustive search on all the possible starting subspaces by focusing on a class of dimension-1 subspaces. We suppose we have an S-box : $\mathbb{K} \to \mathbb{K}$ and we define the set

$$\mathbb{W} = \{W_{i,\alpha} := \{0\}^{i-1} \times \{0, \alpha\} \times \{0\}^{k-i} \mid \alpha \in \mathbb{K}, i \in [1, N]\}.$$

Proposition 7 (Prop. 3 in [9]). *If $U \rightrightarrows_{\mathcal{S}} V$ is a non-trivial subspace trail, and if the S-box satisfies the common property*

$$\forall u \in \mathbb{K}\backslash\{0\}, L_u(\text{S-Box}) \neq \mathbb{K}$$

then there exists $W \in \mathbb{W}$ such that $W \subseteq V$.

With Proposition 7, we can see that if we explicitly compute all the trails starting from $\mathcal{L}(W_{i,\alpha})$ then the length of the longest trail plus 1 gives an upper bound for the length of the longest subspace trail. Indeed, let ℓ be the length of a longest trail, starting from $\mathcal{L}(W_{i,\alpha})$. We have for example $\mathcal{L}(W_{i,\alpha}) \rightrightarrows_{\mathcal{R}} W \rightrightarrows_{\mathcal{R}} \cdots$, with $\dim W > 1$. There might exist a trail $U \rightrightarrows_{\mathcal{S}} V \rightrightarrows_{\mathcal{L}} \mathcal{L}(V) \rightrightarrows_{\mathcal{R}} W$ where $W_{i,\alpha} \subseteq V$. Then the trail starting from U will have length $\ell + 1$, which illustrates why we have to add 1 to the longest trail length for the bound.

4.1.2 Candidate for the longest subspace trail

A question that naturally arises from the study of the algorithm is how to find a candidate for the starting subspace leading to the longest subspace trail without explicitly computing trails. Again from [9], we have this useful property.

Proposition 8 (Lemma 3 in [9]). *If $U \rightrightarrows_{\mathcal{S}} V$ then $V^\perp = \bigcap_{u \in U} L_u^0(\mathcal{S})$.*

Moreover, we have that $L_u^0(\mathcal{S}) \subseteq L_{u_0}(\text{S-box}) \times \ldots \times L_{u_{N-1}}(\text{S-box})$. Starting from $U = \{0, u\}$ of dimension 1, we want to minimize the dimension of V, or equivalently to maximize the dimension of $L_u^0(\mathcal{S}) = V^\perp$. Since

$$\dim L_u^0(\mathcal{S}) \leq \dim L_{u_0}(\text{S-box}) + \ldots + \dim L_{u_{N-1}}(\text{S-box}),$$

— with equality for Midori — maximal dimensions for $L_u^0(\mathcal{S})$ are obtained for $u = (0, \ldots, u_i, \ldots, 0)$ where u_i is such that the dimension of $L_{u_i}^0(\text{S-box})$ is maximal. We can find such u_i with the computation of LS(S-box) presented in Section 2.1.

Unfortunately, I have not been able to find a good criterion when $\dim U$ is bigger than 1.

4.1.3 Infinite trails

When the S-box and the linear layer $\mathcal{L}$ have been poorly chosen, we might encounter the case of an infinite subspace trail. Indeed, if we have :

1. S-box such that $\exists c \in \mathbb{F}_2 : \exists \alpha \in \mathbb{K} : \forall x \in \mathbb{K}, \langle \alpha \mid \text{S-box}(x) + x \rangle = c$,
2. $\mathcal{L}$ such that it only permutes nibbles of $\mathbb{K}^N$ and XORs the nibbles (which is the case of Midori),

we consequently have that $\forall u, x \in \mathbb{K}, \langle \alpha \mid \Delta_u(\text{S-box})(x) \rangle = \langle \alpha \mid u \rangle$. Hence, if we denote $H_\alpha = \ker \langle \alpha \mid . \rangle$, we get the trail $H_\alpha \rightrightarrows_{\text{S-box}} H_\alpha$. In fact having such a trail is equivalent to the first condition on the S-box. Moreover, the second condition implies that we have a trail $H_\alpha^N \rightrightarrows_{\mathcal{L}} H_\alpha^N$ and finally, $H_\alpha^N \rightrightarrows_{\mathcal{R}} H_\alpha^N$. This one-round trail obviously gives an infinite trail when repeated.

However, the first condition is not likely to be verfied by the S-box of a real SPN cipher. Indeed, this condition suggests the cipher would be vulnerable to linear attacks, which are well-known by designers. Hence, designers usually choose the S-box to avoid them.

4.2 Focus on the linear layer of Midori

We have seen in Section 3.2 that, without considering the linear structures of the S-boxes, the maximum number of Midori rounds for which there exist subspace trails that do not increase the dimension is 2, as for the AES (see Section 2.2). This is a bit disappointing because the branch number of Midori's MixColumns is 4 whereas the AES MixColumns has the maximal branch number 5. Interestingly, relaxing the condition on same dimension subspaces along the trail, Algorithm B showed that replacing ShuffleCell by ShiftRows while keeping Midori's M_{MC} gives a longer trail.

We give in this section a criterion for the ShuffleCell permutation to avoid having a longer trail when we keep Midori's MixColumns. Keeping this MixColumns is interesting because its matrix is the only 4×4 binary matrix for which MC has branch number 4, as shown in [8].

To really focus on the linear layer, we will not consider the linear structures of the S-box and we will use the notation $\mathcal{E}_I$ as in Section 2.2. The indexes of the nibbles in the state array are classically given by Figure 1. We then represent $\mathcal{E}_I$ as a state array with a one at index i for all i in I. For example, $\mathcal{E}_{0,5,10,13}$ and the trail $\mathcal{D}'_0 \rightrightarrows_{\mathcal{R}} \mathcal{C}_0$ are respectively represented by

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \rightrightarrows_{\mathcal{R}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

$$\begin{pmatrix} 0 & 4 & 8 & 12 \\ 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \end{pmatrix}$$

Figure 1: Indexes of nibbles

Let us now consider a permutation of indexes Π for the ShuffleCell. Rijmen and Daemen showed in [5] that Π should be *diffusion optimal*, which means that the indexes from the same column should be sent to pairwise distinct columns. We then have this kind of trail for every input $\mathcal{E}_i$, $i \in [0, 15]$.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightrightarrows_{\mathcal{S}\circ\mathcal{R}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \rightrightarrows_{SR} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \rightrightarrows_{\mathcal{S}\circ MC} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

We see here that because of the 0-entries in M_{MC}, a column with three 0 will still have one 0 after MC, just where the single 1 was before MC. Then, for the next round, we do not want to have three 0 in the same column before MC, otherwise the trail will be a bit longer. Since there is still one empty column after two rounds (column 0 in our example), and that after the ShuffleCell the 0-entries of this empty column will be distributed to the four different columns, two 0 from the other columns (at index 7, 10, 13 in our example) should not be sent to the same column after SC. In other words, the images by Π of those three indexes should be in pairwise distinct columns. Moreover, as we have said earlier, those three indexes correspond to the images of indexes from the same column, which leads to the following criterion.

Proposition 9. *For a Midori-like cipher for which the S-box has no non-trivial linear structures and for which ShuffleCell is diffusion optimal, there exist four-round subspace trails if and only if the permutation of indices Π verifies :*

$\exists i, j \in [\![0, 15]\!] : i, j$ *belong to the same column and* $\Pi^2(i)$ *and* $\Pi^2(j)$ *also belong to the same column.*

5 Conclusion

Subspace trails attacks are recent and have given unexpected results on reduced-round AES, actually the best results so far on five-round AES-128 and seven-round AES-192 [3]. Therefore, understanding those attacks becomes more relevant than ever in symmetric cryptography and that was the purpose of the work presented in this report.

First, we have presented a distinguisher of the five-round AES published by Grassi et al. in 2017 [7]. It was one of the first algorithms to reveal the interest to have in subspace trails. It appeared that this distinguisher relied on two different properties that combined very well for the AES.

1. The AES exhibits two-round subspace trails (defined in Definition 11).
2. The function $f(p^0, p^1) = \mathcal{R}(p^0) + \mathcal{R}(p^1)$ restricted to pairs of input (p^0, p^1) such that $p^0 + p^1$ belongs to a certain linear subspace is constant over well-defined equivalence classes.

We saw in Section 1 that the second property can be generalized to any SPN cipher, which has to my knowledge only been proved in this report until now. Furthermore, we gave an example of this fact in Section 3 by applying nearly the same distinguisher to Midori [2], another SPN cipher. However, Section 2 explained that algorithms from [9] searching for subspace trails show that the

AES has no longer subspace trails than the ones already exhibited by Grassi et al. in [7]. We finally tried to push our understanding of subspace trails furthur in Section 4 by studying the influences of the S-box layer and the linear layer. This lead to two modest criteria on each of them to avoid subspace trails for Midori-like ciphers.

The main problems left open by this report are whether it is possible to find more general criteria on the S-box layer and on the linear layer to avoid subspace trails and whether subspace trails can lead to other kinds of attacks, for example by finding a better result than Lemma 6.

Acknowledgements

I would especially like to thank Anne Canteaut for accepting me as an intern in the Project Team SECRET at Inria Paris and Christina Boura for supervising my work there. I also thank the whole team SECRET for its warm welcome. I finally thank the Délégation Générale de l'Armement for its support during this internship.

References

[1] `https://github.com/dnlcog/grassi_disting`.

[2] S. Banik, A. Bogdanov, T. Isobe, K. Shibutani, H. Hiwatari, and T. Akishita. Midori: A block cipher for low energy. In *ASIACRYPT 2015 (2)*, pages 411 – 436, 2015.

[3] A. Bar-On, O. Dunkelman, N. Keller, E. Ronen, and A. Shamir. Cryptology ePrint Archive, Report 2018/527, 2018.

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

[5] J. Daemen and V. Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard.* Springer, 2002.

[6] L. Grassi, C. Rechberger, and S. Rønjom. Subspace trail cryptanalysis and its applications to AES. *IACR Trans. Symmetric Cryptol.*, 2016(2):192–225, 2016.

[7] L. Grassi, C. Rechberger, and S. Rønjom. A new structural-differential property of 5-round AES. In *EUROCRYPT 2017, Part II*, volume 10211 of *LNCS*, pages 289–317. Springer, 2017.

[8] I. Landjev and A. Rousseva. The main conjecture for near-MDS codes. In *WCC2015 - 9th International Workshop on Coding and Cryptography 2015*, Paris, France, 2015.

[9] G. Leander, C. Tezcan, and F. Wiemer. Searching for subspace trails and truncated differentials. *IACR Trans. Symmetric Cryptol.*, 2018(1):74–100, 2018.

[10] NIST. *Specification for the Advanced Encryption Standard (AES)*, November 2001. Federal Information Processing Standard 197.

A Proofs of Section 1.4

Proposition 3

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 $b = 1$, ie $2^{|K|}$ choices. Since we are counting pairs and not tuples, we have $2^{|K|-1+d\sum_{k\notin K} i_k}$ pairs in $\mathfrak{C}$.

$|K| - 1 + d\sum_{k\notin K} i_k$ is minimal for $|K| = h$. Hence $\#\mathfrak{C} \equiv 0 \mod 2^{h-1}$. □

Lemma 5

Proof. Let $P = \{p^0, p^1\}, Q = \{q^0, q^1\} \in \mathcal{P}^2(V + a)$ such that $P \sim Q$. We have with the notation of Definition 4

$$\begin{aligned} p^0 &= \sum_{k=0}^{h-1}\sum_{i\in I_k} p^0_{i,k} g_{i,k} + a \\ &= \sum_{k=0}^{h-1}\sum_{\ell\in J_k}\left(\sum_{i\in I_k} p^0_{i,k}\lambda_{i,k,\ell} + a_{j_k+\ell}\right) f_{j_k+\ell}. \end{aligned}$$

Then

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

and

$$\mathcal{S}(p^0) + \mathcal{S}(p^1) = \sum_{k,\ell}\left[\text{S-box}\left(\sum_{i\in I_k} p^0_{i,k}\lambda_{i,k,\ell} + a_{j_k+\ell}\right) + \text{S-box}\left(\sum_{i\in I_k} p^1_{i,k}\lambda_{i,k,\ell} + a_{j_k+\ell}\right)\right] f_{j_k+\ell} \quad (2)$$

It is now clear with Definition 6 and Equation (2) 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,

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

□

Lemma 6

Proof. The function f from Lemma 5 is constant on the equivalence classes of $\mathcal{P}^2(V + a)$. This justifies the existence of a function $\tilde{f} : \mathcal{P}^2(V + a)/\sim\ \to \mathbb{K}^N$ such that $f = \tilde{f}\circ\Pi$. Since the equivalence classes form a partition of $\mathcal{P}^2(V+a)$, we have that

$$\begin{aligned} n &= \#f^{-1}(\mathcal{E}) \\ &= \sum_{\mathfrak{C}\in\mathcal{P}^2(V+a)\big/\sim} \#(f^{-1}(\mathcal{E})\cap\mathfrak{C}) \\ &= \sum_{\mathfrak{C}\in\mathcal{P}^2(V+a)\big/\sim} 1_{\tilde{f}(\mathfrak{C})\in\mathcal{E}}\#\mathfrak{C} \\ &\equiv 0 \mod 2^{h-1} \quad \text{by proposition 3.} \end{aligned}$$

□

B Algorithms

Algorithm 1 Checks the existence of linear structures

```
procedure Has linear structures($F$)
input : oracle access to $F(x)_i$ when $(x, i) \in \mathbb{F}_2^n \times [0, m-1]$.
output : true if $F$ has linear structures, false otherwise.
    for all $b \in \mathbb{F}_2^m \backslash \{0\}$ do
        $\mathcal{W}_1 \leftarrow$ Walsh transform$(-2\langle b|F\rangle + 1)$
        $\mathcal{W}_2 \leftarrow \mathcal{W}_1 \times \mathcal{W}_1$
        $\mathcal{A} \leftarrow |$Walsh transform$(\mathcal{W}_2)|/2^n$
        for all $a \in \mathbb{F}_2^n \backslash \{0\}$ do
            if $\mathcal{A}(a) == 2^n$ then
                return true
            end if
        end for
    end for
    return false
end procedure
```

Algorithm 2 Find interesting subspace trails

```
procedure Essential trail($\mathcal{L}$)
    input : Matrix of the linear map $\mathcal{L}$ in the basis $(e_i)_{0 \leq i \leq N-1}$.
    output : Table indexed by $I \subseteq [\![0, N-1]\!]$ whose element $J$ at index $I$ is the smallest set such
    that $\mathcal{L}(\mathcal{E}_I) \subseteq \mathcal{E}_J$.

    Create table Tab indexed by subsets of $[\![0, N-1]\!]$
    for all $I \subseteq [\![0, N-1]\!]$ do
        $J \leftarrow \emptyset$
        for all $i \in I$ do
            $x \leftarrow \mathcal{L}(e_i)$
            for all $j \in J$ such that $x_j \neq 0$ do
                $J \leftarrow J \cup \{j\}$
            end for
        end for
        Tab$[I] \leftarrow J$
    end for
    return Tab
end procedure

procedure Find subspace trail($\mathcal{L}, r$)
    input : Matrix of the linear map $\mathcal{L}$ in the basis $(e_i)_{0 \leq i \leq N-1}$ and $r \geq 2$ the number of rounds.
    output : List of sets $(I, d_I)$ such that there exists an $r$-round trail starting from $\mathcal{E}_I$ ending in a
    space of dimension $d_I < N$.

    Create empty list L
    Tab $\leftarrow$ Essential trail($\mathcal{L}$)
    for all $\{0\} \subset I \subset [\![0, N-1]\!]$ do
        $J \leftarrow I$
        for $k = 0$ to $r - 2$ do
            $J \leftarrow Tab[J]$
        end for
        if $J \neq [\![0, N-1]\!]$ then
            Append $(I, |J|)$ to L
        end if
    end for
    return L
end procedure
```