

Weak-Key Distinguishers for AES

Lorenzo Grassi^{1,4}, Gregor Leander², Christian Rechberger¹,
Cihangir Tezcan³ and Friedrich Wiemer²

¹ IAIK, Graz University of Technology, Austria

firstname.lastname@iaik.tugraz.at

² Horst Görtz Institute for IT-Security, Ruhr-Universität Bochum, Germany

firstname.lastname@rub.de

³ Informatics Institute, Department of Cyber Security, CYDES Laboratory, and
Department of Mathematics, Middle East Technical University, Ankara, Turkey

cihangir@metu.edu.tr

⁴ Digital Security Group, Radboud University, Nijmegen

lgrassi@science.ru.nl

Abstract. In this paper, we analyze the security of AES in the case in which the whitening key is a weak key.

After a systematization of the classes of weak-keys of AES, we perform an extensive analysis of weak-key distinguishers (in the single-key setting) for AES instantiated with the original key-schedule and with the new key-schedule proposed at ToSC/FSE'18 (which is faster than the standard key schedule and ensures a higher number of active S-Boxes). As one of the main results, we show that (almost) all the secret-key distinguishers for round-reduced AES currently present in the literature can be set up for a higher number of rounds of AES if the whitening key is a weak-key.

Using these results as starting point, we describe a property for 9-round AES-128 and 12-round AES-256 in the chosen-key setting with complexity 2^{64} without requiring related keys. These new chosen-key distinguishers – set up by exploiting a variant of the multiple-of-8 property introduced at Eurocrypt'17 – improve all the AES chosen-key distinguishers in the single-key setting.

The entire analysis has been performed using a new framework that we introduce here – called “weak-key subspace trails”, which is obtained by combining invariant subspaces (Crypto'11) and subspace trails (FSE'17) into a new, more powerful, attack. Weak-key subspace trails are defined by extending the invariant subspace approach to allow for different subspaces in every round, something that so far only the subspace trail approach and a generalization for invariant subspace and invariant set attacks (Asiacrypt'18) were able to do. For an easier detection, we also provide an algorithm which finds these weak-key subspace trails.

Keywords: AES · Key Schedule · Weak-Keys · Invariant Subspaces · Chosen-Key Distinguisher

Contents

1	Introduction	2
1.1	Our Contribution	2
1.2	Related Work	4
2	Weak-Key (Invariant) Subspace Trails	5
2.1	Subspace Trails	5
2.2	Invariant Subspace Attacks	5
2.3	Weak-Key Subspace Trails	6
3	Preliminary – Subspace Trail Properties of the AES	7
3.1	Subspace Trails of AES	8
3.2	(Weak-Key) Invariant Subspace Trail for AES	9
3.3	Weak-Key Subspace Trail of AES: a Concrete Example	12
4	Weak-Key Secret-Key Distinguishers for AES	12
4.1	Subspace Trail Distinguishers	13
4.2	Weak-Key “Multiple-of- n ” Property for 5- and 6-round AES-128	14
4.3	Practical Experiments	16
5	New Chosen-Key Distinguishers for AES	17
5.1	Open-Key Distinguishers – State of the Art for AES	18
5.2	The “Simultaneous Multiple-of- n ” Property	19
5.3	9-round Chosen-Key Distinguisher for AES-128	19
5.4	Achieving the “Simultaneous Multiple-of- n ” Property Generically	20
5.5	“Simultaneous Properties” for 10-round (full) AES-128	21
A	Algorithmic Detection of Weak-Key Subspace Trails	I
B	Generic Subspace Trail (of length 1) for AES – Proof	II
C	Weak-Key Invariant Subspace Trails of AES-256	III
C.1	AES-256 Key-Schedule	III
C.2	Invariant Subspace – Weak-Keys of AES-256	III
D	Practical Collisions for 7-round AES-256 Compressing Modes	IV
E	Proofs of Results given in Section 4	VI
E.1	Proofs of Section 4.1	VI
E.2	Proofs of Weak-Key “Multiple-of- n ” – Theorem 5	IX
F	Proof of Proposition 3	X
G	On the Difficulty to set up “Multiple-of-n” Open-Key Distinguishers Without Relying on Weak-Keys	XI
H	Chosen-Key Distinguishers and “Simultaneous Property” for AES-128 and AES-256	XII
H.1	“Simultaneous Property” for 10-round AES	XIII
H.2	Chosen-key distinguisher for AES-256	XV

1 Introduction

Block ciphers are certainly among the most important cryptographic primitives. Their design and analysis are well advanced, and with today’s knowledge designing a secure block cipher is a problem that is largely considered solved. Especially with the AES we have at hand a very well analyzed and studied cipher that, after more than 20 years of investigation still withstands all cryptanalytic attacks. However, new results on the AES still appear regularly, especially within the last couple of years (e.g. [BR19; Bar+18; GRR17; Gra18; RBH17]). While those papers do not pose any practical threat to the AES, they do give new insights into the internals of what is arguably the cipher that is responsible for the largest fraction of encrypted data worldwide.

Clearly, security of symmetric crypto is always security against specific attacks. The number of available attacks has increased significantly ever since the introduction of differential [BS90] and linear [Mat94] cryptanalysis in the early 1990. Besides the numerous variations of linear and differential attacks, e.g. truncated differentials [Knu95], impossible differentials [BBS99; Knu98], zero-correlation attacks [BR14], and multidimensional linear cryptanalysis [HCN09] to name only a few, it turned out that in many cases combining two attack vectors might lead to new, more powerful attacks. The most prominent example is the combination of linear and differential cryptanalysis into differential-linear cryptanalysis [LH94].

Another important aspect is that the attacker model is regularly changing. With the introduction of statistical attacks, especially linear and differential cryptanalysis, the attacker was suddenly assumed to be able to retrieve, or even choose, large amounts of plaintext/ciphertext pairs. Later, in the related-key setting, the attacker became even more powerful and was assumed to be able to choose not only plaintexts but also ask for the encryption of chosen messages under a key that is related to the unknown secret key. Finally, in the open-key model, the attacker either knows the key or has the ability to choose the key herself.

While the practical impact of such models is often debatable, they actually might become meaningful when the block cipher is used as a building block for other primitives, in particular for the construction of hash-functions. Moreover, even if those considerations do not pose practical attacks, they still provide very useful insights and observations that strengthen our understanding of block ciphers in general, and on the AES in particular.

Our work builds upon the above in the sense that we combine previously separate attacks to derive new results on the AES, both in the secret- and in the open-key model.

1.1 Our Contribution

A key is said to be “weak” if, used with a specific cipher, it makes the cipher behave in some undesirable way (namely, if it makes the cipher weaker w.r.t. other keys). Even if weak keys usually represent a very small fraction of the overall key-space, it is desirable for a cipher to have no weak keys. The presence of a set of weak keys is usually related to the details of the key-schedule, namely the algorithm that takes as input a master key and outputs so-called round keys that are used in each round to mix the current state with the key. While the concrete security of the AES and other well-known ciphers is well studied, it is not clear what properties a good key schedule has to have. Even if there are some general guidelines on what a key schedule should not look like, these guidelines are rather basic and ensure mainly that trivial guess-and-determine or/and meet-in-the-middle attacks or/and structural attacks (e.g. slide-attacks, symmetries, invariant subspace attacks) are not possible.

Recently, more and more attacks on perfectly good ciphers [BKR11; Guo+16; KS99] – that exploit only weak-keys and key schedule weaknesses – indicate that the research on key schedule design principles is pressing. For the case in which the r -th round-key k_r is

simply defined by the XOR of the whitening key K and a round constant RC_r , that is $k_r = K \oplus RC_r$ (a key-schedule largely used for lightweight ciphers), in [Bei+17] authors show that a proper choice of round constants can easily avoid such properties related to structural attacks. In this paper, we first analyze *the security of AES instantiated with a weak-key* against secret-key distinguishers, both for the case of the AES key-schedule and for the case of a recent proposed key-schedule based only on permutation of the byte positions [Kho+17]. Then, we use these results as starting points in order to construct new chosen-key distinguishers for AES in the single-key model.

1.1.1 Systematization of Knowledge: Weak-key Subspace Trail Cryptanalysis

First of all, we start by recalling the basic set-up of subspace trail cryptanalysis (see [GRR16; GRR17; LTW18]) and invariant subspace attacks (see [LMR15; Lea+11]) in Section 2. Our first main focus is to point out the important differences of these two attacks. As we will explain, those concepts are not generalizations of each other but rather orthogonal attack vectors. From this point of view, a natural step is to *fill this gap, by combining both approaches into a new, more powerful, attack*. This is in line with what was done previously with other attacks as mentioned above.

As invariant subspace attacks are weak-key attacks by nature, the new attack originating from the combination of invariant subspace attacks and subspace trail cryptanalysis is a weak-key attack as well. Here, weak-key refers to the fact that the attacks do not work for any key, but rather only for a fraction of all keys (besides the fact that they heavily depend on details of the key-schedule). Consequently, in Section 2 we coin the new strategy *weak-key subspace trail cryptanalysis*. To be able to detect these trails, we provide an algorithmic way, based on previous search algorithms for invariant subspaces.

1.1.2 Weak-Key Secret-Key Distinguishers for AES

Previously, invariant subspace attacks were only applied to ciphers with very simple key schedule algorithms. As a result, ciphers where the round keys differed not only by round constants seemed secure against this type of attacks. E.g. up to now, it seemed impossible to apply invariant subspace attacks on the AES.

With our new combination of invariant subspace attacks and subspace trail cryptanalysis, we overcome this inherently difficult problem. As a showcase of the increased possibilities of our attack, and as the most important example anyway, in Sections 3.2 and 4 we present several new observations on the AES. Using as starting point the invariant subspace found by our algorithm and presented in Section 3.2, we show that (almost) all the secret-key distinguishers for round-reduced AES currently present in the literature can be set up for a higher number of rounds of AES *if the whitening key is a weak-key*.

In particular, we show that the secret-key distinguisher based on the multiple-of- n property proposed at Eurocrypt 2017 [GRR17] can be extended by one round if the (secret) whitening key is a weak-key. Moreover, we discuss and clarify the relation between “weak-key truncated differential” and “weak-key subspace trail”, and compare it with the strong connection that holds between subspace trails and truncated differentials (as highlighted in [BLN17; GRR16; LTW18]). As a concrete application of such results, in Appendix D we present examples of compression collisions for 6- and 7-round AES-256 used in Davies–Meyer, Miyaguchi–Preneel and Matyas–Meyer–Oseas construction.

As a side-result, we analyze the security of an alternative AES key schedule proposed at FSE/ToSC 2018 [Kho+17], which is defined by a permutation of the byte positions only and that aims to provide resistance against *related key-differential attacks*. In Section 3.2.3, we show the importance of adding random constants at every round in order to prevent the weak-key subspace trail attack proposed here.

1.1.3 Chosen-Key Distinguisher for AES

Building up on those results we are able to construct new chosen-key distinguishers for up to 9-round AES-128 and 12-round AES-256 in the single-key model and based on the multiple-of- n (weak-key) property. This improves all the chosen-key distinguishers for AES in the single-key setting. In particular, in Section 5 we exhibit *a chosen-key distinguisher with complexity 2^{64} for 9-round AES-128* in the single-key model¹, valid for 2^{32} keys.

For these results we combine two weak-key subspace trails in an inside-out manner and, instead of a simple truncated differential property at the plaintexts and ciphertexts, we use a variant of the “multiple-of- n ” property recently shown for AES in [GRR17].

Finally, we discuss the possibility to extend such results on full AES-128 and AES-256, namely we show a non-random property for full AES-128 and AES-256 in the chosen-key setting that seems difficult to produce generically.

1.2 Related Work

Regarding weak-key cryptanalysis, the most famous example of weak-keys is given for the DES. The block cipher DES has a few specific keys termed “weak-keys” and “semi-weak-keys” [MS87]. These are keys that cause the encryption mode of DES to act identically to the decryption mode of DES (albeit potentially that of a different key). Several other examples can be found in the literature, e. g. for Blowfish [KM07; Vau96], PRESENT [Ohk09], or Piccolo [WW16]. Typical “weak-key” attacks (so called as these attacks work only when a key of a special form is used, thus a “weak-key”) are the already mentioned invariant subspace attack [LMR15; Lea+11], the invariant set (or nonlinear invariant) attack [TLS16] and the recent work by Beyne [Bey18] for a generalization of invariant subspace and invariant set attacks.

Weak-keys are much more often a problem where the adversary has some control over what keys are used, such as when a block cipher is used in a mode of operation intended to construct a secure cryptographic hash function. For example, in the Davies-Meyer construction or the Miyaguchi-Preneel one can transform a secure block cipher into a secure compression function. In a hash setting, block cipher security models such as the known-key model (or the chosen-key model) makes sense since in practice the attacker has full access and control over the internal computations. A classical example is the devastating effect on the compression function security of weak-keys for a block cipher [Wei+12], which are usually considered as a minor flaw for a block cipher if the set of these weak-keys is small.

The idea of known-key distinguishers was introduced by Knudsen and Rijmen in [KR07] for their analysis of AES and a class of Feistel ciphers. They examined the security of these block ciphers in a model where the adversary knows the key. To succeed, the adversary has to discover some property of the attacked cipher that e. g. holds with a probability higher than for an ideal cipher, or is generally believed to be hard to exhibit generically. The idea of chosen-key distinguishers was popularized in the attack on the full-round AES-256 [BK09; BKN09] in a related-key setting. This time the adversary is assumed to have a full control over the key. A chosen-key attack was shown on 9-round reduced AES-128 in [FJP13] in the related-key setting, and on 8-round AES-128 in [DFJ12] in the single-key setting. Both the known-key and chosen-key distinguishers are collectively known as *open-key distinguishers*.

¹A 10-round known-key distinguisher for AES has been proposed by Gilbert [Gil14] at Asiacrypt 2014. Echoing Grassi and Rechberger [GR17], in Section 5.1.2 we argue why such distinguisher can be considered artificial. Briefly, the property of this distinguisher does not involve *directly* the plaintexts/ciphertexts, but their encryption/decryption after one round.

2 Weak-Key (Invariant) Subspace Trails

In the course of this section, we develop in a step by step manner the setting of “weak-key subspace trails” in a general and consistent way. Within our systematization, we first recapitulate the ideas and the differences between invariant subspaces and subspace trails.

2.1 Subspace Trails

Subspace trails have been first defined in [GRR16] and later used to attack e.g. reduced round versions of AES [GRR17], PRINCE [GR16] and Simpira [Røn16]. Moreover, the connection between subspace trails and truncated differential attacks has been studied in details in [LTW18]. We refer to [GRR16] for more details about the concept of subspace trails. Our treatment here is however meant to be self-contained.

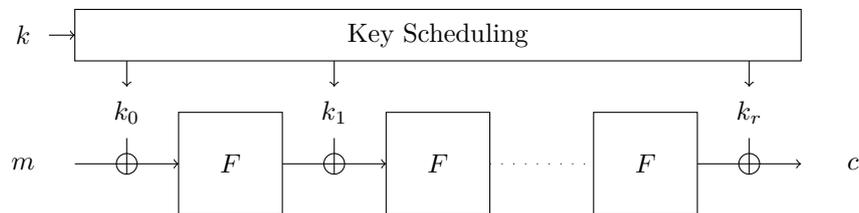
We recall the definition of a subspace trail next. For this, let F denote a round function of a key-alternating block cipher, and let $U \oplus a$ denote a coset of a vector space U . By U^c we denote the complementary subspace of U .

Definition 1 (Subspace Trails). Let $(U_1, U_2, \dots, U_{r+1})$ denote a set of $r + 1$ subspaces with $\dim(U_i) \leq \dim(U_{i+1})$. If for each $i = 1, \dots, r$ and for each a_i , there exists (unique) $a_{i+1} \in U_{i+1}^c$ such that

$$F(U_i \oplus a_i) \subseteq U_{i+1} \oplus a_{i+1},$$

then $(U_1, U_2, \dots, U_{r+1})$ is a *subspace trail* of length r for the function F . If all the previous relations hold with equality, the trail is called a *constant-dimensional subspace trail*.

One important observation is the following. Consider a key-alternating cipher E_k using F as a round function and where the round keys are xored in between the rounds, as depicted below:



In this case, a subspace trail for F will extend to a subspace trail for E_k for any choice of round keys. This is a simple consequence as

$$F(U_i \oplus a_i) \subseteq U_{i+1} \oplus a_{i+1} \text{ implies } F_{k^i}(U_i \oplus a_i) \equiv F(U_i \oplus a_i) \oplus k^i \subseteq U_{i+1} \oplus a'_{i+1}$$

for a suitable $a'_{i+1} = a_{i+1} \oplus k^i$ (where k^i is the i -th subkey). In other words, the key addition changes only the coset of the subspace U_{i+1} , while it does not affect the subspace itself. Thus, not only do subspace trails work for all keys, they are also completely independent of the key schedule. Here, invariant subspace attacks behave very differently.

2.2 Invariant Subspace Attacks

Invariant subspace attacks, which can be seen as a general way of capturing symmetries, have been first introduced in [Lea+11] in an attack on PRINTCIPHER. Later, those attacks have been applied to several other (mainly lightweight) primitives, e.g. in [LMR15], where a generic tool to detect them has been proposed.

As above, denoting by $F_k(\cdot) = F(\cdot) \oplus k$ the round function of a key-alternating block-cipher, let $U \subset \mathbb{F}_2^n$ be a subspace. Then, U is called an invariant subspace if there exist constants $a, b \in \mathbb{F}_2^n$ such that $F_k(U \oplus a) = U \oplus b$. In order to extend the invariant subspace

$U \oplus a_i \mapsto U \oplus a_{i+1}$ to the whole cipher, we need all round keys to be in specific cosets² of U namely, $k_i \in U \oplus (a_{i+1} \oplus b_i)$ (where $F(U \oplus a_i) = U \oplus b_i$):

$$F_k(U \oplus a_i) = F(U \oplus a_i) \oplus k = U \oplus b_i \oplus \underbrace{k}_{\in U \oplus (a_{i+1} \oplus b_i)} = U \oplus a_{i+1}.$$

Definition 2 (Invariant Subspace Trail). Let K_{weak} be a set of keys and $k \in K_{\text{weak}}$, with $k \equiv (k^0, k^1, \dots, k^r)$ where k^j is the j -th round key. For each $k \in K_{\text{weak}}$, the subspace U generates an *invariant subspace trail* of length r for the function $F_k(\cdot) \equiv F(\cdot) \oplus k$ if for each $i = 1, \dots, r$ there exists a non-empty set $A_i \subseteq U^c$ for which the following property holds:

$$\forall a_i \in A_i : \exists a_{i+1} \in A_{i+1} \text{ s.t. } F_{k^i}(U \oplus a_i) \equiv F(U \oplus a_i) \oplus k^i = U \oplus a_{i+1}.$$

All keys in the set K_{weak} are *weak-keys*.

2.3 Weak-Key Subspace Trails

When comparing subspace trail and invariant subspace attacks, two obvious but important differences can be observed. First, subspace trails are clearly much more general as they allow different spaces in the domain and co-domain of F . Second, subspace trails are by far more restrictive, as not only one coset of the subspace has to be mapped to one coset of (a potentially different) subspace, but rather all cosets have to be mapped to cosets. For subspace trails, the later fact is the main reason for allowing arbitrary round keys.

The main idea for weak-key subspace trails is to stick to the property of invariant subspace attacks where only few (even just one) cosets of a subspace are mapped to other cosets of a subspace. However, borrowing from subspace trails, we allow those subspaces to be different for each round. As this will again restrict the choice of round keys that will keep this property invariant to a class of weak-keys we call this combination *weak-key subspace trails* (or simply, weak subspace trails). The formal definition is the following.

Definition 3 (Weak-Key Subspace Trails). Let K_{weak} be a set of keys and $k \in K_{\text{weak}}$ with $k \equiv (k^0, k^1, \dots, k^r)$ where k^j is the j -th round key. Further let $(U_1, U_2, \dots, U_{r+1})$ denote a set of $r + 1$ subspaces with $\dim(U_i) \leq \dim(U_{i+1})$. For each $k \in K_{\text{weak}}$, $(U_1, U_2, \dots, U_{r+1})$ is a *weak-key subspace trail* (WKST) of length r for the function $F_k(\cdot) \equiv F(\cdot) \oplus k$ if for each $i = 1, \dots, r$ there exists a non-empty set $A_i \subseteq U_i^c$ for which the following property holds:

$$\forall a_i \in A_i : \exists a_{i+1} \in A_{i+1} \text{ s.t. } F_{k^i}(U_i \oplus a_i) \equiv F(U_i \oplus a_i) \oplus k^i \subseteq U_{i+1} \oplus a_{i+1}.$$

All keys in the set K_{weak} are *weak-keys*. If all the previous relations hold with equality, the trail is called a *weak-key constant-dimensional subspace trail*.

Usually, the set $A_i \subseteq U_i^c$ reduces to a single element a_i , that is $A_i \equiv \{a_i\}$. Moreover, we can easily see that Definition 3 is a generalization of both Definitions 1 and 2:

- if K_{weak} is equal to the whole set of keys and if $A_i = U_i^c$, then it corresponds to subspace trails;
- if $U_i = U_{i+1}$ for all i , then it corresponds to invariant subspace trails.

²It is not necessary that $a_i = a_{i+1}$ for all i in order to set up an invariant subspace attack.

Security Problem. Clearly, a WKST allows greater freedom for an attacker. In comparison to invariant subspace attacks, WKSTs have the potential of being better applicable to block ciphers with a non trivial key schedule. At the same time, with respect to subspace trails it is not necessary for WKSTs to hold for all possible keys.

Interestingly, proving resistance against invariant subspace (or more generally invariant sets) in the case of identical round keys (up to the addition of round constants) is well understood, see [Bei+17]. However, the situation changes completely when considering WKSTs and/or ciphers with a non-trivial key schedule. In those situations, the analysis of [Bei+17] is no longer applicable and we do not have a generic approach to argue the resistance against WKSTs.

It follows that the concept of WKSTs opens up many new opportunities and raises many new, probably highly non-trivial questions on how to protect against it. In [Appendix A](#), we take a look at how we can find these algorithmically, while in the following we focus on demonstrating the new opportunities by investigating the AES. How to (generically) protect against those attacks is left as an open question for future research.

3 Preliminary – Subspace Trail Properties of the AES

The Advanced Encryption Standard [DR02] is a *Substitution-Permutation network* that supports key sizes of 128, 192 and 256 bits. The 128-bit plaintext initializes the internal state as a 4×4 matrix of bytes as values in the finite field \mathbb{F}_{256} , defined using the irreducible polynomial $x^8 + x^4 + x^3 + x + 1$. Depending on the version of AES, N_r rounds are applied to the state: $N_r = 10$ for AES-128, $N_r = 12$ for AES-192 and $N_r = 14$ for AES-256. One round of AES can be described as $R(x) = K \oplus \text{MC} \circ \text{SR} \circ \text{SB}(x)$, where

- *SubBytes* (SB) – applying the same 8-bit to 8-bit invertible S-Box 16 times in parallel on each byte of the state (it provides non-linearity in the cipher);
- *ShiftRows* (SR) – cyclic shift of each row to the left;
- *MixColumns* (MC) – multiplication of each column by a constant 4×4 invertible matrix M_{MC} (MC and SR provide diffusion in the cipher);
- *AddRoundKey* (ARK) – XORing the state with a 128-bit subkey.

In the first round an additional AddRoundKey operation (using a whitening key) is applied, and in the last round the MixColumns operation is omitted.

Key Schedule AES-128. As we consider only AES with 128-bit key in the main part of this paper, we shall describe only its key schedule algorithm. The key schedule of AES-128 takes the user key and transforms it into 11 subkeys of 128 bits each. The subkey array is denoted by $W[0, \dots, 43]$, where each word of $W[\cdot]$ consists of 4 bytes and where the first 4 words of $W[\cdot]$ are loaded with the user secret key. The remaining words of $W[\cdot]$ are updated according to the following rule:

$$W[i][j] = \begin{cases} W[i][j-4] \oplus \text{SB}(W[i+1][j-1]) \oplus R[i][j/4] & \text{if } j \bmod 4 = 0 \\ W[i][j-1] \oplus W[i][j-4] & \text{otherwise} \end{cases}$$

where $i = 0, 1, 2, 3$, $j = 4, \dots, 43$ and $R[\cdot]$ is an array of predetermined constants.³

³The round constants are defined in $GF(2^8)[X]$ as $R[0][1] = X$, $R[0][r] = X \cdot R[0][r-1]$ if $r \geq 2$ and $R[i][\cdot] = 0$ if $i \neq 0$. For the following, let $R[r] \equiv R[0][r]$.

The Notation used in the Paper. Let x denote a plaintext, a ciphertext, an intermediate state or a key. Then $x_{i,j}$ or $x_{i+4 \times j}$ with $i, j \in \{0, \dots, 3\}$ denotes the byte in the row i and in the column j . We denote by k^r the key of the r -th round. If only one key is used, then we denote it by k to simplify the notation. Finally, we denote by R one round of AES, while we denote r rounds of AES by R^r . We sometimes use the notation R_K instead of R to highlight the round key K . As last thing, in the paper we often use the term ‘‘partial collision’’ (or ‘‘collision’’) when two texts belong to the same coset of a given subspace X .

3.1 Subspace Trails of AES

In this section, we recall the main concepts of the subspace trails of AES presented in [GRR16]. In the following, we only work with vectors and vector spaces over $\mathbb{F}_{2^8}^{4 \times 4}$, and we denote by $\{e_{0,0}, \dots, e_{3,3}\}$ or $\{e_0, \dots, e_{15}\}$ the unit vectors of $\mathbb{F}_{2^8}^{4 \times 4}$ (e.g. $e_{i,j}$ or $e_{i+4 \times j}$ has a single 1 in row i and column j). We also recall that given a subspace X , the cosets $X \oplus a$ and $X \oplus b$ (where $a \neq b$) are equal ($X \oplus a \equiv X \oplus b$) if and only if $a \oplus b \in X$.

Definition 4. The *column spaces* \mathcal{C}_i are defined as $\mathcal{C}_i = \langle e_{0,i}, e_{1,i}, e_{2,i}, e_{3,i} \rangle$.

Definition 5. The *diagonal spaces* \mathcal{D}_i and the *inverse-diagonal spaces* \mathcal{ID}_i are respectively defined as $\mathcal{D}_i = \text{SR}^{-1}(\mathcal{C}_i) \equiv \langle e_{0,i}, e_{1,i+1}, e_{2,i+2}, e_{3,i+3} \rangle$ and $\mathcal{ID}_i = \text{SR}(\mathcal{C}_i) \equiv \langle e_{0,i}, e_{1,i-1}, e_{2,i-2}, e_{3,i-3} \rangle$, where the indexes are taken modulo 4.

Definition 6. The *i -th mixed spaces* \mathcal{M}_i are defined as $\mathcal{M}_i = \text{MC}(\mathcal{ID}_i)$.

Definition 7. For $I \subseteq \{0, 1, 2, 3\}$, let \mathcal{C}_I , \mathcal{D}_I , \mathcal{ID}_I and \mathcal{M}_I be defined as

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

For completeness, we briefly describe the subspace trail notation using a more ‘‘classical’’ one. If two texts t^1 and t^2 are equal except for the bytes in the i -th diagonal⁴ for each $i \in I$, then they belong in the same coset of \mathcal{D}_I . Two texts t^1 and t^2 belong in the same coset of \mathcal{M}_I if the bytes of their difference $\text{MC}^{-1}(t^1 \oplus t^2)$ in the i -th anti-diagonal for each $i \notin I$ are equal to zero. Similar considerations hold for the spaces \mathcal{C}_I and \mathcal{ID}_I .

As shown in detail in [GRR16]:

- for any coset $\mathcal{D}_I \oplus a$ there exists unique $b \in \mathcal{C}_I^\perp$ such that $R(\mathcal{D}_I \oplus a) = \mathcal{C}_I \oplus b$;
- for any coset $\mathcal{C}_I \oplus a$ there exists unique $b \in \mathcal{M}_I^\perp$ such that $R(\mathcal{C}_I \oplus a) = \mathcal{M}_I \oplus b$.

Theorem 1 ([GRR16]). *For each I and for each $a \in \mathcal{D}_I^\perp$, there exists one and only one $b \in \mathcal{M}_I^\perp$ such that*

$$R^2(\mathcal{D}_I \oplus a) = \mathcal{M}_I \oplus b.$$

Observe that if X is a generic subspace, $X \oplus a$ is a coset of X and x and y are two elements of the (same) coset $X \oplus a$, then $x \oplus y \in X$. It follows that:

Lemma 1. *For all x, y and for all $I \subseteq \{0, 1, 2, 3\}$:*

$$\Pr [R^2(x) \oplus R^2(y) \in \mathcal{M}_I \mid x \oplus y \in \mathcal{D}_I] = 1.$$

As demonstrated in [GRR16], we finally recall that for each $I, J \subseteq \{0, 1, 2, 3\}$: $\mathcal{M}_I \cap \mathcal{D}_J = \{0\}$ if and only if $|I| + |J| \leq 4$. It follows that

⁴The i -th diagonal of a 4×4 matrix A is defined as the elements that lie on row r and column c such that $r - c = i \pmod 4$. The i -th anti-diagonal of a 4×4 matrix A is defined as the elements that lie on row r and column c such that $r + c = i \pmod 4$.

Theorem 2 ([GRR16]). *Let $I, J \subseteq \{0, 1, 2, 3\}$ such that $|I| + |J| \leq 4$. For all $x \neq y$:*

$$\Pr [R^4(x) \oplus R^4(y) \in \mathcal{M}_I \mid x \oplus y \in \mathcal{D}_J] = 0.$$

Finally, for the follow-up, we introduce a generic subspace trail of length 1.

Definition 8. Let I be a subset of $\{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$. Let the subspace \mathcal{X}_I be defined as $\mathcal{X}_I = \langle \{e_{i,j}\}_{(i,j) \in I} \rangle \equiv \left\{ \bigoplus_{(i,j) \in I} \alpha_{i,j} \cdot e_{i,j} \mid \forall \alpha_{i,j} \in \mathbb{F}_{2^8} \right\}$.

In other words, \mathcal{X}_I is the set of elements given by linear combinations of $\{e_{i,j}\}_{(i,j) \in I}$, where $e_{i,j} \in \mathbb{F}_{2^8}^{4 \times 4}$ has a single 1 in row i and column j .

Theorem 3. *For each $I \subseteq \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$ and for each $a \in \mathcal{X}_I^\perp$, there exists one and only one $b \in \mathcal{Y}_I^\perp$ such that $R(\mathcal{X}_I \oplus a) = \mathcal{Y}_I \oplus b$, where $\mathcal{Y}_I = \text{MC} \circ \text{SR}(\mathcal{X}_I)$.*

The complete proof⁵ of this Theorem is given in [Appendix B](#). Such subspace trail cannot be extended on two rounds for any generic \mathcal{X}_I , due to the non-linear S-Box operation of the next round (that *can* destroy the linear relations that hold among the bytes).

3.2 (Weak-Key) Invariant Subspace Trail for AES

In this section, we present a subspace \mathcal{IS} which is invariant for a key-less AES round, and a set of weak-keys for AES-128 that allows to set up an invariant subspace trail for 2-round AES-128. Similar results – presented in [Appendix C](#) – can be provided for AES-192 and AES-256. Finally, we discuss a weakness of an alternative linear key-schedule for AES-128 proposed at ToSC/FSE 2018 [Kho+17], based on permutations of the byte positions.

3.2.1 Invariant Subspace \mathcal{IS} for AES

Let the subspace \mathcal{IS} be defined as

$$\mathcal{IS} := \left\{ \begin{bmatrix} a & b & a & b \\ c & d & c & d \\ e & f & e & f \\ g & h & g & h \end{bmatrix} \mid \forall a, b, c, d, \dots, h \in \mathbb{F}_{2^8} \right\} \quad (1)$$

This subspace is invariant under a key-less round $R(\cdot) = \text{MC} \circ \text{SR} \circ \text{SB}(\cdot)$, since

$$\text{SB}(\mathcal{IS}) = \mathcal{IS} \quad \text{SR}(\mathcal{IS}) = \mathcal{IS} \quad \text{MC}(\mathcal{IS}) = \mathcal{IS}.$$

This subspace – already presented and used in e.g. [Cha+17; Guo+13; Le+04] – can easily be found by extending the result of [Algorithm 1](#). It will be our starting point in order to set up a weak-key invariant subspace trail for all versions of AES.

3.2.2 Weak-Keys of AES-128 & Invariant Subspace Trail

In the case of the AES key-schedule, under one of the 2^{32} weak-keys in K_{weak}

$$K_{\text{weak}} := \left\{ \begin{bmatrix} A & A & A & A \\ B & B & B & B \\ C & C & C & C \\ D & D & D & D \end{bmatrix} \mid \forall A, B, C, D \in \mathbb{F}_{2^8} \right\} \quad (2)$$

⁵Here we limit ourselves to highlight that for each $I \subseteq \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$, there exists $J \subseteq \{(i, j)\}_{0 \leq i, j \leq 3}$ such that $\text{SR}(\mathcal{X}_I) = \mathcal{X}_J$ (or equivalently $\text{SR}^{-1}(\mathcal{X}_I) = \mathcal{X}_J$).

the subspace \mathcal{IS} is mapped into a coset of \mathcal{IS} after two complete AES rounds.

In more details, given $k \in K_{\text{weak}}$, let \hat{k} be the corresponding subkey after 2 rounds of the key schedule (where $\hat{k} \notin K_{\text{weak}}$ in general). It follows that

$$\mathcal{IS} \xrightarrow{R_K^2 \circ \text{ARK}(\cdot)} \mathcal{IS} \oplus \hat{k}$$

where $R_K(\cdot) \equiv \text{ARK} \circ \text{MC} \circ \text{SR} \circ \text{SB}(\cdot)$, that is \mathcal{IS} forms a weak invariant subspace of length 2. In order to prove this result, it is sufficient to note that

1. $K_{\text{weak}} \subseteq \mathcal{IS}$, which implies that $\mathcal{IS} \oplus k = \mathcal{IS}$ for all $k \in K_{\text{weak}}$;
2. the first round key derived from the key-schedule of K_{weak} – denoted by K'_w – is a subset of \mathcal{IS}

$$K'_w \equiv \begin{bmatrix} \text{SB}(B) \oplus A \oplus R[1] & \text{SB}(B) \oplus R[1] & \text{SB}(B) \oplus A \oplus R[1] & \text{SB}(B) \oplus R[1] \\ \text{SB}(C) \oplus B & \text{SB}(C) & \text{SB}(C) \oplus B & \text{SB}(C) \\ \text{SB}(D) \oplus C & \text{SB}(D) & \text{SB}(D) \oplus C & \text{SB}(D) \\ \text{SB}(A) \oplus D & \text{SB}(A) & \text{SB}(A) \oplus D & \text{SB}(A) \end{bmatrix}$$

3.2.3 Key Schedules based on Permutation of the Byte Positions

The possibility to set up a weak invariant subspace trail depends on the concrete value of the secret key and of the key schedule details. To better understand this point, here we analyze another key-schedule recently proposed at ToSC/FSE 2018 [Kho+17] *in the case in which no random round-constant is added*. Such a key-schedule is linear and it is based on permutations of the byte positions: each subkey is the result of a particular permutation applied to the whitening key defined as follows

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

Such key-schedule has been proposed with the only goal to provide resistance against *related key-differential attacks*: in [Kho+17, Sect. 6], authors “*emphasize that [their] goal here is to design a simple yet efficient key schedule that maximizes the number of active Sboxes in the related-key model.*”.

In the case in which random round-constants are added, an invariant subspace attack that covers an unlimited number of rounds⁶ is very unlikely, as showed e.g. in [Bei+17] (for the case of other ciphers). Hence, by adding random constants at every round, such key-schedule is perfectly fine and could be a good candidate for future designs.

To emphasize the importance of such random round-constant addition, here we point out that it is always possible to set up an “infinitely-long” weak invariant subspace in the case in which *no random round-constant is added*. Indeed, consider the previous subspace \mathcal{IS} defined in Eq. (1) and *assume that the whitening key belongs to such subspace*. It follows that any subkey generated by the previous permutation belongs to this subspace (due to particular symmetries of the permutation), which implies the possibility to set up an “infinitely-long” weak invariant subspace

$$\mathcal{IS} \xrightarrow{\text{MC} \circ \text{SR} \circ \text{SB}(\cdot)} \mathcal{IS} \xrightarrow{\cdot \oplus k[r]} \mathcal{IS} \quad (4)$$

for a set of 2^{64} weak-keys $K \in \mathcal{IS}$ (for which all round-keys $k[r]$ are in \mathcal{IS}).

⁶Note that even in the case in which random round-constants are added, then it is potentially possible to set up a *finite* weak invariant subspace trail (namely, that covers a limited number of rounds).

Adding a (partial) S-Box Layer. Besides adding random round-constants, another possible way to prevent such invariant subspace attack is by adding non-linear operations in the key-schedule. In [Kho+17, Sect. 6], authors propose to “*tweak this design (without increasing the tracking effort) by adding an S-Box layer every round to the entire first row of the key state*” (with the goal to increment the security against related-key differential attacks). Due to the analysis just proposed and only in the case in which no round-constant is added, here we show that this operation does not improve the security against the presented invariant subspace attack. Indeed, note that the invariant subspace \mathcal{IS} is still mapped into itself *even if* an S-Box layer is applied to the entire first row of the key state:

$$\begin{bmatrix} \text{SB}(a) & \text{SB}(b) & \text{SB}(a) & \text{SB}(b) \\ c & d & c & d \\ e & f & e & f \\ g & h & g & h \end{bmatrix} = \begin{bmatrix} a' & b' & a' & b' \\ c & d & c & d \\ e & f & e & f \\ g & h & g & h \end{bmatrix} \in \mathcal{IS}.$$

We emphasize that this problem can be easily fixed by applying such an S-Box layer every round to the entire (e.g.) first column/diagonal. Indeed, in this last case, such S-Box layer destroys the invariant subspace trail after few rounds. As a result, even in the case in which *no* random round-constant are added, the partial S-Box layer applied every round to the entire first column/diagonal⁷ is sufficient by itself to prevent “*infinitely-long*” weak invariant subspace trails based on \mathcal{IS} .

Follow-Up Works: Key-Schedule based on Permutation. After the initial work [Kho+17], other key-schedules based only on permutations have been recently proposed at SAC 2018 [Der+18]. Here we focus on the one proposed in [Der+18, Theorem 2], and defined by the following byte-permutation:

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

As for the key-schedule proposed by Eq. (3), such key-schedule guarantees more security than the AES one w.r.t. related-key differential attacks (at least 16 S-Boxes are active over 5 rounds). At the same time, with respect to the key-schedule proposed in [Kho+17] and only in the case in which no random round-constant is added, here an “*infinitely-long*” invariant subspace trail can be set up for a set of 2^8 weak keys only, namely the whitening key that belong into

$$\mathcal{IS}^{(trivial)} := \left\{ \begin{bmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix} \mid \forall x \in \mathbb{F}_{2^8} \right\}.$$

This highlights an interesting difference between the key-schedule proposed in [Kho+17] and the follow-up one proposed in [Der+18]. For every AES-like cipher and for every key-schedule based only on permutation of the bytes and for which no random round-constant is added, it is *always* possible to set up a *trivial* “*infinitely-long*” invariant subspace trail for 2^8 weak-keys of the form $\mathcal{IS}^{(trivial)}$. The crucial point is that in the case of the key-schedule proposed in [Kho+17], it is also possible to construct “*infinitely-long*” invariant subspace trails that are *non-trivial* (namely, for the case of 2^{64} weak-keys instead of just 2^8 ones).

Before going on, we remark one more time that both key-schedules are perfectly secure and good candidates for future designs if random round-constants are added.

⁷For completeness, we emphasize that the same result holds in the case of the original AES key-schedule without random constants. In other words, even if one removes the round-constant additions in the original AES key-schedule, the possibility to set up an infinite long invariant subspace trails is prevented by the presence of a partial S-Box layer in the key-schedule.

3.3 Weak-Key Subspace Trail of AES: a Concrete Example

Finally, we present a (proper) weak-key subspace trail for AES. For simplicity, we limit ourselves to focus on an S-Box defined by the inverse operation⁸ namely

$$\forall x \in GF(2^8) : \quad \mathbb{S}(x) = \begin{cases} 1/x \equiv x^{254}, & \text{if } x \neq 0, \\ 0 & \text{otherwise} \end{cases}$$

To achieve our goal, we look for subspaces $V, W \subset GF(2^8)$ of dimension two and/or four s. t.

$$\mathbb{S}(V \oplus v) \subseteq W \oplus w$$

for *certain* (not all) $v, w \in GF(2^8)$, where $V \neq W$ in general.

E. g. the subspace V of dimension four defined as $V = \langle [0x01, 0x0c, 0x50, 0xb0], 0 \rangle = \{x \in GF(2^8) \mid x^{256} \oplus x = 0\}$ is invariant under the S-Box – that is, $\mathbb{S}(V) = V$, since $\mathbb{S}(x)^{256} \oplus \mathbb{S}(x) = [(x^{254})]^{256} \oplus x^{254} = x^{254} \oplus x^{254} = 0$ (remember that $x^{2^n-1} = 1$ for all $x \in GF(2^n)$), while its cosets $V \oplus v$ for $v \neq 0$ are in general not invariant. In [BWP05], several subspaces $V, W \subset GF(2^8)$ of dimension two and four are defined such that $V \neq W$ and $\mathbb{S}(V \oplus v) \subseteq W \oplus w$, like $\mathbb{S}(V \equiv \langle [0x02, 0x18, 0x61, 0xa0], 0 \rangle) = (W \equiv \langle [0x06, 0x28, 0x58, 0x8b], 0 \rangle)$. All these subspaces can be used to set up a weak-subspace trail for 1-round AES, e.g.

$$\mathcal{V} \oplus x \equiv \begin{bmatrix} V & x_{0,1} & x_{0,2} & x_{0,3} \\ x_{1,0} & V & x_{1,2} & x_{1,3} \\ x_{2,0} & x_{2,1} & V & x_{2,3} \\ x_{3,0} & x_{3,1} & x_{3,2} & V \end{bmatrix} \xrightarrow{\text{MC} \circ \text{SR} \circ \mathbb{S}(K^w \oplus \cdot)} \mathcal{W} \oplus y \equiv \begin{bmatrix} W & y_{0,1} & y_{0,2} & y_{0,3} \\ W & y_{1,1} & y_{1,2} & y_{1,3} \\ W & y_{2,1} & y_{2,2} & y_{2,3} \\ W & y_{3,1} & y_{3,2} & y_{3,3} \end{bmatrix}$$

for random value of $x \in \mathcal{D}_0^\perp \equiv \mathcal{D}_{1,2,3}$ (where $y \in \mathcal{C}_{1,2,3}$). Similarly, it is possible to set up different and longer weak-key subspace trails.

The crucial point here is that *this weak-key subspace trail is neither a subspace trail nor an invariant subspace trail*, since:

- initial and final subspaces are different (hence, it's not an invariant subspace trail);
- it is not possible to choose any arbitrary coset of \mathcal{V} , since each byte in the first diagonal must belong to the subspace $\langle [0x02, 0x18, 0x61, 0xa0] \rangle$ (hence, they cannot take any possible value as in the case of a subspace trail);
- the class of weak keys K^w corresponds to the subspace $\mathcal{V} \oplus \mathcal{D}_{1,2,3}$ (where $\mathcal{V} \subset \mathcal{D}_0$), where each byte of the key in the first diagonal belongs to the subspace $\langle [0x02, 0x18, 0x61, 0xa0] \rangle$, while the bytes in the other diagonals can take any possible values (hence, it does not hold for any key as in the case of a subspace trail).

4 Weak-Key Secret-Key Distinguishers for AES

As a first application of the invariant subspaces just found, we are going to show that *under the assumption of weak-keys* it is possible to extend the secret-key distinguishers present in the literature to more rounds. In the following, we present in detail only the results for AES-128 for the encryption/forward direction (analogous results hold also in the decryption/backward direction). Similar results can be obtained also for AES-192 and AES-256, using the corresponding weak-keys and weak-key invariant subspace trails defined in Appendix C. The results – which have been practically tested using a C/C++ implementation – are summarized in Table 1. As concrete example of practical use cases of these results, in Appendix D we present collisions for Matyas-Meyer-Oseas, Davies-Meyer and Miyaguchi-Preneel compressing modes instantiated with 6- and 7-round AES-256.

⁸Analogous results can be obtained for the real AES, since the AES S-Box is affine equivalent to $\mathbb{S}(x)$, that is $\text{AES-S-Box}(x) = M \cdot \mathbb{S}(x) \oplus 0x63$ where M is a 8×8 binary (invertible) matrix.

Table 1: *Secret-key properties for round-reduced AES.* In the following, we list the properties for round-reduced AES which are independent of the secret key, together with the corresponding number of rounds. “Number of keys” denotes the number of keys (with respect to the total space) for which a particular property holds for up to r rounds. Just for simplicity, we do not add the distinguisher complexity (or equivalently, the probability of the exploited property).

Property	Version of AES	Rounds	Number of keys	Reference
Weak-key Subspace Trail	AES-128/256	3	All: $2^{128} / 2^{256}$	folklore
	AES-128	5	All: 2^{128}	[GR18]
	AES-128/256	5	$2^{32} / 2^{128}$	§ 4.1
	AES-128	6	All: 2^{128}	[BGL19; BR19]
	AES-256	6/7/8	$2^{96} / 2^{64} / 2^{32}$	based on App. C.1 & § 4.1
Multiple-of- n	AES-128/256	5	All: $2^{128} / 2^{256}$	[GRR17]
	AES-128/256	6	$2^{32} / 2^{128}$	§ 4.2
	AES-256	7/8/9	$2^{96} / 2^{64} / 2^{32}$	based on App. C.1 & § 4.2

Assumption. *From now on we assume that the secret key is a weak-key* (that is, a key in the set K_{weak} as described previously). We also emphasize that all the following results are independent of the details of the S-Box and of the MixColumns operation.

4.1 Subspace Trail Distinguishers

In the case of AES, it is possible to set up subspace trail distinguishers for 3 and 4-round AES. Both are *independent* of the secret-key, of the details of the S-Box and of the MixColumns matrix (assuming branch number equal to five). In particular, the first one exploits the fact that

$$\Pr [R^3(x) \oplus R^3(y) \in \mathcal{M}_J \mid x \oplus y \in \mathcal{D}_I] = (2^8)^{-4|I|+|I|\cdot|J|}$$

as showed in detail in [GRR16], while for a random permutation Π the previous probability is (approximately) equal to

$$\Pr [\Pi(x) \oplus \Pi(y) \in \mathcal{M}_J \mid x \oplus y \in \mathcal{D}_I] = (2^8)^{-16+4|J|}. \quad (5)$$

In the following, we show that it is possible to extend the previous subspace trail distinguisher for up to 5 rounds in the case of weak-keys. Focusing on the case of AES-128, we have just seen that the subspace \mathcal{IS} is mapped into a coset $\mathcal{IS} \oplus a$ after two rounds if the secret key is a weak-key. In other words, given two plaintexts $x, y \in \mathcal{IS}$, then $R^2(x) \oplus R^2(y) \in \mathcal{IS}$ under a weak-key. By definition of \mathcal{IS} and \mathcal{D}_I , note that⁹

$$\Pr [z \in \mathcal{D}_I \mid z \in \mathcal{IS}] = \begin{cases} 2^{-32} & I \equiv \{0, 2\}, \{1, 3\} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where we assume that $z \notin \mathcal{D}_L$ for all $L \subseteq \{0, 1, 2, 3\}$ s.t. $|L| < |I| < 4$. This is the starting point for our results, together with the fact that $\Pr [z \in \mathcal{D}_{0,2}] = \Pr [z \in \mathcal{D}_{1,3}] = 2^{-64}$ for a generic text z .

⁹Note: the 1st and the 3rd diagonals of each text in \mathcal{IS} are equal, as well as the 2nd and the 4th ones.

4.1.1 Weak-Key Subspace Trail over 4-round AES-128

Since $R^2(\mathcal{D}_I \oplus a) = \mathcal{M}_I \oplus b$ (that is $\Pr[R^2(x) \oplus R^2(y) \in \mathcal{M}_I \mid x \oplus y \in \mathcal{D}_I] = 1$), it follows that for an AES permutation and for a weak-key

$$\Pr[R^4(x) \oplus R^4(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}, k \in K_{\text{weak}}] = 2^{-32} \quad \text{if } I \equiv \{0, 2\}, \{1, 3\},$$

while for a random permutation Π the probability is equal to 2^{-64} (see Eq. (5)).

This fact can also be re-written using the subspace trail notation.

Proposition 1. *Consider 2^{64} plaintexts in the subspace \mathcal{IS} , and the corresponding ciphertexts after 4-rounds AES-128 encrypted under a weak-key $k \in \mathcal{K}_{\text{weak}}$.*

With probability 1 the ciphertexts are distributed as follows:

- *there exist 2^{32} (in 2^{64}) different cosets of $\mathcal{M}_{0,2}$ s.t. each one of them contains exactly 2^{32} ciphertexts;*
- *there exist 2^{32} (in 2^{64}) different cosets of $\mathcal{M}_{1,3}$ s.t. each one of them contains exactly 2^{32} ciphertexts.*

For a random permutation, each one of the previous events is satisfied with probability $\binom{2^{64}}{2^{32}} \cdot \prod_{i=0}^{2^{32}-1} \left[(2^{-64})^{2^{32}-1} \cdot (1 - i \cdot 2^{-64}) \right] \approx 2^{-270}$.

A complete proof of this proposition can be found in [Appendix E.1.1](#).

4.1.2 Weak-Key Subspace Trails over 5-round AES-128

Exploiting the fact that $\Pr[x \in \mathcal{C}_J \mid x \in \mathcal{M}_I] = (2^8)^{-4|I|+|I|\cdot|J|}$ together with Eq. (6), it is possible to set up a 5-round weak-key subspace trail distinguisher on AES.

Proposition 2. *Let $J \subseteq \{0, 1, 2, 3\}$ fixed. Then, the following probability holds:*

$$\Pr[R^5(x) \oplus R^5(y) \in \mathcal{M}_J \mid x, y \in \mathcal{IS}, k \in K_{\text{weak}}] = 2^{-128+32\cdot|J|} + 2^{-95+16\cdot|J|} \quad (7)$$

(with an error of magnitude 2^{-128}).

A complete proof of this proposition just given can be found in [Appendix E.1.2](#). Since for a random permutation Π the probability of the same event is equal to $\frac{2^{32\cdot|J|-1}}{2^{128}-1} \approx 2^{-128+32\cdot|J|}$ (where the approximation holds with an error of magnitude 2^{-128}), it is possible to distinguish the two cases for e.g. $|J| = 3$.

4.2 Weak-Key “Multiple-of- n ” Property for 5- and 6-round AES-128

At Eurocrypt 2017, Grassi et al. [GRR17] presented the first property on 5-round AES (and AES-like ciphers) which is independent of the secret key and of the details of the S-Box and of the MixColumns. The result can be summarized as follows: Given $2^{32\cdot|I|}$ plaintexts in the same coset of a diagonal space \mathcal{D}_I , the number of different pairs of ciphertexts that belong to the same coset of \mathcal{M}_J after 5-round AES is always a multiple of 8. The “multiple-of-8” property is related to the “mixture differential” cryptanalysis presented in [Gra18], and recently re-visited in [BCC19], where authors show that the above property is an immediate consequence of an equivalence relation on the input pairs, under which the difference at the output of the round function is invariant.

In the case of a weak-key, we are able to extend the previous result for up to 6-round AES-128. The obtained results – which hold also in the *decryption* direction – are proposed in the following Theorems:

Theorem 4. Let \mathcal{IS} and \mathcal{M}_I be the subspaces defined as before for a fixed I with $1 \leq |I| \leq 3$. Assume that the whitening key is a weak-key, that is it belongs to the set K_{weak} as defined in Eq. (2). Given 2^{64} plaintexts in \mathcal{IS} , the number n of different pairs¹⁰ of ciphertexts ($c^i = R^5(p^i), c^j = R^5(p^j)$) after 5-round AES for $i \neq j$ that belong to the same coset of \mathcal{M}_I (that is $c^i \oplus c^j \in \mathcal{M}_I$) is a multiple of 128 (that is, $\exists n' \in \mathbb{N}$ such that $n = 128 \cdot n'$), independently of the details of the S-Box and of the MixColumns matrix.

Proof. First of all, since the invariant subspace \mathcal{IS} is mapped into a coset of \mathcal{IS} after 2-round encryption, and similarly a coset of \mathcal{M}_I is mapped into a coset of \mathcal{D}_I after 2-round decryption, that is

$$\forall k \in K_{\text{weak}} : \quad \mathcal{IS} \xrightarrow[\text{prob. 1}]{R^2(\cdot)} \mathcal{IS} \oplus a \xrightarrow{R(\cdot)} \mathcal{D}_I \oplus a' \xrightarrow[\text{prob. 1}]{R^2(\cdot)} \mathcal{M}_I \oplus b'$$

we focus only on the middle round, and we prove the following equivalent result: given 2^{64} plaintexts in a coset of \mathcal{IS} , the number n of different pairs of ciphertexts (c^i, c^j) for $i \neq j$ that belong to the same coset of \mathcal{D}_I (that is $c^i \oplus c^j \in \mathcal{D}_I$) after 1 round is a multiple of 128. This result can be achieved by observing that, given a pair of texts $t^1, t^2 \in \mathcal{IS} \oplus a$, there exist other pair(s) of texts $s^1, s^2 \in \mathcal{IS} \oplus a$ s.t.

- the following property holds with prob. 1:

$$R(t^1) \oplus R(t^2) \in \mathcal{D}_I \Leftrightarrow R(s^1) \oplus R(s^2) \in \mathcal{D}_I;$$

- the texts s^1, s^2 are given by any different combination of the generating variables of t^1, t^2 .

By definition of \mathcal{IS} , let t^1 and t^2 be as $t^i = a \oplus \bigoplus_{j=0}^7 x_j^i \cdot (e_j \oplus e_{j+8})$ where $x_j \equiv x_{r+4 \times c}$ denotes the byte in the r -th row and in the c -th and $(c+2)$ -th columns. For simplicity, let $t^i \equiv (x_0^i, x_1^i, x_2^i, x_3^i, x_4^i, x_5^i, x_6^i, x_7^i)$.

Case: Different Generating Variables. Consider initially the case in which all the generating variables are different, that is $x_j^1 \neq x_j^2$ for $j = 0, 1, \dots, 7$. Let S_{t^1, t^2} be the set of pairs of texts $s^1, s^2 \in \mathcal{IS} \oplus a$ defined by swapping some generating variables of t^1 and t^2 . More formally, the set S_{t^1, t^2} contains all 128 pairs of texts (s^1, s^2) for all $I \subseteq \{0, 1, 2, 3, 4, 5, 6, 7\}$ where

$$s^1 = a \oplus \bigoplus_{j=0}^7 \left\{ \left[\left(x_j^1 \cdot \delta_j(I) \right) \oplus \left(x_j^2 \cdot [1 - \delta_j(I)] \right) \right] \cdot \left(e_j \oplus e_{j+8} \right) \right\}$$

$$s^2 = a \oplus \bigoplus_{j=0}^7 \left\{ \left[\left(x_j^2 \cdot \delta_j(I) \right) \oplus \left(x_j^1 \cdot [1 - \delta_j(I)] \right) \right] \cdot \left(e_j \oplus e_{j+8} \right) \right\}$$

where the pairs (s^1, s^2) and (s^2, s^1) are considered to be equivalent, and where $\delta_x(A)$ is the Dirac measure defined as $\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$. By showing that

$$\forall (s^1, s^2) \in S_{t^1, t^2} : \quad R(t^1) \oplus R(t^2) = R(s^1) \oplus R(s^2), \quad (8)$$

it follows immediately that $R(t^1) \oplus R(t^2) \in \mathcal{D}_I \Leftrightarrow R(s^1) \oplus R(s^2) \in \mathcal{D}_I$ for each $(s^1, s^2) \in S_{t^1, t^2}$. The equivalence Eq. (8) is due to the facts that *the S-Box operation works independently on each byte and that the XOR-sum is commutative*. Since each set S_{t^1, t^2} has cardinality 128, in the case in which one focuses on the pairs of texts with different

¹⁰Two pairs (s, t) and (t, s) are considered to be equivalent.

generating variables, it follows that the multiple-of-128 property previously defined holds.

Generic Case. In the case in which some variables are equal, e.g. $x_j^1 = x_j^2$ for $j \in J \subseteq \{0, \dots, 7\}$ with $|J| \geq 1$, the difference $R(t^1) \oplus R(t^2)$ is independent of the value of $x_j^1 = x_j^2$ for each $j \in J$. Thus, the idea is to consider all the different pairs of texts given by swapping one or more variables x_l^1 and x_l^2 for $l = 0, 1, \dots, 7$, where x_j for $j \in J$ can take any possible value in \mathbb{F}_{2^8} . Note that in the case in which $0 \leq |J| < 8$ variables are equal, it is possible to identify

$$\underbrace{2^{7-|J|}}_{\text{by swapping different gen. variables}} \times \underbrace{2^{8-|J|}}_{\text{due to equal gen. variables}} = 2^{7 \cdot (1+|J|)} = 128^{1+|J|}$$

different texts s^1 and s^2 in $\mathcal{IS} \oplus a$ that satisfy the condition $R(t^1) \oplus R(t^2) = R(s^1) \oplus R(s^2)$. More formally, given t^1 and t^2 , the set S_{t^1, t^2} contains all $2^{7 \cdot (1+|J|)}$ pairs of texts (s^1, s^2) for all $I \subseteq \{0, 1, 2, 3, 4, 5, 6, 7\} \setminus J$ and for all $\alpha_0, \dots, \alpha_{|J|} \in \mathbb{F}_{2^8}$ where

$$s^1 = a \oplus \bigoplus_{j \in \{0, \dots, 7\} \setminus J} \left\{ \left[\left(x_j^1 \cdot \delta_j(I) \right) \oplus \left(x_j^2 \cdot [1 - \delta_j(I)] \right) \right] \cdot \left(e_j \oplus e_{j+8} \right) \right\} \oplus \bigoplus_{j \in J} \alpha_j \cdot \left(e_j \oplus e_{j+8} \right)$$

$$s^2 = a \oplus \bigoplus_{j \in \{0, \dots, 7\} \setminus J} \left\{ \left[\left(x_j^2 \cdot \delta_j(I) \right) \oplus \left(x_j^1 \cdot [1 - \delta_j(I)] \right) \right] \cdot \left(e_j \oplus e_{j+8} \right) \right\} \oplus \bigoplus_{j \in J} \alpha_j \cdot \left(e_j \oplus e_{j+8} \right)$$

In conclusion, given plaintexts in the same coset of \mathcal{IS} , the number of different pairs of ciphertexts that belong to the same coset of \mathcal{D}_I after one round is a multiple of 128. \square

Theorem 5. *Let \mathcal{IS} , \mathcal{M}_J and \mathcal{X}_I be the subspaces defined as before, for an arbitrary $J \subseteq \{0, 1, 2, 3\}$ and arbitrary $I \subset \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$. Assume that the whitening key is a weak-key, i.e. it belongs to the set K_{weak} defined in Eq. (2). Given 2^{64} plaintexts in \mathcal{IS} , the following properties hold independently of the details of the S-Box:*

- for 5-round AES-128, the number n of different pairs of ciphertexts (c^i, c^j) for $i \neq j$ that belong to the same coset of \mathcal{X}_I (i.e. $c^i \oplus c^j \in \mathcal{X}_I$) is a multiple of 2, that is $\exists n' \in \mathbb{N}$ such that $n = 2 \cdot n'$.
- for 6-round AES-128, the number n of different pairs of ciphertexts (c^i, c^j) for $i \neq j$ that belong to the same coset of \mathcal{M}_J (i.e. $c^i \oplus c^j \in \mathcal{M}_J$) is a multiple of 2, that is $\exists n' \in \mathbb{N}$ such that $n = 2 \cdot n'$.

The proof of these properties – similar to the one given in Grassi et al. [GRR17] and to the one already given – is proposed in details in [Appendix E.2](#).

4.3 Practical Experiments

Most of the previous properties have been practically verified¹¹. Here we briefly present the practical results and we compare them with the theoretical ones.

All our distinguishers are based on \mathcal{IS} and their practical verification requires at least 2^{64} reduced-round AES encryptions. For this reason, we performed our experiments on small-scale AES [CMR05], where each word is composed of 4-bit instead of 8 (note that all previous results are independent of the details of the S-Box operation). This implies that the dimension of \mathcal{IS} reduces to 32 bits from 64.

Practical Results. For [Theorem 4](#) and [Theorem 5](#), we performed 5-round and 6-round encryptions of \mathcal{IS} for more than 100 randomly chosen weak-keys in K_{weak} . We counted the

¹¹The source codes of the distinguishers/attacks will be made public with the publication of this paper. It is also part of the submission as supplement material.

Table 2: *AES Chosen-Key Distinguishers.* The computation cost is the cost to generate N -tuples of plaintexts/ciphertexts. “SK” denotes a chosen-key distinguisher in the Single-Key setting, while “RK” denotes a chosen-key distinguisher in the Related-Key setting. For completeness, we mention that the known-key distinguishers presented in Gilbert [Gil14] are excluded from this Table due to the arguments reported in Section 5.1.2.

AES	Rounds	Computations	Property	SK	RK	Reference
AES-128	8	2^{24}	Multiple Diff. Trail	✓		[DFJ12]
	8	$2^{13.4}$	Multiple Diff. Trail	✓		[JNP13]
	9	2^{55}	Multi-Collision Diff.		✓	[FJP13]
	9	2^{64}	Multiple-of-n (2^{32} keys)	✓		§ 5.3
AES-192	9	2^{16}	Multiple Diff. Trail	✓		[ZWZ19]
AES-256	9	2^{24}	Multiple Diff. Trail	✓		[DFJ12]
	12	2^{64}	Multiple-of-n (2^{32} keys)	✓		App. H.2
	14 (full)	2^{120}	Multi-Collision Diff.		✓	[BKN09]

collisions in each of the four inverse diagonals space \mathcal{ID} and observed the multiple-of-128 and multiple-of-2 properties hold for 5-round and 6-round encryptions, respectively. Similar tests have been performed in order to check the multiple-of-2 property on the subspaces \mathcal{X}_I as defined in Definition 8 for each $|I| \leq 4$. Due to increased time and memory complexity, these properties were not verified for $|I| > 4$. The experiment results – also performed in the decryption direction – coincide with the results provided in Tables 1 and 2.

5 New Chosen-Key Distinguishers for AES

In this section we present new chosen-key distinguishers for AES in the single-key setting. In particular, as major results, we are able to present *the first candidate 9-round chosen-key distinguisher for AES-128* and *a 12-round candidate chosen-key distinguisher for AES-256, both in the single-key setting.* All the distinguishers that we present are based on the (practically verified) multiple-of- n property proposed in Section 4.2.

The goal of an open-key distinguisher is to differentiate between a block cipher E which allows to generate plaintext/ciphertext pairs which exhibit a rare relation, even for a small set of keys or a single key, and an ideal cipher Π that does not have such a property. However, this poses a definitional problem as it was shown already in [CGH04] that any concrete implementable cipher (like the AES) can be trivially distinguished from an ideal cipher. To the best of our knowledge, finding a proper formal definition that captures the intuition behind chosen-key distinguishers has been a challenging task for the last fifteen years and is still an open problem.

We do not attempt to address this formalization challenge here, but proceed in the way that is custom in the literature to describe chosen-key distinguisher: (*1st*) describe the rare property (see Section 5.2), (*2nd*) show that it can be efficiently constructed for the block cipher usually using an inside-out approach (see Section 5.3 for 9-round AES-128 and Section 5.5 for 10-round AES-128), and (*3rd*) argue or prove in some model that any generic method is less efficient or has low success probability (see Section 5.4).

Our results are summarized in Table 2: in order to compare the results, note that an attack/distinguisher with *no* key difference is (logically) harder than an attack/distinguisher for which key differences are allowed, since the attacker has less freedom.

As before, in the following we limit ourselves to give all the details for the AES-128 case, and we refer to Appendix H for the corresponding distinguishers on AES-256.

5.1 Open-Key Distinguishers – State of the Art for AES

5.1.1 Chosen-Key Distinguishers – State of the Art for AES

To the best of our knowledge, the first chosen-key distinguisher for AES in the single-key setting has been proposed in [DFJ12]. In there, the chosen-key model asks the adversary to find two plaintexts/ciphertexts pairs and a key such that the two plaintexts are equal in 3 diagonals and the two ciphertexts are equal in 3 anti-diagonals (if the final MixColumns is omitted). Equivalently, using the subspace trail notation, the goal is to find $(p^1, c^1 \equiv R^8(p^1))$ and $(p^2, c^2 \equiv R^8(p^2))$ for $p^1 \neq p^2$ s.t. $p^1 \oplus p^2 \in \mathcal{D}_I$ and $c^1 \oplus c^2 \in \mathcal{M}_J$ for a certain $I, J \subseteq \{0, 1, 2, 3\}$ s.t. $|I| = |J| = 1$.

This problem is equivalent to the one proposed in [GP10; JNPP14] in the known-key scenario. In particular, the main (and only) difference is related to the freedom of choosing the key, which allows to reduce the computational cost. In more details, for the 8-round AES-128 case it is possible to find the required pairs of plaintexts/ciphertexts with 2^{24} computations instead of 2^{44} (\equiv the cost in the known-key scenario), while the computational cost in the case of an ideal cipher is of 2^{64} in both cases. Later on, such distinguisher has been improved in [JNP13]. For completeness, similar results have been proposed for 9-round AES-256.

The chosen-key model has been popularized some years before by Biryukov et al. [BKN09], since a distinguisher in this model has been extended to a related-key attack on full AES-256. A related distinguisher for 9-round AES-128 has been proposed by Fouque et al. [FJP13]. Both the chosen-key distinguisher proposed in these papers are in the related-key setting. Here we briefly recall them, but we emphasize that we do not consider related-keys in this article. In [BKN09], authors show that it is possible to construct a q -multicollision on Davies-Meyer compression function using AES-256 in time $q \cdot 2^{67}$, whereas for an ideal cipher it would require on average $q \cdot 2^{\frac{q-1}{q+1}128}$ time complexity. A similar approach has been exploited in [FJP13] to set up the first chosen-key distinguisher for 9-round AES-128. Here, the chosen-key model asks the adversary to find a pair of keys (k, k') satisfying $k \oplus k' = \delta$ with a *known (fixed)* difference δ , and a pair of messages $(p^1, c^1 \equiv R^9(p^1))$ and $(p^2, c^2 \equiv R^9(p^2))$ conforming to a partially instantiated differential characteristic in the data part.

5.1.2 Gilbert’s Known-Key Distinguisher for AES

For completeness, we mention that a 10-round known-key distinguisher for AES has been proposed by Gilbert [Gil14] at Asiacrypt 2014. In such case, the known-key model asks the adversary to find a set of 2^{64} (plaintext, ciphertext) pairs, that is (p^i, c^i) for $i = 0, \dots, 2^{64} - 1$, and two keys k^0 and k^{10} with the following properties¹²:

1. with respect to the key k^0 , the bytes of $\{R_{k^0}(p^i)\}_i$ are uniformly distributed, or equivalently that the partially encrypted texts $\{R_{k^0}(p^i)\}_i$ are uniformly distributed among the cosets of \mathcal{D}_I for each I with $|I| = 3$;
2. with respect to the key k^{10} , the bytes of $\{R_{k^{10}}^{-1}(c^i)\}_i$ are uniformly distributed, or equivalently that the partially decrypted texts $\{R_{k^{10}}^{-1}(c^i)\}_i$ are uniformly distributed among the cosets of \mathcal{M}_J for each J with $|J| = 3$.

We emphasize that *such properties are not verified directly by the plaintexts and by the ciphertexts but after one round encryption/decryption*, and they involve keys k^0 and k^{10} that can be different from the “real” encryption subkeys derived from k . The probability that 2^{64} (plaintext, ciphertext) generated by a random permutation satisfy the previous

¹²For this distinguisher, we abuse the notation k^r to denote a key of a certain round r . We emphasize that k^r is not necessarily equal to the secret key, that is k^r can be different from the “real” r -th subkey.

property is 2^{-7200} . Thus, given $2^{64} + 2^8$ plaintexts/ciphertexts, the probability to find among them a subset of 2^{64} pairs of texts with the required properties is close to 1.

A distinguisher based on the Gilbert’s technique is different from all the previous distinguishers up to 8 rounds present in the literature. *For all distinguishers up to 8-round (and for the distinguishers proposed in this paper), the property/relation \mathcal{R} – that the N -tuple of (plaintexts, ciphertexts) must satisfy – does not involve any operation of the block cipher E .* On the other hand, *the previous Gilbert’s like distinguishers do not satisfy this requirement, since in such cases the property/relation \mathcal{R} involves and re-uses some operations of E :* indeed, instead of considering properties “directly” on the plaintexts/ciphertexts, the idea is to show the existence of certain keys for which some properties hold after one round encryption/decryption. As a result, even if Gilbert’s known-key distinguisher leads to statements on more rounds of AES than ever before (without related keys), in the same paper (see [Gil14, Abstract]) it is also observed that its “impact on the security of [...] AES when used as a known key primitive, e.g. in a hash function construction, is questionable”.

Moreover, in order to support such a new kind of distinguisher, it is claimed in [Gil14] that *(1st)* it seems technically difficult to use a stronger property than the uniform distribution one to extend an 8-round known-key distinguisher to a 10-round one and *(2nd)* it is impossible to use the same technique in order to extend a distinguisher for more than 2 rounds. Recently, both claims have been disproved in [GR17], in which authors exploit the same technique to propose *(1st)* a distinguisher on 10-round AES based on truncated differential trails and *(2nd)* the *first distinguisher on 12-round AES* obtained by extending an 8-round distinguisher. Hence, the problem to set up a 9 (or more) rounds open-key distinguisher in the single-key setting for AES-128 without exploiting the Gilbert’s technique is still open.

5.2 The “Simultaneous Multiple-of- n ” Property

In our distinguisher, the chosen-key model asks the adversary to find a set of 2^{64} (plaintexts, ciphertexts), that is $(p^i, c^i \equiv R^9(p^i))$ for $i = 0, \dots, 2^{64} - 1$ – where all the plaintexts/ciphertexts are generated by *the same key* – and a key such that the following “*simultaneous multiple-of- n* ” property is satisfied:

- for each $J, I \subseteq \{0, 1, 2, 3\}$, the number of different pairs of ciphertexts that belong to the same coset of \mathcal{M}_J and the number of different pairs of plaintexts that belong to the same coset of \mathcal{D}_I are a multiple of $128 = 2^7$;
- for each $J, I \subset \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$, the number of different pairs of ciphertexts that belong to the same coset of $\text{MC}(\mathcal{X}_I)$ and the number of different pairs of plaintexts that belong to the same coset of \mathcal{X}_J are a multiple of 2.

For the follow-up, we emphasize that the subspaces \mathcal{X} (defined as in Definition 8) are independent, in the sense that e.g. the fact that the multiple-of-2 property is satisfied by \mathcal{X}_I and/or \mathcal{X}_J does not imply anything on $\mathcal{X}_{I \cup J}$ and vice-versa. This is due to the fact that given \mathcal{X}_I and \mathcal{X}_J , then $\mathcal{X}_I \cup \mathcal{X}_J \subsetneq \mathcal{X}_{I \cup J}$ if $\mathcal{X}_{I \cup J} \neq \mathbb{F}_{2^8}^{4 \times 4}$. As a result, any information about the multiple-of- n property on $\mathcal{X}_I, \mathcal{X}_J$ (and so $\mathcal{X}_I \cup \mathcal{X}_J$) is useless to derive information about the multiple-of- n property on $\mathcal{X}_{I \cup J} \setminus (\mathcal{X}_I \cup \mathcal{X}_J)$ – assuming $\mathcal{X}_{I \cup J} \neq \mathbb{F}_{2^8}^{4 \times 4}$.

5.3 9-round Chosen-Key Distinguisher for AES-128

To find a set of 2^{64} plaintexts/ciphertexts with the required “simultaneous multiple-of- n ” property, the distinguisher exploits the fact that *the required property can be fulfilled by starting in the middle with a suitable set of texts*. In particular, the idea is simply to *choose the key such that the subkey of the 4-th round k^4 belongs the subset K_{weak} defined as in Eq. (2)*. Thus, consider the invariant subspace \mathcal{IS} defined as in Eq. (1), and define the

2^{64} plaintexts as the 4-round decryption of \mathcal{IS} and the corresponding ciphertexts as the 5-round encryption of \mathcal{IS} . Due to the secret-key distinguishers just presented, this set satisfies the required “simultaneous multiple-of- n ” property.

In more details, due to the assumption on the key (that is, $k^4 \in K_{\text{weak}} \subseteq \mathcal{IS}$), note that the subspace \mathcal{IS} is mapped into a coset of \mathcal{IS} after two rounds of encryption and one round of decryption, that is

$$\forall k^4 \in K_{\text{weak}} : \quad \mathcal{IS} \oplus \hat{k} \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus \tilde{k}.$$

Due to the results of Section 4.2 and since $k^4 \in K_{\text{weak}}$, the multiple-of- n properties hold with probability 1 on the plaintexts and on the ciphertexts

$$\text{Multiple-of-}n \xleftarrow{R^{-3}(\cdot)} \mathcal{IS} \oplus \hat{k} \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus \tilde{k} \xrightarrow{R^3(\cdot)} \text{Multiple-of-}n$$

It follows that the required set can be constructed using 2^{64} computations. Moreover, we emphasize that our experiments on the secret-key distinguishers of Section 4.2 implies the *practical verification of this distinguisher*. What remains is to give arguments as to why producing that property simultaneously on the plaintext and ciphertext side of an ideal cipher is unlikely to be as efficient.

5.4 Achieving the “Simultaneous Multiple-of- n ” Property Generically

In this case, the adversary faces a family of random and independent *ideal ciphers* $\{\Pi(K, \cdot), K \in \{0, 1\}^k\}$, where $k = 128, 192, 256$ respectively for the cases AES-128/192/256. His goal is to find a key k and a set of 2^{64} plaintexts/ciphertexts $(p^i, c^i = \Pi(k, p^i))$ s.t. the “simultaneous multiple-of- n ” property is satisfied. As we are going to show, *the probability to find a set of 2^{64} plaintexts/ciphertexts pairs (X_i, Y_i) that satisfies the “simultaneous multiple-of- n ” property for a random permutation is upper bounded by $2^{-65\,618}$.*

As first thing, we discuss the freedom to choose the key. Since the adversary does not know the details of the ideal cipher Π , he does not have any advantage to choose a particular key instead of another one. For this reason, in the following we limit to consider the case in which the permutation Π is instantiated by a fixed key chosen at random in the set $\{0, 1\}^k$ – from now on we use the notation $\Pi(p^i) := \Pi(k, p^i)$.

Exploiting the same strategy proposed in [Gil14], here we prove that the success probability of any oracle algorithm of overall time complexity upper bounded by 2^{64} is negligible¹³ – see Appendix F for a complete proof of the following proposition.

Proposition 3. *Given a perfect random permutation Π of $\{0, 1\}^{128}$ (e.g. instantiated by an ideal cipher with a fixed key uniformly chosen at random in $\{0, 1\}^k$), consider $N = 2^{64}$ oracle queries made by any algorithm \mathcal{A} to the perfect random permutation Π or Π^{-1} . Denote this set of 2^{64} plaintexts/ciphertexts pairs by $(X_i, Y_i = \Pi(X_i))$ for $i = 0, \dots, 2^{64} - 1$. The probability that \mathcal{A} outputs a set of 2^{64} plaintexts/ciphertexts pairs $(X_i, Y_i)_{i=0, \dots, 2^{64}-1}$ that satisfies the “simultaneous multiple-of- n ” property is upper bounded by $2^{-65\,618}$.*

What happens if the adversary performs more than 2^{64} computations? To answer this question, we first compute the probability that a *random* set of 2^{64} plaintexts/ciphertexts generated by the same key satisfies the “simultaneous multiple-of- n ” property. As formally showed in Appendix F, the “simultaneous multiple-of- n ” property is satisfied with probability

$$\left[(2^{-1})^{2^{16}-16} \cdot (2^{-7})^{14} \right]^2 = (2^{-65\,618})^2 \simeq 2^{-2^{17}}$$

¹³We highlight that the proof of Proposition 3 is based on a strategy (very) similar to the one proposed in Gilbert [Gil14] to prove that the uniform distribution (on 8- and 10-round AES) is generically hard.

since (1st) there are $\sum_{i=1}^{15} \binom{16}{i} = 2^{16} - 2$ different subspaces \mathcal{X}_I for which the multiple-of-2 property holds, and among them there are 14 subspaces \mathcal{M}_I for which the multiple-of-128 property holds and (2nd) the probability that the number of collisions is a multiple of N is $\approx 1/N$. As a result, given $2^{64} + 2^{12}$ random texts, the player can find a set of 2^{64} texts that satisfy the required property both on the plaintexts and on the ciphertexts, since it is possible to construct $\binom{2^{64}+2^{12}}{2^{64}} \approx \frac{(2^{64})^{2^{12}}}{2^{12}!} \simeq 2^{2^{17.7}}$ different sets of 2^{64} texts (where $n! \simeq (n/e)^n \cdot \sqrt{2\pi n}$ by Stirling's approximation). On the other hand, *the cost to identify the right 2^{64} texts among all the others is in general much higher than 2^{64} computations*: Indeed, to have a chance of success higher than 95%, one must consider approximately $3 \cdot 2^{131\ 236}$ different sets (note that $1 - (1 - 2^{-131\ 236})^{3 \cdot 2^{131\ 236}} \simeq 1 - e^{-3} \equiv 0.95$).

Moreover, consider the following. Given a set of random texts, suppose to change one plaintext in order to modify the number of collisions in the subspace \mathcal{X}_I (or/and \mathcal{D}_I) for a particular I . As a consequence, all the other numbers of collisions in the subspace \mathcal{X}_J (or/and \mathcal{D}_J) for all $J \neq I$ change. Even if it is possible to have control of these numbers, a problem arises since also the numbers of collisions among the ciphertexts in each subspace \mathcal{M}_K and $\text{MC}(\mathcal{X}_K)$ change, and in general it is not possible to predict such change in advance. Indeed, since the the number of pairs of texts with $1 \leq |J| \leq 15$ equal bytes (that is, that belong to the same coset of a particular subspace \mathcal{X}_J) is on average equal to $2^{127} \cdot 2^{-8 \cdot |J|} \geq 2^{127} \cdot 2^{-8 \cdot 15} = 2^7$, it follows that that the change in one text modifies *all* the numbers of collisions in each subspaces \mathcal{X}_J (resp. $\text{MC}(\mathcal{X}_J)$) – analogous for the subspaces \mathcal{D}_I and \mathcal{M}_I . For all these reasons, we conjecture that that *there is no (efficient) strategy* – that does not involve brute force search – *to fulfill the required “simultaneous multiple-of- n ” property* for which the cost is approximately of 2^{64} computations (or lower). The problem to *formally* prove this fact is left for future work.

Remarks. Before going on, we highlight that this claim/result is not true in general if one considers *only* the multiple-of- n property (for $n \leq 8$) in the subspaces \mathcal{D}_I and \mathcal{M}_J , that is, not for the generic subspaces \mathcal{X} . For a broader understanding of the role of the invariant subspace in the previous distinguishers, in Appendix G we discuss the (im)possibility to set up an open-key distinguisher using the multiple-of-8 property proposed in [GRR17] for more than 8-round AES.

5.5 “Simultaneous Properties” for 10-round (full) AES-128

Before going on, we mention that the previous chosen-key distinguisher can be potentially extended to 10-round AES-128, by considering the following two possible approaches:

- *add one round at the beginning (or at the end)* at the previous distinguisher on 9-round + exploit a weaker property on the plaintexts (or on the ciphertexts);
- *add one round in the middle* at previous distinguisher on 9-round + exploit the remaining degrees of freedom in the choice of the key.

As a result, for a given chosen key, both these two strategies allow us to find a set of (plaintexts, ciphertexts) with some particular “simultaneously multiple-of- n ” properties similar to the ones defined for 9-round AES. In any case, we emphasize that *we do not claim anything regarding the possibility to exploit such strategies in order to set up chosen-key distinguishers for 10-round (full) AES-128*, since:

- as showed in Appendix H.1.1, in the case in which one adds one round at the beginning (resp. at the end), one is forced to exploit a (very) weak multiple-of- n property on the plaintexts (resp. on the ciphertexts). As a result, the gap between the cost for the AES case and for the case of an adversary facing a family of random and independent ideal ciphers becomes too small to set up a *confident* distinguisher;

- as showed in [Appendix H.1.2](#), in the case in which one adds one round in the middle by using the remaining degrees of freedom in the choice of the key, one can re-exploit exactly the same “multiple-of- n ” properties proposed for the 9-round case. However, the set up distinguisher over 10-round AES-128 works for just *one* (chosen) key.

Acknowledgement. Authors thank Léo Perrin for useful discussion about weak-key subspace trail. Authors also thank anonymous reviewers for their valuable comments.

References

- [BBS99] Eli Biham, Alex Biryukov, and Adi Shamir. “Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials.” In: *EUROCRYPT 1999*. Vol. 1592. LNCS. 1999, pp. 12–23.
- [BCC19] Christina Boura, Anne Canteaut, and Daniel Coggia. “A General Proof Framework for Recent AES Distinguishers.” In: *IACR Transactions on Symmetric Cryptology 2019.1* (2019), pp. 170–191.
- [BGL19] Zhenzhen Bao, Jian Guo, and Eik List. *Extended Expectation Cryptanalysis on Round-reduced AES*. Cryptology ePrint Archive, Report 2019/622. <https://eprint.iacr.org/2019/622>. 2019.
- [BK09] Alex Biryukov and Dmitry Khovratovich. “Related-Key Cryptanalysis of the Full AES-192 and AES-256.” In: *ASIACRYPT 2009*. Vol. 5912. LNCS. 2009, pp. 1–18.
- [BKN09] Alex Biryukov, Dmitry Khovratovich, and Ivica Nikolić. “Distinguisher and Related-Key Attack on the Full AES-256.” In: *CRYPTO 2009*. Vol. 5677. LNCS. 2009, pp. 231–249.
- [BKR11] Andrey Bogdanov, Dmitry Khovratovich, and Christian Rechberger. “Biclique Cryptanalysis of the Full AES.” In: *ASIACRYPT 2011*. Vol. 7073. LNCS. 2011, pp. 344–371.
- [BLN17] Céline Blondeau, Gregor Leander, and Kaisa Nyberg. “Differential-Linear Cryptanalysis Revisited.” In: *Journal of Cryptology* 30.3 (2017), pp. 859–888.
- [BR14] Andrey Bogdanov and Vincent Rijmen. “Linear hulls with correlation zero and linear cryptanalysis of block ciphers.” In: *Designs, Codes and Cryptography* 70.3 (2014), pp. 369–383.
- [BR19] Navid Ghaedi Bardeh and Sondre Rønjom. “The Exchange Attack: How to Distinguish Six Rounds of AES with $2^{88.2}$ Chosen Plaintexts.” In: *Advances in Cryptology - ASIACRYPT 2019*. Vol. 11923. LNCS. 2019, pp. 347–370.
- [BS90] Eli Biham and Adi Shamir. “Differential Cryptanalysis of DES-like Cryptosystems.” In: *CRYPTO 1990*. Vol. 537. LNCS. 1990, pp. 2–21.
- [BWP05] An Braeken, Christopher Wolf, and Bart Preneel. “Normality of Vectorial Functions.” In: *Cryptography and Coding - 10th IMA International Conference 2005*. Ed. by Nigel P. Smart. Vol. 3796. LNCS. Springer, 2005, pp. 186–200.
- [Bar+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: *CRYPTO 2018*. Vol. 10992. LNCS. 2018, pp. 185–212.
- [Bei+17] Christof Beierle, Anne Canteaut, Gregor Leander, and Yann Rotella. “Proving Resistance Against Invariant Attacks: How to Choose the Round Constants.” In: *CRYPTO 2017*. Vol. 10402. LNCS. 2017, pp. 647–678.

- [Bey18] Tim Beyne. “Block Cipher Invariants as Eigenvectors of Correlation Matrices.” In: *ASIACRYPT 2018, Part I*. LNCS. Springer, Heidelberg, 2018, pp. 3–31.
- [CGH04] Ran Canetti, Oded Goldreich, and Shai Halevi. “The Random Oracle Methodology, Revisited.” In: *Journal ACM* 51.4 (2004), pp. 557–594.
- [CMR05] Carlos Cid, Sean Murphy, and Matthew J. B. Robshaw. “Small Scale Variants of the AES.” In: *FSE 2005*. Vol. 3557. LNCS. 2005, pp. 145–162.
- [Cha+17] Colin Chaigneau, Thomas Fuhr, Henri Gilbert, Jérémy Jean, and Jean-René Reinhard. “Cryptanalysis of NORX v2.0.” In: *IACR Transactions on Symmetric Cryptology* 2017.1 (2017), pp. 156–174.
- [DFJ12] Patrick Derbez, Pierre-Alain Fouque, and Jérémy Jean. “Faster Chosen-Key Distinguishers on Reduced-Round AES.” In: *INDOCRYPT 2012*. Vol. 7668. LNCS. 2012, pp. 225–243.
- [DR02] Joan Daemen and Vincent Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard*. Information Security and Cryptography. Springer, 2002.
- [DR06] Joan Daemen and Vincent Rijmen. “Understanding Two-Round Differentials in AES.” In: *SCN 2006*. Vol. 4116. LNCS. 2006, pp. 78–94.
- [Der+18] Patrick Derbez, Pierre-Alain Fouque, Jérémy Jean, and Baptiste Lambin. “Variants of the AES Key Schedule for Better Truncated Differential Bounds.” In: *SAC 2018*. Vol. 11349. LNCS. 2018, pp. 27–49.
- [FJP13] Pierre-Alain Fouque, Jérémy Jean, and Thomas Peyrin. “Structural Evaluation of AES and Chosen-Key Distinguisher of 9-Round AES-128.” In: *CRYPTO 2013*. Vol. 8042. LNCS. 2013, pp. 183–203.
- [GP10] Henri Gilbert and Thomas Peyrin. “Super-Sbox Cryptanalysis: Improved Attacks for AES-Like Permutations.” In: *FSE 2010*. Vol. 6147. LNCS. 2010, pp. 365–383.
- [GR16] Lorenzo Grassi and Christian Rechberger. “Practical Low Data-Complexity Subspace-Trail Cryptanalysis of Round-Reduced PRINCE.” In: *INDOCRYPT 2016*. Vol. 10095. LNCS. 2016, pp. 322–342.
- [GR17] Lorenzo Grassi and Christian Rechberger. *New and Old Limits for AES Known-Key Distinguishers*. Cryptology ePrint Archive, Report 2017/255. to appear at “Design, Codes and Cryptography” 2020. 2017.
- [GR18] Lorenzo Grassi and Christian Rechberger. *Rigorous Analysis of Truncated Differentials for 5-round AES*. Cryptology ePrint Archive, Report 2018/182. <https://eprint.iacr.org/2018/182>. 2018.
- [GRR16] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. “Subspace Trail Cryptanalysis and its Applications to AES.” In: *IACR Trans. Symm. Cryptol.* 2016.2 (2016). <http://tosc.iacr.org/index.php/ToSC/article/view/571>, pp. 192–225. ISSN: 2519-173X.
- [GRR17] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. “A New Structural-Differential Property of 5-Round AES.” In: *EUROCRYPT 2017, Part II*. Ed. by Jean-Sébastien Coron and Jesper Buus Nielsen. Vol. 10211. LNCS. Springer, Heidelberg, 2017, pp. 289–317.
- [Gil14] Henri Gilbert. “A Simplified Representation of AES.” In: *ASIACRYPT 2014, Part I*. Ed. by Palash Sarkar and Tetsu Iwata. Vol. 8873. LNCS. Springer, Heidelberg, 2014, pp. 200–222.

- [Gra18] Lorenzo Grassi. “Mixture Differential Cryptanalysis: a New Approach to Distinguishers and Attacks on round-reduced AES.” In: *IACR Trans. Symmetric Cryptol.* 2018.2 (2018), pp. 133–160.
- [Guo+13] Jian Guo, Ivica Nikolic, Thomas Peyrin, and Lei Wang. *Cryptanalysis of Zorro*. Cryptology ePrint Archive, Report 2013/713. <http://eprint.iacr.org/2013/713>. 2013.
- [Guo+16] Jian Guo, Jérémy Jean, Ivica Nikolic, Kexin Qiao, Yu Sasaki, and Siang Sim. “Invariant Subspace Attack Against Midori64 and The Resistance Criteria for S-box Designs.” In: *IACR Transactions on Symmetric Cryptology* 2016.1 (2016), pp. 33–56.
- [HCN09] Miia Hermelin, Joo Yeon Cho, and Kaisa Nyberg. “Multidimensional Extension of Matsui’s Algorithm 2.” In: *FSE 2009s*. Vol. 5665. LNCS. 2009, pp. 209–227.
- [JNP13] Jérémy Jean, María Naya-Plasencia, and Thomas Peyrin. “Multiple Limited-Birthday Distinguishers and Applications.” In: *SAC 2013*. Vol. 8282. LNCS. 2013, pp. 533–550.
- [JNPP14] Jérémy Jean, María Naya-Plasencia, and Thomas Peyrin. “Multiple Limited-Birthday Distinguishers and Applications.” In: *SAC 2013*. Vol. 8282. LNCS. 2014, pp. 533–550.
- [Jea16] Jérémy Jean. “Cryptanalysis of Haraka.” In: *IACR Trans. Symm. Cryptol.* 2016.1 (2016). <http://tosc.iacr.org/index.php/ToSC/article/view/531>, pp. 1–12. ISSN: 2519-173X.
- [KM07] Orhun Kara and Cevat Manap. “A New Class of Weak Keys for Blowfish.” In: *FSE 2007*. Vol. 4593. LNCS. 2007, pp. 167–180.
- [KR07] Lars R. Knudsen and Vincent Rijmen. “Known-Key Distinguishers for Some Block Ciphers.” In: *ASIACRYPT 2007*. Vol. 4833. LNCS. 2007, pp. 315–324.
- [KS99] John Kelsey and Bruce Schneier. “Key-Schedule Cryptanalysis of DEAL.” In: *SAC 1999*. Vol. 1758. LNCS. 1999, pp. 118–134.
- [Kho+17] Khoongming Khoo, Eugene Lee, Thomas Peyrin, and Siang Sim. “Human-readable Proof of the Related-Key Security of AES-128.” In: *IACR Transactions on Symmetric Cryptology* 2017.2 (2017), pp. 59–83.
- [Knu95] Lars R. Knudsen. “Truncated and higher order differentials.” In: *FSE 1994*. Vol. 1008. LNCS. 1995, pp. 196–211.
- [Knu98] Lars R. Knudsen. *DEAL - A 128-bit Block Cipher*. Technical Report 151, Department of Informatics, University of Bergen, Norway. 1998.
- [LH94] Susan K. Langford and Martin E. Hellman. “Differential-Linear Cryptanalysis.” In: *CRYPTO 1994*. Vol. 839. LNCS. 1994, pp. 17–25.
- [LMR15] Gregor Leander, Brice Minaud, and Sondre Rønjom. “A Generic Approach to Invariant Subspace Attacks: Cryptanalysis of Robin, iSCREAM and Zorro.” In: *EUROCRYPT 2015*. Vol. 9056. LNCS. 2015, pp. 254–283.
- [LTW18] Gregor Leander, Cihangir Tezcan, and Friedrich Wiemer. “Searching for Subspace Trails and Truncated Differentials.” In: *IACR Trans. Symm. Cryptol.* 2018.1 (2018), pp. 74–100. ISSN: 2519-173X.
- [Lam+15] Mario Lamberger, Florian Mendel, Martin Schläffer, Christian Rechberger, and Vincent Rijmen. “The Rebound Attack and Subspace Distinguishers: Application to Whirlpool.” In: *Journal of Cryptology* 28.2 (2015), pp. 257–296.

- [Le+04] Tri Van Le, Rüdiger Sparr, Ralph Wernsdorf, and Yvo Desmedt. “Complementation-Like and Cyclic Properties of AES Round Functions.” In: *Advanced Encryption Standard - AES, 4th International Conference*. Vol. 3373. LNCS. 2004, pp. 128–141.
- [Lea+11] Gregor Leander, Mohamed Ahmed Abdelraheem, Hoda AlKhzaimi, and Erik Zenner. “A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack.” In: *CRYPTO 2011*. Vol. 6841. LNCS. 2011, pp. 206–221.
- [MS87] Judy H. Moore and Gustavus J. Simmons. “Cycle Structure of the DES with Weak and Semi-Weak Keys.” In: *CRYPTO 1986*. Vol. 263. LNCS. 1987, pp. 9–32.
- [Mat94] Mitsuru Matsui. “Linear Cryptanalysis Method for DES Cipher.” In: *EUROCRYPT 1993*. Vol. 765. LNCS. 1994, pp. 386–397.
- [Ohk09] Kenji Ohkuma. “Weak Keys of Reduced-Round PRESENT for Linear Cryptanalysis.” In: *SAC 2009*. Vol. 5867. LNCS. 2009, pp. 249–265.
- [RBH17] Sondre Rønjom, Navid Ghaedi Bardeh, and Tor Hellesest. “Yoyo Tricks with AES.” In: *ASIACRYPT 2017, Part I*. LNCS. Springer, Heidelberg, 2017, pp. 217–243.
- [Røn16] Sondre Rønjom. *Invariant subspaces in Simpira*. Cryptology ePrint Archive, Report 2016/248. 2016.
- [TLS16] Yosuke Todo, Gregor Leander, and Yu Sasaki. “Nonlinear Invariant Attack - Practical Attack on Full SCREAM, iSCREAM, and Midori64.” In: *ASIACRYPT 2016, Part II*. Ed. by Jung Hee Cheon and Tsuyoshi Takagi. Vol. 10032. LNCS. Springer, Heidelberg, 2016, pp. 3–33.
- [Vau96] Serge Vaudenay. “On the weak keys of blowfish.” In: *FSE 1996*. Vol. 1039. LNCS. 1996, pp. 27–32.
- [WW16] Yanfeng Wang and Wenling Wu. “New Observations on Piccolo Block Cipher.” In: *CT-RSA 2016*. Vol. 9610. LNCS. 2016, pp. 378–393.
- [Wei+12] Lei Wei, Thomas Peyrin, Przemysław Sokołowski, San Ling, Josef Pieprzyk, and Huaxiong Wang. “On the (In)Security of IDEA in Various Hashing Modes.” In: *FSE 2012*. Vol. 7549. LNCS. 2012, pp. 163–179.
- [ZWZ19] Chunbo Zhu, Gaoli Wang, and Boyu Zhu. “Fast Chosen-Key Distinguish Attacks on Round-Reduced AES-192.” In: *ACISP 2019*. Vol. 11547. LNCS. 2019, pp. 573–587.

A Algorithmic Detection of Weak-Key Subspace Trails

Here we take a look at how we can find these algorithmically. To begin with, we recapitulate how the algorithms for invariant subspaces [LMR15] and subspace trails work [LTW18].

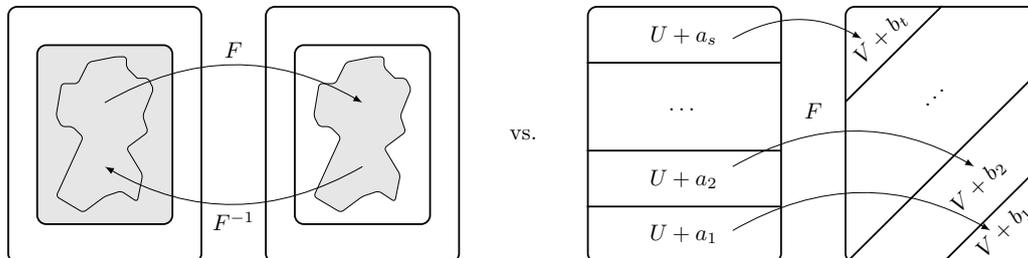


Figure 1: To find invariant subspaces (left part), we iteratively compute the image of the current subspace and map the span of it backwards through the inverse function, until the process stabilizes. For subspace trails (right part), *all* cosets of the starting subspace get mapped to a coset of the ending subspace. This implies that the derivative of the round function is in the ending subspace.

First, Fig. 1 (left part) sketches the idea for invariant subspaces. Given a round function $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$, the algorithm guesses a starting offset a for the affine subspace $U \oplus a$ and then maps $U \oplus a$ forwards and back through F and F^{-1} , every-time computing the span of the image. If the subspace stabilizes, we have found an invariant subspace.

Second, Fig. 1 (right part) illustrates the main idea for subspace trails. The important difference to invariant subspaces is that *every* coset of the starting subspace U is mapped to some coset of the ending subspace V . The implication of this is, see [LTW18, Lemma 1], the images of the derivatives $\Delta_u F(\cdot) := F(\cdot \oplus u) \oplus F(\cdot)$ of the round function F span a subspace of V . In other words, if $U \xrightarrow{F} V$ is a subspace trail, then

$$U \xrightarrow{F} \text{span} \left(\bigcup_{u \in U} \text{Im}(\Delta_u(F)) \right) \subseteq V.$$

We cannot exploit this fact for WKSTs, though. Instead we base the algorithm on the idea for invariant subspaces.

Goal and Details of the Algorithm. Given a round function $R : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ and a key schedule $K_i : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ for $0 \leq i \leq r$ rounds, the goal is to find two subspaces $U, V \subset \mathbb{F}_2^n$ and a subset $S \subseteq \mathbb{F}_2^m$, s. t. every message m chosen from U and every key $k \in S$ get mapped to a ciphertext $c = E_k(m) \in V$, where the encryption uses the round function R and key schedule K_i for the i -th round key. Thus, all master keys in S are weak-keys.

As a starting point, we assume that the zero message $m = 0$ is in our starting subspace U_0 . This is anyway always the case, as we assume all U_i 's to be subspaces. Additionally, we require that a certain key k^{weak} – chosen by the user – is weak, thus in S . Since $k^{\text{weak}} = 0 \in S$ is very often the case if invariant subspace attacks apply, we assume $k^{\text{weak}} = 0$ in the following. In particular, we have the following conditions:

$$\begin{aligned} 0 \in U_i, & & R(U_i) \subseteq U_{i+1}, \\ K_i(S) \subseteq U_i, & & R(K_i(S)) \subseteq U_{i+1}. \end{aligned} \tag{9}$$

Algorithm 1 Compute an initial Weak-key subspace trail

Precondition: A round function $R : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ and a key schedule $K_i : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ for $0 \leq i \leq r$ rounds. An upper bound `max_rnd` on the number of rounds to cover.

Postcondition: A weak-key subspace trail $U_0 \rightarrow \dots \rightarrow U_l$ over l rounds for a set S of weak-keys.

```

1 function WKST( $R, K_i, \text{max\_rnd}$ )
2    $S \leftarrow \{0\}$ 
3    $L \leftarrow [U_0 = \{0, K_0(0)\}]$ 
4   while for the last element  $U_i$  in  $L$ :  $\dim(U_i) < n$  do
5      $U_{i+1} \leftarrow \emptyset$ 
6     for enough  $x \in_R U_i$  do
7        $U_{i+1} \leftarrow U_{i+1} \cup \{R(x)\}$ 
8      $U_{i+1} \leftarrow \text{span}(U_{i+1} \cup \{K_{i+1}(k) \mid \forall k \in S\})$ 
9     append  $U_{i+1}$  to  $L$ 
10    if  $\text{len}(L) \geq \text{max\_rnd}$  then
11      return  $L$ 
12  return ( $L, S$ )

```

Exploiting these conditions and starting at the above mentioned point, we can simply compute the WKST forwards. We may want to check if the resulting trail is invariant, for that we can simply compute the trail backwards at some point. For the complete pseudocode¹⁴ see [Algorithm 1](#).

The runtime of our algorithm depends on the while and for loop. The first loop iterates over the subspaces in our trail and is thus bounded by the length of the WKST. For the later loop, we have to iterate over “enough x ”. Following the same argument as in [LTW18], it follows that sampling $n + 100$ random inputs is enough to compute the following subspace with overwhelming probability.

B Generic Subspace Trail (of length 1) for AES – Proof

Here we give a complete proof regarding the subspace trail of length 1 set up using the generic subspace \mathcal{X} defined in [Section 3.1](#).

First of all, we recall the definition of \mathcal{X} .

Definition 9. Let I a subset of $\{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$. Let the subspace \mathcal{X}_I be defined as $\mathcal{X}_I = \langle \{e_{i,j}\}_{(i,j) \in I} \rangle \equiv \left\{ \bigoplus_{(i,j) \in I} \alpha_{i,j} \cdot e_{i,j} \mid \forall \alpha_{i,j} \in \mathbb{F}_{2^8} \right\}$.

Theorem 6. For each $I \subseteq \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$ and for each $a \in \mathcal{X}_I^\perp$, there exists one and only one $b \in \mathcal{Y}_I^\perp$ such that $R(\mathcal{X}_I \oplus a) = \mathcal{Y}_I \oplus b$, where $\mathcal{Y}_I = MC \circ SR(\mathcal{X}_I)$.

Proof. To prove the Theorem, we simply compute $R(\mathcal{X}_I \oplus a)$. Since SubBytes is bijective and operates on each byte independently, its only effect is to change the coset. In other words, it simply changes the coset $\mathcal{X}_I \oplus a$ to $\mathcal{X}_I \oplus a'$, where $a'_{i,j} = SB(a_{i,j})$ for each $i, j = 0, \dots, 3$. ShiftRows simply moves the bytes of $\mathcal{X}_I \oplus a'$ into $SR(\mathcal{X}_I) \oplus b'$, where $b' = SR(a')$. Since MixColumns is a linear operation, it follows that $MC(SR(\mathcal{X}_I) \oplus b') = MC \circ SR(\mathcal{X}_I) \oplus MC(b') = MC \circ SR(\mathcal{X}_I) \oplus b''$. Key addition then changes the coset to $MC \circ SR(\mathcal{X}_I) \oplus b$. \square

¹⁴Only for simplicity, the update process for the set S is not included in the algorithm. More weak-keys can be found by computing backward from the U_i 's, see [Eq. \(9\)](#) and the Sage implementation.

C Weak-Key Invariant Subspace Trails of AES-256

Here we present the weak-invariant subspace trails for AES-256.

C.1 AES-256 Key-Schedule

In this case, the subkey array is denoted by $W[0, \dots, 59]$, where here the first 8 words of $W[\cdot]$ are loaded with the user secret key. The remaining words of $W[\cdot]$ are updated according to the following rule:

$$W[i][j] = \begin{cases} W[i][j-8] \oplus \text{SB}(W[i+1][j-1]) \oplus R[i][j/8] & \text{if } j \bmod 8 = 0 \\ W[i][j-8] \oplus \text{SB}(W[i][j-1]) & \text{if } j \bmod 8 = 4 \\ W[i][j-1] \oplus W[i][j-8] & \text{otherwise} \end{cases}$$

where $i = 0, 1, 2, 3$, $j = 8, \dots, 59$ and $R[\cdot]$ is an array of predetermined constants.

C.2 Invariant Subspace – Weak-Keys of AES-256

For the case AES-256, a set of 2^{128} weak-keys is given by

$$K_{\text{weak}} := \left\{ \begin{bmatrix} A^0 & A^1 & A^0 & A^1 & E^0 & E^1 & E^0 & E^1 \\ B^0 & B^1 & B^0 & B^1 & F^0 & F^1 & F^0 & F^1 \\ C^0 & C^1 & C^0 & C^1 & G^0 & G^1 & G^0 & G^1 \\ D^0 & D^1 & D^0 & D^1 & H^0 & H^1 & H^0 & H^1 \end{bmatrix} \mid \begin{array}{l} \forall A^i, \dots, H^i \in \mathbb{F}_{2^8} \\ \forall i = 0, 1 \end{array} \right\}$$

Under any of such keys, the subspace \mathcal{IS} is mapped after two complete rounds into a coset of \mathcal{IS} , that is $\mathcal{IS} \xrightarrow{R_K^2 \circ \text{ARK}(\cdot)} \mathcal{IS} \oplus \hat{k}$, where \hat{k} is the corresponding subkey after 2 rounds of the key schedule.

For the follow-up, we also present three subspaces of K_{weak} for which it is possible to construct a longer invariant subspace trail:

3-round: working with any of the 2^{96} keys that satisfy $A^0 = A^1, \dots, D^0 = D^1$, the subspace \mathcal{IS} is mapped after three complete rounds into a coset of \mathcal{IS} , that is $\mathcal{IS} \xrightarrow{R_K^3 \circ \text{ARK}(\cdot)} \mathcal{IS} \oplus \hat{k}'$ where \hat{k}' is the subkey after 3 rounds.

4-round: working with any of the 2^{64} keys that satisfy $A^0 = A^1, \dots, H^0 = H^1$, the subspace \mathcal{IS} is mapped after four complete rounds into a coset of \mathcal{IS} , that is $\mathcal{IS} \xrightarrow{R_K^4 \circ \text{ARK}(\cdot)} \mathcal{IS} \oplus \hat{k}''$ where \hat{k}'' is the subkey after 4 rounds.

5-round: working with any of the 2^{32} keys that satisfy $A^0 = A^1 = B^0 = \dots = D^0 = D^1 = 0$ and $E^0 = E^1, \dots, H^0 = H^1$, the subspace \mathcal{IS} is mapped after five complete rounds into a coset of \mathcal{IS} , that is $\mathcal{IS} \xrightarrow{R_K^5 \circ \text{ARK}(\cdot)} \mathcal{IS} \oplus \hat{k}'''$ where \hat{k}''' is the subkey after 5 rounds.

The complete expressions of the subkeys involved for the previous results are given for completeness in the following.

Details – Sub-Keys involved for AES-256

In order to prove the results proposed for AES-256 given before, we list here the subkeys involved.

(1st) Consider the 2^{96} keys that satisfy

$$A^0 = A^1, \quad B^0 = B^1, \quad C^0 = C^1, \quad D^0 = D^1$$

that is

$$\left\{ \begin{bmatrix} A & A & A & A & E^0 & E^1 & E^0 & E^1 \\ B & B & B & B & F^0 & F^1 & F^0 & F^1 \\ C & C & C & C & G^0 & G^1 & G^0 & G^1 \\ D & D & D & D & H^0 & H^0 & H^0 & H^1 \end{bmatrix} \mid \forall A, B, C, D, \dots, H^0, H^1 \in \mathbb{F}_{2^8} \right\}$$

The next subkey is given by

$$\begin{bmatrix} A \oplus SB(F^1) \oplus R[1] & SB(F^1) \oplus R[1] & A \oplus SB(F^1) \oplus R[1] & SB(F^1) \oplus R[1] \\ B \oplus SB(G^1) & SB(G^1) & B \oplus SB(G^1) & SB(G^1) \\ C \oplus SB(H^1) & SB(H^1) & C \oplus SB(H^1) & SB(H^1) \\ D \oplus SB(E^1) & SB(E^1) & D \oplus SB(E^1) & SB(E^1) \end{bmatrix}$$

(2nd) Consider the 2^{64} keys that satisfy

$$A^0 = A^1, \quad B^0 = B^1, \quad C^0 = C^1, \quad D^0 = D^1, \quad \dots, \quad H^0 = H^1$$

The next subkey is given by

$$\begin{bmatrix} SB(\hat{F} \oplus R[1]) \oplus E & SB(\hat{F} \oplus R[1]) & SB(\hat{F} \oplus R[1]) \oplus E & SB(\hat{F} \oplus R[1]) \\ SB(\hat{G}) \oplus F & SB(\hat{G}) & SB(\hat{G}) \oplus F & SB(\hat{G}) \\ SB(\hat{H}) \oplus G & SB(\hat{H}) & SB(\hat{H}) \oplus G & SB(\hat{H}) \\ SB(\hat{E}) \oplus H & SB(\hat{E}) & SB(\hat{E}) \oplus H & SB(\hat{E}) \end{bmatrix}$$

where

$$\hat{E} := SB(E), \quad \hat{F} := SB(F), \quad \hat{G} := SB(G), \quad \hat{H} := SB(H).$$

(3rd) Consider the 2^{32} keys that satisfy

$$A^0 = A^1 = B^0 = \dots = D^0 = D^1 = 0, \quad E^0 = E^1, \quad F^0 = F^1, \quad \dots \quad H^0 = H^1.$$

Then, the next subkeys is given by

$$\begin{bmatrix} SB(\hat{G}) \oplus \hat{F} \oplus R'[2] & SB(\hat{G}) \oplus R[2] & SB(\hat{G}) \oplus \hat{F} \oplus R'[2] & SB(\hat{G}) \oplus R[2] \\ SB(\hat{H}) \oplus \hat{G} & SB(\hat{H}) & SB(\hat{H}) \oplus \hat{G} & SB(\hat{H}) \\ SB(\hat{E}) \oplus \hat{H} & SB(\hat{E}) & SB(\hat{E}) \oplus \hat{H} & SB(\hat{E}) \\ SB(\hat{F}) \oplus \hat{E} & SB(\hat{F}) & SB(\hat{F}) \oplus \hat{E} & SB(\hat{F}) \end{bmatrix}$$

where

$$\begin{aligned} \hat{\hat{E}} &:= SB(SB(E)), & \hat{\hat{F}} &:= SB(\hat{F} \oplus R[1]), \\ \hat{\hat{G}} &:= SB(SB(G)), & \hat{\hat{H}} &:= SB(SB(H)) \end{aligned}$$

and $R'[2] := R[1] \oplus R[2]$.

D Practical Collisions for 7-round AES-256 Compressing Modes

Many block cipher hashing modes contain XOR of input and output of the cipher. E.g. given an input $x = (x_0, x_1, \dots, x_n)$, the corresponding hash $H = (H_0 \equiv IV, H_1, \dots, H_n)$ can be produced using

- the Davies–Meyer hash function: $H_i = E_{x_i}(H_{i-1}) \oplus H_{i-1}$;

- the Matyas-Meyer-Oseas hash function $H_i = E_{g(H_{i-1})}(x_i) \oplus x_i$;
- the Miyaguchi-Preneel hash function $H_i = E_{g(H_{i-1})}(x_i) \oplus H_{i-1} \oplus x_i$.

In this section, we show how to produce collisions for some of such constructions exploiting our invariant subspace \mathcal{IS} . *Since we assume the attacker is able to choose the initial value IV , we propose our results in the compressing mode.*

Using the result proposed in [Appendix H.2.3](#) and when the first and second round keys (namely, k_1 and k_2) are all zero, we have

$$\mathcal{IS} \oplus k_0 \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^6(\cdot)} \mathcal{IS} \oplus k_6,$$

where k_0 and k_6 are the initial and final round keys.

Since dimension of \mathcal{IS} is 64, we expect to find a collision with (at least) 2^{32} elements in \mathcal{IS} . In fact, since one can construct $\binom{2^{32}}{2} = 2^{32} \cdot (2^{32} - 1)/2 \approx 2^{63}$, the probability to find a collision is approximately $1 - (1 - 2^{-64})^{2^{63}} \approx 1 - e^{-1/2} \approx 39.35\%$.

We performed two experiments by encrypting 2^{32} elements in \mathcal{IS} in an inside out fashion by choosing the AES-256 key as

$$[k_0 \| k_1] = [62636363 \ 00000000 \ 00000000 \ 00000000 \| 00000000 \ 00000000 \ 00000000 \ 00000000],$$

which makes first and second round keys zero. In our first experiment we used the smaller invariant subspace¹⁵ \mathcal{IS}' of dimension 32 where every column is identical

$$\mathcal{IS}' := \left\{ \left[\begin{array}{cccc} a & a & a & a \\ b & b & b & b \\ c & c & c & c \\ d & d & d & d \end{array} \right] \mid \forall a, b, c, d \in \mathbb{F}_{2^8} \right\},$$

which is the same invariant subspace independently used in the cryptanalysis of constructions using the unkeyed AES round permutation, recently e.g. for cryptanalysis of Haraka hash function in [Jea16], and in the second one we chose 2^{32} random elements in \mathcal{IS} .

As a result, we got a 7-round collision in both cases for the Matyas-Meyer-Oseas or Miyaguchi-Preneel compressing functions constructed with 7-round AES-256, where the attacker choose IV ($= H_0$) as k_0 . Note that since AES-256 block size is 128 bits and key size is 256 bits, a $g(\cdot)$ conversion/padding function is used on the output to make it suitable as the key. A very natural function $g(\cdot) : \mathbb{F}_{2^n} \mapsto \mathbb{F}_{2^{2n}}$ that turns out to be good for our purpose is given by

$$g(x) = x \| \underbrace{0 \dots 0}_{n \text{ bit}} \in \mathbb{F}_{2^{2n}}$$

where $\|$ denotes concatenation. Our collisions for 7-round AES-256 hashing modes are provided in Table 3. Moreover, an perhaps a more natural application, these collisions turn into collisions for Davies-Meyer compressing mode where the message block is fixed to $k_0 \| k_1$ and the plaintexts of Table 3 are used as IV s.

To the best of our knowledge, the best known collision attacks on AES compressing modes are the trivial conversion of the Whirlpool attacks of [Lam+15]. They turn into 6-round collision attacks on every key length of AES which require 2^{56} time and 2^{32} memory complexity. Our collisions are on 7 rounds and require 2^{32} time and 2^{32} memory

¹⁵This choice is motivated by the fact that $k_1, k_2, k_3, k_4 \in \mathcal{IS}'$. As a result:

$$\mathcal{IS}' \oplus k_0 \xleftarrow{R^{-1}(\cdot)} \mathcal{IS}' \xrightarrow{R^5(\cdot)} \mathcal{IS}' \oplus k_5$$

where $\mathcal{IS}' \oplus k_5 \subseteq \mathcal{IS}$ (since $k_5 \in \mathcal{IS} \setminus \mathcal{IS}'$). Similar results can be potentially derived working directly on \mathcal{IS} .

Table 3: Examples of compression collisions for 6 and 7-round AES-256 used in Matyas-Meyer-Oseas construction where $k_0 || k_1 = 62636363\ 00000000\ 00000000\ 00000000\ 00000000\ 00000000\ 00000000\ 00000000$. Last round contains the matrix multiplication but not the final key addition, which does not affect the collisions. These plaintexts are also collisions for Miyaguchi-Preneel mode and pseudo-collisions for Davies-Meyer mode.

Plaintext		Hash (i.e., Plaintext \oplus Ciphertext)	
7-round Collisions			
6407503c0664335f	0664335f0664335f	4a2e96618b438711	284df5028b438711
c2e01a46a0837925	a0837925a0837925		
fa8cca8ad93ff889	98efa9e9d93ff889	79b1f1b3c1415dd7	1bd292d0c1415dd7
02cc0aa7b96b44b3	60af69c4b96b44b3		
6-round Collisions			
b1b602e8d3d5618b	d3d5618bd3d5618b	f85752eeb3488419	9a34318db3488419
d0122734b2714457	b2714457b2714457		
e75dd657853eb534	853eb534853eb534	c99eec4ba84135a3	abfd8f28a84135a3
27f4f3b1459790d2	459790d2459790d2		
e00159e982623a8a	82623a8a82623a8a	09315bee8a3b5978	6b52388d8a3b5978
be9c9a2adcfff949	dcfff949dcfff949		
6eed06230c8e6540	0c8e6540c8e6540	a77bf28d6e087b35	c51891ee6e087b35
497163fb2b120098	2b1200982b120098		
345684eb5635e788	5635e7885635e788	c82e26780c32ed63	aa4d451b0c32ed63
5ee850813c8b33e2	3c8b33e23c8b33e2		
439ad67621f9b515	21f9b51521f9b515	2ecfb051888f27dd	4cacd332888f27dd
7e4032701c235113	1c2351131c235113		

complexity where a time-memory tradeoff is also possible. Our attack is also valid for 6 rounds with the same complexities. It may be conceivable that local collision methods from [BKN09] can be adapted to the compression collision setting we consider here. Note however that this approach can not avoid to simultaneously require differences in both the chaining as well as the message input of an AES-256-based compression functions, whereas we only need a difference in one of the two.

E Proofs of Results given in Section 4

E.1 Proofs of Section 4.1

E.1.1 Proof of Prop. 1

As showed in Section 4.1, a subspace \mathcal{IS} is mapped into a coset of \mathcal{IS} after 2 rounds AES-128 under a weak-key. By definition of $\mathcal{IS} \oplus a$, the first and the third diagonals (resp. the second and the fourth) are equal. This means that:

- there are 2^{32} texts that are equal in the first and the third diagonals, and that differ in the second and in the fourth ones. By definition, these 2^{32} texts belong to the same coset of $\mathcal{D}_{1,3}$. It follows that after 2-round encryption, the 2^{64} texts are divided into 2^{32} different cosets of $\mathcal{D}_{1,3}$;

- equivalently, there are 2^{32} texts that are equal in the second and in the fourth diagonals, and that differ first and the third ones. By definition, these 2^{32} texts belong to the same coset of $\mathcal{D}_{0,2}$. It follows that after 2-round encryption, the 2^{64} texts are divided into 2^{32} different cosets of $\mathcal{D}_{0,2}$.

The result follows immediately from the fact that each coset of \mathcal{D}_I is mapped into a coset of \mathcal{M}_I after 2-round AES encryption – see [Theorem 1](#).

In the case of a random permutation, note that

- there are $\binom{2^{64}}{2^{32}}$ different ways to divide 2^{64} texts in sets of 2^{32} elements;
- for each set, 2^{32} texts are equal on two diagonals with prob. $(2^{-64})^{2^{32}-1}$;
- the probability that these two diagonals are different for each set is equal to $\prod_{i=0}^{2^{32}-1} \frac{2^{64}-i}{2^{64}} = \prod_{i=0}^{2^{32}-1} (1 - i \cdot 2^{-64})$.

As a result, the probability for the case of a random permutation is given by

$$\begin{aligned} & \binom{2^{64}}{2^{32}} \cdot \prod_{i=0}^{2^{32}-1} \left[(2^{-64})^{2^{32}-1} \cdot \overbrace{(1 - i \cdot 2^{-64})}^{\leq 1} \right] \leq \binom{2^{64}}{2^{32}} \cdot (2^{-64})^{2^{64}-2^{33}+1} \approx \\ & \approx \frac{1}{\sqrt{2\pi \cdot (2^{32}-1)}} \cdot \frac{(2^{64})^{2^{64}}}{(2^{32})^{2^{32}} \cdot (2^{64}-2^{32})^{2^{64}-2^{32}}} \cdot (2^{-64})^{2^{64}-2^{33}+1} \approx \\ & \approx (2^{32})^{2^{32}} \cdot (1 - 2^{-32})^{2^{64}-2^{32}} \cdot (2^{-64})^{2^{64}-2^{33}+1} \approx 2^{-2^{70}} \end{aligned}$$

using Stirling's approximation $n! \approx n^n \cdot e^{-n} \cdot \sqrt{2\pi \cdot n}$.

E.1.2 Proof of Prop. 2

To compute the previous probability, we first recall the *law of total probability*. Given a finite (or countably infinite) partition B_1, \dots, B_n of a sample space events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ s. t. first $B_i \cap B_j = \emptyset$ for each $i \neq j$ and second $\bigcup_i B_i$ is the entire sample space, then

$$\Pr[A] = \sum_{i=1}^n \Pr[A|B_i] \cdot \Pr[B_i]$$

It follows that for a fixed I :

$$\begin{aligned} & \Pr[R^5(x) \oplus R^5(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}, k \in K_{\text{weak}}] = \\ & = \left\{ \Pr[R^5(x) \oplus R^5(y) \in \mathcal{M}_I \mid R^4(x) \oplus R^4(y) \in \mathcal{M}'] \times \Pr[R^4(x) \oplus R^4(y) \in \mathcal{M}'] \right\} + \\ & + \left\{ \Pr[R^4(x) \oplus R^4(y) \notin \mathcal{M}'] \times \Pr[R^5(x) \oplus R^5(y) \in \mathcal{M}_I \mid R^4(x) \oplus R^4(y) \notin \mathcal{M}'] \right\} = \\ & = 2^{-64+16 \cdot |I|} \cdot 2^{-31} + 2^{-32 \cdot (4-|I|)} \cdot (1 - 2^{-31}) \simeq 2^{-95+16 \cdot |I|} + 2^{-128+32 \cdot |I|} \end{aligned}$$

where $\mathcal{M}' = \mathcal{M}_{0,2} \cup \mathcal{M}_{1,3}$.

Difference w.r.t. the case of a Random Permutation. Since for a random permutation Π the probability of the same event is equal to $\frac{2^{32 \cdot |I|} - 1}{2^{128} - 1} \approx 2^{-128+32 \cdot |I|}$ (where the approximation holds with an error of magnitude 2^{-128}), it is possible to distinguish the two cases. Just to have concrete numbers:

- if $|I| = 3$ (I fixed), then $\Pr [R^5(x) \oplus R^5(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}, k \in K_{\text{weak}}] \simeq 2^{-32} + 2^{-47}$, while for a random permutation $\Pr [\Pi(x) \oplus \Pi(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}] = 2^{-32}$;
- if $|I| = 2$ (I fixed), then $\Pr [R^5(x) \oplus R^5(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}, k \in K_{\text{weak}}] = 3 \cdot 2^{-64}$, while for a random permutation $\Pr [\Pi(x) \oplus \Pi(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}] = 2^{-64}$;
- if $|I| = 1$ (I fixed), then $\Pr [R^5(x) \oplus R^5(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}, k \in K_{\text{weak}}] = 2^{-79}$, while for a random permutation $\Pr [\Pi(x) \oplus \Pi(y) \in \mathcal{M}_I \mid x, y \in \mathcal{IS}] = 2^{-96}$.

E.1.3 “Weak-Key Truncated Differential” versus “Weak-Key Subspace Trail”

Here we emphasize the relation and the difference between a “weak-key truncated differential” and a “weak-key subspace trail”. As highlighted in [BLN17; GRR16; LTW18], there is a strong connection between subspace trails and truncated differentials. Here we clarify the connection and the difference between the “weak-key truncated differential” and the *previous* “weak-key subspace trail”.

As observed in the papers just cited, a truncated differential can be described as affine spaces of differences:

Let K_{weak} be a set of weak-keys and let $E_k(\cdot) : \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$ a cipher. For each key $k \in K_{\text{weak}}$, a “weak-key truncated differential” of probability p is defined by a pair of affine subspaces ($U \subseteq \mathbb{F}_2^n, s \in \mathbb{F}_2^n$) and ($V \subseteq \mathbb{F}_2^m, t \in \mathbb{F}_2^m$) for which:

$$\Pr [E_k(x) \oplus E_k(x \oplus \alpha \oplus s) \in V \oplus t] = p.$$

where $\alpha \in U$ and $x \in \mathbb{F}_2^n$.

The previous probability can be re-written as

$$\Pr [E_k(x) \oplus E_k(y) \in V \oplus t \mid x, y \in \mathbb{F}_2^n \text{ s.t. } x \oplus y \in U \oplus s \text{ and } k \in K_{\text{weak}}] = p.$$

The crucial point here is that (at least) one of the two initial plaintexts can be chosen in a complete arbitrary way. In other words, the plaintext x can be chosen arbitrary in \mathbb{F}_2^n , while the plaintext y must satisfy the condition $x \oplus y \in U \oplus s$ (i.e. $y \in U \oplus (s \oplus x)$).

What about the “weak-key subspace trail” presented in [Section 4.1](#)? Referring to [Section 4.1](#), all probabilities presented there have the form

$$\Pr [E_k(x) \oplus E_k(y) \in V \oplus t \mid x, y \in U \oplus s \text{ and } k \in K_{\text{weak}}] = p.$$

(where $U \oplus s \equiv \mathcal{IS}$ in our case). The crucial difference is that both x and y must belong to the affine space $U \oplus s$. In other words:

- in the case of a weak-key truncated differential, one of the two texts is completely arbitrary ($x \in \mathbb{F}_2^n$), while the other one must belong to a fixed affine subspace (namely, $y \in U \oplus (s \oplus x)$);
- in the case of weak-key subspace trail, both the two texts must belong to the same affine subspace (namely, $x, y \in U \oplus s$).

This means that this second requirement does not involve just the difference between x and y (as before), but the fact that both texts are in the same affine subspace.

In conclusion, *while truncated differentials and subspace trails can be considered to be equivalent, the same is not true for weak-key truncated differentials and weak-key subspace trails*. In particular, a weak-key subspace trail can be considered (a little) weaker than a weak-key truncated differential, since there is less freedom on the choice of the inputs.

E.2 Proofs of Weak-Key “Multiple-of- n ” – Theorem 5

As before, since the invariant subspace \mathcal{IS} is mapped into a coset of \mathcal{IS} after 2-round encryption, since a coset of \mathcal{X}_I is mapped into a coset of $\mathcal{Y}_I = SR^{-1} \circ MC^{-1}(\mathcal{X}_I)$ after 1-round decryption (as showed in Theorem 3) and since a coset of \mathcal{M}_J is mapped into a coset of \mathcal{D}_J after 2-round decryption, that is

$$\begin{aligned} \forall k \in K_{\text{weak}} : \quad & \mathcal{IS} \xrightarrow[\text{prob. 1}]{R^2(\cdot)} \mathcal{IS} \oplus a \xrightarrow{R^2(\cdot)} \mathcal{Y}_I \oplus a' \xrightarrow[\text{prob. 1}]{R(\cdot)} \mathcal{X}_I \oplus b' \\ \forall k \in K_{\text{weak}} : \quad & \mathcal{IS} \xrightarrow[\text{prob. 1}]{R^2(\cdot)} \mathcal{IS} \oplus a \xrightarrow{R^2(\cdot)} \mathcal{D}_J \oplus a' \xrightarrow[\text{prob. 1}]{R^2(\cdot)} \mathcal{M}_J \oplus b' \end{aligned}$$

the idea is to prove an equivalent results that involve only the two middle rounds. Given a pair of texts $t^1, t^2 \in \mathcal{IS} \oplus a$, we prove that there exist other pair(s) of texts $s^1, s^2 \in \mathcal{IS} \oplus a$ such that

$$R^2(t^1) \oplus R^2(t^2) = R^2(s^1) \oplus R^2(s^2)$$

where the texts s^1, s^2 are obtained by swapping the diagonals of t^1, t^2 . As before, this implies that

$$\begin{aligned} R^2(t^1) \oplus R^2(t^2) \in \mathcal{X}_I & \Leftrightarrow R^2(s^1) \oplus R^2(s^2) \in \mathcal{X}_I; \\ R^2(t^1) \oplus R^2(t^2) \in \mathcal{D}_J & \Leftrightarrow R^2(s^1) \oplus R^2(s^2) \in \mathcal{D}_J. \end{aligned}$$

In order to prove the previous claim, we use the “Super-S-Box” notation [DR06], where

$$\text{super-SBox}(\cdot) = \text{SB} \circ \text{ARK} \circ \text{MC} \circ \text{SB}(\cdot). \quad (10)$$

Case: Different Diagonals. As before, in the case in which the diagonals are different (i. e., $[x_0^1, x_5^1, x_2^1, x_7^1] \neq [x_0^2, x_5^2, x_2^2, x_7^2]$ and $[x_1^1, x_4^1, x_3^1, x_6^1] \neq [x_1^2, x_4^2, x_3^2, x_6^2]$), given t^1 and t^2 defined as

$$\text{SR}(t^i) \equiv \left(\underbrace{[x_0^i, x_5^i, x_2^i, x_7^i]}_{\text{1st and 3rd columns}}, \underbrace{[x_1^i, x_4^i, x_3^i, x_6^i]}_{\text{2nd and 4th columns}} \right)$$

where $\text{SR}(\cdot)$ denotes the ShiftRows operation, then $R^2(t^1) \oplus R^2(t^2) = R^2(s^1) \oplus R^2(s^2)$ if s^1 and s^2 are defined as

$$\text{SR}(s^i) \equiv \left(\underbrace{[x_0^{3-i}, x_5^{3-i}, x_2^{3-i}, x_7^{3-i}]}_{\text{1st and 3rd columns}}, \underbrace{[x_1^i, x_4^i, x_3^i, x_6^i]}_{\text{2nd and 4th columns}} \right).$$

To prove the previous fact, we first recall that 2-round encryption can be rewritten using the “super-SBox” notation

$$R^2(\cdot) = \text{ARK} \circ \text{MC} \circ \text{SR} \circ \text{super-SBox} \circ \text{SR}(\cdot).$$

Thus, we are going to prove that

$$\text{super-SBox}(\hat{t}^1) \oplus \text{super-SBox}(\hat{t}^2) = \text{super-SBox}(\hat{s}^1) \oplus \text{super-SBox}(\hat{s}^2) \in \mathcal{W}_I$$

where

$$\hat{t}^i = \text{SR}(t^i) \in \mathcal{IS} \oplus \text{SR}(a) \quad \text{and} \quad \hat{s}^i = \text{SR}(s^i) \in \mathcal{IS} \oplus \text{SR}(a)$$

for $i = 1, 2$ (note that $t^i, s^i \in \mathcal{IS} \oplus a$). Note that the first and the third columns of \hat{t}^i and \hat{s}^i are equal, as well as the second and the fourth columns. Similar to the 5-round case, since the first and the second columns (and so the third and the fourth ones) of \hat{t}^1 and \hat{t}^2 depend on different and independent variables, since the Super-S-Box works independently on each column and since the XOR-sum is commutative, it follows the thesis.

Generic Case. What happens if one diagonal is in common for the two texts, e.g. $[x_0^1, x_5^1, x_2^1, x_7^1] = [x_0^2, x_5^2, x_2^2, x_7^2]$ (analogous for $[x_1^1, x_4^1, x_3^1, x_6^1] = [x_1^2, x_4^2, x_3^2, x_6^2]$)? As before, in this case the difference $R^2(t^1) \oplus R^2(t^2)$ is independent of the values of such diagonal. It follows that the pair of texts s^1 and s^2 can be constructed as

$$\text{SR}(s^i) \equiv \left(\underbrace{[x_0^{3-i}, x_5^{3-i}, x_2^{3-i}, x_7^{3-i}]}_{1st \text{ and } 3rd \text{ columns}}, \underbrace{[\alpha_0, \alpha_5, \alpha_2, \alpha_7]}_{2nd \text{ and } 4th \text{ columns}} \right)$$

or

$$\text{SR}(s^i) \equiv \left(\underbrace{[x_1^i, x_4^i, x_3^i, x_6^i]}_{1st \text{ and } 3rd \text{ columns}}, \underbrace{[\alpha_0, \alpha_5, \alpha_2, \alpha_7]}_{2nd \text{ and } 4th \text{ columns}} \right)$$

where $\alpha_0, \alpha_5, \alpha_2, \alpha_7$ can take any possible values in \mathbb{F}_{2^8} . Note that in this case, it is possible to identify $2 \cdot 2^{32} = 2^{33}$ different texts s^1 and s^2 in $\mathcal{IS} \oplus a$ that satisfy the condition $R^2(t^1) \oplus R^2(t^2) = R^2(s^1) \oplus R^2(s^2)$. In conclusion, given plaintexts in the same coset of \mathcal{IS} , the number of different pairs of ciphertexts that belong to the same coset of \mathcal{X}_I and/or \mathcal{D}_J after two rounds is always a multiple of 2.

F Proof of Proposition 3

Assume all the pairs (X_i, Y_i) result from queries to Π or Π^{-1} . Consider a (random) set of $2^{64} - 1$ plaintexts/ciphertexts pairs $\{(X_i, Y_i)\}_{i=0, \dots, 2^{64}-2}$ such that there exists (at least) one plaintext/ciphertext pair (\hat{X}, \hat{Y}) for which the set $\{(X_i, Y_i)\}_{i=0, \dots, 2^{64}-1}$ satisfies the required multiple-of- n property. By assumption, the player can always find \hat{X}' (resp. \hat{Y}') such that the “simultaneous multiple-of- n ” property is satisfied for the plaintexts (resp. for the ciphertexts). However, the oracle’s answer \hat{Y}' (resp. \hat{X}') is *uniformly drawn* from $\{0, 1\}^{128} \setminus \{Y_1, Y_2, \dots, Y_{2^{64}-1}\}$ (resp. from $\{0, 1\}^{128} \setminus \{X_1, X_2, \dots, X_{2^{64}-1}\}$). Therefore, *the probability that the answer to the N -th query allows the output of \mathcal{A} to satisfy property \mathcal{R} (i. e. multiple-of- n) is upper bounded by $(2^{-1})^{2^{16}-16} \cdot (2^{-7})^{14} = 2^{-65 \cdot 618} \simeq 2^{-2^{16}}$ since*

- there are $\sum_{i=1}^{15} \binom{16}{i} = 2^{16} - 2$ different subspaces \mathcal{X}_I for which the multiple-of-2 property holds, and among them there are 14 subspaces \mathcal{M}_I for which the multiple-of-128 property holds;
- the probability that the number of collisions is a multiple of N is $\approx 1/N$.

In order to prove this second point, we first show that the probabilistic distribution of the number of collisions is a binomial distribution¹⁶.

Given a set of n pairs of texts, consider the event that m pairs belong to the same coset of a subspace \mathcal{X} . As first thing, we show that the probabilistic distribution of number of collisions is simply described by a *binomial distribution*. By definition, a binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p . In our case, given n pairs of texts, each of them satisfies or not the above property/requirement with a certain probability. Thus, this model can be described using a binomial distribution, for which the mean μ and the variance σ^2 are respectively given by $\mu = n \cdot p$ and $\sigma^2 = n \cdot p \cdot (1 - p)$.

In our case, the number of pairs is given by $\binom{2^{64}}{2} \simeq 2^{127}$, the probability that a pair of texts belong to the same coset of \mathcal{X}_I is equal to $2^{-8 \cdot (16 - |I|)}$, while it is equal to $2^{-32 \cdot (4 - |J|)}$ for the subspaces \mathcal{D}_J and \mathcal{M}_J .

¹⁶Note that the fact that “the probability that the number of collisions is a multiple of N is $1/N$ ” is obvious *if* the probabilistic distribution of the number of collisions is a uniform one, which is not the case.

Probability that “the number of collision is even” is (approximately) $1/2$ – Case: subspaces \mathcal{X}_I . The probability that the number of collisions is even is given by

$$\frac{1}{2} + \frac{1}{2} \cdot (1 - 2p)^n$$

where note that n is an even number. In our case, since $n \simeq 2^{127}$ and $2^{-120} \leq p \leq 2^{-8}$ (where the prob. 2^{-120} and 2^{-8} correspond resp. to the cases $|I| = 15$ and $|I| = 1$), the previous probability is well approximated by $1/2 + 1/2 \cdot (1 - 2^{-7})^{2^{127}} \approx 1/2$.

In order to prove the previous result, let X be a binomial distribution $X \sim \mathcal{B}(n, p)$. Combining the facts that

$$\begin{aligned} \Pr[X \text{ even}] + \Pr[X \text{ odd}] &= \sum_{k=0}^n \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = [(1-p) + p]^n = 1 \\ \Pr[X \text{ even}] - \Pr[X \text{ odd}] &= \sum_{k=0}^n \binom{n}{k} \cdot (-p)^k \cdot (1-p)^{n-k} = [(1-p) - p]^n \end{aligned}$$

where

$$\Pr[X \text{ even}] = \sum_{k=0}^{n/2} \binom{n}{2k} \cdot p^{2k} \cdot (1-p)^{n-2k}, \quad \Pr[X \text{ odd}] = \sum_{k=0}^{n/2-1} \binom{n}{2k+1} \cdot p^{2k+1} \cdot (1-p)^{n-2k-1},$$

it follows that $\Pr[X \text{ even}] = \frac{1}{2} + \frac{1}{2} \cdot (1 - 2p)^n$.

Probability that “the number of collision is a multiple of N ” is (approximately) $1/N$ – Case: subspaces \mathcal{M}_J and \mathcal{D}_J . In order to prove this result, we first approximate the binomial distribution with a normal one. De Moivre-Laplace Theorem claims that the normal distribution is a good approximation of the binomial one if the skewness of the binomial distribution – given by $(1 - 2p)/\sqrt{n \cdot p \cdot (1-p)}$ – is close to zero. In our case, since $n \simeq 2^{127}$ and $2^{-96} \leq p \leq 2^{-32}$ (where the prob. 2^{-96} and 2^{-32} correspond resp. to the cases $|J| = 3$ and $|J| = 1$), it follows that $2^{-47.5} \leq \text{skew} \leq 2^{-15.5}$, which means that the normal approximation is sufficiently good. Thus, we approximate the binomial distribution with a normal one $\mathcal{N}(\mu = n \cdot p, \sigma^2 = n \cdot p \cdot (1-p))$, where the probability density function is given by $\varphi(x) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

In this case, what is the probability that the multiple-of- N collisions is satisfied? To answer this question, it is sufficient to sum all the probabilities where the number of collisions is a multiple-of- N (for $N \in \mathbb{N}$ and $N \neq 0$), that is

$$\sum_{x \in \mathbb{Z}} \frac{1}{\sqrt{2\pi \cdot \sigma^2}} e^{-\frac{(N \cdot x - \mu)^2}{2\sigma^2}} = \frac{1}{N} \cdot \underbrace{\sum_{x \in \mathbb{Z}} \frac{1}{\sqrt{2\pi \cdot \tilde{\sigma}^2}} e^{-\frac{(x - \tilde{\mu})^2}{2\tilde{\sigma}^2}}}_{=1 \text{ by definition}} = \frac{1}{N}$$

where $\tilde{\mu} = \mu/N$ and $\tilde{\sigma}^2 = \sigma^2/N^2$. Obviously, if $N = 1$, then this probability is equal to 1.

G On the Difficulty to set up “Multiple-of- n ” Open-Key Distinguishers *Without* Relying on Weak-Keys

In order to better understand the role of the invariant subspace, and hence the dependence on weak-keys, in the previous construction, we briefly discuss the following problem: is it possible to set up a similar distinguisher using the multiple-of-8 property proposed in [GRR17] which holds for any key? We conjecture that this is hard.

Given a coset of a diagonal space \mathcal{D}_I , the multiple-of-8 property holds (1) after 5-round encryption and (2) after 3-round decryption. It follows that given a coset of \mathcal{C}_I in the middle, then

$$\forall k : \quad \text{Multiple-of-8} \xleftarrow{R^{-4}(\cdot)} \mathcal{C}_I \oplus a \xrightarrow{R^4(\cdot)} \text{Multiple-of-8},$$

it is possible to achieve a simultaneous multiple-of-8 property on 8 rounds.

Distinguisher on 8 rounds? First of all, one may ask if this property is strong enough in order to set up a chosen-key distinguisher. In order to answer this question, it is sufficient to consider the case of an adversary faces a family of random and independent *ideal ciphers* $\{\Pi(K, \cdot), K \in \{0, 1\}^k\}$, where $k = 128, 192, 256$ respectively for the cases AES-128/192/256. His goal is to find a key k and a set of 2^{64} plaintexts/ciphertexts $(p^i, c^i = \Pi(k, p^i))$ s.t. the “simultaneous multiple-of- n ” property is satisfied.

Working exactly as in Section 5, a random sets of 2^{64} plaintexts/ciphertexts pairs (X_i, Y_i) satisfies the “simultaneous multiple-of-8” property with prob. $(1/8)^{(2 \cdot 14)} = 2^{-84}$ (since (1) the probability that a number is a multiple of 8 is $1/8$, (2) there are 14 different subspaces \mathcal{D}_I and 14 different subspaces \mathcal{M}_I for $I \subseteq \{0, 1, 2, 3\}$). It follows that given $2^{64} + 2$ random texts, the adversary can construct $\binom{2^{64}+2}{2^{64}} \approx 2^{127}$ different sets of 2^{64} texts. Hence, it seems the simultaneous multiple-of-8 property is not strong enough in order to set up a chosen-key distinguisher.

Extension to 9 rounds. Let’s assume that the previous 8-round distinguisher is valid. In order to extend it to more rounds, a possibility can be to use a coset of $\mathcal{D}_I \oplus \mathcal{M}_J$ in the middle. Here we show why this solution does not work.

First of all, observe that

$$\mathcal{D}_I \oplus \mathcal{M}_J \oplus a \equiv \bigcup_{b \in \mathcal{D}_I \oplus a} \mathcal{M}_J \oplus b \equiv \bigcup_{b \in \mathcal{M}_J \oplus a} \mathcal{D}_I \oplus b$$

Thus, consider 5-round encryption (similar for the decryption direction). The number of collisions between the pairs of ciphertexts whose corresponding plaintexts are in the same coset of \mathcal{D}_I is a multiple of 8 with prob. 1. However, it is not possible to claim anything about the the pairs of ciphertexts whose corresponding plaintexts are in the same coset of \mathcal{M}_J , or for which one plaintext is in $\mathcal{D}_I \oplus a'$ and the other in $\mathcal{M}_J \oplus b'$. As a result, one loses any multiple-of- n property. A similar argumentation works also in the decryption direction.

As we have just seen, the invariant subspace allows to solve this problem in the case of weak-keys. The problem to set up a known-key distinguisher (for which the key does not satisfy any particular property) that exploits the multiple-of- n property for more than 8-round AES is still open.

H Chosen-Key Distinguishers and “Simultaneous Property” for AES-128 and AES-256

In Section 5.3, we have proposed a chosen-key distinguisher for 9-round AES-128. Here we analyze the case in which such a distinguisher is extended to 10-round AES-128. Moreover, using the same strategies, here we present similar results for AES-256. Since the strategies used to set up these distinguishers is similar to the ones proposed for AES-128, we refer to Sections 5.3 and 5.5 for all the details and we highlight here the main differences.

H.1 “Simultaneous Property” for 10-round AES

H.1.1 “Simultaneous Property” for 10-round AES – Exploiting a Weaker Property

In the first approach, one is able to generate a set of 2^{64} (plaintexts, ciphertexts), that is $(p^i, c^i \equiv R^{10}(p^i))$ for $i = 0, \dots, 2^{64} - 1$ – where all the plaintexts/ciphertexts are generated by *the same key* – and a key such that the following “*simultaneous multiple-of- n* ” property is satisfied:

Plaintext: on the plaintexts, we re-use the previous properties:

(1st) for each $J \subseteq \{0, 1, 2, 3\}$, the number of different pairs of plaintexts that belong to the same coset of \mathcal{D}_J is a multiple of $128 = 2^7$;

(2nd) for each $I \subseteq \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$, the number of different pairs of plaintexts that belong to the same coset of \mathcal{X}_I are a multiple of 2;

Ciphertext: for each $J \subseteq \{0, 1, 2, 3\}$, the number of different pairs of ciphertexts that belong to the same coset of \mathcal{M}_J is a multiple of 2.

Choosing one of the 2^{32} keys proposed for the 9-round distinguisher given in Section 5.3, it is possible to construct such set with a computational cost of 2^{64} . In more details:

- due to the assumption on the key (that is, $k^4 \in K_{\text{weak}} \subseteq \mathcal{IS}$), note that the subspace \mathcal{IS} is mapped into a coset of \mathcal{IS} after two rounds encryption and one round decryption, that is

$$\forall k^4 \in K_{\text{weak}} : \quad \mathcal{IS} \oplus \hat{k} \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus \tilde{k};$$

- due to the results of Section 4.2, given $k^4 \in K_{\text{weak}}$, (1st) the multiple-of-128 property (on \mathcal{D}_J) and the multiple-of-2 property (on \mathcal{X}_I) hold on the plaintexts while (2nd) the multiple-of-2 property (on \mathcal{M}_J) holds on the ciphertexts

$$\text{Multiple-of-}n \xleftarrow{R^{-3}(\cdot)} \mathcal{IS} \oplus \hat{k} \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus \tilde{k} \xrightarrow{R^4(\cdot)} \text{Multiple-of-2}$$

What about an adversary facing a family of random and independent *ideal ciphers*?

Due to previous analysis, the property on the plaintexts is satisfied with prob. $2^{-32 \cdot 809} \simeq 2^{-2^{15}}$ while the property on the ciphertexts is satisfied with prob. 2^{-14} , for an overall probability of $2^{-32 \cdot 809} \cdot 2^{-14} = 2^{-32 \cdot 823} \simeq 2^{-2^{15}}$.

In other words, the property on the ciphertexts is much weaker than the property on the plaintexts. This fact can be potentially used to generate a set of 2^{64} plaintexts/ciphertexts with the required properties with a data cost of $3 \cdot 2^{78}$. Indeed, the attacker can easily generate a set of 2^{64} plaintexts that satisfy the “Multiple-of- n ” property as described before (e.g. he can generate such set using the fact that the 4-round AES decryption of \mathcal{IS} – namely $R^4(\mathcal{IS})$ – has the required “Multiple-of- n ” property). Then, he simply asks the oracle for the corresponding ciphertexts, which satisfy the “Multiple-of-2” property with prob. 2^{-14} . By repeating this process $3 \cdot 2^{14}$ times, the probability of success¹⁷ is higher than 95%. The cost of such strategy (which includes both the generation of the texts and the check that the property is satisfied) is *at least* of 2^{78} .

Even if this attack is faster than 2^{128} , its cost is still (much) bigger than 2^{64} , which is the cost to generate the required set of plaintexts/ciphertexts for the case of 10-round AES. Remember that the goal in an open-key distinguisher is indeed to be able to generate the required set of plaintexts/ciphertexts with a *similar* (or even the same) cost for AES (or the studied cipher) and for the ideal cipher. In this case, it is very unlikely that any generic

¹⁷The probability of success is given by $1 - (1 - 2^{-14})^{3 \cdot 2^{14}} \geq 0.95$.

attack can get close to that: *even if we would allow unlimited time, the data complexity of a generic attack would still need to be higher than 2^{64}* . Indeed, working as in the 9-round case, a simple brute force attack requires at least¹⁸ $2^{64} + 2^{11}$ plaintexts/ciphertexts in order to find a set of 2^{64} plaintexts with the required properties. For all these reasons and same as for the 9-round case (see our arguments from Section 5.4), we conjecture that the data/computational cost of an adversary to generate such set is (much) higher than 2^{64} computations.

H.1.2 “Simultaneous Property” for 10-round AES – Exploiting Degrees of Freedom in the Weak-Key

In the second approach, the idea is to extend it to 10-round by *adding one round in the middle* using the remaining degrees of freedom in the choice of the key.

In more details, referring to the 9-round distinguisher proposed in Section 5.3, if the subkey k^4 of the 4-th round belongs in K_{weak} (defined as in Eq. (2) and Section 3.2), it follows that

$$\text{Multiple-of-}n \xleftarrow{R^{-3}(\cdot)} \mathcal{IS} \oplus a \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus b \xrightarrow{R^3(\cdot)} \text{Multiple-of-}n$$

In other words, one exploits the fact that the subspace \mathcal{IS} is mapped into a coset of it after 2-round encryption and 1-round decryption for any subkey in K_{weak} .

By simple computation, there is a key in K_{weak} for which the subspace \mathcal{IS} is mapped into a coset of it after two rounds decryption. In more details, for the key $\hat{k} \in K_{\text{weak}}$ defined by

$$\hat{k} \equiv (A = 0x63 \oplus R[5], B = 0x63, C = 0x63, D = 0x63) \in K_{\text{weak}}$$

it follows that

$$\mathcal{IS} \oplus a \xleftarrow{R^{-2}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus b.$$

To see this, it is sufficient to compute one round of the key schedule

$$\begin{bmatrix} A \oplus 0x63 \oplus R[5] & 0 & 0 & 0 \\ B \oplus 0x63 & 0 & 0 & 0 \\ C \oplus 0x63 & 0 & 0 & 0 \\ D \oplus 0x63 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{1-round Key Schedule}} K_{\text{weak}} \equiv \begin{bmatrix} A & A & A & A \\ B & B & B & B \\ C & C & C & C \\ D & D & D & D \end{bmatrix},$$

and to look for a key in K_{weak} that belongs to \mathcal{IS} one round before. As a result, it follows that for the key $\hat{k} \equiv (A = 0x63 \oplus R[5], B = 0x63, C = 0x63, D = 0x63) \in K_{\text{weak}}$ it is possible to set up a distinguisher on 10 rounds¹⁹ since

$$\text{Multiple-of-}n \xleftarrow{R^{-3}(\cdot)} \mathcal{IS} \oplus a \xleftarrow{R^{-2}(\cdot)} \mathcal{IS} \xrightarrow{R^2(\cdot)} \mathcal{IS} \oplus b \xrightarrow{R^3(\cdot)} \text{Multiple-of-}n$$

Using this observation, one is able to find a set of 2^{64} plaintexts/ciphertexts, i.e. $(p^i, c^i \equiv R^{10}(p^i))$ for $i = 0, \dots, 2^{64} - 1$ – where all the plaintexts/ciphertexts are generated by *the same key* – and a key such that

- for each $J, I \subseteq \{0, 1, 2, 3\}$, the number of different pairs of ciphertexts that belong to the same coset of \mathcal{M}_J and the number of different pairs of plaintexts that belong to the same coset of \mathcal{D}_I are a multiple of $128 \equiv 2^7$;

¹⁸Note that $\binom{2^{64}+2^{11}}{2^{64}} \geq 2^{32823}$.

¹⁹For completeness, we discuss the relevance of a distinguisher that can be constructed for a single key (which this does not mean – in general – that it holds for one key only). A single collision/near-collision/ or similar distinguishing property for a block-cipher based compression function or hash function would be also a property of the cipher that holds (depending on the mode) for a single key. Assume this is found with a non-generic approach. This simple example shows that, in principle, properties even for single keys can be interesting.

- for each $J, I \subseteq \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$, the number of different pairs of ciphertexts that belong to the same coset of $MC(\mathcal{X}_I)$ and the number of different pairs of plaintexts that belong to the same coset of \mathcal{X}_J are a multiple of 2.

Similar to the 9-round case, due to our arguments from Section 5.4 we conjecture that the computational cost of an adversary to generate such set is (much) higher than 2^{64} computations.

H.2 Chosen-key distinguisher for AES-256

Before presenting the distinguisher, we recall that:

- it is possible to set up a weak invariant subspace of length two/three/four/five for $2^{128}/2^{96}/2^{64}/2^{32}$ weak-keys of AES-256;
- due to the argumentation proposed in Section 4.2, it follows that the multiple-of-128 property holds for up to 7-round AES-256, while the multiple-of-2 property holds for up to 9-round AES-256.

H.2.1 Chosen-Key Distinguisher for 12-round AES-256

Similarly, to set up the 12-round distinguisher of AES-256, one exploits the fact that

$$\forall k \in K_{\text{weak}} : \quad \mathcal{IS} \oplus a \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^5(\cdot)} \mathcal{IS} \oplus b$$

for each key in K_{weak} defined in Appendix C.2 where

$$A^0 = A^1 = B^0 = \dots = D^0 = D^1 = 0, \quad E^0 = E^1, F^0 = F^1, \dots, H^0 = H^1.$$

H.2.2 “Simultaneous Property” for 13-round AES-256 – “Weaker” Property

As for AES-128, the simplest way to extend the previous distinguisher to 13-round is to exploit a weaker property on (e.g.) the ciphertexts. As a result, *while the property on the plaintexts is unchanged*, one is able to generate a set of 2^{64} (plaintexts, ciphertexts), that is $(p^i, c^i \equiv R^{13}(p^i))$ for $i = 0, \dots, 2^{64} - 1$ – where all the plaintexts/ciphertexts are generated by the same key – such that for each $J \subseteq \{0, 1, 2, 3\}$, the number of different pairs of ciphertexts that belong to the same coset of \mathcal{M}_I is a multiple of 2.

H.2.3 “Simultaneous Property” for 13-round AES-256 – Freedom of the Key

Another possibility to extend the previous distinguisher to 13-round is to exploit the freedom in the key. In more details, in order to find a set of 2^{64} (plaintexts, ciphertexts) that satisfy the required “simultaneous multiple-of- n ” properties for 13-round AES-256, among the previous weak-keys the idea is to choose the sub-key defined by

$$\hat{k} \equiv (E^0 = E^1 = 0x63 \oplus R[5], F^0 = F^1 = 0x63, \dots, H^0 = 0, H^1 = 0x63) \in K_{\text{weak}}$$

for which

$$\mathcal{IS} \oplus a \xleftarrow{R^{-2}(\cdot)} \mathcal{IS} \xrightarrow{R^5(\cdot)} \mathcal{IS} \oplus b.$$

or

$$\hat{k} \equiv (E^0 = E^1 = F^0 = F^1 = \dots = H^0 = 0, H^1 = 0) \in K_{\text{weak}}$$

for which

$$\mathcal{IS} \oplus a \xleftarrow{R^{-1}(\cdot)} \mathcal{IS} \xrightarrow{R^6(\cdot)} \mathcal{IS} \oplus b.$$

H.2.4 “Simultaneous Property” for *full* AES-256

The previous “simultaneous properties” cover 13 rounds of AES-256. Here we show that it is possible to consider a weaker property (e.g.) on the plaintexts to cover full AES-256 in the single-key setting. In this case, one can generate a set of 2^{64} (plaintexts, ciphertexts), that is $(p^i, c^i \equiv R^{14}(p^i))$ for $i = 0, \dots, 2^{64} - 1$ – where all the plaintexts/ciphertexts are generated by *the same key* – and a key such that the following “*simultaneous multiple-of-n*” property is satisfied:

Plaintext: on the plaintexts, we re-use the previous properties:

(1st) for each $J \subseteq \{0, 1, 2, 3\}$, the number of different pairs of plaintexts that belong to the same coset of \mathcal{D}_J is a multiple of $128 = 2^7$;

(2nd) for each $I \subset \{(0, 0), (0, 1), \dots, (3, 2), (3, 3)\} \equiv \{(i, j)\}_{0 \leq i, j \leq 3}$, the number of different pairs of plaintexts that belong to the same coset of \mathcal{X}_I are a multiple of 2;

Ciphertext: for each $J \subseteq \{0, 1, 2, 3\}$, the number of different pairs of ciphertexts that belong to the same coset of \mathcal{M}_J is a multiple of 2.

Choosing the key as before and due to the same arguments given for AES-128 and AES-192, the computational cost to construct such set is of 2^{64} .

What about an adversary facing a family of random and independent ideal ciphers?

Due to previous analysis, the required properties holds with prob. $2^{-32 \cdot 823} \simeq 2^{-2^{15}}$ for a random set of texts. As before, a simple brute force attack requires at least $2^{64} + 2^{11}$ plaintexts/ciphertexts in order to find a set of 2^{64} plaintexts with the required properties. Due to our argumentations from Section 5.4, we conjecture that the computational cost of an adversary to generate such set is (much) higher than 2^{64} computations.