Weak Linear Layers in Word-Oriented Partial SPN and HADES-Like Ciphers

Lorenzo Grassi[1,2], Christian Rechberger[1] and Markus Schofnegger[1]

[1] IAIK, Graz University of Technology
[2] Digital Security Group, Radboud University, Nijmegen
firstname.lastname@iaik.tugraz.at
l.grassi@cs.ru.nl

Abstract. When designing a classical substitution-permutation network (SPN) permutation, every non-trivial choice of the S-box and of the affine layer provides security after a finite number of rounds. However, this is not necessarily the case for partial SPN (P-SPN) ciphers: Since the nonlinear part does not cover the full state, there may exist highly non-trivial choices of linear layers which, for example, do not provide security against statistical attacks.

Quite surprisingly, this direction has hardly been considered in the literature. For example, LowMC uses different linear layers in each round in order to avoid the problem, but this solution is quite expensive, both computationally and memory-wise. `Zorro`, another construction with an incomplete nonlinear layer, simply reuses the AES matrix, but this introduces weaknesses.

Working from an attacker's perspective and focusing on P-SPN ciphers, in this paper we present conditions which allow to set up attacks based on infinitely long subspace trails – even when using highly non-trivial linear layers. We analyze both the case in which the trail is invariant and the case in which is not invariant (yet still an infinite number of rounds can be covered). In this paper, we consider two scenarios, namely active and inactive S-boxes. For the first case, we finally provide a tool which is able to determine whether a given linear layer matrix is vulnerable against infinitely long invariant subspace trails based on our observations.

Finally, we point out that besides P-SPN ciphers, our results may also have a crucial impact on the HADES design strategy recently presented at Eurocrypt 2020, which mixes rounds with full S-box layers and rounds with partial S-box layers in order to guarantee security and achieve good performance in the target applications.

Keywords: Partial SPN · Linear Layer · Invariant Subspace & Subspace Trail · HADES

Contents

1 Introduction **2**

2 Preliminaries **6**
2.1 SPN and Partial SPN Ciphers 6
2.2 Invariant Subspaces and Subspace Trails 8
2.3 Preliminary Assumptions 9

3 Subspaces Trails for P-SPN Ciphers (No Active S-Boxes) **9**
3.1 Preliminary Results 10
3.2 Linear Layers with Low Multiplicative Order 11
3.3 Infinitely Long Invariant Subspace Trails 12
3.4 Infinitely Long Iterative (Non-Invariant) Subspace Trails 14

4 Practical Tests **15**
4.1 Algorithm to Detect "Weak" Matrices 16
4.2 Percentage of "Weak" Linear Layers 17
4.3 An Open Problem of Finding a Necessary Condition 18

5 Subspace Trails for P-SPN Ciphers with Active S-Boxes **20**
5.1 Subspace Trails and Truncated Differentials 20
5.2 Infinitely Long Subspace Trail with Active S-Boxes 21

6 Open Problems **23**

A Related Works **27**

B 2-Round Iterative Subspace Trail – Details **28**

C Truncated and Impossible Differentials **30**
C.1 Truncated Differentials with Probability < 1 30
C.2 Impossible Differentials 31

D Using our Tool for Starkad and Poseidon Matrices **31**

1 Introduction

Choosing the Linear Layer in SPN and Partial-SPN Ciphers. A substitution-permutation network (SPN) is a popular construction technique for block ciphers. Given a key, such a network transforms a plaintext block into a ciphertext block by applying several alternating *rounds* of substitution boxes and permutations to provide confusion and diffusion. For an SPN cipher over $\mathbb{F}^t$, the substitution layer usually consists of t parallel (independent) nonlinear functions called S-boxes, operating at word level. The permutation layer is in most cases a linear operation that can be described as the multiplication of the state with a $t \times t$ matrix.

Determining a suitable round function and the number of rounds necessary for security is the main objective when designing such a primitive. When choosing a linear layer which provides full diffusion at word level after a finite number of rounds, the corresponding cipher can potentially be secure by choosing a "sufficient" number of rounds.

One of the main approaches to achieve provable security against various statistical attacks is to use the *wide trail strategy* [14], which allows to guarantee security against differential [10, 11] and linear [31] attacks, two of the most powerful cryptanalytic techniques in the literature. Instead of choosing larger S-boxes with strong properties, the wide trail strategy aims to design the linear round transformations in such a way that the minimum number of active S-boxes over multiple rounds is increased. From this point of view, an optimal linear layer can be chosen when using an MDS matrix (that is, a matrix that maximizes the minimum number of active S-boxes in two consecutive rounds). Such an approach was used in e.g. SHARK [32] or AES (together with a word-level permutation).

Driven by various new application areas and settings, a variation of the SPN approach, the so-called partial substitution-permutation network (P-SPN), has been proposed and investigated practically [18, 17]. Replacing part of the substitution layer with an identity mapping can lead to substantial practical advantages in many applications in which the cost of a nonlinear operation is significantly higher than the cost of a linear one. This includes masking and practical applications of secure multi-party computation (MPC), fully homomorphic encryption (FHE), and zero-knowledge proofs (ZK) that use symmetric primitives, where the linear computations are often much cheaper than nonlinear ones.

This approach has been proposed for the first time by Gérard et al. [18] at CHES 2013. A concrete instantiation of their methodology is `Zorro` [18], a 128-bit lightweight AES-like cipher which reduces the number of S-boxes per round from 16 to only 4 (to compensate, the number of rounds has been increased to 24). A similar approach has then been considered by Albrecht et al. [17] in the recent design of a family of block ciphers called LowMC proposed at Eurocrypt 2015. LowMC is a flexible block cipher based on an SPN structure, which combines an incomplete S-box layer with a strong linear layer in order to guarantee security and to be competitive in applications like MPC/FHE/ZK.

Security of Partial-SPN Ciphers. A considerable disadvantage of this approach is that the existing wide trail strategy to rule out large classes of statistical attacks is no longer applicable and has to be replaced by more ad-hoc approaches. In the case of `Zorro`, the heuristic argument proposed by the designers turned out to be insufficient, as Wang et al. [34] found iterative differential and linear characteristics that were missed by the heuristic and used them to break full `Zorro` faster than by exhaustive search. Similarly, the authors of LowMC chose the number of rounds in order to guarantee that no differential or linear characteristic can cover the whole cipher with non-negligible probability. However, they do not provide similarly strong security arguments against other attack vectors including algebraic attacks, and key-recovery attacks on LowMC have thus been found [17].

An automated characteristic search tool and dedicated key-recovery algorithms for SP networks with partial nonlinear layers have been presented in [5]. In there, the authors propose generic techniques for differential and linear cryptanalysis of SP networks with

partial nonlinear layers. Besides obtaining practical attacks on P-SPN ciphers, the authors concluded that even if "*the methodology of building PSP networks based on AES in a straightforward way is flawed, [...] the basic PSP network design methodology can potentially be reused in future secure designs*".

Hades Design Strategy. A possible way to fix the problem regarding the security against statistical attacks has been exploited in the so-called "HADES" design strategy [21], recently proposed at Eurocrypt 2020. The HADES strategy is a high-level design approach for cryptographic permutations and keyed permutations addressing the needs of new applications that emphasize the role of multiplications in these designs, with a focus on simple arguments for its security.

The main ingredient of the HADES strategy is to mix rounds with full S-box layers and rounds with partial S-box layers in order to provide good performance while still being secure. The external rounds with full S-box layers together with the wide trail strategy are used to guarantee security against differential and linear attacks. The main goal of the middle rounds with a single S-box each is to provide security against algebraic attacks by increasing the degree of the overall scheme. In other words, they are not used to provide security against statistical attacks. However, as we are going to recall in the following, a weakness of the linear layer in the rounds with partial S-box layers can impact the security against algebraic attacks as well.

Our Contribution

One of the main problems a designer has to face while designing a partial SPN permutation regards the choice of the linear layer. In the literature, the two possible choices considered so far are using the same linear layers in all rounds (as in `Zorro`), or using different linear layers for each round (as in LowMC).

Even if the second strategy can potentially prevent statistical attacks[1] (as discussed in [17]), it has some drawbacks. First of all, the implementation cost in terms of computation time or memory may become a problem, even when considering the optimizations proposed in [25, 16]. Moreover, the analysis of the security against other attacks may become more complicated, since the linear layer is different in each round. Finally, a poor choice of the linear layers does not guarantee security against such attacks, as shown concretely in [17]. In particular, the fact that the designers of LowMC allow to instantiate it using a pseudo-random source that is not cryptographically secure is risky, since using an over-simplified source for pseudo randomness may give a malicious party additional control over the LowMC instantiation, and may allow finding weak instances much faster than exhaustively searching for them.

For all these reasons, in this paper we focus only on the first strategy: Our goal is to better understand which properties a linear layer matrix has to fulfill in order to prevent the existence of *infinitely long subspace trails* [22, 23], namely the existence of a non-trivial subspace $\mathcal{U} \subseteq \mathbb{F}^t$ of inputs that is mapped into a proper (affine) subspace of the state space over any number of rounds.

In this paper, we do not limit ourselves to work with invariant subspaces, and, to the best of our knowledge for the first time, we also consider subspace trails with active S-boxes. In addition, we present an algorithm and a concrete implemented tool which, given a matrix, can be used to detect subspace trails for the case of inactive S-boxes.

Influence of the Branch Number. Let us focus on a word-oriented partial SPN cipher over $\mathbb{F}^t$, where the linear layer is simply defined as the multiplication with a $t \times t$ MDS matrix. Since such a matrix provides full diffusion at word level, and since a partial

[1] To the best of our knowledge, there is no attack on LowMC based only on differential cryptanalysis.

nonlinear layer is applied, one may expect that after a certain – even huge – number of rounds, the corresponding cipher is secure.

As we are going to show with a concrete example, this is not always the case. Indeed, consider a partial SPN cipher defined over $\mathbb{F}_q^{4\times 4}$, and let the round transformation be

$$R^{(i)}\left((x_1,x_2,x_3,x_4)^T\right) = k^{(i)} + \begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 3 & 1 & 2 \\ 3 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} S(x_1) \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

for a "good" S-box $S : \mathbb{F} \to \mathbb{F}$, where $R^{(i)}$ denotes the i-th round function and $k^{(i)}$ denotes the i-th round key (and where $2 \equiv X$ and $3 \equiv X + 1$ for the Boolean case). Even though the matrix is MDS (and similar to the AES one), an invariant subspace trail generated by the subspace $\mathcal{S} = \left\langle (0,1,0,-1)^T \right\rangle$ (eq., $\left\langle (0,1,0,1)^T \right\rangle$ for the Boolean case) can be set up for an *arbitrary* number of rounds. At the same time, this is not possible any more when considering the MDS matrix used in AES.

Infinitely Long Subspace Trails for Word-Oriented P-SPN Ciphers. This example allows us to conclude that a high branch number alone is not sufficient in the case of word-oriented partial SPN ciphers when compared to the case of (full) SPN ciphers. For this reason, in the following we analyze how the details of the matrix that defines the linear layer influences the security against statistical attacks. Specifically, *working independently of the details of the S-box*, we present a sufficient condition that allows to discard matrices which do not provide security. As we are going to show, this condition is related to the eigenspaces of the matrix that defines the linear layer:

- In the case of inactive S-boxes (Section 3 for details), we directly construct an infinitely long subspace trail via the eigenspaces of the matrix.
- In the case of active S-boxes (Section 5 for details), we show which properties a subspace must satisfy in order to be infinitely long.

We emphasize that we do not only focus on invariant subspace trails (in other words, a non-trivial infinitely long subspace trail is not necessarily invariant). In particular,

(1) such a subspace trail is invariant if it is related to the eigenspaces of M, and

(2) it is not invariant if it is related to the eigenspaces of M^k for $k \geq 2$, as explained in the following.

In both cases, examples are provided to present and support the results. We remark that we do not impose any condition on the matrix M (with the only exception that it is invertible – i.e., we do not limit ourselves to work only with MDS matrices) and that the results are quite different from what is known for the SPN case.

Before going on, we stress that *we are presenting sufficient conditions and not necessary ones.* This means that a "weak" matrix – namely, a matrix for which it is possible to set up infinitely long subspace trails – is not necessarily discarded by the properties presented in our paper. The problem to find a necessary and sufficient condition (or to prove that some of the properties given here are also necessary) is left open for future research.

A Dedicated Tool – Case of Inactive S-Boxes. Together with our theoretical observations, we also provide a practical `Sage` implementation based on our results (Section 4 for details). Given a square matrix, this tool and the underlying algorithm are able to detect the structural vulnerabilities described in this paper in the case of inactive S-boxes.

Impact on Hades. Our results also apply to the HADES strategy. In [21], the authors define the linear layer as a multiplication of the state with a fixed MDS matrix (namely, a matrix with maximum branch number), and no other properties have to be fulfilled by the linear layer. It follows that in the case of a "weak" MDS matrix (namely, a matrix that does not satisfy the properties proposed in this work), an attacker can potentially choose an input space of texts for which no S-box is activated over all rounds with partial S-box layers. In such a case, the security of the corresponding design may potentially be lower. Indeed, if no S-box is active, the degree of the function does not increase in the rounds with partial S-box layers when working with these chosen texts. Consequently, algebraic attacks become possible, as demonstrated in practice in [8].

At the same time, a "strong" linear layer can be used by the designer in order to increase the security against statistical attacks by exploiting the presence of rounds with partial S-box layers [26]. Hence, we suggest[2] that the choice of the MDS matrix that defines the linear layer for such a scheme must not be "weak" with respect to the properties given in this paper. We also point out that currently there is no known key-recovery attack on HADESMiMC exploiting these properties.

Related Work

Relation between Eigenvalues, Eigenvectors, and Invariant Subspace Trails. The relation between the eigenvalues and eigenvectors of the linear layer matrix and the existence of an infinitely long (invariant) subspace trail is already known in the literature. Such a relation was pointed out by Abdelraheem et al. [1], and later on generalized by Beyne in [7]. In more detail, Abdelraheem et al. found such a result by analyzing the invariant subspace trails of PRINTCIPHER (which was presented one year before in [28]), while Beyne found such a result as a generalization and improvement of the nonlinear invariant subspace attack on Midori-64 [33]. In paricular, in [7] *a connection between the eigenvalues of the correlation matrix that defines the round function and the existence of an invariant subspace trail* is made. More details are given in Appendix A.

Both the results presented in [1] and [7] focus on SPN ciphers and on invariant subspaces only. As a consequence, one has to consider the effect of the key (namely, such invariant subspace only holds in the case of weak keys).[3] Here we point out that the situation for partial SPN ciphers is different: The results found for SPN ciphers do not apply to the P-SPN case and vice-versa. First of all, in the P-SPN case, it is possible to set up infinitely long invariant subspaces independently of the choice of the key, of the key schedule, of the round constants, and of the details of the S-box. In other words, for the case of P-SPN ciphers, the existence of an infinitely long invariant subspace trail may depend *only* on the properties of the linear layer, which is not the case for an SPN cipher due to the full nonlinear layer. This has an impact on the subspace trail that can be set up. For SPN ciphers, due to the restriction on the key and on the round constants, it is possible to set up an invariant subspace trail e.g. of the form $R(\mathcal{U}+v)=\mathcal{U}+w$ only in the case in which v is in a subset of $\mathbb{F}^t$. This restriction is not necessary in the P-SPN case. Moreover, for this class of ciphers, following facts hold.

- The subspace trail *does not need to be invariant* in order to be infinitely long (namely, we do not restrict ourselves to the case $R(\mathcal{U}+v)=\mathcal{U}+w$).
- A non-trivial infinitely long invariant subspace trail can potentially be set up both in the case in which *no S-box is active*, and – for the first time – also in the case in which *some (or even all) S-boxes are active*. The crucial point is that we do not

[2] This is also supported by the designers of HADES (private communication).

[3] For completeness, we mention that the existence of such an invariant subspace can be easily prevented by a careful choice of the round constants, as was shown in [6].

need to consider the details of the S-box (namely, we do not require the S-box to fulfill any specific properties), which is not possible for the case of SPN ciphers.

More details about this are given in the following.

Infinitely Long Invariant Subspace Trails for Hades. We note that the idea of considering infinitely long invariant subspace trails for a certain class of linear layer matrices – that is, Cauchy matrices[4] over a Boolean field $\mathbb{F}_{2^n}$ generated in the very specific way given in [19] – has recently been studied independently in [26] and [8]. This particular class of matrices has always low multiplicative order[5] (see [8, Theorem 1] for details). This fact can be exploited in order to set up an infinitely long subspace trail. As a result, in [26] the authors show how to fix this problem by choosing Cauchy matrices that do not have such properties and how to exploit them in order to provide stronger security arguments against statistical attacks. In [8], the authors present a concrete application of such a weakness, by showing zero-sum distinguishers on reduced-round versions of the unkeyed permutations.

While the observations presented in [26] and [19] focus on a small class of (Cauchy) matrices, our results do not make such specific assumptions about the matrices used in the linear layers. This line of research is supported by the fact that our results can be exploited by future versions of cryptographic designs with partial S-box layers. To be more concrete, in the case of HADES-like ciphers, the MDS matrix can also be replaced by a matrix with a smaller (known) branch number (e.g., a near-MDS matrix). The results presented here are naturally relevant in such a case.

Security against Statistical Attacks. In this paper, we present properties which a matrix defining the linear layer *must not* satisfy in order to prevent infinitely long subspace trails. However, in general this does not help in predicting the number of rounds necessary to provide security against statistical attacks. Such a contribution can be found in [5], where the authors present tools which analyze the security of a given partial SPN cipher against statistical attacks. We note that in this paper the authors do not analyze which properties a matrix must satisfy in order to prevent infinitely long subspace trails – as we do here. In this sense, we think that our work and the one proposed in [5] complement each other.

2 Preliminaries

Notation. We denote subspaces with calligraphic letters (e.g., $\mathcal{S}$). Further, we use the superscript notation together with parentheses to differentiate subspaces with similar properties (e.g., $\mathcal{S}^{(i)}$). Matrices are denoted by non-calligraphic letters, and the superscript notation for matrices is used to indicate powers of matrices in their traditional form. The entry of a matrix M in the j-th column of the i-th row is denoted by $M_{i,j}$, where $M_{1,1}$ denotes the entry in the first column of the first row. Finally, we use $a \mid\mid b$ to denote the concatenation between two values a and b.

2.1 SPN and Partial SPN Ciphers

In this paper, we will focus on partial SPN ciphers over $\mathbb{F}_q$, where q is a prime power. These ciphers are similar to classical (full) SPN ciphers, with the only difference being that the S-boxes (i.e., the nonlinear functions of the cipher) are not applied to the whole state.

[4] Cauchy matrices are a class of MDS matrices.

[5] The multiplicative order of a matrix M is the smallest (integer positive) exponent $k \geq 1$ such that $M^k = \mu I$, where $\mu \in \mathbb{F}$ and I is the identity matrix.

SPN Ciphers. We denote the application of r rounds of an SPN cipher by $E_k^r : \mathbb{F}^t \to \mathbb{F}^t$, where $k \in \mathbb{F}^t$ is a fixed secret key and $t \in \mathbb{N}$ denotes the number of cells. For every input $x = (x_1, x_2, \dots, x_t) \in \mathbb{F}^t$ we write $E_k^r(x) = (F_r \circ \cdots \circ F_1 \circ F_0)\,(x + k^{(0)})$, where $F_i : \mathbb{F}^t \to \mathbb{F}^t$ is defined as $F_i(x) = R(x) \oplus k^{(i)}$ for $i \in [1, r]$. The round keys $k^{(0)}, k^{(1)}, \dots, k^{(r)} \in \mathbb{F}^t$ are derived from the master key k using a key schedule (alternatively, they can also be randomly chosen elements). We denote by R the composition of the S-box and the linear layer, i.e., we have $R : \mathbb{F}^t \to \mathbb{F}^t$ with

$$R(x) = (M \circ S)(x) = M(S_1(x), S_2(x), \dots, S_t(x)),$$

where $S_i : \mathbb{F} \to \mathbb{F}$ for $i \in [1, t]$ is a nonlinear polynomial S-box. Finally, $M : \mathbb{F}^t \to \mathbb{F}^t$ denotes an invertible non-trivial linear layer defined by the multiplication with a matrix of the form

$$M(x) = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,t} \\ M_{2,1} & M_{2,2} & \dots & M_{2,t} \\ \vdots & \vdots & \ddots & \vdots \\ M_{t,1} & M_{t,2} & \dots & M_{t,t} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{pmatrix},$$

where $M_{i,j} \in \mathbb{F}$ for $i \in [1, t]$ and $j \in [1, t]$.

Definition 1. A linear layer $M \in \mathbb{F}^{t \times t}$ is *non-trivial* if it ensures full diffusion[6] (in the sense that each word of the output depends on each word of the input and vice versa) after a *finite* number of rounds.

Note that all SPN ciphers can be written in this way. Just to give some examples, if M is an MDS matrix[7], the cipher is similar to SHARK [32]. For AES [15] or AES-like ciphers (where the linear layer is obtained as a combination of a ShiftRows and a MixColumns operation), many words of M are equal to 0.

Partial SPN (P-SPN) Ciphers. The main and only difference to an SPN cipher regards the S-box layer. For the case of partial SPN (P-SPN) ciphers, the round (and so the S-box layer) is defined as

$$R(\cdot) = M \circ (\underbrace{S_1 \,||\, \dots \,||\, S_s \,||\, I_{s+1} \,||\, \dots \,||\, I_t}_{\text{S-box layer}})(\cdot), \tag{1}$$

where $1 \leq s < t$ and where $I_{s+1} = I_{s+2} = \cdots = I_t$ are identity functions. In other words, instead of having a full S-box layer, the nonlinear functions are applied only to a part of the state, while the rest of the state remains unchanged.

In this paper, we assume that the s S-boxes are applied on the first s words. Note that given any partial SPN cipher, it is always possible to find an equivalent representation such that the S-boxes are applied to the first s words.

Hades-Like Ciphers. The recently proposed HADES-strategy [21] combines both SPN and partial SPN ciphers in the following way:

- The initial R_f and the final R_f rounds contain full S-box layers, for a total of $R_F = 2R_f$ rounds with full S-box layers.

[6] The linear layer defined by the multiplication with M provides full diffusion if there exists $r \in \mathbb{N}$ such that the function that describes $[R^r(x)]_i$ (hence, $[M^r \cdot x]_i$) depends on x_j for a state x, where $i \in [1, t]$ and $j \in [1, t]$. For example, the identity matrix does not fulfill this condition.

[7] A matrix $M \in \mathbb{F}^{t \times t}$ is called a maximum distance separable (MDS) matrix iff it has a branch number $\mathcal{B}(M)$ equal to $\mathcal{B}(M) = t + 1$. The branch number of M is defined as $\mathcal{B}(M) = \min_{x \in \mathbb{F}^t} \{\text{wt}(x) + \text{wt}(M(x))\}$, where $\text{wt}(\cdot)$ is the bundle weight in wide trail terminology. A matrix M is MDS if and only if every submatrix of M is invertible.

- In the middle of the construction, R_P rounds with partial S-box layers are used.

Roughly speaking, R_F rounds provide security against statistical attacks, while R_P rounds are exploited in order to increase the overall degree of the encryption/decryption function, in an attempt to provide security against algebraic attacks.

Before going on, let us recall that the middle (partial) rounds are not exploited to increase the security against statistical attacks. Using these partial rounds to provide additional security arguments against this class of attacks has been considered in [26], where the authors are able to improve the lower bounds on the number of active S-boxes for some instantiations of HADES.

2.2 Invariant Subspaces and Subspace Trails

2.2.1 Invariant Subspace Attack

The invariant subspace attack, introduced in [28] and reconsidered e.g. in [29], is based on the possibility to set up an invariant subspace trail, defined as follows.

In order to recall the definition, let F denote the round function of a key-alternating block cipher, and let $\mathcal{U} + a$ denote a coset of a vector space $\mathcal{U}$. By $\mathcal{U}^c$ we denote the complementary subspace of $\mathcal{U}$. Finally, two cosets $\mathcal{U} + a$ and $\mathcal{U} + b$ are considered to be equivalent if and only if $a + b \in \mathcal{U}^c$.

Definition 2. Let K_{weak} be a set of keys and $k \in K_{\text{weak}}$, with $k = (k^{(0)}, k^{(1)}, k^{(2)}, \ldots, k^{(r)})$, where $k^{(j)}$ is the j-th round key. For $k \in K_{\text{weak}}$, the subspace $\mathcal{IS}$ generates an invariant subspace trail of length r for the function $F_k(\cdot) = F(\cdot) + k$ if for each $i \in [1, r]$ there exists a non-empty set $A_i \subseteq \mathcal{IS}^c$ (where $\cdot^c$ denotes the complement) for which

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

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

Let us remark the main difference for invariant subspace attacks when working with partial SPN ciphers instead of SPN ones. In this last case and to the best of our knowledge, the sets A_i are (almost always) non-trivial subsets of $\mathbb{F}^t$. As shown in the following, this restriction is not mandatory for the case of partial SPN ciphers. In particular, we work independently of the details of the S-box, and we assume that for each $i : A_i = \mathbb{F}^t$ and that the set k_{weak} is equal to the set of all possible keys.

2.2.2 Subspace Trail Attack

Subspace trails were first defined in [22], and they are strictly related to truncated differential attacks, as shown in [30]. We refer to [22] for more details about the concept of subspace trails. However, our treatment here is meant to be self-contained.

Definition 3. Let $(\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_{r+1})$ denote a set of $r + 1$ subspaces with $\dim(\mathcal{U}_i) \leq \dim(\mathcal{U}_{i+1})$. If for each $i \in [1, r]$ and for each a_i there exists a (unique) $a_{i+1} \in \mathcal{U}_{i+1}^c$ such that

$$F(\mathcal{U}_i + a_i) \subseteq \mathcal{U}_{i+1} + a_{i+1},$$

then $(\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_{r+1})$ 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*.

Iterative (Constant-Dimensional) Subspace Trails. We now introduce the concept of infinitely long iterative (constant-dimensional) subspace trails.

Definition 4. Let $\{\mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_r\}$ be a constant-dimensional subspace trail for r rounds. We call this subspace trail an *r-round iterative (constant-dimensional) subspace trail* for the considered cipher if it repeats itself an infinite number of times, i.e., if

$$\{\mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_r, \mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_r, \dots, \mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_r, \dots\}$$

is an infinitely long subspace trail.

Obviously, an invariant subspace trail is also an iterated subspace trail for the case of P-SPN ciphers (under the previous assumptions), while not every iterated subspace trail is also an invariant subspace trail.

While, to the best of our knowledge, no example of infinitely long iterative constant-dimensional subspace trails for SPN ciphers is given in the literature, a poor choice of the linear layer matrix allows to find them for the case of P-SPN ciphers.

Weak-Key Subspace Trails. For completeness, we mention that a generalization of the two previous attacks, called "weak-key subspace trail attack", has been proposed in [20] (it basically corresponds to a subspace trail that holds for a class of weak keys only).

2.3 Preliminary Assumptions

Before presenting our results, we make clear the following assumptions that we consider in our work.

"Generic" S-Box: We assume that the S-box has no linear structure. Under this assumption, then one can work independently of the details of the S-box. Indeed, as was shown in [30], there are only two essential subspace trails ($\{0\} \to \{0\}$ and $\mathbb{F} \to \mathbb{F}$) when working at word level if the S-box has no non-trivial linear structure. For example, both the AES S-box and the cube one ($x \mapsto x^3$) satisfy this assumption. In other words, for an S-box S, it is not possible to find $\mathcal{U}, \mathcal{V} \subseteq \mathbb{F}$ such that for each u there exists v such that $S(\mathcal{U} + u) = \mathcal{V} + v$. If this is not the case, the details of the S-box can be exploited in order to set up an invariant subspace attack.

No Weak Keys: We only consider infinitely long constant subspace trails which are independent of the key and of the key schedule. In other words, we assume that the key schedule is designed in order to prevent setting up infinitely long constant subspace trails for a class of weak keys.

P-SPN with $s < \lfloor t/2 \rfloor$: We further assume $s < \lfloor t/2 \rfloor$. This implies that the choice of the linear layer is crucial in order to guarantee that at least one S-box is active after a *finite* number of rounds. Indeed, in the case in which a fixed linear layer matrix M is used, let $2 \leq b \leq t+1$ be its branch number. If $2t - 2s < b$, then at least $b + 2s - 2t \geq 1$ S-boxes are active in every two consecutive rounds. Note that this can never happen if $s < \lfloor t/2 \rfloor$ (equivalently, $s \leq \lfloor t/2 \rfloor - 1$), since $2t - 2s \geq t + 2$ and $b \leq t + 1$.

3 Subspaces Trails for P-SPN Ciphers (No Active S-Boxes)

In the case of SPN ciphers, (weak-key) infinitely long subspace trails can be prevented by carefully choosing the round constants (see [6] for details) and by exploiting the fact that a full S-box layer together with a reasonable linear layer provides full diffusion after a finite number of rounds. In the case of P-SPN ciphers, however, the situation is different.

First of all, due to the fact that the S-box layer is not complete, the details of the round constants (together with a non-trivial linear layer) are not sufficient by themselves

to provide security against the subspace attacks just recalled. In this sense, the linear layer plays a crucial role in order to provide security. Several strategies have been proposed in the literature in order to prevent subspace attacks. Focusing on the linear layer, the designers of LowMC [2] proposed to use different random linear layers in each round in order to prevent this approach and to provide security against statistical attacks. As already mentioned in the introduction, this choice has some drawbacks, and hence we focus on cases where the same linear layer is used in each round – an approach which is taken by the designers of `Zorro` [18], for example.

3.1 Preliminary Results

Due to the fact that the nonlinear layer is only partial in a P-SPN cipher, parts of the state go through the S-box layer unchanged. In particular, if the nonlinear layer is composed of $s \geq 1$ S-boxes and $t - s \geq 1$ identity functions, it is always possible to find an initial subspace such that no S-box is active in the first

$$\left\lfloor \frac{t-s}{s} \right\rfloor \quad \text{rounds.}$$

Indeed, assume that the s S-boxes are applied to the first s words. Thus, by choosing texts in the same coset of $\mathcal{S} = \langle v_1, v_2, \ldots, v_d \rangle$ such that

$$\forall i \in \left\{1, 2, \ldots, \left\lfloor \frac{t-s}{s} \right\rfloor \right\}, \forall j \in [1, d] : [M^i \cdot v_j]_{1,2,\ldots,s} = \underbrace{0 \,||\, 0 \,||\, \cdots \,||\, 0}_{\in \mathbb{F}^s},$$

where $[v]_{1,2,\ldots,s}$ denotes the first s words of a vector $v \in \mathbb{F}^t$, it follows that no S-box is active in the first $\left\lfloor \frac{t-s}{s} \right\rfloor$ rounds. Note that

$$d = \dim(\mathcal{S}) \geq t + 1 - s \cdot \left\lfloor \frac{t}{s} \right\rfloor,$$

where $s < t$. We formalize this result in the following definition.

Definition 5. Consider the case of a P-SPN cipher over $\mathbb{F}^t$ with $1 \leq s \leq t$ S-boxes applied to the first s words. Let $\mathcal{S}^i$ be defined as

$$\mathcal{S}^{(i)} = \left\{ v \in \mathbb{F}^t \;\middle|\; \forall j \leq i : \; [M^j \cdot v]_{1,2,\ldots,s} = \underbrace{0 \,||\, 0 \,||\, \cdots \,||\, 0}_{\in \mathbb{F}^s} \right\}, \tag{2}$$

where $\mathcal{S}^{(0)} = \mathbb{F}^t$, and where $\dim\left(\mathcal{S}^{(i)}\right) \geq t - i \cdot s$. For any $x, y \in \mathbb{F}^t$ s.t. x, y are in the same coset of $\mathcal{S}^{(i)}$, then no S-boxes are active in the first i (consecutive) rounds.

The previous definition can naturally be extended to more rounds, for example in the case in which $\dim\left(\mathcal{S}^{\left(\left\lfloor \frac{t-s}{s} \right\rfloor\right)}\right) \geq s$. Obviously, this depends on the details of the linear layer, and it is formally defined in the following.

Proposition 1. *Consider the case of a P-SPN cipher over $\mathbb{F}^t$ with $1 \leq s \leq t$ S-boxes applied to the first s words, and let $\mathcal{S}^{(i)}$ be defined as before. Let $\mathfrak{R} \geq \left\lfloor \frac{t-s}{s} \right\rfloor$ s.t.*

$$\dim(\mathcal{S}^{(\mathfrak{R})}) \geq 1 \qquad \textit{and} \qquad \dim(\mathcal{S}^{(\mathfrak{R}+1)}) = 0. \tag{3}$$

For each $r \leq \mathfrak{R}$, the collection

$$\left\{ \mathcal{S}^{(r)}, M \cdot \mathcal{S}^{(r)}, M^2 \cdot \mathcal{S}^{(r)}, \ldots, M^{r-1} \cdot \mathcal{S}^{(r)} \right\}$$

forms a subspace trail for the first r rounds (with no active S-boxes).

This well-known result does not require any assumption on the matrix M that defines the linear layer. In the following, we will therefore explore in which cases it is possible to set up an infinitely long subspace trail.

3.2 Linear Layers with Low Multiplicative Order

Here we provide a first sufficient condition that, if satisfied, leads to an infinitely long constant subspace attack.

Proposition 2. *Let $M \in \mathbb{F}^{t\times t}$ be an invertible matrix. If there exists $\mathfrak{k} \in \{2,\dots,\mathfrak{R}\}$, where $\mathfrak{R} \geq \left\lfloor \frac{t-s}{s} \right\rfloor$ is defined as in Eq. (3), and $\mu \in \mathbb{F}\setminus\{0\}$ such that $M^{\mathfrak{k}} = \mu \cdot I$ (equivalently, if M has multiplicative order $\mathfrak{k}$), where $I \in \mathbb{F}^{t\times t}$ is the identity matrix, then it is always possible to find an infinitely long invariant subspace trail.*

Proof. As we have seen before, it is always possible to find an initial subspace such that no S-box is active in the first $\mathfrak{R}$ rounds. This subspace, constructed as in Definition 5, yields a subspace trail of the form

$$\left\{\mathcal{S}^{(\mathfrak{k})}, M\cdot\mathcal{S}^{(\mathfrak{k})}, M^2\cdot\mathcal{S}^{(\mathfrak{k})}, \dots, M^{\mathfrak{k}-1}\cdot\mathcal{S}^{(\mathfrak{k})}\right\}$$

for the first $\mathfrak{k} \leq \mathfrak{R}$ rounds. Here, we only have to show that such a $\mathfrak{k}$-round subspace trail is repeated infinitely. To do this, we compute $M^i \cdot \mathcal{S}^{(\mathfrak{k})}$ for $i \geq \mathfrak{k}$. By definition, there exist $j, h \in \mathbb{N}$ s.t. $i = j\mathfrak{k} + h$, where $h < \mathfrak{k}$. Thus,

$$M^i\cdot\mathcal{S}^{(\mathfrak{k})} = (M^{\mathfrak{k}})^j \cdot M^h \cdot \mathcal{S}^{(\mathfrak{k})} = (\mu\cdot I)^j \cdot M^h \cdot \mathcal{S}^{(\mathfrak{k})} = M^h\cdot\mathcal{S}^{(\mathfrak{k})}. \qquad \square$$

Note that the previous result is independent of the key, the key schedule, and the S-box. Indeed, since no S-box is active, the key only changes the coset of $\mathcal{S}^{(\mathfrak{k})}$: Given two plaintexts whose difference is in $\mathcal{S}^{(\mathfrak{k})}$, the resulting difference after i rounds is in $M^i\cdot\mathcal{S}^{(\mathfrak{k})}$.

Example. In the following, the circulant matrix $\mathrm{circ}^L(c_1, c_2, \dots, c_n)$ is defined as

$$\mathrm{circ}^L(c_1,c_2,\dots,c_n) = \begin{pmatrix} c_1 & c_2 & \dots & c_{n-1} & c_n \\ c_2 & c_3 & \dots & c_n & c_1 \\ \vdots & & \ddots & & \vdots \\ c_n & c_1 & \dots & c_{n-2} & c_{n-1} \end{pmatrix},$$

and we use the notation $\mathrm{circ}^R(c_1, c_2, \dots, c_n)$ if the words are rotated to the right instead. Given a circulant invertible matrix $M = \mathrm{circ}^L(\alpha,\beta,\gamma,\delta) \in \mathbb{F}^{4\times 4}$, we have

$$M^2 = \mathrm{circ}^R(\alpha^2+\beta^2+\gamma^2+\delta^2, \alpha\beta+\beta\gamma+\alpha\delta+\gamma\delta, 2\alpha\gamma+2\beta\delta, \alpha\beta+\beta\gamma+\alpha\delta+\gamma\delta).$$

The condition $M^2 = (\alpha^2+\beta^2+\gamma^2+\delta^2)\cdot I$ is satisfied if and only if $2\alpha\gamma+2\beta\delta = 0$ (which is always satisfied over a binary field) and $\alpha\beta+\beta\gamma+\alpha\delta+\gamma\delta = 0$. A concrete example is given by $M = \mathrm{circ}\left(\beta^2, \beta, 1, -\beta\right)$ for $\beta \neq 0$.[8]

Cauchy Matrices: Recent Results in the Literature

Another concrete example has recently been pointed out by Keller et al. [26] and by Beyne et al. [8]. In these papers, the authors focus on the Cauchy matrix $M \in \mathbb{F}_{2^n}^{t\times t}$ proposed in [19] and defined as

$$M_{i,j} = \frac{1}{x_i + x_j + r},$$

where $x_i = i-1$ for $i \in [1,t]$ and $t \leq r \leq p-t$. Such a matrix is used as the linear layer of some HADES-like permutations, namely STARKAD$^\pi$ and POSEIDON$^\pi$ [19]. In there, they prove that if $t = 2^\tau$, the matrix has a multiplicative complexity equal to 2, namely that M^2 is a multiple of the identity.[9]

[8] Note that in some cases it is also possible to choose β in order to construct an MDS matrix. E.g., working over GF(31), $M = \mathrm{circ}\left(3^2, 3, 1, -3\right) = \mathrm{circ}(9,3,1,28)$ is an MDS matrix and it satisfies $M^2 = 7\cdot I$.

[9] In [8], the authors assume that $\{x_1, x_2, \dots, x_t\}$ forms a closed subgroup of $GF(2^n)$. By definition of x_i, this is always the case for STARKAD$^\pi$ and POSEIDON$^\pi$ if $t < 2^n$ is a power of 2.

3.3 Infinitely Long Invariant Subspace Trails

As we are going to show, the previous condition is sufficient but in general not necessary in order to prevent infinite subspace trails. Namely, there exist linear layer matrices that do not satisfy the previous condition and for which it is still possible to set up an infinitely long subspace trail attack. As a concrete example, consider the matrix

$$M = \mathrm{circ}(\alpha, \beta, \beta) \in \mathbb{F}^{3\times 3}$$

for some particular non-trivial values of $\alpha, \beta \in \mathbb{F}$ such that M is invertible (e.g., $\alpha \neq \beta$ to guarantee that M^{-1} exists). For $s = 1$, the subspace

$$\mathcal{V} = \left\langle (0, 1, -1)^T \right\rangle = \left\{ a \cdot (0, 1, -1)^T \;\middle|\; a \in \mathbb{F} \right\} \subset \mathbb{F}^3$$

yields an infinitely long invariant subspace. Indeed, since $M^i = \mathrm{circ}(\gamma, \delta, \delta)$ for particular γ and δ (which depends on $i \geq 1$ and on α, β), it follows that $\left(M^i \cdot \mathcal{V}\right)_0 = 0\gamma + \delta - \delta = 0$. Choosing $\alpha = 1$ and $\beta = 2$ in $\mathbb{F}_p$ for large p results in $M \neq I$ and $M^2 \neq I$, which implies that the previous condition is not a necessary one.

Starting with this example, here we show a connection between the existence of an infinitely long invariant subspace and the eigenspaces of M. First of all, we recall a preliminary concept.

Definition 6. $M \in \mathbb{F}^{t\times t}$ is a diagonalizable matrix if and only if there exists an (invertible) matrix $P \in \mathbb{F}^{t\times t}$ s.t. $P^{-1} \cdot M \cdot P = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_t)$ is a diagonal matrix. The subspace $\mathcal{P} = \langle \rho_1, \rho_2, \ldots, \rho_d \rangle \in \mathbb{F}^t$ that satisfies the condition $M \cdot \rho_i = \lambda \cdot \rho_i$ for $i \in [1, d]$ is called the (right) eigenspace of M for the eigenvalue λ.

A $t \times t$ matrix M over a field $\mathbb{F}$ is diagonalizable if and only if the sum of the dimensions of its eigenspaces is equal to t.

Theorem 1. *Given a P-SPN cipher with s S-boxes per round, let $M \in \mathbb{F}^{t\times t}$ be an invertible matrix. Let $\lambda_1, \lambda_2, \ldots, \lambda_\tau$ be its eigenvalues and let $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_\tau$ be the corresponding eigenspaces*[10]*. Let*

$$\mathcal{IS} = \langle \mathcal{P}_1 \cap \langle e_{s+1}, e_{s+2}, \ldots, e_t \rangle, \mathcal{P}_2 \cap \langle e_{s+1}, e_{s+2}, \ldots, e_t \rangle, \ldots, \mathcal{P}_\tau \cap \langle e_{s+1}, e_{s+2}, \ldots, e_t \rangle \rangle .$$

If $1 \leq \dim(\mathcal{IS}) < t$, then $\mathcal{IS}$ generates a (non-trivial) infinitely long invariant subspace trail (with no active S-box).

Equivalently, let $\mathcal{IS}$ be defined as

$$\mathcal{IS} = \langle \mathcal{P}'_1, \mathcal{P}'_2, \ldots, \mathcal{P}'_\tau \rangle ,$$

where $\mathcal{P}'_i \subseteq \mathcal{P}_i$ is a subspace of $\mathcal{P}_i$ for $i \in [1, \tau]$, such that

$$\mathcal{IS} \cap \langle e_{s+1}, e_{s+2}, \ldots, e_t \rangle = \mathcal{IS}.$$

If $\dim(\mathcal{IS}) \geq 1$, then an infinitely long invariant subspace trail is generated. This equivalent definition will be used in the following.

Proof. To prove the previous result, we have to show that $M \circ S(x) \in \mathcal{IS}$ for each $x \in \mathcal{IS}$ (we omit the key and constant additions since we deal with differences).

1. Since $\mathcal{IS} \subseteq \mathbb{F}^t \setminus \langle e_1, \ldots, e_s \rangle$, no S-box is active. Hence, only the coset changes through the S-box layer.

[10] We recall that $\dim(\mathcal{P}) \geq 1$.

2. Since the linear layer is linear, it is possible to focus on a single space $\mathcal{P}_i \cap \langle e_{s+1}, e_{s+2}, \dots, e_t \rangle$. W.l.o.g., let $\mathcal{P}_1 = \langle \rho_1, \rho_2, \dots, \rho_d \rangle$, and let $\mathcal{P}_1 \cap \langle e_{s+1}, e_{s+2}, \dots, e_t \rangle = \langle \rho'_1, \rho'_2, \dots, \rho'_\delta \rangle$, where $\delta \leq d$. Since each $\rho' \in \langle \rho'_1, \rho'_2, \dots, \rho'_\delta \rangle$ is a linear combination of eigenvectors with the same eigenvalue, the result is still a multiple of ρ' when applying M. Hence, the result is again in $\mathcal{P}_1 \cap \langle e_{s+1}, e_{s+2}, \dots, e_t \rangle$.

The result follows immediately. □

We highlight that this fact holds since we work independently on the eigenspaces of M. Consider indeed the case in which $\mathcal{P}_1 = \langle v \rangle$, $\mathcal{P}_2 = \langle w \rangle$, and $\langle \mathcal{P}_1, \mathcal{P}_2 \rangle \cap \langle e_{s+1}, e_{s+2}, \dots, e_t \rangle = \langle v + \alpha w \rangle$. Given $x \in \langle \mathcal{P}_1, \mathcal{P}_2 \rangle \cap \langle e_{s+1}, e_{s+2}, \dots, e_t \rangle$, $M \cdot x$ does not belong to such a subspace since $M \cdot (v + \alpha w) = \lambda_v \cdot \left(v + \alpha \cdot \frac{\lambda_w}{\lambda_v} \cdot w \right)$, where $\lambda_w \neq \lambda_v$.

We emphasize that the previous result provides only a sufficient condition. The problem to prove/disprove that it is also necessary (and/or to find a necessary condition) is left open for future research.

SPN Ciphers versus Partial-SPN Ciphers. Let us recall one more time that

(1) in the previous theorem we do not require that M is diagonalizable,

(2) the round keys, constants, and the key schedule do not influence the result (they only change the coset of $\mathcal{IS}$, but not the difference), and that

(3) potentially other invariant subspace trails can be set up if the details of the S-box are taken into account.

These facts emphasize the difference when dealing with invariant subspace trails in the case of SPN ciphers and in the case of P-SPN ones. Indeed, in the first case the condition $R(\mathcal{U} + v) = \mathcal{U} + w$ holds only if v is in a subset of $\mathbb{F}^t$ and the key is a weak key. In the case of P-SPN, no limitation on v is imposed, or, in other words, the invariant subspace trail holds for any initial coset of $\mathcal{U}$.

Example. Besides the example $\text{circ}(\alpha, \beta, \beta)$ just given, let us consider the following invertible 4×4 matrix M over $\mathbb{F}$ (for $s = 1$) defined by

$$M = \begin{pmatrix} M_{1,1} & b & c & d \\ M_{2,1} & M_{2,2} & (-d + M_{2,2} \cdot c)/b & (-c + M_{2,2} \cdot d)/b \\ M_{3,1} & M_{3,2} & M_{3,2} \cdot c/b & (b + M_{3,2} \cdot d)/b \\ M_{4,1} & M_{4,2} & (b + M_{4,2} \cdot c)/b & M_{4,2} \cdot d/b \end{pmatrix},$$

where $b \neq 0$. Properly choosing b, c, d, and the remaining entries of M provides invertibility and full diffusion (at word level after a finite number of rounds) for cryptographic purposes.

Given the subspace

$$\mathcal{V} = \left\langle v_0 := (0, -c, b, 0)^T, v_1 := (0, -d, 0, b)^T \right\rangle,$$

it is not hard to see that $M \cdot v_0 = v_1$ and $M \cdot v_1 = v_0$. Hence, $\mathcal{V}$ yields an infinitely long invariant subspace trail. The connection with the theorem just given is due to the fact that the subspace $\mathcal{V}$ is related to the eigenvalues of M and their corresponding eigenspaces. Indeed,

$$M \cdot (v_0 + v_1) = (v_0 + v_1) \qquad \text{and} \qquad M \cdot (v_0 - v_1) = -(v_0 - v_1).$$

3.4 Infinitely Long Iterative (Non-Invariant) Subspace Trails

Until now, we focused only on the properties of M. However, since we are not working on a closed field, a possible generalization of the previous result can be presented.

Let M be an invertible matrix in $\mathbb{F}^{t\times t}$. If M is diagonalizable, then M^k, where $k \in \mathbb{N}$, is also diagonalizable:

$$P \cdot M \cdot P^{-1} = D \implies P \cdot M^k \cdot P^{-1} = D^k.$$

The other direction is not true in general, as given in the following proposition.

Proposition 3 ([24]). *If M is invertible, $\mathbb{F}$ is algebraically closed, and M^k is diagonalizable for some k that is not an integer multiple of the characteristic of $\mathbb{F}$, then M is diagonalizable.*

Since no finite field can be algebraically closed, it follows that M^k may contain more eigenvalues than M. In other words, if λ is an eigenvalue of M, then λ^k is also an eigenvalue of M^k. The opposite is not true in general: Given an eigenvalue λ of M^k, it is possible that $\lambda^{1/k}$ does not exist, which means that there is no corresponding eigenvalue for M.

This fact has an impact on the existence of infinitely long subspace trails. Indeed, in the case in which there exists $k \geq 2$ s.t. M^k has more eigenvalues than M, it is potentially possible to set up an iterated subspace trail which is not invariant (and for which no S-box is active) for any number of rounds.

Theorem 2. *Given a P-SPN cipher with s S-boxes, let $M \in \mathbb{F}^{t\times t}$ be an invertible matrix. Let $\lambda_1^{(\mathfrak{k})}, \lambda_2^{(\mathfrak{k})}, \ldots, \lambda_\tau^{(\mathfrak{k})}$ be the eigenvalues of $M^{\mathfrak{k}}$ for some $\mathfrak{k} \geq 1$, and let $\mathcal{P}_1^{(\mathfrak{k})}, \mathcal{P}_2^{(\mathfrak{k})}, \ldots, \mathcal{P}_\tau^{(\mathfrak{k})}$ be their corresponding eigenspaces (where $\tau \leq t$). For each $r \geq 1$, let $\mathcal{IS}^{(r)}$ be the subspace defined as*

$$\mathcal{IS}^{(r)} = \left\langle \mathcal{S}^{(r)} \cap \mathcal{P}_1^{(r)}, \mathcal{S}^{(r)} \cap \mathcal{P}_2^{(r)}, \ldots, \mathcal{S}^{(r)} \cap \mathcal{P}_\tau^{(r)} \right\rangle,$$

*where $\mathcal{S}^{(r)}$ is the subspace constructed as in Definition 5 such that no S-box is active in the first r rounds. If $1 \leq \dim\left(\mathcal{IS}^{(r)}\right) < t$, an infinitely long **iterated** subspace trail of the form*

$$\left\{\mathcal{IS}^{(r)}, M \cdot \mathcal{IS}^{(r)}, M^2 \cdot \mathcal{IS}^{(r)}, \ldots, M^{\mathfrak{k}-1} \cdot \mathcal{IS}^{(r)}\right\}$$

is generated.

The proof is equivalent to the one given before (note that no S-box is active due to the construction of $\mathcal{S}^{(r)}$). Moreover, we point out that this result reduces to the previous one if $\mathfrak{k} = 1$, since $\mathcal{S}^{(1)} = \langle e_{s+1}, e_{s+2}, \ldots, e_t \rangle$.

Low Multiplicative Order. This result also includes the case in which the matrix has low multiplicative order, as shown in the following corollary.

Corollary 1. *Theorem 2 implies the result presented in Proposition 2.*

Proof. Assume there exists $\mathfrak{k}$ such that $M^{\mathfrak{k}} = \mu \cdot I$. Then $e_1, e_2, \ldots, e_t$ are all eigenvectors of $M^{\mathfrak{k}}$ with eigenvalue μ (equivalently, the space $\mathbb{F}^t$ is an eigenspace of $M^{\mathfrak{k}}$ w.r.t. the same eigenvalue μ). Moreover, let $\mathcal{S}^{(\mathfrak{k})}$ be the subspace constructed as in Definition 5 such that no S-box is active in the first $\mathfrak{k}$ rounds. Since $\langle e_1, e_2, \ldots, e_t \rangle$ is an eigenspace of $M^{\mathfrak{k}}$ corresponding to the eigenvalue μ, it follows that $\mathcal{S}^{(\mathfrak{k})}$ is an invariant subspace of $M^{\mathfrak{k}}$. Hence, due to the previous considerations,

$$\left\{\mathcal{S}^{(\mathfrak{k})}, M \cdot \mathcal{S}^{(\mathfrak{k})}, M^2 \cdot \mathcal{S}^{(\mathfrak{k})}, \ldots, M^{\mathfrak{k}-1} \cdot \mathcal{S}^{(\mathfrak{k})}\right\}$$

is an infinitely long iterated (constant-dimensional) subspace trail. □

We remark that the two conditions are not equivalent, as shown in the following concrete example.

Example. Consider the circulant matrix $M = \mathrm{circ}(a, b, c, d)$ over $\mathbb{F}^4$. The eigenvalues of M are

$$a + b + c + d, \quad \pm\sqrt{a^2 + b^2 - 2ac + c^2 - 2bd + d^2}, \quad a - b + c - d,$$

while the eigenvalues of M^2 are

$$\begin{aligned}
&(a + b + c + d)^2 = a^2 - 2a(b - c + d) + b^2 - 2b(c - d) + c^2 - 2cd + d^2,\\
&(a - b + c - d)^2 = a^2 + 2a(b + c + d) + b^2 + 2b(c + d) + c^2 + 2cd + d^2,\\
&a^2 + b^2 - 2ac + c^2 - 2bd + d^2, \quad a^2 + b^2 - 2ac + c^2 - 2bd + d^2.
\end{aligned}$$

Since $x \mapsto x^2$ is not a permutation over $\mathbb{F}_p$ for a prime $p \geq 3$ (see Hermite's criterion), there exist a, b, c, d, such that $a^2 + b^2 - 2ac + c^2 - 2bd + d^2$ is not a square. Hence, for certain values of $a, b, c, d \in \mathbb{F}_p$, it is possible that M has two eigenvalues, while M^2 has always four eigenvalues.[11] As shown in details in Appendix B, this fact can be exploited in order to construct a matrix M such that

(1) the corresponding iterated subspace trail $\mathcal{IS}$ is iterated infinitely many times while not being invariant (namely, it is mapped into itself after two rounds, but not after a single round), and

(2) M^2 is not a multiple of the identity.

Given a P-SPN cipher over $\mathbb{F}_p^5$ with $s = 1$, a concrete example of such a matrix is given by

$$\mathfrak{M} = \begin{pmatrix} x & y_0 & y_1 & y_0 & y_1 \\ z_0 & a & b & c & d \\ z_1 & b & c & d & a \\ z_2 & c & d & a & b \\ z_3 & d & a & b & c \end{pmatrix}$$

for particular values of $a, b, c, d, x, y_i, z_j \in \mathbb{F}_p$ (for which the matrix is invertible and $\mathrm{circ}(a, b, c, d)$ has only 2 eigenvalues), where the iterated (non-invariant) subspace trail is given by

$$\left\{\mathcal{IS} = \langle (0, 0, 1, 0, -1)^T \rangle,\ \mathfrak{M} \cdot \mathcal{IS} = \langle (0, b - d, c - a, d - b, a - c)^T \rangle\right\},$$

where $\mathfrak{M}^2 \cdot \mathcal{IS} = \mathcal{IS}$ and where $\mathfrak{M}^2$ is not a multiple of the identity (see Appendix B for concrete examples).

4 Practical Tests

Let M be the matrix that defines the linear layer. If M^k has no eigenvalues and eigenvectors for each integer $k \geq 1$, the previous result does not apply. Hence, choosing a matrix which satisfies this condition could be a possible solution to prevent the previous attack. Unfortunately, this is not the case since

(1) as we are going to show in the following with a concrete example, there exist matrices that do not have any eigenvalues and eigenvectors but for which the previous attack works, and

(2) we may be forced to use a matrix which has eigenvalues and eigenvectors. The crucial point is that, even if M^k has one or more eigenvalues and eigenvectors, the previous attack does not necessarily work. Indeed, the applicability depends on the details of the subspace $\mathcal{V}$ for which no S-box is active in the first r rounds.

[11] For example, choosing $(a, b, c, d) = (1, 1, 2, 3)$, $a^2 + b^2 - 2ac + c^2 - 2bd + d^2$ is a square over $\mathbb{F}_{11}$, but it is not a square over $\mathbb{F}_{13}$.

Algorithm 1: Searching for an (iterative) infinitely long subspace trail without active S-boxes, using Theorem 1 and Theorem 2.

Data: P-SPN cipher over $\mathbb{F}^t$ with $s \in [1, t]$ S-boxes applied to the first s words, where the S-box has no linear structure.
Result: 1 if an (iterative) infinitely long subspace trail exists, 0 otherwise.

$i \leftarrow 0$.
do
  $i \leftarrow i + 1$.
  if $\exists \mu \in \mathbb{F}$ *s.t.* $M^i = \mu \cdot I$ *(where I is the identity matrix)* **then**
    return *1: Discard the matrix M*
  Let $\mathcal{S}^{(i)}$ be the subspace such that no S-box is active in the first i rounds, defined as in Definition 5.
  Let $\{\lambda_1^{(i)}, \lambda_2^{(i)}, \ldots, \lambda_\tau^{(i)}\}$ be all the (linearly independent) eigenvalues of M^i, and let $\{\mathcal{P}_1^{(i)}, \mathcal{P}_2^{(i)}, \ldots, \mathcal{P}_\tau^{(i)}\}$ be the corresponding eigenspaces.
  Note: $\dim(\mathcal{P}_1^{(i-1)}) + \cdots + \dim(\mathcal{P}_\tau^{(i-1)}) = t \implies \mathcal{P}_j^{(i-1)} = \mathcal{P}_j^{(i)}$, where $j \in [1, \tau]$.
  $\mathcal{IS}^{(i)} \leftarrow \left\langle \mathcal{P}_1^{(i)} \cap \mathcal{S}^{(i)}, \mathcal{P}_2^{(i)} \cap \mathcal{S}^{(i)}, \ldots, \mathcal{P}_\tau^{(i)} \cap \mathcal{S}^{(i)} \right\rangle$.
  if $\dim(\mathcal{IS}^{(i)}) \geq 1$ **and** $\mathcal{IS}^{(i)} \neq \mathbb{F}^t$ **then**
    return *1: Discard the matrix M*
while $\dim\left(\mathcal{S}^{(i)}\right) \geq 1$;
return *0: No (iterative) infinitely long subspace trail*

In the following, we first present an algorithm which can be used to find vulnerabilities and/or to discard possible "weak" matrices (w.r.t. the attacks presented before). Secondly, we test several matrices over $\mathbb{F}_p$ and over $\mathbb{F}_{2^n}$ in order to give an idea of the percentage of "weak" matrices.

4.1 Algorithm to Detect "Weak" Matrices

Algorithm 1 can be used to find vulnerabilities and/or to discard possible "weak" matrices (w.r.t. the attacks presented before). Here we present the details of this algorithm.

In order to determine a possible vulnerability, we first ensure that there is no integer $k \geq 1$ which leads to M^k being a multiple of I, where I is the identity matrix. After that, we apply the following steps.

1. We compute the eigenvalues and eigenspaces, denoted by $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_\tau$ (with a basis $\{\rho_1, \rho_2, \ldots, \rho_\tau\}$) of M, and store them.

2. We compute $\mathcal{S}^{(i)}$ such that there is no active S-box in the first i rounds (see Definition 5).

3. Finally, we look for an intersection between $\mathcal{S}^{(i)}$ and $\{\rho_1, \rho_2, \ldots, \rho_\tau\}$. If an intersection exists, M is a weak choice.

Computational Cost of Algorithm 1. The eigendecomposition of a $t \times t$ matrix needs a number of field operations in $\mathcal{O}(t^3)$, which is similar to the cost of solving a system of (linear) equations to find $\mathcal{S}^{(i)}$. Hence, the total runtime cost is in $\mathcal{O}\left(x \cdot t^3\right)$, where x is the number of repetitions of the Do-While loop.

Due to the Do-While condition, it is possible that the algorithm does not finish in a reasonable time. Indeed, note that the number of subspaces may be large, depending on the size of the field $\mathbb{F}^t$. Hence, it is possible that the iterated subspace trail has a large

period which cannot be detected by the algorithm in a reasonable time. In such a case, we suggest to stop the algorithm after a previously chosen number of iterations, and to discard the matrix. This basically corresponds to introducing an additional condition:

> *If there exists a subspace trail with no active S-boxes for more than r rounds, then discard the matrix (where $r \geq \left\lfloor \frac{t-s}{s} \right\rfloor$ is fixed in advance).*

This does not mean that there exists an infinitely long subspace trail for such a matrix. It just corresponds to the idea of using matrices for which we are sure (and we can prove) that the longest subspace trail with no active S-boxes covers at most r rounds (without exploiting the details of the S-box, weak keys, and so on).

Implementation. We make our implementation available online[12]. This tool can be used to detect vulnerabilities of given matrices over prime field or binary fields.

4.2 Percentage of "Weak" Linear Layers

We implemented Algorithm 1 in `Sage` and used it to give an idea of the percentage of matrices that are vulnerable to the attack without active S-boxes presented in Section 3.

Different Classes of Matrices. For concrete use cases, we set $s = 1$ and we focus on two scenarios, namely random invertible matrices and random Cauchy matrices[13]. As the source for randomness we use `Sage`'s random engine in both cases (and for choosing e.g. the prime numbers). In the first scenario, we create a matrix space, sample random matrices, and finally determine if they are invertible. In the second scenario, we generate Cauchy matrices using random (and valid) starting sequences. We tested all matrices using both prime fields and binary fields, focusing on square matrices of order $t \in \{3, 4, 8, 12, 16\}$ and on fields with a size of $n \in \{4, 8, 16\}$ (and $\lceil \log_2(p) \rceil \in \{4, 8, 16\}$ for prime fields). Moreover, we tested our algorithm on the concrete matrices used to instantiate Starkad and Poseidon. We present these results in Appendix D.

Concrete Results. The sample size for all tests was set to 10 000. While a matrix chosen completely at random (or without considering our results) may be vulnerable with a significant probability, it is easy to choose a matrix which is not vulnerable to the attacks presented above. Namely, given the estimated percentages of vulnerable matrices found in the tables above, the probability of finding a "secure" matrix (w.r.t. our results) is already quite high after trying two or more different matrices. In other words, our tool can easily be used to find matrices which are not vulnerable to the attacks presented in Section 3.

Regarding the tables, we can immediately see that the choice of p (or n) has an impact on the number of vulnerable matrices. Specifically, increasing $\lceil \log_2(p) \rceil$ (or n) tends to result in a higher probability for a matrix to be secure against the attacks presented here. We can also see that the cardinality of the field is more important in that regard than the number of cells – indeed, with respect to our observations, increasing t does not seem to have a major impact on the vulnerability.

In addition, we applied Algorithm 1 using different limits for its termination. In particular, we focused on limits of the form $k(t-1)$ for $0 < k \in \mathbb{N}$. We could not observe any major differences between choosing, for example, $k = 2$ or $k = 4$. As mentioned above, if the algorithm is not able to terminate within this fixed number of runs, we suggest to discard the given matrix.

[12] https://extgit.iaik.tugraz.at/krypto/linear-layer-tool

[13] We recall that $M \in \mathbb{F}^{t \times t}$ is a Cauchy matrix if there exists $\{x_i, y_i\}_{i=1}^{t}$ s.t. $M_{i,j} = \frac{1}{x_i + y_j}$, where for each $i \neq j : x_i \neq x_j$, $y_i \neq y_j$, $x_i + y_j \neq 0$. Cauchy matrices are MDS matrices.

Table 1: Percentage of vulnerable matrices for Algorithm 1 and orders t and field sizes $\lceil \log_2(p) \rceil$ when considering prime fields GF(p).

Generation	$\lceil \log_2(p) \rceil$	t	Vulnerable (%)
Random Invertible	4	3	8.74
	8	3	0.48
	4	4	7.30
	8	4	0.45
	8	8	0.57
	8	12	0.43
	8	16	0.55
	16	12	0.01
	16	16	< 0.01
MDS (Random Cauchy)	4	3	7.97
	8	3	0.63
	4	4	5.74
	8	4	0.61
	8	8	0.47
	8	12	0.51
	8	16	0.69
	16	12	< 0.01
	16	16	0.01

Computational Cost in Practice. There are various ways to implement Algorithm 1 in practice. Since we have to recalculate the subspace $\mathcal{S}^{(i)}$ in each run, we store the powers of the input matrix M beforehand, i.e., we compute and store $M, M^2, \ldots, M^l$, where l is the number of iterations. Hence, the memory cost depends on l and is then essentially in $\mathcal{O}\left(l \cdot t^2\right)$ for a $t \times t$ matrix M.[14]

The runtime is dominated by finding a solution to the system of equations and by building the eigendecomposition of a matrix. Both complexities are in $\mathcal{O}\left(t^3\right)$ for $t \times t$ matrices. In practice, we can also observe that running the tests for matrices with smaller t is significantly faster than running the tests for matrices with large t. As an example, and using $n = 16$, the test for a single matrix takes about 50 milliseconds for $t = 4$, while it takes about 1 second for $t = 12$. We used an Intel Xeon E5-2699v4 with a maximum clock frequency of 3.60 GHz for all our tests.

4.3 An Open Problem of Finding a Necessary Condition

Consider the case of a Cauchy matrix M generated as in [19] (recalled in Section 3.2) for $t = 24$ and $\mathbb{F}_{2^n}$, where $n = 63$. As shown in [26], the subspace $\mathcal{S}^{(5)}$ defined as in Definition 5 satisfies

$$M \cdot \mathcal{S}^{(5)} = \mathcal{S}^{(5)}$$

and

$$\forall x \in \mathcal{S}^{(5)} : \quad [M \cdot x]_1 = 0,$$

where $x \in \mathcal{S}^{(5)}$ and where $[v]_1$ denotes the first word (namely, the first n bits) of a vector v (this corresponds to the result already given in [26, Page 20]).

[14] Alternatively, the powers can be recomputed in each run, which increases the runtime cost, but decreases the memory cost to an element in $\mathcal{O}\left(t^2\right)$.

Table 2: Percentage of vulnerable matrices for Algorithm 1 and orders t and field sizes n when considering binary fields $\mathrm{GF}(2^n)$.

Generation	n	t	Vulnerable (%)
Random Invertible	4	3	5.92
	8	3	0.48
	4	4	6.26
	8	4	0.34
	8	8	0.35
	8	12	0.32
	8	16	0.35
	16	12	< 0.01
	16	16	< 0.01
MDS (Random Cauchy)	4	3	6.96
	8	3	0.33
	4	4	5.46
	8	4	0.28
	8	8	0.48
	8	12	0.46
	8	16	0.42
	16	12	< 0.01
	16	16	< 0.01

The reason why we highlight this fact is that it provides an example of a matrix for which our conditions given before are only sufficient but not necessary. In other words, if the previous condition (namely, Theorem 2) is both necessary and sufficient, then the subspace $\mathcal{S}^{(5)}$ must be related to the eigenvalues and eigenvectors of M. However, by simple practical tests, this is not the case since M^j for $j \in [1, 5]$ does not have any eigenvalues and eigenvectors. Hence, we can deduce that our observation provides a sufficient condition, but not a necessary one.

In more details, let d be the dimension of a (generic) invariant subspace $\mathcal{S} = \langle s_1, ..., s_d \rangle$ for a $t \times t$ matrix M. Such a subspace is related to the eigenvectors of M if there exist $\alpha_1, ..., \alpha_d, \mathfrak{A} \in \mathbb{F}$ (with $(\alpha_1, ..., \alpha_d) \neq (0, ..., 0)$ and $\mathfrak{A} \neq 0$) s.t.

$$M \times (\alpha_1 \cdot s_1 + \alpha_2 \cdot s_2 + ... + \alpha_d \cdot s_d) = \mathfrak{A} \cdot (\alpha_1 \cdot s_1 + \alpha_2 \cdot s_2 + ... + \alpha_d \cdot s_d),$$

namely if a system of t equations in $d + 1$ variables is satisfied. It follows that if e.g. $d + 1 < t$, then a solution could not exist[15]: this is exactly the case of the Cauchy matrix M generated as in [19].

A future open problem is to extend the condition to a necessary one. Instead of working over $\mathbb{F}$, one idea could be to work over its algebraic closure[16] $\mathbb{F}^\star$. Indeed, by definition[17], a field $\mathbb{F}$ is algebraically closed if and only if for each natural number n every linear map from $(\mathbb{F})^n$ into itself has some eigenvectors. Hence, instead of working over $\mathbb{F}_p$ or over $\mathbb{F}_{2^n}$, one possibility would be to work over their algebraic closures.

[15] Note that it is still possible that such solution exists: e.g., consider the case $d = 1$.

[16] A field $\mathbb{F}$ is *algebraically closed* if every non-constant polynomial in $\mathbb{F}[X]$ (the univariate polynomial ring with coefficients in $\mathbb{F}$) has a root in $\mathbb{F}$. For example, *no finite field* $\mathbb{F}$ *is algebraically closed*, because if $a_1, a_2, \ldots, a_n$ are all the elements of $\mathbb{F}$, then the polynomial $(x - a_1)(x - a_2)...(x - a_n) + 1$ has no zero in $\mathbb{F}$. By contrast, the field of complex numbers is algebraically closed.

[17] A linear map over a field $\mathbb{F}$ has an eigenvector if and only if its characteristic polynomial has some root. Therefore, when $\mathbb{F}$ is algebraically closed, every linear map of $\mathbb{F}^n$ has some eigenvector.

5 Subspace Trails for P-SPN Ciphers with Active S-Boxes

Until now, we focused on the case in which no S-box is active. Here, we analyze the scenario in which S-boxes are active. We start by presenting a generic result regarding the minimum number of rounds for which it is possible to set up a subspace trail with a probability of 1. Then, for the first time in the literature (to the best of our knowledge), we analyze infinitely long invariant subspace trails in the case of active S-boxes.

5.1 Subspace Trails and Truncated Differentials

Proposition 4. *Given a partial SPN cipher over $\mathbb{F}^t$ with $s \leq t$ S-boxes, it is always possible to set up a subspace trail with probability 1 on at least $2 \cdot \left\lfloor \frac{t-s}{s} \right\rfloor$ rounds, defined by*

$$\left\{ \underbrace{\mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, M \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, \ldots, M^{\lfloor \frac{t-s}{s} \rfloor - 1} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}}_{\text{no active S-boxes}}, \mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \ldots, \mathcal{A}^{(\lfloor \frac{t-s}{s} \rfloor)} \right\}, \tag{4}$$

where $\mathcal{S}^{(\cdot)}$ is defined as in Definition 5, where

$$\forall i \geq 1 : \quad \mathcal{A}^{(i)} := \left\langle M(e_1), M(e_2), \ldots, M(e_s), M \cdot \mathcal{A}^{(i-1)} \right\rangle,$$

and where $\mathcal{A}^{(0)} := M^{\lfloor \frac{t-s}{s} \rfloor - 1} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}$.

As for Proposition 1, this well-known result only depends on the number of S-boxes, and no assumption on the matrix M is made. Moreover, it also includes the case of SPN ciphers (if $t = s$, the subspace trail can be set up for at least one round). As done before and w.l.o.g., in the following we omit the round key and constant additions (since they only change the coset and we deal with differences).

Proof. The subspace trail defined over the first $\left\lfloor \frac{t-s}{s} \right\rfloor$ rounds is already analyzed in Section 3.1. Such a subspace trail cannot be extended for more rounds without activating any S-box since

$$M^{\lfloor \frac{t-s}{s} \rfloor - 1} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)} \not\subseteq \langle e_{s+1}, e_{s+2}, \ldots, e_t \rangle .$$

Hence, at least one S-box would be active after $\left\lfloor \frac{t-s}{s} \right\rfloor$ rounds. It follows that the only way to extend the trail is by increasing the dimension of such a subspace, that is,

$$R\left(M^{\lfloor \frac{t-s}{s} \rfloor} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}\right) \subseteq \mathcal{A}^{(1)} = \left\langle M^{\lfloor \frac{t-s}{s} \rfloor + 1} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, M(e_1), M(e_2), \ldots, M(e_s) \right\rangle .$$

Indeed, the only thing one can do is to consider the biggest subspace for which

$$\text{S-box}\left(M^{(\lfloor \frac{t-s}{s} \rfloor)} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}\right) \subseteq \left\langle \underbrace{e_1, e_2, \ldots, e_s}_{\text{Due to S-boxes}}, \underbrace{M^{\lfloor \frac{t-s}{s} \rfloor} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}}_{\text{Due to identity part}} \right\rangle .$$

In this way, we lose information about the output of the S-box layer (while nothing changes for the part of the identity layer), but we can extend the subspace trail. Working in the same way, it follows that

$$R\left(\mathcal{A}^{(1)}\right) \subseteq \mathcal{A}^{(2)} = \left\langle M \cdot \mathcal{A}^{(1)}, M(e_1), M(e_2), \ldots, M(e_s) \right\rangle$$

and, more generally,

$$R\left(\mathcal{A}^{(r)}\right) \subseteq \mathcal{A}^{(r+1)} = \left\langle M \cdot \mathcal{A}^{(r)}, M(e_1), \ldots, M(e_s) \right\rangle .$$

This operation can be repeated until the dimension of the subspace is smaller than t. Since for a generic cipher the dimension of $\mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}$ is s and the dimension increases by s in each additional round, the dimension remains smaller than t for up to $2 \cdot \left\lfloor \frac{t-s}{s} \right\rfloor$ rounds. □

Similar to the case presented in Section 3.1, depending on the details of the linear layer, a longer subspace trail of dimension 1 can be set up. This happens e.g. for the case of AES, due to the fact that $M = MC \circ SR(\cdot)$ is "sparse" and does not provide full diffusion at word level after a single round (we refer to [22] for more details). In such a case, a 2-round subspace trail with probability 1 can be set up.

Truncated Differentials. Due to the relation between subspace trails and truncated differentials [30], this implies the possibility to set up a truncated differential distinguisher on at least $2 \cdot \left\lfloor \frac{t-s}{s} \right\rfloor$ rounds with probability 1. Moreover, truncated differentials [27] which hold with a probability smaller than 1 and impossible differentials can potentially be set up for more rounds. However, since this is out of the scope of this paper, we refer to Appendix C for more details.

5.2 Infinitely Long Subspace Trail with Active S-Boxes

Working as in Section 3, we now study which properties a linear layer must satisfy in order to set up an infinitely long subspace trail also in the case of active S-boxes. To the best of our knowledge, this is the first time that such an approach is described in the literature.

5.2.1 Invariant Subspace Trail with Active S-Boxes

Using the approach proposed in Section 3.3, we first focus on the case of invariant subspace trails with active S-boxes.

Theorem 3. *Given a P-SPN cipher with s S-boxes, let $M \in \mathbb{F}^{t\times t}$ be an invertible matrix. Let $\lambda_1, \lambda_2, \dots, \lambda_\tau$ be the eigenvalues of M, and let $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_\tau$ be the corresponding eigenspaces (where $\tau \leq t$). Let $I = \{i_1, i_2, \dots, i_{|I|}\} \subseteq \{1, 2, \dots, s\}$ be the indices of the words with active S-boxes, and let $J = \{j_1, j_2, \dots, j_{|J|}\} \subseteq \{1, 2, \dots, \tau\}$ such that*

$$\mathcal{IS} = \left\langle \mathcal{P}'_{j_1}, P'_{j_2}, \dots, \mathcal{P}'_{j_{|J|}} \right\rangle,$$

where $\mathcal{P}' \subseteq \mathcal{P}$ is a non-null subspace. If $1 \leq \dim(\mathcal{IS}) < t$ and if $\mathcal{IS}$ satisfies

(1) $\mathcal{IS} \cap \langle e_{i_1}, e_{i_2}, \dots, e_{i_{|I|}}, e_{s+1}, e_{s+2}, \dots, e_t \rangle = \mathcal{IS}$, and

(2) $\forall i \in I \subseteq \{1, 2, \dots, s\}: \quad \mathcal{IS} \cap \langle e_i \rangle = \langle e_i \rangle$,

then $\mathcal{IS}$ generates an infinitely long invariant subspace trail (with active S-boxes in the case in which $I \neq \emptyset$).

Note that if $|I| = 0$, then this reduces to the previous result.

Proof. The first condition ensures that no l-th word is active, where $l \notin I$. For each i-th active word, where $i \in I$, the second condition implies that the entire space $\langle e_i \rangle$ is included in $\mathcal{IS}$. The consequence is that, when applying the S-box, the subspace remains the same.

As for the results given in the previous sections, since such a subspace is defined via the eigenspaces of M, it remains invariant under the linear layer. Hence, $\mathcal{IS}$ results in an infinitely long invariant subspace trail.

□

Example. Given a P-SPN cipher with $s = 1$, consider the following 4×4 matrix M defined over $\mathbb{F}$:

$$M = \begin{pmatrix} 0 & (1 - M_{02} \cdot v_2 - M_{03} \cdot v_3)/v_1 & M_{02} & M_{03} \\ v_1 & (-M_{12} \cdot v_2 - M_{13} \cdot v_3)/v_1 & M_{12} & M_{13} \\ v_2 & (-M_{22} \cdot v_2 - M_{23} \cdot v_3)/v_1 & M_{22} & M_{23} \\ v_3 & (-M_{32} \cdot v_2 - M_{33} \cdot v_3)/v_1 & M_{32} & M_{33} \end{pmatrix}, \tag{5}$$

where $v_1 \neq 0$. A proper choice of v_1, v_2, v_3 and $M_{\cdot,\cdot}$ provides invertibility and "full diffusion" (at word level after a finite number of rounds) for cryptographic purposes.

The subspace

$$\mathcal{S} = \left\langle e_1 = (1,0,0,0)^T, v = (0, v_1, v_2, v_3)^T \right\rangle$$

is invariant under the round transformation for any number of rounds, since

(1) the first word can take every value and because the S-box is applied only to this word (the S-box is a permutation), $\mathcal{S}$ remains invariant (note that the S-box is active), and

(2) the vectors satisfy $M \cdot e_1 = v$ and $M \cdot v = e_1$.

It follows that this is a concrete example of an infinitely long invariant subspace trails with active S-boxes. As before, v and e_1 are related to the eigenvectors of M, in particular

$$M \cdot (v + e_1) = v + e_1 \qquad M \cdot (v - e_1) = -(v - e_1),$$

where $\mathcal{P}_1 = \langle v + e_1 \rangle$ and $\mathcal{P}_2 = \langle v - e_1 \rangle$ satisfy the conditions given in the previous theorem.

As a last thing, we remark that matrices of the form Eq. (5) are currently used in the literature: For example, the almost-MDS matrix over $\mathbb{F}_{2^n}$ defined as

$$\mathrm{circ}^R(0,1,1,1) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

is used in Midori [4] and QARMA [3].

5.2.2 Computational Cost

While in the case of passive S-boxes we are able to construct the invariant subspace directly, the previous definition is too computationally expensive to exploit in order to construct the invariant subspace trail (with active S-boxes) in the case of large t and/or large $|\mathbb{F}|$. For example, in the case of s S-boxes and in the case of a matrix with t different eigenspaces $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_t$ (each one of dimension 1), there are

$$(2^s - 1) \cdot \left(\sum_{i=1}^{t} \binom{t}{i} \right) = (2^t - 1) \cdot (2^s - 1) \in \mathcal{O}(2^{s+t})$$

different cases the attacker has to check in order to construct such a subspace trail. The situation becomes even worse in the case in which some eigenspaces have dimension greater than 1, due to the fact that the number of possible subspaces grows exponentially, as given in the following proposition.

Proposition 5 ([24]). *Let $0 \leq k \leq n$. Given an n-dimensional vector space $\mathcal{V}$ over $\mathbb{F}_q$, there exist*

$$\frac{(q^n - 1)(q^n - q)(q^n - q^2) \cdots (q^n - q^{k-1})}{(q^k - 1)(q^k - q)(q^k - q^2) \cdots (q^k - q^{k-1})} \in \mathcal{O}\left(q^{k(n-k)}\right)$$

different subspaces of dimension k.

Hence, apart from some special cases, it seems infeasible to determine if an infinitely long subspace trail with active S-boxes can be constructed or not. We leave this as an open problem for future research.

5.2.3 Iterated Subspace Trail with Active S-Boxes

Finally, we mention that the previous result can be generalized by considering iterated (non-invariant) subspace trails with active S-boxes. To do this, the idea is again to consider the eigenspaces of M^k for $k \geq 2$.

Theorem 4. *Given a P-SPN cipher with s S-boxes, let $M \in \mathbb{F}^{t \times t}$ be an invertible matrix. Let $\lambda_1^{(\mathfrak{k})}, \lambda_2^{(\mathfrak{k})}, \dots, \lambda_\tau^{(\mathfrak{k})}$ be the eigenvalues of $M^{\mathfrak{k}}$ for a certain $\mathfrak{k} \geq 1$, and let $\mathcal{P}_1^{(\mathfrak{k})}, \mathcal{P}_2^{(\mathfrak{k})}, \dots, \mathcal{P}_\tau^{(\mathfrak{k})}$ be the corresponding eigenspaces (where $\tau \leq t$). Further, let $\mathfrak{s} \leq s$ be the number of active S-boxes, and let $J = \{j_1, j_2, \dots, j_{|J|}\} \subseteq \{1, 2, \dots, \tau\}$ such that*

$$\mathcal{IS} = \left\langle \mathcal{P}'_{j_1}, \mathcal{P}'_{j_2}, \dots, \mathcal{P}'_{j_{|J|}} \right\rangle,$$

where $\mathcal{P}' \subseteq \mathcal{P}$ is a non-null subspace.

For each $0 \leq j \leq \mathfrak{k} - 1$, let $I_j = \{i_{1,j}, i_{2,j}, \dots, i_{\mathfrak{s},j}\} \subseteq \{1, 2, \dots, s\}$ be the indices of the words with active S-boxes in the x-th round, where $x \mod (\mathfrak{k} - 1) = j$. If $1 \leq \dim(\mathcal{IS}) < t$ and if $\mathcal{IS}$ satisfies

(1) $\forall j \in \{0, 1, \dots, \mathfrak{k} - 1\}: \quad (M^j \cdot \mathcal{IS}) \cap \left\langle e_{i_{1,j}}, e_{i_{2,j}}, \dots, e_{i_{\mathfrak{s},j}}, e_{s+1}, e_{s+2}, \dots, e_t \right\rangle = \mathcal{IS},$

(2) $\forall j \in \{0, 1, \dots, \mathfrak{k} - 1\}$ *and* $\forall i \in I_j \subseteq \{1, 2, \dots, s\}: \quad (M^j \cdot \mathcal{IS}) \cap \langle e_i \rangle = \langle e_i \rangle,$

then an infinitely long iterated subspace (with active S-boxes in the case in which $I \neq \emptyset$) of the form

$$\left\{\mathcal{IS}, M \cdot \mathcal{IS}, M^2 \cdot \mathcal{IS}, \dots, M^{\mathfrak{k}-1} \cdot \mathcal{IS}\right\}$$

is generated.

Note that the active S-boxes do not need to be in fixed position, and it is sufficient to impose $I_j = I_l$ for each $j, l \leq \mathfrak{k} - 1$.

6 Open Problems

In this paper, we presented sufficient conditions that a (highly non-trivial) linear layer must satisfy in order to set up infinitely long subspace trail attacks. As already mentioned in the paper, several problems are still open for future research:

- Is any of the conditions given in this paper both necessary and sufficient? If not, is it possible to find a similar condition which is both necessary and sufficient? For example, what happens if one considers the eigenvalues or eigenspaces of M over the algebraic closure of $\mathbb{F}$?
- In the whole paper, we limit ourselves to work independently of the details of the S-box, since we assume that it is not possible to set up any non-trivial subspace trail for the S-box. However, this is not always the case (e.g., consider the examples given in [30] for PRESENT). As a future open problem, one could extend the result given in this paper in order to take the details of the S-box into account as well.
- Following the previous point, it could make sense to analyze how the key schedule and the existence of weak keys influence the possibility to set up a weak-key infinitely long subspace trail. What is a possible countermeasure that allows to prevent this case? Is the analysis provided in [6] valid also in the case of P-SPN ciphers?
- In the case of active S-boxes, a direct construction of the infinitely long subspace trail (given the details of the matrix that defines the linear layer) is missing. This may be crucial in order to solve the problem regarding the computational cost of constructing it with the current definition. Is it possible to conclude anything, at least in the simplest case in which $s = 1$?

- Given $t, s, \mathbb{F}$, is it possible to give the specification of a regular matrix for which one can easily prove that no infinitely long subspace trail (both with active and inactive S-boxes) exists? Is it possible to conclude anything at least in the case of MDS or near-MDS matrices?

Acknowledgements. The authors thank Reinhard Lüftenegger and Willi Meier for their valuable comments.

References

[1] Abdelraheem, M.A., Ågren, M., Beelen, P., Leander, G.: On the Distribution of Linear Biases: Three Instructive Examples. In: Advances in Cryptology - CRYPTO 2012. LNCS, vol. 7417, pp. 50–67. Springer (2012)

[2] Albrecht, M.R., Rechberger, C., Schneider, T., Tiessen, T., Zohner, M.: Ciphers for MPC and FHE. In: EUROCRYPT 2015. LNCS, vol. 9056, pp. 430–454 (2015)

[3] Avanzi, R.: The QARMA Block Cipher Family. Almost MDS Matrices Over Rings With Zero Divisors, Nearly Symmetric Even-Mansour Constructions With Non-Involutory Central Rounds, and Search Heuristics for Low-Latency S-Boxes. IACR Trans. Symmetric Cryptol. **2017**(1), 4–44 (2017)

[4] Banik, S., Bogdanov, A., Isobe, T., Shibutani, K., Hiwatari, H., Akishita, T., Regazzoni, F.: Midori: A Block Cipher for Low Energy. In: Advances in Cryptology – ASIACRYPT 2015. LNCS, vol. 9453, pp. 411–436 (2015)

[5] Bar-On, A., Dinur, I., Dunkelman, O., Lallemand, V., Keller, N., Tsaban, B.: Cryptanalysis of SP Networks with Partial Non-Linear Layers. In: EUROCRYPT 2015. LNCS, vol. 9056, pp. 315–342 (2015)

[6] Beierle, C., Canteaut, A., Leander, G., Rotella, Y.: Proving Resistance Against Invariant Attacks: How to Choose the Round Constants. In: Advances in Cryptology – CRYPTO 2017. LNCS, vol. 10402, pp. 647–678 (2017)

[7] Beyne, T.: Block Cipher Invariants as Eigenvectors of Correlation Matrices. In: Advances in Cryptology - ASIACRYPT 2018. LNCS, vol. 11272, pp. 3–31 (2018)

[8] Beyne, T., Canteaut, A., Dinur, I., Eichlseder, M., Leander, G., Leurent, G., Naya-Plasencia, M., Perrin, L., Sasaki, Y., Todo, Y., Wiemer, F.: Out of oddity – new cryptanalytic techniques against symmetric primitives optimized for integrity proof systems. Cryptology ePrint Archive, Report 2020/188 (2020), https://eprint.iacr.org/2020/188

[9] Biham, E., Biryukov, A., Shamir, A.: Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. In: EUROCRYPT 1999. LNCS, vol. 1592, pp. 12–23 (1999)

[10] Biham, E., Shamir, A.: Differential Cryptanalysis of DES-like Cryptosystems. Journal of Cryptology **4**(1), 3–72 (1991)

[11] Biham, E., Shamir, A.: Differential Cryptanalysis of the Data Encryption Standard. Springer (1993)

[12] Blondeau, C., Leander, G., Nyberg, K.: Differential-Linear Cryptanalysis Revisited. Journal of Cryptology **30**(3), 859–888 (2017)

[13] Daemen, J., Govaerts, R., Vandewalle, J.: Correlation Matrices. In: Fast Software Encryption 1994 – FSE'94. LNCS, vol. 1008, pp. 275–285 (1994)

[14] Daemen, J., Rijmen, V.: AES and the Wide Trail Design Strategy. In: EUROCRYPT 2002. LNCS, vol. 2332, pp. 108–109 (2002)

[15] Daemen, J., Rijmen, V.: The Design of Rijndael: AES - The Advanced Encryption Standard. Information Security and Cryptography, Springer (2002)

[16] Dinur, I., Kales, D., Promitzer, A., Ramacher, S., Rechberger, C.: Linear equivalence of block ciphers with partial non-linear layers: Application to lowmc. In: Ishai, Y., Rijmen, V. (eds.) EUROCRYPT 2019. Lecture Notes in Computer Science, vol. 11476, pp. 343–372. Springer (2019)

[17] Dinur, I., Liu, Y., Meier, W., Wang, Q.: Optimized Interpolation Attacks on LowMC. In: ASIACRYPT 2015. LNCS, vol. 9453, pp. 535–560 (2015)

[18] Gérard, B., Grosso, V., Naya-Plasencia, M., Standaert, F.X.: Block Ciphers That Are Easier to Mask: How Far Can We Go? In: CHES 2013. LNCS, vol. 8086, pp. 383–399 (2013)

[19] Grassi, L., Kales, D., Khovratovich, D., Roy, A., Rechberger, C., Schofnegger, M.: Starkad and Poseidon: New Hash Functions for Zero Knowledge Proof Systems. Cryptology ePrint Archive, Report 2019/458 (2019)

[20] Grassi, L., Leander, G., Rechberger, C., Tezcan, C., Wiemer, F.: Weak-Key Subspace Trails and Applications to AES. Cryptology ePrint Archive, Report 2019/852 (2019), https://eprint.iacr.org/2019/852

[21] Grassi, L., Lüftenegger, R., Rechberger, C., Rotaru, D., Schofnegger, M.: On a Generalization of Substitution-Permutation Networks: The HADES Design Strategy. Cryptology ePrint Archive, Report 2019/1107 (2019), https://eprint.iacr.org/2019/1107, accepted at Eurocrypt'20

[22] Grassi, L., Rechberger, C., Rønjom, S.: Subspace Trail Cryptanalysis and its Applications to AES. IACR Trans. Symmetric Cryptol. **2016**(2), 192–225 (2016)

[23] Grassi, L., Rechberger, C., Rønjom, S.: A new structural-differential property of 5-round AES. In: Coron, J., Nielsen, J.B. (eds.) EUROCRYPT 2017. LNCS, vol. 10211, pp. 289–317 (2017)

[24] Huppert, B., Willems, W.: Lineare Algebra (2nd edition). Undergraduate texts in mathematics, Vieweg+Teubner, Wiesbaden (2010)

[25] Kales, D., Perrin, L., Promitzer, A., Ramacher, S., Rechberger, C.: Improvements to the Linear Layer of LowMC: A Faster Picnic. Cryptology ePrint Archive, Report 2017/1148 (2017)

[26] Keller, N., Rosemarin, A.: Mind the middle layer: The hades design strategy revisited. Cryptology ePrint Archive, Report 2020/179 (2020), https://eprint.iacr.org/2020/179

[27] Knudsen, L.R.: Truncated and Higher Order Differentials. In: FSE 1994. LNCS, vol. 1008, pp. 196–211 (1994)

[28] Leander, G., Abdelraheem, M.A., AlKhzaimi, H., Zenner, E.: A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack. In: Advances in Cryptology - CRYPTO 2011. LNCS, vol. 6841, pp. 206–221 (2011)

[29] Leander, G., Minaud, B., Rønjom, S.: A Generic Approach to Invariant Subspace Attacks: Cryptanalysis of Robin, iSCREAM and Zorro. In: EUROCRYPT 2015. LNCS, vol. 9056, pp. 254–283 (2015)

[30] Leander, G., Tezcan, C., Wiemer, F.: Searching for Subspace Trails and Truncated Differentials. IACR Trans. Symmetric Cryptol. **2018**(1), 74–100 (2018)

[31] Matsui, M.: Linear Cryptanalysis Method for DES Cipher. In: EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397 (1993)

[32] Rijmen, V., Daemen, J., Preneel, B., Bosselaers, A., Win, E.D.: The Cipher SHARK. In: FSE 1996. LNCS, vol. 1039, pp. 99–111 (1996)

[33] Todo, Y., Leander, G., Sasaki, Y.: Nonlinear Invariant Attack - Practical Attack on Full SCREAM, iSCREAM, and Midori64. In: Advances in Cryptology - ASIACRYPT 2016. LNCS, vol. 10032, pp. 3–33 (2016)

[34] Wang, Y., Wu, W., Guo, Z., Yu, X.: Differential Cryptanalysis and Linear Distinguisher of Full-Round Zorro. In: ACNS 2014. LNCS, vol. 8479, pp. 308–323 (2014)

A Related Works

In order to discuss the results in [1] and [7], and the relation between them and the ones presented in this paper, we first briefly recall the definition of *correlation matrices* [13].

Definition 7. Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be a vectorial Boolean function. The correlation matrix $C^F \in \mathbb{R}^{2^m \times 2^n}$ of F is the representation of the transition matrix of F with respect to the character basis of the algebra $\mathbb{C}[F_2^n]$ and $\mathbb{C}[F_2^m]$. The coordinates of C^F are

$$C^F_{u,v} = \frac{1}{2^n} \cdot \sum_{x \in \mathbb{F}_2^n} (-1)^{u^T \cdot F(x) + v^T \cdot x}.$$

Using these notions, we recall the results presented in the literature.

Proposition 6 (Theorem 5 of [1])**.** *Consider an invertible vectorial Boolean function F, a subspace U, the orthogonal subspace $U^\perp$, and a vector d. Let $C^F_{u,v}$ be the correlation matrix of F, and let $\omega = (\omega_u)_{u \in U^\perp}$, where $\omega_u = (-1)^{d^T \cdot u}$. Then $C^F \cdot \omega^T = \omega^T$ if and only if $F(U + d) = U + d$.*

This result has been generalized by Beyne in [7], who defines a "block cipher invariant" in the following way.

Definition 8 (Definition 2 of [7])**.** A vector $v \in \mathbb{C}^{2^n}$ is an invariant for a block cipher $E_k : \mathbb{F}_2^n \to \mathbb{F}_2^n$ if and only if it is an eigenvector of the correlation matrix C^{E_k}. If v is a multiple of $(1, 0, \ldots, 0)^T$, it will be called a trivial invariant.

For the case of invariant subspace trails, the same approach – *opportunely modified* – can potentially be exploited in order to find the results proposed here. In particular, using the properties of C^F just presented, it follows that in the case of a round function $R_k(\cdot) = k \oplus R(\cdot) = k + \mathfrak{M} \circ \mathfrak{S}(\cdot)$, where $\mathfrak{S}(\cdot) \equiv [S(\cdot) \,||\, \ldots \,||\, S(\cdot) \,||\, I(\cdot) \,||\, \ldots \,||\, I(\cdot)]$ and where $\mathfrak{M}(\cdot) = M \cdot (\cdot)$, it holds that

$$C^{R_k} = C^k C^R = C^k C^{\mathfrak{M}} \cdot C^{\mathfrak{S}} = C^k [C^M] \big([C^S]^{\otimes s} \otimes [C^I]^{\otimes (t-s)} \big),$$

where $C^M_{u,v} = \delta(u + M^T \cdot v)$, $C^I_{u,v} = \delta(u + v)$, and where C^k is a diagonal matrix. In the case studied here, it is not hard to see that if no S-box is active, the eigenvalues and eigenvectors of $C^M_{u,v}$ are strictly related to the eigenvalues and eigenvectors of M, hence the previous result.

At the same time, here we point out the following observations.

1. Both the results [1] and [7] focus on invariant subspaces only. As a consequence, one has to take care of the effect of the key (namely, of C^k) on the eigenvectors of C^R (namely, of the part of the round that is independent of the key).

2. In our case, we look for infinitely long iterative subspace trails of P-SPN ciphers which are independent of the secret key and of the key schedule. Again, this is not possible for an SPN cipher due to the full nonlinear layer.

3. We do not require that the subspace is invariant (namely, we do not restrict ourselves to the case $R(U + v) = U + w$). At the same time, an r-round iterated subspace trail can be seen as an invariant subspace trail for r rounds of the cipher. Hence, the previous result can be adapted in order to include this case.

B 2-Round Iterative Subspace Trail – Details

In this section, we present all the details of the concrete example of an iterated subspace trail that is not invariant given in Section 3.4.

The starting point is given by the circulant matrix $M = \mathrm{circ}(a, b, c, d)$ with elements $a, b, c, d \in \mathbb{F}_p$, which is invertible if and only if its determinant is different from zero:

$$-a^4 + b^4 - 4ab^2c + 2a^2c^2 - c^4 + 4a^2bd + 4bc^2d - 2b^2d^2 - 4acd^2 + d^4 \neq 0 \mod p.$$

Depending on a, b, c, d, such a matrix can have either 2 or 4 eigenvalues and eigenvectors, while M^2 has always 4 eigenvalues and eigenvectors. In particular, the eigenvalues and eigenvectors of M are given by

$$\begin{aligned}
\lambda_0 = a + b + c + d: &\quad (1, 1, 1, 1)^T, \\
\lambda_1 = -\sqrt{a^2 + b^2 - 2ac + c^2 - 2bd + d^2}: &\quad (b - d, -a + c + \lambda_1, d - b, a - c - \lambda_1)^T, \\
\lambda_2 = \sqrt{a^2 + b^2 - 2ac + c^2 - 2bd + d^2}: &\quad (b - d, -a + c + \lambda_2, d - b, a - c - \lambda_2)^T, \\
\lambda_3 = a - b + c - d: &\quad (1, -1, 1, -1)^T,
\end{aligned}$$

while the eigenvalues and eigenvectors of M^2 are given by

$$\begin{aligned}
\Lambda_0 = (\lambda_0)^2 &= a^2 + 2a(b + c + d) + b^2 + 2b(c + d) + c^2 + 2cd + d^2: \quad (1, 1, 1, 1)^T, \\
\Lambda_1 = (\lambda_1)^2 &= a^2 + b^2 - 2ac + c^2 - 2bd + d^2: \quad (1, 0, -1, 0)^T, \\
\Lambda_2 = (\lambda_2)^2 &= a^2 + b^2 - 2ac + c^2 - 2bd + d^2: \quad (0, 1, 0, -1)^T, \\
\Lambda_3 = (\lambda_3)^2 &= a^2 - 2a(b - c + d) + b^2 - 2b(c - d) + c^2 - 2cd + d^2: \quad (1, -1, 1, -1)^T.
\end{aligned}$$

Let $\mathfrak{M}_{t\times t} \in \mathbb{F}^{t\times t}$ be the matrix defined as

$$\mathfrak{M}_{5\times 5} = \begin{pmatrix} x & y_0 & y_1 & y_0 & y_1 \\ z_0 & a & b & c & d \\ z_1 & b & c & d & a \\ z_2 & c & d & a & b \\ z_3 & d & a & b & c \end{pmatrix},$$

$$\mathfrak{M}_{6\times 6} = \begin{pmatrix} x_0 & x_1 & y_0 & y_1 & y_0 & y_1 \\ x_2 & x_3 & y_2 & y_3 & y_2 & y_3 \\ z_0 & z_4 & a & b & c & d \\ z_1 & z_5 & b & c & d & a \\ z_2 & z_6 & c & d & a & b \\ z_3 & z_7 & d & a & b & c \end{pmatrix},$$

and so on, where

(1) the coefficients are chosen in order to provide invertibility and "full diffusion" (at word level after a finite number of rounds) for cryptographic purposes, and

(2) a, b, c, d are chosen such that the corresponding matrix has only 2 eigenvalues, namely

$$\forall x \in \mathbb{F}_p: \qquad a^2 + b^2 - 2 \cdot a \cdot c + c^2 - 2 \cdot b \cdot d + d^2 \neq x^2 \mod p,$$

that is,

$$a \neq c \qquad \text{and} \qquad b \neq d$$

(remember that $x \mapsto x^2$ is not a permutation over $\mathbb{F}_p$ for a prime $p \geq 3$ – see Hermite's criterion).

Note that

$$\textit{(1)} \qquad \begin{pmatrix} a & b & c & d \\ b & c & d & a \\ c & d & a & b \\ d & a & b & c \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} b-d \\ c-a \\ -(b-d) \\ -(c-a) \end{pmatrix},$$

$$\textit{(2)} \qquad \begin{pmatrix} a & b & c & d \\ b & c & d & a \\ c & d & a & b \\ d & a & b & c \end{pmatrix}^2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = (a^2+b^2-2ac+c^2-2bd+d^2) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \text{ and}$$

$$\textit{(3)} \qquad \begin{pmatrix} x & y & x & y \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix}.$$

Working on $\mathbb{F}^5$, and due to these considerations, the subspace $\mathcal{S}$ defined by

$$\mathcal{S} = \left\langle (0,0,1,0,-1)^T \right\rangle$$

is 2-round iterative subspace trail, since

(1) $\mathfrak{M} \cdot \mathcal{S} = \left\langle (0, b-d, c-a, d-b, a-c)^T \right\rangle$, and

(2) $\mathfrak{M}^2 \cdot \mathcal{S} = \mathcal{S}$.

Finally, note that $\mathfrak{M}^2$ is not necessarily equal to a multiple of the identity. A concrete example is given by $(\mathfrak{M}^2_{5\times 5})_{0,4} \neq 0$, where[18]

$$\left(\mathfrak{M}^2_{5\times 5}\right)_{0,4} = xy_0 + y_0 a + y_1 b + y_0 c + y_1 d,$$

which is different from 0 by appropriately choosing the entries.

Other Examples. Note that many other examples can be constructed in a similar way. For example, the matrix $\mathfrak{M}_{8\times 8}$ defined by

$$\mathfrak{M}_{8\times 8} = \begin{pmatrix} \mathrm{circ}(s,z,s,z) & \mathrm{circ}(a,b,c,d) \\ \mathrm{circ}(a,b,c,d) & \mathrm{circ}(u,v,u,v) \end{pmatrix},$$

where a, b, c, d are chosen such that the corresponding circulant matrix has only 2 eigenvalues, admits a 2-round iterative subspace trail defined by

$$\mathcal{S} = \left\langle (0,1,0,-1,0,0,0,0)^T \right\rangle.$$

Indeed,

$$\mathfrak{M}_{8\times 8} \cdot \left\langle (0,1,0,-1,0,0,0,0)^T \right\rangle = \left\langle (0,0,0,0,b-d,c-a,d-b,a-c)^T \right\rangle$$

and

$$(\mathfrak{M}_{8\times 8})^2 \cdot \left\langle (0,1,0,-1,0,0,0,0)^T \right\rangle = \left\langle (0,1,0,-1,0,0,0,0)^T \right\rangle.$$

[18] $M_{x,y}$ denotes the entry of M at row x and column y.

C Truncated and Impossible Differentials

So far, we discussed the possibility to set up truncated differentials with probability 1. However, this does not guarantee security against all other generalizations, precisely truncated differentials with probability smaller than 1 and impossible differentials. Here we briefly focus on this case. However, we point out that we do not discuss the minimum number of rounds necessary to guarantee security against these attacks, since they strongly depend on the details of the linear layer.

Differential attacks [11] exploit the fact that pairs of plaintexts with certain differences yield other differences in the corresponding ciphertexts with a non-uniform probability distribution. A variant of this attack/distinguisher is the truncated differential one [27], in which the attacker can predict only part of the difference between pairs of texts. Using the subspace terminology, given pairs of plaintexts that belong to the same coset of a subspace $\mathcal{X}$, one considers the probability that the corresponding ciphertexts belong to the same coset of a subspace $\mathcal{Y}$ to set up an attack – see e.g. [12] for details. (Truncated) impossible differential distinguishers/attacks [9] exploit differentials that holds with probability 0.

As we are going to show, in the case in which the details of the S-box are not taken into account, then (the "basic" variants of) truncated and/or of impossible differential distinguishers – which are independent of the secret key – can be set up for (at most) $2R$ rounds, where $R \geq 2\left\lfloor \frac{t-s}{s} \right\rfloor$ is the maximum number of rounds for which it is possible to set up a truncated differential with probability 1.

Remark. We stress that the details of the construction (e.g., the S-box, the linear layer, the key schedule) can potentially be used to improve the previous attacks. That is, $2R$ rounds refer only to the "basic" variants of such attacks, and this number must be considered only as lower bound in order to guarantee security.

C.1 Truncated Differentials with Probability < 1

Here we exploit the relation between truncated differentials and subspace trails [22, 30], and the results just given, in order to analyze the minimum number of rounds to prevent these attacks. We recall following proposition from [22].

Proposition 7. *Let* $\left\{\mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, M \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, \ldots, M^{\lfloor \frac{t-s}{s} \rfloor - 1} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, \mathcal{A}^{(1)}, \ldots, \mathcal{A}^{(\lfloor \frac{t-s}{s} \rfloor)}\right\}$ *be a subspace trail of prob. 1 defined as in Eq. (4). For simplicity, let* $\mathfrak{r} = 2 \cdot \lfloor (t-s)/s \rfloor$ *and let*

$$\{V^0, V^1, \ldots, V^{\lfloor (t-s)/s \rfloor - 1}, V^{\lfloor (t-s)/s \rfloor}, \ldots, V^{2 \cdot \lfloor (t-s)/s \rfloor - 2}\} :=$$
$$:= \left\{\mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, M \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, \ldots, M^{\lfloor \frac{t-s}{s} \rfloor - 1} \cdot \mathcal{S}^{(\lfloor \frac{t-s}{s} \rfloor)}, \mathcal{A}^{(1)}, \ldots, \mathcal{A}^{(\lfloor \frac{t-s}{s} \rfloor)}\right\}.$$

If there exist $0 \leq v < u \leq w < \mathfrak{r}$ *s.t.*

$$\frac{\dim(V^v \cap V^u)}{\dim(V^u)} > \frac{\dim(V^w)}{t}$$

(equivalently, s.t. given a text $x \in \mathbb{F}^t$ $P\big(x \in V^v \mid x \in V^u\big) > P(x \in V^w)\big)$*, where* $P(\cdot)$ *denotes the probability), then it is always possible to set up a truncated differential distinguisher for* $w + u - v$ *rounds with prob.* $|\mathbb{F}|^{-\dim(V^u) + \dim(V^v \cap V^u)}$.

The result follows from the fact that for each pair (x, y) of plaintexts, where $x \neq y$,

$$\mathrm{P}\big(E_k(x) \oplus E_k(y) \in V^w \mid x \oplus y \in V^0\big) = \mathrm{P}\big(E_k(x) \oplus E_k(y) \in V^v \mid x \oplus y \in V^u\big) = \frac{|\mathbb{F}|^{\dim(V^v \cap V^u)}}{|\mathbb{F}|^{\dim(V^u)}}$$

independently of the secret key k, due to the fact that

$$\forall a, b: \quad \exists c, d \text{ s.t. } R^u(V^0 + a) \subseteq V^u + b \text{ and } R^{w-v}(V^v + b) \subseteq V^w + d,$$

where $R^x(\cdot)$ denotes the x-round encryption function. For comparison, in the case of a random permutation $\Pi(\cdot)$

$$\mathrm{P}\big(\Pi(x) \oplus \Pi(y) \in V^w \mid x \oplus y \in V^0\big) = \frac{|\mathbb{F}|^{\dim(V^w)}}{|\mathbb{F}|^t}.$$

We finally recall that for each subspace X, Y,

$$\dim(X \cap Y) = \dim(X) + \dim(Y) - \dim(X \cup Y),$$

where $\dim(X \cup Y)$ can be easily computed by using a Gram–Schmidt process on $X \cup Y$.

C.2 Impossible Differentials

(Truncated) impossible differential distinguishers/attacks [9] exploit differential that holds with probability 0.

Proposition 8. *Let $\{V^0, V^1, \ldots, V^{\mathfrak{r}-1}\}$ be a subspace trail of prob. 1 defined as in Proposition 7. If there exist $0 \leq v < u < \mathfrak{r}$ s.t.*

$$P\big(x \in V^v \mid x \in V^u\big) = 0$$

(equivalently, $\dim(V^v \cap V^u) = 0$), then it is always possible to set up an impossible differential distinguisher for $\mathfrak{r} + u - v$ rounds.

The reason of the previous result is analogous to the one given before for truncated differential distinguishers with prob. ≤ 1.

D Using our Tool for Starkad and Poseidon Matrices

In addition to the statistical tests described in Section 4, we also used our tool for the Cauchy matrices using specific starting sequences defined for Starkad and Poseidon [19]. We recall that the matrix M' over $\mathbb{F}_{2^n}$ for Starkad is defined by

$$M'_{i,j} = \frac{1}{x_i \oplus y_j},$$

where $x_i = i$, $y_i = i + t$, and $i \in [0, t-1]$. Similarly, the matrix M'' over $\mathbb{F}_p$ for Poseidon is defined by

$$M''_{i,j} = \frac{1}{x_i + y_j},$$

where again $x_i = i$, $y_i = i + t$, and $i \in [0, t-1]$.

Comparison with Related Results. When using our tool for matrices with various sizes (i.e., different values for t), we can observe that some matrices over $\mathbb{F}_{2^n}$ (i.e., the matrices used for Starkad) are vulnerable to the attacks described in this paper. We can also observe, however, that matrices over $\mathbb{F}_p$ using the same t values are not vulnerable. The detailed results for some instances are shown in Table 3.

These results are not new in the literature, since similar conclusions have been already shown in [26, 8]. Moreover, in [26] authors explain how to modify the choice of x_i and y_j in Appendix D in order to fix this problem. This solution consists in changing the

Table 3: Vulnerable matrices for Algorithm 1 and orders t and field sizes $n = \lceil \log_2(p) \rceil$ when considering the STARKAD and POSEIDON specifications.

Generation	n or $\lceil \log_2(p) \rceil$	t	Vulnerable
STARKAD Specification (over $\mathbb{F}_{2^n}$)	4	3	No
	8	3	No
	4	4	Yes
	8	4	Yes
	8	8	Yes
	8	12	Yes
	8	16	Yes
	16	12	Yes
	16	16	Yes
POSEIDON Specification (over $\mathbb{F}_p$)	4	3	No
	8	3	No
	4	4	No
	8	4	No
	8	8	No
	8	12	No
	8	16	No
	16	12	No
	16	16	No

starting sequences for the Cauchy generation method. For completeness, we also tested our algorithm for matrices suggested in [26]. This solution consists in changing the starting sequences for the Cauchy generation method. As expected, we arrive at the same conclusion that *it is not possible to set up infinite-long subspace trail for such modified Cauchy matrices proposed in [26] (in the case of inactive S-boxes). The problem to extend this conclusion also to the case of active S-boxes is open for future work.*