Pitfalls and Shortcomings for Decompositions and Alignment

Baptiste Lambin[1,2][0000-0002-3040-4503], Gregor Leander[1][0000-0002-2579-8587], and Patrick Neumann[1][0000-0003-1624-4256]

[1] Ruhr University Bochum, Bochum, Germany
{gregor.leander,patrick.neumann}@rub.de
[2] University of Luxembourg, Esch-sur-Alzette, Luxembourg
baptiste.lambin@protonmail.com

Abstract. In this paper we, for the first time, study the question under which circumstances decomposing a round function of a Substitution-Permutation Network is possible *uniquely*. More precisely, we provide necessary and sufficient criteria for the non-linear layer on when a decomposition is unique. Our results in particular imply that, when cryptographically strong S-boxes are used, the decomposition is indeed unique. We then apply our findings to the notion of alignment, pointing out that the previous definition allows for primitives that are both aligned and unaligned simultaneously.
As a second result, we present experimental data that shows that alignment might only have limited impact. For this, we compare aligned and unaligned versions of the cipher PRESENT.

Keywords: Supstitution-Permutation Network · Alignment · PRESENT

1 Introduction

Most of the security analysis of symmetric primitives is actually based on their representation and not on the primitive itself: When arguing about the resistance of ciphers or cryptographic permutations, our arguments are in most cases based on a given decomposition of the cipher, in many cases into a linear layer and a set of mappings that are applied in parallel, i.e. S-boxes. While this is very helpful in many cases, it can lead to wrong results in others, see e.g. [3].

Using those ingredients when designing efficient and secure ciphers or cryptographic permutations has a long standing history. It can be seen as having its roots already in Shannon's seminal ideas on confusion and diffusion [25]. While many alternative design strategies exist, the use of S-boxes and linear layers is arguably dominating today's designs and include AES, SHA-3, and many of the primitives for the final round of the NIST lightweight crypto competition[3].

In this paper we touch upon a very fundamental aspect of this decomposition: its uniqueness.

[3] https://csrc.nist.gov/Projects/lightweight-cryptography/finalists

Decomposing a Round Function. Let us turn the view away from designing a cipher or being given a description of a Substitution-Permutation Network (SPN) but rather to a given round function (or maybe even the composition of several round functions). We can imagine having oracle access to the round function only. Here two natural questions arise. First, how to detect if a given round function is actually an SPN round function, and in a second step, how to find the corresponding decomposition. Those questions have been intensively studied since the last 30 years, starting with the seminal SASAS and ASASA cryptanalysis papers [6,22]. Indeed, there are algorithms that can efficiently find a decomposition into an S-box and a linear layer. In particular, those algorithms answer the existence of a decomposition. Moreover, a natural extension of the decomposition of a round function is the representation of a given block cipher, and having multiple representations of the same cipher can have not only a theoretical impact (some properties might be easier to study using one representation rather than another), but also a practical one, see for example [1,24] where using a different representation to implement the ciphers PRESENT and GIFT can lead to an increase in performances.

However, and this is very surprising for us, while the existence of a decomposition has been studied extensively since almost 30 years, the uniqueness of such a decomposition was, to the best of our knowledge, never studied but always given for granted.

To be very clear, there are obvious, and well known, limitations to a unique decomposition already discussed in [6]. In particular, the order of the S-boxes and their choice up to affine equivalence, i.e. up to composing with affine bijections is clearly not unique. What we are interested in, and what has not been questioned so far, is uniqueness up to those equivalent representations.

Crucially, some security arguments, like counting the number of active S-boxes to bound the probability/absolute correlation of a differential/linear trail, could give different results depending on the decomposition. As we will discuss below, without additional requirements on the S-boxes, a unique decomposition is not guaranteed in general.[4] While this is interesting as a fundamental property of round functions, it also impacts very recent work on alignment.

Alignment - intuitively. Alignment of symmetric primitives is a property that has been initially coined during the SHA-3 competition by the Keccak-team in [5] but actually is an idea that dates back to the wide-trail strategy and the use of super-boxes (or code-concatenation) in order to argue about resistance of block ciphers against differential and linear attacks. Interestingly, while already mentioned more than 10 years ago and since then been used in numerous papers [9,11,13,20], the term was never precisely defined. In all the papers mentioned above the term is used in connection with SPNs and we restrict to those designs here as well.

[4] Thankfully, for the case of counting S-boxes those requirements lead to trivial bounds for the probability/absolute correlation of a differential/linear trail, no matter the decomposition.

Intuitively, and this is common in all those papers, alignment is used when the linear layer of an SPN manipulates the state in words, i.e. is word-oriented, and those words are either identical to S-box inputs or consist of multiple S-box inputs (or outputs). However, there are several flavors of this common intuition. For example, several papers mention strong and weak alignment. In [20] the authors mention that no good bounds on the probabilities of differential trails are known for Keccak as it is weakly aligned. Reciprocally, it is argued e.g. in [13] that strong alignment allows proving strong bounds.

The importance of the property, along with its positive and negative connotations, is reflected in several second-round candidates of the NIST lightweight project. For Subterranean the designers state in [11] as a feature that “In a way, the Subterranean round function is the nec plus ultra of weak alignment“, while the authors of Saturnin [9] advertise their design “the strongly-aligned version of the wide-trail strategy“

So while many researches might have one (or several) more (or less) precise ideas what alignment with respect to designing a substitution permutation network might mean, a formal definition was not given.

Alignment – defined. This only changed very recently with the work of Bordes *et al.* [8] at CRYPTO 2021 where a definition of alignment was given and its impact on several cryptographically relevant criteria was studied.

Even so it is not stated exactly this way, a round function, with a given decomposition into an S-box layer and a linear layer, is aligned, according to [8], if and only if the primitive has a super-box structure. A super-box structure means that two rounds of the primitive decompose (up to linear changes of input and output) into a set of two or more parallel applications of mappings. Those parallel mappings are then refereed to as super-boxes.

Unfortunately, this definition has shortcomings as discussed below.

The first problem is based on the fact that alignment is defined for a round function with a given description as a fixed S-box and linear layer. Ideally, one would hope that alignment, i.e. the existence of super-boxes, is a property that is inherent to the round function and not to its description only.

The second shortcoming of alignment is its impact on security. In [8] several aligned and unaligned ciphers were compared with respect to e.g. the number of linear and differential characteristics of high probability. Within this selection, ciphers being aligned suffered more from clustering effects. However, the exact impact of alignment on those properties remained unclear. The main reason for this is that the ciphers that are used in the comparisons are very different and it is far from clear whether the difference in the observed behaviours is (mainly) due to alignment or (mainly) due to other properties that separate the ciphers.

Our Contribution

We first present our main result on the uniqueness of decomposition, Theorem 1, in Section 2, which states that a decomposition is not unique if and only if (at least) one S-box has maximal differential uniformity and another one has maximal linearity. For better readability, we first present the results and shift the proofs and several more technical insights to Section 5.

With respect to alignment, we show in Section 3, based on the non-uniqueness of maximal decomposition in general, that there exist round functions that have both an aligned and an unaligned description. We also give a non-artificial example, namely the cipher DEFAULT[2], which is aligned and unaligned at the same time.

Furthermore, in Section 4 we present experimental data showing that the impact of alignment with respect to the security criteria studied in [8] may be almost non-existent if we consider alignment as an isolated property. We show this by comparing variants of the same cipher instead of different ciphers. For this we choose the cipher PRESENT. Here, by changing only the original bit-permutation one can nicely create versions that are aligned or unaligned. As we detail, those variants behave very similarly in all aspects, in contrast to the results of [8], see e.g. Fig. 3.

2 Main Results on the Uniqueness of Decompositions

It is easy to see that every round function that is un-keyed (or in which a simple key addition takes place at the end of the round) can be seen as an SPN by simply choosing the non-linear layer to be the round function itself (without the addition of the key in the keyed case) and the linear layer to be the identity. Obviously, this representation of a round function is not very useful. But it already shows that without further restrictions, a decomposition of a round function into non-linear and linear layer cannot be unique. Hence, it is not clear if the properties one infers from one possible decomposition are the same for a different one.

In other words, if we want to infer properties of a round function based on its linear and non-linear layer, we should make sure that it is actually well defined, i. e. it only depends on the round function and not on the choice of decomposition. To give just one example, the arguments made in [3] for the resistance against invariant attacks are only valid for one given decomposition and can be shown to be invalid for another.[5]

Before presenting our results, let us first fix some notations and definitions.

2.1 Preliminaries

Basic Notation. We will denote by $\mathbb{F}_2$ the Galois field with 2 elements and with $+$ the addition in this field, which can be seen as an exclusive or. With $\mathbb{F}_2^n$ we will denote the n dimensional vector space over this field, with x^T the transpose of a (column) vector $x \in \mathbb{F}_2^n$ and with $x^T \cdot y$ the (canonical) inner product of $x, y \in \mathbb{F}_2^n$. Furthermore, we will denote by $\oplus$ the direct sum of vector spaces, i. e. $U \oplus V = W$ if and only if $W = U + V := \{u + v \mid u \in U, v \in V\}$ and

[5] The arguments are mainly based on the rational canonical form of the linear layer and this form might change when using linear equivalent S-boxes and modifying the linear layer accordingly.

$U \cap V = \{0\}$ for vector spaces $U, V \subset W$. Given a direct sum $\bigoplus_i U_i = W$ we will denote by $\pi_i^U : W \to U_i$ the projection onto U_i along $\bigoplus_{l \neq i} U_l$, i. e. π_i^U has kernel $\bigoplus_{l \neq i} U_l$, image U_i and is the identity if restricted to U_i.[6] We will also consider the direct sum $\left(\bigoplus_{i \in I} U_i\right) \oplus \left(\bigoplus_{i \notin I} U_i\right)$ and will write $U_I := \bigoplus_{i \in I} U_i$ as well as $\pi_I^U := \sum_{i \in I} \pi_i^U$ for the projection onto $\bigoplus_{i \in I} U_i$ along $\bigoplus_{i \notin I} U_i$ or simply $U_{i \neq l}$ and $\pi_{i \neq l}^U$ if $I = \{i | i \neq l\}$. Also, we will denote by $F_{|U}$ the restriction of a function F to the subset U of its domain.

General Definitions. Let us give some general definitions that we will use throughout this paper.

Definition 1 (Linear/Affine Equivalence). *We call two functions $F, G : \mathbb{F}_2^n \to \mathbb{F}_2^n$* affine equivalent *if there exist $a, b \in \mathbb{F}_2^n$ as well as invertible matrices $A, B \in \mathbb{F}_2^{n \times n}$ such that $F(x) = b + B \cdot G(A \cdot x + a)$ for all $x \in \mathbb{F}_2^n$. If $a = 0 = b$ we also call F, G* linear equivalent*.*

Definition 2 (Differential Uniformity (cf.[23])). *Let $F : U \to V$ for subspaces $U, V \subset \mathbb{F}_2^n$. Then we call $\max_{\alpha \in U \setminus \{0\}, \beta \in V} |\{x \in U | F(x) + F(x + \alpha) = \beta\}|$ the differential uniformity of F and say that it is maximal if it is equal to $|U|$.*

Definition 3 (Linearity, (see [10] for an in-depth discussion)). *Let $F : U \to V$ for subspaces $U, V \subset \mathbb{F}_2^n$. Then we call $\max_{\alpha \in \mathbb{F}_2^n \setminus V^\perp, \beta \in \mathbb{F}_2^n, c \in \mathbb{F}_2} |\{x \in U | \alpha^T \cdot F(x) = \beta^T \cdot x + c\}|$ the linearity of F, where $V^\perp := \{x \in \mathbb{F}_2^n | x^T \cdot y = 0 \; \forall y \in V\}$, and say that it is maximal if it is equal to $|U|$.*

2.2 Defining a (Maximal) Decomposition

Given a (round) function we would like to be able to find a unique decomposition into non-linear and linear layer(s). Similarly to [8], we require the number of S-boxes in the non-linear layer to be maximal. While in general such a decomposition is not unique, we show conditions under which it is. Hence, under these conditions it is possible to infer properties of the round function (such as alignment[8]) based on the linear and non-linear layer(s) of the corresponding unique decomposition.

A Natural Definition. Let us start with formally defining what we understand by a decomposition. Since a non-linear layer typically consists of independent S-boxes (i. e. functions that don't share input and output bits), one can see the S-box layer as the sum of independent functions. For example, the non-linear layer $N : \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \to \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2}$ with

$$N \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} S_1(x_1) \\ S_2(x_2) \end{pmatrix}$$

[6] In other words, the direct sum enables us to express every element $x \in W$ as $\sum_i x_i$ for unique $x_i \in U_i$. Hence, π_i^U is the mapping defined by $x = \sum_i x_i \mapsto x_i$.

and $S_1 : \mathbb{F}_2^{n_1} \to \mathbb{F}_2^{n_1}$ as well as $S_2 : \mathbb{F}_2^{n_2} \to \mathbb{F}_2^{n_2}$ can be seen as

$$N\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} S_1(x_1) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ S_2(x_2) \end{pmatrix}$$

where the input/output spaces of those two functions are $\mathbb{F}_2^{n_1} \times 0^{n_2}$ resp. $0^{n_1} \times \mathbb{F}_2^{n_2}$. Note that $\mathbb{F}_2^{n_1} \times 0^{n_2} \oplus 0^{n_1} \times \mathbb{F}_2^{n_2} = \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2}$. A linear layer now only changes the input/output spaces of those functions. Hence, we will define a decomposition of a (round) function by the sum of functions defined on subspaces of the input and output spaces.

Definition 4 (Decomposition). *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective. Furthermore, let $U_1, ..., U_d$ and $V_1, ..., V_d$ be non-trivial*[7] *subspaces of $\mathbb{F}_2^n$ with $\bigoplus_i U_i = \mathbb{F}_2^n = \bigoplus_i V_i$, as well as $F_i : U_i \to V_i$ with*

$$F(x) = F\left(\sum_i \pi_i^U(x)\right) = \sum_i F_i \circ \pi_i^U(x).$$

We call $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ a decomposition *of F. If $d > 1$ we call the decomposition* non-trivial*.*

Note that the reason we need to restrict the definition to non-trivial subspaces is that it is always possible to extend the decomposition by $\hat{F} : \{0\} \to \{0\}$, which does not give any additional information about the (round) function. Moreover, as the order of the functions does not matter, we only consider a set of tuples.

We would like to point out that we actually allow two linear layers, i. e. we are decomposing into $L_2 \circ N \circ L_1$ where L_1 and L_2 are linear layers and N is a non-linear layer. But note that the r round iteration $(L_2 \circ N \circ L_1)^r$ is linear equivalent to $(L_1 \cdot L_2 \circ N)^r$, meaning that both (round) functions, $L_2 \circ N \circ L_1$ and $L_1 \cdot L_2 \circ N$, lead to the same cryptographic properties. With that, allowing two linear layers seems actually more natural than restricting to one linear layer only.

Refining Decompositions. It is easy to see that given a (non-trivial) decomposition $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ it is always possible to find another decomposition by combining two (or more) of the F_i, i. e. $\{(\bigoplus_{i \in I_l} U_i, \bigoplus_{i \in I_l} V_i, \sum_{i \in I_l} F_i \circ \pi_i^U) \mid 1 \leq l \leq m\}$ is also a decomposition for every partition of the index space $I_1, ..., I_m \subset \{1, ..., d\}$.

Hence, in order for a decomposition to be unique it is clear that one at least has to maximize the number of S-boxes. To this end, we will now define in which case one decomposition is a refinement of another decomposition, which reminds of [8] while technically being different as [8] focuses on the linear layer only.

Definition 5 (Refinement). *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective. Let further $D = \{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ and $E = \{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two decompositions of F. We call E a* refinement *of D if $e > d$ and for all i there exists a value of j such that $W_i \subset U_j$.*

[7] More precise, we allow the subspaces to be equal to $\mathbb{F}_2^n$ but not to be $\{0\}$.

The intuition behind this is that we further decompose each of the individual S-Boxes. Based on this, we can now define a maximal decomposition.

Definition 6 (Maximal Decomposition). *We call a decomposition D* maximal *if there exists no refinement of D.*

In other words, a decomposition is maximal if we cannot decompose the individual S-boxes any further. It is an interesting question, which we leave open for now, whether two maximal decompositions have to have the same number of S-boxes, or even more, the same size spectrum (see Definition 10 below).

2.3 A Sufficient and Necessary Condition for Unique Decompositions

Knowing the definition of a maximal decomposition, we can now state our main result in context of decompositions.

Theorem 1. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $D = \{(U_i, V_i, F_i)|1 \leq i \leq d\}$ be a maximal decomposition of F. Then D is unique if and only if there exists no pair (i, k) with $i \neq k$ such that F_i has maximal differential uniformity and F_k has maximal linearity.*

We will prove this in detail in Section 5.

Since one of the S-boxes having maximal differential uniformity (resp. maximal linearity) means that the same has to be true for the whole (round) function, we can relax the condition and receive a sufficient, but not necessary, condition.

Corollary 1. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective. If F does not have maximal differential uniformity or maximal linearity then F has a unique maximal decomposition.*

Since differential uniformity and linearity are properties that should already be known for most (if not all) cryptographic primitives, this makes it easy to argue about the uniqueness of their maximal decomposition.

Some Intuition on the Functions without a Unique Maximal Decomposition. Given Theorem 1 it is actually not hard to show (details are given in the full version [17]) that the functions without a unique maximal decomposition are exactly those that are affine equivalent to ones of the form

$$R\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} f(x_1) \\ g(x_1) + x_2 \\ x_3 \\ h\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} \text{S-box(es)} \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} \text{S-box(es)} \end{matrix},$$

where $x_1 \in \mathbb{F}_2^n, x_4 \in \mathbb{F}_2^m$ for integers n, m and $x_2, x_3 \in \mathbb{F}_2$, as well as $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$, $g : \mathbb{F}_2^n \to \mathbb{F}_2$ and $h : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2^m$. The reason is that such functions allow us to "mix" S-boxes without changing it, as we can add x_3 (from the second

S-box) to x_2 (from the first S-box) before the non-linear layer, but also revert this linear transformation after the non-linear layer by just changing the original linear layer (more details are given in Example 2 of the full version [17]).

Given that one typically tries to minimize the differential uniformity and linearity, it may seem like functions without a unique maximal decomposition are not of interest to cryptographers, but there exists at least one widely-used type of round function that actually has no unique maximal decomposition, namely the one of a generalized Feistel network.

In total, we are able to show that whenever cryptographically strong S-boxes are used the representation of its round function can indeed be used for arguing about its properties, while the same is not always true in the opposing case, as we will see next.

3 Re-Aligning Alignment

To give just one example of why the uniqueness of a maximal decomposition can be important, let us take a look at the concept of alignment by Bordes *et al.*[8]. While the intuition should be that a round function is aligned if the primitive has a superbox structure, i. e. the iteration of two rounds exhibits a non-trivial decomposition, the original definition is a bit more involved. Therefore, let us quickly recall the definition of alignment from [8]. For this, we assume that the round function consists of the parallel application of m equally-sized S-boxes, a bijective linear transformation and the addition of a key (resp. constant), i. e. we can write the round function as $L \circ N + c$, where $L \in \mathbb{F}_2^{n \times n}$ is a bijective linear mapping,

$$N = \begin{pmatrix} S_1 \\ \vdots \\ S_m \end{pmatrix}$$

is the non-linear layer, with $S_i : \mathbb{F}_2^{n/m} \to \mathbb{F}_2^{n/m}$ bijective, and $c \in \mathbb{F}_2^n$ is a constant (resp. key). For simplicity, we will use a slightly different but equivalent version of the definition from [8].

Definition 7 (Alignment[8], sub-optimal). *Let $U_i := 0^{(i-1)\cdot n/m} \times \mathbb{F}_2^{n/m} \times 0^{(m-i)\cdot n/m}$ for $i = 1, \dots, m$. A round function $L \circ N + c$ (as above) is called* aligned *if L can be written as $T \circ M$ such that*

- *it exists a permutation $\tau : \{1, \dots, m\} \to \{1, \dots, m\}$ with $T(U_i) = U_{\tau(i)}$ and*
- *it exists $J \subset \{1, \dots, m\}$ (non-trivial) such that $M\left(\bigoplus_{i \in J} U_i\right) = \bigoplus_{i \in J} U_i$ and $M\left(\bigoplus_{i \notin J} U_i\right) = \bigoplus_{i \notin J} U_i$,*

where the split between the linear and nonlinear layer is chosen so as to maximize the number of S-boxes in N. [8]

[8] This version is equivalent to the one from [8] since $T(U_i) = U_{\tau(i)}$ means that T is the composition of a Π_N-Shuffle and a Π_N aligned linear function. As the composition

The problem with this definition is that it is not invariant under affine (or even linear) equivalence (of the round function iterated two times), while the existence of a superbox structure is. One non-artificial example where this is indeed a problem is the cipher DEFAULT[2], which is aligned, but a linear equivalent version would not be aligned.

Before we explain the problem in more detail, let us give an alternative definition that is exactly equivalent to the existence of a superbox structure.

Definition 8 (Alignment). *We call a round function R* aligned *if there exists a non-trivial decomposition of $R \circ R$. In the keyed case we require a non-trivial decomposition for every key.*

Note that whenever a round function is aligned according to the original definition it is indeed aligned, since it implies a non-trivial decomposition of $R \circ R$ with in- and output spaces of a certain form, while a round function that is not aligned according to the original definition could actually be aligned, but the in- and output spaces may not be of the form required in Definition 7.

On the Need of Decomposability for all Keys. On the first glance, requiring decomposibility for all possible key choices may seem as a downside. But note that the original definition does the same (even in the un-keyed case), as it implies a non-trivial decomposition of $R \circ R$ for all possible key/constant additions. Also, there does not seem to be a way around this as the existence of a superbox structure can be key dependent. For instance, let F be self inverse and let $R = F + k$ be the round function for a round key k. Then $R \circ R$ is affine equivalent to $F(F + k)$, which is the identity for $k = 0$ and therefore clearly decomposable, while the same is not necessarily true for $k \neq 0$.

A real world example of a key dependent decomposition is the block cipher CRAFT[4], for which Leander *et al.* show in [19] that there exist some Tweakkeys for which the round function is now very similar to a Feistel network, while originally being an SPN.

That said, it seems to be more in line with [8] to require the existence of a superbox structure for all possible key choices, while clearly a more refined definition would be possible. But, since we show in Section 4 that the impact of alignment may be hugely overestimated, we will settle with a definition most true to the original one for now.

Alignment and Generalized Feistel Networks. While there exist non-artificial examples, like the cipher DEFAULT[2], that are aligned according to the definition from [8], but a linear equivalent cipher is not, showing this for DEFAULT is a bit more involved, but does not give further insights into the problem. Hence,

of a $\Pi_{N'}$ aligned and a Π_N aligned function is obviously $\Pi_{N'}$ aligned if $\Pi_N \leq \Pi_{N'}$, it is then enough to check that M is $\Pi_{N'}$ aligned, with $\Pi_{N'}$ being non-trivial. Since it is only important that $\Pi_{N'}$ is non-trivial, this can be done by checking if $M\left(\bigoplus_{i \in J} U_i\right) = \bigoplus_{i \in J} U_i$ and $M\left(\bigoplus_{i \notin J} U_i\right) = \bigoplus_{i \notin J} U_i$ for some $J \subset \{1, \ldots, m\}$, i. e. checking for all possible $\Pi_{N'}$ with two boxes.

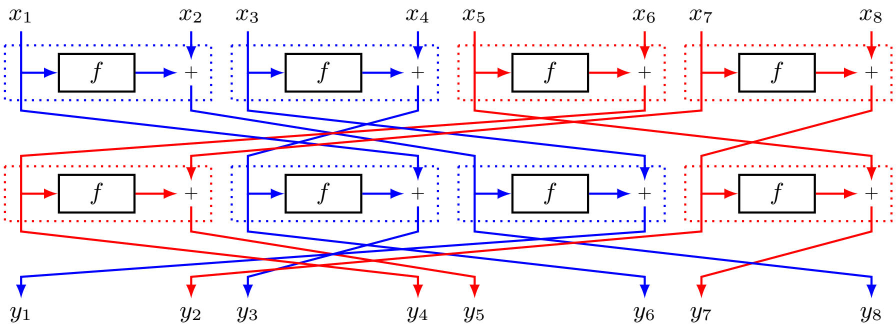


Fig. 1: Example of a round function that is aligned, while a linear equivalent two round composition would not be according to [8]. The colors indicate the alignment resp. superbox structure and the dotted lines the individual S-boxes.

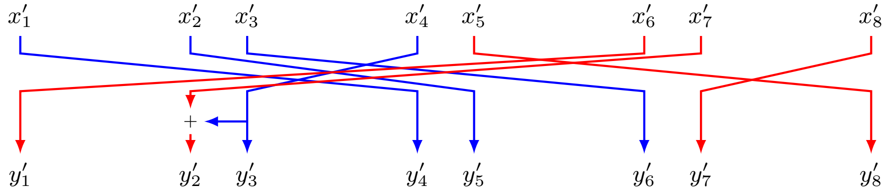


Fig. 2: Linear layer of a (linear) self-equivalent version of the round function from Fig. 1. The colors indicate the initial alignment resp. superbox structure.

we refer to the full version [17] for this and use a more suitable example at this point. Let us have a look at the generalized Feistel network depicted in Fig. 1. Obviously, we can see the permutation of the output as a linear layer and the non-linear transformation of each Feistel branch as an S-box (indicated by the dotted box in the figure). As can be seen, two rounds of this construction result in two independent parts, which shows the existence of a superbox structure. A quick check with the definition from [8] shows that this structure is indeed aligned. But the problem is that the same does not hold true for a linear equivalent version of this two round structure. To see this, note that the linear transformation that adds x_3 to x_2 commutes with the S-box layer. Hence, we can look at a linear equivalent version where we replace the second permutation/linear layer by this linear transformation. Since it commutes with the S-box layer, we end up with the linear layer depicted in Fig. 2 between the two S-box layers. Here we can see that the addition of y'_3 to y'_2 mixes the two colors, and a quick look at the definition from [8] reveals that this version would not be aligned anymore, while the superbox structure is obviously still present. For more details on this, we refer to the full version [17].

4 Aligned and Unaligned Versions of PRESENT

We are interested in looking at how alignment affects the cryptographic properties of very similar primitives, and to do so we chose to look at variants of the block cipher PRESENT where we keep the same S-box but change the permutation. Given the link to central digraphs and the fact that those are fully classified in dimension 16, makes PRESENT a very suitable candidate for our purpose as one can essentially investigate the full spectra of bit permutations and their impact on alignment.

Especially, we will show that alignment can have a very minor effect in how it influences linear and differential trails.[9]

4.1 Digraphs and PRESENT

One round of PRESENT consists of 16 parallel applications of the same 4-bit S-box followed by a bit permutation of the 64-bit state. The permutation is chosen in such a way that full dependency is reached after two rounds, or more precisely 3 applications of the S-box layer interleaved with two permutation layers. In [16] it was shown how any such permutation leads to what is called a central digraph.

Definition 9. *Let $G = (V, E)$ be a directed graph with vertices V and edges E. G is a central digraph if for every pair of nodes $u, v \in V$ it exists a unique $w \in V$ such that $(u, w) \in E$ and $(w, v) \in E$. That is for every pair of vertices there exist a unique path of length 2 between them.*

A PRESENT-like bit permutation P operating on $\{0, \dots, 63\}$ gives rise to a central digraph by identifying the 16 S-boxes with 16 vertices and adding an edge from vertex i to vertex j if there exists an output bit of the i-th S-box that is mapped to the j-th S-box in the next round. Note that when restricting to permutations with full dependency after two rounds (thus leading to a central digraph), there are no duplicate edges on this digraph as, for a given S-box, each of its four output bits needs to be sent to a different S-box on the next round.

All Central Digraphs of Order 16 *up to isomorphism.* There are exactly 3492 central digraphs of order 16 up to graph isomorphism, see [16]. Graph isomorphism, translated to the PRESENT-like structure, correspond to permuting the order of S-boxes. Moreover, many PRESENT-like permutations will end up in the same digraph as the order of input and output bits within each S-box is neglected in the graph representation, but relevant for the cipher. As such, not all properties of the cipher can be deduced from the digraph directly. For example, as shown in [18] the number of (linear) trails might well differ for two permutations that correspond to the same digraph. While some properties cannot be deduced from the digraph only, others – in particular alignment – can.

[9] The code we used to make these experiments is available at: `https://doi.org/10.5281/zenodo.7660387`.

Aligned and Non-Aligned Central Digraphs. As mentioned above there are exactly 3492 central digraphs leading to full diffusion after 2 rounds. We can further group them into several sets using the following definition.

Definition 10 (Size spectrum of a decomposition). *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $D = \{(U_i, V_i, F_i) | 1 \leq i \leq d\}$ be a decomposition of F. We call the multiset $\mathcal{S}_D = \{dim(U_i) | 1 \leq i \leq d\}$ the* size spectrum *of D.*

Moreover, we will say that a digraph G is aligned (resp. unaligned) if the permutations induced from G result in an aligned (resp. unaligned) round function. Recalling the previous sections and in particular Corollary 1, this makes sense as the PRESENT S-box neither has trivial linearity nor differential uniformity and thus its maximal decomposition is unique.

Over all the 3492 central digraphs, their alignment is distributed as follows:

- One single digraph is aligned with a maximal decomposition of size spectrum $\{4, 4, 4, 4\}$. Especially, the original permutation in PRESENT belongs to this class; that is two rounds of PRESENT can be rewritten as consisting of 4 super-boxes of 16 bits each.
- 37 digraphs lead to a maximal decomposition of size spectrum $\{4, 4, 8\}$;
- 1207 digraphs lead to a maximal decomposition of size spectrum $\{4, 12\}$;
- 220 digraphs lead to a maximal decomposition of size spectrum $\{8, 8\}$;
- 2027 digraphs lead to a maximal decomposition of size spectrum $\{16\}$, i.e. all these digraphs are unaligned.

Since our goal is to compare aligned and unaligned versions of PRESENT, we will focus on the two corresponding cases to generate permutations: the (single) case where the maximal decomposition is of size spectrum $\{4, 4, 4, 4\}$, corresponding to the digraph of the original permutation, and the case where the digraph is unaligned. As in [8], we will focus on the linear and differential properties of the resulting variants of PRESENT.

4.2 Linear Cryptanalysis

The idea is to focus only on linear trails with one active S-box per round. This has two advantages. First, it covers the bulk of the correlation used in most of the attacks on PRESENT, cf. [14]. Second, it simplifies the analysis and we can nicely use graph theory and efficient algorithms therein to control the effect of unaligning the original PRESENT permutation.

A linear trail with a single active S-box per round implies for a bit-permutation that has full dependency in two rounds, that only one-to-one bit linear approximations have to be considered through the S-box. There are only 8 possible one-to-one bit transitions with non-zero correlation within the PRESENT S-box, see [7]. For us, the only important point here is that there is no 1-1 transition involving the LSB neither at the input nor at the output of the S-box.

As such, if one were to change the permutation by only modifying the values in the permutation that are affecting the LSB (both at the input and output) of the S-boxes, one would get a new permutation that has the same number of linear trails built from 1-1 transitions. That is, two bit-permutations that only differ in those bits, lead to ciphers with a very similar behaviour with respect to linear attacks. More specifically, this means that an alternative permutation should be built from the following partial permutation, where $\star$ is an undetermined value

$$\begin{aligned} p^\star = (&\star,\star,\star,\star,\star,17,33,49,\star,18,34,50,\star,19,35,51,\star,\star,\star,\star,\star,21,37,53,\star,22,\\ &38,54,\star,23,39,55,\star,\star,\star,\star,\star,25,41,57,\star,26,42,58,\star,27,43,59,\star,\star,\star,\star,\\ &\star,29,45,61,\star,30,46,62,\star,31,47,63). \end{aligned}$$

Extending the partial permutation in an unaligned way. To build the entire permutation, we want it to both have the above structure (i.e. fits the partial permutation) as well as lead to one of the central digraphs that has the alignment we want to study. Luckily, this can nicely be done with graph theory. The idea is that the partial permutation has to correspond to a sub-graph in one of the central digraphs we aim for.

More explicitly, from the partial permutation $p^\star$, one can deduce a digraph $H^\star$. $H^\star$ contains an edge from vertex i to vertex j if there exists an output bit of index a with $p^\star(a) \neq \star$ in the i-th S-box that is mapped to an input bit of index b with $p^\star(b) \neq \star$ in the j-th S-box in the next round. Then there exists a permutation that both fits the partial permutation *and* leads to one of the 2027 unaligned digraphs G if and only if $H^\star$ is subgraph isomorphic to G.

Such a permutation would then be unaligned while still preserving the linear trails built from 1-1 transitions. We used the subgraph isomorphism solver by McCreesh et al. [21] to find all possible subgraph isomorphisms between $H^\star$ and the unaligned digraphs, and over the 2027 unaligned digraphs, there are 346 that lead to at least one subgraph isomorphism.

Thus, to build alternative permutations, we first choose one of these 346 (unaligned) subgraphs, choose one of the subgraph isomorphism and complete the partial permutation according to this isomorphism to obtain a full permutation. Note that several permutations can be built from a given digraph and subgraph isomorphism, however the exact number is rather hard to evaluate.

By building alternative permutations like this, we can deduce the following proposition that takes care of the most prominent linear trails.

Proposition 1. *For any (aligned or unaligned) permutation generated as described above, the number of linear trails in PRESENT built with only 1-1 transitions is exactly the same as when using the original permutation, and all those trails have an absolute bias of* 2^{-3r} *over* r *rounds.*

4.3 Differential Cryptanalysis

Once we obtain a permutation, we can then evaluate different metrics for differential cryptanalysis, and as an echo to [8] we chose to evaluate the number of trails, number of core patterns and number of differentials (i.e. considering the

Table 1: Largest deviations observed on the number of differential trails.

	1 permutation per graph			all permutations in the same class		
Rounds	Original	Alternative	Weight	Min. Alt.	Max. Alt.	Weight
2	$2^{48.35}$	$2^{48.21}$	24	$2^{11.11}$	$2^{11.15}$	6
3	$2^{46.29}$	$2^{43.05}$	32	$2^{8.93}$	$2^{9.15}$	8
4	$2^{8.19}$	$2^{6.32}$	12	$2^{7.09}$	$2^{7.52}$	12

clustering effect of differential trails), which are defined in the following section. We can then repeat the procedure by generating another permutation, possibly also choosing another digraph and/or subgraph isomorphism.

Differential Trails. We made experiments to observe how having an unaligned permutation affects the distribution of differential trails (for a formal definition of *differential trails*, we refer to Definition 13 of the full version [17]), more specifically the number of trails of a given weight (i.e. the $-\log_2$ of the probability).

We computed the number of trails up to a given weight over 2, 3 and 4 rounds for both the original permutation as well as for several permutations generated as described in the previous section. The computation was done as an exhaustive search using a standard Branch & Bound algorithm as well as the convolution technique showed in [8] for the first and last round. For 3 rounds, we adapted the algorithm in Appendix A.3 of [8], with some simplifications and optimizations since the linear layer is only a permutation in our case. For 4 rounds, the same algorithm as for 3 rounds is used to generate trails over 3 rounds, which are then manually extended by adding one round at the end.

We give the detailed results in the full version [17], comparing the original permutation to a batch of variant permutations generated in 2 ways: either one random permutation is generated from one random isomorphism for *each* digraph (thus 346 variants considered), or 346 permutations are generated from one isomorphism and one single graph, to showcase that the number of trails remains rather stable within the same class of permutation. We also give a short summary of the largest deviations in Table 1. Each row starts by the number of rounds, followed by 3 entries giving the number of trails for the original (resp. alternative) permutations with the largest gap and for which weight this gap happens, when the alternative permutations are generated as one permutation for one isomorphism for each graph. The remaining 3 entries of the row showcase the largest gap between any of the alternative permutations, which are generated within the same class, showing that in this case the results are rather stable.

Overall, while the (unaligned) alternative permutations tend to lead to lower numbers of trails than the (aligned) original permutation (but not always, e.g. over 4 rounds, there are $\sim 2^{17.31}$ trails of weight 15 for the original permutation while one alternative permutation leads to $\sim 2^{17.40}$ trails of the same weight), the gap between the original and alternative permutations is rather small and the distribution seems to remain very similar.

Table 2: Largest deviations observed on the number of core patterns. No deviation for 2 rounds when permutations are in the same class.

	1 permutation per graph			all permutations in the same class		
Rounds	Original	Alternative	Active S-boxes	Min. Alt.	Max. Alt.	Active S-boxes
2	$2^{31.48}$	$2^{31.34}$	12	-	-	-
3	$2^{34.18}$	$2^{31.32}$	17	$2^{27.99}$	$2^{28.05}$	15
4	$2^{33.70}$	$2^{31.02}$	18	$2^{31.81}$	$2^{31.92}$	18

Core Patterns. In [8], Bordes et al. look at the influence of alignment for truncated differentials (where one only considers whether or not a given S-box has a non-zero difference) over a few block ciphers. However due to PRESENT's linear layer being a bit-permutation, truncated differentials are not really meaningful to study. Instead, we will consider *core patterns*, which are essentially differential trails without considering the first and last S-box layer. More precisely, core patterns are defined as follows.

Definition 11. *Let $(\alpha_1, \beta_1, \alpha_2, \dots, \alpha_r, \beta_r)$ be a differential trail over r rounds with*

$$\alpha_1 \xrightarrow{S} \beta_1 \xrightarrow{p} \alpha_2 \xrightarrow{S} \dots \xrightarrow{p} \alpha_r \xrightarrow{S} \beta_r$$

Then $(\beta_1, \alpha_2, \dots, \alpha_r)$ is called a core pattern.

While core patterns are still influenced by valid differentials of the S-box for 3 rounds and more, considering them allows us to ignore the effect of the first and last S-box layer that "inflate" the number of trails.

As for differential trails, we computed the number of core patterns for up to a given number of active S-boxes over 2, 3 and 4 rounds, both for the original permutation as well as for a batch of alternative permutations, which are given in details in the full version [17], with a short summary for the largest deviations given in Table 2, structured in the same way as the previous table for differential trails. The algorithm to obtain these results is very similar as for differential trails, except that the bounding step in the Branch & Bound is done over the number of active S-boxes, ignoring the weight of the trails as well as ignoring the first and last S-box layer.

Clustering of Differential Trails. While differential trails are what is usually used to mount differential attacks, it has been shown multiple times (e.g. [12,15,26]) that the actual probability of the underlying differential can deviate quite significantly from the probability of a differential trail because of the *clustering effect*. In short, the probability of a differential (α_1, α_r) is the sum of the probability of all differential trails starting (resp. ending) with α_1 (resp. α_r), see Proposition 2 of the full version [17]. Usually, the *dominant trail hypothesis* is used to argue that the highest probability of any trail fitting a given differential dominates over all other trails, that is, the probability of the differential is about

Table 3: Largest deviations observed on the number of differentials.

	1 permutation per graph			all permutations in the same class		
Rounds	Original	Alternative	Weight	Min. Alt.	Max. Alt.	Weight
2	$2^{25.63}$	$2^{25.08}$	12.9556	$2^{15.99}$	$2^{16.02}$	8.97763
3	$2^{30.15}$	$2^{28.69}$	20.8301	$2^{9.07}$	$2^{9.20}$	8

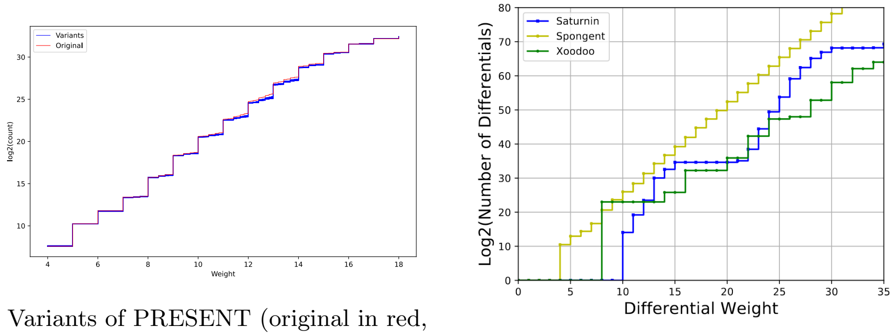
the same as the probability of the trail. However several examples in the literature show that this is not always the case, and as such it is useful to evaluate the actual probability of differentials over a given cipher.

Adapting the algorithm used to compute differential trails, we were able to also compute the probabilities of differentials over 2 and 3 rounds for both the original permutation as well as for several alternative permutations. Again, we give a short summary of the largest observed deviations in Table 3, while the detailed results are given in the full version [17]. As we can see, the distribution for alternative (unaligned) permutations seems quite close to the distribution from the original (aligned) permutation. As an example, in Fig. 3a, we give the *cumulative* histogram of the number of differential of a given weight over 2 rounds, where the red line represents the cumulative histogram for the original permutation, while each alternative permutation is represented by a blue line. Since all the alternative permutations give very close results, their respective lines are all clumped together, which still highlights what we want to show here. Note that while for 2 rounds, the alternative permutations give a very similar curve, for 3 rounds there seems to be a divergence starting around weight 15, while still remaining quite close (we refer to the full version [17] for more details).

To compare this to [8], we give in Fig. 3b the cumulative histogram the authors gave from their analysis of Saturnin, Spongent and Xoodoo, the first two being aligned while the latter being unaligned. Here, the largest gap between Spongent (aligned) and Xoodoo (unaligned) is several orders of magnitude larger than for PRESENT (about 2^{20} near weight 21), and the distributions are very different. Thus, we now have examples of aligned and unaligned ciphers that behave very similar (the PRESENT variants) and very different (Spongent and Xoodoo). Those examples jointly raise doubt about the impact of alignment on its own.

Moreover, the minimal weight for which there is at least one differential is different, while in our experiments when comparing variants of the same cipher, the same minimal weight is achieved both for the (aligned) original permutation as well as for the (unaligned) variants.

While these experiments show that we can find unaligned permutations that are technically better than the original PRESENT permutation, the margin between the two is rather small and the distributions seem to be very close to each other. Overall, the distributions also seem to diverge mostly toward ”larger” weights, while low weight trails/differentials are probably the most important ones to focus on, and it’s worth recalling that all of these alternative permuta-



(a) Variants of PRESENT (original in red, variants in blue)

(b) Saturnin, Spongent and Xoodoo [8]

Fig. 3: Cumulative histogram of the number of differentials of a given weight over 2 rounds.

tions have the exact same linear trails built from 1-1 transitions as the original permutation. At the very least, these experiments clearly show that when comparing aligned vs. unaligned round function of a very similar design, the answer to whether one is clearly better than the other is a lot less clear cut than what was stated in [8].

5 An In Depth Analysis of the Uniqueness of Decompositions

In this section, we will explain the results from Section 2 in greater depth. More precisely, we will give more details on decompositions, how one can (possibly) find refinements based on two decompositions, and finally, prove Theorem 1.

Reconstructing a Decomposition Based on the Input Spaces. While the F_i, which can be interpreted as linear transformed S-boxes, are useful for motivating our definition, they are strictly speaking not needed as long as the actual function F is known. To see this, we first note that they can be recovered by projecting onto the corresponding output space.

Corollary 2. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ be a decomposition of F. Then we have that $\pi_i^V \circ F = F_i \circ \pi_i^U$.*

Proof. Since $F = \sum_j F_j \circ \pi_j^U$ and $\pi_j^V \circ F_i$ is the same as F_i in the case that $i = j$ and zero otherwise, the claim follows from

$$\pi_i^V \circ F = \pi_i^V \circ \left(\sum_j F_j \circ \pi_j^U \right) = \sum_j \pi_i^V \circ F_j \circ \pi_j^U = F_i \circ \pi_i^U . \qquad \square$$

Now, we show that the F_i provide no additional information, since the restriction of F to U_i is identical to F_i plus some constant that can be recovered by projecting $F(0)$ onto $\bigoplus_{l \neq i} V_l$.

Corollary 3. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ be a decomposition of F. Then we have that $F_i = F_{|U_i} + \pi^V_{l \neq i} \circ F(0)$ and we can write $F = \sum_l F \circ \pi^U_l + (d + 1 \bmod 2) \cdot F(0)$.*

Proof. We know that $F = \sum_l F_l \circ \pi^U_l$ by definition. Hence, we have that

$$F \circ \pi^U_i = \sum_l F_l \circ \pi^U_l \circ \pi^U_i = F_i \circ \pi^U_i + \sum_{l \neq i} F_l(0) = F_i \circ \pi^U_i + \sum_{l \neq i} \pi^V_l \circ F(0).$$

Therefore, we get that

$$\begin{aligned} F &= \sum_l F_l \circ \pi^U_l = \sum_l \left(F \circ \pi^U_l + \sum_{i \neq l} \pi^V_i \circ F(0) \right) \\ &= \sum_l F \circ \pi^U_l + \sum_l \big(F(0) + \pi^V_l \circ F(0) \big) = \sum_l F \circ \pi^U_l + (d + 1 \bmod 2) \cdot F(0). \end{aligned}$$

□

Looking at the S-box layer, we quickly see why this is the case. Assume that

$$N \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} S_1(x_1) \\ S_2(x_2) \end{pmatrix}.$$

To isolate S_1, we would fix $x_2 = 0$. But in order to get the second component to be zero for the output space to be a subspace, we have to add $S_2(0)$ to it.

In addition to not having to know the F_i, we also do not need to know the output spaces V_i. But note that it is still important that the output sets are subspaces that are in direct sum.

Corollary 4. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ be a decomposition of F. Then we have that $V_i = F(U_i) + F(0)$.*

Proof. We already know that $V_i = F_i(U_i) = F(U_i) + \pi^V_{l \neq i} \circ F(0)$. The claim now follows from the fact that $\pi^V_i \circ F(0) = F_i(0) \in V_i$. □

Hence, the only thing we need to know for constructing a decomposition are the input spaces. Therefore, if we are able to construct a decomposition of F based on the input spaces U_i by first recovering the output spaces $V_i = F(U_i) + F(0)$, verifying that the V_i are indeed subspaces and in direct sum, and then constructing the F_i as $F_{|U_i} + \pi^V_{l \neq i} \circ F(0)$, validating that $F = \sum_i F_i \circ \pi^U_i$ holds, then we say that the input spaces U_i induce a decomposition of F.

Lemma 1 (Induction of Decomposition). *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective. Let further $U_1, ..., U_d$ be non-trivial subspaces of $\mathbb{F}_2^n$ with $\bigoplus_i U_i = \mathbb{F}_2^n$ and let us define $V_i \coloneqq F(U_i) + F(0)$. If the V_i are subspaces with $\bigoplus_i V_i = \mathbb{F}_2^n$ and if*

$$F = \sum_i F \circ \pi_i^U + (d + 1 \bmod 2) \cdot F(0)$$

then $D = \{(U_i, V_i, F_i) \mid 1 \le i \le d\}$ with $F_i \coloneqq F_{|U_i} + \pi^V_{l \neq i} \circ F(0)$ is a decomposition of F. In this case, we say that $\{U_i \mid 1 \le i \le d\}$ induces *the decomposition D.*

Proof. This follows from the observations above and by definition of the F_i, as

$$\sum_i F_i \circ \pi_i^U = \sum_i \left(F_{|U_i} \circ \pi_i^U + \pi^V_{l \neq i} \circ F(0)\right) = \sum_i F \circ \pi_i^U + (d + 1 \bmod 2) \cdot F(0).$$

□

Note that we could have used this as the definition of decomposition. While this does not need the redundant information of the F_i and V_i, it is way less intuitive. Also, when working with decompositions it can be quite useful to have both the F_i and the V_i at hand, too.

Decompositions of Affine Equivalent Functions. We will now show a one-to-one relationship between the decompositions of two affine equivalent functions, which means that the existence of a unique maximal decomposition is actually invariant under affine equivalence.

Lemma 2. *Let $F, G : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and affine equivalent, i. e. $F = A \circ G\,(B + b) + a$ for invertible matrices $A, B \in \mathbb{F}_2^{n \times n}$ and constants $a, b \in \mathbb{F}_2^n$. Then $\{U_i | 1 \le i \le d\}$ induces a decomposition of F if and only if $\{B \cdot U_i | 1 \le i \le d\}$ induces a decomposition of G.*

Proof. Let $\{(U_i, V_i, F_i) \mid 1 \le i \le d\}$ be the decomposition of F induced by $\{U_i | 1 \le i \le d\}$. It obviously holds that $\bigoplus_i B \cdot U_i = \mathbb{F}_2^n$, as B is invertible. In addition, if we denote by I the identity mapping, we get that

$$\begin{aligned} &A \circ G\,(B + b) + a = F = \sum_i F \circ \pi_i^U + (d + 1 \bmod 2) \cdot F(0) \\ =&\sum_i \left(F \circ \pi_i^U + F \circ \pi^U_{l \neq i} \circ B^{-1}(b) + F \circ \pi_i^U \circ B^{-1}(b) + F \circ B^{-1}(b) + F(0)\right) \\ &+ (d + 1 \bmod 2) \cdot F(0), \end{aligned}$$

since we can decompose $F\circ B^{-1}(b)$ into $F\circ\pi^U_{l\neq i}\circ B^{-1}(b)+F\circ\pi^U_i\circ B^{-1}(b)+F(0)$. But, as $\sum_i F\circ\pi^U_i\circ B^{-1}(b) = F\circ B^{-1}(b) + (d+1 \bmod 2)\cdot F(0)$, this is equal to

$$\begin{aligned}
&\sum_i \left(F\circ\pi^U_i + F\circ\pi^U_{l\neq i}\circ B^{-1}(b) + F(0)\right) + (d+1 \bmod 2)\cdot F\circ B^{-1}(b)\\
=&\sum_i F\left(\pi^U_i + \pi^U_{l\neq i}\circ B^{-1}(b)\right) + (d+1 \bmod 2)\cdot F\circ B^{-1}(b)\\
=&\sum_i F\left(\pi^U_i\left(I+B^{-1}(b)\right) + B^{-1}(b)\right) + (d+1 \bmod 2)\cdot F\circ B^{-1}(b)\\
=&\sum_i \left(A\circ G\left(B\circ\pi^U_i\left(I+B^{-1}(b)\right) + B\cdot B^{-1}(b) + b\right) + a\right)\\
&+ (d+1 \bmod 2)\cdot\left(A\circ G\left(B\circ B^{-1}(b)+b\right)+a\right)\\
=&A\left(\sum_i G\left(B\circ\pi^U_i\left(I+B^{-1}(b)\right)\right) + (d+1 \bmod 2)\cdot G(0)\right) + a.
\end{aligned}$$

In other words, we have that

$$G = \sum_i G\left(B\circ\pi^U_i\circ B^{-1}\right) + (d+1 \bmod 2)\cdot G(0),$$

where $B\circ\pi^U_i\circ B^{-1}$ are the projections onto $B\cdot U_i$. In addition, we get that

$$\begin{aligned}
A\left(G(B\cdot U_i)+G(0)\right) =& F(U_i + B^{-1}\cdot b) + F(B^{-1}\cdot b)\\
=&\sum_l F\circ\pi^U_l(U_i+B^{-1}\cdot b) + (d+1 \bmod 2)\cdot F(0)\\
&+\sum_l F\circ\pi^U_l(B^{-1}\cdot b) + (d+1 \bmod 2)\cdot F(0)\\
=&F\circ\pi^U_i(U_i) + F\circ\pi^U_i(B^{-1}\cdot b)\\
=&F_i\circ\pi^U_i(U_i) + \pi^V_{l\neq i}F(0) + F_i\circ\pi^U_i(B^{-1}\cdot b) + \pi^V_{l\neq i}F(0)\\
=&V_i + F_i\circ\pi^U_i(B^{-1}\cdot b) = V_i.
\end{aligned}$$

As $\bigoplus_i A^{-1}\cdot V_i = \mathbb{F}_2^n$, it now follows from Lemma 1 that $\{B\cdot U_i | 1\leq i\leq d\}$ induces a decomposition of G. The reverse direction follows from switching the roles of F and G, as both affine mappings are bijective. □

This especially shows that a decomposition is invariant (up to changing the F_i) under the addition of constants (e. g. a key), both at the beginning or the end of the round function. Hence, we can ignore such additions.

While the lemma above also enables us to study an affine equivalent version of a function such that one decomposition has a preferable form, e. g. $U_i = 0^{m_i}\times\mathbb{F}_2^{\dim(U_i)}\times 0^{m'_i}$, other decompositions are not necessarily of such a form.

Finding Refinements. In order to judge if a decomposition is maximal we need to know if there exists a refinement. For this, let us try to find a refinement given two decompositions.

Corollary 5. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{U_i \mid 1 \le i \le d\}$ and $\{W_i \mid 1 \le i \le e\}$ both induce decompositions of F. Then we have that*

$$F + (d \cdot e + 1 \bmod 2) \cdot F(0) = \sum_{i,j} F \circ \pi_i^U \circ \pi_j^W = \sum_{i,j} F \circ \pi_j^W \circ \pi_i^U .$$

Proof. We know that we can write $F = \sum_i F \circ \pi_i^U + (d+1 \mod 2) \cdot F(0)$ resp. $F = \sum_j F \circ \pi_j^W + (e+1 \mod 2) \cdot F(0)$. This means that

$$F \circ \pi_j^W = \sum_i F \circ \pi_i^U \circ \pi_j^W + (d+1 \bmod 2) \cdot F(0),$$

which in turn means that

$$\begin{aligned} F &= \sum_j F \circ \pi_j^W + (e+1 \bmod 2) \cdot F(0) \\ &= \sum_j \left(\sum_i F \circ \pi_i^U \circ \pi_j^W + (d+1 \bmod 2) \cdot F(0) \right) + (e+1 \bmod 2) \cdot F(0) \\ &= \sum_{i,j} F \circ \pi_i^U \circ \pi_j^W + (d \cdot e + 1 \bmod 2) \cdot F(0). \end{aligned}$$ □

As we will see, if $\operatorname{Im}(\pi_i^U \circ \pi_j^W) \cap \operatorname{Im}(\pi_l^U \circ \pi_k^W) = \{0\}$ holds for all $i \neq l$ and $j \neq k$ then we have found a refinement. But as soon as there exists a non-trivial intersection, there exist multiple maximal decompositions. To see that the case of a non-trivial intersection is even possible, let us look at the following example.

Example 1. Let R be as in Section 2, i. e.

$$R \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} f(x_1) \\ g(x_1) + x_2 \\ x_3 \\ h\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} \text{S-box} \\ \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} \text{S-box} \end{matrix},$$

where $x_1 \in \mathbb{F}_2^n, x_4 \in \mathbb{F}_2^m$ for integers n, m and $x_2, x_3 \in \mathbb{F}_2$, as well as $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$, $g : \mathbb{F}_2^n \to \mathbb{F}_2$ and $h : \mathbb{F}_2 \times \mathbb{F}_2^m \to \mathbb{F}_2^m$. Because of Lemma 2 we can ignore a linear layer that would usually mix the output of the two S-boxes.

Let us consider the following subspaces

$$\begin{aligned} U_1 &:= \mathbb{F}_2^n \times \mathbb{F}_2 \times 0 \times 0^m, & U_2 &:= 0^n \times 0 \times \mathbb{F}_2 \times \mathbb{F}_2^n, \\ W_1 &:= U_1, & W_2 &:= \{(0, a, a, b)^T \mid a \in \mathbb{F}_2, b \in \mathbb{F}_2^m\}. \end{aligned}$$

It is not too hard to see that both $\{U_1, U_2\}$ and $\{W_1, W_2\}$ induce decompositions (we refer to the full version [17] for more details) while they are obviously not identical, nor is one the refinement of the other. But

$$\operatorname{Im}(\pi_1^U \circ \pi_1^W) \cap \operatorname{Im}(\pi_1^U \circ \pi_2^W) = U_1 \cap 0^n \times \mathbb{F}_2 \times 0 \times 0^m = 0^n \times \mathbb{F}_2 \times 0 \times 0^m \neq \{0\}.$$

5.1 The Case of Trivial Intersections

Based on Corollary 5, one may hope that the set $\{\mathrm{Im}(\pi_i^U \circ \pi_j^W) | i, j\} \setminus \{0\}$ induces a decomposition that is a refinement of both initial decompositions. While it is clear that in case of only trivial intersections of those images they form a direct sum, the same is not directly clear for the corresponding output spaces $\mathrm{Im}(\pi_i^V \circ \pi_j^X \circ F) = \mathrm{Im}(\pi_i^V \circ \pi_j^X)$. Obviously, the intersection can only be non-trivial if $i = l$ as otherwise $\mathrm{Im}(\pi_i^U) \cap \mathrm{Im}(\pi_l^U) = \{0\}$ already holds, which reduces our analysis to the spaces $\mathrm{Im}(\pi_i^U \circ \pi_j^W) \cap \mathrm{Im}(\pi_i^U \circ \pi_k^W)$ resp. $\mathrm{Im}(\pi_i^V \circ \pi_j^X) \cap \mathrm{Im}(\pi_i^V \circ \pi_k^X)$. As the next corollary shows, the input spaces are in direct sum if and only if the output spaces are in direct sum.

Corollary 6. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ and $\{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two decompositions of F. Then for all i, j, k, it holds that*

$$Im(\pi_i^U \circ \pi_j^W) \cap Im(\pi_i^U \circ \pi_k^W) = \{0\} \Leftrightarrow Im(\pi_i^V \circ \pi_j^X) \cap Im(\pi_i^V \circ \pi_k^X) = \{0\}.$$

Proof. From Corollaries 2 and 3 we know that $F \circ \pi_i^U = F_i \circ \pi_i^U + \pi_{l\neq i}^V \circ F(0) = \pi_i^V \circ F + \pi_{l\neq i}^V F(0)$, which means that

$$\begin{aligned} F \circ \pi_i^U \circ \pi_j^W =& \pi_i^V \circ F \circ \pi_j^W + F(0) + \pi_i^V \circ F(0) \\ =& \pi_i^V \left(\pi_j^X \circ F + F(0) + \pi_j^X \circ F(0)\right) + F(0) + \pi_i^V \circ F(0) \\ =& \pi_i^V \circ \pi_j^X \circ F + F(0) + \pi_i^V \circ \pi_j^X \circ F(0) \\ =& \pi_i^V \circ \pi_j^X \left(F + F(0)\right) + F(0). \end{aligned}$$

Let $x \in \mathrm{Im}(\pi_i^U \circ \pi_j^W) \cap \mathrm{Im}(\pi_i^U \circ \pi_k^W)$, i. e. there exist $a, b \in \mathbb{F}_2^n$ such that $\pi_i^U \circ \pi_j^W(a) = x = \pi_i^U \circ \pi_k^W(b)$. This means that

$$F(x) + F(0) = F \circ \pi_i^U \circ \pi_j^W(a) + F(0) = \pi_i^V \circ \pi_j^X \left(F(a) + F(0)\right)$$

and similarly $F(x) + F(0) = \pi_i^V \circ \pi_k^X \left(F(b) + F(0)\right)$, i. e. there exists $c = F(a) + F(0)$ and $d = F(b) + F(0)$ such that $\pi_i^V \circ \pi_j^X(c) = F(x) + F(0) = \pi_i^V \circ \pi_k^X(d)$. As F is a bijection, implying that $F(x) + F(0) = 0$ if and only if $x = 0$ and that the mappings $a \mapsto c$ and $b \mapsto d$ are also bijections, the claim follows. □

Next, we want to show that if $\mathrm{Im}(\pi_i^U \circ \pi_j^W) \cap \mathrm{Im}(\pi_i^U \circ \pi_k^W) = \{0\}$ holds for for all i and $j \neq k$, we either get a refinement or the two decompositions are identical. To do so, we first take a deeper look at compositions of the projections.

Corollary 7. *Let $U_1, ..., U_d$ and $W_1, ..., W_e$ be subspaces of $\mathbb{F}_2^n$ such that $\bigoplus_i U_i = \mathbb{F}_2^n = \bigoplus_i V_i$. Then $Im(\pi_i^U \circ \pi_j^W) \cap Im(\pi_i^U \circ \pi_k^W) = \{0\}$ hold for for all i and $j \neq k$ if and only if $\pi_i^U \circ \pi_j^W = \pi_j^W \circ \pi_i^U$ for all i, j.*

Proof. Let us assume that $\mathrm{Im}(\pi_i^U \circ \pi_j^W) \cap \mathrm{Im}(\pi_i^U \circ \pi_k^W) = \{0\}$ holds for for all i and $j \neq k$. Hence, we know that $\bigoplus_{i,j} \mathrm{Im}(\pi_i^U \circ \pi_j^W) = \mathbb{F}_2^n$, and we have to show that $\pi_i^U \circ \pi_j^W = \pi_j^W \circ \pi_i^U$ for all i, j. Since both $\sum_{l,k} \pi_l^U \circ \pi_k^W$ and $\sum_k \pi_k^W$ are the identity, it holds that

$$0 = \pi_i^U + \pi_i^U = \sum_{l,k} \pi_l^U \circ \pi_k^W \circ \pi_i^U + \sum_k \pi_i^U \circ \pi_k^W \circ \pi_i^U = \sum_{l \neq i,k} \pi_l^U \circ \pi_k^W \circ \pi_i^U .$$

As $0 \in \operatorname{Im}(\pi_l^U \circ \pi_k^W)$ for all l, k, and those images are in direct sum, we know that $\pi_l^U \circ \pi_k^W \circ \pi_i^U = 0$ has to hold for all $l \neq i$ and all k. This shows that $\pi_j^W \circ \pi_i^U = \pi_i^U \circ \pi_j^W$, since

$$\pi_j^W \circ \pi_i^U = \sum_l \pi_l^U \circ \pi_j^W \circ \pi_i^U = \pi_i^U \circ \pi_j^W \circ \pi_i^U ,$$

but also

$$\pi_i^U \circ \pi_j^W = \sum_l \pi_i^U \circ \pi_j^W \circ \pi_l^U = \pi_i^U \circ \pi_j^W \circ \pi_i^U .$$

Now, let us assume that $\pi_i^U \circ \pi_j^W = \pi_j^W \circ \pi_i^U$ for all i, j, which means that

$$\pi_i^U \circ \pi_j^W \circ \pi_i^U \circ \pi_j^W = \pi_i^U \circ \pi_j^W \circ \pi_j^W \circ \pi_i^U = \pi_i^U \circ \pi_j^W \circ \pi_i^U = \pi_i^U \circ \pi_i^U \circ \pi_j^W = \pi_i^U \circ \pi_j^W ,$$

i. e. the $\pi_i^U \circ \pi_j^W$ are projections. If $k \neq j$ then we have that

$$\pi_i^U \circ \pi_j^W \circ \pi_l^U \circ \pi_k^W = \pi_i^U \circ \pi_j^W \circ \pi_k^W \circ \pi_l^U = 0.$$

But also, since $\operatorname{Im}(\pi_i^U \circ \pi_j^W) \cap \operatorname{Im}(\pi_l^U \circ \pi_k^W) \subset \operatorname{Im}(\pi_i^U \circ \pi_j^W)$ and $\pi_i^U \circ \pi_j^W$ is a projection, it has to hold that

$$\begin{aligned}\{0\} \subset \operatorname{Im}(\pi_i^U \circ \pi_j^W) \cap \operatorname{Im}(\pi_l^U \circ \pi_k^W) =& \pi_i^U \circ \pi_j^W \left(\operatorname{Im}(\pi_i^U \circ \pi_j^W) \cap \operatorname{Im}(\pi_l^U \circ \pi_k^W)\right) \\ \subset& \pi_i^U \circ \pi_j^W \left(\operatorname{Im}(\pi_l^U \circ \pi_k^W)\right) = \{0\}. \qquad \square\end{aligned}$$

In other words, if all the intersections are trivial, we not only know that we can find a refinement based on the two decompositions, but also that the order in which we do, first decomposing according to the U_i and then refining according to the W_j or the other way around, does not matter since the resulting projections and therefore the subspaces are identical. This leads us to the following lemma.

Lemma 3. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{U_i \mid 1 \leq i \leq d\}$ and $\{W_i \mid 1 \leq i \leq e\}$ induce the decompositions D and E of F. If for every i and $j \neq k$ it holds that $Im(\pi_i^U \circ \pi_j^W) \cap Im(\pi_i^U \circ \pi_k^W) = \{0\}$ then we can either construct a refinement of D or E induced by the set of input spaces $\{U_i \cap W_j | 1 \leq i \leq d, 1 \leq j \leq e\} \setminus \{\{0\}\}$ or $D = E$.*

Proof. By our reasoning above, we already know that if $\operatorname{Im}(\pi_i^U \circ \pi_j^W) \cap \operatorname{Im}(\pi_i^U \circ \pi_k^W) = \{0\}$ holds for every i and $j \neq k$ the images $\operatorname{Im}(\pi_i^U \circ \pi_j^W)$ form a direct sum, as do the images $\operatorname{Im}(\pi_i^V \circ \pi_j^X)$. Note that Corollary 7 shows that $\pi_i^U \circ \pi_j^W = \pi_j^W \circ \pi_i^U$ are the corresponding projections. Therefore, $U_i \supset \operatorname{Im}(\pi_i^U \circ \pi_j^W) =$

$\mathrm{Im}(\pi_j^W \circ \pi_i^U) \subset W_j$, but also $U_i \cap W_j \subset \mathrm{Im}(\pi_j^W)$ and $U_i \cap W_j \subset \mathrm{Im}(\pi_i^U)$, which means that

$$U_i \cap W_j \subset \mathrm{Im}(\pi_i^U \circ \pi_j^W) \subset U_i \cap W_j.$$

Hence, $\mathrm{Im}(\pi_i^U \circ \pi_j^W) = U_i \cap W_j$ and if we remove the trivial subspaces then Corollary 5 shows that the set of input spaces $\{U_i \cap W_j | 1 \leq i \leq d, 1 \leq j \leq e\} \setminus \{\{0\}\}$ induces a decomposition. Also, it obviously holds that $W_j \cap U_i \subset W_j$ and $U_i \cap W_j \subset U_i$, which means that we either got a refinement of D or E, or D is identical to E. □

5.2 The Case of Non-Trivial Intersections

Now, let us look at the case in which at least one intersection of the images is non-trivial. As we will see, this means that (at least) one S-Box has to have maximal differential uniformity and another one has to have maximal linearity. In order to show that, we need the following lemma.

Lemma 4. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ and $\{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two decompositions of F. Then we have that*

$$F\left(\sum_j \pi_i^U \circ \pi_j^W\right) = \sum_j F \circ \pi_i^U \circ \pi_j^W + (e + 1 \bmod 2) \cdot F(0).$$

Proof. The claim follows from

$$\begin{aligned}
F\left(\sum_j \pi_i^U \circ \pi_j^W\right) &= F \circ \pi_i^U \left(\sum_j \pi_j^W\right) = F \circ \pi_i^U = \pi_i^V \circ F + \pi_{l\neq i}^V \circ F(0)\\
=&\pi_i^V \left(\sum_j F \circ \pi_j^W + (e+1 \bmod 2) \cdot F(0)\right) + \pi_{l \neq i}^V \circ F(0)\\
=&\sum_j \pi_i^V \circ F \circ \pi_j^W + (e + 1 \bmod 2) \cdot \pi_i^V \circ F(0) + \pi_{l\neq i}^V \circ F(0)\\
=&\sum_j \left(F \circ \pi_i^U \circ \pi_j^W + \pi_{l\neq i}^V \circ F(0)\right) + (e+1 \bmod 2) \cdot \pi_i^V \circ F(0) + \pi_{l\neq i}^V \circ F(0)\\
=&\sum_j F \circ \pi_i^U \circ \pi_j^W + (e+1 \bmod 2) \cdot F(0).
\end{aligned}$$

□

Since the inputs to the $F \circ \pi_i^U \circ \pi_j^W$ can be seen as independent as the W_j are in direct sum, this already shows that there has to be some kind of linearity.

Behaviour on the Intersections. Next, we will show that each F_i (resp. the corresponding S-box) has to be affine on each of the subspaces $\mathrm{Im}(\pi_i^U \circ \pi_j^V) \cap \mathrm{Im}(\pi_i^U \circ \pi_k^V)$, which is trivial in the case that those subspaces are $\{0\}$, but gives us more information in the case of a non-trivial intersection. For this, we will first take a deeper look at such intersections.

Corollary 8. *Let $U_1, ..., U_d, W_1, ..., W_e \subset \mathbb{F}_2^n$ be subspaces with $\bigoplus_i U_i = \mathbb{F}_2^n = \bigoplus_i W_i$. Also, let P be a composition of alternating projections π_i^U, π_j^W. Then it holds that*

$$Im(P \circ \pi_j^W) \cap Im(P \circ \pi_{k\neq j}^W) = Im(P \circ \pi_j^W \circ \pi_{l\neq i}^U) = Im(P \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U).$$

Proof. Note that as $\pi_i^W \circ \pi_{k\neq j}^W = 0$, we can assume that P is either of the form $p \circ ... \circ \pi_i^U \circ \pi_j^W \circ \pi_i^U$ with $p \in \{\pi_i^U \circ \pi_j^W, \pi_j^W \circ \pi_i^U\}$ or $P \circ \pi_{k\neq j}^W$ is zero. As in the second case the claim is obviously true, let us assume that P is of the form $p \circ ... \circ \pi_i^U \circ \pi_j^W \circ \pi_i^U$. It holds that

$$\begin{aligned}
&\mathrm{Im}(P \circ \pi_j^W) \cap \mathrm{Im}(P \circ \pi_{k\neq j}^W)\\
=&\{P(x) | x \in W_j\} \cap \{P(x') | x' \in W_{k\neq j}\}\\
=&\{y | P(x) = y = P(x'), x \in W_j, x' \in W_{k\neq j}\}\\
=&\{P \circ \pi_j^W(x + x') | P(x) + P(x') = 0, x \in W_j, x' \in W_{k\neq j}\}\\
=&\{P \circ \pi_j^W(\hat{x}) | \hat{x} \in \ker(P)\}\\
\supset&\mathrm{Im}(P \circ \pi_j^W \circ \pi_{l\neq i}^U)
\end{aligned}$$

as $U_{l\neq i} = \ker(\pi_i^U) \subset \ker(P)$. In addition, we have that

$$\begin{aligned}
\{P \circ \pi_j^W(\hat{x}) | \hat{x} \in \ker(P)\} =&\{P \circ \pi_j^W(\hat{x}) + p \circ P(\hat{x}) | \hat{x} \in \ker(P)\}\\
=&\{P\left(\pi_j^W(\hat{x}) + \pi_j^W \circ \pi_i^U(\hat{x})\right) | \hat{x} \in \ker(P)\}\\
=&\{P \circ \pi_j^W \circ \pi_{l\neq i}^U(\hat{x}) | \hat{x} \in \ker(P)\}\\
\subset&\mathrm{Im}(P \circ \pi_j^W \circ \pi_{l\neq i}^U),
\end{aligned}$$

which, together with $\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U = \pi_i^U \circ \left(I + \pi_j^W\right) \circ \pi_{l\neq i}^U = \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U$, where I denotes the identity, completes the proof. □

In other words, if we set $P = \pi_i^U$, the intersections $\mathrm{Im}(\pi_i^U \circ \pi_j^V) \cap \mathrm{Im}(\pi_i^U \circ \pi_{k\neq j}^V)$ are actually the images of $\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U = \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U$. With this and the lemma above, we can now show that the function has to be affine on those intersection resp. images.

Lemma 5. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ and $\{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two decompositions of F. Then we have that F is affine on $Im(\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U)$ for every i, j.*

Proof. Let $x, y \in \mathrm{Im}(\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U) = \mathrm{Im}(\pi_i^U \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U)$, i.e. there exist $a, b \in \mathbb{F}_2^n$ such that $x = \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a)$ and $y = \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U(b)$. Also, let us define $\hat{W}_1 \coloneqq W_j$ and $\hat{W}_2 \coloneqq W_{k\neq j}$. Note that $\{\hat{W}_1, \hat{W}_2\}$ also induces a decomposition, which, combined with the above lemma, leads to

$$\begin{aligned}F(x+y) =& F\left(\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a) + \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U(b)\right)\\ =& F\left(\sum_{r=1,2} \pi_i^U \circ \pi_r^{\hat{W}}\left(\pi_j^W \circ \pi_{l\neq i}^U(a) + \pi_{k\neq j}^W \circ \pi_{l\neq i}^U(b)\right)\right)\\ =& F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a) + F \circ \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_{l\neq i}^U(b) + F(0)\\ =& F(x) + F(y) + F(0). \qquad \square\end{aligned}$$

A direct consequence of this is that $F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U$ is affine, or equivalent that $F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U + F(0)$ is linear.[10] But as we can iteratively represent F composed with a projection onto an input space as the projection of F onto the corresponding output space (plus a constant), this gives us the following.

Corollary 9. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ and $\{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two decompositions of F. If $\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U \neq 0$ then there exists an $k \neq i$ such that F_k has maximal linearity.*

Proof. We know that $F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U + F(0)$ is linear. But this means that

$$\begin{aligned}F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U + F(0) =& \pi_i^V\left(F \circ \pi_j^W \circ \pi_{l\neq i}^U + F(0)\right)\\ &= \ldots = \pi_i^V \circ \pi_j^X \circ \pi_{l\neq i}^V\left(F + F(0)\right)\\ &= \sum_{l\neq i} \pi_i^V \circ \pi_j^X \circ \pi_l^V\left(F_l \circ \pi_l^U + F_l(0)\right)\end{aligned}$$

is also linear. Since the inputs to the F_l are independent, this shows that $\pi_i^V \circ \pi_j^X \circ \pi_l^V\left(F_l \circ \pi_l^U + F_l(0)\right)$ is linear for every $l \neq i$. As $\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U \neq 0$, we also know that $\pi_i^V \circ \pi_j^X \circ \pi_{l\neq i}^V \neq 0$, and therefore there exists a $k \neq i$ such that $\pi_i^V \circ \pi_j^X \circ \pi_k^V \neq 0$, which in turn means that $\pi_i^V \circ \pi_j^X \circ \pi_k^V\left(F_k + F_k(0)\right) \neq 0$. Hence, we can simply select α^T as a non-zero component of $\pi_i^V \circ \pi_j^X \circ \pi_k^V$ and get that $\alpha^T \cdot F_k + \alpha^T \cdot F_k(0)$ is linear and non-trivial, i. e. F_k has maximal linearity. □

In other words, a non-trivial intersection implies a (non-trivial) affine component of one of the S-boxes. But it even implies more.

Non-Trivial Intersections Imply Maximal Differential Uniformity. Knowing that F is affine on these intersections, we are now able to show that this implies maximal differential uniformity of the corresponding F_i.

Lemma 6. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $\{(U_i, V_i, F_i) \mid 1 \leq i \leq d\}$ and $\{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two decompositions of F. If it holds that $Im(\pi_i^U \circ \pi_j^W) \cap Im(\pi_i^U \circ \pi_{k\neq j}^W) \neq \{0\}$ for some i, j then F_i has maximal differential uniformity.*

[10] Note that in the case of trivial intersections, we have that $\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U = \pi_j^W \circ \pi_i^U \circ \pi_{l\neq i}^U = 0$, which means that $F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U + F(0) = 0$

Proof. Let $\alpha \in \mathrm{Im}(\pi_i^U \circ \pi_j^W) \cap \mathrm{Im}(\pi_i^U \circ \pi_{k\neq j}^W) = \mathrm{Im}(\pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U)$ and $x \in U_i = \mathrm{Im}(\pi_i^U \circ \pi_j^W \circ \pi_i^U) + \mathrm{Im}(\pi_i^U \circ \pi_{k\neq j}^W \circ \pi_i^U)$, i. e. we can write $x = x_1 + x_2$ for $x_1 \in \mathrm{Im}(\pi_i^U \circ \pi_j^W \circ \pi_i^U)$ and $x_2 \in \mathrm{Im}(\pi_i^U \circ \pi_{k\neq j}^W \circ \pi_i^U)$. Furthermore, we can find a, b_1, b_2 such that $\alpha = \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a)$, $x_1 = \pi_i^U \circ \pi_j^W \circ \pi_i^U(b_1)$ and $x_2 = \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_i^U(b_2)$, which, together with Lemma 4, gives us

$$\begin{aligned}
&F(x+\alpha) = F(x_1 + x_2 + \alpha)\\
=&F\left(\pi_i^U \circ \pi_j^W \left(\pi_{l\neq i}^U(a) + \pi_i^U(b_1)\right) + \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_i^U\,(b_2)\right)\\
=&F \circ \pi_i^U \circ \pi_j^W \left(\pi_{l\neq i}^U(a) + \pi_i^U(b_1)\right) + F \circ \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_i^U\,(b_2) + F(0),
\end{aligned}$$

where for the last step we use the fact that $\pi_j^W(\hat{a}) = \pi_j^W\left(\pi_j^W(\hat{a}) + \pi_{k\neq j}^W(\hat{b})\right)$ and $\pi_{k\neq j}^W(\hat{b}) = \pi_{k\neq j}^W\left(\pi_j^W(\hat{a}) + \pi_{k\neq j}^W(\hat{b})\right)$. In addition, it is easy to see that that

$$\begin{aligned}
&F \circ \pi_i^U \circ \pi_j^W \left(\pi_{l\neq i}^U(a) + \pi_i^U(b_1)\right)\\
=&\pi_i^V \circ \pi_j^X \circ F\left(\pi_{l\neq i}^U(a) + \pi_i^U(b_1)\right) + F(0) + \pi_i^V \circ \pi_j^X \circ F(0)\\
=&\pi_i^V \circ \pi_j^X \left(F \circ \pi_{l\neq i}^U(a) + F \circ \pi_i^U(b_1) + F(0)\right) + F(0) + \pi_i^V \circ \pi_j^X \circ F(0)\\
=&F \circ \pi_i^U \circ \pi_j^W \circ \pi_i^U(b_1) + F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a) + F(0).
\end{aligned}$$

If we combine those observations and apply Lemma 4 once more, we get that

$$\begin{aligned}
&F(x+\alpha)\\
=&F \circ \pi_i^U \circ \pi_j^W \circ \pi_i^U(b_1) + F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a) + F(0)\\
&+ F \circ \pi_i^U \circ \pi_{k\neq j}^W \circ \pi_i^U\,(b_2) + F(0)\\
=&F \circ \pi_i^U \left(\pi_j^W \circ \pi_i^U(b_1) + \pi_{k\neq j}^W \circ \pi_i^U\,(b_2)\right) + F \circ \pi_i^U \circ \pi_j^W \circ \pi_{l\neq i}^U(a) + F(0)\\
=&F(x) + F(\alpha) + F(0).
\end{aligned}$$

In other words, $F(x) + F(x+\alpha) = F(\alpha) + F(0)$ holds for every $x \in U_i$, which means that $F_{|U_i}$ and therefore F_i has maximal differential uniformity. □

Obviously, if one of the F_i has maximal differential uniformity, the same has to be true for the whole (round) function.

Maximal Differential Uniformity Together with Maximal Linearity Implies Non-Unique Maximal Decomposition. Next, we will show that one S-box having maximal differential uniformity and another one having maximal linearity implies that there exists no unique maximal decomposition.

Lemma 7. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be bijective and let $D = \{(U_i, V_i, F_i) | 1 \leq i \leq d\}$ be a maximal decomposition of F. If there exist $i \neq k$ such that F_i has maximal differential uniformity and F_k has maximal linearity then D is not unique.*

Proof. Let us assume that there exists $i \neq k$ such that F_i has maximal differential uniformity, i. e. we can find an $a \in U_i$ such that $F_i(x) + F_i(x+a)$ is constant for all $x \in U_i$, and therefore the same as $F_i(a)+F_i(0)$, and F_k has maximal linearity, i. e. there exists an $\alpha \in \mathbb{F}_2^n \setminus V_k^\perp$ and $\beta \in \mathbb{F}_2^n$ such that $\beta^T \cdot x + \alpha^T \cdot F_k(x)$ is constant for all $x \in U_k$, and hence the same as $\alpha^T \cdot F_k(0)$. Let us define $L : \mathbb{F}_2^n \to U_i$ by $L \cdot x := a \cdot \beta^T \cdot \pi_k^U(x)$ for all $x \in \mathbb{F}_2^n$. Then both $\pi_i^U + L$ and $\pi_k^U + L$ are projections, as $\pi_i^U \circ L = L = L \circ \pi_k^U$, $L \circ \pi_i^U = 0 = \pi_k^U \circ L$ and $L^2 = 0$. In addition, their sum is obviously $\pi_i^U + \pi_k^U$, which means that their images together with U_l for $l \notin \{i,k\}$ form a direct sum of $\mathbb{F}_2^n$. Also, as $\operatorname{Im}\left(\pi_i^U + L\right) = U_i + a = U_i$ we have that $F(\operatorname{Im}(\pi_i^U + L)) + F(0) = F(U_i) + F(0) = V_i$ is, by definition, a subspace. To show that $F(\operatorname{Im}(\pi_k^U + L)) + F(0)$ is also a subspace, let us define the projection $P := \pi_k^V + F_i \circ L \circ F_k^{-1} \circ \pi_k^V + F_i(0)$. Note that P is linear, as $F_i(x+a) = F_i(x) + F_i(a) + F_i(0)$ for all $x \in U_i$. Furthermore, P is indeed a projection, since $\pi_k^V \left(\pi_k^V + F_i \circ L \circ F_k^{-1} \circ \pi_k^V + F_i(0)\right) = \pi_k^V$ and therefore $P^2 = P$. We now get that

$$\begin{aligned} F(\operatorname{Im}(\pi_k^U + L)) + F(0) =& \{F\left(\pi_k^U(x) + L \cdot x\right) + F(0) | x \in \mathbb{F}_2^n\} \\ =& \left\{\sum_l \left(F_l \circ \pi_l^U \left(\pi_k^U(x) + L \cdot x\right) + F_l(0)\right) | x \in U_k\right\} \\ =& \{F_k \circ \pi_k^U(x) + F_k(0) + F_i \circ L \cdot x + F_i(0) | x \in U_k\} \\ =& \{y + F_i \circ L \cdot F_k^{-1}(y + F_k(0)) + F_i(0) | y \in V_k\} \\ =& \operatorname{Im}(P), \end{aligned}$$

which means that $F(\operatorname{Im}(\pi_k^U + L)) + F(0)$ is also a subspace. At last, since

$$\begin{aligned} F\left(\pi_i^U + L\right) + F\left(\pi_k^U + L\right) =& \sum_l F_l \circ \pi_l^U \left(\pi_i^U + L\right) + \sum_l F_l \circ \pi_l^U \left(\pi_k^U + L\right) \\ =& F_i\left(\pi_i^U + L\right) + F_k(0) + F_i \circ L + F_k \circ \pi_k^U \\ =& F_i \circ \pi_i^U + F_i \circ L + F_i(0) + F_i \circ L + F_k(0) + F_k \circ \pi_k^U \\ =& F \circ \pi_i^U + F \circ \pi_k^U, \end{aligned}$$

we get that $\sum_{l \notin \{i,k\}} F \circ \pi_l^U + F\left(\pi_i^U + L\right) + F\left(\pi_k^U + L\right) + (d+1 \mod 2) \cdot F(0) = F$. Therefore, $\{U_l | l \neq k\} \cup \operatorname{Im}\left(\pi_k^U + L\right)$ induces a decomposition of F, different than D but with the same number of S-boxes, which means that D is not unique. □

If we combine all the observations above, we can finally prove Theorem 1.

Proof of Theorem 1. First, let $\{(U_i, V_i, F_i) | 1 \leq i \leq d\}$ be a maximal decomposition of F and assume that there exist $i \neq k$ such that F_i has maximal differential uniformity and F_k has maximal linearity. Then we know from Lemma 7 that this maximal decomposition is not unique.

Now, let $\{(U_i, V_i, F_i) | 1 \leq i \leq d\}$ and $\{(W_i, X_i, G_i) \mid 1 \leq i \leq e\}$ be two different maximal decompositions. We know from Lemma 3 that if all the intersections $\operatorname{Im}(\pi_i^U \circ \pi_j^W) \cap \operatorname{Im}(\pi_i^U \circ \pi_k^W)$ were trivial, we could either refine one of the decompositions or they are identical. But since they are both maximal and not identical, we know that some of those intersection have to be non-trivial. Hence, we can follow from Corollary 9 and Lemma 6 that there exists a pair $i \neq k$ such that F_i has maximal differential uniformity and F_k has maximal linearity. □

6 Conclusion

In this paper we discussed the uniqueness of decompositions, as well as the impact of alignment. With respect to the uniqueness of decomposition, we have seen that this very natural and simple question required quite some technical backup to be finally settled. In our opinion, it is of interest to further explore the possible impact of such non-unique decompositions not only on security and security arguments, but also on implementation aspects.

With respect to the impact of alignment, we show, based on the example of aligned and unaligned versions of PRESENT, that the impact of alignment on its own may only be limited. We therefore encourage further work to either find conditions under which alignment has a meaningful impact, or to show that the impact of plain alignment is insignificant in general. Especially, we would like to advocate for further case studies to try to only change the property under scrutiny, as this produces more convincing results. Another future direction is to develop more fine grained notions that do capture the structural alignment in a non-binary manner.

Finally, and this can be seen as the broader scope, we encourage research that investigates the use of representations of a cipher in arguments for its security.

Acknowledgments. This work was funded by the by the project *Analysis and Protection of Lightweight Cryptographic Algorithms* (432878529) and by DFG (German Research Foudation), under Germany's Excellence Strategy - EXC 2092 CASA – 390781972.

References

1. Alejandro Cabrera Aldaya, Cesar Pereida García, and Billy Bob Brumley. From A to Z: projective coordinates leakage in the wild. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2020(3):428–453, 2020.
2. Anubhab Baksi, Shivam Bhasin, Jakub Breier, Mustafa Khairallah, Thomas Peyrin, Sumanta Sarkar, and Siang Meng Sim. DEFAULT: cipher level resistance against differential fault attack. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT 2021*, volume 13091 of *LNCS*, pages 124–156. Springer, 2021.
3. Christof Beierle, Anne Canteaut, Gregor Leander, and Yann Rotella. Proving resistance against invariant attacks: How to choose the round constants. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017*, volume 10402 of *LNCS*, pages 647–678. Springer, 2017.
4. Christof Beierle, Gregor Leander, Amir Moradi, and Shahram Rasoolzadeh. CRAFT: lightweight tweakable block cipher with efficient protection against DFA attacks. *IACR Trans. Symmetric Cryptol.*, 2019(1):5–45, 2019.
5. Guido Bertoni, Joan Daemen, Michael Peeters, and Gilles Van Assche. On alignment in keccak. In *ECRYPT II Hash Workshop*, volume 51, page 122, 2011.

6. Alex Biryukov and Adi Shamir. Structural cryptanalysis of SASAS. In Birgit Pfitzmann, editor, *EUROCRYPT 2001*, volume 2045 of *LNCS*, pages 394–405. Springer, 2001.
7. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: an ultra-lightweight block cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *CHES 2007*, volume 4727 of *LNCS*, pages 450–466. Springer, 2007.
8. Nicolas Bordes, Joan Daemen, Daniël Kuijsters, and Gilles Van Assche. Thinking outside the superbox. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021*, volume 12827 of *LNCS*, pages 337–367. Springer, 2021.
9. Anne Canteaut, Sébastien Duval, Gaëtan Leurent, María Naya-Plasencia, Léo Perrin, Thomas Pornin, and André Schrottenloher. Saturnin: a suite of lightweight symmetric algorithms for post-quantum security. *IACR Trans. Symmetric Cryptol.*, 2020(S1):160–207, 2020.
10. Claude Carlet. *Boolean Functions for Cryptography and Coding Theory.* Cambridge University Press, 2021.
11. Joan Daemen, Pedro Maat Costa Massolino, Alireza Mehrdad, and Yann Rotella. The subterranean 2.0 cipher suite. *IACR Trans. Symmetric Cryptol.*, 2020(S1):262–294, 2020.
12. Joan Daemen and Vincent Rijmen. Plateau characteristics. *IET Inf. Secur.*, 1(1):11–17, 2007.
13. Maria Eichlseder and Daniel Kales. Clustering related-tweak characteristics: Application to mantis-6. *IACR Transactions on Symmetric Cryptology*, 2018(2):111–132, Jun. 2018.
14. Antonio Flórez-Gutiérrez and María Naya-Plasencia. Improving key-recovery in linear attacks: Application to 28-round PRESENT. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020*, volume 12105 of *LNCS*, pages 221–249. Springer, 2020.
15. Mathias Hall-Andersen and Philip S. Vejre. Generating graphs packed with paths estimation of linear approximations and differentials. *IACR Trans. Symmetric Cryptol.*, 2018(3):265–289, 2018.
16. André Kündgen, Gregor Leander, and Carsten Thomassen. Switchings, extensions, and reductions in central digraphs. *J. Comb. Theory, Ser. A*, 118(7):2025–2034, 2011.
17. Baptiste Lambin, Gregor Leander, and Patrick Neumann. Pitfalls and shortcomings for decompositions and alignment (full version). Cryptology ePrint Archive, Paper 2023/240, 2023. `https://eprint.iacr.org/2023/240`.
18. Gregor Leander. On linear hulls, statistical saturation attacks, PRESENT and a cryptanalysis of PUFFIN. In Kenneth G. Paterson, editor, *EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 303–322. Springer, 2011.
19. Gregor Leander and Shahram Rasoolzadeh. Two sides of the same coin: Weak-keys and more efficient variants of CRAFT. *IACR Cryptol. ePrint Arch.*, page 238, 2021.
20. Guozhen Liu, Weidong Qiu, and Yi Tu. New techniques for searching differential trails in keccak. *IACR Transactions on Symmetric Cryptology*, 2019(4):407–437, Jan. 2020.
21. Ciaran McCreesh, Patrick Prosser, and James Trimble. The glasgow subgraph solver: Using constraint programming to tackle hard subgraph isomorphism problem variants. In Fabio Gadducci and Timo Kehrer, editors, *ICGT 2020*, volume 12150 of *LNCS*, pages 316–324. Springer, 2020.

22. Brice Minaud, Patrick Derbez, Pierre-Alain Fouque, and Pierre Karpman. Key-recovery attacks on ASASA. In Tetsu Iwata and Jung Hee Cheon, editors, *ASIACRYPT 2015*, volume 9453 of *LNCS*, pages 3–27. Springer, 2015.
23. Kaisa Nyberg. Differentially uniform mappings for cryptography. In Tor Helleseth, editor, *EUROCRYPT '93*, volume 765 of *LNCS*, pages 55–64. Springer, 1993.
24. Tiago B. S. Reis, Diego F. Aranha, and Julio César López-Hernández. PRESENT runs fast - efficient and secure implementation in software. In Wieland Fischer and Naofumi Homma, editors, *Cryptographic Hardware and Embedded Systems - CHES 2017 - 19th International Conference, Taipei, Taiwan, September 25-28, 2017, Proceedings*, volume 10529 of *LNCS*, pages 644–664. Springer, 2017.
25. Claude E Shannon. A mathematical theory of cryptography. *Mathematical Theory of Cryptography*, 1945.
26. Ling Song, Zhangjie Huang, and Qianqian Yang. Automatic differential analysis of ARX block ciphers with application to SPECK and LEA. In Joseph K. Liu and Ron Steinfeld, editors, *Information Security and Privacy - 21st Australasian Conference, ACISP 2016, Melbourne, VIC, Australia, July 4-6, 2016, Proceedings, Part II*, volume 9723 of *LNCS*, pages 379–394. Springer, 2016.