Algebraic Degree of Strong-Aligned SP-Networks with Low-Degree and Large S-Boxes

Carlos Cid[1,4], Lorenzo Grassi[2], Reinhard Lüftenegger[3], Christian Rechberger[3], and Markus Schofnegger[3]

[1] Information Security Group, Royal Holloway University of London
carlos.cid@rhul.ac.uk
[2] Digital Security Group, Radboud University, Nijmegen
lgrassi@science.ru.nl
[3] IAIK, Graz University of Technology
firstname.lastname@iaik.tugraz.at
[4] Simula UiB, Norway

Abstract. Higher-order differential cryptanalysis and its variants are among the most powerful methods for analyzing iterated cryptographic permutations and hash functions with low algebraic degree over binary extension fields. Predicting the evolution of the algebraic degree (as a function of the number of iterations) is the main obstacle for applying these methods. In this paper, we present a new upper bound on the growth of the algebraic degree in strong-aligned SP-Networks with low-degree and large S-Boxes. Our findings generalize a recent result presented at Asiacrypt 2020, which applies to permutations based on an iterated Even-Mansour construction with low-degree round functions. As a main result, we prove that an initial exponential growth of the algebraic degree is followed by a linear growth until the maximum algebraic degree is reached. Our analysis is particularly relevant for assessing the security of cryptographic permutations designed to be competitive in applications like MPC, FHE, SNARKs, and STARKs, including permutations based on the Hades design strategy. We have verified our findings on small-scale instances and we have compared them against the current best results, showing a substantial improvement for strong-aligned SPN schemes with low-degree and large S-Boxes.

Keywords: Higher-Order Differential Cryptanalysis – SPN – Algebraic Degree

1 Introduction

Motivated by new applications such as secure Multi-Party Computation (MPC), Fully Homomorphic Encryption (FHE) and Zero-Knowledge proofs (ZKP), the need for symmetric encryption schemes with a simple natural algebraic description has become ever more apparent. This is an active area of research, and many dedicated symmetric encryption schemes that aim for simple arithmetization or directly aim for a small number of multiplications in $\mathbb{F}_{2^n}$ or $\mathbb{F}_p$, for large n and prime p (usually, $2^n, p \approx 2^{128}$), have recently been proposed in the literature. They include permutations, block ciphers, and hash functions such as MiMC [3,29], GMiMC [2], HadesMiMC [28] (and its hash variant Poseidon [27]), Jarvis & Friday [6], Vision & Rescue [5] and Ciminion [23]. Many of these proposed schemes with simplified algebraic description use S-Boxes based on a power mapping $x \mapsto x^d$ for a small odd integer $d \geq 3$, or the inversion $x \mapsto x^{-1}$. Due to their particular design based on large-size low-degree S-Boxes, it is likely that some (especially, algebraic) attacks can perform better than what we can expect for "classical" symmetric schemes based on small-size and/or high-degree S-Boxes (e.g. AES-like or Keccak-like designs), as already shown in [1,12]. In other words, these constructions can be considered naturally more vulnerable to algebraic attacks than those which do not exhibit such a clear and simple algebraic structure.

Focusing on the binary case, one of the most powerful cryptanalytic methods for symmetric primitives over $\mathbb{F}_2^n$ with low-degree building blocks is higher-order differential cryptanalysis. In essence, this method allows to distinguish a given Boolean permutation from a random one. More precisely, given an instance of a (keyed or keyless) cryptographic permutation $P : \mathbb{F}_2^n \to \mathbb{F}_2^n$, higher-order differential cryptanalysis exploits the fact that if the algebraic degree of P is strictly smaller than $n-1$ then for any (proper) vector subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$ with dimension strictly greater than the algebraic degree of P and for any $v \in \mathbb{F}_2^n$, we have

$$\bigoplus_{x \in \mathcal{V} \oplus v} P(x) = 0.$$

Since the same property does not, in general, hold for a permutation drawn at random, it can be exploited to distinguish a given (keyed or keyless) permutation from a random permutation. The idea was first introduced by Lai [37], albeit without a concrete application. Knudsen [36] then used higher-order differentials to analyze low-degree ciphers which were deemed secure against standard differential cryptanalysis. Several variants and generalisations of higher-order differential attacks have since been proposed in the literature, including *cube attacks* [22] and the *division property* [39]. More recently, Beyne et al. [12] generalized the higher-order differential attack initially proposed for cryptographic schemes defined over $\mathbb{F}_2^n$ to the case of schemes defined over $\mathbb{F}_q^t$ for $q = p^n$ for a prime number $p \geq 3$ and $n \geq 1$.

The crucial problem in higher-order differential attacks against iterated constructions is the analysis of the growth of the algebraic degree. While the results in [15] provide a generic upper-bound on the algebraic degree for SPN designs defined over $(\mathbb{F}_{2^n})^t$, one can expect that tighter bounds may be obtained by exploiting the details of the analyzed scheme. This is indeed the case for MiMC-like schemes: in [24], the authors exploit the low-degree of the round function defined over $\mathbb{F}_{2^n}$ to derive a preciser estimate on the number of rounds necessary to reach maximum algebraic degree $n-1$.

Nomenclature. Since we do not make any assumption about the round-keys, our results equally apply to keyed and keyless permutations. Thus in this paper we refer to both by using the term “schemes”. In this nomenclature, e.g., an *SPN scheme* is a family of permutations built from an SPN construction parametrized by secret keys or publicly known constants.

1.1 Preventing Higher-Order Differential Attacks – State of the Art

We focus on the case of *iterated* schemes, that is, schemes consisting of several iterations of the same round function parametrized by different round keys. To prevent higher-order differential attacks on schemes over $\mathbb{F}_2^N$, ideally one would like to make a statement such as:

> “After r rounds, no output bit can be represented as a Boolean function
> of algebraic degree strictly smaller than $N-1$.”

To achieve this goal, one needs to estimate the growth of the algebraic degree, which is in general a difficult task. In other words, predicting the evolution of the algebraic degree of the scheme when the number of rounds varies is the main challenge in higher-order differential cryptanalysis.

Theoretical Bounds on the Algebraic Degree. A naive bound for the algebraic degree of the composition of two functions $F, G : \mathbb{F}_2^N \to \mathbb{F}_2^N$ is given by

$$\deg(G \circ F) \leq \deg(G) \cdot \deg(F). \tag{1}$$

This bound leads to a first estimate about the number of rounds *necessary* to reach maximum algebraic degree in SPN schemes. For an SPN scheme defined over $(\mathbb{F}_{2^n})^t$ with S-Box layer of algebraic degree δ, it follows that at least

$$\log_\delta(n \cdot t - 1) \approx \log_\delta(n) + \log_\delta(t)$$

rounds are required to reach maximum degree (note that the affine layer does not increase the algebraic degree).

A Better Estimation for $\deg(G \circ F)$*.* In general, the naive upper bound (1) does not reflect the real growth of the algebraic degree when considering iterated schemes, and the problem of estimating the growth of the algebraic degree has therefore been studied in the literature. After the initial work of Canteaut and Videau [17], a tighter upper bound was presented by Boura, Canteaut, and De Cannière in [15]. In there, the authors deduce a new bound for the algebraic degree of iterated permutations for SPN schemes over $\mathbb{F}_2^N$, which includes functions that have a number of $t \geq 1$ balanced S-Boxes as their non-linear layer. Such bound only relies on the algebraic degree of the S-Box over $\mathbb{F}_2^n$, and no assumption on the linear layer is made. On the contrary, in order to apply the result presented in [15], one has to determine a particular parameter γ, that depends on the details of the S-Box. As we discuss in Section 3, for an S-Box over $\mathbb{F}_{2^n}$ the cost for computing γ is exponential in n. This means, for large S-Boxes (e.g., $n \geq 64$) it is infeasible to determine γ computationally and a further study of the analyzed scheme is necessary. However, theoretically bounding γ is in general a difficult task.

Apart from the bound of Boura, Canteaut and De Cannière, in a follow-up work Boura and Canteaut studied the influence of F^{-1} on the algebraic degree of $\deg(G \circ F)$ [14]. As main result, they discuss how the algebraic degrees of F^{-1} and F affect each other, which subsequently allows them to bound the algebraic degree of $G \circ F$ by means of the degrees of G and F^{-1}. More recently, Carlet [18] presented a bound on the algebraic degree of $G \circ F$ by working with the indicators of the graphs $\mathcal{G}_F$ and $\mathcal{G}_G$ (where $\mathcal{G}_F = \{(x, F(x)) : x \in \mathbb{F}_2^n\}$). In this work, Carlet bounds the algebraic degree of $G \circ F$ via the degree of G and the degree of the indicator function of $\mathcal{G}_F$. However, to the best of our knowledge the bounds in [18] exhibit the same behaviour and limitations as the bound in [15] when applied to the composition of more than two functions and it is unclear if the bounds in [18] practically improve upon the ones in [15] (see [18, Section 5.2.4]). If they do, the improvements seem to be only by one unit ([18, Section 5.2.2]).

Algebraic Degree in MiMC-Like Schemes. MiMC [3,29] is an *iterated Even–Mansour scheme*, i.e., a scheme natively defined over $\mathbb{F}_{2^n}$, where the S-Box is given by the cube function $x \mapsto x^3$. Only recently a new upper bound on the algebraic degree of MiMC-like schemes has been proposed in [24] at Asiacrypt 2020. More precisely, the authors show that when the round function can be described as a low-degree polynomial function over $\mathbb{F}_{2^n}$ of degree at most d, the algebraic degree $\delta(r)$ of r iterations of the round function grows linearly with the number of rounds, i.e.,

$$\delta(r) \leq \log_2(d^r + 1).$$

This observation implies that at least $\log_d(2^{n-1} - 1)$ rounds are required for reaching maximum algebraic degree. As a concrete application, [24] shows that the number of rounds in MiMC needs to be increased by several percent in order to resist all known attacks.

Tool-Based Bounds on the Algebraic Degree. For completeness, we recall other works in which the growth of the algebraic degree is studied. Compared to the works just presented, the following ones do not provide a theoretical estimation of the growth of the degree. Instead, computer tools are used to derive a bound on the algebraic degree. As we are going to highlight, this has approach has some limitations when considering cryptographic schemes with larger S-Boxes over $\mathbb{F}_{2^n}$.

Division Property. A generalization of integral and higher-order differential distinguishers is the division property [39], proposed by Todo at Eurocrypt 2015. Roughly speaking, the division property generalizes integral attacks and higher-order differential distinguishers in the sense that it is a classical higher-order differential distinguisher, but it is exhibited by exploiting the classical properties used in integral attacks together with some algebraic properties related to the degree of several iterations of a nonlinear function. At the current state of the art, the division property can only provide useful bounds on the algebraic degree for *small* n. Indeed, currently it is infeasible to apply the three-subset bit-based division property [40,25,41,33] to large S-Boxes (i.e., of size bigger than 12 bits to the best of our knowledge). Hence, such a tool does not seem to be useful in the case of schemes defined over $(\mathbb{F}_{2^n})^t$ for large n (as targeted in this paper), and a theoretical estimation is hence crucial.

Lower Bounds on the Degree of Block Ciphers. Hebborn, Lambin, Leander, and Todo [31] provide for the first time meaningful lower bounds on the algebraic degree of modern block ciphers. Given an encryption scheme E_k, they are able to show that a certain number of rounds are necessary such that that there exists at least one key and at least one output bit for which the degree is at least equal to a certain value δ. In essence, their result is based on the division property. Therefore it can only be used for small-size S-Boxes, since it is an open problem to model large S-Boxes using the division property. Moreover, it does not apply to the case of unkeyed permutations or hash functions and it does not return the value of the key for which the lower bound is achieved.

1.2 Our Contribution

As discussed above, currently there are only two possible approaches for estimating the growth of the algebraic degree of SPN schemes: a theoretical one based on the results by Boura, Canteaut and De Cannière [15] and a tool-based one using the division property. However, both approaches have inherent limitations when applied to SPN schemes defined over $(\mathbb{F}_{2^n})^t$ for large n (as targeted in this paper and important for MPC-/FHE-/ZKP-friendly schemes): in the first approach, the degree of the S-Box over $\mathbb{F}_{2^n}$ and the alignment of the scheme (for details regarding alignment see Section 2.2) are not taken into account. While this could be an advantage in the sense that such results apply to a large class of schemes, the resulting estimation of the growth of the algebraic degree is far from being optimal when applied to strong-aligned schemes over $(\mathbb{F}_{2^n})^t$ with large and low-degree S-Boxes (as targeted in this paper); in the second approach, the tools cannot tackle large S-Boxes (i.e., $n \geq 12$).

In this paper, we address these limitations and present a new theoretical upper bound on the algebraic degree for strong-aligned SPN schemes over $(\mathbb{F}_{2^n})^t$, see our main result in Theorem 1. In more detail, we consider strong-aligned SPN schemes over $(\mathbb{F}_{2^n})^t$ for $n \geq 3$ and $t \geq 2$, where the S-Boxes are defined via invertible non-linear polynomial functions of univariate degree $d \geq 3$ and algebraic degree $\delta \geq 2$ and where the linear layer is defined via the multiplication with a matrix in $(\mathbb{F}_{2^n})^{t\times t}$. In Section 4.2, we prove Theorem 1 and show that

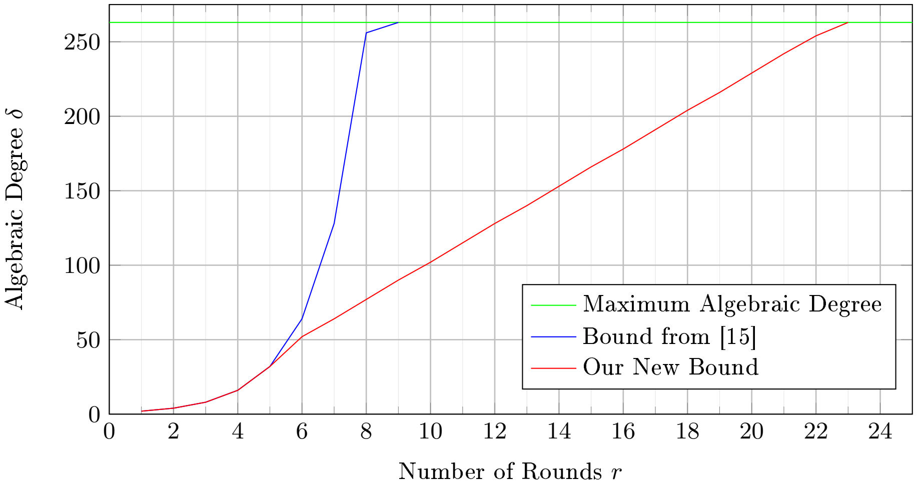

Fig. 1: Comparison between our new bound and the one proposed in [15] for the case of an SPN scheme instantiated over $(\mathbb{F}_{2^{33}})^8$ with a cube S-Box $S(x) = x^3$ (where $n = 33$, $t = 8$, $d = 3, \delta = 2$ and $\gamma = (n + 1)/2 = 17$).

the algebraic degree $\delta(r)$ after r rounds is upper-bounded by

$$\delta(r) \leq \begin{cases} \delta^r & \text{if } r \leq R_{\text{exp}} = 1 + \lfloor \log_\delta(t) \rfloor, \\ t \cdot \log_2\left(\frac{d^r}{t} + 1\right) & \text{if } R_{\text{exp}} < r \leq R_{\text{SPN}}. \end{cases}$$

Here, at least

$$R_{\text{SPN}} = \log_d \left(t \cdot (2^n - 1) - 2^{n-1}\right) \approx \log_d(2^n - 1) + \log_d(t)$$

rounds are necessary to reach maximum algebraic degree $nt - 1$, see Section 4.1. We have practically verified our results on small-scale schemes. Section 5 is devoted to a more detailed discussion of our practical experiments. A concrete comparison between our new bound on the algebraic degree and the one proposed in [15] for an SPN scheme over $(\mathbb{F}_{2^{33}})^8$ with cube S-Box $S(x) = x^3$ is presented in Fig. 1.

This is the first concrete improvement regarding the theoretical estimation of the growth of the algebraic degree for SPN schemes in the last decade. Besides that, our new bound from Theorem 1 substantially improves the bound in [15] for schemes with low-degree and large S-Boxes. In this scenario, the bound in [15] turns out to be almost identical to the naive exponential bound (except, when the algebraic degree is close to its maximum). A concrete example is shown in Fig. 1 and we refer to Proposition 1 for a more detailed discussion of this aspect. We emphasize that, while our bound is applicable for any parameter values of d, n, t and δ, it is most competitive for low-degree ($d \ll 2^n - 1$) and large ($n \gg 1$) S-Boxes.

2 Preliminaries

In this section, we recall the most important results about polynomial representations of Boolean functions and we give and outline of the definition of stronly-aligned SPN and iterated Even–Mansour schemes.

2.1 Polynomial Representations over Binary Extension Fields

We denote addition (and subtraction) in binary extension fields and polynomial rings over binary extension fields by the symbol $\oplus$. For $n, t \in \mathbb{N}$, every function $F : (\mathbb{F}_{2^n})^t \to \mathbb{F}_{2^n}$ can be uniquely represented by a polynomial over $\mathbb{F}_{2^n}$ in t variables with maximum degree $2^n - 1$ in each variable, i.e., as

$$F(X_1, \ldots, X_t) = \bigoplus_{v=(v_1,\ldots,v_t)\in\{0,1,\ldots,2^n-1\}^t} \varphi(v) \cdot X_1^{v_1} \cdot \ldots \cdot X_t^{v_t}, \tag{2}$$

for certain $\varphi(v) \in \mathbb{F}_{2^n}$. We refer to this representation as the *word-level representation*. At the same time, the function F can be written as an n-tuple $(F_1, \ldots, F_n)$ of functions $F_i : \mathbb{F}_2^N \to \mathbb{F}_2$ and thus admits a unique representation as an n-tuple $(F_1, \ldots, F_n)$ of polynomials over $\mathbb{F}_2$ in $N := n \cdot t$ variables with maximum degree 1 in each variable. Here, F_i takes the form

$$F_i(Y_1, \ldots, Y_N) = \bigoplus_{u=(u_1,\ldots,u_N)\in\{0,1\}^N} \rho_i(u) \cdot Y_1^{u_1} \cdot \ldots \cdot Y_N^{u_N}, \tag{3}$$

where the coefficients $\rho_i(u) \in \mathbb{F}_2$ can be computed by the *Moebius transform* with time complexity $\mathcal{O}(n \cdot 2^n)$. We call this alternative description the *bit-level representation* of F. Combining Equations (3), for $1 \leq i \leq n$, into a single polynomial representation leads to a description of F as a single polynomial in $N = n \cdot t$ variables, but now with coefficients in $\mathbb{F}_2^n$, instead of $\mathbb{F}_2$.

Whenever we refer to the degree of a single variable in F (or F_i), we shall speak of the *univariate degree*. In contrast, the degree of F (or F_i) as a multivariate polynomial shall be called its *multivariate degree*, or just its *degree*.

Definition 1. *For $n, t \in \mathbb{N}$, let $F : (\mathbb{F}_{2^n})^t \to \mathbb{F}_{2^n}$ be a function and*

$$F(X_1, \ldots, X_t) = \bigoplus_{v=(v_1,\ldots,v_t)\in\{0,1,\ldots,2^n-1\}^t} \varphi(v) \cdot X_1^{v_1} \cdot \ldots \cdot X_t^{v_t},$$

its word-level representation as defined in Eq. (2)*. The multivariate degree of F over $\mathbb{F}_{2^n}$ is defined as*

$$\max \left\{ \sum_{i=1}^{t} v_i : v = (v_1, \ldots, v_t) \in \{0, 1, \ldots, 2^n - 1\}^t,\ \varphi(v) \neq 0 \right\}.$$

The univariate degree of a one-variable monomial $X_i^{v_i}$ is v_i.

We denote functions $F : \mathbb{F}_2^n \to \mathbb{F}_2$ as *Boolean functions* and hence functions of the form $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$, for $n \in \mathbb{N}$, as *vectorial Boolean functions.* The unique polynomial representation of a Boolean function is called its *algebraic normal form* (ANF), which we emphasize with the following definition.

Definition 2. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function. The algebraic normal form (ANF) of F is the unique representation as a polynomial over $\mathbb{F}_2$ in n variables and with maximum univariate degree 1, i.e., the representation*

$$F(Y_1, \dots, Y_n) = \bigoplus_{u=(u_1,\dots,u_n)\in\{0,1\}^n} \rho(u) \cdot Y_1^{u_1} \cdot \dots \cdot Y_n^{u_n}.$$

The algebraic degree $\delta(F)$ of F is the degree of above representation of F as a multivariate polynomial over $\mathbb{F}_2$.

When the function F is clear from the context, we also write δ instead of $\delta(F)$. If $G : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a vectorial Boolean function and $(G_1, \dots, G_n)$ is its representation as an n-tuple of multivariate polynomials over $\mathbb{F}_2$, then its algebraic degree $\delta(G)$ is defined as the maximal algebraic degree of its coordinate functions G_i, i.e. as $\delta(G) \max_{1\leq i\leq n} \delta(G_i)$. The link between the algebraic degree and the univariate degree of a vectorial Boolean function is well-known, e.g. it is established in [19, Sect. 2.2]: due to the isomorphism of $\mathbb{F}_2$-vector spaces $\mathbb{F}_{2^n} \cong \mathbb{F}_2^n$, every function over $\mathbb{F}_2^n$ can be considered as a function over $\mathbb{F}_{2^n}$ and thus admits a representation as an univariate polynomial over $\mathbb{F}_{2^n}$. Hence, the algebraic degree of a vectorial Boolean function can be computed from its univariate representation. Eq. (4) makes this link explicit: Let $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a function over $\mathbb{F}_{2^n}$ and let $F(X) = \sum_{i=0}^{2^n-1} \varphi_i \cdot X^i$ denote the corresponding univariate polynomial description over $\mathbb{F}_{2^n}$. The algebraic degree $\delta(F)$ of F as a vectorial Boolean function is the maximum over all Hamming weights[5] of exponents of non-vanishing monomials, that is

$$\delta(F) = \max_{0\leq i\leq 2^n-1} \{\mathrm{hw}(i) \,|\, \varphi_i \neq 0\}. \tag{4}$$

Lastly, we recall that the algebraic degree of an invertible function F over $\mathbb{F}_2^n$ is at most $n-1$, while the univariate polynomial representation of an invertible function F over $\mathbb{F}_{2^N}$ has degree at most $2^N - 2$.

2.2 Strong-Aligned SPN Schemes

Here we recall the concept of strong-aligned SPN schemes, and we fix the notation used in the rest of the article. Let $E_k^r : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ denote the application of r rounds of an SPN scheme under a fixed (secret or publicly known) key $k \in (\mathbb{F}_{2^n})^t$ with $n \geq 3$, $t \geq 2$, and $N := n \cdot t$. For every $x = (x_1, \dots, x_t) \in (\mathbb{F}_{2^n})^t$ we write

$$E_k^r(x) := (F_r \circ \dots \circ F_1)\,(x \oplus k_0), \tag{5}$$

[5] Given $x = \sum_{i=0}^{s} x_i \cdot 2^i \in \mathbb{Z}$, for $x_i \in \{0,1\}$, then $\mathrm{hw}(x) = \sum_{i=0}^{s} x_i$.

where $F_i : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ is defined as $F_i(x) \coloneqq R(x) \oplus k_i$, for $1 \leq i \leq r$. The subkeys $k_0, \ldots, k_r \in (\mathbb{F}_{2^n})^t$ may be derived from the master key $k \in (\mathbb{F}_{2^n})^t$ by means of a key schedule, or they may just as well be randomly chosen elements. Here, R denotes the composition of the S-Box and the linear layer, i.e., we have $R : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ with

$$R(x) \coloneqq (M \circ S)(x) \coloneqq M(S_1(x_1), \ldots, S_t(x_t)),$$

where all $S_i : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ are assumed to be invertible non-linear polynomial S-Boxes of degree $d \geq 3$ defined as

$$S_i(x) \coloneqq \bigoplus_{j=0}^{d} c_j^{(i)} \cdot x^j, \tag{6}$$

for $c_j^{(i)} \in \mathbb{F}_{2^n}$ and $c_d^{(i)} \neq 0$. Finally, M denotes an invertible linear layer $M : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ defined by the multiplication with a matrix

$$M(x) \coloneqq \begin{pmatrix} M_{1,1} & M_{1,2} & \ldots & M_{1,t} \\ M_{2,1} & M_{2,2} & \ldots & M_{2,t} \\ \vdots & & \ddots & \vdots \\ M_{t,1} & M_{t,2} & \ldots & M_{t,t} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{pmatrix}, \tag{7}$$

where $M_{i,j} \in \mathbb{F}_{2^n}$ for $i, j = 1, \ldots, t$. Using the terminology introduced in [11], we refer to *weak-aligned* scheme if the matrix that defines the linear layer is defined over $(\mathbb{F}_2)^{N \times N}$ and it does not have any equivalent representation in $(\mathbb{F}_{2^n})^{t \times t}$, and to *strong-aligned* scheme otherwise. We recall that every matrix in $(\mathbb{F}_{2^n})^{t \times t}$ has an equivalent representation over $(\mathbb{F}_2)^{N \times N}$, while vice-versa is not true in general. As a concrete example, AES is a strong-aligned scheme, while Keccak is a weak-aligned scheme.

In the following, we assume that the matrix M ensures *full diffusion after a finite number of rounds*, in the sense that there exists an $r \in \mathbb{N}$ such that every word of the internal state after the application of r rounds depends on every input word $x_1, \ldots, x_t$. E.g., the smallest integer r that satisfies the previous condition for an MDS matrix is 1, for the AES MixLayer it is 2, while it does not exist for a diagonal matrix. We refer to [9,10, App. D] for a more detailed analysis about this concept.

We remark that all strong-aligned SPN schemes over $(\mathbb{F}_{2^n})^t$ can be written as described above. E.g., if the linear layer is defined by an MDS matrix[6], the scheme is similar to SHARK [38]. For AES [21] or AES-like ciphers, where the linear layer is obtained by a combination of the ShiftRows and the MixColumns operations, many elements of the matrix M are equal to 0 (see e.g. [8]).

[6] A matrix $M \in \mathbb{F}^{t \times t}$ is called a maximum distance separable (MDS) matrix iff every $u \times u$ submatrix of M is invertible, where $u \leq t$.

Iterated Even–Mansour Schemes. A particular subclass of SPN schemes are iterated Even–Mansour schemes. An iterated Even–Mansour scheme is an SPN scheme with only one word, i.e., with $t = 1$. In this case, the matrix multiplication would be just the multiplication with a non-zero element in $\mathbb{F}_{2^n}$, which we can omit. A scheme in the literature that qualifies as an iterated Even–Mansour scheme is, e.g., MiMC.

3 Iterative Application of the Bound in [15]

The bounds on the algebraic degree in [15] are stated for the composition of two functions which means that the application to iterated SPN schemes (which often comprise the composition of several dozen functions) requires an ad-hoc analysis of the analyzed scheme. The goal of this section is to provide a closed formuala for the bound in [15, Theorem 2] when extended to the composition of more than two functions. This closed formula provides the basis for our comparisons in Section 5.

The currently best generic upper bound for the algebraic degree of the composition of two functions is given by Boura, Canteaut, and De Cannière in [15, Theorem 2]: Let F be a function from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ corresponding to the concatenation of t smaller balanced[7] S-Boxes $S_1, \ldots, S_t$ defined over $\mathbb{F}_2^n$. Then, for any function G from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$, it holds

$$\deg(G \circ F) \leq N - \frac{N - \deg(G)}{\gamma}, \tag{8}$$

where

$$\gamma := \max_{i=1,\ldots,n-1} \frac{n-i}{n-\delta_i} \leq n - 1, \tag{9}$$

and δ_i is defined as the maximal algebraic degree of the product of any i coordinates of any of the smaller S-Boxes.

We emphasize that γ and δ_i depend on the details of the S-Box. Namely, two S-Boxes with the same algebraic degree can have in general different γ. The result in [14, Theorem 2] uses the algebraic degree of the compositional inverses S_j^{-1}, $1 \leq j \leq t$, for a bound on the algebraic degree of $G \circ F$. Under the same assumptions as above this result leads to the same bound as stated in Eq. (8), with the additional upper bound on γ

$$\gamma \leq \max_{1 \leq j \leq t} \max \left\{ \frac{n-1}{n - \deg(S_j)},\ \frac{n}{2} - 1,\ \deg\left(S_j^{-1}\right) \right\}. \tag{10}$$

Using an upper bound on γ for bounding the algebraic degree of $G \circ F$ in Eq. (8) could lead to a less tight bound on $\deg(G \circ F)$ than using the exact value of γ. However, Eq. (10) has the advantage that it only uses known facts about the

[7] A function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is said to be *balanced* if each element in $\mathbb{F}_2^m$ has exactly 2^{n-m} preimages. For $n = m$, an S-Box is balanced iff it is invertible.

involved functions and thus a bound on $\deg(G \circ F)$ can be computed straight away. The same remark applies to another bound in [14, Corollary 2], which works with the algebraic degree of F^{-1} and is given by

$$\deg(G \circ F) \leq N - \frac{N-1-\deg(G)}{\deg\left(F^{-1}\right)}.$$

In Proposition 1, we derive a *direct* upper bound of the algebraic degree of SPN schemes in the simple but most common case where all S-Boxes are equal. With "direct" upper bound we mean that we iteratively apply (8) to the round functions of an SPN scheme and thus obtain a closed-form statement about the algebraic degree after a certain number of rounds (and not only for the composition of two functions as stated in [15]).

Proposition 1. *Let F be a function from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ corresponding to the concatenation of t copies of a balanced S-Box S over $\mathbb{F}_{2^n}$ with algebraic degree $\delta \geq 2$. For any affine functions $L_1, L_2, \ldots, L_r$ from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ and any integer $r \geq 1$ consider the SPN scheme E_r from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ defined as*

$$E_r := L_r \circ F \circ L_{r-1} \circ F \circ \cdots \circ L_2 \circ F \circ L_1 \circ F.$$

Then the algebraic degree $\delta(r)$ of E after r rounds is upper-bounded by

$$\delta(r) \leq \begin{cases} \delta^r & \text{if } r \leq R_0 := \left\lfloor \log_\delta \left(N \cdot \frac{\gamma-1}{\gamma\cdot\delta-1}\right) \right\rfloor, \\ \frac{\delta^{R_0}}{\gamma^{r-R_0}} + N \cdot \left(1 - \frac{1}{\gamma^{r-R_0}}\right) & \text{if } R_0 < r \leq \mathcal{R}_{[\text{BCD11}]}, \end{cases} \tag{11}$$

independent of the (secret or publicly known) key k, where

$$\mathcal{R}_{[\text{BCD11}]} := \underbrace{\left\lfloor \log_\delta \left(N \cdot \frac{\gamma-1}{\gamma\cdot\delta-1}\right) \right\rfloor}_{=R_0} + \left\lceil \log_\gamma \left(N - \delta^{R_0}\right) \right\rceil \tag{12}$$

is the minimum numbre of rounds for security against higher-order differential distinguishers and where γ is defined as in Eq. (9).

The proof of Proposition 1 can be found in Supplementary Material A. The strategy we adopt to prove Proposition 1 is similar to the one proposed by Biryukov, Khovratovich, and Perrin [13]. In there, authors focused on the case in which all S-Boxes have maximum algebraic degree $\delta = n-1$, while here we do not need this restriction.

Cost of Computing γ. The growth of the degree predicted in (8) depends on the value of γ. Computing γ can be very expensive for large S-Boxes. Indeed, one has to consider all possible combinations of the product of any i coordinates of the given S-Boxes, which implies a cost of order

$$\mathcal{O}\left(\sum_{i=1}^{n} \binom{n}{i}\right) \approx \mathcal{O}(2^n).$$

In the case in which t different S-Boxes are used, the previous cost must be multiplied by t. This means that for large S-Boxes (e.g., $n \geq 64$) it is infeasible to determine γ computationally and a further analysis of the scheme is necessary. Our results in Section 4 do not have this limitation. They depend on known parameters of the scheme and can be computed straight away.

4 Higher-Order Differential Analysis of strong-aligned SPN Schemes

In this section we prove a new upper bound on the growth of the algebraic degree in strong-aligned SPN schemes. We note, our proof proceeds analogously for SPN-derived block ciphers and permutations by assuming fixed and publicly known constants in the latter case and fixed secret keys in the former one.

4.1 Minimum Number of Rounds for Preventing Higher-Order Differential Distinguishers

Here, we provide a minimum number of rounds to reach maximum algebraic degree in strong-aligned SPN schemes. We show, this number matches the minimum number of rounds needed to provide security against the interpolation attack [35].

Proposition 2. *Let $n \geq 3$. Consider r rounds of a strong-aligned SPN scheme E_k^r over $(\mathbb{F}_{2^n})^t$ as defined in Eq. (5), with the additional assumption that all S-Boxes $S_1, \ldots, S_t$ are defined via non-linear polynomial functions with equal univariate degree $d \geq 3$. A lower bound on the number of rounds to prevent higher-order differential distinguishers is given by*

$$\mathcal{R}_{\text{SPN}} \coloneqq \lceil \log_d \left(t \cdot (2^n - 1) - 2^{n-1} \right) \rceil, \tag{13}$$

independent of the (secret or publicly known) key k.

We note that

$$\mathcal{R}_{\text{SPN}} \approx \log_d(2^n - 1) + \log_d(t), \tag{14}$$

especially for $t, n \gg 1$ and small $d \geq 3$.

Proof. To reach maximum algebraic degree $n \cdot t - 1$ the polynomial representation of E_k^r over $\mathbb{F}_{2^n}$ must contain a monomial with algebraic degree n in $t-1$ variables and algebraic degree $n-1$ in one variables. This happens if E_k^r contains a word-level monomial with univariate degree $2^n - 1$ in $t-1$ variables and univariate degree $2^{n-1} - 1$ in one variable. Since the multivariate degree of E_k^r after $r \geq 1$ rounds is upper bounded by d^r, we obtain

$$d^r \geq (t-1) \cdot (2^n - 1) + 2^{n-1} - 1 = t \cdot (2^n - 1) - 2^{n-1}$$

as a necessary condition on the number of rounds to reach maximum algebraic degree $n \cdot t - 1$. Rearranging for r yields $r \geq \log_d \left(t \cdot (2^n - 1) - 2^{n-1} \right)$.

4.2 Algebraic Degree of Strong-Aligned SPN Schemes

As main result of this paper, we prove the following statement.

Theorem 1. *Let $n \geq 3$ and $t \geq 1$. Consider r rounds of a strong-aligned SPN scheme E_k^r over $(\mathbb{F}_{2^n})^t$ as defined in Eq. (5), with the additional assumption that all S-Boxes $S_1, \ldots, S_t$ are defined via the same invertible non-linear function S of univariate degree $d \geq 3$ and algebraic degree $\delta \geq 2$.*

Let $R_{exp} := 1 + \lfloor \log_\delta(t) \rfloor$. Then, the algebraic degree of E_k^r after r rounds, denoted by $\delta(r)$, is upper-bounded by

$$\delta(r) \leq \begin{cases} \delta^r & \text{if } r \leq R_{exp}\,, \\ \min\left\{\delta^r,\ t \cdot \log_2\left(\frac{d^r}{t} + 1\right)\right\} & \text{if } r > R_{exp}, \end{cases} \tag{15}$$

independent of the (secret or publicly known) key k and until the maximum algebraic degree $n \cdot t - 1$ is reached.

Idea of the proof. The roadmap for the proof of Theorem 1 reads as follows:

1. Lemma 1 makes a statement about which monomials can occur in the polynomial representation of the encryption function;
2. In Lemma 2 we prove that the algebraic degree grows as fast as δ^r in the first $R_{\text{exp}} := 1 + \lfloor \log_\delta(t) \rfloor$ rounds; this shows that the naive exponential bound can indeed be achieved;
3. Lemma 3 provides a first upper bound on the algebraic degree by moving from the integer numbers to the real numbers and by involving the logarithmic functions instead of the hamming weights; that is
$$\delta(r) \leq \max\left\{\sum_{i=1}^{t} \log_2(y_i + 1) : (y_1, \ldots, y_t) \in \mathbb{R}^t_{\geq 0},\ \sum_{i=1}^{t} y_i = d^r\right\};$$
4. finally, in Lemma 4, we compute the maximum of the function from Lemma 3 using the method of Lagrange multipliers, resulting in a concrete bound of the form $\delta(r) \leq t \cdot \log_2\left(\frac{d^r}{t} + 1\right)$.

4.3 Proof of Theorem 1: (Initial) Exponential Growth

Lemma 1. *Let $t \geq 1$ and let $d' \geq 3$ be an integer and let $d' = \sum_{i=1}^{\delta} 2^{d_i}$ be the base-2 expansion of d for certain $d_i \in \mathbb{N}$. Given a polynomial $P = \bigoplus_{i \in \{1,\ldots,l\}} c_i \cdot m_i \in \mathbb{F}_{2^n}[X_1, \ldots, X_t]$ that contains the monomials $m_1, m_2, \ldots, m_l \in \mathbb{F}_{2^n}[X_1, \ldots, X_t]$ for a certain $l \geq 1$, the monomials in $P^{d'}$ are of the form*

$$m_{i_1}^{2^{d_1}} \cdot m_{i_2}^{2^{d_2}} \cdot \ldots \cdot m_{i_\delta}^{2^{d_\delta}} \tag{16}$$

where $i_1, i_2, \ldots, i_\delta \in \{1, 2, \ldots, l\}$.

Proof. We obtain

$$\begin{aligned} P^{d'} = \left(\bigoplus_{i \in \{1,\dots,l\}} c_i \cdot m_i \right)^{2^{d_1}+\dots+2^{d_\delta}} &= \prod_{j=1}^{\delta} \left(\bigoplus_{i \in \{1,\dots,l\}} c_i^{2^{d_j}} \cdot m_i^{2^{d_j}} \right) \\ = \bigoplus_{i_1,i_2,\dots,i_\delta \in \{1,2,\dots,l\}} \left(\prod_{j=1}^{\delta} c_{i_j}^{2^{d_j}} \cdot m_{i_j}^{2^{d_j}} \right). \end{aligned}$$

where the second equality holds since $(x \oplus y)^2 = x^2 \oplus y^2$ for each $x, y \in \mathbb{F}_{2^n}$ (based on the fact that $\text{char}(\mathbb{F}_{2^n}) = 2$). Hence, we conclude that only monomial products of the form

$$m_{i_1}^{2^{d_1}} \cdot m_{i_2}^{2^{d_2}} \cdot \ldots \cdot m_{i_\delta}^{2^{d_\delta}}$$

may occur in P^d, where $i_1, i_2, \dots, i_\delta \in \{1, 2, \dots, l\}$. The monomials $m_{i_1}, \dots, m_{i_\delta}$ are not necessarily different, therefore the exponents in Eq. (16) are either powers of 2 or sums of powers of 2. □

The next lemma shows that the naive exponential bound δ^r for the algebraic degree is not only a trivial bound but can indeed be achieved.

Lemma 2. *Let the same conditions as in Theorem 1 hold. Furthermore, let $S(x) = \sum_{i=0}^{d} c_i \cdot x^i$ for $c_i \in \mathbb{F}_{2^n}$, and let d' be a degree for which $hw(d') = \delta$ and $c_{d'} \neq 0$. Let $d' = \sum_{i=1}^{\delta} 2^{d_i}$ be the base-2 expansion of d' for appropriate $d_i \in \mathbb{N}$. In the first $R_{exp} = 1 + \lfloor \log_\delta(t) \rfloor$ rounds the algebraic degree grows as fast as δ^r.*

Proof. The idea is to observe the growth of the algebraic degree with the help of Lemma 1. After the first round, all monomials $X_1^{d'}, \dots, X_t^{d'}$ are present in the polynomial representation of E_k^r and have algebraic degree δ.

According to Lemma 1, after one more round all monomials of the form $(i_1, \dots, i_\delta \in \{1, \dots, t\})$

$$(X_{i_1}^{d'})^{2^{d_1}} \cdot (X_{i_2}^{d'})^{2^{d_2}} \cdot \dots \cdot (X_{i_\delta}^{d'})^{2^{d_\delta}},$$

are present in the encryption polynomial and have algebraic degree δ^2 if $i_1, \dots, i_\delta$ are pairwise different. To see why they have algebraic degree δ^2, we note that: *(a)* raising a (word-level) monomial of E_k^r to the power of 2^k, $k \in \mathbb{N}$, does not change its algebraic degree, and *(b)* if two (word-level) monomials $m_{\alpha_1}, m_{\alpha_2}$ of E_k^r do not contain any shared variable, the algebraic degree of the product $m_{\alpha_1} \cdot m_{\alpha_2}$ is the sum of the respective algebraic degrees.

In the same way as before, after another round, all monomials of the form $(i_1, \dots, i_{\delta^2} \in \{1, \dots, t\})$

$$(X_{i_1}^{d' \cdot 2^{d_1}} \cdots X_{i_\delta}^{d' \cdot 2^{d_\delta}})^{2^{d_1}} (X_{i_{\delta+1}}^{d' \cdot 2^{d_1}} \cdots X_{i_{2\delta}}^{d' \cdot 2^{d_\delta}})^{2^{d_2}} \cdots (X_{i_{\delta^2-(\delta-1)}}^{d' \cdot 2^{d_1}} \cdots X_{i_{\delta^2}}^{d' \cdot 2^{d_\delta}})^{2^{d_\delta}}$$

appear in the encryption polynomial and have algebraic degree δ^3 if $i_1, \dots, i_{\delta^2}$ are pairwise different. Continuing this way, we conclude that the algebraic degree grows as fast as δ^r until all t variables are exhausted, i.e., until $\delta^r = \delta \cdot t$, or equivalently, for the first $\lfloor \log_\delta(\delta \cdot t) \rfloor = 1 + \lfloor \log_\delta(t) \rfloor$ rounds. □

4.4 Proof of Theorem 1: Linear Growth

Lemma 3. *Let the same conditions as in Theorem 1 hold. Then, the algebraic degree of E_k^r after r rounds, denoted by $\delta(r)$, is upper-bounded by*

$$\delta(r) \leq \max \left\{ \sum_{i=1}^{t} \log_2(y_i + 1) : (y_1, \dots, y_t) \in \mathbb{R}_{\geq 0}^t,\ \sum_{i=1}^{t} y_i = d^r \right\}. \tag{17}$$

Proof. Since the multivariate degree of a single output word of E_k^r after r rounds is upper bounded by d^r the algebraic degree $\delta(r)$ of E_k^r after r rounds can be upper bounded by

$$\delta(r) \leq \max_{\{(e_1,\dots,e_t)\in\mathbb{N}^t \,:\, \sum_{i=1}^t e_i \leq d^r\}} \sum_{i=1}^{t} \mathrm{hw}(e_i),$$

where we use the fact that the algebraic degree of a monomial $X_1^{e_1} \cdot \ldots \cdot X_t^{e_t}$ is given by $\sum_{i=1}^t \mathrm{hw}(e_i)$. Let $(e'_1, \dots, e'_t) \in \mathbb{N}^t$ be an integer vector with $\sum_{i=1}^t e'_i \leq d^r$ that maximizes $\sum_{i=1}^t \mathrm{hw}(e_i)$. Then it holds

$$\begin{aligned}
\delta(r) &\leq \max_{\{(e_1,\dots,e_t)\in\mathbb{N}^t \,:\, \sum_{i=1}^t e_i \leq d^r\}} \sum_{i=1}^{t} \mathrm{hw}(e_i) = \sum_{i=1}^{t} \mathrm{hw}(e'_i) \\
&= \sum_{i=1}^{t} \log_2 \left((2^{\mathrm{hw}(e'_i)} - 1) + 1 \right) \\
&\leq \max_{\{(y_1,\dots,y_t)\in\mathbb{R}_{\geq 0}^t \,:\, \sum_{i=1}^t y_i \leq d^r\}} \sum_{i=1}^{t} \log_2(y_i + 1),
\end{aligned}$$

where for the second inequality we note that

$$(2^{\mathrm{hw}(e'_1)} - 1, 2^{\mathrm{hw}(e'_2)} - 1, \dots, 2^{\mathrm{hw}(e'_t)} - 1) \in \left\{ (y_1, \dots, y_t) \in \mathbb{R}_{\geq 0}^t \ : \ \sum_{i=1}^{t} y_i \leq d^r \right\},$$

since $2^x - 1$ is the smallest integer with hamming weight $x \in \mathbb{N}$ and thus

$$\sum_{i=1}^{t} 2^{\mathrm{hw}(e'_i)} - 1 \leq \sum_{i=1}^{t} e'_i \leq d^r.$$

For finding the maximum of the function $\sum_{i=1}^t \log_2(y_i + 1)$ on the set

$$\left\{ (y_1, \dots, y_t) \in \mathbb{R}_{\geq 0}^t \ : \ \sum_{i=1}^{t} y_i \leq d^r \right\},$$

we now prove that it is enough to look on the set

$$\mathcal{Y}_t \coloneqq \left\{ (y_1, \dots, y_t) \in \mathbb{R}_{\geq 0}^t \ : \ \sum_{i=1}^{t} y_i = d^r \right\}.$$

The idea is as follows: for every element $(y_1,\ldots,y_t) \in \mathbb{R}^t_{\geq 0}$ with $\sum_{i=1}^t y_i \leq d^r$ we can find an element $(y'_1,\ldots,y'_t) \in \mathcal{Y}_t$ with

$$\sum_{i=1}^{t} \log_2(y_i+1) \leq \sum_{i=1}^{t} \log_2(y'_i+1).$$

Indeed, for every $(0,\ldots,0) \neq (y_1,\ldots,y_t) \in \mathbb{R}^t_{\geq 0}$ with $\sum_{i=1}^t y_i \leq d^r$ the intersection of the line $(x_1,\ldots,x_t) = \mu \cdot (y_1,\ldots,y_t)$ with the hyperplane $\sum_{i=1}^t x_i = d^r$ yields $\mu = \frac{d^r}{\sum_{i=1}^t y_i} \geq 1$ and therefore

$$\sum_{i=1}^{t} \log_2(y_i+1) \leq \sum_{i=1}^{t} \log_2(\underbrace{\mu \cdot y_i}_{=:y'_i}+1)$$

due to the increasing monotonicity of the logarithmic function on the specified domain. □

Lemma 4. *Let the same conditions as in Theorem 1 hold. Then, the algebraic degree of E^r_k after r rounds, denoted by $\delta(r)$, is upper-bounded by*

$$\delta(r) \leq t \cdot \log_2\left(\frac{d^r}{t}+1\right). \tag{18}$$

Proof. Case: $t=1$. The degree of E^r_k after $r \geq 1$ rounds is upper bounded by d^r, therefore

$$\begin{aligned}\delta(r) &\leq \max_{i\in\mathbb{N},\ 1\leq i\leq d^r} \mathrm{hw}(i) \leq \mathrm{hw}(i') = \log_2\left(2^{\mathrm{hw}(i')}-1+1\right)\\ &\leq \max_{y\in\mathbb{R},\ 1\leq y\leq d^r} \log_2(y+1) = \log_2(d^r+1).\end{aligned}$$

Here, i' is an integer $1 \leq i' \leq d^r$ that maximizes $\mathrm{hw}(i)$ and the third inequality holds since 2^x-1 is the smallest integer with hamming weight $x \in \mathbb{N}$. Therefore

$$2^{\mathrm{hw}(i')}-1 \leq i' \leq d^r,$$

which means the element $2^{\mathrm{hw}(i')}-1$ is in the domain of the last max-expression and hence the last inequality holds. The last equality is due to the increasing monotonicity of the logarithm on the specified domain.

Case: $t \geq 2$. For upper-bounding $\delta(r)$ we use Lemma 3

$$\delta(r) \leq \max_{\{(y_1,\ldots,y_t)\in\mathbb{R}^t_{\geq 0}\ :\ \sum_{i=1}^t y_i = d^r\}} \sum_{i=1}^{t} \log_2(y_i+1)$$

and find the maximum of $\sum_{i=1}^t \log_2(y_i+1)$ on the domain

$$\mathcal{Y}_t = \left\{(y_1,\ldots,y_t) \in \mathbb{R}^t_{\geq 0}\ :\ \sum_{i=1}^{t} y_i = d^r\right\}.$$

For finding the maximum we apply the method of Lagrange multipliers to the function $F(y_1, \dots, y_t) \coloneqq \sum_{i=1}^{t} \log_2(y_i+1)$ under the constraint $d^r - \sum_{i=1}^{t} y_i = 0$ (we recall solving equality constrained optimization problems with Lagrange multipliers in Supplementary Material B). This yields the Lagrangian function

$$L(y_1, \dots, y_t, \lambda) \coloneqq \sum_{i=1}^{t} \log_2(y_i + 1) - \lambda \cdot \left(d^r - \sum_{i=1}^{t} y_i \right).$$

Computing the partial derivatives with respect to $y_1, \dots, y_t, \lambda$ and equating them to zero results in the equation system

$$\frac{1}{\ln(2)(y_1 + 1)} - \lambda = \dots = \frac{1}{\ln(2)(y_t + 1)} - \lambda = d^r - \sum_{i=1}^{t} y_i = 0.$$

Solving for $y_1, \dots, y_t$ gives the candidate solution $y_1 = \dots = y_t = \frac{d^r}{t}$ for a maximum of F on the set $\mathcal{Y}_t$.

To check that it is indeed a maximum we need to compare the candidate solution against the values of F on the boundaries of $\mathcal{Y}_t$. Since F is a symmetric function it suffices to check only one of the t sets that bound $\mathcal{Y}_t$, for which we choose

$$\mathcal{H}_t \coloneqq \{(y_1, \dots, y_t) \in \mathbb{R}^t_{\geq 0} \ : \ y_t = 0, \ \sum_{i=1}^{t-1} y_i = d^r\}.$$

We prove that $y_1 = \dots = y_t = \frac{d^r}{t}$ is indeed a maximum by induction over t. The induction basis is

$$\max_{(y_1,\dots,y_t) \in \mathcal{Y}_t} \sum_{i=1}^{t} \log_2(y_i + 1) = t \cdot \log_2 \left(\frac{d^r}{t} + 1 \right).$$

For $t = 2$ we have the candidate $y_1 = y_2 = \frac{d^r}{2}$ for a maximum and check the boundary via

$$\max_{(y_1,y_2) \in \mathcal{H}_2} \log_2(y_1) + \log_2(y_2) = \log_2(d^r + 1) < 2 \cdot \log_2 \left(\frac{d^r}{2} + 1 \right).$$

For $t \geq 3$ we have the candidate $y_1 = \dots = y_t = \frac{d^r}{t}$ and check the boundary via

$$\begin{aligned} \max_{(y_1,\dots,y_t) \in \mathcal{H}_t} \sum_{i=1}^{t} \log_2(y_i + 1) &= \max_{(y_1,\dots,y_{t-1}) \in \mathcal{Y}_{t-1}} \sum_{i=1}^{t-1} \log_2(y_i + 1) \\ &= (t-1) \cdot \log_2 \left(\frac{d^r}{t-1} + 1 \right) < t \cdot \log_2 \left(\frac{d^r}{t} + 1 \right), \end{aligned}$$

where the second equality uses the induction hypothesis and the last inequality is due to the increasing monotonicity of the function $x \mapsto x \log_2 \left(\frac{1}{x} + 1 \right)$ for all $x > 0$, see, e.g., [30, Theorem 140]. $\square$

4.5 Discussion of Theorem 1

Remarks on implicit assumptions. According to the remark about the connection of forward and backward direction below, it suffices to focus only on one direction of the scheme when attempting to reach maximal algebraic degree. We focus on the forward direction. Furthermore, our analysis is independent of the concrete instantiation of the linear layer, besides the fact that we assume the matrix M ensures full diffusion after a finite number of rounds. Implicitly, our proof assumes the strongest possible linear layer, i.e., a linear layer that guarantees full diffusion after one round. Therefore, depending on the instantiation of the linear layer, the algebraic degree might grow slower than we predict, but never faster. Theorem 1 can easily be generalized to the case in which the S-Boxes are defined via different invertible functions, under the assumption that they all have the same univariate degree d and the same algebraic degree δ.

Forward versus Backward Direction. As originally proved in Corollary 3 of [14], given a fixed key k, the algebraic degrees of E_k^r and its compositional inverse E_k^{-r} are related in a particular way: the algebraic degree of E_k^r is maximal (i.e. $n \cdot t - 1$) if and only if the algebraic degree of E_k^{-r} is maximal. As an immediate consequence we state the following observation: *the number of rounds to reach maximal algebraic degree in encryption and decryption direction is the same.* This fact is particularly surprising if one direction of an SPN scheme is defined via low-degree S-Boxes, while the inverse direction is built from S-Boxes of high degree. For example, for the S-Box function $S(x) = x^3$ over $\mathbb{F}_{2^n}$ the inverse function is given by $S^{-1}(x) = x^{(2^{n+1}-1)/3}$. Here, S has algebraic degree 2, while S^{-1} has algebraic degree $(n+1)/2$.

Relation to Iterated Even–Mansour Schemes. The authors of [24] state in Section 3.3 that for an iterated Even–Mansour scheme whose round function can be described by a low-degree polynomial that

> "*[...] if the round function can be described by a polynomial of low univariate degree d over $\mathbb{F}_{2^n}$, we expect a linear behavior in* [the algebraic degree] *$\delta_{lin}(r)$: $\delta_{lin}(r) \leq \lfloor \log_2(d^r + 1) \rfloor \approx r \cdot \log_2(d)$*".

However, no formal proof of this expectation is given in [24]. Our Theorem 1 comprises this situation as special case $t = 1$; thus we not only prove but also generalize the result in [24]. Indeed, in Theorem 1 the case $t = 1$ corresponds to iterated Even–Mansour schemes and hence the algebraic degree $\delta(r)$ after r rounds is upper bounded by $\log_2(d^r + 1)$.

Minimum Number of Rounds. The minimum number of rounds for preventing higher-order differential distinguishers predicted by Theorem 1 is given by the implicit condition

$$t \cdot \log_2\left(\frac{d^r}{t} + 1\right) \geq n \cdot t - 1 \approx n \cdot t,$$

which gives

$$r \geq \log_d(t) + \log_d(2^n - 1).$$

This lower bound matches the number of rounds given in Eq. (14).

Comparison with the Interpolation Attack. The previous bound on the necessary number of rounds matches the number of rounds needed to guarantee security against the interpolation attack [35] introduced by Jakobsen and Knudsen at FSE 1997. The goal of an interpolation attack is to construct the polynomial that describes the encryption or decryption function. Hence, if the number of monomials is too large, such a polynomial cannot be constructed faster than via a brute force attack. Since the number of monomials can be estimated given the degree of the function, the designers must guarantee that the polynomial that represents the scheme is of maximum degree and full (or at least dense) in order to guarantee security against such attack.

4.6 A Tighter Bound on the Algebraic Degree by Combining [15] and Theorem 1

In some cases it is possible to slightly improve our bound on the algebraic degree from Theorem 1 by combining our result with the one presented in [15, Theorem 2]. In particular, our bound tells us that the algebraic degree after r rounds is upper bounded by $\delta(r) \leq t \cdot \log_2\left(\frac{d^r}{t} + 1\right)$. When $\delta(r)$ is close to its maximum (namely, $n \cdot t - 1$), it is possible that our bound becomes worse than the one from [15]. Working exactly as in Proposition 1, we ask for the maximum number of rounds R' such that our bound is better than $\mathcal{R}_{[\text{BCD11}]}$, yielding the implicit condition

$$t \cdot \log_2\left(\frac{d^{R'}}{t} + 1\right) = N - \frac{N - t \cdot \log_2\left(\frac{d^{R'-1}}{t} + 1\right)}{\gamma}.$$

In other words, for all $r \geq R'$ the bound from [15] is better than our linear bound from Theorem 1. Replacing R_0 and δ^{R_0} in the proof of Proposition 1 by R' and $\varDelta$, respectively, we deduce that the algebraic degree of E_k^r after r rounds, denoted by $\delta(r)$, is upper-bounded by

$$\delta(r) \leq \begin{cases} \delta^r & r \leq R_{\text{exp}}\ , \\ t \cdot \log_2\left(\frac{d^r}{t} + 1\right) & R_{\text{exp}} < r \leq R', \\ \frac{\varDelta}{\gamma^{r-R'}} + N \cdot \left(1 - \frac{1}{\gamma^{r-R'}}\right) & R_{\text{linear}} < r \leq R'_{\text{SPN}}. \end{cases} \tag{19}$$

Here

$$\varDelta := t \cdot \log_2\left(\frac{d^{R'}}{t} + 1\right)$$

is the upper bound for the algebraic degree after R' rounds according to Theorem 1,

$$R'_{\text{SPN}} := R' + \left\lceil \log_\gamma\left(N - \varDelta\right)\right\rceil,$$

is the minimum number of rounds for preventing higher-order differential distinguishers and

$$R_{\exp} = 1 + \lfloor \log_\delta(t) \rfloor.$$

is the number of rounds with exponential growth according to Lemma 2. A concrete example where a combination of our bound and the bound in [15] improves upon our bound is given in Fig. 2 for the case of a strong-aligned SPN scheme over $(\mathbb{F}_{2^{33}})^{32}$ with cubing S-Box $S(x) = x^3$. We observe, when the algebraic degree is close to its maximum value $n \cdot t - 1 = 1055$, the bound from [15] is better than our bound from Theorem 1. As a result, one more round than our bound predicts (that is, 25 instead of 24) is needed to reach the maximum algebraic degree.

4.7 Comparison of Theorem 1 with the Results in [15]

For a better insight when the bound $\mathcal{R}_{\text{SPN}}$ improves upon the one given by $\mathcal{R}_{[\text{BCD11}]}$ we ask the following question: *For which values of n, t, d and δ is*

$$\mathcal{R}_{\text{SPN}} \geq \mathcal{R}_{[\text{BCD11}]}$$

satisfied? Substituting the corresponding expressions we obtain the following inequality

$$\log_d(2^n - 1) + \log_d(t) \geq \left\lfloor \log_\delta\left(N \cdot \frac{\gamma - 1}{\gamma \cdot \delta - 1}\right)\right\rfloor + \left\lceil \log_\gamma\left(N \cdot \frac{\gamma \cdot (\delta - 1)}{\gamma \cdot \delta - 1}\right)\right\rceil.$$

Using the relations $\gamma \cdot \delta - 1 \geq \gamma - 1$ and $\gamma \cdot \delta - 1 \geq \delta - 1$ (note that $\delta \geq 2$), an upper bound for $\mathcal{R}_{[\text{BCD11}]}$ is given by

$$\mathcal{R}_{[\text{BCD11}]} \leq 1 + \lfloor \log_\delta(N) \rfloor + \lceil \log_\gamma(N) \rceil \leq 1 + \lceil \log_\delta(N) \rceil + \lceil \log_2(N) \rceil.$$

Focusing on the case $n \gg 1$, the condition $\mathcal{R}_{\text{SPN}} \geq \mathcal{R}_{[\text{BCD11}]}$ is satisfied if (approximately)

$$\log_d(2^n - 1) + \log_d(t) \approx n \cdot \log_d(2) + \log_d(t) \geq 1 + \log_\delta(n \cdot t) + \log_2(n \cdot t),$$

or to put it another way, if

$$\underbrace{n \cdot \log_d(2) + \log_d(t)}_{\in \mathcal{O}(n)} \geq \underbrace{(\log_2(n) + \log_2(t)) \cdot \left(1 + \frac{1}{\log_2(\delta)}\right) + 1}_{\in \mathcal{O}(\log_2(n))}.$$

It is easy to see that for any fixed values of d, δ, and t, the previous inequality can be satisfied if n is large enough.

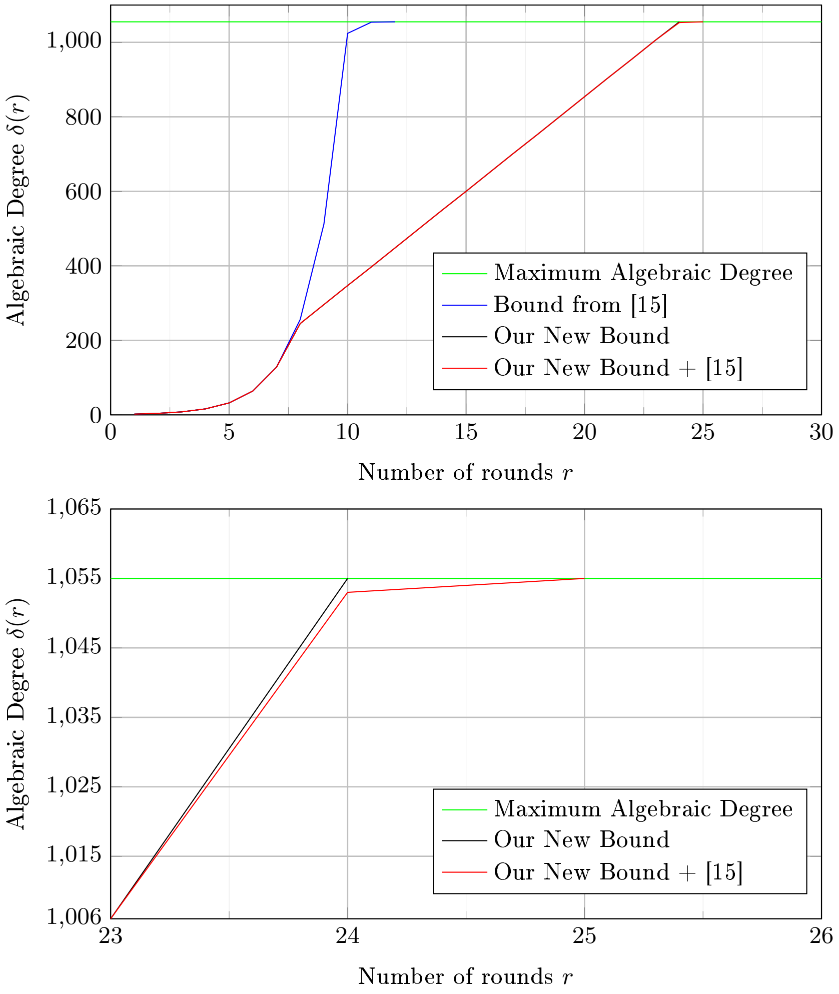

Fig. 2: Comparison between our new bound and the tighter one obtained by combining our bound with the one proposed in [15] for the case of an SPN scheme instantiated over $(\mathbb{F}_{2^{33}})^{32}$ with a cube S-Box $S(x) = x^3$ (where $d = 3, \delta = 2$ and $\gamma = (n+1)/2 = 17$).

A Concrete Example: $S(x) = x^3$*.* As a concrete example, we consider the cube S-Box $S(x) = x^3$. Here, we compute a lower bound for R_0 from Eq. (11), i.e., a lower bound on the number of rounds for which the growth of the algebraic degree is exponential when compared to the estimation given in [15]. By definition of

R_0 it holds

$$\begin{aligned} R_0 &= \log_\delta\left(N \cdot \frac{\gamma-1}{\gamma\cdot\delta-1}\right) = \log_\delta(N) + \log_\delta\left(\frac{\gamma-1}{\gamma\cdot\delta-1}\right) \geq \\ &= \log_\delta(N) + \log_\delta\left(\frac{\gamma-1}{\gamma\cdot\delta}\right) = \log_\delta(N) + \log_\delta\left(\frac{1}{\delta}\right) + \log_\delta\left(\frac{\gamma-1}{\gamma}\right) = \\ &= \log_\delta(N) - 1 + \log_\delta\left(1-\frac{1}{\gamma}\right). \end{aligned}$$

For the case $S(x) = x^3$, we can assume $\gamma = \frac{n+1}{2}$ (see Eq. (21) for a detailed argument supporting this point). It follows that

$$\log_\delta\left(1-\frac{1}{\gamma}\right) = \log_\delta\left(1-\frac{2}{n+1}\right) \geq -\log_\delta(2) = -1$$

where we used that $n \geq 3$ and $\delta = 2$. Hence, the growth of the algebraic degree is exponential for

$$R_0 \geq \log_2(N) - 2 = \log_2(t) + \log_2(n) - 2$$

rounds. According to Eq. (15), our bound suggests that exponential growth occurs for (at least) $R_{\text{exp}} = 1 + \log_2(t)$ rounds.[8]

It is easy to observe that

$$R_0 \geq R_{\text{exp}} + \log_2(n) - 3,$$

which implies that our bound improves the one provided in [15] for a cube S-Box over a very large domain (namely, $n \gg 8$).

5 Practical Results

In this section, we present our practical results on SPN schemes over $(\mathbb{F}_{2^n})^t$ (defined as in Section 4) with low-degree and large S-Boxes. Since the approach we take is the same for all of our tests, we will first describe it.

5.1 Test Methodology

Instead of computing the ANF of a (keyed or keyless) permutation (which is quite expensive already for small field sizes[9]), we evaluate the zero-sum property for multiple random input vector spaces.

[8] Based on our bound, the exponential bound holds for $R_{\text{exp}} + r$ rounds if $2^{1+\lfloor\log_2(t)\rfloor+r} \leq t \cdot \log_2(3^{1+\lfloor\log_2(t)\rfloor+r}/t + 1)$. E.g., the exponential bound holds for two more rounds if (approximately) $8 + \log_2(t) \leq \log_2(3) \cdot (3 + \log_2(t))$, which can be satisfied for *large* t (namely, $t \geq 50$).

[9] For example, the computation of the Möbius transform is exponential in the bit size [7], and other methods (like the symbolic evaluation of the multiplication) are only feasible for small n or large n with small d (i.e., a small number of multiplications).

Algorithm 1: Evaluating the zero sum property of an SPN scheme E_k^r over $(\mathbb{F}_{2^n})^t$ using different input subspaces.

```
Data: SPN scheme $E_k^r$ using $r$ rounds, with S-Box size $n$ and $t$ words,
      dimension $D$ of the subspace, number of tests $n_T$.
Result: True if a zero sum is found in all tests, False otherwise.
for $i \leftarrow 1$ to $n_T$ do
    Randomly distribute $D$ active bits among the $N = n \cdot t$ possible positions,
      resulting in the input vector space $\mathcal{V} \subseteq \mathbb{F}_2^N$.
    Randomly sample round constants $c_1, \ldots, c_r$ and $v$.
    Randomly sample key $k$.
    Fix $E_k^r$ using $c_1, \ldots, c_r$ and $k$.
    $s \leftarrow 0$.
    foreach $x \in \mathcal{V} \oplus v$ do
        $s \leftarrow s \oplus E(x)$.
    if $s \neq 0$ then
        return False.
return True.
```

For this purpose, we wrote a custom program in C that works as follows. [10] For random keys and constants, given an input subspace of dimension $D \leq N - 1$, where $N = n \cdot t$, we look for the minimum number of rounds r for which the corresponding sum of the ciphertexts is different from zero. Such a number corresponds to

(1) the minimum number of rounds for reaching algebraic degree $\delta = D + 1$, and
(2) the minimum number of rounds for preventing higher-order differential distinguishers for $D = N - 1$.

To avoid a bias by weak keys or "bad" round constants, we have repeated the tests multiple times (with new random keys, round constants, and input subspaces).

We illustrate the approach in Algorithm 1 using a keyed permutation.

Number of Subspaces of Dimension D. We emphasize, if the algebraic degree of an SPN scheme E_k^r after r rounds is $\delta(r)$, then summing over all evaluations from any vector space of dimension $D \geq \delta(r) + 1$ always results in a zero sum, i.e., $\bigoplus_{x \in \mathcal{V}} E_k^r(x \oplus v) = 0$ for a generic (fixed) v. However, the converse is not true in general. That is, having a zero sum over a vector space of dimension D, does in general not imply that the algebraic degree is $\delta(r) = D - 1$. Indeed, $\delta(r)$ could be higher, and the zero sum could occur merely due to the specific structure of the vector space and the analyzed function.

Evaluating the zero sum property for all affine subspaces of dimension D is actually infeasible. Indeed, when working over $(\mathbb{F}_p)^N$, for any prime p and

[10] The code we used for the practical tests can be found on GitHub: `https://github.com/IAIK/higher-order-differential`

$N \in \mathbb{N}$, the number of different subspaces of dimension $D \leq N$ is

$$\frac{(p^N - 1) \cdot (p^N - p) \cdot (p^t - p^2) \cdot \dots \cdot (p^N - p^{D-1})}{(p^D - 1) \cdot (p^D - p) \cdot (p^D - p^2) \cdot \dots \cdot (p^D - p^{D-1})} \in \mathcal{O}\left(p^{D \cdot (N-D)}\right)$$

as shown e.g. in [32], which is out of practical range even for small values of p, N, D. For this reason, we have to limit ourselves to evaluate the zero sum property for a limited number of subspaces only. However, in our practical tests we observed that a small number of tests *for each of the possible combinations* of active bits is sufficient to derive a stable number (e.g., around 10 tests for each combination). Indeed, for example, we observed no differences when using an input subspace of dimension $N - 1$ and changing the position of the single inactive bit in multiple tests.

The practical number of rounds to prevent higher-order differential distinguishers we report is *the smallest number of rounds among all tested keys and round constants*. This means that potentially a higher number of rounds can be attacked by choosing the keys and round constants in a particular way.

Randomization of Active Bits. Depending on the position of the active bits, the final results may be very different. For example, significant differences arise when considering a fixed number of active bits in a single word and the same number of active bits split over multiple words. In order to counteract this problem, we choose the input subspaces randomly such that the position of active bits is also randomized. As a concrete example, consider $t = 2$ with $d = 3$ and arbitrary n. Clearly, after one round the algebraic degree is upper-bounded by $\delta = 2$, and indeed, when activating 2 bits in the same word, we do not get a zero sum. However, if we activate one bit in each of the two words (i.e., in total also 2 bits), we do get a zero sum, since only products of at most $\delta = \text{hw}(d) = 2$ bit variables from the *same* word occur in the polynomial representation. Hence, we randomize the input subspaces in our tests.

Computational Cost in Practice. In our practical tests we observed that with very few trials we already reach a stable number for the algebraic degree after a certain number of rounds. It is however crucial to test every possible combination of active words, since this has a significant impact on the final result. Concretely, we fix the number of tests to 100 for "feasible" numbers of active bits (i.e., around 30). For the larger tests, we fix the number to 10. While this may seem like a small sample size, we could not observe any differences when testing more often with lower numbers of bits. As for the concrete runtime, it largely depends on the number of active bits, but also on additional properties like the tested degree. E.g., x^3 can be evaluated faster than x^7 for a given S-Box input x. Practically, a test with 30 active bits can thus take several hours depending on the concrete tested construction.

5.2 Practical Results for S-Boxes of the form $S(x) = x^d$

In our experiments, we focus on a SHARK-like scheme [38] with power maps as S-Box functions. More specifically, we focus on SPN schemes over $(\mathbb{F}_{2^n})^t$ where

Table 1: Theoretical *lower* bound and practical number of rounds for preventing higher-order differential distinguishers on SPN schemes over $(\mathbb{F}_{2^n})^t$ for several values of n and $t \geq 2$ (where $N = n \cdot t$). The chosen S-Box is the cube function $S(x) = x^3$. For the practical number of rounds, we consider both the case of an MDS matrix and the case of a matrix that provides the "worst" possible diffusion (e.g., a sparse matrix as in Eq. (22)). $\mathcal{R}_{[\text{BCD11}]}$ is computed assuming $\gamma = (n+1)/2$.

Parameters			Theoretical # of Rounds		Practical # of Rounds	
N	n	t	$\mathcal{R}_{\text{SPN}}$	$\mathcal{R}_{[\text{BCD11}]}$	MDS matrix	Sparse matrix
35	5	7	5	6	8	15
35	7	5	6	6	8	12
36	9	4	7	6	9	11
33	11	3	8	5	10	10
39	13	3	10	6	11	12
34	17	2	12	6	12	12
38	19	2	13	6	14	14
66	11	6	9	7	-	-
65	13	5	10	6	-	-
60	15	4	11	6	-	-
66	17	4	12	7	-	-
63	21	3	15	6	-	-
66	33	2	22	7	-	-
132	11	12	10	8	-	-
135	15	9	12	8	-	-
133	19	7	14	7	-	-
132	33	4	22	8	-	-
129	43	3	28	7	-	-
130	65	2	42	8	-	-

the S-Box is $S(x) = x^d$ and the mixing layer is defined by the multiplication with an invertible $t \times t$ matrix. The choice of n and d is governed by the requirement $\gcd(d, 2^n - 1) = 1$, ensuring that $S(x) = x^d$ is a permutation of $\mathbb{F}_{2^n}$.

For the S-Box $S(x) = x^3$, we report our results on the minimum number of rounds to prevent higher-order differential distinguishers in Table 1. We observe that the number of rounds that can be covered by a higher-order differential distinguisher is always close to the one predicted by our formula (in some cases a little higher, but never smaller). Moreover, especially when the size of the S-Box is not too small, the round number $\mathcal{R}_{\text{SPN}}$ predicted by our formula is significantly larger than $\mathcal{R}_{[\text{BCD11}]}$. Furthermore, our results of small-scale experiments on the growth of the algebraic degree (according to the test methodology in Section 5.1) for $S(x) = x^3$ and $S(x) = x^7$ are depicted in Fig. 3 and Fig. 4, respectively.

Note that the tests made for Table 1 and, e.g., Fig. 3 use different approaches: in the former case we maximize the number of active bits and see how many rounds we can distinguish, whereas in the latter case we want to estimate the algebraic degree via the number of active bits. For this reason, more test runs

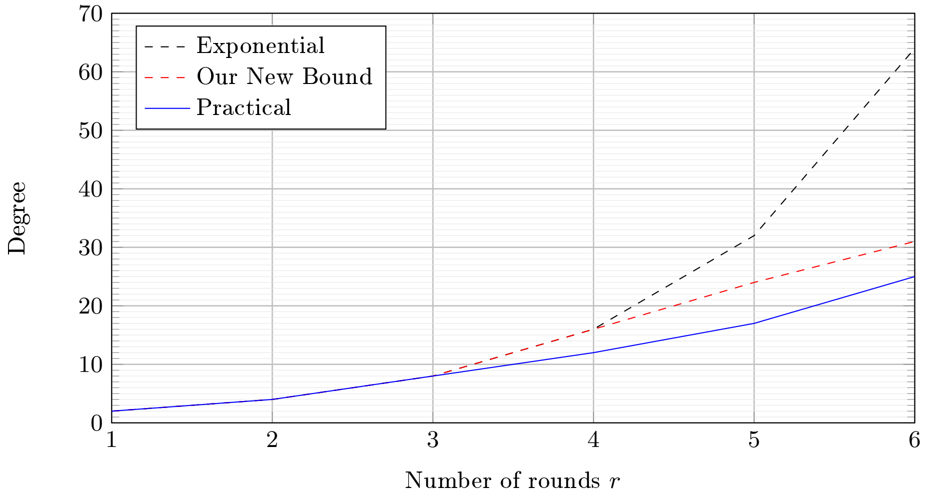

Fig. 3: Degree growth for an SPN scheme over $(\mathbb{F}_{2^{33}})^4$ instantiated with the S-Box $f(x) = x^3$.

are needed to determine the degree growth, especially in order to take care of the different positions of the active bits (where the number of choices is lower for Table 1, since $N-1$ bits are active in all tests).

Determining γ. To use the results from [15] for our comparisons we need to determine the parameter γ (see also Eq. (9)). Since an exact computation of γ is too expensive for most instances we use, we derive an upper bound on γ and use this upper bound as a benchmark. By definition of γ, it holds

$$\begin{aligned}\gamma = \max_{1\leq i\leq n-1} \frac{n-i}{n-\delta_i} &= \max\left\{\max_{1\leq i\leq q} \frac{n-i}{n-\delta_i}, \max_{q+1\leq i\leq n-1} \frac{n-i}{n-\delta_i}\right\} \\ &\leq \max\left\{\max_{1\leq i\leq q} \frac{n-i}{n-i\cdot\delta}, \max_{q+1\leq i\leq n-1} \frac{n-i}{n-(n-1)}\right\} \\ &= \max\left\{\frac{n-q}{n-q\cdot\delta}, n-(q+1)\right\}.\end{aligned} \tag{20}$$

where $q = \lfloor (n-1)/\delta \rfloor$ and $\delta = \mathrm{hw}(d)$ is the algebraic degree of the S-Box. For the particular case $S(x) = x^3$ only odd values for n are allowed (to guarantee $\gcd(2^n-1, 3) = 1$) and thus we obtain $n-1 = q\cdot 2$. Hence,

$$\gamma \leq \max\left\{\frac{n-\frac{n-1}{2}}{n-2\cdot\frac{n-1}{2}}, n-\frac{n-1}{2}-1\right\} = \frac{n+1}{2}. \tag{21}$$

We assume $\gamma = (n+1)/2$ to compute the theoretical values for $\mathcal{R}_{[\mathrm{BCD11}]}$. We also refer to [24, Lemma 3], where authors support this assumption by practical experiments for each odd $n \leq 33$.

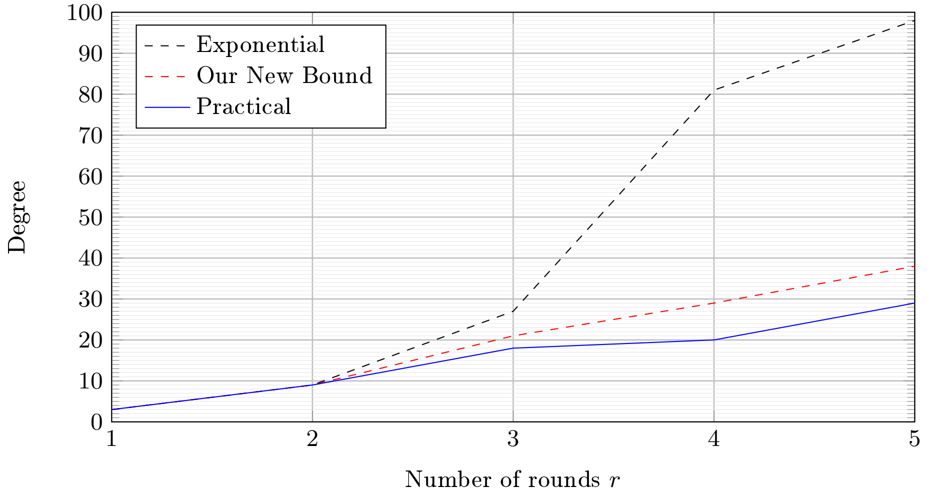

Fig. 4: Degree growth for an SPN scheme over $(\mathbb{F}_{2^{33}})^3$ instantiated with the S-Box $f(x) = x^7$.

Influence of the Linear Layer. In order to understand how the linear layer influences the number of rounds *necessary* to provide security against higher-order differential distinguishers, in our practical tests we consider two extreme cases: *(1)* we evaluate the case in which the linear layer is defined as the multiplication with an MDS matrix (for parameters n and t that allow us to do so[11]), which corresponds to the case of the "strongest" linear layer from a diffusion point of view; *(2)* we also evaluate the case in which the linear layer is "weak", which could happen if it is defined by the multiplication with a matrix containing a large number of zero coefficients. For this second case, we used a $t \times t$ matrix M with coefficients $M_{r,c}$, for $r, c = 0, \ldots, t-1$, given by

$$M_{r,c} = \begin{cases} 1 & \text{if } r = 0 \;\; \mathit{OR} \text{ if } c \equiv r+1 \mod t, \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

We note that by using M we need t rounds to have full diffusion (at word level), instead of just one round as for the MDS case. Hence, especially for large t, we expect that more rounds than previously predicted may be necessary to guarantee security against higher-order differential distinguishers. In Table 1 we report empirical evidence for this expectation: the gap between the number of rounds predicted by our formula and the one found by practical tests in the case of a sparse matrix is close to zero for "small" t, and grows for "large" t.

[11] We recall that a $t \times t$ MDS matrix with elements in $GF(2^n)$ exists if the condition $\log_2(2t+1) \leq n$ is satisfied.

6 Summary, Applications and Open Problems

Our main result of this article is a new upper bound on the growth of the algebraic degree for strong-aligned SPN schemes based on the parallel application of t copies of the same (low-degree and large) S-Box, see Theorem 1. After a decade of stagnation since the last advance in [15] and subsequently in [14], our findings extend the canon of theoretical bounds for the growth of the algebraic degree in SPN schemes by an improved bound for strong-aligned SPN schemes.

In domain specific SPN schemes with their use cases in MPC-/FHE-/ZKP-protocols it is most often algebraic cryptanalysis (such as higher-order differential distinguishers) that dominates the overall security considerations. Thus, a better understanding of the growth of the algebraic degree is not only vital for the security assessment of these schemes but also for navigating design choices towards a more solid theoretical foundation. We have practically verified our findings on small-scale SPN instances. In addition, we have compared our results to the current best bound in [15], showing a substantial improvement for strong-aligned SPN schemes with low-degree and large S-Boxes. In this setting we could observe that the bound from [15] does not provide significant advantage over the naive exponential bound (except when close to maximum algebraic degree) which is why our findings can be considered especially useful tor this area of the design space.

Applications. As a concrete application, HadesMiMC [28] is probably the most suitable candidate to apply our results. The Hades approach [28] combines both SPN and partial SPN schemes 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;
- 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 increase the overall degree of the function in an attempt to prevent algebraic attacks. In particular, if a certain number of rounds with full S-Box layer are needed to reach maximum degree, we can expect that such a number does not decrease if some rounds are replaced by rounds with partial S-Box layer. Hence, in the case in which a scheme based on the Hades strategy is designed over $(\mathbb{F}_{2^n})^t$, our results provide a lower bound on the number of rounds $R_F + R_P$ necessary to provide security in a Hades construction. In particular, note that even both HadesMiMC and Poseidon are designed over $\mathbb{F}_p^t$, there is no reason why a scheme based on the Hades strategy cannot be designed over $(\mathbb{F}_{2^n})^t$. As a concrete example, we refer to Starkad [26], a variant of Poseidon defined over $(\mathbb{F}_{2^n})^t$. In there, authors conjectured/claimed that the number of rounds necessary to provide security interpolation attack is also necessary for providing security against higher-order differential attacks. In this paper, we proved that this is indeed the case.

Moreover, our upper bound for the growth of the algebraic degree plays an important role in higher-order differential distinguishers that do not exploit the biggest non-trivial subspace, but subspaces of smaller dimension than the state size. This is not only of theoretical interest, but it applies to all cases in which the security level is smaller than the size of the full scheme, a scenario that is common for schemes recently proposed for MPC/FHE/ZK applications.

We remark, the growth of the degree also depends on how the active bits are chosen. As a concrete example, let's consider a subspace $\mathcal{V} \subseteq {\mathbb{F}_{2^n}}^t$ of the form $\mathcal{V} := \mathbb{F}_{2^n} \times \{0\} \times ... \times \{0\}$, that is a subspace in which 1 input word is active and $t-1$ are constant. By summing over any such subspace $\mathcal{V}$, we get a sum of a function in which $t-1$ variables are constant:

$$\bigoplus_{x=(x_1,x_2,\ldots,x_t)\in\mathcal{V}\oplus v} E_k^r(x) = \bigoplus_{x_1\in\mathbb{F}_{2^n}} E_k^r(x_1, v_2, v_3, \ldots, v_t),$$

where $v_2, \ldots, v_t \in \mathbb{F}_{2^n}$. Hence the algebraic degree $\delta(r) \leq \log_2(d^r + 1)$ after r rounds (until the maximum algebraic degree $n-1$ is reached).

Open Problems. *Details of the Linear Layer.* Our bounds do not take into account the details of the matrix in $\mathbb{F}_{2^n}^{t\times t}$ that defines the linear layer. As we have seen in Sect. 5.2, we expect that better bounds can be obtained by considering such details, including, e.g., the minimum number of rounds for the linear layer to provide full diffusion.

Weak-Aligned SPN Schemes. Related to the linear layer, another problem regards the generalization of our bounds to the case of weak-aligned SPN schemes. We emphasize that this is not only of theoretical interest, since several schemes proposed for MPC/FHE/ZK applications like Jarvis [6] and Vision [5] are defined in this way.

Extension to Key-Recovery Attacks. A next step could be to extend the higher-order differential distinguishers to key-recovery attacks, as was already done for MiMC in [24].

Conjectured Bound. For a permutation based on an SPN-construction over $(\mathbb{F}_{2^n})^t$ with S-Boxes of univariate degree d and algebraic degree δ (see also Eq. (5)) we conjecture the following lower bound on the number of rounds for reaching maximum algebraic degree $n \cdot t - 1$

$$R_{\text{Conj}} := \log_d(2^n - 1) + \log_\delta(t).$$

Above bound differs from $\mathcal{R}_{\text{SPN}} \approx \log_d(2^n-1) + \log_d(t)$ in the second summand, i.e., it contains $\log_\delta(t)$ instead of $\log_d(t)$. We base this conjecture on the following observation. For the first $1+\log_\delta(t)$ rounds the algebraic degree grows exponentially until there is a monomial that contains all word-level variables. We conjecture, that exponentiating this monomial repeatedly to the power of d gives an upper bound on the growth of the algebraic degree. Since we need algebraic degree n in $t-1$ variables and algebraic degree $n-1$ in one variable to reach maximum algebraic degree, we expect that $\log_d(2^n-1)-1$ more rounds are necessary, giving in total $\log_d(2^n-1)+\log_\delta(t)$ rounds. The problem to find a formal argument for supporting (or disproving) R_{Conj} is left for future research.

Acknowledgment. Authors thank Reviewers for their valuable comments. Lorenzo Grassi is supported by the European Research Council under the ERC advanced grant agreement under grant ERC-2017-ADG Nr. 788980 ESCADA.

References

1. Albrecht, M.R., Cid, C., Grassi, L., Khovratovich, D., Lüftenegger, R., Rechberger, C., Schofnegger, M.: Algebraic Cryptanalysis of STARK-Friendly Designs: Application to MARVELlous and MiMC. In: ASIACRYPT 2019. LNCS, vol. 11923, pp. 371–397 (2019)
2. Albrecht, M.R., Grassi, L., Perrin, L., Ramacher, S., Rechberger, C., Rotaru, D., Roy, A., Schofnegger, M.: Feistel Structures for MPC, and More. In: ESORICS. LNCS, vol. 11736, pp. 151–171 (2019)
3. Albrecht, M.R., Grassi, L., Rechberger, C., Roy, A., Tiessen, T.: MiMC: Efficient Encryption and Cryptographic Hashing with Minimal Multiplicative Complexity. In: ASIACRYPT 2016. LNCS, vol. 10031, pp. 191–219 (2016)
4. 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)
5. Aly, A., Ashur, T., Ben-Sasson, E., Dhooghe, S., Szepieniec, A.: Design of Symmetric-Key Primitives for Advanced Cryptographic Protocols. IACR Trans. Symmetric Cryptol. **2020**(3), 1–45 (2020)
6. Ashur, T., Dhooghe, S.: MARVELlous: a STARK-Friendly Family of Cryptographic Primitives. IACR Cryptology ePrint Archive, Report 2018/1098 (2018)
7. Barbier, M., Cheballah, H., Bars, J.L.: On the computation of the möbius transform. Theor. Comput. Sci. **809**, 171–188 (2020)
8. Barkan, E., Biham, E.: In How Many Ways Can You Write Rijndael? In: ASIACRYPT 2002. LNCS, vol. 2501, pp. 160–175 (2002)
9. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY Family of Block Ciphers and Its Low-Latency Variant MANTIS. In: Advances in Cryptology - CRYPTO 2016. LNCS, vol. 9815, pp. 123–153 (2016)
10. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY Family of Block Ciphers and its Low-Latency Variant MANTIS. Cryptology ePrint Archive, Report 2016/660 (2016), `https://eprint.iacr.org/2016/660`
11. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: On alignment in Keccak. `https://keccak.team/files/KeccakAlignment.pdf`
12. 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. In: CRYPTO 2020. LNCS, vol. 12172, pp. 299–328 (2020)
13. Biryukov, A., Khovratovich, D., Perrin, L.: Multiset-Algebraic Cryptanalysis of Reduced Kuznyechik, Khazad, and secret SPNs. IACR Trans. Symmetric Cryptol. **2016**(2), 226–247 (2016)

14. Boura, C., Canteaut, A.: On the influence of the algebraic degree of f^{-1} on the algebraic degree of $g \circ f$. IEEE Trans. Information Theory **59**(1), 691–702 (2013)
15. Boura, C., Canteaut, A., De Cannière, C.: Higher-Order Differential Properties of Keccak and *Luffa*. In: FSE 2011. LNCS, vol. 6733, pp. 252–269 (2011)
16. Canteaut, A., Lallemand, V., Leander, G., Neumann, P., Wiemer, F.: BISON Instantiating the Whitened Swap-Or-Not Construction. In: EUROCRYPT 2019. LNCS, vol. 11478, pp. 585–616 (2019)
17. Canteaut, A., Videau, M.: Degree of Composition of Highly Nonlinear Functions and Applications to Higher Order Differential Cryptanalysis. In: EUROCRYPT 2002. LNCS, vol. 2332, pp. 518–533 (2002)
18. Carlet, C.: Graph Indicators of Vectorial Functions and Bounds on the Algebraic Degree of Composite Functions. IEEE Trans. Inf. Theory **66**(12), 7702–7716 (2020)
19. Carlet, C., Charpin, P., Zinoviev, V.A.: Codes, Bent Functions and Permutations Suitable For DES-like Cryptosystems. Designs, Codes Cryptography **15**(2), 125–156 (1998)
20. Chase, M., Derler, D., Goldfeder, S., Orlandi, C., Ramacher, S., Rechberger, C., Slamanig, D., Zaverucha, G.: Post-quantum zero-knowledge and signatures from symmetric-key primitives. In: CCS. pp. 1825–1842. ACM (2017)
21. Daemen, J., Rijmen, V.: The Design of Rijndael: AES - The Advanced Encryption Standard. Information Security and Cryptography, Springer (2002)
22. Dinur, I., Shamir, A.: Cube Attacks on Tweakable Black Box Polynomials. In: EUROCRYPT 2009. LNCS, vol. 5479, pp. 278–299 (2009)
23. Dobraunig, C., Grassi, L., Guinet, A., Kuijsters, D.: Ciminion: Symmetric Encryption Based on Toffoli-Gates over Large Finite Fields. Cryptology ePrint Archive, Report 2021/267 (2021), `https://eprint.iacr.org/2021/267` – accepted at EUROCRYPT 2021
24. Eichlseder, M., Grassi, L., Lüftenegger, R., Øygarden, M., Rechberger, C., Schofnegger, M., Wang, Q.: An Algebraic Attack on Ciphers with Low-Degree Round Functions: Application to Full MiMC. In: ASIACRYPT 2020. LNCS, vol. 12491, pp. 477–506 (2020)
25. Funabiki, Y., Todo, Y., Isobe, T., Morii, M.: Improved Integral Attack on HIGHT. In: ACISP 2017. LNCS, vol. 10342, pp. 363–383 (2017)
26. Grassi, L., Khovratovich, D., Rechberger, C., Roy, A., Schofnegger, M.: Starkad and Poseidon: New Hash Functions for Zero Knowledge Proof Systems. Cryptology ePrint Archive, Report 2019/458 (2019), Version: 20200205:104144, `https://eprint.iacr.org/2019/458`
27. Grassi, L., Khovratovich, D., Rechberger, C., Roy, A., Schofnegger, M.: Poseidon: A New Hash Function for Zero-Knowledge Proof Systems. In: 30th USENIX Security Symposium (USENIX Security 21). USENIX Association (2021)
28. Grassi, L., Lüftenegger, R., Rechberger, C., Rotaru, D., Schofnegger, M.: On a Generalization of Substitution-Permutation Networks: The HADES Design Strategy. In: EUROCRYPT 2020. LNCS, vol. 12106, pp. 674–704 (2020)
29. Grassi, L., Rechberger, C., Rotaru, D., Scholl, P., Smart, N.P.: MPC-friendly symmetric key primitives. In: CCS. pp. 430–443. ACM (2016)
30. Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities. Cambridge University Press, 2 edn. (1952)
31. Hebborn, P., Lambin, B., Leander, G., Todo, Y.: Lower Bounds on the Degree of Block Ciphers. In: ASIACRYPT 2020. LNCS, vol. 12491, pp. 537–566 (2020)
32. Hogben, L.: Handbook of Linear Algebra. CRC Press, 2nd edn. (2016)

33. Hu, K., Sun, S., Wang, M., Wang, Q.: An Algebraic Formulation of the Division Property: Revisiting Degree Evaluations, Cube Attacks, and Key-Independent Sums. In: ASIACRYPT 2020. LNCS, vol. 12491, pp. 446–476 (2020)
34. Hughes-Hallett, D., McCallum, W., Flath, D., Gleason, A., Kalaycioglu, S., Lahme, B., Lock, P., Lozano, G., Morris, J., Mumford, D., et al.: Calculus: Multivariable. Wiley, 7 edn. (2017)
35. Jakobsen, T., Knudsen, L.R.: The Interpolation Attack on Block Ciphers. In: FSE 1997. LNCS, vol. 1267, pp. 28–40 (1997)
36. Knudsen, L.R.: Truncated and Higher Order Differentials. In: FSE 1994. LNCS, vol. 1008, pp. 196–211 (1994)
37. Lai, X.: Higher Order Derivatives and Differential Cryptanalysis, pp. 227–233. Springer US (1994)
38. Rijmen, V., Daemen, J., Preneel, B., Bosselaers, A., Win, E.D.: The Cipher SHARK. In: FSE 1996. LNCS, vol. 1039, pp. 99–111 (1996)
39. Todo, Y.: Structural Evaluation by Generalized Integral Property. In: EUROCRYPT 2015. LNCS, vol. 9056, pp. 287–314 (2015)
40. Todo, Y., Morii, M.: Bit-Based Division Property and Application to Simon Family. In: FSE 2016. LNCS, vol. 9783, pp. 357–377. Springer (2016)
41. Wang, Q., Hao, Y., Todo, Y., Li, C., Isobe, T., Meier, W.: Improved Division Property Based Cube Attacks Exploiting Algebraic Properties of Superpoly. In: CRYPTO 2018. LNCS, vol. 10991, pp. 275–305 (2018)

A Proof of Proposition 1

Proof. Applying the naive bound from Eq. (1) and the bound from [15, Theorem 2] (see Eq. (8)) to $E_1 = L_1 \circ F$ yields

$$\deg(L_1 \circ F) \leq \min\left\{\delta, N \cdot \left(1 - \frac{1}{\gamma}\right) + \frac{1}{\gamma}\right\} = \delta.$$

The last equality is justified as follows: for $t = 1$, this is obvious (δ is exactly the degree of 1 round). For $t \geq 2$, this follows from the fact that the non-linear layer has degree δ (since we have parallel independent S-Boxes with algebraic δ) and that the linear layer does not change the algebraic degree.

In other words, for at least one round the naive bound from Eq. (1) for the growth of the algebraic degree is better than the bound in [15]. Therefore, we now look for the maximum number of rounds R_0 with this behavior. This corresponds to solving the following equation for R_0

$$\delta^{R_0} = N \cdot \left(1 - \frac{1}{\gamma}\right) + \frac{\delta^{R_0 - 1}}{\gamma},$$

which gives

$$R_0 = \log_\delta\left(N \cdot \frac{\gamma - 1}{\gamma \cdot \delta - 1}\right).$$

To put it another way, for any number of rounds $r \leq R_0$, the degree of E_r is upper-bounded by δ^r. As a next step, we find the minimum additional number of rounds to prevent higher-order differential attacks, i.e., the minimum additional number of rounds R_1 such that the algebraic degree after $R_0 + R_1$ rounds is $N - 1$ (the biggest non-trivial subspace of $\mathbb{F}_2^N$ has dimension $N - 1$).

For $r > R_0$, the bound in [15] is better than the naive bound, hence, the algebraic degree of E_r after $r = R_0 + 1$ rounds is upper-bounded by

$$\deg\left(E_{R_0+1}\right) \leq \underbrace{N \cdot \left(1 - \frac{1}{\gamma}\right)}_{=:C} + \frac{\delta^{R_0}}{\gamma} = C + \frac{\delta^{R_0}}{\gamma},$$

and after $r = R_0 + 2$ rounds by

$$\deg\left(E_{R_0+2}\right) \leq C + \frac{1}{\gamma} \cdot \left(C + \frac{\delta^{R_0}}{\gamma}\right) = C + \frac{C}{\gamma} + \frac{\delta^{R_0}}{\gamma^2}.$$

Continuing this way, we conclude that after $r = R_0 + s$ rounds, for an integer $s \geq 1$, the algebraic degree is upper bounded by

$$\deg\left(E_{R_0+s}\right) \leq \frac{\delta^{R_0}}{\gamma^s} + C \cdot \sum_{i=0}^{s-1} \frac{1}{\gamma^i} = \frac{\delta^{R_0}}{\gamma^s} + C \cdot \frac{1 - \frac{1}{\gamma^s}}{1 - \frac{1}{\gamma}} = \frac{\delta^{R_0}}{\gamma^s} + N \cdot \frac{\gamma^s - 1}{\gamma^s}.$$

This means, the minimum additional number of rounds R_1 to prevent higher-order differential distinguishers is given by the implicit condition

$$\frac{\delta^{R_0}}{\gamma^{R_1}} + \frac{N \cdot (\gamma^{R_1} - 1)}{\gamma^{R_1}} = N - 1,$$

which gives

$$R_1 = \log_\gamma \left(N - \delta^{R_0} \right).$$

We conclude, the minimum number of rounds $\mathcal{R}_{[\mathrm{BCD11}]}$ to prevent higher-order differential distinguishers is given by

$$\mathcal{R}_{[\mathrm{BCD11}]} = \left\lfloor \log_\delta \left(N \cdot \frac{\gamma - 1}{\gamma \cdot \delta - 1} \right) \right\rfloor + \left\lceil \log_\gamma \left(N - \delta^{R_0} \right) \right\rceil.$$

B Lagrange Multipliers

A classical method of solving equality constrained optimization problems is the method of Lagrange multipliers. It is based on the following result, which can be found in many canonical books about multivariable calculus. We refer to [34, Sec. 15.3] for a more detailed discussion.

Theorem 2. *Let $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be continiuously differentiable functions on an open set D. Let $S := \{x \in D \ : \ g(x) = 0\} \subseteq D$ be the set of all points $x \in D$ satisfying the constraint $g(x) = 0$. If $x_0 \in S$ is a point such that*

(i) f has a local extremum on S at x_0,
(ii) $\nabla g(x_0) \neq 0$,

then there is an element $\lambda \in \mathbb{R}$ with

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

Theorem 2 provides a method to find candidates for local extreme values of f on a specific region of the domain of f. This specific region is defined by an equality constraint, i.e., the zero locus of a *constraint function g*. Often, the function

$$L(x_1, \ldots, x_n, \lambda) := f(x_1, \ldots, x_n) - \lambda \cdot g(x_1, \ldots, x_n)$$

is called the *Lagrangian function* and the variable λ is referred to as the *Lagrange multiplier*. If the constraint set S is a closed and bounded set, then f assumes its (unique) global minimum and maximum values on S. For determining those global extreme values of f on S we can use Theorem 2 and proceed as follows:

1. Find candidates for global extrema of f on S by determining all solutions $(x_1, \ldots, x_n) \in S$ of the equation system

   $$\frac{\partial}{\partial x_1} L(x_1, \ldots, x_n, \lambda) = \cdots = \frac{\partial}{\partial x_n} L(x_1, \ldots, x_n, \lambda) = \frac{\partial}{\partial \lambda} L(x_1, \ldots, x_n, \lambda) = 0.$$

   This brings forth all candidates for which ∇g is well-defined.

2. Check separately if f assumes a global extremum on the *boundary points* of S. The boundary points of S are those points for which ∇g is not well-defined, i.e., the "endpoints" of S.
3. If any, also check the points $(x_1', \ldots, x_n') \in S$ with $\nabla g(x_1', \ldots, x_n') = (0, \ldots, 0)$.

Example 1. – The constraint set $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ has no boundary points.

- The set $\{(x, y) \in \mathbb{R}^2 : x, y \geq 0, x + y = 1\}$ has the boundary points $(0, 1)$ and $(1, 0)$.
- The set $\{(x, y, z) \in \mathbb{R}^3 : x, y, z \geq 0, x + y + z = 1\}$ is a plane that intersects the coordinate planes of the first octant in the 3 line segments

$$S_1 \coloneqq \{(x, y, z) \in \mathbb{R}^3 \ : \ x, y \geq 0, z = 0, x + y = 1\}$$

and

$$S_2 \coloneqq \{(x, y, z) \in \mathbb{R}^3 \ : \ x, z \geq 0, y = 0, x + z = 1\}$$

and

$$S_3 \coloneqq \{(x, y, z) \in \mathbb{R}^3 \ : \ y, z \geq 0, x = 0, y + z = 1\}.$$

Thus the boundary points are given by $S_1 \cup S_2 \cup S_3$.

C Partial-SPN Schemes: A Detailed Analysis

C.1 Observations on Partial-SPN Schemes

Here we briefly present some observations in the case of Partial-SPN schemes, which are a direct application of the just recalled result for iterated Even-Mansour scheme. In partial SPN schemes over $(\mathbb{F}_{2^n})^t$, the round function is defined as

$$R = M(\underbrace{S_1, \ldots, S_s, I, \ldots I}_{t \text{ words}}), \tag{23}$$

for $1 \leq s < t$, where the $S_i : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ denote some non-linear functions on $\mathbb{F}_{2^n}$ and where I denotes the identity function on $\mathbb{F}_{2^n}$. The function $M : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ is defined as the multiplication of the state vector with a $t \times t$-matrix over $\mathbb{F}_{2^n}$. The distinctive feature of a partial SPN scheme is that the non-linear functions S_i are only applied to part of the state, while the rest of the state remains unchanged. To set up higher-order differential distinguishers on partial SPN schemes, it is possible to employ several strategies. The two most extreme ones are:

1. use a proper initial subspace with "maximum" dimension $n \cdot t - 1$ bits;
2. use a proper initial subspace of dimension $n \cdot r$ bits for a certain $1 \leq r \leq t - 1$. If the choice of such an initial subspace is made in a clever way (e.g., as described in the following), it is possible to "skip" at most r initial rounds (in the sense that the input of the S-Box in the first r rounds is always constant, hence the algebraic degree does not change in the first r rounds).

Depending on the details of the scheme, *each of the previous strategies can be the best one*, in the sense that it can cover the highest number of rounds with a higher-order differential distinguisher. For SPN schemes, the first strategy is in general the only possible, since it is not possible to "skip" rounds for free without activating any S-Box. Here, we focus on the second strategy for the particular case $s = 1$, that is, the case in which only a single S-Box is applied in each round (which e.g. corresponds to the instantiation of LowMC used in variants of the Picnic signature scheme [20] and to the partial rounds in the recently proposed HadesMiMC permutation [28,27]). By applying our results from Theorem 1, we can deduce the following proposition.[12]

Proposition 3. *Let $E_k^r : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ be a r-round partial SPN scheme in which the S-Box layer is composed of one S-Box and $t-1$ identity functions. Let us assume that the S-Box is defined as some invertible (low degree) polynomial function over $\mathbb{F}_{2^n}$ of the form $S(x) := \rho_0 + \sum_{i=1}^{d} \rho_i \cdot x^i$ of degree $d \geq 3$ and with $\rho_i \in \mathbb{F}_{2^n}$, $\rho_d \neq 0$. Then at least*

$$\mathcal{R}_{\text{PSPN}} := t - 1 + \lceil \log_d(2^n - 1) \rceil$$

rounds are necessary to prevent higher-order differential distinguishers, independently of the (secret or publicly known) key k. Finally, to prevent higher-order differential distinguishers on $(t-1) \cdot n$ bits (namely, all bits except the n ones corresponding to the position of the S-Box), one more round is necessary.

Proof. Without loss of generality, we assume that the S-Box is applied to the first word[13]. In such a case, it is possible to "skip" $t-1$ rounds (namely, to impose that no S-Box is active in the first $t-1$ rounds) if the initial input $x = (x_0, x_1, \ldots, x_{t-1})$ satisfies the condition

$$\forall i = 0, \ldots, t-2 : \qquad [M^i \cdot x]_0 = \text{ constant}$$

where M^i is just the i-fold product of M, with M^0 being the identity matrix I, and where for $y = (y_0, y_1, \ldots, y_{t-1})$ the expression $[y]_0 := y_0$ denotes the word at position 0. Hence the first $t-1$ rounds do not increase the degree.

After at least t rounds (we remark that depending on the details of the linear layer it is potentially possible to skip more rounds), the S-Box is active. Since only one word is active, we can simply reuse the results for the Even–Mansour case: a necessary condition to guarantee security is that the algebraic degree is at least n, which happens only in the case where a monomial of the form x^{2^n-1} appears in the encryption polynomial. Since the degree of the monomial over $\mathbb{F}_{2^n}$ grows as fast as $d^{r-(t-1)}$, it follows that the number of rounds r must satisfy $d^{r-(t-1)} \geq 2^n - 1$ to prevent higher-order differential distinguishers. $\square$

[12] For completeness, we mention that a similar proof can be obtained by considering the result presented in [12].

[13] If this is not the case, it is always possible to find an equivalent representation – via a different but equivalent linear layer – for which this is the case.

This result is confirmed by our practical tests given in the following. We note, the result in Proposition 3 is less surprising than the corresponding one given for SPN schemes in Proposition 2. This is because the linear growth of the algebraic degree in partial SPN schemes has already been observed in the literature when considering the security of LowMC [4] and Bison [16]. For example, quoting from [16, Sect. 6.2]:

> "[...] the degree of any NLFSR increases linearly with the number of rounds. To the best of our knowledge, this is the first time this have been observed in this generality. We like to add that this is in sharp contrast to how the degree increases for SPN ciphers. For SPN ciphers the degree usually increases exponentially until a certain threshold is reached."

Moreover, note that the number of rounds necessary to prevent a zero sum on *all* $N = n \cdot t$ bits is not in general necessary to prevent a zero-sum on fewer than all bits. For example, consider the extreme case of a function over $\mathbb{F}_2^N$ with ANF $y_0 = x_0 + \prod_{i=1}^{n-1} x_i$ and $y_i = x_i$ for $1 \leq i \leq N-1$. Even if a zero-sum cannot be set up for output bit y_0, it is straightforward to set up a zero-sum over the remaining $N-1$ bits, since the degree of y_i, for $1 \leq i \leq N-1$, is just 1.

C.2 Practical Results

Our practical results for the cubing S-Box are shown in Table 2, and they are obtained using the strategy described before.

We recall that the number of rounds $\mathcal{R}_{\text{PSPN}}$ is computed assuming an initial subspace of dimension n. In our practical tests, we considered the two extreme cases:

- an initial subspace of maximum dimension $N-1$;
- an initial subspace of dimension n chosen in such a way to skip as many initial rounds as possible without increasing the degree.

As can be seen, the number of rounds obtained in practice in this last case matches the theoretical ones.

For the case in which the initial subspace has dimension $N-1$, a bound on the minimum number of rounds necessary for guaranteeing security – denoted by R'_{SPN} – is given by (13). Indeed, consider a SPN scheme and a Partial SPN scheme over over $(\mathbb{F}_{2^n})^t$ instantiated with the same S-Box. The number of rounds necessary for reaching maximum algebraic degree for the SPN scheme cannot be smaller than the number of rounds needed by the Partial SPN scheme. At the same time, the gap for the number of rounds obtained in the case of an initial subspace of maximum dimension $N-1$ is in many cases non-trivial. As a result, in many cases it seems possible to break many more rounds than the ones predicted by R'_{SPN}. This is not a surprise, since the bound R'_{SPN} (derived for SPN scheme) does not take into account the fact that the non-linear layer is composed by $t-1$ identity functions in the case of a Partial SPN scheme. Our

Table 2: Theoretical and practical number of rounds necessary to prevent full higher-order differential distinguishers for P-SPN ciphers/permutations with a single ($s = 1$) S-Box of the form $x \mapsto x^3$. The mixing layer is defined as the multiplication with a matrix that maximizes the branch number, e.g., an MDS matrix. We present practical results with two different initial subspaces, one of dimension $N-1$, and one of dimension n (chosen to skip as many initial rounds as possible). The values of R'_{SPN} are computed using (13). We emphasize that $\mathcal{R}_{\text{PSPN}}$ is computed assuming an initial subspace of dimension n.

Parameters			Theoretical # of rounds		Practical # of rounds	
N	n	t	$\mathcal{R}_{\text{PSPN}}$ (dim. n)	R'_{SPN} (dim. $N-1$)	dimension n	dimension $N-1$
35	5	7	10	5	10	19
35	7	5	9	6	9	15
36	9	4	9	7	9	13
33	11	3	9	8	9	12
39	13	3	11	10	11	14
34	17	2	12	12	12	14
38	19	2	12	13	13	15
65	5	13	16	6	16	-
65	13	5	13	10	13	-
63	7	9	13	7	13	-
63	9	7	12	8	12	-
68	17	4	14	12	14	-
133	7	19	23	7	23	-
133	19	7	18	14	18	-
135	9	15	20	9	20	-
135	15	9	18	12	18	-

understanding of the degree growth for P-SPN schemes is far from complete and we leave this as an open problem. As future work, the goal would be to estimate the growth of the degree in the case in which $N-1$ bits are active.