Higher-Order Differentials of Ciphers with Low-Degree S-Boxes

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

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

Abstract. Higher-order differential attacks are among the most powerful attacks against low-degree ciphers and hash functions. Predicting the evolution of the algebraic degree of the cipher (as a function of the number of rounds) is the main obstacle in assessing the feasibility of these attacks. For an SPN cipher over a finite field $\mathbb{F}$ of characteristic 2 with round function of algebraic degree δ, it is a common belief that the degree of the cipher grows almost exponentially with δ. However, for an iterated Even–Mansour cipher whose round function can be described as an invertible low-degree polynomial over $\mathbb{F}_{2^n}$ it has recently been shown that the algebraic degree grows linearly with the number of rounds, and not exponentially.
In this paper we generalise these results for SPN ciphers, showing that the growth of the algebraic degree is often linear for SPN ciphers with low-degree S-Boxes as well. We prove that the initial exponential growth of the degree turns into a linear growth after a certain number of rounds. Our analysis includes iterated Even–Mansour and MiMC-like ciphers as a special case, but most notably it also applies to SPN ciphers designed to be competitive for recent applications like MPC, FHE, SNARKs, and STARKs (e.g., HadesMiMC). Our findings have been practically verified on small-scale ciphers.

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

1 Introduction

One of the most powerful cryptanalytic methods in the literature for low-degree symmetric primitives working over $\mathbb{F}_2^n$ is higher-order differential cryptanalysis. In essence, this method allows to distinguish a given Boolean function from a random one. More precisely, given an instance of a block cipher $E_k : \mathbb{F}_2^n \to \mathbb{F}_2^n$ under a fixed but unknown secret key k, higher-order differential cryptanalysis exploits the fact that for any vector subspace $\mathcal{V} \subseteq \mathbb{F}_2^n$ with dimension strictly greater than the algebraic degree of E_k and any $c \in \mathbb{F}_2^n$

$$\sum_{x \in \mathcal{V}} E_k(x + c) = 0.$$

Since the same property does not, in general, hold for a permutation drawn at random, it can be exploited to set up distinguishers and/or key-recovery attacks. The idea was first introduced by Lai [Lai94], albeit without a concrete application. Knudsen [Knu94] then used

higher-order differentials to break low-degree ciphers which were deemed secure against standard differential cryptanalysis. Several generalisations of higher-order differential attacks have since been proposed in the literature, including *cube attacks* [DS09] and the *division property* [Tod15].

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

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

> "After r rounds, there is no output bit with 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 cipher when the number of rounds varies is the main challenge in higher-order differential cryptanalysis. A trivial 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 allows to derive a first estimate about the number of rounds *necessary* to reach the full algebraic degree in SPN ciphers. For an SPN cipher defined over $(\mathbb{F}_{2^n})^t$ with S-Box layer of algebraic degree δ it follows that at least $\lceil \log_\delta(n \cdot t - 1) \rceil$ rounds are necessary to prevent higher-order differential attacks (the affine layer does not increase the algebraic degree).

A Better Estimation for $\deg(G \circ F)$. In general, the upper bound (1) does not reflect the real growth of the algebraic degree when considering iterated ciphers, 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 [CV02], a tighter upper bound was presented by Boura, Canteaut, and De Cannière in [BCD11]. In there, the authors show how to deduce a new bound for the algebraic degree of iterated permutations for a special category of SP networks over $(\mathbb{F}_{2^n})^t$, which includes functions that have a number of $t \geq 1$ balanced S-Boxes as their non-linear layer. As a consequence, the number of rounds necessary to prevent higher-order differential attacks is in general higher than the one obtained using the trivial bound in (1). Apart from the bounds of Boura, Canteaut and De Cannière, Boura and Canteaut studied the influence of F^{-1} on the algebraic degree of $\deg(G \circ F)$ [BC13]. 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}.

MiMC-Like Ciphers. MiMC [AGR+16] is an *iterated Even–Mansour cipher*, i.e. a cipher 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 growth of MiMC-like ciphers has been proposed in [EGL+20]. 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 $\delta(r)$ grows linearly with the number of rounds, and not (almost) exponentially as previously believed, i.e.

$$\delta(r) \leq \min\{\log_2(d^r), n-1\}.$$

Consequently, this observation implies that roughly $(n-1) \cdot \log_d(2)$ rounds are necessary for security against higher-order differential distinguishers. As a concrete application, the authors in [EGL+20] were able to exploit this result for setting up the first higher-order differential attack on MiMC better than exhaustive search.

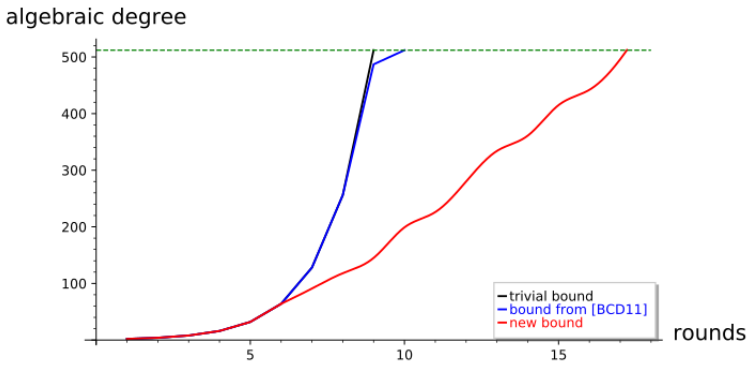

Figure 1: Growth of upper bounds of the algebraic degree for an SPN cipher with S-Box $x \mapsto x^3$ over $(\mathbb{F}_{2^{19}})^{27}$.

1.2 Our Contribution

Results for SPN Ciphers. In this work, we generalise the result presented in [EGL+20] and extend it to SPN ciphers, where one round consists of the parallel application of $t \geq 2$ invertible low-degree polynomial functions over $\mathbb{F}_{2^n}$. As such, we work with permutations over $(\mathbb{F}_{2^n})^t$. In particular, we improve on the best currently known estimation for the growth of the algebraic degree published in [BCD11]. We show in Section 4 that for all SPN ciphers with the parallel application of t copies of the same low-degree S-Box, the algebraic degree grows almost linearly with the number of rounds, and not exponentially as previously believed. More precisely, if d denotes the degree of the S-Boxes as a polynomial over $\mathbb{F}_{2^n}$ and δ denominates the corresponding algebraic degree, we prove that

$$n \cdot \log_d(2) + \log_\delta(t)$$

rounds are necessary to provide security against higher-order differential attacks. This result improves on the belief that around

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

rounds are necessary, which is based on the assumption of an exponential growth of the algebraic degree. A concrete comparison of our bound and other upper bounds of the algebraic degree of a SHARK-like cipher [RDP+96] over $(\mathbb{F}_{2^{19}})^{27}$ (with the cubing function $x \mapsto x^3$ as S-Boxes and an MDS matrix as mixing layer) is depicted in Fig. 1.

Preliminary Results for Feistel and Partial SPN Ciphers. Finally, we mention that our results apply to the case of Feistel schemes and partial SPN ciphers as well (that is, ciphers with a partial non-linear layers), by combining the results on SPN ciphers and the fact that the non-linear layers are not full. Also in this case, it is possible to show that for low-degree polynomial functions the algebraic degree of the cipher grows linearly. More details are given in Section 6.

2 Preliminaries

In this section, we recall the most important results about polynomial representations of Boolean functions and summarize the currently best known results regarding bounding the algebraic degree in the context of SP networks. We start with a clarification about what we mean by the *necessary* and the *sufficient* number of rounds to provide security against an attack. Informally, 'necessary' means *at least* and 'sufficient' can be read as *at most*.

Definition 1. Given an iterated cipher and a certain number of rounds $\mathcal{R} \geq 1$, we say $\mathcal{R}$ rounds are *necessary* to prevent a certain attack if this attack can be set up for each number of rounds r with $r < \mathcal{R}$. Similarly, a certain number of rounds $\mathcal{R} \geq 1$ is *sufficient* to prevent an attack if the attack cannot be set up for all r with $r \geq \mathcal{R}$.

While upper-bounding the algebraic degree (i.e., lower-bounding the number of rounds) is more important from an attacker's point of view, lower-bounding the algebraic degree (i.e., upper-bounding the number of rounds) is more relevant when arguing about the security from a designer's viewpoint. However, at the current state of the art and to the best of our knowledge, it seems very hard to find such a lower bound for a given cipher without investigating concrete instances experimentally. An experimental approach of course limits the scope of any analysis. We emphasize that in general it is only possible to provide a *necessary* number of rounds to provide security against higher-order differential attacks.

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

$$F(X_1, \ldots, X_t) = \bigoplus_{u=(u_1,\ldots,u_t)\in\{0,1,\ldots,2^n-1\}^t} \varphi(u) \cdot X_1^{u_1} \cdot \cdots \cdot X_t^{u_t}, \tag{2}$$

for certain $\varphi(u) \in \mathbb{F}_{2^n}$. We refer to this representation as the *word-level representation*. At the same time, the function F 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

$$F_i(X_1, \ldots, X_N) = \bigoplus_{u=(u_1,\ldots,u_N)\in\{0,1\}^N} \varphi_i(u) \cdot X_1^{u_1} \cdot \ldots \cdot X_N^{u_N}, \tag{3}$$

where the coefficients $\varphi_i(u) \in \mathbb{F}_2$ 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 Eq. (3) into a single polynomial representation leads to a description of F as a single polynomial in $n \cdot t$ 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*. We denote functions $F : \mathbb{F}_2^n \to \mathbb{F}_2$ as *Boolean functions* and hence functions of the form $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, for $n, m \in \mathbb{N}$, as *vectorial Boolean functions*. If not explicitly stated otherwise, we work with vectorial Boolean functions where $m = n$. 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

$$F(X_1, \ldots, X_n) = \bigoplus_{u=(u_1,\ldots,u_n)\in\{0,1\}^n} \varphi(u) \cdot X_1^{u_1} \cdot \ldots \cdot X_n^{u_n}.$$

The *algebraic degree* $\delta(F)$ of F is the degree of the 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 $(G_1, \ldots, G_n)$

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 [CCZ98, 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 $\mathbb{F}_{2^n}$. Thus, the algebraic degree of a vectorial Boolean function can be computed from its univariate representation. Lemma 1 makes this link explicit.

Lemma 1. *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[1] of exponents of non-vanishing monomials, that is*

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

2.2 Higher-Order Differentials and SPN Ciphers – State of the Art

The currently best-known generic upper bound for the algebraic degree of the composition of two functions is given by Boura, Canteaut, and De Cannière.

Proposition 1 ([BCD11])**.** *Let F be a function from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ corresponding to the concatenation of t smaller balanced[2] 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*

$$\deg(G \circ F) \leq \min\left\{\deg(F)\cdot\deg(G), N - \frac{N-\deg(G)}{\gamma}\right\}, \tag{4}$$

where

$$\gamma := \max_{i=1,\ldots,n-1} \frac{n-i}{n-\delta_i} \leq n-1, \tag{5}$$

and δ_i is defined as the maximal 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 have in general different γ. Exploiting relation (4), we present a *direct* upper bound of the algebraic degree after a certain number of rounds in the simplest but most common case of an SPN cipher where all S-Boxes are equal. With direct upper bound we mean that we iteratively apply (4) to the round functions of an SPN cipher and thus obtain a statement about the algebraic degree after a certain number of rounds. We refer to Appendix A for the details of the proof.

Proposition 2. *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 $L_1, L_2, \ldots, L_{r-1}$ from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ and any integer $r \geq 1$ consider the function E from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ defined as*

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

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

$$\delta(r) \leq \begin{cases} \delta^r & \text{if } r \leq \mathfrak{R} := \left\lfloor \log_\delta\left(N\cdot\frac{\gamma-1}{\gamma\cdot\delta-1}\right)\right\rfloor, \\ N - \gamma^{-r}\cdot\gamma^{\mathfrak{R}}\cdot(N-\delta^{\mathfrak{R}}) & \text{if } \mathfrak{R} < r \leq \mathcal{R}^{[\mathrm{BCD11}]}, \end{cases} \tag{6}$$

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

[2] 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.

where $\mathcal{R}^{[\mathrm{BCD11}]}$ is the number of rounds necessary to prevent secret-key zero-sum distinguishers defined by

$$\mathcal{R}^{[\mathrm{BCD11}]} := \underbrace{\left\lfloor \log_\delta\left(N \cdot \frac{\gamma - 1}{\gamma \cdot \delta - 1}\right)\right\rfloor}_{=:R_0} + \left\lceil \log_\gamma\left(N - \delta^{R_0}\right)\right\rceil \tag{7}$$

and where γ is defined as in Eq. (5).

A proof how to derive $\mathcal{R}^{[\mathrm{BCD11}]}$ can be found in Appendix A. We note that a similar result has been proposed in [BKP16] for the particular case in which $\gamma = n - 1$, meaning that all S-Boxes have maximum algebraic degree, that is, $\delta = n - 1$.

3 Higher-Order Differential Analysis of Iterated Even–Mansour Ciphers

As a main result, in [EGL+20] the authors show that the algebraic degree of iterated Even–Mansour ciphers may grow much slower than what is commonly suggested in the literature. More precisely, they show that in some cases the algebraic degree grows linearly with the number of rounds and not exponentially. In this section, we briefly recall the results published in [EGL+20].

We start by considering an *iterated Even–Mansour cipher* $EM_k^r : \mathbb{F}_{2^N} \to \mathbb{F}_{2^N}$ defined as

$$EM_k^r(x) := k_r \oplus (\cdots R(k_1 \oplus R(k_0 \oplus x)) \cdots) \tag{8}$$

for $r \geq 1$ rounds, where $k_0, \ldots, k_r \in \mathbb{F}_{2^N}$ are derived from a master key $k \in \mathbb{F}_{2^N}$ using a certain key schedule, and where each round function $R : \mathbb{F}_{2^N} \to \mathbb{F}_{2^N}$ is simply defined as some invertible polynomial function

$$R(x) := \rho_0 \oplus \bigoplus_{i=1}^{d} \rho_i \cdot x^i \tag{9}$$

of degree $d \geq 3$ and with $\rho_i \in \mathbb{F}_{2^N}$, $\rho_d \neq 0$. A cipher in the literature that falls into this category is e.g. MiMC. In [EGL+20], the authors derive a necessary condition on the number of rounds to prevent higher-order differential attacks against iterated Even–Mansour ciphers, see Proposition 3. In the remarks following Proposition 3 we discuss its scope in more detail.

Proposition 3 ([EGL+20])**.** *Let R be the round function of an iterated Even–Mansour cipher EM_k^r with degree d defined as in Eq. (9). The number of rounds[3] $\mathcal{R}^{Linear}$ necessary to prevent a secret-key higher-order differential distinguisher is given by*

$$\mathcal{R}^{Linear} = \left\lceil \log_d\left(2^{N-1} - 1\right)\right\rceil \approx (N - 1) \cdot \log_d(2). \tag{10}$$

The idea of the proof is simple: to prevent higher-order differential attacks, the algebraic degree of EM_k^r must reach its maximum value $N - 1$. Due to the relation between the word-level degree and the algebraic degree (see Section 2.1), EM_k^r has algebraic degree $N - 1$ if at least one monomial with exponent $2^N - 2^j - 1$ (for $0 \leq j < N$) is present in the univariate polynomial representation. Since the smallest exponent of this form is $2^{N-1} - 1$, the number of rounds r must satisfy $r \geq \lceil \log_d(2^{N-1} - 1)\rceil$.

[3] We use the notation $\mathcal{R}^{\mathrm{Linear}}$ to indicate that the algebraic degree grows almost linearly.

Forward versus Backward Direction. As recalled in [EGL+20] and originally proved in Corollary 3 of [BC13], given a fixed key k, the algebraic degrees of EM_k^r and its compositional inverse EM_k^{-r} are related in a particular way: the algebraic degree of EM_k^r is maximal (i.e. $n-1$) if and only if the algebraic degree of EM_k^{-r} is maximal. As an immediate consequence we state the following observation for iterated Even–Mansour ciphers:

> "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 iterated Even–Mansour cipher has a low-degree round function, while the inverse direction is built from a round function of high degree. For example, when $R(x) = x^3$, the inverse round function is given by $R^{-1}(x) = x^{(2^{N+1}-1)/3}$. Here, R has algebraic degree 2, while R^{-1} has algebraic degree $(n+1)/2$.

Only a Necessary Condition. We stress once more, the condition on the number of rounds in Proposition 3 is only a *necessary* condition, not a sufficient one. This comes from the (complicated) cancellation behaviour of powers of polynomials over finite fields: even if the univariate degree satisfies $d^r \geq 2^{N-1} - 1$ after a certain number of rounds r, it is not guaranteed that monomials of algebraic degree $N-1$ will be present in the encryption polynomial. A deeper analysis and concrete examples of this behaviour are given in [EGL+20, Sect. 3.1] and in Section 5.2.

Full versus Partial Zero-Sums. Another point to keep in mind is that the result of Proposition 3 only focuses on security against *full* zero-sums, i.e. zero-sums over all outputs bits. We mention that this does not provide security against *partial* zero sums, i.e. zero-sums only in some particular output 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 distinguisher over the remaining $N-1$ bits, since the degree of y_i, for $1 \leq i \leq N-1$, is just 1.

4 Higher-Order Differential Analysis of SPN Ciphers

In this section we prove a new upper bound on the growth of the algebraic degree in SPN ciphers with low-degree polynomial S-Boxes. In the following, let $E_k^r : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ denote the application of r rounds of an SPN cipher under a fixed but unknown secret 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

$$E_k^r(x) := (F_r \circ \cdots \circ F_1)(x \oplus k_0), \tag{11}$$

where $F_i : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ is defined as $F_i(x) := R(x) \oplus k_i$, for $1 \leq i \leq r$. The subkeys $k_0, \dots, k_r \in (\mathbb{F}_{2^n})^t$ 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

$$R(x) := (M \circ S)(x) := M(S_1(x_1), \dots, S_t(x_t)),$$

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

$$S_i(x) := \bigoplus_{j=0}^{d} \sigma_j^{(i)} \cdot x^j, \tag{12}$$

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

$$M(x) := \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,t} \\ M_{2,1} & M_{2,2} & \dots & M_{2,t} \\ \vdots & & \ddots & \vdots \\ M_{t,1} & M_{t,2} & \dots & M_{t,t} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{pmatrix}, \tag{13}$$

where $M_{i,j} \in \mathbb{F}_{2^n}$ for $i, j = 1, \dots, t$. The notion of a non-trivial linear layer is made precise in the following definition.

Definition 3. A linear layer M is *non-trivial* if it ensures full diffusion (in the sense that each word of the output depends on each word of the input and vice versa) after a *finite* number of rounds.

We remark that all SPN ciphers over $(\mathbb{F}_{2^n})^t$ can be written as described above. Just to give some examples, if the linear layer is defined by an MDS matrix[4], the cipher is similar to SHARK [RDP+96]. For AES [DR02] 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. [BB02]).

4.1 New Bound on the Necessary Number of Rounds to Prevent Higher-Order Differential Attacks

The main result in this section is the following.

Theorem 1. *Let $n \geq 3$ and $t \geq 2$. Consider r rounds of an SPN cipher E_k^r over $(\mathbb{F}_{2^n})^t$ as defined in Eq. (11), with the additional assumption that all S-Boxes $S_1, \dots, S_t$ are defined via the same function S of degree $d \geq 3$ and of algebraic degree $\delta \geq 2$. Then the algebraic degree after r rounds, denoted by $\delta(r)$, is upper-bounded by*

$$\delta(r) \leq \begin{cases} \delta^r & \text{if } r \leq \mathfrak{R} := 1 + \lceil \log_\delta(t) \rceil \ , \\ t \cdot \delta + \left\lceil t \cdot \log_2(d^{r-\mathfrak{R}}) \right\rceil & \text{if } \mathfrak{R} < r \leq \mathcal{R}^{Linear}, \end{cases} \tag{14}$$

where

$$\mathcal{R}^{Linear} := \lceil n \cdot \log_d(2) \rceil + \lceil \log_\delta(t) \rceil \tag{15}$$

is the number of rounds necessary to prevent a higher-order differential distinguisher.

We emphasize that we are speaking of a *necessary* number of rounds to prevent higher-order differential distinguishers. While we do *not* claim the above number of rounds to be *sufficient* for E_k^r to have maximum algebraic degree, the finesse of our new bound is that it is considerably closer to a sufficient bound than the currently best known results in the literature.

4.1.1 Idea of the Proof

Before we prove Theorem 1, we give a brief overview of the proof itself and explain some of the terminology and general assumptions.

[4] 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 $u \leq t$.

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

1. Lemma 2 makes a statement about which monomials can occur in the polynomial representation of the encryption function.
2. In Lemma 3 we tightly upper-bound the number of rounds with exponentially growing algebraic degree.
3. Finally, we use this observation to prove Proposition 4, which is a slightly reformulated equivalent to Theorem 1.

Terminology. We recall part of our terminology relevant for this section. For an unknown but fixed secret key $k = (k_1, \dots, k_t) \in (\mathbb{F}_{2^n})^t$ let $E_k^r = (E_{k,1}^r, \dots, E_{k,t}^r)$ denote the representation of an SPN cipher $E_k^r : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ after r rounds as a t-tuple of polynomials over $\mathbb{F}_{2^n}$ with maximum univariate degree $2^n - 1$. The variables $X_1, \dots, X_t$ shall represent n-bit words. Furthermore, let $m_\alpha := X_1^{\alpha_0} \cdots X_t^{\alpha_t}$ denote the monomial with exponent vector $\alpha = (\alpha_1, \dots, \alpha_t) \in \{0, \dots, 2^n - 1\}^t$. The word-level degree of the S-Box $S(X) = X^d$ is denominated by d, while the algebraic degree of S is denoted by $\delta := \mathrm{hw}(d)$. The base-2 expansion of d is written as $d = \sum_{i=1}^{\delta} 2^{d_i}$, for appropriate $d_i \in \mathbb{N}$. When we speak of monomials to be *expected*[5] in the encryption polynomial we allude to the fact that the actual number of non-zero coefficients in the polynomial representation of E_k^r also depends on the secret key k, and as a result, the coefficient of a monomial may be zero under a specific key. We presume to be sloppy on this point and occasionally just speak of monomials that appear.

General Assumptions. For simplicity, and since this is the most common case in the literature, we assume that all S-Boxes $S_1, \dots S_t$ are defined via the same *monomial* function $S : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ with $S(x) := S_1(x) = \dots = S_t(x) = x^d$. We remark that our results can be extended to the case where the S-Boxes are defined via more general *polynomial* mappings as described in Eq. (12).

According to the remark about the connection of forward and backward direction in Section 3, it suffices to focus only on one direction of the cipher when attempting to reach the 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 to be non-trivial (see Definition 3). However, depending on the instantiation of the linear layer the algebraic degree might grow slower than we predict, but never faster because a linear function does not increase the algebraic degree. Therefore our analysis focuses on the polynomial representation of a single output word E_j for a certain $1 \leq j \leq t$. Hence, when we refer to *the* encryption polynomial, it is one of the polynomials E_j^r.

4.1.2 Details of the Proof

Here we provide the details of the proof, following the strategy just presented.

Lemma 2. *Let $r \geq 1$ and $d = \sum_{i=1}^{\delta} 2^{d_i}$ be the base-2 expansion of d. Let $m_{\alpha_1}, m_{\alpha_2}, \dots, m_{\alpha_\delta}$ be monomials in the encryption polynomial after r rounds. Only those monomials given by*

$$m_{\alpha_1}^{2^{d_1}} \cdot m_{\alpha_2}^{2^{d_2}} \cdot \ldots \cdot m_{\alpha_\delta}^{2^{d_\delta}}$$

are to be expected in the encryption polynomial after $r + 1$ rounds.

[5] This notion of expectation has nothing to do with statistical expectation.

Proof. For each $j = 1, \ldots, t$, the expression

$$E_j^{r+1} = M_{j,1} \cdot (E_1^r \oplus k_1)^d \oplus M_{j,2} \cdot (E_2^r \oplus k_2)^d \oplus \cdots \oplus M_{j,t} \cdot (E_t^r \oplus k_t)^d$$

describes the encryption polynomial of word j after $r+1$ rounds. If we write

$$E_j^r \oplus k_j = \bigoplus_{\alpha \in (0,1,\ldots,2^n-1)^t} \varphi_\alpha \cdot m_\alpha,$$

and use the fact that $\mathrm{char}(\mathbb{F}_{2^n}) = 2$, we obtain

$$\begin{aligned}(E_j^r \oplus k_j)^d &= \left(\bigoplus_{\alpha \in (0,1,\ldots,2^n-1)^t} \varphi_\alpha \cdot m_\alpha \right)^{2^{d_1}+\cdots+2^{d_\delta}} \\ &= \prod_{i=1}^{\delta} \left(\bigoplus_{\alpha \in (0,1,\ldots,2^n-1)^t} \varphi_\alpha^{2^{d_i}} \cdot m_\alpha^{2^{d_i}} \right) = \bigoplus_{\alpha_1,\ldots,\alpha_\delta \in (0,1,\ldots,2^n-1)^t} \left(\prod_{i=1}^{\delta} \varphi_{\alpha_i}^{2^{d_i}} \cdot m_{\alpha_i}^{2^{d_i}} \right).\end{aligned}$$

Hence, we conclude that only monomial products of the form

$$m_{\alpha_1}^{2^{d_1}} \cdot m_{\alpha_2}^{2^{d_2}} \cdot \ldots \cdot m_{\alpha_\delta}^{2^{d_\delta}} \tag{16}$$

are expected to occur in the encryption polynomial after $r+1$ rounds. The monomials $m_{\alpha_1}, \ldots, m_{\alpha_\delta}$ are not necessarily different, therefore the exponents in Eq. (16) are either powers of 2 or sums of powers of 2. □

Lemma 3. *Let $d = \sum_{i=1}^{\delta} 2^{d_i}$ for appropriate $d_i \in \mathbb{N}$. Only for the first $1 + \lceil \log_\delta(t) \rceil$ rounds the algebraic degree of the encryption polynomial grows as fast as δ^r.*

Proof. The idea of the proof is to observe the growth of the algebraic degree with the help of Lemma 2. After the first round, all monomials of the form $X_1^d, \ldots, X_t^d$ appear in the encryption polynomial. The algebraic degree of each of these monomials is δ. According to Lemma 2, after one more round the monomial

$$(X_1^d)^{2^{d_1}} \cdot (X_2^d)^{2^{d_2}} \cdot \ldots \cdot (X_\delta^d)^{2^{d_\delta}}$$

appears in the encryption polynomial, and it has algebraic degree δ^2. To see why it has algebraic degree δ^2, we note that: *(a)* raising a (word-level) monomial m to the power of 2^k, $k \in \mathbb{N}$, does not change its algebraic degree, and *(b)* if two (word-level) monomials m_1, m_2 do not contain any shared variable, the algebraic degree of the product $m_1 \cdot m_2$ is the sum of the algebraic degrees. In the same way as before, after another round, the monomial

$$\underbrace{(X_1^{d2^{d_1}} \cdots X_\delta^{d2^{d_\delta}})^{2^{d_1}}}_{\text{algebr. degree } \delta^2} \underbrace{(X_{\delta+1}^{d2^{d_1}} \cdots X_{2\delta}^{d2^{d_\delta}})^{2^{d_2}}}_{\text{algebr. degree } \delta^2} \cdots \underbrace{(X_{(\delta-1)\delta+1}^{d2^{d_1}} \cdots X_{\delta\delta}^{d2^{d_\delta}})^{2^{d_\delta}}}_{\text{algebr. degree } \delta^2}$$

has algebraic degree δ^3 and, again, appears in the encryption polynomial. Continuing this way, we conclude that the algebraic degree grows as fast as δ^r until all t variables are exhausted, i.e., until $\delta^r = \delta \cdot t$.

It remains to show that $\lceil \log_\delta(\delta \cdot t) \rceil = 1 + \lceil \log_\delta(t) \rceil$ is the maximum number of rounds with exponential growth. To see this, we argue with the connection between the word-level degree and the bit-level degree (which is the algebraic degree). After $1 + \lceil \log_\delta(t) \rceil$ rounds, we have a monomial that uses all t variables $X_1, \ldots, X_t$. Therefore, the algebraic degree cannot be increased any further by multiplying bit variables that belong to different words, but only by multiplying bit variables that belong to the same word. As we have discussed in Section 3, in this case the algebraic degree in each variable X_i grows at most as fast as $\log_2(d) \cdot r$, and hence in total at most by $t \cdot \log_2(d) \cdot r$. □

Proposition 4. *Only after at least*

$$\lceil n \cdot \log_d(2) \rceil + \lceil \log_\delta(t) \rceil$$

rounds is the encryption polynomial expected to have the maximum algebraic degree $n \cdot t - 1$.

Proof. As shown in Lemma 3, only for the first $1 + \lceil \log_\delta(t) \rceil$ rounds the algebraic degree grows exponentially with the number of rounds and, eventually, adds up to $t \cdot \delta$. The idea of the proof is to give the minimum additional number of rounds until the maximum algebraic degree $n \cdot t - 1$ is reached.

As discussed at the end of the proof of Lemma 3, from now on it suffices to observe the algebraic degree *in a single* variable X_i. Hence, to reach algebraic degree n in a single variable and maximum overall algebraic degree $n \cdot t - 1$ respectively, it takes at least

$$\lceil n \cdot \log_d(2) \rceil - 1$$

more rounds, because the algebraic degree in each variable is already δ and we need degree n in $t - 1$ variables and degree $n - 1$ in one variable to reach maximum overall degree. □

4.2 Comparison with Related Work in the Literature

4.2.1 Linear Growth versus Exponential Growth

We compare the growth of the degree predicted by our formula with the currently best known results in the literature. Let us focus on the case in which $n \cdot \log_d(2)$ is large compared to $\log_\delta(t)$, namely in the case of *large* S-Boxes (i.e., $n \geq t$) described by a *low-degree* polynomial (i.e., $d \ll n$). In such a case, after some initial rounds for which the growth is exponential, the growth of the degree predicted by our formula is linear in r, that is according Eq. (14)

$$\delta(r) \approx t \cdot \delta + t \cdot \log_2(d^{r-\mathfrak{R}}) = r \cdot t \cdot \log_2(d) + \underbrace{t \cdot \delta - \mathfrak{R} \cdot t \cdot \log_2(d)}_{\text{constant}}.$$

For a concrete example we refer to Fig. 2. There we compare the *upper* bound predicted by our formula and the one proposed in [BCD11] in the case of an SPN cipher with the cube S-Box $S(x) = x^3$ over $(\mathbb{F}_{2^n})^t$ for $n = 63$ and $t = 16$.

4.2.2 When is $\mathcal{R}^{\text{Linear}} \geq \mathcal{R}^{[\text{BCD11}]}$?

For a better insight when the bound $\mathcal{R}^{\text{Linear}}$ 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*

$$\mathcal{R}^{Linear} \geq \mathcal{R}^{[\text{BCD11}]}$$

satisfied? Substitung the corresponding expression we obtain the following inequality

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

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

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

Thus, the condition $\mathcal{R}^{\text{Linear}} \geq \mathcal{R}^{[\text{BCD11}]}$ is satisfied if

$$\mathcal{R}^{\text{Linear}} = \lceil \log_d(2^n - 1) \rceil + \lceil \log_\delta(t) \rceil \geq 1 + \log_\delta(n \cdot t) + \log_2(n \cdot t),$$

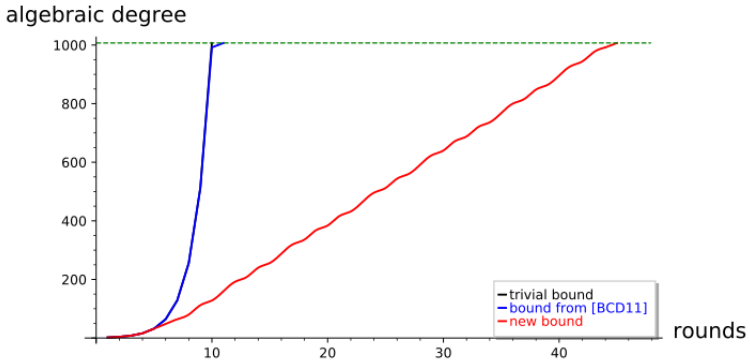

Figure 2: Growth of upper bounds of the algebraic degree for an SPN cipher over $(\mathbb{F}_{2^{63}})^{16}$ with S-Box $x \mapsto x^3$. Trivial bound in black corresponds to the case $\delta(r) = 2^r$. The line in blue indicates the bound obtained by exploiting (6) , which almost identical to the trivial bound, except at $r = 10$. Our new bound from relation (14) is depicted by the red line. After a first exponential growth up to $(5, 32)$, the growth of the algebraic degree is basically linear. The green line indicates the degree sufficient to prevent higher-order differential attacks.

or to put it another way, if

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

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.

5 Practical Results

In this section, we present our practical results on SPN ciphers over $(\mathbb{F}_{2^n})^t$ (defined as in Section 4) with low-degree S-Boxes. The practical tests have been performed in the same way as described in [EGL+20]: Instead of computing the ANF of a keyed permutation (which is quite expensive already for small field sizes), we evaluate the zero-sum property for a specific input vector space. Namely, for random keys and constants, given an input subspace of dimension $N - 1$, 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 the number of rounds necessary to prevent higher-order differential distinguishers. To avoid a bias by weak keys or "bad" round constants, we have repeated the tests multiple times (with new random keys and round constants). The code we used for the practical tests can be found on GitHub:

https://github.com/IAIK/higher-order-differential.

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

Practical Results on SPN ciphers with $S(x) = x^3$. For our practical results, we focus on a SHARK-like cipher [RDP+96], namely an SPN cipher over $(\mathbb{F}_{2^n})^t$, where the S-Box

Table 1: Theoretical and practical round numbers *necessary* to guarantee security against secret-key zero-sum distinguishers on SPN ciphers 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 $S(x) = x^3$. To better understand the influence of the linear layer, we consider both the case of a matrix that provides full diffusion after one round (e.g., an MDS matrix, if possible) – denoted by "Practical $\mathcal{R}$ (MDS)" – and the case of a matrix that provides the "worst" possible diffusion (e.g., a sparse matrix as in Eq. (17)) – denoted by "Practical $\mathcal{R}$ (Sparse)". $\mathcal{R}^{[\text{BCD11}]}$ are computed assuming $\gamma = (n+1)/2$.

Param.			Theoretical		Practical	
N	n	t	$\mathcal{R}^{\text{Linear}}$	$\mathcal{R}^{[\text{BCD11}]}$	Practical $\mathcal{R}$ (MDS)	Practical $\mathcal{R}$ (Sparse)
35	5	7	7	6	8	15
35	7	5	8	5	8	12
36	9	4	8	6	9	11
33	11	3	9	5	10	10
39	13	3	11	6	11	12
34	17	2	12	6	12	12
38	19	2	13	6	14	14
66	11	6	9	6	-	-
65	13	5	11	6	-	-
60	15	4	12	6	-	-
66	17	4	13	7	-	-
63	21	3	15	6	-	-
66	33	2	22	7	-	-
132	11	12	10	8	-	-
135	15	9	14	8	-	-
133	19	7	14	7	-	-
132	33	4	23	8	-	-
129	43	3	29	7	-	-
130	65	2	43	8	-	-

is $S(x) = x^3$ and the mixing layer is defined by the multiplication with an invertible $t \times t$ matrix. Our practical results are reported in Table 1. As in [EGL+20], the theoretical values for $\mathcal{R}^{[\text{BCD11}]}$ are computed assuming $\gamma = (n+1)/2$ (we refer to [EGL+20, Lemma 2] for a detailed argument supporting this point). We observe that the number of rounds that can be covered by a zero-sum distinguisher is (almost) always equal 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 numbers $\mathcal{R}^{\text{Linear}}$ predicted by our formula is significantly larger than $\mathcal{R}^{[\text{BCD11}]}$.

Test Methodology. In order to derive the results shown in Table 1, we wrote a custom script in C. We then searched for zero sums with various round numbers, using different random input subspaces of dimension $N-1$, together with random round constants and random round keys for each test. The round numbers we found for full zero sums were then verified again using additional test runs.

5.1 Influence of the Linear Layer

In order to understand how the linear layer influences the number of rounds *necessary* to provide security against zero-sum 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[6]), 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

$$M_{r,c} = \begin{cases} 1 & \text{if } r = 0 \text{ OR if } c \equiv r+1 \mod t, \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

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 zero-sum 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. An *open problem* for future research would be to take into account the details of the linear layer for a more accurate prediction of the number of rounds that can be covered by a zero-sum distinguisher.

5.2 Practical Results for $d = 2^{d'} \pm 1$

Besides the effect of the linear layer, we also tried to better understand the impact of the details of the S-Box. As already mentioned in [EGL+20, Sect. 3.1], whenever the output of the S-Box is sparse, or full, this fact will affect the gap between the practical number of rounds needed for security and the one predicted by our formula (which is just a lower bound since it is only a necessary condition). In particular, focusing on an S-Box defined by a power map $x \mapsto x^d$ there are two "extreme" cases, namely $d = 2^{d'} \pm 1$ for some $d' \in \mathbb{N}$. For these cases it is possible to observe that for $d = 2^{d'} + 1$ the output of a single round

$$(x+y)^{2^{d'}+1} = x^{2^{d'}+1} + x^{2^{d'}} \cdot y + y^{2^{d'}} \cdot x + y^{2^{d'}+1}$$

is sparse. Note that it contains only 4 terms instead of $d+1 = 2^{d'} + 2$. While for $d = 2^{d'} - 1$ the output of a single round

$$(x+y)^{2^{d'}-1} = \sum_{i=0}^{2^{d'}-1} x^i \cdot y^{2^{d'}-1-i}$$

contains much more monomials. In order to better understand this fact, we performed practical tests on a MiMC-like cipher with different S-Boxes $x \mapsto x^d$ of the form $d = 2^{d'} + 1$ and $d = 2^{d'} - 1$. We emphasize that the choice of working on MiMC-like ciphers and not on SPN ciphers has been made to prevent the influence of the linear layer. Our practical results are presented in Appendix B, see Table 4 and Table 5 for more details. As expected, we found that the polynomial that describes the encryption function is in general dense in the case $d = 2^{d'} - 1$. As a result, for $d = 2^{d'} - 1$ the sufficient number of rounds that provides security against secret-key zero-sum distinguishers is close to $\mathcal{R}^{\text{Linear}}$ previously given, while the gap between these two figures is in general larger for the case $d = 2^{d'} + 1$. These observations match the theoretical expectation expressed before, and the results concerning "Variants of MiMC" presented in [AGR+16].

[6] 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 Preliminary Results for Partial-SPN and Feistel Ciphers

Finally, we point out that the results presented in the previous sections, and from [EGL+20] (for the Even–Mansour case), may be also exploited to derive preliminary results for the case of Partial-SPN ciphers over $(\mathbb{F}_{2^n})^t$ or for Feistel schemes over $(\mathbb{F}_{2^n})^2$ in the case in which:

(1) the S-Box/round function can be described by a low-degree polynomial over $\mathbb{F}_{2^n}$,

(2) the input space has dimension n.

6.1 Partial-SPN Ciphers

First of all, here we show how to extend the previous results on Even–Mansour and SPN ciphers to analyze the growth of the algebraic degree of Partial-SPN ciphers. The main difference to SPN ciphers regards the S-Box layer. In Partial-SPN ciphers over $(\mathbb{F}_{2^n})^t$, the round function is defined as

$$R = M \circ (\underbrace{S_1, \dots, S_s, I, \dots I}_{t \text{ branches}}), \tag{18}$$

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 $M : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ can be regarded as the multiplication of the state vector with a $t \times t$-matrix over $\mathbb{F}_{2^n}$. In other words, the distinctive feature of a Partial-SPN cipher is that the non-linear functions S_i are only applied to part of the state, while the rest of the state remains unchanged by the S-Box layer. In the case of Partial-SPN ciphers, the attacker can employ several strategies to set up a higher-order differential zero-sum distinguisher. 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 cipher, 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 zero-sum distinguisher. For SPN ciphers, the first strategy is in general always the best, since it is not possible to "skip" rounds for free without activating any S-Box.

For simplicity, we limit ourselves to consider the case in which $s = 1$, that is one single S-Box is applied in each round (which e.g. corresponds to the instantiation of LowMC used in the Picnic signature scheme [CDG+17, CDG+19] and to the partial rounds in the recently proposed HadesMiMC permutation [GLR+20] permutation).

Proposition 5. *Let $E_k : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ be a Partial-SPN cipher 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 $S(x) := \rho_0 \oplus \bigoplus_{i=1}^{d} \rho_i \cdot x^i$ of degree $d \geq 3$ and with $\rho_i \in \mathbb{F}_{2^n}$, $\rho_d \neq 0$. Then at least*

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

rounds are necessary to prevent higher-order differential distinguishers. In particular, 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[7]. 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 presented 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 zero-sum distinguishers. □

6.1.1 Related Work

We note that the result just presented is less surprising than the corresponding one given for SPN ciphers. Indeed, the linear growth of the algebraic degree in case of Partial-SPN ciphers had already been observed in the literature when considering the security of LowMC [ARS+15] and Bison [CLL+19]. For example, we quote from Section 6.2 of [CLL+19]:

> "[...] 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."

A formal analysis of the growth of the algebraic degree in a Partial-SPN cipher has been given e.g. for LowMC in [ARS+15].

Proposition 6 ([ARS+15])**.** *Let F be a function that corresponds to the parallel application of s balanced n-bit S-Boxes and an identity function of width $l = N - s \cdot n$. Thus, F is a mapping from $\mathbb{F}_2^{n \cdot s + l}$ to $\mathbb{F}_2^{n \cdot s + l}$. Let δ_k be the maximum algebraic degree of the product of any k output bits of the S-Box. Then for any function G from $\mathbb{F}_2^{n \cdot s + l}$ to $\mathbb{F}_2^N$, we have*

$$\deg(G \circ F) \leq \min\big\{\deg(G) \cdot \deg(F), \beta \cdot s + \deg(G)\big\},$$

where $\beta = \max_{1 \leq i \leq n}(\delta_i - i)$.

As before, focusing on the case where only one S-Box is applied in each round ($s = 1$), it is possible to obtain a closed formula that describes the growth of the degree.

Proposition 7. *Let F be a function on $(\mathbb{F}_{2^n})^t$ corresponding to the concatenation of $s = 1$ balanced S-Box defined over $\mathbb{F}_{2^n}$ and the identity function on $(\mathbb{F}_{2^n})^{t-1}$. Moreover, assume that the S-Box has algebraic degree $\delta \geq 2$, and let $\beta := \max_{1 \leq i \leq n}(\delta_i - i)$, where δ_k is the maximum algebraic degree of the product of any k output bits of the S-Box. For any affine functions $L_1, L_2, \ldots, L_{r-1}$ on $\mathbb{F}_{2^n}^t$, consider the encryption function $E^{(r)} : (\mathbb{F}_{2^n})^t \to (\mathbb{F}_{2^n})^t$ defined as*

$$E^{(r)} = \underbrace{F}_{r\text{-th round}} \circ \underbrace{L_{r-1} \circ F}_{(t-1)\text{-th round}} \circ \cdots \circ \underbrace{L_2 \circ F}_{2\text{-nd round}} \circ \underbrace{L_1 \circ F}_{1\text{-st round}}$$

[7] 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.

Table 2: Theoretical and practical round numbers necessary to prevent (full) zero-sum distinguishers on N bits for a P-SPN cipher with a single ($s = 1$) S-Box of the form $x \mapsto x^3$. The mixing layer is defined as the multiplication with an MDS matrix (or, in the case in which it does not exist, with a matrix that maximizes the branch number). We give the practical results obtained by using an initial subspace of dimension $N - 1$, and one of dimension n (chosen in order to skip as many initial rounds as possible). The numbers $\mathcal{R}^{[ARS+15]}$ are computed under the assumption $\beta = (n-1)/2$ and $\gamma = (n+1)/2$.

Param.			Theoretical		Practical $\mathcal{R}$	
N	n	t	$\mathcal{R}^{\text{Linear}}$	$\mathcal{R}^{[ARS+15]}$	dimension n	dimension $N-1$
35	5	7	10	18	10	19
35	7	5	9	12	9	15
36	9	4	10	10	9	13
33	11	3	9	8	9	12
39	13	3	11	8	11	14
34	17	2	12	7	12	14
38	19	2	12	7	13	15
65	5	13	16	33	16	-
65	13	5	13	12	13	-
63	7	9	13	22	13	-
63	9	7	12	17	12	-
68	17	4	14	11	14	-
133	7	19	23	45	23	-
133	19	7	18	17	18	-
135	9	15	20	35	20	-
135	15	9	18	21	18	-

for a certain number of rounds $r \geq 1$. Then, the minimum number of rounds $\mathcal{R}^{[ARS+15]}$ rounds of $E^{(r)}$ necessary to prevent higher-order differential distinguishers on $N = n \cdot t$ bits is given by

$$\mathcal{R}^{[ARS+15]} \geq 1 + \left\lfloor \log_\delta\left(\frac{\beta}{\delta - 1}\right)\right\rfloor + \left\lfloor \frac{N}{\beta} - \frac{1}{\delta - 1} - \frac{\gamma}{\gamma - 1}\right\rfloor + \left\lceil \log_\gamma\left(\frac{\beta}{\gamma - 1}\right)\right\rceil, \quad (19)$$

where γ and β are defined as before. In particular, to prevent higher-order differential distinguishers on $(t-1) \cdot n$ bits one more round with respect to the number above is necessary.

The proof can be found in Appendix C. Note that a *full* zero sum on N bits after the S-Box layer implies always (at least) a *partial* zero sum on $n \cdot (t-1)$ bits after the next S-Box layer.[8]

6.1.2 Practical Results and Open Problems

Our practical results for the cubing S-Box are shown in Table 2, and they are obtained using the strategy described before. We considered the two extreme cases, namely an initial subspace of maximum dimension $N - 1$ and an initial subspace of dimension n chosen in order 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.

[8] Indeed, a full zero sum on N bits after the S-Box layer implies a full zero-sum on N bits after the mixing layer (since it is linear). The result follows from the fact that the non-linear layer is partial (in our case, just 1 S-Box and $t - 1$ identity functions are applied).

On the other hand, 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. This choice of the initial subspace seems to allow an attacker to break more rounds than by using an initial subspace of dimension n. As a result, in many cases it seems it is possible to break many more rounds than the ones predicted by $\mathcal{R}^{\text{Linear}}$ and/or $\mathcal{R}^{[ARS+15]}$. We conjecture that this result is due to the fact that several cancellations of monomials in the encryption polynomial occur. Thus, our understanding of the degree growth for P-SPN ciphers 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.

6.2 Feistel Ciphers over $(\mathbb{F}_{2^n})^2$

Similar results can be proposed for Feistel schemes over $(\mathbb{F}_{2^n})^2$ in the case in which one branch is active (i.e., n bits).

Proposition 8. *Let $E_k : (\mathbb{F}_{2^n})^2 \to (\mathbb{F}_{2^n})^2$ be a Feistel cipher for which the i-th round function R_{k_i} is defined as*

$$(x, y) \mapsto (y \oplus R(x \oplus k_i), x),$$

where $R : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is simply defined as some invertible (low-degree) polynomial function $R(x) := \rho_0 \oplus \bigoplus_{i=1}^{d} \rho_i \cdot x^i$ of degree $d \geq 3$ and with $\rho_i \in \mathbb{F}_{2^N}$, $\rho_d \neq 0$. Here $k_i \in \mathbb{F}_{2^n}$ denotes the i-th subkey. In the case in which n bits (i.e., one branch) are active, at least

$$1 + \lceil \log_d(2^n - 1) \rceil$$

rounds are necessary to guarantee security against higher-order differential distinguishers.

Proof. Let us consider the case of n active bits in an initial structure of the form (C, A), where C denotes a constant word and A denotes an active word. The first round does not activate any S-Box. The output of the r-th round for $r \geq 2$ can be described as a polynomial of degree at most d^{r-1}. Working in the same way as in the Even–Mansour case, a monomial of degree $2^n - 1$ is expected to appear if the number of rounds r satisfies $d^{r-1} \geq 2^n - 1$. In case where more bits are active, the number of rounds necessary to guarantee security cannot be smaller. □

6.2.1 Practical Results

We consider the setting of a balanced Feistel network, where the number of branches is $t = 2$. Each of the branches is n bits long (i.e. the block size is $2n$). In our practical tests, we considered two scenarios:

1. We set the number of active bits to only n (namely, 1 branch) in the first approach.
2. We set the number of active bits to $2n - 1$ in the second one.

We summarize our results in Table 3. They were obtained using the same testing methodology as described in Section 5 (i.e., random subspaces[9], random round constants, random keys). For the first case (one active branch – n active bits), we provide a comparison between the practical results and the ones obtained using the theoretical result given in the previous proposition. As expected, the theoretical results for the number of rounds necessary were always lower than what we observed in practice. Moreover, as for the Even–Mansour case or for the SPN/P-SPN cases, the gap between the two cases decreases by increasing the size of the branch n.

[9] Namely, one branch is active and the remaining $n-1$ active bits are distributed at random in the other branch. Our practical results are (almost) not influenced by this distribution.

Table 3: Practical number of rounds *necessary* to prevent zero-sum distinguishers for the case of Feistel ciphers over $\mathbb{F}_{2^n}$, where the round function is of the form $R(x) = x^3$. Note that we explicitly try to avoid zero sums, and we ignore subspaces and round numbers resulting in constant-sum distinguishers.

Param.	Theoretical	Practical $\mathcal{R}$	
n	dimension n	dimension n	dimension $2n-1$
3	3	5	6
5	5	7	10
7	6	8	12
9	7	9	14
11	8	9	16
13	10	11	17

6.2.2 Related Work

The theoretical result just presented for the case of one active branch is analogous to the analysis presented in [BCD+20] for the case of a Feistel scheme over $\mathbb{F}_p$. In such a case, the authors exploit the fact that for any polynomial function f over $\mathbb{F}_p$

$$\deg(f) \leq p-2 \implies \sum_{x \in \mathbb{F}_p} f(x) = 0$$

in order to estimate the number of rounds that can be broken via a natural generalisation of higher-order differential attacks over $\mathbb{F}_p$ (see [BCD+20, Prop. 1]). Since our previous result is derived without exploiting the fact that the Feistel scheme is defined over a Boolean field (we indeed assume that the entire branch is active), it is not a surprise that the results are equivalent.

7 Summary and Open Problems

7.1 Summary

Our results on the security of SPN ciphers against higher-order differential attacks can be summarized as follows: The number of rounds necessary to prevent *full* secret-key zero-sum distinguishers on SPN ciphers over $(\mathbb{F}_{2^n})^t$ defined as in Eq. (11) is given by

$$\mathcal{R}^{\text{Linear}} = \lceil \log_d(2^n - 1) \rceil + \lceil \log_\delta(t) \rceil \,,$$

where all S-Boxes are defined via the same low-degree polynomial function S. We refer to the remark about full versus partial zero sums in Section 3 for a clarification of our terminology in this regard.

7.2 Open Problems

Below we list the main open problems we identified during our research, and which we leave for future work.

1. One may take into account the details of the linear layer for improving the bounds on the growth of the degree presented in this paper.

2. HadesMiMC [GLR+20] is probably the most suitable candidate in order to apply our results. A natural question arises immediately: by combining the results for SPN ciphers and partial SPN ciphers presented here, how many rounds of this scheme

is it possible to break by higher-order differential attacks? Moreover, is it possible to break the full scheme by exploiting the slow growth of the algebraic degree in the case of a weak linear layer [KR20, BCD+20] (namely, a linear layer M for which $M^2 = \mu I$ for a certain $\mu \in \mathbb{F}$ where I is the identity matrix)?

3. The analysis of the growth of the degree for partial SPN ciphers and Feistel schemes is far from being complete. Our current results are similar to the ones obtained for the $\mathbb{F}_p$-case in [BCD+20]. However, we expect that better results can be obtained for the Boolean cases, due to the larger number of subspaces that exist in $(\mathbb{F}_{2^n})^t$.

4. Finally, a next step would be to extend the higher-order differential distinguishers presented in this paper to e.g. key-recovery attacks, as was already done for MiMC in [EGL+20].

Acknowledgements. The authors thank Willi Meier for his valuable comments on an intermediate version of this paper.

References

[AGR+16] Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. MiMC: Efficient Encryption and Cryptographic Hashing with Minimal Multiplicative Complexity. In *ASIACRYPT 2016*, volume 10031 of *LNCS*, pages 191–219, 2016.

[ARS+15] Martin R. Albrecht, Christian Rechberger, Thomas Schneider, Tyge Tiessen, and Michael Zohner. Ciphers for MPC and FHE. In *EUROCRYPT 2015*, volume 9056 of *LNCS*, pages 430–454, 2015.

[BB02] Elad Barkan and Eli Biham. In How Many Ways Can You Write Rijndael? In *Advances in Cryptology - ASIACRYPT 2002*, volume 2501 of *LNCS*, pages 160–175, 2002.

[BC13] Christina Boura and Anne Canteaut. 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.

[BCD11] Christina Boura, Anne Canteaut, and Christophe De Cannière. Higher-Order Differential Properties of Keccak and *Luffa*. In *FSE 2011*, volume 6733 of *LNCS*, pages 252–269, 2011.

[BCD+20] Tim Beyne, Anne Canteaut, Itai Dinur, Maria Eichlseder, Gregor Leander, Gaëtan Leurent, María Naya-Plasencia, Léo Perrin, Yu Sasaki, Yosuke Todo, and Friedrich Wiemer. Out of Oddity – New Cryptanalytic Techniques against Symmetric Primitives Optimized for Integrity Proof Systems. Cryptology ePrint Archive, Report 2020/188, 2020. https://eprint.iacr.org/2020/188.

[BKP16] Alex Biryukov, Dmitry Khovratovich, and Léo Perrin. Multiset-Algebraic Cryptanalysis of Reduced Kuznyechik, Khazad, and secret SPNs. *IACR Trans. Symmetric Cryptol.*, 2016(2):226–247, 2016.

[CCZ98] Claude Carlet, Pascale Charpin, and Victor A. Zinoviev. Codes, Bent Functions and Permutations Suitable For DES-like Cryptosystems. *Designs, Codes Cryptography*, 15(2):125–156, 1998.

[CDG+17] Melissa Chase, David Derler, Steven Goldfeder, Claudio Orlandi, Sebastian Ramacher, Christian Rechberger, Daniel Slamanig, and Greg Zaverucha. Post-quantum zero-knowledge and signatures from symmetric-key primitives. In *CCS*, pages 1825–1842. ACM, 2017.

[CDG+19] Melissa Chase, David Derler, Steven Goldfeder, Jonathan Katz, Valdimir Kolesnikov, Claudio Orlandi, Sebastian Ramacher, Christian Rechberger, Daniel Slamanig, Xiao Wang, and Greg Zaverucha. The picnic signature scheme design document (version 2), 2019.

[CLL+19] Anne Canteaut, Virginie Lallemand, Gregor Leander, Patrick Neumann, and Friedrich Wiemer. BISON Instantiating the Whitened Swap-Or-Not Construction. In *EUROCRYPT 2019*, volume 11478 of *LNCS*, pages 585–616, 2019.

[CV02] Anne Canteaut and Marion Videau. Degree of Composition of Highly Nonlinear Functions and Applications to Higher Order Differential Cryptanalysis. In *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 518–533, 2002.

[DR02] Joan Daemen and Vincent Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard*. Information Security and Cryptography. Springer, 2002.

[DS09] Itai Dinur and Adi Shamir. Cube Attacks on Tweakable Black Box Polynomials. In *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 278–299, 2009.

[EGL+20] Maria Eichlseder, Lorenzo Grassi, Reinhard Lüftenegger, Morten Øygarden, Christian Rechberger, Markus Schofnegger, and Qingju Wang. An Algebraic Attack on Ciphers with Low-Degree Round Functions: Application to Full MiMC. Cryptology ePrint Archive, Report 2020/182, 2020. https://eprint.iacr.org/2020/182.

[GLR+20] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a Generalization of Substitution-Permutation Networks: The HADES Design Strategy, 2020.

[Knu94] Lars R. Knudsen. Truncated and Higher Order Differentials. In *FSE 1994*, volume 1008 of *LNCS*, pages 196–211, 1994.

[KR20] Nathan Keller and Asaf Rosemarin. Mind the Middle Layer: The HADES Design Strategy Revisited. Cryptology ePrint Archive, Report 2020/179, 2020. https://eprint.iacr.org/2020/179.

[Lai94] Xuejia Lai. *Higher Order Derivatives and Differential Cryptanalysis*, pages 227–233. Springer US, 1994.

[RDP+96] Vincent Rijmen, Joan Daemen, Bart Preneel, Antoon Bosselaers, and Erik De Win. The Cipher SHARK. In *FSE 1996*, volume 1039 of *LNCS*, pages 99–111, 1996.

[Tod15] Yosuke Todo. Structural Evaluation by Generalized Integral Property. In *EUROCRYPT 2015*, volume 9056 of *LNCS*, pages 287–314, 2015.

A Closed Formula for SPN Ciphers – Proof

Proposition 9. *Let F be a function from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ corresponding to the concatenation of t balanced S-Boxes $S_1, \ldots, S_t$ defined over $\mathbb{F}_2^n$. Moreover, assume that all S-Boxes are equal, that is $S := S_1 = \cdots = S_t$, with algebraic degree $\delta \geq 2$. Then, for any sequence of affine functions $L_1, L_2, \ldots, L_{R-1}$ from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$, consider the function $E^{(R)}$ from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ defined as*

$$E^{(R)} := \underbrace{F}_{R\text{-}th\ round} \circ \underbrace{L_{R-1} \circ F}_{(R-1)\text{-}th\ round} \circ \cdots \circ \underbrace{L_2 \circ F}_{2\text{-}nd\ round} \circ \underbrace{L_1 \circ F}_{1\text{-}st\ round}$$

*for a certain number of rounds $R \geq 1$. The minimum number of rounds $\mathcal{R}^{[\mathrm{BCD11}]}$ **necessary** to prevent (secret-key) zero-sum distinguishers is given by*

$$\mathcal{R}^{[\mathrm{BCD11}]} := \underbrace{\left\lfloor \log_\delta \left(N \cdot \frac{\gamma - 1}{\gamma \cdot \delta - 1} \right) \right\rfloor}_{\equiv R_0} + \left\lceil \log_\gamma \left(N - \delta^{R_0} \right) \right\rceil, \tag{20}$$

where γ is defined as in Proposition 1.

Proof. First of all, note that since neither a linear function nor an affine function can increase the algebraic degree, it follows that the algebraic degree of the function $E^{(R)}$ is trivially upper-bounded by

$$\deg\left(E^{(R)}\right) \leq \delta^R.$$

In order to prove the result, we assume the bound Eq. (4) of $\deg(G \circ F)$ is as given in Proposition 1:

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

Note that we are not considering a generic function G , but rather are interested in a concrete encryption function $E^{(R)}$, and thus G is the composition of R rounds $L \circ F(\cdot)$, where F is the non-linear layer and L is the affine layer.

For our specific case, the previous formula can be rewritten as an iterative sequence. Let x_i denote the degree of $E^{(R)}$ at round $i = R \geq 1$. It follows that

$$x_{i+1} \leq \min\left\{ \delta \cdot x_i; \frac{N \cdot (\gamma - 1)}{\gamma} + \frac{x_i}{\gamma} \right\},$$

where in the following $f(x) = \delta \cdot x$ and $g(x) = \frac{N \cdot (\gamma - 1)}{\gamma} + \frac{x}{\gamma}$. By simple computation, note that

$$\left[f'(x) = \delta \right] > \left[g'(x) = \frac{1}{\gamma} \right]$$

since $\delta \geq 2$ and $\gamma \geq 1$. Given x_0, let i be the minimum positive index s.t. $f(x_i) \geq g(x_i)$. It follows that

$$\forall j \geq i: \qquad x_{j+1} \leq \min\{f(x_j); g(x_j)\} = g(x_j).$$

In order to apply the bound Eq. (4), we work by induction. In the first step, $G(\cdot)$ is just an identity function $I(\cdot)$ (of degree 1). This means that

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

Due to the previous considerations, we now look for the maximum number of rounds R_0 such that $\deg(G) \cdot \deg(F) \leq \frac{N \cdot (\gamma - 1)}{\gamma} + \frac{\deg(G)}{\gamma}$.

In our case, this corresponds to

$$\delta^{R_0+1} \leq \frac{N \cdot (\gamma - 1)}{\gamma} + \frac{\delta^{R_0}}{\gamma}.$$

By simple computation, we obtain that $\delta^{R_0} \cdot (\gamma \cdot \delta - 1) \leq N \cdot (\gamma - 1)$, that is,

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

In other words, for any number of rounds $R \leq R_0$, the degree of $E^{(R)}$ is upper-bounded by the left term of Eq. (4).

Next, we need to find the minimum number of rounds in order to prevent higher-order differential attacks. Remember that given a function $f(\cdot)$ of algebraic degree δ, the sum over the outputs of the function applied to all elements of a vector space $\mathcal{V}$ of dimension $\geq \delta + 1$ is zero. Since we only consider attacks that use less than the full code book, the biggest subspace of $\mathbb{F}_{2^N}$ has dimension $N - 1$. As a result, a zero-sum distinguisher can be set up for at most R rounds, where

$$\deg\left(E^{(R)}\right) \leq N - 2.$$

By simple computation (using the left term of Eq. (4)), it follows that the degree of $E^{(R)}$ after $R = R_0 + R_1 > R_0$ rounds (for a certain $R_1 > 0$) is upper-bounded by[10]

$$\deg\left(E^{(R)}\right) \leq N \cdot (\gamma - 1) \cdot \sum_{j=1}^{R_1} \left(\frac{1}{\gamma}\right)^j + \frac{\delta^{R_0}}{\gamma^{R_1}} = \frac{N \cdot (\gamma^{R_1} - 1)}{\gamma^{R_1}} + \frac{\delta^{R_0}}{\gamma^{R_1}},$$

where δ^{R_0} is the degree of $E^{(R)}$ after R_0 rounds. Indeed, after $R_0 + 1$ rounds the degree is upper-bounded by

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

after $R_0 + 2$ rounds it is upper-bounded by

$$\frac{N \cdot (\gamma - 1)}{\gamma} + \frac{1}{\gamma} \cdot \left(\frac{N \cdot (\gamma - 1)}{\gamma} + \frac{\delta^{R_0}}{\gamma} \right) = N \cdot (\gamma - 1) \cdot \left(\frac{1}{\gamma} + \frac{1}{\gamma^2} \right) + \frac{\delta^{R_0}}{\gamma^2},$$

and so on. As a result, it follows that $\deg(E^{(R)}) \geq N - 1$ if

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

that is, $N - \delta^{R_0} \leq \gamma^{R_1}$, or equivalently

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

In conclusion, The number of rounds $\mathcal{R}^{[\text{BCD11}]}$ **necessary** to prevent zero-sum distinguishers is given by

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

$$\approx \left\lfloor \log_\delta \left(N \cdot \frac{\gamma - 1}{\gamma \cdot \delta - 1} \right) \right\rfloor + \left\lceil \log_\gamma \left(N \cdot \frac{\gamma \cdot (\delta - 1)}{\gamma \cdot \delta - 1} \right) \right\rceil.$$

□

[10] Remember that for each $0 \leq X < 1$: $\sum_{j=1}^{n} X^j = \frac{X - X^{n+1}}{1 - X}$.

Table 4: Theoretical and practical round numbers *necessary* to prevent zero-sum distinguishers for MiMC-like ciphers over $\mathbb{F}_{2^N}$, in the case of a round function of the form $R(x) = x^d$ for several values of $d = 2^{d'} + 1$.

Param.		Theoretical		Practical
d	N	$\mathcal{R}^{\text{Linear}}$	$\mathcal{R}^{[\text{BCD11}]}$	Practical $\mathcal{R}$
5	13	6	5	7
	17	8	5	8
	33	15	6	15
	65	28	7	-
	129	56	8	-
9	13	5	5	6
	17	6	5	7
	33	11	6	12
	65	21	7	-
	129	41	8	-
17	13	4	5	6
	17	5	5	7
	33	9	6	10
	65	16	7	-
	129	32	8	-
33	13	3	5	6
	17	4	5	7
	33	7	6	9
	65	13	7	-
	129	26	8	-
65	13	4	5	7
	17	3	5	7
	33	6	6	9
	65	11	7	-
	129	22	8	-

B Practical Tests on Iterated Even–Mansour Ciphers

In this section, we present our practical results for Even–Mansour Ciphers instantiated with S-Boxes $x \mapsto x^d$ for which the exponent is of the form $d = 2^{d'} + 1$ and $d = 2^{d'} - 1$ (see Table 4 and Table 5 for more details). We recall that $x \mapsto x^d$ is a permutation in $\mathbb{F}_{2^N}$ iff $\gcd(d, 2^N - 1) = 1$. Before going on, we recall that we work on Even–Mansour Ciphers rather than SPN ciphers in order to get results that are independent of the linear layer.

Estimation for $\mathcal{R}^{[\text{BCD11}]}$ and Bound for γ. In the case in which $d = 2^{d'} + 1$, the algebraic degree of the round function $x \mapsto x^d$ is always 2. This means that we can reuse the bound for γ given in [EGL+20, Sect. 3.1], namely $\gamma = (n+1)/2$.

In the case in which $d = 2^{d'} - 1$, the algebraic degree of the round function $x \mapsto x^d$ is equal to d'. Hence, it is not hard to show that

$$\gamma \leq \max\left\{ \frac{N - \lfloor (N-1)/d' \rfloor}{N - \delta \cdot \lfloor (N-1)/d' \rfloor}, n - \left\lfloor \frac{N-1}{d'} \right\rfloor - 1 \right\}. \tag{21}$$

Practical Results. As we have already explained in Section 3, we expect that the gap between the theoretical and the practical results can be larger for $d = 2^{d'} + 1$ than for the case $d = 2^{\delta} - 1$. Again, this is due to the fact that in the case $d = 2^{d'} + 1$, the encryption polynomial is in general far from being full (or at least dense for large d').

Table 5: Theoretical and practical round numbers *necessary* to prevent zero-sum distinguishers for the case of MiMC-like ciphers over $\mathbb{F}_{2^N}$, in the case of a round function of the form $R(x) = x^d$ for several values of $d = 2^{d'} - 1$.

Param.		Theoretical		Practical
d	N	$\mathcal{R}^{\text{Linear}}$	$\mathcal{R}^{[\text{BCD11}]}$	Practical $\mathcal{R}$
7	13	5	3	5
	17	7	3	7
	33	12	4	14
	65	24	4	-
	128	46	5	-
15	13	4	2	4
	17	5	3	5
	33	9	3	9
	65	17	4	-
	129	34	4	-
31	13	3	2	4
	17	4	2	5
	33	7	3	8
	64	13	3	-
	129	27	3	-
63	13	3	2	3
	17	3	2	6
	35	6	2	8
	65	11	3	-
	129	22	3	-

Our practical results confirm this analysis. In more details, we found that – especially in the case in which d is smaller than (or comparable to) the size of the field – the polynomial that describes the encryption function is in general dense (even if it is not always full) in the case $d = 2^{d'} - 1$. In the case in which d is bigger than (or comparable to) the size of the field, a possible gap can occur between the practical round numbers necessary to provide security and the predicted theoretical ones. We found that the reason of this is due to the fact that the encryption polynomial is in general sparse[11].

As a result, for $d = 2^{d'} - 1$ the real number of rounds that provides security against (secret-key) zero-sum distinguishers is close to $\mathcal{R}^{\text{Linear}}$ previously given (even if it can be a little bigger), while the gap between these two numbers is in general bigger for the case $d = 2^{d'} + 1$. This fits with the theoretical prediction made before (and with the results regarding "Variants of MiMC" presented in [AGR+16]). For completeness, we point out that $\mathcal{R}^{\text{Linear}}$ is similar (or even equal) for $d = 2^{d'} + 1$ and $d = 2^{d'} - 1$, since

$$\log_{2^{d'}+1}(2^N - 1) \approx \log_{2^{d'}-1}(2^N - 1) \approx \frac{N}{d'}.$$

However, we remark one more time that the gap between $\mathcal{R}^{\text{Linear}}$ and the (practical/real) number of rounds necessary to provide security is in general bigger for $d = 2^{d'} + 1$ than for $d = 2^{d'} - 1$.

[11] See the discussion given in Section 3 for $d = 3$, where we show that even if the polynomial is full after the first round, it may not be full after the next rounds.

C Closed Formula for Partial-SPN Ciphers– Proof

Proposition 10. *Let F be a function from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ corresponding to the concatenation of 1 smaller balanced S-Box S defined over $\mathbb{F}_2^n$ and an identity function over $\mathbb{F}_{2^{N-n}} \equiv \mathbb{F}_{2^{n\cdot(t-1)}}$. Moreover, assume that the S-Box S has algebraic degree $\delta \geq 2$, and let $\beta = \max_{1\leq i\leq n}(\delta_i - i)$, where δ_k is the maximum algebraic degree of the product of any k output bits of the S-Box.*

Then, for any affine functions $L_1, L_2, \ldots, L_{t-1}$ from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$, consider the encryption function E from $\mathbb{F}_2^N$ to $\mathbb{F}_2^N$ defined as

$$E^{(r)}(\cdot) = \underbrace{F}_{R\text{-}th\ round} \circ \underbrace{L_{R-1}\circ F}_{(t-1)\text{-}th\ round} \circ \cdots \circ \underbrace{L_2 \circ F}_{2\text{-}nd\ round} \circ \underbrace{L_1 \circ F}_{1\text{-}st\ round} (\cdot)$$

*for a certain number of rounds $r \geq 1$. The minimum number of rounds $\mathcal{R}^{[ARS+15]}$ of $E^{(\mathcal{R}^{[ARS+15]})}$ **necessary** to prevent (secret-key) zero-sum distinguishers on N bits is defined as in Eq. (19). [12] In a similar way, in order to prevent zero-sum distinguisher on $t-1$ words, (approximately) $1+\mathcal{R}^{[ARS+15]}$ rounds of $E^{(\mathcal{R}^{[ARS+15]})}$ – where $\mathcal{R}^{[ARS+15]}$ is defined as before – are needed.*

Proof. The proof of this proposition is similar to the one given in Appendix A. For this reason, here we limit ourselves to highlight the main differences.

As before, let

$$f(x) = \delta \cdot x, \qquad h(x) = \beta \cdot m + x, \qquad g(x) = \frac{N\cdot(\gamma-1)}{\gamma} + \frac{x}{\gamma},$$

where δ, β, m, γ are constants. It follows that

$$\left[f'(x) = \delta\right] > \left[h'(x) = 1\right] > \left[g'(x) = \frac{1}{\gamma}\right],$$

since $\delta \geq 2$ and $\gamma \geq 1$.

1st Part. Due to the bounds already recalled, we know that

$$\deg(G \circ F(\cdot)) \leq \min\left\{\deg(G)\cdot\deg(F), \beta\cdot m + \deg(G), \frac{N\cdot(\gamma-1)}{\gamma} + \frac{\deg(G)}{\gamma}\right\},$$

where β is defined as before.

First of all, note that we are not considering a generic function G, but a concrete encryption function $E^{(R)}$. Since the growth of the degree is independent of any affine/linear function, it follows that the degree of such an encryption function $E^{(R)}$ is upper-bounded by

$$\deg\left(E^{(R)}\right) \leq \delta^R$$

(note that we are working with S-Boxes with equal algebraic degree).

In order to apply the bound from Eq. (4), the first step is to find the maximum number of rounds R_0 such that $\deg(G)\cdot\deg(F) \leq \beta + \deg(G)$. In our case, this corresponds to

$$\delta^{R_0+1} \leq \beta + \delta^{R_0},$$

[12] We observe that the result is meaningful since

1*st*) $\left\lfloor \log_\delta\left(\frac{\beta}{\delta-1}\right)\right\rfloor \geq 0$ since $\beta \geq \delta - 1$;

2*nd*) $\left\lfloor \frac{N}{\beta} - \frac{1}{\delta-1} - \frac{\gamma}{\gamma-1}\right\rfloor \geq 0$ since $\frac{1}{\delta-1} + \frac{\gamma}{\gamma-1} = 1 + \frac{1}{\delta-1} + \frac{1}{\gamma-1} \leq 3$ and since $\frac{N}{\beta} \geq \frac{N}{n-2} \geq t\cdot(2+\frac{2}{n-2}) \geq 2t \geq 4$ (where $\beta \leq n-2$);

3*rd*) $\left\lfloor \log_\gamma\left(\frac{\beta}{\gamma-1}\right)\right\rfloor \geq 0$ since $\beta \geq \gamma - 1$.

which implies that

$$R_0 = \left\lfloor \log_\delta\left(\frac{\beta}{\delta-1}\right)\right\rfloor.$$

Note that $\beta \geq \delta - 1$ by definition. As a result, $R_0 \geq 0$.

Next, we find the maximum number of rounds R_1 such that $\beta + \deg(G) \leq \frac{N\cdot(\gamma-1)}{\gamma} + \frac{\deg(G)}{\gamma}$. This corresponds to

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

which implies

$$R_1 = \left\lfloor \frac{N}{\beta} - \frac{\delta^{R_0}}{\beta} - \frac{\gamma}{\gamma-1}\right\rfloor.$$

Finally, we need to find the minimum number of rounds in order to prevent higher-order differential attacks. Remember that given a function $f(\cdot)$ of degree d, the sum over the outputs of the function applied to all elements of a vector space $\mathcal{V}$ of dimension $\geq \delta + 1$ is zero. Since we only consider attacks that use less than the full code book, the biggest subspace of $\mathbb{F}_{2^N}$ has dimension $N-1$. As a result, a zero-sum distinguisher can be set up for at most $\mathcal{R}^{[ARS+15]}$ rounds, where

$$\deg\left(E^{(\mathcal{R}^{[ARS+15]})}\right) \leq N-1.$$

For this last step, we can simply reuse the computation proposed in the previous proof. It follows that $\deg\left(E^{(R)}\right) \leq N-1$ if

$$\frac{N\cdot(\gamma^{R_2}-1)}{\gamma^{R_2}} + \frac{\beta\cdot R_1 + \delta^{R_0}}{\gamma^{R_2}} \leq N-1,$$

which means that

$$R_2 = \left\lfloor \log_\gamma\left(N - \beta\cdot R_1 - \delta^{R_0}\right)\right\rfloor.$$

In conclusion, the minimum number of rounds $\mathcal{R}^{[ARS+15]}$ of $E^{\left(\mathcal{R}^{[ARS+15]}\right)}$ necessary to prevent (secret-key) zero-sum distinguishers on N bits is given by

$$\begin{aligned}\mathcal{R}^{[ARS+15]} &\geq \underbrace{\left\lfloor \log_\delta\left(\frac{\beta}{\delta-1}\right)\right\rfloor}_{\equiv R_0} + \underbrace{\left\lfloor \frac{N}{\beta} - \frac{\delta^{R_0}}{\beta} - \frac{\gamma}{\gamma-1}\right\rfloor}_{\equiv R_1} + \left\lceil \log_\gamma\left(N - \beta\cdot R_1 - \delta^{R_0}\right)\right\rceil \approx \\ &\approx 1 + \left\lfloor \log_\delta\left(\frac{\beta}{\delta-1}\right)\right\rfloor + \left\lfloor \frac{N}{\beta} - \frac{1}{\delta-1} - \frac{\gamma}{\gamma-1}\right\rfloor + \left\lceil \log_\gamma\left(\frac{\beta}{\gamma-1}\right)\right\rceil,\end{aligned}$$

where the last inequality is an equality if

$$\log_\delta\left(\frac{\beta}{\delta-1}\right) \qquad \text{and} \qquad \frac{N}{\beta} - \frac{\delta^{R_0}}{\beta} - \frac{\gamma}{\gamma-1}$$

are integer numbers.

2nd Part. The previous formula is obtained by considering the case in which the attacker works with a subspace of size $N-1$ bits. However, another strategy can be better in general. Due to the partial S-Box layer, the attacker can skip some initial rounds by exploiting a clever choice of the input subspace set. In particular, by skipping x rounds,

the subspace that the attacker can exploit has dimension at most $n \times (t-x)$. It follows that

$$\begin{aligned}\mathcal{R}^{[ARS+15]} \geq \max_{0\leq x\leq t-1}\Bigg\{x + \underbrace{\left\lfloor \log_\delta\left(\frac{\beta}{\delta-1}\right)\right\rfloor}_{\equiv R_0} + \underbrace{\left\lfloor \frac{n\times(t-x)}{\beta} - \frac{\delta^{R_0}}{\beta} - \frac{\gamma}{\gamma-1}\right\rfloor}_{\equiv R_1} + \\ + \left\lceil \log_\gamma\left(n\times(t-x) - \beta\cdot R_1 - \delta^{R_0}\right)\right\rceil\Bigg\} \approx \\ \approx \max_{0\leq x\leq t-1}\left\{x + 1 + \left\lfloor \log_\delta\left(\frac{\beta}{\delta-1}\right)\right\rfloor + \left\lfloor \frac{n\times(t-x)}{\beta} - \frac{1}{\delta-1} - \frac{\gamma}{\gamma-1}\right\rfloor + \right. \\ \left. + \left\lceil \log_\gamma\left(\frac{\beta}{\gamma-1}\right)\right\rceil\right\}.\end{aligned}$$

3rd Part: Conclusion. What is the best choice for x (equivalently, how many rounds x does it make sense to "skip")? By simple computation, note that

$$\frac{\mathrm{d}}{\mathrm{d}\,x}\mathcal{R}^{[ARS+15]} \simeq 1 - \frac{n}{\beta}.$$

By definition of β, it follows that

$$1 \leq \beta = \max_{1\leq i\leq n}(\delta_i - i) \leq n-1$$

since $-i \leq -1$ and since $\delta_i \leq n$ (the product of i coordinates of an S-Box over n cannot be of degree bigger than n). As a result

$$\frac{\mathrm{d}}{\mathrm{d}\,x}\mathcal{R}^{[ARS+15]} \simeq 1 - \frac{n}{\beta} \leq 1 - \frac{n}{n-1} = -\frac{1}{n-1} < 0,$$

where remember that $n \geq 3$. This means that the maximum is taken for $x = 0$, that is, the best choice is $x = 0$. □