The ABC of Symmetric Primitives over Integer Rings: Milk Before Meat
(Extended Version)

Tim Beyne[1] and Lorenzo Grassi[2] and Morten Øygarden[3] and Berenika Richterová[2] and Arne Sandrib[4]

[1] KU Leuven, Belgium
[2] Eindhoven University of Technology, Netherlands
[3] University of Bergen and Simula UiB, Norway
[4] University of Bergen and Nasjonal Sikkerhetsmyndighet, Norway
tim.beyne@esat.kuleuven.be, l.grassi@tue.nl, m.oygarden@uib.no, b.richterova@tue.nl, arne.sandrib@student.uib.no

Abstract. Designing a secure symmetric-key cipher over a vector space over a field $\mathbb{F}_{p^n}^t$ is well known and understood by the cryptographic community. Even if the attacks are continuously improving, our current understanding regarding the design and security of the majority of the symmetric-key primitives has not fundamentally changed in the last 20 years.
How does this picture change when we move to an integer ring $\mathbb{Z}_{p^n}^t$? Although the question is easy to state, it turns out to be far harder to answer. Indeed, there is a significant difference between the arithmetics of $\mathbb{F}_{p^n}^t$ and $\mathbb{Z}_{p^n}^t$ and attack vectors do not apply/translate directly between the two. As a case in point, a few ciphers have already been designed over integer rings, yet their initial versions have already been broken.
In this paper, we lay the foundations for a more rigorous approach to designing ciphers over integer rings, noting that this is not only of theoretical interest, but also has concrete applications. We analyze how existing statistical and algebraic attacks will behave for these ciphers and also present new attacks that take into account that not all functions over integer rings admit a polynomial representation. Based on this, we discuss possible design strategies, in which we analyze the security effect of having/not having polynomial S-boxes. In particular, we introduce new properties for the non-polynomial S-boxes that measure their resistance against the attacks presented in this paper. Finally, we discuss how to design such non-polynomial S-boxes, presenting two concrete constructions, and one based on the "digit manipulation".

1 Introduction

The vast majority of the symmetric primitives have been designed over fields. Whether these are binary or primitive fields is typically determined by the application that the designers of the symmetric primitive have in mind. The binary field allows for good performance in software and hardware applications,

while recent applications such as Multi-Party Computation (MPC), Homomorphic Encryption (HE) and Zero-Knowledge (ZK) usually require specific prime fields, and even more recently, also integer rings.

Integer Rings in Applications. The areas of MPC, ZK, HE, and their associated symmetric primitives are constantly evolving. While traditionally defined over fields such as $\mathbb{F}_p$ and $\mathbb{F}_{p^n}$, there has recently been a surge of new MPC-protocols [27,38,32,37,40,63], ZK-protocols [9,10,39,64] and HE-protocols [34,25,20,21,22,23] defined over integer rings $\mathbb{Z}_{p^n}$. It is natural to expect that the efficiency of these protocols may be maximized with dedicated symmetric primitives defined directly over their respective native ring $\mathbb{Z}_{p^n}$, as we have seen for prime fields. This has indeed also been the case for the HE-friendly schemes Elisabeth [26], FRAST [24] and Rubato [51].

An example beyond the setting of MPC, ZK and HE is the tweakable wide block cipher Mystrium [1] defined over $\mathbb{Z}_{2^{32}}^t$, which is currently the fastest primitive on low-end processors that lack dedicated AES or other cryptographic instructions, such as ARM Cortex-A7.

Overall, the design of a symmetric primitive over an integer ring $\mathbb{Z}_{p^n}^t$ is not only of theoretical interest, but also of practical relevance.

From Fields to Integer Rings. In both the binary and prime setting, the choice of working with a field guarantees two crucial factors:

1. every non-zero element admits a multiplicative inverse;
2. every function over a field admits a polynomial representation.

However, both these two properties are lost when working over an integer ring $\mathbb{Z}_{p^n}$, since only non-zero elements co-prime with p admit a multiplicative inverse, and only a particular class of functions admits a polynomial representation. An open question regards the impact of these facts on the security of symmetric ciphers natively defined over an integer ring $\mathbb{Z}_{p^n}$.

In this paper, we face this problem, and more general, we provide a first answer to the following question: *How to design a secure symmetric primitive over an integer ring* $\mathbb{Z}_{p^n}^t$*?* Although the question may seem simple at first glance, no clear answer is provided in the literature.

1.1 Security of Ciphers over $\mathbb{Z}_{p^n}^t$

The properties of working over finite fields have crucial consequences for the design of symmetric primitives. Most symmetric algorithms are iterated schemes, in the sense that they are defined via the repeated application of a fixed round function. The number of these iterations is chosen to prevent all possible attacks known in the literature, including a certain security margin to hedge against future cryptanalytic improvements. The main attack vectors against symmetric cryptographic schemes can broadly be classified as statistical and algebraic attacks. As the name suggests, statistical attacks make use of "statistical properties" that hold for the attacked scheme with a non-negligible probability. Examples include linear [67] and differential [15] attacks. Algebraic attacks are based

on characteristics related to the polynomial expression that defines the scheme. Examples include interpolation [54], higher-order differential [57] and Gröbner basis attacks. By definition, algebraic attacks require the existence of a polynomial representation to work. At the current state of the art, it is therefore not clear to what extent this class of attacks can be a threat for ciphers over $\mathbb{Z}_{p^n}$ that do not admit a polynomial representation.

The uncertain state of cryptanalysis is reflected in the (few) symmetric ciphers recently designed over an integer ring $\mathbb{Z}_q$ for an integer q, including Elisabeth, FRAST and Rubato:

- Elisabeth does not admit a polynomial representation per se, but properties of the underlying non-linear S-boxes still allow for instantiations of algebraic attacks over $\mathbb{F}_2$, as shown in [42,4];
- FRAST admits a statistical distinguisher that may lead to an attack [5];
- Rubato admits a polynomial representation with error terms, and was broken for specific integers q in [44]. As a consequence of this, in a later update, the authors of Rubato restrict to prime q, which mitigates this attack.

Both Elisabeth and FRAST argue resistance against algebraic attacks by analogy with $\mathbb{F}_{p^n}$ – see [24, Section 5.1]: "*We consider an XOR-variant of FRAST with addition and constant multiplication over $\mathbb{Z}_{16}$ replaced by XOR and multiplication over $\mathbb{F}_{2^4}$. Such an approach has also been taken for the algebraic analysis of the Elisabeth cipher*". Thus the current best approach to arguing security for ciphers over $\mathbb{Z}_{p^n}$ can be summarized by the following three steps:

1. consider an equivalent version of the cipher defined over $\mathbb{F}_{p^n}$ (or $\mathbb{F}_p^n$);
2. show that this version defined over $\mathbb{F}_{p^n}$ is secure;
3. *claim* that the original cipher over $\mathbb{Z}_{p^n}$ is secure as well.

While this may be a pragmatic first step to argue security against algebraic attacks, it remains a heuristic one at best. In particular, we remark that there is no evidence to support the correctness of this approach.

From this current state of cryptanalysis, we conclude that the cryptographic community is not able to sufficiently judge (and hence guarantee) the security of a symmetric primitive over $\mathbb{Z}_q$, and it is not advisable to consider ciphers over integer rings for concrete applications and use-cases. The aim of this paper is to start filling this gap.

1.2 Outline of the Paper and Our Contribution

The goal of this paper is to present concrete criteria for designing a secure cipher over an integer ring $\mathbb{Z}_{p^n}^t$. We start in Section 2 by recalling known results about functions over $\mathbb{Z}_{p^n}$. Next, in Section 3, we outline the high-level design strategy for the ciphers that we will consider in this paper. As the majority of the ciphers over $\mathbb{F}_{p^n}^t$ (including $p = 2$) are based on the SHARK/AES [70,29,31], we decided to focus our attention on the design of a SPN cipher that resembles SHARK/AES. This means that our round functions will consist of an S-box layer, a matrix multiplication and round-key addition.

For the linear layer, we show that it is possible to deduce bounds on the probability of the differential characteristic and linear correlations by applying the wide-trail design strategy to the $\mathbb{Z}_{p^n}^t$ case, keeping in mind that Maximum Distance Separable (MDS) matrices may not exist over, e.g. $\mathbb{Z}_{2^n}^t$. For the choice of non-linear layer, there are three main directions to take: i) a layer of S-boxes that can be represented by polynomials over $\mathbb{Z}_{p^n}$; ii) S-boxes without such representation; or iii) a mix of these two. Each of these options has its own advantages and disadvantages, which we explore in the rest of the paper.

Are Polynomial Ciphers Secure? In Section 4, we re-visit the algebraic and the statistical attacks against ciphers that admit a polynomial representation over $\mathbb{Z}_{p^n}^t$, first presented in [44]. The idea is to work iteratively modulo $p, p^2, \ldots, p^n$, by lifting the secret key via this polynomial structure. As expected, a polynomial cipher cannot be secure if p^t is too small, as this approach allows for iterated exhaustive searches.

New Attacks for Non-Polynomial Ciphers. In practice, the parameters p, n and t are determined by the considered application. If the goal is to design a cipher over a particular integer ring $\mathbb{Z}_{p^n}^t$, non-polynomial components may be required to achieve security. It is then crucial to understand how to choose these non-polynomial components, and how they affect the security of the cipher. To answer this question, we need to understand how such a non-polynomial cipher can be attacked. We address this by presenting multiple attacks, including:

- probabilistic algebraic attacks that exploit a good polynomial approximation of the cipher in Section 6;
- a new truncated differential attack that works modulo p^ℓ for $\ell < n$ in Section 7, exploiting the fact that a polynomial function preserves the zero-differences modulo p^ℓ, while this happens with different probability depending on the considered non-polynomial function.

To understand the complexities of these attacks, we introduce new properties in Section 5 to measure how "good" a non-polynomial S-box is in terms of resisting these attacks. These properties include the *ℓ-Polynomial-Correlation* ($\texttt{PC}^{(\ell)}$) and the *Congruence-Preservation Probability* modulo p^ℓ ($\texttt{CPP}^{(\ell)}$). The former can be seen as a measure on how "close" the S-box is to a polynomial function, while the latter measures the probability of zero-difference being preserved modulo p^ℓ.

Designing Optimal Non-Polynomial S-Boxes over $\mathbb{Z}_{p^n}$. Putting all together, in Section 8 we list the criteria that optimal S-boxes over $\mathbb{Z}_{p^n}$ should satisfy. This includes minimizing both $\texttt{PC}^{(\ell)}$ and $\texttt{CPP}^{(\ell)}$. We conclude in Section 9 with concrete examples of non-polynomial S-boxes over $\mathbb{Z}_{p^n}$ that behave well with respect to these criteria, taking inspiration from the the "cut and sew" approach [44] and the "`Kintsugi`" one [46] already used in Reinforced Concrete [45], and its follow-up works Monolith [46] and Skyscraper [18].

Remark 1. Although designing a dedicated symmetric primitive over $\mathbb{Z}_{p^n}^t$ for any of the applications listed above is out of the scope of this paper, we do believe that the results proposed here will provide a crucial step towards this goal.

2 Preliminaries

For an integer $q = p_1^{n_1} \cdots p_z^{n_z}$, where $p_1, \ldots, p_z$ are distinct primes, we recall that an element in $\mathbb{Z}_q$ (e.g., part of a secret key) can be recovered from its values modulo $\mathbb{Z}_{p_i^{n_i}}$, by the Chinese Remainder Theorem. Consequently, we restrict ourselves to work over $\mathbb{Z}_{p^n}$ throughout the paper. At times – when there is no likelihood of confusion – we will also adopt the convenient abuse of notation of identifying a congruence class in $\mathbb{Z}_{p^n}$ by its unique integer in $\{0, \ldots, p^n - 1\}$.

2.1 Background

We start by introducing some crucial notions for the following.

Definition 1. *For each $i \geq 0$, a falling factorial polynomial $(x)_i$ is defined as $(x)_i \coloneqq (x - (i-1)) \cdot (x)_{i-1}$, where $(x)_0 = 1$. Equivalently, $(x)_i = \prod_{j=0}^{i-1}(x - j)$.*

Definition 2. *The $v_p(\cdot)$ denotes the p-adic valuation function defined as $v_p(m) \coloneqq \max\{i \in \mathbb{Z}_{\geq 0} \mid m = 0 \mod p^i\}$ if $m \neq 0$, and ∞ otherwise. The $\mu(\cdot)$ denotes the Smarandache function defined as $\mu(q) \coloneqq \min\{i \in \mathbb{N} \mid i! \bmod q = 0\}$.*

Based on these two definitions, note that $\mu(p^n)$ is the smallest integer for which $v(\mu(p^n)!) \geq n$. We limit ourselves to recall that $\mu(q) \leq q$ and $\mu(q) = q$ if and only if q is a prime or $q \in \{1, 4\}$, as proved in [74]. Moreover, we provide concrete upper bounds for $\mu(p^n)$ in Appendix A.1, showing in particular that $\mu(p^n) \leq n \cdot p$.

2.2 Polynomial Functions over $\mathbb{Z}_{p^n}$

We start with the following definition.

Definition 3. *If a function $\mathsf{F} \colon \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}$ admits a polynomial representation, i.e., if there exists a polynomial function $\mathfrak{f}$ with coefficients in $\mathbb{Z}_{p^n}$ such that $\mathfrak{f}(x) \bmod p^n = \mathsf{F}(x) \bmod p^n$ for each $x \in \mathbb{Z}_{p^n}^t$, then F is said to be a polyfunction.*

As it is well known, not every function is a polyfunction modulo p^n. We show a necessary condition for being a polyfunction in the following. For the particular case of invertible polyfunctions, a stronger result can be proved.

Proposition 1. *Let $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}$ be a polyfunction. Then,*

$$x \bmod p^i = y \bmod p^i \qquad \Longrightarrow \qquad \mathsf{F}(x) \bmod p^i = \mathsf{F}(y) \bmod p^i \tag{1}$$

for all $i \in \{0, 1, \ldots, n-1\}$. Let $\mathsf{F} \equiv (\mathsf{F}_0, \mathsf{F}_1, \ldots, \mathsf{F}_{t-1}) : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}^t$ be a permutation polyfunction. Then, for all $i \in \{0, 1, \ldots, n-1\}$, we have

$$x \bmod p^i = y \bmod p^i \qquad \Longleftrightarrow \qquad \mathsf{F}(x) \bmod p^i = \mathsf{F}(y) \bmod p^i \,. \tag{2}$$

For $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}$ being polynomial, the statement (1) is a straightforward generalization of the analogous statement for functions over $\mathbb{Z}_{p^n}$ from [19]. The proof for the second part is given in Appendix A.2.

We emphasize that preserving congruences in this manner is not a sufficient condition to be a polyfunction, as seen in the following example.

Example 1. The function F defined over $\mathbb{Z}_{2^4}$ as

input	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
output	3	8	5	14	7	4	9	2	11	0	13	6	15	12	1	10

($\mathsf{F} = [3, 8, 5, 14, 7, 4, 9, 2, 11, 0, 13, 6, 15, 12, 1, 10]$ in a compact notation) preserves the modulo congruences modulo 2, 4, and 8, but does not admit any polynomial representation modulo 16.

The following result is proved in Appendix A.3.

Proposition 2. *Consider two polyfunctions F, G over $\mathbb{Z}_{p^n}$ for $n > 1$. If there exists $a \in \mathbb{Z}_{p^n}$ such that $\mathsf{F}(x \bmod p^n) = \mathsf{G}(x \bmod p^n)$ for each $x \in \{a, a+1, \ldots, a+\mu(p^n)-1\}$, then $\mathsf{F} = \mathsf{G}$ on the entire domain.*

Criteria for Being a Polyfunction & Null Polyfunctions. Without going into details, it is possible to check whether a function is a polyfuction using the results given in [41], which we briefly recall in Appendix A.4.

The fact that not all functions admit a polynomial representation is strictly related to the existence of null polyfunctions.

Definition 4. *A polyfunction O over $\mathbb{Z}_{p^n}$ is a null polynomial function if $\mathsf{O}(x) = 0 \mod p^n$ for each $x \in \mathbb{Z}_{p^n}$.*

Theorem 1 ([74,41]). *A polyfunction O over $\mathbb{Z}_{p^n}$ is a null polynomial modulo p^n of degree at most d if and only if there exist $a_0, a_1, \ldots, a_d \in \mathbb{Z}_{p^n}$ such that $\mathsf{O}(x) \coloneqq \prod_{i=0}^{d} a_i \cdot \mathsf{O}_i(x)$ with $\mathsf{O}_i(x) = p^{\max\{n-\upsilon_p(i!),0\}} \cdot (x)_i$. In this equation, the exponent of p equals 0 if $i \geq \mu(p^n)$.*

We highlight that the evaluation of $(x)_i$ at any integer is divisible by $p^{\upsilon_p(i!)}$. Hence, each term in $\mathsf{O}_i(x)$ evaluated at any integer is divisible by $p^{\max\{n,\upsilon_p(i!)\}} \geq p^n$, that is, it is zero modulo p^n.

Polynomial Representation. Due to the previous result, it follows that the representation of any polyfunction is never unique. Indeed, the polyfunctions F and $\mathsf{F} + \mathsf{O}$ for any null polynomial function O represent the same function.

In particular, the following result holds.

Proposition 3 ([56]). *Each polyfunction F over $\mathbb{Z}_{p^n}$ has a representation of degree strictly less than $\mu(p^n)$.*

The proof is related to the fact that $(x)_{\mu(p^n)}$ is the monic null polynomial function of lowest degree due to Theorem 1. Hence, $\mathsf{F}(x) \equiv \mathsf{F}(x) \mod (x)_{\mu(p^n)}$ for each $x \in \mathbb{Z}_{p^n}$. Although this guarantees a representation of degree less than $\mu(p^n)$, the obtained representation is not necessarily the one with minimum degree. In order to guarantee the lowest possible degree, one has to consecutively divide by $\mathsf{O}_i(x)$ defined as before for $i \in \{\mu(p^n)-1, \ldots, 0\}$. This leads to the following canonical representation.

Theorem 2 ([56,74]). *Let* $\mathsf{F} \in \mathbb{Z}_{p^n}$ *be a polyfunction, then there exists a unique canonical representation* $\mathsf{F}(x) = \sum_{i=0}^{\mu(p^n)-1} c_i \cdot (x)_i \equiv \sum_{i=0}^{\mu(p^n)-1} b_i \cdot x^i$ *with* $0 \leq c_i, b_i < p^{n-v_p(i!)} \equiv p^n/\gcd(i!, p^n)$.

Number of Polyfunctions over $\mathbb{Z}_{p^n}$. In Appendix A.5, we obtain a result on the number of distinct polyfunctions. We also recall the number of null polynomials of degree less than $\mu(p^n)$. For comparison, in Appendix A.7, we provide an estimate of the number of functions that preserve the congruences.

Polyfunctions over $\mathbb{Z}_{p^n}^t$. We limit ourselves to mention that the previous results can be extended to the case of polyfunctions over $\mathbb{Z}_{p^n}^t$ as well, by keeping in mind that the evaluation of $(x)_{\mu(p^n)}$ at any integer is equal to zero modulo p^n. Hence, the following result holds:

Proposition 4. *Each polyfunction* $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}$ *has a representation of degree strictly less than* $t \cdot (\mu(p^n) - 1) + 1$.

As before, the obtained representation of the polyfunction does not have a minimal degree in general. By iterating the process just described for $p^{n-v_p(i)} \cdot (x)_i$, it is possible to guarantee the lowest possible degree, and so a canonical form. Hence, the result proposed in Theorem 2 holds for a polyfunction over $\mathbb{Z}_{p^n}^t$ as well.

Polynomial Permutations over $\mathbb{Z}_{p^n}$**.** Hensel's Lemma (recalled in Proposition 13, Appendix A.6) allows to check if a given polynomial function over $\mathbb{Z}_{p^n}$ is invertible or not. The conditions provided by this result can be rewritten in the following form, which can be efficiently verified. In the following, $\mathsf{F}'(x) \coloneqq \sum_{i=1}^{d} i \cdot a_i \cdot x^{i-1}$.

Corollary 1. *Let* F *be a polyfunction over* $\mathbb{Z}_{p^n}$. *It permutes the elements of* $\mathbb{Z}_{p^n}$ *for* $n > 1$ *if and only if*

1. $\mathsf{F}(\cdot) \bmod p$ *permutes the elements of* $\mathbb{Z}_p$, *and*
2. *the formal derivative* $\mathsf{F}'(x) \bmod p \neq 0$ *for every integer* $x \in \mathbb{Z}_p$.

In Appendix A.6, we further recall several useful results regarding the polyfunctions over $\mathbb{Z}_{p^n}$, including a detailed analysis for the special case $\mathbb{Z}_{2^n}$.

2.3 Cryptanalysis over $\mathbb{Z}_{p^n}$

Linear and differential cryptanalysis are the most widespread techniques used to analyze the security of primitives over $\mathbb{F}_2^n$. Both of these methods have been generalized to arbitrary finite Abelian groups such as the additive group of the ring $\mathbb{Z}_{p^n}$. However, the specific structure of $\mathbb{Z}_{p^n}$ has interesting consequences that have not been systematically discussed in the literature.

Linear Cryptanalysis. Linear cryptanalysis [67] is a general form of cryptanalysis based on finding affine approximations to the action of a cipher. Introduced by Matsui, it was initially generalized to finite Abelian groups by Baignères,

Stern and Vaudenay [3] in the key-averaged setting. A description on par with the theory for $\mathbb{F}_2^n$ that makes no assumptions about key-dependence was given in [11, Section 3.2]. In this paper, we follow the latter work.

A linear approximation of a function $\mathsf{F}\colon \mathbb{Z}_{p^n}^s \to \mathbb{Z}_{p^n}^t$ is a pair (ψ, χ) of characters of the groups $\mathbb{Z}_{p^n}^s$ and $\mathbb{Z}_{p^n}^t$ respectively. It is possible to identify the characters of $\mathbb{Z}_{p^n}^s$ with the elements of $\mathbb{Z}_{p^n}^s$ (though, in more than one way). Explicitly, let ω be a p^n-th primitive root of unity such as $\omega = \exp(2\pi\sqrt{-1}/p^n)$. The character corresponding to u is the map $\chi_u\colon \mathbb{Z}_{p^n}^s \to \mathbb{C}^\times$ defined by $\chi_u(x) = \omega^{u \cdot x}$ with $u \cdot x = \sum_{i=1}^s u_i x_i$ for some u in $\mathbb{Z}_{p^n}^s$. Up to this identification, the linear approximation (χ_u, χ_v) is the same as a pair (u, v), with u in $\mathbb{Z}_{p^n}^s$ and v in $\mathbb{Z}_{p^n}^t$. Its correlation is equal to

$$C_{v,u}^{\mathsf{F}} = \frac{1}{p^{n \cdot s}} \sum_{x \in \mathbb{Z}_{p^n}^s} \chi_v(\mathsf{F}(x))\, \chi_u(-x)\,.$$

A linear approximation for F with correlation c allows distinguishing F from a uniform random function or permutation with data-complexity proportional to $1/|c|^2$ when the success probability and false-positive rate are constant.

The matrix C^{F} with coordinates $C_{v,u}^{\mathsf{F}}$ is called the correlation matrix of F [11, Definition 3.3]. The basic properties of correlations matrices, such as $C^{\mathsf{F}_1 \circ \mathsf{F}_0} = C^{\mathsf{F}_1} \times C^{\mathsf{F}_0}$, lead to linear trails and their propagation rules. In particular, a linear trail for a function $\mathsf{F} = \mathsf{F}_{r-1} \circ \cdots \circ \mathsf{F}_1 \circ \mathsf{F}_0$ is a sequence $(u_0, u_1, \ldots, u_r)$. By the multiplicative property of correlation matrices, $C_{u_r,u_0}^{\mathsf{F}} = \sum_{u_1,\ldots,u_{r-1}} \prod_{i=0}^{r-1} C_{u_{i+1},u_i}^{\mathsf{F}_i}$, where the product $\prod_{i=0}^{r-1} C_{u_{i+1},u_i}^{\mathsf{F}_i}$ is called the correlation of the trail $(u_0, u_1, \ldots, u_r)$.

Differential Cryptanalysis. Differential cryptanalysis [15] is a general form of cryptanalysis that studies how differences in input can affect the resulting difference at the output. Initially introduced by Biham and Shamir, follow-up generalizations of differential cryptanalysis to $\mathbb{Z}_{2^n}$ go back to the early work on collision attacks, such as the work of Dobbertin [36]. A differential of a function $\mathsf{F}\colon \mathbb{Z}_{p^n}^s \to \mathbb{Z}_{p^n}^t$ is a pair (δ_I, δ_O) of differences in $\mathbb{Z}_{p^n}^s$ and $\mathbb{Z}_{p^n}^t$, respectively. The probability of a differential (δ_I, δ_O) is equal to

$$\mathrm{DP}_{\delta_O,\delta_I}^{\mathsf{F}} \coloneqq \frac{\left|\left\{x \in \mathbb{Z}_{p^n}^s \mid \mathsf{F}(x + \delta_I) = \mathsf{F}(x) + \delta_O\right\}\right|}{p^{n \cdot s}}\,.$$

A differential for F with probability ρ allows distinguishing F from a uniform random function or permutation with data-complexity proportional to $1/\rho$ when the success probability and false-positive rate are constant.

The matrix DP^{F} with coordinates $\mathrm{DP}_{\delta_O,\delta_I}^{\mathsf{F}}$ will be called the difference-distribution matrix of F. While it is *not* true that $\mathrm{DP}^{\mathsf{F}_1 \circ \mathsf{F}_0} = \mathrm{DP}^{\mathsf{F}_1} \times \mathrm{DP}^{\mathsf{F}_0}$, this does hold on the average for key-alternating ciphers with respect to independent and uniform random round keys. We limited ourselves to mention that this assumption can be avoided by replacing DP^{F} with the quasi-differential transition matrix of F [12]. Similar to the case of linear cryptanalysis, a differential

trail for a function $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$ is a sequence of differences $(\delta_0, \delta_1, \ldots, \delta_r)$. Under the previous assumption of independent and uniform random round keys, $\mathrm{DP}^{\mathsf{F}}_{\delta_0,\delta_r} = \sum_{\delta_1,\ldots,\delta_{r-1}} \prod_{i=0}^{r-1} \mathrm{DP}^{\mathsf{F}}_{\delta_{i+1},\delta_i}$.

Integral Cryptanalysis. For primitives defined over large finite fields, integral cryptanalysis often leads to more significant results than linear and differential cryptanalysis. Traditionally, integral cryptanalysis is used to find affine subspaces $\mathcal{S}$ such that $\sum_{x\in\mathcal{S}} \mathsf{F}(x) = 0$ for a coordinate function F of a cryptographic primitive. This property can be obtained either by generalizing the approach initially introduced in the "square attack" [28] (i.e., by considering the propagation of structured sets with some constant and some saturated parts), or via a purely algebraic approach based on higher-order derivatives, as proposed by Knudsen [57]. Later, Todo [76] partially consolidated the two approaches by introducing the "division property", which characterizes structured sets algebraically.

More recently, integral cryptanalysis has been generalized to rings that are isomorphic to a product of finite fields by Beyne and Verbauwhede [13,14] (the so called, ultrametric integral cryptanalysis). However, the ring $\mathbb{Z}_{p^n}$ does not satisfy this condition. Hence, unlike for linear and differential cryptanalysis, there is currently no theory of integral cryptanalysis over the ring $\mathbb{Z}_{p^n}$. This issue was already pointed out in [44, Prop. 1], and it is closely related to the fact that not all functions on $\mathbb{Z}_{p^n}$ have a polynomial representation.

3 SHARK-/AES-like Ciphers over $\mathbb{Z}_{p^n}$

In this section, we introduce a generic SHARK-/AES-like cipher over $\mathbb{Z}^t_{p^n}$ which will be the analysis-target of this paper. While we limit ourselves to work on SPN ciphers to simplify the description of the analysis and the attacks presented in this paper, we highlight that our results work for other primitives as well including, e.g., Feistel.

Parameters and Key-Schedule. Consider a prime p, integers $n \geq 2$, $r, t \geq 1$ and let $k \in \mathbb{Z}^\kappa_{p^n}$ be the secret key. Moreover, let $\mathsf{K} : \mathbb{Z}^\kappa_{p^n} \to (\mathbb{Z}^t_{p^n})^*$ be a key-schedule where $\mathsf{K}(k) = (k_0, k_1, \ldots, k_r, \ldots)$. The ciphers we consider in this work will then be functions $\mathbb{Z}^t_{p^n} \to \mathbb{Z}^t_{p^n}$ given as the composition of r round functions.

Round Function. The i-th round function R_{k_i} of the iterated SPN cipher over $\mathbb{Z}^t_{p^n}$ is defined as

$$\mathsf{R}_{k_i}(x_0, x_1, \ldots, x_{t-1}) \coloneqq k_i + M \times (\mathsf{S}_{i,0}(x_0), \mathsf{S}_{i,1}(x_1), \ldots, \mathsf{S}_{i,t-1}(x_{t-1}))\,,$$

where for each $j \in \{0, 1, \ldots, t-1\}$:

- $\mathsf{S}_{i,j} : \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ is an invertible S-box function;
- $M \in \mathbb{Z}^{t\times t}_{p^n}$ is the invertible matrix that – combined with the non-linear layer – provides full diffusion in a finite number of rounds;
- $k_i \in \mathbb{Z}^t_{p^n}$ is the round key (which also includes the round-constant).

For an illustration of R_{k_i}, see Fig. 2 in Appendix B.1. In the following, we refer to $\mathsf{R}_i(\cdot)$ as the round function without the key addition, that is, $\mathsf{R}_{k_i}(\cdot) = k_i + \mathsf{R}_i(\cdot)$.

Remark 2. We emphasize that other strategies can be considered for constructing a cipher over $\mathbb{Z}_{p^n}^t$. For example, the matrix multiplication M can be e.g. replaced with a function over a Galois Ring $\mathrm{GR}(p^n, t)$ that provides diffusion. We leave the problem to study this case open for future work.

3.1 Choice of the Non-Linear Layer

As recalled earlier, not every function over $\mathbb{Z}_{p^n}$ is a polyfunction. Due to this fact, several instantiations of a symmetric primitive are possible. Here, we list three possible cases, highlighting possible advantages and disadvantages based on the analysis proposed in the following.

Class I: Polynomial S. As a first possibility, one may instantiate all the non-linear functions S as polynomials. However, as already shown in [44, Section 2.2] (and recalled in Section 4), the corresponding symmetric primitive over $\mathbb{Z}_{p^n}^t$ can be broken in time $\mathcal{O}(n \cdot p^t)$ via a brute force attack. Roughly speaking, this means that p^t must be sufficiently large in order to guarantee security, a condition that strongly depends on the considered application. In other words, *if this condition is not satisfied, it is not possible to design a secure symmetric primitive over $\mathbb{Z}_{p^n}^t$ instantiated with polyfunctions only.*

Class II: Non-Polynomial S. At the opposite end of the scale, one may ensure that none of the functions S admits a polynomial representation. One immediate consequence of this choice is that the algebraic attacks over $\mathbb{F}_{p^n}^t$ known in the literature would not work directly, as there is no immediate system of polynomial equations that succinctly describes the cipher.

That said, the elements of $\mathbb{Z}_{p^n}^t$ correspond one-to-one with the elements of $\mathbb{F}_{p^n}^t$ (and $\mathbb{F}_p^{nt}$), and it is always possible to construct a polynomial representation over the latter. However, the arithmetics of these domains do not coincide for $n \geq 2$, and we generally expect an algebraic representation over $\mathbb{F}_{p^n}^t$ to be dense and of high degree (for both linear and non-linear layers), making this attack approach infeasible.[1]

Further still, there are new attacks dedicated to work over $\mathbb{Z}_{p^n}^t$ that are possible in this setting, as we will discuss in Section 6 and 7. Since not all non-polynomial ciphers are equally secure against these attacks, we will discuss how to set up the non-polynomial components of the cipher in order to achieve security and efficiency.

Class III: Mixing Polynomial and Non-Polynomial S. Non-polynomial functions seem to be a good choice for arguing security against algebraic attacks, but for statistical attacks the picture is not so clear. In order to argue security against these attacks, designers usually try to maximize the number of active non-linear functions, and to choose non-linear S-boxes that have a small linear

[1] Special cases exists, as in the case of the attack against the Elisabeth-4 cipher over $\mathbb{Z}_{16}$ proposed in [42,4]. As the least significant bit of $\mathsf{S}(x)$ does not depend on the most significant bit of x, some monomials never appear in its algebraic normal form (ANF), a fact that can be exploited to break the full Elisabeth-4 cipher.

correlation/maximum differential probability. For non-polynomial functions the only way to measure the linear correlation/maximum differential probability of an S-box is through a brute force approach, which becomes infeasible for large p^n. Using polynomial functions can then be beneficial, as these two parameters can be upper bounded by considering the degree of the polynomial function itself (see Section 3.3 and 4 for more details). Thus the third class of ciphers we consider will combine both polynomial and non-polynomial S-boxes. We note that a similar design approach has also been adapted in the symmetric primitives Reinforced Concrete [45], Tip5 [75], Monolith [46] and Skyscraper [18].

3.2 The Linear Layer M

In the case of an SPN cipher, the linear layer is the only responsible for the diffusion, which is measured by the branch number [30,29]. We let $\mathrm{hw}(\cdot)$ denote the Hamming weight (or bundle weight, in the wide trail terminology).

Definition 5. *The* branch number *of a matrix $M \in \mathbb{Z}_{p^n}^{t\times t}$ is defined as $\mathcal{B}(M) = \min_{x\in\mathbb{Z}_{p^n}^t\setminus\{0\}}\{hw(x) + hw(M \times x)\}$. A matrix $M \in \mathbb{Z}_{p^n}^{t\times t}$ is called* Maximum Distance Separable *(MDS) if $\mathcal{B}(M) = t + 1$.*

An MDS matrix satisfies the following property.

Proposition 5. *$M \in \mathbb{Z}_{p^n}^{t\times t}$ is MDS if and only if every submatrix is invertible.*

The proof is analogous to that of the field case [66, page 321]. It is immediate that if $M \in \mathbb{Z}_{p^n}^{t\times t}$ is MDS, then every submatrix $M' \in \mathbb{Z}_{p^n}^{m\times m}$ of M is MDS over $\mathbb{Z}_{p^n}^m$. This provides a method for constructing MDS matrices over $\mathbb{Z}_{p^n}^t$.

Lemma 1. *Let $M \in \mathbb{Z}_{p^n}^{t\times t}$ be a matrix, and let $M' \in \mathbb{F}_p^{t\times t}$ be defined as $M' = M \mod p$ (that is, $M'_{i,j} = M_{i,j} \mod p$ for each $i, j \in \{0, 1, \ldots, t-1\}$). Then M is invertible over $\mathbb{Z}_{p^n}$ if and only if M' is invertible over $\mathbb{F}_p$.*

The proof is given in Appendix B.2. By combining Lemma 1 and Proposition 5, we get the following result.

Proposition 6. *Let $M \in \mathbb{Z}_{p^n}^{t\times t}$ be a matrix, and let $M' \in \mathbb{F}_p^{t\times t}$ be defined as $M' = M \mod p$. Then M is MDS over $\mathbb{Z}_{p^n}$ if and only if M' is MDS over $\mathbb{F}_p$.*

This result can be used to construct MDS matrices over $\mathbb{Z}_{p^n}^{t\times t}$ from MDS matrices over $\mathbb{F}_p^{t\times t}$, e.g., Cauchy matrices [72], sequential matrices [49,50], or Vandermonde matrices [59]. Indeed, for an MDS matrix $M' \in \mathbb{F}_p^{t\times t}$, we need only to ensure that each entry in $M \in \mathbb{Z}_{p^n}^{t\times t}$ coincides modulo p with that of M'.

Case: $\mathbf{p = 2}$. The previous result does not cover the case $p = 2$, as no MDS matrix exists in $\mathbb{F}_2^{t\times t}$. More generally, due to Proposition 6, it is possible to prove that no MDS matrix exists in $\mathbb{Z}_q^{t\times t}$ for q multiple of 2.

Corollary 2. *Let q be divisible by 2. There exists no MDS matrix $M \in \mathbb{Z}_q^{t\times t}$.*

Although no MDS matrix exists in $\mathbb{Z}_{2^n}^{t\times t}$, it is possible to construct *near-MDS* matrices, that is, matrices with branch number equal to t (instead of $t+1$). For instance, the circulant matrix $\mathrm{circ}(2,1,1,1)$ is a near-MDS invertible matrix over $\mathbb{Z}_{2^n}$. The authors of [62] propose a way to construct near-MDS matrices over $\mathbb{F}_{2^n}^t$ for $n \in \{4,8\}$ and $t \geq 5$. We limit ourselves to mention that these strategies can be adapted to the case $\mathbb{Z}_{2^n}^t$ as well.

3.3 Linear and Differential Cryptanalysis

If the SHARK-construction is instantiated with strong S-boxes, then there are known bounds for the key-averaged probabilities of differentials and key-averaged squared correlations of linear approximations. Results of this type were first obtained by Hong et al. [53] for the SHARK-like ciphers over $\mathbb{F}_2^t$, and their proofs readily generalize to arbitrary finite fields. It is less straightforward that these results remain valid over $\mathbb{Z}_{p^n}^t$, but we nonetheless have the following result.

Theorem 3. *Consider two rounds of a SHARK-like cipher over $\mathbb{Z}_{p^n}^t$. If the linear layer is MDS, and if the maximum squared correlation (respectively, differential probability) of the S-boxes is at most ε, then the key-averaged squared correlation (probability) of any nontrivial linear approximation (differential) with respect to uniform random round keys is at most ε^t.*

The proof is provided in Appendix B.3. The main reason the proof goes through is because Proposition 5 still holds for MDS matrices over $\mathbb{Z}_{p^n}$.

We emphasize that the result given in Theorem 3 is tight in the sense that if most effective differentials over the S-boxes have probability ε, then there exist roughly $1/\varepsilon$ differential characteristics that activate $t+1$ S-boxes and consequently have key-averaged probability ε^{t+1}.

4 Pitfalls of Polynomial S-Boxes

This section focuses on the security of ciphers in Class I, where each component can be described by a polynomial over $\mathbb{Z}_{p^n}$.

4.1 Algebraic Attacks on Polynomial Ciphers over $\mathbb{Z}_{p^n}^t$

Consider a cipher over $\mathbb{Z}_{p^n}^t$ that admits a polynomial representation. That is, for some parameter $\kappa \geq 1$ we let $k \in \mathbb{Z}_{p^n}^\kappa$ denote the secret key, and for each pair of plaintext/ciphertext $(x,y) \in \mathbb{Z}_{p^n}^t$, it is possible to construct a system of equations

$$\forall i \in \{0,1,\ldots,t-1\}: \qquad \mathsf{F}_i(x_0,x_1,\ldots,x_{t-1},k) = y_i \tag{3}$$

where $\mathsf{F}_i : \mathbb{Z}_{p^n}^t \times \mathbb{Z}_{p^n}^\kappa \to \mathbb{Z}_{p^n}$ is a polynomial function for each $i \in \{0,1,\ldots,t-1\}$. Given *known* pairs of plaintexts/ciphertexts, the attacker can now solve this system of equations to recover the unknown value k (i.e., the secret key).

As already shown in [44, Section 2.2], the corresponding symmetric primitive over $\mathbb{Z}_{p^n}^t$ can be broken in time $\mathcal{O}(n \cdot p^t)$ via brute force and Hensel lifting for increasing moduli $p, p^2, \ldots, p^n$. The brute force procedure can also be substituted for polynomial system solving, as we briefly discuss in the following.

Equation Solving over $\mathbb{Z}_{p^n}$. Based on Proposition 1, if k is a solution of $\mathsf{F}_h(x, k) = y_h$, then $k \bmod p^j$ is a solution of $\mathsf{F}_h(x, k) \bmod p^j = y_h \bmod p^j$. This suggests the following solving approach (also described in [44, Section 2.2]):

1. Re-write the system of equations modulo p, and solve it to find $k \bmod p$;
2. for each $1 \leq j \leq n-1$, given $k \bmod p^j$ (known due to the previous steps), re-write the system of equations modulo p^{j+1}, and solve it in order to find $k \bmod p^{j+1}$.

More concretely, we have that each key element $k_i \in \mathbb{Z}_{p^n}$ can be re-written as $k_i = \sum_{j=0}^{n-1} k_{i,j} \cdot p^j$. The iterations of this solving method then recovers $k_{1,j}, \ldots, k_{\kappa,j}$ for increasing moduli j. In particular, this approach then reduces to solving n polynomial systems over $\mathbb{Z}_p = \mathbb{F}_p$. For completeness, we offer a brief overview of the main methods in the algorithmic zoo for solving multivariate polynomial systems in Appendix C.

4.2 Linear and Differential Cryptanalysis

The following results show that if an S-box $\mathsf{S}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ is a polyfunction, then it usually does not have good linear and differential properties.

Linear Properties. We start with the linear properties.

Theorem 4. *Let $\mathsf{S}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ be a polyfunction. Let $\mathsf{S}'\colon \mathbb{Z}_{p^m} \to \mathbb{Z}_{p^m}$ be defined as $\mathsf{S}'(x) \coloneqq \mathsf{S}(x + p^m\mathbb{Z}_{p^n}) + p^m\mathbb{Z}_{p^n} \pmod{p^m}$. Then,*

$$C^{\mathsf{S}}_{p^{n-m}v',\, p^{n-m}u'} = C^{\mathsf{S}'}_{v',u'}\,,$$

for all masks u' and v' in $\mathbb{Z}_{p^m}$.

Proof. The function $x \mapsto v' \cdot \mathsf{S}(x) - u' \cdot x \pmod{p^m}$ is well-defined on $\mathbb{Z}_{p^n}/p^m\mathbb{Z}_{p^n}$, which is essentially the same as $\mathbb{Z}_{p^m}$. Hence, for each x in $\mathbb{Z}_{p^m}$, there are p^{n-m} values in $\mathbb{Z}_{p^n}$ on which the function takes the same value. It follows that

$$\begin{aligned} C^{\mathsf{S}}_{p^{n-m}v',\, p^{n-m}u'} &= \frac{1}{p^{nt}} \sum_{x \,\in\, \mathbb{Z}_{p^n}} \omega^{p^{n-m}(v' \cdot \mathsf{S}(x) - u' \cdot x)} \\ &= \frac{p^{(n-m)t}}{p^{nt}} \sum_{x \,\in\, \mathbb{Z}_{p^m}} \omega^{p^{n-m}(v' \cdot \mathsf{S}'(x) - u' \cdot x)}. \end{aligned}$$

The right-hand side above is equal to $C^{\mathsf{S}'}_{v',u'}$. □

The implication of Theorem 4 is that polyfunctions have linear approximations with high absolute correlation for masks divisible by p. Indeed, the function S' is a function on $\mathbb{Z}_{p^m}$, and in the best case has linear approximations of correlation $p^{-mt/2}$. This is much higher than $p^{-nt/2}$, which is the best case for a function on $\mathbb{Z}_{p^n}$. As a consequence of this, the following lower bound holds *independently of n and of the details of* S.

Corollary 3. *Let $\mathsf{S}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ be a polyfunction. Then, the maximum linear correlation of S – denoted by $C_{\max}(\mathsf{S})$ – is lower bounded by $1/p$.*

Proof. By definition, $C_{\max}(\mathsf{S}) = \max_{u\in\mathbb{Z}_{p^n}\setminus\{0\}, v\in\mathbb{Z}_{p^n}} C^{\mathsf{F}}_{v,u}$. Let S' be defined as in Theorem 4. Hence, it follows that

$$C_{\max}(\mathsf{S}) \geq \max_{u'\in\mathbb{Z}_p\setminus\{0\}, v'\in\mathbb{Z}_p} C^{\mathsf{S}}_{p^{n-1}\cdot v', p^{n-1}\cdot u'} = \max_{u'\in\mathbb{Z}_p\setminus\{0\}, v'\in\mathbb{Z}_p} C^{\mathsf{S}'}_{v',u'} \geq \frac{1}{p},$$

where the last inequality follows from the fact that S'_δ is a function over $\mathbb{F}_p$. □

We emphasize that *no analogous result* holds for a function defined over $\mathbb{F}_{p^n}$.

Differential Properties. A similar result holds for differentials.

Theorem 5. *Let $\mathsf{S}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ be a polyfunction. Given $\delta \in \mathbb{Z}_{p^m}$, let $\mathsf{S}'_\delta\colon \mathbb{Z}_{p^m} \to \mathbb{Z}_{p^m}$ be the function defined by $\mathsf{S}'_\delta(x) \coloneqq (\mathsf{S}(x+p^{n-m}\,\delta) - \mathsf{S}(x))/p^{n-m} \pmod{p^m}$. For all differences δ and Δ in $\mathbb{Z}_{p^m}$,*

$$\mathsf{DP}^{\mathsf{S}}_{p^{n-m}\Delta,\, p^{n-m}\delta} = \frac{|\{x \in \mathbb{Z}_{p^m} \mid S'_\delta(x) = \Delta\}|}{p^m}.$$

Proof. The function S'_δ is well-defined, because the coefficients of the polynomial $\mathsf{S}(x + p^{n-m}\delta) - \mathsf{S}(x)$ are all divisible by p^{n-m}. The number of solutions of the equation $S(x + p^{n-m}\delta) - \mathsf{S}(x) \equiv p^{n-m} \pmod{p^n}$ is equal to the number of solutions to $\mathsf{S}'_\delta(x) = \Delta$ in $\mathbb{Z}_{p^m}$ times p^{n-m}. This yields the result. □

Theorem 5 shows that polyfunctions have high-probability differentials for differences divisible by p. Indeed, the function S'_δ is a function on $\mathbb{Z}_{p^m}$, and typically satisfies $\mathsf{S}'(\delta) = \Delta$ with probability p^{-mt}, which is much higher than p^{-nt} (i.e., the best case for a function on $\mathbb{Z}_{p^n}$). As a consequence of this, the following lower bound holds *independently of n and of the details of* S.

Corollary 4. *Let $\mathsf{S}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ be a polyfunction. Then, the maximum differential probability of S – denoted by $DP_{\max}(\mathsf{S})$ – is lower bounded by $1/p$.*

The proof, similar to the proof of Corollary 3, is given in Appendix D. We emphasize that *no analogous result* holds for a function defined over $\mathbb{F}_{p^n}$.

Instead, in the case of differences that are not multiple of p, the differential probabilities can easily be shown to be small for low-degree S. For example, if S is of degree d, then $\mathsf{S}(x + \delta) \equiv \mathsf{S}(x) \pmod{p}$ has at most d solutions. If none of these solutions has multiplicity greater than one, then these solutions lift to unique solutions modulo p^n by Hensel's lemma.

Example 2. Consider $x \mapsto x^d$ in $\mathbb{Z}_{p^n}$, with d coprime to p and $d-1$ coprime to $p-1$. Let δ and Δ be differences, with δ not divisible by p. The equation $(x+\delta)^d - x^d = \Delta$ in $\mathbb{Z}_{p^n}$ has the same number of solutions as $(x+1)^d - x^d = \Delta/\delta^d$. The number of solutions modulo p is at most $d-1$. Let x be a solution. One can check that, under the stated assumptions, the derivative $d(x+1)^{d-1} - dx^{d-1}$ has no zeros over $\mathbb{Z}_{p^n}$. It follows that all solutions modulo p lift uniquely to solutions in $\mathbb{Z}_{p^n}$, so the probability of the differential (δ, Δ) is at most $(d-1)/p^n$.

Example 2 and the discussion above demonstrate that, for low-degree polynomials, *most* differentials have low probability. Similar arguments can be given for linear cryptanalysis, based on exponential sum estimates. Unfortunately, the linear layer preserves the property that every coordinate of a difference is divisible by p. Hence, at least some of the S-boxes must be non-polynomial to resist linear and differential cryptanalysis.

5 How to Classify Non-Polynomial Functions?

Not all non-polynomial functions are the same. How do we distinguish the "good" ones from the "bad" ones for cryptographic purposes? In this section, we introduce a list of properties that will help us to answer this question.

5.1 Distance From the "Best" Polynomial Approximations

We first consider parameters to measure how "close" a function is to a polyfunction, by considering the best possible polyfunction approximation.

ℓ-Maximum Polynomial Approximation. As one may expect, a first parameter is related to the possibility to approximate a function via a polynomial. For this reason, we introduce the following definitions.

Definition 6. *Let F (resp. $\bar{\mathsf{F}}$) be a function (resp. polyfunction), such that $\mathsf{F}, \bar{\mathsf{F}} \colon \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$. Given $1 \leq \ell \leq n$, the ℓ-Polynomial-Correlation of F via $\bar{\mathsf{F}}$ (in short, $\mathtt{PC}^{(\ell)}(\mathsf{F}, \bar{\mathsf{F}})$ or just $\mathtt{PC}^{(\ell)}$) is defined as*

$$\mathtt{PC}^{(\ell)}(\mathsf{F}, \bar{\mathsf{F}}) \coloneqq \frac{|\{x \in \mathbb{Z}_{p^n}^{t_1} \mid \mathsf{F}(x) = \bar{\mathsf{F}}(x) \mod p^\ell\}|}{p^{n \cdot t_1}}.$$

Definition 7. *Let $\mathsf{F} \colon \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$ be a function. The maximum ℓ-Polynomial-Correlation is defined as $\mathtt{PC}_{max}^{(\ell)} \coloneqq \max_{\bar{\mathsf{F}} \in \mathcal{F}} \mathtt{PC}^{(\ell)}(\mathsf{F}, \bar{\mathsf{F}})$, where $\mathcal{F}$ denotes the set of polyfunctions from $\mathbb{Z}_{p^n}^{t_1}$ to $\mathbb{Z}_{p^n}^{t_2}$. A polyfunction $\bar{\mathsf{F}}$ that satisfies $\mathtt{PC}_{max}^{(\ell)} = \mathtt{PC}^{(\ell)}(\mathsf{F}, \bar{\mathsf{F}})$ is said to be an ℓ-Optimal Polynomial Approximation (in short, ℓ-OPA) of F.*

If F is a polyfunction then $\mathtt{PC}_{\max}^{(\ell)} = 1$. We also remark that an ℓ-OPA is not necessarily unique, as shown in Example 3. Moreover, $\bar{\mathsf{F}}$ being ℓ-OPA for F does not imply $\bar{\mathsf{F}}$ being $(\ell-1)$-OPA for F, as shown again in Example 3.

Lemma 2. *For each $1 \leq i < j \leq n$, $\mathtt{PC}_{max}^{(i)} \geq \mathtt{PC}_{max}^{(j)}$.*

Proof. If the j-optimal polynomial approximation $\bar{\mathsf{F}}$ of F satisfies $\mathsf{F}(x) = \bar{\mathsf{F}}(x) \mod p^j$ for some x, then we also have $\mathsf{F}(x) = \bar{\mathsf{F}}(x) \mod p^{j-1}$. □

We highlight that there are functions for which the equality does not hold, as shown in the following example.

Example 3. Consider the invertible function G over $\mathbb{Z}_{2^4}$ defined as $\mathsf{G} = [2, 4, 7, 12, 3, 10, 11, 6, 5, 9, 15, 8, 0, 1, 14, 13]$. G is a non-polynomial function with $\mathtt{PC}^{(1)}_{\max}$ equal to 10/16 and $\mathtt{PC}^{(3)}_{\max}$ equal to 7/16. Both $x \mapsto 5x^3 + 6x^2 + x + 13$ and $x \mapsto \bar{\mathsf{G}}(x) = x^5 + 4x^3 + 6x^2 + x + 13$ over $\mathbb{Z}_{2^4}$ are 3-OPA. Moreover, $\bar{\mathsf{G}}$ is also a 2-OPA, but it is not a 1-OPA of G.

5.2 Congruence-Preservation

One of the more problematic aspects of polyfunctions is that two distinct elements x, y such that $x = y \mod p^\ell$ will map to the same point modulo p^ℓ. This effect is related to, but not entirely captured by polynomial correlation (Definition 6), as we do not differentiate between the points on which F and $\bar{\mathsf{F}}$ coincide. Instead, the following notions aim to capture this effect more precisely.

Congruence Equality. This concept focuses on the equalities along specific input-output pairs (a, b) modulo p^ℓ.

Definition 8 (Congruence Equality). *Consider a function $\mathsf{F} : \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$, and an integer $1 \le \ell \le n$. For each pair $(a, b) \in \mathbb{Z}_{p^\ell}^{t_1} \times \mathbb{Z}_{p^\ell}^{t_2}$, we define the* congruence equality *modulo p^ℓ of F with respect to the input a and the output b as*

$$\mathtt{CE}^{(\ell)}(a, b) \coloneqq |\{x \in \mathbb{Z}_{p^n}^{t_1} \mid x = a \mod p^\ell \text{ and } \mathsf{F}(x) = b \mod p^\ell\}|.$$

Note that we have $\mathtt{CE}^{(\ell)}(a, b) \le p^{t_1(n-\ell)}$ by counting the number of x satisfying $x = a \mod p^\ell$. For cryptographic designs, we are interested in functions that do not exhibit a particular bias modulo p^ℓ. That is, we would like these $p^{t_1(n-\ell)}$ inputs modulo p^ℓ to be distributed among the $p^{t_2\ell}$ outputs as evenly as possible. This motivates the following definition.

Definition 9 (Balanced Functions). *Let $\mathsf{F} : \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$ be a function. We say that F is* ℓ-balanced *for an integer $1 \le \ell \le n$ if*

$$\forall\ (a, b) \in \mathbb{Z}_{p^\ell}^{t_1} \times \mathbb{Z}_{p^\ell}^{t_2} : \qquad \mathtt{CE}^{(\ell)}(a, b) = p^{t_1 \cdot (n-\ell) - t_2 \ell}.$$

Proposition 7. *Suppose that the function $\mathsf{F}\colon \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$ is ℓ-balanced. Then it is also k-balanced for every integer $1 \le k < \ell$.*

Proof. It suffices to show that if F is ℓ-balanced for some $\ell \ge 2$, then it is also $(\ell - 1)$-balanced. To this end, given a pair $(a, b) \in \mathbb{Z}_{p^{\ell-1}}^{t_1} \times \mathbb{Z}_{p^{\ell-1}}^{t_2}$, note that it can be lifted to a pair $(a + x \cdot p^{\ell-1}, b + y \cdot p^{\ell-1}) \in \mathbb{Z}_{p^\ell}^{t_1} \times \mathbb{Z}_{p^\ell}^{t_2}$ modulo p^ℓ by choosing elements $x \in \mathbb{Z}_p^{t_1}$ and $y \in \mathbb{Z}_p^{t_2}$. Thus, the pair (a, b) is a descendant

from $p^{t_1+t_2}$ different pairs in $\mathbb{Z}_{p^\ell}^{t_1} \times \mathbb{Z}_{p^\ell}^{t_2}$, each of which will have a congruence-equality (modulo p^ℓ) equal to $p^{t_1 \cdot (n-\ell)-t_2\ell}$ since F is ℓ-modulo uniform. This means that we can count the congruence-equivalence of F as $p^{t_1+t_2} \cdot p^{t_1(n-\ell)-t_2\ell} = p^{t_1(n-(\ell-1))-t_2(\ell-1)}$, which shows that F is also $(\ell-1)$-balanced. □

Due to this result, we need only focus on the largest value ℓ for which a function F is ℓ-balanced. Note that F has $p^{t_1 n}$ possible inputs and there are $p^{n(t_1+t_2)}$ possible choices for (a, b). It follows that F can only be ℓ-balanced for some $\ell \leq t_1 \cdot n/(t_1+t_2)$. We will refer to functions where this upper bound holds with equality as *well-balanced*.

A link between $\mathtt{PC}_{\max}$ and $\mathtt{CE}^{(\ell)}$ is given in the following proposition.

Proposition 8. *Let $\mathsf{F} : \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$ be a function. Then:*

$$\mathtt{PC}_{max} \leq \frac{1}{p^{n \cdot t_1}} \cdot \min_{1 \leq \ell \leq n} \left(\sum_{a \in \mathbb{Z}_{p^\ell}^{t_1}} \max_{b \in \mathbb{Z}_{p^\ell}^{t_2}} \mathtt{CE}^{(\ell)}(a, b) \right) .$$

Proof. Due to Proposition 1, we know that if F is a polynomial function, then $\mathsf{F}(x) \bmod p^\ell = b \in \mathbb{Z}_{p^\ell}^{t_2}$ for each $x \bmod p^\ell = a \in \mathbb{Z}_{p^\ell}^{t_1}$.

Hence, given F, the maximum number of inputs x for which a polynomial P can be equal to F is given by the number of maximum inputs x for which F satisfies the condition just given. The claim follows immediately. □

For balanced functions, we have the following upper bound on $\mathtt{PC}_{\max}$, which follows immediately from Proposition 7 and from Proposition 8.

Lemma 3. *Given an ℓ-balanced function $\mathsf{F} : \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$, $\mathtt{PC}_{max} \leq p^{-t_2 \cdot \ell}$.*

Congruence-Preservation Probability modulo p^ℓ. The parameter $\mathtt{CE}^{(\ell)}(a, b)$ just introduced depends on the concrete values of the input and the output modulo p^ℓ. Here we introduce the "congruence-preservation probability", which only takes into account the (zero-)differences modulo p^ℓ in input/output.

Definition 10. *Let $\mathsf{F} : \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$ be a function, and let $1 \leq \ell < n$. Let $\mathcal{S}^{(\ell)} \subseteq (\mathbb{Z}_{p^n}^{t_1})^2$ be the set defined as $\mathcal{S}^{(\ell)} \coloneqq \{(x, y) \in (\mathbb{Z}_{p^n}^{t_1})^2 \mid x = y \mod p^\ell \ \wedge \ x \neq y\}$. We define the* congruence-preservation probability *modulo p^ℓ of F – denoted as $\mathtt{CPP}^{(\ell)}(\mathsf{F})$ – as follows*

$$\mathtt{CPP}^{(\ell)} \coloneqq \frac{|\{(x, y) \in \mathcal{S}^{(\ell)} \mid \mathsf{F}(x) = \mathsf{F}(y) \mod p^\ell\}|}{p^{n \cdot t_1} \cdot (p^{(n-\ell) \cdot t_1} - 1)} .$$

Note that $\mathtt{CPP}^{(\ell)} = 0$ implies $\mathtt{CPP}^{(j)} = 0$ for each $j \geq \ell$.

For ℓ-balanced functions, we obtain a concrete result on the value of their $\mathtt{CPP}^{(\ell)}$, which is independent of the specific function.

Proposition 9. *Suppose that the function* $\mathsf{F}\colon \mathbb{Z}_{p^n}^{t_1} \to \mathbb{Z}_{p^n}^{t_2}$ *is* ℓ*-balanced. Then,*

$$\mathtt{CPP}^{(\ell)} = \frac{p^{(n-\ell)\cdot t_1 - \ell \cdot t_2} - 1}{p^{(n-\ell)\cdot t_1} - 1}\,.$$

We refer to this value as the *optimal* CPP. Note that in general it is well approximated by $p^{-\ell \cdot t_2}$ except in special cases (e.g., $t_1 = t_2$ and F being $n/2$-balanced for even n implies $\mathtt{CPP}^{(n/2)} = 0$). The proof of Proposition 9 is given in Appendix E and is based on the relation between $\mathtt{CPP}^{(\ell)}$ and $\mathtt{CE}^{(\ell)}$.

Having said that, it remains an open problem for future work to find a relation (if it exists) between $\mathtt{PC}_{\max}^{(\ell)}$ and $\mathtt{CPP}^{(\ell)}$. In particular, the next example shows that both $\mathtt{PC}_{\max}^{(\ell)} < \mathtt{CPP}^{(\ell)}$ and $\mathtt{PC}_{\max}^{(\ell)} > \mathtt{CPP}^{(\ell)}$ can occur, depending on the considered function. We only note that if $\mathtt{PC}_{\max}^{(\ell)} = 1$, then also $\mathtt{CPP}^{(\ell)} = 1$.

Example 4. Consider the invertible functions F and G over $\mathbb{Z}_{2^3}$ defined as $\mathsf{F} = [1, 2, 4, 3, 5, 6, 0, 7]$ and $\mathsf{G} = [0, 1, 4, 6, 3, 5, 2, 7]$. For F, $\mathtt{CPP}^{(2)}$ is 1 and $\mathtt{PC}_{\max}^{(2)}$ equals to $1/2$ with a 2-OPA given by $x \mapsto x + 1$ over $\mathbb{Z}_{2^3}$. For G, $\mathtt{CPP}^{(2)}$ equals to $1/4$ and $\mathtt{PC}_{\max}^{(2)}$ is equal to $5/8$ with identity over $\mathbb{Z}_{2^3}$ as its 2-OPA.

6 Non-Polynomial Ciphers: Attacks via Polynomial Approximations

In Section 4 we focused on algebraic attacks against ciphers over $\mathbb{Z}_{p^n}^t$ that admit a polynomial representation. In this section, we consider ciphers of class II and III (see Section 3.1) that cannot be fully described by polynomials over $\mathbb{Z}_{p^n}^t$. The attack strategy we follow is instead to find good polynomial approximations of the given cipher and exploit them. We quantify the quality of these approximations using the notion of polynomial correlation, PC, introduced in Section 5.

The idea behind this attack strategy is not completely new in the literature. For example, in linear cryptanalysis [67], the goal of the attacker is to find the best *linear* approximation of a given cipher, in order to break it. A generalization is found in *nonlinear* cryptanalysis [58], in which the linear approximation is replaced by a low-degree approximation.

Remark 3. We restrict ourselves to polynomial approximations over $\mathbb{Z}_{p^\ell}$ in this section. Still, one may wonder whether there are other mathematical structures that are more suitable to work with. A particularly interesting candidate may be the field of *p-adic numbers*, $\mathbb{Q}_p$. One reason is the fact that *every* function over $\mathbb{Z}_{p^n}$ can still be represented as a polynomial in $\mathbb{Q}_p[x]$. We discuss approximations over other fields and why they are *unlikely* to work in Appendix G.

6.1 Optimal Polynomial Approximations: Connections to Coding Theory and Impacts on the Interpolation Attack

Let the "Optimal Polynomial Approximation" (OPA) of a function F over $\mathbb{Z}_{p^n}$ be defined as in Definition 7. When p^n is small, such an OPA can easily be found

by brute forcing all possible $\prod_{i=0}^{\mu(p^n)-1} \frac{p^n}{\gcd(i!, p^n)}$ polyfunctions. Finding OPAs (or even just decent polynomial approximations) when p^n is large appears to be a highly non-trivial problem in general. As only the S-box is non-polynomial in our construction, here we restrict ourselves to functions $\mathsf{F} : \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$, but remark that the results could be generalized to also hold for the $\mathbb{Z}_{p^n}^t$ case.

Connection to Coding Theory. Let $\mathcal{F}$ denote the set of polyfunctions on $\mathbb{Z}_{p^n}$. Then the code $\mathcal{C}_{p^n}$, defined by

$$\mathcal{C}_{p^n} = \{(\mathsf{P}(0), \mathsf{P}(1), \ldots, \mathsf{P}(p^n - 1)) \mid \mathsf{P}(x) \in \mathcal{F}\} \subseteq \mathbb{Z}_{p^n}^{p^n},$$

is the polynomial evaluation code where polynomials are evaluated at *all* points of the ring. For a function P over $\mathbb{Z}_{p^n}$, we write $\mathbf{c}_\mathsf{P} = (\mathsf{P}(i))_{i \in \mathbb{Z}_{p^n}}$ and have $\mathbf{c}_\mathsf{P} \in \mathcal{C}_{p^n}$ if and only if $\mathsf{P} \in \mathcal{F}$. Note that the definition of $\mathcal{C}_{p^n}$ is similar to that of the (classic) Reed-Solomon (RS) code, with the important distinction that we work over a ring instead of a field.

Observe that $\mathcal{C}_{p^n}$ is both linear and cyclic. Indeed, consider a polynomial P and its codeword $\mathbf{c}_\mathsf{P}$. Then $\hat{\mathsf{P}}(x) = \mathsf{P}(x-1)$ gives rise to its cyclic shift $\mathbf{c}_{\hat{\mathsf{P}}}$. Hence, the codewords can be identified with an ideal of the ring $\mathbb{Z}_{p^n}[x]/(x^{\mu(p^n)} - 1)$. This ideal may, however, have a rather complicated description as $\mathbb{Z}_{p^n}[x]$ is in general not a principal ideal ring.

The problem of finding an OPA for a general function F over $\mathbb{Z}_{p^n}$ is now equivalent to finding the closest codeword $\mathbf{c}_{\tilde{\mathsf{F}}} \in \mathcal{C}_{p^n}$ to the word $(\mathsf{F}(0), \ldots \mathsf{F}(p^n - 1)) \in \mathbb{Z}_{p^n}^{p^n}$. In other words, $\mathbf{c}_{\tilde{\mathsf{F}}}$ is the codeword defined by the polynomial $\tilde{\mathsf{F}}$ that agrees with F in the highest number of points. Knowledge of the codeword makes recovering the polynomial $\tilde{\mathsf{F}}(x) \in \mathcal{F}$ corresponding to $\mathbf{c}_{\tilde{\mathsf{F}}}$ easy. We remark that there may be many *closest* polynomials, and list decoders may be relevant.

Remark 4. RS codes can also be generalized over rings (see e.g. [69]), but in these definitions the evaluation is only performed over a small subset of $\mathbb{Z}_{p^n}$. The algorithms of e.g. [69] are not applicable for $\mathcal{C}_{p^n}$, since most standard algorithms require that the evaluation points constitute a so-called *subtractive set* (i.e., the difference of any two evaluation points must be a unit in the ring).

Parameters of $\mathcal{C}_{p^n}$. We now briefly examine some of the code parameters of $\mathcal{C}_{p^n}$. Recall that the *minimum distance* of a code $\mathcal{C}$ is $d_{\min} = \min_{\mathbf{c} \neq \hat{\mathbf{c}} \in \mathcal{C}} \mathrm{wt}(\mathbf{c} - \hat{\mathbf{c}})$, where wt denotes the Hamming weight. By construction of $\mathcal{C}_{p^n}$, we note that the minimum distance is $d_{\min} = p^n(1 - \max_{\tilde{\mathsf{F}} \in \mathcal{F}} \mathsf{PC}^*_{\max})$, where $\mathsf{PC}^*_{\max}$ is defined as in Definition 7, but using the punctured set $\mathcal{F} \setminus \{\tilde{\mathsf{F}}\}$. Using that $\mathcal{C}_{p^n}$ is a linear code, we establish its minimum distance in the following result, which is proved in Appendix F.1. To state it, we define the *base-p digit sum* of a positive integer a as $\sigma_p(a) \coloneqq \sum_{i \geq 0} a_i$, where $a = \sum_{i \geq 0} a_i \cdot p^i$ for $0 \leq a_i < p$.

Proposition 10. *Given a prime p, an integer $n > 1$ and let $m = \max\{\sigma_p(k) \mid k < \mu(p^n)\}$. Write $m = a(p-1) + b$ with $0 \leq b < p - 1$. Then the minimum distance of $\mathcal{C}_{p^n}$ is $(p-b)p^{n-a-1}$. In particular, for any two distinct polyfunctions F, G over $\mathbb{Z}_{p^n}$ we have $\mathsf{PC}(\mathsf{F}, \mathsf{G}) \leq \frac{p^n - (p-b)p^{n-a-1}}{p^n} = 1 - \frac{p-b}{p^{a+1}}$.*

We remark that there are words in the space $\mathbb{Z}_{p^n}^{p^n}$ that are very far away from any codeword. One example is the class of well-balanced functions.

Corollary 5 (of Lemma 3). *Given a polyfunction* F *on* $\mathbb{Z}_{p^n}$ *and an* ℓ*-balanced function* G *for* $\ell \leq \lfloor n/2 \rfloor$*. Then* $d(\mathbf{c}_\mathsf{F}, \mathbf{c}_\mathsf{G}) \geq p^n - p^{n-\ell}$.

Corollary 5 gives a bound on the *covering radius* of $\mathcal{C}_{p^n}$, i.e., the maximal distance between some word in $\mathbb{Z}_{p^n}^{p^n}$ and its nearest codeword in $\mathcal{C}_{p^n}$. We give an overview of the code parameters for $p = 2$ for various n in Table 2 – Appendix F.2.

A Decoding Algorithm for $\mathcal{C}_{p^n}$. We outline a decoding algorithm correcting a fixed number of s errors. Let $\tilde{\mathsf{F}}$ be a polyfunction, and let $\mathsf{F} = \tilde{\mathsf{F}} + \mathsf{E}$ be the function for which we are looking for an OPA, where E is any function on $\mathbb{Z}_{p^n}$ with support size s (equivalently, E is different from zero in $\leq s$ points). By Proposition 2, if there is a *window* of $\mu(p^n)$ consecutive points on which $\tilde{\mathsf{F}}$ and F agree then $\tilde{\mathsf{F}}$ is uniquely determined by the outputs of F on these points. The simplest way of finding such a *window* is by considering the cyclic shifts of F and interpolating the initial $\mu(p^n)$ points. A pseudo-code of the corresponding algorithm is given in Appendix F.3, along with further discussion and an analysis of its complexity. In particular, we show that the algorithm will always terminate if $\mathrm{wt}(\mathbf{c}_\mathsf{E}) < \frac{p^n - \mu(p^n)+1}{\mu(p^n)}$.

Practical Use Cases. Since our algorithm requires a relatively high polynomial correlation to guarantee decoding, it is not suitable for approximating completely random or well-balanced functions. At the same time, the S-box of a given cipher may have been designed to appear random despite having a precise algebraic structure, as in the case of Streebog and Kuznyechik [68]. Our algorithm may allow for recovery of the original function in such cases.

Interpolation Attack via Polynomial Approximation. A possibility to exploit a polynomial approximation of a cipher would be to set up an interpolation attack [54]. If the cipher admits a polynomial representation modulo p^ℓ for any $\ell \in \{1, 2, \ldots, n\}$, the attacker guesses the last round key(s), (partially) decrypts the given known ciphertexts, and constructs an interpolation polynomial modulo p^ℓ connecting the given plaintexts and the (partially) decrypted ciphertexts (the variable in the polynomial is the plaintext). Finally, an extra pair of plaintext/ciphertext is used to check if the guessed last round key is correct.

Instead, if the cipher does not admit a polynomial representation modulo p^ℓ for any ℓ, and based on the considerations just given, it seems hard for the attacker to find a "sufficiently good" polynomial approximation for the (almost) entire cipher in an efficient way. Hence, we conjecture that the interpolation attack is in general not efficient in the case of a cipher of class II/III.

6.2 Solving the Associated System of Equations obtained via the Polynomial Approximation

Rather than finding a good polynomial approximation for the entire cipher, the attacker can focus only on the non-polynomial parts of the cipher, that is, the

S-boxes. Since these are usually defined over a smaller space than the entire cipher, it is generally easier to find a good polynomial approximation for them. By doing this, we derive a round-based system of equations that well approximate the cipher to break, in which the unknown variables are the round keys.

Attack Strategy. A first attempt may be for the attacker to approximate the cipher modulo p^n and attempt to recover the secret key using multiple plaintext/ciphertext pairs (for completeness, details are provided in Appendix H).

However, a more efficient method is to work with increasing moduli, essentially resulting in a probabilistic variant of the method in Section 4.1. An extra trick is that we can use information at smaller moduli to filter plaintext/ciphertext pairs (step 3 below), recalling that a polynomial approximation that fails modulo $p^{\ell-1}$ will also fail modulo p^ℓ. Let $\mathfrak{P}$ denote a (pre-sampled) set of known plaintext/ciphertext pairs. Then for increasing $1 \leq \ell \leq n$:

1. the attacker first finds a polynomial approximation of the non-polynomial functions modulo p^ℓ (which is independent of the secret key);
2. working round by round, the attacker sets up the system of equations modulo p^ℓ of the approximated cipher, by replacing the non-polynomial functions with their corresponding polynomial approximations;
3. If $\ell > 1$, the attacker filters the pairs in $\mathfrak{P}$, keeping only the pairs where (the found) $k \bmod p^{\ell-1}$ is a valid solution of the associated system of equations modulo $p^{\ell-1}$;
4. exploiting the knowledge of $k \bmod p^{\ell-1}$ for $\ell > 1$, the attacker solves the system of equations for different, filtered pairs from $\mathfrak{P}$. The solution that appears the most often among these systems is our candidate for $k \bmod p^\ell$;
5. When $k \bmod p^\ell$ is found (with suitably high probability), the process is repeated for $\ell \leftarrow \ell + 1$.

Note that if the polynomial approximations at step ℓ reduce to their respective approximations modulo $\ell - 1$, then we only need to do the filtering in step 3 over the previously filtered set. If the approximations differ, we filter over $\mathfrak{P}$.

Estimating the Probability of the Polynomial Approximation. A crucial ingredient for analyzing the complexity of this attack approach consists in the probability of setting up a polynomial approximation for an iterated cipher over $\mathbb{Z}_{p^n}^t$. Following the structure of the cipher from Section 3, let $\mathsf{P}_{i,j}$ over $\mathbb{Z}_{p^n}$ be an optimal polynomial approximation of the S-box $\mathsf{S}_{i,j}$ modulo p^ℓ. By replacing all S-boxes $\mathsf{S}_{i,j}$ with $\mathsf{P}_{i,j}$, we obtain a polynomial approximation A of the cipher C. Assuming that C is a Markov cipher, the rounds of C are independent, and given r the number of rounds, the probability ρ_ℓ that the cipher and the polynomial approximation A coincide modulo p^ℓ is given by

$$\rho_\ell \coloneqq \frac{1}{p^{n\cdot\kappa}} \sum_{k \in \mathbb{Z}_{p^n}^\kappa} \mathsf{PC}^{(\ell)}(\mathsf{C}_k(\cdot), \mathsf{A}(\cdot, k)) \geq \frac{1}{p^{n\cdot\kappa}} \sum_{k \in \mathbb{Z}_{p^n}^\kappa} \prod_{i=0}^{r-1} \mathsf{PC}^{(\ell)}(\mathsf{R}_{k_i}, \tilde{\mathsf{R}}_{k_i})$$
$$\geq \frac{1}{p^{n\cdot\kappa}} \sum_{k \in \mathbb{Z}_{p^n}^\kappa} \prod_{i=0}^{r-1} \prod_{j=0}^{t-1} \mathsf{PC}^{(\ell)}(\mathsf{S}_{i,j}, \mathsf{P}_{i,j}) = \prod_{i=0}^{r-1} \prod_{j=0}^{t-1} \mathsf{PC}^{(\ell)}(\mathsf{S}_{i,j}, \mathsf{P}_{i,j}) \,,$$

for R_{k_i} the i-th round of C, $\tilde{\mathsf{R}}_{k_i}$ the i-th round of A, and the i-th round key k_i. The last inequality is obtained by the fact that for any k_i, the output of R_{k_i} and $\tilde{\mathsf{R}}_{k_i}$ coincides if all their S-boxes do, hence with probability at least $\prod_{j=0}^{t-1} \mathsf{PC}^{(\ell)}(\mathsf{S}_{i,j}, \mathsf{P}_{i,j})$.

Lower Bound of ρ_ℓ. If $\mathsf{S}_{i,j} = \mathsf{S}$ fixed with its ℓ-OPA P, then the lower bound of ρ_ℓ – denoted by $\rho_\ell^{\min}$ – reduces to

$$\rho_\ell^{\min} := \prod_{i=0}^{r-1} \prod_{j=0}^{t-1} \mathsf{PC}^{(\ell)}(\mathsf{S}_{i,j}, \mathsf{P}_{i,j}) = \left(\mathsf{PC}^{(\ell)}(\mathsf{S}, \mathsf{P})\right)^{r \cdot t}.$$

We emphasize that $\rho_\ell^{\min}$ is a lower bound in general. Indeed, it is possible that $\mathsf{C}_k(x) = \mathsf{A}(x, k)$ even if $\mathsf{S}_{i,j}(x') \neq \mathsf{P}_{i,j}(x')$ for some i, j and an intermediate S-box input x'. Equivalently, it is possible that A is a good approximation of C even for the inputs x for which the local polynomial approximation $\mathsf{S}_{i,j}(x') = \mathsf{P}_{i,j}(x')$ does not hold. This is due to the fact that the combination of two non-polyfunctions can be equal to a poly function, which is illustrated in Example 5.

Example 5. Consider the invertible function F over $\mathbb{Z}_{2^4}$ defined as $\mathsf{F} = [0, 3, 4, 9, 2, 15, 6, 5, 8, 11, 10, 1, 12, 7, 14, 13]$. It is equal to the function $x \mapsto \mathsf{G}(x) = 2 \cdot x^3 + x$ except for $x \in \{2, 4\}$ (indeed, $\mathsf{G}(2) = \mathsf{F}(4) = 2$ and $\mathsf{G}(4) = \mathsf{F}(2) = 4$). As we saw before, F is not a polyfunction. Still, $\mathsf{F} \circ \mathsf{F}(x) = \mathsf{G} \circ \mathsf{G}(x) = 8 \cdot x^7 + 12 \cdot x^5 + 4 \cdot x^3 + x$ is a polyfunction over $\mathbb{Z}_{2^4}$.

6.3 Complexity of the Attack from Section 6.2

In the case of the attack strategy from Section 6.2, as the attacker cannot test the key candidates mod p^ℓ directly, it is crucial to determine how many pairs of plaintext/ciphertext are needed to detect the correct secret key mod p^ℓ for each $1 \leq \ell \leq n$. Moreover, for simplicity, we assume $\kappa = t$.

Probabilities of the Right versus the Wrong Key Candidate. As we just saw, at each step $\ell \geq 1$, the attacker aims to determine the correct extension $k_{\ell-1} \in \mathbb{Z}_p$ of the secret key mod $p^{\ell-1}$ to mod p^ℓ using the knowledge of $k_0, \ldots, k_{\ell-2} \in \mathbb{Z}_p$. In order to do this, we assume that each such system mod p^ℓ has only one solution $(\hat{k}_0, \ldots, \hat{k}_{\ell-1})$.[2]

Starting with $\ell = 1$, for a random plaintext/ciphertext pair $(x, y = \mathsf{C}_k(x))$, the attacker solves a system of polynomial equations $\mathsf{A}_1(x, \cdot) = y \mod p$ to determine the correct k_0. It is simple to observe that the right key extension k_0 satisfies the system with probability ω_1 given by $\rho_1 \geq \rho_1^{\min} =: \omega_1^{\min}$. Any other key satisfies the system with probability ζ_1 given by

$$\rho_1 \cdot 0 + (1 - \rho_1) \cdot \frac{1}{(p^t - 1)} = (1 - \rho_1) \cdot \frac{1}{(p^t - 1)} \leq (1 - \rho_1^{\min}) \cdot \frac{1}{(p^t - 1)} =: \zeta_1^{\max}.$$

[2] We point out that this is a realistic assumption, as for the real cipher, we should obtain a random system of polynomial equations. Still, more solutions are possible. If this is the case, the attacker can simply brute-force the candidates.

For $\ell \geq 2$, using the filtering of plaintexts, we know that the remaining plaintexts satisfy $\mathsf{A}_\ell(x, k_0, \ldots, k_{\ell-2}) = y \mod p^{\ell-1}$. The attacker now uses the knowledge of $k_0, \ldots, k_{\ell-2}$ and solves a system of polynomial equations given by $\mathsf{A}_\ell(x, k_0, \ldots, k_{\ell-2}, \cdot) = y \mod p^\ell$:

- the right key extension $k_{\ell-1}$ satisfies the system with probability ω_ℓ where

$$\begin{aligned}\omega_\ell &= \mathrm{Prob}(\mathsf{A}_\ell(x, k_0, \cdots, k_{\ell-1}) = y \mod p^\ell \mid \mathsf{A}_\ell(x, k_0, \cdots, k_{\ell-2}) = y \mod p^{\ell-1}) \\ &= \frac{\rho_\ell}{\rho_{\ell-1}} \geq \rho_\ell^{\min} =: \omega_\ell^{\min},\end{aligned}$$

  which has been computed using Bayes' theorem;
- any other key satisfies the system with probability ζ_ℓ given by

$$\zeta_\ell = (1 - \omega_\ell) \cdot \frac{1}{(p^t - 1)} \leq (1 - \rho_\ell^{\min}) \cdot \frac{1}{(p^t - 1)} =: \zeta_\ell^{\max}.$$

Data Complexity. Not surprisingly, the distribution that the right key satisfies the given system of polynomial equations for N_ℓ random plaintext/ciphertext pairs can be reasonable approximated by a normal distribution $\mathcal{N}(N_\ell \cdot \omega_\ell, N_\ell \cdot \omega_\ell \cdot (1 - \omega_\ell))$. Following the same argument just given, the probability distribution for the wrong key is approximated by a normal distribution with parameters $\mathcal{N}(N_\ell \cdot \zeta_\ell, N_\ell \cdot \zeta_\ell \cdot (1 - \zeta_\ell))$.

For each $1 \leq \ell \leq n$, the attacker solves N_ℓ polynomial systems in order to identify the correct key mod p^ℓ. Following [67,73], an approximation of the number of plantext/ciphertext pairs N_ℓ necessary to set up the attack is given by

$$N_\ell \approx \frac{\omega_\ell}{|\omega_\ell - \zeta_\ell|^2} \cdot \left(\Phi^{-1}(\Omega_\ell^{(S)}) + \Phi^{-1}(1 - 2^{-\alpha_\ell - 1})\right)^2,$$

where $\Omega_\ell^{(S)}$ is the success probability of the attack, and where α_ℓ is the bit-advantage of the attack introduced in [73], which corresponds to the probability for the attacker to identify a set of $p^{t-\alpha_\ell}$ possible extensions of the correct key k from modulus $p^{\ell-1}$ to p^ℓ that contains the correct extension $k_{\ell-1}$ of k mod $p^{\ell-1}$.

Note that if $\omega_\ell \gg \zeta_\ell$, then $|\omega_\ell - \zeta_\ell|^2 \approx \omega_\ell^2$, which implies

$$N_\ell \geq \frac{1}{\omega_\ell} \cdot \left(\Phi^{-1}(\Omega_\ell^{(S)}) + \Phi^{-1}(1 - 2^{-\alpha_\ell - 1})\right)^2,$$

that is, N_ℓ is proportional to the probability ω_ℓ, as one may expect.

For simplicity, let us assume that the attacker uses the same polynomial approximation for all $1 \leq \ell \leq n$. Note that in round ℓ, the ratio of plaintexts that pass Step 3 of the attack strategy from Section 6.2, i.e., filtering the remaining data according to the last found key-extension, is ω_ℓ. Hence, assuming a success probability $\prod_{\ell=1}^n \Omega_\ell^{(S)}$, the data complexity of the attack is approximated by $N \approx \max(N_1, N_2/\omega_1, N_3/\omega_2, \ldots, N_n/\omega_{n-1})$. This ensures that even after filtering, there remain enough plaintexts so it is possible to distinguish the correct key-extension for each ℓ, given the success probabilities $\Omega_\ell^{(S)}$. This data complexity can be easily generalized to the case when the attacker selects a different polynomial approximation A_ℓ for some ℓ.

Remark 5. We limit ourselves to mention that the data complexity is computed in practice with $\rho_l^{\min}$ instead of ρ_l. While $\rho_\ell^{\min}$ provides a good approximation for the data complexity, the real complexity may be lower than what we expect.

7 Non-Polynomial Ciphers: Truncated Differential Attacks

A possible drawback of the attacks presented in the previous section is the fact that the attacker has to find a *good* polynomial approximation of the given function. As we saw, this is not possible in many cases. In this section, we propose another attack strategy for ciphers over $\mathbb{Z}_{p^n}^t$ based on Proposition 1, which does not require the knowledge of such a polynomial approximation.

7.1 Truncated Differential Attacks modulo p^ℓ: Overall Strategy

The attack we are going to discuss can be classified as a truncated differential one. Truncated differential attacks [57] are a generalization of differential cryptanalysis, in which the attacker only partially knows the difference between two elements. In most of the attacks over $\mathbb{F}_{p^n}^t$ (e.g., truncated differential against AES), the attacker only knows if a particular $\mathbb{F}_{p^n}$-word is active or not (equivalently, if the difference in a particular $\mathbb{F}_{p^n}$-word is zero or not). We generalize this approach by considering zero-difference modulo p^ℓ.

Polyfunction Components. Consider an iterated cipher $\mathsf{C}_k(\cdot)$ over $\mathbb{Z}_{p^n}^t$ in which every component admits a polynomial representation. Given $\ell \in \{0, 1, \ldots, n-1\}$, choose two plaintexts x, y such that $x \bmod p^\ell = y \bmod p^\ell$, or equivalently, $(x-y) \bmod p^\ell = 0$:

- the round-key addition does not change this zero difference, as $x \bmod p^\ell = y \bmod p^\ell$ holds if and only if $(x+k) \bmod p^\ell = (y+k) \bmod p^\ell$ for any $k \in \mathbb{Z}_{p^n}^t$;
- due to Proposition 1, this zero-difference is invariant under any polyfunction F, that is, $x \bmod p^\ell = y \bmod p^\ell$ implies $\mathsf{F}(x) \bmod p^\ell = \mathsf{F}(y) \bmod p^\ell$.

As a consequence, given the corresponding ciphertexts $\mathsf{C}_k(x)$ and $\mathsf{C}_k(y)$, we have that $x \bmod p^\ell = y \bmod p^\ell$ if and only if $\mathsf{C}(x) \bmod p^\ell = \mathsf{C}(y) \bmod p^\ell$. This is a powerful secret-key distinguisher that is independent of the secret key.

Non-Polyfunction Components. As next step, assume that some components S (for simplicity, over $\mathbb{Z}_{p^n}$) of the cipher do not admit a polynomial representation. In such a case, given x, y as before (hence, $x \bmod p^\ell = y \bmod p^\ell$) and $x \neq y$, we have that

$$\mathrm{Prob}(\mathsf{S}(x) \bmod p^\ell = \mathsf{S}(y) \bmod p^\ell \mid x = y \mod p^\ell) = \mathtt{CPP}^{(\ell)}(\mathsf{S}).$$

In the case where $\mathtt{CPP}^{(\ell)}(\mathsf{S}) \neq 1/p^\ell$, the attacker can set up a distinguisher and a key-recovery attack by analyzing how the zero-difference modulo p^ℓ propagates through the cipher. As before, the crucial point to keep in mind is that the key

does not affect the zero difference (as $x \bmod p^\ell = y \bmod p^\ell$ holds if and only if $(x+k) \bmod p^\ell = (y+k) \bmod p^\ell$ for any $k \in \mathbb{Z}_{p^n}^t$), which implies the possibility to compute the probability over the entire cipher by considering the probability that the zero-difference modulo p^ℓ is preserved by all its components.

Remark 6. Compared to the attacks from Section 6.2, where the attacker has to find polynomial approximations of the non-polynomial parts and compute their corresponding $\mathtt{PC}^{(\ell)}$, the truncated differential attack requires to compute the $\mathtt{CPP}^{(\ell)}$ of the non-polynomial S-boxes. This computation does not depend on the specific polynomial approximations found for the non-polynomial components of the cipher, and so it may be easier to compute.

Before going on, we emphasize that the truncated differential modulo p^ℓ just described over $\mathbb{Z}_{p^n}^t$ does not have any equivalent for ciphers over $\mathbb{F}_{p^n}^t$.

7.2 Complexity of Truncated Differential Attacks modulo p^ℓ

In order to estimate the complexity of a truncated differential modulo p^ℓ as just defined, we first compute the probability of a truncated differential modulo p^ℓ of an iterated r-round cipher over $\mathbb{Z}_{p^n}^t$ as defined in Section 3. Here, in order to simplify the discussion, we assume that all the round functions are identical (a part from the value of the round-key), which corresponds to the majority of the symmetric ciphers in the literature.

Proposition 11. *Let C_k be an iterated r-round cipher over $\mathbb{Z}_{p^n}^t$, instantiated with a fixed invertible round function R_{k_i} for a round-key $k_i \in \mathbb{Z}_{p^n}^\kappa$.*

Let $1 \leq \ell < n$. Assume that

$$\mathrm{Prob}\left(\mathsf{R}_{k_i}(x) = \mathsf{R}_{k_i}(y) \mod p^\ell \mid x = y \mod p^\ell \wedge x \neq y\right) = \frac{p^{(n-\ell)\cdot t} - 1}{p^{n\cdot t} - 1} + \varepsilon,$$

for $|\varepsilon| \geq \frac{1}{p^{n\cdot t}}$ and $0 < \frac{p^{(n-\ell)\cdot t}-1}{p^{n\cdot t}-1} + \varepsilon < 1$, with x, y random. Under the usual assumptions that C_k is a Markov cipher and the rounds are independent, the truncated differential modulo p^ℓ over $r \geq 1$ rounds has probability

$$\frac{1}{p^{\ell\cdot t}} + \left(\frac{p^{\ell\cdot t}}{p^{\ell\cdot t} - 1}\right)^{r-1} \cdot \varepsilon^r + \mathcal{O}\left(\frac{1}{p^{n\cdot t}}\right). \tag{4}$$

The proof is given in Appendix I, and it exploits the fact that $\mathsf{R}_{k_i}(x) - \mathsf{R}_{k_i}(y) = (k_i + \mathsf{R}(x)) - (k_i + \mathsf{R}(y)) = \mathsf{R}(x) - \mathsf{R}(y)$, so this difference is independent of the value of the key. As expected, the probability (4) tends to $1/p^{t\cdot\ell}$ for $r \gg 1$. In the case $p = 2$ and $\ell = t = 1$, the previous probability corresponds to $\frac{1}{2} + 2^{r-1} \cdot \varepsilon^r$. Interestingly, this is exactly the result of the *piling-up lemma* [67,52], crucial in linear cryptanalysis for estimating the probability that a sum of binary variables is equal to zero.

Based on previous assumptions, in the following we denote by

$$\mathtt{CPP}^{(\ell)}(\mathsf{C}) \coloneqq \frac{1}{p^{n\cdot\kappa}} \sum_{k \in \mathbb{Z}_{p^n}^\kappa} \mathtt{CPP}^{(\ell)}(\mathsf{C}_k) = \frac{1}{p^{\ell\cdot t}} + \left(\frac{p^{\ell\cdot t}}{p^{\ell\cdot t} - 1}\right)^{r-1} \cdot \varepsilon^r + \mathcal{O}\left(\frac{1}{p^{n\cdot t}}\right).$$

Turning the Secret-Key Distinguisher into a Key-Recovery Attack. As usual in the literature, this secret-key distinguisher can be turned into a key-recovery attack, under the assumption of the *wrong-key randomization hypothesis* [16,2]. To be precise, we exploit the probability that, given two *different* inputs that are equal modulo p^ℓ, the outputs of the random oracle are equal modulo p^ℓ. For an ideal permutation, that this event occurs with probability $\frac{p^{(n-\ell)\cdot t}-1}{p^{n\cdot t}-1} = \frac{1}{p^{\ell\cdot t}} + \mathcal{O}\left(\frac{1}{p^{n\cdot t}}\right)$.

Working in the same way as in Section 6.3, the distribution of the number of plaintext/ciphertext couples that satisfy the truncated differential for the right key can be approximated by the normal distribution $\mathcal{N}(N_C \cdot \mathtt{CPP}^{(\ell)}(\mathsf{C}), N_C \cdot \mathtt{CPP}^{(\ell)}(\mathsf{C}) \cdot (1 - \mathtt{CPP}^{(\ell)}(\mathsf{C})))$ (note that each couple is composed of two plaintext/ciphertext pairs). Working as usual in the literature, an approximation of the number of couples N_C necessary to set up the attack is given by

$$N_C \approx \frac{\mathtt{CPP}^{(\ell)}(\mathsf{C})}{\left|\mathtt{CPP}^{(\ell)}(\mathsf{C}) - p^{-\ell\cdot t}\right|^2} \cdot \left(\Phi^{-1}(\Omega_\ell^{(S)}) + \Phi^{-1}(1 - 2^{-\alpha_\ell - 1})\right)^2,$$

that is, approximately $\sqrt{2 \cdot N_C}$ plaintext/ciphertext pairs (as $\binom{\sqrt{2\cdot N_C}}{2} = N_C$).

7.3 Truncated Differential against AES-like Ciphers over $\mathbb{Z}_{p^n}^{4\times 4}$

The previous data estimation does not take into account the possibility of the attacker to choose the plaintexts. Still, choosing the plaintexts could have a crucial impact on the possibility to set up the attack. We show this concretely by discussing truncated differential attacks against an AES-like cipher over $\mathbb{Z}_{p^n}^{4\times 4}$, instantiated with the same linear layer defined as the composition of a ShiftRows and a MixColumns operation with a 4×4 MDS matrix. We assume that *the S-boxes over $\mathbb{Z}_{p^n}^{4\times 4}$ do **not** admit a polynomial representation, but they preserve the equivalence modulo p^ℓ for some $\ell \in \{1, 2, \ldots, n-1\}$*. We recall that such S-boxes exist, as for the case of F over $\mathbb{Z}_{2^3}$ in Example 4 (which preserves the equivalence modulo 4 as $\mathtt{CPP}^{(2)} = 1$, and does not admit any polynomial representation).

In such a case, attacking such an AES-like cipher with the truncated differentials modulo p^ℓ is the best possible approach, as (i) the attacks proposed in Section 4 would not work, since such S-boxes do not admit a polynomial representation, and (ii) the attacks presented in Section 6 work but are not optimal, as they are not deterministic. Exploiting the fact that equivalence modulo p^ℓ is preserved is the crucial characteristic of the truncated differentials described here.

Results from AES over $\mathbb{F}_{2^8}^{4\times 4}$. In this scenario, it is trivial to observe that the well-known (impossible) truncated differentials [31,47] for 1, 2, 3 and 4 rounds AES over $\mathbb{F}_{2^8}^{4\times 4}$ simply apply to our 1, 2, 3 and 4 rounds AES-like cipher over $\mathbb{Z}_{p^n}^{4\times 4}$. Indeed, consider the case of one active $\mathbb{F}_p^n$-word, and all other words equal modulo p^ℓ. In such a case, the differences modulo p^ℓ are preserved both by the S-box by assumption, and by the matrix multiplication (as it is a linear

operation, and so admits a polynomial representation). Moreover, as the attacker works modulo $p^\ell \leq p^n$, the data complexity is expected to be smaller than the corresponding attack against round-reduced AES over $\mathbb{F}_{p^n}^{4\times 4}$ (for which the attacker cannot work modulo p^ℓ).

In a similar way, it is not hard to check that other properties such as the "multiple-of-8" one [48] and the mixture differential one [43] (and similar) would apply to this AES-like cipher over $\mathbb{Z}_{p^n}^{4\times 4}$ as well. Indeed, such properties combine a truncated differential with the fact SS-box(x) = SS-box(y) holds if and only if SS-box(x') = SS-box(y') holds as well, where $x', y' \in \mathbb{Z}_{p^n}^{4\times 4}$ are obtained by *properly* "mixing the $\mathbb{Z}_{p^n}$-words" of $x, y \in \mathbb{Z}_{p^n}^{4\times 4}$ (for SS-box$(\cdot)$ being the super-Sbox over $\mathbb{Z}_{p^n}^4$ in our case).

Infinitely-Long Invariant Subspace Trails. As already observed in Section 4.2, the subring $\mathfrak{R}_\ell$ defined as

$$\mathfrak{R}_\ell \coloneqq \{p^\ell \cdot x \mid x \in \mathbb{Z}_{p^n}^t\}$$

for $\ell \in \{1, 2, \ldots, n-1\}$ is invariant through matrix multiplication. It is not hard to check that both a constant addition and the S-box considered here would only change the coset of such subspace. Hence, the AES-like cipher over $\mathbb{Z}_{p^n}^{4\times 4}$ considered here admits an infinitely-long invariant subspace trail, as defined in [60,61,47]. Such a powerful distinguisher shows that a non-polynomial S-box is not sufficient for guaranteeing security.

Before going on, we emphasize that a similar result may hold also in the case of e.g. "almost polynomial" S-boxes that preserve such subspace differences with high probability.

8 Design Criteria for Constructing Non-Polynomial S-Boxes over $\mathbb{Z}_{p^n}$

As we have seen in the previous sections, the security of a cipher depends heavily on the properties of its S-boxes. To avoid the attacks discussed so far, we recommend the following criteria when designing S-boxes over $\mathbb{Z}_{p^n}$.

- **Design Criterion 0.** As every cipher over $\mathbb{Z}_{p^n}^t$ admits a representation over $\mathbb{F}_{p^n}^t \equiv \mathbb{F}_p^{n\cdot t}$ as well, full diffusion must be achieved over $\mathbb{F}_p$ (that is, every $\mathbb{F}_p$-word in output depends on all the $\mathbb{F}_p$-words in input).
- **Design Criterion 1.** In Section 6.2, we have seen that if an S-box is too close to a polyfunction, then the attacker can leverage it to find a secret key using algebraic attacks. The probability of these attacks depends on $\mathtt{PC}_{\max}^{(\ell)}$. Thus, to prevent them, the S-box should be far from any polynomial approximation, i.e., $\mathtt{PC}_{\max}^{(\ell)}$ should be as small as possible.
- **Design Criterion 2.** In Section 7, we have seen that if the outputs of the S-box mod p^ℓ are not uniformly distributed for inputs that are congruent mod p^ℓ, the attacker can use a truncated differential attack to obtain a part of the secret key. To prevent this, the S-boxes should satisfy $\mathtt{CPP}^{(\ell)} \approx 1/p^\ell$ for each $\ell < n/2$, and otherwise be (close to being) well-balanced (c.f. Proposition 9).

– **Design Criterion 3.** To prevent the attacks in Section 2.3 and 3.3, the cipher should have good statistical properties, that is, the maximal differential probability and maximal linear correlation should be sufficiently small. To achieve this goal, it is crucial to keep in mind that the maximum linear correlation and the maximum differential probability of a polyfunction over $\mathbb{Z}_{p^n}$ are *never* smaller than $1/p$ independently of n (c.f. Section 4.2).

We remark that here we limit ourselves to focus on the security aspects. In practice, the choice of the S-box also depends on other considerations, including the cost metric the designers aim to minimize.

9 Construction of Non-Polynomial S-Boxes

Given the design criteria from the previous section, one may ask if there exists any S-box over $\mathbb{Z}_{p^n}$ that satisfy them. Here, we discuss a potential way to construct it.

A naive approach to set up such a non-polynomial function may be obtained via a "perturbation" of a polynomial function, that is, by modifying a polynomial function in a subset of $\mathbb{Z}_{p^n}$. As a concrete example, given a polynomial function P over $\mathbb{Z}_{p^n}$, we can obtain a non-polynomial F as

$$\mathsf{F}(x) = \begin{cases} \mathsf{P}((x+1) \bmod p) & \text{if } 0 \leq x < p\,, \\ \mathsf{P}(x) & \text{otherwise (i.e., } p \leq x < p^n)\,. \end{cases}$$

While this approach indeed allows us to set up a non-polynomial function, (at least) one problem arises: *how to determine the subset of $\mathbb{Z}_{p^n}$ over which the polynomial function should be perturbed?* Indeed, if this set is too small, then the obtained function is too close to a being a polynomial, while if it is large, it may be expensive to implement in practice.

In this section, we propose another concrete strategy for constructing non-polynomial invertible S-boxes, which resembles the "cut and sew" approach [44] and the "`Kintsugi`" one [46]. This strategy – called *"digit manipulation" strategy* – consists in

1. splitting the input in $\mathbb{Z}_{p^n}^t$ as an element of $\mathbb{Z}_{p^{m_0}}^t \times \mathbb{Z}_{p^{m_1}}^t \times \ldots \times \mathbb{Z}_{p^{m_{l-1}}}^t$ for $l \geq 2$ and $m_0, m_1, \ldots, m_{l-1} \in \mathbb{N}$ with $\sum_{i=0}^{l-1} m_i = n$;
2. applying a non-trivial function over $\mathbb{Z}_{p^{m_0}}^t \times \mathbb{Z}_{p^{m_1}}^t \times \ldots \times \mathbb{Z}_{p^{m_{l-1}}}^t$;
3. re-composing the $\mathbb{Z}_{p^n}^t$ element.

We limit ourselves to present the "digit manipulation" construction for functions with the same input and output spaces, i.e., $\mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}^t$. Recall that such functions can be at most $\lfloor n/2 \rfloor$-balanced, and that we refer to them as *well-balanced* if this is the case.

Remark 7. We emphasize that the main goal of this final section is to show that it is possible to efficiently construct an S-box over $\mathbb{Z}_{p^n}$ that satisfy all the design

criteria given in Section 8. The open problem to find other efficient strategies to construct such good S-boxes is left for future work. Moreover, we highlight that while the design strategy considered here is not new in the literature, the security properties we aim to achieve in the case of an integer ring are *different* than the ones required in a field.

9.1 Digit Manipulation: Characterizing and Constructing Balanced Permutations

We start by introducing some notation that will be used throughout this section. Write $\ell_1 = \lfloor n/2 \rfloor$ and $\ell_2 = \lceil n/2 \rceil$. Elements $z \in \mathbb{Z}_{p^n}^t$ will be (uniquely) represented as $z = z_2 p^{\ell_1} + z_1$, for a pair $(z_1, z_2) \in \mathbb{Z}_{p^{\ell_1}}^t \times \mathbb{Z}_{p^{\ell_2}}^t$. General functions over this ring will then be expressed as $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}^t$ defined as

$$\mathsf{F}(z_2 \cdot p^{\ell_1} + z_1) = \mathsf{F}_2(z_1, z_2) \cdot p^{\ell_1} + \mathsf{F}_1(z_1, z_2), \tag{5}$$

for $\mathsf{F}_i : \mathbb{Z}_{p^{\ell_1}}^t \times \mathbb{Z}_{p^{\ell_2}}^t \equiv \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^{\ell_i}}^t$. We highlight that this strategy follows the three steps just described.

Proposition 12. *Consider a function $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}^t$ and, using notation as in (5), the associated restricted function $\mathsf{F}_1(a, \cdot) : \mathbb{Z}_{p^{\ell_2}}^t \to \mathbb{Z}_{p^{\ell_1}}^t$, for a fixed $a \in \mathbb{Z}_{p^{\ell_1}}^t$. If $0 < \ell_1 \leq \ell_2$ and $\ell_1 + \ell_2 = n$, then F is p^{ℓ_1}-balanced if and only if $\mathsf{F}_1(a, \cdot)$ is a $p^{t(\ell_2 - \ell_1)}$-to-1 function for all $a \in \mathbb{Z}_{p^{\ell_1}}^t$.*

Proof. Proposition 7 implies that F is well-balanced if and only if it is p^{ℓ_1}-balanced. Hence, for F as in (5), this property depends solely on F_1. Indeed, by construction we now have that the congruence equality modulo p^{ℓ_1} (Definition 8) can be written $\mathtt{CE}^{(\ell_1)}(a, b) = |\{z_2 \in \mathbb{Z}_{p^{\ell_2}}^t \mid \mathsf{F}_1(a, z_2) = b\}|$. The statement follows from Definition 9 by noting that $\mathtt{CE}^{(\ell_1)}(a, b) = p^{t(\ell_2 - \ell_1)}$ for all $(a, b) \in \mathbb{Z}_{p^{\ell_1}}^t \times \mathbb{Z}_{p^{\ell_1}}^t$ if and only if $\mathsf{F}_1(a, \cdot)$ is a $p^{t(\ell_2 - \ell_1)}$-to-1 function for all $a \in \mathbb{Z}_{p^{\ell_1}}^t$. □

As a special case, note that if n is even, i.e., $\ell_1 = \ell_2$, then F is well-balanced if and only if $\mathsf{F}_1(a, \cdot)$ is a permutation for all $a \in \mathbb{Z}_{p^{\ell_1}}^t$. We moreover have the following bi-conditional statement on whether F is a permutation. Let $\mathsf{F}_1^{-1}(c)$ denote the preimage of $c \in \mathbb{Z}_{p^{\ell_1}}^t$ under F_1, and $im\left(\mathsf{F}_2|_{\mathsf{F}_1^{-1}(c)}\right)$ the image of F_2 restricted to this set.

Lemma 4. *A function $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}^t$ with notations as in (5) is a permutation if and only if $im\left(\mathsf{F}_2|_{\mathsf{F}_1^{-1}(c)}\right) = \mathbb{Z}_{p^{\ell_2}}^t$ for each $\forall c \in \mathbb{Z}_{p^{\ell_1}}^t$.*

Proof. Observe, by counting, that $\text{im}\left(\mathsf{F}_2|_{\mathsf{F}_1^{-1}(c)}\right) = \mathbb{Z}_{p^{\ell_2}}^t$ for each $\forall c \in \mathbb{Z}_{p^{\ell_1}}^t$ implies $|\mathsf{F}_1^{-1}(c)| = p^{t \cdot \ell_2}$ for all $c \in \mathbb{Z}_{p^{\ell_1}}^t$. In particular, $\mathsf{F}_2|_{\mathsf{F}_1^{-1}(c)}$ must be injective for all c, hence F is a permutation. Conversely, if $\text{im}\left(\mathsf{F}_2|_{\mathsf{F}_1^{-1}(c)}\right) = \mathbb{Z}_{p^{\ell_2}}^t$ for each $\forall c \in \mathbb{Z}_{p^{\ell_1}}^t$ does not hold, then there must exist a c such that $|\mathsf{F}_1^{-1}(c)| > p^{t \cdot \ell_2}$, which guarantees that there is a collision in the output of F. □

Proposition 12 and Lemma 4 suggest a constructive method for building permutations over $\mathbb{Z}_{p^n}^t$ that are well-balanced, by fixing permutations over $\mathbb{Z}_{p^{\ell_1}}^t$ when n is even (resp. permutations over $\mathbb{Z}_{p^{\ell_1}}^t$ and p^t-to-1 maps when n is odd). For completeness, we write this out in Appendix J, including the number of all well-balanced permutations in Corollary 10. A concrete example of the construction is also provided in Section 9.3.

9.2 Concrete Examples over $\mathbb{Z}_{p^n}$ for Small p^n

The previous section provides the theoretical framework for constructing balanced permutations, yielding optimal permutations with respect to $\mathtt{CE}$ and $\mathtt{CPP}$. However, the statistical properties may still be far from optimal. In the following we demonstrate by an exhaustive search on the balanced permutations that their statistical properties may be both good and bad. In effect, we argue that one should consider multiple properties when choosing a strong S-box.

Exhaustive Search over Balanced Permutations for $p^n = 16$. By Proposition 12, each well-balanced permutation can be parametrized by choosing $2 \cdot 2^{n/2}$ permutations of $2^{n/2}$ elements, yielding a total of $(2^{n/2}!)^{2^{n/2}+1}$ permutations. For $n = 4$, this results in $24^8 \approx 2^{36}$ balanced permutations, that we searched through. Our results are summarized in Appendix K. We remark that for every considered well-balanced permutation $\mathsf{F}\colon \mathbb{Z}_{16} \to \mathbb{Z}_{16}$, then $\mathtt{CE}^{(2)}(a, b) = 1$ by construction, which is the optimal in terms of balance mod 4. Furthermore, there is *no* tuple $(x, y) \in \mathcal{S}^{(2)}$ such that $\mathsf{F}(x) = \mathsf{F}(y) \mod p^\ell$, giving $\mathtt{CPP}^{(2)}(\mathsf{F}) = 0$. Finally, all considered permutations agree with some polynomial in at most 4 points, giving $\mathtt{PC}_{\max}(\mathsf{F}) = 1/4$, as predicted by Corollary 5.

Concrete Examples over $\mathbb{Z}_{16}$. We provide two examples of well-balanced permutations with opposite statistical properties, namely $\mathsf{G}, \mathsf{H}\colon \mathbb{Z}_{16} \to \mathbb{Z}_{16}$ defined as following: $\mathsf{G} = [0, 12, 4, 10, 1, 2, 14, 8, 7, 11, 3, 13, 6, 5, 9, 15]$, and $\mathsf{H} = [14, 0, 4, 8, 3, 2, 15, 13, 5, 11, 9, 6, 12, 1, 10, 7]$. Despite having differential uniformity 3 in $\mathbb{Z}_{16}$, G achieves maximal differential uniformity over $\mathbb{F}_2^4$ (namely, 16, or equivalently $\mathrm{DP}_{\max} = 1$), the function H combines all the desirable properties. Indeed, it has $\mathtt{CE}^{(2)}(\mathsf{H}) = 1$, $\mathtt{CPP}^{(2)}(\mathsf{H}) = 0$, $\mathtt{PC}_{\max} = 1/4$, while its differential uniformity over $\mathbb{Z}_{16}$ is 3, over $\mathbb{F}_2^4$ is 4, nonlinearity over $\mathbb{F}_4$ is 4, and the algebraic degree over $\mathbb{F}_2^4$ is 3.

9.3 Digit Manipulation: Concrete S-Boxes over $\mathbb{Z}_{p^n}$

We conclude with two concrete examples of concrete S-boxes over $\mathbb{Z}_{p^n}$ set up with the digit manipulation.

1st Example. Let $n = 2\ell$ for some integer $\ell \geq 1$, and let $\mathsf{L} : \mathbb{Z}_{p^\ell} \times \mathbb{Z}_{p^\ell} \to \mathbb{Z}_{p^\ell}$ be the function $\mathsf{L}(x_0, x_1) \coloneqq x_1 + \mathsf{G}(x_0)$ where $\mathsf{G} : \mathbb{Z}_{p^\ell} \to \mathbb{Z}_{p^\ell}$ is invertible. Given an invertible polynomial function P over $\mathbb{Z}_{p^n}$, we define the function F over $\mathbb{Z}_{p^n}$ as

$$\mathsf{F}(x) \coloneqq \mathsf{P}(\mathsf{L}(x_0, x_1) + p^\ell \cdot x_1),$$

for $x = x_0 + p^\ell \cdot x_1$ with $x_0, x_1 \in \mathbb{Z}_{p^\ell}$.

First of all, it is trivial to check that F is invertible. Indeed, given $y = \mathsf{F}(x)$, note that $\mathsf{L}(x_0, x_1) + p^\ell \cdot x_1 = \mathsf{P}^{-1}(y)$, which allows us to find x_1, and then x_0.

Next, note that F is well-balanced. Indeed, due to Proposition 12, it is sufficient to prove that $\mathsf{L}(a, \cdot)$ is a permutation for each $a \in \mathbb{Z}_{p^\ell}$, which follows immediately by the definition of L. This ensures that F is optimal with respect to e.g. the $\mathtt{CPP}^\ell$ parameters introduced before.

Finally, we point out that it is possible to provide good bounds for the statistical properties of F thanks to P being a polynomial. Indeed, consider the probability of a differential $\Delta_I \mapsto \Delta_O$ over $\mathbb{Z}_{p^n}$. Let $\Delta_I = \delta_0 + p^\ell \cdot \delta_1$ for $\delta_0, \delta_1 \in \mathbb{Z}_{p^\ell}$. As P is a polynomial, we have that $\left(\mathsf{F}\left((x + \Delta_I) \bmod p^\ell\right) - \mathsf{F}\left((x) \bmod p^\ell\right)\right) = \Delta_O \bmod p^\ell$, that is, $\left(\mathsf{P}\left(x_1 + \delta_1 + \mathsf{G}(x_0 + \delta_0)\right) - \mathsf{P}\left(x_1 + \mathsf{G}(x_0)\right)\right) = \Delta_O \bmod p^\ell$, which admits at most $p^\ell \cdot (\deg(\mathsf{P}) - 1)$ solutions. It follows that $\mathrm{DP}_{\max}(\mathsf{F}) \leq p^\ell \cdot (\deg(\mathsf{P}) - 1)/p^n = (\deg(\mathsf{P}) - 1)/p^\ell$. Similar result can be obtained for the linear correlation.

2nd Example. Next, we generalize the previous construction by re-considering the "cut and sew" approach [44] and the "$\mathtt{Kintsugi}$" one [46]

Let m be a factor of n, that is, $n = m \cdot s$. Then:

1. decompose $x \in \mathbb{Z}_{p^n}$ as $x_0, x_1, \ldots, x_{s-1} \in \mathbb{Z}_{p^m}$, where $x_0, x_1, \ldots, x_{s-1} \in \mathbb{Z}_{p^m}$ are the unique elements such that $x = \sum_{i=0}^{s-1} x_i \cdot p^{m \cdot i}$;
2. given F an invertible function over $\mathbb{Z}_{p^m}^s$, let
$$(y_0, y_1, \ldots, y_{s-1}) = \mathsf{F}(x_0, x_1, \ldots, x_{s-1})$$
3. recompose the element $y \in \mathbb{Z}_{p^n}$ as $x = \sum_{i=0}^{s-1} y_i \cdot p^{m \cdot i}$.

A possible way to instantiate F is via a r-round SPN scheme over $\mathbb{Z}_{p^m}^s$ described already in Section 3. In the case p^m being small, it is possible to set up the S-box S over $\mathbb{Z}_{p^m}$ via the functions proposed in Section 9.2. In the case $m = 1$, hence, $\mathbb{Z}_{p^1} \equiv \mathbb{F}_p$ being a prime field, we recall that every function S admits a polynomial representation over $\mathbb{Z}_p$.

Due to the one-to-one correspondence between $\mathbb{Z}_{p^m}^s$ and $\mathbb{Z}_{p^n}$, it is possible to deduce the cryptographic properties of the obtained function over $\mathbb{Z}_{p^n}$ via its representation over $\mathbb{Z}_{p^m}^s$. Consider for example the case of $\mathtt{CPP}^\ell$, that is, the probability that given two elements $x, y \in \mathbb{Z}_{p^n}$ that are equal modulo p^ℓ, their images are equal given modulo p^ℓ. Hence:

- if ℓ is a multiple of m (which is always the case for $m = 1$), this corresponds to the probability of the truncated differential $[*, *, \ldots, *, 0, 0, \ldots 0] \mapsto [*, *, \ldots, *, 0, 0, \ldots 0]$ for R over $\mathbb{Z}_{p^m}^s$, where $[0, 0, \ldots 0] \in \mathbb{Z}_{p^m}^{\ell/m}$;
- if ℓ is not a multiple of m, it is trivial to derive an upper bound by working with truncated differentials with $[0, 0, \ldots 0] \in \mathbb{Z}_{p^m}^{\lceil \ell/m \rceil}$.

Hence, in order to obtain a well-balanced function, it is sufficient to e.g. instantiate F as a 2-round SPN scheme, with the linear layer being a MDS matrix, and S being non-linear over $\mathbb{Z}_{p^m}$. We limit ourselves to emphasize that a similar argument holds for the statistical properties over $\mathbb{Z}_{p^n}$, which can be computed by analyzing the statistical properties of F over $\mathbb{Z}_{p^m}^s$.

Acknowledgment. Lorenzo Grassi and Berenika Richterová were supported by the European Research Council (ERC), grant number 101160608 "SYMPZON". Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

References

1. Amiri-Eliasi, P., Ghosh, K., Daemen, J.: Mystrium: Wide Block Encryption Efficient on Entry-Level Processors. In: Security and Cryptography for Networks - SCN 2024, Part II. LNCS, vol. 14974, pp. 71–96. Springer (2024). https://doi.org/10.1007/978-3-031-71073-5_4
2. Ashur, T., Beyne, T., Rijmen, V.: Revisiting the Wrong-Key-Randomization Hypothesis. J. Cryptol. **33**(2), 567–594 (2020). https://doi.org/10.1007/S00145-020-09343-2
3. Baignères, T., Stern, J., Vaudenay, S.: Linear Cryptanalysis of Non Binary Ciphers. In: Selected Areas in Cryptography - SAC 2007. LNCS, vol. 4876, pp. 184–211. Springer (2007). https://doi.org/10.1007/978-3-540-77360-3_13
4. Bak, A.: A Note on the Rank Defect Phenomena in The Linearization Attack on Elisabeth-4. Cryptology ePrint Archive, Paper 2025/1133 (2025), https://eprint.iacr.org/2025/1133
5. Bak, A., Ghosh, S., Liu, F., Meier, W., Ni, J., Perrin, L.: Cryptanalysis of TFHE-Friendly Cipher FRAST. IACR Trans. Symmetric Cryptol. **2026**(1), 119–147 (2026). https://doi.org/10.46586/TOSC.V2026.I1.119-147, https://doi.org/10.46586/tosc.v2026.i1.119-147
6. Bardet, M., Faugère, J.C., Salvy, B.: On the complexity of the F_5 Gröbner basis algorithm. Journal of Symbolic Computation **70**, 49–70 (2015)
7. Bardet, M., Faugère, J.C., Salvy, B., Spaenlehauer, P.J.: On the complexity of solving quadratic Boolean systems. Journal of Complexity **29**(1), 53–75 (2013)
8. Bardet, M., Faugère, J.C., Salvy, B., Yang, B.Y.: Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In: Proc. of MEGA. vol. 5, pp. 2–2 (2005)
9. Baum, C., Braun, L., Munch-Hansen, A., Razet, B., Scholl, P.: Appenzeller to brie: Efficient zero-knowledge proofs for mixed-mode arithmetic and z2k. In: CCS '21: 2021 ACM SIGSAC Conference on Computer and Communications Security. pp. 192–211. ACM (2021). https://doi.org/10.1145/3460120.3484812
10. Baum, C., Braun, L., Munch-Hansen, A., Scholl, P.: Moz$\mathbb{Z}_{2^k}$arella: Efficient Vector-OLE and Zero-Knowledge Proofs over $\mathbb{Z}_{2^k}$. In: Advances in Cryptology - CRYPTO 2022, Part IV. LNCS, vol. 13510, pp. 329–358. Springer (2022). https://doi.org/10.1007/978-3-031-15985-5_12
11. Beyne, T.: A Geometric Approach to Linear Cryptanalysis. In: Advances in Cryptology - ASIACRYPT 2021, Part I. LNCS, vol. 13090, pp. 36–66. Springer (2021). https://doi.org/10.1007/978-3-030-92062-3_2
12. Beyne, T., Rijmen, V.: Differential Cryptanalysis in the Fixed-Key Model. In: Advances in Cryptology - CRYPTO 2022, Part III. LNCS, vol. 13509, pp. 687–716. Springer (2022). https://doi.org/10.1007/978-3-031-15982-4_23
13. Beyne, T., Verbauwhede, M.: Ultrametric Integral Cryptanalysis. In: Advances in Cryptology - ASIACRYPT 2024, Part VII. LNCS, vol. 15490, pp. 392–423. Springer (2024). https://doi.org/10.1007/978-981-96-0941-3_13

14. Beyne, T., Verbauwhede, M.: Integral Cryptanalysis in Characteristic p. In: Advances in Cryptology - ASIACRYPT 2025, Part I. LNCS, vol. 16245, pp. 66–96. Springer (2025). https://doi.org/10.1007/978-981-95-5018-0_3
15. Biham, E., Shamir, A.: Differential Cryptanalysis of DES-like Cryptosystems. J. Cryptol. **4**(1), 3–72 (1991). https://doi.org/10.1007/BF00630563
16. Bogdanov, A., Tischhauser, E.: On the Wrong Key Randomisation and Key Equivalence Hypotheses in Matsui's Algorithm 2. In: Fast Software Encryption - FSE 2013. LNCS, vol. 8424, pp. 19–38. Springer (2013). https://doi.org/10.1007/978-3-662-43933-3_2
17. Bouillaguet, C., Chen, H.C., Cheng, C.M., Chou, T., Niederhagen, R., Shamir, A., Yang, B.Y.: Fast Exhaustive Search for Polynomial Systems in $\mathbb{F}_2$. In: International Workshop on Cryptographic Hardware and Embedded Systems. pp. 203–218. Springer (2010)
18. Bouvier, C., Grassi, L., Khovratovich, D., Koschatko, K., Rechberger, C., Schmid, F., Schofnegger, M.: Skyscraper: Fast Hashing on Big Primes. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2025**(2), 743–780 (2025). https://doi.org/10.46586/TCHES.V2025.I2.743-780
19. Carlitz, L.: Functions and polynomials (mod p^n). Acta Arithmetica **9**(1), 67–68 (1964)
20. Cheon, J.H., Coron, J., Kim, J., Lee, M.S., Lepoint, T., Tibouchi, M., Yun, A.: Batch Fully Homomorphic Encryption over the Integers. In: Advances in Cryptology - EUROCRYPT 2013. LNCS, vol. 7881, pp. 315–335. Springer (2013). https://doi.org/10.1007/978-3-642-38348-9_20
21. Cheon, J.H., Kim, A., Kim, M., Song, Y.S.: Homomorphic Encryption for Arithmetic of Approximate Numbers. In: Advances in Cryptology - ASIACRYPT 2017, Part I. LNCS, vol. 10624, pp. 409–437. Springer (2017). https://doi.org/10.1007/978-3-319-70694-8_15
22. Chillotti, I., Gama, N., Georgieva, M., Izabachène, M.: TFHE: Fast Fully Homomorphic Encryption Over the Torus. J. Cryptol. **33**(1), 34–91 (2020). https://doi.org/10.1007/S00145-019-09319-X
23. Cho, J., Ha, J., Kim, S., Lee, B., Lee, J., Lee, J., Moon, D., Yoon, H.: Transciphering Framework for Approximate Homomorphic Encryption. In: Advances in Cryptology - ASIACRYPT 2021, Part III. LNCS, vol. 13092, pp. 640–669. Springer (2021). https://doi.org/10.1007/978-3-030-92078-4_22
24. Cho, M., Chung, W., Ha, J., Lee, J., Oh, E., Son, M.: FRAST: TFHE-Friendly Cipher Based on Random S-Boxes. IACR Trans. Symmetric Cryptol. **2024**(3), 1–43 (2024). https://doi.org/10.46586/TOSC.V2024.I3.1-43
25. Coron, J., Mandal, A., Naccache, D., Tibouchi, M.: Fully Homomorphic Encryption over the Integers with Shorter Public Keys. In: Advances in Cryptology - CRYPTO 2011. LNCS, vol. 6841, pp. 487–504. Springer (2011). https://doi.org/10.1007/978-3-642-22792-9_28
26. Cosseron, O., Hoffmann, C., Méaux, P., Standaert, F.: Towards Case-Optimized Hybrid Homomorphic Encryption - Featuring the Elisabeth Stream Cipher. In: Advances in Cryptology - ASIACRYPT 2022, Part III. LNCS, vol. 13793, pp. 32–67. Springer (2022). https://doi.org/10.1007/978-3-031-22969-5_2
27. Cramer, R., Damgård, I., Escudero, D., Scholl, P., Xing, C.: SPD$\mathbb{Z}_{2^k}$: Efficient MPC mod 2^k for Dishonest Majority. In: Advances in Cryptology - CRYPTO 2018, Part II. LNCS, vol. 10992, pp. 769–798. Springer (2018). https://doi.org/10.1007/978-3-319-96881-0_26

28. Daemen, J., Knudsen, L.R., Rijmen, V.: The Block Cipher Square. In: Fast Software Encryption - FSE 1997. LNCS, vol. 1267, pp. 149–165. Springer (1997). https://doi.org/10.1007/BFB0052343
29. Daemen, J., Rijmen, V.: AES and the Wide Trail Design Strategy. In: Advances in Cryptology - EUROCRYPT 2002. LNCS, vol. 2332, pp. 108–109. Springer (2002). https://doi.org/10.1007/3-540-46035-7_7
30. Daemen, J., Rijmen, V.: Security of a Wide Trail Design. In: Progress in Cryptology - INDOCRYPT 2002. LNCS, vol. 2551, pp. 1–11. Springer (2002). https://doi.org/10.1007/3-540-36231-2_1
31. Daemen, J., Rijmen, V.: The Design of Rijndael - The Advanced Encryption Standard (AES), Second Edition. Information Security and Cryptography, Springer (2020). https://doi.org/10.1007/978-3-662-60769-5
32. Dalskov, A.P.K., Escudero, D., Nof, A.: Fully Secure MPC and zk-FLIOP over Rings: New Constructions, Improvements and Extensions. In: Advances in Cryptology - CRYPTO 2024, Part VIII. LNCS, vol. 14927, pp. 136–169. Springer (2024). https://doi.org/10.1007/978-3-031-68397-8_5
33. Delsarte, P., Goethals, J., Mac Williams, F.: On generalized reed-muller codes and their relatives. Information and Control **16**(5), 403–442 (1970). https://doi.org/https://doi.org/10.1016/S0019-9958(70)90214-7, https://www.sciencedirect.com/science/article/pii/S0019995870902147
34. van Dijk, M., Gentry, C., Halevi, S., Vaikuntanathan, V.: Fully Homomorphic Encryption over the Integers. In: Advances in Cryptology - EUROCRYPT 2010. LNCS, vol. 6110, pp. 24–43. Springer (2010). https://doi.org/10.1007/978-3-642-13190-5_2
35. Dinur, I.: Cryptanalytic Applications of the Polynomial Method for Solving Multivariate Equation Systems over GF(2). In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 374–403. Springer (2021)
36. Dobbertin, H.: Cryptanalysis of MD4. J. Cryptol. **11**(4), 253–271 (1998). https://doi.org/10.1007/S001459900047
37. Escudero, D., Song, Y., Wang, W.: Perfectly-Secure Multiparty Computation with Linear Communication Complexity over Any Modulus. In: Advances in Cryptology - ASIACRYPT 2024, Part VI. LNCS, vol. 15489, pp. 102–132. Springer (2024). https://doi.org/10.1007/978-981-96-0938-3_4
38. Escudero, D., Xing, C., Yuan, C.: More Efficient Dishonest Majority Secure Computation over $\mathbb{Z}_{2^k}$ via Galois Rings. In: Advances in Cryptology - CRYPTO 2022, Part I. LNCS, vol. 13507, pp. 383–412. Springer (2022). https://doi.org/10.1007/978-3-031-15802-5_14
39. Ganesh, C., Nitulescu, A., Soria-Vazquez, E.: Rinocchio: Snarks for ring arithmetic. J. Cryptol. **36**(4), 41 (2023). https://doi.org/10.1007/S00145-023-09481-3
40. Garreta, A., Waldner, H., Vlasov, I., Hristova, K., Dall'Ava, L., Cupic, M., Klein, M.: sfZinc: Succinct Arguments with Small Arithmetization Overheads from IOPs of Proximity to the Integers. In: Advances in Cryptology - CRYPTO 2025, Part VII. LNCS, vol. 16006, pp. 259–291. Springer (2025). https://doi.org/10.1007/978-3-032-01907-3_9
41. Geelen, R., Iliashenko, I., Kang, J., Vercauteren, F.: On Polynomial Functions Modulo p^e and Faster Bootstrapping for Homomorphic Encryption. In: Advances in Cryptology - EUROCRYPT 2023, Part III. LNCS, vol. 14006, pp. 257–286. Springer (2023). https://doi.org/10.1007/978-3-031-30620-4_9

42. Gilbert, H., Boissier, R.H., Jean, J., Reinhard, J.: Cryptanalysis of Elisabeth-4. In: Advances in Cryptology - ASIACRYPT 2023, Part III. LNCS, vol. 14440, pp. 256–284. Springer (2023). https://doi.org/10.1007/978-981-99-8727-6_9
43. Grassi, L.: Mixture Differential Cryptanalysis: a New Approach to Distinguishers and Attacks on round-reduced AES. IACR Transactions on Symmetric Cryptology **2018**(2), 133–160 (2018). https://doi.org/10.13154/tosc.v2018.i2.133-160
44. Grassi, L., Ayala, I.M., Hovd, M.N., Øygarden, M., Raddum, H., Wang, Q.: Cryptanalysis of Symmetric Primitives over Rings and a Key Recovery Attack on Rubato. In: Advances in Cryptology - CRYPTO 2023, Part III. LNCS, vol. 14083, pp. 305–339. Springer (2023). https://doi.org/10.1007/978-3-031-38548-3_11
45. Grassi, L., Khovratovich, D., Lüftenegger, R., Rechberger, C., Schofnegger, M., Walch, R.: Reinforced Concrete: A Fast Hash Function for Verifiable Computation. In: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security - CCS 2022. pp. 1323–1335. ACM (2022). https://doi.org/10.1145/3548606.3560686
46. Grassi, L., Khovratovich, D., Lüftenegger, R., Rechberger, C., Schofnegger, M., Walch, R.: Monolith: Circuit-Friendly Hash Functions with New Nonlinear Layers for Fast and Constant-Time Implementations. IACR Trans. Symmetric Cryptol. **2024**(3), 44–83 (2024). https://doi.org/10.46586/TOSC.V2024.I3.44-83
47. Grassi, L., Rechberger, C., Rønjom, S.: Subspace Trail Cryptanalysis and its Applications to AES. IACR Trans. Symmetric Cryptol. **2016**(2), 192–225 (2016). https://doi.org/10.13154/TOSC.V2016.I2.192-225
48. Grassi, L., Rechberger, C., Rønjom, S.: A New Structural-Differential Property of 5-Round AES. In: Advances in Cryptology - EUROCRYPT 2017, Part II. LNCS, vol. 10211, pp. 289–317 (2017). https://doi.org/10.1007/978-3-319-56614-6_10
49. Guo, J., Peyrin, T., Poschmann, A.: The PHOTON Family of Lightweight Hash Functions. In: Advances in Cryptology - CRYPTO 2011. LNCS, vol. 6841, pp. 222–239. Springer (2011). https://doi.org/10.1007/978-3-642-22792-9_13
50. Guo, J., Peyrin, T., Poschmann, A., Robshaw, M.J.B.: The LED Block Cipher. In: Cryptographic Hardware and Embedded Systems - CHES 2011. LNCS, vol. 6917, pp. 326–341. Springer (2011). https://doi.org/10.1007/978-3-642-23951-9_22
51. Ha, J., Kim, S., Lee, B., Lee, J., Son, M.: Rubato: Noisy Ciphers for Approximate Homomorphic Encryption. In: Advances in Cryptology - EUROCRYPT 2022, Part I. LNCS, vol. 13275, pp. 581–610. Springer (2022). https://doi.org/10.1007/978-3-031-06944-4_20
52. Harpes, C., Kramer, G., Massey, J.L.: A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma. In: Advances in Cryptology - EUROCRYPT 1995. LNCS, vol. 921, pp. 24–38. Springer (1995). https://doi.org/10.1007/3-540-49264-X_3
53. Hong, S., Lee, S., Lim, J., Sung, J., Cheon, D.H., Cho, I.: Provable Security against Differential and Linear Cryptanalysis for the SPN Structure. In: Fast Software Encryption – FSE 2000. LNCS, vol. 1978, pp. 273–283. Springer (2000). https://doi.org/10.1007/3-540-44706-7_19, https://doi.org/10.1007/3-540-44706-7_19
54. Jakobsen, T., Knudsen, L.R.: The Interpolation Attack on Block Ciphers. In: Fast Software Encryption - FSE 1997. LNCS, vol. 1267, pp. 28–40. Springer (1997). https://doi.org/10.1007/BFB0052332
55. Joux, A., Vitse, V.: A Crossbred Algorithm for Solving Boolean Polynomial Systems. In: International Conference on Number-Theoretic Methods in Cryptology. pp. 3–21. Springer (2017)

56. Keller, G., Olson, F.R.: Counting polynomial functions (mod p^n). Duke Mathematical Journal **35**(4), 835–838 (1968)
57. Knudsen, L.R.: Truncated and Higher Order Differentials. In: Fast Software Encryption - FSE 1994. LNCS, vol. 1008, pp. 196–211 (1994). https://doi.org/10.1007/3-540-60590-8_16
58. Knudsen, L.R., Robshaw, M.J.B.: Non-Linear Approximations in Linear Cryptanalysis. In: Advances in Cryptology - EUROCRYPT 1996. LNCS, vol. 1070, pp. 224–236. Springer (1996). https://doi.org/10.1007/3-540-68339-9_20
59. Lacan, J., Fimes, J.: Systematic MDS erasure codes based on Vandermonde matrices. IEEE Commun. Lett. **8**(9), 570–572 (2004). https://doi.org/10.1109/LCOMM.2004.833807
60. Leander, G., Abdelraheem, M.A., Alkhzaimi, H., Zenner, E.: A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack. In: Advances in Cryptology - CRYPTO 2011. LNCS, vol. 6841, pp. 206–221. Springer (2011). https://doi.org/10.1007/978-3-642-22792-9_12
61. Leander, G., Minaud, B., Rønjom, S.: A Generic Approach to Invariant Subspace Attacks: Cryptanalysis of Robin, iSCREAM and Zorro. In: Advances in Cryptology - EUROCRYPT 2015, Part I. LNCS, vol. 9056, pp. 254–283. Springer (2015). https://doi.org/10.1007/978-3-662-46800-5_11
62. Li, C., Wang, Q.: Design of Lightweight Linear Diffusion Layers from Near-MDS Matrices. IACR Trans. Symmetric Cryptol. **2017**(1), 129–155 (2017). https://doi.org/10.13154/TOSC.V2017.I1.129-155
63. Li, Z., Xing, C., Yao, Y., Yuan, C.: Efficient Pseudorandom Correlation Generators over $\mathbb{Z}/p^k\mathbb{Z}$. In: Advances in Cryptology - CRYPTO 2025, Part IV. LNCS, vol. 16003, pp. 207–240. Springer (2025). https://doi.org/10.1007/978-3-032-01884-7_7
64. Lin, F., Xing, C., Yao, Y.: More Efficient Zero-Knowledge Protocols over $\mathbb{Z}_{2^k}$ via Galois Rings. In: Advances in Cryptology - CRYPTO 2024, Part IX. LNCS, vol. 14928, pp. 424–457. Springer (2024). https://doi.org/10.1007/978-3-031-68400-5_13
65. Lokshtanov, D., Paturi, R., Tamaki, S., Williams, R., Yu, H.: Beating Brute Force for Systems of Polynomial Equations over Finite Fields. In: Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 2190–2202. SIAM (2017)
66. MacWilliams, F., Sloane, N.: The Theory of Error-Correcting Codes. North-holland Publishing Company, 2nd edn. (1978)
67. Matsui, M.: Linear Cryptanalysis Method for DES Cipher. In: Advances in Cryptology - EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397. Springer (1993). https://doi.org/10.1007/3-540-48285-7_33
68. Perrin, L.: Partitions in the S-Box of Streebog and Kuznyechik. IACR Transactions on Symmetric Cryptology **2019**(1), 302–329 (Mar 2019). https://doi.org/10.13154/tosc.v2019.i1.302-329, https://tosc.iacr.org/index.php/ToSC/article/view/7405
69. Quintin, G., Barbier, M., Chabot, C.: On Generalized Reed–Solomon Codes Over Commutative and Noncommutative Rings. IEEE Transactions on Information Theory **59**(9), 5882–5897 (2013). https://doi.org/10.1109/TIT.2013.2264797
70. Rijmen, V., Daemen, J., Preneel, B., Bosselaers, A., Win, E.D.: The Cipher SHARK. In: Fast Software Encryption - FSE 1996. LNCS, vol. 1039, pp. 99–111. Springer (1996). https://doi.org/10.1007/3-540-60865-6_47
71. Rivest, R.L.: Permutation Polynomials Modulo 2^w. Finite Fields and Their Applications **7**(2), 287–292 (2001). https://doi.org/https://doi.org/10.1006/ffta.2000.0282

72. Roth, R.M., Lempel, A.: On MDS codes via Cauchy matrices. IEEE Trans. Inf. Theory **35**(6), 1314–1319 (1989). https://doi.org/10.1109/18.45291
73. Selçuk, A.A.: On Probability of Success in Linear and Differential Cryptanalysis. J. Cryptol. **21**(1), 131–147 (2008). https://doi.org/10.1007/S00145-007-9013-7
74. Singmaster, D.: On polynomial functions (mod m). Journal of Number Theory **6**(5), 345–352 (1974)
75. Szepieniec, A., Lemmens, A., Sauer, J.F., Threadbare, B., Al-Kindi: The Tip5 Hash Function for Recursive STARKs. Cryptology ePrint Archive, Paper 2023/107 (2023), https://eprint.iacr.org/2023/107
76. Todo, Y.: Structural Evaluation by Generalized Integral Property. In: Advances in Cryptology - EUROCRYPT 2015, Part I. LNCS, vol. 9056, pp. 287–314. Springer (2015). https://doi.org/10.1007/978-3-662-46800-5_12
77. Vasiliev, N.N., Kanzheleva, O.: Polynomial Interpolation over the Residue Rings $\mathbb{Z}_n$. Journal of Mathematical Sciences **209**(6), 845–850 (2015)

SUPPLEMENTARY MATERIAL

A Results on Polyfuctions over Integer Rings

A.1 Smarandache function

Here, we provide concrete upper bounds for $\mu(p^n)$.

Lemma 5. *$\mu(p^i) \leq \mu(p^{i+1})$.*

Proof. Let j be the smallest integer such that p^{i+1} divides $j!$, that is, $j! \bmod p^{i+1} = 0$. Then, it follows that $j! \bmod p^i = 0$ as well. Hence, the claim. □

Lemma 6. *$\mu(p^n) \leq n \cdot p$.*

Proof. Since $(n \cdot p)! = \prod_{i=1}^{n \cdot p} i$, it follows that $j \cdot p$ divides $(n \cdot p)!$ for each $j \in \{1, 2, \ldots, n\}$. Hence, $\prod_{j=1}^{n} j \cdot p = p^n \cdot \prod_{j=1}^{n}$ divides $(n \cdot p)!$, which implies that p^n divides $(n \cdot p)!$. The claim follows from the definition of $\mu(p^n)$, as $\mu(p^n) = i$ if and only if i is the smallest integer such that $i! \bmod p^n = 0$. □

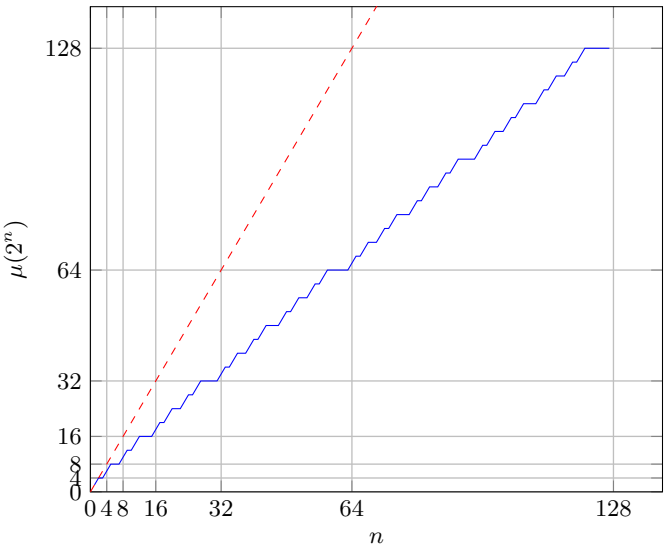


Fig. 1: Values of $\mu(2^n)$ compared to the bound in red.

A graphical representation of $\mu(2^n)$ is given in Fig. 1. As it is possible to observe, the bound just given is not tight in general (see red dashed line). A closed formula for $\mu(p^n)$ is known only for some values of n.[3]

Lemma 7. *$\mu(p^{p^l}) = p^l \cdot (p-1) + p$ for $l \geq 1$.*

[3] https://mathworld.wolfram.com/SmarandacheFunction.html

A.2 Proof of Proposition 1

As already mentioned, for $\mathsf{F} : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}$ being polynomial, the statement (1) is a simple generalization of the analogous statement for functions over $\mathbb{Z}_{p^n}$ from [19].

Now, for $\mathsf{F} \equiv (\mathsf{F}_0, \mathsf{F}_1, \ldots, \mathsf{F}_{t-1}) : \mathbb{Z}_{p^n}^t \to \mathbb{Z}_{p^n}^t$, the implication $\Longrightarrow$ follows immediately. For the implication $\Longleftarrow$, we will show that a permutation polyfunction over $\mathbb{Z}_{p^n}^t$ is also a permutation for $\mathbb{Z}_{p^i}^t$ for any $i \leq n$. Then, the claim follows from the injectivity of this permutation.

For contradiction, assume that F is not invertible over $\mathbb{Z}_{p^{n-1}}^t$. Hence, there must exist two different inputs $x, x' \in \mathbb{Z}_{p^{n-1}}^t$ such that $y = \mathsf{F}(x) = \mathsf{F}(x') \mod p^{n-1}$. It also follows that

$$y = \mathsf{F}(x + p^{n-1} \cdot v) \mod p^{n-1} \qquad \text{and} \qquad y = \mathsf{F}(x' + p^{n-1} \cdot v) \mod p^{n-1}$$

for each $v \in \{0, 1, \ldots, p-1\}^t$. This gives $2 \cdot p^t$ inputs that F maps to a value congruent to y modulo p^{n-1}. However, there can only be p^t such values if F is a permutation polyfunction. Thus $\mathsf{F} \bmod p^{n-1}$ is invertible over $\mathbb{Z}_{p^{n-1}}^t$ and, by repeating the procedure, $\mathsf{F} \bmod p^i$ is invertible over $\mathbb{Z}_{p^i}^t$ for each $i \leq n$.

□

A.3 Proof of Proposition 2

The values that a polyfunction takes on the inputs $0, 1, \ldots, \mu(p^n) - 1$ determines the entire function by [41, Theorem 3]. If $\mathsf{F}(x \bmod p^n) = \mathsf{G}(x \bmod p^n)$ holds for each $x \in \{a, a+1, \ldots, a + \mu(p^n) - 1\}$ for some a, then $\hat{\mathsf{F}}(x) = \mathsf{F}(x - a)$ and $\hat{\mathsf{G}}(x) = \mathsf{G}(x - a)$ agree on $0, 1, \ldots, \mu(p^n) - 1$. Hence $\hat{\mathsf{F}} = \hat{\mathsf{G}}$, implying that $\mathsf{F} = \mathsf{G}$.

□

A.4 Criteria of Polyfunctions over $\mathbb{Z}_{p^n}$

Here, we recall a result from [41] to check whenever a function is a polyfunction or not. For doing this, we first need to recall the definition of the i-th forward difference.

Definition 11. *The reduced i-th forward difference of a function F over $\mathbb{Z}_{p^n}$ is recursively defined as*

$$\Delta^i \mathsf{F}(x) \coloneqq \begin{cases} \mathsf{F}(x) & \text{if } i = 0 \\ \left(\Delta^{i-1}\mathsf{F}(x+1) - \Delta^{i-1}\mathsf{F}(x)\right) \bmod p^n & \text{otherwise (if } i > 0) \end{cases}.$$

In the following, let $\Delta^i \mathsf{F}(x) \mid_{x=j}$ be the i-th reduced forward difference of F evaluated at $j \in \mathbb{Z}_{p^n}$.

Theorem 6 ([41, Theorem 3]). *A function F over $\mathbb{Z}_{p^n}$ is a polyfunction if and only if the following two criteria are satisfied:*

1. *for all $i < \mu(p^n)$: $v_p(\Delta^i \mathsf{F}(x) \mid_{x=0}) \geq v(i!)$;*
2. *for all $i < p^n - \mu(p^n)$: $\Delta^{\mu_{p^n}} \mathsf{F}(x) \mid_{x=i} = 0$.*

We refer to [41] for a complete proof.

A.5 Null Polynomials

Corollary 6 ([74]). *If* O *is a monic null polyfunction, then* $\deg(\mathsf{O}) \geq \mu(p^n)$. *The number* $\#_{\mathsf{O}}(p^n)$ *of null polynomials over* $\mathbb{Z}_{p^n}$ *of degree strictly less than* $\mu(p^n)$ *is given by*

$$\#_{\mathsf{O}}(p^n) = \prod_{i=0}^{\mu(p^n)-1} \gcd(i!, p^n)\,.$$

Corollary 7 ([74, Corollary 9]). *The number* $\#_{\mathsf{F}}(p^n)$ *of distinct polynomial functions (equivalently, of polyfunctions in canonical form) over* $\mathbb{Z}_{p^n}$ *is given by*

$$\#_{\mathsf{F}}(p^n) = \prod_{i=0}^{\mu(p^n)-1} \frac{p^n}{\gcd(i!, p^n)} = \prod_{i=0}^{\mu(p^n)-1} p^{n-\upsilon_p(i!)} = \frac{p^{n\cdot\mu(p^n)}}{\#_{\mathsf{O}}(p^n)}\,.$$

For comparison, the total number of functions over $\mathbb{Z}_{p^n}$ is given by $(p^n)^{p^n}$. It follows that the number of polyfunctions is in general much smaller than the total number of functions over $\mathbb{Z}_{p^n}$, since $p^{n-\upsilon_p(i!)} \leq p^n$ and $\mu(p^n) - 1 < p^n$.

A.6 Polynomial Permutations over $\mathbb{Z}_{p^n}$

First of all, we recall the Hensel's lemma.

Proposition 13 (Hensel's Lemma). *Let* $\mathsf{F}(x) = \sum_{i=0}^{d} a_i \cdot x^i$ *be a non-zero polynomial function over* $\mathbb{Z}_{p^n}$. *Suppose there exists an integer* $\alpha' \in \mathbb{Z}_p$ *such that*

$$\mathsf{F}(\alpha') = 0 \mod p \qquad \textit{and} \qquad \mathsf{F}'(\alpha') \neq 0 \mod p\,.$$

Then there exists a unique integer $\alpha \in \mathbb{Z}_{p^n}$ *such that*

$$\mathsf{F}(\alpha) = 0 \mod p^n \qquad \textit{and} \qquad \mathsf{F}(\alpha) = \mathsf{F}(\alpha') \mod p\,.$$

We highlight that if a polyfunction is invertible, then its inverse admits polynomial representation.

Proposition 14. *If a polyfunction* $\mathsf{F}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$ *is invertible, then its inverse admits polynomial representation.*

Proof. Now $\mathsf{F} \in \mathrm{Sym}(\mathbb{Z}_{p^n})$, so it has finite order. In particular, there is an $m > 0$ such that $\mathsf{F}^m = \mathsf{F} \circ \mathsf{F} \circ \cdots \circ \mathsf{F} = \mathrm{Id}_{\mathbb{Z}_{p^n}}$. Hence, $\mathsf{F}^{-1} = \mathsf{F}^{m-1}$ is polynomial since the composition of polyfunctions is a polyfunction. □

We limit ourselves to mention that Vasiliev and Kanzheleva [77] proposed a recursive formula for computing the number of polynomial functions over an integer ring $\mathbb{Z}_{p^n}$.

Special Case: Polynomial Permutations over $\mathbb{Z}_{2^n}$. Here, we discuss the case of polynomial permutations over $\mathbb{Z}_{2^n}$.

The following condition given by Rivest [71] allows us to check if a polyfunction over $\mathbb{Z}_{2^n}$ is invertible or not.

Theorem 7 ([71]). *Let $\mathsf{P}(x) = \sum_{i\geq 0} a_i \cdot x^i$ be a polynomial function over $\mathbb{Z}_{2^n}$. P is invertible (i.e., a polynomial permutation) if and only if (1st) a_1 is odd, (2nd) $\sum_{i\geq 1} a_{2i}$ is even, and (3rd) $\sum_{i\geq 1} a_{2i+1}$ is even.*

As a direct consequence, the following result holds:

Corollary 8. *The number $\#_{\mathsf{P}}(2^n)$ of distinct polynomial invertible functions (equivalently, of polyfunctions in canonical form) over $\mathbb{Z}_{2^n}$ is given by*

$$\#_{\mathsf{P}}(2^n) = \frac{2^{\mu(2^n)\cdot n-3}}{\#_{\mathsf{O}}(2^n)}\,,$$

where note that $\#_{\mathsf{O}}(2^n)$ is always a power of 2 as it is a product of gcd's of 2^n and i!.

Proof. The proof follows by a simple counting argument.

Obviously, we consider polynomials of degree smaller than $\mu(2^n)$. Hence, there are $\eta = (\mu(2^n) - 2)/2$ coefficients for even degree terms, excluding a_0. There is an equal number of coefficients for odd degree terms, excluding a_1.

The number of possible ways to choose odd degree terms (excluding a_1) is given by:

$$\begin{aligned}\sum_{i=0}^{\frac{\eta-1}{2}} \binom{\eta}{2i} \cdot (2^{n-1})^{2i} \cdot (2^{n-1})^{\eta-2i} &= 2^{\eta\cdot(n-1)} \cdot \sum_{i=0}^{\frac{\eta-1}{2}} \binom{\eta}{2i} = 2^{\eta\cdot(n-1)} \cdot \frac{\sum_{i=0}^{\eta}\binom{\eta}{i}}{2} \\ &= 2^{\eta\cdot(n-1)} \cdot 2^{\eta-1} = 2^{\eta\cdot n-1}\,,\end{aligned}$$

due to the symmetry of the binomial coefficient and since η is odd. Similarly, the number of ways to choose odd degree coefficients is the same, that is, $2^{\eta\cdot n-1}$. The number of ways of choosing a_0 is 2^n and, since a_1 must be odd, it can be chosen 2^{n-1} different ways.

By combining everything together, we end up with the following

$$2^n \cdot 2^{n-1} \cdot (2^{\eta\cdot n-1})^2 = 2^{2\eta\cdot n+2n-3} = 2^{\mu(2^n)\cdot n-3}\,.$$

The claim follows immediately. □

Corollary 9. $\#_{\mathsf{F}}(2^n) = 8 \cdot \#_{\mathsf{P}}(2^n)$.

To the best of our knowledge, our result is the first closed formula for $\mathbb{Z}_{2^n}$. A comparison of the number of functions, polyfunctions, and invertible polyfunctions over $\mathbb{Z}_{2^n}$ is given in Table 1. As it is possible to observe, $\#_{\mathsf{P}}(2^3) = \#_{\mathsf{IC}}(2^3)$, which implies that every invertible function over $\mathbb{Z}_{2^3}$ that preserves the modulo congruences is also a polynomial invertible function. This is not the case anymore for $n \geq 4$.

Table 1: Comparison between the number of null polynomials $\#_{\mathsf{O}}(2^n)$ of degree strictly less than $\mu(2^n)$, permutation polyfunctions $\#_{\mathsf{P}}(2^n)$ (in canonical form), polyfunctions $\#_{\mathsf{F}}(2^n)$ (in canonical form), (invertible) congruence-preserving functions $\#_{\mathsf{C}}(2^n)$ and $\#_{\mathsf{IC}}(2^n)$, and invertible functions $\#_{\mathsf{I}}(2^n) = 2^n!$ over $\mathbb{Z}_{2^n}$. Note that the null polynomials can have degree larger than $\mu(2^n)$. We limit ourselves to recall that the total number of functions is given by $2^{n \cdot 2^n}$.

n	$\#_{\mathsf{O}}(2^n)$	$\#_{\mathsf{P}}(2^n)$	$\#_{\mathsf{F}}(2^n)$	$\#_{\mathsf{IC}}(2^n)$	$\#_{\mathsf{C}}(2^n)$	$\#_{\mathsf{I}}(2^n)$
3	2^2	2^7	2^{10}	2^7	2^{14}	$2^{15.3}$
4	2^8	2^{13}	2^{16}	2^{15}	2^{30}	$2^{44.3}$
5	2^{16}	2^{21}	2^{24}	2^{31}	2^{62}	$2^{117.7}$
6	2^{16}	2^{29}	2^{32}	2^{63}	2^{126}	$2^{296.0}$
7	2^{16}	2^{37}	2^{40}	2^{127}	2^{254}	$2^{716.2}$
8	2^{30}	2^{47}	2^{50}	2^{255}	2^{510}	$2^{1684.0}$
16	2^{118}	2^{167}	2^{170}	$2^{65\,535}$	$2^{131\,070}$	$2^{954\,036.9}$
32	2^{478}	2^{607}	2^{610}	$2^{4\,294\,967\,295}$	$2^{8\,589\,934\,590}$	$2^{131\,243\,000\,000.0}$

A.7 Number of Congruence-Preserving Functions

We estimate the number of functions that preserve the congruences.

Lemma 8. *Let $\#_{\mathsf{C}}(p^n)$ be the number of functions over $\mathbb{Z}_{p^n}$ that preserve the congruences mod p^i. Let $\#_{\mathsf{IC}}(p^n)$ be the number of invertible functions over $\mathbb{Z}_{p^n}$ that preserve the congruences mod p^i. Then:*

$$\#_{\mathsf{C}}(p^n) = p^{\frac{p\cdot(p^n-1)}{p-1}}, \qquad \#_{\mathsf{IC}}(p^n) = (p!)^{\frac{p^n-1}{p-1}}.$$

Proof. Let F be a function over $\mathbb{Z}_{p^n}$. Working mod p, we have p^p choices for the values $\mathsf{F}(x) \bmod p$ for each $x \in \{0, 1, \ldots, p-1\} \subseteq \mathbb{Z}_{p^n}$. Note that this fixes the values of $\mathsf{F}(y) \bmod p$ for each $y \in \mathbb{Z}_{p^n}$.

Working mod p^2, we have p^{p^2} choices for the values $\mathsf{F}(x) \bmod p^2$ for each $x \in \{0, 1, \ldots, p^2-1\} \subseteq \mathbb{Z}_{p^n}$. Indeed, note that $\mathsf{F}(x) \bmod p^2 = x_1 \cdot p + x_0$ for a certain $x_1, x_0 \in \{0, 1, \ldots, p-1\} \subseteq \mathbb{Z}_{p^n}$. However, x_0 is already fixed, due to the previous case.

By iterating this process, we have

$$\#_{\mathsf{C}}(p^n) = \prod_{i=1}^{n} p^{p^i} = p^{\sum_{i=1}^{n} p^i} = p^{\sum_{i=0}^{n} p^i - 1} = p^{\frac{p^{n+1}-1}{p-1}-1} = p^{\frac{p\cdot(p^n-1)}{p-1}}.$$

An analogous proof holds for the case in which the function is invertible. □

B Details of Section 3

B.1 SPN Cipher

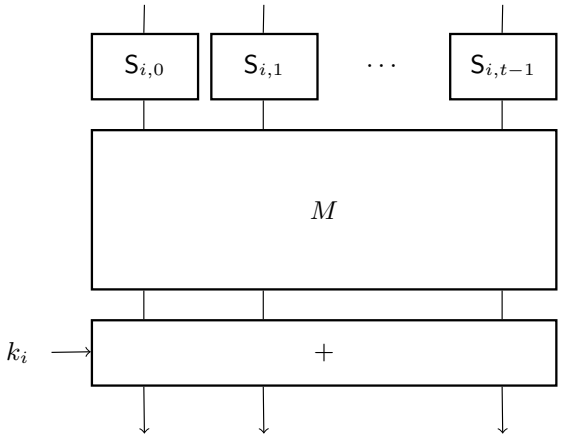


Fig. 2: The i-th round function R_{k_i} of the iterated SPN cipher over $\mathbb{Z}_{p^n}^t$.

B.2 Proof of Lemma 1

Assume M to be invertible. Then, $\det(M) \neq 0 \mod p^n$, and $\gcd(\det(M), p^n) = 1$. It is obvious that $\gcd(\det(M), p^n) = \gcd(\det(M), p) = 1$, which implies that $\det(M) \bmod p \neq 0$ as well. Since $M' = M \mod p$, it follows that $\det(M') = \det(M) \mod p$, which implies M' to be invertible.

Vice-versa, assume M' to be invertible. It follows that $\det(M') \neq 0 \mod p$ and $\gcd(\det(M'), p) = 1$. Since $M' = M \mod p$, it follows that $\det(M) = \det(M') + p \cdot a$ for a certain $a \in \mathbb{Z}_{p^n}$. Hence, since $\det(M') \neq 0 \mod p$, then $(\det(M') + p \cdot a) \mod p^n \neq 0$ as well. Similarly, $\gcd(\det(M), p^n) = \gcd(\det(M') + p \cdot a, p^n) = 1$ as $\gcd(\det(M'), p) = 1$. The claim follows immediately. □

B.3 Proof of Theorem 3

As the proofs of the linear and differential versions of Theorem 3 are nearly identical, we limit ourselves to the differential case.

Since the second linear layer does not affect the probability of differentials, we focus on the composition of an S-box layer, the linear layer, and a second S-box layer. The key-averaged probability of a differential trail $(\alpha, \beta, \gamma, \delta)$ with $\gamma = M \times \beta$ for this construction is $\sum_{\beta \in \mathbb{Z}_{p^n}^t} \prod_{i=1}^t \mathrm{DP}^{\mathsf{S}}_{\beta_i,\alpha_i} \cdot \mathrm{DP}^{\mathsf{S}}_{\delta_i,\gamma_i}$. W.l.o.g., due to the symmetry of the argument below, we can assume that only the first x coordinates of α are nonzero, and similarly only the last y coordinates of δ are

nonzero. The intermediate difference β is then restricted by the following system of linear equations:

$$\underbrace{\left[\begin{array}{c|c} M_1 & M_2 \\ \hline M_3 & M_4 \end{array}\right]}_{M} \underbrace{\left[\begin{array}{c} * \\ \vdots \\ * \\ \hline 0 \\ \vdots \\ 0 \end{array}\right]}_{\beta} = \underbrace{\left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline * \\ \vdots \\ * \end{array}\right]}_{\gamma}.$$

Here, M_1 is a $(t-y) \times x$ sub-block of M. Since the branch number of an MDS matrix is $t+1$, this system of equations only has a nontrivial solution if $x+y > t$ or equivalently $t-y < x$. After deleting $x+y-t$ columns of M_1, we obtain a square submatrix of M. By Proposition 5, this submatrix is invertible. It follows that if $\beta_1, \ldots, \beta_{x+y-t}$ are fixed to arbitrary nonzero values, then $\beta_{x+y-t+1}, \ldots, \beta_{t-y}$ and $\gamma_{x+1}, \ldots, \gamma_t$ are uniquely determined. Hence, the key-averaged probability of the differential (α, δ) is equal to

$$\sum_{\beta_1,\ldots,\beta_{x+y-t} \in \mathbb{Z}_{p^n}} \prod_{i=1}^{t} \mathrm{DP}^{\mathsf{S}}_{\beta_i,\alpha_i} \cdot \mathrm{DP}^{\mathsf{S}}_{\delta_i,\gamma_i} \leq \varepsilon^{(t-y)-(x+y-t)+(t-x)} \prod_{i=1}^{x+y-t} \underbrace{\sum_{\beta_i \in \mathbb{Z}_{p^n}} \mathrm{DP}^{\mathsf{S}}_{\beta_i,\alpha_i}}_{\leq 1}.$$

The inequality follows from the upper bound on the probability of nontrivial differentials, and because $\beta_{x+y-t+1}, \ldots, \beta_{t-y}$ and $\gamma_{x+1}, \ldots, \gamma_t$ must all be nonzero to contribute to the sum. The result follows from $(t-y)-(x+y-t)+(t-x) = t$. □

C Methods for Polynomial System Solving

This appendix provides a brief overview of methods for solving polynomial systems over a finite field $\mathbb{F}_p$, which features in Section 4.1. For ease of exposition, we restrict ourselves to the case of m polynomial equations of degree d in κ variables, where $m \geq \kappa$.

Linearization. For fixed d and κ, the problem generally gets easier with a larger m. This holds up to the point where m is about the same size as the number of possible monomials, $m \approx \binom{\kappa+d}{d}$, in which case the value of each monomial can be recovered by linear algebra methods in $\mathcal{O}\left(\binom{\kappa+d}{d}^{\omega}\right)$ $\mathbb{F}_p$-operations, where $2 \leq \omega \leq 3$ is the linear algebra constant.

Gröbner Basis Methods. For a smaller number of equations, $\kappa \leq m < \binom{\kappa+d}{d}$, the polynomial solving problem is typically dealt with through techniques

relying on Gröbner bases, or related ideas. The complexity of the state-of-the-art algorithms for computing Gröbner bases is dominated by the reduction of huge matrices [6], as was also the case for the linearization approach described above. However, the matrices are expected to be even larger in this latter setting. We limit ourselves to note that while there exists good bounds on the complexity of Gröbner basis computation for *generic* polynomial systems [8], the structured systems we work with in cryptography often fail to adhere to this notion of genericness. It follows that bounding the complexity of Gröbner basis attacks is a challenge that often requires dedicate analysis for particular ciphers.

Other Methods. Further solving methods can take advantage of a small field size. Examples range from methods that combine guessing a set of variables with algebraic procedures [7,55], to dedicated brute-force methods [17]. A different class of solving algorithms follows the *polynomial method* [65,35], which is particularly competitive when $m \approx \kappa$ (for both small and large field sizes).

D Proof of Corollary 3

By definition, $\mathrm{DP}_{\max}(\mathsf{S}) = \max_{\delta_I \in \mathbb{Z}_{p^n} \setminus \{0\}, \delta_O \in \mathbb{Z}_{p^n}} \mathrm{DP}^{\mathsf{S}}_{\delta_O, \delta_I}$. Let S'_δ be defined as in Theorem 5. Hence, it follows that

$$\begin{aligned}\mathrm{DP}_{\max}(\mathsf{S}) &\geq \max_{\delta \in \mathbb{Z}_p \setminus \{0\}, \Delta \in \mathbb{Z}_p} \mathrm{DP}^{\mathsf{S}}_{p^{n-1}\Delta,\, p^{n-1}\delta} \\ &= \max_{\delta \in \mathbb{Z}_p \setminus \{0\}, \Delta \in \mathbb{Z}_p} \frac{|\{x \in \mathbb{Z}_p \mid S'_\delta(x) = \Delta\}|}{p} \geq \frac{1}{p}\,,\end{aligned}$$

where the last inequality follows from the fact that S'_δ is a function over $\mathbb{F}_p$. □

E Proof of Proposition 9

We obtain Proposition 9 immediately from the following relation between $\mathtt{CPP}^{(\ell)}$ and $\mathtt{CE}^{(\ell)}$, simply using the fact that F is ℓ-balanced.

Lemma 9. *The relation between* $\mathtt{CPP}^{(\ell)}$ *and* $\mathtt{CE}^{(\ell)}(a,b)$ *is given by:*

$$\mathtt{CPP}^{(\ell)} = \frac{\sum_{a \in \mathbb{Z}_{p^\ell}^{t_1}} \sum_{b \in \mathbb{Z}_{p^\ell}^{t_2}} \left(\mathtt{CE}^{(\ell)}(a,b)\right)\left(\mathtt{CE}^{(\ell)}(a,b) - 1\right)}{p^{n \cdot t_1} \cdot (p^{(n-\ell)\cdot t_1} - 1)}\,, \qquad \textit{and}$$

$$\mathtt{CPP}^{(\ell)} = \frac{\sum_{x \in \mathbb{Z}_{p^n}^{t_1}} \mathtt{CE}^{(\ell)}(x \bmod p^\ell, \mathsf{F}(x) \bmod p^\ell) - p^{n \cdot t_1}}{p^{n \cdot t_1} \cdot (p^{(n-\ell)\cdot t_1} - 1)}\,.$$

We note that the first equality is more efficient to compute for small ℓ than the second one.

Proof. We start with the first equality. For fixed $a \in \mathbb{Z}_{p^\ell}^{t_1}$ and $b \in \mathbb{Z}_{p^\ell}^{t_2}$, the elements $x, y \in \mathbb{Z}_{p^n}^{t_1}$ counted for $\mathtt{CE}^{(\ell)}(a,b)$ are both congruent modulo p^ℓ, as $x = a = y \mod p^\ell$, and preserve the congruence, as they satisfy $\mathsf{F}(x) = b = \mathsf{F}(y) \mod p^\ell$. It is easy to see that such an element x satisfies both congruences only with elements of $\mathbb{Z}_{p^n}^{t_1}$ that are also counted for $\mathtt{CE}^{(\ell)}(a,b)$. Thus, it now suffices to count the number of possible pairs of such elements satisfying $x \neq y$.

Now, we turn to the second equality. Starting from the definition of $\mathtt{CPP}^{(\ell)}$, we can write

$$\begin{aligned}
&|\{(x,y) \in \mathcal{S}^{(\ell)} \mid \mathsf{F}(x) = \mathsf{F}(y) \mod p^\ell\}| \\
&= \sum_{x \in \mathbb{Z}_{p^n}^{t_1}} \left(|\{y \in \mathbb{Z}_{p^n}^{t_1} \setminus \{x\} \mid y = x \mod p^\ell \text{ and } \mathsf{F}(y) = \mathsf{F}(x) \mod p^\ell\}|\right) \\
&= \sum_{x \in \mathbb{Z}_{p^n}^{t_1}} \left(|\{y \in \mathbb{Z}_{p^n}^{t_1} \mid y = x \mod p^\ell \text{ and } \mathsf{F}(y) = \mathsf{F}(x) \mod p^\ell\}| - 1\right) \\
&= \sum_{x \in \mathbb{Z}_{p^n}^{t_1}} \mathtt{CE}^{(\ell)}(x \bmod p^\ell, \mathsf{F}(x) \bmod p^\ell) - p^{n \cdot t_1} \,.
\end{aligned}$$

Then, the claim follows immediately. □

F Details of Section 6.1

F.1 Proof of Proposition 10

Recall the generalized Reed-Muller code: given a finite field $\mathbb{F}_q$, $k \geq 0$ and $n \geq 1$, then the codewords of the generalized Reed-Muller code $\mathrm{RM}_q(k,n)$ is given by

$$\mathrm{RM}_q(k,n) = \{(\mathsf{F}(x))_{x \in \mathbb{F}_q^n} \mid \mathsf{F} \in \mathbb{F}_q[x_1, \ldots, x_n]_{\leq k}\}.$$

The minimum distance of this code is well known, and was first proved in [33].

Theorem 8. *Let q be a prime power and $k \geq 0$ and $n \geq 1$ integers such that $0 \leq k < n(q-1)$. Write $k = a(q-1)+b$ with $a \geq 0$ and $0 \leq b < q-1$. Then the generalized Reed-Muller code $\mathrm{RM}_q(k,n)$ has minimum distance $d(\mathrm{RM}_q(k,n)) = (q-b)q^{n-a-1}$.*

We give the following lemma to ensure that Theorem 8 applies to our case.

Lemma 10. *Let p be a prime and $n > 1$. Then*

$$0 \leq \max_{k < \mu(p^n)} \sigma_p(k) \leq n(p-1)$$

with equality in the upper bound if and only if $p = n = 2$.

Proof. The lower bound is obviously true. On the other hand, when $n > 1$ and $(p, n) \neq (2, 2)$, it is well known that $\mu(p^n) < p^n$ (see Section 2). Then $k < p^n - 1$, so its p-ary expansion can be written in at most n base-p digits. To maximize $\sigma_p(k)$ with $k < p^n - 1$ all the digits except for one must be chosen to $p-1$, which is at most $p - 2$. This yields $\sigma_p(k) < n(p - 1)$. However, when $(p, n) = (2, 2)$ we know that $\mu(4) = 4$ and hence $\max_{k<4} \sigma_p(k) = 2 = 2(2 - 1)$. □

In the following, let $\mathrm{wt}(\mathsf{F}) \coloneqq \mathrm{wt}(\mathbf{c}_\mathsf{F})$ for some function $\mathsf{F} \colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$. Finally we prove Proposition 10:

Proof. First we establish the lower bound on the minimum distance by bounding the Hamming weight of any nonzero codeword. Assume that $\mathsf{F}(x) \in \mathbb{Z}_{p^n}[x]$ is a nonzero polynomial written $\mathsf{F}(x) = \sum_{k=0}^{\mu(p^n)-1} \mathsf{F}_k x^k$. Let $e \geq 0$ be the highest integer such that $p^e \mid \mathsf{F}(x)$ for every $x \in \mathbb{Z}_{p^n}$. Define

$$\mathsf{G}(x) \coloneqq \mathsf{F}(x)/p^e \mod p^{n-e}\,.$$

While $\mathsf{G}(x)$ is integer valued, it is not necessarily a polyfunction. Observe that

$$\mathsf{G}(x) = \sum_{k=0}^{\mu(p^n)-1} \frac{\mathsf{F}_k k!}{p^e} \binom{x}{k} \mod p^{n-e}\,.$$

We show that $\mathsf{G}_k = \mathsf{F}_k k!/p^e$ is an integer for every k by the following induction argument.

Since $\mathsf{F}(0) = \mathsf{F}_0$, we have $p^e \mid \mathsf{F}_0 \cdot 0!$. Suppose $p^e \mid \mathsf{F}_j \cdot j!$ for all $j < k$. Observe that

$$\mathsf{F}(k) = \sum_{i=0}^{\mu(p^n)-1} \mathsf{F}_i i! \binom{k}{i} = \mathsf{F}_k k! + \sum_{i=0}^{k-1} \mathsf{F}_i i! \binom{k}{i} \in \mathbb{Z}\,,$$

where $p^e \mid \sum_{i=0}^{k-1} \mathsf{F}_i i! \binom{k}{i}$ by the induction hypothesis. Since $p^e \mid \mathsf{F}(k)$, also $p^e \mid \mathsf{F}_k k!$.

Finally, define $\mathsf{H}(x) = \mathsf{G}(x) \mod p$. Note that $\mathsf{H}(x) \neq 0 \implies \mathsf{F}(x) \neq 0$. Furthermore, by the maximality of e above, there is some x such that $\mathsf{G}(x) \neq 0 \mod p$. Hence we have $0 < \mathrm{wt}(\mathsf{H}) \leq \mathrm{wt}(\mathsf{F})$. Using Lucas' theorem, we get

$$\mathsf{H}(x) = \sum_{k=0}^{\mu(p^n)-1} (\mathsf{G}_k \bmod p) \left[\prod_{i=0}^{n-1} \binom{x_i}{k_i} \bmod p \right] \mod p$$

where $x = \sum_{i=0}^{n-1} x_i p^i$ and $k = \sum_{i=0}^{n-1} k_i p^i$, and we may in fact view $\mathsf{H}(x)$ as a function $\mathbb{F}_p^n \to \mathbb{F}_p$, specifically as a polynomial over $\mathbb{F}_p[x_0, \ldots, x_{n-1}]$, where the p-digits of x play the role of the indeterminates. Note that $\binom{x_i}{k_i} \neq 0$ if and only if $k_i \leq x_i$, in which case $\binom{x_i}{k_i}$ is a degree k_i polynomial. For any x, k, if $k_i \leq x_i$ for all i, the degree of the polynomial $\prod \binom{x_i}{k_i}$ is equal to the sum of the p-digits of k; i.e., $\deg \prod_i \binom{x_i}{k_i} = \sigma_p(k)$. Hence the degree $\mathsf{H}(x)$ is upper bounded by $m = \max_{0 \leq k < \mu(p^n)} \sigma_p(k)$. Furthermore, since $\mathsf{H}(x)$ is just a linear combination

of polynomials in $\mathbb{F}_p[x_0, \ldots, x_{n-1}]$, we have that $\mathsf{H}(x) \in \mathbb{F}_p[x_0, \ldots, x_{n-1}]$. Thus, it constitutes a codeword in the generalized Reed-Muller code $\mathrm{RM}_p(m, n)$ where $m < n(p-1)$ by Lemma 10, when $(p, n) \neq (2, 2)$ (we consider this case at the end of the proof). Using Theorem 8, we may write $m = a(p-1) + b$ for $0 \leq b < p-1$, it has minimum distance $d_{\min} = (p-b)p^{n-a-1}$, implying that $\mathrm{wt}(\mathsf{F}) \geq \mathrm{wt}(\mathsf{H}) \geq (p-b)p^{n-a-1}$.

To prove the upper bound, we construct a polynomial which achieves the bound. Define $\mathsf{M}_k(x) = p^{n-v_p(k!)-1}(x)_k$. Then observe that

$$\mathsf{M}_k(x) = p^{n-1}k!\binom{x}{k}$$

which is zero if and only if $\binom{x}{k} = 0 \mod p$. I.e, whenever $\binom{x}{k} \equiv \prod_{i=0}^{n-1}\binom{x_i}{k_i} \equiv 0 \mod p$ where $x = \sum_{i=0}^{n-1} x_i p^i$ and $k = \sum_{i=0}^{n-1} k_i p^i$. The weight of M_k is then given by $\mathrm{wt}(\mathsf{M}_k) = \prod_{i=0}^{n-1}(p - k_i)$. It is minimized when the number of base-p digits of k chosen to be $p-1$ is maximized, and any possible remaining digit is chosen to be $k \bmod (p-1)$. In effect, $\mathrm{wt}(\mathsf{M}_k) = (p-(p-1))^a \cdot (p-b) \cdot p^{n-a-1}$, so M_k indeed reaches the lower bound.

Finally, when $p = n = 2$ we get an upper bound on minimum distance of 1 which in turn has to be the true minimum distance of $\mathcal{C}_{2^2}$. This is in accordance with the provided formula. □

F.2 Overview of the Code Parameters for $p = 2$ and various n

Table 2: Parameters of the $\mu(2^n)$-dimensional code $\mathcal{C}_{2^n}$ for various n, where $|\mathcal{C}_{2^n}|$ is the number of distinct codewords, t is number of errors that can be uniquely corrected and $\rho_{\min}$ a lower bound for the covering radius. Finally, $\mathtt{PC}_{\min}$ is the lowest polynomial correlation for which we *always* can recover the unique OPA.

n	$k = \mu(2^n)$	$\lvert\mathcal{C}_{2^n}\rvert$	$d_{\min}$	t	$\rho_{\min}$	$\mathtt{PC}_{\min}$
4	6	2^{16}	4	1	12	0.94
8	10	2^{50}	32	15	240	0.94
16	18	2^{170}	4096	2047	65280	0.97
32	34	2^{610}	2^{27}	$2^{26}-1$	$2^{31.99998}$	0.98
64	66	2^{2274}	2^{58}	$2^{57}-1$	$2^{63.9999999997}$	0.99

F.3 Decoding Algorithms From Section 6.1

Algorithm 1 Algorithm Finding a Polynomial Approximation

Input: Parameters p, n and a function $\mathsf{F}\colon \mathbb{Z}_{p^n} \to \mathbb{Z}_{p^n}$, and a bound for tolerable error t.

Output: a polyfunction $\hat{\mathsf{F}}(x) \in \mathbb{Z}_{p^n}[x]$ having $d(\mathbf{c}_{\hat{\mathsf{F}}}, \mathbf{c}_{\mathsf{F}}) \leq t$ or $\perp$.

for r in $0, 1, \ldots, p^n - 1$ **do**
  $\mathsf{G}(x \bmod p^n) \leftarrow \mathsf{F}((x + r) \bmod p^n)$ for all x
  $\hat{\mathsf{G}} \leftarrow \text{interpolate}(\mathsf{G}, p, n)$
  if $\hat{\mathsf{G}} \neq \perp$ **then**
    $\Delta \leftarrow d(\mathbf{c}_{\mathsf{G}}, \mathbf{c}_{\hat{\mathsf{G}}})$
    if $\Delta \leq t$ **then**
      $\mathsf{H}(x) \leftarrow \hat{\mathsf{G}}((x - r) \bmod p^n)$ for all x
      return H
    end if
  end if
end for
return $\perp$

Here, interpolate(F, p, n) is a black-box algorithm interpolating F in its initial $\mu(p^n)$ points, returning a polyfunction $\hat{\mathsf{F}}$ if successful and $\perp$ otherwise.

Complexity of Decoding. Interpolation requires the computation of the $\mu(p^n)$ coefficients $\Delta^k(\mathsf{F}(x))|_{x=0}$ for $k = 0, 1, \ldots, \mu(p^n) - 1$, resulting in $1 + 2 + \cdots + \mu(p^n) = \mathcal{O}(\mu(p^n)^2)$ ring operations. Each time this process is repeated, the Hamming distance between $\mathbf{c}_{\mathsf{F}}$ and the candidate for the decoded word is computed. This is the comparison of p^n points of two functions. Hence, for each r, we perform at most $\mathcal{O}(\mu(p^n)^2 + p^n)$ operations. Since r runs through p^n values, the worst case complexity becomes $\mathcal{O}(\mu(p^n)^2 p^n + p^{2n})$.

The complexity can be improved precomputing all forward differences of F, which has complexity $\mathcal{O}(\mu(p^n)p^n)$, reducing the complexity of the main algorithm to $\mathcal{O}(\mu(p^n)p^n + p^{2n})$, but increasing the memory cost.

Finally, for generic cases, the process of computing the Hamming distance of the input and the candidate OPA can be terminated after about t comparisons since having the wrong candidate OPA makes it unlikely that it agrees with the input function on many points, where $t = \lfloor d_{\min} - 1 \rfloor / 2$, giving an average complexity estimate of $\mathcal{O}(t\mu(p^n) + tp^n)$.

Conclusion. To conclude, we prove the following.

Proposition 15. *The algorithm will always terminate if*

$$\mathrm{wt}(\mathbf{c}_{\mathsf{E}}) < \frac{p^n - \mu(p^n) + 1}{\mu(p^n)}. \tag{6}$$

Proof. We require a *window* of $\mu(p^n)$ consecutive points of agreement between $\hat{\mathsf{F}}$ and F, so we require the ceil of the average window length to be greater than or equal to $\mu(p^n)$. Hence, we solve $\left\lceil \frac{p^n - \mathrm{wt}(\mathbf{c}_\mathsf{E})}{\mathrm{wt}(\mathbf{c}_\mathsf{E})+1} \right\rceil \geq \mu(p^n)$ for $\mathrm{wt}(\mathbf{c}_\mathsf{E})$. □

We remark that, provided $\mathrm{wt}(\mathbf{c}_\mathsf{E}) \leq t = \lfloor d_{\min} - 1 \rfloor / 2$, then calculation for various $n > 1$ indicate that Eq. (6) holds for $p = 2, 3$ and any n. For larger primes, it holds only for certain n. Furthermore, unless the function in interest is designed with *errors* that are uniformly distributed, there is a very high chance that the algorithm terminates with success:

Example 6. Suppose F is a function on $\mathbb{Z}_{2^{16}}$ with $\mathtt{PC}_{\max} = 60\%$ whose OPA is $\tilde{\mathsf{F}}$ such that $\mathsf{F} = \tilde{\mathsf{F}} + \mathsf{E}$, where E is nonzero in $0.40 \cdot 2^{16}$ inputs, which are chosen uniformly at random. Then one can compute the probability of the existence of a length $\mu(2^{16}) = 18$ window to be $\approx 93\%$. Increasing $\mathtt{PC}_{\max} = 65\%$ makes the probability of the existence of such a practically equal to 1. In effect, the algorithm in discussion would fail to recover $\tilde{\mathsf{F}}$ with astronomically low probability.

Experimental verification. The algorithm was tested for $p = 2$ on a MacBook Pro M2, and the cases $n = 4, \ldots, 16$ all terminated in seconds.

Practical Use Cases. Because our algorithm requires a relatively high polynomial correlation to guarantee decoding, it is not suitable for approximating completely random or well-balanced functions. At the same time, the S-box of a given cipher may have been designed to appear random despite having a precise algebraic structure. In e.g. [68], Perrin identified nontrivial algebraic substructures of the S-boxes of Streebog and Kuznyechik, suggesting intentional design choices. Analogously, one may encounter a function over $\mathbb{Z}_{p^n}^t$ which is based one or multiple polyfunctions, even though its truth table appears random. The function(s) on which it was based may be recoverable by the algorithm presented in this section, or modifications of it.

G Polynomial Representations over $\mathbb{Q}_p$

In Section 6, we have focused on polynomials over $\mathbb{Z}_{p^\ell}$, but one may also wonder whether there are other mathematical structures that are more suitable to work with. A particularly interesting candidate may be the field of *p-adic numbers*, $\mathbb{Q}_p$, due to its connections with Diophantine equations. Indeed, we recall that $\mathbb{Z}$ is a subring of $\mathfrak{I}_p$, the *p-adic integers*[4] (which, in turn, is a subring of $\mathbb{Q}_p$), hence we can view $\mathbb{Z}_{p^\ell}$ as the quotient of a particular substructure of $\mathbb{Q}_p$.

More concretely, we recall that elements in the $\mathfrak{I}_p$ can be viewed as formal power series of the form $\sum_{i=0}^{\infty} a_i p^i$ where each a_i is an integer in $\{0, 1, \ldots, p-1\}$, and arithmetic is performed in the usual manner for power series. Elements of

[4] Works in number theory tend to use $\mathbb{Z}_p$ for denoting the p-adic integers. We adopt the non-standard notation $\mathfrak{I}_p$ to avoid confusion with the integers modulo p.

$\mathbb{Q}_p$ can be seen by expanding the series to also allow a finite number of terms $a_j \cdot p^{-j}$ where $-j$ is a negative integer.

As $\mathbb{Q}_p$ is a field, it is always possible to interpolate our finite-valued function to a polynomial function in $\mathbb{Q}_p[x]$ (e.g., through Newton interpolation). Two scenarios are possible:

1. polynomials with coefficients in $\mathfrak{I}_p$, and
2. polynomial with some coefficients in $\mathbb{Q}_p \setminus \mathfrak{I}_p$.

The first case is not really interesting for us, as working with polynomials over $\mathfrak{I}_p$ does not give us anything "new" that we cannot find when working over $\mathbb{Z}_{p^\ell}$, in the sense of the following lemma.

Lemma 11. *Let $\mathfrak{I}_p$ denote the p-adic integers, and consider a polynomial $\mathfrak{f} \in \mathfrak{I}_p[x]$. Then for any integer $\ell \geq 1$, there exists a polynomial $\mathsf{F} \in \mathbb{Z}_{p^\ell}[x]$ such that $\mathfrak{f}$ and F agree modulo p^ℓ when evaluated on any integer.*

The proof is trivial, and it follows on the fact that $\mathsf{F}(x) = \mathfrak{f}(x) \mod p^\ell$.

More interesting for us is the second case: even if a cipher over $\mathbb{Z}_{p^n}$ cannot be represented as a polynomial with coefficients in $\mathfrak{I}_p$, it can still be represented as a polynomial in $\mathbb{Q}_p[x]$ (with at least one coefficient in $\mathbb{Q}_p \setminus \mathfrak{I}_p$). In this case we have to be careful in our handling of the congruence classes of $\mathbb{Z}_{p^n}$, as the resulting polynomial will map different integers of the same residue class to different elements modulo p^n. As an example, consider the function G over $\mathbb{Z}_4$ whose polynomial representation over $\mathbb{Q}_2$ is

$$\mathsf{G}(x) = \frac{2 \cdot x^3 - x^2 + 3 \cdot x}{2},$$

which results in $\mathsf{G}(x) = \{0, 1, 2, 3\}$ for $x \in \mathbb{Z}_4$. It is easy to check that $\mathsf{G}(4) = 2 \neq 0 = \mathsf{G}(0)$.

A possible way to deal with it consists in identifying each congruence class in $\mathbb{Z}_{p^n}$ with its unique representative in $\{0, \ldots, p^n - 1\} \subset \mathbb{Q}_p$ and proceed from there. Although the resulting polynomial in $\mathbb{Q}_p$ could, in principal, be used for algebraic attacks, we have no guarantee that it is of low degree. On the contrary, since this interpolated polynomial no longer respects the usual arithmetic in $\mathbb{Z}_{p^n}$, the linear layer and key addition must also be interpolated as polynomials that are generally non-linear. As such, the resulting polynomial system for the entire cipher will likely have too high degree to be useful.

Remark 8. Finally, we remark that the argument proposed for this second case does not exploit any crucial property of $\mathbb{Q}_p$ beyond the fact that it is a field. Hence we expect a similar reasoning (and conclusion) to hold when interpolating over other fields, e.g., $\mathbb{Q}$ or $\mathbb{F}_{p^n}$.

H Alternative Attack Strategy from Section 6.2

As discussed in Section 6.2, the following attack can be set up

1. the attacker first finds a polynomial approximation of the non-polynomial functions modulo p^n (note that they are independent of the secret key);
2. the attacker sets up the system of equations modulo p^n (3) of the approximated cipher, by replacing the non-polynomial functions with their corresponding polynomial approximation;
3. the attacker solves the system of equations for one plaintext/ciphertext pair, and tests the found value with other plaintext/ciphertext pairs, until the secret key is found.

Given the estimated probability of the polynomial approximation mod p^ℓ, we now derive the data complexity of this attack strategy, assuming $\kappa = t$ for simplicity.

Probabilities of the Right versus the Wrong Key Candidate. Let $k \in \mathcal{K}$ be the secret key, and consider a plaintext/ciphertext pair $(x, y = \mathsf{C}_k(x))$. We assume that the system of polynomial equations given by $\mathsf{A}(x, \tilde{k}) = y$ has only one solution and that when $\mathsf{A}(x, k) \neq y$, then any other key candidate $\tilde{k}$ satisfies the system with probability $1/(p^{n \cdot t} - 1)$. We point out that this is a realistic assumption, as for the real cipher, we should obtain a random system of polynomial equations.

Given ρ_n and its lower bound $\rho_n^{\min}$ as before, the following scenario occurs:

– the secret key satisfies the system of equations with probability ω_n given by

$$\rho_n \geq \rho_n^{\min} =: \omega_n^{\min} ,$$

as the right key satisfies the system only when $\mathsf{A}(x, k) = \mathsf{C}_k(x) \bmod p^n$;
– any other key satisfies the system with probability ζ_n given by

$$\rho_n \cdot 0 + (1 - \rho_n) \cdot \frac{1}{p^{n \cdot t} - 1} = (1 - \rho_n) \cdot \frac{1}{p^{n \cdot t} - 1} \leq (1 - \rho_n^{\min}) \cdot \frac{1}{p^{n \cdot t} - 1} .$$

Data Complexity. For this strategy, we solve systems of equations, each given by a plaintext/ciphertext pair, until we obtain the correct secret key k. Thus, for a success probability of the attack $\Omega^{(S)}$ using N plaintext/ciphertext pairs, we have

$$\begin{aligned} \Omega^{(S)} &\geq 1 - (\mathrm{Prob}(\mathsf{C}_k(x) \neq \mathsf{A}(x, k)))^N \\ &= 1 - (1 - \omega_n)^N \approx 1 - e^{-N \cdot \omega_n}. \end{aligned}$$

As $\omega_n \geq \omega_n^{\min}$, we obtain $\Omega^{(S)} \geq 1 - e^{-N \cdot \omega_n^{\min}}$. It follows that $N \geq \frac{3}{\omega_n^{\min}}$ for $\Omega^{(S)} \geq 0.95$.

I Proof of Proposition 11

We prove the result by induction. We will use the notation $x \equiv_{p^\ell} y$ to denote $(x-y) \bmod p^\ell = 0$. We also denote the condition $x \equiv_{p^\ell} y \ \wedge \ x \neq y$ by CO and the value $\frac{p^{(n-\ell)\cdot t}-1}{p^{n\cdot t}-1}$ by ρ_π. Starting over 2 rounds, we have that

$$\begin{aligned}
&\mathrm{Prob}\left(\mathsf{R}^2(x) \equiv_{p^\ell} \mathsf{R}^2(y) \mid \mathsf{CO}\right) \\
=&\mathrm{Prob}\left(\mathsf{R}^2(x) \equiv_{p^\ell} \mathsf{R}^2(y) \mid \mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \ \wedge \ \mathsf{CO}\right) \cdot \mathrm{Prob}\left(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid \mathsf{CO}\right) \\
&+ \mathrm{Prob}\left(\mathsf{R}^2(x) \equiv_{p^\ell} \mathsf{R}^2(y) \mid \mathsf{R}(x) \not\equiv_{p^\ell} \mathsf{R}(y) \ \wedge \ \mathsf{CO}\right) \cdot \mathrm{Prob}\left(\mathsf{R}(x) \not\equiv_{p^\ell} \mathsf{R}(y) \mid \mathsf{CO}\right).
\end{aligned}$$

We now compute all four probabilities. Using Markov cipher assumption, independence of the rounds and the fact that R is invertible, the first probability is equal to $\rho_\pi + \varepsilon$. By definition, the second and fourth probabilities are equal to $\rho_\pi + \varepsilon$ and $(1-(\rho_\pi+\varepsilon))$, respectively. To continue, note that

$$\begin{aligned}
&\mathrm{Prob}(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \equiv_{p^\ell} y) \\
=&\mathrm{Prob}(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \equiv_{p^\ell} y \ \wedge \ x \neq y) \cdot \mathrm{Prob}(x \neq y \mid x \equiv_{p^\ell} y) \\
&+ \mathrm{Prob}(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \equiv_{p^\ell} y \ \wedge \ x = y) \cdot \mathrm{Prob}(x = y \mid x \equiv_{p^\ell} y) \\
=&\,(\rho_\pi + \varepsilon) \cdot \left(1 - \frac{p^{\ell\cdot t}}{p^{n\cdot t}}\right) + \frac{p^{\ell\cdot t}}{p^{n\cdot t}}.
\end{aligned}$$

It remains to compute the following

$$\begin{aligned}
&\mathrm{Prob}\left(\mathsf{R}^2(x) \equiv_{p^\ell} \mathsf{R}^2(y) \mid \mathsf{R}(x) \not\equiv_{p^\ell} \mathsf{R}(y) \ \wedge \ x \equiv_{p^\ell} y \ \wedge \ x \neq y\right) \\
=&\mathrm{Prob}\left(\mathsf{R}^2(x) \equiv_{p^\ell} \mathsf{R}^2(y) \mid \mathsf{R}(x) \not\equiv_{p^\ell} \mathsf{R}(y) \ \wedge \ \mathsf{R}(x) \neq \mathsf{R}(y)\right) \\
=&\mathrm{Prob}\left(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \not\equiv_{p^\ell} y \ \wedge \ x \neq y\right) \\
=&\mathrm{Prob}\left(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \not\equiv_{p^\ell} y\right) \\
=&\frac{\mathrm{Prob}(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y)) - \mathrm{Prob}(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \equiv_{p^\ell} y) \cdot \mathrm{Prob}(x \equiv_{p^\ell} y)}{\mathrm{Prob}(x \not\equiv_{p^\ell} y)} \\
=&\frac{1/p^{\ell\cdot t} - \mathrm{Prob}(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \equiv_{p^\ell} y) \cdot 1/p^{\ell\cdot t}}{(p^{\ell\cdot t}-1)/p^{\ell\cdot t}} \\
=&\rho_\pi - \varepsilon \cdot \frac{p^{n\cdot t} - p^{\ell\cdot t}}{p^{n\cdot t} \cdot (p^{\ell\cdot t}-1)}.
\end{aligned}$$

Finally, we obtain for two rounds

$$\begin{aligned}
&\mathrm{Prob}\left(\mathsf{R}^2(x) \equiv_{p^\ell} \mathsf{R}^2(y) \mid \mathsf{CO}\right) \\
=&\,(\rho_\pi + \varepsilon)^2 + \left(\rho_\pi - \varepsilon \cdot \frac{p^{n\cdot t} - p^{\ell\cdot t}}{p^{n\cdot t} \cdot (p^{\ell\cdot t}-1)}\right) \cdot (1 - (\rho_\pi + \varepsilon)) \\
=&\,\rho_\pi + \left(1 + \frac{1}{p^{\ell\cdot t}-1}\right) \cdot \left(1 - \frac{1}{p^{n\cdot t}}\right).
\end{aligned}$$

Let us denote the condition $\mathsf{R}^{r-1}(x) \equiv_{p^\ell} \mathsf{R}^{r-1}(y)$ by EQ^{r-1}. Working over r rounds, we have that

$$
\begin{aligned}
&\mathrm{Prob}\left(\mathsf{R}^r(x) \equiv_{p^\ell} \mathsf{R}^r(y) \mid \mathsf{CO}\right) \\
=&\mathrm{Prob}\left(\mathsf{R}^r(x) \equiv_{p^\ell} \mathsf{R}^r(y) \mid \mathsf{EQ}^{r-1} \wedge \mathsf{CO}\right) \cdot \mathrm{Prob}\left(\mathsf{EQ}^{r-1} \mid \mathsf{CO}\right) \\
&+ \mathrm{Prob}\left(\mathsf{R}^r(x) \equiv_{p^\ell} \mathsf{R}^r(y) \mid \neg\mathsf{EQ}^{r-1} \wedge \mathsf{CO}\right) \cdot \mathrm{Prob}\left(\neg\mathsf{EQ}^{r-1} \mid \mathsf{CO}\right) \\
=&\mathrm{Prob}\left(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid \mathsf{CO}\right) \cdot \mathrm{Prob}\left(\mathsf{EQ}^{r-1} \mid \mathsf{CO}\right) \\
&+ \mathrm{Prob}\left(\mathsf{R}(x) \equiv_{p^\ell} \mathsf{R}(y) \mid x \not\equiv_{p^\ell} y \wedge x \neq y\right) \cdot \left(1 - \mathrm{Prob}\left(\mathsf{EQ}^{r-1} \mid \mathsf{CO}\right)\right) \\
=&\left(\rho_\pi + \varepsilon\right) \cdot \left(\rho_\pi + \varepsilon^{r-1} \cdot \left(\left(1 + \frac{1}{p^{\ell \cdot t} - 1}\right) \cdot \left(1 - \frac{1}{p^{n \cdot t}}\right)\right)^{r-2}\right) \\
&+ \left(\rho_\pi - \varepsilon \cdot \frac{p^{n \cdot t} - p^{\ell \cdot t}}{p^{n \cdot t} \cdot (p^{\ell \cdot t} - 1)}\right) \\
&\cdot \left(1 - \rho_\pi - \varepsilon^{r-1} \cdot \left(\left(1 + \frac{1}{p^{\ell \cdot t} - 1}\right) \cdot \left(1 - \frac{1}{p^{n \cdot t}}\right)\right)^{r-2}\right) \\
=&\rho_\pi + \varepsilon^r \cdot \left(1 + \frac{1}{p^{\ell \cdot t} - 1}\right)^{r-1} \cdot \left(1 - \frac{1}{p^{n \cdot t}}\right)^{r-1} .
\end{aligned}
$$

The statements then follows from the fact that

$$
\frac{p^{(n-\ell) \cdot t} - 1}{p^{n \cdot t} - 1} = \left(\frac{1}{p^{\ell \cdot t}} - \frac{1}{p^{n \cdot t}}\right) \cdot \left(1 - \frac{1}{p^{n \cdot t}}\right)^{-1} .
$$

J Constructing/Counting Well-Balanced Permutations over $\mathbb{Z}^t_{p^n}$

Proposition 12 and Lemma 4 suggest a constructive method for building permutations over $\mathbb{Z}^t_{p^n}$ that are well-balanced:

- we can first construct F_1 by defining $\mathsf{F}_1(a, \cdot)$ for each $a \in \mathbb{Z}^t_{p^{\ell_1}}$ according to Proposition 12. Concretely, this amounts to choosing $p^{t \cdot \ell_1}$ permutations over $\mathbb{Z}^t_{p^{\ell_2}}$ when n is even, and $p^{t \cdot \ell_1}$ p^t-to-1 functions when n is odd.
- then, we construct F_2 according to Lemma 4 for each $c \in \mathbb{Z}^t_{p^{\ell_1}}$. It follows from the proof of this lemma that we may identify each preimage $\mathsf{F}_1^{-1}(c)$ with $\mathbb{Z}^t_{p^{\ell_2}}$. Hence, the construction of F_2 can be realized by choosing $p^{t \cdot \ell_1}$ permutations over $\mathbb{Z}^t_{p^{\ell_2}}$.

We summarize this construction in the following statement.

Proposition 16. *Consider positive integers $n, , t$, $\ell_1 = \lfloor n/2 \rfloor$, $\ell_2 = \lceil n/2 \rceil$, and a positive prime p. Then constructing a well-balanced permutation $\mathsf{F} : \mathbb{Z}^t_{p^n} \to \mathbb{Z}^t_{p^n}$ is equivalent to constructing:*

1) $2p^{t \cdot \ell_1}$ permutations over $\mathbb{Z}^t_{p^{\ell}_1}$ if n is even.

2) $p^{t\cdot\ell_1}$ *permutations over* $\mathbb{Z}_{p^{\ell_2}}^t$ *and* $p^{t\cdot\ell_1}$ p^t*-to-1 functions* $\mathbb{Z}_{p^{\ell_2}}^t \to \mathbb{Z}_{p^{\ell_1}}^t$ *if* n *is odd.*

This result immediately provides us with an easy method for counting the number of well-balanced permutations, as given in Corollary 10.

Corollary 10. *The number of permutations over* $\mathbb{Z}_{p^n}^t$ *that are well-balanced is*

i) $\left(p^{t\cdot\ell_1}!\right)^{2p^{t\cdot\ell_1}}$ *if* n *is even,*

ii) $\left(\left(p^{t\cdot\ell_2}!\right)\prod_{i=0}^{p^{t\cdot\ell_1}-1}\binom{p^{t\cdot\ell_2}-ip^t}{p^{t\cdot\ell_1}}\right)^{p^{t\cdot\ell_1}}$ *if* n *is odd.*

Proof. The first statement follows directly from Proposition 16 1) by counting the ways of choosing $2p^{t\cdot\ell_1}$ permutations over $\mathbb{Z}_{p_1^\ell}^t$.

When n is odd we also have to consider p^t-to-1 mappings $\mathbb{Z}_{p^{\ell_2}}^t \to \mathbb{Z}_{p^{\ell_1}}^t$. To count the number of these functions, we can imagine that the first element in the codomain will have $\binom{p^{t\cdot\ell_2}}{p^{t\cdot\ell_1}}$ choices of valid preimages. The second element will now have $\binom{p^{\ell_2}-p}{p^{\ell_1}}$ choices for preimages, and so on. This results in

$$\underbrace{\binom{p^{t\cdot\ell_2}}{p^{t\cdot\ell_1}}\binom{p^{t\cdot\ell_2}-p^t}{p^{t\cdot\ell_1}}\cdots\binom{p^{t\cdot\ell_1}}{p^{t\cdot\ell_1}}}_{p^{t\cdot\ell_1}} = \prod_{i=0}^{p^{t\cdot\ell_1}-1}\binom{p^{t\cdot\ell_2}-ip^t}{p^{t\cdot\ell_1}}$$

possible p^t-to-1 mappings. Counting these, along with the permutations listed in Proposition 16 2) yields the result. □

K Statistical Properties of Well-Balanced Permutations over $\mathbb{Z}_{16}$

We provide the results of the search outlined in Section 9.2. In the following, $N \approx 2^{36}$ is the number of well-balanced permutations over $\mathbb{Z}_{16}$.

For each permutation, we computed the difference distribution table of the function in $\mathbb{Z}_{2^n}$ and as a function over $\mathbb{F}_2^n$ and finally computing the algebraic degree of the function over $\mathbb{F}_2^n$. See Table 3 for a side by side comparison of the differential uniformity in $\mathbb{Z}_{16}$ and $\mathbb{F}_2^4$, respectively. They provide the total number of well-balanced permutations for every value of differential uniformity. Table 4 provides the distribution on algebraic degree on well-balanced permutations.

Finally, Table 5 provides data for various joint properties. In particular, we see that every perfectly balanced function which is 3-uniform in $\mathbb{Z}_{2^4}$ and 4-uniform $\mathbb{F}_2^4$ has algebraic degree 3. Also, a function can be 3-uniform and balanced in $\mathbb{Z}_{2^4}$ but still achieve the worst possible differential uniformity when XOR-ing. The converse is not true.

Δ_f	# of permutations	Share of N
3	4,120,657,920	3.7435%
4	46,113,730,560	41.8929%
5	41,012,219,904	37.2583%
6	14,319,329,280	13.0087%
7	3,608,223,744	3.2780%
8	727,990,272	0.6616%
9	138,960,896	0.1263%
10	28,844,032	0.0262%
11	4,980,736	0.0045%
12	376,832	0.0003%

(a) Differential uniformity in $\mathbb{Z}_{2^4}$.

Δ_f	# of permutations	Share of N
4	3,001,024,512	2.7263%
6	43,637,538,816	39.6433%
8	46,636,204,032	42.3675%
10	12,697,731,072	11.5355%
12	3,902,275,584	3.5451%
16	200,540,160	0.1823%

(b) Differential uniformity in $\mathbb{F}_2^4$.

Table 3: Number of well-balanced permutation with given differential uniformity: left table uses subtraction in $\mathbb{Z}_{16}$; right table uses XOR in $\mathbb{F}_2^4$.

Algebraic degree	# of balanced permutations	Share of N
1	147,456	0.0001%
2	136,691,712	0.1242%
3	109,938,475,008	99.8757%

Table 4: Number of well-balanced permutations with given algebraic degree over $\mathbb{F}_2^4$.

Property	# of permutations	Share of N
3-uniform (in $\mathbb{Z}_{16}$) and 4-uniform (in $\mathbb{F}_2^4$)	169,302,016	0.1538 %
and algebraic degree 1	0	0.0000 %
and algebraic degree 2	0	0.0000 %
and algebraic degree 3	169,302,016	0.1538 %
3-uniform in $\mathbb{Z}_{16}$ but $\mathbb{F}_2$-linear	4,397,312	0.0040 %
4-uniform in $\mathbb{F}_2$ but $\mathbb{Z}_{16}$-linear	0	0.0000 %

Table 5: Various joint properties. In this context $\mathbb{F}_2$ and $\mathbb{Z}_{16}$-linear (resp.) means that the differential uniformity of the function is the highest achievable over $\mathbb{F}_2$ ($\mathbb{Z}_{16}$ resp.).