Invertible Quadratic Non-Linear Layers for MPC-/FHE-/ZK-Friendly Schemes over $\mathbb{F}_p^n$

Application to Poseidon

Lorenzo Grassi[1], Silvia Onofri[2], Marco Pedicini[3], Luca Sozzi[4]

[1] Radboud University, Nijmegen, the Netherlands
[2] Scuola Normale Superiore di Pisa, Pisa, Italy
[3] Università Roma Tre, Roma, Italy
[4] Università degli Studi di Milano, Milano, Italy
lgrassi@science.ru.nl, silvia.onofri@sns.it,
marco.pedicini@uniroma3.it, luca.sozzi2@studenti.unimi.it

Abstract. Motivated by new applications such as secure Multi-Party Computation (MPC), Fully Homomorphic Encryption (FHE), and Zero-Knowledge proofs (ZK), many MPC-, FHE- and ZK-friendly symmetric-key primitives that minimize the number of multiplications over $\mathbb{F}_p$ for a large prime p have been recently proposed in the literature. This goal is often achieved by instantiating the non-linear layer via power maps $x \mapsto x^d$. In this paper, we start an analysis of new non-linear permutation functions over $\mathbb{F}_p^n$ that can be used as building blocks in such symmetric-key primitives. Given a local map $F : \mathbb{F}_p^m \to \mathbb{F}_p$, we limit ourselves to focus on S-Boxes over $\mathbb{F}_p^n$ for $n \geq m$ defined as $\mathcal{S}(x_0, x_1, \ldots, x_{n-1}) = y_0 \| y_1 \| \ldots \| y_{n-1}$ where $y_i := F(x_i, x_{i+1}, \ldots, x_{i+m-1})$. As main results, we prove that

- given any quadratic function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$, the corresponding S-Box $\mathcal{S}$ over $\mathbb{F}_p^n$ for $n \geq 3$ is never invertible;
- similarly, given any quadratic function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$, the corresponding S-Box $\mathcal{S}$ over $\mathbb{F}_p^n$ for $n \geq 5$ is never invertible.

Moreover, for each $p \geq 3$, we present (1st) generalizations of the Lai-Massey construction over $\mathbb{F}_p^n$ defined as before via functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for each $n = m \geq 2$ and (2nd) (non-trivial) quadratic functions $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ such that $\mathcal{S}$ over $\mathbb{F}_p^n$ for $n \in \{3, 4\}$ is invertible. As an open problem for future work, we conjecture that for each $m \geq 1$ there exists a *finite* integer $n_{\max}(m)$ such that $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as before via a quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ is *not* invertible for each $n \geq n_{\max}(m)$. Finally, as a concrete application, we propose Neptune, a variant of the sponge hash function Poseidon, whose non-linear layer is designed by taking into account the results presented in this paper. We show that this variant leads to a concrete multiplication reduction with respect to Poseidon.

Keywords: Multiplicative Complexity · Non-Linear Layer · MPC/FHE/ZK-Friendly Schemes · Poseidon

Contents

1 Introduction

Due to the development of new applications such as Secure Multi-Party Computation (MPC), Fully Homomorphic Encryption (FHE), and Zero-Knowledge proofs (ZK), several symmetric cryptographic schemes have been recently proposed in the literature to minimize the number of field multiplications in their natural algorithmic description, often referred to as the *multiplicative complexity*. Today, many of the mentioned applications operate on $\mathbb{F}_p \equiv \mathrm{GF}(p)$ for a large prime $p \geq 3$ (usually, p is of order $2^{64}, 2^{128}$ or even bigger), hence having cryptographic schemes that have a natural description over $\mathbb{F}_p$ is desirable. MPC-, FHE- and ZK-friendly symmetric-key primitives defined over $\mathbb{F}_p$ include MiMC [AGR+16], GMiMC [AGP+19], HadesMiMC [GLR+20], *Rescue* [AAB+20], Poseidon [GKR+21], Masta [HKC+20], Ciminion [DGGK21], Pasta [DGH+21] and *Grendel* [Sze21]. As designing symmetric-key primitives in this domain is relatively new and not well-understood, many of these schemes share some common features. In particular, the non-linear function used in almost all of them is a simple power map, that is $x \mapsto x^d$.[1] The only exceptions are Ciminion (whose permutation is based on a Feistel construction, whose non-linear function is defined as $(x, y) \in \mathbb{F}^2 \mapsto x \cdot y \in \mathbb{F}$) and Masta, whose non-linear layer resembles the chi-function introduced in [Wol85], which constitutes a prototype for the construction of the new non-linear functions we study in this paper. We start a research of new non-linear permutation functions over $\mathbb{F}_p^n$ that can be used as building blocks in MPC-, FHE- and ZK-friendly symmetric-key primitives.

1.1 The Round Function and the Non-Linear Layer

Symmetric cryptographic schemes including ciphers, permutations and hash functions are typically designed by iterating an efficiently implementable round function a sufficient number of times in order to guarantee the desired security level. Focusing on Substitution-Permutation Network (SPN) schemes, this round function is usually composed of two layers, a non-linear one and a linear one. In more details, consider a SPN scheme over $\mathbb{F}_p^t$ for a prime $p \geq 3$ and $t \geq 1$. The round function is usually defined as

$$x \quad \mapsto \quad c + \mathcal{M} \times \text{S-Box}(x) \tag{1}$$

for each $x \in \mathbb{F}_p^t$, where

- S-Box : $\mathbb{F}_p^t \to \mathbb{F}_p^t$ is the non-linear layer (or substitution layer);
- $\mathcal{M} \in \mathbb{F}_p^{t \times t}$ is an invertible matrix;
- $c \in \mathbb{F}_p^t$ is a round constant or a secret key.

Focusing on the non-linear layer, it is usually composed of parallel independent non-linear functions. Let $1 \leq n \leq t$ be a divisor of t, and let $\mathcal{S} : \mathbb{F}_p^n \to \mathbb{F}_p^n$ be an invertible non-linear function. Given $x = (x_0, x_1, \ldots, x_{t-1}) \in \mathbb{F}_p^t$, the substitution layer is usually defined as

$$\text{S-Box}(x) := \mathcal{S}(x_0, \ldots, x_{n-1}) \parallel \mathcal{S}(x_n, \ldots, x_{2n-1}) \parallel \ldots \parallel \mathcal{S}(x_{t-n}, \ldots, x_{t-1}), \tag{2}$$

[1] We recall that *Grendel* uses $x \mapsto x \cdot L_p(x)$ as its S-Box, where $L_p(\cdot) : \mathbb{F}_p \to \{-1, 0, 1\}$ is the Legendre definition. However, by definition, $L_p(x) = x^{(p-1)/2}$ is a power map.

where $\cdot \| \cdot$ denotes concatenation.

For each $z \in \mathbb{F}_p^n$, the S-Box $\mathcal{S}(z) = y_0 \| y_1 \| \dots \| y_{n-1} \in \mathbb{F}_p^n$ is defined as

$$\mathcal{S}(z) := F_0(z) \| F_1(z) \| \dots \| F_{n-1}(z)$$

where $F_0, F_1, \dots, F_{n-1} : \mathbb{F}_p^n \to \mathbb{F}_p$ are potentially distinct functions. In this paper, we limit ourselves to consider the case in which each value $y_i \in \mathbb{F}_p$ is specified according to a single local map $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for a certain $m \leq n$. More formally:

Definition 1. Let $p \geq 3$ be a prime integer. Let $1 \leq m \leq n$, and let $F : \mathbb{F}_p^m \to \mathbb{F}_p$ be a non-linear function. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ is defined as

$$\mathcal{S}(x_0, x_1, \dots, x_{n-1}) := y_0 \| y_1 \| \dots \| y_{n-1} \tag{3}$$

where

$$y_i = F(x_i, x_{i+1}, \dots, x_{i+m-1}) \tag{4}$$

for each $i \in \{0, 1, \dots, n-1\}$, where the sub-indexes are taken modulo n.

In the following, we sometimes use the notation $\mathcal{S}_F$ to highlight the local function F that defines $\mathcal{S}$.

One of the most well known examples of this kind of non-linear layer is the chi-function over $\mathbb{F}_2^n$ defined via the local map $\chi : \mathbb{F}_2^3 \to \mathbb{F}_2$

$$\chi(x_0, x_1, x_2) := x_0 + (x_1 + 1) \cdot x_2 \mod 2, \tag{5}$$

first introduced by Wolfram [Wol85] and then re-considered and analyzed by Daemen [Dae95]. It is used as a building component in many designs, including Keccak [BPVA+11,BDPA13], Rasta [DEG+18], Subterranean [DMMR20], among many others. Other examples of local maps $F : \mathbb{F}_2^m \to \mathbb{F}_2$ for which the corresponding $\mathcal{S}$ over $\mathbb{F}_2^n$ defined as in Def. 1 is invertible are listed in [Dae95, App. A.3].

In this paper, we only focus on quadratic functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for a prime $p \geq 3$, studying the properties and the multiplicative cost of the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1.

1.2 Our Contributions

Related Works: Invertible Functions over $\mathbb{F}_p^n$. Well known examples of invertible functions over $\mathbb{F}_p^n$ are recalled in Sect. 3, and include – among others – the power maps $x \mapsto x^d$ (whose invertibility is proved using Hermite's criterion) and their generalizations, the Dickson polynomials.

The Legendre symbol $L_p : \mathbb{F}_p \to \{-1, 0, 1\}$ defined as $L_p(x) := x^{(p-1)/2}$ (recalled in Def. 5) is a particular example of power maps that can be exploited in order to construct invertible functions over $\mathbb{F}_p$ for $p \geq 3$. Examples of invertible functions over $\mathbb{F}_p^n$ based on the Legendre symbol include $x \mapsto x \cdot (\alpha + L_p(x))$ where $L_p(\alpha^2 - 1) = 1$ introduced by Shallue [Sha12], $x \mapsto x^d \cdot L_p(x)$ where $\gcd(d + (p-1)/2, p-1) = 1$ introduced by Szepieniec [Sze21], and their generalization.

Probably, the most well known example of a function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for which the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^n$ for $n = m$ is a permutation is the Lai-Massey construction [LM90]. In the case $n = m = 2$, the function F is e.g. of the form $F(x_0, x_1) = x_0 + (x_0 - x_1)^2$. In Prop. 5, we present generalizations of such function over $\mathbb{F}_p^m$ for *even* $m = n$, that is, $F(x_0, x_1, \dots, x_{n-1}) = \sum_{i=0}^{n-1} \gamma_i \cdot x_i + (\sum_{i=0}^{n-1} (-1)^i \cdot x_i)^2$ and $F(x_0, x_1, \dots, x_{n-1}) = \sum_{i=0}^{n-1} \gamma_i \cdot x_i + \sum_{i=0}^{n-1} (x_i - x_{i-1})^2$, which we prove to be invertible if the matrix $\mathrm{circ}(\gamma_0, \gamma_1, \dots, \gamma_{n-1}) \in \mathbb{F}_p^{n \times n}$ is invertible.

Invertible Quadratic Functions. Even if the Lai-Massey constructions just presented can be efficiently computed (from the point of view of the multiplicative complexity), a cryptographic scheme based only on such non-linear functions can be potentially broken using e.g. an invariant subspace attack [Vau99] if e.g. the linear layer is not chosen appropriately. For this reason, we look for other quadratic functions as possible building block of a MPC-/ FHE-/ZK-friendly symmetric-key primitive, and we find the following:

- $F(x_0, x_1, x_2) = \sum_{i=0}^{2} \psi_i \cdot x_i + (x_0 + x_1 + x_2) \cdot (\alpha \cdot x_0 + \beta \cdot x_1 + \gamma \cdot x_2)$ for which $\mathcal{S}$ over $\mathbb{F}_p^3$ defined as in Def. 1 is invertible if $p = 2 \mod 3$ by carefully choosing $\psi_i, \alpha, \beta, \gamma$ as given in Prop. 9;
- $F(x_0, x_1, x_2) = \alpha \cdot (x_0 - x_1)^2 + \beta \cdot (x_1 - x_2)^2 + \gamma \cdot (x_2 - x_0)^2 + \varepsilon \cdot x_0 + \varepsilon' \cdot (x_0 + x_1 + x_2)$ for which $\mathcal{S}$ over $\mathbb{F}_p^3$ defined as in Def. 1 is invertible if $p = 1 \mod 3$ by carefully choosing $\alpha, \beta, \gamma, \varepsilon, \varepsilon'$ as given in Prop. 10.

These two functions cover all possible values of $p \geq 3$, and they can be computed via only three $\mathbb{F}_p$-multiplications, that is, t $\mathbb{F}_p$-multiplications per round. For comparison, a non-linear layer instantiated via the power map $x \mapsto x^d$ – where $d \geq 3$ so that $\gcd(d, p-1) = 1$ – requires $t \cdot (\lfloor \log_2(d) \rfloor + \mathrm{hw}(d) - 1) \geq 2 \cdot t$ $\mathbb{F}_p$-multiplications,[2] which is at least a double the cost required for functions in the two families just proposed.

Non-Existence Results. As main results of this paper:

- in Theorem 3, we *prove* that there is **no** quadratic function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ such that the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^n$ for $n \geq 3$ defined as in Def. 1 is a permutation;
- in Theorem 4, we *prove* that there is **no** quadratic function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ such that the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^n$ for $n \geq 5$ defined as in Def. 1 is a permutation.

Both results are also supported by our practical experiments, as given in Sect. 5.3.2. Regarding the case $m = n = 2$, in Prop. 8 we *prove* that the only quadratic function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ for which $\mathcal{S}$ over $\mathbb{F}_p^2$ defined as in Def. 1 is invertible is a Lai-Massey function of the form $F(x_0, x_1) = \alpha \cdot x_0 + \beta \cdot x_1 + \gamma \cdot (x_0 - x_1)^2$ for $\alpha \neq \pm\beta$.

Focusing on the case $m = 3$, it is some-way surprising when comparing the binary case and the prime case. Indeed, while e.g. the function $\mathcal{S}$ over $\mathbb{F}_2^n$ defined as in Def. 1 instantiated via the local map χ defined as in (5) is known to be a permutation for each odd $n \geq 3$, here we prove that there is no equivalent of the chi-function when working with a quadratic function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ for a prime integer $p \geq 3$.

As an open problem for future work, we conjecture that for each $m \geq 1$ there exists a *finite* integer $n_{\max}(m)$ such that $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 via a quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ is *not* invertible for each $n \geq n_{\max}(m)$ (see Conjecture 1 for details). Our results and observations suggest that if such conjecture is true, then $n_{\max}(m)$ growths linearly with m (more specifically, $n_{\max}(m) = 2 \cdot m - 1$).

Neptune as a Concrete Application. Estimating the impact of quadratic non-linear layers in the design of a generic MPC-/FHE-/ZK-friendly iterative symmetric scheme is in general very hard, since many factors play a crucial role in determining the performance of the scheme in the target applications (e.g., the number of rounds required for its security – and so the overall multiplicative complexity – does not depend only on the details of the non-linear layer, but also on the details of the linear layer, on the possible attack scenarios, on the security level, and so on). For this reason, we focus on Poseidon – a

[2] Given $d = \sum_{i=0}^{\lfloor \log_2(d) \rfloor} d_i \cdot 2^i$ for $d_i \in \{0, 1\}$, evaluating $x \mapsto x^d$ can require computing x^{2^j} for each $j \in \{0, 1, \ldots, \lfloor \log_2(d) \rfloor\}$ for a cost of $\lfloor \log_2(d) \rfloor$ multiplications, plus other $\mathrm{hw}(d) - 1$ multiplications to get $x \mapsto x^d$ (where $\mathrm{hw}(\cdot)$ is the Hamming weight).

sponge hash function [BDPV07, BDPA08] recently proposed for ZK applications – and we show a possible way to modify it based on the non-linear layers presented in this paper in order to reduce its multiplicative complexity.

POSEIDON is a sponge hash function, whose internal permutation is based on the Hades design strategy [GLR+20], proposed at Eurocrypt 2020. Its main feature and novelty regards the use of both rounds with *full* S-Box layer and rounds with *partial* S-Box layer in order to achieve both security and good performance. *Here, we take this concept to its extremes.* Instead of limiting ourselves to consider an uneven distribution of the S-Boxes, we propose to use two different round functions, one for the internal part and one for the external one. In Sect. 7, we propose a new sponge hash function called NEPTUNE over $\mathbb{F}_p^t$, a variant of the hash function POSEIDON in which

- the power maps $x \mapsto x^d$ in the external full rounds are replaced by a concatenation of independent S-Boxes defined over $\mathbb{F}_p^2$ via the Lai-Massey construction;
- the power map $x \mapsto x^d$ in the internal partial rounds remains unchanged, but the matrix that instantiates the linear layer of the internal partial rounds is different from the one proposed for the external full rounds.

As we show in there, these changes have the effect of (largely) reducing the multiplicative complexity of POSEIDON in the case of large $t \gg 1$.

2 Preliminary

Notation. Let p be a prime number (unless specified otherwise, we always assume $p \geq 3$). Let $\mathbb{F}_p$ denote the field of integer numbers modulo p. We use small letters to denote either parameters/indexes or variables and greek letters to denote fixed elements in $\mathbb{F}_p$. Given $x \in \mathbb{F}_p^n$, we denote by x_i its i-th component for each $i \in \{0, 1, \ldots, n-1\}$, that is, $x = (x_0, x_1, \ldots, x_{n-1})$ or $x = x_0 \| x_1 \| \ldots \| x_{n-1}$, where $\cdot\|\cdot$ denotes concatenation. We use capital letters to denote functions from $\mathbb{F}_p^m$ to $\mathbb{F}_p$ for $m \geq 1$, e.g., $F : \mathbb{F}_p^m \to \mathbb{F}_p$ and the calligraphic font to denote functions over $\mathbb{F}_p^n$ for $n > 1$, e.g., $\mathcal{S} : \mathbb{F}_p^n \to \mathbb{F}_p^n$. We use the fraktur font (e.g., $\mathfrak{X}$) to denote sets of elements, where $|\mathfrak{X}|$ denotes the cardinality of the set $\mathfrak{X}$. We denote by $\mathsf{e}_i \in \mathbb{F}_p^n$ the vector with 1 in the i-th component (for $i \in \{0, 1, \ldots, n-1\}$), and 0 in all others. We denote by $\mathrm{circ}(\mu_0, \mu_1, \ldots, \mu_{n-1}) \in \mathbb{F}_p^{n \times n}$ a circulant matrix

$$\mathrm{circ}(\mu_0, \mu_1, \ldots, \mu_{n-1}) := \begin{bmatrix} \mu_0 & \mu_1 & \ldots & \mu_{n-2} & \mu_{n-1} \\ \mu_{n-1} & \mu_0 & \ldots & \mu_{n-3} & \mu_{n-2} \\ \vdots & & & & \vdots \\ \mu_1 & \mu_2 & \ldots & \mu_{n-1} & \mu_0 \end{bmatrix}.$$

2.1 Class of Equivalence

First, we introduce a relation for classifying functions with similar properties.

Definition 2 (Class of Equivalence)**.** Let $q = p^r$ where $p \geq 2$ is a prime and $r \geq 1$. Let $F, F' : \mathbb{F}_q^m \to \mathbb{F}_q$ be two functions. F and F' are similar – denoted as $F \sim F'$ – if and only if $F'(x) = \omega \cdot F(\mathcal{M} \times x + \nu) + \psi$ for each $x \in \mathbb{F}_p^m$, where

- $\mathcal{M} = \mu \cdot \mathrm{diag}(1, 1, \ldots, 1) \in \mathbb{F}_q^{m \times m}$ with $\mu \in \mathbb{F}_q \setminus \{0\}$ is an invertible matrix ;
- $\nu = \nu' \| \nu' \| \ldots \| \nu' \in \mathbb{F}_q^m$ for $\nu' \in \mathbb{F}_q$;
- $\omega \in \mathbb{F}_q \setminus \{0\}$ and $\psi \in \mathbb{F}_q$.

Proposition 1. *Let $F, F' : \mathbb{F}_q^m \to \mathbb{F}_q$ be two similar functions. Let $\mathcal{S}_F, \mathcal{S}_{F'} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be two functions defined as in Def. 1 induced respectively by F and F'. Then, $\mathcal{S}_F$ is invertible if and only if $\mathcal{S}_{F'}$ is invertible.*

Proof. By definition of F' and $\mathcal{S}_{F'}$, we have that $[\mathcal{S}_{F'}(x_0, x_1, \ldots, x_{n-1})]_i = F'(x_i, x_{i+1}, \ldots, x_{i+m-1})$, where the sub-indexes are taken modulo n. Since $F'(x) = \omega \cdot F(\mathcal{M} \times x + \nu) + \psi$ for each $x \in \mathbb{F}_p^m$, it follows that

$$\mathcal{S}_{F'}(x) = \mathrm{diag}(\omega, \omega, \ldots, \omega) \times \mathcal{S}_F(\mathcal{M} \times x + \nu) + \bar{\psi}$$

where $\mathrm{diag}(\omega, \omega, \ldots, \omega) \in \mathbb{F}_q^{n \times n}$ is an invertible matrix and where $\bar{\psi} = (\psi, \psi, \ldots, \psi) \in \mathbb{F}_q^n$. Since the two diagonal matrices are invertible, then $\mathcal{S}_{F'}$ is equal to $\mathcal{S}_F$ pre-composed and post-composed with two invertible affine functions. This implies that $\mathcal{S}_{F'}$ is invertible if and only if $\mathcal{S}_F$ is invertible. □

Note that this result is not true if one changes the equivalence class defined in Def. 2 by considering generic matrices $\mathcal{M} \in \mathbb{F}_q^{m \times m}$ and/or generic $\nu \in \mathbb{F}_q^m$.

2.2 A Necessary Condition for Inverbility

As the next step, we recall a necessary condition that a function F has to satisfy for $\mathcal{S}$ to be invertible.

Definition 3 (Balanced Function)**.** Let $q = p^r$ where $p \geq 2$ is a prime and $r \geq 1$. Let $F : \mathbb{F}_q^m \to \mathbb{F}_q$. We say that F is **balanced** if and only if the pre-image of every element in $\mathbb{F}_q$ has the same cardinality, i.e. $|\{x \in \mathbb{F}_q^m \mid F(x) = y\}| = q^{n-1}$ for each $y \in \mathbb{F}_q$.

Proposition 2. *Let $q = p^r$ where $p \geq 2$ is a prime and $r \geq 1$. Given $F : \mathbb{F}_q^m \to \mathbb{F}_q$, let $\mathcal{S}$ over $\mathbb{F}_q^n$ defined as in Def. 1. If F is not balanced, then $\mathcal{S}$ is not invertible.*

The proof of this well known result is given in App. A. A concrete application of it is given in the following proposition:

Proposition 3. *Let $p \geq 2$ be a prime number. Let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be defined as in (7). If $\alpha_{2,0} = \alpha_{0,2} = 0$, then F is **not** a balanced function.*

Proof. Let $F(x_0, x_1) = \alpha_{2,0} \cdot x_0^2 + \alpha_{1,1} \cdot x_0 \cdot x_1 + \alpha_{0,2} \cdot x_1^2 + \alpha_{1,0} \cdot x_0 + \alpha_{0,1} \cdot x_1 + \alpha_{0,0}$. W.l.o.g., let's assume $\alpha_{1,1} = 1$ and $\alpha_{0,0} = 0$ due to Prop. 1 based on the class of equivalence defined in Def. 2. In order to prove the result, we analyse separately two cases: (1st) $\alpha_{1,0} = \alpha_{0,1} = 0$ and (2nd) $\alpha_{0,1} \neq 0$ (the proof is analogous for $\alpha_{1,0} \neq 0$):

- If $\alpha_{0,1} = \alpha_{1,0} = 0$, then $F(x_0, x_1) = 0$ if $x_0 = 0$ or $x_1 = 0$. It follows that $|F^{-1}(0)| \geq 2p - 1 \gneq p$, hence F is not balanced;
- For the latter, we re-write $F(x_0, x_1) = (x_0 + \alpha_{0,1}) \cdot x_1 + \alpha_{1,0} \cdot x_0$. If $x_0 = -\alpha_{0,1}$, then $F(-\alpha_{0,1}, x_1) = -\alpha_{0,1} \cdot \alpha_{1,0}$ for all $x_1 \in \mathbb{F}_q$. Moreover, $F(0, -\alpha_{1,0}) = -\alpha_{0,1} \cdot \alpha_{1,0}$. Since $\alpha_{0,1} \neq 0$ by assumption, it follows that $|F^{-1}(-\alpha_{0,1} \cdot \alpha_{1,0})| \geq p + 1 \gneq p$, which means that F is not balanced. □

Since $x^2 = x$ for each $x \in \mathbb{F}_2$, we have that $\alpha_{2,0} = \alpha_{0,2} = 0$, this means that there is no quadratic function $F : \mathbb{F}_2^2 \to \mathbb{F}_2$ such that $\mathcal{S}$ over $\mathbb{F}_2^n$ for $n \geq 2$ defined as in Def. 1 is invertible, in accordance with the results given in [Dae95].

Corollary 1. *Let $F : \mathbb{F}_2^2 \to \mathbb{F}_2$ be a quadratic function. Then, the function $\mathcal{S}$ over $\mathbb{F}_2^n$ for $n \geq 2$ defined as in Def. 1 is **not** invertible.*

3 Related Works

3.1 Hermite's Criterion and Known Permutation Polynomials (PPs) over $\mathbb{F}_p$

Given a non-linear function $F(x) = \sum_{i=0}^{d} \alpha_i \cdot x^i$ of degree $d \geq 2$, a characterization of which F is or not a permutation polynomial is given by Hermite's criterion.

Theorem 1 (Hermite's Criterion [MP13])**.** *Let $q = p^r$, where $p \geq 2$ is a prime and r is a positive integer. Then a polynomial $F \in \mathbb{F}_q[x]$ is a Permutation Polynomial (PP) of $\mathbb{F}_q$ if and only if the following two conditions hold:*

1. *the reduction of $(F(x))^{q-1} \mod (x^q - x)$ is monic polynomial of degree $q-1$;*
2. *for each integer t with $1 \leq t \leq q-2$ and $t \neq 0 \mod p$, the reduction of $(F(x))^t \mod (x^q - x)$ has degree $\leq q-2$.*

Applying the previous criteria on a generic function over $\mathbb{F}_q$ in order to establish if it is a PP or not is in general computational demanding. However, for certain special classes of polynomials, including the power maps and the Dickson polynomials, this question is easy to answer.

Power Maps. As we have already mentioned in the introduction, non-linear functions of many cryptographic schemes over $\mathbb{F}_p$ are power maps $x \mapsto x^d$.

Theorem 2 ([MP13, Sect. 8])**.** *Let $q = p^r$, where $p \geq 2$ is a prime and r is a positive integer. The function $F(x) = x^d$ where d is a positive integer is a PP if and only if $\gcd(d, q-1) = 1$.*

As it is well known, this means that the choice of the exponent d depends on the prime field if one aims to guarantee invertibility. Obviously, this also implies that no quadratic function over $\mathbb{F}_p$ for $p \geq 3$ is invertible. Indeed, consider the generic quadratic function $F(x) = \alpha \cdot x^2 + \beta \cdot x + \gamma$. Via the change of variable $y = x - \beta/(2\alpha)$, we obtain $F(y) = \alpha y^2 + \gamma$, which is not invertible since $F(y) = F(-y)$ for each $y \in \mathbb{F}_p$.

Dickson Polynomials. Dickson polynomials generalize power maps. Let $q = p^r$, where $p \geq 3$ is a prime and r is a positive integer. Let $\alpha \in \mathbb{F}_q$ fixed. The Dickson polynomial $\mathcal{D}_{d,\alpha}(x)$ of degree d with parameter α over $\mathbb{F}_q$ is defined as

$$\mathcal{D}_{d,\alpha}(x) := \sum_{j=0}^{\lfloor d/2 \rfloor} \frac{d}{d-j} \cdot \binom{d-j}{j} \cdot (-\alpha)^j \cdot x^{d-2j}.$$

Note that Dickson polynomials reduce to power maps for $\alpha = 0$. As proved e.g. in [MP13], the Dickson polynomial $\mathcal{D}_{d,\alpha}(x)$ is a PP of $\mathbb{F}_q$ if and only if $\gcd(d, q^2-1) = 1$.

3.2 Permutation Polynomials via the Legendre Symbol

Here we recall some properties of the Legendre symbol used in the following.

Definition 4. Let $p \geq 3$ be a prime number. An integer α is a quadratic residue modulo p if it is congruent to a perfect square modulo p, and it is a quadratic non-residue modulo p otherwise.

Definition 5. The Legendre symbol $L_p(\cdot)$ is a function $L_p : \mathbb{F}_p \to \{-1, 0, 1\}$ defined as $L_p(x) := x^{\frac{p-1}{2}} \mod p \in \{-1, 0, 1\}$, or equivalently $L_p(0) = 0$ and

$$L_p(x) := \begin{cases} 1 & \text{if } x \text{ is a non-zero quadratic residue modulo } p, \\ -1 & \text{if } x \text{ is a quadratic non-residue modulo } p \end{cases}.$$

Proposition 4 ([Nag51]). *The Legendre symbol has the following properties:*

1. *if* $x = y \mod p$, *then* $L_p(x) = L_p(y)$;
2. $L_p(x \cdot y) = L_p(x) \cdot L_p(y)$.

Moreover, particular identities include:

- $L_p(-1) = 1$ *if* $p = 1 \mod 4$, *while* $L_p(-1) = -1$ *if* $p = 3 \mod 4$;
- $L_p(-3) = 1$ *if* $p = 1 \mod 3$, *while* $L_p(-3) = -1$ *if* $p = 2 \mod 3$;
- $L_p(2) = 1$ *if* $p = 1, 7 \mod 8$, *while* $L_p(2) = -1$ *if* $p = 3, 5 \mod 8$.

For completeness, we point out that some permutations based on the Legendre symbol have been proposed in the literature:

- in [Sha12, Theorem 1.20], Shallue proved that the function $x \mapsto x \cdot (L_p(x) + \alpha)$ is invertible if and only if $\alpha \in \mathbb{F}_p \setminus \{\pm 1\}$ such that $L_p(\alpha^2 - 1) = 1$;
- in [Sze21], Szepieniec proved that $x \mapsto x^d \cdot L_p(x) = x^{d+(p-1)/2}$ is invertible if and only if $\gcd(d + (p-1)/2, p-1) = 1$ for a positive integer d;
- other permutations $\mathcal{S}_F$ over $\mathbb{F}_p^n$ via a function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ that involve the Legendre symbol have been recently proposed in [GKRS21].

3.3 (Generalized) Lai-Massey Functions $\mathcal{S} : \mathbb{F}_p^n \to \mathbb{F}_p^n$

Other classes of invertible functions over $\mathbb{F}_p^n$ include the generalization of the Lai-Massey construction

$$(x_0, x_1) \mapsto (y_0, y_1) = (x_0 + H(x_0 - x_1), x_1 + H(x_0 - x_1))$$

proposed in [LM90], whose invertibility relies on the fact that $y_0 - y_1 = x_0 - x_1$.

Proposition 5. *Let* $p \geq 2$ *be a prime integer. Let* $n = m \geq 2$ *such that either* n *is a multiple of* p *(i.e.,* $n = 0 \mod p$*) or* n *is even (i.e.,* $n = 2n'$*), Given*

$$F(x_0, x_1, \ldots, x_{n-1}) = \sum_{i=0}^{n-1} \mu_i \cdot x_i + H\big(\omega_0 \cdot x_0 + \omega_1 \cdot x_1 + \ldots + \omega_{n-1} \cdot x_{n-1}\big) \tag{6}$$

where

1. $\omega_i = 1$ *for each* $i \in \{0, 1, \ldots, n-1\}$ *if* $n = 0 \mod p$, *and* $\omega_i = (-1)^i$ *for each* $i \in \{0, 1, \ldots, n-1\}$ *if* $n = 0 \mod 2$;
2. *if* $n = 0 \mod 2$: $H : \mathbb{F}_p \to \mathbb{F}_p$ *is an even function (that is,* $H(x) = H(-x)$*);*
3. *the circulant matrix* $\mathcal{M} = circ(\mu_0, \ldots, \mu_{n-1}) \in \mathbb{F}_p^{n \times n}$ *is invertible;*

then the function $\mathcal{S} : \mathbb{F}_p^n \to \mathbb{F}_p^n$ *defined as in Def. 1 is invertible.*

Proof. Let $y = \mathcal{S}(x)$. By definition of $\mathcal{S}$ and since H is an even function for $n = 0 \mod 2$:

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} = \mathcal{M} \times \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} + \begin{bmatrix} H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i\right) \\ H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i\right) \\ \vdots \\ H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i\right) \end{bmatrix} = \mathcal{M} \times \begin{bmatrix} x_0 + \frac{1}{\mu'} \cdot H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i\right) \\ x_1 + \frac{1}{\mu'} \cdot H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i\right) \\ \vdots \\ x_{n-1} + \frac{1}{\mu'} \cdot H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i\right) \end{bmatrix}$$

where $\mu' := \sum_i \mu_i \neq 0$ since $\mathcal{M}$ is invertible by assumption.

Let $z := \mathcal{M}^{-1} \times y \in \mathbb{F}_p^n$. The overall costruction is invertible since

$$\sum_{i=0}^{n-1} \omega_i \cdot z_i = \sum_{i=0}^{n-1} \omega_i \cdot x_i$$

where

$$\sum_{i=0}^{n-1} \left(\frac{\omega_i}{\mu'} \cdot H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i \right) \right) = \frac{1}{\mu'} \cdot H\left(\sum_{i=0}^{n-1} \omega_i \cdot x_i \right) \cdot \underbrace{\sum_{i=0}^{n-1} \omega_i}_{=0} = 0\,. \qquad \square$$

The simplest example of a function that satisfies the previous assumptions is obtained by choosing $H(x) = \beta \cdot x^2 + \gamma$ for $\beta, \gamma \in \mathbb{F}_p$ and $\mathcal{M} = \text{circ}(1, 0, \ldots, 0) \in \mathbb{F}_p^n$. In such a case, computing $\mathcal{S}$ over $\mathbb{F}_p^n$ requires just one $\mathbb{F}_p$-multiplication.

Proposition 6. *Let $p \geq 2$ be a prime integer. Let*

$$F(x_0, x_1, \ldots, x_{n-1}) = \sum_{i=0}^{n-1} \mu_i \cdot x_i + \gamma \cdot \sum_{i=0}^{n-1} H\big(x_i - x_{i+1}\big)$$

where $H : \mathbb{F}_p \to \mathbb{F}_p$ is an even function, where $\mathcal{M} = \ circ(\mu_0, \ldots, \mu_{n-1}) \in \mathbb{F}_p^{n \times n}$ is an invertible matrix and where $\gamma \in \mathbb{F}_p \setminus \{0\}$. Then, the function $\mathcal{S} : \mathbb{F}_p^n \to \mathbb{F}_p^n$ defined as in Def. 1 is invertible.

Proof. Let $y = \mathcal{S}(x)$. Working as before, note that

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} = \mathcal{M} \times \begin{bmatrix} x_0 + \frac{\gamma}{\mu'} \cdot \sum_{i=0}^{n-1} H\big(x_i - x_{i+1}\big) \\ x_1 + \frac{\gamma}{\mu'} \cdot \sum_{i=0}^{n-1} H\big(x_i - x_{i+1}\big) \\ \vdots \\ x_{n-1} + \frac{\gamma}{\mu'} \cdot \sum_{i=0}^{n-1} H\big(x_i - x_{i+1}\big) \end{bmatrix}$$

since H is an even function, where $\mu' := \sum_i \mu_i \neq 0$ since $\mathcal{M}$ is invertible. The overall construction is invertible by noting that $x_i - x_{i+1} = z_i - z_{i+1}$, where $z := \mathcal{M}^{-1} \times y$. $\square$

If $n \geq 3$, then evaluating $\mathcal{S}$ costs n $\mathbb{F}_p$-multiplications (and just one multiplication for the case $n = 2$).

4 Function $\mathcal{S}_F$ over $\mathbb{F}_p^n$ via Quadratic Functions $F : \mathbb{F}_p^2 \to \mathbb{F}_p$

In this section, we study functions $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 instantiated via a quadratic polynomial function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$.

A Necessary Condition for the Invertibility. We start by providing a necessary condition that a quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ defined as

$$F(x_0, x_1, \ldots, x_{m-1}) := \sum_{\substack{0 \leq i_0, \ldots, i_{m-1} \leq 2 \text{ s.t.} \\ i_0 + \ldots + i_{m-1} \leq d}} \alpha_{i_0, \ldots, i_{m-1}} \cdot x_0^{i_0} \cdot \cdots \cdot x_{m-1}^{i_{m-1}}. \tag{7}$$

must satisfy in order to guarantee that $\mathcal{S}$ can be a permutation.

Proposition 7. *Let $p \geq 3$ be a prime integer. Let $F : \mathbb{F}_p^m \to \mathbb{F}_p$ be defined as in (7). Let $\alpha^{(2)}, \alpha^{(1)} \in \mathbb{F}_p$ be the sum of the coefficients of the monomials of degree l, that is*

$$\alpha^{(l)} := \sum_{\substack{0 \leq i_0, \ldots, i_{m-1} \leq l \ s.t. \\ i_0 + \ldots + i_{m-1} = l}} \alpha_{i_0, \ldots, i_{m-1}}. \tag{8}$$

*for each $l \in \{1, 2\}$. If $\alpha^{(2)} = \alpha^{(1)} = 0$ or $\alpha^{(2)} \neq 0$, the function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **not** a permutation for each $n \geq m$.*

Proof. In order to prove that $\mathcal{S}$ is not a permutation, we look for collisions, that is, we look for two *different* elements $y, z \in \mathbb{F}_p^n$ such that $\mathcal{S}(y) = \mathcal{S}(z)$ and $y \neq z$. In order to do this, we work with elements of the form $w = (\hat{w}, \hat{w}, \ldots, \hat{w}) \in \mathbb{F}_p^n$, that is $w_i = w_j$ for each $i, j \in \{0, 1, \ldots, n\}$. Note that

$$F(x, x, \ldots, x) = \alpha^{(2)} \cdot x^2 + \alpha^{(1)} \cdot x + \alpha_{0,0,\ldots,0}.$$

If $\alpha^{(2)} = \alpha^{(1)} = 0$, then $F(x, x, \ldots, x) = \alpha_{0,0,\ldots,0}$ for each $x \in \mathbb{F}_p$. It follows that $\mathcal{S}(z \equiv (x, x, \ldots, x)) = (\alpha_{0,0,\ldots,0}, \ldots, \alpha_{0,0,\ldots,0})$ for each $x \in \mathbb{F}_p$. If $\alpha^{(2)} \neq 0$:

- if $\alpha^{(2)} \neq 0$ and $\alpha^{(1)} \neq 0$, then $F(x, x, \ldots, x) = x \cdot (\alpha^{(2)} \cdot x + \alpha^{(1)}) + \alpha_{0,0,\ldots,0} = \alpha_{0,0,\ldots,0}$ admits two different solutions, which are $x = 0$ and $x = -\alpha^{(1)}/\alpha^{(2)}$;
- if $\alpha^{(2)} \neq 0$ and $\alpha^{(1)} = 0$, then $F(x, x, \ldots, x) = \alpha^{(2)} \cdot x^2 + \alpha_{0,0,\ldots,0} = \alpha^{(2)} \cdot \beta^2 + \alpha_{0,0,\ldots,0}$ for $\beta \in \mathbb{F}_p \setminus \{0\}$ has two different solutions, which are $x = \pm\beta$. □

As a result, $\mathcal{S}$ *can* be a permutation only in the case in which $\alpha^{(2)} = 0$ and $\alpha^{(1)} \neq 0$. In the following, we study this case depending on the value of $m \geq 2$.

4.1 Analysis of the Case $n = 2$

Here we prove that the only quadratic function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ for which $\mathcal{S}$ is invertible over $\mathbb{F}_p^2$ is $F(x_0, x_1) = \gamma_0 \cdot x_0 + \gamma_1 \cdot x_1 + \gamma_2 \cdot (x_0 - x_1)^2$ where $\gamma_0 \neq \pm\gamma_1$.

Proposition 8. *Let $p \geq 3$ be a prime integer, and let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be a quadratic function. The non-linear function $\mathcal{S}$ defined as in Def. 1 is invertible over $\mathbb{F}_p^2$ if and only if*

$$F(x_0, x_1) = \gamma_0 \cdot x_0 + \gamma_1 \cdot x_1 + \gamma_2 \cdot (x_0 - x_1)^2$$

where $\gamma_0 \neq \pm\gamma_1$.

Proof. Consider a generic function $G(x_0, x_1) = \alpha_{1,0} \cdot x_0 + \alpha_{0,1} \cdot x_1 + \alpha_{2,0} \cdot x_0^2 + \alpha_{0,2} \cdot x_1^2 + \alpha_{1,1} \cdot x_0 \cdot x_1$. First of all, if $\alpha_{2,0} + \alpha_{0,2} + \alpha_{1,1} \neq 0$, then the function is not invertible (see Prop. 7). Hence, we assume $\alpha_{1,1} = -\alpha_{2,0} - \alpha_{0,2}$, which means

$$G(x_0, x_1) = \alpha_{1,0} \cdot x_0 + \alpha_{0,1} \cdot x_1 + (\alpha_{2,0} \cdot x_0 - \alpha_{0,2} \cdot x_1) \cdot (x_0 - x_1)$$

If $\alpha_{2,0} = \alpha_{0,2}$ and $\alpha_{1,0} \neq \alpha_{0,1}$, then this corresponds to a generalization of the Lai-Massey construction, which is invertible by observing the following. Given $(y_0, y_1) = (G(x_0, x_1), G(x_1, x_0))$, then $y_0 - y_1 = (\alpha_{1,0} - \alpha_{0,1}) \cdot (x_0 - x_1)$, which implies that

$$\begin{bmatrix} \alpha_{1,0} & \alpha_{0,1} \\ \alpha_{0,1} & \alpha_{1,0} \end{bmatrix} \times \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} y_0 - \frac{\alpha_{2,0}}{(\alpha_{1,0} - \alpha_{0,1})^2} \cdot (y_0 - y_1)^2 \\ y_1 - \frac{\alpha_{2,0}}{(\alpha_{1,0} - \alpha_{0,1})^2} \cdot (y_0 - y_1)^2 \end{bmatrix}$$

admits a solution if $\alpha_{1,0} \neq \pm\alpha_{0,1}$.

Let's now consider the case $\alpha_{2,0} \neq \alpha_{0,2}$. We show that the function $\mathcal{S}_G$ is never invertible by looking for a collision $G(x_0, x_1) = G(y_0, y_1)$ and $G(x_1, x_0) = G(y_1, y_0)$. Via

the change of variables $d_i = x_i - y_i$ and $s_i = x_i + y_i$ for $i \in \{0, 1\}$, the system to solve becomes

$$\begin{bmatrix} \alpha_{2,0} \cdot d_0 - \frac{\alpha_{2,0}+\alpha_{0,2}}{2} \cdot d_1 & \alpha_{0,2} \cdot d_1 - \frac{\alpha_{2,0}+\beta_{0,2}}{2} \cdot d_0 \\ \alpha_{0,2} \cdot d_0 - \frac{\alpha_{2,0}+\alpha_{0,2}}{2} \cdot d_1 & \alpha_{2,0} \cdot d_1 - \frac{\alpha_{2,0}+\beta_{0,2}}{2} \cdot d_0 \end{bmatrix} \times \begin{bmatrix} s_0 \\ s_1 \end{bmatrix} = - \begin{bmatrix} \alpha_{1,0} \cdot d_0 + \alpha_{0,1} \cdot d_1 \\ \alpha_{1,0} \cdot d_1 + \alpha_{0,1} \cdot d_0 \end{bmatrix}$$

The determinant of the matrix is equal to

$$(\alpha_{2,0} - \beta_{0,2}) \cdot (\alpha_{2,0} + \alpha_{0,2}) \cdot (d_0 - d_1)^2.$$

If $\alpha_{2,0} \neq \pm\alpha_{0,2}$, it is sufficient to choose $d_0 \neq d_1$ in order to find a collision. The only remaining case to analyze is $\alpha_{0,2} = -\alpha_{2,0}$, for which the system of equation reduces to

$$\begin{aligned} \alpha_{2,0} \cdot (d_0 \cdot s_0 - d_1 \cdot s_1) &= \alpha_{1,0} \cdot d_0 + \alpha_{0,1} \cdot d_1, \\ d_0 \cdot (\beta_{1,0} - \beta_{0,1}) &= d_1 \cdot (\beta_{1,0} - \beta_{0,1}). \end{aligned}$$

By choosing $d_0 = d_1 \neq 0$ and s_0, s_1 such that $\alpha_{2,0} \cdot (s_0 - s_1) = \alpha_{1,0} + \alpha_{0,1}$ (note that $\alpha_{2,0} \neq 0$, otherwise G is linear), it is possible to find a collision. This concludes the proof. □

4.2 Analysis of the Case $n \geq 3$

As one of the main results of this paper, we prove that given any quadratic function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$, then the function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **never** invertible for each $n \geq 3$.

Theorem 3. *Let $p \geq 3$ be a prime integer. Let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be a function of degree 2. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **never** a permutation for each $n \geq 3$.*

Proof. Due to the results given in Prop. 7, here we limit ourselves to consider the case

$$\alpha_{2,0} + \alpha_{1,1} + \alpha_{0,2} = 0. \tag{9}$$

We first prove the result for the case $n = 3$. Our goal is to prove that for each function $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ of degree 2 defined as in (7), it is possible to find a collision, that is, two different inputs $x, y \in \mathbb{F}_p^3$ such that $\mathcal{S}(x) = \mathcal{S}(y)$:

$$F(x_0, x_1) = F(y_0, y_1), \qquad F(x_1, x_2) = F(y_1, y_2), \qquad F(x_2, x_0) = F(y_2, y_0).$$

In order to generalize this result for $n \geq 4$, we assume $x_0 = y_0 = \hat{z}$. Note that one such collision is found for $n = 3$, a collision for $n \geq 4$ can be easily set up by working with $x, y \in \mathbb{F}_p^n$ where $x_i = y_i = \hat{z}$ for each $i \geq 3$. Indeed, this implies that $F(x_i, x_{i+1}) = F(y_i, y_{i+1}) = 0$ for each $i \in \{3, \ldots, n-1\}$. Just for simplicity, we fix $\hat{z} = 0$.

Condition $F(0, x_1) = F(0, y_1)$ and $F(x_2, 0) = F(y_2, 0)$. The condition $F(0, x_1) = F(0, y_1)$ implies

$$(x_1 - y_1) \cdot (\alpha_{0,2} \cdot (x_1 + y_1) + \alpha_{0,1}) = 0\,,$$

which is satisfied by (1st) $x_1 = y_1$ or by (2nd) $x_1 = -y_1 - \alpha_{0,1}/\alpha_{0,2}$ (assuming $\alpha_{0,2} \neq 0$). Working in the same way on $F(x_2, 0) = F(y_2, 0)$, we get the conditions (1st) $x_2 = y_2$ or (2nd) $x_2 = -y_2 - \alpha_{1,0}/\alpha_{2,0}$ (assuming $\alpha_{2,0} \neq 0$).

Condition $F(x_1, x_2) = F(y_1, y_2)$. Regarding the condition $F(x_1, x_2) = F(y_1, y_2)$, that is,

$$(x_1 - y_1) \cdot (\alpha_{2,0} \cdot (x_1 + y_1) + \alpha_{1,0}) + (x_2 - y_2) \cdot (\alpha_{0,2} \cdot (x_2 + y_2) + \alpha_{0,1}) + \alpha_{1,1} \cdot (x_1 \cdot x_2 - y_1 \cdot y_2) = 0\,,$$

we consider the following cases separately:

1. if $\alpha_{0,2}, \alpha_{2,0}, \alpha_{1,1} \neq 0$, we present a collision by working with $x_1 = y_1$;
2. if $\alpha_{0,2} = 0$, we present a collision by working with $x_1 = y_1$;
3. if $\alpha_{2,0} = 0$, we present a collision by working with $x_2 = y_2$;
4. finally, if $\alpha_{1,1} = 0$, we present a collision by working with

$$x_1 = -y_1 - \frac{\alpha_{0,1}}{\alpha_{0,2}} \qquad \text{and} \qquad x_2 = -y_2 - \frac{\alpha_{1,0}}{\alpha_{2,0}}\,,$$

where $\alpha_{1,1} = 0$ and Eq. (9) implies $\alpha_{0,2} = -\alpha_{2,0} \neq 0$.

Case: $\alpha_{0,2}, \alpha_{2,0}, \alpha_{1,1} \neq 0$. By choosing $x_1 = y_1$, we have $x_2 = -(\alpha_{1,1} \cdot x_1 + \alpha_{0,1})/\alpha_{0,2} - y_2$ that implies $F(x_1, x_2) = F(y_1 = x_1, y_2)$. The condition $F(x_2, 0) = F(y_2, 0)$ always holds true if $x_1 = \frac{\alpha_{0,2} \cdot \alpha_{1,0}}{\alpha_{1,1} \cdot \alpha_{2,0}} - \frac{\alpha_{0,1}}{\alpha_{1,1}}$. Hence:

$$\forall x_2 \in \mathbb{F}_p : \qquad \mathcal{S}\left(0, \frac{\alpha_{0,2} \cdot \alpha_{1,0}}{\alpha_{1,1} \cdot \alpha_{2,0}} - \frac{\alpha_{0,1}}{\alpha_{1,1}}, x_2\right) = \mathcal{S}\left(0, \frac{\alpha_{0,2} \cdot \alpha_{1,0}}{\alpha_{1,1} \cdot \alpha_{2,0}} - \frac{\alpha_{0,1}}{\alpha_{1,1}}, -x_2 - \frac{\alpha_{1,0}}{\alpha_{2,0}}\right).$$

Case: $\alpha_{0,2} = 0$. Observe that $\alpha_{0,2} = 0$ implies $\alpha_{1,1} = -\alpha_{2,0} \neq 0$ due to Eq. (9). Under the assumption $x_1 = y_1$, the condition $F(x_1, x_2) = F(y_1, y_2)$ implies that $(x_2 - y_2) \cdot (\alpha_{1,1} \cdot x_1 + \alpha_{0,1}) = 0$ always holds true if $x_1 = -\alpha_{0,1}/\alpha_{1,1}$. Hence:

$$\forall x_2 \in \mathbb{F}_p : \qquad \mathcal{S}\left(0, -\frac{\alpha_{0,1}}{\alpha_{1,1}}, x_2\right) = \mathcal{S}\left(0, -\frac{\alpha_{0,1}}{\alpha_{1,1}}, -x_2 - \frac{\alpha_{1,0}}{\alpha_{2,0}}\right).$$

Case: $\alpha_{2,0} = 0$. Observe that $\alpha_{2,0} = 0$ implies that $\alpha_{1,1} = -\alpha_{0,2} \neq 0$ due to Eq. (9). Under the assumption $x_2 = y_2$, the condition $F(x_1, x_2) = F(y_1, y_2)$ implies that $(x_1 - y_1) \cdot (\alpha_{1,1} \cdot x_2 + \alpha_{1,0}) = 0$ always holds true if $x_2 = -\alpha_{1,0}/\alpha_{1,1}$. Hence:

$$\forall x_1 \in \mathbb{F}_p : \qquad \mathcal{S}\left(0, x_1, -\frac{\alpha_{1,0}}{\alpha_{1,1}}\right) = \mathcal{S}\left(0, -x_1 - \frac{\alpha_{0,1}}{\alpha_{0,2}}, -\frac{\alpha_{1,0}}{\alpha_{1,1}}\right).$$

Case: $\alpha_{1,1} = 0$. Observe that $\alpha_{1,1} = 0$ implies that $\alpha_{2,0} = -\alpha_{0,2} \neq 0$ due to Eq. (9). Choosing $y_1 = -x_1 - \alpha_{0,1}/\alpha_{0,2}$ and $y_2 = -x_2 - \alpha_{1,0}/\alpha_{2,0}$, the condition $F(x_1, x_2) = F(y_1, y_2)$ implies $x_1 + x_2 = \frac{\alpha_{1,0} - \alpha_{0,1}}{2 \cdot \alpha_{0,2}}$. Hence $\forall x_2 \in \mathbb{F}_p$:

$$\mathcal{S}\left(0, -x_2 + \frac{\alpha_{1,0} - \alpha_{0,1}}{2 \cdot \alpha_{0,2}}, x_2\right) = \mathcal{S}\left(0, x_2 - \frac{\alpha_{1,0} + \alpha_{0,1}}{2 \cdot \alpha_{0,2}}, -x_2 - \frac{\alpha_{1,0}}{\alpha_{2,0}}\right). \qquad \square$$

5 Function $\mathcal{S}_F$ over $\mathbb{F}_p^n$ via Quadratic Functions $F : \mathbb{F}_p^3 \to \mathbb{F}_p$

In this section, we set up permutations $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 via quadratic polynomial functions $F : \mathbb{F}_p^3 \to \mathbb{F}_p$.

5.1 Analysis of the Case $n = 3$

Here we present two non-trivial quadratic functions $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ (one for $p = 2 \mod 3$ and one for $p = 1 \mod 3$) for which the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^3$ is a permutation.

5.1.1 Case: $p = 2 \mod 3$

First, we present a family of functions $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ for which the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^3$ is a permutation if $p = 2 \mod 3$. Three $\mathbb{F}_p$-multiplications are sufficient for computing $\mathcal{S}$.

Proposition 9. *Let $p \geq 5$ be a prime integer such that $p = 2 \mod 3$. Let $\alpha, \beta \in \mathbb{F}_p$ and $\psi_0, \psi_1, \psi_2 \in \mathbb{F}_p$ such that (1st) $\sum_i \psi_i \neq 0$ and (2nd) one of the following conditions is satisfied*

- *$\psi_1 = \psi_2$ and $\psi_0 \neq \psi_1$ and $2 \cdot \alpha + \beta \neq 0$;*
- *$\psi_0 = \psi_2$ and $\psi_0 \neq \psi_1$ and $2 \cdot \beta + \alpha \neq 0$;*
- *$\psi_0 = \psi_1$ and $\psi_0 \neq \psi_2$ and $\alpha - \beta \neq 0$.*

Let $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ be defined as

$$F(x_0, x_1, x_2) = \sum_{i=0}^{2} \psi_i \cdot x_i + (x_0 + x_1 + x_2) \cdot (\alpha \cdot x_0 + \beta \cdot x_1 - (\alpha + \beta) \cdot x_2).$$

The function $\mathcal{S}$ defined as in Def. 1 over $\mathbb{F}_p^3$ is invertible.

Proof. Note that $\alpha = \beta = 0$ is never possible. Given $y_i = F(x_i, x_{i+1}, x_{i+2})$ for each $i \in \{0, 1, 2\}$ (where the sub-indexes are taken modulo 3), note that

$$y_0 + y_1 + y_2 = (\psi_0 + \psi_1 + \psi_2) \cdot (x_0 + x_1 + x_2).$$

Let $\hat{y} := (y_0 + y_1 + y_2)/(\psi_0 + \psi_1 + \psi_2)$, where $\hat{y} = 0$ if and only if $x_0 + x_1 + x_2 = y_0 + y_1 + y_2 = 0$. The system of equations for $\mathcal{S}(x) = y$ becomes

$$\begin{bmatrix} \alpha \cdot \hat{y} + \psi_0 & \beta \cdot \hat{y} + \psi_1 & -(\alpha + \beta) \cdot \hat{y} + \psi_2 \\ -(\alpha + \beta) \cdot \hat{y} + \psi_2 & \alpha \cdot \hat{y} + \psi_0 & \beta \cdot \hat{y} + \psi_1 \\ \beta \cdot \hat{y} + \psi_1 & -(\alpha + \beta) \cdot \hat{y} + \psi_2 & \alpha \cdot \hat{y} + \psi_0 \end{bmatrix} \times \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}.$$

By simple computation, the matrix is invertible if

$$(\psi_0 + \psi_1 + \psi_2) \cdot \Big(3 \cdot (\alpha^2 + \alpha \cdot \beta + \beta^2) \cdot \hat{y}^2 + 3 \cdot (\alpha \cdot (\psi_0 - \psi_2) + \beta \cdot (\psi_1 - \psi_2)) \cdot \hat{y} + \\ + (\psi_0^2 + \psi_1^2 + \psi_2^2 - \psi_0 \cdot \psi_1 - \psi_1 \cdot \psi_2 - \psi_0 \cdot \psi_2) \Big) \neq 0,$$

where $\psi_0 + \psi_1 + \psi_2 \neq 0$. If $\hat{y} = 0$, then the matrix is invertible, since $\psi_0^2 + \psi_1^2 + \psi_2^2 - \psi_0 \cdot \psi_1 - \psi_1 \cdot \psi_2 - \psi_0 \cdot \psi_2$ is different from zero by assumption on ψ_0, ψ_1, ψ_2.

If $\hat{y} \neq 0$, first note that the coefficient $\alpha^2 + \alpha \cdot \beta + \beta^2$ of $\hat{y}^2$ is always different from zero for each α, β since -3 is not a square modulo p due to the assumption $p = 2 \mod 3$ (see Prop. 4). Indeed, $\alpha^2 + \alpha \cdot \beta + \beta^2 = 0$ for $\beta \neq 0$ is equivalent to $z^2 + z + 1 = 0$ for $z = \alpha/\beta$, which admits as solutions $(-2 \pm \sqrt{-3})/2$. Since -3 is a quadratic non-residue modulo p, then no solution exists. Hence, assuming $\psi_1 = \psi_2$ (analogous for the others), the determinant is equal to zero if and only if

$$\hat{y} = \frac{-3 \cdot \alpha \cdot (\psi_0 - \psi_2) \pm \sqrt{-3 \cdot (\psi_0 - \psi_2)^2 \cdot (\alpha + 2 \cdot \beta)^2}}{6 \cdot (\alpha^2 + \alpha \cdot \beta + \beta^2)},$$

which does not admit any solution since -3 is a quadratic non-residue modulo p (note that $\psi_0 \neq \psi_2$ and $\alpha \neq -2\beta$). It follows that $\mathcal{S}(x) = y$ is invertible. □

5.1.2 Case: $p = 1 \mod 3$

Next, we present a family of functions $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ for which the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^3$ is a permutation if $p = 1 \mod 3$. Again, three $\mathbb{F}_p$-multiplications are sufficient for computing $\mathcal{S}$.

Proposition 10. *Let $p \geq 7$ be a prime integer such that $p = 1 \mod 3$. Let $\alpha, \beta, \gamma, \varepsilon, \varepsilon' \in \mathbb{F}_p$ such that*

1. *$\varepsilon \neq 0$, $\varepsilon + 3 \cdot \varepsilon' \neq 0$*
2. *$\alpha \neq \gamma$, $\alpha \cdot \beta \neq \gamma^2$, and $\beta \in \{\beta_+, \beta_-\}$ where*

$$\beta_\pm = \frac{\alpha \cdot (1 \pm \sqrt{-3}) - \gamma \cdot (-1 \pm \sqrt{-3})}{2}. \tag{10}$$

Let $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ be defined as

$$F(x_0, x_1, x_2) = \alpha \cdot (x_0 - x_1)^2 + \beta \cdot (x_1 - x_2)^2 + \gamma \cdot (x_2 - x_0)^2 + \varepsilon \cdot x_0 + \varepsilon' \cdot (x_0 + x_1 + x_2).$$

The function $\mathcal{S}$ defined as in Def. 1 over $\mathbb{F}_p^3$ is invertible.

Note that the case $\alpha = \beta = \gamma$ has been already analyzed in Prop. 6.

Proof. Let's define $\omega := \alpha \cdot \beta - \gamma^2$, $\psi := \beta \cdot \gamma - \alpha^2$, $\tau := \alpha \cdot \gamma - \beta^2$. First of all, note that $\omega, \tau, \psi \neq 0$ and that the following equations

$$\begin{aligned} \omega \cdot \alpha + \psi \cdot \beta + \tau \cdot \gamma &= 0 \\ \omega \cdot \beta + \psi \cdot \gamma + \tau \cdot \alpha &= 0 \\ \omega \cdot \gamma + \psi \cdot \alpha + \tau \cdot \beta &= 0 \end{aligned}$$

are always satisfied. In particular, focusing on the last one, we have that

$$\omega \cdot \gamma + \psi \cdot \alpha + \tau \cdot \beta = (\alpha + \beta + \gamma) \cdot (\alpha^2 + \beta^2 + \gamma^2 - \alpha \cdot \beta - \beta \cdot \gamma - \alpha \cdot \gamma) = 0,$$

where $\alpha^2 + \beta^2 + \gamma^2 - \alpha \cdot \beta - \beta \cdot \gamma - \alpha \cdot \gamma = \omega + \tau + \psi = 0$ due to the assumption on β. Indeed, the solutions of this last equality are $\{\beta_+, \beta_-\}$ as defined in (10), recalling that -3 is a quadratic residue modulo p for $p = 1 \mod 3$ (see Prop. 4).

In order to prove the result, we show how to invert $\mathcal{S}(x) = y$. Given $y_i = F(x_i, x_{i+1}, x_{i+2})$ for each $i \in \{0, 1, 2\}$ (where the sub-indexes are taken modulo 3), note that $\omega \cdot y_0 + \tau \cdot y_1 + \psi \cdot y_2 = \varepsilon \cdot (\omega \cdot x_0 + \tau \cdot x_1 + \psi \cdot x_2)$, which is satisfied by

$$x_0 = \frac{\hat{y} - \tau \cdot x_1 - \psi \cdot x_2}{\omega}, \qquad \text{where} \qquad \hat{y} := \frac{\omega \cdot y_0 + \tau \cdot y_1 + \psi \cdot y_2}{\varepsilon}$$

and where $\omega, \varepsilon \neq 0$ by assumption.

By taking the difference between y_1 and y_2 and by substituting x_0, we obtain:

$$\underbrace{\left(\omega^2 \cdot (\alpha - \gamma) + \tau^2 \cdot (\beta - \alpha) + \psi^2 \cdot (\gamma - \beta)\right)}_{=0} \cdot (x_1 - x_2)^2 + \\ + \omega^2 \cdot \varepsilon \cdot (x_1 - x_2) + (\gamma - \alpha) \cdot \hat{y}^2 - \omega^2 \cdot (y_1 - y_2) = 0$$

where the coefficient of $(x_1 - x_2)^2$ is equal to zero due to assumption on $\beta = \beta_\pm$.

As a result, we have a linear equation in x_1 with $\omega^2 \cdot \varepsilon \neq 0$ as coefficient, hence we have

$$x_1 = x_2 + \frac{(\gamma - \alpha) \cdot \hat{y}^2 - \omega^2 \cdot (y_1 - y_2)}{\omega^2 \cdot \varepsilon}.$$

By substituing x_0, x_1 in e.g. the third equation $y_2 = \varepsilon \cdot x_2 + \varepsilon' \cdot (x_0 + x_1 + x_2) + \alpha \cdot (x_2 - x_0)^2 + \beta \cdot (x_0 - x_1)^2 + \gamma \cdot (x_1 - x_2)^2$, we get a linear equation in x_2 of the form:

$$(\varepsilon + 3\varepsilon') \cdot x_2 + G(y_0, y_1, y_2) = 0$$

where

$$G(y_0, y_1, y_2) = \left(\frac{\alpha+\beta}{\omega^2} + \gamma\right) \cdot \left(\frac{(\gamma-\alpha)\cdot\hat{y}^2 - \omega^2 \cdot (y_1 - y_2)}{\omega^2 \cdot \varepsilon}\right)^2 - y_2 + \frac{\varepsilon'}{\omega} \cdot \hat{y}$$
$$+ \hat{y}^2 \cdot \frac{\alpha+\beta}{\omega^2} + \left(\frac{\varepsilon'}{\omega} \cdot (\omega - \tau) + \frac{2\hat{y}}{\omega^2} \cdot (\psi - \tau)\right) \cdot \frac{(\gamma-\alpha)\cdot\hat{y}^2 - \omega^2 \cdot (y_1 - y_2)}{\omega^2 \cdot \varepsilon} .$$

Since the coefficient $\varepsilon + 3 \cdot \varepsilon'$ of x_2 is different from zero by assumption, the system of equations has a unique solution for any given y_1, y_2, y_3 and $\mathcal{S}$ is invertible.

□

5.2 An Example for the Case $n = 4$

Here we limit ourselves to present an example of a quadratic function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ for which $\mathcal{S}$ over $\mathbb{F}_p^4$ is invertible. Such function is constructed based on the following result.

Proposition 11. *Let $q = p^r$ where $p \geq 2$ is a prime and r is a positive integer. Given $2 \leq g \leq h$, let $G : \mathbb{F}_q^g \to \mathbb{F}_q$ be a function for which $\mathcal{S}_G$ defined over $\mathbb{F}_q^h$ as in Def. 1 (that is, $\mathcal{S}_G(x_0, \ldots, x_{h-1}) = G(x_0, \ldots, x_{g-1}) \| G(x_1, \ldots, x_g) \| \ldots \| G(x_{h-1}, x_0, \ldots, x_{g-2})$, where the sub-indexes are taken modulo h) is invertible.*

Let $m := (g-1)\cdot(z+1)+1$ and $n := h \cdot (z+1)$ for any integer $z \geq 0$. Let $F : \mathbb{F}_q^m \to \mathbb{F}_q$ be defined as

$$F(x_0, \ldots, x_{m-1}) := G(x_0, x_{z+1}, x_{2\cdot(z+1)}, \ldots, x_{(g-1)\cdot(z+1)}) .$$

(Note that F depends only on the variables x_i for which the sub-index i is a multiple of $z+1$.) The function $\mathcal{S}_F$ defined over $\mathbb{F}_q^n$ as in Def. 1 is invertible.

Proof. The result is obviously true for $z = 0$ (for which $m = g$ and $n = h$). So, let's assume $z \geq 1$. Let $y = \mathcal{S}_F(x)$. The system of n equations $y_i = F(x_i, x_{i+1}, \ldots, x_{i+m-1}$ for each $i \in \{0, 1, \ldots, n-1\}$ can be split into $z+1$ *independent* systems, each one consisting of h equations of the form

$$\begin{cases} y_i = G(x_i, x_{i+(z+1)}, x_{i+2\cdot(z+1)}, \ldots, x_{i+(g-1)\cdot(z+1)}) \\ y_{i+(z+1)} = G(x_{i+(z+1)}, x_{i+2\cdot(z+1)}, \ldots, x_{i+(g-1)\cdot(z+1)}, x_{i+g\cdot(z+1)}) \\ y_{i+2\cdot(z+1)} = G(x_{i+2\cdot(z+1)}, x_{i+3\cdot(z+1)}, \ldots, x_{i+g\cdot(z+1)}, x_{i+(g+1)\cdot(z+1)}) \\ \vdots \\ y_{i+(h-1)\cdot(z+1)} = G(x_{i+(h-1)\cdot(z+1)}, x_i, x_{i+(z+1)}, \ldots, x_{i+(g-2)\cdot(z+1)}) \end{cases}$$

for each $i \in \{0, 1, \ldots, z\}$, that is,

$$\forall i \in \{0, 1, \ldots, z\} : \quad (y_i, y_{i+(z+1)}, \ldots, y_{i+(h-1)\cdot(z+1)}) = \mathcal{S}_G(x_i, x_{i+(z+1)}, \ldots, x_{i+(h-1)\cdot(z+1)}) .$$

The invertibility of each one of these sub-systems follows from the fact that $\mathcal{S}_G$ is invertible by assumption. □

The following corollary follows immediately.

Corollary 2. *Let $p \geq 3$ be a prime integer, and let $m \geq 2$. Let $G : \mathbb{F}_p^2 \to \mathbb{F}_p$ be a function for which $\mathcal{S}_G$ over $\mathbb{F}_p^2$ defined as in Def. 1 is invertible. Let $F : \mathbb{F}_p^m \to \mathbb{F}_p$ be defined as $F(x_0, \ldots, x_{m-1}) := G(x_0, x_{m-1})$. Then $\mathcal{S}_F$ over $\mathbb{F}_p^{2\cdot(m-1)}$ defined as in Def. 1 is invertible.*

Based on these results, given $F(x_0, x_1, x_2) = \gamma_0 \cdot x_0 + \gamma_2 \cdot x_2 + (x_0 - x_2)^2$ defined over $\mathbb{F}_p^3$ such that $\gamma_0 \neq \pm\gamma_2$, then $\mathcal{S}_F$ defined over $\mathbb{F}_q^4$ as in Def. 1 is invertible.

5.3 Analysis of the Case $n \geq 5$

As a main result of this work, we prove that given any quadratic function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$, then $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is never invertible for $n \geq 5$.

Theorem 4. *Let $p \geq 3$ be a prime integer. Let $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ be a function of degree 2. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **never** a permutation for each $n \geq 5$.*

As highlighted in the introduction, this result is quite surprising if compared to the $\mathbb{F}_2$ case, for which it is well known that the function $\mathcal{S}$ over $\mathbb{F}_2^n$ instantiated via the local map χ defined as in (5) is a permutation for each odd $n \geq 3$.

5.3.1 The Roadmap for the Proof of Theorem 4

The detailed proof of Theorem 4 is given in Sect. 6. Here we limit ourselves to present the roadmap of such proof.

In order to prove Theorem 4, we consider separately the following cases:

1. the function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ depends on at most two input variables (equivalently, it is independent of at least one variable): due to the result given in Theorem 3, we know that the corresponding $\mathcal{S}$ is never invertible (note that the case $F(x_0, x_1, x_2) = G(x_0, x_2)$ reduces to the one studied in Theorem 3);
2. the function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ only contains monomials that depend on a single variable, that is $\alpha_{1,1,0} = \alpha_{1,0,1} = \alpha_{0,1,1} = 0$;
3. the function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ is linear in one variable, e.g., $F(x_0, x_1, x_2) = x_0 + G(x_1, x_2)$ where $G : \mathbb{F}_p^2 \to \mathbb{F}_p$ is a function of degree 2;
4. for each variable x_i, there is at least one monomial of degree two that depends on it.

The second case is studied in Lemma 1, the third case is studied in Lemma 2, while the last case is studied in Lemma 3. *Since the proofs of these Lemmas are similar to the one given for Theorem 3*, here we present a sketch of the proofs for each one of the cited Lemmas, and we refer to Sect. 6 for all the details.

Lemma 1. *Let $p \geq 3$ be a prime integer. Let $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ be a function of degree 2 defined as*

$$F(x_0, x_1, x_2) = \alpha_{2,0,0} \cdot x_0^2 + \alpha_{0,2,0} \cdot x_1^2 + \alpha_{0,0,2} \cdot x_2^2 + \alpha_{1,0,0} \cdot x_0 + \alpha_{0,1,0} \cdot x_1 + \alpha_{0,0,1} \cdot x_2,$$

*that is, $\alpha_{1,1,0} = \alpha_{1,0,1} = \alpha_{0,1,1} = 0$. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **never** a permutation for each $n \geq 5$.*

We refer to Sect. 6.1 for the proof. The idea of the proof is the following. In order to find a collision, we study separately the cases (1st) $\alpha_{0,0,2}, \alpha_{2,0,0} \neq 0$ and (2nd) $\alpha_{0,0,2} = 0$ or $\alpha_{2,0,0} = 0$ (note that at least two terms among $\alpha_{0,0,2}, \alpha_{0,2,0}, \alpha_{2,0,0}$ are different from zero, since $\alpha^{(2)} = 0$):

- in the first case, we show that the result is true for $n = 5$ by finding two different inputs $x, y \in \mathbb{F}_p^5$ such that $x_0 = y_0 = x_1 = y_1 = \hat{z} \in \mathbb{F}_p$ and $\mathcal{S}(x) = \mathcal{S}(y) \in \mathbb{F}_p^5$ and $x \neq y$. This is done by solving a system of (linear) equations. The collision over $\mathbb{F}_p^n$ for $n \geq 6$ is obtained by working with $x' = x\|\hat{z}\|\hat{z}\| \ldots \|\hat{z}$ and $y' = y\|\hat{z}\|\hat{z}\| \ldots \|\hat{z} \in \mathbb{F}_p^n$;
- in the second case, we construct a collision directly over $\mathbb{F}_p^n$. The condition for the collision corresponds on a system of linear equation in $(x_i + y_i)$. By choosing in an appropriate way the differences $(x_i - y_i)$, it is possible to find a non-trivial collision for such system of equations.

Lemma 2. *Let $p \geq 3$ be a prime integer. Let $G : \mathbb{F}_p^2 \to \mathbb{F}_p$ be a function of degree 2, and let $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ be a function of degree 2 defined as*

1. $F(x_0, x_1, x_2) = \alpha_{1,0,0} \cdot x_0 + G(x_1, x_2)$, *or*
2. $F(x_0, x_1, x_2) = \alpha_{0,0,1} \cdot x_2 + G(x_0, x_1)$.

*The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **never** a permutation for each $n \geq 5$.*

We refer to Sect. 6.2 for the proof. The idea of the proof is the following. First of all, the case $F(x_0, x_1, x_2) = \alpha_{0,1,0} \cdot x_1 + G(x_0, x_2)$ is included in Lemma 3.[3] In the other cases, in order to find a collision, we study separately the cases (1st) $n = 2n' + 1 \geq 5$ odd and (2nd) $n = 2n'' + 2 \geq 6$ even:

- in the first case, we consider two inputs $x, y \in \mathbb{F}_p^n$ such that $x_i = y_i$ for each i odd and $x_j \neq y_j$ for each j even. The collision is found by solving a system of (linear) equations in x_i for i odd and in $(x_j + y_j)$ for j even;
- the strategy for the second case is similar. The only difference regards the choice of the input $x, y \in \mathbb{F}_p^n$ which are defined as $x_i = y_i$ for each $i \neq n - 1$ odd and $x_j \neq y_j$ for each j even and $j = n - 1$.

Lemma 3. *Let $p \geq 3$ be a prime integer. Let $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ be a function of degree 2 defined as in (7), such that*

$$(\alpha_{2,0,0}, \alpha_{1,0,1}, \alpha_{1,1,0}), (\alpha_{0,2,0}, \alpha_{0,1,1}, \alpha_{1,1,0}) \neq (0, 0, 0),$$

*that is, either $F(x_0, x_1, x_2) = \alpha_{0,1,0} \cdot x_1 + G(x_0, x_2)$ where $G : \mathbb{F}_p^2 \to \mathbb{F}_p$ is a quadratic function or for each variable x_0 and x_2 there is at least one monomial of degree two that depends on it. The function $\mathcal{S}_F$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **never** a permutation for each $n \geq 5$.*

The idea of the proof – given in details in Sect. 6.3 – is the following. First of all, we show that the result is true for $n = 5$ by finding two different inputs $x, y \in \mathbb{F}_p^5$ such that $x_0 = y_0 = x_1 = y_1 = \hat{z} \in \mathbb{F}_p$ and $\mathcal{S}(x) = \mathcal{S}(y) \in \mathbb{F}_p^5$ and $x \neq y$. In particular:

- if $\alpha_{1,0,1} \neq 0$, then it is sufficient that x and y are different by a single component in order to find a collision. In this case, the condition for the collision corresponds to a system of linear equation in $\hat{z}$ and the two variables that are equal in x and y;
- if $\alpha_{1,0,1} = 0$, then at least two components of x and y should be different in order to find a collision. Again, the collision is found by solving a system of linear equations.

The collision over $\mathbb{F}_p^n$ for $n \geq 6$ is obtained by working with $x' = x \| \hat{z} \| \hat{z} \| \dots \| \hat{z}$ and $y' = y \| \hat{z} \| \hat{z} \| \dots \| \hat{z} \in \mathbb{F}_p^n$.

5.3.2 Practical Verification

The theoretical results just given are supported by our practical verification, for which **no** quadratic function F that induces an invertible $\mathcal{S}$ is found. For our practical tests, we limit ourselves to consider balanced functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ under the class of equivalence defined in Def. 2. The practical results are reported in App. B – Table 3 for the case of quadratic functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for $m = 2, 3$. Those include the number and the percentage of balanced functions, the maximum value of $n \geq 3$ tested and the total runtime.

[3] Besides that, note that e.g. a function $F(x_0, x_1, x_2) = x_0 + \alpha \cdot x_1 + G(x_2)$ that is linear in two variables cannot generate a permutation $\mathcal{S}$ due to Prop. 7 ($\alpha^{(2)} \neq 0$).

6 Proof of Theorem 4

As we already mentioned, we prove that the function $\mathcal{S}$ over $\mathbb{F}_p^n$ is not a permutation for any quadratic function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ and $n \geq 5$ by constructing collisions, which means we find two distinct n-uples $x, y \in \mathbb{F}_p^n$ such that $\mathcal{S}(x) = \mathcal{S}(y)$ and $x \neq y$ or equivalently:

$$\forall i \in \{0, 1, 2, \ldots, n-1\} : \qquad F(x_i, x_{i+1}, x_{i+2}) = F(y_i, y_{i+1}, y_{i+2}),$$

where the sub-indexes are taken modulo n. For reaching this goal, we introduce new variables $s, d \in \mathbb{F}_p^n$, respectively the sum and the difference:

$$s := x + y \qquad \text{and} \qquad d := x - y.$$

Clearly a pair (s_i, d_i) is equivalent to a pair x_i, y_i, since $x = (s + d)/2$ and $y = (s - d)/2$. For the follow-up, note that the equality $2 \cdot (x_i \cdot x_j - y_i \cdot y_j) = s_i \cdot d_j + s_j \cdot d_i$ always holds.

6.1 Proof of Lemma 1

Due to the results given in Prop. 7, here we limit ourselves to consider the case $\alpha^{(2)} = 0$ and $\alpha^{(1)} \neq 0$. We study separately the cases

1. $\alpha_{0,0,2}, \alpha_{2,0,0} \neq 0$;
2. $\alpha_{0,0,2} = 0$ or $\alpha_{2,0,0} = 0$.

Note that at least two terms among $\alpha_{0,0,2}, \alpha_{0,2,0}, \alpha_{2,0,0}$ are different from zero (since $\alpha^{(2)} = 0$).

6.1.1 Case: $\alpha_{0,0,2}, \alpha_{2,0,0} \neq 0$

First of all, we show that the result is true for $n = 5$ by finding two different inputs $x, y \in \mathbb{F}_p^5$ such that $\mathcal{S}(x) = \mathcal{S}(y) \in \mathbb{F}_p^5$ and $x \neq y$. These inputs satisfy an additional condition, namely $x_0 = y_0 = x_1 = y_1 = \hat{z} \in \mathbb{F}_p$. This allows us to generalize the found collision for each $n \geq 6$. Indeed, exactly as in the case of Theorem 3, given $x, y \in \mathbb{F}_p^5$ as before, note that

$$x' = x \| \hat{z} \| \hat{z} \| \ldots \| \hat{z}, \; y' = y \| \hat{z} \| \hat{z} \| \ldots \| \hat{z} \in \mathbb{F}_p^n$$

implies a collision $\mathcal{S}(x) = \mathcal{S}(y) \in \mathbb{F}_p^n$, since F is defined over $\mathbb{F}_p^3$ and since $x_0 = y_0 = x_1 = y_1 = \hat{z}$. W.l.o.g., we fix $\hat{z} = 0$ in the following.

The condition $\mathcal{S}(x) = \mathcal{S}(y)$ over $\mathbb{F}_p^5$ is equivalent to:

$$\begin{cases} d_2 \cdot (\alpha_{0,0,2} \cdot s_2 + \alpha_{0,0,1}) = 0 \\ d_4 \cdot (\alpha_{2,0,0} \cdot s_4 + \alpha_{1,0,0}) = 0 \\ d_2 \cdot (\alpha_{0,2,0} \cdot s_2 + \alpha_{0,1,0}) + d_3 \cdot (\alpha_{0,0,2} \cdot s_3 + \alpha_{0,0,1}) = 0 \\ d_3 \cdot (\alpha_{2,0,0} \cdot s_3 + \alpha_{1,0,0}) + d_4 \cdot (\alpha_{0,2,0} \cdot s_4 + \alpha_{0,1,0}) = 0 \\ d_2 \cdot (\alpha_{2,0,0} \cdot s_2 + \alpha_{1,0,0}) + d_3 \cdot (\alpha_{0,2,0} \cdot s_3 + \alpha_{0,1,0}) + d_4 \cdot (\alpha_{0,0,2} \cdot s_4 + \alpha_{0,0,1}) = 0 \end{cases}.$$

Since $\alpha_{0,0,2}$ and $\alpha_{2,0,0}$ are different from zero, then $s_2 = -\frac{\alpha_{0,0,1}}{\alpha_{0,0,2}}$ and $s_4 = -\frac{\alpha_{1,0,0}}{\alpha_{2,0,0}}$ from the first two equations. The sum of the last three equations gives $d_3 = -d_2 - d_4$ since

$$d_2 \cdot (s_2 \cdot (\alpha_{0,2,0} + \alpha_{2,0,0}) + \alpha_{1,0,0} + \alpha_{0,1,0}) \cdot d_3 \cdot (\underbrace{\alpha^{(2)}}_{=0} \cdot s_3 + \alpha^{(1)})$$
$$+ d_4 \cdot (s_4 \cdot (\alpha_{0,2,0} + \alpha_{0,0,2}) + \alpha_{0,0,1} + \alpha_{0,1,0}) = 0$$

where

$$s_2 \cdot (\alpha_{0,2,0} + \alpha_{2,0,0}) + \alpha_{1,0,0} + \alpha_{0,1,0} = -\alpha_{0,0,2} \cdot s_2 + \alpha_{1,0,0} + \alpha_{0,1,0} = \alpha^{(1)}$$

(analogous for the coefficient of d_4). Since $d_3 \neq 0$, the third equation becomes

$$s_3 = \frac{1}{\alpha_{0,0,2}} \cdot \left(\frac{d_2 \cdot (\alpha_{0,2,0} \cdot s_2 + \alpha_{0,1,0})}{d_2 + d_4} - \alpha_{0,0,1} \right)$$

and substituting this value in the fourth equation, we get

$$(d_4 - d_2) \cdot (\alpha_{2,0,0}^2 \cdot \alpha_{0,0,1} + \alpha_{2,0,0} \cdot \alpha_{0,0,2} \cdot \alpha_{0,1,0} + \alpha_{0,0,2}^2 \cdot \alpha_{1,0,0}) = 0$$

which is always satisfied if $d_2 = d_4 \neq 0$ and $d_3 = -2 \cdot d_2$. Now taking d_i and s_i as stated, the solution of the system corresponds to a collision for $\mathcal{S}$.

6.1.2 Case: $\alpha_{0,0,2} = 0$ (analogous for $\alpha_{2,0,0} = 0$)

If $\alpha_{0,0,1} = 0$, then F depends only on x_0 and x_1 and the result follows from Theorem 3. If $\alpha_{0,0,1} \neq 0$, we highlight that it is not possible to find any collision under the condition $x_0 = y_0 = x_1 = y_1$ (using the previous strategy, we would have $d_3 = d_4 = d_5 = 0$). Here we construct a collision directly over $\mathbb{F}_p^n$.

First of all, note that since $\alpha^{(2)} = 0$ and $\alpha_{0,0,2} = 0$, it follows that $\alpha_{0,2,0} = -\alpha_{2,0,0}$. Working under the class of equivalence defined in Def. 2, we assume $\alpha_{2,0,0} = -\alpha_{0,2,0} = 1$, hence

$$F(x_0, x_1, x_2) = x_0^2 - x_1^2 + A(x_0, x_1, x_2)$$

where $A(x_0, x_1, x_2) = \alpha_{1,0,0} \cdot x_0 + \alpha_{0,1,0} \cdot x_1 + \alpha_{0,0,1} \cdot x_2$.

Note

$$F(x_i, x_{i+1}, x_{i+2}) - F(y_i, y_{i+1}, y_{i+2}) = s_i \cdot d_i - s_{i+1} \cdot d_{i+1} + A(d_i, d_{i+1}, d_{i+2})$$

where the sub-indices are taken modulo n. Since $F(x_i, x_{i+1}, x_{i+2}) = F(y_i, y_{i+1}, y_{i+2})$, note that

$$\sum_{i=0}^{n-1} (F(x_i, x_{i+1}, x_{i+2}) - F(y_i, y_{i+1}, y_{i+2})) = 0 \qquad \text{implies} \qquad \sum_{i=0}^{n-1} A(d_i, d_{i+1}, d_{i+2}) = 0$$

since

$$\begin{aligned}
&\sum_{i=0}^{n-1} (F(x_i, x_{i+1}, x_{i+2}) - F(y_i, y_{i+1}, y_{i+2})) \\
=&\sum_{i=0}^{n-1} (s_i \cdot d_i - s_{i+1} \cdot d_{i+1} + A(d_i, d_{i+1}, d_{i+2})) \\
=&\underbrace{\sum_{i=0}^{n-1} s_i \cdot d_i - \sum_{i=0}^{n-1} s_{i+1} \cdot d_{i+1}}_{=0} + \sum_{i=0}^{n-1} A(d_i, d_{i+1}, d_{i+2}) \\
=&\sum_{i=0}^{n-1} A(d_i, d_{i+1}, d_{i+2}).
\end{aligned}$$

Moreover:

$$\sum_{i=0}^{n-1} A(d_i, d_{i+1}, d_{i+2}) \qquad \text{implies} \qquad \sum_{i=0}^{n-1} d_i = 0,$$

that is, $d_{n-1} = -\sum_{i=0}^{n-2} d_i$. Indeed:

$$\sum_{i=0}^{n-1} A(d_i, d_{i+1}, d_{i+2}) = \sum_{i=0}^{n-1} (\alpha_{1,0,0} \cdot d_i + \alpha_{0,1,0} \cdot d_{i+1} + \alpha_{0,0,1} \cdot d_{i+2}) = \alpha^{(1)} \cdot \sum_{i=0}^{n-1} d_i$$

where $\alpha^{(1)} \neq 0$.

The condition $\mathcal{S}(x) = \mathcal{S}(y)$ is equivalent to

$$\begin{bmatrix} d_0 & -d_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & d_1 & -d_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & d_2 & -d_3 & 0 & \dots & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & d_{n-3} & -d_{n-2} & 0 \\ 0 & 0 & \dots & 0 & 0 & d_{n-2} & -d_{n-1} \\ -d_0 & 0 & \dots & 0 & 0 & 0 & d_{n-1} \end{bmatrix} \times \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_{n-3} \\ s_{n-2} \\ s_{n-1} \end{bmatrix} = \begin{bmatrix} A(d_0, d_1, d_2) \\ A(d_1, d_2, d_3) \\ A(d_2, d_3, d_4) \\ \vdots \\ A(d_{n-3}, d_{n-2}, d_{n-1}) \\ A(d_{n-2}, d_{n-1}, d_0) \\ A(d_{n-1}, d_0, d_1) \end{bmatrix}.$$

Note that the determinant of the l.h.s. matrix is zero, since the sum of all the rows is equal to zero. At the same time, the determinant of the first $(n-1) \times (n-1)$ submatrix is $\prod_{i=0}^{n-2} d_i$. Hence, let's fix $d_i \neq 0$ for each $i \in \{0, 1, \dots, n-2\}$ (e.g., $d_i = 1$), and also fix $d_{n-1} = -\sum_{i=0}^{n-2} d_i$ (due to the previous consideration). Working on the first $n-1$ equations in the first $n-1$ variables, it is possible to find $s_0, s_1, \dots, s_{n-2}$ that solve the first $n-1$ equations for the given $d_i \neq 0$. Finally, s_{n-1} is given by $d_{n-2} \cdot s_{n-2} - d_{n-1} \cdot s_{n-1} = A(d_{n-2}, d_{n-1}, d_0)$ (the last equation is obviously satisfied). □

6.2 Proof of Lemma 2

Due to the results given in Prop. 7, here we limit ourselves to consider the case $\alpha^{(2)} = 0$ and $\alpha^{(1)} \neq 0$. Furthermore, due to the equivalence class defined in Def. 2 in Sect. 2.1 we can assume *w.l.o.g.* that $\alpha_{1,0,0} = 1$ in the case $F(x_0, x_1, x_2) = \alpha_{1,0,0} \cdot x_0 + G(x_1, x_2)$, and $\alpha_{0,0,1} = 1$ in the case $F(x_0, x_1, x_2) = \alpha_{0,0,1} \cdot x_2 + G(x_0, x_1)$.

Observe that proving the Lemma for the first function $F(x_0, x_1, x_2) = \alpha_{1,0,0} \cdot x_0 + G(x_1, x_2)$ implies the proof for the second case $F(x_0, x_1, x_2) = \alpha_{0,0,1} \cdot x_2 + G(x_0, x_1)$ as well. Indeed, if $x = (x_0, x_1, \dots, x_{n-1})$ and $y = (y_0, y_1, \dots, y_{n-1})$ generate a collision for the function $\mathcal{S}_F$ defined via $F(x_0, x_1, x_2) = x_0 + G(x_1, x_2)$, then $x' = (x_{n-1}, x_{n-2}, \dots, x_1, x_0)$ and $y' = (y_{n-1}, y_{n-2}, \dots, y_1, y_0)$ generate a collision for the function $\mathcal{S}_{F'}$ defined via $F'(x_0, x_1, x_2) = x_2 + G(x_0, x_1)$. Hence, from now on we assume $F(x_0, x_1, x_2) = x_0 + G(x_1, x_2)$.

We study separately the cases

1. $\alpha_{0,0,2} = \alpha_{0,2,0} \neq 0$;
2. n odd and $\alpha_{0,0,2} \neq \alpha_{0,2,0}$;
3. n even and $\alpha_{0,0,2} \neq \alpha_{0,2,0}$.

6.2.1 Case: $\alpha_{0,0,2} = \alpha_{0,2,0} \neq 0$

The following proof holds for every $n \geq 5$. Observe that $\alpha^{(2)} = 0$ implies $\alpha_{0,1,1} = -2\alpha_{0,2,0}$. Working under the class of equivalence defined in Def. 2, we also assume $\alpha_{0,2,0} = 1$.

The condition for the collision is given by

$$\begin{aligned} &F(x_i, x_{i+1}, x_{i+2}) - F(y_i, y_{i+1}, y_{i+2}) \\ =&(d_{i+1} - d_{i+2}) \cdot s_{i+1} + (d_{i+2} - d_{i+1}) \cdot s_{i+2} + A(d_i, d_{i+1}, d_{i+2}) = 0, \end{aligned}$$

where $A(z_0, z_1, z_2) : \mathbb{F}_p^3 \to \mathbb{F}_p$ is the linear function $A(z_0, z_1, z_2) = z_0 + \alpha_{0,1,0} \cdot z_1 + \alpha_{0,0,1} \cdot z_2$.

The first condition $F(x_0, x_1, x_2) = F(y_0, y_1, y_2)$ is always satisfied if $d_1 = d_2$ and $d_0 + (\alpha_{0,1,0} + \alpha_{0,0,1}) \cdot d_2 = 0$, that is, $d_0 = -(\alpha_{0,1,0} + \alpha_{0,0,1}) \cdot d_2$. Note that $d_0 \neq d_2$ if $\alpha_{0,1,0} + \alpha_{0,0,1} \neq -1$, which is always the case since $\alpha^{(1)} = 1 + \alpha_{0,1,0} + \alpha_{0,0,1} \neq 0$ (remember

that we are assuming $\alpha_{1,0,0} = 1$). In such a case and by fixing $s_0 = 0$, the condition $\mathcal{S}(x) = \mathcal{S}(y)$ becomes

$$\begin{bmatrix} 0 & d_2 - d_3 & d_3 - d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 - d_4 & d_4 - d_3 & \dots & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & d_{n-2} - d_{n-1} & d_{n-1} - d_{n-2} \\ 0 & 0 & \dots & 0 & 0 & d_{n-1} - d_0 \\ d_1 - d_0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{n-3} \\ s_{n-2} \\ s_{n-1} \end{bmatrix} =$$
$$- \left[A(d_1, d_2, d_3), A(d_2, d_3, d_4), \dots, A(d_{n-2}, d_{n-1}, d_0), A(d_{n-1}, d_0, d_1)\right]^T ,$$

where $\cdot^T$ denotes the transpose matrix. By simple computation, the determinant of the l.h.s. matrix is

$$(-1)^n \cdot (d_0 - d_1) \cdot \prod_{i=2}^{n-1} (d_i - d_{i+1}),$$

which is always different from zero by choosing $d_i \neq d_{i+1}$ for each $i = 3, \dots, n-2$ (where $d_1 = d_2$ and $d_0 = (1 - \alpha^{(1)}) \cdot d_1$). As a result, given d_i as just defined and $s_0 = 0$, the solution of the previous system corresponds to two distinct $x, y \in \mathbb{F}_p^n$ such that $\mathcal{S}(x) = \mathcal{S}(y)$.

6.2.2 Case: n odd and $\alpha_{0,0,2} \neq \alpha_{0,2,0}$

Let $n = 2n' + 1$. Our strategy for finding a collision is to work with two inputs $x, y \in \mathbb{F}_p^n$ defined as

- $x_i = y_i$ for each $i \in \{1, 3, \dots, 2i' + 1, \dots, n-2\}$ odd (equivalently, $d_i = 0$ for each i odd);
- $x_i \neq y_i$ for each $i \in \{0, 2, \dots, 2i', \dots, n-1\}$ even.

The condition $\mathcal{S}(x) = \mathcal{S}(y)$ implies:

- if the first sub-index is even and if $j \in \{0, 1, \dots, (n-3)/2\}$ (equivalently, $2j + 2 \leq n - 1$):

$$\begin{aligned} &F(x_{2j}, x_{2j+1}, x_{2j+2}) - F(y_{2j}, y_{2j+1}, y_{2j+2}) \\ =&\alpha_{0,0,2} \cdot s_{2j+2} \cdot d_{2j+2} + \alpha_{0,1,1} \cdot s_{2j+1} \cdot d_{2j+2} + d_{2j} + \alpha_{0,0,1} \cdot d_{2j+2} = 0; \end{aligned}$$

- if the first sub-index is odd and if $j \in \{0, 1, \dots, (n-5)/2\}$ (equivalently, $2j+3 \leq n-1$):

$$\begin{aligned} &F(x_{2j+1}, x_{2j+2}, x_{2j+3}) - F(y_{2j+1}, y_{2j+2}, y_{2j+3}) \\ =&\alpha_{0,2,0} \cdot s_{2j+2} \cdot d_{2j+2} + \alpha_{0,1,1} \cdot s_{2j+3} \cdot d_{2j+2} + \alpha_{0,1,0} \cdot d_{2j+2} = 0; \end{aligned}$$

- finally, since $n - 1$ is even:

$$\begin{aligned} &F(x_{n-1}, x_0, x_1) - F(y_{n-1}, x_0, x_1) \\ =&\alpha_{0,2,0} \cdot s_0 \cdot d_0 + \alpha_{0,1,1} \cdot s_1 \cdot d_0 + d_{n-1} + \alpha_{0,1,0} \cdot d_0 = 0; \end{aligned}$$

  and

$$\begin{aligned} &F(x_{n-2}, x_{n-1}, x_0) - F(y_{n-2}, x_{n-1}, x_0) = \alpha_{0,2,0} \cdot s_{n-1} \cdot d_{n-1} + \alpha_{0,0,2} \cdot s_0 \cdot d_0 \\ &+ \alpha_{0,1,1} \cdot \frac{s_0 \cdot d_{n-1} + s_{n-1} \cdot d_0}{2} + \alpha_{0,1,0} \cdot d_{n-1} + \alpha_{0,0,1} \cdot d_0 = 0; \end{aligned}$$

It follows that $\mathcal{S}(x) = \mathcal{S}(y)$ corresponds to

$$\begin{bmatrix} 0 & \mathcal{M}_{0,1} & \mathcal{M}_{0,2} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \mathcal{M}_{1,2} & \mathcal{M}_{1,3} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathcal{M}_{2,3} & \mathcal{M}_{2,4} & \dots & 0 & 0 & 0 \\ \vdots & & & & \ddots & \ddots & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & \mathcal{M}_{n-3,n-2} & \mathcal{M}_{n-3,n-1} \\ \mathcal{M}_{n-2,0} & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \mathcal{M}_{n-2,n-1} \\ \mathcal{M}_{n-1,0} & \mathcal{M}_{n-1,1} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} s_0 \\ x_1 \\ s_2 \\ x_3 \\ s_4 \\ \vdots \\ s_{n-3} \\ x_{n-2} \\ s_{n-1} \end{bmatrix} =$$

$$- \left[A(d_0, 0, d_2), A(0, d_2, 0), A(d_2, 0, d_4), \dots, A(d_{n-3}, 0, d_{n-1}), A(0, d_{n-1}, d_0), A(d_{n-1}, d_0, 0)\right]^T ,$$

where $A(z_0, z_1, z_2) : \mathbb{F}_p^3 \to \mathbb{F}_p$ is the linear function $A(z_0, z_1, z_2) = z_0 + \alpha_{0,1,0} \cdot z_1 + \alpha_{0,0,1} \cdot z_2$ and where the coefficients of $\mathcal{M}$ are defined as following:

$$(\mathcal{M}_{j,j+1}; \mathcal{M}_{j,j+2}) := \begin{cases} (\alpha_{0,1,1} \cdot d_{j+2}; \alpha_{0,0,2} \cdot d_{j+2}) & \text{if } j \in \{0, 2, 4, \dots, n-3\} \\ (\alpha_{0,2,0} \cdot d_{j+1}; \alpha_{0,1,1} \cdot d_{j+1}) & \text{if } j \in \{1, 3, 5, \dots, n-4\} \\ \left(\alpha_{0,0,2} \cdot d_0 + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}; \alpha_{0,2,0} \cdot d_{n-1} + \frac{\alpha_{0,1,1} \cdot d_0}{2}\right) & \text{if } j = n-2 \\ (\alpha_{0,2,0} \cdot d_0; \alpha_{0,1,1} \cdot d_0) & \text{if } j = n-1 \end{cases} .$$

Let's analyze the determinant of the r.h.s matrix $\mathcal{M}$:

- if $\alpha_{0,2,0} = 0$ (hence, $\alpha_{0,0,2} = -\alpha_{0,1,1} \neq 0$): the determinant of the matrix is

$$\det(\mathcal{M}) = (\alpha_{0,0,2} \cdot \alpha_{0,1,1})^{n'} \cdot \left(\alpha_{0,0,2} \cdot d_0 + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}\right) \cdot \prod_{i=0}^{n'} d_{2i}^2 ,$$

  which is different from zero by choosing $\alpha_{0,0,2} \cdot d_0 \neq -\frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}$. Analogous if $\alpha_{0,0,2} = 0$;

- if $\alpha_{0,2,0}, \alpha_{0,0,2} \neq 0$ and $\alpha_{0,0,2} \neq \alpha_{0,2,0}$: by choosing $\frac{d_{n-1}}{d_0} = -\frac{\alpha_{0,1,1}}{2\alpha_{0,2,0}}$, the determinant becomes

$$\det(\mathcal{M}) = (\alpha_{0,2,0} \cdot \alpha_{0,1,1})^{n'} \cdot \left(\alpha_{0,0,2} \cdot d_0 + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}\right) \cdot \prod_{i=0}^{n'} d_{2i}^2 ,$$

  which is different from zero if $\alpha_{0,0,2} \neq \alpha_{0,2,0}$. Indeed:

$$\alpha_{0,0,2} + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2 \cdot d_0} = \frac{4\alpha_{0,0,2} \cdot \alpha_{0,2,0} - \alpha_{0,1,1}^2}{4\alpha_{0,2,0}} = -\frac{(\alpha_{0,0,2} - \alpha_{0,2,0})^2}{4\alpha_{0,2,0}} \neq 0 .$$

6.2.3 Case: n even and $\alpha_{0,0,2} \neq \alpha_{0,2,0}$

Let $n = 2n'' + 2$. The proof is similar to the one for n odd. Our strategy for finding a collision is to work with two inputs $x, y \in \mathbb{F}_p^n$ defined as

- $x_i = y_i$ for each $i \in \{1, 3, \dots, 2i'+1, \dots, n-3\}$ odd, but not for $i = n-1$ (equivalently, $d_i = 0$ for each i odd $\neq n-1$);
- $x_i \neq y_i$ for each $i \in \{0, 2, \dots, 2i', \dots, n-2\}$ even and $i = n-1$.

The condition $\mathcal{S}(x) = \mathcal{S}(y)$ implies:

- if the first sub-index is even and if $j \in \{0, 1, \dots, (n-4)/2\}$ (equivalently, $2j + 2 \leq n - 2$):

$$\begin{aligned} &F(x_{2j}, x_{2j+1}, x_{2j+2}) - F(y_{2j}, y_{2j+1}, y_{2j+2}) \\ =&\alpha_{0,0,2} \cdot s_{2j+2} \cdot d_{2j+2} + \alpha_{0,1,1} \cdot x_{2j+1} \cdot d_{2j+2} + \alpha_{1,0,0} \cdot d_{2j} + \alpha_{0,0,1} \cdot d_{2j+2} = 0 ; \end{aligned}$$

- if the first sub-index is odd and if $j \in \{0, 1, \dots, (n-6)/2\}$ (equivalently, $2j+3 \leq n-2$):

$$\begin{aligned} &F(x_{2j+1}, x_{2j+2}, x_{2j+3}) - F(y_{2j+1}, y_{2j+2}, y_{2j+3}) \\ =&\alpha_{0,2,0} \cdot s_{2j+2} \cdot d_{2j+2} + \alpha_{0,1,1} \cdot x_{2j+3} \cdot d_{2j+2} + \alpha_{0,1,0} \cdot d_{2j+2} = 0\,; \end{aligned}$$

- finally, since $n-2$ is even:

$$\begin{aligned} &F(x_{n-3}, x_{n-2}, x_{n-1}) - F(y_{n-3}, x_{n-2}, x_{n-1}) = \alpha_{0,2,0} \cdot s_{n-2} \cdot d_{n-2} + \alpha_{0,0,2} \cdot s_{n-1} \cdot d_{n-1} \\ &+ \alpha_{0,1,1} \cdot \frac{s_{n-2} \cdot d_{n-1} + s_{n-1} \cdot d_{n-2}}{2} + \alpha_{0,1,0} \cdot d_{n-2} + \alpha_{0,0,1} \cdot d_{n-1} = 0\,; \end{aligned}$$

and

$$\begin{aligned} &F(x_{n-2}, x_{n-1}, x_0) - F(y_{n-2}, x_{n-1}, x_0) = \alpha_{0,2,0} \cdot s_{n-1} \cdot d_{n-1} + \alpha_{0,0,2} \cdot s_0 \cdot d_0 \\ &+ \alpha_{0,1,1} \cdot \frac{s_0 \cdot d_{n-1} + s_{n-1} \cdot d_0}{2} + \alpha_{1,0,0} \cdot d_{n-2} + \alpha_{0,1,0} \cdot d_{n-1} + \alpha_{0,0,1} \cdot d_0 = 0\,; \end{aligned}$$

and

$$\begin{aligned} &F(x_{n-1}, x_0, x_1) - F(y_{n-1}, x_0, x_1) \\ =&\alpha_{0,2,0} \cdot s_0 \cdot d_0 + \alpha_{0,1,1} \cdot x_1 \cdot d_0 + \alpha_{1,0,0} \cdot d_{n-1} + \alpha_{0,1,0} \cdot d_0 = 0\,; \end{aligned}$$

By working as in the case n odd, the determinant of the matrix $\mathcal{M}$ of the obtained linear system is given by

$$\begin{aligned} \det(\mathcal{M}) = (\alpha_{0,1,1})^{n''} \cdot \Bigg[&(\alpha_{0,0,2})^{n''} \cdot \left(\alpha_{0,0,2} \cdot d_0 + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}\right) \cdot \left(\alpha_{0,0,2} \cdot d_{n-1} + \frac{\alpha_{0,1,1} \cdot d_{n-2}}{2}\right) \\ - &(\alpha_{0,2,0})^{n''} \cdot \left(\frac{\alpha_{0,1,1} \cdot d_{n-1}}{2} + \alpha_{0,2,0} \cdot d_{n-2}\right) \cdot \left(\frac{\alpha_{0,1,1} \cdot d_0}{2} + \alpha_{0,2,0} \cdot d_{n-1}\right) \Bigg] \cdot \prod_{i=0}^{n''} d_{2i}^2\,. \end{aligned}$$

Let's analyze the determinant of $\mathcal{M}$:

- if $\alpha_{0,2,0} = 0$ (hence, $\alpha_{0,0,2} = -\alpha_{0,1,1} \neq 0$): the determinant of the matrix is

$$(\alpha_{0,1,1} \cdot \alpha_{0,0,2})^{n''} \cdot \left(\alpha_{0,0,2} \cdot d_0 + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}\right) \cdot \left(\alpha_{0,0,2} \cdot d_{n-1} + \frac{\alpha_{0,1,1} \cdot d_{n-2}}{2}\right) \cdot \prod_{i=0}^{n''} d_{2i}^2\,,$$

which is different from zero by choosing $\alpha_{0,0,2} \cdot d_0 \neq -\frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}$ and $\alpha_{0,0,2} \cdot d_{n-1} \neq -\frac{\alpha_{0,1,1} \cdot d_{n-2}}{2}$. Analogous if $\alpha_{0,0,2} = 0$;

- if $\alpha_{0,2,0}, \alpha_{0,0,2} \neq 0$ and $\alpha_{0,0,2} \neq \alpha_{0,2,0}$: by choosing $\frac{d_{n-2}}{d_{n-1}} = -\frac{\alpha_{0,1,1}}{2\alpha_{0,2,0}}$, the determinant becomes

$$(\alpha_{0,1,1} \cdot \alpha_{0,0,2})^{n''} \cdot \left(\alpha_{0,0,2} \cdot d_0 + \frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}\right) \cdot \left(\alpha_{0,0,2} \cdot d_{n-1} + \frac{\alpha_{0,1,1} \cdot d_{n-2}}{2}\right) \cdot \prod_{i=0}^{n''} d_{2i}^2\,,$$

which is different from zero by choosing $\alpha_{0,0,2} \cdot d_0 \neq -\frac{\alpha_{0,1,1} \cdot d_{n-1}}{2}$ since $\alpha_{0,0,2} \neq \alpha_{0,2,0}$. Indeed:

$$\alpha_{0,0,2} + \frac{\alpha_{0,1,1} \cdot d_{n-2}}{2 \cdot d_{n-1}} = \frac{4\alpha_{0,0,2} \cdot \alpha_{0,2,0} - \alpha_{0,1,1}^2}{4\alpha_{0,2,0}} = -\frac{(\alpha_{0,0,2} - \alpha_{0,2,0})^2}{4\alpha_{0,2,0}} \neq 0\,.$$

□

6.3 Proof of Lemma 3

We first prove the result for the case $n = 5$. Our goal is to prove that for each function $F : \mathbb{F}_p^3 \to \mathbb{F}_p$ of degree 2 defined as in (7), it is possible to find two different inputs $x, y \in \mathbb{F}_p^5$ such that $\mathcal{S}(x) = \mathcal{S}(y)$, or equivalently:

$$\forall i \in \{0, 1, 2, 3, 4\} : \qquad F(x_i, x_{i+1}, x_{i+2}) = F(y_i, y_{i+1}, y_{i+2}),$$

where the sub-indexes are taken modulo n. As before, we assume $x_0 = y_0 = x_1 = y_1 = \hat{z} \in \mathbb{F}_p$.

Once such a collision is found, a collision for $n \geq 6$ is trivially set up by working with $x, y \in \mathbb{F}_p^n$ where

$$\forall i \geq 5 : \qquad x_i = y_i = \hat{z}\,,$$

which implies that $F(x_i, x_{i+1}, x_{i+2}) = F(y_i, y_{i+1}, x_{i+2}) = 0$ for each $i \in \{5, \ldots, n-1\}$.

6.3.1 Initial Considerations

The condition $F(\hat{z}, \hat{z}, x_2) = F(\hat{z}, \hat{z}, y_2)$ implies

$$(x_2 - y_2) \cdot (\alpha_{0,0,2} \cdot (x_2 + y_2) + (\alpha_{1,0,1} + \alpha_{0,1,1}) \cdot \hat{z} + \alpha_{0,0,1}) = 0,$$

which is satisfied either by $d_2 = 0$ or by

$$\alpha_{0,0,2} \cdot (x_2 + y_2) + (\alpha_{1,0,1} + \alpha_{0,1,1}) \cdot \hat{z} + \alpha_{0,0,1} = 0.$$

In a similar way, the condition $F(x_4, \hat{z}, \hat{z}) = F(y_4, \hat{z}, \hat{z})$ is satisfied by

$$\alpha_{2,0,0} \cdot (x_4 + y_4) + (\alpha_{1,1,0} + \alpha_{1,0,1}) \cdot \hat{z} + \alpha_{1,0,0} = 0.$$

The equality $F(\hat{z}, x_2, x_3) = F(\hat{z}, y_2, y_3)$

$$\begin{aligned}&(x_3 - y_3) \cdot (\alpha_{0,0,2} \cdot (x_3 + y_3) + \alpha_{1,0,1} \cdot \hat{z} + \alpha_{0,1,1} \cdot y_2 + \alpha_{0,0,1}) \\ =& -(x_2 - y_2) \cdot (\alpha_{0,2,0} \cdot (x_2 + y_2) + \alpha_{1,1,0} \cdot \hat{z} + \alpha_{0,1,1} \cdot x_3 + \alpha_{0,1,0})\end{aligned}$$

admits as possible solutions:

- $x_2 = y_2$ and $x_3 = y_3$;
- $x_2 = y_2$ and $\alpha_{0,0,2} \cdot (x_3 + y_3) + \alpha_{1,0,1} \cdot \hat{z} + \alpha_{0,1,1} \cdot x_2 + \alpha_{0,0,1} = 0$;
- $x_3 = y_3$ and $\alpha_{0,2,0} \cdot (x_2 + y_2) + \alpha_{1,1,0} \cdot \hat{z} + \alpha_{0,1,1} \cdot x_3 + \alpha_{0,1,0} = 0$;

and so on. Similar considerations hold for the equality $F(x_3, x_4, \hat{z}) = F(y_3, y_4, \hat{z})$, that is

$$\begin{aligned}&(x_3 - y_3) \cdot (\alpha_{2,0,0} \cdot (x_3 + y_3) + \alpha_{1,0,1} \cdot \hat{z} + \alpha_{1,1,0} \cdot y_4 + \alpha_{1,0,0}) \\ =& -(x_4 - y_4) \cdot (\alpha_{0,2,0} \cdot (x_4 + y_4) + \alpha_{0,1,1} \cdot \hat{z} + \alpha_{1,1,0} \cdot x_3 + \alpha_{0,1,0}).\end{aligned}$$

In order to find a collision, let's start by fixing $d_2 = d_4 = 0$ and $d_3 \neq 0$. The conditions for the collision are:

$$\begin{bmatrix} \alpha_{1,0,1} & \alpha_{0,1,1} & 0 \\ 0 & \alpha_{1,1,0} & \alpha_{0,1,1} \\ \alpha_{1,0,1} & 0 & \alpha_{1,1,0} \end{bmatrix} \times \begin{bmatrix} \hat{z} \\ x_2 \\ x_4 \end{bmatrix} = - \begin{bmatrix} \alpha_{0,0,2} \cdot s_3 + \alpha_{0,0,1} \\ \alpha_{0,2,0} \cdot s_3 + \alpha_{0,1,0} \\ \alpha_{2,0,0} \cdot s_3 + \alpha_{1,0,0} \end{bmatrix},$$

which clearly admits a non-trivial solution (hence, a collision) if $\alpha_{1,0,1} \neq 0$ and $\alpha_{1,1,0}^2 + \alpha_{0,1,1}^2 \neq 0$.

Similarly, by fixing $d_2 = d_3 = 0$ and $d_4 \neq 0$, the conditions for the collision are:

$$\begin{bmatrix} \alpha_{1,1,0} + \alpha_{1,0,1} & 0 & 0 \\ \alpha_{0,1,1} & 0 & \alpha_{1,1,0} \\ 0 & \alpha_{1,0,1} & \alpha_{0,1,1} \end{bmatrix} \times \begin{bmatrix} \hat{z} \\ x_2 \\ x_3 \end{bmatrix} = - \begin{bmatrix} \alpha_{2,0,0} \cdot s_4 + \alpha_{1,0,0} \\ \alpha_{0,2,0} \cdot s_4 + \alpha_{0,1,0} \\ \alpha_{0,0,2} \cdot s_4 + \alpha_{0,0,1} \end{bmatrix},$$

which has a non-trivial solution (hence, a collision) if $\alpha_{1,1,0} \neq -\alpha_{1,0,1}$ and $\alpha_{1,1,0}, \alpha_{1,0,1} \neq 0$.

Working in the same way and by fixing $d_3 = d_4 = 0$ and $d_2 \neq 0$, the conditions for the collision are:

$$\begin{bmatrix} \alpha_{1,0,1} + \alpha_{0,1,1} & 0 & 0 \\ \alpha_{1,1,0} & \alpha_{0,1,1} & 0 \\ 0 & \alpha_{1,1,0} & \alpha_{1,0,1} \end{bmatrix} \times \begin{bmatrix} \hat{z} \\ x_3 \\ x_4 \end{bmatrix} = - \begin{bmatrix} \alpha_{0,0,2} \cdot s_2 + \alpha_{0,0,1} \\ \alpha_{0,2,0} \cdot s_2 + \alpha_{0,1,0} \\ \alpha_{2,0,0} \cdot s_2 + \alpha_{1,0,0} \end{bmatrix},$$

which always admits a non-trivial solution (hence, a collision) if $\alpha_{0,1,1} \neq -\alpha_{1,0,1}$ and $\alpha_{0,1,1}, \alpha_{1,0,1} \neq 0$.

In summary, if only a single difference d_i is non-zero, the cases in which it is *not* possible to find a collisions by using the strategy just proposed are

1. $\alpha_{1,0,1} = 0$;
2. $\alpha_{1,0,1} \neq 0$ and
$$\begin{cases} \alpha_{1,1,0}^2 + \alpha_{0,1,1}^2 = 0; & \\ \alpha_{1,0,1} = -\alpha_{1,1,0} & \textit{or} \quad \alpha_{1,1,0} = 0; \\ \alpha_{1,0,1} = -\alpha_{0,1,1} & \textit{or} \quad \alpha_{0,1,1} = 0. \end{cases}$$
Indeed, if $\alpha_{1,0,1} \neq 0$, it is sufficient that one of three conditions given in the system is not fulfilled in order to find a collision using the previous results.

Let's analyze them in details.

6.3.2 Case: $\alpha_{1,0,1} \neq 0$

Case: $\alpha_{1,0,1} \neq 0$ and $\alpha_{1,1,0} = 0$. Note that in this case, $\alpha_{1,0,1} = -\alpha_{1,1,0}$ cannot occur since $\alpha_{1,0,1} \neq 0$. Due to $\alpha_{1,1,0}^2 + \alpha_{0,1,1}^2 = 0$, we have $\alpha_{0,1,1} = 0$. Let's consider separately the cases $\alpha_{0,2,0} \neq 0$ and $\alpha_{0,2,0} = 0$:

- if $\alpha_{0,2,0} \neq 0$, then working with $d_2 = d_3 = 0$ and $d_4 \neq 0$, we get new conditions of the form:
$$\begin{bmatrix} \alpha_{1,0,1} & 0 & \alpha_{2,0,0} \\ 0 & 0 & \alpha_{0,2,0} \\ 0 & \alpha_{1,0,1} & \alpha_{0,0,2} \end{bmatrix} \times \begin{bmatrix} \hat{z} \\ x_2 \\ s_4 \end{bmatrix} = - \begin{bmatrix} \alpha_{1,0,0} \\ \alpha_{0,1,0} \\ \alpha_{0,0,1} \end{bmatrix},$$
which admits a solution (namely, a collision), since the determinant of the l.h.s. matrix is non-zero.

- If $\alpha_{0,2,0} = 0$, then if $\alpha_{0,1,0} = 0$ the result follows from Theorem 3. If $\alpha_{0,2,0} = 0$ and $\alpha_{0,1,0} \neq 0$, then we can assume $\alpha_{0,1,0} = 1$ due to the equivalence class defined in Def. 2 in Sect. 2.1, that is,

  $$F(x_0, x_1, x_2) = \alpha_{2,0,0} \cdot x_0^2 + \alpha_{0,0,2} \cdot x_2^2 + \alpha_{1,0,1} \cdot x_0 \cdot x_2 + \alpha_{1,0,0} \cdot x_0 + x_1 + \alpha_{0,0,1} \cdot x_2 \,.$$

  Observe that the condition $\alpha^{(2)} = 0$ implies that at least one between $\alpha_{2,0,0}$ and $\alpha_{0,0,2}$ is different from zero. Suppose $\alpha_{2,0,0} \neq 0$ and choose $d_2 = 0$, note that if

$\alpha_{2,0,0} = 0$ then $\alpha_{0,0,2} \neq 0$ and the result is analogous choosing $d_4 = 0$. The condition $\mathcal{S}(x) = \mathcal{S}(y)$ becomes:

$$\begin{cases} \alpha_{0,0,2} \cdot s_3 + \alpha_{1,0,1} \cdot \hat{z} + \alpha_{0,0,1} = 0 \\ \alpha_{2,0,0} \cdot s_4 + \alpha_{1,0,1} \cdot \hat{z} + \alpha_{1,0,0} = 0 \\ \alpha_{0,0,2} \cdot s_4 \cdot d_4 + \alpha_{1,0,1} \cdot x_2 \cdot d_4 + d_3 + \alpha_{0,0,1} \cdot d_4 = 0 \\ \alpha_{2,0,0} \cdot s_3 \cdot d_3 + \alpha_{1,0,1} \cdot \hat{z} \cdot d_3 + \alpha_{1,0,0} \cdot d_3 + d_4 = 0 \end{cases} .$$

By isolating $\hat{z}$ in the first equation and s_4 in the second, and by substituting $\hat{z}$ in the last equation, we get

$$(\alpha_{2,0,0} - \alpha_{0,0,2}) \cdot s_3 \cdot d_3 + (\alpha_{1,0,0} - \alpha_{0,0,1}) \cdot d_3 + d_4 = 0 \,.$$

If $\alpha_{2,0,0} \neq \alpha_{0,0,2}$ or $\alpha_{1,0,0} \neq \alpha_{0,0,1}$ there exist a pair (s_3, d_4) such that this equality holds and $d_4 \neq 0$. In this case choose x_2 such that the third equation always holds true, and these conditions correspond to a collision.

It remains to find a collision if $\alpha_{2,0,0} = \alpha_{0,0,2}$ and $\alpha_{1,0,0} = \alpha_{0,0,1}$. In this case choose $\hat{z} = 0$, $d_2, d_3, d_4 \neq 0$ and the system becomes:

$$\begin{cases} \alpha_{2,0,0} \cdot s_2 + \alpha_{1,0,0} = 0 \\ \alpha_{2,0,0} \cdot s_4 + \alpha_{1,0,0} = 0 \\ \alpha_{2,0,0} \cdot s_3 \cdot d_3 + d_2 + \alpha_{1,0,0} \cdot d_3 = 0 \\ \alpha_{2,0,0} \cdot s_3 \cdot d_3 + \alpha_{1,0,0} \cdot d_3 + d_4 = 0 \\ \alpha_{2,0,0}(s_2 \cdot d_2 + s_4 \cdot d_4) + \frac{\alpha_{1,0,1}}{2}(s_2 \cdot d_4 + s_4 \cdot d_2) + \alpha_{1,0,0} \cdot (d_2 + d_4) + d_3 = 0 \end{cases} .$$

The first two equations hold if $s_2 = s_4 = -\frac{\alpha_{1,0,0}}{\alpha_{2,0,0}}$. From the difference between the third and fourth $d_2 = d_4$ and choosing $s_3 = -\frac{d_2 + \alpha_{1,0,0} \cdot d_3}{\alpha_{2,0,0} \cdot d_3}$ both equations always hold true. Now rewrite the last equation as

$$(2\alpha_{2,0,0} + \alpha_{1,0,1}) \cdot s_2 \cdot d_2 + 2\alpha_{1,0,0} \cdot d_2 + d_3 = 0 \,.$$

Furthermore $2\alpha_{2,0,0} + \alpha_{1,0,1} = \alpha^{(2)} = 0$, hence the last equation holds true if $d_3 = -2\alpha_{1,0,0} \cdot d_2$ and choosing $d_2 \neq 0$ all the conditions for the collision are verified.

Case: $\alpha_{1,0,1} \neq 0$ and $\alpha_{1,1,0} \neq 0$. This case cannot occur. Indeed, due to the second condition, we have $\alpha_{1,0,1} = -\alpha_{1,1,0}$, and due to the third condition, we have $\alpha_{1,0,1} = -\alpha_{0,1,1}$ (note that $\alpha_{0,1,1} = 0$ cannot occur, since this would imply $\alpha_{1,1,0} = 0$ due to the first condition). Hence, $-\alpha_{1,0,1} = \alpha_{1,1,0} = \alpha_{0,1,1}$. The first equation becomes $2 \cdot \alpha_{1,1,0}^2 = 0$, that is, $\alpha_{1,1,0} = 0$, which is not possible.

6.3.3 Case: $\alpha_{1,0,1} = 0$

Case: $\alpha_{1,0,1} = 0$ and "$\alpha_{0,1,1} = 0$ or/and $\alpha_{1,1,0} = 0$". If $\alpha_{1,1,0}, \alpha_{0,1,1} = 0$, then the result follows immediately from Lemma 1. Hence, let's suppose that at least one term among $\alpha_{1,1,0}$ and $\alpha_{0,1,1}$ is different from zero. In such a case, let's re-write the conditions for $d_2 = d_4 = 0$ and $d_3 \neq 0$ as:

$$\begin{bmatrix} \alpha_{0,0,2} & \alpha_{0,1,1} & 0 \\ \alpha_{0,2,0} & \alpha_{1,1,0} & \alpha_{0,1,1} \\ \alpha_{2,0,0} & 0 & \alpha_{1,1,0} \end{bmatrix} \times \begin{bmatrix} s_3 \\ x_2 \\ x_4 \end{bmatrix} = - \begin{bmatrix} \alpha_{0,0,1} \\ \alpha_{0,1,0} \\ \alpha_{1,0,0} \end{bmatrix} .$$

We study separately the two cases:

- $\alpha_{0,1,1} = 0$ or $\alpha_{1,1,0} = 0$;
- $\alpha_{0,1,1}, \alpha_{1,1,0} \neq 0$.

If $\alpha_{0,1,1} = 0$ and if $\alpha_{0,0,2} \neq 0$ the determinant is non-zero and the system has solution, hence there is a collision. Otherwise if $\alpha_{0,0,2} = 0$ the result follows from Lemma 2, because all the monomials of degree two with x_2 as a factor have coefficients equal to zero in F. Analogous for $\alpha_{1,1,0} = 0$.

Case: $\alpha_{1,0,1} = 0$ and $\alpha_{0,1,1}, \alpha_{1,1,0} \neq 0$ This is the last remaining case to analyze; here we need at least two terms d_i to be non-zero. Let's start with $d_2 = 0$, $d_3, d_4 \neq 0$ and define $A(x_0, x_1, x_2) = \alpha_{1,0,0} \cdot x_0 + \alpha_{0,1,0} \cdot x_1 + \alpha_{0,0,1} \cdot x_2$ as before. The condition for a collision is:

$$\begin{bmatrix} \alpha_{0,0,2} & 0 & \alpha_{0,1,1} & 0 \\ 0 & \alpha_{2,0,0} & 0 & \alpha_{1,1,0} \\ \alpha_{0,2,0} \cdot d_3 + \frac{\alpha_{0,1,1} \cdot d_4}{2} & \alpha_{0,0,2} \cdot d_4 + \frac{\alpha_{0,1,1} \cdot d_3}{2} & \alpha_{1,1,0} \cdot d_3 & 0 \\ \alpha_{2,0,0} \cdot d_3 + \frac{\alpha_{1,1,0} \cdot d_4}{2} & \alpha_{0,2,0} \cdot d_4 + \frac{\alpha_{1,1,0} \cdot d_3}{2} & 0 & \alpha_{0,1,1} \cdot d_4 \end{bmatrix} \times \begin{bmatrix} s_3 \\ s_4 \\ x_2 \\ \hat{z} \end{bmatrix} = - \begin{bmatrix} A(0,0,1) \\ A(1,0,0) \\ A(0,d_3,d_4) \\ A(d_3,d_4,0) \end{bmatrix}.$$

The determinant of the matrix is equal to

$$\begin{aligned} &\left(\alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0} - \alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} - \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0}\right) \cdot \left(\alpha_{1,1,0} \cdot d_3^2 + \alpha_{0,1,1} \cdot d_4^2\right) \\ + &\left(\alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{2,0,0} \cdot \alpha_{0,0,2} - \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0} \cdot \alpha_{0,2,0} - \alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} \cdot \alpha_{0,2,0}\right) \cdot d_3 \cdot d_4. \end{aligned}$$

Note that

- if $\alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} + \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0} - \alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0} \neq 0$ (namely, the coefficient that multiplies $\alpha_{1,1,0} \cdot d_3^2 + \alpha_{0,1,1} \cdot d_4^2$), then it is possible to choose $(d_3, d_4) \neq (0,0)$ such that the determinant is non-zero (e.g., $d_3 = 0$ and $d_4 \neq 0$);
- if $\alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} + \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0} - \alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0} = 0$ and if $\alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{2,0,0} \cdot \alpha_{0,0,2} - \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0} \cdot \alpha_{0,2,0} - \alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} \cdot \alpha_{0,2,0} \neq 0$ (namely, the coefficient that multiplies $d_3 \cdot d_4$), then it is possible to choose $(d_3, d_4) \neq (0,0)$ such that the determinant is non-zero (e.g., $d_3 \neq 0$ and $d_4 \neq 0$).

Similarly, considering now the case $d_3 = 0$ and $d_2, d_4 \neq 0$, there is a collision if $\alpha_{0,0,2} \cdot \alpha_{2,0,0} \cdot (\alpha_{1,1,0}^2 - \alpha_{0,1,1}^2) + \alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0} \cdot (\alpha_{0,0,2} - \alpha_{2,0,0}) \neq 0$. Hence, the only case that remains to analyze is

$$\begin{cases} \alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0} - \alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} - \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0} = 0 \\ \alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{2,0,0} \cdot \alpha_{0,0,2} - \alpha_{0,1,1}^2 \cdot \alpha_{2,0,0} \cdot \alpha_{0,2,0} - \alpha_{1,1,0}^2 \cdot \alpha_{0,0,2} \cdot \alpha_{0,2,0} = 0 \\ \alpha_{0,0,2} \cdot \alpha_{2,0,0} \cdot (\alpha_{1,1,0}^2 - \alpha_{0,1,1}^2) + \alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0} \cdot (\alpha_{0,0,2} - \alpha_{2,0,0}) = 0 \end{cases}.$$

By taking the sum between the first equation and the second one multiplied by $\alpha_{0,2,0}$, it follows that $\alpha_{0,2,0}^2 = \alpha_{0,0,2} \cdot \alpha_{2,0,0}$. By replacing $\alpha_{1,1,0} \cdot \alpha_{0,1,1} \cdot \alpha_{0,2,0}$ in the last equation via the first one, we get $(\alpha_{0,0,2} \cdot \alpha_{1,1,0})^2 = (\alpha_{2,0,0} \cdot \alpha_{0,1,1})^2$, which implies that the previous conditions become:

$$\begin{cases} \alpha_{0,0,2} \cdot \alpha_{1,1,0} = \pm \alpha_{2,0,0} \cdot \alpha_{0,1,1} \\ \alpha_{0,2,0}^2 = \alpha_{0,0,2} \cdot \alpha_{2,0,0} \end{cases}.$$

This system is satisfied if and only if one of the two following events happens:

1. $\alpha_{2,0,0} = \alpha_{0,2,0} = \alpha_{0,0,2} = 0$;
2. $\beta^2 = \pm \frac{\alpha_{0,1,1}}{\alpha_{1,1,0}}$ is a quadratic residue *and* $\alpha_{0,0,2} = \alpha_{2,0,0} \cdot \beta^2$, $\alpha_{0,2,0} = \pm \alpha_{2,0,0} \cdot \beta$.

SubCase: $\alpha_{2,0,0} = \alpha_{0,2,0} = \alpha_{0,0,2} = 0$. We have that $\alpha_{0,1,1} = -\alpha_{1,1,0}$ due to $\alpha^{(2)} = 0$. Fix $d_4 = 0$ and $d_2, d_3 \neq 0$, the conditions for the collision are:

$$\begin{cases} (\alpha_{0,1,1} \cdot \hat{z} + \alpha_{0,0,1}) \cdot d_2 = 0 \\ (-\alpha_{0,1,1} \cdot x_4 + \alpha_{1,0,0}) \cdot d_3 = 0 \\ -\alpha_{0,1,1} \cdot \hat{z} \cdot d_2 + \frac{\alpha_{0,1,1}}{2} \cdot (d_2 \cdot s_3 + d_3 \cdot s_2) + \alpha_{0,1,0} \cdot d_2 + \alpha_{0,0,1} \cdot d_3 = 0 \\ -\frac{\alpha_{0,1,1}}{2} \cdot (d_2 \cdot s_3 + d_3 \cdot s_2) + \alpha_{0,1,1} \cdot x_4 \cdot d_3 + \alpha_{1,0,0} \cdot d_2 + \alpha_{0,1,0} \cdot d_3 = 0 \end{cases}.$$

Since the sum of the four equations is $\alpha^{(1)} \cdot (d_2 + d_3) = 0$, it follows that $d_2 = -d_3$. By substituting this in the third equation, we get $s_2 = s_3 + 2\frac{\alpha_{0,1,0}}{\alpha_{0,1,1}}$. By choosing $d_3 \neq 0$ and a proper s_3, the system admits solution, which corresponds to a collision.

SubCase: $\pm\frac{\alpha_{0,1,1}}{\alpha_{1,1,0}}$ *is a quadratic residue and* $\alpha_{0,0,2} = \alpha_{2,0,0} \cdot \beta^2$, $\alpha_{0,2,0} = \pm\alpha_{2,0,0} \cdot \beta$. We assume $\alpha_{2,0,0} \neq 0$ (since $\alpha_{2,0,0} = 0$ would reduce this case to the previous one). Let's choose $\hat{z} = 0$, $d_2 = d_4 = 0$ and $d_3 \neq 0$. In such a case, the condition for having a collision becomes:

$$\begin{bmatrix} \pm\beta^2 \cdot \alpha_{1,1,0} & \beta^2 \cdot \alpha_{2,0,0} & 0 \\ \alpha_{1,1,0} & \pm\beta^2 \cdot \alpha_{2,0,0} & \pm\beta^2 \cdot \alpha_{1,1,0} \\ 0 & \alpha_{2,0,0} & \alpha_{1,1,0} \end{bmatrix} \times \begin{bmatrix} x_2 \\ s_3 \\ x_4 \end{bmatrix} = - \begin{bmatrix} \alpha_{0,0,1} \\ \alpha_{0,1,0} \\ \alpha_{1,0,0} \end{bmatrix}.$$

The determinant of the matrix is $-\alpha_{1,1,0}^2 \cdot \alpha_{2,0,0} \cdot (1 \pm \beta + \beta^2)$. If $1 \pm \beta + \beta^2 \neq 0$, then the determinant is different from zero and a collision can be found (remember that we are working in the case $\alpha_{1,1,0}, \alpha_{2,0,0} \neq 0$).

Let's focus on the case $1 \pm \beta + \beta^2 = 0$. Note that

$$\alpha^{(2)} = \alpha_{2,0,0} \cdot (1 \pm \beta + \beta^2) + \alpha_{1,1,0} + \alpha_{0,1,1}.$$

Since $\alpha^{(2)} = 0$, we have that $\alpha_{1,1,0} = -\alpha_{0,1,1}$, which implies $\beta^2 = \pm 1$. Hence:

- if $\beta^2 = -1$, then $0 = 1 \pm \beta + \beta^2 = \pm\beta$, which is never satisfied ($0 \neq \pm 1$);
- if $\beta^2 = 1$, then $0 = 1 \pm \beta + \beta^2 = \pm\beta + 2$ hence $\beta = \pm 2$, where $\beta^2 = 4 = 1$ if and only if $p = 3$.

The only remaining case to analyze for concluding the proof is $p = 3$ and $\beta = \pm 1$, that is

$$F(x_0, x_1, x_2) = x_0^2 + x_1^2 + x_2^2 \pm x_0 \cdot x_1 \mp x_1 \cdot x_2 + \alpha_{1,0,0} \cdot x_0 + \alpha_{0,1,0} \cdot x_1 + \alpha_{0,0,1} \cdot x_2\,.$$

Let's focus on the case $\alpha_{1,1,0} = 1$ and $\alpha_{0,1,1} = -1$ (analogous for the other case). By fixing $d_4 = 0$ and $d_2, d_3 \neq 0$, the collision occurs if

$$\begin{cases} (s_2 - \hat{z} + \alpha_{0,0,1}) \cdot d_2 = 0 \\ (s_3 + x_4 + \alpha_{1,0,0}) \cdot d_3 = 0 \\ s_2 \cdot d_2 + s_3 \cdot d_3 + \hat{z} \cdot d_2 + s_2 \cdot d_3 + s_3 \cdot d_2 + \alpha_{0,1,0} \cdot d_2 + \alpha_{0,0,1} \cdot d_3 \\ s_2 \cdot d_2 + s_3 \cdot d_3 - s_2 \cdot d_3 - s_3 \cdot d_2 + x_4 \cdot d_3 + \alpha_{1,0,0} \cdot d_2 + \alpha_{0,1,0} \cdot d_3 \end{cases}.$$

Similar to before, the sum of all the equations is $\alpha^{(1)} \cdot (d_2 + d_3) = 0$, hence $d_3 = -d_2$. By substituting d_3 in the third equation and $\hat{z}$ with the value given by the first equation, the system of equations has a solution if $s_2 = -\alpha_{0,1,0}$. This corresponds to a collision for $\mathcal{S}$.

□

7 Neptune: a Concrete Application

As final step, we present Neptune, a sponge hash function [BDPV07, BDPA08] instantiated with the Neptune$^\pi$ permutation. Neptune$^\pi$ resembles the permutation Poseidon$^\pi$ [GKR+21], and takes into account the results proposed in this paper. In the following, after recalling Poseidon and presenting Neptune as its variant, we discuss its design rationale and its security. Next, we compare the multiplicative complexity of Neptune with the one of Poseidon.

7.1 Poseidon and the Hades Design Strategy

Poseidon is a sponge hash function over $\mathbb{F}_p^t$. Its internal permutation is based on the Hades design strategy [GLR+20], recently proposed at Eurocrypt 2020. The main feature of Hades schemes is the use of two different non-linear layers, namely a *full* one (composed of t power maps $x \mapsto x^d$ for odd $d \geq 3$) in the external rounds, and a *partial* one (composed of a single power map $x \mapsto x^d$ and $t-1$ identity functions) in the internal rounds. This particular structure allows to provide security against both statistical and algebraic attacks, and at the same time to achieve a low multiplicative complexity. In particular, the external rounds aim to prevent statistical attacks as the classical and truncated differential attacks, linear attacks, rebound attacks and so on. The main goal of the partial rounds is to increase the overall degree of the permutation. Together with the external rounds, they provide security against Gröbner basis attacks.

Let $p > 2^{63}$ be a prime number and let $\kappa \in [80, 256]$ be the security level. Let $t \geq 2$ be such that $p^t \geq 2^{3\cdot\kappa}$.[4] Let $d \geq 3$ be the smallest integer such that $\gcd(d, p-1) = 1$. The Poseidon permutation $\widehat{\mathcal{P}} : \mathbb{F}_p^t \to \mathbb{F}_p^t$ is defined as

$$\widehat{\mathcal{P}}(\cdot) = \underbrace{\mathcal{F}^{(7)} \circ \cdots \circ \mathcal{F}^{(4)}}_{=4 \text{ rounds}} \circ \underbrace{\mathcal{P}^{(R_P-1)} \circ \cdots \circ \mathcal{P}^{(0)}}_{=R_P \text{ rounds}} \circ \underbrace{\mathcal{F}^{(3)} \circ \cdots \circ \mathcal{F}^{(0)}}_{=4 \text{ rounds}}(\cdot),$$

where

$$\mathcal{F}^{(j)} = c^{(F,j)} + M \times \mathcal{S}^{(F)}(\cdot) \qquad \text{and} \qquad \mathcal{P}^{(j)} = c^{(P,j)} + M \times \mathcal{S}^{(P)}(\cdot)$$

and where $c^{(F,j)}, c^{(P,j)}$ are (random) round constants, $M \in \mathbb{F}_p^{t\times t}$ is a MDS matrix and $\mathcal{S}^{(F)}, \mathcal{S}^{(P)} : \mathbb{F}_p^t \to \mathbb{F}_p^t$ are defined as

$$\mathcal{S}^{(F)}(x_0, \ldots, x_{t-1}) = x_0^d \| x_1^d \| \ldots \| x_{t-1}^d, \quad \mathcal{S}^{(P)}(x_0, \ldots, x_{t-1}) = x_0^d \| x_1 \| \ldots \| x_{t-1}.$$

The number of full rounds is $R_F = 8$ and the number of partial rounds is $R_P = \lceil 1.125 \cdot \lceil \log_d(2) \cdot (\min\{\kappa, \log_2(p)\} - 8) + \log_d(t) \rceil \rceil$

In [BCD+20], distinguishers for this reduced-round permutation were presented, which lead to collision attacks on the sponge hash function instantiated with the reduced-round permutation $\widehat{\mathcal{P}}$. Moreover, in the same paper, authors were able to set up preimage attacks on the sponge hash function instantiated with the full-round permutation $\widehat{\mathcal{P}}$ in the case of *a weak MDS matrix M such that M^2 is a multiple of the identity*, and so, for which an invariant subspace trail that covers all the internal rounds with probability 1 exists (see also [KR21]). In [GRS21], Grassi et al. showed how to properly choose the MDS matrix M in order to prevent this (and similar) attack(s).

7.2 Neptune

Let $\kappa \in [80, 256]$ be the security level, and let $p > 2^{63}$ be a prime number. Let $t = 2t' \in \{2, 4, \ldots, 24\}$ be an even integer. Since Neptune is intended to be used as the internal

[4] Given $t = c + r$, the capacity c and the rate r satisfy respectively $p^c \geq 2^{2\cdot\kappa}$ and $p^r \geq 2^\kappa$ for κ bits of security.

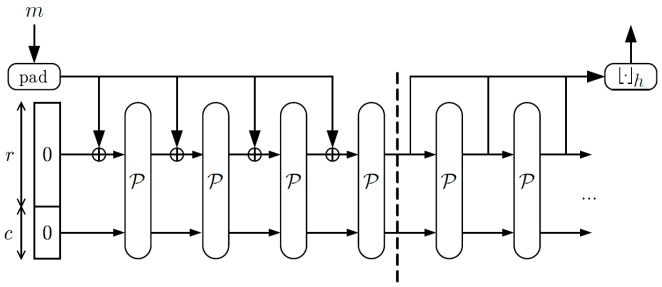

Figure 1: A sponge hash function instantiated with a permutation $\mathcal{P}$.

permutation of a sponge hash function, the parameters p, κ and t have to satisfy (1st) $p^c \geq 2^{2\cdot\kappa}$ and (2nd) $p^r \geq 2^\kappa$, where r and c are respectively the rate and the capacity such that $t = c + r$. A sponge hash function instantiated by a generic permutation $\mathcal{P}$ is shown in Fig. 1. About the *padding*, we suggest to use the same one proposed e.g. in POSEIDON. That is, (1st) the message m is padded with 0^* until the size of $m\|0^*$ is a multiple of r and (2nd) the inner part is initially instantiated as $\texttt{IV} = |m|\|\texttt{IV}' \in \mathbb{F}_p^c$, where $|m|$ is the size of the input message m as an element of $\mathbb{F}_p$ and where $\texttt{IV}' \in \mathbb{F}_p^{c-1}$ is a fixed initial value.

The NEPTUNE permutation $\widehat{\mathcal{N}} : \mathbb{F}_p^t \to \mathbb{F}_p^t$ is defined as[5]

$$\widehat{\mathcal{N}}(\cdot) = \underbrace{\mathcal{E}^{(5)} \circ \mathcal{E}^{(4)}}_{=2 \text{ rounds}} \circ \underbrace{\mathcal{I}^{(R_I-1)} \circ \cdots \circ \mathcal{I}^{(0)}}_{=R_I \text{ rounds}} \circ \underbrace{\mathcal{E}^{(3)} \circ \cdots \circ \mathcal{E}^{(0)}}_{=4 \text{ rounds}} (M^{(E)} \times \cdot),$$

where

$$\mathcal{E}^{(j)} = c^{(E,j)} + M^{(E)} \times \mathcal{S}^{(E)}(\cdot) \qquad \text{and} \qquad \mathcal{I}^{(j)} = c^{(I,j)} + M^{(I)} \times \mathcal{S}^{(I)}(\cdot)$$

and where $c^{(E,j)}, c^{(I,i)} \in \mathbb{F}_p^t$ are (random) round constants.

About the External Rounds $\mathcal{E}$. The non-linear $\mathcal{S}^{(E)} : \mathbb{F}_p^t \to \mathbb{F}_p^t$ is defined as

$$\mathcal{S}^{(E)} = \mathcal{S}'(x_0, x_1)\|\mathcal{S}'(x_2, x_3)\| \ldots \|\mathcal{S}'(x_{t-2}, x_{t-1})$$

where $\mathcal{S}' : \mathbb{F}_p^2 \to \mathbb{F}_p^2$ is defined as $\mathcal{S}'(x_{2i}, x_{2i+1}) = y_{2i}\|y_{2i+1}$ for $i \in \{0, 1, \ldots, t'-1\}$ where

$$y_{2i} = \alpha^2 \cdot (2x_{2i} + x_{2i+1}) + 3\alpha \cdot (x_{2i} - x_{2i+1})^2 + \left(\gamma + \alpha \cdot (x_{2i} - 2x_{2i+1}) - (x_{2i} - x_{2i+1})^2\right)^2$$
$$y_{2i+1} = \alpha^2 \cdot (x_{2i} + 3x_{2i+1}) + 4\alpha \cdot (x_{2i} - x_{2i+1})^2 + \left(\gamma + \alpha \cdot (x_{2i} - 2x_{2i+1}) - (x_{2i} - x_{2i+1})^2\right)^2$$

for fixed $\alpha, \gamma \in \mathbb{F}_p \setminus \{0\}$ (e.g., $\alpha = 1$ and $\gamma \neq 0$).

Let $M', M'' \in \mathbb{F}_p^{t' \times t'}$ be two MDS matrices such that

1. $M' \neq \mu \cdot M''$ for each $\mu \in \mathbb{F}_p$;
2. for each $i, j \in \{0, 1, \ldots, t'-1\}$: $M'_{i,j} \neq M''_{i,j}$.

The matrix $M^{(E)} \in \mathbb{F}_p^{t \times t}$ is defined as

$$M^{(E)}_{i,j} = \begin{cases} M'_{i',j'} & \text{if } (i,j) = (2i', 2j') \\ M''_{i'',j''} & \text{if } (i,j) = (2i''+1, 2j''+1) \\ 0 & \text{otherwise} \end{cases},$$

[5] In [GLR+20, GKR+21], authors use the nomenclature "Full" and "Partial" rounds for referring respectively to the "External" and the "Internal" rounds.

that is,

$$M^{(E)} = \begin{bmatrix} M'_{0,0} & 0 & M'_{0,1} & 0 & \dots & M'_{0,t'-1} & 0 \\ 0 & M''_{0,0} & 0 & M''_{0,1} & \dots & 0 & M''_{0,t'-1} \\ M'_{1,0} & 0 & M'_{1,1} & 0 & \dots & M'_{1,t'-1} & 0 \\ 0 & M''_{1,0} & 0 & M''_{1,1} & \dots & 0 & M''_{1,t'-1} \\ \vdots & & & & \ddots & & \vdots \\ M'_{t'-1,0} & 0 & M'_{t'-1,1} & 0 & \dots & M'_{t'-1,t'-1} & 0 \\ 0 & M''_{t'-1,0} & 0 & M''_{t'-1,1} & \dots & 0 & M''_{t'-1,t'-1} \end{bmatrix}.$$

About the Internal Round $\mathcal{I}$. The internal round $\mathcal{I}$ is defined via a Partial-SPN scheme as in Poseidon, where

$$\mathcal{S}^{(I)}(x_0, x_1, \dots, x_{t-2}, x_{t-1}) = x_0^d \| x_1 \| \dots \| x_{t-2} \| x_{t-1}$$

where $d \geq 3$ is the *smallest* integer such that $\gcd(d, p-1) = 1$, and where $M^{(I)} \in \mathbb{F}_p^{t \times t}$ is an invertible matrix that

1. must prevent arbitrary-long subspace trails for the Partial-SPN scheme $\mathcal{I}^{(R_I-1)} \circ \dots \circ \mathcal{I}^{(0)}$, as explained in [GRS21];
2. can be computed via $\mathcal{O}(t)$ affine operations.

A *possible* example of a matrix $M^{(I)}$ that satisfies such conditions is

$$M^{(I)} = \begin{bmatrix} M^{(I)}_{0,0} & 1 & 1 & \dots & 1 & 1 \\ 1 & M^{(I)}_{1,1} & 1 & \dots & 1 & 1 \\ 1 & 1 & M^{(I)}_{2,2} & \dots & 1 & 1 \\ \vdots & & & \ddots & & \vdots \\ 1 & 1 & 1 & \dots & M^{(I)}_{t-2,t-2} & 1 \\ 1 & 1 & 1 & \dots & 1 & M^{(I)}_{t-1,t-1} \end{bmatrix}$$

where $M^{(I)}_{i,i} \in \mathbb{F}_p \setminus \{0\}$ are chosen in order to guarantee the previous requirements, for a cost of t multiplications with constants.

Number of Rounds. The number of rounds are $R_F = 6$ for the external ones (that is, 4 for at the beginning and 2 at the end) and

$$R_I = \lceil 1.125 \cdot \lceil \log_d(2) \cdot (\min\{\kappa, \log_2(p)\} - 6) + 3 + t + \log_d(t) \rceil \rceil$$

for the internal ones (where we add 12.5% of security margin, as in Poseidon).

7.3 Design Rationale

By a simple computation, the number of $\mathbb{F}_p$-multiplications required to evaluate Poseidon is

$$(\lfloor \log_2(d) \rfloor + \mathrm{hw}(d) - 1) \cdot (8 \cdot t + R_P),$$

that is $\mathcal{O}(16 \cdot t)$ for $d = 3$ and $\mathcal{O}(24 \cdot t)$ for $d = 5$ (where $d = 3, 5$ are the two most common values used in ZK protocols). In order to design Neptune, we decided to focus only on the external full rounds, since we noticed that the number of internal partial rounds is almost constant with respect to t. For this reason, we decided not to modify them. Regarding the external rounds and in order to make use of the results proposed in this paper, the goals we tried to achieve were:

1. having a full round that does not cost more than t $\mathbb{F}_p$-multiplications;
2. be able to guarantee security against statistical attacks via a small number of full external rounds.

As a result, instead of limiting ourselves to consider an uneven distribution of the S-Boxes, we propose two different round functions, one for the internal part and one for the external one.

Open Conjectures for Future Work. As we have already seen, given any quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for $m = 2, 3$, the corresponding function $\mathcal{S}$ defined over $\mathbb{F}_p^n$ as in Def. 1 is not invertible for $n \geq 3$ and $n \geq 5$ respectively. We conjecture that the same occurs for bigger values of m. More formally:

Conjecture 1. *Let $p \geq 3$ be a prime integer, and let $1 \leq m \leq n$. For each m, there exists a **finite** integer $n_{max}(m)$ such that given any quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$, the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 is **not** invertible for any $n \geq n_{max}(m)$.*

E.g., if $m = 1$, then $n_{\max} = 1$; if $m = 2$, then $n_{\max} = 3$; if $m = 3$, then $n_{\max} = 5$. Moreover, based on the result proposed in Sect. 3.3, $n_{\max}(m) \geq m + 1$ for each $m \geq 2$. Indeed, given a quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$, the Lai-Massey functions defined over $\mathbb{F}_p^m$ as in Sect. 3.3 are invertible.

If the conjecture is true, it would be interesting to analyze how fast $n_{\max}(m)$ grows. The current results for $m \in \{1, 2, 3\}$ suggest that

$$n_{\max}(m) = 2 \cdot m - 1 \,.$$

By applying Corollary 2 on a generic m, we can construct an invertible function $\mathcal{S}_F$ over $\mathbb{F}_p^{2\cdot(m-1)}$ via a quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ (e.g., $F(x_0, x_1, \ldots, x_{m-1}) = x_0 + (x_0 - x_{m-1})^2$). Such a result is not in conflict with $n_{\max}(m) = 2 \cdot m - 1$ just given. The same happens when applying Prop. 11 to the results proposed in this paper. E.g., in the case $g = h \geq 2$ (which include both the Lai-Massey constructions proposed in Sect. 3.3, as well as the functions proposed in Prop. 10 and in Prop. 9 for $g = h = 3$), we get $m = g+(g-1)\cdot z$ and $n = g \cdot (z+1)$ for any $z \geq 0$, where $2 \cdot (m-1) = 2 \cdot (z+1) \cdot (g-1) \geq n$ for each $g \geq 2$, which is again not in conflict with $n_{\max}(m) = 2 \cdot m - 1$ just given.

Conjecture 2. *Let $n_{max}(m)$ be defined as in Conjecture 1. Then, $n_{max}(m) = 2 \cdot m - 1$.*

If the conjecture "$n_{max}(m) = 2 \cdot m - 1$" is true, this implies that given a local quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$, it is not possible to set up an invertible function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Def. 1 for $n \gg m$.

Concatenation of Independent S-Boxes. At the current state, we do not know any (non-trivial) quadratic function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for which it is possible to set up an invertible function $\mathcal{S}$ over $\mathbb{F}_p^n$ as in Def. 1 for $n \gg m$. For this reason, we are "forced" to set up the non-linear layer of the external rounds as a concatenation of independent quadratic S-Boxes defined either over $\mathbb{F}_p^2$ or over $\mathbb{F}_p^3$.

Based on our results, possible options for $F : \mathbb{F}_p^m \to \mathbb{F}_p$ include:

- Lai-Massey constructions, as $F(x_0, x_1) = x_0 + (x_0 - x_1)^2$ or $F(x_0, x_1, x_2) = x_0 + \sum_{j=0}^{2}(x_j - x_{j+1})^2$;
- if $p = 1 \mod 3$: $F(x_0, x_1, x_2) = x_0 + \alpha \cdot (x_0 - x_1)^2 + \beta \cdot (x_1 - x_2)^2 + \gamma \cdot (x_2 - x_0)^2$ as in Prop. 10; otherwise, if $p = 2 \mod 3$: $F(x_0, x_1, x_2) = x_0 + (x_0 + x_1 + x_2) \cdot (\alpha \cdot x_0 + \beta \cdot x_1 - (\alpha + \beta) \cdot x_2)$ as in Prop. 9.

We decided to discard the last two functions, since they would force us to consider separately the case $p = 1 \mod 3$ from the case $p = 2 \mod 3$. Regarding the first option, the two Lai-Massey functions admit invariant subspaces, that is, there exists a subspace $\mathfrak{X} \subset \mathbb{F}_p^m$ which is invariant through the non-linear function. E.g., $\langle [1,1]^T \rangle$ is invariant for the case $m = 2$, while $\langle [1,1,0]^T \rangle, \langle [1,0,1]^T \rangle, \langle [0,1,1]^T \rangle$ (and their linear combinations) are invariant for the case $m = 3$. We opted for the smallest m, since it also allows to cover a larger range of values of t, besides the fact that it admits a smaller number of invariant subspaces.

Let $F(x_0, x_1) = \alpha \cdot x_0 + (x_0 - x_1)^2$ for $\alpha \in \mathbb{F}_p \setminus \{0\}$, and let $\mathcal{S}_F$ over $\mathbb{F}_p^2$ be defined as in Def. 1. Due to the presence of the invariant subspace $\langle [1,1]^T \rangle$, we do not use $\mathcal{S}_F$ directly, but we consider $\mathcal{S}'(x_i, x_{i+1})$ defined as

$$\mathcal{S}'(x_i, x_{i+1}) = \begin{bmatrix} -\alpha \cdot \gamma \\ 0 \end{bmatrix} + \mathcal{S}_F \circ \left(\begin{bmatrix} \gamma \\ 0 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \times \mathcal{S}_F(x_i, x_{i+1}) \right). \tag{11}$$

The invertible matrix $[2, 1; 1, 3]$ and the vector $[\gamma; 0]$ (for $\gamma \neq 0$) have been chosen in order to destroy the invariant subspace $\langle [1,1]^T \rangle$. Note that $\mathcal{S}'$ over $\mathbb{F}_p^2$ costs 2 $\mathbb{F}_p$-multiplications, which implies that $\mathcal{S}^{(E)}$ over $\mathbb{F}_p^t$ costs t $\mathbb{F}_p$-multiplications.

The Linear Layer $M^{(E)}$. The S-Box $\mathcal{S}'$ over $\mathbb{F}_p^2$ mixes two $\mathbb{F}_p$-words in a non-linear way. Hence, it is not necessary to instantiate the linear layer with e.g. a $t \times t$ MDS matrix in order to achieve both full diffusion and a high number of active S-Boxes over two consecutive rounds. Indeed, it is not hard to check that such goal can be achieved by mixing only the first output components of the S-Boxes among them via a MDS matrix M', and independently only the second output components of the S-Boxes among them via a different MDS matrix M''. This is exactly the definition of $M^{(E)}$, for which half of the components are equal to zero. Moreover, it is not hard to check that $M^{(E)}$ is invertible.

Low-Degree Inverse. By considering only the external rounds, a concrete drawback of the quadratic Lai-Massey function regards the fact that its degree is low both in the forward and in the backward direction. For this reason, the partial rounds instantiated with an invertible power map – which has low degree in e.g. the forward direction and high degree in the backward one – play a crucial role in order to stop Meet-in-the-Middle (MitM) attacks. Indeed, we recall that the inverse $x \mapsto x^{d'}$ of $x \mapsto x^d$ satisfies $(d \cdot d' - 1) \mod (p - 1) = 0$ (due to Fermat's little theorem $x^p = x \mod p$ for each $x \in \mathbb{F}_p \setminus \{0\}$), which implies that d' is of approximately the same order of p (for small values of d).

Initial Matrix Multiplication. With respect to Poseidon, we emphasize that the input of Neptune$^\pi$ is multiplied by $M^{(E)}$ before the first S-Box layer is applied. This could make a difference in the case of algebraic attacks. Indeed, remember that the *invertible* S-Box layer is defined via the concatenation of independent non-linear functions. If no initial diffusion/matrix multiplication takes place, one can ignore the first S-Box layer (by replacing the initial value `IV` with the corresponding output via the S-Box layer), with the result of making the attack independent of the first S-Box layer, and so of its degree. Once a solution is found at the output of the first S-Box layer, it is possible to invert it in order to find the corresponding solution at the input of the permutation and so of the hash function. A similar scenario could occur at the end of the permutation if no full diffusion takes place.

7.4 Security Analysis

Due to the similarities between Poseidon and Neptune, we emphasize that (almost) all the attacks work in the same way for the two schemes. This means that we are going to adapt the security analysis of Poseidon to Neptune.

7.4.1 (Invariant) Subspace Trails for the Internal Rounds

As already pointed out in [BCD+20, KR21], there exist several subspaces of $\mathbb{F}_p^t$ that are invariant through the internal rounds of Poseidon and so of Neptune. The matrix $M^{(I)}$ plays a crucial role in order to destroy them.

Definition 6 (*(Invariant) Subspace Trail* [LAAZ11, LMR15, GRR16])**.** Let $(\mathfrak{U}_0, \ldots, \mathfrak{U}_l)$ denote a set of $l+1$ subspaces of $\mathbb{F}_p^t$ with $\dim(\mathfrak{U}_i) \leq \dim(\mathfrak{U}_{i+1})$. $(\mathfrak{U}_0, \ldots, \mathfrak{U}_l)$ is a *subspace trail* of length l with respect to the function $\mathcal{R}$ defined over $\mathbb{F}_p^t$ if for each $i \in \{0, \ldots, l\}$ and for each $\alpha_i \in \mathbb{F}_p^t$ there exists $\alpha_{i+1} \in \mathbb{F}_p^t$ such that

$$\mathcal{R}(\mathfrak{U}_i + \alpha_i) := \{\mathcal{R}(x) \mid x \in \mathfrak{U}_i + \alpha_i\} \subseteq \mathfrak{U}_{i+1} + \alpha_{i+1}\,.$$

If $\mathfrak{U}_i = \mathfrak{U}_j$ for each $i, j = 0, \ldots, l$ (that is, the subspace is invariant), the trail is called an *invariant subspace trail*.

Following Poseidon, for each $i \geq 0$, let's define the subspace $\mathfrak{X}_i \subseteq \mathbb{F}_p^t$ as

$$\mathfrak{X}_i := \left\{ x \in \mathbb{F}_p^t \mid \forall j \leq i : \quad \left((M^{(I)})^j \times x \right)_0 \in \mathbb{F}_p \right\}.$$

As shown in [GRS21, GSW+21], the matrix $M^{(I)}$ must be chosen in order to guarantee that no subspace $\mathfrak{X}_i$ is invariant for an arbitrary number of internal rounds, and more generally, that no subspace trail can cover any arbitrary number of internal rounds. We suggest to use the tool presented in [GRS21] in order to properly choose the matrix $M^{(I)}$ for this goal. This implies that e.g. no more than $t-1$ internal rounds can be covered without activating any S-Box $x \mapsto x^d$.

7.4.2 Statistical Attacks

The external rounds aim to provide security against statistical attacks. Working as in HadesMiMC or as in Poseidon (see [GLR+20, Sect. 4.2] for details), the idea is that the permutation composed of the external rounds only (that is, with the internal rounds replaced by an invertible linear layer) resists statistical attacks. Here we focus on (truncated) differential and rebound attacks. As in Poseidon, the security against these attacks implies the security against other statistical attacks, as the linear one [Mat93], impossible differential [Knu98, BBS99], integral one [DKR97], zero-correlation linear one [BR11, BR14], multiple-of-n/mixture differential [GRR17, Gra18], and so on.

Differential Attacks. Given pairs of inputs with some fixed input differences, differential cryptanalysis [BS93] considers the probability distribution of the corresponding output differences produced by the cryptographic primitive. Let $\delta, \Delta \in \mathbb{F}_p^n$ be respectively the input and the output differences through a permutation $\mathcal{P}$ over $\mathbb{F}_p^n$. The differential probability (DP) of having a certain output difference Δ given a particular input difference δ is equal to

$$\mathrm{Prob}_{\mathcal{P}}(\delta \to \Delta) = \frac{|\{x \in \mathbb{F}_p^n \mid \mathcal{P}(x+\delta) - \mathcal{P}(x) = \Delta\}|}{p^n}.$$

In the case of iterated schemes, a cryptanalyst searches for ordered sequences of differences over any number of rounds that are called differential characteristics/trails. Assuming the independence of the rounds, the DP of a differential trail is the product of the DPs of its one-round differences.

Definition 7. Let $\mathcal{P}$ be a permutation over $\mathbb{F}_{p^n} \equiv \mathbb{F}_p^n$. Its maximum differential probability is defined as $\mathrm{DP}_{\max} = \max_{\delta, \Delta \in \mathbb{F}_p^n \setminus \{0\}} \mathrm{Prob}_{\mathcal{P}}(\delta \to \Delta)$.

As it is well known, the maximum differential probability of the function $x \mapsto x^d$ is $(d-1)/p$. Regarding the function $\mathcal{S}'$, we prove the following result.

Lemma 4. *Let $p \geq 3$, and let $\mathcal{S}' : \mathbb{F}_p^2 \to \mathbb{F}_p^2$ be defined as in Def. 11. Let $\delta \equiv (\delta_0, \delta_1) \in \mathbb{F}_p^2 \setminus \{(0,0)\}$ and $\Delta \equiv (\Delta_0, \Delta_1) \in \mathbb{F}_p^2 \setminus \{(0,0)\}$ be respectively the input and the output differences. Then:*

$$\frac{|\{x \in \mathbb{F}_p^2 \mid \mathcal{S}'(x+\delta) - \mathcal{S}'(x) = \Delta\}|}{p^2} = \begin{cases} p^{-2} & \text{if } \delta_0 \neq \delta_1 \textbf{ and } \Delta_0 \neq \Delta_1 \\ p^{-1} & \text{if } \delta_0 = \delta_1 \textbf{ or } \Delta_0 = \Delta_1 \\ 0 & \text{if } \delta_0 = \delta_1 \textbf{ and } \Delta_0 = \Delta_1 \end{cases}.$$

In other words, its maximum differential probability is p^{-1}. The proof is given in App. C.1.

Working over two consecutive rounds, the minimum number of active S-Boxes is $t'+1$, due to the fact that (1st) both M' and M'' are MDS matrices (with branch number equal to $t'+1 = t/2+1$) and (2nd) they are "independent", in the sense that they work over independent t' $\mathbb{F}_p$-words. This means that the overall probability of each differential trail over two consecutive rounds per three times is at most

$$p^{-3(t'+1)} = p^{-3t/2-3} \leq p^{-3} \cdot 2^{-9\kappa/2} \ll 2^{-4\cdot\kappa}.$$

since $t = 2t'$ and $p^t = p^c \cdot p^r \geq 2^{3\kappa}$. As a result, when targeting a security level of κ bits, two consecutive rounds per three times are sufficient for preventing classical differential attacks.

By considering the internal rounds as well (as suggested in e.g. [KR21]), we point out that the probability of every differential trail is even smaller, more precisely it is at most

$$p^{-3(t'+1)} \cdot \left(\frac{d-1}{p}\right)^{\left\lfloor \frac{R_I}{t} \right\rfloor}$$

(where $\left\lfloor \frac{R_I}{t} \right\rfloor \geq 1$) due the fact that at least one S-Box $x \mapsto x^d$ is active every t internal rounds.

Truncated Differential and Rebound Attacks. Truncated differential [Knu94] is a variant of classical differential attack in which the attacker can specify only part of the difference between pairs of texts. In the particular case of an hash function, truncated differentials can be exploited in order to set up rebound attacks [MRST09]. The goal of this attack is to find two (input, output) pairs such that the two inputs satisfy a certain (truncated) input difference and the corresponding outputs satisfy a certain (truncated) output difference.

Due to the choice of the matrix $M^{(E)}$ and working as in Poseidon (see [GKR+21, Sect. 5.5.1] for details), no truncated differential (equivalently, subspace trail) with probability 1 can cover more than a single round. In particular, while the S-Box $\mathcal{S}'$ is defined over $\mathbb{F}_{p^2} \equiv \mathbb{F}_p^2$, we point out that the matrix $M^{(E)}$ does *not* admit an equivalent representation over $\mathbb{F}_{p^2}^{t' \times t'}$. Indeed, consider the field $\mathbb{F}_{p^2} = \mathrm{GF}(p)[x]/P(x)$, where P is an irreducible polynomial of the form $P(x) = x^2 - \eta$ where $L_p(\eta) = -1$. The product of two elements $a \cdot x + b$ and $c \cdot x + d$ is given by

$$(a \cdot x + b) \cdot (c \cdot x + d) = ac \cdot x^2 + (ad + bc) \cdot x + bd \equiv (ad + bc) \cdot x + (bd + \eta \cdot ac),$$

that is

$$\begin{bmatrix} ad + bc \\ bd + \eta \cdot ac \end{bmatrix} = \begin{bmatrix} b & a \\ \eta \cdot a & b \end{bmatrix} \times \begin{bmatrix} c \\ d \end{bmatrix}. \tag{12}$$

It is simple to observe that each 2×2 sub-matrix of $M^{(E)}$

$$\begin{bmatrix} M'_{i,j} & 0 \\ 0 & M''_{i,j} \end{bmatrix}$$

is of the form Eq. (12) if and only $M'_{i,j} = M''_{i,j}$, which can never hold due to the definition of M', M''.[6]

Due to these facts and working as in Poseidon (for which both the S-Boxes and the matrix multiplications are defined over the same field $\mathbb{F}_p$), we conjecture that six external rounds are sufficient for preventing rebound attacks.

7.4.3 Algebraic Attacks

Interpolation Attacks. The interpolation attack [JK97] aims to construct an interpolation polynomial that describes the function. Such polynomial can be used in order to set up a distinguisher and/or an attack on the symmetric scheme. The attack does not work if the number of unknown monomials is sufficiently large (e.g., larger than the data available for the attack). In the MitM scenario, the attacker construct two polynomials, that is, one that involves the input(s) and one that involve the output(s), that must match in the middle.

Due to the presence of the map $x \mapsto x^d$ in the internal rounds, the final two full rounds combined with three internal rounds ensure maximum degree in the backward direction (remember that $1/d \equiv d'$ such that $(d' \cdot d - 1) \mod (p-1) = 0$, so d' is of the same order of p). Working as in Poseidon (see [GKR+21, Sect. 5.5.2] for details) and in order to guarantee security against the interpolation attack, the number of internal rounds R_I must satisfy

$$4^3 \cdot d^{R_I - 3} \geq \min\{p, 2^\kappa\} \quad \rightarrow \quad R_I \geq 3 + \log_d(2) \cdot (\min\{\kappa, \log_2(p)\} - 6),$$

where (1st) the two final rounds and 3 internal rounds are necessary for reaching maximum degree in the backward direction and (2nd) the first round is not taken into account, since no full diffusion is achieved. Finally, we add t internal rounds due to the possibility to cover them with an invariant subspace trails (which would imply no degree growth), and $\log_d(t)$ additional internal rounds in order to ensure that the polynomial is dense.

Before going on, we recall that the security against interpolation attack implies security against higher-order differential attack [Lai94, Knu94], due to the results presented in [BCD+20, Prop. 1].

Factorization and Gröbner Basis Attacks. Polynomial factorization can be used to solve a single univariate equation $F(x) = 0$ for a polynomial F over $\mathbb{F}_p$. E.g., in the case $r \geq 1$, factorization can be used to find a pre-image of $h \in \mathbb{F}_p$, by solving $F(x) = \left[\widehat{\mathcal{N}}(x \| \hat{v} \| \texttt{IV})\right]_0 - h = 0$ for a fixed $\hat{v} \in \mathbb{F}_p^{r-1}$, where $\texttt{IV} \in \mathbb{F}_p^c$ is the initial value that instantiates the inner part. In such a case, it is actually not necessary to find the full factorization of the polynomial, since one root is sufficient for setting up the attack. The cost of finding a root is proportional to the degree Δ of the polynomial F, more precisely

$$\Delta \cdot (\log_2(\Delta))^2 \cdot (\log_2(\Delta) + \log_2(p)) \cdot (1 + 64 \cdot \log_2(\log_2(\Delta)))$$

as shown in [vzGG13]. It is easy to check that security against interpolation attack implies security against this attack as well.

Gröbner basis [Buc76] generalizes factorization, and it allows to solve a system of non-linear equations that describe the function. As we explain in App. C.2, the cost of

[6] For completeness, we mention that also the matrix $[2, 1; 1, 3]$ that defines the S-Boxes does not admit an equivalent representation in $\mathbb{F}_{p^2}$.

Table 1: Comparison between Poseidon and Neptune – both instantiated with $d = 3$ – for the case $p \approx 2^{256}$, $\kappa = 128$ and several values of t.

	t	R_F	R_P & R_I	Multiplicative Complexity
Poseidon ($d = 3$)	4	8	87	238 (+ 10.2 %)
Neptune ($d = 3$)	4	6	96	**216**
Poseidon ($d = 3$)	8	8	88	304 (+ 21.6 %)
Neptune ($d = 3$)	8	6	101	**250**
Poseidon ($d = 3$)	12	8	88	368 (+ 29.6 %)
Neptune ($d = 3$)	12	6	106	**284**
Poseidon ($d = 3$)	16	8	89	434 (+ 36.5 %)
Neptune ($d = 3$)	16	6	111	**318**
Poseidon ($d = 3$)	20	8	89	498 (+ 42.3 %)
Neptune ($d = 3$)	20	6	115	**350**
Poseidon ($d = 3$)	24	8	89	562 (+ 46.4 %)
Neptune ($d = 3$)	24	6	120	**384**

Table 2: Comparison between Poseidon and Neptune – both instantiated with $d = 5$ – for the case $p \approx 2^{256}$, $\kappa = 128$ and several values of t.

	t	R_F	R_P & R_I	Multiplicative Complexity
Poseidon ($d = 5$)	4	8	60	276 (+ 21.0 %)
Neptune ($d = 5$)	4	6	68	**228**
Poseidon ($d = 5$)	8	8	60	372 (+ 40.1 %)
Neptune ($d = 5$)	8	6	72	**264**
Poseidon ($d = 5$)	12	8	61	471 (+ 53.9 %)
Neptune ($d = 5$)	12	6	78	**306**
Poseidon ($d = 5$)	16	8	61	567 (+ 64.3 %)
Neptune ($d = 5$)	16	6	83	**345**
Poseidon ($d = 5$)	20	8	61	663 (+ 74.0 %)
Neptune ($d = 5$)	20	6	87	**381**
Poseidon ($d = 5$)	24	8	61	759 (+ 80.7 %)
Neptune ($d = 5$)	24	6	92	**420**

such an attack depends both on the number and on the degree of the equations, on the number of variables, but also on the fact that the equations to solve are dense or not. In [GKR+21, Sect. 5.5.2], authors showed that the security of Poseidon against the interpolation attack implies the security against Gröbner basis attacks. As one may expect, in App. C.2, we show that the same conclusion holds for Neptune as well, due to the similarity between the internal rounds of Neptune and the ones of Poseidon.

7.5 Multiplicative Complexity: Poseidon versus Neptune

With these results in mind, we finally compare the multiplicative complexity between Poseidon and Neptune. By simple computation:

- Neptune requires

$$(6 + (\lfloor \log_2(d) \rfloor + \mathrm{hw}(d) - 1)) \cdot t + (\lfloor \log_2(d) \rfloor + \mathrm{hw}(d) - 1) \cdot (R_I - t)$$

  $\mathbb{F}_p$-multiplications, where $(R_I - t)$ is almost constant with respect to t;

- Poseidon requires

$$(\lfloor \log_2(d) \rfloor + \mathrm{hw}(d) - 1) \cdot (8 \cdot t + R_P)$$

  $\mathbb{F}_p$-multiplications, where again R_P is almost constant with respect to t.

Note that $R_P \approx R_I - t$. In the case of large $t \gg 1$ and for $d = 3$, Neptune requires $\mathcal{O}(8 \cdot t)$ $\mathbb{F}_p$-multiplications versus $\mathcal{O}(16 \cdot t)$ $\mathbb{F}_p$-multiplications required by Poseidon. Similarly, in the case of large $t \gg 1$ and for $d = 5$, Neptune requires $\mathcal{O}(9 \cdot t)$ $\mathbb{F}_p$-multiplications versus $\mathcal{O}(24 \cdot t)$ $\mathbb{F}_p$-multiplications required by Poseidon. More concretely, a comparison between the two schemes for small values of t is proposed in Table 2 for the case $p \approx 2^{256}$. As it is possible to observe, Neptune has always a smaller multiplicative complexity with respect to Poseidon.[7]

Besides that, we point out that

- the matrix multiplication of each external round of Neptune costs $t^2/2$ multiplications with constants[8] *versus* t^2 multiplications with constants in the case of Poseidon (besides the fact that Neptune has two external/full rounds less than Poseidon);
- in Poseidon, the same matrix M is used for the full/external rounds and for the partial/internal ones. Since such matrix must prevent arbitrary-long subspace trails with probability 1 for the partial/internal rounds, it cannot be instantiated with e.g. a circulant matrix. Vice-versa, the MDS matrices M', M'' in the external rounds of Neptune do not have to satisfy such requirement. Hence, they can be instantiated with e.g. $\mathrm{circ}(2, 1, 1)$ or $\mathrm{circ}(3, 2, 1, 1)$ for $t' \in \{3, 4\}$ respectively;
- both Neptune and Poseidon admit an equivalent representation in which the matrix multiplication of each internal/partial round costs $2 \cdot t$ multiplications with constants (for more details, we refer to [GLR+19, GLR+20, App. C]). However, in such representation, the matrix of the internal/partial round is not fixed, that is, changes at every round. Without using such equivalent representation, the matrix multiplication of each internal round of Neptune can cost only t multiplications with constants, besides being fixed.

These facts could represent an advantage of Neptune with respect to Poseidon for the plain performance point.

Acknowledgments. Lorenzo Grassi is supported by the European Research Council under the ERC advanced grant agreement under grant ERC-2017-ADG Nr. 788980 ESCADA.

References

[AAB+20] Abdelrahaman Aly, Tomer Ashur, Eli Ben-Sasson, Siemen Dhooghe, and Alan Szepieniec. Design of Symmetric-Key Primitives for Advanced Cryptographic Protocols. *IACR Transactions on Symmetric Cryptology*, 2020(3):1–45, 2020.

[AGP+19] Martin R. Albrecht, Lorenzo Grassi, Léo Perrin, Sebastian Ramacher, Christian Rechberger, Dragos Rotaru, Arnab Roy, and Markus Schofnegger. Feistel Structures for MPC, and More. In *ESORICS 2019*, volume 11736 of *LNCS*, pages 151–171, 2019.

[7] We point out that a similar result holds even in the case in which Neptune is instantiated with 8 external rounds.

[8] Note that $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix}$ can be computed via five additions only.

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

[BBS99] Eli Biham, Alex Biryukov, and Adi Shamir. Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. In *EUROCRYPT 1999*, volume 1592 of *LNCS*, pages 12–23, 1999.

[BCD+20] Tim Beyne, Anne Canteaut, Itai Dinur, Maria Eichlseder, Gregor Leander, Gaëtan Leurent, María Naya-Plasencia, Léo Perrin, Yu Sasaki, Yosuke Todo, and Friedrich Wiemer. Out of Oddity - New Cryptanalytic Techniques Against Symmetric Primitives Optimized for Integrity Proof Systems. In *CRYPTO 2020*, volume 12172 of *LNCS*, pages 299–328, 2020.

[BDPA08] Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. On the Indifferentiability of the Sponge Construction. In *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 181–197, 2008.

[BDPA13] Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. Keccak. In *EUROCRYPT 2013*, volume 7881 of *LNCS*, pages 313–314, 2013.

[BDPV07] G. Bertoni, J. Daemen, M. Peeters, and G. Van Assche. Sponge functions, 2007. In: Ecrypt Hash Workshop 2007, http://www.csrc.nist.gov/pki/HashWorkshop/PublicComments/2007_May.html.

[BPVA+11] Guido Bertoni, Michaël Peeters, Gilles Van Assche, et al. The Keccak reference, 2011. https://keccak.team/files/Keccak-reference-3.0.pdf.

[BR11] Andrey Bogdanov and Vincent Rijmen. Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Cryptology ePrint Archive, Report 2011/123, 2011. https://ia.cr/2011/123.

[BR14] Andrey Bogdanov and Vincent Rijmen. Linear hulls with correlation zero and linear cryptanalysis of block ciphers. *Des. Codes Cryptogr.*, 70(3):369–383, 2014.

[BS93] Eli Biham and Adi Shamir. *Differential Cryptanalysis of the Data Encryption Standard*. Springer, 1993.

[Buc76] Bruno Buchberger. A theoretical basis for the reduction of polynomials to canonical forms. *SIGSAM Bull.*, 10(3):19–29, 1976.

[Dae95] Joan Daemen. *Cipher and hash function design, strategies based on linear and differential cryptanalysis, PhD Thesis*. K.U.Leuven, 1995. https://cs.ru.nl/~joan/.

[DEG+18] Christoph Dobraunig, Maria Eichlseder, Lorenzo Grassi, Virginie Lallemand, Gregor Leander, Eik List, Florian Mendel, and Christian Rechberger. Rasta: A Cipher with Low ANDdepth and Few ANDs per Bit. In *CRYPTO 2018*, volume 10991 of *LNCS*, pages 662–692, 2018.

[DGGK21] Christoph Dobraunig, Lorenzo Grassi, Anna Guinet, and Daniël Kuijsters. Ciminion: Symmetric Encryption Based on Toffoli-Gates over Large Finite Fields. In *EUROCRYPT 2021*, volume 12697 of *LNCS*, pages 3–34, 2021.

[DGH+21] Christoph Dobraunig, Lorenzo Grassi, Lukas Helminger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Pasta: A Case for Hybrid Homomorphic Encryption. Cryptology ePrint Archive, Report 2021/731, 2021. https://ia.cr/2021/731.

[DKR97] Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The block cipher Square. In *FSE 1997*, volume 1267 of *LNCS*, pages 149–165, 1997.

[DMMR20] Joan Daemen, Pedro Maat Costa Massolino, Alireza Mehrdad, and Yann Rotella. The Subterranean 2.0 Cipher Suite. *IACR Trans. Symmetric Cryptol.*, 2020(S1):262–294, 2020.

[GKR+21] Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. Poseidon: A New Hash Function for Zero-Knowledge Proof Systems. In *USENIX Security 2021*. USENIX Association, 2021.

[GKRS21] Lorenzo Grassi, Dmitry Khovratovich, Sondre Rønjom, and Markus Schofnegger. The Legendre Symbol and the Modulo-2 Operator in Symmetric Schemes over $(\mathbb{F}_p)^n$. Cryptology ePrint Archive, Report 2021/1533, 2021. https://ia.cr/2021/1533.

[GLR+19] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a Generalization of Substitution-Permutation Networks: The HADES Design Strategy. Cryptology ePrint Archive, Report 2019/1107, 2019. https://ia.cr/2019/1107.

[GLR+20] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a Generalization of Substitution-Permutation Networks: The HADES Design Strategy. In *EUROCRYPT 2020*, volume 12106 of *LNCS*, pages 674–704, 2020.

[Gra18] Lorenzo Grassi. Mixture Differential Cryptanalysis: a New Approach to Distinguishers and Attacks on round-reduced AES. *IACR Trans. Symmetric Cryptol.*, 2018(2):133–160, 2018.

[GRR16] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. Subspace Trail Cryptanalysis and its Applications to AES. *IACR Trans. Symmetric Cryptol.*, 2016(2):192–225, 2016.

[GRR17] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. A New Structural-Differential Property of 5-Round AES. In *EUROCRYPT 2017*, volume 10211 of *LNCS*, pages 289–317, 2017.

[GRS21] Lorenzo Grassi, Christian Rechberger, and Markus Schofnegger. Proving Resistance Against Infinitely Long Subspace Trails: How to Choose the Linear Layer. *IACR Trans. Symmetric Cryptol.*, 2021(2):314–352, 2021.

[GSW+21] Chun Guo, François-Xavier Standaert, Weijia Wang, Xiao Wang, and Yu Yu. Provable Security of SP Networks with Partial Non-Linear Layers. *IACR Transactions on Symmetric Cryptology*, 2021(2):353–388, 2021.

[HKC+20] Jincheol Ha, Seongkwang Kim, Wonseok Choi, Jooyoung Lee, Dukjae Moon, Hyojin Yoon, and Jihoon Cho. Masta: An HE-Friendly Cipher Using Modular Arithmetic. *IEEE Access*, 8:194741–194751, 2020.

[JK97] Thomas Jakobsen and Lars R. Knudsen. The Interpolation Attack on Block Ciphers. In *FSE 1997*, volume 1267 of *LNCS*, pages 28–40, 1997.

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

[Knu98] Lars R. Knudsen. DEAL - A 128-bit Block Cipher. *Technical Report, Department of Informatics, Bergen, Norway*, 1998.

[KR21] Nathan Keller and Asaf Rosemarin. Mind the Middle Layer: The HADES Design Strategy Revisited. In *EUROCRYPT 2021*, volume 12697 of *LNCS*, pages 35–63, 2021.

[LAAZ11] Gregor Leander, Mohamed Ahmed Abdelraheem, Hoda AlKhzaimi, and Erik Zenner. A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack. In *CRYPTO 2011*, volume 6841 of *LNCS*, pages 206–221, 2011.

[Lai94] X. Lai. Higher order derivatives and differential cryptanalysis. Communications and Cryptography: Two Sides of One Tapestry, 1994.

[LM90] Xuejia Lai and James L. Massey. A Proposal for a New Block Encryption Standard. In *EUROCRYPT 1990,*, volume 473 of *LNCS*, pages 389–404, 1990.

[LMR15] Gregor Leander, Brice Minaud, and Sondre Rønjom. A Generic Approach to Invariant Subspace Attacks: Cryptanalysis of Robin, iSCREAM and Zorro. In *EUROCRYPT 2015*, volume 9056 of *LNCS*, pages 254–283, 2015.

[Mat93] Mitsuru Matsui. Linear Cryptanalysis Method for DES Cipher. In *EUROCRYPT 1993*, volume 765 of *LNCS*, pages 386–397, 1993.

[MP13] Gary L. Mullen and Daniel Panario. *Handbook of Finite Fields*. Chapman & Hall/CRC, 1st edition, 2013.

[MRST09] Florian Mendel, Christian Rechberger, Martin Schläffer, and Søren S. Thomsen. The Rebound Attack: Cryptanalysis of Reduced Whirlpool and Grøstl. In *FSE 2009*, volume 5665 of *LNCS*, pages 260–276, 2009.

[Nag51] T. Nagell. Euler's Criterion and Legendre's Symbol. Introduction to Number Theory, 1951.

[Sha12] Christopher J. Shallue. Permutation Polynomials of Finite Fields. arXiv, ePrint: 1211.6044, 2012.

[Sze21] Alan Szepieniec. On the Use of the Legendre Symbol in Symmetric Cipher Design. Cryptology ePrint Archive, Report 2021/984, 2021. https://ia.cr/2021/984.

[Vau99] Serge Vaudenay. On the Lai-Massey Scheme. In *ASIACRYPT 1999*, volume 1716 of *LNCS*, pages 8–19, 1999.

[vzGG13] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra (3. ed.)*. Cambridge University Press, 2013.

[Wol85] Stephen Wolfram. Cryptography with Cellular Automata. In *CRYPTO 1985*, volume 218 of *LNCS*, pages 429–432, 1985.

A Proof of Proposition 2

Given $F : \mathbb{F}_q^m \to \mathbb{F}_q$ and $y \in \mathbb{F}_q$, here we use the notation $F^{-1}(y)$ to denote $F^{-1}(y) := \{x \in \mathbb{F}_q^m \mid F(x) = y\}$. Without loss of generality (W.l.o.g.), let's assume that $\left|F^{-1}(0)\right| \gtrsim q^{m-1}$ (analogous for the other cases). Let's define the sets $\mathfrak{A}, \mathfrak{B} \subseteq \mathbb{F}_q^n$ as:

$$\mathfrak{A} := \{(x_0, x_1, \ldots, x_{n-1}) \in \mathbb{F}_q^n \mid x_0 = 0\}$$
$$\mathfrak{B} := \{(x_0, x_1, \ldots, x_{n-1}) \in \mathbb{F}_q^n \mid (x_0, x_1, \ldots x_{m-1}) \in F^{-1}(0)\}.$$

In particular:

$$\mathfrak{B} = \bigcup_{(x_m,\ldots,x_{n-1}) \in \mathbb{F}_q^{n-m}} \mathfrak{B}_{(x_m,\ldots,x_{n-1})}, \qquad \text{where}$$
$$\mathfrak{B}_{(x_m,\ldots,x_{n-1})} := \{(x_0, \ldots, x_{m-1}, x_m, \ldots, x_{n-1}) \in \mathbb{F}_q^n \mid (x_0, \ldots, x_{m-1}) \in F^{-1}(0)\}.$$

Note that:

- $|\mathfrak{A}| = q^{n-1}$
- $\mathcal{S}|_{\mathfrak{B}} \subseteq \mathfrak{A}$ (otherwise, $F(x_0, x_1, \ldots, x_{m-1}) \neq 0 \in \mathbb{F}_q$);
- $\left|\mathcal{B}_{(x_m,\ldots,x_{n-1})}\right| \gtrsim q^{m-1}$ for each $(x_m, \ldots, x_{n-1}) \in \mathbb{F}_q^{n-m}$ (by the hypothesis of F being not balanced and $\left|F^{-1}(0)\right| \gtrsim q^{m-1}$), which implies that

$$|\mathfrak{B}| = \sum_{(x_m,\ldots,x_{n-1}) \in \mathbb{F}_q^{n-m}} |\mathfrak{B}_{(x_m,\ldots,x_{n-1})}| \gtrsim q^{m-1} \cdot q^{n-m} = q^{n-1} = |\mathfrak{A}|.$$

By meanings of cardinality of $\mathcal{S}$ restricted on $\mathfrak{B}$ cannot be injective, hence $\mathcal{S}$ is not injective, which implies that $\mathcal{S}$ is not invertible. □

B Practical Verification for Quadratic Functions

In this section, we describe the practical experiments we performed in order to support our theoretical results. Supplemental material including the source code in C++ can be found in a zip-file at

> https://drive.google.com/file/d/1tYn3OaAxQVY0IRIPxIS3TAXeHM_fmr-j/view?usp=sharing

containing the following files:

- source code `main.cpp`,
- `Makefile` to compile and run the code,
- a `readme.md` file containing instructions,
- a description of practical experiments (`experiments.pdf`),
- directories with log files of our practical tests.

Compiling with make and running by setting in the makefile the variables for the case to run. A standard C++ compiler should work (we have used `g++` with gcc version 7.5.0 and the GNU multiprecision library `libgmp` version 6.2.0). The code is not optimised although offers a rudimentary form of parallelization, which becomes necessary to run some cases we report on.

Algorithm 1: Pseudo-code for finding functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for which the corresponding function $\mathcal{S}$ over $\mathbb{F}_p^n$ is invertible.

Data: Input: $p \geq 3, m \geq 2, n \geq m$
Result: Output: $F : \mathbb{F}_p^m \to \mathbb{F}_p$ such that $\mathcal{S}$ over $\mathbb{F}_p^n$ is invertible
let $\mathfrak{X} = \emptyset$ be the set of functions $\mathbb{F}_p^m \to \mathbb{F}_p$;
for *each function F defined as in (7)* **do**
    `// 1st Step: check if` F `is balanced`
    let $a = 0 \in \mathbb{N}^p$ and $b = 0 \in \{0,1\}^{p^n}$;
    for *all* $x \in \mathbb{F}_p^m$ **do**
        $a_{F(x)} \leftarrow a_{F(x)} + 1$;
        if $a_{F(x)} > p^{m-1}$ **then**
            Break: F is not balanced, hence discard it;
    `// 2nd Step: given` F `balanced, check if` $\mathcal{S}_F$ `is invertible`
    for *all* $x \in \mathbb{F}_p^n$ **do**
        if $b_{\mathcal{S}(x)} = 0$ **then**
            $b_{\mathcal{S}(x)} \leftarrow 1$;
        else
            Break: $\mathcal{S}$ is not a permutation, hence discard F;
    $\mathfrak{X} \leftarrow \mathfrak{X} \cup \{F\}$;
return $\mathfrak{X}$

B.1 Brute Force Research

Here we propose a pseudo-code of the algorithm that we used for our tests for the case of polynomial functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ of degree $d \geq 2$, defined as (7).

Given $p \geq 3$ and $n \geq m \geq 2$, Algorithm 1 consists of two steps:

1. checking if a function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ is balanced or not;
2. if a function is balanced, checking if $\mathcal{S}$ over $\mathbb{F}_p^n$ is invertible or not.

The check is done making use of hash tables.

Reducing the Search Space. As first step, we show how to use the necessary conditions given in Sect. 2 in order to reduce the cost:

- first of all, the coefficient of the monomial of degree zero can be fixed equal to zero (that is, $\alpha_{0,\ldots,0} = 0$); indeed, just choose $\psi = -\alpha_{0,\ldots,0} \cdot \omega$;
- the coefficient of one monomial of degree one and one of degree two can be chosen in $\{0,1\}$ (e.g., $\alpha_{1,0,\ldots,0}, \alpha_{2,0,\ldots,0} \in \{0,1\}$); indeed, if they are both equal to zero the result is obvious, if only one of them is different from zero just choose ω as the inverse of the non-zero one. If both $\alpha_{1,0,\ldots,0} \neq 0$ and $\alpha_{2,0,\ldots,0} \neq 0$, take $\mu = \frac{\alpha_{1,0,\ldots,0}}{\alpha_{2,0,\ldots,0}}$, $\omega = \frac{\alpha_{2,0,\ldots,0}}{(\alpha_{1,0,\ldots,0})^2}$ and $\nu = 0$.

In this way, the number of quadratic functions $F : \mathbb{F}_p^m \to \mathbb{F}_p$ reduces as following

$$p^{1+2m+\binom{m}{2}} = p^{\frac{m^2+3m+2}{2}} \quad \to \quad 2^2 \cdot p^{\frac{m^2+3m-4}{2}}.$$

Memory and Computational Costs. Let's analyze the cost of the algorithm. First of all, the memory cost is given by $\mathcal{O}(\max\{p \cdot \lceil \log_2(p^{m-1}) \rceil, p^n\}) = \mathcal{O}(p^n)$ bits. Indeed, since

Table 3: Summary of our practical results for $d = 2$ and $m \in \{2, 3\}$. For each $p \geq 3$, we report the maximum value of n tested, the number of balanced *quadratic* functions with respect to the total number of functions F (with $\alpha_{0,0,0} = 0$, $\alpha_{2,0,0}, \alpha_{0,0,1} \in \{0, 1\}$) and the total runtime in hours/days.

	Case: $m = 2$ **and** $n \geq 3$			
p	# balanced F	percentage	max n	runtime
3	19	17.5%	31	1.5 hours
5	69	13.8%	10	3.6 hours
7	151	11.0%	7	0.9 hours
11	411	7.7%	7	9.6 hours
13	589	6.7%	5	1.0 hours
17	1 041	5.2%	5	3.7 hours
19	1 315	4.8%	5	6.3 hours
23	1 959	4.0%	5	16.0 hours
	Case: $m = 3$ **and** $n \geq 5$			
p	# balanced F	percentage	max n	runtime
3	2 175	24.9%	13	9.8 hours
5	53 725	17.2%	7	5.3 hours
7	426 139	12.9%	7	6.0 days
11	2 464 657	3.2%	5	46.8 days

the first step stops when one entry of $a \in \mathbb{N}^p$ is bigger than p^{m-1}, we need $\lceil \log_2(p^{m-1}) \rceil$ bits for each entry of such array.

Regarding the computational cost, for each function $F : \mathbb{F}_p^m \to \mathbb{F}_p$:

- we have to test p^m different inputs in order to check if F is balanced;
- if the function F is balanced, we have to test p^n different inputs in order to check if $\mathcal{S}$ is invertible.

This requires $\mathcal{O}(p^m \cdot 2^2 \cdot p^{(m^2+3m-4)/2} \cdot p^n)$ steps (namely, memory access, evaluation of the function F, etc.). Note that these are just rough estimations, since several functions F are e.g. discarded in the first step if they are not balanced.

B.2 Practical Results

In order to carry out the practical experiments, we implemented the brute-force collision-search algorithm described in Algorithm 1: for each *quadratic* function $F : \mathbb{F}_p^m \to \mathbb{F}_p$ we look for a collision in the domain of the corresponding function $\mathcal{S}_F$ (as defined in Def. 1) over $\mathbb{F}_p^n$ for $n \geq m$. We aim to practically verify that no invertible function $\mathcal{S}_F$ exists for the cases (1st) $m = 2$ and $n \geq 3$ (as proved in Theorem 3) and (2nd) $m = 3$ and $n \geq 5$ (as proved in Theorem 4). We verify it practically just for small values of p and n, while the theoretical proofs confirm that the behavior that occurs for small values is also valid for all $p \geq 3$.

The tests have been done on a Intel 40-cores Xeon E5-2698 v4 @ 2.20GHz. The results of the practical experiments are given in Table 3, describing for each $p \geq 3$:

- the number of balanced *quadratic* functions with respect to the total number of functions F;
- the maximum value of n tested (denoted as "max n");
- the total runtime in hours/days.

We restrict the domain of functions F by using the equivalent classes introduced in Sect. B.1 (that is, $\alpha_{0,0,0} = 0$, $\alpha_{2,0,0}, \alpha_{0,0,1} \in \{0, 1\}$).

As described in Algorithm 1, tests are divided into two main phases: (1st) the balanced testing and (2nd) the collision search. The time each step requires depends on the case considered:

- $d = 2$ and $m = 2$: the balanced testing takes just the 0.1% of the total runtime, while the collision search takes most of the time spent on the tests.

- $d = 2$ and $m = 3$: runtimes for balanced testing and collision search depend on p, e.g. for $p = 3$ the balanced testing takes the 0.1% of the total runtime, while for $p = 11$ it takes the 88%.

Anyway, the balanced testing and collisions search runtimes depend strongly on the number of iterations that the program requires in order to establish if a function is balanced or, respectively, invertible (i.e., to find the first collision), since the program works iteratively, testing for each value whether its image has already been evaluated as the image of another value.

C Details about the Security Analysis of Neptune

C.1 Maximum Differential Probability of $\mathcal{S}'$

Let $p \geq 3$ be a prime integer, and let $\mathcal{S}' : \mathbb{F}_p^2 \to \mathbb{F}_p^2$ be defined as in (11). Here we prove that its maximum differential probability is p^{-1}.

In order to do this, we proceed in two steps:

- first, we compute the maximum differential probability of $\mathcal{S}_{F'}$ over $\mathbb{F}_p^2$ defined as in Def. 1 via $F'(x_0, x_1) = \alpha \cdot x_0 + \beta \cdot (x_0 - x_1)^2$;

- based on this result, we compute the maximum differential probability of $\mathcal{S}'$.

Maximum Differential Probability of $\mathcal{S}_{F'}$. Given input/output differences $(\delta_0, \delta_1), (\Delta_0, \Delta_1) \in \mathbb{F}_p^2 \setminus \{(0, 0)\}$, we first analyze the number of solutions (x_0, x_1) of the following system

$$\begin{aligned} \alpha \cdot \delta_0 + \beta \cdot (\delta_0 - \delta_1)^2 + 2\beta \cdot (\delta_0 - \delta_1) \cdot (x_0 - x_1) &= \Delta_0 \\ \alpha \cdot \delta_1 + \beta \cdot (\delta_0 - \delta_1)^2 + 2\beta \cdot (\delta_0 - \delta_1) \cdot (x_0 - x_1) &= \Delta_1 \,, \end{aligned}$$

which corresponds to

$$\alpha \cdot (\delta_0 - \delta_1) = \Delta_0 - \Delta_1 \tag{13}$$

and

$$(\delta_0 - \delta_1) \cdot (x_0 - x_1) = \frac{\Delta_0 - \alpha \cdot \delta_0 - \beta \cdot (\delta_0 - \delta_1)^2}{2\beta}.$$

It follows that:

- if $\delta_0 \neq \delta_1$, such system of equations admits exactly p solutions;

- if $\delta_0 = \delta_1$ (hence, $\Delta_0 = \Delta_1$), then the equations are always satisfied if $\Delta_0 = \alpha \cdot \delta_0$.

Maximum Differential Probability of $\mathcal{S}'$. Given $(\delta_0, \delta_1), (\Delta_0, \Delta_1) \in \mathbb{F}_p^2 \setminus \{(0,0)\}$, the maximum differential probability of $\mathcal{S}'$ is given by

$$\sum_{(\varepsilon_0,\varepsilon_1)\in\mathbb{F}_p^2} \text{Prob}\left(\begin{bmatrix}\delta_0\\ \delta_1\end{bmatrix} \to \begin{bmatrix}\varepsilon_0\\ \varepsilon_1\end{bmatrix}\right) \times \text{Prob}\left(\left(\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix} \times \begin{bmatrix}\varepsilon_0\\ \varepsilon_1\end{bmatrix}\right) \to \begin{bmatrix}\Delta_0\\ \Delta_1\end{bmatrix}\right). \tag{14}$$

In our case, condition (13) becomes:

$$\alpha \cdot (\delta_0 - \delta_1) = \varepsilon_0 - \varepsilon_1 \qquad \text{and} \qquad \alpha \cdot (\varepsilon_0 - 2 \cdot \varepsilon_1) = \Delta_0 - \Delta_1,$$

that is

$$\varepsilon_0 = 2 \cdot \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha} \qquad \text{and} \qquad \varepsilon_1 = \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha}.$$

Hence, the probability given in (14) reduces to

$$\text{Prob}\left(\begin{bmatrix}\delta_0\\ \delta_1\end{bmatrix} \to \begin{bmatrix}2 \cdot \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha}\\ \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha}\end{bmatrix}\right) \times \text{Prob}\left(\begin{bmatrix}5 \cdot \alpha \cdot (\delta_0 - \delta_1) - 3 \cdot \frac{\Delta_0 - \Delta_1}{\alpha}\\ 5 \cdot \alpha \cdot (\delta_0 - \delta_1) - 4 \cdot \frac{\Delta_0 - \Delta_1}{\alpha}\end{bmatrix} \to \begin{bmatrix}\Delta_0\\ \Delta_1\end{bmatrix}\right).$$

Such probability is never bigger than p^{-1}, since:

- if $\delta_0 = \delta_1$, then

  $$2 \cdot \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha} = \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha}.$$

  This implies that the first probability is equal to 1. If $\delta_0 \neq \delta_1$, then the first probability is $1/p$;

- if $\Delta_0 = \Delta_1$, then

  $$5 \cdot \alpha \cdot (\delta_0 - \delta_1) - 3 \cdot \frac{\Delta_0 - \Delta_1}{\alpha} = 5 \cdot \alpha \cdot (\delta_0 - \delta_1) - 4 \cdot \frac{\Delta_0 - \Delta_1}{\alpha}.$$

  This implies that the second probability is equal to 1. If $\Delta_0 \neq \Delta_1$, then the first probability is $1/p$;

- if $\delta_0 = \delta_1$ **and** $\Delta_0 = \Delta_1$, then

  $$\begin{aligned} &2 \cdot \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha} = \alpha \cdot (\delta_0 - \delta_1) - \frac{\Delta_0 - \Delta_1}{\alpha} \\ =&5 \cdot \alpha \cdot (\delta_0 - \delta_1) - 3 \cdot \frac{\Delta_0 - \Delta_1}{\alpha} = 5 \cdot \alpha \cdot (\delta_0 - \delta_1) - 4 \cdot \frac{\Delta_0 - \Delta_1}{\alpha} = 0. \end{aligned}$$

  In such a case, the overall probability is equal to zero, since we cannot have a zero difference in the middle when the input/output differences are non-zero (remember that the construction is invertible).

It follows that the probability is maximum when either $\delta_0 = \delta_1$ **or** $\Delta_0 = \Delta_1$, and in such a case it is equal to $1/p$.

C.2 Gröbner Basis Attacks on Neptune

The cost of the Gröbner basis attack depends on the system of equations that describes Neptune. As usually done in the literature, instead of considering (collision or/and preimage) attacks on the sponge hash function, we focus on the CICO problem on the permutation that instantiates Neptune.

Definition 8. The invertible function $\mathcal{G} : \mathbb{F}_p^t \to \mathbb{F}_p^t$ is κ-secure against the CICO (t_1, t_2)-problem (where $t_1, t_2 < t$) if there is no algorithm with expected complexity smaller than 2^κ that for given $i_1 \in \mathbb{F}_p^{t_1}$ and $o_1 \in \mathbb{F}_p^{t_2}$ finds i_2, o_2 such that $\mathcal{G}(i_1 \| i_2) = o_1 \| o_2$.

We consider two approaches:

- working on the relation between the input and the output of the entire permutation;
- working at round level.

Preliminary. Gröbner basis attack consists of three steps:

1. first, the attacker needs to set up the equation system and compute a Gröbner basis for it;
2. secondly, they perform a change of term ordering for the basis, usually going to a term order which makes it easier to eliminate variables and find the solutions;
3. finally, the attacker uses the system obtained in the second step in order to start solving for the variables.

As is usually done in the literature, here we focus on the complexity of the first step (i.e., computing a Gröbner basis), which can be estimated by

$$C_{\text{GB}} = \mathcal{O}\left(\binom{D_{\text{reg}} + n_v}{n_v}^\omega\right)$$

operations, where D_{reg} is the degree of regularity, n_v is the number of variables, and $2 \leq \omega < 3$ is a constant representing the complexity of a matrix multiplication. Let n_e denotes the number of equations in the polynomial system and d_i is the degree of the i-th equation. Directly computing D_{reg} is hard in general, but an estimate for *regular sequences* (namely, in the case $n_e = n_v$) is given by

$$D_{\text{reg}} = 1 + \sum_{i=1}^{n_e} (d_i - 1).$$

C.2.1 Working on the Input and the Output

Let's first consider the input and the output of the permutation, focusing on the case in which the number of unknown input variables x is equal to the number of known output variables. In such a case, we get x equations of degree $4^6 \cdot d^{R_I} = 2^{12+R_I \cdot \log_2(d)}$ (we assume that $R_F = 6$ is fixed) in x variables. Hence, we have that

$$D_{\text{reg}} = 1 + x \cdot (2^{12+R_I \cdot \log_2(d)} - 1)$$

which implies a cost of approximately

$$\mathcal{O}\left(\binom{x \cdot 2^{12+R_I \cdot \log_2(d)}}{x}^\omega\right)$$

assuming a semi-regular system (as done for Poseidon). Since $\omega \geq 2$ (the best scenario for the attacker), we have that

$$\binom{x \cdot 2^{12+R_I \cdot \log_2(d)}}{x}^\omega \geq \left(\frac{\left(1 + x \cdot (2^{12+R_I \cdot \log_2(d)} - 1)\right)^x}{x!}\right)^2$$

$$\geq \left(\frac{1 + x \cdot (2^{12+R_I \cdot \log_2(d)} - 1)}{x}\right)^{2x} \approx (2^{12+R_I \cdot \log_2(d)})^{2x},$$

where $x! \leq x^x$ for each $x \geq 1$. In order to guarantee κ bits of security:

$$(2^{12+R_I \cdot \log_2(d)})^{2x} \geq \min\{2^\kappa, p^x\}.$$

The maximum is obtained for $x = 1$, which implies

$$R_I \geq \log_d(2) \cdot \left(\frac{\min\{\kappa, \log_2(p)\}}{2} - 12 \right),$$

which is always satisfied by the number of rounds required to prevent the interpolation attack.

C.2.2 Working at Round Level

Another possibility for setting up the Gröbner basis attack consists of working at round level. In such a case:

- every internal round can be described as a single equation of degree d;
- every external round can be described via t equations of degree 2. Indeed, assuming for simplicity $\alpha = \beta = 1$, note that given $(y_0, y_1) = \mathcal{S}'(x_0, x_1)$, we have

$$\begin{bmatrix} y_0 + (y_0 - y_1)^2 \\ y_1 + (y_0 - y_1)^2 \end{bmatrix} = \begin{bmatrix} \gamma \\ 0 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \times \begin{bmatrix} x_0 + (x_0 - x_1)^2 \\ x_1 + (x_0 - x_1)^2 \end{bmatrix}.$$

It follows that we have

- R_I equations of degree d;
- $R_F \cdot t - c$ equations of degree 2 (note that the final c $\mathbb{F}_p$-elements are truncated)

in $R_F \cdot t - c + R_I$ variables (note that the inner part is instantiated with a fixed initial value). Assuming a semi-regular system and $R_F = 6$, we have that

$$D_{\text{reg}} = 1 + 6 \cdot t - c + R_I \cdot (d - 1).$$

As in the case of Poseidon, the number of rounds necessary for preventing the interpolation attack satisfies the inequality

$$\binom{1 + 12 \cdot t - 2 \cdot c + R'_I \cdot d}{6 \cdot t - c + R'_I}^2 \geq \min\{2^\kappa, p^x\},$$

where $R'_I = R_I - t$ in order to take into account the fact that (up to) t internal rounds can be skipped via an invariant subspace.