The Legendre Symbol and the Modulo-2 Operator in Symmetric Schemes over $\mathbb{F}_p^n$

Preimage Attack on Full *Grendel*

Lorenzo Grassi[1], Dmitry Khovratovich[2], Sondre Rønjom[3] and Markus Schofnegger[4]

[1] Digital Security Group, Radboud University, Nijmegen
[2] Dusk Network, Luxembourg
[3] University of Bergen, Norway
[4] IAIK, Graz University of Technology, Austria
l.grassi@cs.ru.nl
khovratovich@gmail.com
sondre.ronjom@uib.no
markus.schofnegger@tugraz.at

Abstract. Motivated by modern cryptographic use cases such as multi-party computation (MPC), homomorphic encryption (HE), and zero-knowledge (ZK) protocols, several symmetric schemes that are efficient in these scenarios have recently been proposed in the literature. Some of these schemes are instantiated with low-degree nonlinear functions, for example low-degree power maps (e.g., MiMC, HadesMiMC, Poseidon) or the Toffoli gate (e.g., Ciminion). Others (e.g., *Rescue*, *Vision*, *Grendel*) are instead instantiated via high-degree functions which are easy to evaluate in the target application. A recent example for the latter case is the hash function *Grendel*, whose nonlinear layer is constructed using the Legendre symbol.

In this paper, we analyze high-degree functions such as the Legendre symbol or the modulo-2 operation as building blocks for the nonlinear layer of a cryptographic scheme over $\mathbb{F}_p^n$. Our focus regards the security analysis rather than the efficiency in the mentioned use cases. For this purpose, we present several new invertible functions that make use of the Legendre symbol or of the modulo-2 operation.

Even though these functions often provide strong statistical properties and ensure a high degree after a few rounds, the main problem regards their small number of possible outputs, that is, only three for the Legendre symbol and only two for the modulo-2 operation. By guessing them, it is possible to reduce the overall degree of the function significantly. We exploit this behavior by describing the first preimage attack on full *Grendel*, and we verify it in practice.

Keywords: Legendre Symbol · Modulo-2 Operator · Grendel · Preimage Attack

1 Introduction

Recently, modern cryptographic applications such as multi-party computation (MPC), homomorphic encryption (HE), and zero-knowledge (ZK) protocols have motivated the design of specific cryptographic schemes. These are often defined over large prime fields $\mathbb{F}_p$ with the aim to increase efficiency in the above-mentioned use cases. Examples of these schemes include MiMC [AGR+16], GMiMC [AGP+19], HadesMiMC/Poseidon [GLR+20, GKR+21], Masta [HKC+20], Pasta [DGH+21], and Ciminion [DGGK21], all of which use low-degree power maps $x \mapsto x^d$ or Toffoli gates $(x, y) \mapsto x \cdot y$.

However, it is also possible to achieve efficiency with high-degree functions. E.g., the cost of some of ZK protocols is the cost of proving/verifying a given statement $y = F(x)$ for a certain function over $\mathbb{F}^n$. In such a case, instead of showing $y = F(x)$ directly, it is also possible to prove an *equivalent* relation $G(x, y) = c$ for some constant c. In some cases, proving/verifying this second equivalent equality could be much cheaper than proving/verifying $y = F(x)$. This approach is used in Friday and Jarvis [AD18], *Rescue*/*Vision* [AAB+20], and more recently in *Grendel* [Sze21]. In the first two, the function F is defined either as $F(x) = 1/x$ (hence, $G(x, y) = x \cdot y$) or as $F(x) = x^{1/d}$ for a small $d \ll p$ (hence, $G(x, y) = y^d - x$). In *Grendel*, F is defined via the Legendre symbol, which is a function from $\mathbb{F}_p$ to $\{-1, 0, 1\}$ and which returns ±1 if the input is a (nonzero) quadratic residue or not (and zero otherwise). Besides being potentially efficient in ZK protocols [Sze21], the Legendre symbol as a PRF is competitive both in MPC applications as showed in [GRR+16] and in digital signature schemes as showed in [BdSG20].

In this paper, we analyze the advantages and disadvantages of using the Legendre symbol and the modulo-2 operation in the design of symmetric schemes over $\mathbb{F}_p^n$. We emphasize that we focus on the security aspect of the symmetric schemes, rather than their efficiency on MPC-/FHE-/ZK-applications. As a first contribution, we propose several invertible nonlinear layers over $\mathbb{F}_p^n$ for $n \geq 1$ constructed either via the Legendre symbol or the modulo-2 operation, together with a detailed analysis of their statistical and algebraic properties. Secondly, we propose an attack strategy of a scheme instantiated via any of these nonlinear layers. We use this strategy to present a preimage attack on a sponge hash function instantiated with the full *Grendel* permutation.

1.1 The Legendre Symbol and the Modulo-2 Operation

Several works have been presented in the literature asserting that the Legendre symbol exhibits high pseudo-randomness. In 1997, Mauduit and Sárközy [MS97] introduced several metrics to measure the pseudo-randomness of binary sequences, concluding that "*Legendre symbol sequences are the most natural candidate for pseudo-randomness*". The high linear complexity of Legendre symbol sequences has been confirmed later on by Ding et al. [DHS98]. Tóth [Tó07] and Gyarmati et al. [GMS14] introduced new pseudo-randomness measures (avalanche effect and cross-correlation) and asserted high values of those in Legendre symbol sequences. Modulo operator is also used for similar cryptographic purposes, e.g., in the construction of pseudo-random number generators. Well known examples include the Lehmer pseudo-random number generator [Leh54] (whose formula is $x_i = \alpha \cdot x_{i-1} \mod m$, where the modulus m is a power of a prime number, the multiplier α is a primitive root modulo m and the seed x_0 is coprime to m) and the Blum Blum Shub pseudo-random number generator [BBS86] (whose formula is $x_i = x_{i-1}^2 \mod m$, where $m = p \cdot q$ for two large primes p, q).

In the case of a cryptographic symmetric scheme, the Legendre symbol may be used as *part of an invertible S-box over* $\mathbb{F}_p^n$, that is, it can be used in order to build a high-degree invertible nonlinear permutation over $\mathbb{F}_p^n$, as pointed out in [Sze21]. As we are going to show in Sections 3 and 4, similar argument holds for the modulo-2 operation.

Related Work [Sha12, Sze21]. Let $L_p(\cdot)$ be the Legendre symbol. Two examples proposed in the literature are

- the map $x \mapsto x \cdot (L_p(x) + \alpha)$ [Sha12], which is invertible if $L_p(\alpha^2 - 1) = 1$, and
- the map $x \mapsto x^d \cdot L_p(x)$ [Sze21], which is invertible if $\gcd(d + (p - 1)/2, p - 1) = 1$.

As our first contribution, we generalize the first function to $x \mapsto x^{d'} \cdot (L_p(x) + \alpha)$, and we provide the values of d that satisfy the condition for the second one for each p. We show that when $d', d \geq 2$, both of these two functions have differential and linear properties [BS90, Mat93]. By the definition of the Legendre symbol, they are of high degree (respectively, $d' + (p-1)/2$ and $d + (p-1)/2$), but their algebraic representation is sparse.

Modulo-2 Operator. Secondly, in Section 4.2 we propose new invertible functions that are built via the modulo-2 operation. While no quadratic map $x \mapsto x^2$ is invertible over $\mathbb{F}_p$ ($x^2 = (-x)^2$ for each $x \in \mathbb{F}_p$), in Section 4.2 we show how to slightly modify it via the modulo-2 operation in order to construct a permutation over $\mathbb{F}_p$. In particular, given a non-quadratic residue $\alpha \in \mathbb{F}_p$ (i.e., $\alpha \neq \beta^2$ for each $\beta \in \mathbb{F}_p$), we show that $x \mapsto \alpha^{x \mod 2} \cdot x^2$ is a permutation over $\mathbb{F}_p$. For example, if $p = 3 \mod 4$, then $x \mapsto (-1)^{x \mod 2} \cdot x^2 \equiv (-1)^x \cdot x^2$ is a permutation over $\mathbb{F}_p$. Similar to the previous functions based on the Legendre symbol, this map has good statistical properties. It also has a high degree and its algebraic representation is dense.

Local Maps and Legendre Symbol. In Section 4.3, we present other invertible nonlinear layers over $\mathbb{F}_p^n$ for $n \geq 2$ defined via a *local map*. Let $1 < m \leq n$ and let $F : \mathbb{F}_p^m \to \mathbb{F}_p$ be a local map. Let $\mathcal{S}_F$ over $\mathbb{F}_p^n$ be defined as $\mathcal{S}_F(x_0, x_1, \ldots, x_{n-1}) = y_0 \| y_1 \| \ldots \| y_{n-1}$ where

$$y_i := F(x_i, x_{i+1}, \ldots, x_{i+m-1}),$$

where the index is taken modulo n. A well-known example of an invertible function $\mathcal{S}_F$ over $\mathbb{F}_2^n$ for odd n is the χ-function [Wol85, DGV92, Dae95], whose local map is $F(x_0, x_1, x_2) = x_1 \cdot x_2 + x_0 + x_2$. In this paper, we focus on the prime case, and we show how to construct F via the Legendre symbol such that $\mathcal{S}_F$ is invertible. Moreover, we prove that $x \mapsto x^{d_+} \cdot (1 + L_p(x)) + x^{d_-} \cdot (1 - L_p(x))$ is invertible if $\gcd(d_+, p-1) = \gcd(d_-, p-1) = 1$, but we point out that is vulnerable to e.g. side-channel attacks.

1.2 Attack on Full *Grendel*

Invertible functions over $\mathbb{F}_p^n$ constructed via the Legendre symbol or the modulo-2 operation are promising components for the nonlinear layer of a symmetric scheme over $\mathbb{F}_p^n$. Indeed, since they have good statistical (both linear and differential) properties, by combining them with a good linear layer (e.g., by using an invertible matrix with a high branch number [DR01, DR02]), it may be possible to ensure security against statistical attacks after only a few rounds. In a similar way, it is possible to achieve a maximum-degree and dense polynomial description of the scheme after a few rounds (in the case of the modulo-2 operation, one round seems to be sufficient for that). We also refer to [Sze21, Sect. 2-3] for a detailed analysis of this.

At the same time, a potential problem arises in the case of large prime numbers $p \gg 3$. This is related to the fact that the output space of both the Legendre symbol and of the modulo-2 operation has a much smaller size than the input space. In particular, the output space of the Legendre symbol contains only three elements (one of which occurs with low probability) and the output space of the modulo-2 operation contains only two elements. Hence, by guessing all Legendre symbols or modulo-2 operations, the algebraic representation of the scheme may be of low degree.

In Section 5, we exploit this strategy in order to set up preimage attacks on a sponge hash function instantiated with the *full Grendel* permutation. *Grendel* is a SHARK-like scheme [RDP$^+$96] defined over $\mathbb{F}_p^n$. Its nonlinear layer is defined as the concatenation of invertible maps $x \mapsto x^d \cdot L_p(x)$ for $d \geq 2$, and its linear layer is defined as the multiplication by an MDS matrix. Given a hash value $y \in \mathbb{F}_p^h$, our attack works as follows:

1. We guess all the Legendre symbols, and we construct the system of equations that link the input of the sponge hash function to the given output.
2. We solve this system by a root-finding technique.

3. We determine if the solution found satisfies the Legendre symbols guessed. If this is not the case, we repeat the procedure.

Using our new insights, we further estimate the new minimum number of rounds for security, and show that in some cases the original number of rounds must be increased significantly. We practically verified the attack and make the code available online.

As a result, while the Legendre symbol and the modulo-2 operation have good properties for invertible cryptographic schemes, special attention must be paid to the fact that they return a small number of outputs compared to the size of the input space. As we concretely show in this paper, this is a vulnerability that can be exploited in attacks.

1.3 Related Work

We recall related work in the literature, focusing on the security analysis of the Legendre symbol used as a PRF and other Gröbner basis and/or factorization attacks recently proposed for MPC-/HE-/ZK-friendly schemes.

Security Analysis of the Legendre Symbol PRF. Several works have been proposed in the literature regarding the security of the PRF $x \mapsto L_p^k(x) := L_p(x+k)$, where k is a secret key. In [Kho19], Khovratovich proposed a birthday bound attack that exploits the property $L_p^k(x) = L_p^{k+\delta}(x-\delta)$ for any $\delta \in \mathbb{F}_p$. This attack has later been improved by Beullens et al. [BBUV20] and Kaluderovic et al. [KKK20]. More recently, Seres et al. [SHB21] show that key-recovery attacks against the Legendre PRF are equivalent to solving a specific family of multivariate quadratic equation systems over a prime field.

In the quantum setting, Frixons and Schrottenloher [FS21] investigated the quantum security of the Legendre PRF without quantum random access to an oracle. While they presented two new attacks in this setting, both of them remain impractical for key recovery. To the best of our knowledge, if the oracle can only be queried classically, then no efficient quantum algorithm is known.

Factorization and Gröbner Basis Attacks. Algebraic attacks are usually the most powerful ones for MPC-/HE-/ZK-schemes, since such schemes usually present a simple algebraic structure. Here we recall some attacks proposed in the literature that are based on these techniques.

- In [ACG+19], Albrecht at al. present a security analysis of STARK-friendly designs based on Gröbner basis attacks. They propose preimage attacks on the Friday hash function and key-recovery attacks on full Jarvis.

- In [BCD+20], Beyne et al. set up a preimage attack on a sponge hash function instantiated with full Poseidon via Gröbner basis for a particular class of instances using weak linear layers.

- In [RAS20], Roy et al. set up collision and (second) preimage attacks on a sponge hash function instantiated with (reduced-round) members of the GMiMC family via a factorization technique.

- In [EGL+20], Eichlseder et al. propose a distinguisher on almost full MiMC over $\mathbb{F}_{2^n}$ based on a improved estimation of the degree growth. They also exploit this distinguisher for a key-recovery attack via a factorization technique.

2 Preliminaries

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 $x = x_0 \mid\mid x_1 \mid\mid \cdots \mid\mid x_{n-1}$, where $\cdot \mid\mid \cdot$ denotes concatenation. We use capital letters to denote round numbers and 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., $\mathcal{S} : \mathbb{F}_p^n \to \mathbb{F}_p^n$).

2.1 The Legendre Symbol and the Hash Function *Grendel*

2.1.1 The Legendre Symbol

First we recall the definition of the Legendre symbol and some of its properties.

Definition 1. Let $p \geq 3$ be a prime. An integer α is a quadratic residue modulo p if it is congruent to a perfect square modulo p, and a quadratic non-residue modulo p otherwise.

Definition 2. 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

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

Proposition 1 ([Nag51]). *Let $p, q \geq 3$ be two prime integers. The Legendre symbol has the following properties.*

1. *If $x = y \mod p$, then $L_p(x) = L_p(y)$.*
2. *$L_p(x \cdot y) = L_p(x) \cdot L_p(y)$.*
3. *$L_p(q) \cdot L_q(p) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}$.*
4. *If $p = 3 \mod 4$, then $\pm x^{\frac{p+1}{4}}$ are the square roots of the quadratic residue x.*

Moreover, particular identities include the following.

- *$L_p(-1) = 1$ 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 $p = 3, 5 \mod 8$.*

2.1.2 The Hash Function *Grendel*

A sponge hash function [BDPA08] is an iterated construction for building a function with inputs and outputs of variable length, based on a function or permutation operating on a state with a fixed size. Let $\mathcal{P}$ be a permutation over $\mathbb{F}^n$ (for a certain field $\mathbb{F}$), and let $n = r + c$, where c denotes the capacity and r the rate. For a security level of κ bits, c and r must satisfy $|\mathbb{F}|^c \geq 2^{2 \cdot \kappa}$ and $|\mathbb{F}|^r \geq 2^{\kappa}$. An input message $m \in \mathbb{F}^*$ is first injectively padded and split into $\mathbb{F}^r$-blocks $m_0, m_1, \ldots, m_\mu$. Then, the message blocks are compressed sequentially into an n-element state, i.e.,

$$h_{i+1} = \mathcal{P}(h_i + (m_i \mid\mid 0^c)) \quad \text{for } i = 0, \ldots, \mu,$$

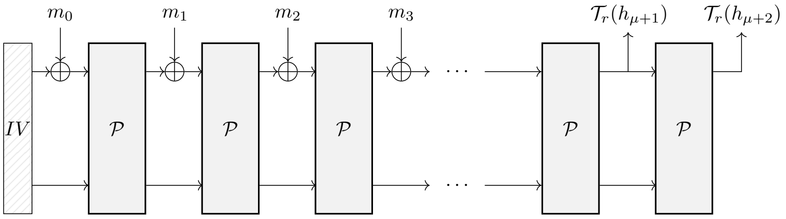

Figure 1: A sponge hash function with a two-element output, where $\oplus$ denotes the element-wise addition of two vectors in $\mathbb{F}^r$.

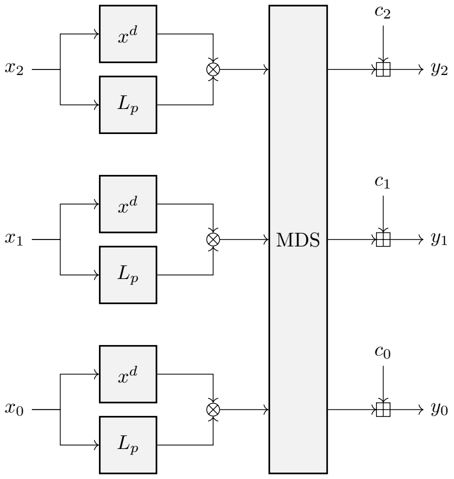

Figure 2: The round function of *Grendel* over $\mathbb{F}_p^n$, where $n = 3$ and c_i are round constants.

where $h_0 = IV \in \mathbb{F}^n$ is an initial value. After the absorption of the last message block, the output is of the form

$$\mathcal{T}_r(h_{\mu+1}) \mid\mid \mathcal{T}_r(h_{\mu+2}) \mid\mid \mathcal{T}_r(h_{\mu+3}) \mid\mid \cdots ,$$

where $h_{i+1} = \mathcal{P}(h_i)$ for $i \geq \mu+1$, and where $\mathcal{T}_r(x) = x_0 \mid\mid x_1 \mid\mid \cdots \mid\mid x_{r-1}$ is the truncation operation. A sponge hash function is shown in Fig. 1.

Let $p \geq 3$ be a prime number, and let $n \geq 2$ be an integer. The permutation *Grendel* [Sze21] defined over $\mathbb{F}_p^n$ resembles a Shark-like one, using independent S-boxes as its nonlinear layer and an MDS matrix as its linear layer. The main feature of *Grendel* regards its nonlinear layer, which is defined as $\mathcal{S}(x_0, x_1, \ldots, x_{n-1}) = S(x_0) \mid\mid S(x_1) \mid\mid \cdots \mid\mid S(x_{n-1})$, where

$$S(x) = x^d \cdot L_p(x).$$

Here, $L_p : \mathbb{F}_p \to \{-1, 0, 1\}$ is the Legendre symbol and $d \geq 2$ is the smallest integer which satisfies $\gcd((2d + p - 1)/2, p - 1) = 1$. The round function of *Grendel* is shown in Fig. 2. As we will prove in Corollary 2,

(1) if $p = 3 \mod 4$, then $d = 2$, and

(2) if $p = 1 \mod 4$, then $d \geq 3$ is the smallest integer that satisfies $\gcd(d, p - 1) = 1$.

For a security of κ bits (where $p^n \geq 2^{3\kappa}$, since $p^c \geq 2^{2\kappa}$ and $p^r \geq 2^\kappa$), the number of rounds $R \geq 1$ of *Grendel* is defined as

$$R = \max\left\{ \left\lceil \frac{2.5 \cdot \kappa}{\log_2(p) - \log_2(d) - 1} \right\rceil, R' \right\},$$

where $R' \geq 1$ is the smallest positive integer that satisfies

$$\max\left\{ \binom{2R'n - 2c + \frac{1+(R'n-c)(d+3)}{8}}{2R'n - 2c}^2 ;\; 2^{R'n-c} \times \binom{R'n - c + \frac{1+(R'n-c)(d-1)}{9}}{R'n - c}^2 \right\} \geq 2^{1.25\kappa}. \tag{1}$$

These complexities are derived from various attack vectors analyzed in the original paper, and we refer to [Sze21, Table 1] and [Sze21, Section 5.7] for more details.

2.2 Overview of Differential/Linear and Algebraic Cryptanalysis

Apart from presenting invertible nonlinear functions over $\mathbb{F}_p^n$ using the Legendre symbol or the modulo-2 operation, we also analyze their differential, linear, and algebraic properties. Here we recall the concepts of maximum differential probability useful in the context of differential cryptanalysis, and the main weaknesses exploited in algebraic attacks. Linear cryptanalysis is recalled in Appendix A.

Differential Cryptanalysis. Given pairs of inputs with fixed input differences, differential cryptanalysis [BS90, BS93] is based on the probability distribution of the corresponding output differences produced by the cryptographic primitive.

Let $\Delta_I, \Delta_O \in \mathbb{F}_p^n$ be respectively the input and the output differences through a function $\mathcal{F}$ over $\mathbb{F}_p^n$. The differential probability (DP) of having a certain output difference Δ_O given a particular input difference Δ_I is equal to

$$\text{Prob}_{\mathcal{F}}(\Delta_I \to \Delta_O) = \frac{|\{x \in \mathbb{F}_p^n \mid \mathcal{F}(x + \Delta_I) - \mathcal{F}(x) = \Delta_O\}|}{p^n}.$$

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

Definition 3. Let $\mathcal{P}$ be a permutation over $\mathbb{F}_{p^n} \equiv \mathbb{F}_p^n$. Its maximum differential probability is defined as $\text{DP}_{\max} = \max_{\Delta_I, \Delta_O \in \mathbb{F}_p^n \setminus \{0\}} \text{Prob}_{\mathcal{P}}(\Delta_I \to \Delta_O)$.

Remark. In the following, we will use $\text{DP}_{\max}(x \mapsto x^d) \leq (d-1)/p$. However, we point out that this is just an upper bound. For example, consider the case $d = -1$, which corresponds to $d = p-2$. By the previous assumption, we get $\text{DP}_{\max}(x \mapsto x^{-1}) = (p-2)/p$, while $\text{DP}_{\max}(x \mapsto x^{-1}) = 4/p$. As another example, assume that $d = 1/d'$ where $d' \ll d$. In such a case, $\text{DP}_{\max}(x \mapsto x^{1/d'}) = \text{DP}_{\max}(x \mapsto x^{d'}) = (d'-1)/p$, which is much smaller than $(d-1)/p$.

Algebraic Attacks. Algebraic attacks exploit the algebraic description of the attacked cryptographic schemes in terms of polynomials. Examples of algebraic attacks include the interpolation attack [JK97], the higher-order differential attack [Lai94, Knu94], the factorization/GCD attacks and Gröbner basis attacks [Buc76, SS21], among others.

In the case of an interpolation attack, the attacker constructs the polynomial describing the scheme. This polynomial can be used for a key-recovery attack or in a forgery attack. The cost of constructing the polynomial depends on the number of its monomials, and this number depends both on the degree and the density of the polynomial. Knowing the maximum degree of the scheme allows to set up a zero-sum distinguisher, used in a higher-order differential attack. In the case of the factorization/GCD attacks and Gröbner basis attacks, the attacker solves a system of equations describing the scheme in order to find the secret key in the case of a cipher or a preimage or collision in the case of a hash function. The cost of the attack depends on several factors, including the number of equations, the number of variables, and the degrees and the density of the equations.

2.3 Solving Algebraic Equations

For the follow-up, we present the details of the algebraic techniques that we consider in our attacks on *Grendel*.

2.3.1 Univariate Factorization and Root Finding

Polynomial factorization can be used to solve a single univariate equation. More formally, setting a polynomial $F(x) \in \mathbb{F}_p[x]$ to zero, with factorization we are able to solve this polynomial for x. This problem is well-studied in literature, and has an estimated complexity in $\mathcal{O}(D^3n^2 + Dn^3)$ for a polynomial $F(x) \in \mathbb{F}_{p^n}[x]$ of degree D [Gen07]. This algorithm is based on matrix and vector operations and split into different steps. In total, the hidden constant is relatively small and thus the estimate can mostly be used directly.

While a full factorization can yield all roots of a polynomial, this will not be needed for our purposes. In our case, one root is sufficient for setting up the attack. In order to find it, it is sufficient to compute the GCD between $F(x)$ and $x^p - x$. Considering a polynomial $F(x) \in \mathbb{F}_p[x]$, $(x - x^\star)$ divides $\gcd(F(x), x^p - x)$ for an *existing* root $x^\star$ of F with probability 1, since $x^p - x = 0$ for all $x \in \mathbb{F}_p$ (indeed, $x \cdot (x^{p-1} - 1) = 0$ since either $x = 0$ or $x^{p-1} = 1$ due to Fermat's little theorem). We expect $\deg(\gcd(F(x), x^p - x)) \ll \deg(F(x))$, and a final more efficient (low-degree) factorization of $\gcd(F(x), x^p - x)$ is sufficient to recover the solution $x^\star$.

The complexity of computing GCDs is an element of $\mathcal{O}(D \log^2 D)$, and hence this method may be more efficient than the straight-forward factorization approach with degree-D polynomials. However, $x^p - x$ is of high degree. To avoid this, observe that

$$\gcd(F(x), x^p - X) = \gcd((x^p \mod F(x)) - X, F(x)),$$

since

$$\gcd(F(x), G(x)) = \gcd(a_1 \cdot F(x) + b_1 \cdot G(x), a_2 \cdot F(x) + b_2 \cdot G(x)),$$

where $a_1 \cdot b_2 \neq a_2 \cdot b_1$. Hence, the degree of $(x^p \mod F(x)) - x$ is much lower than that of $x^p - x$ if $\deg(F) \ll p$.

This root-finding approach using GCD computations is given in [vzGG13], and the final complexity of the algorithm is estimated by

$$D \cdot (\log_2(D))^2 \cdot (\log_2(D) + \log_2(p)) \cdot (1 + 63.43 \cdot \log_2(\log_2(D))), \tag{2}$$

which includes the complexity for *(1st)* polynomial multiplication resulting from repeated squaring to obtain $x^p \mod F(x)$, *(2nd)* GCD computations, and *(3rd)* the complexity for equal-degree factorization into low-degree factors. We note that this algorithm was also used in a recent attack [RAS20].

2.3.2 Gröbner Bases

We will use Gröbner basis computations in cases where the number of equations and variables is greater than 1. A Gröbner basis attack consists of three steps.

1. First, the attacker sets up the equation system and compute a Gröbner basis for it.

2. Secondly, they perform a change of term ordering for the basis (e.g., choosing a term order which makes it easier to eliminate variables and find the solutions).

3. Finally, the attacker finds the solution of the system obtained in the second step.

The cost of computing a Gröbner basis for an input system of n_e equations in n_v unknowns is estimated to be an element of

$$\mathcal{O}\left(\binom{D_{\text{reg}} + n_v}{n_v}^{\omega}\right), \tag{3}$$

where $2 < \omega \leq 3$ is the linear algebra constant representing the cost of matrix multiplication [BFP12]. The constants hidden by $\mathcal{O}(\cdot)$ are relatively small [ACG⁺19], and hence we use Eq. (3) directly. In this representation, D_{reg} is the *degree of regularity*, which for regular sequences [BFSY05] (that is, $n_v = n_e$) can be estimated by $1 + \sum_{i=1}^{n_e} \deg(f_i) - 1$.

For performance reasons, the Gröbner Basis is typically computed in the *degrevlex* term order, which is a monomial ordering, and by using algorithms such as [Fau02]. However, eliminating variables and finding the final solutions is in general more efficient for e.g. the *lex* term order [Tra00], which is why the second step usually consists of converting the Gröbner basis to a different monomial ordering. This can be done by algorithms such as FGLM [FGLM93]. After that, the final step in the attack consists of solving the system, which is done by first finding a solution to (one of) the univariate equation(s) in the system.

In our analysis, we will focus on the first step, i.e., the process of computing a Gröbner basis. For a more detailed description and analysis of the other steps involved in a Gröbner basis attack, we refer to [ACG⁺19, Section 3.3].

3 Related Work (and More)

3.1 Hermite's Criterion and Invertible Maps over $\mathbb{F}_{p^n} \equiv \mathbb{F}_p^n$

Let $p \geq 2$ be a prime number and let $n \geq 1$. Given a nonlinear polynomial function $F(x) = \sum_{i=0}^{d} \alpha_i \cdot x^i$ over $\mathbb{F}_{p^n}$ of degree $d \geq 2$, Hermite's criterion provides a necessary and sufficient condition for F to be a permutation.

Theorem 1 (Hermite's Criterion [MP13])**.** *Let $q = p^n$ for a prime $p \geq 2$ and a positive integer n. A polynomial $F \in \mathbb{F}_q[x]$ is a permutation polynomial (PP) of $\mathbb{F}_q$ if and only if*

(1) the reduction of $(F(x))^{q-1} \mod (x^q - x)$ is a monic polynomial of degree $q - 1$, and

(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.

Power Maps and Dickson Polynomials. For certain classes of polynomials, including the power maps $x \mapsto x^d$ and the Dickson polynomials, this question is easy to answer.

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

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

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

for fixed $\alpha \in \mathbb{F}_q$. We recall that a Dickson polynomial $\mathcal{D}_{d,\alpha}$ is invertible if $\gcd(d, p^2 - 1) = 1$.

Invertible Functions Over $\mathbb{F}_{p^n}$ via Linearized Polynomials. Another class of permutations over $\mathbb{F}_{p^n}$ is the class of linearized polynomials.

Definition 4. A linearized polynomial $\mathcal{L}(x)$ is a polynomial of the form $\mathcal{L}(x) := \sum_{i=0}^{d} \lambda_i \cdot x^{p^i}$ for some fixed $\lambda_0, \lambda_1, \ldots, \lambda_d \in \mathbb{F}_{p^n}$ and for a fixed $d \le n-1$.

The trace function $\mathrm{Tr}(x) = \sum_{i=0}^{n-1} x^{p^i}$ is an example of a linearized polynomial.

Proposition 2 ([MP13]). *Let $p \ge 3$ be a prime integer number. The linearized polynomial $\mathcal{L}(x) = \sum_{i=0}^{n-1} \lambda_i \cdot x^{p^n} \in \mathbb{F}_{p^n}[x]$ is a PP of $\mathbb{F}_{p^n}[x]$ if and only if $\det(M) \neq 0$, where $M_{i,j} := (\lambda_{i-j})^{p^j}$ for $i, j \in \{0, 1, \ldots, n-1\}$.*

Since linearized polynomials are linear, they satisfy $\mathcal{L}(x+y) = \mathcal{L}(x) + \mathcal{L}(y)$ for each $x, y \in \mathbb{F}_{p^n}$. They can be used as a starting point to construct nonlinear permutations over $\mathbb{F}_{p^n}$. Concrete examples are given by Tu et al. [TZLH15] and by Li, Helleseth and Tang [LHT13], who studied the class of permutation polynomials on $\mathbb{F}_{p^n}$ of the form $\left(x^{p^l} - x + \delta\right)^s + \mathcal{L}(x)$, for $l, s \in \mathbb{N}$ and for fixed $\delta \in \mathbb{F}_{p^n}$. They proved that $\left(x^{p^l} - x + \delta\right)^{\frac{p^n+1}{2}} + x^{p^l} + x$ is a PP over $\mathbb{F}_{p^n}$ for each $\delta \in \mathbb{F}_{p^n}$.

3.2 The Permutation $x \mapsto x^d \cdot (L_p(x) + \alpha)$ over $\mathbb{F}_p$

Theorem 3 ([Sha12]). *Given a prime number $p \ge 3$, let $\alpha \in \mathbb{F}_p \setminus \{\pm 1\}$ be such that $L_p(\alpha^2 - 1) = 1$. The function $F(x) = x \cdot (L_p(x) + \alpha)$ over $\mathbb{F}_p$ is a permutation.*

Here we show a variant of this function that is still a permutation.

Proposition 3. *Given a prime number $p \ge 3$, let $\alpha \in \mathbb{F}_p \setminus \{\pm 1\}$ be such that $L_p(\alpha^2 - 1) = 1$ and let $d \ge 1$ be such that $\gcd(d, p-1) = 1$. The function*

$$F(x) = x^d \cdot (L_p(x) + \alpha) \equiv x^{d+(p-1)/2} + \alpha \cdot x^{(p-1)/2}$$

over $\mathbb{F}_p$ is a permutation.

Proof. Given $y = F(x)$, $x = 0$ if and only if $y = 0$. Assuming $x, y \neq 0$ note that

$$L_p(y) = L_p(x^d) \cdot L_p(L_p(x) + \alpha) = L_p(x) \cdot L_p(L_p(x) + \alpha) = L_p(x) \cdot L_p(\alpha \pm 1),$$

where $L_p(x^d) = (L_p(x))^d = L_p(x)$ due to Proposition 1 and since d is an odd integer (note that $\gcd(d, p-1) = 1$, where $p-1$ is even). Since $L_p(\alpha - 1) = L_p(\alpha + 1)$, $L_p(x) = L_p(y)/L_p(\alpha \pm 1)$, and due to Theorem 2 we have that $x = \left(\frac{y \cdot L_p(\alpha \pm 1)}{\alpha \cdot L_p(\alpha \pm 1) + L_p(y)}\right)^{1/d}$. □

By noting that $L_p(1) = L_p(-1)$ for $p = 1 \mod 4$ (hence, by setting $\alpha = 0$ in the previous proposition), we get the following corollary.

Corollary 1. *Let $p \ge 3$ be a prime, and let $d \ge 1$ be such that $\gcd(d, p-1) = 1$. If $p = 1 \mod 4$, then $x \mapsto x^d \cdot L_p(x)$ is a permutation.*

Differential Property. Before going on, we study the differential properties of the permutation just proposed. The linear ones can be found in Appendix A.1.

Lemma 1. *Let $p \ge 3$ be a prime and let $F(x) = x \cdot (L_p(x) + \alpha)$. For $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$,*

$$|\{x \in \mathbb{F}_p \mid F(x + \Delta_I) - F(x) = \Delta_O\}| \le \begin{cases} 0 & \text{if } \Delta_O \neq (\pm 1 + \alpha) \cdot \Delta_I, \\ \frac{p+3}{4} & \text{if } \Delta_O = (\pm 1 + \alpha) \cdot \Delta_I \text{ and } p = 1 \mod 4, \\ \frac{p+1}{4} & \text{if } \Delta_O = (\pm 1 + \alpha) \cdot \Delta_I \text{ and } p = 3 \mod 4. \end{cases}$$

Proof. For each fixed $(\Delta_I, \Delta_O) \in \mathbb{F}_p^2 \setminus \{(0,0)\}$, we analyze the number of solutions x of

$$(x + \Delta_I) \cdot (L_p(x + \Delta_I) + \alpha) - x \cdot (L_p(x) + \alpha) = \Delta_O. \tag{4}$$

We separately analyze the cases

(1) $L_p(x + \Delta_I) = 0$ or $L_p(x) = 0$, and

(2) $L_p(x + \Delta_I) = \pm 1$ and $L_p(x) = \pm 1$.

Let us first focus on $L_p(x + \Delta_I) = 0$ or $L_p(x) = 0$. Obviously, $L_p(x + \Delta_I) = 0$ if and only if $x = -\Delta_I$, which implies $\Delta_I \cdot (L_p(-\Delta_I) + \alpha) = \Delta_O$. If this equality is satisfied, then $x = -\Delta_I$ is a solution. A similar result holds for $L_p(x) = 0$ (i.e., $x = 0$), which implies $\Delta_I \cdot (L_p(\Delta_I) + \alpha) = \Delta_O$. Again, if this equality is satisfied, then $x = 0$ is a solution. Note that $\Delta_I \cdot (L_p(-\Delta_I) + \alpha) = \Delta_O$ and $\Delta_I \cdot (L_p(\Delta_I) + \alpha) = \Delta_O$ cannot hold simultaneously.

We now consider the case $L_p(x + \Delta_I) = \pm 1$ and $L_p(x) = \pm 1$. We analyze the case $L_p(x + \Delta_I) = -L_p(x)$ and the case $L_p(x + \Delta_I) = L_p(x)$ separately.

1. If $L_p(x + \Delta_I) = -L_p(x) = \omega \in \{-1, +1\}$, then Eq. (4) reduces to

$$2\omega \cdot x + \Delta_I \cdot (\omega + \alpha) = \Delta_O \qquad \Longrightarrow \qquad x = \frac{\Delta_O - \Delta_I \cdot (\omega + \alpha)}{2 \cdot \omega}.$$

   This is a valid solution if $L_p\left(\frac{\Delta_O - \Delta_I \cdot (\omega+\alpha)}{2\cdot\omega} + \Delta_I\right) = -L_p\left(\frac{\Delta_O - \Delta_I \cdot (\omega+\alpha)}{2\cdot\omega}\right) = \omega$.

2. If $L_p(x + \Delta_I) = L_p(x) = \omega \in \{-1, +1\}$, then Eq. (4) reduces to $\Delta_I \cdot (\omega + \alpha) = \Delta_O$, which is satisfied independently of x.

Focusing on the second case and assuming $\Delta_I \cdot (\pm 1 + \alpha) = \Delta_O$, we want to find the number of solutions of $L_p(x + \Delta_I) = L_p(x)$. First of all, the number of x which satisfy $L_p(x) = L_p(x + \Delta_I)$ is equal to the number of x' which satisfy $L_p(x') = L_p(x' + 1)$, where $x' = x/\Delta_I$. By [GU82, Corollary 1.9], we have the following.

- If $p = 1 \mod 4$, the number of x' which satisfy $L_p(x') = L_p(x'+1) = 1$ is $\frac{p-5}{4}$, while the number of x' which satisfy $L_p(x') = L_p(x' + 1) = -1$ is $\frac{p-1}{4}$.
- If $p = 3 \mod 4$, the number of x' which satisfy $L_p(x') = L_p(x' + 1) = \pm 1$ is $\frac{p-3}{4}$.

In both cases, one additional solution could be $x = 0$ or $x = -\Delta_I$ (see before), and the final result follows immediately by adding it. □

Lemma 2. *Let* $p \geq 3$ *be a prime and let* $d \geq 3$ *be such that* $\gcd(d, p-1) = 1$*. Let* $F(x) = x^d \cdot (L_p(x) + \alpha)$*. For each* $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$*,*

$$|\{x \in \mathbb{F}_p \mid F(x + \Delta_I) - F(x) = \Delta_O\}| \leq 2d.$$

Proof. Since $F(x+\Delta_I) - F(x) = \Delta_O$ corresponds to $(x+\Delta_I)^d \cdot L_p(x+\Delta_I) - x^d \cdot L_p(x) = \Delta_O$, we have the following.

- If $L_p(x) = 0$, the equation is satisfied if $\Delta_I^d \cdot L_p(\Delta_I) = \Delta_O$. In a similar way, if $L_p(x + \Delta_I) = 0$, then the equality is satisfied if $(-\Delta_I)^d \cdot L_p(\Delta_I) = \Delta_O$. As before, the two cases cannot hold simultaneously.
- If $L_p(x) = L_p(x + \Delta_I) \in \{-1, 1\}$, the equation has degree $d - 1$, and it admits at most $d - 1$ solutions.
- If $L_p(x) = -L_p(x + \Delta_I) \in \{-1, 1\}$, the equation has degree d, and it admits at most d solutions.

It follows that its $\text{DP}_{\max}$ is equal to $(1 + d - 1 + d)/p = (2d)/p$. □

3.3 *Grendel*'s Nonlinear S-Box $x \mapsto x^d \cdot L_p(x)$

Proposition 4 ([Sze21])**.** *Let $p \geq 3$ be a prime and $d \geq 1$ be an integer such that* $\gcd(d + (p-1)/2, p-1) = 1$*. The map $x \mapsto x^d \cdot L_p(x) = x^{d+(p-1)/2}$ is invertible over $\mathbb{F}_p$.*

The proof is based on the Hermite's criterion, that is, a function $x \mapsto x^{d'}$ is invertible if $\gcd(d', p-1) = 1$. Corollary 2 corollary follows immediately.

Corollary 2. *Let $p \geq 3$ be a prime integer. The function $x \mapsto x^d \cdot L_p(x) = x^{d+(p-1)/2}$ is a permutation over $\mathbb{F}_p$ if $d = 1$ and $p = 1 \mod 4$ or if $d = 2$ and $p = 3 \mod 4$.*

Proof. If $p = 1 \mod 4$ and $d = 1$, then the result follows from Corollary 1.

Otherwise, if $p = 3 \mod 4$ and $d = 2$, then $\gcd(2 + (p-1)/2, p-1) = 1$ as shown in [Sze21, Sect. 3]. This implies that $x \mapsto x^2 \cdot L_p(x)$ is invertible due to Hermite's criterion. □

Moreover, let $d' \geq 1$ be such that $\gcd(d', p-1) = 1$. The function $x \mapsto x^{d \cdot d'} \cdot L_p(x)$ is a permutation over $\mathbb{F}_p$ if $p = 1 \mod 4$ and $d = 1$ or if $p = 3 \mod 4$ and $d = 2$. The result follows immediately since d' is an odd integer and $(\pm 1)^{d'} = \pm 1$.

Differential Property. The differential and linear properties of $x \mapsto x^d \cdot L_p(x)$ are similar to the ones of the function $x \mapsto x^d \cdot (L_p(x) + \alpha)$. Here, we focus on the differential property only, in order to point out a different approach to study them with respect to the one proposed in [Sze21].

In order to study the differential properties of $F(x) = x^d \cdot L_p(x)$, in [Sze21] it is proposed to use the square operation in order to cancel the Legendre symbol that depends both on the input x and on the input difference Δ_I, since

$$\begin{aligned}&(x + \Delta_I)^d \cdot L_p(x + \Delta_I) = x^d \cdot L_p(x) + \Delta_O \\ &\qquad \Longrightarrow \; (x + \Delta_I)^{2d} = x^{2d} + \Delta_O^2 + 2\Delta_O \cdot L_p(x) \cdot x^d.\end{aligned}$$

In this way, assuming $d \geq 2$ and $x \neq -\Delta_I$ (for which $L_p(x + \Delta_I) = 0 \neq 1$), we get an equation of degree $2d - 1$, which can have at most $2d - 1$ solutions. Since there are two nonzero values for $L_p(x)$, we get $2 \cdot (2d-1) = 4d - 2$ solutions (one more solution could be $x = 0$). However, half of these solutions are "false" solutions, i.e., they are not solutions of the original differential equation. For this reason, here we propose a more specific analysis.

Lemma 3. *Let $p \geq 3$ be a prime number such that $p = 3 \mod 4$. Let $d' \geq 1$ be such* $\gcd(d', p-1) = 1$*. The maximum differential probability of $F(x) = x^{2 \cdot d'} \cdot L_p(x)$ is $(4 \cdot d')/p$.*

Lemma 4. *Let $p \geq 3$ be a prime number such that $p = 1 \mod 4$. Let $d' \geq 1$ be such* $\gcd(d', p-1) = 1$*, and let $F(x) = x^{d'} \cdot L_p(x)$. Then*

$$|\{x \in \mathbb{F}_p \mid F(x + \Delta_I) - F(x) = \Delta_O\}| \leq \begin{cases} \frac{p+3}{4} & \textit{if } d' = 1, \\ 2d' & \textit{otherwise.} \end{cases}$$

The results follow from the previous analysis for the case $F(x) = x^d \cdot (L_p(x) + \alpha)$.

4 New Permutations over $\mathbb{F}_p^n$ via the Legendre Symbol and the Modulo-2 Operation

4.1 $F(x) = x^{d_+} \cdot (1 + L_p(x)) + x^{d_-} \cdot (1 - L_p(x))$

Proposition 5. *Let $p \geq 3$ be a prime integer. Let $d_+, d_- \geq 1$ be integers such that* $\gcd(d_+, p-1) = \gcd(d_-, p-1) = 1$. *Then, the function*

$$F(x) = \frac{x^{d_+} \cdot (1 + L_p(x)) + x^{d_-} \cdot (1 - L_p(x))}{2} = \begin{cases} x^{d_+} & \textit{if } L_p(x) = 1, \\ x^{d_-} & \textit{otherwise} \end{cases}$$

is invertible over $\mathbb{F}_p$.

Proof. If $d_+ = d_- = d$, this is obvious, since it reduces to $x \mapsto x^d$. Assume $d_+ \neq d_-$. Given $2y = x^{d_-} \cdot (1 - L_p(x)) + x^{d_+} \cdot (1 + L_p(x))$, its inverse is

$$x = \begin{cases} 0 & \text{if } L_p(y) = y = 0, \\ y^{1/d_-} & \text{if } L_p(y) = -1, \\ y^{1/d_+} & \text{if } L_p(y) = 1. \end{cases} \implies 2x = (1 - L_p(y)) \cdot y^{1/d_-} + (1 + L_p(y)) \cdot y^{1/d_+},$$

where note that $L_p(x) = L_p(y)$. Indeed, $x = 0$ implies $y = L_p(y) = 0$ (and vice-versa). If $x, y \neq 0$ and $L_p(x) = 1$ (analogous for $L_p(x) = -1$), then $y = x^{d_+}$, which implies $L_p(y) = L_p(x^{d_+}) = L_p(x)^{d_+} = L_p(x)$ since d_+ is odd (similarly for $x \neq 0$ and $L_p(x) = -1$). □

Multiplicative Complexity and Side Channel Attacks. Even though the function just given is invertible, it has some undesirable properties for cryptographic purposes. Depending on the value of x, the number of multiplications required to compute $x \mapsto x^{d_+} \cdot (1 + L_p(x)) + x^{d_-} \cdot (1 - L_p(x))$ varies. Besides making cost estimations difficult (i.e., the scheme can either be cheap or expensive), this may allow to set up side-channel attacks [Koc96, KJJ99] (i.e., attacks enabled by the leakage of information from a physical cryptosystem). For this reason, we do not encourage to use it, and we do not propose an analysis of its linear and differential properties.

4.2 Permutations over $\mathbb{F}_p$ via the Modulo-2 Operation

4.2.1 $F(x) = \alpha^{(x \mod 2)} \cdot x^2$

As we have seen before, the quadratic power function $x \mapsto x^2$ is never a PP over $\mathbb{F}_p$ for $p \geq 3$. Here we present a variant of such function that is a permutation, and that is defined via the modulo-2 operation.

Theorem 4. *Let $p \geq 3$ be a prime number. Let $\alpha \in \mathbb{F}_p$ be a quadratic non-residue modulo p, as defined in Definition 1. The function*

$$F(x) = \alpha^{(x \mod 2)} \cdot x^2 = \begin{cases} x^2 & \textit{if } x = 0 \mod 2, \\ \alpha \cdot x^2 & \textit{if } x = 1 \mod 2 \end{cases}$$

is a permutation over $\mathbb{F}_p$.

The function also admits an equivalent representation via $x \mapsto (-1)^x$ instead of the modulo-2 operation, that is,

$$F(x) = \left(\frac{\alpha \cdot (1 - (-1)^x) + 1 + (-1)^x}{2} \right) \cdot x^2.$$

Proof. We prove that the function is injective (that is, $F(x) = F(y)$ implies $x = y$). Then, since it is defined over a finite field, it follows that the function is invertible.

First of all, note that there is no $x, y \in \mathbb{F}_p \setminus \{0\}$ such that $\alpha \cdot x^2 = y^2$. Indeed, $\alpha \cdot x^2$ is a quadratic non-residue modulo p (since $L_p(\alpha \cdot x^2) = L_p(\alpha) \cdot L_p(x^2) = L_p(\alpha) = -1$), while y^2 is a quadratic residue modulo p. It follows that

$$\alpha^{(x \mod 2)} \cdot x^2 = \alpha^{(y \mod 2)} \cdot y^2 \qquad \Longrightarrow \qquad x \mod 2 = y \mod 2.$$

Hence, only two scenarios can occur, either $x^2 = y^2$ or $\alpha \cdot x^2 = \alpha \cdot y^2$. In both cases, the only solutions are $x = \pm y$. Note that

$$x = 0 \mod 2 \qquad \text{if and only if } -x = 1 \mod 2,$$

since $-x = p - x$, where p is odd. In conclusion, $F(x) = F(y)$ implies $x = y$ ($x = -y$ is not possible since the condition $x \mod 2 = y \mod 2$ is not satisfied). □

Before going on, we study the case $\alpha = -1$ in more detail.

Corollary 3. *Let $p \geq 3$ be a prime number such that $p = 3 \mod 4$. The function $F(x) = (-1)^x \cdot x^2$ is a permutation over $\mathbb{F}_p$.*

Proof. This is a direct application of Theorem 4, where -1 is a quadratic non-residue modulo p if and only if $p = 3 \mod 4$ (see Proposition 1) and where $(-1)^x = (-1)^{(x \mod 2)}$, since $(x \mod 2) \mod (p-1) = (x \mod (p-1)) \mod 2$ for each $x \in \{0, 1, 2, \ldots, p-1\}$ (remember that the exponent is modulo $p - 1$). □

Differential Property. Here we study the differential properties of the permutation just proposed. The linear ones can be found in Appendix A.2.

Lemma 5. *Let $p \geq 3$ be a prime number, and let $\alpha \in \mathbb{F}_p$ be such that $L_p(\alpha) = -1$. Let $F(x) = (\alpha)^{x \mod 2} \cdot x^2$. For each $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$*

$$|\{x \in \mathbb{F}_p \mid F(x + \Delta_I) - F(x) = \Delta_O\}| \leq 6.$$

If $\alpha = -1$ and $p = 3 \mod 4$,

$$|\{x \in \mathbb{F}_p \mid F(x + \Delta_I) - F(x) = \Delta_O\}| \leq \begin{cases} 2 & \text{if } \Delta_I \mod 2 = 0, \\ 4 & \text{otherwise.} \end{cases}$$

Proof. Given $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$, we analyze the number of solutions x of

$$\alpha^{((x+\Delta_I) \mod 2)} \cdot (x + \Delta_I)^2 - \alpha^{(x \mod 2)} \cdot x^2 = \Delta_O$$

by separately studying the cases $\alpha^{((x+\Delta_I) \mod 2)} = \alpha^{(x \mod 2)}$ and $\alpha^{((x+\Delta_I) \mod 2)} = \alpha^{(x \mod 2)}$.

- If $\alpha^{((x+\Delta_I) \bmod 2)} = \alpha^{(x \bmod 2)}$, the equation admits at most two solutions, i.e.,

$$\forall k \in \{0, 1\} : \qquad x = \frac{\Delta_O - \Delta_I^2 \cdot (\alpha)^k}{2\Delta_I \cdot (\alpha)^k}.$$

- If $\alpha^{((x+\Delta_I) \bmod 2)} \neq \alpha^{(x \bmod 2)}$, the equation admits at most four solutions, i.e., at most two for $(\alpha - 1) \cdot x^2 + 2\alpha \cdot \Delta_I \cdot x + \alpha \cdot \Delta_I^2 = \Delta_O$ and at most two for $(1 - \alpha) \cdot x^2 + 2\Delta_I \cdot x + \Delta_I^2 = \Delta_O$.

Hence, the total number of solutions is at most six.

Finally, in the case in which $\alpha = -1$ and $p = 3 \mod 4$, note that $(-1)^{(x+\Delta_I)} = (-1)^x$ if and only if $\Delta_I = 0 \mod 2$. □

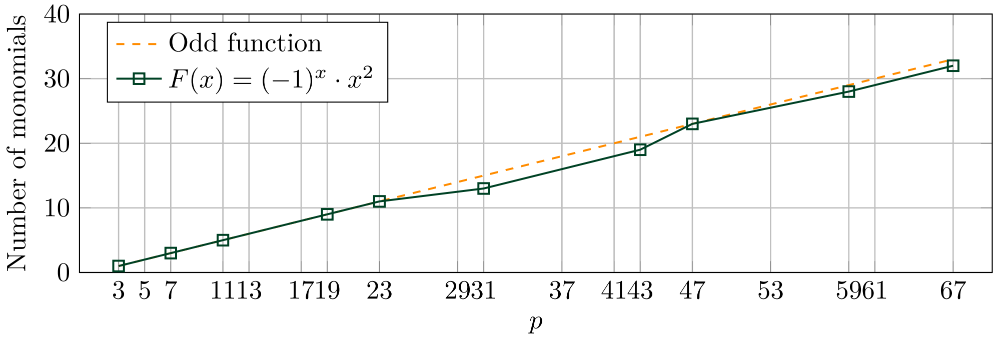

Figure 3: Comparison of the number of monomials for a generic odd function and for $F(x) = (-1)^x \cdot x^2$ for several values of p.

Algebraic Property. Regarding the algebraic properties, we first prove the following.

Proposition 6. *Let $p \geq 3$ be a prime number, and let $\alpha \in \mathbb{F}_p$ be such that $L_p(\alpha) = -1$. There exist $\varphi_1, \varphi_3, \ldots, \varphi_{p-2} \in \mathbb{F}_p$ such that*

$$\alpha^{(x \mod 2)} \cdot x^2 = \frac{1+\alpha}{2} \cdot x^2 + \sum_{i=0}^{(p-3)/2} \varphi_{2i+1} \cdot x^{2i+1}.$$

Moreover, $\sum_{i=0}^{(p-3)/2} \varphi_{2i+1} = \frac{\alpha-1}{2}$.

Note that if $\alpha = -1$, then $\frac{1+\alpha}{2} = 0$. It follows that $F(x) = (-1)^x \cdot x^2$ (for $p = 3 \mod 4$) is an odd function.

Proof. In order to prove the result, it is sufficient to show that $F(x) - \frac{1+\alpha}{2} \cdot x^2 = \alpha^{(x \mod 2)} \cdot x^2 - \frac{1+\alpha}{2} \cdot x^2$ is an odd function. We recall that a function G is odd if and only if $G(-x) = -G(x)$ for each $x \in \mathbb{F}_p$. In our case, the equality

$$F(-x) - \frac{1+\alpha}{2} \cdot (-x)^2 = \left(\alpha^{(-x \mod 2)} - \frac{1+\alpha}{2}\right) \cdot (-x)^2$$
$$= -\left(\alpha^{(x \mod 2)} - \frac{1+\alpha}{2}\right) \cdot x^2 = -\left(F(x) - \frac{1+\alpha}{2} \cdot x^2\right)$$

is always satisfied since $x \mod 2 = 0$ if and only if $-x \mod 2 = 1$. Hence, $1 - \frac{1+\alpha}{2} = \frac{1-\alpha}{2} = -\frac{\alpha-1}{2} = -\alpha + \frac{1+\alpha}{2}$. Finally, $\sum_{i=0}^{(p-3)/2} \varphi_{2i+1} = F(1) - \frac{1+\alpha}{2} = \frac{\alpha-1}{2}$. □

By practical experiments, we found that $F(x) = \alpha^{(x \mod 2)} \cdot x^2$ (where $L_p(\alpha) = -1$) and its inverse are functions of maximum degree. For their density, we found the following.

- Fig. 3 for $\alpha = -1$ (hence, $p = 3 \mod 4$): $F(x) = (-1)^x \cdot x^2$ is an odd function, as expected.

- Fig. 4 for $\alpha = 2$ (hence, $p = 3, 5 \mod 8$) and Fig. 5 for $\alpha = -3$ (hence, $p = 2 \mod 3$): We found that $F(x)$ is of the form $\varphi_2 \cdot x^2 + \sum_{i \geq 0} \varphi_{2i+1} \cdot x^{2i+1}$, as expected.

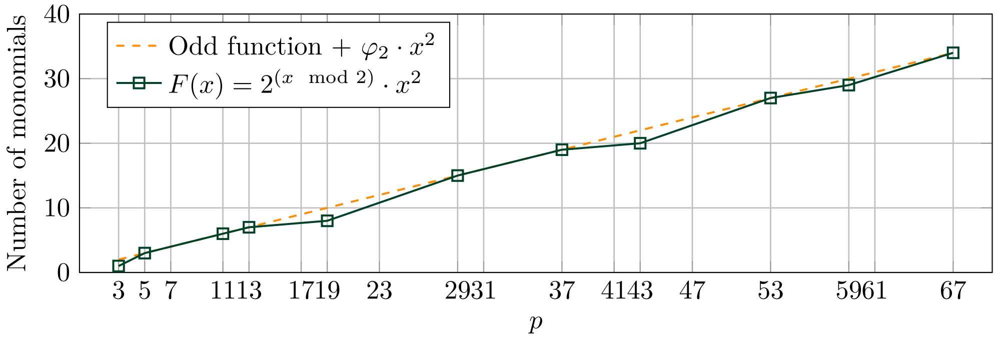

Figure 4: Comparison of the number of monomials for a generic function (with zero constant) and for $F(x) = 2^{(x \mod 2)} \cdot x^2$ for several values of p.

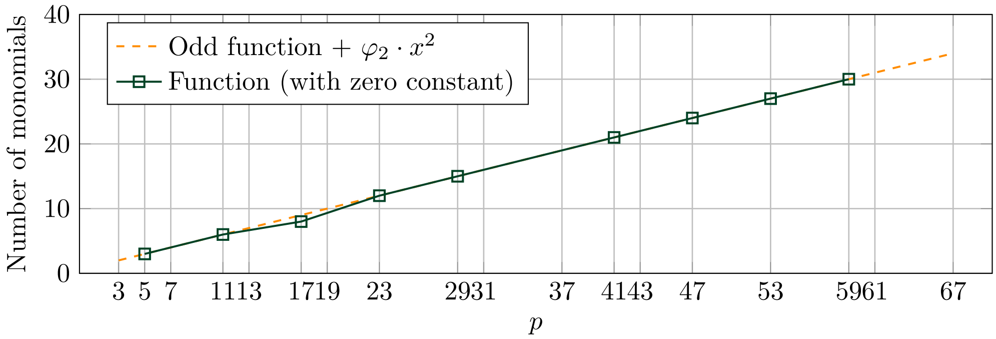

Figure 5: Comparison of the number of monomials for a generic function (with zero-constant) and for $F(x) = (-3)^{(x \mod 2)} \cdot x^2$ for several values of p.

4.2.2 $F(x) = x \cdot (1 - 2 \cdot (x^2 \mod 2))$

A function based on the modulo-2 operation that is always invertible for each p is $F(x) = x \cdot (1 - 2 \cdot (x^2 \mod 2))$.

Lemma 6. *Let $p \geq 3$ be a prime number. The function*

$$F(x) = x \cdot (1 - 2 \cdot (x^2 \mod 2)) \equiv (-1)^{x^2} \cdot x$$

is invertible over $\mathbb{F}_p$.

Proof. Note that $1 - 2 \cdot (x^2 \mod 2) \in \{-1, 1\}$ for each $x \in \mathbb{F}_p$. Hence, given $y = F(x) = x \cdot (1 - 2 \cdot (x^2 \mod 2))$, note that $y^2 = x^2$. Hence, $x = y \cdot (1 - 2 \cdot (y^2 \mod 2))$. □

Corollary 4. *Let $p \geq 3$ be a prime, and let $d \geq 1$ such that $\gcd(d, p-1) = 1$. Then $F(x) = (-1)^{x^2} \cdot x^d$ is invertible.*

Just observe that (1st) d is odd and (2nd) $(-1)^{x^2} \cdot x^d = ((-1)^{x^2} \cdot x)^d$ since $(\pm 1)^d = \pm 1$ for each odd d.

Differential Property. Here we study the differential properties of the permutation just proposed. The linear ones can be found in Appendix A.3.

Lemma 7. *Let $p \geq 3$ be a prime number and let $F(x) = (-1)^{x^2} \cdot x$. For each $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$*

$$|\{x \in \mathbb{F}_p \mid F(x+\Delta_I) - F(x) = \Delta_O\}| \leq \begin{cases} 0 & \textit{if } \Delta_I^2 = 1 \mod 2 \textit{ and } \Delta_I \neq \pm\Delta_O; \\ 2 & \textit{if } \Delta_I^2 = 0 \mod 2 \textit{ and } \Delta_I = \pm\Delta_O; \\ \frac{p+1}{2} & \textit{otherwise.} \end{cases}$$

Proof. Given $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$, we analyze the number of solutions x of $(-1)^{(x+\Delta_I)^2} \cdot (x+\Delta_I) - (-1)^{x^2} \cdot x = \Delta_O$, that is

$$x \cdot ((-1)^{\Delta_I^2} - 1) = (-1)^{x^2} \cdot \Delta_O - (-1)^{\Delta_I^2} \cdot \Delta_I.$$

Let's consider separately the case $\Delta_I^2 = 0 \mod 2$ and $\Delta_I^2 = 1 \mod 2$:

- If $\Delta_I^2 = 1 \mod 2$, we have $x = -\frac{(-1)^{x^2} \cdot \Delta_O + \Delta_I}{2}$, which admits at most two possible solutions, that is $-\frac{\Delta_I \pm \Delta_O}{2}$;
- If $\Delta_I^2 = 0 \mod 2$, we have $(-1)^{x^2} \cdot \Delta_O = \Delta_I$, which admits solution if and only if $\Delta_I = \pm\Delta_O$. In this case, there are at most $(p+1)/2$ solutions (since x^2 can take at most $(p+1)/2$ different values). □

Lemma 8. *Let $p \geq 3$ be a prime number and let $d \geq 3$ be such that* $\gcd(d, p-1) = 1$. *Let $F(x) = (-1)^{x^2} \cdot x^d$. For each $\Delta_I, \Delta_O \in \mathbb{F}_p \setminus \{0\}$:*

$$|\{x \in \mathbb{F}_p \mid F(x+\Delta_I) - F(x) = \Delta_O\}| \leq \begin{cases} 2(d-1) & \textit{if } \Delta_I^2 \mod 2 = 0 \\ 2d & \textit{otherwise} \end{cases}.$$

Proof. We have

$$(-1)^{(x+\Delta_I)^2} \cdot (x+\Delta_O)^d - (-1)^{x^2} \cdot x^d = \Delta_O,$$

where $(-1)^{(x+\Delta_I)^2} = (-1)^{x^2} \cdot (-1)^{\Delta_I^2}$. Hence:

- if $(-1)^{\Delta_I^2} = 1$, we have at most $2(d-1)$ solutions, that is, $d-1$ for the case $(-1)^{x^2} = 1$ and $d-1$ for the case $(-1)^{x^2} = -1$;
- if $(-1)^{\Delta_I^2} = -1$, we have at most $2d$ solutions. □

Algebraic Property (Case: $d = 1$). Regarding the algebraic properties, we first point out that the function $F(x) = (-1)^{x^2} \cdot x$ (and its inverse) is an odd function. Indeed, $F(-x) = (-1)^{(-x)^2} \cdot (-x) = -(-1)^{x^2} \cdot x = -F(x)$. By practical experiments (see Fig. 6), it turns out that both such function and its inverse are of maximum degree, but they are full/dense only for some values of p. In particular, if $p = 3 \mod 4$ (for which $L_p(-1) = -1$), then the algebraic representation of $F(x) = (-1)^{x^2} \cdot x$ is dense/full. Vice-versa, if $p = 1 \mod 4$ (for which $L_p(-1) = 1$), we found that

$$(-1)^{x^2} \cdot x = \sum_{i=0}^{(p-5)/4} \gamma_{4 \cdot i+3} \cdot x^{4 \cdot i+3} \tag{5}$$

for certain $\gamma_3, \gamma_7, \ldots, \gamma_{p-2} \in \mathbb{F}_p$, that is, the monomials with exponents $4 \cdot i + 1$ do not appear. E.g., the function $F(x) = (-1)^{x^2} \cdot x$ is equal to $-x^{(p+1)/2} \equiv -x \cdot L_p(x)$ for $p \in \{5, 13\}$. We leave the problem to formally prove Eq. (5) for future work.

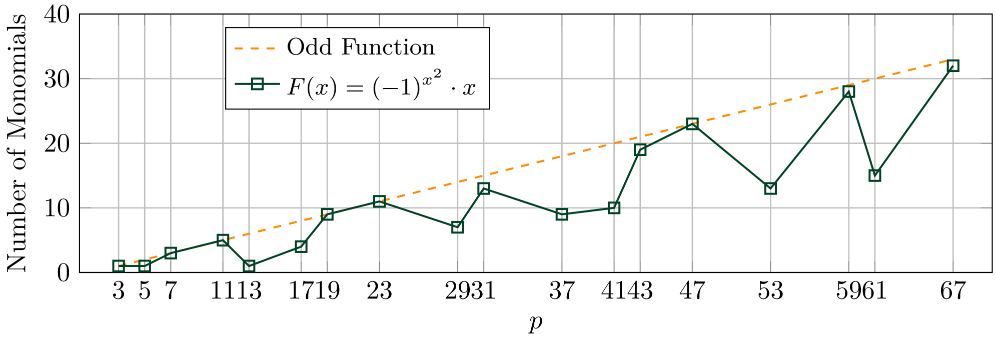

Figure 6: Comparison of the number of monomials for a generic odd function and for $F(x) = (-1)^{x^2} \cdot x$ for several values of p.

4.3 Nonlinear Layer over $\mathbb{F}_p^n$ via a Local Map

A nonlinear function $\mathcal{F}$ over $\mathbb{F}_p^n$ is defined

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

where

$$y_i = F_i(x_0, x_1, \dots, x_{n-1})$$

for n (potentially different) nonlinear functions $F_i : \mathbb{F}_p^n \to \mathbb{F}_p$ for $i \in \{0, 1, \dots, n-1\}$. Here, we limit ourselves to focus on the case in which the functions F_i are defined as $F_i(\cdot) = F \circ \mathcal{T}(\cdot, i)$ where

- F is a *local map* of the form $F : \mathbb{F}_p^m \to \mathbb{F}_p$ for a certain m of the form $2 \leq m \leq n$;
- $\mathcal{T}(\cdot, i)$ is a *translation* function $\mathcal{T} : \mathbb{F}_p^n \times \{0, 1, \dots, n-1\} \to \mathbb{F}_p^m$ defined as

  $$\mathcal{T}(x, i) = x_i \| x_{i+1} \| \dots \| x_{i+m-1},$$

  where the indexes are computed modulo n.

More formally:

Definition 5. 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 nonlinear function. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ is defined as

$$\mathcal{S}_F(x_0, x_1, \dots, x_{n-1}) := y_0 \| y_1 \| \dots \| y_{n-1} \tag{6}$$

where

$$y_i = F(x_i, x_{i+1}, \dots, x_{i+m-1}) \tag{7}$$

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

As already mentioned in the introduction, probably one of the most well known example of invertible functions of this form over $\mathbb{F}_2^n$ (for odd n) is the chi-function [Wol85, DGV92, Dae95], whose local map is $F(x_0, x_1, x_2) = x_1 \cdot x_2 + x_0 + x_2$.

Examples: Generalization of $F(x) = x^d \cdot (L_p(x) + \alpha)$ and of $F(x) = (-1)^{x^2} \cdot x^d$.

Lemma 9. *Given $p \geq 3$ be a prime number, let $d \geq 1$ be such that $\gcd(d, p-1) = 1$. Let $\alpha \in \mathbb{F}_p \setminus \{\pm 1\}$ such that $L_p(\alpha^2 - 1) = 1$. Let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be defined as $F(x_0, x_1) = x_0^d \cdot (\alpha + L_p(x_1))$. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Definition 5 is invertible.*

Proof. The proof is equivalent to the one given for Proposition 3. Given $y_i = F(x_i, x_{i+1})$, we get $L_p(y_i) = L_p(x_i) \cdot L_p(\alpha \pm 1)$, or equivalently $L_p(x_i) = \frac{L_p(y_i)}{L_p(\alpha\pm 1)}$, which implies $x_i = \frac{y_i \cdot L_p(\alpha\pm1)}{L_p(\alpha\pm1)\cdot\alpha + L_p(y_{i+1})}$. □

In a similar way, we get the following

Lemma 10. *Given $p \geq 3$ be a prime number and let $d \geq 1$ be such that $\gcd(d, p-1) = 1$. Let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be defined as $F(x_0, x_1) = (-1)^{x_0^2} \cdot x_1^d$. The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Definition 5 is invertible.*

The proof is similar to the one given for Lemma 6.

4.3.1 $F(x_0, x_1) = x_0 \cdot \big(x_1 \cdot (1 + L_p(x_1)) + (1 - L_p(x_1)\big)$

Theorem 5. *Let $p \geq 3$ be a prime number such that $p = 3 \mod 4$. Let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be defined as*

$$F(x_0, x_1) = x_0 \cdot \big(x_1 \cdot (1 + L_p(x_1)) + (1 - L_p(x_1)\big) = \begin{cases} 2 \cdot x_0 & \text{if } L_p(x_1) = -1 \\ x_0 & \text{if } x_1 = 0 \\ 2 \cdot x_0 \cdot x_1 & \text{if } L_p(x_1) = 1 \end{cases}.$$

The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Definition 5 is invertible for each odd $n \geq 3$.

Proof. Given $\mathcal{S}(x) = y$, we recursively construct the inverse $x \in \mathbb{F}_p^n$. Firstly, note that $\mathcal{S}(x) = 0 \in \mathbb{F}_p^n$ if and only if $x = 0 \in \mathbb{F}_p^n$, since $F(x_0, x_1) = 0$ if and only if $x_0 = 0$. We consider separately the two following cases:

1. there exists $i \in \{0, 1, \ldots, n-1\}$ such that $y_i = 0$;
2. $y_i \neq 0$ for each $i \in \{0, 1, \ldots, n-1\}$.

Given $y_i = 0$, one can immediately deduce that $x_i = 0$ due to the argument just provided. By working recursively for each $j < i$:

$$x_j = \begin{cases} y_j & \text{if } x_{j+1} = y_{j+1} = 0 \\ \frac{y_j}{x_{j+1}\cdot(1+L_p(x_{j+1}))+(1-L_p(x_{j+1}))} & \text{otherwise} \end{cases}, \tag{8}$$

where all sub-indexes are taken modulo n.

In the second case, assume that $y_i \neq 0$ for each $i \in \{0, 1, \ldots, n-1\}$. First of all, note that

$$\forall i \in \{0, 1, \ldots, n-1\} : \qquad L_p(x_i) = (L_p(2))^{-1} \cdot L_p(y_i) = L_p(2) \cdot L_p(y_i) = L_p(2 \cdot y_i),$$

where $(L_p(z))^{-1} = L_p(z)$ for each $z \neq 0$. Indeed, since $x_i \neq 0$ for each $i \in \{0, 1, \ldots, n-1\}$:

- if $L_p(x_{i+1}) = -1$, then $y_i = 2 \cdot x_i$, which implies $L_p(x_i) = (L_p(2))^{-1} \cdot L_p(y_i)$;
- if $L_p(x_{i+1}) = 1$, then $y_i = 2 \cdot x_i \cdot x_{i+1}$, which implies $L_p(x_i) = (L_p(2))^{-1} \cdot L_p(y_i) \cdot (L_p(x_{i+1}))^{-1} = L_p(x_i) = (L_p(2))^{-1} \cdot L_p(y_i)$.

This means that

$$y_i = x_i \cdot (x_{i+1} \cdot (1 + L_p(2 \cdot y_{i+1})) + (1 - L_p(2 \cdot y_{i+1}))\,,$$

Again, we have to consider two cases:

2.a there exists $i \in \{0, 1, \ldots, n-1\}$ such that $L_p(2 \cdot y_{i+1}) = -1$;

2.b $L_p(2 \cdot y_i) = 1$ for each $i \in \{0, 1, \dots, n-1\}$.

In the first case, if there exists $i \in \{0, 1, \dots, n-1\}$ such that $L_p(2 \cdot y_{i+1}) = -1$, then $x_i = y_i/2$. Given x_{i-1}, it is then possible to find all x_j working recursively as given in Eq. (8).

Vice-versa, assume that $L_p(2 \cdot y_i) = 1$ for each $i \in \{0, 1, \dots, n-1\}$. In such a case, we have that $y_i = 2 \cdot x_i \cdot x_{i+1}$ for each $i \in \{0, 1, \dots, n-1\}$. Working recursively with $x_i = \frac{y_i}{2 \cdot x_{i+1}}$ and since n is odd, we have that

$$x_0^2 = \frac{1}{2} \cdot \prod_{j=0}^{n-1} y_i^{(-1)^i} \qquad \rightarrow \qquad x_0 = \sigma \cdot \left(\frac{1}{2} \cdot \prod_{j=0}^{n-1} y_i^{(-1)^i} \right)^{(p+1)/4} ,$$

where $\sigma \in \{-1, +1\}$. Note that $z^{(p+1)/4}$ is the square root of the quadratic residue z modulo p since $p = 3 \mod 4$ (see Proposition 1 for details). The value of σ must satisfy $L_p(x_0) = L_p(2 \cdot y_0) = 1$ (since we are in the case in which $L_p(2 \cdot y_i) = 1$ for each $i \in \{0, 1, \dots, n-1\}$). Note that $L_p(-z) = -L_p(z)$ for each $z \in \mathbb{F}_p$ since $p = 3 \mod 4$, which means that σ is uniquely defined. Finally, given x_0, it is then possible to find all x_j working recursively as given before. □

4.3.2 $F(x_0, x_1) = x_0 \cdot (1 + x_1^2 - x_1^{p-1})$

Theorem 6. *Let $p \geq 3$ be a prime number. Let $F : \mathbb{F}_p^2 \to \mathbb{F}_p$ be defined as*

$$F(x_0, x_1) = x_0 \cdot (1 + x_1^2 - x_1^{p-1}) = \begin{cases} x_0 & \text{if } x_1 = 0 \\ x_0 \cdot x_1^2 & \text{otherwise} \end{cases} .$$

The function $\mathcal{S}$ over $\mathbb{F}_p^n$ defined as in Definition 5 is invertible for each odd $n \geq 3$ such that $\gcd(2^n + 1, p - 1) = 1$.

Note that $(L_p(x_1))^2 = x_1^{p-1}$ for each $x_1 \in \mathbb{F}_p$.

Proof. First of all, note that $\mathcal{S}(x) = 0 \in \mathbb{F}_p^n$ if and only if $x = 0 \in \mathbb{F}_p^n$. Given $y \in \mathbb{F}_p^3$, we show how to compute $x \in \mathbb{F}_p^3$ s.t. $\mathcal{S}(x) = y$. For doing this, we consider separately the two cases:

- there exists $i \in \{0, 1, \dots, n-1\}$ such that $y_i = 0$;
- $y_i \neq 0$ for each $i \in \{0, 1, \dots, n-1\}$.

In the first case, assume that there exists $i \in \{0, 1, \dots, n-1\}$ such that $y_i = 0$. By definition of F, it follows that $x_i = 0$ and that $x_{i-1} = y_{i-1}$. Working recursively for each $j < i$, we have $x_j = y_j/(1 + x_{j+1}^2 - x_{j+1}^{p-1})$ where the sub-indexes are taken modulo n.

In the second case, by simple computation, note that

$$x_i^{2^n+1} = \prod_{j=0}^{n-1} (y_{i+j})^{(-2)^j}$$

where the sub-indexes are taken modulo n. Since $x \mapsto x^{2^n+1}$ is a permutation, it is possible to find x_i. Given x_i, it is possible to find x_j for each j as before. □

5 Preimage Attack on Full *Grendel*

In this section, we present a preimage attack on a sponge hash function instantiated with the full *Grendel* permutation and a new Gröbner basis analysis based on the same ideas. The code for the practical verification of the proposed attack is available online.[1]

[1] https://github.com/mschof/grendel-analysis

5.1 High-Level Idea of the Attack

Security Analysis given in [Sze21]. We first start by recalling the security analysis given in [Sze21]. One advantage of using a function based on the Legendre symbol is that the corresponding scheme reaches its maximum degree after only a small number of rounds, which, together with density, helps against interpolation attacks. The main algebraic attack then becomes the Gröbner basis one recalled in Section 2.3.2. In [Sze21], it is proposed to set up the equations system to solve in two different ways, namely

(1) working at round level without guessing the Legendre symbols, or

(2) first guessing the Legendre symbols and then working at round level.

Regarding the first strategy, it is proposed to rewrite $x \mapsto y = x \cdot L_p(x)$ as

$$(\psi - 1) \cdot y = (x - z^2) - (\psi \cdot x - z^2) \qquad \text{and} \qquad (x - z^2) \cdot (\psi \cdot x - z^2) = 0,$$

where z is introduced as a helper variable. In the second strategy, the attacker simply guesses all the Legendre symbols of the scheme. For each guess, they solve the corresponding system of equations and then determine if the guesses are correct using the obtained solution. The complexity of this attack hence increases by factor of around 2 for each Legendre symbol, since guessing it correctly has a probability of around 1/2.

Regarding this second approach, since the S-boxes in *Grendel* (after guessing all Legendre symbols) are of low degree, it is not strictly necessary to introduce intermediate variables in each round in order to reduce the degree growth. Indeed, the attacker can avoid adding any intermediate variables and instead solve a higher-degree system of equations, which we focus on in the next section. As we are going to show, this approach can outperform the original analysis given in [Sze21], and it can even break the full scheme including the security margin.

Our Attack Strategy. Based on the strategy just presented, we are going to show a preimage attack on a sponge hash function instantiated with full *Grendel*. Let κ be the security level, and let p be the prime that defines the field. We limit ourselves to focus on the case in which $p \geq 2^\kappa$. As we have seen in Section 2.1.2, this implies that $r \geq 1$ is allowed, and that the hash function can output a single element from $\mathbb{F}_p$.

Hence, given $h \in \mathbb{F}_p$, we are looking for a preimage. If $r \geq 2$, we first fix $r - 1$ input elements. In contrast to the analysis given in the original paper, we will not introduce any intermediate variables. Instead, we will guess all Legendre symbols and work with polynomials of degree d^R, where R is the number of attacked rounds. In short, the attack for an R-round construction consists of three steps, namely

(1) guessing all Legendre symbols,

(2) solving the resulting univariate equation to find a preimage, and

(3) determining if the solution is a valid one.

For the second step, we use a root-finding approach to solve the single univariate polynomial. We point out that the capacity c does not play any role in the attack. A pseudo code is given in Algorithm 1. The particular steps will now be explained in detail.

5.2 Details and Cost of the Attack

In the following, we provide the details of the attack, and we estimate its cost.

Algorithm 1: Finding a preimage for a sponge hash function instantiated with R-round *Grendel* over $\mathbb{F}_p^n$. For simplicity and w.l.o.g., $IV = 0 \mid\mid \cdots \mid\mid 0 \in \mathbb{F}_p^n$.

```
Data: Hash output $h \in \mathbb{F}_p$.
Result: Preimage $x^\star \mid\mid v \in \mathbb{F}_p^r$.
Generate random $v \in \mathbb{F}_p^{r-1}$.
$\mathfrak{L} \leftarrow \emptyset$.
while $|\mathfrak{L}| < 2^{n(R-1)+1}$ do
    Guess all the Legendre symbols $\lambda \notin \mathfrak{L}$ for $R$-round Grendel (only 1 or $-1$).
    $\mathfrak{L} \leftarrow \mathfrak{L} \cup \lambda$.
    Symbolically compute $F(x) = [Grendel(x \mid\mid v \mid\mid 0 \mid\mid \cdots \mid\mid 0)]_0 \in \mathbb{F}_p[x]$ using $\lambda$.
    Find a root $x^\star$ of $F(x) - h$.
    if $Grendel(x^\star \mid\mid v \mid\mid 0 \mid\mid \cdots \mid\mid 0)$ satisfies all guessed Legendre symbols $\lambda$ then
        return $x^\star \mid\mid v$.
return No solution found, try with different $v$.
```

5.2.1 Guessing Legendre Symbols

Correctly guessing a single Legendre symbol has a probability of around $1/2$ when ignoring the edge case of the symbol being 0. More precisely, the probability of ± 1 is $(p-1)/(2p) \approx 1/2$, while the probability of 0 is $1/p$. Hence, the probability that l symbols are different from zero is $(1-(1/p))^l$, which is greater than 99.99% even for a large number of rounds if $p \approx 2^{32}$. However, we still conservatively acknowledge this by adding an additional factor of 2 to the attack complexity.

From now, we assume to be in the case in which the Legendre symbol can only output 1 or -1. The probability of correctly guessing all l Legendre symbols is 2^{-l}. Hence, if we target a success probability of 95%, we need x tests such that

$$1 - (1 - 2^{-l})^x \geq 0.95,$$

which is fulfilled when choosing $x \geq 3 \cdot 2^l$ and added as an additional cost to the final attack. In our factorization attack, the number l is equal to

$$l = n \cdot R - (n-1) = n(R-1) + 1,$$

where $t-1$ Legendre symbols in the first round can also be computed deterministically, since there is no linear layer before the first application of the S-boxes. Thus, to take care of the guessing part, we add a cost of $6 \cdot 2^l = 6 \cdot 2^{t \cdot (R-1)+1}$. Note that this specifically includes repeating the attack until the correct set of Legendre symbols is found.[2]

We finally point out that this guessing approach also works for other variations, such as the modulo-2 operation presented in Section 4.2.

5.2.2 Finding the Preimage

We focus on the case where the number of hash output elements is 1, which occurs to be the case if $p \geq 2^\kappa$. Note that since the final hash value consists of a single element, we also have only one squeezing step in the sponge. Guessing all Legendre symbols, we only have a single unknown (the input variable) and a single equation of degree at most d^R in the end. W.l.o.g., we assume $IV = 0 \mid\mid \cdots \mid\mid 0 \in \mathbb{F}_p^n$ for the sponge initialization vector. The polynomial equation is then given by

$$F(x) = [Grendel(x \mid\mid v \mid\mid 0 \mid\mid \cdots \mid\mid 0)]_0 - h \quad \in \mathbb{F}_p[x], \tag{9}$$

[2] We note that an attacker can abort the respective trial as soon as their guess does not match the correct symbol. However, the cost advantage is negligible compared to finding the roots.

where x is the unknown variable, $v \in \mathbb{F}_p^{r-1}$ is chosen randomly and fixed, and $h \in \mathbb{F}_p$ is the hash output we want to find a preimage for. Our equation system hence consists of only one univariate equation, and we can apply a root-finding algorithm to this equation in order to solve for x. We point out that a similar approach was also shown in [RAS20] when attacking GMiMCHash, a sponge hash function built using the symmetric primitive GMiMC [AGP+19].

5.2.3 Complexity of Finding the Roots and Verifying the Solution

We apply a root-finding algorithm to the polynomial F given in Eq. (9). The complexity of this step is estimated in Eq. (2), where in our case $D = d^R$. In the first round, $n-1$ Legendre symbols can be computed deterministically. Moreover, after finding the roots, we need to practically verify whether the given solution is a valid one. For this purpose, we take the solution found and we determine if the given instance is a valid one (i.e., if all computed Legendre symbols match the guesses for this trial). Note that we can abort the verification step as soon as we find an inconsistency between the computed symbol using our solution and the guess, which then makes the trial invalid. Considering that we have to compute only the first Legendre symbol in each instance with a probability of 50%, only the first two with a probability of 25%, etc., the *average* number of Legendre symbols we expect to compute for each trial before finding an inconsistency is well approximated by[3]

$$\sum_{i\geq 1} \frac{i}{2^i} = \frac{2^{-1}}{(1-2^{-1})^2} = 2.$$

The complexity of computing a Legendre symbol [BZ10] is estimated as

$$\mathcal{O}\left(u \cdot (\log_2(u))^2 \cdot \log_2(\log_2(u))\right), \tag{10}$$

for $u = \log_2(p)$ and a small hidden constant. Consequently, this instance can be attacked if

$$(C_\text{R} + 2 \cdot C_\text{V}) \cdot \left(6 \cdot 2^{n(R-1)+1}\right) \leq 2^\kappa, \tag{11}$$

where we assume the $n-1$ Legendre symbols of the fixed part in the first round to be known, and where

$$\begin{aligned} C_\text{R} &= D \cdot (\log_2(D))^2 \cdot \log_2(D \cdot p) \cdot (1 + 63.43 \log_2(\log_2(D))), \\ C_\text{V} &= u \cdot (\log_2(u))^2 \cdot \log_2(\log_2(u)) \end{aligned}$$

from Eq. (2) and Eq. (10), and $u = \log_2(p)$.

The complexity of the root-finding algorithm does not change when increasing n. Indeed, only the number of Legendre symbols to be guessed is larger. As shown in Table 1 for $p \approx 2^{256}$ and $\kappa = 128$, this attack breaks more rounds than the approaches given in [Sze21, Table 1]. Similar attacks can be set up for other values of p and κ.

5.2.4 Practical Verification

We practically implemented and verified the preimage attack on *Grendel* in Sage. As expected, we could observe that the probability of having a zero at the output of the Legendre symbol is negligibly low, especially when using larger primes and with the low number of Legendre symbols involved. Moreover, a solution is found in around half of the attempts (i.e, we may have to choose different random values $v \in \mathbb{F}_p^{r-1}$ for some preimages), which results in a high success probability when repeating the attack 6 times, as suggested by our complexity estimate.

[3] We recall that $\sum_{i\geq 1} i \cdot x^i = \frac{x}{(1-x)^2}$ if $|x| < 1$.

Table 1: Number of attacked rounds of *Grendel* when considering the univariate root-finding approach. Bold numbers indicate that the attack breaks the full-round versions of *Grendel* considering the round numbers proposed in the original paper [Sze21]. In this table, $\kappa = 128$ and $p \approx 2^{256}$.

Instance (d, n)	R [Sze21]	Attacked R Eq. (11)
$(2,3)$	28	25
$(2,4)$	21	20
$(2,8)$	11	**12**
$(2,12)$	7	**8**
$(3,3)$	22	21
$(3,4)$	16	**18**
$(3,8)$	8	**11**
$(3,12)$	6	**8**
$(5,3)$	16	**18**
$(5,4)$	12	**16**
$(5,8)$	6	**10**
$(5,12)$	4	**7**

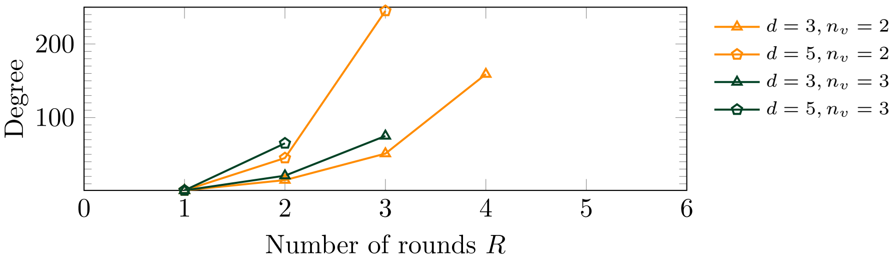

Figure 7: The values for D_{reg} in practice, where $d \in \{3, 5\}$, $n_v \in \{2, 3\}$, and $n = n_v + 1$. We did not encounter any differences for $n > n_v + 1$ in these tests.

5.3 Restoring Security and Final Considerations

Avoiding the attack shown in the previous section is not sufficient in order to argue security. Indeed, if $p < 2^\kappa$, the number of output elements has to be greater than 1. Consequently, finding a preimage is done with an equation system of $n_e \geq 2$ equations in $n_v = n_e$ variables. While the previous approach is not applicable any more, a Gröbner basis approach might.

5.3.1 Gröbner Basis Attack

Here we focus on the complexity of computing the Gröbner basis in order to find the minimum number of rounds of *Grendel* which avoids this attack. Note that from a designer's point of view, it is sufficient to argue that the complexity of computing the basis is prohibitively high.

In contrast to the original analysis given in [Sze21], here we evaluate the complexity of building a Gröbner basis without introducing any intermediate variables. In other words, we do the same as in the root-finding approach before, with the exception of having more than a single input variable and more than a single output equation.

For our purpose, assume an equation system with n_e equations in n_v unknown variables. From Section 2.3.2 we have an estimated degree of regularity of $1 + n_e \cdot (d^R - 1)$. Our practical tests regarding the actual degrees reached in the computation are given in Fig. 7,

Table 2: Adjusted round numbers for *Grendel* for $\kappa = 128$ and $p \approx 2^{256}$ considering our new attacks. We also list the recommended number of rounds from the original paper [Sze21], together with the incremental percentage.

Instance (d, n)	R from [Sze21] (with security margin)	New minimum R (with security margin)
$(2, 3)$	28	**34** (+21.43%)
$(2, 4)$	21	**28** (+33.33%)
$(2, 8)$	11	**16** (+45.45%)
$(2, 12)$	7	**12** (+71.43%)
$(3, 3)$	22	**30** (+36.36%)
$(3, 4)$	16	**25** (+56.25%)
$(3, 8)$	8	**15** (+87.50%)
$(3, 12)$	6	**11** (+83.33%)
$(5, 3)$	16	**26** (+62.50%)
$(5, 4)$	12	**22** (+83.33%)
$(5, 8)$	6	**14** (+133.33%)
$(5, 12)$	4	**11** (+175.00%)

and they suggest that this bound is reasonable, with the actual degree being

$$D_{\text{est}} = n_v \cdot d^R - d \cdot (n_v - 1)$$

in our tests. Using this estimation, we obtain

$$6 \cdot 2^{n(R-1)+n_v} \cdot \binom{D_{\text{est}} + n_v}{n_v}^2 = 6 \cdot 2^{n(R-1)+n_v} \cdot \binom{n_v \cdot (d^R + 1) - d \cdot (n_v - 1)}{n_v}^2 \tag{12}$$

for the cost of the Gröbner basis computation, including the cost of guessing the Legendre symbols (n_v Legendre symbols are known in the first round). Note that we set $\omega = 2$ (optimistic from an attacker's point of view). We emphasize that for some instances (especially with increasing n), this Gröbner basis strategy is more efficient than the strategies proposed in the original paper.

5.3.2 Security Against our Attacks

In order to restore the security of *Grendel* with respect to the attacks presented in this section, the recommended number of rounds must be adjusted for some parameters. For this purpose, we adopt the method used in the original paper [Sze21]. In particular, we first evaluate the security of each known attack (including our new ones) and we then ensure that the estimated complexity is larger than or equal to $2^{1.25\kappa}$ for a security of κ bits, where 1.25 is chosen as the security margin. Further, we use only the complexity of the Gröbner basis computation as the estimate for the multivariate attack (i.e., the complexity given in Eq. (12)), as is also done in [Sze21].

Following our attacks, in addition to the inequalities given in [Sze21, Table 1], the number of rounds of each *Grendel* instance must satisfy

$$R \geq \max \left\{ \left\lceil \frac{2.5 \cdot \kappa}{\log_2(p) - \log_2(d) - 1} \right\rceil, R', R^{''} \right\},$$

where $R' \geq 1$ is the smallest positive integer that satisfies Eq. (1), while $R^{''} \geq 1$ is the smallest positive integer that satisfies

$$2^{n(R''-1)+1} \cdot \max\Bigg\{ 2^{n_v-t+1} \cdot \binom{n_v \cdot (d^{R''}+1) - d \cdot (n_v - 1)}{n_v}^2 ;\\ d^{R''} \cdot \left(\log_2\left(d^{R''}\right)\right)^2 \cdot \log_2\left(d^{R''} \cdot p\right) \cdot \left(1 + 63.43 \cdot \log_2\left(\log_2\left(d^{R''}\right)\right)\right) \Bigg\} \geq 2^{1.25\kappa}$$

for a security level of κ bits, where n_v denotes the number of elements in the hash output. The updated round numbers for various instances are given in Table 2, where we also list the recommended number of rounds from the original paper. Note that we omit the factor 6 in Eq. (11) and Eq. (12) (we assume an attacker finds a solution immediately), we also omit the complexity of verifying a solution, and we use $\omega = 2$ in Eq. (3) (pessimistic from the designer's point of view).

5.3.3 Attack on *Grendel* instantiated with $d = 1$

Finally, even though $d \geq 2$ in *Grendel*, we study the case $d = 1$ from an algebraic point of view (we recall that *Grendel* would be invertible for $d = 1$ if $p = 1 \mod 4$). Indeed, the only nonlinear component in the description would then be the Legendre symbol. If we guess all of them, as we did before, the construction is affine, and for a single element in the hash output the challenge would then be to solve a single affine equation. Since the cost of this would be negligible, the modified construction can be attacked if

$$6 \cdot 2^{n(R-1)+1} < 2^\kappa \qquad \Longrightarrow \qquad R < \frac{\kappa - 2 - \log_2(3)}{n} + 1.$$

The situation is only slightly different when considering $r \geq 2$ and multiple hash outputs. Indeed, for n_e outputs we have n_e linear equations in $n_v = r$ input variables. The solving complexity is then in $\mathcal{O}({n_e}^\omega)$, and for an attack ${n_e}^\omega \cdot 6 \cdot 2^{n(R-1)+n_v} < 2^\kappa$, where pessimistically $\omega = 3$.

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.

[ACG+19] Martin R. Albrecht, Carlos Cid, Lorenzo Grassi, Dmitry Khovratovich, Reinhard Lüftenegger, Christian Rechberger, and Markus Schofnegger. Algebraic Cryptanalysis of STARK-Friendly Designs: Application to MARVELlous and MiMC. In *ASIACRYPT 2019*, volume 11923 of *LNCS*, pages 371–397, 2019.

[AD18] Tomer Ashur and Siemen Dhooghe. MARVELlous: a STARK-Friendly Family of Cryptographic Primitives. *IACR Cryptol. ePrint Arch.*, page 1098, 2018.

[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 (2)*, volume 11736 of *LNCS*, pages 151–171. Springer, 2019.

[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.

[BBS86] Lenore Blum, Manuel Blum, and Mike Shub. A Simple Unpredictable Pseudo-Random Number Generator. *SIAM J. Comput.*, 15(2):364–383, 1986.

[BBUV20] Ward Beullens, Tim Beyne, Aleksei Udovenko, and Giuseppe Vitto. Cryptanalysis of the Legendre PRF and Generalizations. *IACR Trans. Symmetric Cryptol.*, 2020(1):313–330, 2020.

[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.

[BdSG20] Ward Beullens and Cyprien Delpech de Saint Guilhem. LegRoast: Efficient Post-quantum Signatures from the Legendre PRF. In *Post-Quantum Cryptography - PQCrypto 2020*, volume 12100 of *LNCS*, pages 130–150, 2020.

[BFP12] Luk Bettale, Jean-Charles Faugère, and Ludovic Perret. Solving polynomial systems over finite fields: improved analysis of the hybrid approach. In *ISSAC*, pages 67–74. ACM, 2012.

[BFSY05] M Bardet, JC Faugere, B Salvy, and BY Yang. Asymptotic behaviour of the index of regularity of quadratic semi-regular polynomial systems. In *The Effective Methods in Algebraic Geometry Conference (MEGA)*, pages 1–14, 2005.

[BS90] Eli Biham and Adi Shamir. Differential Cryptanalysis of DES-like Cryptosystems. In *CRYPTO 1990*, volume 537 of *LNCS*, pages 2–21, 1990.

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

[BSV07] Thomas Baignères, Jacques Stern, and Serge Vaudenay. Linear Cryptanalysis of Non Binary Ciphers. In *Selected Areas in Cryptography – SAC 2007*, volume 4876 of *LNCS*, pages 184–211. Springer, 2007.

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

[BZ10] Richard P. Brent and Paul Zimmermann. An $O(M(n)\log n)$ Algorithm for the Jacobi Symbol. In *ANTS*, volume 6197 of *LNCS*, pages 83–95. Springer, 2010.

[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/.

[DGGK21] Christoph Dobraunig, Lorenzo Grassi, Anna Guinet, and Daniël Kuijsters. Ciminion: Symmetric Encryption Based on Toffoli-Gates over Large Finite Fields. In *EUROCRYPT (2)*, volume 12697 of *LNCS*, pages 3–34. Springer, 2021.

[DGH+21] Christoph Dobraunig, Lorenzo Grassi, Lukas Helminger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Pasta: A Case for Hybrid Homomorphic Encryption. *IACR Cryptol. ePrint Arch.*, page 731, 2021.

[DGV92] Joan Daemen, René Govaerts, and Joos Vandewalle. A Hardware Design Model for Cryptographic Algorithms. In *ESORICS 1992*, volume 648 of *LNCS*, pages 419–434, 1992.

[DHS98] Cunsheng Ding, Tor Helleseth, and Weijuan Shan. On the Linear Complexity of Legendre Sequences. *IEEE Trans. Inf. Theory*, 44(3):1276–1278, 1998.

[DR01] Joan Daemen and Vincent Rijmen. The Wide Trail Design Strategy. In *Cryptography and Coding, 8th IMA International Conference*, volume 2260 of *LNCS*, pages 222–238, 2001.

[DR02] Joan Daemen and Vincent Rijmen. Security of a wide trail design. In *Progress in Cryptology - INDOCRYPT 2002*, volume 2551 of *LNCS*, pages 1–11, 2002.

[EGL+20] Maria Eichlseder, Lorenzo Grassi, Reinhard Lüftenegger, Morten Øygarden, Christian Rechberger, Markus Schofnegger, and Qingju Wang. An Algebraic Attack on Ciphers with Low-Degree Round Functions: Application to Full MiMC. In *ASIACRYPT 2020*, volume 12491 of *LNCS*, pages 477–506, 2020.

[Fau02] Jean Charles Faugère. A New Efficient Algorithm for Computing GröBner Bases without Reduction to Zero (F5). In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, ISSAC '02, page 75–83, New York, NY, USA, 2002. Association for Computing Machinery.

[FGLM93] Jean-Charles Faugère, Patrizia M. Gianni, Daniel Lazard, and Teo Mora. Efficient computation of zero-dimensional gröbner bases by change of ordering. *J. Symb. Comput.*, 16(4):329–344, 1993.

[FS21] Paul Frixons and André Schrottenloher. Quantum Security of the Legendre PRF. *MathCrypto*, 1(1):1–18, 2021.

[Gen07] Giulio Genovese. Improving the algorithms of berlekamp and niederreiter for factoring polynomials over finite fields. *J. Symb. Comput.*, 42(1-2):159–177, 2007.

[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 Symposium*, pages 519–535. USENIX Association, 2021.

[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.

[GMS14] Katalin Gyarmati, Christian Mauduit, and András Sárközy. The cross-correlation measure for families of binary sequences. In *Applied Algebra and Number Theory*, number Theory, pages 126–143. 2014.

[GRR+16] Lorenzo Grassi, Christian Rechberger, Dragos Rotaru, Peter Scholl, and Nigel P. Smart. MPC-Friendly Symmetric Key Primitives. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 430–443. ACM, 2016.

[GU82] L. Guerra and E. Ughi. On the distribution of Legendre symbols in Galois fields. *Discrete Mathematics*, 42(2-3):197–208, 1982.

[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.

[Kho19] Dmitry Khovratovich. Key recovery attacks on the Legendre PRFs within the birthday bound. Cryptology ePrint Archive, Report 2019/862, 2019. https://ia.cr/2019/862.

[KJJ99] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential Power Analysis. In *Advances in Cryptology - CRYPTO 1999*, volume 1666 of *LNCS*, pages 388–397, 1999.

[KKK20] Novak Kaluđerović, Thorsten Kleinjung, and Dusan Kostic. Improved key recovery on the Legendre PRF. *ANTS XIV: Proceedings of the Fourteenth Algorithmic Number Theory Symposium*, 4(1):267–282, 2020.

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

[Koc96] Paul C. Kocher. Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems. In *Advances in Cryptology - CRYPTO 1996*, volume 1109 of *LNCS*, pages 104–113, 1996.

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

[Leh54] D. H. Lehmer. Random number generation on the BRL highspeed computing machines. *Math. Rev.*, 15:559, 1954.

[LHT13] Nian Li, Tor Helleseth, and Xiaohu Tang. Further results on a class of permutation polynomials over finite fields. *Finite Fields Their Appl.*, 22:16–23, 2013.

[Mat93] Mitsuru Matsui. Linear Cryptanalysis Method for DES Cipher. In *Advances in Cryptology - 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.

[MS97] Christian Mauduit and András Sárközy. On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol. *Acta Arithmetica*, 82(4):365–377, 1997.

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

[RAS20] Arnab Roy, Elena Andreeva, and Jan Ferdinand Sauer. Interpolation cryptanalysis of unbalanced feistel networks with low degree round functions. In *SAC*, volume 12804 of *LNCS*, pages 273–300. Springer, 2020.

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

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

[SHB21] István András Seres, Máté Horváth, and Péter Burcsi. The Legendre Pseudorandom Function as a Multivariate Quadratic Cryptosystem: Security and Applications. Cryptology ePrint Archive, Report 2021/182, 2021. https://ia.cr/2021/182.

[SS21] Jan Ferdinand Sauer and Alan Szepieniec. SoK: Gröbner Basis Algorithms for Arithmetization Oriented Ciphers. Cryptology ePrint Archive, Report 2021/870, 2021. https://ia.cr/2021/870.

[Sze21] Alan Szepieniec. On the Use of the Legendre Symbol in Symmetric Cipher Design. *IACR Cryptol. ePrint Arch.*, page 984, 2021.

[Tra00] Quoc-Nam Tran. A fast algorithm for gröbner basis conversion and its applications. *J. Symb. Comput.*, 30(4):451–467, 2000.

[TZLH15] Ziran Tu, Xiangyong Zeng, Chunlei Li, and Tor Helleseth. Permutation polynomials of the form $(x^{p^m} - x + \delta)^s + L(x)$ over the finite field $\mathbb{F}_{p^{2m}}$ of odd characteristic. *Finite Fields Their Appl.*, 34:20–35, 2015.

[Tó07] Viktória Tóth. Collision and avalanche effect in families of pseudorandom binary sequences. *Periodica Mathematica Hungarica*, 55(2):185–196, 2007.

[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 Linear Cryptanalysis

Preliminaries. Linear cryptanalysis [Mat93, BSV07] exploits large correlations (in absolute value) between linear combinations of input elements and linear combinations of output elements of a permutation. While differential cryptanalysis is a chosen (plaintext, ciphertext) approach, linear cryptanalysis is a known (plaintext, ciphertext) one.

For each $x, y \in \mathbb{F}_p^n$, let $\langle x, y\rangle = \sum_{j=0}^{t-1} x_j \cdot y_j$. Let $\Psi, \Phi \in \mathbb{F}_p^n$ be the input and the output masks. The bias $\varepsilon_F(\Psi, \Phi)$ of a function F over $\mathbb{F}_p^n$ is defined as

$$\varepsilon_F(\Psi, \Phi, \Omega) = \left| \text{Prob}\left(\langle \Psi, x\rangle + \langle \Phi, \mathcal{F}(x)\rangle = \Omega\right) - \frac{1}{p} \right|$$

for $x \in \mathbb{F}_p^n$ and a constant $\Omega \in \mathbb{F}_p^n$. In order to make linear cryptanalysis successful, the attacker has to choose Ψ and Φ in order to make $\varepsilon_F(\Psi, \Phi)$ as large as possible, since such a linear approximation can then be used as a distinguishing property.

Remark: In the following, we limit ourselves to consider the case $\Omega = 0$. However, the analysis can easily be generalized for the case $\Omega \neq 0$ (which corresponds to an affine approximation).

A.1 Linear Property of $x \mapsto x^d \cdot (L_p(x) + \alpha)$

Lemma 11. *Let $p \geq 3$ be a prime and let $F(x) = x \cdot (L_p(x) + \alpha)$. For each $\Psi, \Phi \in \mathbb{F}_p \setminus \{0\}$,*

$$|\{x \in \mathbb{F}_p \mid \Psi \cdot x = \Phi \cdot F(x)\}| \leq \begin{cases} 1 & \textit{if } \Psi \neq \Phi \cdot (\alpha \pm 1), \\ (p+1)/2 & \textit{otherwise} . \end{cases}$$

Proof. The equality $\Psi \cdot x = \Phi \cdot x \cdot (L_p(x) + \alpha)$ is satisfied by $x = 0$. If $x \neq 0$, it becomes $L_p(x) = -\frac{\Psi}{\Phi} + \alpha$, which

- does not have any solution if $\Psi \neq \Phi \cdot (\alpha \pm 1)$ and $\Psi \neq \Phi \cdot \alpha$,
- admits 0 as solution if $\Psi = \Phi \cdot \alpha$ (note: solution already included),
- admits $(p-1)/2$ nonzero solutions if $\Psi = \Phi \cdot (\alpha \pm 1)$. □

Lemma 12. *Let $p \geq 3$ be a prime and let $d \geq 3$ be such that* $\gcd(d, p-1) = 1$. *Let $F(x) = x^d \cdot (L_p(x) + \alpha)$. For each $\Psi, \Phi \in \mathbb{F}_p \setminus \{0\}$,*

$$|\{x \in \mathbb{F}_p \mid \Psi \cdot x = \Phi \cdot F(x)\}| \leq 2d - 1.$$

Proof. The equality $\Psi \cdot x = \Phi \cdot x^d \cdot (L_p(x) + \alpha)$ is satisfied by $x = 0$. If $x \neq 0$, it becomes $\Phi \cdot x^{d-1} \cdot (L_p(x) + \alpha) - \Psi = 0$. For each value of $L_p(x) \in \{-1, 1\}$, this equation admits at most $d - 1$ solutions, for a total of $2 \cdot (d-1) + 1 = 2d - 1$. □

A.2 Linear Property of $x \mapsto \alpha^{(x \mod 2)} \cdot x^2$

Lemma 13. *Let $p \geq 3$ be a prime number such that $p = 3 \mod 4$. Let $F(x) = \alpha^{(x \mod 2)} \cdot x^2$, where $L_p(\alpha) = -1$. For each $\Psi, \Phi \in \mathbb{F}_p \setminus \{0\}$,*

$$|\{x \in \mathbb{F}_p \mid \Psi \cdot x = \Phi \cdot F(x)\}| \leq 3.$$

Proof. Obviously, 0 is a solution of $\Psi \cdot x = \Phi \cdot \alpha^{(x \mod 2)} \cdot x^2$. If $x \neq 0$, the solutions of this equation can be $x \in \left\{\frac{\Psi}{\Phi}, \frac{\alpha \cdot \Psi}{\Phi}\right\}$. □

A.3 Linear Property of $x \mapsto (-1)^{x^2} \cdot x^d$

Lemma 14. *Let $p \geq 3$ be a prime number, and let $d \geq 1$ be such that* $\gcd(d, p-1) = 1$. *Let $F(x) = (-1)^{x^2} \cdot x^d$. For each $\Psi, \Phi \in \mathbb{F}_p \setminus \{0\}$,*

$$|\{x \in \mathbb{F}_p \mid \Psi \cdot x = \Phi \cdot F(x)\}| \leq \begin{cases} 1 & \textit{if } d = 1 \textit{ and } \Psi \neq \pm\Phi, \\ (p+1)/2 & \textit{if } d = 1 \textit{ and } \Psi = \pm\Phi, \\ 2d - 1 & \textit{if } d \geq 3. \end{cases}$$

Proof. Obviously, 0 is a solution of $\Psi \cdot x = \Phi \cdot (-1)^{x^2} \cdot x^d$. For $x \neq 0$,

- if $d = 1$, then the equation reduces to $\Psi = \Phi \cdot (-1)^{x^2}$, which does not admit any solution if $\Psi \neq \pm\Phi$, and which admits at most $(p-1)/2$ solutions otherwise,
- if $d \geq 3$, then the equation admits at most $d - 1$ solutions for each value of $(-1)^{x^2} \in \{-1, 1\}$.

□