The Legendre Pseudorandom Function as a Multivariate Quadratic Cryptosystem: Security and Applications

István András Seres[1], Máté Horváth[2], and Péter Burcsi[1]

[1] Eötvös Loránd University, Faculty of Informatics, 3in Research Group
[2] Budapest University of Technology and Economics, CrySyS Lab

Abstract. Sequences of consecutive Legendre and Jacobi symbols as pseudorandom bit generators were proposed for cryptographic use in 1988. Since then, they were mostly forgotten in the applications. However, recently revived interest has been shown towards pseudorandom functions (PRF) based on the Legendre and power residue symbols, due to their efficiency in the multi-party setting and their conjectured post-quantum security. The lack of provable security results hinders the deployment of PRFs based on power residue symbols. On the other hand, the security of these PRFs do not seem to be related to standard cryptographic assumptions, e.g. discrete logarithm or factoring.
In this work, we show that key-recovery attacks against the Legendre PRF are equivalent to solving a specific family of multivariate quadratic (MQ) equation system over a finite prime field. This new perspective sheds some light on the complexity of key-recovery attacks against the Legendre PRF and allows us to take the first steps in settling the provable security of the Legendre PRF and other variants. We do this by conducting extensive algebraic cryptanalysis on the resulting MQ instance. We show that the currently best-known techniques and attacks fall short in solving these sparse quadratic equation systems. Another benefit of viewing the Legendre PRF as an MQ instance is that it facilitates new applications, such as verifiable random function or oblivious (programmable) pseudorandom function. These new applications can be used in cryptographic protocols, such as state-of-the-art proof-of-stake consensus algorithms or private set intersection protocols.

Keywords: Pseudorandom functions · Multivariate quadratic cryptography · Post-quantum cryptography · MPC-friendly primitives.

1 Introduction

Zero-knowledge proofs (ZKP) and secure multi-party computation (MPC) protocols are eating the crypto-world. These advanced cryptographic tools are applied and deployed in countless applications, for instance, in privacy-preserving cryptocurrencies, threshold cryptography and secure instant-messaging. The widespread

adoption of ZKPs and MPC protocols necessitates novel symmetric-key primitives [30]. Traditional symmetric-key primitives, like AES or SHA-3, cause significant overhead in ZKPs or MPC due to their immense multiplicative complexity.

Therefore, recently, revived interest has been shown towards algebraic symmetric key primitives with low multiplicative depth [30]. Lately, several novel algebraic MACs [21,13], hash functions [1,29] or algebraic pseudorandom functions [17] have been proposed for cryptographic use. New algebraic constructions with low multiplicative complexity are especially attractive due to their distinguished efficiency properties in ZKPs or MPC protocols. However, this new algebraic design paradigm possibly opens up new venues for attacks [4]. The cryptanalysis of these new symmetric-key primitives is an active research field with notable published works. For instance, Albrecht et al. conducted an algebraic cryptanalysis of MARVELlous [5] and MiMC hash functions [2], while Li and Preneel refined interpolation attacks on low algebraic degree cryptosystems [44]. One of the most promising cryptosystems for use in ZKPs and MPC protocols is a pseudorandom function (PRF) that is based on quadratic and power residue symbols. Recall that if p is a prime, the Legendre symbol $\left(\frac{a}{p}\right)$ is 1 if a is a square modulo p and -1 otherwise (the symbol of zero modulo p is 0 by convention). In this work, we focus on the cryptographic security of a PRF family, called the Legendre PRF, and its extensions that are derived from the evaluation of the Legendre symbol.

There exists vast mathematics literature asserting that Legendre and power residue symbols are particularly well suited to be applied in pseudorandom functions since they exhibit high pseudorandomness. One of the first results is due to Pólya and Vinogradov [56]. They assert that character sums behave like independent fair coin tosses, i.e. $\sum_{a=M+1}^{M+N} \left(\frac{a}{p}\right) \leq \sqrt{p} \log p$. In the case of Legendre symbols, Peralta extended this result by showing that any n-grams of Legendre symbols are asymptotically equally distributed [51]. Mauduit and Sárközy introduced several metrics to measure the pseudorandomness of binary sequences and argued that "Legendre symbol sequences are the most natural candidate for pseudorandomness" [46]. Ding et al. confirmed the high linear complexity of Legendre-symbol sequences [20]. Tóth and Gyarmati et al. introduced new pseudorandomness measures (avalanche effect and cross-correlation) and asserted high values of those in Legendre symbol sequences [55,32].

Related work. In spite of the above results, surprisingly, the security guarantees of the Legendre PRF from a cryptographic standpoint are poorly understood. The quantum case is settled whenever a quantum oracle is available for the attacker as polynomial quantum algorithms are known to recover the key of a Legendre PRF [16,53]. However, if the oracle can only be queried classically, then no efficient quantum algorithm is known. In concurrent and independent work, Frixons and Schrottenloher [26] 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,

strengthening the security intuition. On the other hand, in the classical setting, only exponential key-recovery algorithms are known due to Khovratovich [39], Beullens et al. [7] and Kaluderovic et al. [37]. One might ask, whether there could be sub-exponential key-recovery attacks on the Legendre PRF. Damgård in 1988 proposed as an open problem to assess the security and complexity of predicting Legendre or Jacobi symbols. He was contemplating on reducing well-known number-theoretic assumptions to the problem of predicting Legendre or Jacobi symbol sequences [17]. This approach in the last decades has been eluding researchers. Thus, in this paper, we show connections of the Legendre and Jacobi sequences to a different branch of cryptography, namely, multivariate quadratic cryptography. This study is essential to establish the security of various cryptographic applications derived from the Legendre PRF, e.g. the digital signature scheme by Beullens et al. [8].

Our contributions. In this work, we make the following contributions.

Legendre PRF as an MQ instance. We show that key-recovery attacks against the Legendre PRF are equivalent to solving a specific family of sparse multivariate quadratic equation system over a finite field. Moreover, the weak unpredictability of the PRF is reducible to the decidability of the aforementioned equation system.

Algebraic cryptanalysis. We conduct the first algebraic cryptanalysis on the MQ instance induced by the Legendre PRF. We find that the Legendre PRF is immune to interpolation, direct (Gröbner-basis) and rank attacks. We also present algebraic geometric arguments to support the complexity of finding solutions in these sparse MQ instances over a finite field.

Novel cryptographic applications of the Legendre PRF. Expressing the Legendre PRF as an MQ instance facilitates novel cryptographic applications of the PRF. Namely, we can construct efficient verifiable random functions, oblivious (programmable) pseudorandom functions from the Legendre PRF. Thanks to their efficiency, these novel extensions can be applied in several cryptographic protocols, such as state-of-the-art private set intersection (PSI) protocols.

Organisation. The rest of this paper is organised as follows. In Section 2, we provide the necessary background on Legendre symbols and related hard cryptographic problems. In Section 3, we show that key-recovery attacks against the Legendre PRF are equivalent to solving a specific MQ instance. In Section 4, we analyze the security of the MQ instance induced by the Legendre PRF. In Section 5, we describe several extensions to the Legendre PRF. Finally, we conclude our paper in Section 6 by pointing out promising future directions.

2 Preliminaries

Notations. Whenever we sample x from set S uniformly at random we write $x \in_R S$. Let p be an odd prime and secret key $K \in_R \mathbb{F}_p$. Vectors of group elements

are denoted in bold. In the following, n, m denote the number of variables and equations, respectively. Throughout this work, we will work in the multivariate polynomial ring $\mathbb{F}_p[x_1, \ldots, x_n]$ over a finite field $\mathbb{F}_p$. For the ease of exposition we use $[x]$ to denote a secret share of the value $x \in \mathbb{F}_p$.

Background on the Legendre PRF. Damgård proposed using the sequence of consecutive Legendre symbols with respect to a large prime p for "pseudorandom bit generation" [17].

Definition 1 (Sequential Legendre PRF). *Let p be a prime, depending on the security parameter λ, then let $\{a\}_K$ denote the following sequence:*

$$\{a\}_K := \left(\frac{K}{p}\right), \left(\frac{K+1}{p}\right), \ldots, \left(\frac{K+a-1}{p}\right).$$

Damgård conjectured that the sequence is pseudorandom, when starting at a secret K. Sometimes, it is easier to work with bits, rather than the original Legendre symbols themselves, therefore the Legendre PRF is defined with Boolean output (for a key- and input-space $\mathbb{F}_p$).

Definition 2 (Legendre pseudorandom function). *The function $L_K(x)$ is defined by mapping the corresponding Legendre-symbol to the set $\{0,1\}$, i.e.*

$$L_K(x) = \left\lfloor \frac{1}{2}\left(1 - \left(\frac{K+x}{p}\right)\right)\right\rfloor.$$

Assumptions. Grassi et al. formulated the following problem that underpins the security of the Legendre PRF [30].

Definition 3 (Shifted Legendre Symbol (SLS) Problem). *Let K be uniformly sampled from $\mathbb{F}_p$, and define $\mathcal{O}_{Leg}$ to be an oracle that takes $x \in \mathbb{F}_p$ and outputs $\left(\frac{K+x}{p}\right)$. Then the Shifted Legendre Symbol (SLS) problem is to find K given oracle access to $\mathcal{O}_{Leg}$ with non-negligible probability.*

It is conjectured that no classical adversary running in sub-exponential time could recover the hidden shift K. One might also consider generalisations of the problem, such as changing the linear polynomial to a secret degree-d polynomial in the Legendre symbol evaluations or changing the quadratic symbol to an rth power residue symbol. For more details, see Appendix A.

Definition 4 (Multivariate Quadratic (MQ) problem). *Given random quadratic polynomials $\mathbf{f} = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \in \mathbb{F}[x_1, \ldots, x_n]^m$, find a common zero $\mathbf{x} \in \mathbb{F}^n$ of the polynomials $f_1, \ldots, f_m$.*

We note that the MQ problem is **NP**-hard for any choice of field $\mathbb{F}$. In cryptographic applications, $\mathbb{F}$ is often $\mathbb{F}_2$ or an extension of it. However, throughout this work, we consider MQ problems over $\mathbb{F}_p$, for some large prime p. The MQ problem is one of the major candidates on which post-quantum secure cryptosystems can be based. Currently, there are no known sub-exponential algorithms to solve the MQ problem.

3 The Legendre PRF as an MQ instance

Hereby, we describe how to express the sequential Legendre PRF, cf. Definition 1, as a multivariate quadratic equation system. We remark that in a similar fashion, all the variants (higher-degree) and extensions (power-residue and Jacobi PRF) of the sequential Legendre PRF could be expressed as a suitable MQ instance. Most of our results and observations can be easily ported to those MQ instances as well. Therefore, in this work, we solely focus on the linear Legendre PRF.

3.1 The Ideal

Let us fix an arbitrary quadratic non-residue r in $\mathbb{Z}_p^*$. Furthermore, let us assume that we are given $\{a\}_K$, for $a \approx \log(p)$. Let $b_i := \left(\frac{K+i}{p}\right)$ and x_i be the corresponding unknown. We think of the unknown x_i as the square root of $K+i$ if $b_i = 1$, otherwise x_i denotes the square root of $r(K+i)$, which is a quadratic residue. Therefore, for each pair of neighboring Legendre symbols (b_i, b_{i+1}), we define a unique quadratic equation. If $b_i = b_{i+1} = 1$, then we know that $x_{i+1}^2 = K+i+1$ and $x_i^2 = K+i$, hence

$$x_{i+1}^2 - x_i^2 = 1. \tag{1}$$

If $b_i = b_{i+1} = -1$, then we have that $x_{i+1}^2 = r(K+i+1)$ and $x_i^2 = r(K+i)$, hence

$$x_{i+1}^2 - x_i^2 = r. \tag{2}$$

Finally if $b_i = 1 = -b_{i+1}$ or $b_i = -1 = -b_{i+1}$ then we obtain the following two quadratic equations:

$$x_{i+1}^2 - r x_i^2 = r, \qquad x_{i+1}^2 - r^{-1} x_i^2 = 1. \tag{3}$$

Altogether, this allows us to efficiently transform any Legendre symbol sequence into an equivalent multivariate quadratic equation system. If we have n symbols, then we obtain $m = n-1$ independent equations in n variables, hence the MQ instance is underdefined. Note, that the equation system is rather sparse.

Example 1. We consider the following example to illustrate the quadratic equation system induced by the Legendre PRF. Let $p =$ `0xffffffffffffffffffffdd` and $K =$ `0x27aaa97c746c22e12d10`. The smallest quadratic non-residue modulo p is 2. We display the MQ instance induced by the evaluation of the linear Legendre PRF, $\{5\}_K = (1, 1, -1, -1, 1)$. Each consecutive Legendre-symbol pairs define an equation. The ideal corresponding to $\{5\}_K$ has the following form:

$$\langle x_1^2 - x_0^2 - 1, x_2^2 - 2x_1^2 - 2, x_3^2 - x_2^2 - 2, x_4^2 - 2^{-1}x_3^2 - 1 \rangle$$

Let $I := \langle f_1, f_2, \ldots, f_m \rangle$ be the ideal generated by the quadratic polynomials defined by Equations 1, 2 and 3. We are interested in solving simultaneously this equation system, i.e. finding points in the variety $V(I)$. If the sequence of

Legendre-symbols is long enough, namely $\mathcal{O}(\log p)$, then there are $\mathcal{O}(1)$ solutions (only considering solutions where $x_i \in [0, \frac{p-1}{2}]$ for all i) and one of them corresponds to the secret key K of the Legendre PRF. Given our previous discussion, the following lemma is obvious.

Lemma 1. *A successful Legendre key-recovery attack is equivalent to solving the MQ system defined by the ideal I. On the other hand, the weak unpredictability of the Legendre PRF is equivalent to the decidability of the induced MQ instance over the finite prime field.*

We highlight again the extreme sparsity of the induced MQ instance. This is in contrast with most MQ public-key cryptosystems, where the MQ instance is generated uniformly at random by the signer or encryptor. Typically, a random MQ instance has many non-zero coefficients resulting in large public keys. Contrarily, in the case of the Legendre PRF, the MQ instances exhibit a very specific structure (cf. Example 1) stemming from the multiplicative group of the field $\mathbb{F}_p$. Interestingly, if a single coefficient in the Legendre MQ instance became 0, then the whole equation system suddenly would be trivially solvable by "back-substitution". The Legendre MQ instance seems to be the smallest possible, yet still secure MQ instance. In Section 4, we turn our attention to assessing the security of the MQ instance induced by the Legendre PRF.

3.2 The Gröbner-basis

To better understand the variety $V(I)$, first we describe the Gröbner basis of I. Interestingly, we can easily compute the Gröbner basis of I regardless of the size of p or the length of the Legendre sequence $\{a\}_K$.

Theorem 1. *Given a Legendre symbol sequence $\{n\}_K = (b_0, \ldots, b_{n-1})$ and its corresponding ideal $I = \langle f_1, f_2, \ldots, f_m \rangle$, where $m = n - 1$ as defined by the Equations 1, 2 and 3, its Gröbner basis with respect to the (graded) lexicographic ordering, consists of the polynomials g_i, for $i \in [0, n-2]$ such that,*

$$g_i = \begin{cases} x_i^2 - x_{n-1}^2 + (n-i), \text{if } b_{n-1} = 1 \wedge b_i = 1 \\ x_i^2 - r x_{n-1}^2 + r(n-i), \text{if } b_{n-1} = 1 \wedge b_i = -1 \\ x_i^2 - r^{-1} x_{n-1}^2 + (n-i), \text{if } b_{n-1} = -1 \wedge b_i = 1 \\ x_i^2 - x_{n-1}^2 + r(n-i), \text{if } b_{n-1} = -1 \wedge b_i = -1 \end{cases} \tag{4}$$

Specifically, $I = \langle g_0, \ldots, g_{n-2} \rangle$ and $G := (g_i)_{i=0}^{n-2}$ is a reduced Gröbner-basis.

A proof of Theorem 1 is sketched in Appendix B.2. We remark that one can view the resulting equation system as a simultaneous Pell-equation system over $\mathbb{F}_p$. Each polynomial in the Gröbner-basis is quadratic bi-variate and has $p-1$ solutions in $\mathbb{F}_p$. Put differently, seemingly no elimination ideal turns out to be helpful in finding a common zero.

In the following, we want to assess the complexity of solving the particular equation system induced by the Legendre PRF. According to [34], in order to

prove the security of the Legendre PRF, it suffices to show that the family of MQ instances $\mathbf{f}$ induced by the PRF is hard to solve. This is because then the distributions $D_1 = (\mathbf{f}, \mathbf{f}(x_0, x_1, \ldots, x_{n-1}))$ and $D_2 = (\mathbf{f}, U_m)$ are computationally indistinguishable, where U_m is a uniform distribution over $\mathbb{F}_p^m$ [34]. First, we observe that the polynomials in I lack any special internal structure, i.e. the only relations holding are the trivial ones. More formally, the multivariate quadratic polynomials define a regular ideal.

Lemma 2. *I is a regular ideal.*

A proof of Lemma 2 is sketched in Appendix B.3.

3.3 The Overdetermined Cases of Legendre PRFs

As we have seen in Section 3.2, the Legendre key-recovery attack is equivalent to solving an undertedermined MQ instance. However, when $p \equiv 3 \mod 4$ or $p \equiv 5 \mod 8$, we can decrease the complexity of solving the resulting MQ instance by adding new independent equations. Observe that in these cases, we can express the modular square root function $\mathsf{sqrt}_p : \mathbb{F}_p^* \to \mathbb{F}_p^*$ as a polynomial function:

$$\mathsf{sqrt}_p(x) : y = \begin{cases} \pm x^{\frac{p+1}{4}} \mod p, \text{ if } p \equiv 3 \mod 4 \\ \pm x(2x)^{\frac{p-5}{8}}(4x^{\frac{p-1}{4}} - 1) \mod p, \text{ if } p \equiv 5 \mod 8. \end{cases} \tag{5}$$

If $p \equiv 1 \mod 8$, it is not possible to express easily the $\mathsf{sqrt}_p(\cdot)$ function as a polynomial function, since in that case the root-finding Tonelli-Shank algorithm is a probabilistic algorithm. Nevertheless, we can obtain $\mathcal{O}(\log^2 p)$ new polynomials in the other cases, one for each quadratic term $x_i x_j$:

$$x_i x_j = \mathsf{sqrt}_p(r^{L_0(x_i)+L_0(x_j)}(K+i)(K+j)). \tag{6}$$

In a similar fashion, we can add new polynomials involving the linear terms of the unknowns for every $i \neq j$:

$$x_i = \mathsf{sqrt}_p(r^{L_0(x_i)-L_0(x_j)}(x_j^2 - r^{L_0(x_j)}(j-i))). \tag{7}$$

Note, that all polynomials in Equations 6 and 7 have almost full degree, i.e. they have degree $\approx p$. Therefore, the addition of each of those polynomials incur the inclusion of $\approx \log p$ new quadratic equations in $\approx \log p$ new variables in order to break down the almost full degree polynomials to quadratic polynomials. All in all, we end up with an equation system in n variables and $m = n+k$ equations, where $m, n \in \mathcal{O}(\log^3 p)$ and $k \approx \log^2 p$.

4 Security of the Legendre PRF as MQ instances

In this section, we evaluate the complexity of a key recovery attack on the Legendre PRF as an MQ instance. We find that direct attacks, solvers and other traditional algebraic attacks (interpolation attacks, MinRank etc.) do not improve on the state-of-the-art classical attack due to Kaluderovic et al [37].

4.1 Interpolation Attacks

Interpolation attacks aim to interpolate a cryptosystem's polynomial without knowing its secret key [35]. In a single party setting, the Legendre PRF is typically evaluated more than once for a particular key K, i.e. $\{a\}_K$ is used as a pseudo-random bit-string, where $a > 0$. In these cases, the resulting bit-string is mapped to integers, for instance, in the following way,

$$F_K(a) = \sum_{i=0}^{a-1} 2^{a-1-i}(K+i)^{\frac{p-1}{2}} \mod p \tag{8}$$

Note that $deg(F_K(a)) = \frac{p-1}{2}$, i.e. the degree of the polynomial representing the Legendre PRF has almost full degree over $\mathbb{F}_p$, that is exponential in the security parameter. The polynomial is dense (all possible monomials appear) and no coefficient is dependent on the key K. These properties make interpolation attacks infeasible as they would require at least $\frac{p-1}{2}+1$ pairs of keys and pseudo-random field elements to interpolate $F_K(a)$.

4.2 Direct Algebraic Attacks

Direct algebraic attacks, such as Gröbner-basis [11], F_5 [24], XL [14] aim to directly solve the cryptosystem's underlying MQ instance. The computational complexity of these attacks is equivalent to that of computing the Gröbner-basis [54], which in turn depends on the *degree of regularity* of the MQ instance at hand. Therefore, it is of great interest to compute the degree of regularity of an MQ cryptosystem. However, in many cases, this is not possible without actually calculating the Gröbner-basis itself. For m equations of degree at most d in n variables, the arithmetic complexity of Gröbner-basis computation are $2^{2^{\mathcal{O}(n)}}$ in general and $\mathcal{O}\left(m \cdot \binom{n+d_{reg}-1}{n}^{\omega}\right)$ in case of 0-dimensional regular systems (just like the Legendre PRF MQ instance, see Lemma 2), where $2 \leq \omega \leq 3$ is the linear algebra constant of matrix multiplication.

In the underdetermined case, we saw in Section 3.1 that we can compute efficiently the Gröbner-basis. The resulting Gröbner-basis seemingly does not facilitate direct solving of the Legendre MQ instance. In the overdetermined case, we empirically confirmed for small instances that the induced MQ instance of the Legendre PRF behaves as a random system in terms of degree of regularity, cf. Table C.1 in the Appendix. It is reasonable to expect that this similarity to random MQ instances remains as the parameters of the Legendre PRF increase.

4.3 MinRank Attacks

The MinRank attack is a powerful and ubiquitous tool in the cryptanalysis of multivariate cryptography. MinRank attacks broke numerous multivariate cryptosystems, such as the cryptanalysis of HFE due to Kipnis and Shamir [40] or the cryptanalysis of SRP encryption system [52]. In the following, we show that the Legendre PRF has high Q-rank, therefore it is immune to MinRank attacks. For the complete calculation the reader is referred to Appendix C.1.

4.4 Group Structure of the Legendre PRF MQ Instances' Solutions

In Section 3.1, it was shown, that the PRF seed lies in the intersection of multiple Pell-conics. It is well known, that the solutions of a single Pell-equation over a finite field form a cyclic Abelian-group over $\mathbb{F}_p$, cf. [18]. These groups were previously suggested for use in cryptography by Lemmermeyer as it is believed that the discrete logarithm problem is hard in these groups [43]. A single Pell conic has 0 genus. The intersection of two Pell-conics yields a nonsingular elliptic curve with genus 1. Therefore, if one wants to find every secret key K that results in a 3-long specific binary sequence produced by the Legendre PRF, e.g. $(1, -1, 1)$, then every satisfying secret key K is a rational point on a sequence-specific elliptic curve. For a concrete example on how to obtain the corresponding curve equation, see Appendix E.1.

However, if one considers longer sequences, then the resulting curve has a genus greater than 1, cf. Figure 1. This implies, that the solutions of those algebraic curves *do not have an Abelian group structure equipped with them*. The computation of the genus of the high-degree surfaces induced by the Legendre PRF in the general case can be found in Appendix C.2.

m	1	2	3	4	5	6	7	8	9	10
genus	0	1	1	5	13	33	81	193	449	1025

Fig. 1: The genus of the algebraic curves containing the solutions corresponding to a Legendre symbol sequence of length $m + 1$.

5 Extensions of the Legendre PRF

In this section, we construct various extensions of the Legendre PRF and compare them with other state-of-the-art constructions. We build verifiable random functions in Section 5.1 and oblivious (programmable) pseudorandom functions from the Legendre PRF in Sections 5.2 and 5.3.

5.1 Verifiable Random Functions from the Legendre PRF

Verifiable random functions (VRFs) are natural extensions of PRFs due to Micali, Rabin and Vadhan [47]. In a VRF, the PRF evaluator can produce a publicly verifiable short proof about the correct evaluation of the PRF $F_K(x)$ given the PRF input x, the output $F_K(x) = y$ and a public key pk, without revealing anything about the secret key K. In many applications, in addition to the efficient production of pseudorandom strings, one also needs to prove the correctness of those pseudorandom objects, e.g. proof-of-stake consensus algorithms [27].

We start off by observing that one of the main advantages of the Legendre PRF arithmetization as an MQ instance, is that it allows to model the PRF as a low-degree polynomial equation system, namely as a multivariate quadratic equation system. This low-degree arithmetization easily facilitates the construction of efficient Legendre VRFs. By contrast, if one models the Legendre PRF as a high-degree $\frac{p-1}{2}$ univariate polynomial by Euler's criterion, then it hinders applying efficient proof systems for the correct evaluation statement. More formally, the Legendre PRF evaluator wants to prove that the following binary relation $\mathcal{R} : \{0,1\}^* \times \{0,1\}^*$ holds:

$$\mathcal{R}_{PRF} = \left\{ \left(\{n\}_K, K \right) : \{n\}_K = \left(\left(\frac{K}{p} \right), \left(\frac{K+1}{p} \right), \ldots, \left(\frac{K+n-1}{p} \right) \right) \right\}, \quad (9)$$

which is equivalent to the relation:

$$\mathcal{R}^*_{PRF} = \left\{ \left(\{n\}_K, \mathbf{x} \right) : (f_1(\mathbf{x}) = 0, f_2(\mathbf{x}) = 0, \ldots, f_m(\mathbf{x}) = 0) \right\}, \quad (10)$$

where the multivariate quadratic polynomials $(f_i)_{i=1}^m$ are defined in Section 3.1. Note that, for the relation $\mathcal{R}_{PRF}$, it suffices for the PRF evaluator to prove that she knows the roots of $m = n - 1$ quadratic equations. The arithmetic circuit $\mathcal{C}_n$ expressing the relation $\mathcal{R}^*_{PRF} = \{\{n\}_K, \mathbf{x}\}$ can be characterized with the following metrics. The arithmetic circuit $\mathcal{C}_n$ has a constant circuit depth 3 (two layers of multiplication gates and one layer of subtraction (addition) gates), circuit width of $2n$, multiplication complexity of $\approx 1.5n$ (on average, since every $(1,-1)$ or $(-1,1)$ pair induces an extra multiplication gate in comparison with the $(1,1)$ and $(-1,-1)$ Legendre symbol pairs) and witness complexity of $n\lambda$ bits, i.e. n group elements. For an illustrative example, see Appendix D.1.

To prove in zero-knowledge the computational integrity of the arithmetic circuit evaluation, one might choose from several off-the-shelf zero-knowledge proof systems. Still, as of time of writing, the state-of-the-art zkSNARK proof system is due to Groth [31]. It provides proofs of size 3 group elements and verifier complexity of 3 pairings and n group operations and last but not least significant developer tooling. However, this proof system does not provide post-quantum security and furthermore, it would require a trusted setup, which is undesirable or even unattainable in many applications.

The most important proof system family of zero-knowledge succinct transparent arguments of knowledge was pioneered by the work of Ben-Sasson et al. [6]. STARK proof systems, on top of being succinct and zero-knowledge, provide post-quantum security and does not rely on trusted setups. The performance evaluation of [6] shows, that the proof of a Legendre PRF statement with 2^{21} multiplication gates, i.e. verifying $\approx 2^{19}$ Legendre-symbols, can be generated in less than a second, while can be verified in 100ms. The proof size is $\approx$ 100KB.

5.2 Oblivious PRF from the Legendre PRF

An oblivious PRF (OPRF) [48,25] is a two-party secure computation protocol (2PC) to evaluate a PRF $F(\cdot,\cdot)$ in an oblivious fashion. Specifically, it allows a

sender and a receiver with inputs K and x, respectively, to compute $F(K, x)$ such that the sender does not learn anything new from the protocol messages, while the receiver can output $F(K, x)$ without obtaining information about the used key K. Grassi et al. [30] showed an efficient protocol to evaluate the Legendre PRF in the multi-party setting. In the sequel, we adapt their original multi-party protocol to the OPRF setting and show the beneficial properties of the resulting Legendre OPRF. The protocol can be divided into online and offline parts, where the latter one is also called preprocessing phase that is entirely independent of the inputs of the participants and as secret shares of random squares, and multiplication triples are generated (for secret share multiplication, denoted with $\boxdot$, in the online phase). The idea is that in the online phase, the parties can secret share their inputs with each other and then they only have to prepare a masked value that has the same Legendre symbol as the sum of key K and input x. More formally, they have to compute $[c] = [s^2] \boxdot ([K] + [x])$ for $s \in_R \mathbb{F}_p$ such that the receiver reconstructs c at the end and computes $L_p(c) = L_p(K + x)$.

While the addition of secret shares is for free, i.e. corresponds to ordinary local addition, $\boxdot$ consumes one multiplication triple and requires one round of interaction and 2 group elements of communication. The resulting online part then consists of three rounds of interaction and 5 group elements of communication. For the formal definition of OPRF and the details of the Legendre OPRF, we refer to Appendix D.2. We note that the described protocol is only statistically correct as with probability $1/p = \Pr(s^2 = 0)$ the output is necessarily zero. For perfect correctness, we need to rule out $s^2 = 0$ in the preprocessing phase that is possible in expected constant (1) rounds. The security of the protocol can be reduced to the SLS problem following the blueprint of [30].

Our efficiency comparisons in Table D.2 shows that in terms of both message size and computational complexity, the Legendre OPRF is the most promising candidate for a post-quantum OPRF.

5.3 Oblivious Programmable PRF from the Legendre PRF

The notion of oblivious *programmable* PRF (O**P**PRF) was introduced in [42]. A PRF is said to be O**P**PRF if it is in addition to being an OPRF, also allows the sender to program the output of the OPRF at certain evaluation points. O**P**PRF is the corner-stone of the state-of-the-art multi-party private set intersection protocol of Kolesnikov et al. [42]. They formulated three *generic* O**P**PRF constructions, that can turn any OPRF into an O**P**PRF. These generic constructions provide different trade-offs, cf. Table D.3, and form the basis of state-of-the-art PSI protocols [42].

Programming the Legendre PRF We show how one can program efficiently the output of the Legendre PRF by carefully choosing the prime modulus. The naïve way to program the Legendre PRF would be to generate primes randomly and hope that the PRF outputs match the desired values y_i at the programmed

points x_i. This certainly works for small number of programmed points, however, this naïve PRF programming method incurs an exponential time-complexity in the number of programmed points.

To circumvent the exponential time-complexity of the programming of the Legendre PRF, we take a different approach, cf. Figure 4b. We note, however, that the "programmability" of the Legendre PRF is rather space-inefficient, since $p \approx \prod_{i=1}^{n} x_i$. Therefore, the number of programmed points is somewhat limited in the algorithm proposed in Figure 4b. The main ideas of this programming algorithm were already proposed in a different context (secure comparison protocols) by Yu [57]. In a similar fashion, one could generalize our approach in Figure 4b to power residue symbols, i.e. programming power residue symbol PRFs. This was already achieved by Cascudo et al. [12]. However, finding concrete applications of their protocol was proposed as an open question. We note that their methods can be applied to program power residue symbol OPRFs.

Our novel non-generic Legendre PRF programming methods minimize the necessary auxiliary information to evaluate the O**P**PRF. Hence, they incur minimal online communication costs in the applications. This could be favorable in PSI protocols.

6 Future Directions

We perceive three main areas for future work. There is still quite some work to be done on the *provable security* part of the Legendre PRF. It would be fascinating to find new connections to other post-quantum secure cryptographic assumptions, e.g. LWE. For instance, note that in Equation 13, the probability distribution of the coefficients of the quadratic terms in the induced MQ instance follows a discrete Gaussian distribution. Could one reframe the MQ instance as an LWE instance for a suitable change in the variables? Moreover, it would be fruitful to establish concrete and asymptotic lower bounds on the degree of regularity of the Legendre PRF's MQ instances. That would pave the path for settling the provable security of this PRF.

It is quintessential to improve on existing key-recovery attacks or find new, more performant cryptanalytic approaches. It would allow us to better estimate the *bit-security* of the Legendre PRF and other variants.

We foresee many more *novel cryptographic applications* of the Legendre PRF due to its homomorphic properties and MPC-friendliness. For instance, it seems accessible to prove the existence of related-key secure PRFs, verifiable OPRFs or key-homomorphic PRFs from quadratic and power residue symbol PRFs.

Acknowledgements

We are grateful for the insightful conversations to Gergő Zábrádi.

References

1. Albrecht, M., Grassi, L., Rechberger, C., Roy, A., Tiessen, T.: Mimc: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 191–219. Springer (2016)
2. Albrecht, M.R., Cid, C., Grassi, L., Khovratovich, D., Lüftenegger, R., Rechberger, C., Schofnegger, M.: Algebraic cryptanalysis of stark-friendly designs: application to marvellous and mimc. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 371–397. Springer (2019)
3. Albrecht, M.R., Davidson, A., Deo, A., Smart, N.P.: Round-optimal verifiable oblivious pseudorandom functions from ideal lattices. IACR Cryptol. ePrint Arch. **2019**, 1271 (2019)
4. Aly, A., Ashur, T., Ben-Sasson, E., Dhooghe, S., Szepieniec, A.: Design of symmetric-key primitives for advanced cryptographic protocols. IACR Transactions on Symmetric Cryptology pp. 1–45 (2020)
5. Ashur, T., Dhooghe, S.: Marvellous: a stark-friendly family of cryptographic primitives. IACR Cryptol. ePrint Arch. **2018**, 1098 (2018)
6. Ben-Sasson, E., Bentov, I., Horesh, Y., Riabzev, M.: Scalable, transparent, and post-quantum secure computational integrity. IACR Cryptol. ePrint Arch. **2018**, 46 (2018)
7. Beullens, W., Beyne, T., Udovenko, A., Vitto, G.: Cryptanalysis of the legendre prf and generalizations. IACR Transactions on Symmetric Cryptology pp. 313–330 (2020)
8. Beullens, W., de Saint Guilhem, C.D.: Legroast: Efficient post-quantum signatures from the legendre prf. In: International Conference on Post-Quantum Cryptography. pp. 130–150. Springer (2020)
9. Boneh, D., Gentry, C., Lynn, B., Shacham, H.: Aggregate and verifiably encrypted signatures from bilinear maps. In: International Conference on the Theory and Applications of Cryptographic Techniques. pp. 416–432. Springer (2003)
10. Boneh, D., Kogan, D., Woo, K.: Oblivious pseudorandom functions from isogenies. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 520–550. Springer (2020)
11. Buchberger, B.: Ein algorithmus zum auffinden der basiselemente des restklassenringes nach einem nulldimensionalen polynomideal. PhD thesis, Universitat Insbruck (1965)
12. Cascudo, I., Schnyder, R.: A note on secure multiparty computation via higher residue symbol techniques. IACR Cryptol. ePrint Arch. **2020**, 183 (2020)
13. Chase, M., Meiklejohn, S., Zaverucha, G.: Algebraic macs and keyed-verification anonymous credentials. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. pp. 1205–1216 (2014)
14. Courtois, N., Klimov, A., Patarin, J., Shamir, A.: Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In: International Conference on the Theory and Applications of Cryptographic Techniques. pp. 392–407. Springer (2000)
15. Cox, D., Little, J., OShea, D.: Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra. Springer Science & Business Media (2013)
16. van Dam, W., Hallgren, S., Ip, L.: Quantum algorithms for some hidden shift problems. SIAM J. Comput. **36**(3), 763–778 (2006)

17. Damgård, I.B.: On the randomness of legendre and jacobi sequences. In: Conference on the Theory and Application of Cryptography. pp. 163–172. Springer (1988)
18. Déchene, I.: Generalized Jacobians in cryptography. ProQuest (2007)
19. Demmler, D., Schneider, T., Zohner, M.: ABY - A framework for efficient mixed-protocol secure two-party computation. In: NDSS. The Internet Society (2015)
20. Ding, C., Hesseseth, T., Shan, W.: On the linear complexity of legendre sequences. IEEE Transactions on Information Theory **44**(3), 1276–1278 (1998)
21. Dodis, Y., Kiltz, E., Pietrzak, K., Wichs, D.: Message authentication, revisited. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 355–374. Springer (2012)
22. Dodis, Y., Yampolskiy, A.: A verifiable random function with short proofs and keys. In: International Workshop on Public Key Cryptography. pp. 416–431. Springer (2005)
23. Esgin, M.F., Kuchta, V., Sakzad, A., Steinfeld, R., Zhang, Z., Sun, S., Chu, S.: Practical post-quantum few-time verifiable random function with applications to algorand. IACR Cryptol. ePrint Arch **2020**, 1222 (2020)
24. Faugere, J.C.: A new efficient algorithm for computing gröbner bases without reduction to zero (f 5). In: Proceedings of the 2002 international symposium on Symbolic and algebraic computation. pp. 75–83 (2002)
25. Freedman, M.J., Ishai, Y., Pinkas, B., Reingold, O.: Keyword search and oblivious pseudorandom functions. In: TCC. Lecture Notes in Computer Science, vol. 3378, pp. 303–324. Springer (2005)
26. Frixons, P., Schrottenloher, A.: Quantum security of the legendre prf. Cryptology ePrint Archive, Report 2021/149 (2021), https://eprint.iacr.org/2021/149
27. Gilad, Y., Hemo, R., Micali, S., Vlachos, G., Zeldovich, N.: Algorand: Scaling byzantine agreements for cryptocurrencies. In: Proceedings of the 26th Symposium on Operating Systems Principles. pp. 51–68 (2017)
28. Goldberg, S., Naor, M., Papadopoulos, D., Reyzin, L., Vasant, S., Ziv, A.: Nsec5: Provably preventing dnssec zone enumeration. In: NDSS (2015)
29. Grassi, L., Khovratovich, D., Rechberger, C., Roy, A., Schofnegger, M.: Poseidon: A new hash function for zero-knowledge proof systems. In: Proceedings of the 30th USENIX Security Symposium. USENIX Association (2020)
30. Grassi, L., Rechberger, C., Rotaru, D., Scholl, P., Smart, N.P.: Mpc-friendly symmetric key primitives. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 430–443. ACM (2016)
31. Groth, J.: On the size of pairing-based non-interactive arguments. In: Annual international conference on the theory and applications of cryptographic techniques. pp. 305–326. Springer (2016)
32. Gyarmati, K., Mauduit, C., Sárközy, A.: The cross-correlation measure for families of binary sequences. (2014)
33. Hartshorne, R.: Algebraic geometry, vol. 52. Springer Science & Business Media (2013)
34. Huang, Y.J., Liu, F.H., Yang, B.Y.: Public-key cryptography from new multivariate quadratic assumptions. In: International Workshop on Public Key Cryptography. pp. 190–205. Springer (2012)
35. Jakobsen, T., Knudsen, L.R.: The interpolation attack on block ciphers. In: International Workshop on Fast Software Encryption. pp. 28–40. Springer (1997)
36. Jarecki, S., Kiayias, A., Krawczyk, H.: Round-optimal password-protected secret sharing and T-PAKE in the password-only model. In: ASIACRYPT (2). Lecture Notes in Computer Science, vol. 8874, pp. 233–253. Springer (2014)

37. Kaluderovic, N., Kleinjung, T., Kostic, D.: Improved key recovery on the legendre prf. IACR Cryptol. ePrint Arch. **2020**, 98 (2020)
38. Keller, M., Pastro, V., Rotaru, D.: Overdrive: Making SPDZ great again. In: EUROCRYPT (3). Lecture Notes in Computer Science, vol. 10822, pp. 158–189. Springer (2018)
39. Khovratovich, D.: Key recovery attacks on the legendre prfs within the birthday bound. Cryptology ePrint Archive, Report 2019/862 (2019), https://eprint.iacr.org/2019/862
40. Kipnis, A., Shamir, A.: Cryptanalysis of the hfe public key cryptosystem by relinearization. In: Annual International Cryptology Conference. pp. 19–30. Springer (1999)
41. Kolesnikov, V., Kumaresan, R., Rosulek, M., Trieu, N.: Efficient batched oblivious PRF with applications to private set intersection. In: CCS. pp. 818–829. ACM (2016)
42. Kolesnikov, V., Matania, N., Pinkas, B., Rosulek, M., Trieu, N.: Practical multiparty private set intersection from symmetric-key techniques. In: CCS. pp. 1257–1272. ACM (2017)
43. Lemmermeyer, F.: Conics-a poor man's elliptic curves. arXiv preprint math/0311306 (2003)
44. Li, C., Preneel, B.: Improved interpolation attacks on cryptographic primitives of low algebraic degree. In: International Conference on Selected Areas in Cryptography. pp. 171–193. Springer (2019)
45. Liang, B., Banegas, G., Mitrokotsa, A.: Statically aggregate verifiable random functions and application to e-lottery. Cryptography **4**(4), 37 (2020)
46. Mauduit, C., Sárközy, A.: On finite pseudorandom binary sequences i: Measure of pseudorandomness, the legendre symbol. Acta Arithmetica **82**(4), 365–377 (1997)
47. Micali, S., Rabin, M., Vadhan, S.: Verifiable random functions. In: 40th annual symposium on foundations of computer science (cat. No. 99CB37039). pp. 120–130. IEEE (1999)
48. Naor, M., Reingold, O.: Number-theoretic constructions of efficient pseudo-random functions. In: FOCS. pp. 458–467. IEEE Computer Society (1997)
49. Ospina, D.E.E.: Groebner bases and applications to the security of multivariate public key cryptosystems. Ph.D. thesis, Ph. D. dissertation, Escuela de Matemáticas, Univ. Nacional de Colombia ... (2016)
50. Papadopoulos, D., Wessels, D., Huque, S., Naor, M., Včelák, J., Reyzin, L., Goldberg, S.: Making nsec5 practical for dnssec. Cryptology ePrintArchive, Report 2017/099 (2017)
51. Peralta, R.: On the distribution of quadratic residues and nonresidues modulo a prime number. Mathematics of Computation **58**(197), 433–440 (1992)
52. Perlner, R., Petzoldt, A., Smith-Tone, D.: Total break of the srp encryption scheme. In: International Conference on Selected Areas in Cryptography. pp. 355–373. Springer (2017)
53. Russell, A., Shparlinski, I.E.: Classical and quantum function reconstruction via character evaluation. Journal of Complexity **20**(2-3), 404–422 (2004)
54. Sugita, M., Kawazoe, M., Imai, H.: Relation between xl algorithm and gröbner bases algorithms, iacr eprint server (2004)
55. Tóth, V.: Collision and avalanche effect in families of pseudorandom binary sequences. Periodica Mathematica Hungarica **55**(2), 185–196 (2007)
56. Vinogradov, I.M.: Elements of number theory. Courier Dover Publications (2016)
57. Yu, C.H.: Sign modules in secure arithmetic circuits. IACR Cryptol. ePrint Arch. **2011**, 539 (2011)

A Background

For completeness, we define possible generalisations of the Legendre PRF.

Definition 5 (Higher-degree Legendre PRF). *In case of the Higher-degree Legendre PRF with a secret polynomial $f \in_R \mathbb{F}_p[x]$, let $\{a\}_f$ denote the following sequence:*

$$\{a\}_f := \left(\frac{f(0)}{p}\right), \left(\frac{f(1)}{p}\right), \ldots, \left(\frac{f(a-1)}{p}\right).$$

Definition 6 (rth power residue function). *Let $p \equiv 1 \mod r$ and $g \in \mathbb{F}_p^\times$ a generator. The rth power residue function $l^{(r)} : \mathbb{F}_p \to \mathbb{Z}_r$ is defined as*

$$l^{(r)}(a) := \begin{cases} k, & \text{if } a \not\equiv 0 \mod p \wedge a/g^k \text{is an } r\text{th power} \mod p \\ 0, & \text{if } a \equiv 0 \mod p. \end{cases}$$

Similarly to Definitions 1 and 5, we might introduce the power residue PRF and its higher-degree variants, relying on the power residue function. Once again, we note that our results and observations can be generalized to the higher-degree and other variants of the Legendre PRF.

B The MQ Instance Induced by the Legendre PRF

B.1 An Alternative View

We view the resulting equation system globally and assess the probability distribution of each coefficient to appear in the MQ instance. Adjacent pairs of Legendre symbols are asymptotically equi-distributed [51]. Therefore we can easily describe the discrete probability distribution of the coefficients in the induced equation system. Let $X_q^{(i,j)}, X_l^{(i)}, X_c$ be the random discrete variables corresponding to the ith unknown's quadratic, linear and constant terms. For the equation system's coefficients, we have the following discrete probability distributions given Equations 1, 2 and 3. For the constant terms, we have that

$$\Pr[X_c = 1] = \Pr[X_c = r] = \frac{1}{2}. \tag{11}$$

Every linear term is zero, namely,

$$\Pr[X_l^{(i)} = 0] = 1, \forall i \in [1, n]. \tag{12}$$

Finally, the quadratic terms' coefficients have the following probability distribution. The $\Pr[X_q^{(i,j)} = 0] = 1$, if $i \neq j$,. Otherwise, we have that

$$\begin{gathered} \Pr[X_q^{(i,i)} = 1] = \frac{1}{n}, \quad \Pr[X_q^{(i,i)} = -1] = \frac{1}{2n}, \\ \Pr[X_q^{(i,i)} = -r] = \Pr[X_q^{(i,i)} = -r^{-1}] = \frac{1}{4n}, \quad \Pr[X_q^{(i,i)} = 0] = 1 - \frac{2}{n}. \end{gathered} \tag{13}$$

We remark that the discrete probability distribution of the quadratic terms is reminiscent of a discrete normal Gaussian distribution with average 0, whenever n goes to infinity. If the linear terms, cf. Equation 12, would follow a uniformly random distribution after a suitable change in the variables, the resulting MQ instance could be seen asymptotically as a learning with errors (LWE) instance. We leave this as an interesting future direction to investigate further connections to other post-quantum secure assumptions.

B.2 Proof of Theorem 1

Proof. By Buchberger-criterion, we only need to verify that for all i, j, it holds that the S-polynomial $S(g_i, g_j)$ divided by the Gröbner-basis has no remainder, i.e. $\overline{S(g_i, g_j)}^G = 0$. We let $i < j$ and hereby solely consider the case when $b_i = b_j = b_{n-1} = 1$. The rest of the cases result in a similar calculation. By the definition of the S-polynomials, we have $S(g_i, g_j) = x_j^2 g_i - x_i^2 g_j$. First, we divide $S(g_i, g_j)$ by g_i. We observe that the remainder of the polynomial division is $g_j(x_{n-1}^2 - (n-i))$, which is divisible by g_j. Therefore, indeed $\overline{S(g_i, g_j)}^G = 0$. Hence, the polynomials in G indeed form a Gröbner-basis.

Example 2. The Gröbner-basis of the polynomials corresponding to the Legendre symbol sequence $\{5\}_K$, from Example 1, consists of the following quadratic bivariate polynomials:

$$\langle x_0^2 - x_4^2 + 4, x_1^2 - x_4^2 + 3, x_2^2 - 2x_4^2 + 4, x_3^2 - 2x_4^2 + 2 \rangle.$$

B.3 Proof of Lemma 2

Proof. Let $I = \langle f_1, \ldots, f_m \rangle$ be the ideal induced by the Legendre PRF, and we assume that f_i forms a reduced Gröbner-basis. For a homogeneous sequence of polynomials $(f_1, \ldots, f_m)$ being regular, we need to show that if for all $i \in [1, m]$ and g such that $gf_i \in \langle f_1, \ldots, f_{i-1} \rangle$, then $g \in \langle f_1, \ldots, f_{i-1} \rangle$. An affine sequence of polynomials $(f_1, \ldots, f_m)$ is regular by definition, if the homogeneous sequence $(f_1^h, \ldots, f_m^h)$ is regular, where f_i^h is the homogeneous part of f_i of highest degree with respect to the (graded) lexicographic monomial ordering. In our case $(f_1^h, f_2^h, \ldots, f_m^h) = (x_1^2, x_2^2, \ldots, x_m^2)$.

Since $f_i^h = x_i^2$, in our case for every i, therefore the ideal $I_{i-1} := \langle f_1^h, \ldots, f_{i-1}^h \rangle$ is a monomial ideal. If $gf_i^h \in I_{i-1}$, then gf_i^h is divisible by a generator of I_{i-1}, since I_{i-1} is a monomial ideal [15]. Since $(f_i, f_j) = 1$, for every $j \in [1, i-1]$, thus it is necessary that g is divisible by some $f_j^h = x_j^2 \in I_{i-1}$, for $j \leq i-1$. Namely $g = x_j^2 g' \in I_{i-1}$, for some polynomial g'. This completes the proof.

C Algebraic Cryptanalysis of the Legendre PRF

C.1 Computing the Q-rank of the Legendre PRF

The Q-rank of a MQ cryptosystem plays a crucial role in cryptanalysis. Every multivariate quadratic equation system $\mathbf{f}$ can be lifted to a quadratic form $\mathcal{Q}$ in

m	n	d_{reg} Random MQ	d_{reg} Legendre MQ
7	7	3	3
8	8	4	4
9	9	4	4
10	10	5	5
11	11	5	5

Table C.1: Degree of regularity for a random MQ system and a Legendre PRF MQ instance for various small parameters of m and n. The corresponding prime p was chosen to be 32003. Since $p \equiv 3 \mod 4$, we are in the (over)determined case. Adding a single high-degree equation, cf. Section 3.3, causes the equation system to behave like a random system in terms of degree of regularity.

an extension field. Let $\mathbb{E}$ denote an extension field over $\mathbb{F}_p$. Informally, Q-rank is the rank of the quadratic form $\mathcal{Q}$ as a matrix over the field $\mathbb{E}$. Low Q-rank is detrimental, since it facilitates successful cryptanalysis (key-recovery, decryption etc.) [40,52].

Definition 7 (Q-rank). *The Q-rank of a multivariate quadratic map* $\mathbf{f} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ *over the finite field* $\mathbb{F}_q$ *is the rank of the quadratic form* $\mathcal{Q}$ *on the extension field* $\mathbb{E}[X_0, \ldots, X_{n-1}]$ *defined by* $Q(X_0, \ldots, X_{n-1}) = \phi \circ \mathbf{f} \circ \phi^{-1}(X, X^q, \ldots, X^{q^{n-1}})$, *under the identification* ϕ*:* $X_0 = X, X_1 = X^q, \ldots, X_{n-1} = X^{q^{n-1}}$.

We compute now the Q-rank (cf. Definition C.1) of the Legendre PRF equation system [49]. We rewrite each generator polynomial f_i in the ideal $I = \langle f_1, \ldots, f_m \rangle$ induced by the Legendre PRF, as folllows:

$$f_i(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c = \mathbf{x}^T A_i \mathbf{x} + B_i \mathbf{x} + c, \tag{14}$$

where $\mathbf{x} = [x_1, \ldots, x_n]^T$, $A_i \in \mathcal{M}_{n \times n}(\mathbb{F})$ is the matrix $[a_{ij}]_{ij}$ and $B_i \in \mathcal{M}_{1 \times n}(\mathbb{F})$ is the matrix $[b_i]_{1i}$. We note, that in the case of the Legendre PRF, $B_i = \mathbf{0}$. Each polynomial f_i can be represented in the extension field, in the following form:

$$\mathcal{F}_i(X) = \sum_{i,j=1}^{n} \alpha_{ij} X^{q^{i-1}+q^{j-1}} + \sum_{i=1}^{n} \beta_i X^{q^{i-1}} + \gamma = \mathbf{X}^T M_i \mathbf{X} + N_i \mathbf{X} + \gamma, \tag{15}$$

where $\mathbf{X} = [X^{q^0}, \ldots, X^{q^{n-1}}]^T$, $M_i \in \mathcal{M}_{n \times n}(\mathbb{E})$ is the matrix $[\alpha_{ij}]_{ij}$ and $B \in \mathcal{M}_{1 \times n}(\mathbb{F})$ is the matrix $[\beta_i]_{1i}$. It is well-known that a quadratic polynomial equation system F defined by the generating polynomials f_i of I, can be lifted to the extension field by

$$\mathsf{Lft}(F)(X) = \phi^{-1} \circ \mathcal{F} \circ \phi(X) = \mathbf{X}^T M \mathbf{X} + N \mathbf{X} + \gamma, \tag{16}$$

where $\mathbf{x} = \phi(X)$. Our goal is to establish the rank of the matrix $M \in \mathcal{M}_{n\times n}(\mathbb{E})$. We start off by defining $\mathbf{X} = \Delta\cdot\phi(X)$, where Δ is the following invertible matrix,

$$\Delta = \begin{bmatrix} y^0 & y^1 & \dots & y^{n-2} & y^{n-1} \\ (y^0)^{q^1} & (y^1)^{q^1} & \dots & (y^{n-2})^{q^1} & (y^{n-1})^{q^1} \\ (y^0)^{q^2} & (y^1)^{q^2} & \dots & (y^{n-2})^{q^2} & (y^{n-1})^{q^2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (y^0)^{q^{n-1}} & (y^1)^{q^{n-1}} & \dots & (y^{n-2})^{q^{n-1}} & (y^{n-1})^{q^{n-1}} \end{bmatrix} \tag{17}$$

Equipped with all this, we can now define $M \in \mathcal{M}_{n\times n}(\mathbb{F}), N \in \mathcal{M}_{1\times n}(\mathbb{F})$ and $\gamma \in \mathbb{E}$ from the lifting Equation 16. We define $\gamma = c_1 + c_2 y + \dots + c_n y^{n-1}$ and the matrices as,

$$M = (\Delta^T)^{-1}\left(\sum_{i=1}^{n} y^{i-1}A_i\right)\Delta^{-1} \qquad \textit{and} \qquad N = \left(\sum_{i=1}^{n} y^{i-1}B_i\right)\Delta^{-1}. \tag{18}$$

Note that in case of the Legendre PRF MQ instance, $N = 0$, since $B_i = \mathbf{0}$ for all i. The second term in matrix M, $\sum y^{i-1}A_i$ is a double diagonal non-singular matrix. Hence, matrix M has full rank, since it is the product of non-singular matrices.

C.2 Computing the genus

We want to calculate the genus of the algebraic curve containing the solutions of a Legendre PRF key-recovery attack. More formally, we want to compute $1 - P(0)$, where $P(\cdot)$ is the Hilbert-polynomial of the curve defined by the intersection of several Pell conics. Let $(f_1, f_2, \dots, f_m)$ be the given Pell conics in variables $x_0, x_1, \dots, x_n$ and I the corresponding ideal generated by them. Note that n denotes the length of the given Legendre sequence. For $N \gg 0$, we have that $P(N)$ is the dimension over $\mathbb{F}_p$ of the degree-N homogenous part of I in $\mathbb{F}_p[x_0, \dots, x_n]/I$ [33]. This is a linear polynomial. Since for all $i, j, i \neq j$ we have $(f_i, f_j) = 1$, we obtain the following inclusion-exclusion type equation,

$$P_n(N) = g_n(N) - \binom{n-1}{1} g_n(N-2) + \binom{n-1}{2} g_n(N-4) - \binom{n-1}{3} g_n(N-6) + \dots, \tag{19}$$

where $g_n(N)$ denotes the number of N-degree monomials in $\mathbb{F}_p[x_0, \dots, x_n]$. Therefore $g_n(N) = \binom{N+n}{n}$. For the sake of concreteness and as a simple example let us consider the case of four intersecting Pell-conics, i.e. Legendre-sequences of length five. We have the following expression for the Hilbert-polynomial, when $n = 4$:

$$P_4(N) = \binom{N+4}{4} - 3\binom{N+2}{4} + 3\binom{N}{4} - \binom{N-2}{4}. \tag{20}$$

By substituting $N = 0$, we obtain that $P_4(0) = -4$, namely the arithmetic genus is $1 - P_4(0) = 5$.

A lengthy but straightforward computation shows that in the general case, we can obtain the following closed formula for the Hilbert-polynomial:

Lemma 3. $P_n(N) = 2^{(n-1)} \cdot N - (n-1) \cdot 2^{(n-2)}$.

D Applications of the Legendre PRF

D.1 Verifiable Random Function

Figure 2 illustrates the low multiplicative complexity of the statement a Legendre PRF evaluator needs to prove in zero-knowledge to obtain a VRF from the Legendre PRF. Table D.1 compares the proposed VRF to the state of the art.

		Time complexity		
	$\lvert\pi\rvert$	Prove	Verify	Assumption
[28]	$1\mathbb{G}$	$1\mathsf{H} + 1\mathbb{G}$	$1\mathsf{H} + 1\mathbb{G}$	Factoring
[50]	$1\mathbb{G} + 2\mathbb{F}_p$	$3\mathsf{H} + 2\mathbb{G}$	$3\mathsf{H} + 4\mathbb{G}$	EC-DDH
[9]	$1\mathbb{G}$	$2\mathsf{H} + 1\mathbb{G}$	$1P$	co-DH
[22]	$1\mathbb{G}$	$1\mathbb{G} + 1\mathbb{F}_p$	$2\mathbb{G} + 2P$	q-DBDHI
[45]	$1\mathbb{G}$	$1\mathbb{G}$	$1P$	q-DDHE
[23]†	$\mathcal{O}(k+l)$	$\mathcal{O}(kl)$	$\mathcal{O}(kl)$	Module-SIS
§5.1+SNARK	$3\mathbb{G}$	$9n\mathbb{G}$	$n\mathbb{G} + 3P$	SLS, KEA
§5.1+STARK	$\mathcal{O}(\log(n))\mathbb{G}$	$\mathcal{O}(n\log(n))\mathbb{G}$	$\mathcal{O}(\log(n))\mathbb{G}$	SLS

Table D.1: Overview of various VRF constructions. Hashing, group operations, exponentiation and pairings are denoted as $\mathsf{H}, \mathbb{G}, \mathbb{F}_p, P$ respectively. Note that [23] only provides a few-time VRF. Module-SIS and module-LWE ranks are denoted as k and l, respectively. In case of the Legendre VRF, n is the length of the Legendre-symbol sequence being proved. Assumptions written in red are not post-quantum secure, while assumptions in green are post-quantum secure.

D.2 Oblivious Pseudorandom Function

For completeness, we formally define the ideal functionalities of O(P)PRF (see Figures 3b–3c) and also describe the proposed protocols in more details.

For simplicity, we abstract away the underlying details of preprocessing and use the necessary operations in a black-box manner through the ideal functionality of Figure 3a. Potential realizations of $\mathcal{F}_{\text{prepr}}$ is possible using several 2PC or MPC frameworks, e.g. ABY by [19] in the semi-honest or Overdrive [38] in the malicious setting.

For one bit outputs, Π_{Legendre} requires the precomputation of a random square and a Beaver multiplication triple $[a], [b], [ab]$. Using these $[x] \boxdot [y] = [xy]$ can be computed by revealing $(x+a)$ and $(y+b)$ (that does not disclose information about x and y, because a, b are random), then $(x+a) \cdot (y+b) - (x+a) \cdot [b] - (y+b) \cdot [a] + [ab] = [xy]$ can be evaluated. As already noted, the security of

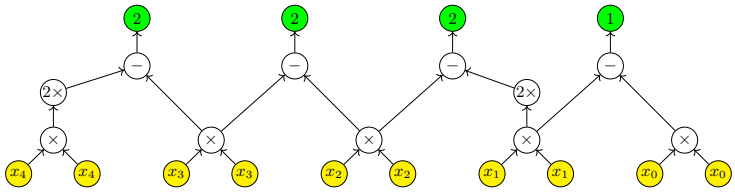

Fig. 2: Arithmetic circuit representation of the ZKP statement that proves the relation $\mathcal{R}_{PRF} = \{\{5\}_K = (1, 1, -1, -1, 1), K\}$ from Example 1 where 2 is the least quadratic non-residue. Applying our arithmetization the PRF evaluator proves that it knows the zeros of the following polynomials $(2x_4^2 - x_3^2 = 2, x_3^2 - x_2^2 = 2, x_2^2 - x_1^2 = 2, x_1^2 - x_0^2 = 1)$. Secret input nodes are colored with yellow, while public output nodes are colored with green. Nodes with $2x$ denote a multiplication gate, where one of the inputs is the constant quadratic non-residue 2. Note, that for any Legendre PRF statement $\mathcal{R}^*_{PRF}$ the arithmetic circuit has a constant multiplicative depth of two.

Π_{Legendre} of Figure 4a can be reduced to the SLS problem in the $\mathcal{F}_{\text{prepr}}$-hybrid model following the blueprint of [30].

For comparisons to other OPRF protocols, we refer to Table D.2.

Functionality $\mathcal{F}_{\text{prepr}}$

RandSquare: Sample $s \in_R \mathbb{F}_p$ and output shares $[s^2]$.
TripleGen: Sample $a, b \in_R \mathbb{F}_p$ and output shares $[a], [b], [ab]$.

(a) Ideal functionality of preprocessing.

Functionality $\mathcal{F}_{\text{OPRF}}$

Participants: sender $\mathcal{S}$, receiver $\mathcal{R}$.
Parameters: a PRF $F : \mathcal{K} \times \mathcal{X} \to \{0,1\}$, for key-space $\mathcal{K}$ and input-space $\mathcal{X}$.
Input:
- $\mathcal{S}$: $K \in \mathcal{K}$,
- $\mathcal{R}$: $x \in \mathcal{X}$.

Output:
- $\mathcal{S}$ obtains nothing,
- $\mathcal{R}$ obtains $F(K, x)$.

(b) Ideal functionality of OPRF.

Functionality $\mathcal{F}_{\text{OPPRF}}$

Participants: sender $\mathcal{S}$, receiver $\mathcal{R}$.
Parameters: a PRF $F : \mathcal{K} \times \mathcal{X} \to \{0,1\}$, for key-space $\mathcal{K}$ and input-space $\mathcal{X}$, and n the number of programmed points.
Input:
- $\mathcal{S}$: $K \in \mathcal{K}, x'_1, \ldots, x'_n \in \mathcal{X}$ and random $y'_1, \ldots, y'_n \in \{0,1\}$.
- $\mathcal{R}$: $x \in \mathcal{X}$.

Output:
- $\mathcal{S}$ obtains nothing,
- $\mathcal{R}$ obtains $F(K, x)$ that is y'_i if $x = x'_i \; \forall i \in \{1, \ldots, n\}$.

(c) Ideal functionality of OP**P**RF.

Fig. 3: Ideal functionalities that we use in this work.

D.3 Oblivious Programmable Pseudorandom Function

We first review the additional algorithms an O**P**PRF consists of on top of being an OPRF:

Protocol Π_{Legendre}

Participants: sender $\mathcal{S}$, receiver $\mathcal{R}$.
Preprocessing:
- execute RandSquare,
- execute TripleGen.

Input:
- $\mathcal{S}$: $K \in \mathbb{F}_p$,
- $\mathcal{R}$: $x \in \mathbb{F}_p$.

Evaluation:
1. $\mathcal{S}$, $\mathcal{R}$ share $[K], [x]$ with each other,
2. both compute $[c] = [s^2] \boxdot ([K] + [x])$,
3. $\mathcal{S}$ sends $[c]$ to $\mathcal{R}$,
4. $\mathcal{R}$ outputs $L_p(c) = L_p(K + x)$.

(a) Legendre OPRF based on [30]

Algorithm $\mathsf{Lprog}(\lambda, (x_1, y_1), \ldots, (x_n, y_n)) \to p$
- Find a p prime, s.t.$\forall i \in [0, n)$: $y_i = \left(\frac{x_i}{p}\right) = \left(\frac{p}{x_i}\right)(-1)^{\frac{(p-1)(x_i-1)}{4}}$.
- Without loss of generality search p in the form $p \equiv 1 \mod 4$.
- Compute $y_i(-1)^{\frac{(p-1)(x_i-1)}{4}} = \left(\frac{p}{x_i}\right)$.
- Identify $m_i \in \mathbb{Z}_{x_i}$, s.t. $\left(\frac{m_i}{x_i}\right) = y_i(-1)^{\frac{(p-1)(x_i-1)}{4}}$.
- $\forall i$ let M_i be $M_i = \left\{m | m \in \mathbb{Z}_{x_i} \wedge b_i(-1)^{\frac{(p-1)(x_i-1)}{4}} = \left(\frac{m}{x_i}\right)\right\}$.
- If $m \in M_i$, then p can be sought as $p \equiv m \mod x_i$.
- Note, p is a solution of $p \equiv m_i \mod x_i$, for all $i \in [0, n)$, where $m_i \in M_i$. Solve this by the Chinese-Remainder Theorem.

Output: p

(b) Programming the Legendre OPRF of Figure 3b by appropriate parameter selection. For ease of exposition, we assume that all the programmed points x_i are primes.

Fig. 4: Legendre OPRF and the algorithm to extend it to be an OP**P**RF.

- $\mathsf{KeyGen}(1^\lambda, \mathcal{P}) \to (K, \mathsf{hint})$: Given a security parameter and set of points $\mathcal{P} = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ with distinct x_i-values, generates a PRF key K and (public) auxiliary information hint.
- $F(K, \mathsf{hint}, x) \to y$: Evaluates the PRF on input x, yielding output y.

We require from an O**P**PRF the following high-level security notions to hold:

Correctness: whenever $(x, y) \in \mathcal{P} \wedge \big((k, \mathsf{hint}) \leftarrow \mathsf{KeyGen}(\mathcal{P})\big) \implies F(k, \mathsf{hint}, x) = y$.

$(n, t)-$**security**: No efficient adversary should be able to distinguish the n programmed points from non-programmed points given oracle access to the PRF using t queries. Note that this definition implies that unprogrammed PRF outputs (i.e., those not set by the input to KeyGen) are pseudorandom.

For the formal security definitions, the reader is referred to [42].

We observe that the Legendre PRF can be natively and efficiently programmed using elementary number-theoretic techniques, see Figure 4b.

E Group Structure of the Solutions of a Legendre PRF key-recovery attack

In Section 4.4, we showed that if there exists a probabilistic polynomial-time algorithm that breaks the SLS problem, then it could be used to find solutions of high order algebraic curves over $\mathbb{F}_p$. This is essentially an equivalent restatement of viewing the Legendre PRF as an MQ instance.

Moreover, the resulting algebraic curves have a genus greater than 1, implying that the solutions lying on the curve lack an Abelian group structure. However,

OPRF	Comm. Complexity			Comp. Complexity		Model	Assumption
	Rounds	Msg. Size	Concr. eff.	Client	Server		
RSA-OPRF	2	2 $\mathbb{G}$	0.77KB	1H + 2 $\mathbb{G}$	1 $\mathbb{G}$	ROM	1-more-RSA-inv
[36]	2	2 $\mathbb{G}$	64 byte	1H + 2 $\mathbb{G}$	1 $\mathbb{G}$	ROM/Standard	EC-DDH
[41]†	5	2λ bits	256 bits	1H + 2XOR	2H + 2XOR	ROM	OT*
[3]	2	$\mathcal{O}(\lambda^c)$ $\mathbb{F}_p$	$\approx$ 1MB	$\mathcal{O}(\lambda^c)$ $\mathbb{F}_p$	$\mathcal{O}(\lambda^c)$ $\mathbb{F}_p$	QROM	RLWE
[10]	2	$\mathcal{O}(\lambda)$ $\mathbb{G}$	$\approx$ 2MB	$\mathcal{O}(\lambda)$ $\mathbb{G}$	$\mathcal{O}(\lambda)$ $\mathbb{G}$	ROM	SIDH
Section 5.2	3	5λ $\mathbb{G}$	13.44KB	17λ $\mathbb{G}$	17λ $\mathbb{G}$	ROM	SLS, OT*

Table D.2: Comparing the online costs of various Oblivious PRF protocols. In the columns of communication and computation complexity $\mathbb{G}$ denotes a group element or group operation, while H denotes a hashing operation. Concrete efficiency of obtaining λ pseudorandom bits with the corresponding OPRFs were computed with $\lambda = 128$ bit-security. (Q)ROM stands for the (quantum) random oracle model. Note, that the PRF of [41] is only a relaxed PRF. SIDH stands for the Supersingular Isogeny Diffie-Hellman assumption, while RLWE is the abbreviation for the ring-learning with errors assumption. Oblivious transfer (OT) can be instantiated both with classic and post-quantum security. Non post-quantum secure assumptions are written in red, while assumptions written in green are secure even against quantum attackers.

in the case of shorter sequences, e.g. Legendre sequences of length three, all the points that result in a specific Legendre symbol sequence of length three lie on a sequence-specific non-singular elliptic curve. In the sequel, we show how to obtain the Legendre-sequence specific elliptic curve equation by elementary methods.

E.1 The Case of Consecutive Legendre-symbol triplets

Let us suppose that one wants to generate key candidates $K^{'}$, whose subsequent Legendre symbols match the first three symbols of a sequence, i.e. $\left(\left(\frac{K^{'}}{p}\right), \left(\frac{K^{'}+1}{p}\right), \left(\frac{K^{'}+2}{p}\right)\right) = (b_0, b_1, b_2)$. Hereby, we show that such key candidates can be obtained as solutions of an elliptic curve over $\mathbb{F}_p$. One might generalise this approach to potentially speed up key-recovery attacks against the Legendre PRF and reduce its security to finding rational points on higher order algebraic curves over $\mathbb{F}_p$.

For the sake of concreteness, let us assume that $(b_0, b_1, b_2) = (1, 1, 1)$. Similar techniques apply for other bit-sequence patterns. Put it differently, the shifted Legendre sequence starts with 3 quadratic residues. Let us denote the corresponding square roots as $a, b, c \mod p$. Therefore we wish to solve the following equations:

$$c^2 - b^2 = b^2 - a^2 = 1$$

OPPRF	Programming complexity	Hint size	Online communication complexity	Constraint on no. of programmed points	No. of evaluations
Lagrange interpol.	$O(n^2)$	$O(n)$	$(n+kn)$ $\mathbb{G}$	space-efficiency	any
Garbled Bloom Filter	$O(n\lambda_{\mathsf{BF}})$	$n\lambda_{\mathsf{BF}}$	$(60n+kn)$ $\mathbb{G}$	space-efficiency	any
Table-based	$O(n)$	$O(n)$	$(n+kn)$ $\mathbb{G}$	space-efficiency	1
Legendre 5.3	$O(n \log n)$	1	$\mathcal{O}(n)$ $\mathbb{G}$	depends on λ	any
Legendre bruteforce	$O(2^n)$	1	1 $\mathbb{G}$	time-efficiency	any

Table D.3: Comparison of the generic OPPRF constructions of [42] (these are all built from an OPRF, e.g. that of [41]) and the Legendre OPRF that was shown to be programmable in Section 5.3. The number of programmed input positions is denoted as n, λ_{BF} is the soundness parameter of the Bloom filter, while k denotes the number of base-OTs, typically $k \approx 4\lambda$.

We introduce the following notation: $s := b - a$, $\frac{1}{s} := b + a$ and $\frac{c-b}{b-a} = \lambda$. We have that $2b = s + \frac{1}{s}$ and $2b = \frac{1}{s\lambda} - s\lambda$. This implies the following:

$$s + \frac{1}{s} = \frac{1}{s\lambda} - s\lambda$$

$$s^2\lambda + \lambda = 1 - s^2\lambda^2$$

$$s^2 = \frac{1-\lambda}{\lambda^2+\lambda}$$

$$s^2(1+\lambda)^2\lambda^2 = (1-\lambda)(1+\lambda)\lambda \tag{21}$$

By denoting the left hand side of Equation 21. as t^2, we finally obtain the following nonsingular elliptic curve of genus 1:

$$t^2 = \lambda^3 - \lambda.$$

4-symbol case (sketch): Now, let us assume we have an additional $b_3 = 1$. Let d be the square-root of $K + 3$. Furhtermore, let $r := c - b$ and $\mu := \frac{d-c}{c-b}$. Given Equation 21, we also have that

$$r^2(1+\mu)^2\mu^2 = (1-\mu)(1+\mu)\mu \tag{22}$$

Since, $r = s\lambda$ we can squeeze Equation 21 and Equation 22 into a single two-variable quartic equation:

$$\lambda^2\mu^2 + \lambda^2\mu - \lambda\mu^2 - \lambda\mu + \lambda - \mu - \lambda\mu + 1 = 0$$