Applicable Algebra in Engineering, Communication and Computing
https://doi.org/10.1007/s00200-023-00599-2

ORIGINAL PAPER

The Legendre pseudorandom function as a multivariate quadratic cryptosystem: security and applications

István András Seres[1] · Máté Horváth[2] · Péter Burcsi[1]

Received: 15 August 2022 / Revised: 4 November 2022 / Accepted: 1 February 2023

Abstract
Sequences of consecutive Legendre and Jacobi symbols as pseudorandom bit generators were proposed for cryptographic use in 1988. Major interest has been shown towards pseudorandom functions (PRF) recently, based on the Legendre and power residue symbols, due to their efficiency in the multi-party setting. The security of these PRFs is not known to be reducible to standard cryptographic assumptions. 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. We conduct algebraic cryptanalysis on the resulting MQ instance. We show that the currently known techniques and attacks fall short in solving these sparse quadratic equation systems. Furthermore, we build novel cryptographic applications of the Legendre PRF, e.g., verifiable random function and (verifiable) oblivious (programmable) PRFs.

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

✉ István András Seres
seresistvanandras@gmail.com

Máté Horváth
horvath@uni-wuppertal.de

Péter Burcsi
bupe@inf.elte.hu

[1] Eötvös Loránd University, Budapest, Hungary

[2] Bergische Universität Wuppertal, Wuppertal, Germany

Published online: 01 March 2023

1 Introduction

Zero-knowledge proofs (ZKP) and secure multi-party computation (MPC) protocols are ubiquitous in cryptography. These advanced cryptographic tools are applied and deployed in many applications, e.g., privacy-preserving cryptocurrencies, threshold cryptography and secure instant-messaging. The widespread adoption of ZKPs and MPC protocols necessitates novel symmetric-key primitives [43]. Traditional symmetric-key primitives, e.g., AES, cause significant overhead in ZKPs or MPC due to their vast multiplicative complexity.

Therefore, recently, revived interest has been shown towards algebraic symmetric key primitives with low multiplicative depth [43]. Lately, several novel algebraic MACs [22, 30], hash functions [6, 38] or algebraic pseudorandom functions [24] 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 avenues for attacks [1]. 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 [3] and MiMC hash functions [2], while Li and Preneel refined interpolation attacks on low algebraic degree cryptosystems [64]. 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 0 mod 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 (1918), and later Davenport (1931) cf. [25, 79]. 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 for any fixed n, n-grams of Legendre symbols are asymptotically equally distributed [70]. Mauduit and Sárközy [67] introduced several metrics to measure the pseudorandomness of binary sequences and argued that "Legendre symbol sequences are the most natural candidate for pseudorandomness". Ding et al. [29] confirmed the high linear complexity of Legendre symbol sequences. Tóth and Gyarmati et al. [39] introduced new pseudorandomness measures and asserted high values of those in Legendre symbol sequences [76].

1.1 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 [74, 78]. However, if the oracle can only be queried classically, then no efficient quantum algorithm is known. In concurrent and independent work, Frixons and Schrottenloher [35] 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 [54], Beullens et al. [8] and Kaluderovic et al. [56]. 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 [24]. 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 useful in establishing the security of various cryptographic applications derived from the Legendre PRF, e.g. the digital signature scheme by Beullens et al. [11].

1.2 Our contributions

In this work, we make the following contributions.

Legendre PRF as an MQ instance We show that key-recovery attacks on 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. These connections naturally extend to higher-degree Legendre PRFs and power residue symbol PRFs.

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. However, all these standard cryptanalytic tools from multivariate cryptography do not improve the state of the art key recovery attacks against the Legendre PRF [8, 54, 56]. On the other hand, we find that the induced MQ instances behave like random MQ instances in terms of degree of regularity, i.e., the corresponding ideals are semi-regular. This observation might be interpreted as evidence of the difficulty of breaking the Legendre PRF.

Novel cryptographic applications of the Legendre PRF Besides assessing the security of the Legendre PRF, we utilise its special properties to apply it in various

cryptographic tasks. Expressing the Legendre PRF as an MQ instance facilitates novel cryptographic applications, i.e., verifiable random functions. Moreover, we exploit its multiplicativity to construct (verifiable) oblivious (programmable) pseudorandom functions. Due to their efficiency, these novel extensions can be applied in several cryptographic protocols, such as state-of-the-art private set intersection (PSI) protocols.

1.3 Organisation

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

2 Preliminaries

2.1 Notations

Whenever we sample x from set S uniformly at random we write $x \in_R S$. Let p be an odd prime and let $K \in_R \mathbb{F}_p$ be a secret key. The modular square root algorithm mod p is denoted as $\mathsf{sqrt}_p(\cdot)$. 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 $\mathbb{F}_p$. $\mathsf{LT}(I)$ denotes the ideal generated by the leading terms of the ideal I. For the ease of exposition we use $[x]$ to denote a secret share of the value $x \in \mathbb{F}_p$.

2.2 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" [24].

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 {0,1}, i.e.,

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

Definition 3 (Weak Unpredictability) A pseudo-random bit-generator $\mathcal{X}_\lambda(s) : \{0,1\}^\lambda \to \{0,1\}^{l(\lambda)}$, where s is a seed and $l(\cdot)$ is an expansion factor, is next bit unpredictable (sometimes weakly unpredictable) if for all probabilistic polynomial time algorithm $\mathcal{A}$, there is a negligible function $\mathtt{negl}(\lambda)$ such that

$$\Pr[\mathcal{A}(x_1, x_2, \ldots, x_{l(\lambda)-1}) = x_{l(\lambda)}] \leq \frac{1}{2} + \mathtt{negl}(\lambda),$$

where the sequence $X = x_1 x_2 \ldots x_{l(\lambda)}$ is generated by $\mathcal{X}_\lambda(s)$ with $s \in_R \{0,1\}^\lambda$.

Assumptions. Grassi et al. formulated the following problem that underpins the security of the Legendre PRF [43].

Definition 4 (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.

Definition 5 (Multivariate Quadratic (MQ) problem) Given random quadratic polynomials over a finite field, i.e., $(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \in \mathbb{F}[x_1, \ldots, x_n]^m$, find a common zero $\mathbf{x} \in \mathbb{F}^n$ of the polynomials $f_1, \ldots, f_m$.

It is well-known that the MQ problem is **NP**-hard for any choice of finite field $\mathbb{F}$ [37]. 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.

2.3 NIZK arguments

Since in our VRF proposal we make use of non-interactive zero-knowledge (NIZK) arguments, we recall the relevant syntax following [12] and for the details and exact security requirements we refer to [12]. NIZK arguments consist of four PPT algorithms that are defined with respect to a relation generator algorithm $\mathcal{R}\text{-}\mathsf{Gen}(1^\lambda)$ that, upon receiving some security parameter λ, outputs a polynomial time decidable relation $\mathcal{R} : \{0,1\}^* \times \{0,1\}^*$ for which in our case $\{(\phi, \mathsf{w}) \in \mathcal{R} \mid \phi(\mathsf{w}) = 0\}$, where the statement ϕ is a MQ equation system over $\mathbb{F}_p$ and a valid witness w is a solution of the system.

- $\mathsf{NIZK.Setup}(\mathcal{R}) \to (\sigma, \tau)$. For the relation $\mathcal{R}$ the setup produces a common reference string σ and a simulation trapdoor τ.
- $\mathsf{NIZK.Prove}(\mathcal{R}, \sigma, \phi, \mathsf{w}) \to \pi$. Upon the $(\phi, \mathsf{w}) \in \mathcal{R}$ and the common reference string σ, the prover returns an argument π.
- $\mathsf{NIZK.Vfy}(\mathcal{R}, \sigma, \phi, \pi) \to \{0,1\}$. Upon the common reference string σ, the statement ϕ and an argument π the verification algorithm returns 0 or 1.
- $\mathsf{NIZK.Sim}(\mathcal{R}, \tau, \phi) \to \pi$. Using the simulation trapdoor τ and statement ϕ the simulator returns an argument π.

Definition 6 (Perfect NIZK argument [12]) We say that a NIZK is a perfect NIZK argument for $\mathcal{R}$ if it has perfect completeness, perfect zero-knowledge and computational soundness as defined in [12].

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 sequential Legendre PRF.

3.1 The ideal

Let us fix an arbitrary quadratic non-residue $r \in \mathbb{Z}_p^*$. Furthermore, it is assumed that we are given $\{a\}_K$, often $a \approx \log(p)$. 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 Legendre symbols, then we obtain $m = n - 1$ independent equations in n variables, hence the MQ instance is underdefined. Note, that the equation system is extremely sparse.

Example 1 We consider the following example to illustrate the quadratic equation system induced by the Legendre PRF. Let $p = \texttt{0xffffffffffffffffffffffdd}$ and $K = \texttt{0x27aaa97c746c22e12d10}$. The smallest quadratic non-residue modulo p is 2. We display the MQ instance induced by the evaluation of the sequential 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 Eqs. 1, 2 and 3. We want to solve simultaneously this equation system, i.e., finding points in the variety $V(I)$. If the sequence of Legendre symbols is long enough, heuristically $\mathcal{O}(\log p)$, then there are $\mathcal{O}(1)$ solutions in $\mathbb{F}_p$ (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 PRF. Note that $V(I)$ might contain additional solutions when considered above the algebraic closure $\overline{\mathbb{F}}_p$.

3.2 The Gröbner basis

To better understand the variety $V(I)$, first we describe the Gröbner basis of I [17]. 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 Eqs.* 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*.

Proof With a case distinction one can show that G generates I. For instance, if $b_i = b_j = b_{n-1} = 1$, then $g_i - g_j = f_i$. The other cases are similar. Thus $I \subset \langle G \rangle$.

By the 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., $\overline{S(g_i, g_j)}^G = 0$. This follows from Buchberger's product criterion but we include the following simple proof for completeness. 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.

G is reduced, since all of its basis polynomials have a leading coefficient one. Moreover, $\langle \mathsf{LT}(g_i) \rangle = \langle \mathsf{LT}(I) \rangle$ and no trailing term of any $g_i \in G$ lies in $\langle \mathsf{LT}(I) \rangle$. □

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

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

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, bivariate 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.

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 $m = n-1$ multivariate quadratic polynomials of I in n variables define a *regular ideal*, i.e., $V(I)$ is a 1-dimensional variety, namely, it contains an infinite number of solutions in $\overline{\mathbb{F}}_p$. The proof of the following lemma is in Appendix A.

Lemma 1 *I is a regular ideal*.

3.3 The field equations

As we have seen previously the corresponding variety $V(I)$ of the ideal I has dimension 1. However, in the cryptanalysis of the Legendre PRF, we wish to obtain a 0-dimensional variety that contains the secret key K of the PRF. As we show, this can be achieved by adding the field equations to the ideal I.

A sequence $\{n\}_K$ can be described with polynomials in $\mathbb{F}_p[x_0, x_1, \ldots, x_n]$. Let us define I_{FE} as follows:

$$I_{\mathsf{FE}} = I + \{x_i^p - x_i | i \in [0, n]\}. \tag{5}$$

Example 3 We illustrate the ideal I_{FE} complemented with the field equations with parameters $p = 191$ and $\{9\}_{45} = (1, 1, -1, 1, 1, 1, 1, 1, -1)$. The smallest quadratic non-residue is $r = 7 \bmod 191$.

$$\begin{aligned} I_{\mathsf{FE}} = \langle & -x_0^2 + x_1^2 - 1, -7x_1^2 + x_2^2 - 7, -x_2^2 + 7x_3^2 - 7, -x_3^2 + x_4^2 - 1, \\ & -x_4^2 + x_5^2 - 1, -x_5^2 + x_6^2 - 1, -x_6^2 + x_7^2 - 1, -7x_7^2 + x_8^2 - 7, \\ & x_0^{191} - x_0, x_1^{191} - x_1, x_2^{191} - x_2, x_3^{191} - x_3, x_4^{191} - x_4, \\ & x_5^{191} - x_5, x_6^{191} - x_6, x_7^{191} - x_7, x_8^{191} - x_8 \rangle. \end{aligned}$$

The corresponding Gröbner basis has the following form,

$$\langle x_0^2 - 45, x_1^2 - 46, x_2^2 + 53, x_3^2 - 48, x_4^2 - 49, x_5^2 - 50, x_6^2 - 51, x_7^2 - 52, x_8^2 + 11 \rangle.$$

Note how helpful the Gröbner bases are in obtaining the secret key K. In addition, one can also read off all the evaluated points from the Gröbner bases. If the variable x_i corresponds to a residue, then x_i^2 is one of the evaluated points in the PRF. Alternatively, if x_i corresponds to a non-residue, then $r^{-1}x_i^2 \bmod p$ is the evaluated point in the PRF.

Using the intuition of the Example 3, we can show in general the structure of the Gröbner basis of I_{FE}.

Theorem 2 *Let $\{n\}_K = (b_0, \ldots, b_{n-1})$ be a Legendre symbol sequence for which there exists a unique key K. We consider its corresponding ideal complemented with the field equations $I_{\mathsf{FE}} = \langle f_1, f_2, \ldots, f_m \rangle$, where $m = 2(n-1) + 1$ as defined by Eq. 5. Then the Gröbner basis of I_{FE} with respect to the (graded) lexicographic ordering, consists of the polynomials g_i, for $i \in [0, n-1]$ such that,*

$$g_i = \begin{cases} x_i^2 - (K + i), \text{if } b_i = 1 \\ x_i^2 - r(K + i), \text{if } b_i = -1 \end{cases} \tag{6}$$

Moreover, $G := (g_i)_{i=0}^{n-1}$ is a reduced Gröbner basis.

Proof G generates the ideal I_{FE}, since each f_i can be expressed by using the generators g_i. The generating polynomials f_i of the ideal I can be expressed as $f_i = r^{L_0(K+i+1)}g_{i+1} - r^{L_0(K+i)}g_i$. The field polynomials can be also expressed using the generators of G. Specifically, let us denote the modular square roots of $r^{L_0(K+i)}(K+i)$ as b and c. Then, $x_i^p - x_i = g_i\Pi_{a\neq b,c}(x-a)$. Hence, $I_{\mathsf{FE}} \subset \langle G\rangle$. By the uniqueness of K, we also have that $\langle G\rangle \subset I_{\mathsf{FE}}$, since the corresponding varieties are equal above the algebraic closure.

Next, we verify that the Buchberger-criterion holds for the polynomials in G. In this case, $S(g_i, g_j) = x_j^2 g_i - x_i^2 g_j$. Depending on the residuosity of b_i, b_j we have four cases, but for the sake of simplicity we only consider here the case of $b_i = b_j = 1$. The other cases follow similarly. The S-polynomial is divisible by G, since $S(g_i, g_j) = x_j^2(x_i^2 - (K+i)) - x_i^2(x_j^2 - (K+j)) = -(K+i)x_j^2 + (K+j)x_i^2 = (K+j)g_i - (K+i)g_j$, that is clearly divisible by the polynomials of G. G is clearly a reduced Gröbner basis as each leading coefficient is one and no monomial of g_i lies in $\langle \mathsf{LT}(G \setminus g_i)\rangle$. □

In Sect. 4, we evaluate empirically the time complexity of computing the Gröbner basis of MQ instances (the I_{FE} ideal) induced by Legendre PRF sequences. The ideal I_{FE} cannot be regular as it contains more polynomials than variables. However, the Gröbner basis of I_{FE} allows us to observe easily that in I_{FE} there are no internal dependencies between the ideal's generating polynomials. More precisely, we prove the following lemma in Appendix A.

Lemma 2 *I_{FE} is a semi-regular ideal, if the conditions of Theorem 2 are met.*

The asymptotic behavior of *the degree of regularity* of semi-regular ideals is well understood [13]. The degree of regularity d_{reg} of an ideal is a measure to assess the theoretical complexity of computing the Gröbner basis of an ideal. For a precise definition, the reader is referred to [21]. Finally, we show the usefulness of I_{FE} in connection with the Legendre PRF.

Lemma 3 *A successful Legendre key-recovery attack is equivalent in polynomial time to solving the MQ system defined by the ideal I_{FE}. 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.*

Proof Let us define the variety V and ideal I defined by the Legendre PRF evaluation $\{n\}_K$. More precisely, we fix a quadratic non-residue $r \in \mathbb{F}_p$. In polynomial-time, we construct $V^* = \{(x_0, x_1, \ldots, x_n) | x_i = \pm\mathsf{sqrt}_p(r^{L_K(i)}(K+i)), i \in [0, n-1]\}$. The corresponding ideal is denoted as I^*. We show that $V^* = V(I_{\mathsf{FE}})$. First, $V^* \subset V(I_{\mathsf{FE}})$, because this is how the polynomials in I_{FE} are constructed, such that all the points in V^* vanish on the polynomials of I_{FE}. The other inclusion is trivial by the construction of the polynomials of I_{FE}. I_{FE} is a radical ideal, since every ideal that contains its

field equations is a radical ideal [77, Lemma 2.2.3.]. Hence, I_{FE} is the smallest ideal that vanishes on V^*.

As for the unpredictability of the Legendre PRF, if the MQ system corresponding to a purported PRF evaluation is not solvable, then it is sure that the psuedorandom sequence is not obtained by evaluating the Legendre PRF. □

We highlight again the 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 specific structure (cf. Example 1, 3) stemming from the multiplicative group of $\mathbb{F}_p$. 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".

In Sect. 4, we turn our attention to assessing the security of the MQ instance induced by the Legendre PRF. In particular, we assess the complexity of solving the particular equation systems. According to [46], in order to prove the security of a multivariate 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 $\mathbb{F}_p^m$ [46].

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 [56].

4.1 Algebraic cryptanalytic attempts

4.1.1 Interpolation attacks

Interpolation attacks aim to interpolate a cryptosystem's polynomial without knowing its secret key [48]. 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 pseudorandom bit-string, where $a > 0$. In these cases, the resulting bit-string is mapped to integers, for instance, in the following way,

$$F_K(a) = \sum_{i=0}^{a-1} 2^{a-1-i}(K+i)^{\frac{p-1}{2}} \mod p \tag{7}$$

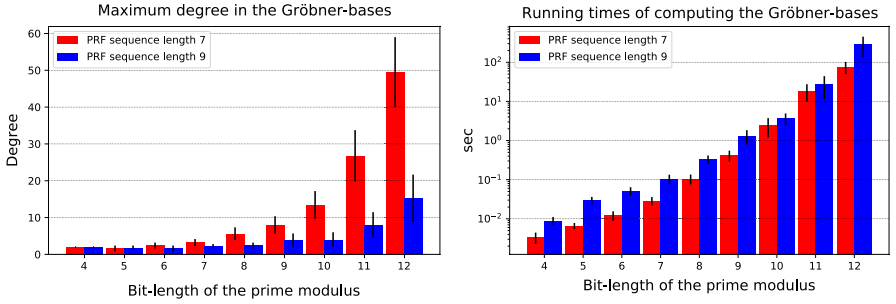


Fig. 1 The maximum degree in the Gröbner basis (left) and the exponential time complexity of computing the Gröbner bases (right) for the ideals I_{FE} defined by the Legendre PRF

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 pseudorandom field elements to interpolate $F_K(a)$.

4.1.2 Direct algebraic attacks

Direct algebraic attacks, i.e., computing the Gröbner basis [17], 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 [75], which in turn depends on the *degree of regularity*, d_{reg}, of the MQ instance at hand. Hence, it is of great interest to compute d_{reg} 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, where $2 \leq \omega \leq 3$ is the linear algebra constant of matrix multiplication.

We empirically evaluated the performance of computing the Gröbner basis for the ideal I_{FE} induced by the PRF evaluations, see Fig. 1. We sampled random small primes with a given bit-length and evaluated the Legendre PRF for a sequence of length seven and nine. We computed and recorded the time it takes to compute the Gröbner basis of the corresponding ideal I_{FE}. We repeated the experiment 10 times. We observe that computing the Gröbner basis takes exponential time in the bit-length of the prime modulus. We expect that launching key-recovery against the Legendre PRF using Gröbner basis methods is hopeless for cryptographic parameter sets, i.e., for primes of size $\approx 2^{128}$. Attaining lower and upper bounds for d_{reg} to

m	1	2	3	4	5	6	7	8	9	10
genus	0	1	1	5	17	49	129	321	769	1793

Fig. 2 The genus of the algebraic curves containing the solutions corresponding to a Legendre symbol sequence of length $m + 1$

assess the exact complexity of the Gröbner basis computation of I_{FE} is an interesting open problem.

4.1.3 MinRank attacks

The MinRank attack is a powerful tool in the cryptanalysis of multivariate cryptography. MinRank attacks broke numerous multivariate cryptosystems, such as the cryptanalysis of HFE due to Kipnis and Shamir [61] or the cryptanalysis of SRP encryption system [71]. 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.

4.2 Group atructure of the Legendre PRF MQ instances'solutions

We give an algebraic-geometric argument on the security of the Legendre PRF. In Sect. 3.1, we showed that the PRF seed lies in the intersection of multiple Pell-conics. The solutions of a single Pell-equation over $\mathbb{F}_p$ form a cyclic Abelian-group [26]. These groups were previously suggested for use in cryptography as it is believed that the discrete logarithm problem is hard in these groups [63]. A single Pell conic has genus 0. The intersection of two Pell-conics yields a nonsingular elliptic curve with genus 1. Specifically, 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. However, if one considers longer sequences, then the resulting curve has a genus greater than 1, cf. Fig. 2. Hence, the solutions of those algebraic curves *do not have an Abelian group structure equipped with them*. In the following, we compute the genus of the high-degree surfaces induced by the Legendre PRF in the general case.

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 $(f_1, f_2, \ldots, f_m)$ be the given Pell conics in variables $x_0, x_1, \ldots, x_n$ 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 homogeneous part of $\mathbb{F}_p[x_0, \ldots, x_n]/I$ [44]. 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,

$$P_n(N) = g_n(N) - \binom{n-1}{1} g_n(N-2) + \binom{n-1}{2} g_n(N-4) - \dots, \tag{8}$$

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 concreteness and as an 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 $n = 4$:

$$P_4(N) = \binom{N+4}{4} - 3\binom{N+2}{4} + 3\binom{N}{4} - \binom{N-2}{4}. \tag{9}$$

By substituting $N = 0$, we have that $P_4(0) = -4$, namely the arithmetic genus is $1 - P_4(0) = 5$. We obtain the following closed formula for the Hilbert-polynomial:

Lemma 4 $P_n(N) = 2^{(n-1)} \cdot N - (n-3) \cdot 2^{(n-2)}$.

Proof The proof is enclosed in Appendix 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 Sect. 5.1, oblivious pseudorandom functions (OPRF) in Sect. 5.2 and verifiable OPRF in Appendix E.

5.1 Verifiable random functions from the Legendre PRF

Verifiable random functions (VRFs) are natural extensions of PRFs [66]. In a VRF, the PRF evaluator can produce a publicly verifiable 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 verification key, 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 bits, e.g., proof-of-stake consensus algorithms [36].

An advantage of the Legendre PRF arithmetization as an MQ instance, is that it allows to model the PRF as a low-degree polynomial equation system. This 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. Building on this observation and using NIZK with the Legendre PRF (following the high-level approach sketched in [66]), we propose a new

VRF that admits post-quantum secure instantiations with comparable performance to the state of the art.

5.1.1 Syntax and security of VRFs

Definition 7 A VRF is comprised of the polynomial-time algorithms $\mathcal{VRF} = (\mathsf{VRF.PPGen}, \mathsf{VRF.Gen}, \mathsf{VRF.Eval}, \mathsf{VRF.Vfy})$ with the following functionality:

- $\mathsf{VRF.PPGen}(1^\lambda) \to \mathsf{pp}_{\mathsf{vrf}}$. Upon the security parameter λ, the algorithm samples the public parameters $\mathsf{pp}_{\mathsf{vrf}}$.
- $\mathsf{VRF.Gen}(\mathsf{pp}_{\mathsf{vrf}}) \to (\mathsf{sk}, \mathsf{vk})$. Upon $\mathsf{pp}_{\mathsf{vrf}}$, the algorithm samples secret and verification keys $(\mathsf{sk}, \mathsf{vk})$.
- $\mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, X) \to (Y, \pi)$. This algorithm evaluates a PRF $F : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^\lambda$ using the public parameters $\mathsf{pp}_{\mathsf{vrf}}$, secret key sk and PRF input X and outputs the PRF value Y and a proof of honest evaluation π.
- $\mathsf{VRF.Vfy}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{vk}, X, Y, \pi) \to \{0,1\}$. Upon the public parameters $\mathsf{pp}_{\mathsf{vrf}}$, verification key vk, PRF input–output pair X, Y and proof π, the verification algorithm either outputs 1 (accept) or 0 (reject).

Furthermore, the following requirements must hold:

1. *Correctness:* $\forall \lambda \in \mathbb{N}$, $\mathsf{pp}_{\mathsf{vrf}} \leftarrow\$ \mathsf{VRF.PPGen}(1^\lambda)$, input $X \in \{0,1\}^\lambda$, keys $(\mathsf{vk}, \mathsf{sk}) \leftarrow\$ \mathsf{VRF.Gen}(\mathsf{pp}_{\mathsf{vrf}})$, and $(Y, \pi) \leftarrow\$ \mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, X)$ it must hold that $\mathsf{VRF.Vfy}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{vk}, X, Y, \pi) = 1$.
2. $\boxed{\mathit{Trusted}^1}$[1] computational[2] unique provability: $\forall \lambda \in \mathbb{N}, X \in \{0,1\}^\lambda$ and PPT adversary $\mathcal{A}$, there exists a negligible function $\mathtt{negl}(\lambda)$ s.t.

$$\Pr\left[\begin{array}{l}\mathsf{VRF.Vfy}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{vk}, X, Y_0, \pi_0) = \\ \mathsf{VRF.Vfy}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{vk}, X, Y_1, \pi_1) = 1\end{array}\middle|\begin{array}{c}\boxed{\mathsf{pp}_{\mathsf{vrf}} \leftarrow\$ \mathsf{VRF.PPGen}(1^\lambda)} \\ (\mathsf{vk}, X, Y_0, Y_1, \pi_0, \pi_1) \leftarrow\$ \mathcal{A}(\mathsf{pp}_{\mathsf{vrf}}) \\ Y_0 \neq Y_1\end{array}\right] \leq \mathsf{negl}(\lambda) \tag{10}$$

3. *Pseudorandomness:* Let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be an attacker with oracle access to $\mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, \cdot)$ in the following pseudoramndomness game:

[1] Unique provability requires uniqueness to hold even when all the values are maliciously generated by the adversary. [73] proposed the relaxation of requiring uniqueness to hold only when some values are assumed to be generated honestly. While we use this approach, it is important to emphasize that we only assume that public system parameters ($\mathsf{pp}_{\mathsf{vrf}}$) are generated honestly, while e.g., [73] assumed this for the verification key that is a stronger assumption than ours.

[2] We say that the unique provability requirement holds unconditionally if the probability in the requirement is equal to zero even if $\mathcal{A}$ is not computationally bounded. The relaxation we use is due to [20] and it was first formulated by [41].

$\mathcal{G}_{\mathcal{A}}^{\mathcal{VRF}}(1^\lambda)$

$\mathsf{pp}_{\mathsf{vrf}} \leftarrow\$ \mathsf{VRF.PPGen}(1^\lambda)$

$(\mathsf{vk}, \mathsf{sk}) \leftarrow\$ \mathsf{VRF.Gen}(\mathsf{pp}_{\mathsf{vrf}}), \rho_{\mathcal{A}} \leftarrow\$ \{0,1\}^\lambda$

$(X^*, \mathsf{st}) \leftarrow\$ \mathcal{A}_1^{\mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, \cdot)}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{vk}, \rho_{\mathcal{A}})$

$(Y_0, \pi) := \mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, X^*)$

$Y_1 \leftarrow\$ \mathcal{Y}$

$b \leftarrow\$ \{0,1\}$

$b' := \mathcal{A}_2^{\mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, \cdot)}(Y_b, \mathsf{st})$

return $b == b'$

Denoting the oracle queries of $\mathcal{A}$in the game with $\mathcal{Q} = (X_1, \ldots, X_Q)$, we say that $\mathcal{A}$is legitimate if for any random coin choices $\rho_{\mathcal{A}} \in \{0,1\}^\lambda$ of $\mathcal{A}$, there exists no $i \in [Q]$ for which $X_i = X^*$ would hold. We say that a $\mathcal{VRF}$is pseudorandom, if for all legitimate $\mathcal{A}$, its advantage in game $\mathcal{G}_{\mathcal{A}}^{\mathcal{VRF}}(1^\lambda)$ is at most negligible, i.e., $\left|\Pr\mathcal{G}_{\mathcal{A}}^{\mathcal{VRF}}(1^\lambda) = 1 - \frac{1}{2}\right| \leq \mathtt{negl}(\lambda)$.

5.1.2 Construction

We proceed with the construction of the Legendre VRF.

Intuition We face two challenges in creating a Legendre VRF. First, we need a verification key vk. For $\mathsf{sk} = K \in_R \mathbb{F}_p$, we let $\mathsf{vk} = \{c \cdot \log p\}_K$. Heuristic arguments imply that a long enough symbol sequence is unique if its length is roughly $\log p$ [70]. Hence, a unique symbol sequence acts as a "commitment" to sk. Second, we need to verify efficiently the correct evaluation of the Legendre PRF. We can leverage NIZK argument systems, since we can express the correct PRF evaluation statement as a low-degree polynomial equation system.

- $\mathsf{VRF.PPGen}(1^\lambda) \to \mathsf{pp}_{\mathsf{vrf}}$. On receiving the security parameter 1^λ, the public parameter generation algorithm runs $(\mathcal{R}, \mathsf{aux}) \leftarrow \mathcal{R}\text{-Gen}$ and $(\sigma, \tau) \leftarrow \mathsf{NIZK.Setup}(\mathcal{R})$ and output $\mathsf{pp}_{\mathsf{vrf}} = (\sigma, \mathcal{R})$.
- $\mathsf{VRF.Gen}(\mathsf{pp}_{\mathsf{vrf}}) \to (\mathsf{vk}, \mathsf{sk})$. Using the public parameters $\mathsf{pp}_{\mathsf{vrf}}$, the key generation algorithm samples random $\mathsf{sk} = K \in_R \mathbb{F}_p$, compute the Legendre sequence $\mathsf{vk} := \{c \cdot \log p\}_K$ that serves as a "commitment" to K (for a fixed constant c).
- $\mathsf{VRF.Eval}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{sk}, X) \to (Y, \pi)$. The evaluation of the VRF takes the public parameters $\mathsf{pp}_{\mathsf{vrf}}$, the secret key $\mathsf{sk} = K$ and an input X to the PRF. Let Y be λ consecutive Legendre symbols, i.e., $Y = \{\lambda\}_{K+X\lambda}$, so that for all X we evaluate the symbol on disjoint intervals (we constrain $X \leq p/\lambda$). Disjointness is used to ensure the pseudorandomness of the VRF, see the proof in Appendix D. Let $\pi \leftarrow \mathsf{NIZK.Prove}(\mathcal{R}, \sigma, \phi, \mathsf{w})$, where the witness $\mathsf{w} = \mathsf{sk}$ and ϕ corresponds to a MQ equation system that consists of

  - quadratic equations corresponding to the evaluation of the Legendre PRF as defined in Sect. 3.1,
  - similar equations showing the relation of sk and $\mathsf{sk} + X\lambda$, i.e., the ith bits of vk and Y correspond to Legendre symbols of values with distance $X\lambda$. For instance, in case of two quadratic residues, we have $x_i^2 - x_{\mathsf{vk}_i}^2 = X\lambda$, cf. Equation 1. The equations corresponding to the other cases can be similarly adapted from the quadratic equations of Sect. 3.1.

  The algorithm outputs (Y, π).
- $\mathsf{VRF.Vfy}(\mathsf{pp}_{\mathsf{vrf}}, \mathsf{vk}, X, Y, \pi) \rightarrow \{0, 1\}$. On receiving the public parameters $\mathsf{pp}_{\mathsf{vrf}} = (\mathcal{R}, \sigma)$, verification key vk, a VRF input–output pair X, Y with a proof π, the verification algorithm first determines ϕ based on vk, X, Y, and $|Y| = n$, then runs $\mathsf{NIZK.Vfy}(\mathcal{R}, \sigma, \phi, \pi)$ and returns its output.

The following theorem, which we prove in Appendix D, formalizes the security of the Legendre VRF.

Theorem 3 *Assuming the hardness of the SLS problem* (*Definition 1*) *the Legendre VRF is secure according to Definition 7, if the underlying NIZK argument fulfils the perfect completeness, perfect zero-knowledge and computational soundness requirements* (*defined in* [12]).

5.1.3 Instantiations and performance

We instantiate our VRF with the state of the art succinct NIZK [42]. However, it does not provide post-quantum security. Another proof system family of zero-knowledge succinct transparent arguments of knowledge (zkSTARK) was pioneered by the work of Ben-Sasson et al. [16]. STARK proof systems provide post-quantum security and does not rely on trusted setups. The performance evaluation of [16] 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 $\approx$ 50KB. An even more efficient VRF instantiation can be obtained by applying the NIZK of Beullens and Delpech de Saint [11]. In Table 1, we compare the proposed VRF to the state of the art. The Legendre VRF is a potential contender for being the most efficient post-quantum secure VRF in terms of proof size, prover and verifier complexity.

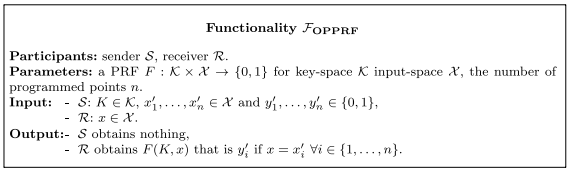
5.2 Oblivious PRFs from the Legendre PRF

An oblivious PRF (OPRF) [34, 68] 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

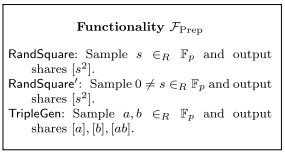
Table 1 Overview of various VRF constructions

	Proof size		Time complexity		
	$\|\pi\|$	($\lambda = 128$)	Prove	Verify	Assumption
[40]	$1\mathbb{G}$	0.34KB	$1\mathsf{H} + 1\mathbb{G}$	$1\mathsf{H} + 1\mathbb{G}$	**Factoring**
[73]	$1\mathbb{G} + 2\mathbb{F}_p$	768 bits	$3\mathsf{H} + 2\mathbb{G}$	$3\mathsf{H} + 4\mathbb{G}$	**EC-DDH**
[15]	$1\mathbb{G}$	377 bits	$2\mathsf{H} + 1\mathbb{G}$	$1P$	**co-DH**
[32]	$1\mathbb{G}$	377 bits	$1\mathbb{G} + 1\mathbb{F}_p$	$2\mathbb{G} + 2P$	**q-DBDHI**
[62]	$1\mathbb{G}$	377 bits	$1\mathbb{G}$	$1P$	**q-DDHE**
[33]†	$\mathcal{O}(k + l)$	5KB	$\mathcal{O}(kl)$	$\mathcal{O}(kl)$	*Module-SIS*
[10] (SL-VRF)	$\widetilde{\mathcal{O}}(\|\mathcal{C}\|)$	40KB	$\mathcal{O}(\|\mathcal{C}\|)$	$\mathcal{O}(\|\mathcal{C}\|)$	*LowMC*, ROM
§5.1+SNARK	$3\mathbb{G}$	209 bytes	$9n\mathbb{G}$	$n\mathbb{G} + 3P$	*SLS*, **KEA**
§5.1+STARK	$\mathcal{O}(\log(n))\mathbb{G}$	≈ 50KB	$\mathcal{O}(n \log(n))\mathbb{G}$	$\mathcal{O}(\log(n))\mathbb{G}$	*SLS*, ROM
§5.1+ [11]	$\mathcal{O}(n)$	≈ 30KB	$\mathcal{O}(n)$	$\mathcal{O}(\lambda)$	*SLS*, ROM

Hashing, group operations, exponentiation and pairings are denoted as $\mathsf{H}, \mathbb{G}, \mathbb{F}_p, P$, respectively. Note that [33] only provides a few-time VRF. Module-SIS and module-LWE ranks are denoted as k and l, respectively. $|\mathcal{C}|$ denotes the number of AND gates of the LowMC [7] PRF applied in [10]. Here n is the length of the Legendre symbol sequence being proved. Assumptions written in textit are post-quantum secure, while those written in textbf are not



(a) The ideal OPRF functionality. Together with the extensions in blue, we get the OPPRF ideal functionality.

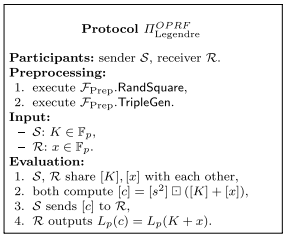


(b) Ideal preprocessing functionality.

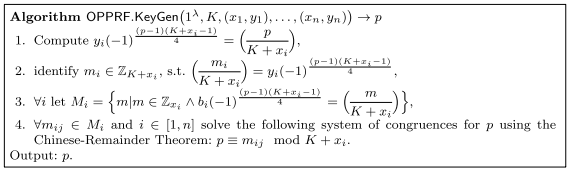
Fig. 3 Ideal functionalities

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. In this section, we show how to build an OPRF relying on the hardness of the SLS problem and also extend this result to two variants of OPRFs, namely to programmable and to verifiable OPRFs (denoted as OPPRF and VOPRF respectively).

These protocols are extensively used in various tasks. A non-exhaustive list of OPRF applications include secure keyword search [34], private set intersection (PSI) [45, 52, 57, 58], secure deduplicated storage [53], password-protected secret sharing [50], password-authenticated key exchange [51]. OPPRFs were used to build two-party PSI [55, 72], multi-party PSI [59] and circuit-PSI that enables secure function evaluation on the intersection of sets [18]. Finally, VOPRF is the cornerstone of Privacy Pass, a privacy-preserving lightweight authentication mechanism [28] and password-protected secret sharing [49]. The importance of (V)OPRF is also indicated by the ongoing effort to standardize them [27].



(a) Legendre OPRF based on [GRR+16].



(b) Programming the Legendre OPRF of Figure 3a 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 O**P**PRF

5.2.1 The Legendre OPRF

Motivated by the wide range of applications, our goal is to present a novel pathway to the realization of OPRFs that we formally define in Fig. 3.

We observe that the distributed protocol for evaluating the Legendre PRF of [43] yields an OPRF. For completeness, we include their protocol presented in the language of OPRFs. The key ingredient—that was used in [43] for the secure computation of the Legendre PRF in the multi-party setting—is that the key of the PRF can be masked without changing the PRF value by utilizing the multiplicative property of the Legendre symbol. Namely, if we choose a random square and multiply it with some number, the Legendre symbol of the resulting value will be equal to the symbol of the original number. This fact gives rise to the arithmetic sharing-based[3] OPRF protocol $\Pi^{OPRF}_{\text{Legendre}}$, depicted in Fig. 4. The protocol is divided into online and offline parts. In an offline preprocessing phase the parties can compute the shares of the previously mentioned random square and a so-called Beaver multiplication triple $[a]$, $[b]$, $[ab]$ (for some random a, b) both of which operations are entirely independent of the inputs of the participants. 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 Fig. 3. The realization of $\mathcal{F}_{\text{Prep}}$ is possible using a 2PC framework in the semi-honest model, such as ABY by [31].

After exchanging secret shares of their inputs, both participants execute the same computation on their shares in the online phase. While the addition of secret shares is for free, i.e., corresponds to ordinary local addition, share multiplication, which we denote with $\boxdot$, consumes one multiplication triple and requires one round of interaction and 2 group elements of communication. Concretely, $[x] \boxdot [y] = [xy]$ 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. The

[3] We denote secret shares in square brackets, i.e., $[x]_1 = r \in_R \mathbb{F}_p$ and $[x]_2 = x - r$ so $[x]_1 + [x]_2 = x$. For simplicity, we omit the lower indices denoting the owner of the given secret share, when this does not cause confusion.

Table 2 Comparing the online costs of various Oblivious PRF protocols

OPRF	Comm. complexity			Comp. complexity		Model	Assumption
	Rounds	Msg. aize	Concr. eff.	Client	Server		
RSA-OPRF	2	2 $\mathbb{G}$	0.77KB	1H + 2 $\mathbb{G}$	1 $\mathbb{G}$	ROM	**1-more-RSA-inv**
[49]	2	2 $\mathbb{G}$	64 byte	1H + 2 $\mathbb{G}$	1 $\mathbb{G}$	ROM/Standard	**EC-DDH**
[57]†	5	2λ bits	256 bits	1H + 2XOR	2H + 2XOR	ROM	*OT**
[4]	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*
Fig. 4	3	5λ $\mathbb{G}$	13.44KB	17λ $\mathbb{G}$	17λ $\mathbb{G}$	ROM	*SLS*, *OT**

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 [57] is only a relaxed PRF. 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 bold, while assumptions written in italics are secure even against quantum attackers

resulting online part then consists of three rounds of interaction and 5 group elements of communication.

Theorem 4 *The protocol* $\Pi^{OPRF}_{\text{Legendre}}$ *securely computes the functionality* $\mathcal{F}_{\text{OPRF}}$ *in the* $\mathcal{F}_{\text{Prep}}$*-hybrid model, if the SLS problem is hard.*

For brevity, we omit the proof since it follows the blueprint of the proof of [43, Theorem 2.]. We note that $\Pi^{OPRF}_{\text{Legendre}}$ is only statistically correct as with probability $1/p = \Pr(s^2 = 0)$ the output is necessarily zero. For perfect correctness, we need to use RandSquare′ in the preprocessing phase to rule out $s^2 = 0$ the cost of which appears in the round complexity, resulting in *expected* constant (one) round. Our efficiency comparisons in Table 2 show that in terms of both message size and computational complexity, the Legendre OPRF is a promising candidate for a post-quantum OPRF since the underlying SLS problem is not known to be vulnerable to post-quantum attacks.

5.2.2 OPPRF: programming the Legendre OPRF

The notion of oblivious *programmable* PRF (O**P**PRF) was introduced by Kolesnikov et al. [59]. A PRF is an 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 (see Fig. 3). Kolesnikov et al. [59] formulated three *generic* O**P**PRF constructions, that can turn any OPRF into an O**P**PRF. We follow the terminology of these generic constructions and introduce two algorithms that aims to turn an OPRF into an O**P**PRF:

- $\mathsf{OPPRF.KeyGen}(1^\lambda, \mathcal{P}) \to (K, \mathsf{hint})$: 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.
- $\mathsf{OPPRF.Eval}\big(F(K, x), \mathsf{hint}\big) \to y$: Using the hint turns the OPRF output into the O**P**PRF output y.

We require from an O**P**PRF the following high-level security notions to hold (for the formal security definitions, the reader is referred to [59]):

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

(n, t)-security: No efficient adversary is 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 $\mathsf{OPPRF.KeyGen}$) are pseudorandom.

Programming the Legendre OPRF We show how one can program efficiently the output of the Legendre PRF by carefully choosing the prime modulus, which defines our $\mathsf{OPPRF.KeyGen}$ algorithm. This strategy already highlights the strength of the resulting O**P**PRF: it does not require an explicit hint beyond the prime modulus that is a public parameter anyway. Moreover, the $\mathsf{OPPRF.Eval}$ algorithm can simply return the output of the Legendre OPRF.

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 for a given key K. 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, we take a different approach, cf. Figure 4. The goal of the algorithm is to find a prime p, such that

$$i \in [0, n) \,:\, y_i = \left(\frac{K + x_i}{p}\right) = \left(\frac{p}{K + x_i}\right)(-1)^{\frac{(p-1)(K+x_i-1)}{4}}.$$

Without loss of generality, we search p in the form $p \equiv 1 \mod 4$. Moreover, we assume that the programmed points $K + x_i$ are prime numbers. This assumption is natural and eases our exposition. This is because programming the PRF output at a composite $K + x_i$ is reducible to programming the PRF output at the prime factors of $K + x_i$ due to the multiplicativity of the Legendre symbol. For each $K + x_i$ the value $\left(\frac{p}{K + x_i}\right)$ establishes possible residue classes for $p \bmod K + x_i$. The appropriate modulus p can be obtained via the Chinese remainder theorem. Therefore, the

Table 3 Comparison of the generic O**P**PRF constructions of [59] (which can be based on an OPRF, e.g. that of [57]) and the Legendre OPRF that was shown to be programmable in Sect. 5.2.2

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 (Fig. 4)	$O(n \log n)$	1	$\mathcal{O}(n)$ $\mathbb{G}$	Depends on λ	Any
Legendre brute-force	$O(2^n)$	1	1 $\mathbb{G}$	Time-efficiency	Any

The number of programmed input positions is denoted as n, λ_{BF} is the soundness parameter of the Bloom filter, and k denotes the number of base-OTs, typically $k \approx 4\lambda$

"programmability" of the Legendre PRF is rather space-inefficient, since $p \approx \prod_{i=1}^{n} K + x_i$. Hence, the number of programmed points is somewhat limited with our algorithm. We note that the main ideas of this programming method were already proposed in a different context (secure comparison protocols) by Yu [80]. In a similar fashion, one could generalize the approach of Fig. 4 to power residue symbols, i.e., programming power residue symbol PRFs. Such generalization was shown recently by Cascudo et al. [23] who proposed as an open question to find concrete applications for their protocol. We note that their methods can be applied to program power residue symbol OPRFs.

Hint size and batch OPPRFs As our novel programming methods—specifically designed for the Legendre OPRF—minimize the necessary auxiliary information for the O**P**PRF evaluation, it outperforms all existing solutions in this metric. For a detailed comparison, we refer to Table 3. Finally, we note that [72] uses a so-called "Batch O**P**PRF" that—informally—invokes independent O**P**PRF instances with a total number of programmed points σ (the number of programmed points per instance may vary but has to remain hidden) and only uses a single hint with size linear in σ. Since the hint size of the Legendre O**P**PRF is independent of the number of programmed points, it naturally fulfils the requirement of Batch O**P**PRFs.

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 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 or key-homomorphic PRFs from quadratic and power residue symbol PRFs.

A Proofs from Sect. 3

Lemma 5 *I is a regular ideal.*

Proof Let $I = \langle f_1, \dots, f_m \rangle$ 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 $(f_1, \dots, f_m)$ being regular, we need to show that if for all $i \in [1, m]$ and g such that $gf_i \in \langle f_1, \dots, f_{i-1} \rangle$, then $g \in \langle f_1, \dots, f_{i-1} \rangle$. An affine sequence of polynomials $(f_1, \dots, f_m)$ is regular by definition, if the homogeneous sequence $(f_1^h, \dots, f_m^h)$ 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 $(f_1^h, f_2^h, \dots, f_m^h) = (x_1^2, x_2^2, \dots, x_m^2)$.

Since $f_i^h = x_i^2$, in our case for every i, therefore the ideal $I_{i-1} := \langle f_1^h, \dots, f_{i-1}^h \rangle$ 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 [21]. 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. □

Lemma 6 *I_{FE} is a semi-regular ideal, if the conditions of Theorem 2 are met.*

Proof The proof's blueprint is the same as that of Lemma 1. We consider the generating set for I_{FE} provided by the Gröbner basis, i.e., $I_{\mathsf{FE}} = (f_1, \dots, f_m)$. By definition, a homogeneous sequence of polynomials $(f_1, \dots, f_m)$ is semi-regular if for all $i = 1, \dots, m$ and g such that $gf_i \in \langle f_1, \dots, f_{i-1} \rangle \wedge deg(gf_i) < d_{reg}$ then g is also in $\langle f_1, \dots, f_{i-1} \rangle$. An affine sequence of polynomials $(f_1, \dots, f_m)$ is semi-regular if the sequence $(f_1^h, \dots, f_m^h)$ is semi-regular, where f_i^h is the homogeneous part of f_i of highest degree. In our case $(f_1^h, \dots, f_m^h)) = (x_1^2, \dots, x_m^2)$. Previously in the proof of Lemma 1, we saw why $(x_1^2, \dots, x_m^2)$ forms a regular ideal. □

B Adding more polynomials to the ideal of the PRF

As we have seen in Sect. 3.3, the Legendre key-recovery attack is equivalent to solving an overtedermined MQ instance. However, when $p \equiv 3 \mod 4$ or $p \equiv 5 \mod 8$, we might decrease the complexity of solving the resulting MQ instance by adding new equations. Observe that in these cases, we can express the modular square roots as follows:

$$\mathsf{sqrt}_p(x) : y = \begin{cases} \pm x^{\frac{p+1}{4}} \mod p, \text{ if } p \equiv 3 \mod 4 \\ \pm x(2x)^{\frac{p-5}{8}}(4x^{\frac{p-1}{4}} - 1) \mod p, \text{ if } p \equiv 5 \mod 8. \end{cases} \tag{11}$$

If $p \equiv 1 \mod 8$, it is not possible to express easily the $\mathsf{sqrt}_p(\cdot)$ algorithm 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 $x_i x_j$:

$$x_i x_j = \mathsf{sqrt}_p(x_i^2 x_j^2). \tag{12}$$

Similarly, we can add new polynomials to the system involving the linear terms of the unknowns for every $i \neq j$,

$$x_i = \mathsf{sqrt}_p(r^{L_0(x_i)-L_0(x_j)}(x_j^2 - r^{L_0(x_j)}(j-i))). \tag{13}$$

All polynomials in Eqs. 12 and 13 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 $k \approx \log^2 p$. We leave it as future work to analyze the independence of the newly introduced polynomials of Eqs. 12 and 13 from the polynomials of the ideal I_{FE}. We suspect that adding these high-degree polynomials to the ideal does not significantly speed up the Gröbner basis computation. Hence, these new polynomials might not have cryptanalytic relevance.

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 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.) [61, 71].

Definition 8 (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 $\mathcal{Q}(X_0, \ldots, X_{n-1}) = \phi \circ \mathbf{f} \circ \phi^{-1}(X, X^q, \ldots, X^{q^{n-1}})$, 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 8) of the Legendre PRF equation system [69]. 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,

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

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 Eq. 16. We define $\gamma = c_1 + c_2 y + \cdots + c_n y^{n-1}$ 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, M has full rank, since it is the product of non-singular matrices.

C.2 Group Structure of the Legendre PRF MQ Instances' Solutions

Lemma 7 $P_n(N) = 2^{(n-1)} \cdot N - (n-3) \cdot 2^{(n-2)}$.

Proof We first determine the linear coefficient by considering the difference polynomial $Q_n(N) = P_n(N+1) - P_n(N)$, which is a constant by the linearity of P_n. Using the inclusion–exclusion argument again, we see that $Q_n(N)$ is also a Hilbert-polynomial. To obtain an ideal with $Q_n(N)$ as its Hilbert polynomial, take an $(n-1)$-variable ring and $n-1$ polynomials, each of which is quadratic in a distinct single variable. The ideal generated by these polynomials is zero-dimensional, and therefore has a constant Hilbert polynomial whose value is the size of the corresponding variety,

i.e., 2^{n-1}. For the constant term, first note that for any real value of x, $\binom{x}{n} = (-1)^n \binom{-x+n-1}{n}$. Therefore, by substituting $N = (n-3)/2$ into (8), the terms $g_n(N-2k)\binom{n-1}{k}$ and $g_n(N-2(n-k))\binom{n-1}{n-k}$ cancel, and the middle term (for odd n) is 0, hence $P_n(n-3/2) = 0$, which gives the constant term. □

D Proof of Theorem 3

Next, we sketch the security proof of the Legendre VRF.

Proof To prove the theorem, we show that the requirements of Definition 7 are fulfilled by the Legendre VRF. *Correctness* directly follows from the prefect correctness of NIZK. To see that *pseudorandomness* holds, notice that game $\mathcal{G}_{\mathcal{A}}^{\mathcal{VRF}}(1^\lambda)$ is indistinguishable from the pseudorandomness game for PRFs as long as the honestly evaluated π in the answers for $\mathcal{A}$'s evaluation queries can be substituted by simulated π. Indeed, the game knows $\tau, \mathcal{R}, \phi$ for such simulation. Since the perfect zero-knowledge property of NIZK guarantees this, the proposed VRF is pseudorandom if the Legendre PRF is pseudorandom, i.e., assuming the hardness of the SLS problem.

We prove *trusted computational unique provability* indirectly. Therefore, let us assume that there exists a PPT $\mathcal{A}$ for which the probability in Eq. (10) is greater than $\mathtt{negl}(\lambda)$. As the values vk, X, Y determine ϕ, it follows that for the above $\mathcal{A}$

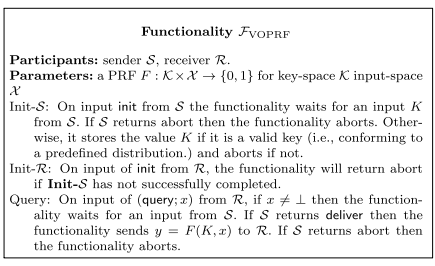
$$\Pr\left[\begin{array}{l}\mathsf{NIZK.Vfy}(\mathcal{R},\sigma,\phi_0,\pi_0) = \\ \mathsf{NIZK.Vfy}(\mathcal{R},\sigma,\phi_1,\pi_1) = 1\end{array}\middle|\begin{array}{c}\mathsf{pp_{vrf}} \leftarrow\$\ \mathsf{VRF.PPGen}(1^\lambda) \\ (\mathsf{vk}, X, Y_0, Y_1, \pi_0, \pi_1) \leftarrow\$\ \mathcal{A}(\mathsf{pp_{vrf}}) \\ Y_0 \neq Y_1\end{array}\right] > \mathsf{negl}(\lambda).$$

This only holds if either $\mathsf{NIZK.Vfy}$ accepts false statements with non-negligible probability or both statements are true. As the first option would contradict with the assumed computational soundness of NIZK, both statements has to be true, i.e., $(\phi_0, \mathsf{w}_0), (\phi_1, \mathsf{w}_1) \in \mathcal{R}$. Two Legendre sequences of the same length are equal if their starting points are equal, so $Y_0 \neq Y_1 \implies K_0 + X\lambda \neq K_1 + X\lambda \implies \mathsf{sk}_0 = K_0 \neq K_1 = \mathsf{sk}_1$. However, both statements ϕ_0 and ϕ_1 ensures that $\{c \cdot \log p\}_{K_0} = \{c \cdot \log p\}_{K_1} = \mathsf{vk}$ implying that the values of these different Legendre sequences must collide with non-negligible probability. This is contradiction since we know from [70] that the probability of such collision is $1/2^{c \log p} = 1/2^{c\lambda} < \mathtt{negl}(\lambda)$. □

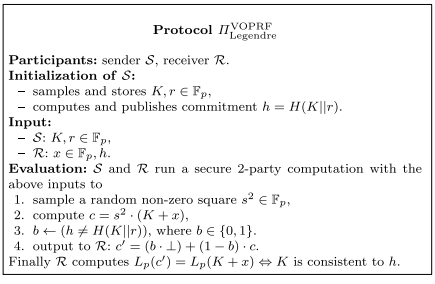
E The Legendre verifiable OPRF

In Sect. 5.2, we built an OPRF relying on semi-honest 2PC that clearly cannot prevent the participants from deviating the protocol. What is even more problematic in practice is that sometimes the server is supposed to behave consistently in multiple OPRF evaluations, namely, it is assumed to use the same key. To check this on the receiver side—without obtaining information about the key—active security alone is not enough, but in an initialization phase the sender has to commit to the key(s) it wishes to use. Such commitments can then be published (as a "public key") to enable the receiver the verification of whether distinct OPRF evaluations happened under the same or different keys. OPRF protocols that guarantee such verifiability are called verifiable OPRFs (VOPRFs). In Fig. 5, we recall the ideal functionality as defined in [5], for the precise security definition we also refer to this work. We note that different formalizations of VOPRF exist, e.g. [49] considered in the concurrent setting when defining the universal composable VOPRF.

Turning our attention to the realization, it seems obvious that special purpose protocols beat general ones in all efficiency metrics. Indeed, known realizations [5, 27, 49] try to avoid generic tools such as 2PC that leads to efficient solutions in case of constructions using pre-quantum assumptions but not when aiming protocols that offer post-quantum security. Besides their theoretical post-quantum solutions, Albrecht et al. [5] mention an alternative pathway towards post-quantum VOPRFs that has comparable efficiency with their lattice-based solutions. This solution consists of a hash commitment to a key K, and an actively secure MPC evaluation of the AES circuit on inputs K and x (from $\mathcal{S}$ and $\mathcal{R}$ respectively) together with comparison of the hash of the used key with the committed key, after which $\mathcal{R}$ receives output iff the check goes through. At this point, one may recall the Legendre OPRF of Fig. 4 that requires a single multiplication in the online phase for one bit output (or 128 multiplications for 128 bits). This is in contrast to the 960 multiplication of the AES circuit evaluation [43]. This observation motivates our Legendre VOPRF protocol, that is described in details in Fig. 5.



(a) Ideal functionality for VOPRF adapted from [ADDS21].



(b) Legendre VOPRF based on actively secure 2PC and collision-resistant hash H.

Fig. 5 Legendre VOPRF

Theorem 5 (Informal) *When instantiated with actively secure 2PC, protocol* $\Pi^{\text{VOPRF}}_{\text{Legendre}}$ *securely realizes* $\mathcal{F}_{\text{VOPRF}}$ *under the SLS assumption and the assumptions which the 2PC protocol relies on and if H is a collision-resistant hash.*

The generality of the utilized 2PC protocol leads to various instantiation opportunities causing that the above result can have several different flavours. We mention some of these. [60] showed that actively secure 2PC in the standard model requires 5 rounds of interaction. With some relaxations, namely by allowing the simulator to run in superpolynomial time while the adversary is still restricted to polynomial time (a.k.a. SPS security), actively secure non-interactive secure computation (NIZK) is possible in the plain model under the subexponential security of the LWE assumption [9, 14] leading to a VOPRF realization under the same assumptions. Leaving the plain model, it is also possible to instantiate our VOPRF utilizing NIZK built on oblivious transfer (OT) in the OT-hybrid model [47], in the common reference string model [65] or in the global random oracle model [19].

Acknowledgements We are grateful for the insightful conversations to Gergő Zábrádi. The first and the third author were supported by the Ministry of Innovation and Technology and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary.

Funding Open access funding provided by Eötvös Loránd University.

References

1. Aly, A., Ashur, T., Ben-Sasson, E., Dhooghe, S., Szepieniec, A.: Design of symmetric-key primitives for advanced cryptographic protocols. IACR Trans. Symmetric Cryptol. **2020**, 1–45 (2020). https://www.iacr.org/cryptodb/data/paper.php?pubkey=30562
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: ASIACRYPT (3), vol. 11923 of Lecture Notes in Computer Science, pp. 371–397. Springer (2019)
3. Ashur, T., Dhooghe, S.: Marvellous: a stark-friendly family of cryptographic primitives. IACR Cryptol. ePrint Arch. **2018**, 1098 (2018)
4. 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)
5. Albrecht, M.R., Davidson, A., Deo, A., Smart, N.P.: Round-optimal verifiable oblivious pseudorandom functions from ideal lattices. In: Public Key Cryptography (2), vol. 12711, Lecture Notes in Computer Science, pp. 261–289. Springer (2021)

6. 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)
7. Albrecht, M.R., Rechberger, C., Schneider, T., Tiessen, T., Zohner, M.: Ciphers for MPC and FHE. In: EUROCRYPT (1), vol. 9056, Lecture Notes in Computer Science, pp. 430–454. Springer (2015)
8. Beullens, W., Beyne, T., Udovenko, A., Vitto, G.: Cryptanalysis of the Legendre PRF and generalizations. IACR Trans. Symmetric Cryptol. **2020**, 313–330 (2020)
9. Brakerski, Z., Döttling, N.: Two-message statistically sender-private OT from LWE. In: TCC (2), vol. 11240, Lecture Notes in Computer Science, pp. 370–390. Springer (2018)
10. Buser, M., Dowsley, R., Esgin, M.F., Kermanshahi, S.K., Kuchta, V., Liu, J.K., Phan, R., Zhang, Z.: Post-quantum verifiable random function from symmetric primitives in pos blockchain. IACR Cryptol. ePrint Arch. 302 (2021)
11. 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)
12. Blum, M., Feldman, P., Micali, S.: Non-interactive zero-knowledge and its applications. In: Providing sound foundations for cryptography: on the work of Shafi Goldwasser and Silvio Micali, pp. 329–349. (2019)
13. Bardet, M., Faugere, J.-C., Salvy, B., Yang, B.-Y.: Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In: Proceedings of MEGA, vol. 5 (2005)
14. Badrinarayanan, S., Garg, S., Ishai, Y., Sahai, A., Wadia, A.: Two-message witness indistinguishability and secure computation in the plain model from new assumptions. In: ASIACRYPT (3), vol. 10626, Lecture Notes in Computer Science, pp. 275–303. Springer (2017)
15. 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)
16. Ben-Sasson, E, Bentov, I, Horesh, Y, Riabzev, M: Scalable, transparent, and post-quantum secure computational integrity. IACR Cryptol. ePrint Arch. **2018**, 46 (2018)
17. Buchberger, B.: Ein algorithmus zum auffinden der basiselemente des restklassenringes nach einem nulldimensionalen polynomideal. PhD thesis, Universitat Insbruck (1965)
18. Chandran, N., Gupta, D., Shah, A.: Circuit-PSI with linear complexity via relaxed batch opprf. In: 22nd Privacy Enhancing Technologies Symposium (PETS 2022), June (2022)
19. Canetti, R., Jain, A., Scafuro, A.: Practical UC security with a global random oracle. In: CCS, pp. 597–608. ACM (2014)
20. Chase, M., Lysyanskaya, A.: Simulatable VRFS with applications to multi-theorem NIZK. In: CRYPTO, vol. 4622, Lecture Notes in Computer Science, pp. 303–322. Springer (2007)
21. Cox, D., Little, J., OShea, D.: Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra. Springer Science & Business Media (2013)
22. Chase, M., Meiklejohn, S., Zaverucha, G.: Algebraic MACs and keyed-verification anonymous credentials. In: CCS, pp. 1205–1216. ACM (2014)
23. Cascudo, I, Schnyder, R: A note on secure multiparty computation via higher residue symbol techniques. IACR Cryptol. ePrint Arch. **2020**, 183 (2020)
24. Damgård, I.: On the randomness of Legendre and Jacobi sequences. In: CRYPTO, vol. 403, Lecture Notes in Computer Science, pp. 163–172. Springer (1988)
25. Davenport, Harold: On the distribution of quadratic residues (mod p). J. Lond. Math. Soc. **1**(1), 49–54 (1931)
26. Déchene, I.: Generalized Jacobians in cryptography. ProQuest (2007)
27. Davidson, A., Faz-Hernández, A., Sullivan, N., Wood, C.: Oblivious pseudorandom functions (OPRFs) using prime-order groups (2021). https://datatracker.ietf.org/doc/draft-irtf-cfrg-voprf/
28. Davidson, Alex, Goldberg, Ian, Sullivan, Nick, Tankersley, George, Valsorda, Filippo: Privacy pass: bypassing internet challenges anonymously. Proc. Priv. Enhanc. Technol. **2018**(3), 164–180 (2018)
29. Ding, Cunsheng, Helleseth, Tor, Shan, Weijuan: On the linear complexity of Legendre sequences. IEEE Trans. Inf. Theory **44**(3), 1276–1278 (1998)
30. Dodis, Y., Kiltz, E., Pietrzak, K., Wichs, D.: Message authentication, revisited. In: EUROCRYPT, vol. 7237, Lecture Notes in Computer Science, pp. 355–374. Springer (2012)
31. Demmler, D, Schneider, T, Zohner, M: ABY— a framework for efficient mixed-protocol secure two-party computation. The Internet Society, In NDSS (2015)

32. Dodis, Y., Yampolskiy, A.: A verifiable random function with short proofs and keys. In: Public Key Cryptography, vol. 3386 of Lecture Notes in Computer Science, pp. 416–431. Springer (2005)
33. 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)
34. Freedman, M.J., Ishai, Y., Pinkas, B., Reingold, O.: Keyword search and oblivious pseudorandom functions. In: TCC, vol. 3378, Lecture Notes in Computer Science, pp. 303–324. Springer (2005)
35. Frixons, P., Schrottenloher, A.: Quantum security of the Legendre PRF. IACR Cryptol. ePrint Arch. 149 (2021)
36. Gilad, Y., Hemo, R., Micali, S., Vlachos, G., Zeldovich, N.: Algorand: scaling byzantine agreements for cryptocurrencies. In: SOSP, pp. 51–68. ACM (2017)
37. Garey, M.R., Johnson, D.S.: Computers and Intractability, San Francisco. W.H. Freeman, CA (1979)
38. Grassi, L., Khovratovich, D., Rechberger, C., Roy, A., Schofnegger, M.: Poseidon: a new hash function for zero-knowledge proof systems. In: USENIX Security Symposium, pp. 519–535. USENIX Association (2021)
39. Gyarmati, Katalin, Mauduit, Christian, Sárközy, András: The cross-correlation measure for families of binary sequences. Applied Algebra and Number Theory, pp. 126–143. Cambridge University Press (2014)
40. Goldberg, S., Naor, M., Papadopoulos, D., Reyzin, L., Vasant, S., Ziv, A.: Nsec5: provably preventing DNSSEC zone enumeration. In: The Network and Distributed System Security (NDSS) Symposium, CA, San Diego (2015)
41. Goldberg, S., Naor, M., Papadopoulos, D., Reyzin, L.: NSEC5 from elliptic curves: provably preventing dnssec zone enumeration with shorter responses. Cryptol. ePrint Arch. Report 2016/083 (2016). https://ia.cr/2016/083
42. Groth, J.: On the size of pairing-based non-interactive arguments. In: EUROCRYPT (2), vol. 9666, Lecture Notes in Computer Science, pp. 305–326. Springer (2016)
43. Grassi, L., Rechberger, C., Rotaru, D., Scholl, P., Smart, N.P.: MPC-friendly symmetric key primitives. In: CCS, pp. 430–443. ACM (2016)
44. Hartshorne, R.: Algebraic Geometry, vol. 52. Springer Science & Business Media (2013)
45. Hazay, Carmit, Lindell, Yehuda: Efficient protocols for set intersection and pattern matching with security against malicious and covert adversaries. J. Cryptol. **23**(3), 422–456 (2010)
46. 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)
47. Ishai, Y., Kushilevitz, E., Ostrovsky, R., Prabhakaran, M., Sahai, A.: Efficient non-interactive secure computation. In: EUROCRYPT, vol. 6632, Lecture Notes in Computer Science, pp. 406–425. Springer (2011)
48. Jakobsen, T., Knudsen, L.: The interpolation attack on block ciphers. In: International Workshop on Fast Software Encryption, pp. 28–40. Springer (1997)
49. Jarecki, S., Kiayias, A., Krawczyk, H.: Round-optimal password-protected secret sharing and T-PAKE in the password-only model. In: ASIACRYPT (2), vol. 8874, Lecture Notes in Computer Science, pp. 233–253. Springer (2014)
50. Jarecki, S., Kiayias, A., Krawczyk, H., Xu, J.: Highly-efficient and composable password-protected secret sharing (or: how to protect your bitcoin wallet online). In: 2016 IEEE European Symposium on Security and Privacy (EuroS &P), pp. 276–291. IEEE (2016)
51. Jarecki, S., Krawczyk, H., Xu, J.: Opaque: an asymmetric PAKE protocol secure against pre-computation attacks. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 456–486. Springer (2018)
52. Jarecki, S., Liu, X.: Efficient oblivious pseudorandom function with applications to adaptive OT and secure computation of set intersection. In: TCC, vol. 5444, Lecture Notes in Computer Science, pp. 577–594. Springer (2009)
53. Keelveedhi, S., Bellare, M., Ristenpart, T.: Dupless: server-aided encryption for deduplicated storage. In: 22nd {USENIX} Security Symposium ({USENIX} Security 13), pp. 179–194. (2013)
54. Khovratovich, D.: Key recovery attacks on the Legendre PRFS within the birthday bound. Cryptol. ePrint Arch. Report 2019/862 (2019)
55. Karakoç, F., Küpçü, A.: Linear complexity private set intersection for secure two-party protocols. In: CANS, vol. 12579, Lecture Notes in Computer Science, pp. 409–429. Springer (2020)

56. Kaluderovic, N, Kleinjung, T, Kostic, D: Improved key recovery on the Legendre PRF. IACR Cryptol. ePrint Arch. **2020**, 98 (2020)
57. 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)
58. Kiss, Á., Liu, J., Schneider, T., Asokan, N., Pinkas, B.: Private set intersection for unequal set sizes with mobile applications. Proc. Priv. Enhanc. Technol. **4**, 177–197 (2017)
59. Kolesnikov, V., Matania, N., Pinkas, B., Rosulek, M., Trieu, N.: Practical multi-party private set intersection from symmetric-key techniques. In: CCS, pp. 1257–1272. ACM (2017)
60. Katz, J., Ostrovsky, R.: Round-optimal secure two-party computation. In: CRYPTO, vol. 3152 Lecture Notes in Computer Science, pp. 335–354. Springer (2004)
61. Kipnis, A., Shamir, A.: Cryptanalysis of the HFE public key cryptosystem by relinearization. In: Annual International Cryptology Conference, pp. 19–30. Springer (1999)
62. Liang, B., Banegas, G., Mitrokotsa, A.: Statically aggregate verifiable random functions and application to e-lottery. Cryptography **4**(4), 37 (2020)
63. Lemmermeyer, F.: Conics-a poor man's elliptic curves. arXiv:math/0311306. (2003)
64. 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)
65. Mohassel, P., Rosulek, M.: Non-interactive secure 2pc in the offline/online and batch settings. In: EUROCRYPT (3), vol. 10212, Lecture Notes in Computer Science, pp. 425–455. (2017)
66. 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)
67. Mauduit, Christian, Sárközy, András: On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol. Acta Arith. **82**(4), 365–377 (1997)
68. Naor, M., Reingold, O.: Number-theoretic constructions of efficient pseudo-random functions. In: FOCS, pp. 458–467. IEEE Computer Society (1997)
69. 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)
70. Peralta, Rene: On the distribution of quadratic residues and nonresidues modulo a prime number. Math. Comput. **58**(197), 433–440 (1992)
71. 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)
72. Pinkas, B., Schneider, T., Tkachenko, O., Yanai, A.: Efficient circuit-based PSI with linear communication. In: EUROCRYPT (3), vol. 11478, Lecture Notes in Computer Science, pp. 122–153. Springer (2019)
73. Papadopoulos, D., Wessels, D., Huque, S., Naor, M., čelák, J.V., Reyzin, L., Goldberg, S.: Making NSEC5 practical for DNSSEC. Cryptol. ePrint Arch. Report 2017/099 (2017)
74. Russell, A., Shparlinski, I.E.: Classical and quantum function reconstruction via character evaluation. J. Complex. **20**(2—–3), 404–422 (2004)
75. Sugita, M., Kawazoe, M., Imai, H.: Relation between XL algorithm and Gröbner bases algorithms. IACR eprint Server (2004)
76. Tóth, Viktória: Collision and avalanche effect in families of pseudorandom binary sequences. Period. Math. Hung. **55**(2), 185–196 (2007)
77. Ullah, E.: New techniques for polynomial system solving (Doctoral dissertation Universität Passau) (2012)
78. van Dam, W., Hallgren, S., Ip, L.: Quantum algorithms for some hidden shift problems. SIAM J. Comput. **36**(3), 763–778 (2006)
79. Vinogradov, I.M..: Elements of Number Theory. Courier Dover Publications (2016)
80. Ching-Hua, Y.: Sign modules in secure arithmetic circuits. IACR Cryptol. ePrint Arch. **2011**, 539 (2011)