POLYNOMIAL INTERPOLATION AND IDENTITY TESTING FROM HIGH POWERS OVER FINITE FIELDS

GÁBOR IVANYOS, MAREK KARPINSKI, MIKLOS SANTHA, NITIN SAXENA, AND IGOR E. SHPARLINSKI

Abstract. We consider the problem of recovering (i.e. interpolating) and identity testing of a "hidden" monic polynomial f, given an oracle access to $f(x)^e$ for $x \in \mathbb{F}_q$ (extension fields access is not permitted). The naive interpolation algorithm needs $O(e \deg f)$ queries and thus requires $e \deg f < q$. We design algorithms that are asymptotically better in certain cases; requiring only $e^{o(1)}$ queries to the oracle. In the randomized (and quantum) setting, we give a substantially better interpolation algorithm, that requires only $O(\deg f \log q)$ queries. Such results have been known before only for the special case of a linear f, called the *hidden shifted power* problem.

We use techniques from algebra, such as effective versions of Hilbert's Nullstellensatz, and analytic number theory, such as results on the distribution of rational functions in subgroups and character sum estimates.

1. Introduction

Let $\mathbb{F}_q$ be a finite field of q elements. Here we consider several problems of recovering and identity testing of a "hidden" monic polynomial $f \in \mathbb{F}_q[X]$, given $\mathfrak{O}_{e,f}$ an oracle that on every input $x \in \mathbb{F}_q$ outputs $\mathfrak{O}_{e,f}(x) = f(x)^e$ for some large positive integer $e \mid q-1$.

More precisely, we consider the following problem *Interpolation from Powers*:

> given an oracle $\mathfrak{O}_{e,f}$ for some unknown monic polynomial $f \in \mathbb{F}_q[X]$, recover f.

We also consider the following two versions of the *Identity Testing from Powers*:

> given an oracle $\mathfrak{O}_{e,f}$ for some unknown monic polynomial $f \in \mathbb{F}_q[X]$ and another known polynomial $g \in \mathbb{F}_q[X]$, decide whether $f = g$,

and

given two oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$ for some unknown monic polynomials $f, g \in \mathbb{F}_q[X]$, decide whether $f = g$.

In particular, for a linear polynomial $f(X) = X + s$, with a 'hidden' $a \in \mathbb{F}_q$, we denote $\mathfrak{O}_{e,f} = \mathcal{O}_{e,s}$. We remark that in this case there are two naive algorithms that work for linear polynomials:

- One can query $\mathcal{O}_{e,s}$ at $e+1$ arbitrary points and then using a fast interpolation algorithm, see [vzGG13], get a deterministic algorithm of complexity $e(\log q)^{O(1)}$ (as in [vzGG13], we measure the complexity of an algorithm by the number of bit operations in the standard RAM model).
- For probablistic testing one can query $\mathcal{O}_{e,s}$ (and $\mathcal{O}_{e,t}$) at randomly chosen elements $x \in \mathbb{F}_q$ until the desired level of confidence is achieved (note that the equation $(x+s)^e = (x+t)^e$ has at most e solutions $x \in \mathbb{F}_q$).

These naive algorithms have been improved by Bourgain, Garaev, Konyagin and Shparlinski [BGKS12] in several cases (with respect to both the time complexity and the number of queries).

Furthermore, in the case when a quantum version of the oracle $\mathcal{O}_{e,s}$ is given, van Dam, Hallgren and Ip [vDHI06] have given a polynomial time quantum algorithm which recovers s, see also [vD02].

For non-linear polynomials $f \in \mathbb{F}_q[X]$ some classical and quantum algorithms are given by Russell and Shparlinski [RS04]. However they do not reach the level of those of [BGKS12, vD02, vDHI06] due to several additional obstacles which arise for non-linear polynomials. For example, we note that both the interpolation and random sampling algorithms fail if $e \deg f > q$. Indeed, note that queries from the extension field are not permitted, and $\mathbb{F}_q$ may not have enough elements to make these algorithms correct.

Here we consider both classical and quantum algorithms. In particular, we extend the results of [BGKS12, Section 3.3] to arbitrary monic polynomials $f \in \mathbb{F}_p[X]$ for a prime p. These results are based on some bounds of character sums and also new results about the order of multiplicative group generated by the values of a rational function on several consecutive integers.

Further, we also consider quantum algorithms. However, our setting is quite different from those of [vD02, vDHI06] as we do not assume that the values of f are given by a quantum oracle, rather the algorithm works with the classical oracle $\mathfrak{O}_{e,f}$.

The above questions appear naturally in understanding the pseudorandomness of the *Legendre symbol* $\left(\frac{f(x)}{p}\right)$. In particular, this has applications in the cryptanalysis of certain homomorphic cryptosystems. See [BM84, BL96, Dam90, MvOV10] for further details.

Note that the above questions are closely related to the general problem of oracle (also sometimes called "black-box") polynomial interpolation and identity testing for arbitrary polynomials (though forbidding the use of field extensions makes the problems harder), see [Sax09, Sax14, SY10] and the references therein.

Throughout the paper, any implied constants in the symbols O, $\ll$ and $\gg$ may occasionally, where obvious, depend on the degree d of the polynomial f (& an integer parameter ν), and are absolute otherwise. We recall that the notations $U = O(V)$, $U \ll V$ and $V \gg U$ are all equivalent to the statement that the inequality $|U| \le cV$ holds with some constant $c > 0$.

2. Identity Testing on Classical Computers

2.1. **Main results.** Here we consider the identity testing case of two unknown *monic* polynomials $f, g \in \mathbb{F}_q[X]$ of degree d given the oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$. We remark that if f/g is an $(q-1)/e$-th power of a non-constant rational function over $\mathbb{F}_q$ then it is impossible to distinguish between f and g from the oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$. We write $f \sim_e g$ in this case, and $f \not\sim_e g$ otherwise.

We note that it is shown in the proof of [RS04, Theorem 6] that the Weil bound of multiplicative character sums (see [IK04, Theorem 11.23]) implies that given two oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$ for some unknown monic polynomials $f, g \in \mathbb{F}_q[X]$ with $f \not\sim_e g$ one can decide whether $f = g$ in time $q^{1/2+o(1)}$. Note that the result of [RS04] is stated only for prime fields $\mathbb{F}_p$ but it can be extended to arbitrary fields at the cost of only typographical changes. The same holds for here the results of Section 3 but the results of Section 2 hold only for prime fields.

For "small" values of e, over prime fields $\mathbb{F}_p$, we have a stronger result.

Theorem 1 (Small e). *For a prime p and a positive integer $e \mid p-1$, with $e \le p^{\delta}$ for some fixed $\delta > 0$, given two oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$ for some unknown monic polynomials $f, g \in \mathbb{F}_p[X]$ of degree d with $f \not\sim_e g$, there is a deterministic algorithm to decide whether $f = g$ in $e^{c_0(d)\delta^{1/(2d-1)}}$ queries to the oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$, where $c_0(d)$ depends only on d.*

In particular, we see from (the proof of) Theorem 1 that if $e = p^{o(1)}$ and $e \to \infty$ then we can test whether $f = g$ in time $e^{o(1)}(\log p)^{O(1)}$ in $e^{o(1)}$ oracle calls.

For intermediate values of e, the following result complements both Theorem 1 and the result of [RS04]. We, however, have to assume that the polynomials f and g are *irreducible*.

Theorem 2 (Medium e). *For a prime p and a positive integer $e \mid p-1$, with $e \le p^{\eta-\delta}$ for some fixed $\delta > 0$, given two oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$ for some unknown monic polynomials $f, g \in \mathbb{F}_p[X]$ of degree $d \ge 1$ with $f \not\sim_e g$, there is a deterministic algorithm to decide whether $f = g$ in $e^{\kappa+\delta}$ queries to the oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$, where*

$$\eta = \frac{4d-1}{4d^2(d+1)^2} \qquad \text{and} \qquad \kappa = \frac{2d}{4d-1}.$$

The proofs of Theorems 1 and 2 are given below in Sections 2.5 and 2.6, respectively.

2.2. **Background from arithmetic algebraic geometry.** Our argument makes use of a slight modification of [BGKS12, Lemma 23], which is based on a quantitative version of effective Hilbert's Nullstellensatz given by D'Andrea, Krick and Sombra [DKS13], which improved the previous estimates due to Krick, Pardo and Sombra [KPS01].

As usual, we define the *logarithmic height* of a nonzero polynomial $P \in \mathbb{Z}[Z_1, \ldots, Z_n]$ as the maximum logarithm of the largest (by absolute value) coefficient of P.

The next statement is essentially [BGKS12, Lemma 23], however we now use [DKS13, Theorem 2] instead of [KPS01, Theorem 1].

Lemma 3. *Let $P_1, \ldots, P_N \in \mathbb{Z}[Z_1, \ldots, Z_n]$ be $N \ge 2$ polynomials in n variables of degree at most $D \ge 3$ and of logarithmic height at most H and let $R \in \mathbb{Z}[Z_1, \ldots, Z_n]$ be a polynomial in n variables of degree at most $d \ge 3$ and of logarithmic height at most h such that R vanishes on the variety*

$$P_1(Z_1, \ldots, Z_n) = \ldots = P_N(Z_1, \ldots, Z_n) = 0.$$

There are polynomials $Q_1, \ldots, Q_N \in \mathbb{Z}[Z_1, \ldots, Z_n]$ and positive integers A and r with

$$\log A \le 2(n+1)dD^n H + 3D^{n+1}h + C(d, D, n, N),$$

such that

$$P_1 Q_1 + \ldots + P_N Q_N = AR^r,$$

where $C(d, D, n, N)$ depends only on d, D, n and N.

We note that using Lemma 3 in the argument of [BGKS12] allows to replace ν^{-4} with ν^{-3} in [BGKS12, Lemma 35]. In turn, this allows us to replace $\delta^{1/3}$ with $\delta^{1/2}$ in [BGKS12, Lemma 38 and Theorem 51].

We now define the logarithmic height of an algebraic number $\alpha \neq 0$ as the logarithmic height of its minimal polynomial.

We need a slightly more general form of a result of Chang [Cha03]. In fact, this is exactly the statement that is established in the proof of [Cha03, Lemma 2.14], see [Cha03, Equation (2.15)].

Lemma 4. *Let $P_1, \ldots, P_N, R \in \mathbb{Z}[Z_1, \ldots, Z_n]$ be $N+1 \geq 2$ polynomials in n variables of degree at most D and of logarithmic height at most $H \geq 1$. If the zero-set*

$$P_1(Z_1, \ldots, Z_n) = \ldots = P_N(Z_1, \ldots, Z_n) = 0 \quad \textit{and} \quad R(Z_1, \ldots, Z_n) \neq 0$$

is not empty then it has a point $(\beta_1, \ldots, \beta_n)$ in an extension $\mathbb{K}$ of $\mathbb{Q}$ of degree $[\mathbb{K} : \mathbb{Q}] \leq C_1(D, n)$ such that its logarithmic height is at most $C_2(D, n, N)H$, where $C_1(D, n)$ depends only on D, n and $C_2(D, n, N)$ depends only on D, n and N.

2.3. **Product sets in number fields.** For a set $\mathcal{A}$ in an arbitrary semi-group, we use $\mathcal{A}^{(\nu)}$ to denote the ν-*fold product set*, that is

$$\mathcal{A}^{(\nu)} = \{a_1 \ldots a_\nu \ : \ a_1, \ldots, a_\nu \in \mathcal{A}\}.$$

We recall the following result given in [BGKS12, Lemma 29], which in turn generalises [BKS08, Corollary 3].

Corollary 5. *Let $\mathbb{K}$ be a finite extension of $\mathbb{Q}$ of degree $D = [\mathbb{K} : \mathbb{Q}]$. Let $\mathcal{C} \subseteq \mathbb{K}$ be a finite set with elements of logarithmic height at most $H \geq 2$. Then we have*

$$\#\mathcal{C}^{(\nu)} > \exp\left(-c(D, \nu)\frac{H}{\sqrt{\log H}}\right)(\#\mathcal{C})^{\nu},$$

where $c(D, \nu)$ depends only on D and ν.

2.4. **Product sets of consecutive values of rational functions in prime fields.** We now show that for a nontrivial rational function $f/g \in \mathbb{F}_p(X)$ and an integer $h \geq 1$, the set formed by h consecutive values of f/g cannot be all inside a small multiplicative subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$. For the linear fractional function $(X + s)/(X + t)$ this has been obtained in [BGKS12, Lemma 35].

Lemma 6. *Let $\nu \geq 1$ be a fixed integer. Assume that for some sufficiently large positive integer h and prime p we have*

$$h < p^{c(d)\nu^{-2d}},$$

where $c(d)$ depends only on d. For two distinct monic polynomials $f, g \in \mathbb{F}_p$ of degrees d, we consider the set

$$\mathcal{A} = \left\{ \frac{f(x)}{g(x)} \ : \ 1 \le x \le h \right\} \subseteq \mathbb{F}_p.$$

Then

$$\#\mathcal{A}^{(\nu)} > \exp\left(-c(d,\nu)\frac{\log h}{\sqrt{\log\log h}}\right) h^{\nu},$$

where $c(d,\nu)$ depends only on ν and d.

Proof. We closely follow the proof of [BGKS12, Lemma 35]. Let

$$f(X) = X^d + \sum_{k=0}^{d-1} a_{d-k} x^k \qquad \text{and} \qquad g(X) = X^d + \sum_{\ell=0}^{d-1} b_{d-\ell} X^{\ell}.$$

The idea is to move from the finite field to a number field, where we are in a position to apply Corollary 5.

We consider the collection $\mathcal{P} \subseteq \mathbb{Z}[\mathbf{U}, \mathbf{V}]$, where

$$\mathbf{U} = (U_1, \ldots, U_d) \qquad \text{and} \qquad \mathbf{V} = (V_1, \ldots, V_d),$$

of polynomials

$$\begin{aligned} P_{\mathbf{x},\mathbf{y}}(\mathbf{U}, \mathbf{V}) = &\prod_{i=1}^{\nu} \left(x_i^d + \sum_{k=0}^{d-1} U_{d-k} x_i^k \right) \left(y_i^d + \sum_{\ell=0}^{d-1} V_{d-\ell} y_i^{\ell} \right) \\ &- \prod_{i=1}^{\nu} \left(x_i^d + \sum_{\ell=0}^{d-1} V_{d-\ell} x_i^{\ell} \right) \left(y_i^d + \sum_{k=0}^{d-1} U_{d-k} y_i^k \right), \end{aligned}$$

where $\mathbf{x} = (x_1, \ldots, x_\nu)$ and $\mathbf{y} = (y_1, \ldots, y_\nu)$ are integral vectors with entries in $\mathcal{I} := [1, h]$ and such that

$$P_{\mathbf{x},\mathbf{y}}(x_1, \ldots, x_d, y_1, \ldots, y_d) \equiv 0 \pmod p.$$

Note that

$$P_{\mathbf{x},\mathbf{y}}(a_1, \ldots, a_d, b_1, \ldots, b_d) \equiv \prod_{i=1}^{\nu} f(x_i) g(y_i) - \prod_{i=1}^{\nu} f(y_i) g(x_i) \pmod p.$$

Clearly if $P_{\mathbf{x},\mathbf{y}}$ is identical to zero then, by the uniqueness of polynomial factorisation in the ring $\mathbb{F}_p[\mathbf{U}, \mathbf{V}]$, the components of $\mathbf{y}$ are permutations of those of $\mathbf{x}$. So in this case we obviously obtain

$$\#\mathcal{A}^{(\nu)} \ge \frac{1}{\nu!} \left(\#f(\mathcal{I})\right)^{\nu} \gg H^{\nu}.$$

Hence, we now assume that $\mathcal{P}$ contains non-zero polynomials.

Clearly, every $P \in \mathcal{P}$ is of degree at most 2ν and of logarithmic height $O(\log h)$.

We take a family $\mathcal{P}_0$ containing the largest possible number

$$N \le (\nu+1)^{2d} - 1$$

of linearly independent polynomials $P_1, \ldots, P_N \in \mathcal{P}$, and consider the variety

$$\mathcal{V}: \ \{(\mathbf{U}, \mathbf{V}) \in \mathbb{C}^{2d} \ : \ P_1(\mathbf{U}, \mathbf{V}) = \ldots = P_N(\mathbf{U}, \mathbf{V}) = 0\}.$$

Clearly $\mathcal{V} \neq \emptyset$ as it contains the diagonal $\mathbf{U} = \mathbf{V}$.

We claim that $\mathcal{V}$ contains a point outside of the diagonal, that is, there is a point $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ with $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{C}^d$ and $\boldsymbol{\beta} \neq \boldsymbol{\gamma}$.

Assume that $\mathcal{V}$ does not contain a point outside of the diagonal. Then for every $k = 1, \ldots, d$, the polynomial

$$R_k(U_1, \ldots, U_d, V_1, \ldots, V_d) = U_k - V_k$$

vanishes on $\mathcal{V}$.

Then by Lemma 3 we see that there are polynomials $Q_{k,1}, \ldots, Q_{k,N} \in \mathbb{Z}[\mathbf{U}, \mathbf{V}]$ and positive integers A_k and r_k with

$$\log A_k \le c_0 d (2\nu)^{2d} \log h \tag{1}$$

for some absolute constant c_0 (provided that h is large enough) and such that

$$P_1 Q_{k,1} + \ldots + P_N Q_{k,N} = A_k (U_k - V_k)^{r_k}. \tag{2}$$

Since $f \neq g$, there is $k \in \{1, \ldots, d\}$ for which $a_k \not\equiv b_k \pmod p$. For this k we substitute

$$(\mathbf{U}, \mathbf{V}) = (a_1, \ldots, a_d, b_1, \ldots, b_d)$$

in (2). Recalling the definition of the set $\mathcal{P}$ we now derive that $p \mid A_k$. Taking

$$c(d) = \frac{1}{c_0 d 2^d + 1}$$

in the condition of the lemma, we see from (1) that this is impossible.

Hence the set

$$\mathcal{U} = \mathcal{V} \cap [\mathbf{U} - \mathbf{V} \neq 0]$$

is nonempty. Applying Lemma 4 we see that it has a point $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ with components of logarithmic height $O(\log h)$ in an extension $\mathbb{K}$ of $\mathbb{Q}$ of degree $[\mathbb{K} : \mathbb{Q}] \le \Delta(d, \nu)$, where $\Delta(d, \nu)$ depends only on d and ν.

Consider the maps $\Phi: \ \mathcal{I}^\nu \to \mathbb{F}_p$ given by

$$\Phi: \ \mathbf{x} = (x_1, \ldots, x_\nu) \mapsto \prod_{j=1}^{\nu} \frac{f(x_j)}{g(x_j)}$$

and $\Psi : \mathcal{I}^\nu \to \mathbb{K}$ given by

$$\Psi : \ \mathbf{x} = (x_1, \ldots, x_\nu) \mapsto \prod_{j=1}^{\nu} \frac{F_{\boldsymbol\beta}(x_j)}{G_{\boldsymbol\gamma}(x_j)},$$

where

$$F_{\boldsymbol\beta}(X) = X^d + \sum_{k=0}^{d-1} \beta_{d-k} x^k \qquad \text{and} \qquad G_{\boldsymbol\gamma}(X) = X^d + \sum_{\ell=0}^{d-1} \gamma_{d-\ell} X^\ell.$$

By construction of $(\boldsymbol\beta, \boldsymbol\gamma)$ we have that $\Psi(\mathbf{x}) = \Psi(\mathbf{y})$ if $\Phi(\mathbf{x}) = \Phi(\mathbf{y})$. Hence

$$\#\mathcal{A}^{(\nu)} \ge \mathrm{Im}\Psi = \#\mathcal{C}^{(\nu)},$$

where $\mathrm{Im}\Psi$ is the image set of the map Ψ and

$$\mathcal{C} = \left\{ \frac{F_{\boldsymbol\beta}(x)}{G_{\boldsymbol\gamma}(x)} \ : \ 1 \le x \le h \right\} \subseteq \mathbb{K}.$$

Using Corollary 5, we derive the result. □

We also recall the following bound which is a special case of a more general result from [GPS15, Theorem 7].

Lemma 7. *If for two relatively prime monic polynomials $f, g \in \mathbb{F}_p$ of degree $d \ge 1$, a positive integer h and a multiplicative subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ we have*

$$\left\{ \frac{f(x)}{g(x)} \ : \ 1 \le x \le h \right\} \subseteq \mathcal{G}.$$

Then

$$\#\mathcal{G} \gg \min\{h^{2(1-\tau)+o(1)}, h^{2(1-\rho-\tau)+o(1)} p^{2\vartheta}\},$$

where

$$\vartheta = \frac{1}{2d(d+2)}, \qquad \rho = \frac{(d+1)^2}{2(d+2)}, \qquad \tau = \frac{1}{4d},$$

and the implied constant depends on d.

Proof. By [GPS15, Theorem 7], applied with $d = e$ (and thus with $k = d(d+1)^2$, $s = d^2 + 2d$ and hence the above values of ϑ, ρ and τ), we have

$$\#\left(\left\{ \frac{f(x)}{g(x)} \ : \ 1 \le x \le h \right\} \bigcap \mathcal{G} \right) \le \left(1 + h^\rho p^{-\vartheta}\right) h^{\tau+o(1)} T^{1/2}$$

where $T = \#\mathcal{G}$. Under the condition of the lemma we have

$$\#\left(\left\{ \frac{f(x)}{g(x)} \ : \ 1 \le x \le h \right\} \bigcap \mathcal{G} \right) = h$$

and the result follows. □

2.5. **Proof of Theorem 1.** We set

$$\nu = \left\lfloor \left(\frac{c(d)}{2\delta}\right)^{2d-1} \right\rfloor \qquad \text{and} \qquad h = \left\lfloor e^{1/\nu} \right\rfloor + 1,$$

where $c(d)$ is the constant of Lemma 6. We note that

$$\frac{2\delta}{\nu} \le \frac{c(d)}{\nu^{2d}}$$

so as $e \to \infty$ we have

$$e^{1/\nu} < h = e^{1/\nu + o(1)} \le e^{2/\nu} \le p^{2\delta/\nu} \le p^{c(d)/\nu^{2d}}. \tag{3}$$

We now query the oracles $\mathfrak{O}_{e,f}$ and $\mathfrak{O}_{e,g}$ for $x = 1, \ldots, h$.

If the oracles return two distinct values then clearly $f \ne g$. Now assume

$$f(x)^e = g(x)^e, \qquad x = 1, \ldots, h.$$

Therefore, the values $f(x)/g(x)$, $x = 1, \ldots, h$ belong to the subgroup $\mathcal{G}_e$ of $\mathbb{F}_p^*$ of order e. Hence for the set

$$\mathcal{A} = \left\{ \frac{f(x)}{g(x)} \ : \ 1 \le x \le h \right\} \subseteq \mathbb{F}_p \tag{4}$$

for any integer $\nu \ge 1$ we have

$$\mathcal{A}^{(\nu)} = \{a_1 \ldots a_\nu \ : \ a_1, \ldots, a_\nu \in \mathcal{A}\} \subseteq \mathcal{G}_e. \tag{5}$$

We see from (3) that Lemma 6 applies which contradicts (5) as we have $h^\nu > e$ for the above choice of the parameters. This concludes the proof.

2.6. **Proof of Theorem 2.** We fix some $\varepsilon > 0$ and set

$$h = \left\lceil e^{(1+\varepsilon)/(2-2\tau))} \right\rceil.$$

We also note that for the above choice of h and for

$$e^{1+\varepsilon} \le e^{(1-\rho-\tau)(1+\varepsilon)/(1-\tau)} p^{\vartheta} \tag{6}$$

we have

$$\min\{h^{2(1-\tau)}, h^{2(1-\rho-\tau)} p^{2\vartheta}\} \ge e^{1+\varepsilon}.$$

Therefore, under the condition (6), we derive from Lemma 7 that for the set $\mathcal{A}$ given by (4) we have $\mathcal{A} \not\subseteq \mathcal{G}_e$. Proceeding as in the proof of Theorem 1, we obtain an algorithm that requires h queries.

Clearly, for the above choice of h, the condition (6) is satisfied if

$$e^{(1+\varepsilon)\rho/(1-\tau)} \le p^{\vartheta}. \tag{7}$$

Taking

$$\eta = \frac{\vartheta(1-\tau)}{\rho} \qquad \text{and} \qquad \kappa = \frac{1}{2-2\tau}$$

we see that the condition (7) is equivalent to $e \le p^{\eta/(1+\varepsilon)}$, under which we get an algorithm which requires $h = O\left(e^{(1+\varepsilon)\kappa}\right)$ queries. Since $\varepsilon > 0$ is arbitrary, the result now follows.

3. Quantum and Randomized Interpolation

3.1. **Main results.** Here we present a quantum algorithm for the interpolation problem of finding an unknown monic polynomial $f \in \mathbb{F}_q[X]$ of degree d given the oracle $\mathfrak{O}_{e,f}$. We emphasise the difference between our settings where the oracle is classical and only the algorithm is quantum and the settings of [vD02, vDHI06] which employ the quantum analogue of the oracle $\mathfrak{O}_{e,f}$.

We recall that the oracle $\mathfrak{O}_{e,f}$ does not accept queries from field extensions of $\mathbb{F}_q$, and therefore, if $de > q$, we *cannot* interpolate f^e from queries to $\mathfrak{O}_{e,f}$.

Theorem 8. *Given an oracle $\mathfrak{O}_{e,f}$ for some unknown monic polynomial f of degree at most d, for any $\varepsilon > 0$ there is a quantum algorithm to find with probability $1-\varepsilon$ a polynomial g such that $g \sim_e f$ in time $e^{d/2}\left(d \log q \log(1/\varepsilon)\right)^{O(1)}$ and $O\left(d \log q \log(1/\varepsilon)\right)$ calls to $\mathfrak{O}_{e,f}$.*

Replacing quantum parts of the algorithm above with classical (randomized) methods, we obtain the following.

Theorem 9. *Given an oracle $\mathfrak{O}_{e,f}$ for some unknown monic polynomial f of degree at most d, for any $\varepsilon > 0$ there is a* randomized *algorithm to find with probability $1-\varepsilon$ a polynomial g such that $g \sim_e f$ in time $e^{d}\left(d \log q \log(1/\varepsilon)\right)^{O(1)}$ and $O\left(d \log q \log(1/\varepsilon)\right)$ calls to $\mathfrak{O}_{e,f}$.*

The proofs of Theorems 8 and 9 are given below in Sections 3.3 and 3.4, respectively.

3.2. **Coincidences among eth powers of polynomials.** The following result is immediate from the Weil bound on multiplicative character sums, see [IK04, Theorem 11.23].

Lemma 10. *Let $g_1, g_2 \in \mathbb{F}_q[X]$ be two monic polynomials of degree at most d with $g_1 \not\sim_e g_2$. Then*

$$\#\{x \in \mathbb{F}_q \ : \ g_1(x)^e = g_2(x)^e\} = \frac{q}{e} + O(dq^{1/2}).$$

We now immediately conclude.

Corollary 11. *Let $g_1, g_2 \in \mathbb{F}_q[X]$ be two monic polynomials of degree $o(q^{1/2})$ with $g_1 \not\sim_e g_2$. Then for any $e \le (q-1)/2$ and a sufficiently large q*

$$\#\{x \in \mathbb{F}_q \ : \ g_1(x)^e \neq g_2(x)^e\} \ge \frac{1}{3}q.$$

3.3. **Proof of Theorem 8.** Let $\mathcal{S}$ stand for the monic polynomials of degree at most d. By Corollary 11, a random choice of elements $x \in \mathbb{F}_q$ gives with probability at least 0.99 a set T of size $O(\log |\mathcal{S}|) = O(d \log q)$ such that for every pair $f, g \in \mathcal{S}$ we have $f(a)^e = g(a)^e$ for every $a \in T$ if and only if $f \sim_e g$.

We continue with picking d different elements $a_1, \ldots, a_d$ and use the oracle $\mathfrak{O}_{e,f}$ to obtain the values $b_j = f(a_j)^e$, $j = 1, \ldots, d$, as well as to get the values $b(a) = f(a)^e$ for every $a \in T$.

Using Shor's order finding and discrete logarithm algorithms [Sho97] we can also compute a generator ζ_e for the multiplicative subgroup $\{u \in \mathbb{F}_q : u^e = 1\}$ and for every j an element $z_j \in \mathbb{F}_q$ such that $z_j^e = b_j$.

The cost of the steps performed so far is polynomial in $\log q$ and d. Let $E = \{0, \ldots, e-1\}$. For a tuple $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ from E^d, let $f_{\boldsymbol{\alpha}}$ be the monic polynomial of degree at most d such that $f_{\boldsymbol{\alpha}}(a_j) = z_j \zeta_e^{\alpha_j}$, $j = 1, \ldots, d$. For any specific tuple α, the polynomial $f_{\boldsymbol{\alpha}}$ can be computed by simple interpolation in time polynomial in $d \log q$.

We use Grover's search [Gro96] over E^d to find a tuple $\boldsymbol{\alpha}$ with probablity at least 0.99 such that $f_{\boldsymbol{\alpha}}^e(a) = b(a)$ for every $a \in T$. The cost of this part is bounded by $O(e^{d/2})$ times a polynomial in $\log q$ and d. Repeating the whole procedure $O(\log(1/\varepsilon))$ times we achieve the desired probability level, which concludes the proof.

3.4. **Proof of Theorem 9.** Observe that a generator for the group $\{u \in \mathbb{F}_q : u^e = 1\}$ as well as elements z_j with $z_j^e = b_j$ can be found by simple classical algorithms of complexity bounded by $e^{1/2}(\log q)^{O(1)}$, that is, even within the complexity bound of Theorem 8. Indeed, assume that for every prime r diving e we have an element $g_r \in \mathbb{F}_q$ which is not an rth power of an $\mathbb{F}_q$ element. Such elements can be found in time $(\log q)^{O(1)}$ using random choices. The product of appropriate powers of the elements g_r is a generator for the group of the eth roots of unity.

For computing an eth roots of b_j it is sufficient to be able to take rth root of an arbitrary field element y for every prime divisor r of e. This task can be accomplished in time $\sqrt{r}(\log q)^{O(1)}$ as in the algorithm of Adleman, Manders and Miller [AMM77] instead of the brute force one that uses Shanks' baby step-giant step method for computing discrete logarithms in groups of order r, see [CP01, Section 5.3].

Therefore, if we replace Grover's search [Gro96] over E^d with a classical search we obtain a classical randomised algorithm of complexity $e^d(d \log q \log(1/\varepsilon))^{O(1)}$.

3.5. **Further Remarks.** Under Generalised Riemann Hypothesis we can derandomize the proof of Theorem 9. If $q = p$ is a prime then a generator for the group of eth roots of unity can be found in deterministic polynomial time. If, furthermore, $e \le p^{\delta}$ or $e \le p^{\eta-\delta}$ for some fixed $\delta > 0$, then we could use the test of Theorem 1 or Theorem 2 to obtain a deterministic algorithm of complexity $e^{d+c_0(d)\delta^{1/(2d-1)}}(d \log p)^{O(1)}$ or $e^{d+\kappa+o(1)}(d \log p)^{O(1)}$, respectively.

4. Comments and open problems

One can obtain analogues of Theorems 1 and 2 in the settings of high degree extensions of finite fields. More precisely, if $q = p^n$ for a fixed p and growing n, we write $\mathbb{F}_q \cong \mathbb{F}_p[X]/\langle\psi(X)\rangle$ for a fixed irreducible polynomial $\psi \in \mathbb{F}_p[X]$ of degree n. Then one can attempt to transfer the technique used in the proofs of Theorems 1 and 2 to this case where a role of a short interval of length h is now played by the set of polynomials of degree at most h. This approach has been used in [CS13, Shp14] for several related problems. We also note that a version of effective Hilbert's Nullstellensatz for function fields, which is needed for this approach, has recently been given by D'Andrea, Krick and Sombra [DKS13].

We remark that we do not know how to take any advantage of actually knowing g, and get stronger version of Theorems 1 and 2 in this case, like, for example, in [BGKS12, Section 3.2].

References

[AMM77] Leonard Adleman, Kenneth Manders, and Gary Miller, *On taking roots in finite fields*, 2013 IEEE 54th Annual Symposium on Foundations of Computer Science, IEEE, 1977, pp. 175–178.

[BGKS12] Jean Bourgain, Moubariz Z. Garaev, Sergei V. Konyagin, and Igor E. Shparlinski, *On the hidden shifted power problem*, SIAM Journal on Computing **41** (2012), no. 6, 1524–1557.

[BKS08] Jean Bourgain, Sergei V. Konyagin, and Igor E. Shparlinski, *Product sets of rationals, multiplicative translates of subgroups in residue rings, and fixed points of the discrete logarithm*, International Mathematics Research Notices **2008** (2008), rnn090.

[BL96] Dan Boneh and Richard J. Lipton, *Algorithms for black-box fields and their application to cryptography*, Advances in Cryptology CRYPTO 96, Springer, 1996, pp. 283–297.

[BM84] Manuel Blum and Silvio Micali, *How to generate cryptographically strong sequences of pseudorandom bits*, SIAM journal on Computing **13** (1984), no. 4, 850–864.

[Cha03] Mei-Chu Chang, *Factorization in generalized arithmetic progressions and application to the erdős-szemerédi sum-product problems*, Geometric And Functional Analysis **13** (2003), no. 4, 720–736.

[CP01] Richard Crandall and Carl Pomerance, *Prime numbers: A computational perspective*, New York, 2001.

[CS13] Javier Cilleruelo and Igor Shparlinski, *Concentration of points on curves in finite fields*, Monatshefte für Mathematik **171** (2013), no. 3-4, 315–327.

[Dam90] Ivan B. Damgård, *On the randomness of legendre and jacobi sequences*, Advances in Cryptology CRYPTO 88, Springer, 1990, pp. 163–172.

[DKS13] Carlos D'Andrea, Teresa Krick, and Martin Sombra, *Heights of varieties in multiprojective spaces and arithmetic nullstellensätze*, Annales Sci. de l'ENS **46** (2013), 549–627.

[GPS15] Domingo Gómez-Pérez and Igor E. Shparlinski, *Subgroups generated by rational functions in finite fields*, Monat. Math. **176** (2015), 241–253.

[Gro96] Lov K. Grover, *A fast quantum mechanical algorithm for database search*, Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, ACM, 1996, pp. 212–219.

[IK04] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, vol. 53, American Mathematical Society, Providence, 2004.

[KPS01] Teresa Krick, Luis Miguel Pardo, and Martín Sombra, *Sharp estimates for the arithmetic nullstellensatz*, Duke Mathematical Journal **109** (2001), no. 3, 521–598.

[MvOV10] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone, *Handbook of applied cryptography*, CRC press, 2010.

[RS04] Alexander Russell and Igor E. Shparlinski, *Classical and quantum function reconstruction via character evaluation*, Journal of Complexity **20** (2004), no. 2, 404–422.

[Sax09] Nitin Saxena, *Progress on polynomial identity testing*, Bulletin of the EATCS **99** (2009), 49–79.

[Sax14] ———, *Progress on polynomial identity testing - 2*, arXiv Preprint, 2014, `http://arxiv.org/abs/1401.0976`.

[Sho97] Peter W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Journal on Computing **26** (1997), no. 5, 1484–1509.

[Shp14] Igor E. Shparlinski, *Products with variables from low-dimensional affine spaces and shifted power identity testing in finite fields*, Journal of Symbolic Computation **64** (2014), 35–41.

[SY10] Amir Shpilka and Amir Yehudayoff, *Arithmetic circuits: A survey of recent results and open questions*, Foundations and Trends in Theoretical Computer Science **5** (2010), no. 3-4, 207–388.

[vD02] Wim van Dam, *Quantum algorithms for weighing matrices and quadratic residues*, Algorithmica **34** (2002), no. 4, 413–428.

[vDHI06] Wim van Dam, Sean Hallgren, and Lawrence Ip, *Quantum algorithms for some hidden shift problems*, SIAM Journal on Computing **36** (2006), no. 3, 763–778.

[vzGG13] Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, Cambridge university press, 2013.

Institute for Computer Science and Control, Hungarian Academy of Sciences, H-1111 Budapest, Hungary
E-mail address: gabor.ivanyos@sztaki.mta.hu

Department of Computer Science, Bonn University, 53113 Bonn, Germany
E-mail address: marek@cs.uni-bonn.de

CNRS, Université Paris Diderot, 75013 Paris, France and CQT, National University of Singapore, 117543 Singapore
E-mail address: miklos.santha@liafa.univ-paris-diderot.fr

Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur, UP 208016, India
E-mail address: nitin@cse.iitk.ac.in

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052 Australia
E-mail address: igor.shparlinski@unsw.edu.au