

A NEW APPROACH TO THE DISCRETE LOGARITHM PROBLEM WITH AUXILIARY INPUTS

JUNG HEE CHEON AND TAECHAN KIM

ABSTRACT. The discrete logarithm problem with auxiliary inputs is to solve α for given elements $g, g^\alpha, \dots, g^{\alpha^d}$ of a cyclic group $G = \langle g \rangle$ of prime order p . The best-known algorithm, proposed by Cheon in 2006, solves α in the case of $d|(p \pm 1)$ with running time of $O(\sqrt{p/d} + d^i)$ group exponentiations ($i = 1$ or $1/2$ depending on the sign). There have been several attempts to generalize this algorithm in the case of $\Phi_k(p)$ for $k \geq 3$, but it has been shown, by Kim, Cheon and Lee, that they cannot have better complexity than the usual square root algorithms.

We propose a new algorithm to solve the DLPwAI. The complexity of the algorithm is determined by a chosen polynomial $f \in \mathbb{F}_p[x]$ of degree d . We show that the proposed algorithm has a running time of $\tilde{O}(\sqrt{p/\tau_f} + d)$ group exponentiations, where τ_f is the number of absolutely irreducible factors of $f(x) - f(y)$. We note that it is always smaller than $\tilde{O}(p^{1/2})$.

To obtain a better complexity of the algorithm, we investigate an upper bound of τ_f and try to find polynomials that achieve the upper bound. We can find such polynomials in the case of $d|(p \pm 1)$. In this case, the algorithm has a running time of $\tilde{O}(\sqrt{p/d} + d)$ group operations which corresponds with the lower bound in the generic group model. On the contrary, we show that no polynomial exists that achieves the upper bound in the case of $d|\Phi_3(p) = p^2 + p + 1$.

As an independent interest, we present an analysis of a non-uniform birthday problem. Precisely, we show that a collision occurs with a high probability after $O(\frac{1}{\sqrt{\sum_k w_k^2}})$ samplings of balls, where the probability w_k of assigning balls to the bin k is arbitrary.

1. INTRODUCTION

1.1. Discrete logarithm problem with auxiliary inputs. The discrete logarithm problem with auxiliary inputs (DLPwAI) in a group G of prime order p is: Given $g, g^\alpha, \dots, g^{\alpha^d} \in G$, to compute α . A number of variants of the DLP such as Weak Diffie-Hellman Problem (WDHP) [16], Strong Diffie-Hellman Problem (SDHP) [2], Bilinear Diffie-Hellman Inversion Problem (BDHIP) [1] and Bilinear Diffie-Hellman Exponent Problem (BDHEP) [3] ask to determine some values encoded by the discrete logarithm α for the given $g, g^\alpha, \dots, g^{\alpha^d} \in G$, so solving the DLPwAI implies to solve these problems. These problems arise in a number of contexts, for example, a traitor tracing [16], short signatures [2], an ID-based encryption [1], a broadcast encryption [3] and so on.

2000 *Mathematics Subject Classification.* Primary 68Q25; Secondary 11Y16.

Key words and phrases. discrete logarithm problem, Cheon's algorithm, birthday problem.

A part of the work appears in PhD dissertation of the second author.

The state-of-the-art algorithm for this was proposed by Cheon [5, 6] and Brown and Gallant [4]. It has a running time of $O(\sqrt{p/d} + \sqrt{d})$ group exponentiations in the case of $d|(p-1)$ and $O(\sqrt{p/d} + d)$ in the case of $d|(p+1)$. The idea of Cheon's algorithm is to embed the discrete logarithm α into the finite fields \mathbb{F}_p or \mathbb{F}_{p^2} . He exploits the fact that α^d can be embedded into an element of a small subgroup of \mathbb{F}_p or \mathbb{F}_{p^2} when d is a divisor of $p \pm 1$.

After then, there have been several generalizations that attempt to solve the problem when d is a divisor of $\Phi_k(p)$ for the k -th cyclotomic polynomial $\Phi_k(x)$ [20, 14, 7]. Satoh [20] generalized the algorithm using the embedding of $\alpha \in \mathbb{F}_p$ into the general linear group $GL_k(\mathbb{F}_p)$. However, its complexity for $k \geq 3$ was not clearly understood. Recently, Kim, Cheon and Lee [14] realized that Satoh's generalization is essentially the same as the embedding of \mathbb{F}_p into \mathbb{F}_{p^k} , and clarified the complexity of the algorithm. Unfortunately, their result suggests that the complexity of this generalization is not faster than the current square-root complexity algorithm such as Pollard's rho algorithm [19] for $k \geq 3$.

Cheon, Kim and Song [8] recently presented an algorithm to solve α when neither $p+1$ nor $p-1$ has an appropriate small divisor d . Precisely, they solve α when $g^{\alpha^{k_1}}, \dots, g^{\alpha^{k_d}}$ are given where the set of k_i 's forms a multiplicative subgroup of \mathbb{Z}_{p-1}^\times . However, currently the reduction of the DLPwAI to this problem is not known.

1.2. Our contributions. We present a new algorithm to solve the DLPwAI. The proposed algorithm has a running time less than $\tilde{O}(p^{1/2})$ in any case of $d < p^{1/2}$. We briefly describe the algorithm in the followings. One chooses a proper polynomial f of degree d . Then one randomly chooses elements r_i and s_j from \mathbb{F}_p and computes two lists (the value of m will be determined later)

$$L_1 := \{g^{f(r_i\alpha)} : r_i \in \mathbb{F}_p, 1 \leq i \leq m\} \text{ and } L_2 := \{g^{f(s_j)} : s_j \in \mathbb{F}_p, 1 \leq j \leq m\}.$$

If the two lists have an element in common, say $f(r_{i_0}\alpha) = f(s_{j_0})$, then finding roots of $\hat{f}(\alpha) := f(r_{i_0}\alpha) - f(s_{j_0})$ in \mathbb{F}_p gives d candidates for the desired solution.

For the refinement of the complexity, we consider several problems. As a first step, computing the list L_1 can be considered as the problem to compute multipoint evaluation when coefficients of a polynomial are exponentiated, that is, to compute $g^{q(r_1)}, \dots, g^{q(r_m)}$ for given g^{q_0}, \dots, g^{q_d} , where $q(x) := q_0 + q_1x + \dots + q_dx^d$. The naive approach would take $O(m \cdot d)$ operations in G . However, we obtain a fast multipoint evaluation method to compute it in $\tilde{O}(m+d)$ group operations using the usual fast multipoint evaluation method. A similar result was proposed in [17] using the fast Fourier transform multiplication. We note that the technique is also extended to the Schönhage-Strassen multiplication algorithm.

If the size of the image of f is N , then the birthday paradox (under the assumption that f is a random function) suggests that the lists L_1 and L_2 have a collision with high probability for $m = O(N^{1/2})$. To obtain a more precise collision probability, we consider a non-uniform birthday problem. Suppose that there exist N bins and a randomly sampled ball is assigned to the bin $k \in \{1, 2, \dots, N\}$ with a probability w_k . Our analysis shows that the probability of a bin containing at least two different balls after r samplings is non-negligible for $r \geq \frac{1}{\sqrt{\sum_{k=1}^N w_k^2}}$.

Applying the result to our case, the two lists have a collision with high probability after $O\left(\frac{1}{\sqrt{\sum_{k=1}^N w_k^2}}\right)$ samplings.

There have been several contexts [9, 18, 21] dealing with the birthday problem of non-uniform distributions. Although they precisely determine the expected number of trials until a collision, their results only apply when the probability w_k is bounded by c/N for some constant c independent with N . We remark that our analysis applies even when w_k is not well-bounded, e.g. $w_k = N^{-O(1)}$.

Let ρ_f be the number of rational points over \mathbb{F}_p on the curve defined by $f(x) - f(y) = 0$. Then, as in [15, 10], we can see that $\sum_k w_k^2 = \frac{\rho_f}{p^2}$. From this, we derive that the overall complexity of the proposed algorithm is given by $\tilde{O}\left(\sqrt{p^2/\rho_f} + d\right)$ group exponentiations. By Weil's theorem, we have $\rho_f = \tau_f p \pm O(d^2\sqrt{p})$, where τ_f is the number of absolutely irreducible (that is both irreducible over \mathbb{F}_p and its algebraic closure) factors of $f(x) - f(y)$. To obtain a better complexity, we need to find a polynomial $f \in \mathbb{F}_p[x]$ with the large number of τ_f as possible.

We show that τ_f is at most $\sum_{D|d} \frac{\varphi(D)}{\text{ord}_D(p)}$, where $\text{ord}_D(p)$ is the multiplicative order of p modulo D . In particular, in the case of $d|\Phi_k(p)$ for the prime k , we have $\tau_f \leq \frac{d - \gcd(d, p-1)}{k} + \gcd(d, p-1)$. When $d|(p-1)$, one has $\tau_f = d$ for the polynomial $f(x) = x^d$ and $f(x) - f(y)$ factors into all linear factors. In the case of $d|(p+1)$, one has $\tau_D = \frac{d - \gcd(d, p-1)}{2} + \gcd(d, p-1)$ for the Dickson polynomial $D(x)$ and $D(x) - D(y)$ factors into all quadratics except one or two linear(s). Applying the proposed algorithm, it takes $\tilde{O}(\sqrt{p/d} + d)$ group exponentiations to compute the discrete logarithm α . In the case of $d|\Phi_3(p) = (p^2 + p + 1)$, we show that $f(x) - f(y)$ cannot have an absolutely irreducible cubic factor for any polynomial f , thus it is impossible to achieve the upper bound of τ_f in the case.

The rest of the paper is organized as follows: We begin with the description of the algorithm and present complexity analysis in Section 2. In Section 3, we present a fast multipoint evaluation method on exponents. The analysis of the birthday problem with a non-uniform distribution is presented in Section 4. In Section 5, we discuss on the choices of polynomials that attain the proposed upper bound of τ_f . We summarize the results and suggest future problems in Section 6.

2. THE MAIN ALGORITHM

In this section, we present an algorithm to solve the DLPwAI with a function defined by a polynomial $f \in \mathbb{F}_p[x]$. Throughout the paper, $\Phi_k(x)$ denotes the k -th cyclotomic polynomial and $\varphi(k)$ is the Euler-totient function.

2.1. Algorithm description. Let $G = \langle g \rangle$ be a group of prime order p . The problem is to solve α for given $g, g^\alpha, \dots, g^{\alpha^d} \in G$. Cheon's algorithm and its generalizations use an embedding of the discrete logarithm $\alpha \in \mathbb{F}_p$ to auxiliary groups such as extension fields of \mathbb{F}_p . However, by the recent result of Kim, Cheon and Lee [14], the complexity of the several generalizations in the case of $d|\Phi_k(p)$ for $k \geq 3$ [20, 7] is always greater than $p^{1/2}$. Thus we need to consider a different approach to solve the DLPwAI.

To begin with, we choose a polynomial $f \in \mathbb{F}_p[x]$ of degree d . The proposed algorithm uses a map defined by the polynomial f . While the previous algorithms

require an algebraic structures of the auxiliary groups, we solely concentrate on the value set of the polynomial f . The brief description of the algorithm is as follows.

Step 1: For given $f \in \mathbb{F}_p[x]$ and $g, g^\alpha, \dots, g^{\alpha^d} \in G$, one computes two lists

$$L_1 := \{g^{f(r_i\alpha)} : r_i \in \mathbb{F}_p, 1 \leq i \leq m\} \text{ and } L_2 := \{g^{f(s_j)} : s_j \in \mathbb{F}_p, 1 \leq j \leq m\},$$

where r_i and s_j are randomly chosen from \mathbb{F}_p and m is a positive integer determined later.

Step 2: Find a non-empty intersection between L_1 and L_2 , if it exists. If not, repeat Step 1.

Step 3: One recovers α by finding roots of $\tilde{f}(\alpha) := f(r_{i_0}\alpha) - f(s_{j_0})$ in \mathbb{F}_p and using (g, g^α) to identify α .

We take a close look into the complexity of the proposed algorithm in the next subsection.

2.2. Complexity analysis. Consider a naive analysis of the algorithm. First, suppose that the value set $V(f) := \{f(x) : x \in \mathbb{F}_p\}$ is of size N . Assume that the map $x \mapsto f(x)$ behaves as a random function. Then by the birthday paradox, we expect the lists L_1 and L_2 have an element in common for $m = O(N^{1/2})$ with high probability. Next, a naive approach to compute L_1 would take $O(md)$ exponentiations in G . Overall, the complexity of the algorithm is at least $\Omega(N^{1/2})$. However, for a random polynomial of degree d , the size of the value set of f is about $\left(1 - \frac{1}{2!} + \dots + \frac{(-1)^{d-1}}{d!}\right) \cdot p \approx \left(1 - \frac{1}{e}\right) \cdot p$ on average [22], where e denotes the base of the natural logarithm. Thus the complexity is already greater than $\Omega(p^{1/2})$.

To obtain the better complexity of the algorithm, we should consider several problems. The following theorem shows how to compute L_1 in $\tilde{O}(m)$ exponentiations in G .

Theorem 2.1. *Suppose that an algorithm that multiplies two polynomials of degree less than d has a running time of $M(d)$ operations in \mathbb{F}_p . Let $f(x) = a_0 + a_1x + \dots + a_{d-1}x^{d-1} \in \mathbb{F}_p[x]$. Given $g^{f(x)} := (g^{a_0}, g^{a_1}, \dots, g^{a_{d-1}})$ and random $r_1, \dots, r_d \in \mathbb{F}_p$, one computes $g^{f(r_0)}, \dots, g^{f(r_{d-1})}$ in $O(M(d) \log d)$ operations in G .*

Proof. The sketch of the proof will be given in Section 3. □

We also observe that the map defined by the polynomial is not a random function in general. In the case, as opposed to the random case, the values of the map are non-uniformly distributed. Intuitively, one might have more collisions for value sets in the non-uniform case. This leads us to consider the birthday problem with non-uniform distribution. The following result is a simpler than the problem studied in the papers [9, 18, 21], but it is enough for our applications.

Theorem 2.2. *For a positive integer N and $k \in \{1, 2, \dots, N\}$, let w_k be the probability that a randomly sampled ball is put into the bin k . Let S_r be the probability that a collision occurs in r trials. Assume that $W = \max_k \{w_k\} \leq \frac{1}{4}$. Let $D = \frac{1}{\sum_k w_k^2}$.*

If $r \geq \sqrt{D + \frac{1}{4}} + \frac{1}{2}$, then $S_r \geq \frac{1}{64}$.

Proof. The proof will be given in Section 4. □

Let $V(f) := \{f(x) : x \in \mathbb{F}_p\} = \{y_1, \dots, y_N\}$ be the value set of a polynomial $f \in \mathbb{F}_p[x]$ of degree d . Let $R_i := |\{y \in V(f) : |f^{-1}(y)| = i\}|$ and $\rho_f := |\{(x, y) \in \mathbb{F}_p \times \mathbb{F}_p : f(x) = f(y)\}|$, then we have the following equalities,

$$p = \sum_{i=1}^d iR_i, |V(f)| = \sum_{i=1}^d R_i, \text{ and } \rho_f = \sum_{i=1}^d i^2 R_i,$$

and we can see that $p \leq \rho_f \leq dp$.

Using the above theorems, we can obtain our main theorem. It shows that the proposed algorithm has a running time of $\tilde{O}\left(\frac{p}{\sqrt{\rho_f}} + d\right)$ group operations that is always less than $\tilde{O}(p^{1/2})$.

Theorem 2.3. *Let notations described above. Let $f(x) = a_0 + a_1x + \dots + a_dx^d \in \mathbb{F}_p[x]$ be a polynomial of degree d . Let $G = \langle g \rangle$ be a cyclic group of prime order p . Suppose that an algorithm that multiply two polynomials less than degree d has a running time of $M(d)$ operations in \mathbb{F}_p . For given $f(x)$ and $g^\alpha, \dots, g^{\alpha^d}$, one computes α in an expected number of $O\left(\frac{p}{\sqrt{\rho_f}} \cdot \frac{M(d) \log d}{d} + d\right)$ operations in G and an expected number of $O(M(d) \log d \log(dp))$ operations in \mathbb{F}_p .*

Proof. Suppose that one evaluates f at a random element $r_i \in \mathbb{F}_p$ repeatedly until a collision, $f(r_i) = f(r_j)$ for $i \neq j$, occurs. This event can be regarded as a non-uniform birthday problem in the sense that the ball r_i is put into the bin $f(r_i)$. Let w_k be the probability that a ball is thrown into the bin k . Then each probability is given by (after proper reordering)

$$(w_{y_1}, \dots, w_{y_N}) = \left(\underbrace{\frac{1}{p}, \dots, \frac{1}{p}}_{R_1}, \underbrace{\frac{2}{p}, \dots, \frac{2}{p}}_{R_2}, \dots, \underbrace{\frac{d}{p}, \dots, \frac{d}{p}}_{R_d} \right).$$

By Theorem 2.2, one obtains a collision with high probability after $O\left(\sqrt{1/\sum_k w_k^2}\right) = O(\sqrt{p^2/\rho_f})$ random elements are sampled. The overall complexity is analyzed in the following steps.

- (1) Given $f(x) = a_0 + a_1x + \dots + a_dx^d$ and $g, g^\alpha, \dots, g^{\alpha^d}$, one computes $g^{a_0}, (g^\alpha)^{a_1}, \dots, (g^{\alpha^d})^{a_d}$ in $(d+1)$ group exponentiations. Given $f_\alpha(x) := f(\alpha x) = a_0 + (\alpha \cdot a_1)x + (\alpha^2 \cdot a_2)x^2 + \dots + (\alpha^d \cdot a_d)x^d$, we denote by $g^{f_\alpha(x)} = (g^{a_0}, (g^\alpha)^{a_1}, \dots, (g^{\alpha^d})^{a_d})$.
- (2) Let $m := O(\sqrt{p^2/\rho_f})$. One chooses random elements $r_1, \dots, r_m \in \mathbb{F}_p$, and computes $g^{f_\alpha(r_1)}, \dots, g^{f_\alpha(r_m)}$ in $\lceil \frac{m}{d} \rceil \cdot O(M(d) \log d)$ group operations in G by Theorem 2.1.
- (3) One chooses random elements $s_1, \dots, s_m \in \mathbb{F}_p$ and computes $f(s_1), \dots, f(s_m)$ using the standard fast multipoint evaluation method in $\lceil \frac{m}{d} \rceil \cdot O(M(d) \log d)$ operations in \mathbb{F}_p .
- (4) One raises to the power of $f(s_i)$ for each $i = 1, \dots, m$ to obtain $g^{f(s_1)}, \dots, g^{f(s_m)}$ requiring m group exponentiations.
- (5) One finds indices i and j satisfying $g^{f_\alpha(r_i)} = g^{f(s_j)}$, which is guaranteed by Theorem 2.2 with non-negligible probability. If there is no such element then one repeats the steps from 2 to 4.

- (6) One computes at most d candidates for α by finding roots in \mathbb{F}_p of

$$\begin{aligned}\tilde{f}(\alpha) &:= f_\alpha(r_i) - f(s_j) \\ &= (a_1 r_i) \alpha + (a_1 r_i^2) \alpha^2 + \cdots + (a_d r_i^d) \alpha^d - (a_1 s_j + \cdots + a_d s_j^d).\end{aligned}$$

It takes an expected number of $O(M(d) \log d \log(dp))$ operations in \mathbb{F}_p using the root finding algorithm [23, Corollary 14.16].

- (7) One identifies the exact solution α from d candidates exhaustively using the information of (g, g^α) which takes d group operations. □

Remark 2.4. The multiplication cost $M(d)$ is $O(d \log d)$ when using the FFT method and $O(d \log d \log d)$ when using the Schönhage-Strassen method. In both cases, the complexity of the proposed algorithm is bounded by $\tilde{O}\left(\sqrt{p^2/\rho_f} + d\right)$ operations in G without the log factors $\log d \log \log d$ for the FFT method (or, $\log^2 d \log \log d$ for the SS method).

In the following sections, we will present the omitted proofs and discuss on several polynomials suited for the proposed algorithm.

3. FAST MULTIPOINT EVALUATION ON EXPONENTS

In this section, we discuss on the polynomial evaluation method when the coefficients of a polynomial are exponentiated, that is, to compute $g^{f(r_1)}, \dots, g^{f(r_d)}$ when $g^{f(x)} := (g^{a_0}, \dots, g^{a_{d-1}})$ is given for a polynomial $f(x) = a_{d-1}x^{d-1} + \cdots + a_1x + a_0 \in \mathbb{F}_p[x]$.

Given $g^{f(x)}$ and $h(x)$, where $f(x)$ and $h(x)$ are polynomials of degree less than d , one can compute $g^{f(x)h(x)}$ and $g^{f(x)+h(x)}$ in $O(d^2)$ and $O(d)$ exponentiations in G . Furthermore, one can apply the fast multiplication method such as the fast Fourier transform (FFT) method or the Schönhage-Strassen (SS) method to compute $g^{f(x)h(x)}$ in $\tilde{O}(d)$ exponentiations in G .

From the observations, it is easy to obtain a fast multipoint evaluation method on the exponentiated elements. A similar method is used in [17], where the coefficients of the polynomial being encrypted by an additive homomorphic encryption scheme. He showed that the evaluation costs $O(M(d) \log d)$ homomorphic operations (additions and scalar multiplications), where $M(d)$ is the computational cost of the FFT multiplication. It follows from the observation that the FFT multiplication algorithm analogously applies to compute $\text{Enc}(f \cdot \tilde{f})$ for given $\text{Enc}(f)$ and \tilde{f} in $M(d)$ homomorphic operations. Here, Enc is the additive homomorphic encryption and $\text{Enc}(f) := (\text{Enc}(a_0), \dots, \text{Enc}(a_{d-1}))$. The technique also applies to our case simply replacing $\text{Enc}(a_i)$ with g^{a_i} .

The FFT multiplication, however, only works when \mathbb{F}_p contains a d -th root of unity, i.e. $d|(p-1)$. In our application, $(p-1)$ does not necessarily have a proper divisor d , so we need to remark that the multipoint evaluation on the exponentiated elements is also possible using the SS multiplication method. In the SS multiplication, the field \mathbb{F}_p can be arbitrary.

We briefly describe the algorithm in the following.

3.1. Schönhage-Strassen multiplications. Suppose that $\deg(fh) \leq d = 2^k$ for $m = 2^{\lfloor k/2 \rfloor}$ and $t = d/m$. Write down the polynomial as $f(x) = A_0(x) + A_1(x)x^m + \dots + A_{t-1}(x)x^{m(t-1)}$ where $A_i \in \mathbb{F}_p[x]$ with degree less than m and let $\bar{f}(x, y) := A_0(x) + A_1(x)y + \dots + A_{t-1}(x)y^{t-1} \in \mathbb{F}_p[x, y]$ so that $\bar{f}(x, x^m) = f(x)$.

Consider the ring $D := \mathbb{F}_p[x]/(x^{2m} + 1)$ and let $\zeta := x \bmod (x^{2m} + 1) \in D$ be an element corresponding to x in $\mathbb{F}_p[x]/(x^{2m} + 1)$. Then we can regard $f^*(y) := \bar{f}(\zeta, y) = A_0(\zeta) + A_1(\zeta)y + \dots + A_{t-1}(\zeta)y^{t-1}$ as a polynomial in y with the coefficients in D . For two polynomials f and h , the SS multiplication computes $f^*(y)h^*(y) \bmod y^t + 1$ that is equivalent to $f(x)h(x) \bmod x^d + 1$.

Since $\zeta^{2m} = -1$, ζ is a $4m$ -th primitive root of unity in D , so $\eta = \zeta^2$ (or $\eta = \zeta$) is a primitive $2t$ -th root of unity in D , when $t = m$ (or $t = 2m$, respectively). Now $f^*(y)h^*(y) \bmod (y^t + 1)$ is equivalent to compute $f^*(\eta y)h^*(\eta y) \bmod (y^t - 1)$. It can be done by the fast Fourier transform method with the t -th primitive root of unity $\omega = \eta^2$ in D . The multiplication in D can be done recursively with polynomials of degree less than $2m$. We simply write $g^{f(x)} = (g^{a_0}, g^{a_1}, \dots, g^{a_{d-1}}) = (g^{A_0}, \dots, g^{A_{t-1}})$, where $g^{A_i} = (g^{a_{mi}}, g^{a_{mi+1}}, \dots, g^{a_{mi+(m-1)}})$.

Algorithm 1 Schönhage-Strassen Multiplication (in exponential form)

Input: $d = 2^k \in \mathbb{N}$, an element g of order p , $(g^{a_0}, g^{a_1}, \dots, g^{a_{d-1}})$ and (b_0, \dots, b_{d-1}) where $f(x) = a_0 + \dots + a_{d-1}x^{d-1}$ and $h(x) = b_0 + \dots + b_{d-1}x^{d-1}$ with $\deg(fh) < d$

Output: $g^{f(x)h(x)} := (g^{c_0}, g^{c_1}, \dots, g^{c_{d-1}}) \in G^d$

- (1) $m \leftarrow 2^{\lfloor k/2 \rfloor}$, $t \leftarrow d/m$
 let $g^{f(x)} = (g^{A_0}, \dots, g^{A_{t-1}})$ and $h(x) = (B_0, \dots, B_{t-1})$ so that $f(x) = \sum_{i=0}^{t-1} A_i(x)x^{mi}$, $h(x) = \sum_{i=0}^{t-1} B_i(x)x^{mi}$ where $A_i, B_j \in \mathbb{F}_p[x]$ of degree less than m .
 - (2) let $D = \mathbb{F}_p[x]/(x^{2m} + 1)$ and $\zeta \leftarrow x \bmod (x^{2m} + 1)$
 if $t = 2m$ then $\eta \leftarrow \zeta$, otherwise $\eta \leftarrow \zeta^2$ (η is a primitive $2t$ -th root of unity)
 compute $g^{c^*(\eta y)} = g^{f^*(\eta y)h^*(\eta y) \bmod (y^t - 1)}$ with a t -th root of unity η^2 using the fast Fourier transform method as described in [17]
 call the algorithm 1 recursively to compute multiplications in D
 - (3) return $g^{c^*(y)} = (g^{C_0}, \dots, g^{C_{t-1}})$
-

sketch proof of Theorem 2.1. The analysis of the complexity easily follows by replacing the addition/multiplication in the field \mathbb{F}_p with the multiplication/exponentiation in the group G . In the case of using FFT multiplication, we refer to [17]. The original SS multiplication takes $O(d \log d \log \log d)$ operations in \mathbb{F}_p , so the SS multiplication in the exponential form requires $O(d \log d \log \log d)$ operations in G . The multipoint evaluation method in the exponential form using SS multiplication takes $O(d \log^2 d \log \log d)$ operations in G . \square

4. GENERALIZED BIRTHDAY PROBLEM: NON-UNIFORM DISTRIBUTION

Consider a function $f(x)$ on \mathbb{F}_p with image size N . If one evaluates $f(x)$ at a random point repeatedly, one eventually has a collision $f(x_i) = f(x_j)$ for $i \neq j$ since its image is finite. Assuming that f behaves like a random function, the collision occurs in $O(\sqrt{N})$ steps with high probability by the birthday paradox.

The randomness of the function requires the preimage of f to be equi-sized. It is not always the case, if the function is given by a random polynomial of degree d . For the efficiency of our algorithm, we hope to find a collision faster than $O(\sqrt{N})$. It leads us to consider a birthday problem that applies when the sampling probability is not uniformly distributed.

Suppose that we have N bins numbered from 1 to N . For $k = 1, 2, \dots, N$, let w_k be a probability that a randomly sampled ball is put into the bin k . We are interested in finding the probability of a bin containing at least two different balls.

A number of contexts discuss on this kind of problem [9, 18, 21]. They precisely determine the expected number of the trials until a collision, it is given by $\sqrt{\frac{\pi}{2 \sum_k w_k^2}} + O(N^{1/4})$. However, their analysis only applies when the probability w_k is bounded by c/N , where c is a constant independent of N . For our case, the probability w_k can be up to d/N , where $d = N^{1/3}$. Thus we present an analysis of a non-uniform birthday problem of which the probabilities are not well bounded. Our analysis shows that a collision occurs with non-negligible probability in $O\left(\sqrt{1/\sum_k w_k^2}\right)$ samplings for any probability distribution of w_k .

Let S_r be the probability that a collision occurs in r trials. Define $E_k^{(r)}$ by an event that a collision occurs in the bin k after r trials. Then, by Bonferroni inequality and the inclusion-exclusion principle, we have

$$\begin{aligned} S_r &= \Pr(E_1^{(r)} \cup \dots \cup E_N^{(r)}) = \sum_{i=1}^N (-1)^{i+1} \sum_{1 \leq k_1 < k_2 < \dots < k_i \leq N} \Pr(E_{k_1}^{(r)} \cap \dots \cap E_{k_i}^{(r)}) \\ &\geq \sum_{k=1}^N \Pr(E_k^{(r)}) - \sum_{1 \leq k < \ell \leq N} \Pr(E_k^{(r)} \cap E_\ell^{(r)}). \end{aligned}$$

Unless there is no ambiguity, we will omit the superscript (r) in $E_k^{(r)}$. We first determine a lower bound for $\Pr(E_k)$.

Definition 4.1. For $k \in \{1, 2, \dots, N\}$, $B_{r,k}^{(i)}$ is the set of vectors $\vec{b} = (b_1, \dots, b_r) \in \{1, 2, \dots, N\}^r$ such that the number of b_j 's satisfying $b_j = k$ is equal to i .

Lemma 4.2. For a positive integer N and $k \in \{1, 2, \dots, N\}$, let w_k be the probability that a randomly sampled ball is put into the bin k . Let E_k be the event that the bin k contains at least two different balls after $r \geq 2$ samplings. Then the probability of E_k is lower bounded by

$$\Pr(E_k) \geq \frac{(r-1)r}{2} \cdot w_k^2 \left\{ 1 - (r-1) \left(1 - \frac{2}{r} \right) w_k \right\}.$$

Proof. With the notations in Definition 4.1, we have

$$\Pr(E_k) = \sum_{i \geq 2} \sum_{\vec{b} \in B_{r,k}^{(i)}} w_{b_1} \cdots w_{b_r} = 1 - \left(\sum_{\vec{b} \in B_{r,k}^{(1)}} w_{b_1} \cdots w_{b_r} + \sum_{\vec{b} \in B_{r,k}^{(0)}} w_{b_1} \cdots w_{b_r} \right)$$

The summation $\sum_{\vec{b} \in B_{r,k}^{(1)}} w_{b_1} \cdots w_{b_r}$ is the probability that only one ball is put into the bin k until r trials, so we have $\sum_{\vec{b} \in B_{r,k}^{(1)}} w_{b_1} \cdots w_{b_r} = r \cdot w_k \cdot (1 - w_k)^{r-1}$. Similarly, we have $\sum_{\vec{b} \in B_{r,k}^{(0)}} w_{b_1} \cdots w_{b_r} = (1 - w_k)^r$ since it is the probability that

no ball is thrown to the bin k until r trials. It follows that

$$\Pr(E_k) = 1 - (r \cdot w_k \cdot (1 - w_k)^{r-1} + (1 - w_k)^r) = 1 - (1 - w_k)^{r-1} \cdot (1 + (r-1)w_k).$$

On the other hand, for $r \geq 2$, we have

$$\begin{aligned} 1 - (1 - w_k)^{r-1} \cdot (1 + (r-1)w_k) &\geq 1 - \left(1 - (r-1)w_k + \binom{r-1}{2} w_k^2\right) \cdot (1 + (r-1)w_k) \\ &\geq \frac{(r-1)r}{2} \cdot w_k^2 - \frac{(r-1)^2(r-2)}{2} \cdot w_k^3 \\ &= \frac{(r-1)r}{2} \cdot w_k^2 \left\{1 - (r-1) \left(1 - \frac{2}{r}\right) w_k\right\}. \end{aligned}$$

In the first inequality, we used that $(1-x)^n \leq 1 - nx + \binom{n}{2}x^2$ for $0 \leq x \leq 1$ and $n \geq 1$. \square

Now let us consider an upper bound of $\Pr(E_k \cap E_\ell)$.

Lemma 4.3. *Let the notations as in Lemma 4.2. Then we have*

$$\Pr(E_k \cap E_\ell) = \sum_{i,j \geq 2} \sum_{\vec{b} \in B_{r,k}^{(i)} \cap B_{r,\ell}^{(j)}} w_{b_1} \cdots w_{b_r} \leq \binom{r}{2} \cdot \binom{r-2}{2} \cdot w_k^2 \cdot w_\ell^2.$$

Proof. For $\vec{b} = (b_1, \dots, b_r) \in B_{r,k}^{(i)} \cap B_{r,\ell}^{(j)}$ with $i \geq 2, j \geq 2$, there exist $i_1 \neq i_2$ and $j_1 \neq j_2$ such that $b_{i_1} = b_{i_2} = k$ and $b_{j_1} = b_{j_2} = \ell$. In that case, we have $w_{b_1} \cdots w_{b_r} \leq w_k^2 \cdot w_\ell^2$. The value $\binom{r}{2}$ indicates the possible number of two positions for k and $\binom{r-2}{2}$ stands for the possible number of the other two positions for ℓ . \square

From the above results, we prove Theorem 2.2.

proof of Theorem 2.2. Assume that $r \leq \frac{1}{2W}$. Then we have $(r-1) \left(1 - \frac{2}{r}\right) w_k \leq (r-1)w_k \leq \frac{r-1}{2r} \leq \frac{1}{2}$ for all k . It yields $\Pr(E_k) \geq \frac{(r-1)r}{4} \cdot w_k^2$ using Lemma 4.2. Let $B(r) := \frac{(r-1)r}{2} \sum_{k=1}^N w_k^2$. Then we have the followings

(4.1)

$$\begin{aligned} S_r &\geq \sum_{1 \leq k \leq N} \Pr(E_k) - \sum_{1 \leq k < \ell \leq N} \Pr(E_k \cap E_\ell) \geq \frac{B(r)}{2} - \frac{r^2(r-1)^2}{4} \cdot \sum_{1 \leq k < \ell \leq N} w_k^2 w_\ell^2 \\ &= \frac{B(r)}{2} - \frac{r^2(r-1)^2}{8} \cdot \left\{ \left(\sum_{1 \leq k \leq N} w_k^2 \right)^2 - \left(\sum_{1 \leq k \leq N} w_k^4 \right) \right\} \geq \frac{B(r)}{2} - \frac{B(r)^2}{2}. \end{aligned}$$

The last term, $\frac{B(r)}{2}(1 - B(r))$, is maximized by $\frac{1}{8}$ when $r = r_0$ such that $B(r_0) = \frac{1}{2}$, i.e. $r_0(r_0 - 1) = D$, equivalently, $r_0 = \sqrt{D + \frac{1}{4}} + \frac{1}{2}$. If $r_0 \leq \frac{1}{2W}$, then the above inequality holds, so we have $S_{r_0} \geq \frac{1}{8}$. Since S_r increases as r grows, we have $S_r \geq \frac{1}{8}$ for $r \geq r_0$.

On the contrary, if $\frac{1}{2W} \leq r_0$, we have

$$\frac{1}{2} \geq [B(r)]_{r=1/2W} = \left(\frac{1}{8W^2} - \frac{1}{4W} \right) \cdot \sum_k w_k^2 \geq \left(\frac{1}{8W^2} - \frac{1}{4W} \right) \cdot W^2 = \frac{1}{8} - \frac{W}{4} \geq \frac{1}{16},$$

where the first inequality comes from $B(r) \leq \frac{1}{2}$ if and only if $r(r-1) \leq \frac{1}{\sum_k w_k^2}$ if and only if $\frac{1}{2} - \sqrt{D + \frac{1}{4}} \leq r \leq \frac{1}{2} + \sqrt{D + \frac{1}{4}}$. Since $\frac{1}{2}B(1-B)$ is increasing for $B \leq \frac{1}{2}$, we have $S_{1/2W} \geq [\frac{1}{2}B(1-B)]_{r=1/2W} \geq \frac{1}{2} \cdot \frac{1}{16} (1 - \frac{1}{16}) \geq \frac{1}{64}$ (Note that the inequality (4.1) also holds when $r = \frac{1}{2W}$). Thus, we have $S_{r_0} \geq S_{1/2W} \geq \frac{1}{64}$. \square

The above theorem shows that a collision occurs with non-negligible probability after about $\frac{1}{\sqrt{\sum_k w_k^2}}$ trials although the probabilities are arbitrarily distributed. We consider several examples in the followings.

Example 4.4. In the case of $w_k = O(\frac{1}{N})$ for all k , a collision occurs with non-negligible probability after $\Omega(\sqrt{N})$ trials as the usual birthday paradox.

Theorem 2.2 asserts that a collision occurs with high probability after $O\left(\sqrt{\frac{1}{\sum_k w_k^2}}\right)$ trials. Precisely, in [9], they derived the expected number of trials until a collision when $w_k = O(\frac{1}{N})$, it is $\sqrt{\frac{\pi}{2\sum_k w_k^2}} + O(N^{1/4})$ as $N \rightarrow \infty$.

Example 4.5. The proposed theorem also applies even when $w_k = \Omega(\frac{1}{N})$ for some k . In the proof of Theorem 2.2, we have shown that $S_{1/2W} \geq \frac{1}{64}$ for $W = \max_k\{w_k\}$. It is meaningful, when $W = \Omega\left(\frac{1}{\sqrt{N}}\right)$, since a collision is guaranteed in $O\left(\frac{1}{W}\right)$ trials which is faster than the usual expectation of the birthday paradox.

Example 4.6. Consider the birthday problem given by a polynomial $f \in \mathbb{F}_p[x]$ as in Theorem 2.3. Suppose that the probabilities are given by

$$(w_1, \dots, w_v) = \left(\underbrace{\frac{1}{p}, \dots, \frac{1}{p}}_{R_1}, \underbrace{\frac{2}{p}, \dots, \frac{2}{p}}_{R_2}, \dots, \underbrace{\frac{d}{p}, \dots, \frac{d}{p}}_{R_d} \right).$$

The size of the value set of f is $\sum_i R_i$, and rough estimation of the birthday paradox suggests a collision occurs in $O\left(\sqrt{\sum_i R_i}\right)$, but a collision can be found in $O\left(\frac{N}{\sqrt{\sum_i i^2 R_i}}\right) \leq O\left(\sqrt{\sum_i R_i}\right)$ by Theorem 2.2. The inequality comes from the Cauchy-Schwartz inequality.

5. POLYNOMIALS FOR THE PROPOSED ALGORITHM

In the rest of the paper, we assume that d is relatively prime to p .

5.1. Substitution polynomials. Let $f(x, y) \in \mathbb{F}[x, y]$ be an irreducible bivariate polynomial defined over a field \mathbb{F} . The polynomial f is said to be *absolutely irreducible* if it is also irreducible over the algebraic closure. For a polynomial $f(x)$, one defines *substitution polynomial of f* as the bivariate polynomial $f(x) - f(y)$.

For the efficiency of the algorithm, one needs a polynomial f with large value of ρ_f as possible. In the following lemma, we observe that ρ_f is closely related to the number of absolutely irreducible factors of the substitution polynomial $f(x) - f(y)$.

Lemma 5.1 (Weil's bound [24]). *Let $f \in \mathbb{F}_p[x]$ be a polynomial of degree d . Let τ_f be the number of absolutely irreducible factors of the substitution polynomial of f . Then we have $\tau_f p - d^2 \sqrt{p} \leq \rho_f \leq \tau_f p + d^2 \sqrt{p}$.*

By Weil's bound, for $d < p^{1/4}$, the complexity of the proposed algorithm becomes

$$\tilde{O}\left(\sqrt{p^2/\rho_f + d}\right) \sim \tilde{O}\left(\sqrt{p/\tau_f + d}\right).$$

Thus, Lemma 5.1 reduces the proposed algorithm of finding a polynomial whose substitution polynomial has many absolutely irreducible factors as possible. We shall discuss on the upper bound of the number τ_f and try to find polynomials that attain this bound in the following subsections.

5.2. An upper bound of the number of absolutely irreducible factors. We observe that there exist polynomials of which the substitution polynomial factors into all linear absolutely irreducible factors (i.e. $\tau_f = d$) when $d|\Phi_1(p) = (p-1)$, or all quadratics but one or two linear(s) (i.e. $\tau_f \approx \frac{d}{2}$) when $d|\Phi_2(p) = (p+1)$ [10, 15]. From this observation, we attempt to find a polynomial, in the case of $d|\Phi_k(p)$, whose substitution polynomial factors into all k -degree factors except a few small degree factors. We show that the substitution polynomial of any polynomial cannot have an absolutely irreducible cubic factors in the case of $k = 3$ using the same idea used in the previous papers such as [13, 10, 11, 12]. This shows that we cannot achieve $\tau_f \approx \frac{d}{3}$ in the case of $d|\Phi_3(p) = (p^2 + p + 1)$.

Assume that the factorization of $f(x) - f(y)$ into irreducibles over \mathbb{F}_p be given by

$$f(x) - f(y) = g_1(x, y) \cdots g_s(x, y).$$

Let $g_i(x, y) = h_{i,d_i} + h_{i,d_i-1} + \cdots + h_{i,1} + h_{i,0}$, where $h_{i,j} \in \mathbb{F}_p[x, y]$ is the homogenous part of degree j in $g_i(x, y)$, and d_i denotes the highest degree of $g_i(x, y)$.

Throughout this section, we simply write τ instead of τ_f . As an independent interest, we also give another proofs of Lemma 5.2 and Theorem 5.3 in Appendix A.

Lemma 5.2. *Let d be a positive integer dividing $\Phi_k(p)$ for prime k . Let ζ be a primitive d -th root of unity in \mathbb{F}_{p^k} . Then we have the followings: either $\zeta^i \in \mathbb{F}_p \setminus \mathbb{F}_p$ for all $i \not\equiv 0 \pmod{d}$ if $d \equiv 1 \pmod{k}$, or only $\zeta^{(i/k) \cdot d}$ for $i = 0, 1, \dots, k-1$ are in \mathbb{F}_p if $d \equiv 0 \pmod{k}$. Note that there exists no positive integer d dividing $\Phi_k(p)$ if $d \not\equiv 0, 1 \pmod{k}$.*

Proof. Note that $\zeta^i \in \mathbb{F}_p$ if and only if $\zeta^{i(p-1)} = 1$ if and only if $i(p-1) \equiv 0 \pmod{d}$. The number of such i is equal to $\gcd(d, p-1)$. The value of $\gcd(d, p-1)$ divides

$$\gcd(\Phi_k(p), p-1) = \gcd(p^{k-1} + \cdots + p + 1, p-1) = \gcd(p-1, k)$$

which only can be 1 or k for prime k .

If $d \equiv 0 \pmod{k}$, then $\gcd(d, p-1) = k$ and all the k -th roots of unity, $\zeta^{(i/k) \cdot d}$, lie in \mathbb{F}_p . If $d \not\equiv 0 \pmod{k}$, then $\gcd(d, p-1) = 1$. Thus only $\zeta^0 = 1$ lie in \mathbb{F}_p in that case and all $\zeta^i \in \mathbb{F}_{p^k} \setminus \mathbb{F}_p$ for $i \neq 0$ must form conjugate k -tuples

$$\left(\zeta^i, (\zeta^i)^p, (\zeta^i)^{p^2}, \dots, (\zeta^i)^{p^{k-1}}\right),$$

which is possible only when $d-1 \equiv 0 \pmod{k}$. Otherwise, d cannot divide $\Phi_k(p)$ if $d \not\equiv 0, 1 \pmod{k}$. \square

In the following theorem, we give an upper bound of τ .

Theorem 5.3. *Let $f \in \mathbb{F}_p[x]$ be a polynomial of degree d . Assume that $d|\Phi_k(p)$. Let τ be the number of absolutely irreducible factors in the factorization of $f(x) -$*

$f(y)$. Then we have $\tau \leq \sum_{D|d} \frac{\varphi(D)}{\text{ord}_D(p)}$. In particular, when k is the prime, we have either

$$\tau \leq \frac{d-1}{k} + 1 \quad \text{for } d \equiv 1 \pmod{k} \quad \text{or} \quad \tau \leq \frac{d-k}{k} + k \quad \text{for } d \equiv 0 \pmod{k}.$$

Proof. Consider

$$f(x) - f(y) = (x^d - y^d) + a_{d-2}(x^{d-2} - y^{d-2}) + \cdots + a_2(x^2 - y^2) + a_1 = g_1 \cdots g_s.$$

Comparing the highest homogeneous term gives

$$x^d - y^d = h_{1,d_1} \cdots h_{s,d_s}, \text{ where } h_{i,d_i} \in \mathbb{F}_p[x, y].$$

Since $x^d - y^d = \prod_{D|d} \Phi_D(x, y)$ and $\Phi_D(x, y)$ factorizes into $\frac{\varphi(D)}{\text{ord}_D(p)}$ distinct irreducibles of degree $\text{ord}_D(p)$, we have at most $\sum_{D|d} \frac{\varphi(D)}{\text{ord}_D(p)}$ absolutely irreducible factors. Here, $\text{ord}_D(p)$ denotes the multiplicative order of p modulo D .

Let ζ be a primitive d -th root of unity in \mathbb{F}_{p^k} . For the prime k , $x^d - y^d$ has irreducible factors (over \mathbb{F}_p) of either linear factor, $(x - \zeta^i y)$ for $\zeta^i \in \mathbb{F}_p$, or degree- k factor,

$$(x - \zeta^i y) (x - \zeta^{i+p} y) \cdots (x - \zeta^{i+p^{k-1}} y),$$

for $\zeta^i \in \mathbb{F}_{p^k} \setminus \mathbb{F}_p$. Thus, by Lemma 5.2, the number of irreducible factors of $x^d - y^d$ is either $\frac{d-1}{k} + 1$ for $d \equiv 1 \pmod{k}$ or $\frac{d-k}{k} + k$ for $d \equiv 0 \pmod{k}$. Since the factor g_i is determined by its highest degree term h_{i,d_i} , the number of absolutely irreducible factors is less than the number of irreducible factors of $x^d - y^d$. \square

5.3. Several Examples. In this section, we present several polynomials of degree d that achieve the upper bound of τ_f in the case of $d|\Phi_1(p)$ and $d|\Phi_2(p)$. In the case of $d|\Phi_3(p)$, no polynomial can attain the upper bound.

5.3.1. Case 1: $d|\Phi_1(p) = (p-1)$. In this case, the possible number of the irreducible factors is at most d . Consider $f(x) = x^d$ with $d|(p-1)$. Then a primitive d -th root of unity ζ exists in \mathbb{F}_p and $f(x) - f(y)$ has d absolutely irreducible linear factors over \mathbb{F}_p , since the factorization is given by

$$f(x) - f(y) = \prod_{i=1}^d (x - \zeta^i y).$$

For a fixed non-zero y , $f(x) = f(y)$ if and only if $x = \zeta^i y$ for each $i = 1, \dots, d$, so the map $x \mapsto f(x)$ is a d -to-1 function except on $x = 0$. Finally, $\rho_f = R_1 + d^2 \cdot R_d = 1 + d^2 \cdot \frac{p-1}{d} = 1 + d(p-1)$.

Remark 5.4. Applying Theorem 2.3 with the polynomial $f(x) = x^d$ such that $d|(p-1)$, one solves the discrete log α in $\tilde{O}(\sqrt{p/d} + d)$ group operations which can be lowered by $\tilde{O}(p^{1/3})$ when $d = p^{1/3}$. Note that the polynomial of form $f(x) = a(x+b)^d + c$ suggests the same asymptotic complexity since the value set does not change by translations.

5.3.2. **Case 2:** $d|\Phi_2(p) = (p+1)$. In the case, the possible number of the absolutely irreducible factors is at most $\lfloor \frac{d+2}{2} \rfloor$. Consider the Dickson polynomial of degree d . For a nonzero $a \in \mathbb{F}_p$, the Dickson polynomial is defined by

$$D_d(x, a) = \sum_{k=0}^{\lfloor d/2 \rfloor} \frac{d}{d-k} \binom{d-k}{k} (-a)^k x^{d-2k}.$$

The following lemma shows that the substitution polynomial of the Dickson polynomial has exactly $\lfloor \frac{d+2}{2} \rfloor$ absolutely irreducible factors and presents the exact number of ρ_f .

Lemma 5.5 ([10, 15]). *Assume that $d|(p+1)$ and ζ be a primitive d -th root of unity in \mathbb{F}_{p^2} . Then*

$$D_d(x, a) - D_d(y, a) = (x^t - y^t) \prod_{i=1}^{\lfloor (d-1)/2 \rfloor} (x^2 - (\zeta^i + \zeta^{-i})xy + y^2 + a(\zeta^{2i} + \zeta^{-2i} - 2)),$$

where $t = 1$ for odd d and $t = 2$ for even d . For the number ρ_f , one has

$$\rho_f = \frac{(d+1)p}{2} + O(d^2).$$

Remark 5.6. Applying Theorem 2.3 with the Dickson polynomial $D_d(x, a)$ with $d|(p+1)$, the discrete log α can be recovered within $\tilde{O}(\sqrt{\frac{p}{2d}} + d)$ group operations for $d < p^{1/2}$. It can be lowered to $\tilde{O}(p^{1/3})$ when $d = p^{1/3}$.

5.3.3. **Case 3:** $d|\Phi_3(p) = (p^2 + p + 1)$. In this case, the possible r is at most $\frac{d-1}{3} + 1$ for $d \equiv 1 \pmod{3}$ and $\frac{d-3}{3} + 3$ for $d \equiv 0 \pmod{3}$. This kind of polynomial appears only when the factorization of $f(x) - f(y)$ is given by

$$f(x) - f(y) = (x^t - y^t) \prod_{i=1}^{s-1} g_i(x, y),$$

where each of g_i is an absolutely irreducible cubic factor ($t = 1$ or $t = 3$ depending on the residue class of d modulo 3). In the next section, however, we show that such a polynomial does not exist.

Theorem 5.7. *Let ζ be a primitive d -th root of unity in \mathbb{F}_{p^3} . Assume that $f(x) \in \mathbb{F}_p[x]$ is a polynomial of degree d . Then $f(x) - f(y)$ cannot have an absolutely irreducible cubic factor.*

5.4. **Proof of Theorem 5.7.** Let $f(x) = x^d + a_{d-1}x^{d-1} + \dots + a_1x + a_0 \in \mathbb{F}_p[x]$ be a monic polynomial of degree d . Since the value set of a polynomial does not vary by translations, we always assume that $a_{d-1} = 0$ by taking $f(x - \frac{a_{d-1}}{d})$ instead of $f(x)$.

To prove Theorem 5.7, we need the following lemmas. The next lemma determines the highest term of an irreducible cubic factor of $\frac{f(x)-f(y)}{x-y}$.

Lemma 5.8. *Let d be a positive integer dividing $\Phi_3(p)$ and $F(x, y) = \frac{f(x)-f(y)}{x-y}$ for a polynomial $f(x)$ of degree d . If $g_i(x, y)$ is an irreducible cubic factor of $F(x, y)$, then*

$$h_{i,3}(x, y) = (x - \xi y)(x - \xi^p y)(x - \xi^{p^2} y),$$

for some d -th root of unity ξ not in \mathbb{F}_p .

Proof. Let ζ be a primitive d -th root of unity in \mathbb{F}_{p^3} . Consider

$$F(x, y) = \frac{x^d - y^d}{x - y} + a_{d-2} \frac{x^{d-2} - y^{d-2}}{x - y} + \cdots + a_2 \frac{x^2 - y^2}{x - y} + a_1.$$

Comparing the highest homogeneous term gives

$$\frac{x^d - y^d}{x - y} = h_{1,d_1} \cdots h_{s-1,d_{s-1}} = (x - \zeta y)(x - \zeta^2 y) \cdots (x - \zeta^{d-1} y).$$

Assume that g_1 is cubic without loss of generality. Since $h_{1,3} \in \mathbb{F}_p[x, y]$ has a factor $x - \zeta^i y$ for some $\zeta^i \in \mathbb{F}_{p^3} \setminus \mathbb{F}_p$ (note that $h_{1,3}$ cannot have a factor $x - \zeta^i y$ for $\zeta^i \in \mathbb{F}_p$, since there exist only two linear factors over \mathbb{F}_p by Lemma 5.2), it also contains the homogenization of the minimal polynomial of $\xi := \zeta^i$, i.e. $h_{1,3} = (x - \xi y)(x - \xi^p y)(x - \xi^{p^2} y)$ for some d -th root of unity ξ which does not lie in \mathbb{F}_p . \square

The next lemma recovers a whole form of the absolutely irreducible cubic factors of $\frac{f(x)-f(y)}{x-y}$, if it exists.

Lemma 5.9. *Let $f(x) = x^d + a_{d-1}x^{d-1} + a_{d-2}x^{d-2} + \cdots + a_1x + a_0$ be a polynomial over \mathbb{F}_p with $d|\Phi_3(p) = (p^2 + p + 1)$. Let $F(x, y) = \frac{f(x)-f(y)}{x-y} = g_1(x, y) \cdots g_{s-1}(x, y)$ be the factorization over \mathbb{F}_p . Assume that $a_{d-1} = 0$. If $F(x, y)$ has an absolutely irreducible cubic factor, then $a_{d-2} \neq 0$ and it must be of form*

$$(5.1) \quad (x - \zeta y)(x - \zeta^p y)(x - \zeta^{p^2} y) + \frac{(\text{tr}(\zeta) - 3)(\text{tr}(\zeta) + 1)}{d} \cdot a_{d-2} \cdot (x - y),$$

where $\zeta \in \mathbb{F}_{p^3} \setminus \mathbb{F}_p$ is a (not necessarily primitive) d -th root of unity and $\text{tr}(\zeta)$ denotes the trace of ζ .

Proof. Comparing the first two terms in $F(x, y) = g_1 \cdots g_{s-1}$, one has the identities

$$\frac{x^d - y^d}{x - y} = h_{1,d_1} \cdots h_{s-1,d_{s-1}} \quad \text{and} \quad a_{d-1} \cdot \frac{x^{d-1} - y^{d-1}}{x - y} = \sum_{i=1}^{s-1} \left(\prod_{j \neq i} h_{j,d_j} \right) \cdot h_{i,d_i-1}.$$

Dividing the second equality by the first equality yields $a_{d-1} \cdot \frac{x^{d-1} - y^{d-1}}{x^d - y^d} = \sum_{i=1}^{s-1} \frac{h_{i,d_i-1}}{h_{i,d_i}}$. Since $a_{d-1} = 0$, we have $h_{i,d_i-1} = 0$ for all i by the uniqueness of the partial fractional decomposition. Similarly, since $h_{i,d_i-1} = 0$, we have

$$(5.2) \quad a_{d-2} \cdot \frac{x^{d-2} - y^{d-2}}{x^d - y^d} = \sum_{i=1}^{s-1} \frac{h_{i,d_i-2}}{h_{i,d_i}} \quad \text{and} \quad a_{d-3} \cdot \frac{x^{d-3} - y^{d-3}}{x^d - y^d} = \sum_{i=1}^{s-1} \frac{h_{i,d_i-3}}{h_{i,d_i}}.$$

Now consider the partial fractional decomposition of $\frac{x^{d-k} - y^{d-k}}{x^d - y^d}$, for $k = 1, 2, \dots, d$, regarded as a rational polynomial with the coefficients in $\overline{\mathbb{F}_p}[y]$. Then we have the followings in the usual way,

$$(5.3) \quad \frac{x^{d-k} - y^{d-k}}{x^d - y^d} = \sum_{i=1}^d \frac{A_{k,i}}{x - \zeta^i y} \quad \text{for} \quad A_{k,i} = \frac{(\zeta^i y)^{d-k} - y^{d-k}}{d(\zeta^i y)^{d-1}}.$$

From the equations (5.2) and (5.3) with $k = 3$, we have

$$\sum_{i=1}^d \frac{A_{3,i}}{x - \zeta^i y} = \frac{1}{a_{d-3}} \cdot \sum_{i=1}^{s-1} \frac{h_{i,d_i-3}}{h_{i,d_i}} \quad \text{for} \quad A_{3,i} = \frac{1}{dy^2} \cdot (\zeta^{-2i} - \zeta^i).$$

Let $g_1 = (x - \zeta y)(x - \zeta^p y)(x - \zeta^{p^2} y) + (\alpha x + \beta y) + \gamma$ for α, β and $\gamma \in \mathbb{F}_p$. If $a_{d-3} \neq 0$, then

$$\frac{\gamma}{(x - \zeta y)(x - \zeta^p y)(x - \zeta^{p^2} y)} = \frac{a_{d-3}}{dy^2} \cdot \left(\frac{\zeta^{-2} - \zeta}{x - \zeta y} + \frac{\zeta^{-2p} - \zeta^p}{x - \zeta^p y} + \frac{\zeta^{-2p^2} - \zeta^{p^2}}{x - \zeta^{p^2} y} \right).$$

Comparing the numerators in both sides, one has the followings:

$$(5.4) \quad (\zeta^{-2} - \zeta) + (\zeta^{-2p} - \zeta^p) + (\zeta^{-2p^2} - \zeta^{p^2}) = 0,$$

$$(5.5) \quad (\zeta^{-2} - \zeta)(\zeta^p + \zeta^{p^2}) + (\zeta^{-2p} - \zeta^p)(\zeta^{p^2} + \zeta) + (\zeta^{-2p^2} - \zeta^{p^2})(\zeta + \zeta^p) = 0,$$

and

$$(5.6) \quad (\zeta^{-2} - \zeta) \cdot \zeta^{p+p^2} + (\zeta^{-2p} - \zeta^p) \cdot \zeta^{1+p^2} + (\zeta^{-2p^2} - \zeta^{p^2}) \cdot \zeta^{1+p} = \frac{d}{a_{d-3}} \gamma.$$

Since d is a divisor of $p^2 + p + 1$, we have $\zeta^{1+p+p^2} = 1$. Let $(x - \zeta y)(x - \zeta^p y)(x - \zeta^{p^2} y) = x^3 - ax^2y + bxy^2 - y^3$, i.e. $a = \zeta + \zeta^p + \zeta^{p^2}$ and $b = \zeta\zeta^p + \zeta^p\zeta^{p^2} + \zeta^{p^2}\zeta$. From Equation (5.4), we have $a = b^2 - 2a$, and Equation (5.5) yields $a(b^2 - 2a) - b = 2b$. These two equations yield

$$(a, b) = (0, 0) \text{ or } (a, b) = (3, 3) \text{ or } (a, b) = (3\zeta_3, 3\zeta_3^2)$$

for a primitive third root of unity ζ_3 . Then from Equation (5.6), we have $b(b^2 - 3a) = \frac{d}{a_{d-3}} \gamma$ yielding $\gamma = 0$. Consequently, if $a_{d-2} = 0$, then $h_{i, d_i-2} = 0$, so only possible g_1 is of form

$$x^3 - y^3 \text{ or } (x - y)^3 \text{ or } (x - \zeta_3 y)^3.$$

If $a_{d-2} \neq 0$, then we have the following,

$$\sum_{i=1}^d \frac{A_{2,i}}{x - \zeta^i y} = \frac{1}{a_{d-2}} \cdot \sum_{i=1}^{s-1} \frac{h_{i, d_i-2}}{h_{i, d_i}} \quad \text{for } A_{2,i} = \frac{1}{dy} (\zeta^{-i} - \zeta^i).$$

The similar argument gives the following equations

$$(5.7) \quad (\zeta^{-1} - \zeta) + (\zeta^{-p} - \zeta^p) + (\zeta^{-p^2} - \zeta^{p^2}) = 0,$$

$$(5.8) \quad (\zeta^{-1} - \zeta)(\zeta^p + \zeta^{p^2}) + (\zeta^{-p} - \zeta^p)(\zeta + \zeta^{p^2}) + (\zeta^{-p^2} - \zeta^{p^2})(\zeta + \zeta^p) = -\frac{d}{a_{d-2}} \cdot \alpha$$

and

$$(5.9) \quad (\zeta^{-1} - \zeta) \cdot \zeta^{p+p^2} + (\zeta^{-p} - \zeta^p) \cdot \zeta^{1+p^2} + (\zeta^{-p^2} - \zeta^{p^2}) \cdot \zeta^{1+p} = \frac{d}{a_{d-2}} \cdot \beta.$$

The equation (5.7) follows $a = b$. From the third equation, we have $b^2 - 2a - 3 = \frac{d}{a_{d-2}} \beta$, thus together with $a = b$, we have

$$\beta = \frac{(a-3)(a+1)}{d} \cdot a_{d-2}.$$

The equation (5.8) deduces that $ab - 3 - 2b = a^2 - 2a - 3 = -\frac{d}{a_{d-2}} \alpha$ which follows the results

$$\alpha = -\beta = -\frac{(a-3)(a+1)}{d} \cdot a_{d-2}.$$

If $a = b = 3$, then $g_1 = (x-y)^3$, and if $a = b = 0$, then $g_1 = (x^3 - y^3) + \frac{3}{d} \cdot a_{d-2} \cdot (x-y)$. In any case, g_1 is not absolutely irreducible. In the case of $a_{d-3} = 0$ and $a_{d-2} \neq 0$, we have

$$g_1 = (x - \zeta y)(x - \zeta^p y)(x - \zeta^{p^2} y) - \frac{(a-3)(a+1)}{d} \cdot a_{d-2}(x-y).$$

□

Now we can prove Theorem 5.7 using Lemma 5.9.

proof of Theorem 5.7. Suppose that g_1 is an absolutely irreducible cubic factor. Then by Lemma 5.9, $g_1 = (x - \zeta y)(x - \zeta^p y)(x - \zeta^{p^2} y) + \frac{(\text{tr}(\zeta)-3)(\text{tr}(\zeta)+1)}{d} a_{d-2}(x-y)$ for some (not necessarily primitive) d -th root of unity $\zeta \in \mathbb{F}_{p^3} \setminus \mathbb{F}_p$. Since g_1 has no constant term, the least degree of $F(x, y)$ is at least one. Let $\ell + 1$ be the least degree of $F(x, y)$ for $\ell \geq 0$. Let h_{i, ℓ_i} be the least homogeneous term in g_i . Consider the least degree term in $F(x, y) = g_1 \cdots g_{s-1}$,

$$a_{\ell+2} \frac{x^{\ell+2} - y^{\ell+2}}{x-y} = \frac{(\text{tr}(\zeta)-3)(\text{tr}(\zeta)+1)}{d} \cdot a_{d-2} \cdot (x-y) \cdot h_{2, \ell_2} \cdots h_{s-1, \ell_{s-1}}.$$

Regarding without the constant multiple, we must have

$$\frac{x^{\ell+2} - y^{\ell+2}}{x-y} = (x-y)(x^\ell + a_1 \cdot x^{\ell-1}y + a_2 \cdot x^{\ell-2}y^2 + \cdots + a_\ell \cdot y^\ell),$$

which is impossible for $\ell \geq 0$. □

Corollary 5.10. *If a polynomial $f(x)$ is of degree $d \mid \Phi_3(p)$, then one has $\tau \leq \frac{d-t}{6} + t$ where $t = 1$ if $d \equiv 1 \pmod{3}$ and $t = 3$ if $d \equiv 0 \pmod{3}$.*

Proof. In the proof of Lemma 5.8, one observes that h_{i, d_i} must be a multiple of $(x - \xi y)(x - \xi^p y)(x - \xi^{p^2} y)$, so the most of irreducible factors of $\frac{f(x)-f(y)}{x-y}$ are at least sextic, since it cannot have a cubic factors by Theorem 5.7. Then the result follows in the analogous way. □

Remark 5.11. After writing up the paper, the authors has been informed that the same result can be directly obtained from the result of [12]. In that paper, the author describes all the possible forms of the absolutely irreducible cubic factors for arbitrary d . By Lemma 6 in [12], the coefficient of y^3 in the cubic factors must be equal to 1 which contradicts to the fact the coefficient is -1 in $(x - \xi y)(x - \xi^p y)(x - \xi^{p^2} y)$.

Remark 5.12. One might attempt to construct a polynomial whose substitution polynomial contains many absolutely irreducible cubic factors using the result of [12] for arbitrary d . However, it appears that it is not so attractable for several reasons. Using again Lemma 6 in [12], $x^d - y^d$ must have factors of the form $\prod_{i=1}^r (x^3 - c_i y^3)$ for $c_i \neq 1 \in \mathbb{F}_p$. If c_i is a cubic in \mathbb{F}_p , then $x^3 - c_i y^3$ factors into a linear and a quadratic. The linear factor then must be of the form $x - \xi y$ for $\xi \in \mathbb{F}_p$, a power of the d -th primitive root ζ . The number of such ξ is at most $\gcd(d, p-1)$ and so $\gcd(d, p-1)$ must be large to have many such elements. This case is not so fruitful because it is already well covered by Cheon's $p-1$ algorithm. Otherwise, $x^3 - c_i y^3$ is irreducible for non-cubic c_i . It must be also a homogenization of the minimal polynomial of some $\xi \in \mathbb{F}_{p^3}$ and the trace of ξ must be zero. It seems unlikely to happen, though.

6. CONCLUSION

We have proposed a new algorithm to solve the DLPwAI. The algorithm has a running time of $\tilde{O}\left(\frac{p}{\sqrt{\rho_f}} + d\right) = \tilde{O}\left(\sqrt{\frac{p}{\tau_f}} + d\right)$ group exponentiations for a chosen polynomial $f \in \mathbb{F}_p[x]$ of degree d , therefore it reduces the DLPwAI to find polynomials with the large value of ρ_f or τ_f .

It remains open to find a polynomial with large τ_f so that the proposed algorithm has the complexity $O\left(\sqrt{p/d}\right)$ as the lower bound in the generic group model. For example, we have such polynomials in the case of $d|(p \pm 1)$.

For the birthday problem, it would be interesting to determine the expected number of trials until a collision for arbitrary probability distribution.

ACKNOWLEDGEMENT

The authors would like to thanks to Steven Galbraith, Hansol Ryu, Jae Hong Seo, Yong Soo Song, Mehdi Tibouchi, Michael Zieve for their useful discussions.

REFERENCES

- [1] Dan Boneh and Xavier Boyen, *Efficient selective-ID secure identity-based encryption without random oracles*, Advances in Cryptology - EUROCRYPT 2004 (Christian Cachin and Jan Camenisch, eds.), Lecture Notes in Computer Science, vol. 3027, Springer, 2004, pp. 223–238.
- [2] ———, *Short signatures without random oracles*, Advances in Cryptology - EUROCRYPT 2004 (Christian Cachin and Jan Camenisch, eds.), Lecture Notes in Computer Science, vol. 3027, Springer, 2004, pp. 56–73.
- [3] Dan Boneh, Craig Gentry, and Brent Waters, *Collusion resistant broadcast encryption with short ciphertexts and private keys*, Advances in Cryptology - CRYPTO 2005 (Victor Shoup, ed.), Lecture Notes in Computer Science, vol. 3621, Springer, 2005, pp. 258–275.
- [4] Daniel R. L. Brown and Robert P. Gallant, *The static Diffie-Hellman problem*, IACR Cryptology ePrint Archive (2004), <http://eprint.iacr.org/2004/306>.
- [5] Jung Hee Cheon, *Security analysis of the strong Diffie-Hellman problem*, Advances in Cryptology - EUROCRYPT 2006 (Serge Vaudenay, ed.), Lecture Notes in Computer Science, vol. 4004, Springer, 2006, pp. 1–11.
- [6] ———, *Discrete logarithm problems with auxiliary inputs*, J. Cryptology **23** (2010), no. 3, 457–476.
- [7] Jung Hee Cheon and Taechan Kim, *Discrete logarithm with auxiliary inputs*, MSJ-KMS Joint Meeting 2012, 2012.
- [8] Jung Hee Cheon, Taechan Kim, and Yong Soo Song, *A Group action on F_p^\times and the generalized dlp with auxiliary inputs*, Selected Areas in Cryptography 2013.
- [9] Steven D. Galbraith and Mark Holmes, *A non-uniform birthday problem with applications to discrete logarithms*, Discrete Applied Mathematics **160** (2012), no. 10-11, 1547–1560.
- [10] J. Gomez-Calderon and D. J. Madden, *Polynomials with small value set over finite fields*, Journal of Number Theory **28** (1988), 167–188.
- [11] Javier Gomez-Calderon, *On the cardinality of value set of polynomials with coefficients in a finite field*, **68** (1992), no. 10, 338–340.
- [12] ———, *The third-order factorable core of polynomials over finite fields*, **74** (1998), no. 1, 16–19.
- [13] D. R. Hayes, *A geometric approach to permutation polynomials over a finite field*, **34** (1967), no. 2, 293–305.
- [14] Minkyu Kim, Jung Hee Cheon, and In-Sok Lee, *Analysis on a generalized algorithm for the strong discrete logarithm problem with auxiliary inputs*, Mathematics of Computation **to appear**.
- [15] D. A. Mit'kin, *Polynomials with minimal set of values and the equation $f(x) = f(y)$ in a finite prime field*, Matematicheskie Zametki **38** (1985), no. 1, 3–14.

- [16] Shigeo Mitsunari, Ryuichi Sakai, and Masao Kasahara, *A new traitor tracing*, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences **E85-A** (2002), no. 2, 481–484.
- [17] Payman Mohassel, *Fast computation on encrypted polynomials and applications*, CANS (Dongdai Lin, Gene Tsudik, and Xiaoyun Wang, eds.), vol. 7092, Lecture Notes in Computer Science, 2011, pp. 234–254.
- [18] Kazuo Nishimura and Masaaki Sibuya, *Occupancy with two types of balls*, Annals of the Institute of Statistical Mathematics **40** (1988), no. 1, 77–91.
- [19] J. M. Pollard, *Monte carlo methods for index computation (mod p)*, Mathematics of Computation **32** (1978), no. 143, pp. 918–924.
- [20] Takakazu Satoh, *On generalization of Cheon's algorithm*, IACR Cryptology ePrint Archive (2009), <http://eprint.iacr.org/2009/058>.
- [21] B. I. Selivanov, *On waiting time in the scheme of random allocation of coloured particles*, Discrete. Math. Appl. **5** (1955), no. 1, 73–82.
- [22] Saburo Uchiyama, *Note on the mean value of $v(f)$* , Proc. Japan Acad **31** (1955), 199–201.
- [23] J. von zur Gathen and J. Gerhard, *Modern computer algebra*, Cambridge University Press, 2003.
- [24] Andre Weil, *Sur les courbes algebriques et les variétés qui s'en deduissent*, Publ. Inst. Math. Strasbourg, vol. 1041, 1948.

APPENDIX A. THE OMITTED PROOFS

A.1. Another proof of Lemma 5.2 and Theorem 5.3. We give another proof of Lemma 5.2 and Theorem 5.3. We begin with several notations. Assume that a primitive d -th root of unity ζ lies in \mathbb{F}_{p^k} , where k is the smallest integer satisfying the condition. For $\tilde{k}|k$, we define

- $D(\tilde{k})$ be the number of $i \in \{1, 2, \dots, d\}$ satisfying ζ^i lies in $\mathbb{F}_{p^{\tilde{k}}}$.
- $N(\tilde{k})$ be the number of $i \in \{1, 2, \dots, d\}$ satisfying ζ^i exactly lies in $\mathbb{F}_{p^{\tilde{k}}}$ not in any proper subfield.

proof of Lemma 5.2 and Theorem 5.3. It suffices to find the number of irreducible factors of $x^d - 1$ over \mathbb{F}_p . If ζ^i is in $\mathbb{F}_{p^{\tilde{k}}}$ and not in any proper subfield, then the minimal polynomial of ζ^i is of degree \tilde{k} . Thus $x^d - 1$ factors into $\sum_{\tilde{k}|k} \frac{N(\tilde{k})}{\tilde{k}}$ irreducibles.

Now we can easily check that $D(\tilde{k}) = \gcd(d, p^{\tilde{k}} - 1)$, since $\zeta^i \in \mathbb{F}_{p^{\tilde{k}}}$ if and only if $\zeta^{i(p^{\tilde{k}}-1)} = 1$ if and only if $i(p^{\tilde{k}} - 1) \equiv 0 \pmod{d}$. From the definitions, $D(\tilde{k}) = \sum_{\ell|\tilde{k}} N(\ell)$, so the Möbius inversion formula suggests that

$$N(\tilde{k}) = \sum_{\ell|\tilde{k}} \mu\left(\frac{\tilde{k}}{\ell}\right) \cdot D(\tilde{k}) = \sum_{\ell|\tilde{k}} \mu\left(\frac{\tilde{k}}{\ell}\right) \cdot \gcd(d, p^{\tilde{k}} - 1),$$

where $\mu(n)$ is the Möbius function.

For the prime k , we have

$$\begin{aligned} \sum_{\tilde{k}|k} \frac{N(\tilde{k})}{\tilde{k}} &= N(1) + \frac{N(k)}{k} = \gcd(d, p - 1) + \frac{\gcd(d, p^k - 1) - \gcd(d, p - 1)}{k} \\ &= \gcd(d, p - 1) + \frac{d - \gcd(d, p - 1)}{k}. \end{aligned}$$

Since $N(k) = d - \gcd(d, p - 1)$ must be a multiple of k and $\gcd(d, p - 1)$ only can be either 1 of k , d modulo k only can be either 1 or 0. \square

DEPARTMENT OF MATHEMATICAL SCIENCES, SEOUL NATIONAL UNIVERSITY
E-mail address: `jhcheon@snu.ac.kr`

DEPARTMENT OF MATHEMATICAL SCIENCES, SEOUL NATIONAL UNIVERSITY
E-mail address: `yoshiki1@snu.ac.kr`