

# On the discrete logarithm problem for prime-field elliptic curves

**Citation for published version (APA):**

Amadori, A. G., Pintore, F., & Sala, M. (2018). On the discrete logarithm problem for prime-field elliptic curves. *Finite Fields and their Applications*, 51, 168-182. [2017/609]. <https://doi.org/10.1016/j.ffa.2018.01.009>

**DOI:**

[10.1016/j.ffa.2018.01.009](https://doi.org/10.1016/j.ffa.2018.01.009)

**Document status and date:**

Published: 01/05/2018

**Document Version:**

Accepted manuscript including changes made at the peer-review stage

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# On the discrete logarithm problem for prime-field elliptic curves

Alessandro Amadori<sup>1</sup>, Federico Pintore<sup>2</sup> \*\*, Massimiliano Sala<sup>2</sup>

<sup>1</sup> Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, The Netherlands

<sup>2</sup> Department of Mathematics, University of Trento, Italy

21 June 2017

## Abstract

In recent years several papers have appeared investigating the classical discrete logarithm problem for elliptic curves by means of the multivariate polynomial approach based on the celebrated summation polynomials, introduced by Semaev in 2004. However, with a notable exception by Petit *et al.* in 2016, all numerous papers have investigated only the composite-field case, leaving apart the laborious prime-field case. In this paper we propose a variation of Semaev's original approach for the prime-field case. Our proposal outperforms both the original Semaev's method and Petit *et al.* specialized algorithm. The improvement is reached by reducing the necessary Groebner basis computations to only one basis calculation.

**Keywords:** Elliptic curve, Discrete logarithm problem (DLP), Prime field, Summation polynomials, Groebner basis.

## 1 Introduction

Several cryptographic schemes base their security upon the hardness of the discrete logarithm problem for elliptic curves (ECDLP) [12,14]. For an elliptic curve  $E$  defined over a finite field  $\mathbb{F}_q$ , an instance of the ECDLP is the following:

given  $P, Q \in E(\mathbb{F}_q)$ , compute an integer  $w$ , if it exists, s.t.  $Q = wP$ .

The best known algorithms for the ECDLP are algorithms that work on arbitrary cyclic groups - like Pollard's Rho algorithm [18], which runs in time  $\mathcal{O}(\sqrt{|E(\mathbb{F}_q)|})$  if  $|E(\mathbb{F}_q)|$  is prime - exception made for algorithms that are specific for some families of weak curves (as, for example, [13]).

In 2004 I. Semaev introduced [19] a family of polynomials, named *summation polynomials*, proposing their exploitation for an index calculus algorithm for elliptic curves. The Index Calculus is originally a subexponential algorithm to compute discrete logarithms in the multiplicative groups of finite fields. However,

---

\*\* This work was supported by CARITRO Foundation

it is customary to use the name *index calculus algorithm* to refer to any algorithm that computes discrete logarithms in a cyclic group  $G$  by collecting linear relations first and using linear algebra afterwards. Following [6] and restricting to the case  $G = E(\mathbb{F}_q)$ , with  $r = |E(\mathbb{F}_q)|$  a prime integer, the simplest version of index calculus algorithm consists of two main steps: the *relation collection* step and the *linear algebra* step. In the *relation collection* step:

1. a factor base  $\mathcal{F} \subset E(\mathbb{F}_q)$  is defined;
2. for random integers  $u, v$ , the point  $R = uP + vQ$  is computed;
3. if possible,  $R$  is written as a sum of multiples of points of  $\mathcal{F}$ :

$$R = uP + vQ = \sum_{F \in \mathcal{F}} \ell_F F, \quad (1)$$

with the integers  $\ell_F$ 's ranging in a small coefficient set;

4.  $u, v$  and the vector  $(\ell_F)$  are stored as a row of a matrix  $M$ ;
5. the procedure from item 2 to item 4 is repeated until at least  $|\mathcal{F}|$  points  $R$  as in (1) are found.

After the collection of a large enough number of relations, the *linear algebra* step solves the discrete logarithm problem:

1. using linear algebra on  $M$ , a linear dependency of points  $R$  is computed, obtaining the relation  $\lambda P + \mu Q = \infty$  with  $\lambda, \mu \in \mathbb{Z}$ ;
2.  $w$  is recovered from the linear congruence  $\lambda + \mu w = 0 \pmod{r}$ , which is solvable except the extremely unlikely case when  $\mu = 0$ .

The complexity of an index calculus algorithm mainly lies in the decomposition of a point  $R$  as sum of multiples of points in  $\mathcal{F}$ , usually known as the *point decomposition problem* (PDP). The solution of the PDP must be efficient (including the decision on the decomposability of a point  $R$ ) and with a high success rate. Both these features are deeply affected by how  $\mathcal{F}$  is defined and by its size.

In the third section of Semaev's paper [19], he proposed to reduce the PDP to the problem of finding specific solutions of a multivariate polynomial equation deduced from summation polynomials. Semaev sketched his proposal in the case of elliptic curves defined over prime fields, suggesting to define  $\mathcal{F}$  as the set of rational points whose  $x$ -coordinates are *small* when taken as integers in the standard complete residue system  $[0, \dots, p-1]$ . Furthermore, in Remark 2 of [19], Semaev also outlined a variation of his proposal to binary elliptic curves defined over extension fields. This variation involves the Weil descent, which cannot be applied to prime fields.

Semaev's paper triggered a deep interest in summation polynomials. Starting from them, Diem claimed that the ECDLP for some elliptic curves could be solved with a subexponential-time complexity (rather than exponential) in

his 2006 talk [2]. However, the first papers developing Semaev's approach appeared only in 2009, by Gaudry [8], and then in 2011 by Diem [1]. In [8] Gaudry shows that Semaev's proposal can be used to solve, in heuristic time  $\tilde{\mathcal{O}}(q^{2-\frac{2}{n}})$ , the ECDLP for elliptic curves defined over fields of size  $q = q^n$  (from now on,  $\text{ECDLP}(q,n)$ ), with  $q$  prime or a prime power and  $n \geq 2$ . However, Gaudry's results cannot be extended to the prime-field case. Diem [1] proved that the same problem can be solved in an expected time of  $e^{\mathcal{O}(\max(\log q, n^2))}$ . Unfortunately, when  $q = 2$  and  $n$  is large (one of the two common cases in leading cryptographic standards), the above works do not lead to an algorithm more effective than Pollard's Rho algorithm. In 2012 Faugère, Perret, Petit and Renault [5] claimed that, under a heuristic assumption, the ECDLP over any binary field  $\mathbb{F}_{2^n}$  can be solved in time  $\mathcal{O}(2^{\omega t})$ , with  $t \approx n/2$  and where  $2.376 \leq \omega \leq 3$  is the *linear algebra constant*. Petit and Quisquater [16] revisited polynomial systems proposed by Faugère *et al.* in [5], introducing a heuristic assumption, named *first fall degree assumption*, under which the  $\text{ECDLP}(2,n)$  can be solved in time  $\mathcal{O}(2^{cn^{2/3} \log n})$ , where  $c$  is a constant smaller than 2.

Further relevant results were then accomplished exploiting symmetries in the works of Huang, Petit, Shinohara and Takagi [9], of Faugère, Gaudry, Huot and Renault [4], and of Galbraith and Gebregiyorgis [7]. In the latter paper, the authors propose to use disjoint factor basis  $\mathcal{F}_i$  and require  $P_i \in \mathcal{F}_i$  in (1).

In 2015, Semaev [20] and Karabina [11] had independently the idea to lower the degrees of polynomial equation systems, appearing in the PDP, at the cost of a larger number of variables. Both [20] and [11] provided experimental results suggesting that their idea can lead to smaller running times.

There is no doubt that the listed results (and many others not cited here) determined a huge improvement over the index calculus algorithms for elliptic curves. However, the question whether the index calculus algorithm proposed by Semaev in 2004 can lead to a subexponential algorithm for binary elliptic curves is nowadays still open.

Even if the original proposal in Semaev's paper directly considered the case of elliptic curves defined over prime fields, only one recent paper ([17], 2016) takes such demanding case into account. Among the main obstacles that one may encounter when facing this case, it is worth mentioning the difficulty to endow  $\mathcal{F}$  with an algebraic structure, and the impossibility to exploit the Weil descent in the PDP. Nevertheless, elliptic curves defined over prime fields are nowadays widely spread in cryptographic applications, as for example the Bitcoin system [15]. In [17] Petit *et al.* deal with the identification of a suitable  $\mathcal{F}$  by introducing rational (or algebraic) maps  $L$ , obtained by a preprocessing specific for a given curve  $E$ , and defining  $\mathcal{F}$  as the set of rational points  $\{(x, y) \in E(\mathbb{F}_p) | L(x) = 0\}$ .

In this paper we propose a new variant of index calculus algorithm for the prime-field case. At the beginning of our algorithm (item 1 in the *relation collection* step), we define the size of  $\mathcal{F}$  but not  $\mathcal{F}$  itself, which instead will be defined

on the run. Indeed, every point  $R$  produced in item 2 in the *relation collection* step is appended to  $\mathcal{F}$ , giving a trivial decomposition of  $R$  in item 3. We proceed in this fashion until  $\mathcal{F}$  contains the predetermined size. At this point a linear dependency among points in  $\mathcal{F}$  is obtained solving a point decomposition problem by means of summation polynomials. Consequently, our proposal reduces the index calculus algorithm for elliptic curves over prime fields to the solution of a single multivariate polynomial equation system.

We compare the complexity of our algorithm with that by Petit *et al.* showing that we improve the complexity of the algorithm in [17]. Nevertheless, more work is needed in order to reach a possible enhancement on Pollard's Rho algorithm. We hope that our proposal can encourage further discussions on the prime-field case which has been somewhat neglected so far.

The paper is organized as follows. In Section 2 we provide some background on summation polynomials and the relevant index calculus algorithm for elliptic curves. In Section 3 we introduce our index calculus algorithm for the prime-field case, providing some complexity evaluations and experimental results. Section 4 concludes the paper with comments and future perspectives.

## 2 Summation polynomials and the index calculus algorithm for elliptic curves

In 2004 Semaev introduced summation polynomials in order to use them to build an index calculus algorithm for elliptic curves. Let  $\mathbb{K}$  be a field of any characteristic. Given an elliptic curve  $E$  defined over  $\mathbb{K}$ , for each  $m \geq 2$  the  $m$ -th summation polynomial  $S_m$  is defined as follows:

**Definition 1** *Let  $\overline{\mathbb{K}}$  be the algebraic closure of the field  $\mathbb{K}$ . For any integer  $m \geq 2$ , the  $m$ -th summation polynomial  $S_m$  is an element of  $\mathbb{K}[X_1, \dots, X_m]$  and it is such that, given  $x_1, \dots, x_m \in \overline{\mathbb{K}}$ , then  $S_m(x_1, \dots, x_m) = 0$  if and only if there exist  $y_1, \dots, y_m \in \overline{\mathbb{K}}$  for which  $(x_1, y_1), \dots, (x_m, y_m) \in E(\overline{\mathbb{K}})$  and*

$$(x_1, y_1) + \dots + (x_m, y_m) = \infty$$

The existence of summation polynomials is proved by providing recursive formulae along the following properties.

**Theorem 2 ([19])** *Let  $E$  be an elliptic curve defined over a field  $\mathbb{K}$ , with  $\text{char}(\mathbb{K}) \neq 2, 3$ , by the Weierstrass equation*

$$y^2 = x^3 + Ax + B$$

*The summation polynomials for  $E$  are given as follows:*

$$\begin{aligned} S_2(X_1, X_2) &= X_1 - X_2, \\ S_3(X_1, X_2, X_3) &= (X_1 - X_2)^2 X_3^2 - 2[(X_1 + X_2)(X_1 X_2 + A) + 2B]X_3 + \\ &\quad + (X_1 X_2 - A)^2 - 4B(X_1 + X_2) \end{aligned}$$

and, for all  $m \geq 4$  and  $1 \leq k \leq m - 3$ , it holds

$$S_m(X_1, \dots, X_m) = \text{Res}_X(S_{m-k}(X_1, \dots, X_{m-k-1}, X), S_{k+2}(X_{m-k}, \dots, X_m, X))$$

where  $\text{Res}_X$  denotes the resultant polynomial of  $S_{m-k}$  and  $S_{k+2}$  with respect to the variable  $X$ . Moreover, the polynomial  $S_m$  is symmetric of degree  $2^{m-2}$  in each variable  $X_i$ , it is absolutely irreducible and

$$S_m(X_1, \dots, X_m) = S_{m-1}^2(X_1, \dots, X_{m-1})X_m^{2^{m-2}} + \dots$$

for any  $m \geq 3$ .

In the case of elliptic curves defined over fields of characteristics 2 or 3, it turns out that summation polynomials exist and have a similar form (and the same properties, see [19], [1]).

### 2.1 Index calculus algorithm for elliptic curves

In [19], Semaev proposed to use summation polynomials for an index calculus algorithm solving the discrete logarithm problem for elliptic curves. Semaev's idea was firstly developed by Gaudry [8] and Diem [1], and then many other papers on this topic followed. We describe here the current approach to these ideas. As we pointed out previously, all works subsequent to [19], except one, take into account only elliptic curves defined over extension fields.

Consider an elliptic curve  $E$  defined over a finite field  $\mathbb{F}_{q^n}$ , where  $q$  is a prime or a prime power. Let  $P \in E(\mathbb{F}_{q^n}) \setminus \{\infty\}$  and  $Q \in \langle P \rangle$ , where  $\langle P \rangle$  denotes the cyclic group generated by  $P$ . It is standard to restrict attention to points  $P$  of prime order  $r$ . Concerning the *relation collection* step, the general procedure is the following:

1. An  $\mathbb{F}_q$ -vector subspace  $V$  of  $\mathbb{F}_{q^n}$  is fixed, with  $\#V = q^\ell$  where  $1 \leq \ell < n$ . Then the factor base  $\mathcal{F}$  is defined as:

$$\mathcal{F} = \{F \in E(\mathbb{F}_{q^n}) \mid x(F) \in V\} \tag{2}$$

where  $x(F)$  denotes the  $x$ -coordinate of  $F$ .

2. Two random integers  $u, v$  are generated and the point  $R = uP + vQ$  is computed.
3. For a fixed integer  $m$ , points  $P_i \in \mathcal{F}$  with  $i \in \{1, \dots, m\}$  and such that

$$R = P_1 + \dots + P_m \tag{3}$$

are searched. Such  $P_i$ 's are obtained by solving the polynomial equation

$$S_{m+1}(X_1, \dots, X_m, x(R)) = 0 \tag{4}$$

and considering solutions of the form  $(x_1, \dots, x_m) \in V^m$ . Hence the Point Decomposition Problem is reduced to the solution of a multivariate polynomial equation thanks to summation polynomials. If  $R$  is not decomposable as a sum of points of  $\mathcal{F}$ , it is replaced by a new  $R$ .

4. To every  $x \in V$ , one point in  $\mathcal{F}$  having  $x$  as abscissa is chosen. Consequently,  $R$  can be written as  $R = \pm P_1 \pm \dots \pm P_m$ , where  $P_i \in \mathcal{F}$  and with signs depending on the selected points. A matrix  $M$  with  $|V| + 2$  columns is then constructed: each of the first  $|V|$  columns corresponds to a single element of  $V$ , while the two last columns correspond to the coefficients  $u$  and  $v$ , respectively. Given the relation  $R = \pm P_1 \pm \dots \pm P_m$ , a row of  $M$  is filled with the coefficients of the chosen points and the coefficients  $u, v$  such that  $R = uP + vQ$ .

5. The three steps above are repeated until at least  $|V|$  relations have been computed, i.e. at least  $|V|$  rows of  $M$  have been filled.

At this point the integer  $w$  such that  $Q = wP$  can be computed by executing the *linear algebra* step:

1. via Gauss' elimination, a linear dependency among some  $M$ 's rows is obtained, and then its corresponding linear dependency of points  $R$ . This leads to the sought-after relation

$$\lambda P + \mu Q = \infty \quad (5)$$

with  $\lambda, \mu \in \mathbb{Z}$ .

2.  $w$  is recovered from the modular equation  $\lambda + \mu w = 0 \pmod{r}$  if it is solvable.

As observed, while considering elliptic curves defined over extension fields, the subset  $V \subset \mathbb{F}_{q^n}$  is usually chosen as an  $\mathbb{F}_q$ -vector subspace of  $\mathbb{F}_{q^n}$ . Consequently, the Weil descent can be exploited to replace the system of the PDP with an equivalent one, which have polynomial equations over  $\mathbb{F}_q$ . Systems that arise are then usually solved with Groebner basis methods [5], [16], [10], [21], [4].

## 2.2 The prime-field case

The case when the considered elliptic curve  $E$  is defined over a prime field  $\mathbb{F}_p$  was originally taken into account by Semaev in [19], where he defines  $V$  as the subset  $\{x \in \mathbb{F}_p \mid x < p^{1/m+\delta}\}$  for some small  $\delta > 0$ . However, Semaev did not provide an algorithm to solve the systems arising in the point decomposition problem and this case was somewhat neglected so far.

In 2016, Petit *et al.* were the first, after Semaev, to face the prime-field case. In [17], they primarily observe that the constraints  $x_i \in V$  of item 3 in the above *relation collection* step are equivalent to the constraints  $L(X_i) = 0$  where  $L(z)$  is a polynomial defined as:

$$L = \prod_{v \in V} (z - v)$$

This observation can be exploited to transform Semaev's prime-field proposal into an actual algorithm. Furthermore, Petit *et al.* generalise it considering *suitable* algebraic (or rational) maps  $L$  over  $\mathbb{F}_p$  (i.e. maps  $L$  that are composition of low degrees maps  $L_1, \dots, L_t$ ) and defining  $V$  as the set  $\{x \in \mathbb{F}_p \mid L(x) = 0\}$ . Accordingly, the factor base  $\mathcal{F}$  is then defined as  $\{F \in E(\mathbb{F}_p) \mid x(F) \in V\}$ , where  $x(F)$  is the abscissa of  $F$ . To obtain a suitable map  $L$ , they introduce a pre-computation, which is specific for every elliptic curve  $E$ .

The index calculus algorithm considered by Petit *et al.* to solve the ECDLP in the prime-field case is the same described in the previous section. However, given a point  $R = uP + vQ \in E(\mathbb{F}_p)$ , in order to solve the point decomposition problem for  $R$  with respect to  $\mathcal{F}$ , they suggest to solve the following polynomial equation system:

$$\begin{cases} S_{m+1}(X_{11}, \dots, X_{m1}, x(R)) = 0 \\ X_{i,j+1} = L_j(X_{i,j}) & i = 1, \dots, m; j = 1, \dots, t-1 \\ L_t(X_{i,t}) = 0 & i = 1, \dots, m \end{cases} \quad (6)$$

where  $x(R)$  is the  $x$ -coordinate of  $R$ . In their work, Petit *et al.* provide some experimental results and partial complexity analysis, with which we are going to compare.

### 3 Our proposal

In this section we introduce a new variant, for elliptic curves defined over prime fields, of the index calculus algorithm based on summation polynomials. Our proposal differs significantly from the algorithm reported in Section 2.2, since after the *relation collection* step we do not have a *linear algebra* step, but rather a *relation solving* step, as we detail in the following.

Let us consider an elliptic curve  $E$  defined over a prime field  $\mathbb{F}_p$  and having a prime number of rational points, i.e.  $\#E(\mathbb{F}_p) = r$  with  $r$  prime integer. With  $P$  and  $Q$  we denote again two points of  $E(\mathbb{F}_p)$  such that  $Q = wP$ .

#### *Relation collection step*

1. The factor base  $\mathcal{F}$  is not fixed, rather it is initialized as the empty set. The final size  $s$  of  $\mathcal{F}$  is fixed together with an integer  $m$ , named *decomposition constant*. As usual (see for example [17, pag.7]) we take  $s^m \approx p$ .
2. For random integers  $u, v$ , the point  $R = uP + vQ$  is computed and added to the factor base  $\mathcal{F}$ .
3.  $R$  does not need to be written as a sum of multiples of points in  $\mathcal{F}$ , since  $R$  itself is a point of  $\mathcal{F}$ .
4. The procedure in item 2 is repeated until at least  $s$  different points are added to  $\mathcal{F}$ .



*Relation solving step*

1. Using summation polynomials, a linear dependency among points in  $\mathcal{F}$  is determined, but not with linear algebra. To be more precise, we consider the multivariate polynomial equation  $S_m(X_1, \dots, X_m) = 0$ , where  $S_m$  is the  $m$ -th summation polynomial for  $E$ , and we search solutions of the form  $(x_1, \dots, x_m) \in V^m$  where:

$$V = \{x \mid x = x(R) \text{ for some } R \in \mathcal{F}\}. \quad (7)$$

Once that such a solution has been found, it is trivial to determine  $m$  points  $P_i$  such that  $-P_i \in \mathcal{F}$  or  $P_i \in \mathcal{F}$ , and

$$P_1 + \dots + P_m = \infty. \quad (8)$$

Hence, the linear congruence  $\lambda P + \mu Q = \infty$ , with  $\lambda, \mu \in \mathbb{Z}$ , is deduced.

2. The discrete logarithm  $w$  of  $Q$  with respect to  $P$  is recovered solving the modular equation  $\lambda + \mu w = 0 \pmod{r}$ .

The algorithm described above differs in two key points from the standard index calculus algorithm usually considered when exploiting summation polynomials. First,  $\mathcal{F}$  is not fixed deterministically at the beginning of the algorithm. Instead, it is constructed step by step adding the random points  $R$  progressively. This prevents the computation of many relations (i.e. resolutions of the PDP), which traditionally requires expensive Grobner basis computations. Second, in our second step (which is traditionally the *linear algebra* step) we actually do not execute linear algebra. On the other hand, the linear dependency among points  $R$  is obtained by looking for suitable solutions of the equation  $S_m(X_1, \dots, X_m) = 0$ . To do that, we do need to compute one Groebner basis, but it is just one. We now specify which is the ideal basis that we use for our Groebner basis computation.

### 3.1 The system to be solved

In order to deduce a linear relation among points in  $\mathcal{F}$  we solve the polynomial equation  $S_m(x_1, \dots, x_m) = 0$  considering only solutions of the form  $(x_1, \dots, x_m) \in V^m$ . Formally, this is reached solving the polynomial equation system

$$\begin{cases} S_m(X_1, \dots, X_m) = 0 \\ f(X_1) = 0 \\ \dots \\ f(X_m) = 0 \end{cases} \quad (9)$$

where  $f(z)$  is the polynomial generating the vanishing ideal of  $V \subset \mathbb{F}_p$ . In particular:

$$f(z) = \prod_{v \in V} (z - v)$$

Although we would be content with any solution of  $S_m(X_1, \dots, X_m) = 0$  of the form

$$(x_1, \dots, x_m) \in V^m,$$

we add further constraints on the corresponding variables  $X_1, \dots, X_m$  in order to lower the degrees in system (9). Indeed, given any partition  $\{\mathcal{F}_i \mid i = 1, \dots, m\}$  of  $\mathcal{F}$ , the sets

$$V_i = \{x \mid x = x(R) \text{ for some } R \in \mathcal{F}_i\} \quad i = 1, \dots, m$$

can be trivially deduced and the following polynomials constructed:

$$f_i(X_i) = \prod_{v \in V_i} (X_i - v) \quad i = 1, \dots, m. \quad (10)$$

The idea of partitioning  $\mathcal{F}$  is similar to what Galbraith and Gebregiyorgis proposed in [7] and it allows us to limit the degrees of the univariate polynomials in  $X_1, \dots, X_m$ . Therefore, the system to be solved in item 1 of the *Relation solving step* of our algorithm is the following:

$$\begin{cases} S_m(X_1, \dots, X_m) = 0 \\ f_1(X_1) = 0 \\ \dots \\ f_m(X_m) = 0. \end{cases} \quad (11)$$

The constraints on  $X_1, \dots, X_m$  of system (11) break the symmetry, since  $m$  disjoint factor bases are used and only  $P_i \in \mathcal{F}_i$  (or  $-P_i \in \mathcal{F}_i$ ) are then allowed in (8). This will obviously impact the probability of finding a solution of system (11): a complexity analysis is reported in the next subsection.

### 3.2 Complexity analysis

The whole complexity of the proposed algorithm relies on the solution of a single polynomial equation system. Let  $T(E, m, V)$  be the computational cost of solving the system (11). Under the assumption that the sizes of the  $\mathcal{F}_i$ 's are about the same (which is in practice what we always enforce), the expected number of linear combinations  $P_1 + \dots + P_m$  (with  $P_i$  or  $-P_i$  in  $\mathcal{F}_i$ ) is about

$$\left(\frac{s}{m}\right)^m.$$

Then the probability that a random point of  $E(\mathbb{F}_p)$  could be written as a sum  $P_1 + \dots + P_m$  (with  $P_i$  or  $-P_i$  in  $\mathcal{F}_i$ ) is nearly

$$\frac{s^m}{p m^m}.$$

Therefore, the total cost of the algorithm is given by:

$$\frac{p m^m}{s^m} T(E, m, V).$$

In order to reach an improvement with respect to algorithms for generic cyclic groups, this complexity should be smaller than  $p^{1/2}$ . We do not claim this result. Indeed our complexity analysis is partial, since we are unable to estimate the complexity of solving our polynomial equation system. We compare our complexity with that of the algorithm of Petit *et al.* which, following the notation of our paper, can be expressed as:

$$\frac{p \cdot m!}{s^{m-1}} T'(E, m, V)$$

where  $T'(E, m, V)$  is the cost to solve one of the systems proposed in [17], which are of the form of system (6). We can assume

$$T(E, m, V) \approx T'(E, m, V).$$

As a consequence, for a fixed  $m$  our algorithm outperforms that of Petit *et al.* when  $p$  increases. In particular, the outperforming happens when

$$s(m!) > m^m, \quad \text{that is,} \quad \sqrt[m]{p} > \frac{m^m}{m!}$$

Considering  $m \in \{3, 4, \dots, 8\}$  ( $S_8$  is nowadays the summation polynomial with the highest index that has been computed [3]), the following table shows from which bit size of  $p$  our algorithm would outrun that in [17].

$m$	3	4	5	6	7	8
bit size	7	14	24	37	52	70

We underline that the bit sizes of  $p$  reported in the above table are all smaller than the bit sizes used in cryptographic applications.

The systems arising in our algorithm need to be analyzed more deeply in order to reach a more precise complexity evaluation and it should be also inspected whether algorithms different from Grobner basis methods are more suitable to tackle such systems. Hence, proving theoretically the assumption  $T(E, m, V) \approx T'(E, m, V)$  is a challenging issue. However, the experimental results reported in next section show that the assumption has heuristic evidence.

### 3.3 Experimental results

We report some computer experiments, executed on elliptic curves defined over prime fields  $\mathbb{F}_p$  with the bit size of  $p$  lying in  $\{11, 12, \dots, 22\}$ . The possible values for the *decomposition constant*  $m$  that we considered belong to the set  $\{3, 4, 5\}$ . Following [11] and [20], for  $m = 4$  we substituted  $S_4(X_1, X_2, X_3, X_4) = 0$  with

$$\begin{cases} S_3(X_1, X_2, Y) = 0 \\ S_3(X_3, X_4, Y) = 0 \end{cases}$$

introducing the variable  $Y$  and obtaining the following system in place of (11):

$$\begin{cases} S_3(X_1, X_2, Y) = 0 \\ S_3(X_3, X_4, Y) = 0 \\ f_1(X_1) = \dots = f_4(X_4) = 0. \end{cases}$$

Similarly, for  $m = 5$  the equation  $S_5(X_1, X_2, X_3, X_4, X_5) = 0$  was substituted with

$$\begin{cases} S_3(X_1, X_2, Y) = 0 \\ S_4(X_3, X_4, X_5, Y) = 0 \end{cases}$$

introducing the variable  $Y$  and obtaining the following system in place of (11):

$$\begin{cases} S_3(X_1, X_2, Y) = 0 \\ S_4(X_3, X_4, X_5, Y) = 0 \\ f_1(X_1) = \dots = f_5(X_5) = 0. \end{cases}$$

For every possible values of  $m$  and the bit size of  $p$ , we considered ten different elliptic curves. For each of them, we use our algorithm to solve ten instances of the discrete logarithm problem, taking the average execution time. From the ten average times obtained (one for every elliptic curve), we considered the worst one. All experiments were performed with MAGMA on a CPU with an Intel Xeon Process 5460 at 3.16 GHz with a cache of 6 MB. The collected data are showed in the following table, together with an extra column with the experimental results reported in [17] and concerning the average time necessary to solve one of the systems arising in Petit *et al.* algorithm. The data for even bit sizes are relatively to primes of a specific form; the ones for odd bite sizes are relatively to generic primes. All of them were obtained exploiting a computational power similar to ours.

bit size	$m = 3$	$m = 4$	$m = 5$	Petit <i>et al.</i> ( $m = 3$ )
11	0.018	0.032	0.153	0.02
12	0.034	0.059	0.185	0.15
13	0.079	0.077	0.408	0.13
14	0.146	0.118	0.256	1.16
15	0.401	0.350	0.383	1.14
16	1.248	0.931	0.742	9.08
17	3.777	1.669	2.895	9.09
18	7.452	6.298	7.783	51.87
19	23.291	17.983	45.291	59.87
20	60.135	85.127	139.006	438.57
21	447.245	619.166	594.751	454.79
22	600.324	1808.291	1607.601	5163.46

Comparing our results and those of Petit *et al.* for the case  $m = 3$ , it is easy to observe that they are quite similar, heuristically corroborating the assumption

$T(E, m, V) \approx T'(E, m, V)$  that we considered in the complexity analysis. However, our results concern the whole resolution of the discrete logarithm problem (as executing our algorithm just one system needs to be solved), while Petit *et al.* results concern the time needed to tackle one out of  $s$  systems that need to be solved in the execution of their algorithm. So the times reported in the last column should be multiplied by  $s$  to obtain the time needed by Petit *et al.* algorithm to solve the whole discrete logarithm problem. For example, if  $p$  has a bit size equal to 11, our algorithm is at least 10 times faster than their. The same remark holds considering Semaev’s original proposal in [19].

We performed many preliminary computations with several monomial orderings available in MAGMA. Since we did not find any which would clearly outperforms the others with our systems, we decided to use only the “graded reverse lexicographic order” for the computations reported in the table.

## 4 Conclusions

We have presented a new index calculus algorithm that exploits summation polynomials for solving the discrete logarithm problem in elliptic curves defined over prime fields. This algorithm significantly differs from the algorithm, for the same case, proposed by Petit *et al.* at the beginning at 2016 and it reduces to one the number of polynomial equation systems to be solved during its execution. A preliminary complexity analysis suggests that our algorithm improves the one of Petit *et al.* when  $m$  is fixed and  $p$  increases. However, the complexity of the single system to be tackled in our algorithm is far to be well understood and it needs to be deeply analyzed. Furthermore, it should be investigated whether other algorithms different from Grobner basis algorithms are better to solve such systems.

## References

1. C. Diem. On the discrete logarithm problem in elliptic curves. *Compositio Mathematica*, 147:75–104, 2011.
2. C. Diem. On the Subexponentiality of the Elliptic Curve Discrete Logarithm Problem over Extension Fields. *Workshop: Grobner Bases in Cryptography, Coding Theory, and Algebraic Combinatorics organized by Mikhail Klin, Ludovic Perret, Massimiliano Sala*, May 2006.
3. J. C. Faugère, L. Huot, A. Joux, G. Renault, and V. Vitse. Symmetrized summation polynomials: using small order torsion points to speed up elliptic curve index calculus. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Springer Berlin Heidelberg:40–57, 2014.
4. J. C. Faugère, P. Gaudry, L. Huot, and G. Renault. Using symmetries in the index calculus for elliptic curves discrete logarithm. *Journal of Cryptology*, 27(4):595–635, 2014.
5. J. C. Faugère, L. Perret, C. Petit, and G. Renault. Improving the complexity of index calculus algorithms in elliptic curves over binary fields. *Annual International*

- Conference on the Theory and Applications of Cryptographic Techniques*, Springer Berlin Heidelberg, 2012.
6. S. D. Galbraith and P. Gaudry. Recent progress on the elliptic curve discrete logarithm problem. *Designs, Codes and Cryptography*, 78.1:51–72, 2016.
  7. S. D. Galbraith and S. W. Gebregiyorgis. Summation polynomial algorithms for elliptic curves in characteristic two. *International Conference in Cryptology in India*, Springer International Publishing:409–427, , December 2014.
  8. P. Gaudry. Index calculus for abelian varieties of small dimension and the elliptic curve discrete logarithm problem. *Journal of Symbolic Computation*, 44(12):1690–1702, 2009.
  9. Y. J. Huang, C. Petit, N. Shinohara, and T. Takagi. Improvement of Faugere et al.’s Method to Solve ECDLP. *International Workshop on Security*, Springer Berlin Heidelberg:115–132, November 2013.
  10. A. Joux and V. Vitse. Elliptic curve discrete logarithm problem over small degree extension fields. *Journal of Cryptology*, 2013: 1-25.
  11. K. Karabina. Point decomposition problem in binary elliptic curves. *International Conference on Information Security and Cryptology*, Springer International Publishing, 2015.
  12. N. Koblitz. Elliptic curves cryptosystems. *Mathematics of Computation*, 48 (177):203–209, 1987, January.
  13. A. Menezes, S. Vanstone, and T. Okamoto. Reducing Elliptic Curve Logarithms to Logarithms in a Finite Field. *IEEE Transactions on information Theory* 39.5: 1639-1646, 1993.
  14. V. S. Miller. Use of elliptic curves in cryptography. *Conference on the Theory and Application of Cryptographic Techniques*, Springer Berlin Heidelberg, 1985.
  15. Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.
  16. C. Petit and J. J. Quisquater. On polynomial systems arising from a Weil descent. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 451–466. Springer, 2012.
  17. Christophe Petit, Michiel Kusters, and Ange Messeng. Algebraic Approaches for the Elliptic Curve Discrete Logarithm Problem over Prime Fields. *IACR International Workshop on Public Key Cryptography*, Springer Berlin Heidelberg:3–18, 2016, March.
  18. J. M. Pollard. Monte Carlo methods for index computation mod  $p$ . *Mathematics of computation*, 32.143:918–924, 1978.
  19. I. Semaev. Summation polynomials and the discrete logarithm problem on elliptic curves. *IACR Cryptology ePrint Archive 2004*, (2004): 31.
  20. I. Semaev. New algorithm for the discrete logarithm problem on elliptic curves. *arXiv preprint arXiv:1504.01175*, 2015.
  21. M. Shantz and E. Teske. Solving the elliptic curve discrete logarithm problem using Semaev polynomials, Weil descent and Groebner basis methods. An experimental study. *Number Theory and Cryptography*, Springer Berlin Heidelberg, 94-107, 2013.