PQCrypto 2023

Fast Enumeration Algorithm for Multivariate Polynomial Systems over General Finite Fields

Hiroki Furue and Tsuyoshi Takagi

Department of Mathematical Informatics,
The University of Tokyo

August 16th, 2023

Outline

1. **Background**
2. Results of Bouillaguet et al. (in $\mathbb{F}_2$)
3. Proposed Algorithm (in $\mathbb{F}_q$)
4. Conclusion

MPKC

- Multivariate Public Key Cryptosystems (MPKC)
  - based on the difficulty of **MQ problem**
  - candidates for **post quantum cryptosystems**
  - mainly used for **digital signature**

MQ (Multivariate Quadratic equations) problem

Given $\mathcal{F} = (f_1, \ldots, f_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ with $\deg f_i = 2$,

find *one* solution $(a_1, \ldots, a_n) \in \mathbb{F}_q^n$ such that

$$\mathcal{F}(a_1, \ldots, a_n) = \mathbf{0} \in \mathbb{F}_q^m.$$

Enumeration Problem

The security of MPKC depends on the difficulty of some algebraic problems. (MinRank problem etc.)

Enumeration problem (exhaustive search)

・ $n, d \in \mathbb{N}$, q: a prime power

Given $f \in \mathbb{F}_q[x_1, \ldots, x_n]$ with degree d,

find $f(a_1, \ldots, a_n)$ for all inputs $a_1, \ldots, a_n \in \mathbb{F}_q$.

・ a fundamental tool of some algebraic attacks

(claw finding attack, hybrid approach, crossbred, polynomial XL)

Classical Approach

- For any inputs $a_1, \dots, a_n$, directly substitute $a_1, \dots, a_n$ for f.

$$f(x_1, \dots, x_n) = x_1 \cdot f_1(x_1, \dots, x_n) + f_2(x_2, \dots, x_n)$$

n var., d deg. $\quad$ n var., $d-1$ deg. $\quad$ $n-1$ var., d deg.

$$\vdots$$

Complexity: $O\left(\binom{n+d}{d}\right)$ $\quad$ ※ $\binom{n+d-1}{d-1} + \binom{n-1+d}{d} = \binom{n+d}{d}$

Complexity of the enumeration problem

$$O\left(\binom{n+d}{d} \cdot q^n\right)$$

Fast Enumeration Algorithm

The result in $\mathbb{F}_2$ [Bouillaguet et al., CHES 2010]

In $\mathbb{F}_2$, the enumeration problem can be solved

by $O(d \cdot 2^n)$ operations over $\mathbb{F}_2$ after initialization phase.

(using **Gray codes** and **derivatives**)

Our Proposed Algorithm

In $\mathbb{F}_q$ with a prime number q,

the enumeration problem can be solved

by $O(d \cdot q^n)$ operations over $\mathbb{F}_q$ after initialization phase.

※ This algorithm is applicable to $\mathbb{F}_{p^r}$.

Outline

1. Background
2. **Results of Bouillaguet et al. (in $\mathbb{F}_2$)**
3. Proposed Algorithm (in $\mathbb{F}_q$)
4. Conclusion

Gray Code

Gray Code:

an ordering of the binary vector space such that two successive values differ in only one bit.

Standard base-2 system	Gray code
000	000
001	001
010	011
011	010
100	110
101	111
110	101
111	100

Derivatives in $\mathbb{F}_2$

$\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{F}_2^n,\ \boldsymbol{e}_i = (0, \dots, 0, \underset{\uparrow}{1}, 0, \dots, 0) \in \mathbb{F}_2^n$

i

For $1 \leq i \leq n$,

$$\partial_i f(\boldsymbol{x}) \coloneqq f(\boldsymbol{x} + \boldsymbol{e}_i) + f(\boldsymbol{x})$$

$$\Rightarrow \quad f(\boldsymbol{x} + \boldsymbol{e}_i) = f(\boldsymbol{x}) + \partial_i f(\boldsymbol{x})$$

ex) $\partial_i 1 = 0,\ \partial_i x_i = 1$ ※ $x_i^2 = x_i$ in $\mathbb{F}_2$

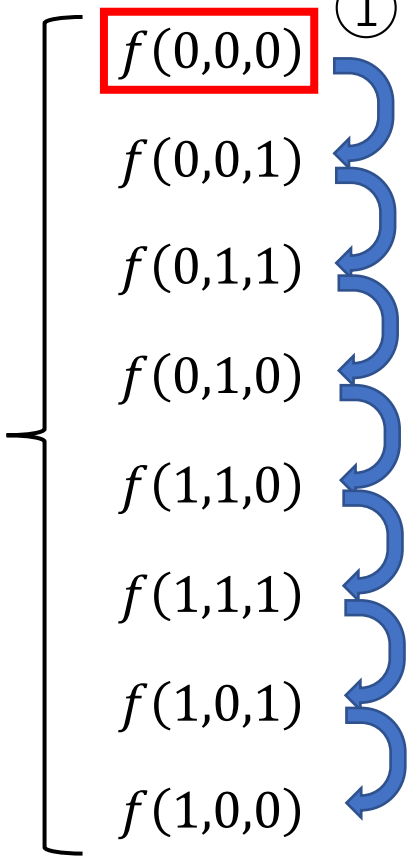
$\deg f = d \quad \Rightarrow \quad \deg \partial_i f \leq d - 1$

Case of $d = 1$

$\deg f = 1$ ➡ $\partial_1 f = c_1, \ldots, \partial_n f = c_n$ **constant**

$(n = 3)$

In the order of Gray code



① Evaluate f at initial point $(0,0,0)$.

differ in only the 3rd bit

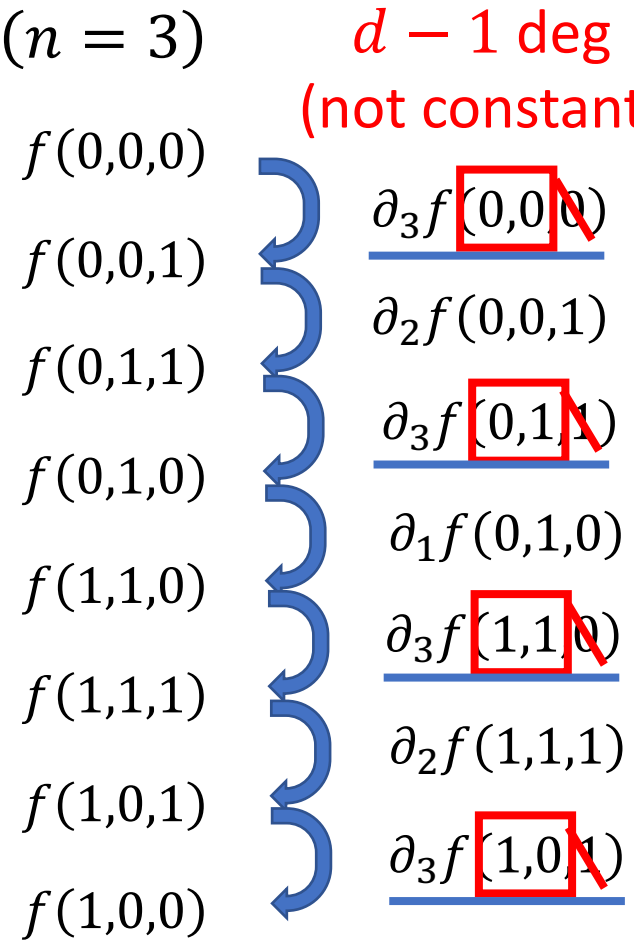
$$f(0,0,1) = f(0,0,0) + \underbrace{\partial_3 f(0,0,0)}_{= c_3}$$

② Compute successive outputs by adding c_i.

Case of $d \geq 2$

$(n = 3)$

$d - 1$ deg (not constant)



In $\mathbb{F}_2$, $x_i^2 = x_i$ and thus $\partial_i f$ does not include x_i.

$\partial_3 f$ (with degree $d - 1$) is evaluated in the order of the Gray code for the 1st and 2nd variables.

Derivatives can be enumerated recursively in the order of Gray codes.

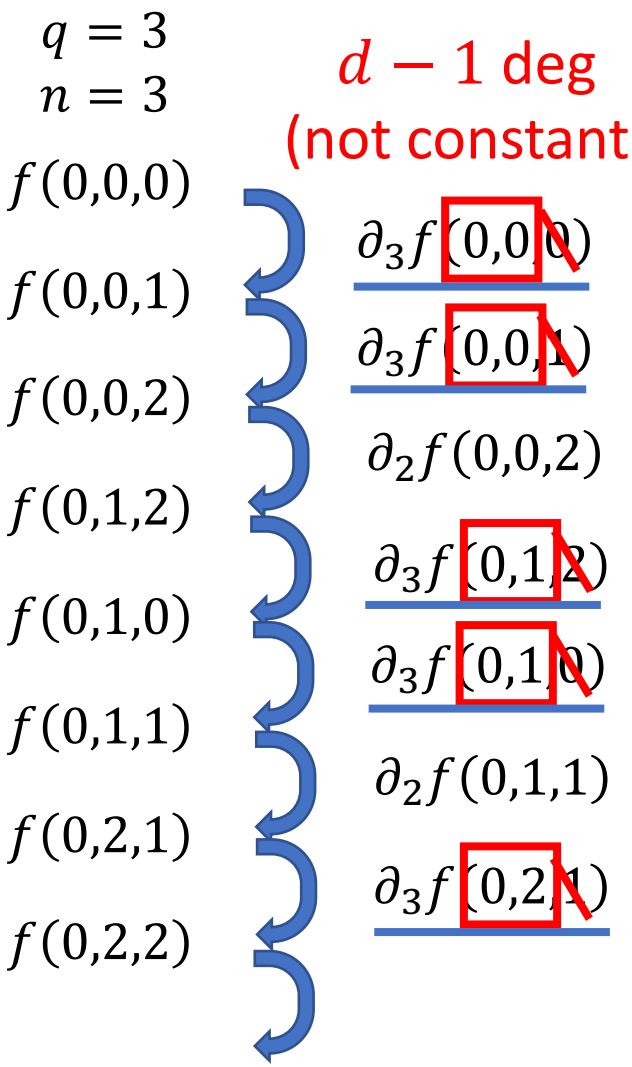
Time Complexity:

$$O(d \cdot 2^n)\left(+O\left(n^{2d}\right)\right)$$

Case of $\mathbb{F}_q$ $(q \neq 2)$

$q = 3$
$n = 3$

$d - 1$ deg
(not constant)



There exist Gray codes over $\mathbb{F}_q$
(q-ary gray code)

$x_i^2 \neq x_i$ in $\mathbb{F}_q$

$\partial_i f$ cannot be evaluated
in the order of Gray codes.

Outline

1. Background
2. Results of Bouillaguet et al. (in $\mathbb{F}_2$)
3. **Proposed Algorithm (in $\mathbb{F}_q$)**
4. Conclusion

Main Idea

- $\mathbb{F}_q$ with a prime q
- We also use **derivatives**. $\partial_i f(\boldsymbol{x}) \coloneqq f(\boldsymbol{x} + \boldsymbol{e}_i) - f(\boldsymbol{x})$

  ex) $\partial_1 1 = 0, \partial_1 x_1 = 1, \partial_1 x_1^2 = 2x_1 + 1$
- We use a lexicographic order instead of Gray codes.

Gray codes

$f(0,0)$
↓
$f(0,1)$
↓
$f(1,1)$
↓
$f(1,0)$

branching structure

Our order $(q = 3)$

$f(0,0) \rightarrow f(0,1) \rightarrow f(0,2)$
↓
$f(1,0) \rightarrow f(1,1) \rightarrow f(1,2)$
↓
$f(2,0) \rightarrow f(2,1) \rightarrow f(2,2)$

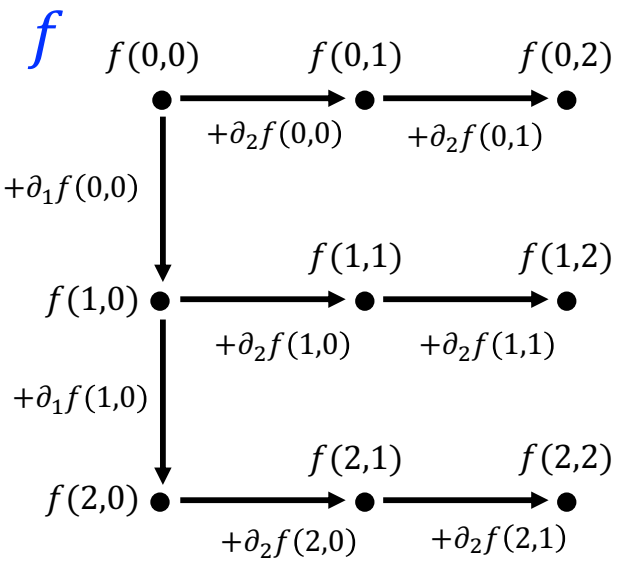
Formal Description

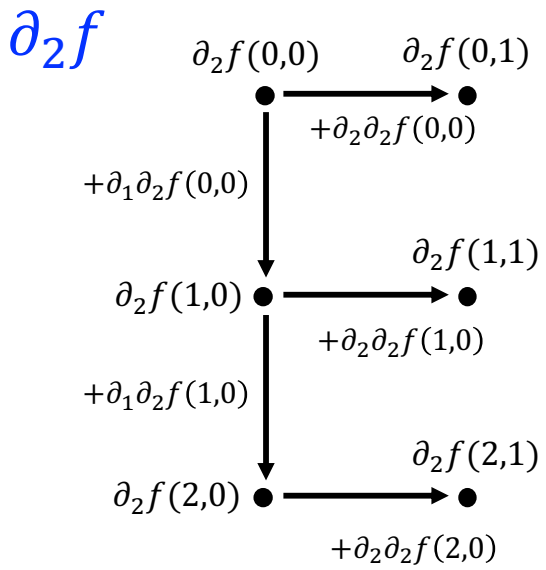
For $\boldsymbol{x} = (x_1, \dots, x_{k-1}, \underset{\neq 0}{x_k}, 0, \dots, 0) \in \mathbb{F}_q^n$ and $\boldsymbol{a} = (a_1, \dots, a_n) \in \mathbb{Z}_{\geq 0}^n$,

$\partial^{\boldsymbol{a}} f(\boldsymbol{x}) = \partial_1^{a_1} \cdots \partial_n^{a_n} f(\boldsymbol{x})$ is computed as follows:

$$\partial^{\boldsymbol{a}} f(x_1, \dots, x_{k-1}, x_k, 0, \dots, 0) =$$

$$\partial^{\boldsymbol{a}} f(x_1, \dots, x_{k-1}, x_k - 1, 0, \dots, 0) + \partial^{\boldsymbol{a}+\boldsymbol{e}_k} f(x_1, \dots, x_{k-1}, x_k - 1, 0, \dots, 0)$$





Algorithm Design

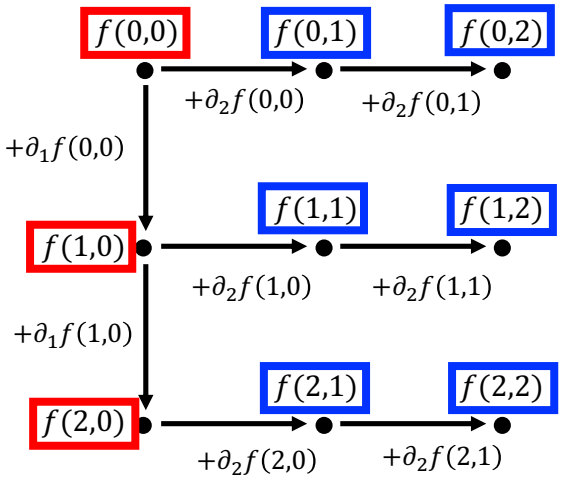
To realize the branching structure,

we store some outputs for each $\partial^{\boldsymbol{a}} f$.

・ $D[\boldsymbol{a}, k]$ stores the output $\partial^{\boldsymbol{a}} f(* \cdots * \underset{\underset{k}{\uparrow}}{0} \cdots 0)$.

$D[(0,0), 1]$

$D[(0,0), 2]$



※ small cost of memory consumption compared with [BFS10]

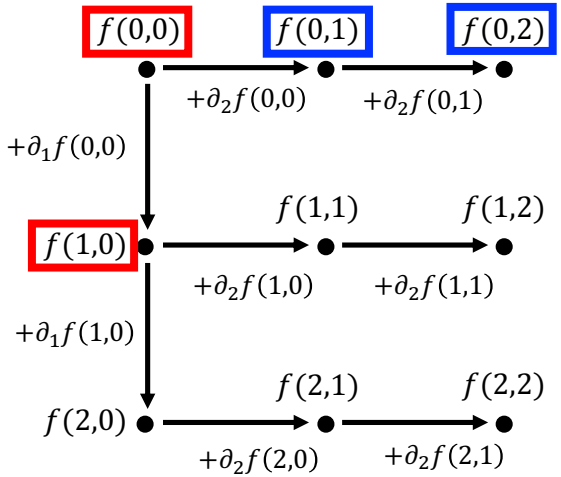
Example

Input: $f = 2x_1^2 + x_1x_2 + x_2 + 1 \in \mathbb{F}_3[x_1, x_2]$

$(q = 3, n = 2, d = 2)$

$D[(0,0), 1]$

$D[(0,0), 2]$



$D[(0,0), 1] = 1$ ⇨ $D[(0,0), 1] = 1$, $D[(0,0), 2] = 2$ ($D[(0,0), 1] + \partial_2 f(0,0) = 1$) ⇨ $D[(0,0), 1] = 1$, $D[(0,0), 2] = 0$ ($D[(0,0), 2] + \partial_2 f(0,1) = 1$) ⇨ $D[(0,0), 1] = 3$, $D[(0,0), 2] = 0$ ($D[(0,0), 1] + \partial_1 f(0,0) = 2$)

Complexity

For each input, the update requires $O(d)$ manipulations.

Time Complexity: $O(d \cdot q^n)\left(+O\left(\binom{n+d}{d}^2\right)\right)$

negligible initial phase

※ [BFS10]: $O(d \cdot 2^n)$ in $\mathbb{F}_2$

- In the case of $\mathbb{F}_{p^r}$ with a prime p,

  f: n variables in $\mathbb{F}_{p^r}$

  ⇓

  $f_1, \ldots, f_r$: $r \cdot n$ variables in $\mathbb{F}_p$

  the time complexity is given as $O(r \cdot d \cdot p^{r \cdot n})$.

Outline

1. Background
2. Results of Bouillaguet et al. (in $\mathbb{F}_2$)
3. Proposed Algorithm (in $\mathbb{F}_q$)
4. **Conclusion**

Conclusion

- We construct a new enumeration algorithm for a single polynomial over finite fields $\mathbb{F}_q$.
- Our algorithm achieves an equivalent efficiency to the result of Bouillaguet et al. in $\mathbb{F}_2$.
- The proposed algorithm is applicable to general finite fields with a small cost.
- The proposed algorithm can be used to solve the MQ problem more efficiently compared with the natural exhaustive search.