

Improved preprocessing for the Crossbred algorithm and application to the MQ problem

Damien Vidal

joint work with Claire Delaplace and Sorina Ionica

11/08/2026

Workshop AM-PQC



Multivariate polynomials in cryptography

MQ problem

$\mathcal{F} = \{f_1, f_2, \dots, f_m\}$ s.t. $f_i \in \mathbb{F}_q[x_1, \dots, x_n]$ and $\deg(f_i) = 2$.

Find $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{F}_q^n$ such that :

$$f_1(\mathbf{a}) = f_2(\mathbf{a}) = \dots = f_m(\mathbf{a}) = 0.$$

Multiple schemes based on the hardness of MQ : UOV and variants (MAYO, SNOVA, QR-UOV), MQOM

We focus on the case $q = 2$.

Known methods

- Gröbner basis
- Linearisation
- Exhaustive search

In practice : a mix of all of the above.

Macaulay matrix and its sub-matrix

$$\begin{array}{c} f_1 \\ f_2 \\ x_3 f_1 \\ x_2 f_1 \\ x_1 f_1 \\ x_3 f_2 \\ x_2 f_2 \\ x_1 f_2 \end{array} \begin{pmatrix} x_1 x_2 x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 & x_1 & x_2 & x_3 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$Mac_{\leq D, m}(\mathcal{F})$

Rows: Polynomials uf_i st. $\deg(uf_i) \leq D$.

Columns: Monomials M st. $\deg(M) \leq D$.

Macaulay matrix and its sub-matrix

$$\begin{array}{c}
 f_1 \\
 f_2 \\
 x_3 f_1 \\
 x_2 f_1 \\
 x_1 f_1 \\
 x_3 f_2 \\
 x_2 f_2 \\
 x_1 f_2
 \end{array}
 \begin{pmatrix}
 x_1 x_2 x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 & x_1 & x_2 & x_3 & 1 \\
 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\
 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0
 \end{pmatrix}$$

$Mac_{\leq 3, \geq 2, 2}^2(\mathcal{F})$

$$Mac_{\leq D, \geq d, m}^k(\mathcal{F})$$

Rows: Polynomials uf_i st. $\deg(uf_i) \leq D$ and $\deg_k(u) \geq d - 1$

Columns: Monomials M st. $\deg(M) \leq D$.

A bit of context

Proposed by Antoine Joux and Vanessa Vitse in 2017.
Its implementation breaks records.

A bit of context

Proposed by Antoine Joux and Vanessa Vitse in 2017.
Its implementation breaks records.

The algorithm has parameters D , d and k and takes $\mathcal{F}$ as input.

A bit of context

Proposed by Antoine Joux and Vanessa Vitse in 2017.
Its implementation breaks records.

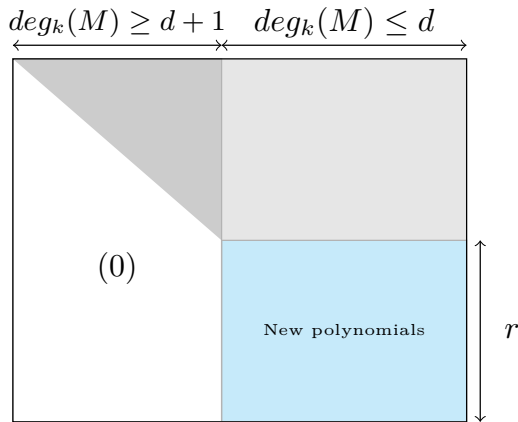
The algorithm has parameters D , d and k and takes $\mathcal{F}$ as input.

Preprocessing

- With linear algebra on $Mac_{\leq D, \geq d, m}^k(\mathcal{F})$, generate new polynomials $p_1, \dots, p_r$ of degree at most d over the first k variables.

Preprocessing step

Computation of polynomials f in $Mac_{\leq D, \geq d, m}^k(\mathcal{F})$ such that $f = \sum_i u_i f_i$ with $\deg(u_i) \leq D - 2, \deg_k(u_i) \geq d - 1, \deg_k(f) \leq d$.



Resolution

- 1 Specify the last $n - k$ variables in $\mathbb{F}_2^{n-k}$.
- 2 Check if the specified system $\mathcal{F}^* \cup \{p_1^*, \dots, p_r^*\}$ has a solution. If not, go to next iteration.

We assume that the specified system $\mathcal{F}^* \cup \{p_1^*, \dots, p_r^*\}$ is solved with linearisation.

Admissible parameters

(D, d, k) admissible iff $M^* = \text{Mac}_{\leq d, m}(\mathcal{F}^* \cup \{p_1^*, \dots, p_r^*\})$ has more independent rows than columns.

Experiments

We note:

- r the number of generated polynomials.
- N the number of columns of $Mac_{\leq d,m}(\mathcal{F}^*)$.

m	n	k	D	d	r	N
63	30	17	4	1	2930	18
59	28	17	4	1	2591	18
49	23	18	4	1	1944	19
69	33	22	4	2	5945	254
63	30	20	4	2	6429	211
43	20	14	3	2	238	106
47	22	14	3	2	294	106
79	38	27	4	3	7019	3304
83	40	27	4	3	8097	3304

Question

What is the minimal value possible of r to get N independant polynomials in each specialization ?

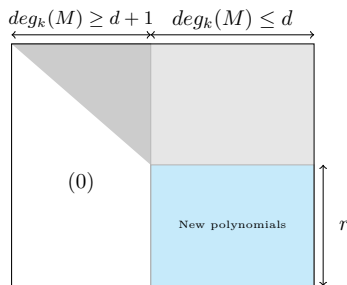
- Inefficient preprocessing (we generate far more polynomials than necessary).

- Inefficient preprocessing (we generate far more polynomials than necessary).
- New criterion to **remove redundant rows** during the preprocessing.
- Determine the **minimal numbers** of polynomials required.

Redundancy in generated polynomials

Vector space of newly generated polynomials

$$\mathcal{K}_{\leq D, \geq d}^k = \left\{ f = \sum_i u_i f_i \mid \deg(u_i) \leq D-2, \deg_k(u_i) \geq d-1, \deg_k(f) \leq d \right\}$$



Redundancy in generated polynomials

Vector space of newly generated polynomials

$$\mathcal{K}_{\leq D, \geq d}^k = \left\{ f = \sum_i u_i f_i \mid \deg(u_i) \leq D - 2, \deg_k(u_i) \geq d - 1, \deg_k(f) \leq d \right\}$$

We have

$$\begin{aligned} \mathcal{K}_{\leq D, \geq d}^k &\subset \mathcal{K}_{\leq D+1, \geq d}^k, \\ \sum_{i \in \{k+1, \dots, n\}} x_i \mathcal{K}_{\leq D, \geq d}^k &\subset \mathcal{K}_{\leq D+1, \geq d}^k. \end{aligned}$$

Redundancy in generated polynomials

We have

$$\mathcal{K}_{\leq D, \geq d}^k \subset \mathcal{K}_{\leq D+1, \geq d}^k,$$
$$\sum_{i \in \{k+1, \dots, n\}} x_i \mathcal{K}_{\leq D, \geq d}^k \subset \mathcal{K}_{\leq D+1, \geq d}^k.$$

Example

We consider a system of 49 polynomials and 23 variables and fix $d = 2$ and $k = 18$. We need to generate $N = 172$ polynomials. We have:

- $\dim(\mathcal{K}_{\leq 3, \geq 2}^{18}) = 66$
- $\dim(\mathcal{K}_{\leq 4, \geq 2}^{18}) = 3608$

Redundancy in generated polynomials

We have

$$\begin{aligned} \mathcal{K}_{\leq D, \geq d}^k &\subset \mathcal{K}_{\leq D+1, \geq d}^k, \\ \sum_{i \in \{k+1, \dots, n\}} x_i \mathcal{K}_{\leq D, \geq d}^k &\subset \mathcal{K}_{\leq D+1, \geq d}^k. \end{aligned}$$

Example

We consider a system of 49 polynomials and 23 variables and fix $d = 2$ and $k = 18$. We need to generate $N = 172$ polynomials. We have:

- $\dim(\mathcal{K}_{\leq 3, \geq 2}^{18}) = 66$
- $\dim(\mathcal{K}_{\leq 4, \geq 2}^{18}) = 3608$

However, $\sum_{i \in \{19, \dots, 23\}} x_i \mathcal{K}_{\leq 3, \geq 2}^{18} \subset \mathcal{K}_{\leq 4, \geq 2}^{18}$ and

$$\dim\left(\sum_{i \in \{19, \dots, 23\}} x_i \mathcal{K}_{\leq 3, \geq 2}^{18}\right) = 330.$$

Let $R_k = \mathbb{F}_2[x_{k+1}, \dots, x_n]$.

We denote:

- $\tilde{\mathcal{K}}$ the R_k -module $R_k \mathcal{K}_{\leq D, \geq d}^k$
- $\mu(\tilde{\mathcal{K}})$ the minimal number of generators of $\tilde{\mathcal{K}}$
- N the number of columns of $Mac_{\leq d, m}(\mathcal{F}^*)$

Preprocessing with modules

Let $R_k = \mathbb{F}_2[x_{k+1}, \dots, x_n]$.

We denote:

- $\tilde{\mathcal{K}}$ the R_k -module $R_k \mathcal{K}_{\leq D, \geq d}^k$
- $\mu(\tilde{\mathcal{K}})$ the minimal number of generators of $\tilde{\mathcal{K}}$
- N the number of columns of $Mac_{\leq d, m}(\mathcal{F}^*)$

We have:

- $\mu(\tilde{\mathcal{K}}) \leq N$ for any $1 \leq d < D$ and k .
- $\mu(\tilde{\mathcal{K}}) \geq N$ if (D, d, k) are admissible parameters.

Theorem (V.-Delaplace-Ionica)

Under genericity assumptions, it suffices to consider $\mu(\tilde{\mathcal{K}}) + (n - k)$ elements of $\tilde{\mathcal{K}}$ to get $\mu(\tilde{\mathcal{K}})$ independent polynomials for each specialization of the last $n - k$ variables.

Theorem (V.-Delaplace-Ionica)

Under genericity assumptions, it suffices to consider $\mu(\tilde{\mathcal{K}}) + (n - k)$ elements of $\tilde{\mathcal{K}}$ to get $\mu(\tilde{\mathcal{K}})$ independent polynomials for each specialization of the last $n - k$ variables.

Idea of the proof

We generate $\mu(\tilde{\mathcal{K}}) + n - k$ polynomials.

We consider these polynomials as linear combinations with coefficients in R_k in terms of a minimal set of generators of $\tilde{\mathcal{K}}$ and obtain a polynomial matrix with coefficients in R_k .

Theorem (V.-Delaplace-Ionica)

Under genericity assumptions, it suffices to consider $\mu(\tilde{\mathcal{K}}) + (n - k)$ elements of $\tilde{\mathcal{K}}$ to get $\mu(\tilde{\mathcal{K}})$ independent polynomials for each specialization of the last $n - k$ variables.

Idea of the proof

We generate $\mu(\tilde{\mathcal{K}}) + n - k$ polynomials.

We consider these polynomials as linear combinations with coefficients in R_k in terms of a minimal set of generators of $\tilde{\mathcal{K}}$ and obtain a polynomial matrix with coefficients in R_k .

Under genericity assumptions, the determinantal ideal given by the minors of this matrix has dimension 0 (theorem of Faugère, Safey and Spaenlehauer)

This implies that we get $\mu(\tilde{\mathcal{K}})$ independent polynomials after specialization.

Matrix size for MQOM

	Level	I	III	V
	m	160	240	320
	n	160	240	320
	h	3	4	3
	k	31	44	56
	D	13	20	27
Original preprocessing	$\log_2(\#Rows(Mac_{\leq D, \geq 1, m}^k(\mathcal{F})))$	61.91	96.45	131.08
	$\log_2(\#Cols(Mac_{\leq D, \geq 1, m}^k(\mathcal{F})))$	61.69	95.52	129.68
	$\log_2(r)$	60.34	95.50	130.41
Improved preprocessing	$\log_2(\#Rows(Mac_{\leq D, \geq 1, m}^k(\mathcal{F})))$	61.27	95.15	129.03
	$\log_2(\#Cols(Mac_{\leq D, \geq 1, m}^k(\mathcal{F})))$	61.66	95.3	129.1
	$\log_2(r)$	7.3	7.89	8.31
	% of rows removed	35.73	59.3	75.7
	% of cols removed	2.2	14.16	33.19

Complexity estimation for MQOM

Level	I	III	V
m	160	240	320
n	160	240	320
h	3	4	3
k	31	44	56
D	13	20	27
Time (with original preprocessing)	177.64	276.34	382.05
Time (with improved preprocessing)	151.6	244.12	337.18

Check our preprint

- 1 More details on the results.
- 2 Proofs of the results showed in this presentation
- 3 An analysis of the complexity