Security study of a MQ-Commitment

Julia Sauvage

Sorbonne Université, CNRS, LIP6

March 31, 2025

Goal of our Work

MQ commitment

- Post-quantum commitment.
- possibility to produce zero knowledge proof on the message with MPC-in-the-head methods.
- Security relies on the problem of solving multivariate quadratic polynomials (MQ problem).

$\Rightarrow$ Zero knowledge proof better than $\mathit{Commit}(\mu, r) = \mathrm{SHA256}(\mu \| r)$.

My work

- Cryptanalysis of this commitment.
- Study of specifics instances of the MQ problem.
- Finding optimal parameters.

MQ Problem

- Let F a random quadratic map from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$, i.e. $F = f_1, \ldots, f_m$ in n variables $\mathbf{x} = x_1, \ldots, x_n$ in $\mathbb{F}_q^n$ with f_i quadratic polynomials.

- The quadratic polynomials are denoted by:

$$f_i(\mathbf{x}) = \mathbf{x}^T A_i \mathbf{x} + \mathbf{b_i}^T \mathbf{x} + c_i$$

  with $A_i \in \mathbb{F}_q^{n \times n}, \mathbf{b_i} \in \mathbb{F}_q^n$ and $c_i \in \mathbb{F}_q$.

- **MQ problem:** Find a $\mathbf{x} \in \mathbb{F}_q^n$ such that $F(\mathbf{x}) = 0$

Our Commitment

Let $\mathbb{F}_q$ be a finite field and k, n and m positive integers. The $(q, k, n, m)-$MQ commitment is defined as follows:

- **Setup:** Sample two random quadratic maps F (resp. G) from $\mathbb{F}_q^n$ to $\mathbb{F}_q^k$ (resp. $\mathbb{F}_q^m$).
  Public parameters $\rightarrow (q, k, n, m, F, G)$.
- **Commit:** Given a message $\mu \in \mathbb{F}_q^k$, the commit is $c \rightarrow (\mu + F(\mathbf{r}), G(\mathbf{r}))$ with $\mathbf{r} \xleftarrow{\$} \mathbb{F}_q^n$.
- **Verification:** We recompute the commitment.

Parameters examples

$q = 256, k = 246, n = 115, m = 32$

MQ Commitment - Security Properties

Commitment

$Commit(\mu, \mathbf{r}) = (\mu + F(\mathbf{r}), G(\mathbf{r}))$

Properties of our Commitment Scheme

- **Hiding:** Let μ and μ' two messages chosen by the adversary and $c = (c_1, c_2)$ the commitment of one of these messages.
  The adversary needs to find if c is the commitment for μ or μ'.
- **Binding:** The adversary needs to find a commitment c and two messages μ and μ' such that c is a valid commitment for μ and μ'.

Statistically binding: $\rightarrow$ Very low probability ($2^{-\lambda}$) of the existence of a collision.

Computationally binding: $\rightarrow$ Finding a collision is hard (2^{λ} operations) With λ the level of security

Computationally Hiding

$c = (c_1, c_2)$ commitment of μ or μ' ?

Best known attack

- We try to find $\mathbf{r}$ such that:

$$\mu + F(\mathbf{r}) - c_1 = 0 \text{ and } G(\mathbf{r}) - c_2 = 0$$

  If we find a solution, μ is the message, else it is μ'.
- MQ problem with a random quadratic map from $\mathbb{F}_q^n$ to $\mathbb{F}_q^{k+m}$.
- Well studied complexity.

- Formal proof with important security loss.
- Parameter with $q = 256$ for 128 bits of security:

Hiding	Provable		Heuristic	
	n	m	n	m
	912	1872	115	278

Breaking the Binding Property

Finding a collision on the commitement

- If we have $\mathbf{x} \in \mathbb{F}_q^n$ and Δ with $\Delta \neq 0$ such that:

$$G(\mathbf{x}) = G(\mathbf{x} + \Delta) \tag{1}$$

- Let's be $\mu \in \mathbb{F}_q^k$ and $\mu' \leftarrow \mu + F(\mathbf{x}) - F(\mathbf{x} + \boldsymbol{\Delta})$ and we have:

$$\begin{aligned} Commit(\mu', \mathbf{x} + \Delta) &= (\mu' + F(\mathbf{x} + \Delta), G(\mathbf{x} + \boldsymbol{\Delta})) \\ &= (\mu + F(\mathbf{x}) - F(\mathbf{x} + \boldsymbol{\Delta}) + F(\mathbf{x} + \boldsymbol{\Delta}), G(\mathbf{x})) \\ &= (\mu + F(\mathbf{r}), G(\mathbf{r})) \\ &= Commit(\mu, \mathbf{x}) \end{aligned}$$

- Breaking the binding property is equivalent to finding a solution for (1)

Statistically Binding

Injective quadratic map

G quadratic map from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$.
If $m \ggg n$, then we expect G to be **injective** with high probability;
In this case $G(x) = G(x + \Delta)$ implies that $\Delta = 0$.
$\rightarrow$ Statistically binding.
Example: For $\mathbb{K} = 257$ and a security level of 128 bits we need $\rightarrow m > 2 * n + 16$

Case $m > 2n$:

- **Statistically binding**.
- If $m \leq 2n$, G is not injective.
  $\rightarrow$ Goal: obtaining smaller commitments with **computational binding**.

Computational Binding

Finding a collision on G

- Studied cases : $m \leq 2n$.
- We have to solve **structured** polynomials.

First study of the structure of our system

We want $\mathbf{x}$ and Δ such that $\Delta \neq 0$ and for $1 \leq i \leq m$:

$$g_i(\mathbf{x} + \Delta) - g_i(\mathbf{x}) = 0$$

With $g_i(\mathbf{x}) = \mathbf{x}^T A_i \mathbf{x} + \mathbf{b_i}^T \mathbf{x} + c_i$

$$\begin{aligned} g_i(\mathbf{x} + \Delta) - g_i(\mathbf{x}) &= (\mathbf{x} + \Delta)^T A_i (\mathbf{x} + \Delta) - \mathbf{x}^T A_i \mathbf{x} + \mathbf{b_i}^T \Delta \\ &= \Delta^T A_i \mathbf{x} + \mathbf{x}^T A_i \Delta + \Delta^T A \Delta + \mathbf{b_i}^T \Delta \end{aligned}$$

$\rightarrow$ **Linear in x**

Computationally Binding

Finding a collision on G- Easy case: $m \leq n$

- We choose random values for the entire Δ.
- We have now a random **linear** system of m equations in n variables.
- If $m \leq n$, this linear system will have a solution with great probability.

If $m \leq n$

- We just have a linear system to solve.
- $\rightarrow m^3$ operations.
- The problem is easy.
- **Unusable parameters.**

Computationally Binding - Naive Algorithm

Studied case

- $n \leq m \leq 2n$
- We want to solve $G(\mathbf{x} + \Delta) = G(\mathbf{x})$ with $\Delta \neq 0$.

Naive algorithm

1. We set the n variables of Δ to random values.
2. We have m random linear equations in n variables.
   $\rightarrow$ This system has a solution with probability $q^{-(m-n)}$
3. We try to solve this system

We have to repeat this in average $q^{(m-n)}$ to find a solution.
$\rightarrow q^{(m-n)} n^3$ operations in average.

Computationally Binding - Algebraic methods

Studied case

- $n \leq m \leq 2n$
- We want to solve $G(\mathbf{x} + \Delta) = G(\mathbf{x})$ with $\Delta \neq 0$.

Reduction to a bilinear system

We want for $1 \leq i \leq n$:

$$g_i(\mathbf{x} + \Delta) - g_i(\mathbf{x}) = 0$$

And so:

$$\begin{aligned} g_i(\mathbf{x} + \Delta) - g_i(\mathbf{x}) &= (\mathbf{x} + \Delta)^T A_i (\mathbf{x} + \Delta) - \mathbf{x}^T A_i \mathbf{x} + \Delta^T \mathbf{b_i} \\ &= \Delta^T A_i \mathbf{x} + \mathbf{x}^T A_i \Delta + \Delta^T A_i \Delta + \Delta^T \mathbf{b_i} \\ &= (\Delta + 2\mathbf{x})^T A_i \Delta + \Delta^T \mathbf{b_i} \end{aligned}$$

Bilinear system !

Only if A is a symmetric matrix $\rightarrow q \neq 2^k$.

Reduction to Bilinear systems

Bilinear Systems

m **bilinear** polynomials $F = (f_1, \ldots, f_m)$ in $n_x + n_y$ variables $\mathbf{x} = x_1, \ldots, x_{n_x}$ and $\mathbf{y} = y_1, \ldots, y_{n_y}$
with $f_i(\mathbf{x}) = \mathbf{x}^T A_i \mathbf{y} + \mathbf{b_i^T x} + \mathbf{c_i^T y} + e_i$

Reduction to a bilinear system

We have:

$$g_i(\mathbf{x} + \Delta) - g_i(\mathbf{x}) = (\Delta + 2\mathbf{x})^T A \Delta + \Delta^T \mathbf{b_i}$$

Let $\mathbf{y} = 2\mathbf{x} + \Delta$ and $\Delta_0 = 1$:

$$g_i(\mathbf{x} + \Delta) - g_i(\mathbf{x}) = \mathbf{y}^T A_{i,\{1,n\}} \Delta_{1,n} + \mathbf{b_{i,\{1,n\}}} \Delta_{1,n}^T + A_{i,0} \mathbf{y}^T + b_{i,0}$$

Solving Bilinear Systems

Bilinear Systems

m **bilinear** polynomials $F = (f_1, \ldots, f_m)$ in $n_x + n_y$ variables $\mathbf{x} = x_1, \ldots, x_{n_x}$ and $\mathbf{y} = y_1, \ldots, y_{n_y}$
with $f_i(\mathbf{x}) = \mathbf{x}^T A_i \mathbf{y} + \mathbf{b_i^T x} + \mathbf{c_i^T y} + e_i$

$A_i, \mathbf{b_i}, \mathbf{c_i}$ and e_i are uniformly random on $\mathbb{F}_q$.

- $n_x + n_y = m$: Known complexity [Faugère et al., 2011].
- $n_x + n_y \leq m$: Open problem
  Intuition : We have a lower bound on the complexity with given n_x, n_y and m.

New Algorithm for Finding a Collision

Studied case

- $n \leq m \leq 2n$
- We want to solve $G(\mathbf{x} + \Delta) = G(\mathbf{x})$ with $\Delta \neq 0$.

Algebraic algorithm

1. We set the $2n - m$ variables of Δ to random values.
2. We have m random bilinear equations in m variables.
3. We try to solve this system with algebraic algorithm.
   $\rightarrow$ This system has a solution with great probability.

Known complexity !

Hybrid Method

Let F be a quadratic map from $\mathbb{F}_q^m$ to $\mathbb{F}_q^m$
$F = (f_1, \ldots, f_m)$ in m variables.

Hybrid method [Bettale et al., 2012]

1. We set k variables to random values.
2. We have m quadratics equations in $m - k$ variables.
3. We try to solve this system with algebraic algorithm.
   $\rightarrow$ This system has a solution with probability q^{-k}.

We have to repeat this operation q^k times in average.

Hybrid method for our case

Hybrid algorithm

1. We set the $2n - m + k$ variables of Δ to random values.
2. We have m random bilinear equations in $m - k$ variables.
3. We try to solve this system with algebraic algorithm.
   $\rightarrow$ This system has a solution with probability q^{-k}.

We have to repeat this operation q^k times in average.

Claim: Lower bound on the complexity.

Macaulay matrix

$$f_0(\mathbf{x}) = x_0^2 + 100x_0x_1 - 11x_1^2 - 121x_0x_2 + 23x_1x_2 - 104x_2^2 + 101x_0 - 22x_1 + 101x_0 - 36$$
$$f_1(\mathbf{x}) = x_0x_1 - 13x_1^2 - 38x_0x_2 - 19x_1x_2 + 19x_2^2 - 86x_0 + 33x_1 - 24x_0 - 45$$

Macaulay matrix $d = 2$

$$\begin{array}{c} \\ f_0 \\ f_1 \end{array} \begin{array}{c} \begin{matrix} x_0^2 & x_0x_1 & x_1^2 & x_0x_2 & x_1x_2 & x_2^2 & x_0 & x_1 & x_0 & 1 \end{matrix} \\ \begin{pmatrix} 1 & 100 & -11 & -121 & 23 & -104 & 101 & -22 & 101 & -36 \\ 0 & 1 & -13 & -38 & -19 & 19 & -86 & 33 & -24 & -45 \end{pmatrix} \end{array}$$

Macaulay matrix

$$f_0(\mathbf{x}) = x_0^2 + 100x_0x_1 - 11x_1^2 - 121x_0x_2 + 23x_1x_2 - 104x_2^2 + 101x_0 - 22x_1 + 7x_2 - 36$$
$$f_1(\mathbf{x}) = x_0x_1 - 13x_1^2 - 38x_0x_2 - 19x_1x_2 + 19x_2^2 - 86x_0 + 33x_1 - 24x_2 - 45$$

Macaulay matrix $d = 3$

$$\begin{array}{c} \\ f_0 \\ f_1 \\ x_0f_0 \\ x_1f_0 \\ \vdots \\ x_2f_1 \end{array}
\begin{array}{c} \begin{array}{cccccccccccccc} x_0^3 & x_0^2x_1 & \cdots & x_2^3 & x_0^2 & x_0x_1 & x_1^2 & x_0x_2 & x_1x_2 & x_2^2 & x_0 & x_1 & x_2 & 1 \end{array} \\
\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 100 & -11 & -121 & 23 & -104 & 101 & -22 & 7 & -36 \\ 0 & 0 & \cdots & 0 & 0 & 1 & -13 & -38 & -19 & 19 & -86 & 33 & -24 & -45 \\ 1 & 100 & \cdots & 0 & 101 & 33 & 0 & 7 & 0 & 0 & -36 & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 101 & 33 & 0 & 7 & 0 & 0 & -36 & 0 & 0 \\ & & \cdots & & & & & & & & & & & \\ 0 & 0 & \cdots & 19 & 0 & 0 & 0 & -86 & 33 & -24 & 0 & 0 & -45 & 0 \end{pmatrix} \end{array}$$

XL algorithm

Requirement

System with one or zero solution.

$\Rightarrow$ Square or overdetermined system.

Algorithm

1. We compute the Macaulay matrix of degree i for $i \in \mathbb{N}$
2. Until **full rank** (as many linearly independant rows as columns)
3. Search **a** solution to the linear system (Block-Wiedemann)

If quadratic system has:

- 1 solution: We found the only solution
- 0 solution: linear system $\rightarrow$ no solution

Goal : Knowing the degree denote d for given parameters

$\Rightarrow$ Known complexity

XL algorithm

Problem: Linear dependencies

- # lines of Macaulay matrix $\rightarrow$ known
- **but** linear dependencies.
- Example $f_0 = x_0^2 + x_2$ and $f_1 = x_1 x_2 + 1$

$$x_1 x_2 f_0 + f_0 - x_0^2 f_1 - x_2 f_1 = f_1 f_0 - f_0 f_1 = 0$$

$\Rightarrow$ Linear dependence in the degree 4 Macaulay matrix.

Random systems

- F5 criterium.
- We known exaclty how many linearly interdependent rows we have at any degree.
- d smallest degree.
- Known complexity.

Macaulay matrix on bilinear systems

$$f_0(\mathbf{x}) = x_0x_1 - 121x_0y_0 + 23x_1y_0 + 101x_0 - 22x_1 + 7y_0 - 36$$
$$f_1(\mathbf{x}) = x_0y_0 - 19x_1y_0 - 86x_0 + 33x_1 - 24y_0 - 45$$

Macaulay matrix $d = 3$

	x_0^3	$x_0^2x_1$	$\cdots$	y_0^3	x_0^2	x_0x_1	x_1^2	x_0y_0	x_1y_0	y_0^2	x_0	x_1	y_0	1
f_0	0	0	$\cdots$	0	0	1	0	−121	23	0	101	−22	7	−36
f_1	0	0	$\cdots$	0	0	0	0	1	−19	0	−86	33	−24	−45
x_0f_0	0	1	$\cdots$	0	101	0	0	7	0	0	−36	0	0	0
x_1f_0	0	0	$\cdots$	0	0	101	−22	7	0	0	0	−36	0	0
$\vdots$	$\ddots$		$\cdots$											
y_0f_1	0	0	$\cdots$	0	0	0	0	−86	33	−24	0	0	−45	0

No x_i^d and y_i^d monomials.

XL on bilinear systems

Square bilinear systems

- Less monomials and more linear dependencies than random systems.
- Specific criterium for **square** random bilinear systems [Faugère et al., 2011].
- known d.

Overdetemined bilinear systems

- **More** linear dependencies than square systems.
- Criterium from [Faugère et al., 2011] don't get them all.
- Intuition: we have **less** linearly independent rows than expected.

$\Rightarrow$ Expected d is smaller than real degree.

$\Rightarrow$ Lower bound on the complexity.

Optimal parameters

Goal : small m (optimal commitment size), $q \sim 257$

Studied case

- $n \leq m \leq 2n$
- We want to solve $G(\mathbf{x} + \Delta) = G(\mathbf{x})$ with $\Delta \neq 0$.

Naive algorithm

$$q^{(m-n)} n^3$$

Optimal in our case.

XL algorithm

$$\binom{m+2}{2} \binom{2m-n}{m-n}^2$$

Optimal when exhaustive search on $\mathbb{F}_q$ is too costly.

Hybrid XL algorithm

$$q^k \binom{m-k+2}{2} \binom{m-k+d}{d}^2$$

d is a lower bound.
We choose k to be optimal.

Summary and Work in Progress

Summary for binding security study

	$m \leq n$	$n \leq m \leq 2n$	$m \geq 2n$
Binding security	No	Computational	Statistical
Time complexity	m^3	$q^{m-n}n^3$	

With $2 \leq \omega \leq 3$

Work in progress

- Proof for our assumption.
- Study the possible application of the Hybrid method on bilinear systems.

Summary and Work in Progress

Summary for binding security study

	$m \leq n$	$n \leq m \leq 2n$	$m \geq 2n$
Binding security	No	Computational	Statistical
Time complexity	m^3	$q^{m-n}n^3$	

With $2 \leq \omega \leq 3$

Work in progress

- Proof for our assumption.
- Study the possible application of the Hybrid method on bilinear systems.

Thank you for your attention !