Improving Support-Minors rank attacks: applications to GeMSS and Rainbow

Pierre Briaud, joint work with J. Baena, D. Cabarcas, R. Perlner, D. Smith-Tone and J. Verbel

Crypto 2022

Inria Paris & Sorbonne Université

NIST Round 3 Multivariate signatures

Public key. Map $\mathcal{P} = (p_1, \ldots, p_m)$ hard to invert.
Secret key. Central map $\mathcal{F} = (f_1, \ldots, f_{m'})$ easy to invert, linear maps $\mathcal{T}$, $\mathcal{U}$ s.t.

$$\mathcal{P} = \mathcal{T} \circ \mathcal{F} \circ \mathcal{U}.$$

- **GeMSS** (alternate) : HFE-like.
- **Rainbow** (finalist) : UOV-like.

Rank attacks

Key-recovery by solving MinRank.

Input. $d \in \mathbb{N}$ and $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_K \in \mathbb{F}_q^{n_r \times n_c}$.

Output. $x_1, x_2, \ldots, x_K \in \mathbb{L} \supseteq \mathbb{F}_q$ not all zero s.t.

$$\operatorname{rank}\left(\sum_{i=1}^{K} x_i \boldsymbol{M}_i\right) \leq d$$

Recent work on NIST candidates. "Weaker" instances:

- Rectangular MinRank attack on Rainbow. [Beu21]
- MinRank on HFE variants. [TPD21]

[Beu21] Beullens. "Improved Cryptanalysis of UOV and Rainbow". *Advances in Cryptology – EUROCRYPT 2021*.

[TPD21] Tao, Petzoldt, and Ding. "Efficient Key Recovery for All HFE Signature Variants". *Advances in Cryptology – CRYPTO 2021*.

Support-Minors ("SM") [Bar+20]

New algebraic modeling: unknowns $x_i \in \mathbb{L}$, $\boldsymbol{D} \in \mathbb{L}^{n_r \times d}$, $\boldsymbol{C} \in \mathbb{L}^{d \times n_c}$,

$$\boldsymbol{M} := \sum_{i=1}^{K} x_i \boldsymbol{M}_i = \boldsymbol{D}\boldsymbol{C}.$$

For $1 \leq j \leq n_r$, $\boldsymbol{r}_j := \boldsymbol{M}_{j,*}$ j-th row of $\boldsymbol{M}$. The matrix $\boldsymbol{C}_j := \begin{bmatrix} \boldsymbol{r}_j \\ \boldsymbol{C} \end{bmatrix}$ is of rank $\leq d$.

Maximal minors vanish

$$\mathcal{Q} := \left\{ f = 0 \;\middle|\; f \in \text{MaxMinors}(\boldsymbol{C}_j),\ 1 \leq j \leq n_r \right\}.$$

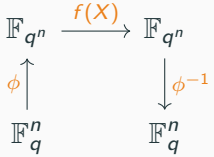
How to solve $\mathcal{Q}$?

$\approx$ Linearize (or dedicated XL) $\rightarrow$ kernel of matrix $\boldsymbol{M}(\mathcal{Q})$.

SM attack on HFE variants

HFE central map

Big-field: $\phi : \mathbb{F}_q^n \to \mathbb{F}_{q^n}$ isomorphism, $f \in \mathbb{F}_{q^n}[X]$ of degree D, $\mathcal{F} = \phi^{-1} \circ f \circ \phi$.



HFE variant

GeMSS uses HFEv- : HFE with modifiers $a \geq 0$, $v \geq 0$.

MinRank of [TPD21]

Let $d := \left\lceil \log_q(D) \right\rceil$. Matrices $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_{n+v} \in \mathbb{F}_q^{(n-a)\times(n+v)}$ built from public key.

"Weaker" MinRank on HFEv-

$$\exists \boldsymbol{u} \in \mathbb{F}_{q^n}^{n+v},\ \boldsymbol{u} \neq \boldsymbol{0},\ \text{rank}\left(\sum_{i=1}^{n+v} u_i \boldsymbol{M}_i\right) \leq d.$$

Matrices over $\mathbb{F}_q$, sols over $\mathbb{F}_{q^n}$. Frobenius: $\forall j,\ \text{rank}\left(\sum_{i=1}^{n+v} u_i^{q^j} \boldsymbol{M}_i\right) \leq d.$

Fix $u_1 = 1$? Still, $\#\text{sols} = n \gg 1$! Can't apply XL !

Step 0: SM equations

Rank d matrix $\sum_{i=1}^{n+v} u_i \boldsymbol{M}_i^{\mathsf{T}} := \boldsymbol{DC}$ restricted to $m' \leq n+v$ cols.

SM is bilinear [Bar+20]

$$\mathcal{M} := \{u_i c_T, 1 \leq i \leq n+v, T, \#T = d\},\ c_T := |\boldsymbol{C}|_{*,T} \text{ (maximal minors)}.$$

For GeMSS we have $\#\mathcal{Q} > \#\mathcal{M}$: $\Rightarrow \dim_{\mathbb{F}_q}(\langle \mathcal{Q} \rangle) < \#\mathcal{Q}$.

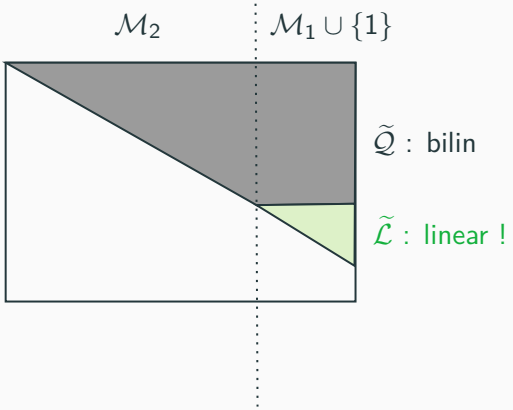
As $\#\text{sols} = n$, we assume

$$\dim_{\mathbb{F}_q}(\langle \mathcal{Q} \rangle) = \#\mathcal{M} - n.$$

Step 1: Linear combinations

We fix $u_1 = 1$ and $c_{\{1..d\}} = 1$. $\mathcal{M}$ becomes $\{1\} \cup \mathcal{M}_1 \cup \mathcal{M}_2$ (affine bilinear).

We have $\dim_{\mathbb{F}_q}(\langle \mathcal{Q} \rangle) > \#\mathcal{M}_2$: linear polys from ech. form !



Step 2: Gröbner Basis (but "easy")

$\widetilde{\mathcal{L}}$ large enough to kill all c_T variables !

- Plug $\widetilde{\mathcal{L}}$ into $\widetilde{\mathcal{Q}} \rightarrow$ new system $\mathcal{S}$, quadratic in u_i's, very overdefined.

Gröbner Bases on $\mathcal{S}$ is easy

GB found by simple Gaussian elimination, *i.e.* in deg. 2, as long as $\binom{m'}{d} \geq n$.

We can make sure that $\binom{m'}{d} \geq n$ for GeMSS.

Impact on GeMSS

Step 1 (find $\widetilde{\mathcal{L}}$) + Step 2 (GB of $\mathcal{S}$). Step 1 is dominant.

Scheme	Minors [TPD21]	SM (conj.) [TPD21]	**Improved SM**
GeMSS128	139	118	**72**
BlueGeMSS128	119	99	**65**
RedGeMSS128	86	72	**49**
GeMSS192	154	120	**75**
BlueGeMSS192	132	101	**67**
RedGeMSS192	95	75	**51**
GeMSS256	166	121	**75**
BlueGeMSS256	141	103	**68**
RedGeMSS256	101	76	**62**

Memory access costs for SM attacks

Does [Beu21] really break Rainbow ?

- Yes . . . in RAM model !
- In 2D nearest neighbor model ?
  - No . . . assuming **random memory access patterns**: Rainbow Response [The20].

But what if they are **not random** ?

[The20] The Rainbow Team. *Response to Recent Paper by Ward Beullens.*

2D nearest neighbor cost model

From NTRUPrime submission, then reused by Rainbow team in [The20].

$\rightarrow$ Memory locations evenly distributed over 2D surface.

Memory overhead:

- $\propto$ **total distance covered by bits of data**.
- $+$ **distance traveled by memory addresses**.

Wiedemann on SM matrix $M(\mathcal{Q})$

Matrix-vector products $\boldsymbol{M}(\mathcal{Q}) \cdot \boldsymbol{v}$:

- large but sparse $\boldsymbol{M}(\mathcal{Q})$: entries cheaply generated on the fly, not stored.
- vector $\boldsymbol{v}$ stored in memory, main overhead is to store and access coefs.

Compute $\boldsymbol{M}(\mathcal{Q}) \cdot \boldsymbol{v}$ via several "row $\times$ $\boldsymbol{v}$".

Naïve approach. For each "row $\times$ $\boldsymbol{v}$" (sum of $\mathbb{F}_q \cdot \mathbb{F}_q$ terms), **central processor** ...

- finds $\neq 0$ coefs in row.
- sends read request to associated memory locations. Query to memory of size $|\boldsymbol{v}|$!

Our strategy: use SM structure

1. Memory broken into **local partitions** $\Pi = (\pi_i)_i$.
   - Each "row $\times$ $\boldsymbol{v}$" involves small number of π_i in Π.
   - Partial sum $\text{row}_{\pi_i} \times \boldsymbol{v}_{\pi_i}$ computed by **local processor** (short distance).
   - **Central accumulator** computes final "row $\times$ $\boldsymbol{v}$" from $\text{row}_{\pi_i} \times \boldsymbol{v}_{\pi_i}$'s: few partial sums sent long distance.

2. Some rows with same $\neq 0$ positions $\rightarrow$ **same memory access patterns**.
   - Group them in a **batch**.
   - Routing information only sent once per batch.

Under the carpet

- Include $\mathcal{P} = 0$ eqs from [Beu21] in the model (standard way).
- Induced costs, but also neglected costs that we analyze in the paper.

Cost comparison for [Beu21]

Set	RAM model	2D (Naïve)	2D (Our strategy)	Security Target
Rainbow-I	127	152.3	**139.5**	143
Rainbow-III	177	216.2	**201.2**	207
Rainbow-V	226	276.2	**260.9**	272

New [Beu22] breaks Rainbow anyway  ...

→ **Still, analysis applies to other SM-based MinRank attacks.**

[Beu22] Beullens. *Breaking Rainbow Takes a Weekend on a Laptop.*