MinRank Gabidulin encryption scheme on matrix codes

Adrien Vinçotte

In collaboration with Nicolas Aragon, Alain Couvreur, Victor Dyseryn and Philippe Gaborit

McEliece frame

Consists on masking a structured code used for both encryption and decryption.

- **Advantage:** Very small ciphertexts (especially if the code has strong decoding capacity)
- **Drawback:** Structured secret code, very large public key

The McEliece scheme is the only encryption scheme with so small ciphertexts. We provide an alternative to McEliece with better parameters (ciphertext and public key sizes).

Our contribution

Gabidulin codes have strong decoding capacity, which implies small parameters. However, their strong structure makes them easy to characterize.

New masking: Turn a Gabidulin code into a matrix code $\mathcal{C}_{mat}$ with coefficients on the base field $\mathbb{F}_q$, which breaks the $\mathbb{F}_{q^m}$-linearity. After hiding $\mathcal{C}_{mat}$, use a McEliece-like encryption frame adapted to matrix codes.

Performances of our scheme

Scheme	pk	ct
Our scheme	98 kB	65 B
Classic McEliece	261 kB	96 B
ROLLO I	696 B	696 B
KYBER	800 B	768 B
RQC-Block-NH-MS-AG	312 B	1118 B
BIKE	1540 B	1572 B
RQC-NH-MS-AG	422 B	2288 B
RQC	1834 B	3652 B
HQC	2249 B	4481 B

Figure: Comparison of different schemes for 128 bits of security

γ-expansion

Let $\gamma = (\gamma_1, \ldots, \gamma_m) \in \mathcal{B}(\mathbb{F}_{q^m})$.

For every $x \in \mathbb{F}_{q^m}$, there exists an only vector $(x_1, ..., x_m) \in \mathbb{F}_q^m$ such that $x = \sum_{i=1}^m x_i \gamma_i$.

We can define γ-expansion as an application:

$$\Psi_\gamma : x \in \mathbb{F}_{q^m} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{F}_q^m$$

From vectors to matrices

Ψ_γ extends naturally to a vector $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ and turns it into a matrix $\Psi_\gamma(\boldsymbol{x}) \in \mathbb{F}_q^{m\times n}$:

$$\Psi_\gamma : \boldsymbol{x} = (x_1, ..., x_n) \longrightarrow \begin{pmatrix} \\ \Psi_\gamma(x_1) & \cdots & \Psi_\gamma(x_n) \\ \\ \end{pmatrix} \in \mathbb{F}_q^{m\times n}$$

Definition: Rank metric

The *support* of $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ is the the $\mathbb{F}_q$-vector space spanned by its coordinates. The *rank* of $\boldsymbol{x}$ is the dimension of its support.

$$\mathrm{Supp}(\boldsymbol{x}) \stackrel{\text{def}}{=} \langle x_1, ..., x_n \rangle_q$$
$$\|\boldsymbol{x}\| \stackrel{\text{def}}{=} \dim(\langle x_1, ..., x_n \rangle_q) = \mathrm{rank}(\Psi_\gamma(\boldsymbol{x}))$$

Weight of a vector: independent of the basis γ.

For two bases β and γ, if we denote $\boldsymbol{P}$ the transition matrix between β and γ, we get:

$$\Psi_\gamma(\boldsymbol{x}) = \boldsymbol{P}\,\Psi_\beta(\boldsymbol{x})$$

Matrix codes

Definition: Matrix code

A matrix code $\mathcal{C}_{mat}$ is an $\mathbb{F}_q$-subspace of $\mathbb{F}_q^{m \times n}$ endowed with the rank metric.

Matrix codes

Definition: Matrix code

A matrix code $\mathcal{C}_{mat}$ is an $\mathbb{F}_q$-subspace of $\mathbb{F}_q^{m \times n}$ endowed with the rank metric.

Let $\mathcal{C}_{vec}$ be an $\mathbb{F}_{q^m}$-linear vector code of parameters $[n, k]_{q^m}$.
Turn $\mathcal{C}_{vec}$ into a matrix code:

$$\mathcal{C}_{mat} \stackrel{\text{def}}{=} \Psi_\gamma(\mathcal{C}_{vec}) = \{\Psi_\gamma(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{C}_{vec}\}.$$

$\mathcal{C}_{mat}$ is a matrix code of parameters $[m \times n, mk]_q$

- **Size of matrices:** $m \times n$ by definition of Ψ_γ.
- **Dimension:** $\mathcal{C}_{vec}$ is $\mathbb{F}_{q^m}$-linear, then for every $\boldsymbol{x} \in \mathcal{C}_{vec}$ and $\alpha \in \mathbb{F}_{q^m}$, we have $\Psi_\gamma(\alpha \boldsymbol{x}) \in \mathcal{C}_{mat}$. Then $\mathcal{C}_{mat}$ has dimension mk.

Encoding and Decoding in matrix codes

Let $\mathcal{C}_{mat}$ be a $[m \times n, K]_q$ matrix code of basis $(\boldsymbol{M}_1, ..., \boldsymbol{M}_K)$.
To encode $\boldsymbol{x} \in \mathbb{F}_q^K$, sample an matrix $\boldsymbol{E} \in \mathbb{F}_q^{m \times n}$ of rank at most r and compute:

$$\boldsymbol{Y} = \sum_{i=1}^{K} x_i \boldsymbol{M}_i + \boldsymbol{E}$$

Encoding and Decoding in matrix codes

Let $\mathcal{C}_{mat}$ be a $[m \times n, K]_q$ matrix code of basis $(\boldsymbol{M}_1, ..., \boldsymbol{M}_K)$.
To encode $\boldsymbol{x} \in \mathbb{F}_q^K$, sample an matrix $\boldsymbol{E} \in \mathbb{F}_q^{m \times n}$ of rank at most r and compute:

$$\boldsymbol{Y} = \sum_{i=1}^{K} x_i \boldsymbol{M}_i + \boldsymbol{E}$$

The decoding problem is exactly the well-known MinRank problem.

MinRank(q, m, n, K, r) problem

Given as input matrices $\boldsymbol{Y}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K \in \mathbb{F}_q^{m \times n}$, the problem asks to find $x_1, \ldots, x_K \in \mathbb{F}_q$ and $\boldsymbol{E} \in \mathbb{F}_q^{m \times n}$ with $\operatorname{rank} \boldsymbol{E} \leq r$ such that $\boldsymbol{Y} = \sum_{i=1}^{K} x_i \boldsymbol{M}_i + \boldsymbol{E}$.

Folding

Unfold: turns a matrix to a vector.

$$\text{Unfold}: \begin{array}{ccc} \mathbb{F}_q^{m\times n} & \longrightarrow & \mathbb{F}_q^{mn} \\ \begin{pmatrix} v_1^{(1)} & v_1^{(n)} \\ \vdots & \vdots \\ v_m^{(1)} & v_m^{(n)} \end{pmatrix} & \longrightarrow & (v_1^{(1)}, ..., v_m^{(1)}, ..., v_1^{(n)}, ..., v_m^{(n)}) \end{array}$$

Fold: inverse map which turns a vector into a matrix

Vectorial representation of a matrix code

Let $(\boldsymbol{M}_1, ..., \boldsymbol{M}_K)$ a basis of a $[m \times n, K]_q$ matrix code $\mathcal{C}_{mat}$. We can define $\mathcal{C}_{mat}$ with an only generator matrix:

$$\boldsymbol{G} = \begin{pmatrix} \mathsf{Unfold}(\boldsymbol{M}_1) \\ \mathsf{Unfold}(\boldsymbol{M}_2) \\ \vdots \\ \mathsf{Unfold}(\boldsymbol{M}_K) \end{pmatrix} \in \mathbb{F}_q^{K \times mn}.$$

The parity check-matrix $\boldsymbol{H} \in \mathbb{F}_q^{(mn-K) \times mn}$ is the matrix whose lines are orthogonal to the lines of $\boldsymbol{G}$ for canonical scalar product $\mathbb{F}_q^{mn}$. Allows to define the dual code $\mathcal{C}_{mat}^{\perp}$.

q-polynomials

Let $x \in \mathbb{F}_{q^m}$. We define: $x^{[i]} = x^{q^i}$.

Definition: q-polynomial

q-polynomial of q-degree r:

$$P(X) = \sum_{i=0}^{r} p_i X^{[i]} \in \mathbb{F}_{q^m}[X] \quad \text{with } p_r \neq 0$$

We denote q-degree by $\deg_q$.

Gabidulin codes

Definition: Gabidulin code

Let $k, m, n \in \mathbb{N}$, such that $k \leq n \leq m$. Let $\boldsymbol{g} = (g_1, \ldots, g_n) \in \mathbb{F}_{q^m}^n$ a vector of $\mathbb{F}_q$-linearly independent elements of $\mathbb{F}_{q^m}$. The Gabidulin code $\mathcal{G}_{\boldsymbol{g}}(n, k, m)$ is the vector code of parameters $[n, k]_{q^m}$ defined by:

$$\mathcal{G}_{\boldsymbol{g}}(n, k, m) = \left\{ P(\boldsymbol{g}) \middle| \ \deg_q P < k \right\},$$

where $P(\boldsymbol{g}) = (P(g_1), \ldots, P(g_n))$ and P is a q-polynomial.

Decoding capacity $= \left\lfloor \frac{n-k}{2} \right\rfloor$

Our masking: Random Rows and Columns Matrix Code transformation

Let $\mathcal{B} = (\boldsymbol{A}_1, ..., \boldsymbol{A}_K)$ a basis a matrix $\mathcal{C}_{mat}$ of size $m \times n$ and dimension K. How we propose to hide $\mathcal{C}_{mat}$:

- Add ℓ_1 rows and ℓ_2 columns of random coefficients: represented by matrices $\boldsymbol{R}_i$, $\boldsymbol{R}'_i$ and $\boldsymbol{R}''_i$.
- Scrambler matrices: multiply by invertible matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$.

$$\mathcal{B}' = \left(\boldsymbol{P} \left(\begin{array}{c|c} \boldsymbol{A}_1 & \boldsymbol{R}_1 \\ \hline \boldsymbol{R}'_1 & \boldsymbol{R}''_1 \end{array} \right) \boldsymbol{Q}, \dots, \boldsymbol{P} \left(\begin{array}{c|c} \boldsymbol{A}_K & \boldsymbol{R}_K \\ \hline \boldsymbol{R}'_K & \boldsymbol{R}''_K \end{array} \right) \boldsymbol{Q} \right).$$

Trapdoor: relies on MinRank and Code Equivalence problems.

Enhanced Gabidulin matrix code

Let $\mathcal{G}_{\boldsymbol{g}}$ a Gabidulin code $[n, k, r]$ on $\mathbb{F}_{q^m}$, γ a $\mathbb{F}_q$-basis of $\mathbb{F}_{q^m}$.

Enhanced Gabidulin code: matrix code $\Psi_\gamma(\mathcal{G}_{\boldsymbol{g}})$ on which we apply the Random Rows and Columns matrix code transformation.

It follows $\mathcal{E}\mathcal{G}_{\boldsymbol{g}}(n, k, m, \ell_1, \ell_2)$: a matrix code of size $(m + \ell_1) \times (n + \ell_2)$ and dimension km.

Application of the McEliece frame to matrix codes with our masking

We apply the MinRank-McEliece frame to matrix Gabidulin codes, using the RRCMC previously defined.

KeyGen (1^λ):

- Select an $[m,k]_{q^m}$ Gabidulin code $\mathcal{G}$, capable of decoding up to $r = \left\lfloor \frac{m-k}{2} \right\rfloor$ errors.

- Sample a basis $\gamma \xleftarrow{\$} \mathcal{B}(\mathbb{F}_{q^m})$ and compute a basis of the code $\mathcal{C}_{mat} = \Psi_\gamma(\mathcal{G})$.

- Apply the RRCMC transformation to $\Psi_\gamma(\mathcal{G})$, by sampling random matrices $\boldsymbol{R}_i, \boldsymbol{R}'_i, \boldsymbol{R}''_i$, and invertible matrices $\boldsymbol{P}, \boldsymbol{Q}$. Let be $\mathcal{C}'_{mat}$ the resulting matrix code.

- **Return:** $\mathsf{pk} = \mathcal{B}$ a basis of $\mathcal{C}'_{mat}$, $\mathsf{sk} = (\mathcal{G}, \gamma, \boldsymbol{P}, \boldsymbol{Q})$

Figure: EGMC-McEliece encryption scheme: KeyGen

EGMC-McEliece encryption scheme: Encryption

The encryption relies on coding the message μ with the public code $\mathcal{C}_{mat}$.

Takes in input: $\mathsf{pk} = (\boldsymbol{M}_1, ..., \boldsymbol{M}_{km})$ a basis of $\mathcal{C}'_{mat}$, $\mu \in \mathbb{F}_q^{km}$.

Sample uniformly at random a matrix $\boldsymbol{E} \in \mathbb{F}_q^{(m+\ell_1)\times(m+\ell_2)}$ such that $\operatorname{rank} \boldsymbol{E} \leq r$.

Return the ciphertext:

$$\boldsymbol{Y} = \sum_{i=1}^{km} \mu_i \boldsymbol{M}_i + \boldsymbol{E}$$

EGMC-McEliece encryption scheme: Decryption

Compute:

$$\boldsymbol{P}^{-1}\boldsymbol{Y}\boldsymbol{Q}^{-1} = \sum_{i=1}^{km} \mu_i \left(\begin{array}{c|c} \boldsymbol{A}_i & \boldsymbol{R}_i \\ \hline \boldsymbol{R}'_i & \boldsymbol{R}''_i \end{array} \right) + \boldsymbol{P}^{-1}\boldsymbol{E}\boldsymbol{Q}^{-1}$$

Truncate the ℓ_1 last rows and ℓ_2 last columns, be $\boldsymbol{M} \in \mathbb{F}_q^{m \times m}$ the resulting matrix.

The first m coordinates of $\Psi_\gamma^{-1}(\boldsymbol{M}) \in \mathbb{F}_{q^m}^m$ form a noisy codeword of $\mathcal{G}$. Its decoding algorithm allow to retrieve the vector error $\boldsymbol{e}$.

By computing $\Psi_\gamma(\boldsymbol{e})$, we can consider the system $\boldsymbol{Y} = \sum_{i=1}^{km} \mu_i \boldsymbol{M}_i + \boldsymbol{E}$, whose unknowns are the (μ_i) and some coefficients of $\boldsymbol{E}$.

EGMC-Niederreiter encryption scheme

Niederreiter frame: rather than give the generator matrix $\boldsymbol{G}$, we can give a parity check matrix $\boldsymbol{H}$ as public key.

The message to encrypt is also a matrix $\boldsymbol{E}$ of rank less than r, and the ciphertext its associated syndrome, that is shorter than the message.

Advantage: smaller ciphertext size for an equal security.

OW-CPA security

Under the assumption that:

- there exists no PPT algorithm to solve the MinRank problem with non negligible probability
- there exists no PPT distinguisher for the problem which consists on distinguish a valid public key and a random matrix code with non negligible advantage

then the scheme is OW-CPA.

Attacks on the message: solve the MinRank problem

Main attacks on an instance MinRank(q, m, n, K, r):

- Kernel attack: combinatorial attack which consists on sampling vectors, hoping they are in the kernel of $\boldsymbol{E}$, and deducing a linear system of equations. Complexity:
$$O(q^{r\lceil \frac{K}{m} \rceil} K^{\omega})$$
- Support minors attack: rank $\left(\boldsymbol{Y} - \sum_{i=1}^{K} \mu_i \boldsymbol{M}_i\right) \leq r$. All the minors of size more than r are equal to zero. We deduce a system of equations whose unknowns are the (μ_i).

Attacks on the key: Stabilizer algebra

Left Stabilizer algebra

$$\mathrm{Stab}_L(\mathcal{C}_{mat}) \stackrel{\text{def}}{=} \left\{ \boldsymbol{P} \in \mathbb{F}_q^{m \times m} \mid \boldsymbol{P}\mathcal{C}_{mat} \subseteq \mathcal{C}_{mat} \right\}$$

We similarly define the Right Stabilizer algebra.

For every $\mathbb{F}_{q^m}$-linear $\mathcal{C}_{vec} \subseteq \mathbb{F}_{q^m}^n$, the code $\Psi_\gamma(\mathcal{C}_{vec})$ has a non trivial stabilizer algebra:

$$\dim \mathrm{Stab}_L(\Psi_\gamma(\mathcal{C}_{vec})) \geq m$$

Combinatorial distinguisher against the $\mathbb{F}_{q^m}$-linear structure

Non scrambled version of the code, denoted $\mathcal{C}_0$, spanned by the basis:

$$\mathcal{B}_0 = \left(\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{R}_1 \\ \boldsymbol{R}'_1 & \boldsymbol{R}''_1 \end{pmatrix}, \dots, \begin{pmatrix} \boldsymbol{A}_{km} & \boldsymbol{R}_{km} \\ \boldsymbol{R}'_{km} & \boldsymbol{R}''_{km} \end{pmatrix} \right)$$

where $(\boldsymbol{A}_i)_i$ is a $\mathbb{F}_q$-basis of $\Psi_\gamma(\mathcal{G}_{\boldsymbol{g}}(n, k, m))$.

Idea: apply a projection map on both the row and columns spaces of $\mathcal{C}_{pub}$ in order to get rid of the contributions of the matrices $\boldsymbol{R}_i, \boldsymbol{R}'_i$ and $\boldsymbol{R}''_i$.

Choose two matrices:

- $\boldsymbol{U} \in \mathbb{F}_q^{m\times(m+\ell_1)}$ such that $\boldsymbol{U} = (\boldsymbol{U}_0 \mid \boldsymbol{0})$, with $\boldsymbol{U}_0 \in \mathbf{GL}_m(\mathbb{F}_q)$
- $\boldsymbol{V} \in \mathbb{F}_q^{(n+\ell_2)\times n'}$ such that
  $\boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}_0 \\ \boldsymbol{0} \end{pmatrix}$, with $\boldsymbol{V}_0 \in \mathbb{F}_q^{n\times n'}$ of full rank and $k < n' \leq n$

Observation: the code $\boldsymbol{U}\mathcal{C}_0\boldsymbol{V}$ spanned by the $(\boldsymbol{U}_0\boldsymbol{A}_i\boldsymbol{V}_0)_i$.

Consequently: $\boldsymbol{U}\mathcal{C}_0\boldsymbol{V} = \Psi_{\gamma\boldsymbol{U}_0}(\mathcal{G}_{\boldsymbol{g}\boldsymbol{V}_0}(n', k, m))$

Number of choices for $\boldsymbol{U}, \boldsymbol{V}$ is $\approx q^{m^2+nn'}$. Being minimal when $n' = k+1$.

The public code $\mathcal{C}_{pub}$ is spanned by:

$$\mathcal{B}' = \left(\boldsymbol{P} \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{R}_1 \\ \boldsymbol{R}_1' & \boldsymbol{R}_1'' \end{pmatrix} \boldsymbol{Q}, \dots, \boldsymbol{P} \begin{pmatrix} \boldsymbol{A}_{km} & \boldsymbol{R}_{km} \\ \boldsymbol{R}_{km}' & \boldsymbol{R}_{km}'' \end{pmatrix} \boldsymbol{Q} \right) = \boldsymbol{P} \mathcal{B}_0 \boldsymbol{Q}$$

The same reasoning can be made replacing $\boldsymbol{U}$ by $\boldsymbol{U}' \stackrel{\text{def}}{=} \boldsymbol{U}\boldsymbol{P}^{-1}$ and $\boldsymbol{V}$ by $\boldsymbol{V}' \stackrel{\text{def}}{=} \boldsymbol{Q}^{-1}\boldsymbol{V}$.

The number of choices for $\boldsymbol{U}', \boldsymbol{V}'$ is still $\approx q^{m^2+n(k+1)}$.

The distinguisher consists in:

- Guess the pair $\boldsymbol{U}', \boldsymbol{V}'$ with $\boldsymbol{U}' \in \mathbb{F}_q^{m\times(m+\ell_1)}$ and $\boldsymbol{V}' \in \mathbb{F}_q^{(n+\ell_2)\times(k+1)}$,
- Compute the left stabilizer algebra of $\boldsymbol{U}'\mathcal{C}_{pub}\boldsymbol{V}'$,

until get a stabilizer algebra of dimension $\geq m$.

Probability of finding a valid pair $\boldsymbol{U}', \boldsymbol{V}'$ is

$$\mathbb{P} \approx \frac{q^{m^2+n(k+1)}}{q^{m(m+\ell_1)+(n+\ell_2)(k+1)}} = q^{-(m\ell_1+(k+1)\ell_2)}$$

which yields a complexity of $\widetilde{O}(q^{m\ell_1+(k+1)\ell_2})$.

Resulting parameters

Sec.	q	k	m	ℓ_1	ℓ_2	r	Message	Structu.	pk	ct
128	2	17	37	3	3	10	170	165	76 kB	121 B
	2	25	37	3	3	6	150	189	78 kB	84 B
	2	35	43	2	2	4	145	158	98 kB	65 B
	2	47	53	2	2	3	147	202	166 kB	66 B
192	2	51	59	2	2	4	209	222	268 kB	89 B
256	2	23	47	3	3	12	271	284	191 kB	177 B
	2	37	53	3	2	8	290	273	274 kB	139 B
	2	71	79	2	2	4	289	302	667 kB	119 B

Figure: Reference parameters for the EGMC-Niederreiter encryption scheme

Thank you for your attention