Rank Metric Code Based Cryptography

J.-P. Tillich
Inria, team-project Cosmiq

October 2, 2020

1. Code based cryptography

Difficult problem in coding theory

Problem 1. [Decoding]
Input: n, k, t *with* $k < n$, *generator matrix* $\boldsymbol{G} \in \mathbb{F}_q^{k \times n}$ *of the code* $\mathcal{C} \stackrel{\text{def}}{=} \{\boldsymbol{uG} : \boldsymbol{u} \in \mathbb{F}_q^k\}$, $\boldsymbol{y} \in \mathbb{F}_q^n$
Question: $\exists$? $\boldsymbol{e} \in \mathbb{F}_q^n$ *and* $\boldsymbol{u} \in \mathbb{F}_q^k$ *such that*

$$\begin{cases} \underbrace{\boldsymbol{uG}}_{\in \mathcal{C}} + \boldsymbol{e} & = & \boldsymbol{y} \\ |\boldsymbol{e}| & \leqslant & t \end{cases}$$

where $|\boldsymbol{e}|$ = *Hamming weight of* $\boldsymbol{e} = \#\{i \in [\![1, n]\!], e_i \neq 0\}$.

Problem NP-complete

Syndrome decoding

Problem 2. [Decoding]
Input: n, k, t *with* $k < n$, *parity-check matrix* $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$ *of the code* $\mathcal{C} \stackrel{def}{=} \{\boldsymbol{c}\mathbb{F}_q^n : \boldsymbol{H}\boldsymbol{c}^\intercal = 0\}$, $\boldsymbol{s} \in \mathbb{F}_q^{n-k}$
Question: $\exists ?$ $\boldsymbol{e} \in \mathbb{F}_q^n$ *such that* $\begin{cases} \boldsymbol{H}\boldsymbol{e}^\intercal & = & \boldsymbol{s}^\intercal \\ |\boldsymbol{e}| & \leqslant & t \end{cases}$.

equivalent version of the decoding problem:

$$\begin{array}{rcl} \boldsymbol{y} & = & \underbrace{\boldsymbol{c}}_{\in \mathcal{C}} + \boldsymbol{e} \\ \Rightarrow \boldsymbol{s}^\intercal \stackrel{\text{def}}{=} \boldsymbol{H}\boldsymbol{y}^\intercal & = & \boldsymbol{H}\boldsymbol{e}^\intercal \end{array}$$

Rank Metric

Difficult problem in coding theory

Problem 3. [Decoding]
Input: n, k, t *with* $k < n$, *generator matrix* $\boldsymbol{G} \in \mathbb{F}_q^{k \times n}$ *of the code* $\mathcal{C} \stackrel{\text{def}}{=} \{\boldsymbol{u}\boldsymbol{G} : \boldsymbol{u} \in \mathbb{F}_q^k\}$, $\boldsymbol{y} \in \mathbb{F}_q^n$
Question: $\exists$? $\boldsymbol{e} \in \mathbb{F}_q^n$ *and* $\boldsymbol{u} \in \mathbb{F}_q^k$ *such that*

$$\begin{cases} \underbrace{\boldsymbol{u}\boldsymbol{G}}_{\in \mathcal{C}} + \boldsymbol{e} & = & \boldsymbol{y} \\ |\boldsymbol{e}|_R & \leqslant & t \end{cases}$$

where $|\boldsymbol{e}|_R =$ *rank weight of* $\boldsymbol{e}$.

Randomized reduction [Gaborit-Zemor2014] of the previous problem to it.

Rank metric

- $(\beta_1 \dots \beta_m)$ basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$

$$\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{F}_{q^m}^n \rightarrow \mathbf{Mat}(\boldsymbol{x}) = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \in \mathbb{F}_q^{m \times n}$$

where $x_j = \sum_{i=1}^{m} x_{ij} \beta_i$.

- Rank metric = viewing an element of $\mathbb{F}_{q^m}^n$ as an $m \times n$ matrix.

$$|\boldsymbol{x} - \boldsymbol{y}|_r \stackrel{\text{def}}{=} \mathbf{Rank}\left(\mathbf{Mat}(\boldsymbol{x}) - \mathbf{Mat}(\boldsymbol{y})\right).$$

Rank/Hamming/Euclidean metric

Ambient space $\mathbb{F}_q^{n^2}$

	Euclidean metric	Hamming metric	Rank metric
# levels	$O\left(q^2n^2\right)$	n^2+1	$n+1$

A very rigid metric

- Projection in Hamming space, $I \subset \{1, \cdots, n\}$, $|I| = p$

$$\begin{array}{rcl} \pi_I : \mathbb{F}_q^n & \to & \mathbb{F}_q^p \\ \boldsymbol{x} & \mapsto & \boldsymbol{x}_I = (x_i)_{i \in I} \\ \text{typically } |\pi_I(\boldsymbol{x})|_{\mathsf{Ham}} & \approx & \dfrac{p}{n} |\boldsymbol{x}|_{\mathsf{Ham}} \end{array}$$

  Phenomenon used in ISD

- Projection in rank metric, associated to a full-rank matrix $\boldsymbol{P} \in \mathbb{F}_q^{p \times m}$:

$$\begin{array}{rcl} \pi : \mathbb{F}_q^{m \times n} & \to & \mathbb{F}_q^{p \times n} \\ \boldsymbol{M} & \mapsto & \boldsymbol{P}\boldsymbol{M} \\ \text{typically } |\pi(\boldsymbol{M})|_{\mathsf{Rank}} & \approx & |\boldsymbol{M}|_{\mathsf{Rank}} \text{ if } |\boldsymbol{M}|_{\mathsf{Rank}} \leqslant p \end{array}$$

  No weight reduction

MinRank

Problem 4. [MinRank]
Input: m, n, K, t, $\boldsymbol{M}_1, \cdots, \boldsymbol{M}_K, \boldsymbol{Y} \in \mathbb{F}_q^{m \times n}$
Question: $\exists?$ $\boldsymbol{E} \in \mathbb{F}_q^{m \times n}$ *and* $\boldsymbol{u} \in \mathbb{F}_q^K$ *such that*

$$\begin{cases} \underbrace{\sum_{i=1}^{K} u_i \boldsymbol{M}_i}_{\in \mathcal{C} \overset{def}{=} \langle \boldsymbol{M}_1, \cdots, \boldsymbol{M}_K \rangle_{\mathbb{F}_q}} + \boldsymbol{E} = \boldsymbol{Y} \\ \operatorname{rank} |\boldsymbol{E}| \leqslant t \end{cases}$$

Decoding in Hamming metric reduces to solving MinRank.

$$\boldsymbol{Y} = \begin{bmatrix} y_1 & 0 & \dots & 0 \\ 0 & y_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & y_n \end{bmatrix}$$

Decoding $\mathbb{F}_{q^m}$ linear codes reduces to MinRank

Code $\mathcal{C}$ generated by $\boldsymbol{G} = \begin{bmatrix} \boldsymbol{g}_1 \\ \dots \\ \boldsymbol{g}_k \end{bmatrix}$, of dimension k over $\mathbb{F}_{q^m}$:

$$\begin{aligned} \mathcal{C} &= \{u_1\boldsymbol{g}_1 + \cdots + u_k\boldsymbol{g}_k, u_i \in \mathbb{F}_{q^m}\} \\ &= \langle \boldsymbol{g}_1, \cdots, \boldsymbol{g}_k \rangle_{\mathbb{F}_{q^m}} \end{aligned}$$

Corresponding matrix code $\mathcal{C}'$:

$$\begin{aligned} \mathcal{C}' &\overset{\text{def}}{=} \mathbf{Mat}(\mathcal{C}) = \{\mathbf{Mat}(\boldsymbol{c}) : \boldsymbol{c} \in \mathcal{C}\} \\ &= \langle \mathbf{Mat}(\alpha^i \boldsymbol{g}_j) : i \in \{0, \cdots, m-1\},\ j \in \{1, \cdots, k\} \rangle_{\mathbb{F}_q} \end{aligned}$$

$\mathcal{C}'$ matrix code of dimension $K = mn$ over $\mathbb{F}_q$.

decoding $\mathcal{C}$ for the rank metric $\Leftrightarrow$ solving MinRank for $\mathcal{C}'$

The complexity picture

Hamming-Decoding $\leqslant_r$ Rank-Decoding $\leqslant$ MinRank

Hamming-Decoding $\leqslant$ MinRank

Rank-decoding rather than MinRank in code-based cryptography

- public key m times shorter!

	public key	size
rank-dec$[m,n,k,t]$	$\boldsymbol{g}_1,\cdots,\boldsymbol{g}_k \in \mathbb{F}_{q^m}^n$	$kmn\log q$
MinRank$[m,n,k,t]$	$\begin{array}{ccc} \mathbf{Mat}(\boldsymbol{g}_1), & \cdots & \mathbf{Mat}(\alpha^{m-1}\boldsymbol{g}_1) \\ \vdots & \vdots & \vdots \\ \hline \mathbf{Mat}(\boldsymbol{g}_k), & \cdots & \mathbf{Mat}(\alpha^{m-1}\boldsymbol{g}_k) \end{array}$	$km^2n\log q$

- Very similar to quasi-cyclic codes in code-based cryptography

$$\begin{array}{rcl} \text{homomorphism } M:\mathbb{F}_{q^m} & \rightarrow & \mathbb{F}_q^{m\times m} \\ M(\alpha\beta) & = & M(\alpha)M(\beta) \\ \text{for an } \mathbb{F}_{q^m} \text{ linear code } \mathcal{C} & : & \mathbf{Mat}(\mathcal{C}) \text{ is invariant by left. mult. by } M(\mathbb{F}_{q^m}^{\times}) \\ \mathbf{Mat}(\alpha\boldsymbol{c}) & = & M(\alpha)\mathbf{Mat}(\boldsymbol{c}),\ \forall\alpha\in\mathbb{F}_{q^m} \\ M(\alpha)\mathbf{Mat}(\mathcal{C}) & = & \mathbf{Mat}(\mathcal{C}),\ \forall\alpha\in\mathbb{F}_{q^m}^{\times} \end{array}$$

Codes with a decoding algorithm

- Gabidulin codes = rank metric analogues of Reed-Solomon codes
- LRPC codes = structured rank metric analogues of LDPC/MDPC codes

2. LRPC codes

[Gaborit, Murat, Ruatta, Zémor 2013]

Definition 1. *An LRPC code over $\mathbb{F}_{q^m}$ of weight w has a parity-check matrix with entries h_{ij} that span an $\mathbb{F}_q$ space of dimension w.*

$$|\boldsymbol{x}|_r = \dim\langle x_1, \ldots, x_n\rangle_{\mathbb{F}_q}$$

$\Rightarrow$ all rows of $\boldsymbol{H}$ have weight $\leqslant w$.

- Correct t errors when $tw \leqslant n - k$.

LDPC codes

Definition 2. *An LDPC code over* $\mathbb{F}_{q^m}$ *of weight* w *is a code* $\mathcal{C}$ *that admits an* $(n-k) \times n$ *parity-check matrix* $\boldsymbol{H}$ *whose rows have Hamming weight* $\leqslant w$.

The notion of support

Definition 3. [Hamming Support] *The (Hamming) support* $\boldsymbol{Supp}_H(\boldsymbol{x})$ *of a vector* $\boldsymbol{x}$ *is the set of positions* i *where* $x_i \neq 0$*:*

$$\boldsymbol{Supp}(\boldsymbol{x}) \stackrel{def}{=} \{i : x_i \neq 0\}$$

$$\boldsymbol{Supp}(\mathcal{C}) \stackrel{def}{=} \bigcup_{\boldsymbol{c}\in\mathcal{C}} \boldsymbol{Supp}(\boldsymbol{c})$$

Definition 4. [Rank Support] *The column rank support (resp. row rank support)* $\boldsymbol{Supp}(\boldsymbol{X})$*, resp.* $\boldsymbol{Supp}_r(\boldsymbol{X})$*, of a matrix* $\boldsymbol{X} \in \mathbb{F}_q^{m\times n}$ *is the subspace of* $\mathbb{F}_q^m$ *generated by the columns of* $\boldsymbol{X}$*, resp. by the rows of* $\boldsymbol{X}$*.*

$$\boldsymbol{Supp}_c(\boldsymbol{x}) \stackrel{def}{=} \boldsymbol{Supp}_c(\boldsymbol{Mat}(\boldsymbol{x}))$$

$$\boldsymbol{Supp}_c(\mathcal{C}) \stackrel{def}{=} \bigoplus_{\boldsymbol{c}\in\mathcal{C}} \boldsymbol{Supp}_c(\boldsymbol{c})$$

LRPC/LDPC

A parity check matrix $\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}_1 \\ \dots \\ \boldsymbol{h}_{n-k} \end{bmatrix}$ whose entries H_{ij} are all in a subspace V of dimension w

$$\begin{aligned} \mathcal{C}^{\perp} &= \langle \boldsymbol{h}_1, \cdots, \boldsymbol{h}_{n-k} \rangle_{\mathbb{F}_{q^m}} \\ \mathcal{C}' &= \langle \boldsymbol{h}_1, \cdots, \boldsymbol{h}_{n-k} \rangle_{\mathbb{F}_q} \\ \textbf{Supp}(\mathcal{C}') &\subseteq V \end{aligned}$$

$\Rightarrow q^{n-k}$ codewords in $\mathcal{C}^{\perp}$of rank $\leqslant w$

Corresponds to an LDPC code whose dual contains a space of subcode of dimension $n-k$ whose support is of size w.

Decoding algorithm for LRPC codes [Gaborit, Murat, Ruatta, Zémor, 2013]

Definition 5. [product space] *E and F two subspaces of $\mathbb{F}_{q^m}$.*

$$E \cdot F = \langle ef,\ e \in E,\ f \in F \rangle_{\mathbb{F}_q}$$

$$\dim E \cdot F \leqslant \dim E \dim F$$

Decoding algorithm for LRPC codes [Gaborit, Murat, Ruatta, Zémor, 2013] (II)

$$
\begin{array}{rcl}
\boldsymbol{y} & = & \boldsymbol{c} + \boldsymbol{e} \\
\boldsymbol{s}^{\intercal} = \boldsymbol{H}\boldsymbol{y}^{\intercal} & = & \boldsymbol{H}\boldsymbol{e}^{\intercal} \\
e_i & \in & E \\
\dim E & \leqslant & t \\
H_{ij} & \in & F \\
\dim F & \leqslant w & \\
\langle s_1, \cdots s_{n-k} \rangle_{\mathbb{F}_q} & =_{\text{likely if } n-k \geqslant tw!} & E \cdot F \\
A & = & \underbrace{E}_{\text{unknown}} \cdot F
\end{array}
$$

Decoding algorithm for LRPC codes [Gaborit, Murat, Ruatta, Zémor, 2013] (III)

$$
\begin{array}{lcl}
A & = & \underbrace{E}_{\text{unknown}} \cdot F \\
F & = & \langle f_1, \cdots, f_w \rangle_{\mathbb{F}_q} \\
E & \subset & f_i^{-1} A \\
E & =_{\text{likely}} & \bigcap_{i=1}^{w} f_i^{-1} A
\end{array}
$$

$$
\left\{ \begin{array}{lcr} E & = & \textbf{Supp}(\boldsymbol{e}) \\ \boldsymbol{H}\boldsymbol{e}^{\intercal} & = & \boldsymbol{s}^{\intercal} \end{array} \right. \Rightarrow \boldsymbol{e} \text{ by solving a linear system if } nt \leqslant m(n-k)
$$

Cyclicity

$$
\begin{aligned}
\boldsymbol{H} &= \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 \end{bmatrix} \\
\boldsymbol{H}_i &: \text{circulant matrix } p \times p \text{ matrix} \\
\boldsymbol{H}_i &= \begin{bmatrix} h_0 & h_1 & \cdots & h_{p-1} \\ h_{p-1} & h_0 & \cdots & h_{p-2} \\ \vdots & \ldots & \ddots & \vdots \\ h_1 & h_2 & \ldots & h_0 \end{bmatrix} \\
&\equiv h_0 + h_1 X + \cdots + h_{p-1} X^{p-1} \\
\{\text{circulant matrices in } \mathbb{F}_q^{p \times p}\} &\simeq \mathbb{F}_q[X]/(X^p - 1)
\end{aligned}
$$

If the first row has all its entries in V then so do the other rows.

NTRU-MDPC-LRPC

	NTRU	MDPC	LRPC
ambient space E	$\mathbb{Z}_q[X]/(X^p-1)$	$\mathbb{F}_2[X]/(X^p-1)$	$\mathbb{F}_{q^m}[X]/(X^p-1)$
metric	$\lvert\lvert f\rvert\rvert_\infty \stackrel{\text{def}}{=} \sup_i \lvert f_i\rvert$	$\lvert f\rvert_{\mathsf{H}} \stackrel{\text{def}}{=} \#\{i : f_i \neq 0\}$	$\lvert f\rvert_{\mathsf{R}} \stackrel{\text{def}}{=} \dim_{\mathbb{F}_q} < f_i >$
public key	$(1,h) \in E^2$	$(1,h) \in E^2$	$(1,h) \in E^2$
message	$\mu \in E$ $\lvert\mu\rvert_\infty \leqslant t_1$	$\mu \in E$ $\lvert\mu\rvert_{\mathsf{H}} \leqslant t_2$	$\mu \in E$ $\lvert\mu\rvert_{\mathsf{R}} \leqslant t_3$
random	$r \in E$ $\lvert r\rvert_\infty \leqslant t_1$	$r \in E$ $\lvert r\rvert_{\mathsf{H}} \leqslant t_2$	$r \in E$ $\lvert(m,r)\rvert_{\mathsf{R}} \leqslant t_3$
ciphertext	$rh+\mu$	$rh+\mu$	$rh+\mu$
private key	$(f,g) \in E^2$ $\lvert f\rvert_\infty, \lvert g\rvert_\infty \leqslant w_1$	$(f,g) \in E^2$ $\lvert f\rvert_H, \lvert g\rvert_H \leqslant w_2$	$(f,g) \in E^2$ $\lvert f,g\rvert_R \leqslant w_3$
constraint	$\sqrt{pw_1t_1} \leqslant q$	$2t_2w_2 \leqslant p$	$t_3w_3 \leqslant \min(m,p)$
the point	$h = \frac{p'f}{g}$ p' small	$h = \frac{f}{g}$	$h = \frac{f}{g}$

3. The RSL problem

Problem 5. [RSL]
Input: n, k, t, ℓ, *(parity-check) matrix* $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$ $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_\ell \in \mathbb{F}_q^{n-k}$
Promise: $\exists$ *subspace* V *of* $\mathbb{F}_{q^m}$ *and* $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_\ell$ *with* $\mathbf{Supp}_c(\boldsymbol{e}_i) = V$ *and* $\boldsymbol{H}\boldsymbol{e}_i{}^\intercal = \boldsymbol{s}_i{}^\intercal$
Question: *Find* V

Simultaneous decoding problem of ℓ errors sharing the same column support

An authentication scheme

Lyubashevsky's "Fiat-Shamir with aborts"

- Public matrix $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$
- prover has a secret matrix $\boldsymbol{S} \in \mathbb{F}_q^{\ell\times n\times}$ of ℓ small row vectors.
- $\boldsymbol{T} = \boldsymbol{H}\boldsymbol{S}^\intercal$ is public
- Prover wants to prove his knowledge of S (he knows how to decode ℓ instances of the decoding problem)

Prover		Verifier
Chooses small random $\boldsymbol{e} \in \mathbb{F}_q^n$	$\xrightarrow{\boldsymbol{s}^\intercal = \boldsymbol{H}\boldsymbol{e}^\intercal}$	
		chooses a small $\boldsymbol{c} \in \mathbb{F}_q^\ell$
	$\xleftarrow{\boldsymbol{c}}$	
$\boldsymbol{z} = \boldsymbol{e} + \boldsymbol{c}\boldsymbol{S}$		
	$\xrightarrow{\boldsymbol{z}}$	
		checks $\begin{cases} \lVert\boldsymbol{z}\rVert & \leqslant & \lVert\boldsymbol{e}\rVert + \lVert\boldsymbol{c}\boldsymbol{S}\rVert \\ \boldsymbol{H}\boldsymbol{z}^\intercal & = & \boldsymbol{s}^\intercal + \boldsymbol{T}\boldsymbol{c}^\intercal \end{cases}$

Verification

$$\begin{aligned} \boldsymbol{T} &= \boldsymbol{H}\boldsymbol{S}^{\mathsf{T}} \\ \boldsymbol{H}\boldsymbol{z}^{\mathsf{T}} &= \boldsymbol{H}\left(\boldsymbol{e}^{\mathsf{T}} + \boldsymbol{S}^{\mathsf{T}}\boldsymbol{c}^{\mathsf{T}}\right) \\ &= \boldsymbol{s}^{\mathsf{T}} + \boldsymbol{T}\boldsymbol{c}^{\mathsf{T}} \end{aligned}$$

The ideas

- If c and S are small, then cS is small.
- Adding a small random e to a small cS can make it random and "washes" out the information contained on S brought by cS

How to do this for the rank metric ?

- If $\boldsymbol{c}$ and $\boldsymbol{S}$ are small, then $\boldsymbol{cS}$ is small.

$$\begin{aligned}
\boldsymbol{S} &= \begin{bmatrix} \boldsymbol{e}_1 \\ \cdots \\ \boldsymbol{e}_\ell \end{bmatrix} \\
\textbf{Supp}(\boldsymbol{e}_i) &\subseteq E \quad \text{(RSL condition!)} \\
\dim E &= t \\
|\boldsymbol{c}| &= w \\
\textbf{Supp}(\boldsymbol{c}) &= F \\
\textbf{Supp}(\boldsymbol{cS}) &\subseteq E \cdot F \\
|\boldsymbol{cS}| &\leqslant tw
\end{aligned}$$

Does not completely work like this...

$$
\begin{aligned}
\boldsymbol{c} &= (c_1, \cdots, c_\ell) \\
\textbf{Supp}(\boldsymbol{c}) &= \langle f_1, \cdots, f_w \rangle = F \\
\textbf{Supp}(\boldsymbol{cS}) &\subseteq E \cdot F \\
E \cdot F \subset E \cdot F + \textbf{Supp}(\boldsymbol{e}) &=_{typically!} \textbf{Supp}(\boldsymbol{z}) = \textbf{Supp}(\boldsymbol{e} + \boldsymbol{cS}) \\
\Rightarrow E &=_{\text{typically!}} \bigcap_{i=1}^{w} f_i^{-1} \textbf{Supp}(\boldsymbol{z})
\end{aligned}
$$

Durandal, IBE both based on RSL

- IBE scheme [Gaborit-Hauteville-Phan-Tillich/CRYPTO 2017] in rank metric based on RSL
- signature scheme Durandal [Aragon-Blazy-Gaborit-Hauteville-Ruatta-Zemor/EUROCR 2019] based on RSL

4. Complexity of the best known algorithms

Solving the decoding problem Dec$[m, n, k, t]$

- Algebraic attacks (MinRank)
- Combinatorial attacks $\tilde{O}\left(q^{t(k+1)-m}\right)$ when $m = n$.

Decoding $\Leftrightarrow$ finding a low weight codeword

$$
\begin{aligned}
\boldsymbol{y} &= \underbrace{\boldsymbol{c}}_{\in \mathcal{C}} + \boldsymbol{e}, \quad |\boldsymbol{e}| = t \\
\mathcal{C}' &\stackrel{\text{def}}{=} \mathcal{C} + \langle \boldsymbol{y} \rangle_{\mathbb{F}_{q^m}} \\
\boldsymbol{e} \in \mathcal{C}' &\Rightarrow d_{\min}(\mathcal{C}') \leqslant t
\end{aligned}
$$

Decoding t errors in $\mathcal{C} \leftrightarrow$ finding a codeword of weight t in $\mathcal{C}'$.

RSL $\Leftrightarrow$ finding a subcode of small support

$$
\begin{array}{rcl}
\boldsymbol{y}_1 & = & \underbrace{\boldsymbol{c}_1}_{\in \mathcal{C}} + \boldsymbol{e}_1 \\
\boldsymbol{y}_2 & = & \underbrace{\boldsymbol{c}_2}_{\in \mathcal{C}} + \boldsymbol{e}_2 \\
\cdots & = & \cdots \\
\boldsymbol{y}_\ell & = & \underbrace{\boldsymbol{c}_\ell}_{\in \mathcal{C}} + \boldsymbol{e}_\ell \\
\textbf{Supp}_c(\boldsymbol{e}_i) & \subseteq & E \text{ with } \dim E = t \\
\mathcal{C}' & \stackrel{\text{def}}{=} & \mathcal{C} + \langle \boldsymbol{y}_1, \cdots, \boldsymbol{y}_\ell \rangle_{\mathbb{F}_q} \\
\mathcal{C}'' & \stackrel{\text{def}}{=} & \langle \boldsymbol{e}_1, \cdots, \boldsymbol{e}_\ell \rangle_{\mathbb{F}_q}
\end{array}
$$

$\mathcal{C}''$ subcode of $\mathcal{C}'$ with support of size $\leqslant t$
$\Rightarrow q^\ell$ codewords in $\mathcal{C}'$ of rank weight $\leqslant t$.

The influence of the structure

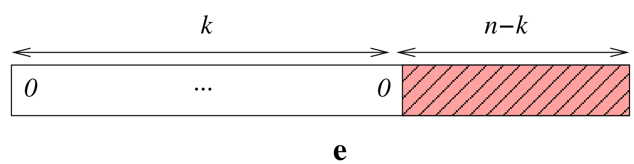
- Finding a codeword of weight t in a matrix code $\in \mathbb{F}_q^{n \times n}$ of dimension $K = kn$ with a combinatorial approach $\tilde{O}\left(q^{tk}\right)$
- Finding a codeword of weight t in an $\mathbb{F}_{q^n}$ linear code $[n, k]_{\mathbb{F}_{q^n}}$ with a combinatorial approach $\tilde{O}\left(q^{(t-1)k}\right)$: q^n codewords of weight t!
- Finding a codeword of weight t in a double-circulant code $\mathbb{F}_{q^n}$ linear code $[2k, k]_{\mathbb{F}_{q^n}}$: $\tilde{O}\left(q^{(t-2)k}\right)$: q^{n+k} codewords of weight t!

The basic principle of combinatorial attacks : rank analogue of the simplest information set decoder

$$\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n},\ \boldsymbol{e} \in \mathbb{F}_q^n,\ \boldsymbol{s} \in \mathbb{F}_q^{n-k}$$

$$\begin{aligned} \boldsymbol{H}\boldsymbol{e}^{\intercal} &= \boldsymbol{s}^{\intercal} \\ |\boldsymbol{e}|_{\text{Ham}} &= t \end{aligned} \qquad (1)$$

Basic principle : hope to be lucky $e_i = 0$ on k positions



$$(1) \Rightarrow \begin{cases} n-k \text{ equations} \\ n-k \text{ unknowns} \end{cases}$$

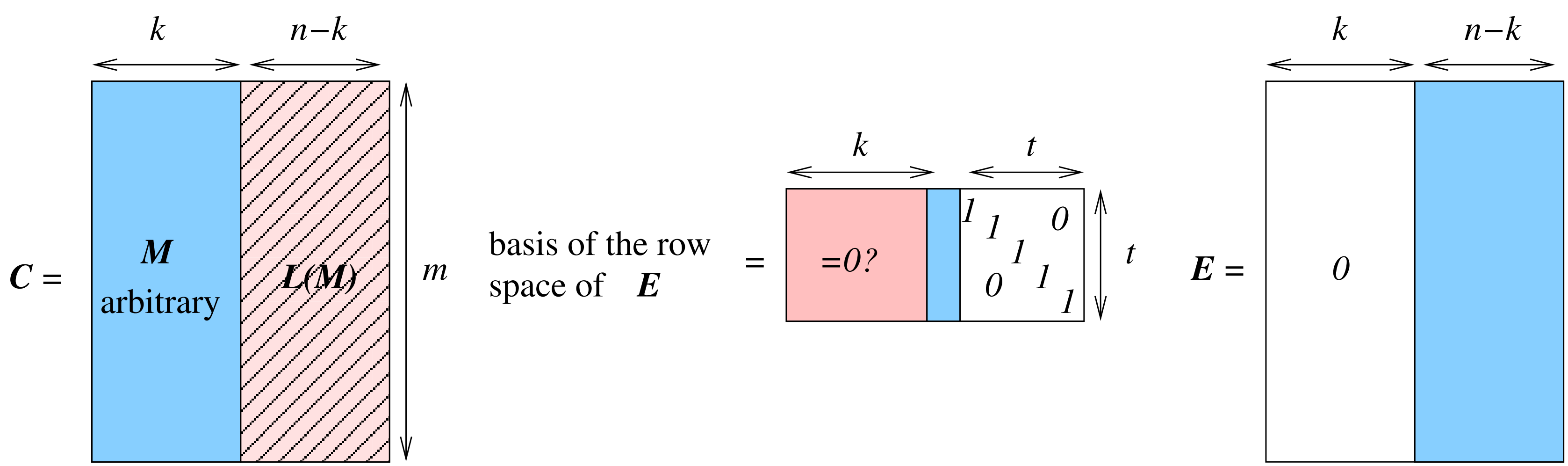
Complexity : $\approx \frac{1}{\text{Prob}(e_i=0,\ \forall i \in I)}$ for a random I of size k.

Rank analogue

Matrix code $\mathcal{C}$ over $\mathbb{F}_q^{m\times n}$ of dimension $K = km$.

$$\boldsymbol{Y} = \boldsymbol{C} + \boldsymbol{E}$$

- Principle 1: (generally) we can choose arbitrarily km entries of $\boldsymbol{C}$ in a codeword $\boldsymbol{C}$ of $\mathcal{C}$ and the rest are linear functions of these entries.
- Principle 2: we hope that the first k columns of the error $\boldsymbol{E}$ are zero (more generally we hope that the first k columns of $\boldsymbol{EP}$ are zero), prob.$= O\left(q^{-kt}\right)$



Scaling ?

	QC-MDPC	QC-LRPC
condition for correct decoding	$wt = O(n)$	$wt = O(n)$
keysize K	$O(n)$	$O(n^2)$
security (bits) λ	$\min(w, t)$	$\min(wn/2, tn/2)$
scaling $(w = t)$	$\lambda = O\left(\sqrt{K}\right)$	$\lambda = O\left(K^{3/4}\right)$

5. Algebraic attacks

- For some time the most efficient approach for solving the rank decoding problem were combinatorial approaches. Parameters of the NIST submissions computed with this belief
- Bardet-Briaud-Bros-Gaborit-Neiger-Ruatta-Tillich/EUROCRYPT 2020] changed this : modelling the problem with an algebraic system and solving with a dedicated Gröbner basis approach is more efficient!
- Bardet-Bros-Cabarcas-Gaborit-Perlner-Smith-Tone-Tillich-Verbel/ASIACRYPT 2020] changed this : modelling the problem with an algebraic system and solving a suitable linear system is more efficent!

The last approach can really be seen as "extracting" the useful computations from the Gröbner basis approach.

Several approaches for solving the MinRank problem

Problem 6. [MinRank (homogeneous)]
Input: m, n, K, t, $\boldsymbol{M}_1, \cdots, \boldsymbol{M}_K \in \mathbb{F}_q^{m\times n}$
Question: $\exists$? $\boldsymbol{x} \in \mathbb{F}_q^K$ *such that*

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

- Kipnis-Shamir approach : bilinear system
- Support modelling : bilinear system
- Minor modelling : system of degree $t+1$ by writing that all minors of size $(t+1)\times(t+1)$ of $\sum_{i=1}^{K} x_i \boldsymbol{M}_i$ are zero.

Setting up the linear system

- Decoding $\boldsymbol{y} = \boldsymbol{c} + \boldsymbol{e}$ with $|\boldsymbol{e}| = t$ reduced to finding a word of weight t in $\tilde{\mathcal{C}} \stackrel{\text{def}}{=} \mathcal{C} + \langle \boldsymbol{y} \rangle$: $q^m - 1$ solutions : $\alpha \boldsymbol{e}$ with $\alpha \in \mathbb{F}_{q^m}^{\times}$

$$\tilde{\mathcal{C}} = \{\boldsymbol{c} \in \mathbb{F}_{q^m}^n : \boldsymbol{c}\tilde{\boldsymbol{H}}^{\intercal} = 0\}.$$

$$\begin{aligned} \boldsymbol{e} &= \begin{pmatrix} 1 & \alpha & \dots & \alpha^{m-1} \end{pmatrix} \boldsymbol{S}\boldsymbol{C} \\ \boldsymbol{S} &\in \mathbb{F}_q^{m \times t} \\ \boldsymbol{C} &\in \mathbb{F}_q^{tx \times n} \end{aligned}$$

unknowns : entries of $\boldsymbol{S}$ and entries of $\boldsymbol{C}$
Columns of $\boldsymbol{S}$ = basis of the support of $\boldsymbol{e}$

The algebraic system

$$\begin{pmatrix} 1 & \alpha & \ldots & \alpha^{m-1} \end{pmatrix} \boldsymbol{S}\boldsymbol{C}\tilde{\boldsymbol{H}}^{\intercal} = \mathbf{0}_{n-k-1} \qquad (2)$$

- Approach 1: solving the bilinear system (2) by computing a Gröbner basis for it. At degree $t+1$: degree fall we obtain new equations of degree t involving only the entries of $\boldsymbol{C}$...
- Approach 2: constructing directly these equations and deduce directly the C_{ij} by solving a (huge) linear system

The point: (2) $\Rightarrow \boldsymbol{C}\tilde{\boldsymbol{H}}^{\intercal}$ is of rank $< t$

Proposition 1. *The maximal minors of the* $t \times (n-k-1)$ *matrix* $\boldsymbol{C}\tilde{\boldsymbol{H}}^{\intercal}$ *are all equal to* 0*.*

- $\binom{n-k-1}{t}$ equations of degree t in the C_{ij}'s

The Cauchy-Binet Formula

$$\begin{aligned} \boldsymbol{A} &\in \mathbb{F}_q^{m\times n} \\ \boldsymbol{B} &\in \mathbb{F}_q^{n\times m} \\ \det(\boldsymbol{A}\boldsymbol{B}) &= \sum_{S\subseteq\{1,\cdots,n\}:|S|=m} \det(\boldsymbol{A}_{*,S})\det(\boldsymbol{B}_{S,*}) \end{aligned}$$

$$c_T \stackrel{\text{def}}{=} \det(\boldsymbol{C}_{*,T}) \text{ for } T\subseteq\{1,\cdots,n\} \text{ and } |T|=t$$

$\Rightarrow$ The maximal minors of $\boldsymbol{C}\tilde{\boldsymbol{H}}^{\intercal}$ are linear combinations of the c_T
$\Rightarrow$ linear system with $\binom{n-k-1}{t}$ equations with coefficients in $\mathbb{F}_{q^m}$ involving $\binom{n}{t}$ variables (the c_T's)
$\Rightarrow$ linear system with $m\binom{n-k-1}{t}$ equations with coefficients in $\mathbb{F}_q$ involving $\binom{n}{t}$ variables ($c_T \in \mathbb{F}_q$!)

Specifying some entries in C

- If $(\boldsymbol{S}, \boldsymbol{C})$ solution of

$$\begin{bmatrix}1 & \alpha & \dots & \alpha^{m-1}\end{bmatrix} \boldsymbol{S}\boldsymbol{C}\boldsymbol{H}^{\intercal} = \mathbf{0}_{n-k-1}$$

so is $(\boldsymbol{S}\boldsymbol{A}, \boldsymbol{A}^{-1}\boldsymbol{C})$ for any $\boldsymbol{A}$ invertible in $\mathbb{F}_q^{t\times t}$. Therefore we may assume that

$$\boldsymbol{C} = \begin{bmatrix}1 & 0 & \dots & 0 & * & * & \dots & * \\ 0 & 1 & \ddots & 0 & * & * & \dots & * \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & * & * & \dots & *\end{bmatrix}$$

-

$$C_{ij} = (-1)^{t+i} c_{\{1,\cdots,t\}\{i\}\cup\{j\}}$$

$\Rightarrow$ solving the linear system in the c_T's yields directly the C_{ij}'s. Once we know $\boldsymbol{C}$ we obtain $\boldsymbol{S}$ by solving a linear system.

should work when $\quad m\binom{n-k-1}{t} \geqslant \binom{n}{t} - 1$

(condition verified for many initial parameters of the rank based submissions).

Further improvements

- Puncturing trick to reduce the number of variables when $m\binom{n-k-1}{t} \geqslant \binom{n}{t} - 1$
- Exhaustive search on a subset of variables to reduce to the previous case

One step beyond

Using the MinRank formulation (with $K = (k+1)m$)

$$\boldsymbol{SC} = \sum_{j=1}^{K} x_j \boldsymbol{M}_j \tag{3}$$

$$\boldsymbol{r}_i \stackrel{\text{def}}{=} i\text{-th row of } \sum_{j=1}^{K} x_j \boldsymbol{M}_j$$

$$\begin{aligned}(3) \;&\Rightarrow\; r_i \text{ belongs to the rowspace of } \boldsymbol{C}\\ &\Rightarrow\; \begin{bmatrix}\boldsymbol{r}_i\\ \boldsymbol{C}\end{bmatrix} \text{ is of rank } \leqslant t\\ &\Rightarrow\; \text{all maximal minors are } = 0\\ &\Rightarrow\; m\binom{n}{t+1} \text{ linear eq. in the } x_i c_T\text{'s}\end{aligned}$$

Solve (3) when $\underbrace{m\binom{n}{t+1}}_{\#\text{ eq.}} \geqslant \underbrace{K\binom{n}{t}}_{\#\text{ var.}} - 1$

A step further

$m\binom{n}{t+1}\binom{K+b-2}{b-1}$ equations of the form

$$x_{i_1}\cdots x_{i_{b-1}}\det\begin{bmatrix}\boldsymbol{r}_i\\ \boldsymbol{C}\end{bmatrix}_{*,S}=0$$

where S is a subset of $\{1,\cdots,n\}$ of size $t+1$.
Cofactor expansion $\Rightarrow$ equations in the $x_{i_1}\cdots x_{i_b}c_T$'s.

Problem: not all equations are independent:

$$\det\begin{bmatrix}\boldsymbol{r}_i\\ \boldsymbol{r}_j\\ \boldsymbol{C}\end{bmatrix}_{*,S}+\det\begin{bmatrix}\boldsymbol{r}_j\\ \boldsymbol{r}_0\\ \boldsymbol{C}\end{bmatrix}_{*,S}=0$$

$\Rightarrow$ linear relation between these equations.

Linearization

$$\begin{aligned}
\#\text{ of lin. indep. relations } D &= \sum_{i=1}^{b}(-1)^{i+1}\binom{n}{t+i}\binom{m+i-1}{i}\binom{K+b-i-1}{b-i} \\
\#\text{ of variables} &= \underbrace{\binom{n}{t}}_{\#\, c_T}\ \underbrace{\binom{K+b-1}{b}}_{\#\text{ of mon. of degree } b}
\end{aligned}$$

We expect to solve by linearization when # of lin. independent relations $\geqslant$ # of variables -1.

Results

	(m,n,k,r)	$\frac{m\binom{n-k-1}{r}}{\binom{n}{r}-1}$	a	p	b	**complexity (bits)**
Loidreau	$(128,120,80,4)$	1.28	0	43	0	**65**
ROLLO-I-128	$(79,94,47,5)$	1.97	0	9	0	**71**
ROLLO-I-192	$(89,106,53,6)$	1.06	0	0	0	**87**
ROLLO-I-256	$(113,134,67,7)$	0.67	3	0	1	**151**
ROLLO-II-128	$(83,298,149,5)$	2.42	0	40	0	**93**
ROLLO-II-192	$(107,302,151,6)$	1.53	0	18	0	**111**
ROLLO-II-256	$(127,314,157,7)$	0.89	0	6	1	**159**
ROLLO-III-128	$(101,94,47,5)$	2.52	0	12	0	**70**
ROLLO-III-192	$(107,118,59,6)$	1.31	0	4	0	**88**
ROLLO-III-256	$(131,134,67,7)$	0.78	0	0	1	**131**
RQC-I	$(97,134,67,5)$	2.60	0	18	0	**77**
RQC-II	$(107,202,101,6)$	1.46	0	10	0	**101**
RQC-III	$(137,262,131,7)$	0.93	3	0	0	**144**

Multivariate schemes

						Complexity		
GeMSS(D, n, Δ, v)	n/m	K	r	n'	b	New	Previous	Type
GeMSS128$(513, 174, 12, 12)$	174	162	34	61	2	**154**	522	MinRank
GeMSS192$(513, 256, 22, 20)$	265	243	52	94	2	**223**	537	MinRank
GeMSS256$(513, 354, 30, 33)$	354	324	73	126	3	**299**	1254	MinRank
RedGeMSS128$(17, 177, 15, 15)$	177	162	35	62	2	**156**	538	MinRank
RedGeMSS192$(17, 266, 23, 25)$	266	243	53	95	2	**224**	870	MinRank
RedGeMSS256$(17, 358, 34, 35)$	358	324	74	127	3	**301**	1273	MinRank
BlueGeMSS128$(129, 175, 13, 14)$	175	162	35	63	2	**158**	537	MinRank
BlueGeMSS192$(129, 265, 22, 23)$	265	243	53	95	2	**224**	870	MinRank
BlueGeMSS256$(129, 358, 34, 32)$	358	324	74	127	3	**301**	1273	MinRank

Rainbow$(GF(q), v_1, o_1, o_2)$	n	K	r	n'	b	New	Previous	Best / Type
Ia$(GF(16), 32, 32, 32)$	96	33	64	82	3	155	161	145 / RBS
IIIc$(GF(256), 68, 36, 36)$	140	37	104	125	5	**208**	585	215 / DA
Vc$(GF(256), 92, 48, 48)$	188	49	140	169	5	**272**	778	275 / DA

Conclusion

NIST :

Despite the development of algebraic attacks, NIST believes rank-based cryptography should continue to be researched. The rank metric cryptosystems offer a nice alternative to traditional hamming metric codes with comparable bandwidth.