LRPC codes with multiple syndromes: near ideal-size KEMs without ideals

Carlos Aguilar-Melchor, Nicolas Aragon, **Victor Dyseryn**,
Philippe Gaborit, Gilles Zémor

XLIM, Université de Limoges, France

PQCrypto - September 28, 2022

Summary

Summary

Error-correcting codes

Let $\mathbb{F}_q$ be a finite field. We equip $(\mathbb{F}_q)^n$ with a weight w (and a distance $\delta(\boldsymbol{x}, \boldsymbol{y}) = w(\boldsymbol{x} - \boldsymbol{y})$).

Definition (Error-correcting code)

An error-correcting code $\mathcal{C}$ of length n and dimension k is a subspace of $(\mathbb{F}_q)^n$ (of dimension k).
The minimal distance d of a code $\mathcal{C}$ is the smallest weight of the non-zero vectors in $\mathcal{C}$.

$$d := \min_{\boldsymbol{x} \in \mathcal{C} \setminus \{\boldsymbol{0}\}} \{w(\boldsymbol{x})\}$$

A code is given by either :

- a generating matrix $\boldsymbol{G} \in \mathbb{F}_q^{k \times n}$ whose rows form a basis of $\mathcal{C}$.
- a parity-check matrix $\boldsymbol{H} \in \mathbb{F}_q^{(n-k) \times n}$ such that $\mathcal{C} = \{\boldsymbol{x} \in (\mathbb{F}_q)^n | \boldsymbol{H}\boldsymbol{x}^\top = \boldsymbol{0}\}$

Hamming matric codes

Definition (Hamming weight)

The hamming weight of a word $\boldsymbol{x} = (x_1, ..., x_n)$ is the number of non-zero coordinates

$$w_h(\boldsymbol{x}) = \#\{i \in [1, n] \mid x_i \neq 0\}$$

Definition (Hamming support)

The support of a word $\boldsymbol{x} = (x_1, ..., x_n) \in (\mathbb{F}_q)^n$ is the set of indexes of its non-zero coordinates

$$\mathrm{Supp}(\boldsymbol{x}) = \{i \in [1, n] \mid x_i \neq 0\}$$

Rank metric codes

In rank metric, we consider $\mathbb{F}_{q^m}$-linear codes ($\mathbb{F}_{q^m}$ is a field extension of $\mathbb{F}_q$ of degree m).

Definition (Rank weight)

An element $\boldsymbol{x} = (x_1, ..., x_n) \in (\mathbb{F}_{q^m})^n$ can be unfold against an $\mathbb{F}_q$-basis of $\mathbb{F}_{q^m}$ in a matrix

$$\mathcal{M}(\boldsymbol{x}) = \begin{pmatrix} x_{1,1} & \cdots & x_{n,1} \\ \vdots & & \vdots \\ x_{1,m} & \cdots & x_{n,m} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{F}_q)$$

The rank weight of $\boldsymbol{x}$ is defined as the rank of this matrix (which does not depend on the choice of the basis).

$$w_r(\boldsymbol{x}) = \text{Rank } \mathcal{M}(\boldsymbol{x}) \in [0, \min(m, n)]$$

Example

Let $\mathbb{F}_8 = \mathbb{F}_{2^3}$ and let α such that $\mathbb{F}_8 \simeq \mathbb{F}_2[\alpha] = \mathit{Vect}(1, \alpha, \alpha^2)$.

Example

$$\boldsymbol{x} = (1, \alpha, \alpha^2 + 1, \alpha + 1) \in \mathbb{F}_8^4$$

$$\mathcal{M}(\boldsymbol{x}) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$w_r(\boldsymbol{x}) = 3$$

Support in rank metric

Definition (Rank support)

The support of a word $\boldsymbol{x} = (x_1, ..., x_n) \in (\mathbb{F}_{q^m})^n$ is the subspace of $\mathbb{F}_{q^m}$ generated by its coordinates :

$$\text{Supp}(\boldsymbol{x}) = \langle x_1, ..., x_n \rangle_{\mathbb{F}_q} \subset \mathbb{F}_{q^m}$$

Hamming metric : $w_h(\boldsymbol{x}) = \#(\text{Supp}(\boldsymbol{x}))$

Rank metric : $w_r(\boldsymbol{x}) = dim(\text{Supp}(\boldsymbol{x}))$

Difficult problems in code-based cryptography

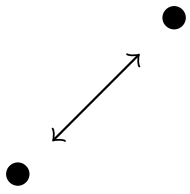
Definition (Syndrome Decoding $\mathtt{SD}(n, k, w)$)

Given a random parity check matrix $\boldsymbol{H} \in \mathcal{M}_{n-k,n}(\mathbb{F}_q)$ and a syndrome $\boldsymbol{s} = \boldsymbol{H}\boldsymbol{e}$ for $\boldsymbol{e}$ an error of Hamming weight $w_h(\boldsymbol{e}) = w$, find $\boldsymbol{e}$.

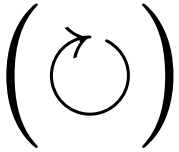
Definition (Rank Syndrome Decoding $\mathtt{RSD}(m, n, k, w)$)

Given a random parity check matrix $\boldsymbol{H} \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$ and a syndrome $\boldsymbol{s} = \boldsymbol{H}\boldsymbol{e}$ for $\boldsymbol{e}$ an error of rank weight $w_r(\boldsymbol{e}) = w$, find $\boldsymbol{e}$.

Variations in code-based crypto



Metric

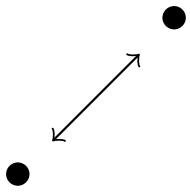


Structure

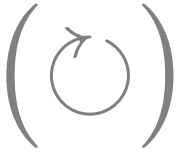
MGP

Masking

Variations in code-based crypto



Metric

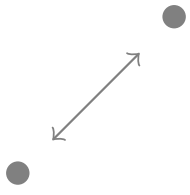


Structure

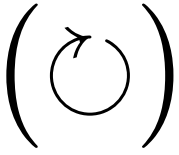
MGP

Masking

Variations in code-based crypto



Metric



Structure

Masking

Structuration

To reduce the memory footprint of the public key, we add structure to the codes.

Definition (Double circulant code)

A double circulant code is a code $\mathcal{C}[2n, n]$ which admits a double circulating matrix as a generating matrix :

$$\boldsymbol{G} = \left(\begin{array}{cccc|cccc} a_0 & a_1 & \dots & a_{n-1} & b_0 & b_1 & \dots & b_{n-1} \\ a_{n-1} & a_0 & \ddots & a_{n-2} & b_{n-1} & b_0 & \ddots & b_{n-2} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_0 & b_1 & b_2 & \dots & b_0 \end{array}\right)$$

Structuration

Lattices	$\Longrightarrow$	Module or Ring structure
Hamming codes	$\Longrightarrow$	Quasi-cyclic structure
Rank codes	$\Longrightarrow$	Ideal structure

Difficult problems in a structured context

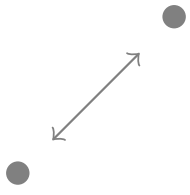
Definition (Ideal Rank Syndrome Decoding $\mathtt{IRSD}(m, n, k, w)$)

Given an ideal random parity check matrix $\boldsymbol{H} \in \mathcal{M}_{n-k,n}(\mathbb{F}_q)$ and a syndrome $\boldsymbol{s} = \boldsymbol{H}\boldsymbol{e}$ for $\boldsymbol{e}$ an error of rank weight $w(\boldsymbol{e}) = w$, find $\boldsymbol{e}$.

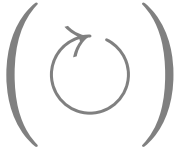
Problematic with the structure :

- Quantum attacks [1]
- Potential weaknesses

Variations in code-based crypto



Metric



Structure

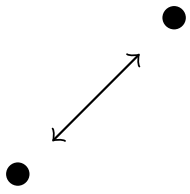
MGP

Masking

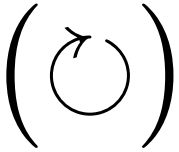
To mask or not to mask

	No masking	With masking
Key	$\mathsf{sk} = \boldsymbol{e}$ $\mathsf{pk} = (\boldsymbol{H}, \boldsymbol{y})$ where $\boldsymbol{y} = \boldsymbol{He}$	$\mathsf{sk} = \boldsymbol{H}$ $\mathsf{pk} = \boldsymbol{H}'$ where $\boldsymbol{H}'$ is a masked version of $\boldsymbol{H}$ (usually $\boldsymbol{H}' = \boldsymbol{MHP}$ with $\boldsymbol{M}$ a random invertible matrix and $\boldsymbol{P}$ an isometry matrix)
	(R)SD	(R)SD and disting. problems

Variations in code-based crypto



Metric



Structure

MGP

Masking

Panorama of code-based cryptography

	Hamming metric		Rank metric	
	Structured	Unstructured	Structured	Unstructured
No masking	HQC		RQC	
With masking	BIKE LEDAcrypt	Classic McEliece	ROLLO	

Panorama with recent results

	Hamming metric		Rank metric	
	Structured	Unstructured	Structured	Unstructured
No masking	HQC		RQC Multi-RQC [2]	Multi-UR [2]
With masking	BIKE LEDAcrypt	Classic McEliece	ROLLO *this work*	*this work*

Summary

Low Rank Parity Check Codes

Definition (Homogenous matrix)

An homogeneous matrix of weight d is a full-rank matrix $\boldsymbol{H} = (h_{ij})_{\substack{1 \leqslant i \leqslant k \\ 1 \leqslant j \leqslant n}} \in \mathbb{F}_{q^m}^{k \times n}$ whose coordinates generate an $\mathbb{F}_q$-subspace of dimension d :

$$\dim(\langle h_{ij} \rangle_{\mathbb{F}_q}) = d$$

Low Rank Parity Check Codes

Definition (Homogenous matrix)

An homogeneous matrix of weight d is a full-rank matrix $\boldsymbol{H} = (h_{ij})_{\substack{1 \leqslant i \leqslant k \\ 1 \leqslant j \leqslant n}} \in \mathbb{F}_{q^m}^{k \times n}$ whose coordinates generate an $\mathbb{F}_q$-subspace of dimension d :

$$\dim(\langle h_{ij} \rangle_{\mathbb{F}_q}) = d$$

Definition (LRPC codes)

An LRPC code of dual weight d is a code $\mathcal{C}$ which admits an homogeneous matrix of small weight d as a parity-check matrix.

LRPC decoding

Problem (LRPC decoding)

Let $E = \langle e_1, ..., e_r \rangle$ an (unknown) subspace of $\mathbb{F}_{q^m}$ of dimension r and $F = \langle f_1, ..., f_d \rangle$ a (given) subspace of $\mathbb{F}_{q^m}$ of dimension d. Given an LRPC matrix $\boldsymbol{H} \in F^{n-k \times n}$ and $\boldsymbol{s} = \boldsymbol{He}$ where $\boldsymbol{e} \in E^n$, find E.

LRPC decoding

Problem (LRPC decoding)

Let $E = \langle e_1, ..., e_r \rangle$ an (unknown) subspace of $\mathbb{F}_{q^m}$ of dimension r and $F = \langle f_1, ..., f_d \rangle$ a (given) subspace of $\mathbb{F}_{q^m}$ of dimension d. Given an LRPC matrix $\boldsymbol{H} \in F^{n-k \times n}$ and $\boldsymbol{s} = \boldsymbol{He}$ where $\boldsymbol{e} \in E^n$, find E.

Proposition ([3])

There exists a polynomial algorithm RSR which, on input $\boldsymbol{H}$ and $\boldsymbol{s} = \boldsymbol{He}$, returns E.
The Decoding Failure Rate of RSR is bounded from above by :

$$q^{rd-(n-k)-1} + q^{-(d-1)(m-rd-r)}$$

LRPC decoding

Problem (LRPC decoding)

Let $E = \langle e_1, ..., e_r \rangle$ an (unknown) subspace of $\mathbb{F}_{q^m}$ of dimension r and $F = \langle f_1, ..., f_d \rangle$ a (given) subspace of $\mathbb{F}_{q^m}$ of dimension d. Given an LRPC matrix $\boldsymbol{H} \in F^{n-k \times n}$ and $\boldsymbol{s} = \boldsymbol{He}$ where $\boldsymbol{e} \in E^n$, find E.

Proposition ([3])

There exists a polynomial algorithm RSR which, on input $\boldsymbol{H}$ and $\boldsymbol{s} = \boldsymbol{He}$, returns E.
The Decoding Failure Rate of RSR is bounded from above by :

$$q^{rd-(n-k)-1} + q^{-(d-1)(m-rd-r)}$$

Application of LRPC to cryptography

Definition (Key generation)

Let $\boldsymbol{U} = (\boldsymbol{A}|\boldsymbol{B})$ an ideal LRPC matrix of weight d and size $k \times 2k$.

$$\begin{cases} pk &= \boldsymbol{H} = (\boldsymbol{I}|\boldsymbol{A}^{-1}\boldsymbol{B}) \\ sk &= \boldsymbol{U} \end{cases}$$

Application of LRPC to cryptography

Definition (Key generation)

Let $\boldsymbol{U} = (\boldsymbol{A}|\boldsymbol{B})$ an ideal LRPC matrix of weight d and size $k \times 2k$.

$$\begin{cases} pk &= \boldsymbol{H} = (\boldsymbol{I}|\boldsymbol{A}^{-1}\boldsymbol{B}) \\ sk &= \boldsymbol{U} \end{cases}$$

Definition (Encaps)

Choose an error support E of dimension r. Pick a random error $\boldsymbol{e}$ in E^n and send ciphertext $\boldsymbol{c} = \boldsymbol{H}\boldsymbol{e}$. The shared secret is $Hash(E)$.

Application of LRPC to cryptography

Definition (Key generation)

Let $\boldsymbol{U} = (\boldsymbol{A}|\boldsymbol{B})$ an ideal LRPC matrix of weight d and size $k \times 2k$.

$$\begin{cases} pk &= \boldsymbol{H} = (\boldsymbol{I}|\boldsymbol{A}^{-1}\boldsymbol{B}) \\ sk &= \boldsymbol{U} \end{cases}$$

Definition (Encaps)

Choose an error support E of dimension r. Pick a random error $\boldsymbol{e}$ in E^n and send ciphertext $\boldsymbol{c} = \boldsymbol{H}\boldsymbol{e}$. The shared secret is $Hash(E)$.

Definition (Decaps)

Compute $\boldsymbol{s} = \boldsymbol{A}\boldsymbol{c} = \boldsymbol{U}\boldsymbol{e}$ and use RSR algorithm to find E.

ROLLO-I parameters

Instance	pk size	ct size	Security	DFR
ROLLO-I-128	696	696	128	2^{-28}
ROLLO-I-192	958	958	192	2^{-34}
ROLLO-I-256	1371	1371	256	2^{-33}

FIGURE: Parameters for ROLLO-I. Sizes are in bytes and security is expressed in bits.

ROLLO-II parameters

Instance	pk size	ct size	Security	DFR
ROLLO-II-128	1941	2089	128	2^{-134}
ROLLO-II-192	2341	2469	192	2^{-130}
ROLLO-II-256	2559	2687	256	2^{-136}

FIGURE: Parameters for ROLLO-II. Sizes are in bytes and security is expressed in bits.

Summary

Idea

Definition (Key generation)

Let $\boldsymbol{U} = (\boldsymbol{A}|\boldsymbol{B})$ an LRPC matrix of weight d.

$$\begin{cases} pk &= \boldsymbol{H} = (\boldsymbol{I}|\boldsymbol{A}^{-1}\boldsymbol{B}) \\ sk &= \boldsymbol{U} \end{cases}$$

Idea

Definition (Key generation)

Let $\boldsymbol{U} = (\boldsymbol{A}|\boldsymbol{B})$ an LRPC matrix of weight d.

$$\begin{cases} pk & = & \boldsymbol{H} = (\boldsymbol{I}|\boldsymbol{A}^{-1}\boldsymbol{B}) \\ sk & = & \boldsymbol{U} \end{cases}$$

Definition (Encaps)

Choose an error support E of dimension r. Pick ℓ random errors $\boldsymbol{e}_i$ in E^n for $1 \leq i \leq \ell$ and send ciphertexts $\boldsymbol{c}_i = \boldsymbol{H}\boldsymbol{e}_i$. The shared secret is $Hash(E)$.

Idea

Definition (Key generation)

Let $\boldsymbol{U} = (\boldsymbol{A}|\boldsymbol{B})$ an LRPC matrix of weight d.

$$\begin{cases} pk &= \boldsymbol{H} = (\boldsymbol{I}|\boldsymbol{A}^{-1}\boldsymbol{B}) \\ sk &= \boldsymbol{U} \end{cases}$$

Definition (Encaps)

Choose an error support E of dimension r. Pick ℓ random errors $\boldsymbol{e}_i$ in E^n for $1 \leq i \leq \ell$ and send ciphertexts $\boldsymbol{c}_i = \boldsymbol{H}\boldsymbol{e}_i$. The shared secret is $Hash(E)$.

Definition (Decaps)

Compute $\boldsymbol{s}_i = \boldsymbol{A}\boldsymbol{c}_i = \boldsymbol{U}\boldsymbol{e}_i$ and use RSR algorithm with multiple syndromes to find E.

New failure probability

The LRPC decoding algorithm has now several syndromes [1] as inputs

$$\boldsymbol{s}_i = \boldsymbol{U}\boldsymbol{e}_i$$

Proposition

The Decoding Failure Rate of algorithm RSR with multiple syndromes is bounded from above by :

$$(n-k)q^{rd-(n-k)\ell} + q^{-(d-1)(m-rd-r)}$$

1. can also be seen as decoding an interleaved LRPC code [4]

Parameters with an ideal structure

Instance	pk size	ct size	Security	DFR
ROLLO-II-128	1,941	2,089	128	2^{-134}
ROLLO-II-192	2,341	2,469	192	2^{-130}
ROLLO-II-256	2,559	2,687	256	2^{-136}

FIGURE: Parameters for ROLLO-II. Sizes are in bytes and security is expressed in bits.

$\Downarrow$

Instance	pk size	ct size	Security	DFR
ILRPC-MS-128	488	1,951	128	2^{-126}
ILRPC-MS-192	846	3,384	192	2^{-198}

FIGURE: Parameters for ILRPC-MS. Sizes are in bytes and security is expressed in bits.

Parameters without an ideal structure

Instance	pk size	ct size	Security	DFR
LRPC-MS-128	$4,083$	$3,122$	128	2^{-126}
LRPC-MS-192	$7,663$	$5,474$	192	2^{-190}

FIGURE: Parameters for LRPC-MS. Sizes are in bytes and security is expressed in bits.

Comparison to other KEMs

Instance	128 bits	192 bits
LRPC-MS	**7,205**	**12,445**
Loong.CCAKEM-III [5]	18,522	N/A
FrodoKEM	19,336	31,376
Loidreau cryptosystem [6]	36,300	N/A
Classic McEliece	261,248	524,348

FIGURE: Comparison of sizes of unstructured post-quantum KEMs. The sizes represent the sum of public key and ciphertext expressed in bytes.

Instance	128 bits	192 bits
ILRPC-MS	**2,439**	**4,230**
BIKE	3,113	6,197
ROLLO-II	4,030	4,810
HQC	6,730	13,548

FIGURE: Comparison of sizes of structured code-based KEMs. The sizes represent the sum of public key and ciphertext expressed in bytes.

Specificity to rank metric

- Sending errors with the same support is less efficient in Hamming metric
- Additional information given by multiple syndromes can be specifically leveraged by LRPC decoding algorithm

IND-CPA proof

Definition (LRPC indistinguishability)

Given a matrix $\boldsymbol{H} \in \mathbb{F}_{q^m}^{(n-k)\times k}$, distinguish whether the code $\mathcal{C}$ with the parity-check matrix $(\boldsymbol{I}_{n-k}|\boldsymbol{H})$ is a random code or an LRPC code of weight d.

Definition (Rank Support Learning $\mathtt{RSL}(m, n, k, w, \ell)$ [7])

Given a random parity check matrix $\boldsymbol{H} \in \mathcal{M}_{n-k,n}(\mathbb{F}_q)$ and ℓ syndromes $\boldsymbol{s}_i = \boldsymbol{H}\boldsymbol{e}_i$ for $\boldsymbol{e}_i$ errors of same support E a subspace of dimension w, find E.

Summary

Conclusion

- New rank metric based cryptosystem with competitive parameters and no ideal structure
- Probabilistic result on the support of the product of two random matrices
- Additional idea to make m down by 10 %
- The approach can generalize to RQC but is less efficient in that case [2]

Thank you for your attention !

References I


Ronald Cramer, Léo Ducas, and Benjamin Wesolowski.
Mildly short vectors in cyclotomic ideal lattices in quantum polynomial time.
Journal of the ACM (JACM), 68(2) :1–26, 2021.


Loïc Bidoux, Pierre Briaud, Maxime Bros, and Philippe Gaborit.
Rqc revisited and more cryptanalysis for rank-based cryptography.
arXiv preprint arXiv :2207.01410, 2022.


Nicolas Aragon, Olivier Blazy, Jean-Christophe Deneuville, Philippe Gaborit, Adrien Hauteville, Olivier Ruatta, Jean-Pierre Tillich, Gilles Zémor, Carlos Aguilar Melchor, Slim Bettaieb, Loïc Bidoux, Bardet Magali, and Ayoub Otmani.
ROLLO (merger of Rank-Ouroboros, LAKE and LOCKER).

References II

Second round submission to the NIST post-quantum cryptography call, March 2019.


Julian Renner, Thomas Jerkovits, and Hannes Bartz.
Efficient decoding of interleaved low-rank parity-check codes.
In *2019 XVI International Symposium "Problems of Redundancy in Information and Control Systems"(REDUNDANCY)*, pages 121–126. IEEE, 2019.


Li-Ping Wang.
Loong : a new ind-cca-secure code-based kem.
In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 2584–2588. IEEE, 2019.

References III

Ba Duc Pham.
Étude et conception de nouvelles primitives de chiffrement fondées sur les codes correcteurs d'erreurs en métrique rang.
PhD thesis, Rennes 1, 2021.

P. Gaborit, A. Hauteville, D. H. Phan, and J.-P. Tillich.
Identity-based encryption from rank metric.
In *Advances in Cryptology - CRYPTO*, 2017.