Introduction to rank based cryptography

Philippe Gaborit (XLIM, Université de Limoges)

Summer School SAC - Ottawa, August 24, 2026

Summary

1. Rank Codes : definition and basic properties
2. Decoding in rank metric
3. Complexity issues of Rank Decoding : decoding random rankcodes
4. Encryption/key exchange LRPC/RQC
5. Signature with codes RYDE
6. The case of MinRank : new recent approaches

Rank Codes : definition and basic properties

Rank metric codes

The rank metric is defined in finite extensions.

- $GF(q)$ a finite field with q a power of a prime.
- $GF(q^m)$ an extension of degree m of $GF(q)$.
- $B = (b_1, ..., b_m)$ a basis of $GF(q^m)$ over $GF(q)$.

$GF(q^m)$ can be seen as a vector space on $GF(q)$.

- $\mathcal{C}$ a linear code over $GF(q^m)$ of dimension k and length n.
- G a $k \times n$ generator matrix of the code $\mathcal{C}$.
- H a $(n-k) \times n$ parity check matrix of $\mathcal{C}$, $G.H^t = 0$.
- H a dual matrix, $x \in GF(q^m)^n \rightarrow$ syndrome of $x = H.x^t \in GF(q^m)^{n-k}$

Words of the code $\mathcal{C}$ are n-uplets with coordinates in $GF(q^m)$.

$$v = (v_1, \ldots, v_n)$$

with $v_j \in GF(q^m)$.
Any coordinate $v_j = \sum_{i=1}^{m} v_{ij} b_i$ with $v_{ij} \in GF(q)$.

$$v(v_1, ..., v_n) \rightarrow V = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{pmatrix}$$

Definition (Rank weight of word)

v has rank $r = Rank(v)$ iff the rank of $V = (v_{ij})_{ij}$ is r.
equivalently $Rank(v) = r <=> v_j \in V_r \subset GF(q^m)^n$ with dim(V_r)=r.

the determinant of V does not depend on the basis

Definition (Rank distance)

Let $x, y \in GF(q^m)^n$, the rank distance between x and y is defined by $d_R(x, y) = Rank(x - y)$.

Definition (Minimum distance)

Let C be a $[n, k]$ rank code over $GF(q^m)$, the minimum rank distance d of C is $d = min\{d_R(x, y)|x, y \in C, x \neq y\}$;

Theorem (Unique decoding)

Let $C[n, k, d]$ be a rank code over $GF(q^m)$. Let e an error vector with $r = Rank(e) \leq \frac{d-1}{2}$, and $c \in C$:
if $y = c + e$ then there exists a unique element $c' \in C$ such that $d(y, c') = r$. Therefore $c' = c$.

proof : same as for Hamming, distance property.

Rank isometry

Notion of **isometry** : weight preservation

- Hamming distance : $n \times n$ permutation matrices
- Rank distance : $n \times n$ invertible matrices over $GF(q)$

proof : multiplying a codeword $x \in GF(q^m)^n$ by an $n \times n$ invertible matrix **over the base field GF(q)** does not change the rank (see x as a $m \times n$ matrix over $GF(q)$).

remark : for any $x \in GF(q^m)^n$: $Rank(x) \leq w_H(x)$: potential linear combinations on the x_i may only decrease the rank weight.

Support analogy

An important insight between Rank and Hamming distances tool : support analogy

- **support of a word of $GF(q)^n$ in Hamming metric**
  $x(x_1, x2, \cdots, x_n)$: set of positions $x_i \neq 0$
- **support of a word of $GF(q)^n$ in rank metric**
  $x(x_1, x2, \cdots, x_n)$: the subspace over $GF(q)$, $E \subset GF(q^m)$ generated by $\{x_1, \cdots, x_n\}$
- in both cases if the order of size of the support is small, knowing the support of x and syndrome $s = H.x^t$ permits to recover the complete coordinates of x.

Sphere packing bound

Counting the number of possible supports for length n and dimension t

- Hamming : number of sets with t elements in sets of n elements : Newton binomial $\binom{n}{t}$ $(\leq 2^n)$
- Rank : number of subspaces of dimension t over $GF(q)$ in the space of dimension n $GF(q^m)$: Gaussian binomial $\begin{bmatrix} n \\ t \end{bmatrix}_q (\sim q^{t(n-t)})$

For r=RankGV (e.g. n=m, k=n/2, r=0.29n) $q^{r(m-r)} = q^{O(n^2)}$, quadratic exponential coefficient.

Sphere packing bound

Theorem (Sphere packing bound)

Let $C[n,k,d]$ be a rank code over $GF(q^m)^n$, the parameters n,k,d and d satisfy : $q^{mk} B(n,m,q,\lfloor \frac{d-1}{2} \rfloor) \leq q^{nm}$

Theorem (Singleton bound)

Let $C[n,k,d]$ be a rank code over $GF(q^m)^n$, the parameters n,k and d satisfy : $d \leq 1 + \lfloor \frac{(n-k)m}{n} \rfloor$

The **rank Gilbert-Varshamov (GVR) bound** for a $C[n,k]$ rank code over $GF(q^m)^n$ with dual matrix H corresponds to the average value of the minimum distance of a random $[n,k]$ rank code.

asymptotically : in the case $m = n$: $\frac{GVR(n,k,m,q)}{n} \sim 1 - \sqrt{\frac{k}{n}}$

Decoding in rank metric

Families of decodable codes in rank metric

There exists 3 main families of decodable codes in rank metric

- Gabidulin codes (1985) (analog of Reed-Solomon codes with rank metric and q-polynomials)
- Simple codes (2017)
- LRPC codes (2013)
- Augmented Gabidulin codes (2022) (EG 2009)

These codes have different properties, a lot of attention was given to rank metric and especially to subspace metric with the development of Network coding in the years 2000's.

q-polynomials

Definition (Ore 1933)

A q-polynomial of q-degree r over $GF(q^m)$ is a polynomal of the form $P(x) = \sum_{i=0}^{r} p_i x^{q^i}$ with $p_r \neq 0$ et $p_i \in GF(q^m)$.

Proposition (GF(q)-linearity)

Let P be a q-polynomial and $\alpha, \beta \in GF(q)$ and $x, y \in GF(q^m)$ then :

$$P(\alpha x + \beta y) = \alpha P(x) + \beta P(y)$$

proof : for any $x, y \in GF(q^m), \alpha \in GF(q) : (x+y)^q = x^q + y^q$ and $(\alpha x)^q = \alpha x^q$

Relation between q-polynomials and subspaces

Proposition

The set of zeros of a q-polynomial of q-degree r is a $GF(q)$-linear space of dimension $\leq r$.

Proposition (annulator polynomial)

For any linear subspace E of dimension $r \leq m$ of $GF(q^m)$, there exists a unique monic q-polynomial of q-degree r such that :

$$\forall x \in E, P(x) = 0$$

proof : by construction from Ore paper.

Gabidulin codes

They are a natural analog of Reed-Solomon codes.

Define $\mathcal{P}_k$ = the set of q-polynomials of q-degree $\leq k$

Definition

Let $GF(q^m)$ be an extension field of $GF(q)$ and let $\{x_1, \cdots, x_n\}$ be a set of independent elements of $GF(q^m)$ over $GF(q)$ and let $(k \leq n \leq m)$, a Gabidulin code $[n, k]$ over $GF(q^m)$ is :

$$Gab[n, k] = \{c(p) = (p(x_1), p(x_2), \cdots, p(x_n)) | p \in \mathcal{P}_{k-1}\}$$

Theorem

The codes $Gab[n, k, r]$ are $[n, k, n - k + 1]$ rank codes over $GF(q^m)$. They are MRD codes.

A simple optimal family of codes[Gaborit et al. 2017]

Consider codes $[n, k]$ over $GF(q^m)$ defined by their $(n-k) \times n$ dual matrices :

$$H = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & \mathrm{R} \end{pmatrix}$$

for R random matrix over $GF(q^m)$;

Suppose one wants to decode $y = x + e$ with e a random error of rank weight r and error support E one computes the syndrome $s = H.e^t$

idea : when computing the first r coordinates of the syndrome one obtains, r elements of E, hence :

with a good probability the first r coordinates of the dual fix the error support for the receiver.

Rank metric : you have to think GLOBAL : coordinates are not independent !

LRPC codes

LDPC : dual with low weight (ie : small support)

$\rightarrow$ equivalent for rank metric : dual with small rank support

Definition (GMRZ13)

A Low Rank Parity Check (LRPC) code of rank d, length n and dimension k over F_{q^m} is a code such that the code has for parity check matrix, a $(n-k) \times n$ matrix $H(h_{ij})$ such that the vector space F of F_{q^m} generated by its coefficients h_{ij} has dimension at most d. We call this dimension the weight of H.

In other terms : all coefficients h_{ij} of H belong to the same 'low' vector space $F < F_1, F_2, \cdots, F_d >$ of F_{q^m} of dimension d.

Decoding LRPC codes

Idea : as usual recover the support and then deduce the coordinates values.

Let $e(e_1, ..., e_n)$ be an error vector of weight r, ie : $\forall e_i : e_i \in E$, and dim(E)=r. Suppose $H.e^t = s = (s_1, ..., s_{n-k})^t$.

$$e_i \in E < E_1, ..., E_r >, h_{ij} \in F < F_1, F_2, \cdots, F_d >$$

$$\Rightarrow s_k \in < E_1F_1, .., E_rF_d >$$

$\Rightarrow$ if $n - k$ is large enough, it is possible to recover the product space $< E_1F_1, .., E_rF_d >$

Decoding LRPC codes

Syndrome $s(s_1, .., s_{n-k}) : S =< s_1, .., s_{n-k} > \subset < E_1 F_1, .., E_r F_d >$

Suppose $S =< E.F > \Rightarrow$ possible to recover E.

Let $S_i = F_i^{-1}.S$, since

$$S =< E.F > = < F_i E_1, F_i E_2, .., F_i E_r, ... > \Rightarrow E \subset S_i$$

$$\mathrm{E} = \mathrm{S}_1 \cap \mathrm{S}_2 \cap \cdots \cap \mathrm{S}_\mathrm{d}$$

General decoding of LRPC codes

Let $y = xG + e$

1. **Syndrome space computation**
   Compute the syndrome vector $H.y^t = s(s_1, \cdots, s_{n-k})$ and the syndrome space $S =< s_1, \cdots, s_{n-k} >$.
2. **Recovering the support E of the error**
   $S_i = F_i^{-1}S$, $E = S_1 \cap S_2 \cap \cdots \cap S_d$,
3. **Recovering the error vector** e Write $e_i (1 \leq i \leq n)$ in the error support as $e_i = \sum_{i=1}^{n} e_{ij} E_j$, solve the system $H.e^t = s$.
4. **Recovering the message** x
   Recover x from the system $xG = y - e$.

Decoding of LRPC

- **Conditions of success**

  - $S =< F.E >\Rightarrow$ rd $\leq$ n-k.

  - possibility that $dim(S) \neq n-k \Rightarrow$ probabilistic decoding with error failure in $q^{-(n-k-rd)}$

  - if $d = 2$ can decode up to $(n-k)/2$ errors.

- **Complexity of decoding** : very fast symbolic matrix inversion $O(m(n-k)^2)$ write the system with unknowns : $e_E = (e_{11}, ..., e_{nr})$: rn unknowns in $GF(q)$, the syndrome s is written in the symbolic basis $\{E_1F_1, ..., E_rF_d\}$, H is written in $h_{ij} = \sum h_{ijk}F_k$, $\rightarrow nr \times m(n-k)$ matrix in $GF(q)$, can do precomputation.

- Decoding Complexity $O(m(n-k)^2)$ op. in $GF(q)$

- **Comparison with Gabidulin codes :** probabilistic, decoding failure, but as fast.

Complexity issues : decoding random rankcodes

Rank syndrome decoding

For cryptography we are interested in difficult problems, in the case of rank metric the problem is :

Definition (Rank Syndrome Decoding problem (RSD))

Instance : a $(n-k) \times n$ matrix H over $GF(q^m)$, a syndrome s in $GF(q^m)^{n-k}$ and an integer w
Question : does there exist $x \in GF(q^m)^n$ such that $H.x^t = s$ and $w_R(x) \leq w$?

Definition (Syndrome Decoding problem (SD))

Instance : an $r \times n$ matrix $H = [h_1, h_2, \ldots, h_n]$ over a field $GF(q)$, a column vector $s \in GF(q)^r$, an integer w
Question : does there exist $x = (x_1, \ldots, x_n) \in GF(q)^n$ of Hamming weight at most w such that $H^t x = \sum_{i=1}^{n} x_i h_i = s$?

Problem SD proven NP-complete by Berlekamp et al. in 1978.

Computational complexity of RSD : solved in 2014 (G.,Zemor 2014)

Definition (embedding strategy)

Let $m \geq n$ and $Q = q^m$. Let $\alpha = (\alpha_1, \ldots \alpha_n)$ be an n-tuple of elements of $GF(Q)$. Define the embedding of $GF(q)^n$ into $GF(Q)^n$

$$\begin{array}{rccc} \psi_\alpha : & GF(q)^n & \rightarrow & GF(Q)^n \\ & x = (x_1, \ldots, x_n) & \mapsto & \mathrm{x} = (x_1\alpha_1, \ldots x_n\alpha_n) \end{array}$$

and for any $GF(q)$-linear code C in $GF(q)^n$, define $\mathcal{C} = \mathcal{C}(C, \alpha)$ as the $GF(Q)$-linear code generated by $\psi_\alpha(C)$, i.e. the set of $GF(Q)$-linear combinations of elements of $\psi_\alpha(C)$.

A randomized reduction

General idea of the embedding :

$$(1,0,0,1,0,1) \rightarrow (\alpha_1,0,0,\alpha_4,0,\alpha_6)$$

Theorem

Let C be a random code over GF(q) and α random, then for convenient m, with a very strong probability :

$$d_H(C) = d_R(\mathcal{C})$$

.

Theorem

If there exists a polynomial time algorithms which solves RSD then P=RP.

Adaptable for Decisional-RSD proble G.,Hauteville,Tillich PQCrypto '16

Best known attacks

There are two types of attacks on the RSD problem :

- Combinatorial attacks
- Algebraic attacks

Depending on type of parameters, the efficiency varies a lot.

Combinatorial attacks

- first attack Chabaud-Stern '96 : basis enumeration
- improvements A.Ourivski and T.Johannson '02
  - Basis enumeration : $\leq (k+r)^3 q^{(r-1)(m-r)+2}$ (amelioration on polynomial part of Chabaud-Stern '96)
  - Coordinates enumeration : $\leq (k+r)^3 r^3 q^{(r-1)(k+1)}$
- improvement : G. et al. '16
  - Support attack : $\mathcal{O}(q^{(r-1)\frac{\lfloor (k+\mathbf{1})m \rfloor}{n}})$
  - improvement Aragon,G.,Hauteville,Tillich ISIT '18 (GRS+) : $\mathcal{O}((nm)^3 q^{r\lceil \frac{km}{n} - m \rceil})$
  - Quantum Speed Up : Grover's algorithm directly applies to GRS+ $\implies$ exponent divided by 2.

Algebraic attacks for rank metric

- Levy-Perret '06 : Taking error support as unknown $\rightarrow$ quadratic setting
- Kipnis-Shamir '99 (FLP '08) and others..) : Kernel attack, $(r+1) \times (r+1)$ minors $\rightarrow$ degree $r+1$
- G. et *al.* '16 : annulator polynomial $\rightarrow$ degree q^r
- Max Minor and Support Minor approaches

- Bardet, Bros, Cabarcas, Gaborit, Perlner, Smith-Tone, Tillich, Verbel : Improvements of Algebraic Attacks for Solving the Rank Decoding and MinRank Problems. ASIACRYPT (1) 2020 : 507-536

- Bardet, Briaud, Bros, Gaborit, Neiger, Ruatta, Tillich : An Algebraic Attack on Rank Metric Code-Based Cryptosystems. EUROCRYPT (3) 2020 : 64-93

$\rightarrow$ Strong impact on parameters , especially when $r = O(\sqrt{(n)})$

Estimator : https ://estimators.crypto.tii.ae/

Attack with q-polynomial

Reformulation :

$$c + e = y$$

with c a word of $\mathcal{C}$, e a word of weight r and y known.
There exists a polynomial P of q-degree r such that

$$P(c - y) = 0$$

moreover there exists x such that $c = xG$, which gives :

$$(\sum_{i=0}^{r} p_i(xG_1 - y_1)^{q^i}, \ldots, \sum_{i=0}^{r} p_i(xG_n - y_n)^{q^i})$$

with $x \in {\mathbb{F}_{q^m}}^k$, G_j the j-ith column of G and $y \in {\mathbb{F}_{q^m}}^n$ known.

Attack with q-polynomials

Advantages : less unknowns, sparse equations
Disadvantages : higher degree equations $q^r + 1$
Attack is linear whenever $(k+1)(r+1) \leq n$

Three methods to solve :

- Linearization
- Grobner basis
- Hybrid approach : partial enumeration of unknowns

The promise of rank metric

The hardest case for attacker corresponds to the case where $m = n$ $k = n/2$

in that case $d_{RGV} \approx 0.29n$

exponential part of best attacks : $q^{kr} = q^{O(n^2)}$

which explains the advantage of rank metric when no structure :

- description of matrix n^3

- $\lambda =$ security value

- $n = \sqrt{\lambda}$, size in $O(\lambda^{\frac{3}{2}})$

when $O(\lambda^2)$ for Hamming metric.

$\rightarrow$ **Potential smaller keysize than HAmming metric**

ENCRYPTION/Key Exchange IN RANK METRIC

- Gabidulin *et al.* '91 : first encryption scheme based on rank metric

- adaptation of McEliece scheme, many variations :

- Parameters
  - B $\{b_1, \cdots, b_m\}$ a basis of $GF(q^m)$ over $GF(q)$
- Private key
  - G generates a Gabidulin code $Gab(m, k), r = \frac{m-k}{2}$
  - S a random $k \times t_1$ matrix in $GF(q^m)$
  - P a random matrix in $GL_{m+t_1}(GF(q))$
  - S a random invertible $k \times k$ matrix in $GF(q^m)$
- Public key

$$G_{pub} = S(G|Z)P$$

- Encryption

$$y = xG_{pub} + e, \quad Rank(e) \leq r$$

- Decryption
  - Compute $yP^{-1} = x(G|Z) + eP^{-1}$
  - Puncture the last t_1 columns and decode

Other variations : G Gabidulin matrix, H : dual matrix

Masking	public matrix	authors
Scrambling matrix	$SG + X$	GPT '91
Right scrambling	$S(G\|Z)P$	Gabi. Ouriv. '01
Subcodes	$\begin{pmatrix} \mathrm{H} \\ \mathrm{A} \end{pmatrix}$	Ber. Loi. '02
Rank Reducible	$\begin{pmatrix} G_1 & 0 \\ \mathrm{A} & G_2 \end{pmatrix}$	[OGHA03],[BL04]
Gabidulin-LRPC	G.H(LRPC)	Loidreau '17

Overbeck's structural attack

Overbeck's attack '06

- general idea : if one consider G=Gab[n,k] and one applies the frobenius : $x \to x^q$ to each coordinate of G then G^q and G have $k-1$ rows in common !
- starting from $G_{pub} = S(G|Z)P$, one can prove there is a rank default in :

$$\begin{pmatrix} G_{pub} \\ \vdots \\ G_{pub}^{q^{n-k-1}} \end{pmatrix}$$

the matrix is a $k(n-k) \times (n+t_1)$ matrix, first n columns part : rank $n-1$ and not n !

- Overbeck uses this point to break parameters of all presented GPT-like systems at that time (generalization G.,Otmani,Tale DCC '18).

The NTRU-like family

- NTRU
  - double circulant matrix $(A|B) \rightarrow (I|H)$
  - A and B : cyclic with 0 and 1, over Z/qZ (small weight) (q=256), $N \sim 300$
- MDPC
  - double circulant matrix $(A|B) \rightarrow (I|H)$
  - A and B : cyclic with 0 and 1, 45 1, (small weight) $N \sim 4500$
- LRPC
  - double circulant matrix $(A|B) \rightarrow (I|H)$
  - A and B : cyclic with small weight (small rank)

Ideal structure

To reduce the memory footprint of a generator matrix, we define ideal 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 :

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

Ideal structure

Definition (Ideal matrix)

Let $P(X)$ a polynomial in $\mathbb{F}_q[X]$ of degree n. A square matrix M of size $n \times n$ is ideal modulo P generated by $f(X)$ when it is of the form :

$$\boldsymbol{M} = \begin{pmatrix} f(X) \bmod P \\ Xf(X) \bmod P \\ \vdots \\ X^{n-1}f(X) \bmod P \end{pmatrix}.$$

Definition (Ideal code)

An ideal code is a code $\mathcal{C}[2n, n]$ having $\boldsymbol{G} = (G_1|G_2)$ as a generator matrix where G_1 and G_2 are two ideal matrices.

Original LRPC scheme [GMRZ13]

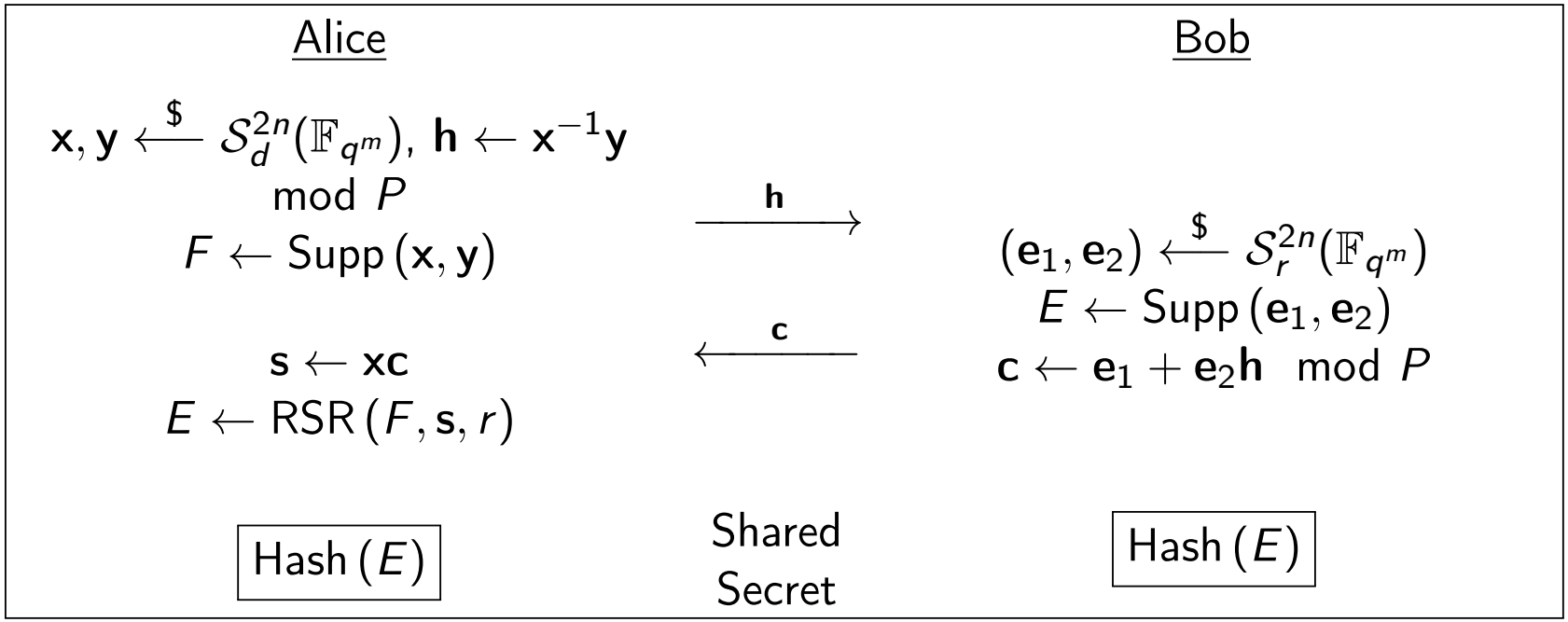

Figure – Informal description of ROLLO-I. $\mathbf{h}$ constitutes the public key.

Original RQC scheme [ABDGZ16]

Vectors $\mathbf{x}$ of $\mathbb{F}_{q^m}^n$ seen as elements of $\mathbb{F}_{q^m}[X]/(P)$ for some polynomial P.

$\mathcal{S}_w^n(\mathbb{F}_{q^m}) = \left\{\mathbf{x} \in \mathbb{F}_{q^m}^n \text{ such that } \omega(\mathbf{x}) = w\right\}$

- Public Data : $\mathbf{G}$ is a generator matrix of some public code $\mathcal{C}$
- Secret key $\mathsf{sk} = (\mathbf{x}, \mathbf{y})$, Public key : $\mathsf{pk} = (\mathbf{h}, \mathbf{s} = \mathbf{x} + \mathbf{h}.\mathbf{y})$

Alice		Bob
$\mathbf{h} \xleftarrow{\$} \mathbb{F}_{q^m}^n$		
$(\mathbf{x}, \mathbf{y}) \xleftarrow{\$} \mathcal{S}_\omega^{2n}(\mathbb{F}_{q^m}), \mathbf{s} \leftarrow \mathbf{x} + \mathbf{h}.\mathbf{y}$	$\xrightarrow{\mathbf{h},\mathbf{s}}$	
		$\mathbf{r}_1, \mathbf{r}_2, \mathbf{e} \xleftarrow{\$} \mathcal{S}_{\omega_r}^{3n}(\mathbb{F}_{q^m})$
		$\mathbf{u} \leftarrow \mathbf{r}_1 + \mathbf{h}.\mathbf{r}_2, \quad \mathbf{v} \leftarrow \mathbf{mG} + \mathbf{s}.\mathbf{r}_2 + \mathbf{e}$
	$\xleftarrow{\mathbf{u},\mathbf{v}}$	
$\mathbf{m} \leftarrow \mathcal{C}.\mathsf{Decode}(\mathbf{v} - \mathbf{u}.\mathbf{y})$		

Adaptation of HQC scheme in rank metric.
No need to mask a code.

Multi-syndrome approach [BBBG23]

The Multi-syndrome approach consists on giving to the decoder several syndromes associated to errors of same support.

Definition : RSL Problem [GHHT17]

Given $(\mathbf{H}, \mathbf{S}) \in \mathbb{F}_{q^m}^{(n-k)\times n} \times \mathbb{F}_{q^m}^{N\times(n-k)}$, the Rank Support Learning Problem RSL(n, k, r, N) asks to compute a subspace $E \subset \mathbb{F}_{q^m}$ of dimension r for which there exists a matrix $\mathbf{V} \in E^{\ell\times n}$ such that $\mathbf{H}\mathbf{V}^\mathsf{T} = \mathbf{S}^\mathsf{T}$.

Increases capacity of decoding when used for cryptography.

Augmented-Gabidulin codes [BBBG23]

Definition : Augmented-Gabidulin code

Let $k \leq n' \leq m < n$ integers. Let $\mathbf{g} = (g_1, ..., g_{n'}) \in \mathbb{F}_{q^m}^{n'}$ an $\mathbb{F}_q$-linearly independent family.
The Gabidulin code $[n, k]_{q^m}$ is defined by :

$$\mathcal{G}_{\mathbf{g}}(n, n', k, m) = \left\{(P(g_1), ..., P(g_{n'}), 0, ..., 0) \mid \deg_q(P) < k\right\}$$

Advantage :
We immediatly deduce from the $n - n'$ last coordinates a part of the support (called the support erasure). If it has dimension ε , the decoding capacity is equal to $\left\lfloor \frac{n-k+\varepsilon}{2} \right\rfloor$.

Blockwise errors [SZHW23]

Definition : Blockwise ℓ-error

Let $\mathbf{n} = (n_1, ..., n_\ell) \in \mathbb{N}^\ell$, $\mathbf{r} = (r_1, ..., r_\ell) \in \mathbb{N}^\ell$ and $n \stackrel{\text{def}}{=} \sum_{i=1}^{\ell} n_i$.
An error $\mathbf{e} \in \mathbb{F}_{q^m}^n$ is said to be an ℓ-error with parameters $\mathbf{n}$ and $\mathbf{r}$ if $\mathbf{e} = (\mathbf{e}_1 \mid \cdots \mid \mathbf{e}_\ell)$ such that :

- for all $i \in \{1, ..., \ell\}$, $\|\mathbf{e}_i\| = r_i$,
- for all $i \neq j$, $\mathrm{Supp}(\mathbf{e}_i) \cap \mathrm{Supp}(\mathbf{e}_j) = \{0\}$.

We talk about a homogeneous error in the case of 1-error, i.e. standard error, close to SumRank metric.

Definition : ℓ-LRPC code

An ℓ-LRPC code is a code such that its parity check matrix $\mathbf{H} = (\mathbf{H}_1 \mid \cdots \mid \mathbf{H}_\ell) \in \mathbb{F}_{q^m}^{(n-k)\times n}$ is such that :

- $\mathbf{H}_i \in \mathbb{F}_{q^m}^{(n-k)\times n_i}$ has its coefficients in a subspace F_i, with $d_i = \dim F_i$.
- For $i \neq j$: $F_i \cap F_j = \{0\}$.

If $\mathbf{e} = (\mathbf{e}_1 \mid \cdots \mid \mathbf{e}_\ell)$ is an ℓ-error, then :

$$\mathbf{s}^\mathsf{T} = \mathbf{H}\mathbf{e}^\mathsf{T} = \sum_{i=1}^{\ell} \mathbf{H}_i \mathbf{e}_i^\mathsf{T}$$

Advantage : For a 2-LRPC code $[2n, n]$, with $r_1 = r_2 = \frac{r}{2}$ and $d_1 = d_2 = \frac{d}{2}$.

- With 2-errors : syndrome of weight $r_1 d_1 + r_2 d_2 = \frac{rd}{2}$
- Classical LRPC with homogeneous error of weight r and LRPC code of dual weight d : syndrome of weight rd

Have block errors of smaller weight has a strong impact on decoding capacity.

Description of RQC-MS-AG scheme[ABDGV24]

$\mathcal{S}^{(n_2,n_2)}_{(r_x,r_y)}(\mathbb{F}_{q^m})$: set of 2-errors of size (n_2, n_2) and weight (r_x, r_y)

- Public Data : $\mathbf{G}$ is a generator matrix of some public Augmented-Gabidulin code $\mathcal{G}$
- Secret key $\mathsf{sk} = (\mathbf{x}, \mathbf{y})$, Public key : $\mathsf{pk} = (\mathbf{h}, \mathbf{s} = \mathbf{x} + \mathbf{h}.\mathbf{y})$

Alice		Bob
$\mathbf{h} \xleftarrow{\$} \mathbb{F}_{q^m}^n$		
$(\mathbf{x}, \mathbf{y}) \xleftarrow{\$} \mathcal{S}^{(n_2,n_2)}_{(r_x,r_y)}(\mathbb{F}_{q^m}), \mathbf{s} \leftarrow \mathbf{x} + \mathbf{h}.\mathbf{y}$	$\xrightarrow{\mathbf{h},\mathbf{s}}$	$(\mathbf{R}_1, \mathbf{R}_2, \mathbf{E}) \xleftarrow{\$} \mathcal{S}^{n_1 \times 3n_2}_{(r_1,r_2,r_e)}(\mathbb{F}_{q^m})$
		$\mathbf{U} \leftarrow \mathbf{R}_1 + \mathbf{h} \cdot \mathbf{R}_2,$
	$\xleftarrow{\mathbf{U},\mathbf{V}}$	$\mathbf{V} \leftarrow \mathsf{Fold}(\mathbf{mG}) + \mathbf{s} \cdot \mathbf{R}_2 + \mathbf{E}$
$\mathbf{m} \leftarrow \mathcal{G}.\mathsf{Decode}(\mathsf{Unfold}(\mathbf{V} - \mathbf{y} \cdot \mathbf{U}))$		

$$\mathbf{V} - \mathbf{y} \cdot \mathbf{U} = \mathbf{mG} + \text{Unfold}(\mathbf{x} \cdot \mathbf{R}_2 - \mathbf{y} \cdot \mathbf{R}_1 + \mathbf{E})$$

Structure of errors :

- $(\mathbf{x}|\mathbf{y})$: 2-error of weight $\mathbf{r} = (r_x, r_y)$ and size $\mathbf{n} = (n_2, n_2)$.
- $(\mathbf{R}_1|\mathbf{R}_2|\mathbf{E})$: collection of 3-errors of weight $\mathbf{r} = (r_1, r_2, r_e)$ and size $\mathbf{n} = (n_2, n_2, n_2)$.

The weight of error to decode is equal to $r_x r_2 + r_y r_1 + r_e$.

Assume that $r_x = r_y = r_1 = r_2 = r_e = r$.

As blockwise errors, the error to correct would have weight $2r^2 + r$.

If $(\mathbf{x}|\mathbf{y})$ and $(\mathbf{R}_1|\mathbf{R}_2|\mathbf{E})$ were considered as homogeneous errors, their weight would be $2r$ and $3r$, and the error to correct would have weight $6r^2$.

Description of ILRPC-block-MS scheme[ABDGV24]

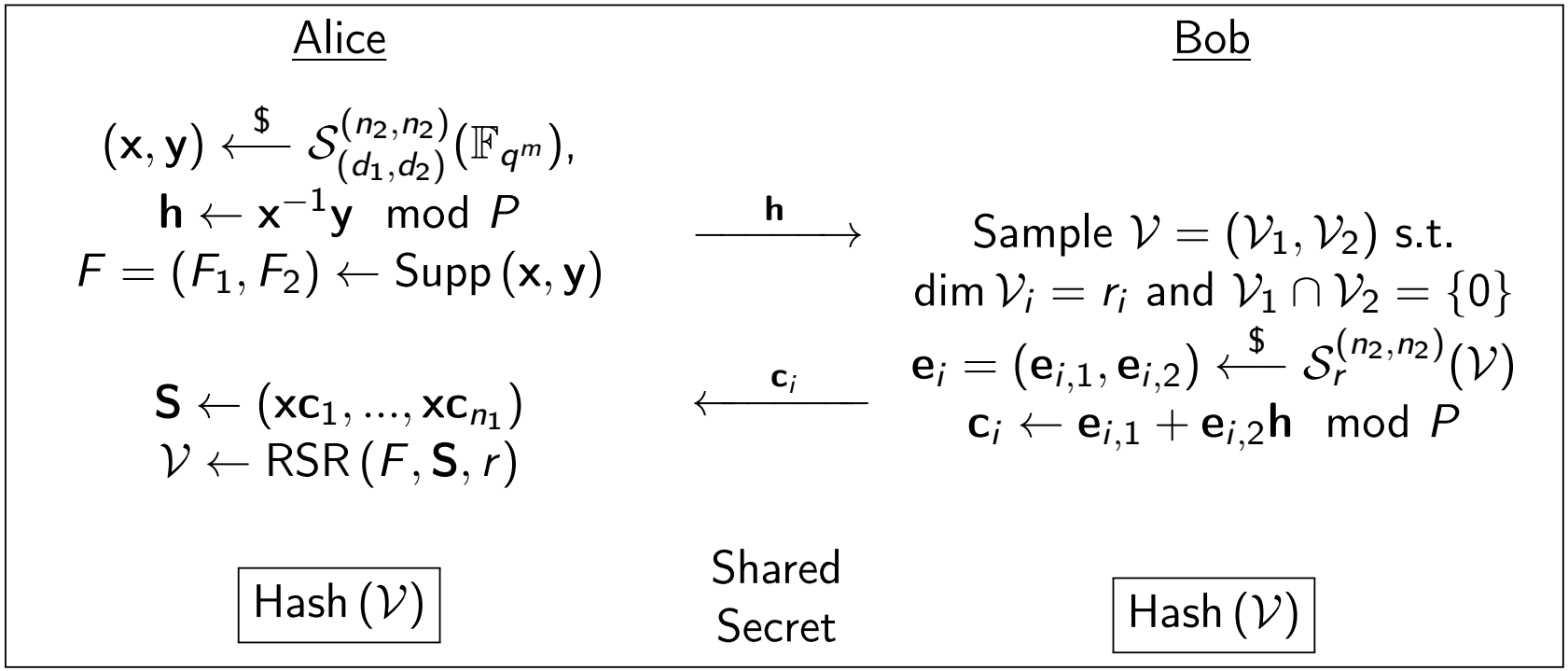

Figure – Informal description of ROLLO-I. $\mathbf{h}$ constitutes the public key.

Comparaison of parameters of different RQC schemes

Scheme	DFR	pk + ct
RQC-Block-MS-AG-128	-145	**1.4 kB**
RQC-Block-128	-	2.5 kB
RQC-NH-MS-AG-128	-158	2.7 kB
RQC-128	-	5.3 kB
RQC-Block-MS-AG-192	-206	**2.8 kB**
RQC-Block-192	-	5.3 kB
RQC-NH-MS-AG-192	-238	4.7 kB
RQC-192	-	8.3 kB

AG : Augmented Gabidulin
MS : Multiple Syndrome
Block : Blockwise errors

Scheme	DFR	pk + ct
ILRPC-Block-xMS-128	-128	1.7 kB
ILRPC-Block-xMS-192	-194	3.3 kB

Figure – Parameters of ILRPC schemes

Scheme	128 bits	192 bits
RQC-Block-MS-AG	**1.4 kB**	**2.8 kB**
KYBER	1.5 kB	2.2 kB
ILRPC-Block-MS	**1.7 kB**	**3.3 kB**
BIKE	3.1 kB	6.2 kB
HQC	6.7 kB	13.5 kB
Classic McEliece	261.2 kB	624.3 kB

Figure – Comparaison of different post-quantum encryption schemes

The sizes represent the sum of the key and the ciphertext.

Signature in rank metric

Different approaches for signature

1. HAsh-and-sign
   1. unique pre-image : RSA,CFS
   2. several pre-images : NTRUSign, GGH, GPV, WAVE, RankSign (broken), Falcon,Miranda
2. Signature by proof of knowledge
   1. by construction : Schnorr, DSA, Lyubashevski, Di-Lithium, Durandal, SPhincs
   2. generic : Fiat-Shamir paradigm
3. one-time signatures : KKS '97, Lyubashevski '07

Fiat-Shamir paradigm

1. Starts from a ZK authentication scheme (cheating proba p) and turns into a signature scheme
2. Idea : the prover replaces the verifier by a hash function.
   1. Fix a security level and t rounds
   2. Prepare t commitments at once $\rightarrow C$
   3. consider a sequence of challenges $b(b_1, ..., b_t)$ as $b \leftarrow hash(C)$
   4. compute the sequence of answers $A = (A_1, A_2, ..., A_t)$
   5. the signature is : (C, A).
   6. verification : the verifier checks that A is the lists of answers regarding C

$\Rightarrow$ usually p is 1/2 or 2/3 hence the signature is long 80$\times$cost of communications, for FS : 160.000b, for Stern 200,000 but it works, small public keys

Evolutions of ZKP

- Many evolutions of zero-knowledge proofs in codes :
  - Stern protocol soundness error of $\frac{2}{3}$, uses permutations [Stern94] ;
  - AGS protocol improvement of Stern, $\frac{1}{2}$ [AGS11] ;
  - Gaborit,Schrek,Zemor RSDc version of Stern protocol (2011)
  - Shared Permutation : soundness error down to $\frac{1}{N} \rightarrow$ now depends of a chosen parameter [FJR21] ;
  - Protocol with helper : $\frac{1}{N}$, more efficient [BG23] ;
  - MPC-in-the-Head : Additive secret sharing, $\frac{1}{N}$ too but more efficient [FJR22] ;
  - Threshold-Computation-in-the-Head and VOLE-in-the-Head : Shamir secret sharings, $\frac{1}{N}$ much more efficient [FR24], [VOLEitH].

• Before MPC : ZKP depended on properties of given problem

- Fiat-Shamir : $a \to ka \pmod N$ for masking

- Stern : $e \to e + y$ (random), $\to \sigma(e)$

$\sigma(e + y) = \sigma(e) + \sigma(y)$

• with MPC : abstraction of proof : prove a certain mutivariate relation, the higher the degree, more expensive the cost.

• Best case : quadratic relations

• new goal : consider your favorite problem and find a modeling which describes the most efficiently you problem in terms of degree and size of data

• the best modeling may vary with the MPCith approach and also for a given type of parameters

Several modelings for RSD and MinRank

- q-polynomials ;
- Kipnis-Shamir
- Rank decomposition ;
- New modeling : dual support decomposition.

- Crucial to keep a degree 2 modeling $\rightarrow$ bigger sizes otherwise.

q-polynomial

A q-polynomial of q-degree r is a polynomial in $\mathbb{F}_{q^m}[X]$ of the form :

$$P(X) = X^{q^r} + \sum_{i=0}^{r-1} p_i \cdot X^{q^i} \qquad \text{with } p_i \in \mathbb{F}_{q^m}.$$

- $\boldsymbol{H}\boldsymbol{x}^\top = \boldsymbol{y}^\top$ and $\forall i\ , P_{\boldsymbol{x}}(x_i) = 0.$
  - For RSD : send $\boldsymbol{x}_B \in \mathbb{F}_{q^m}^k$, $P_{\boldsymbol{x}} \rightarrow \mathbb{F}_{q^m}^r$;
  - For MinRank : send $\boldsymbol{x} \in \mathbb{F}_q^k$, $P_{\boldsymbol{E}} \rightarrow \mathbb{F}_{q^m}^r$.

Rank decomposition

- $\boldsymbol{H}\boldsymbol{x}^\top = \boldsymbol{y}^\top$ and $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{R}$,
- with $\boldsymbol{T} \in \mathbb{F}_q^{m \times r}$ and $\boldsymbol{R} \in \mathbb{F}_q^{r \times n}$.
- $\boldsymbol{E} = \boldsymbol{M} + \sum_{i=1}^{k} x_i \boldsymbol{M}_i$.
- $\boldsymbol{E} = \boldsymbol{T}\boldsymbol{R}$.
  - For RSD : send $\boldsymbol{x}_B \in \mathbb{F}_{q^m}^k$, $\boldsymbol{T} \in \mathbb{F}_q^{m \times r}$, $\boldsymbol{R} \in \mathbb{F}_q^{r \times n}$;
  - For MinRank : send $\boldsymbol{x} \in \mathbb{F}_q^k$, $\boldsymbol{T} \in \mathbb{F}_q^{m \times r}$, $\boldsymbol{R} \in \mathbb{F}_q^{r \times n}$.

Dual support decomposition

- New modeling introduced at Asiacrypt 2024 (BFGNR).
- For RSD : improvement of the Rank Decomposition modeling, Shamir's secret sharing $\rightarrow$ easier multiplications.
- Inputs :
  - $\mathrm{Supp}(\boldsymbol{x}) = \langle 1, x_2, \ldots, x_r \rangle$;
  - $\boldsymbol{C} \in \mathbb{F}_q^{r \times (n-r)}$ such that $(1, x_2, \ldots, x_r) \cdot \begin{pmatrix} \boldsymbol{I_r} & \boldsymbol{C} \end{pmatrix} = (1, x_2, \ldots, x_n) = \boldsymbol{x}$.
- $\boldsymbol{H}\boldsymbol{x}^\top = \boldsymbol{y}^\top$ where $\boldsymbol{x} := (1, x_2, \ldots, x_r) \cdot \begin{pmatrix} \boldsymbol{I_r} & \boldsymbol{C} \end{pmatrix}$.
- But here : no need to transmit $\boldsymbol{x}$ because of polynomial constraints and Shamir's sharing.

Comparison of the modelings

Modeling	**Witness size**	**Parameters for $\lambda = 128$**	
		(q, m, n, k, r)	**Size**
Rank Decomposition	$[\,km + (r-1)m + r(n-r)\,] \cdot \log_2(q)$	$(2, 31, 33, 15, 10)$	122 B
q-polynomial	$[\,km + (r-1)m\,] \cdot \log_2(q)$	$(2, 31, 33, 15, 10)$	93 B
Kipnis-Shamir	$[\,km + (r-1)(n-r)\,] \cdot \log_2(q)$	$(2, 31, 33, 15, 10)$	86 B
Dual Support Decomp.	$[\,(r-1)m + r(n-r)\,] \cdot \log_2(q)$	$(2, 53, 53, 45, 4)$	45 B

Table – Witness size for the RSD problem.

Parameters and performances

Security	Trade-off	Framework	τ	Signature	Estimated time (MCycles)
NIST I	Short	TCitH	12	2 937 B	16.0
		VOLEitH	11	2 851 B	14.9
	Fast	TCitH	20	3 708 B	5.0
		VOLEitH	16	3 450 B	2.7
NIST III	Short	TCitH	18	6 713 B	54.3
		VOLEitH	16	6 566 B	40.6
	Fast	TCitH	30	8 454 B	33.3
		VOLEitH	24	8 207 B	8.0
NIST V	Short	TCitH	25	12 371 B	79.8
		VOLEitH	22	12 682 B	50.1
	Fast	TCitH	39	14 926 B	60.8
		VOLEitH	32	14 768 B	11.8

Table – Parameters and performance - RSD

- Will be used for RYDE - 2nd round.

NIST competition

- RYDE, MIRATH went to 2nd round, not third.
- the case of MEDS : an equivalent of LESS in rank metric based on code equivalence problem.
- Comparison to Sphincs,
- Fast, secure and ... resilient
- comparison to Di-Lithium, MPC and lattices

The MinRank case

Matrix codes

Definition : Matrix code

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

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

$$CM(C) = \{M(\boldsymbol{x}) : \boldsymbol{x} \in C\}.$$

Encoding and Decoding in matrix codes

Let C 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 t 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 (proven NPC by N. Courtois in 2001).

Dual of a matrix code

Given a $[m \times n, k]_q$-code $\mathcal{C}$, its *dual* is defined as

$$\mathcal{C}^{\perp} \stackrel{\text{def}}{=} \left\{ B \in \mathbb{F}_q^{m \times n} : \ \forall C \in \mathcal{C}, \ \mathit{Trace}\left(CB^{\top}\right) = 0 \right\}.$$

It defines a $[m \times n, mn - k]_q$-code.
If $B_1, \ldots, B_{mn-k}$ denotes a basis of $\mathcal{C}^{\perp}$, then

$$\mathcal{C} = \left\{ C \in \mathbb{F}_q^{m \times n} : \ \forall i \in [1, mn - k], \ \left(CB_i^{\top}\right) = 0 \right\}.$$

Relation between Rank Codes and Matrix-codes

• A rank codes can be trivially turned into a matrix code, by turning all vectors into matrices

• Rank codes can be seen as structured matrix-codes through the action of the F_{q^m}-linearity.

• for scalar product, turn $m \times n$ matrices into code of length $mn \rightarrow$ usual scalar product

• Description size for Matrix-codes are multiplied by a factor m, naturally bigger but less structured

• Decoding : Gabidulin codes can be considered in a matrix code form and used to decode matrix codes.

MinRank and cryptography

- First used for attacks on MQ
- First scheme : ZK signature Courtois AC '01
- Improvements of this approach led to MIRATH = merge MIRITH/MIRA (2nd round NIST 2023) 3kB signature
- Other approach : MEDS (NIST 1st round 2023) based on Matrix-Code equivalence.
- Recently (2024) used to attack McEliece

Potential Promises of MinRank

- Trapdoor Encryption : constant size advantage in parameters compared to McEliece, greater than Rank Codes
- ZK Signature : no advantage compared to Rank codes or Hamming except constant , signature $\sim 3kB$
- FDH Signature : in codes only WAVE (GPV-like with rejection sampling), no very small size signature
- Good potential for encryption without trapdoor only Aleknovich or Rank-PKE

Trapdoor Encryption with MinRank (à la McEliece)

Nicolas Aragon, Alain Couvreur, Victor Dyseryn, Philippe Gaborit, Adrien Vincotte : MinRank Gabidulin Encryption Scheme on Matrix Codes. ASIACRYPT (4) 2024

• Gabidulin codes have a strong potential not fully exploited with rank codes because of attacks derived from F_{q^m}-linearity

• **idea** : consider Gabidulin codes in matrix-code form

$\rightarrow$ increasing of public key size but potential to have **very small** ciphertext size like MacEliece scheme

• Limited analogy : Goppa codes are binary codes derived from codes over an extension F_{2^m} (Reed-Solomon codes).

Main Idea : Random Rows and Columns Matrix Code transformation

Let $\mathcal{B} = (A_1, ..., A_K)$ a basis a matrix C of size $m \times n$ and dimension K. How we propose to hide :

- Add l_1 rows and l_2 columns of random coefficients : represented by matrices $\mathcal{R}_i$, $\mathcal{R}'_i$ and $\mathcal{R}''_i$.
- Scrambler matrices : multiply by invertible matrices P and Q.
- Give a random basis B' of the hidden matrix code

$$\mathcal{B}' = \left(P \left(\begin{array}{c|c} A_1 & R_1 \\ \hline R'_1 & R''_1 \end{array} \right) Q, \ldots, P \left(\begin{array}{c|c} A_K & R_K \\ \hline R'_K & R''_K \end{array} \right) Q \right).$$

Trapdoor : relies on MinRank and Code Equivalence problems.

Application with Gabidulin codes

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 (small increase needed from recent attacks)

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

Code-based FDH signatures

Consider the one-way function :

$$f : E \in \mathcal{S}_t \longrightarrow eH^\top.$$

Need to find parity-check matrices $H \in \mathbb{F}_2^{(n-k)\times n}$ such that the solution $\vec{e}$ of smallest weight of the equation

$$eH^\top = s$$

could be found for a non-negligible proportion of all $s \in \mathbb{F}_2^{n-k}$.

Density of codes

Density $= \frac{\text{Number of decodable words}}{\text{Total number of words}}$ = probability that a random element is decodable.

- Goppa Codes $[2^m, 2^m - mt, 2t+1]$, density $\sim \frac{1}{t!}$
- Reed-Solomon codes $[n, k, n-k+1]$, $t = \frac{n-k}{2}$, density $\sim \frac{1}{t!}$
- Gabidulin codes $[m, k, m-k+1]$ over F_{q^m}, $r = \frac{m-k}{2}$, density $\sim q^{-r^2}$

Details Gabidulin codes : Number of syndromes : $q^{m(m-k)}$
Number of decodable codewords= Number of supports $q^{(m-r)r}$, number of coordinates q^{mr}

MIRANDA : new MinRank FDH approach

A. Couvreur, T. Debris-Alazard, P. Gaborit, A. Vincotte, to appear AsiaCrypt 2026

Take advantage of the best of both CFS and GPV approaches without their drawbacks.
Initial Idea : Consider a code $\mathcal{C}$ with a strong density that we know how to decode. Then add a random code to consider :

$$\mathcal{D} \stackrel{\text{def}}{=} \mathcal{C} + \mathcal{A}.$$

Remark : It is possible to decode in $\mathcal{D}$ a random element of the space with the same density than decoding $\mathcal{C}$: just decode with $\mathcal{C}$ without considering $\mathcal{A}$.

$\rightarrow$ we will use this masking remark with a matrix-code version of GAbidulin codes

Add-and-Remove matrix-codes transformation

Let $\mathcal{F}$ be a set of $[m \times n, k]_q$-codes. The Add-and-Remove construction AddRemove $(\mathcal{F}, \ell_a, \ell_s)$ is a family of codes defined as follows :

- Let $\mathcal{C} \in \mathcal{F}$
- Let $\mathcal{C}_s$ be a subcode of $\mathcal{C}$ of codimension ℓ_s
- Let $\mathcal{A}$ be an $[m \times n, \ell_a]_q$-code such that $\mathcal{C} \cap \mathcal{A} = \{\vec{0}_{m \times n}\}$.
- We define the 'Add-and-Remove" code $\mathcal{D}$:

$$\mathcal{D} \stackrel{\text{def}}{=} \mathcal{C}_s \oplus \mathcal{A} \ .$$

It defines a $[m \times n, k - \ell_s + \ell_a]_q$-code.

Decoding capacity : $r = \lfloor \frac{m-k}{2} \rfloor$.

m	k	r	ℓ_a	Dens.	Forge	Struc.	Dist.	σ (B)	pk (B)
79	71	4	189	-16	145	143	-87	55	346 K
113	107	3	220	-11	143	143	-105	59	734 K
116	110	3	225	-9	145	144	-109	60	792 K
127	121	3	250	-9	154	158	-121	64	1032 K
223	219	2	330	-4	147	143	-163	73	3493 K

Table – $\lambda = 128$, $\omega = 2.8$

Non optimized signature time : 200ms - 1s depending on parameters.

Comparison with other schemes

Scheme	σ	pk
miranda	60 B	792 kB
UOV	96 B	412 kB
Mayo	186 B	4912 B
Falcon	666 B	897 B
Wave	822 B	3677 kB

Table – Comparison of different schemes for 128 bits of security

Encryption without trapdoor à la Aleknovich for matrix codes - T. Debris,P. Gaborit, R. Neveu, O. Ruatta EuroCrypt 2026

• To obtain a result a la Aleknovich : we want a situation where we want to use the fact that a sequence of $e =< x, y >$ is small for x and y small for a given scalar product

- **Hamming** : possible to extract a bias of probability for the smallness of e for **one** position
- **Euclidean** : possible to extract a bias for e for **one** position e small compared to size of field
- **Rank metric** (Rank PKE) : e $\in F_{q^m}$ not possible for one position $\rightarrow$ necessity to introduce correlated errors $\rightarrow$ Rank Support Learning problem but (n,1) and $(\sqrt{n}, \sqrt{n})$ are possible , relatively compatible with the scalar product. .
- **Matrix Codes and rank metric** : not directly possible, not enough information from e (just one bit), does not make sense, not directly compatible with the (binary) scalar product, the case (n,1) is not possible.

New approach, main idea : a curious theorem !

In our case : objects $n \times n$ matrices with ONE bit scalar product. How to link these (ONE bit) scalar products together to get a small rank matrix to apply Aleknovich ? ?

- Sample $V_1 \subset \mathbb{F}_q^m$ and $V_2 \subset \mathbb{F}_q^n$ of dimension r and d
- Sample $\underbrace{E_1, \ldots, E_{\ell_1}}_{\text{Same column space}} \in {V_1}^n$, $\underbrace{F_1^\top, \ldots, F_{\ell_2}^\top}_{\text{Same column space}} \in {V_2}^m$.

$$\underbrace{\begin{bmatrix} \text{Trace}(E_1 {F_1}^\top) & \cdots & \text{Trace}(E_1 {F_{\ell_2}}^\top) \\ \vdots & \ddots & \vdots \\ \text{Trace}(E_{\ell_1} {F_1}^\top) & \cdots & \text{Trace}(E_{\ell_1} {F_{\ell_2}}^\top) \end{bmatrix}}_{\ell_2} \Bigg\} \ell_1$$

Rank of the matrix

- This matrix has a rank $\leq \min(rd, \ell_1, \ell_2)$.

Summary of the scheme MinRankPKE

What it achieves :

- Follows closely Alekhnovich and Regev frameworks $\rightarrow$ Does not hide a structured code
- Gabidulin codes for decoding
- Relies only on MinRank

Good performances (following Alekhnovich, HQC etc) :

Level	Scheme	PK size (Bytes)	CT size (Bytes)
I	Our scheme	14 700	14 158
	FrodoKEM	9 616	9 752

A more in-depth comparison : security assumptions

Scheme	Metric	No ideal structure	No masking of a code	No extension	No large field	pk	ct
Kyber	Euclidean	x	o	o	x	0.8 kB	0.8 kB
HQC	Hamming	x	o	o	o	2.2 kB	4.4 kB
RQC	Rank	x	o	x	o	0.3 kB	1.1 kB
LowMS	Rank	o	x	x	o	4.77 kB	1.14 kB
Loidreau	Rank	o	x	x	o	34.5 kB	1.8 kB
Generalization of Loidreau	Rank	o	x	x	o	9.5 kB	0.94 kB
McEliece	Hamming	o	x	o	o	261 kB	96 B
MinRank Gabidulin	Rank	o	x	o	o	33-78 kB	207-84 B
FrodoKEM	Euclidean	o	o	o	x	9.6 kB	9.7 kB
Multi-UR-AG	Rank	o	o	x	o	4.1 kB	6.9 kB
Injective Rank Trapdoor	Rank	o	o	x	o	203 kB	1663 kB
Alekhnovich	Hamming	o	o	o	o	$\sim 700kB$*	$\sim$ 700kB
MinRankPKE-I	Rank	o	o	o	o	**14.7 kB**	**14.1 kB**

SUMMARY RANK METRIC

- Situation vs Hamming metric can be related to discrete log elliptic curves vs very discrete log Z/pZ
- Rather stabilized situation now that algebraic attacks have been scrutinized
- Encryption : RQC/LRPC smaller alternatives to HQC/BIKE
- Signature : efficient ZK : RYDE, MIRATH
- New direction of research with recent MinRank based crypto, which could permit to fully consider the potential of Gabidulin codes (very small ciphertexts or signatures).
- Side channel attacks need to be more studied
- Still difficulties for advanced encryption, but may be more possible to obtain result (see Aguilar-Dyseryn-Gaborit on "close" result to homomorphic encryption)
- Funny mathematical stuff ! ! !

QUESTIONS ?