Rank based Cryptography: a credible post-quantum alternative to classical cryptography

Philippe Gaborit, Olivier Ruatta, Julien Schreck, Jean-Pierre Tillich, Gilles Zémor

Rank based cryptography : a credible post-quantum alternative to classical cryptography

- submission to NIST Workshop on Cybersecurity in a Post - Quantum World 2015 -

Philippe Gaborit (contact author) * Olivier Ruatta † Julien Schrek ‡ Jean-Pierre Tillich §
Gilles Zémor¶

Abstract

In this paper we present rank-based cryptography and discuss its interest relatively to other post-quantum families of cryptosystems like for instance code (Hamming)-based cryptography or lattice (Euclidean)-based cryptography. We also survey recent interesting results for this metric and show that rank based cryptography has many advantages and is a credible alternative to classical cryptography, when not so many alternatives are possible. In particular we present recent cryptosystems like the encryption LRPC scheme or the RankSign signature algorithm which are both fast algorithm with small size of key of only a few hundred bits. We also consider the security of the general cryptographic problem, the Rank Decoding problem and highlights upcoming results on its security both theoretical and practical.

Keys words: Public key cryptosystem, rank metric, error-correcting codes

*XLIM, Université de Limoges, 123,av. A. Thomas, 87000 Limoges, France. **Email: gaborit@unilim.fr**

†XLIM, Université de Limoges, 123,av. A. Thomas, 87000 Limoges, France. **Email: ruattat@unilim.fr**

‡TelecomBrest, 29000 Brest, France. **Email: julien.schrek@laposte.net**

§INRIA Rocquencourt, projet SECRET, 78153 Le Chesnay, France. **Email: jean-pierre.tillich@inria.fr**

¶Université Bordeaux, 351, av. de la Libération, 33405 Talence, France. **Email: Gilles.Zemor@math.u-bordeaux1.fr**

1 Introduction

In recent year there has been a burst of activities regarding post-quantum cryptography, the interest of such a field has become even more obvious since the recent attacks on the discrete logarithm problem in small characteristic [2]. These attacks show that finding new attacks on classical cryptographic systems is always a possibility and that it is important not have all its eggs in the same basket.

Among potential candidate for alternative cryptography, lattice-based and code-based cryptography are strong candidates. Rank-based cryptography relies on the difficulty of decoding error-correcting codes embedded with the rank metric (codes over extension fields of type $GF(q^m)$), when code-based cryptography relies on difficult problems related to error-correcting codes embedded with the Hamming metric (often over small fields $GF(q)$) and when lattice-based cryptography is mainly based on the study of q-ary lattices, which can be seen as codes over rings of type Z/qZ (for large q), embedded with the Euclidean distance.

1.1 A unified presentation for post-quantum cryptography

Before going into more details on rank-based cryptography, let us try to have a more general point of view on problems involved in post-quantum cryptography: lattice-based cryptography, code-based cryptography or multivariate based cryptography.

Consider H a random $(n-k) \times n$ matrix over a cer-

tain structure K which can be a field or a ring for instance. Consider also a syndrome s, a random element of K^{n-k} and suppose one wants to solve the following prolem:

Is it easy to find x in K^n such that $H.x^t = s$?

The answer is generally 'yes', indeed in order to find such an x there may be several approaches. A first approach consists in solving the set of equations induced by the problem $H.x^t = s$ with $x(x_1, \cdots, x_n)$ as unknowns, in general since H is random there will be solution with a very strong probability. Another approach would consist in picking a random $(n-k) \times (n-k)$ submatrix H' of H by picking $n-k$ random columns of H and then try to invert H', which is invertible with a good probability since H is random. Then a particular solution x' can be computed as $x' = H'^{-1}$ and zeros are added to complete x' (which has length $n-k$) into x of length n.

Hence we see that solving this problem is an easy task which can be done in polynomial time in n. Now what happens when we slightly modify the problem ? We can do this for instance in two different ways. A first way consists in adding some constraints: suppose we want not only $H.x^t = s$ but also that x has a weight constraint, like for instance x has a small weight (the distance of x to the null vector) for a particular distance over K^n, like for instance the Hamming distance, the Euclidean distance or the rank distance. In that case the problem becomes more difficult since the approaches we described previously do not deal well with weight constraints and in fact for the three cited distance the problem becomes NP-hard, and depending on the distance considered we obtain the main problems on which are based lattice-based cryptography (Euclidean distance), code-based cryptography (Hamming distance) and rank-based cryptography (rank distance).

Another way to modify the problem would be to consider not linear equations as in $H.x^t = s$ but quadratic equations, then again the problem becomes difficult to solve and corresponds to the main problem used for multivariate cryptography the MQ problem.

Hence we saw how the main problems used for post-quantum cryptography can be seen as simple variations on the same unique simple problem: solving a linear system.

The point of view proposed here, also shows that lattice-based cryptography and code-based cryptography problems can be seen as particular cases of some distance based cryptography problems depending on which distance is used. In that sense rank-based cryptography which solely consists in considering another type of distance, the rank distance, is at the same level that lattice-based and code-based cryptography. Eventually the only meaningful differences between these types of cryptographical problems are the properties induced by the considered distances (Euclidean, Hamming or Rank), which have their advantages and drawbacks. In particular the main difference between distances is that in the case of Hamming or rank distance the distance between two words is bounded by the length of a code, when for Euclidean distance there is no limitation and if one considers a very large alphabet it is possible to obtain distance exponential in the dimension of a lattice. This latter property of the distance induces all the different and subtile possible type of approximations, which are not possible with rank or Hamming distance and makes the variety of lattice-based cryptography. In the last section of this paper we will consider some advantages and drawbacks of rank-based cryptography.

1.2 Historical background on rank metric

Besides the notion of Hamming distance for error-correcting codes and the notion of Euclidean distance for lattices, the concept of rank metric was introduced in 1951 by Loo-Keng Hua [25] as an "arithmetic distance" for matrices over a field $\mathbb{F}_q$. Given two $n \times n$ matrices A and B over a finite field $\mathbb{F}_q$, the rank distance between A and B is defined as $d_R(A, B) = Rank(A - B)$. In 1978, Delsarte defined [9] the notion of rank distance on the set of bilinear forms (which can also be seen as the set of rectangular matrices). He derived a Singleton-type bound for these codes and gave a construction of optimal *matrix codes* meeting this bound. A matrix code over $\mathbb{F}_q$ for the rank metric is defined as the set of $\mathbb{F}_q$-linear combinations of a set $\mathcal{M}$ of $m \times n$ matrices over $\mathbb{F}_q$. Such codes are linear over $\mathbb{F}_q$ and the number k of independent matrices in $\mathcal{M}$, is bounded from above

by nm. Then in 1985, Gabidulin introduced in [13] the notion of rank codes in *vector representation* (as opposed to *matrix* representation) over an extension field $\mathbb{F}_Q$ of $\mathbb{F}_q$ (for $Q = q^m$). A rank code $C[n, k]$ of length n and dimension k over $\mathbb{F}_Q$ in vector representation is defined as a subspace over $\mathbb{F}_Q$ of dimension k of $\mathbb{F}_Q^n$.

Since then, rank metric codes have been used in many applications: for coding theory and space-time codes and also for cryptography. Until now the main tool for rank based cryptography relied on masking the Gabidulin codes [14] in different ways and using the McEliece (or Niederreiter) setting with these codes. Meanwhile most of the systems were broken by using structural attacks which use the particular structure of the Gabidulin codes ([29], [12], [5], [23] - see also [12]). A similar situation exists in the Hamming case for which most of cryptosystems based on the Reed-Solomon have been broken for the same reason: the Reed-Solomon are so structured that their structure is difficult to mask and there is always structural information leaking.

Recently a new family of decodable codes in rank metric was proposed in [15], namely the LRPC codes. The basic idea to decode these codes is similar to the NTRU decryption method [21] or the LDPC approach and only uses the fact of knowing a dual matrix with low weight in order to decode (or decrypt) the code. The advantage of this approach is that the induced structure is very poor compared to very structured codes like Reed-Solomon or Gabidulin codes. For instance, the NTRU cryptosystem has never really been attacked for almost 20 years (nor has the more recent MDPC cryptosystem [26], an analog of the NTRU cryptosystem with LDPC codes for Hamming distance and with small public keys of $4,800$ bits).

In this paper we first review in Section 2, basic facts about rank metric and the difficulty of decoding rank metric codes, then we consider LRPC codes in Section 3 and their application to cryptography in Section 4, sections 5 and 6 deal with authentication and signature with rank metric, at last section 7 consider the advantages and potential limitations of rank-based cryptography.

2 Background on rank metric codes and cryptography

2.1 Definitions and notation

Let q be a power of a prime p, m an integer and let V_n be a n dimensional vector space over the finite field $\mathrm{GF}(q^m) = F_{q^m}$. Let $\beta = (\beta_1, \ldots, \beta_m)$ be a basis of $GF(q^m)$ over $GF(q)$. Let $\mathcal{F}_i$ be the map from $GF(q^m)$ to $GF(q)$ where $\mathcal{F}_i(x)$ is the i-th coordinate of x in the basis β. To any $v = (v_1, \ldots, v_n)$ in V_n we associate the matrix $\overline{v} \in \mathcal{M}_{m,n}(F_q)$ in which $\overline{v}_{i,j} = \mathcal{F}_i(v_j)$. The rank weight of a vector v can be defined as the rank of the associated matrix $\overline{v}$. If we name this value $\mathrm{rank}(v)$ we can have a distance between two vectors x, y using the formula $\mathrm{rd}(x, y) = \mathrm{rank}(x - y)$. We refer to [24] for more details on codes for the rank distance.

A rank code C of length n and dimension k over $GF(q^m)$ is a subspace of dimension k of $GF(q^m)^n$ embedded with rank metric. The minimum rank distance of the code C is the minimum rank of non-zero vectors of the code. One also considers the usual inner product which allows to define the notion of dual code. An important notion which differs from the Hamming distance, is the notion of support. Let $x = (x_1, x_2, \cdots, x_n) \in GF(q^m)^n$ be a vector of rank r. We denote by $E := < x_1, x_2, \cdots, x_n >$, the $GF(q)$-sub vector space of $GF(q^m)$ generated by the coordinates of x, $\{x_1, x_2, \cdots, x_n\}$. The vector space E is called the **support** of x. In the following, $\mathcal{C}$ is a rank metric code of length n and dimension k over $GF(q^m)$. The matrix G denotes a $k \times n$ generator matrix of $\mathcal{C}$ and H one of its parity check matrix. At last, the notion of **isometry** which in Hamming metric corresponds to the action on the code of $n \times n$ permutation matrices, is replaced for rank metric by the action of $n \times n$ invertible matrices over the base field $GF(q)$.

2.2 Difficult problem for rank-based cryptography

The main problem used for rank codes in the cryptographic context is the generalization of the classical syndrome decoding problem with Hamming distance in the case of rank metric:

Rank syndrome decoding problem (RSD) Let H be a $((n-k) \times n)$ matrix over $GF(q^m)$ with $k \leq n$, $s \in GF(q^m)^k$ and r an integer. The problem is to find x such that $\text{rank}(x) = r$ and $Hx^t = s$

The computational complexity of the RSD problem has been an open problem for more than 30 years and has been proven NP-hard with a randomized reduction in [19]. The general idea of the approach is to plunge a random binary code (for the Hamming distance) into a rank code over a larger field K by multiplying all the columns of the binary code by random elements of K. The proof then shows that decoding the plunged code for rank distance permits to decode (with a very good probability) the binary code.

2.3 Complexity of the rank decoding problem

The complexity of practical attacks grows very fast with the size of parameters, there is a structural reason for this: for Hamming distance a key notion in the attacks is counting the number of words of length n and support size t, which corresponds to the notion of Newton binomial coefficient $\binom{n}{t}$, whose value is exponential and upper bounded by 2^n. In the case of rank metric, counting the number of possible supports of size r for a rank code of length n over $GF(q^m)$ corresponds to counting the number of subspaces of dimension r in $GF(q^m)$: **the Gaussian binomial coefficient** of size roughly q^{rm}, whose value is also exponential but with a quadratic term in the exponent.

There exist two types of generic attacks on the problem:

- **combinatorial attacks**: these attacks are usually the best ones for small values of q (typically $q = 2$) and when n and k are not too small, when q increases, the combinatorial aspect makes them less efficient. The first non-trivial attack on the problem was proposed by Chabaud and Stern [6] in 1996, then in 2002 Ourivski and Johannson [28] improved the previous attack and proposed a new attack, meanwhile these two attacks did not take account of the value of n in the exponent. Very recently the two previous attacks were generalized in [18] by Gaborit et *al.* in $(n-k)^3 m^3 q^{(r-1)\lfloor \frac{(k+1)m}{n} \rfloor})$ and took the value of n into account and were used to break some repaired versions of the GPT cryposystem. The idea of the latter new approach is to use the notion of support of a word in rank metric and then applying the classical Information Set Decoding [3] approach with this generalized notion of support.

- **algebraic attacks**: the particular nature of rank metric makes it a natural field for algebraic attacks and solving by Groebner basis, since these attacks are largely independent of the value of q and in some cases may also be largely independent on m. These attacks are usually the most efficient ones when q increases. There exist different type of algebraic equations settings to try to solve a multivariate system with Groebner basis. The algebraic context proposed by Levy and Perret [22] in 2006 considers a quadratic setting over $GF(q)$ by taking as unknowns the support E of the error and the error coordinates regarding E. It is also possible to consider the Kernel attack by [10] and the minor approach [11] which give multivariate equations of degree $r+1$ over $GF(q)$ obtained from minors of matrices At last, more recently the annulator setting by Gaborit et *al.* in [18] (which is valid on certain type of parameters but may not be independent on m) give multivariate sparse equations of degree q^{r+1} but on the large field $GF(q^m)$ rather than on the base field $GF(q)$. The latter attack is based on the notion of q-polynomial [27] and is particularly efficient when r is small. Moreover all these attacks can be declined in an hybrid approach where some unknowns are guessed.

3 Low Rank Parity Check Codes and their decoding

3.1 Definition of Low Rank Parity Check codes

The idea of these codes is to generalize the classical LDPC codes approach for Hamming distance to the rank metric. There is a natural analogy between low density matrices and matrices with low rank.

A Low Rank Parity Check (LRPC) code of rank d, length n and dimension k over $GF(q^m)$ is a code

such that the code has for parity check matrix, a $(n-k) \times n$ matrix $H(h_{ij})$ such that the sub-vector space of $GF(q^m)$ generated by its coefficients h_{ij} has dimension at most d. We call this dimension the weight of H. Denoting F the sub-vector space of $GF(q^m)$ generated by the coefficients h_{ij} of H, we denote by $\{F_1, F_2, \cdots, F_d\}$ one of its basis.

In practice it means that for any $1 \leq i \leq n-k, 1 \leq j \leq n$, there exist $h_{ijl} \in GF(q)$ such that $h_{ij} = \sum_{l=1}^{d} h_{ijl} F_l$. Naturally the LRPC codes can also be considered in a quasi-cyclic context in which the matrix H is double circulant (a concatenation of two circulant matrices), which permits to dramatically decrease the size of the description of the matrix H.

3.2 Decoding algorithm for LRPC codes

The general idea of the algorithm is to use the fact that all coordinates of the parity check matrix H belong to the same vector space F of small dimension d. Given an error e with associated support E of dimension r and its associated syndrome $s = H.e^t$, we take advantage that when r and d are such that $rd \leq n-k$, the $GF(q)$-vector space S generated by the coordinates of the syndrome $S =< s_1, \ldots, s_{n_k} >$ permits to recover, with a strong probability, the whole product space $P =< E.F >$ (generated by the $E_i F_j$, $1 \leq i \leq r, 1 \leq i \leq d$), of the error support E and the LRPC small vector space F. Then knowing the *whole* product space $P =< E.F >= S$ and the space F, allows to recover E by a simple intersection of subspaces of the form: S times the inverse of the elements of a basis of F. Once the support E of the error e is recovered, it is easy to compute the exact value of each coordinate of e by solving a linear system.

Consider a $[n, k]$ LRPC code C of low weight d over $GF(q^m)$, with generator matrix G and dual $(n-k) \times n$ matrix H, such that all the coordinates h_{ij} of H belong to a space F of rank d with basis $\{F_1, \cdots, F_d\}$.

Suppose the received word to be $y = xG + e$ for x and e in $GF(q^m)^n$, and where $e(e_1, \cdots, e_n)$ is the error vector of rank r, which means that for any $1 \leq i \leq n$, $e_i \in E$, a vector space of dimension r with basis (say) $\{E_1, \ldots, E_r\}$. The Fig. 1 describes a general probabilistic algorithm which decodes an error e up to rank distance $r = (n-k)/d$.

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**

   Define $S_i = F_i^{-1} S$, the subspace where all generators of S are multiplied by F_i^{-1}. Compute the support of the error $E = S_1 \cap S_2 \cap \cdots \cap S_d$, and compute a basis $\{E_1, E_2, \cdots, E_r\}$ of E.

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 $H.e^t = s$, where the equations $H.e^t$ and the syndrome coordinates s_i are written as elements of the product space $P =< E.F >$ in the basis $\{F_1 E_1, \cdots, F_1 E_r, \cdots, F_d E_1, \cdots, F_d E_r\}$. The system has nr unknowns (the e_{ij}) in F_q and $(n-k).rd$ equations from the syndrome.

4. **Recovering the message x**

   Recover x from the system $xG = y - e$.

Figure 1: Algorithm 1:a general decoding algorithm for LRPC codes

3.2.1 Correctness, probability of failure and complexity of the decoding

The decoding algorithm is probabilistic since the probability to recover a set of maximal independant elements of the syndrome space is probabilistic and also since there is very small probability that Step 2 recovers a greater space than E, but all these probabilities can be easily evaluated. In term of complexity of decoding it is possible to use a formal description of the matrix H to compute the inversion of the coefficients of the error vector e with only a quadratic complexity (see [15] for details). Overall we have the following

theorem:

[[15]] Let H be a $(n-k) \times n$ dual matrix of a LRPC codes with low rank $d \geq 2$ over $GF(q^m)$, then algorithm 1 decodes a random error e of low rank r such that $rd \leq n-k$, with failure probability $q^{-(n-k+1-rd)}$ and complexity $r^2(4d^2m + n^2)$.

4 Application of LRPC codes to cryptography: the LRPC cryptosystem

4.1 The LRPC cryptosystem

The LRPC cryptosystem consists in applying a McEliece-like or a Niederreiter-like encryption setting to the LRPC family of decodable codes: Figure 2 presents the LRPC cryptosytem in a McEliece setting.

1. **Key creation** Choose a random $[n, k]$ LRPC code C over $GF(q^m)$, with low rank support F of weight d, which corrects errors of rank r and with parity check matrix, a $(n-k) \times n$ matrix H. Let G be a generator matrix of the LRPC code C.

   • **Secret key**: the LRPC $(n-k) \times n$ dual matrix H

   • **Public key**: a $k \times n$ generator matrix G of the LRPC code C

2. **Encryption**

   Translate the information vector M into a word x, choose a random error e of rank r on $GF(q^m)$. The encryption of M is $c = xG + e$.

3. **Decryption**

   Compute the syndrome $s = H.c^t$ and recover the error vector e by decoding s with the LRPC code, then compute $xG = c - e$ and recover x.

Figure 2: The LRPC cryptosystem

The system works for any LRPC code C and considering G in systematic form permits to decrease a little the size of the public key. The case of double circulant LRPC codes (DC-LRPC) is of particular interest since it permits to dramatically decrease the size of the public key: in that case the matrix H can be written $(A|B)$ where A and B are two circulant invertible LRPC matrices of low rank d and G can be written $G = ((A^{-1}B)^t|I)$ (for A^t the transposed matrix of A).

• General parameters of the LRPC cryptosystem:

Writing the matrices of the system in systematic form we obtain:

1. Size of public key (bits): LRPC: $(n-k)kmLog_2(q)$ / DC-LRPC: $\frac{nm}{2}Log_2(q)$
2. Size of secret key (bits): a seed can be used to recover the different parameters
3. Size of message: LRPC: $nmLog_2(q)$ / DC-LRPC: $nmLog_2(q)$
4. Encryption rate: LRPC: $\frac{k}{n}$ / DC-LRPC: $\frac{1}{2}$

For decryption and encryption, the computational cost is dominated by the matrix-vector multiplication and the cost of syndrome computing. In the case of DC-LRPC , one can use the double-circulant structure to improve computations. The cost of a multiplication in the extension field $GF(q^m)$, in binary operations, is $mLog_2(m)Log_2(Log_2(m))$ ([20]). The system in Fig. 2 is presented in a McEliece setting, in that case the size of the message is larger than for the Niederreiter setting but more can be proven regarding semantic security.

4.2 Security of the LRPC cryptosystem

4.2.1 Attacks on the LRPC cryptosystem

There are two type of attacks. The first type of attacks are **direct attacks on the message**, in which the attacker tries to recover directly the message by finding the error e of rank r with classical attacks described in Section 2.3. For the type of considered parameters, the recent combinatorial attacks or algebraic attacks of [15] are the most efficient ones.

It is also possible to consider **structural attacks** and try to attack directly the structure of the public key to recover the secret key. In particular one can use the fact that all the elements of the dual LRPC matrix H belong to the same subspace F of rank d. Let D be the dual code generated by H. Denote by $H_i (1 \leq i \leq n-k)$ the $n-k$ rows of H and consider a word x of D obtained from linear combinations in the small field $GF(q)$: $x = \sum_{i=1}^{n-k} a_i H_i$ for $a_i \in GF(q)$. All the coordinates of x belongs to F, now since F has dimension d, fixing d variables a_i in $GF(q)$ can allow to put to zero a coordinate of x, overall since there are $n-k$ variables a_i one can put to zero (with a good probability depending of the matrix H), $\lfloor (n-k)/d \rfloor$ coordinates positions of x. Therefore with a good probability the dual code D contains a word x with all coordinates in F and with $\lfloor (n-k)/d \rfloor$ coordinates to zero which can be the first $\lfloor (n-k)/d \rfloor$ coordinates without loss of generality. Hence the attacker can attack the subcode D' of D of all the words of D which are zero on the first $\lfloor (n-k)/d \rfloor$ coordinates. This code D' is a $[n - \lfloor (n-k)/d \rfloor, n-k-\lfloor (n-k)/d \rfloor]$ code which, by the previous discussion, contains a word of rank d.

The previous structural attack uses deeply the structure of the LRPC matrix so that the attacker has only to attack a smaller code which contains at least one word of rank d. This exponential attacks slightly reduces the computational cost of the attack on the system and can be easily handled. This attack has an equivalent attack for NTRU [21] and for MDPC codes [26] in which the attacker uses the cyclicity to decrease slightly the number of columns of the attacked matrix: by removing columns corresponding to zeros of a small weight vector of the secret key.

4.3 Examples of parameters

We give three examples of parameters for the DC-LPRC case: an example with security 2^{80} operations which optimizes the size of the public key at 1680 bits with a decryption probability of 2^{-22}, an example with security 2^{128}, and at last an example with decryption failure probability of 2^{-80}.

These parameters update the parameters from [15] after the weak structural attack described in Section 4.4.

In the table 'failure' stands for probability of 'decryption failure', the size of the public key is in bits, the security is in bits. We give parameters for different level of security, but also for different decryption failure, in particular it is possible to reach a 2^{-80} easily at the cost of doubling the size of the key. Notice that the parameters are very versatile. Although no special attack is known for non prime number we choose to consider prime numbers in general. The complexity of decryption for the first set of parameters is 2^{20} bit operations. In particular in terms of computation cost the LRPC cryptosystem seems to compare very well with the MDPC cryptosystem.

n	k	m	q	d	r	failure	public key	security
82	41	41	2	5	4	-22	1681	80
106	53	53	2	6	5	-24	2809	128
74	37	23	2^4	4	4	-88	3404	110

5 Zero-knowledge authentication with rank metric

5.1 Previous work and definitions

In 1995 Chen proposed in [7] a 5-pass zero-knowledge protocol based on rank metric with cheating probability 1/2, the protocol was in the spirit of the Stern SD protocol [30] meanwhile it turned out that the protocol was not correct and was subsequently broken in [17]. The main reason was that the zero-knowledge proof was false, especially since the author failed to construct an equivalent notion of permutation for Hamming distance which would associate any word of rank weight r to any particular given word of rank r. Indeed let x be a word of length n and rank r with support E, then for any $n \times n$ random invertible matrix P in the small field $GF(q)$, the word xP has also rank r but the support of x and xP are the same so that information leaks if one tries to hide x only by turning it into xP.

The definition of the product "$*$" allows to obtain such a property for rank metric. With the notation of Section 2.1: for a given basis β, we denote Φ_β the inverse of the function $V_n \to \mathcal{M}_{m,n}(\mathrm{GF}(q)) : x \to \overline{x}$ computed with the basis β. [product] Let Q be in $\mathcal{M}_{m,m}(\mathrm{GF}(q))$,

$v \in V_n$ and β a basis. We define the product $Q*v$ such that $Q * v = \Phi_\beta(Q\overline{v})$, where $\overline{v}$ is constructed from the basis β.

Then one can prove the following proposition which gives the equivalent notion of permutation for Hamming distance but in a rank metric context:

[[17]] For any $x, y \in V_n$ and $\mathrm{rank}(x) = \mathrm{rank}(y)$, it is possible to find $P \in \mathcal{M}_{n,n}(\mathrm{GF}(q))$ and $Q \in \mathcal{M}_{m,m}(\mathrm{GF}(q))$ such that $x = Q * yP$.

5.2 Description of the protocol

The previous definition of the "*" product permits to obtain a rank metric adaptation of the Stern protocol [30] which was presented in [17]: the masking of a codeword by a permutation is replaced by the masking $x \to Q * xP$ which has the same property in terms of rank distance as a permutation for a codeword with Hamming distance, since it can transform any given x with given rank to *any* element with the same rank. In the following the notation $(a|b)$ corresponds to the concatenation of a and b. The notation $hash(a)$ is the hash value of a. A given basis β is fixed and known in advance for the "*" product.

For the protocol a public $k \times n$ matrix over $GF(q^m)$ H is fixed. The **secret key** is a vector s of $V_n (= (GF(q^m)^n)$ with rank r. The **public key** consists of the matrix H, the syndrome $i = Hs^t$ and the rank r of s. The protocol is described in Fig. 3. For the protocol the small base field is $GF(2)$, (ie: $q = 2$). It is proven in [17] that the protocol described in Fig. 3, is a 3-pass zero-knowledge protocol with cheating probability 2/3.

1. [Commitment step] The prover $\mathcal{P}$ chooses $x \in V_n$, $P \in GL_n(\mathrm{GF}(q))$ and $Q \in GL_m(q)$. He sends c_1, c_2, c_3 such that :

$$c_1 = hash(Q|P|Hx^t), c_2 = hash(Q * xP),$$

$$c_3 = hash(Q * (x + s)P)$$

2. [Challenge step] The verifier $\mathcal{V}$ sends $b \in \{0, 1, 2\}$ to P.

3. [Answer step] there are three possibilities :
   - if $b = 0$, $\mathcal{P}$ reveals x and $(Q|P)$
   - if $b = 1$, $\mathcal{P}$ reveals $x + s$ and $(Q|P)$
   - if $b = 2$, $\mathcal{P}$ reveals $Q * xP$ and $Q * sP$

4. [Verification step] there are three possibilities :
   - if $b = 0$, $\mathcal{V}$ checks c_1 and c_2.
   - if $b = 1$, $\mathcal{V}$ checks c_1 and c_3.
   - if $b = 2$, $\mathcal{V}$ checks c_2 and c_3 and that $rank(Q * sP) = r$.

Figure 3: Rank-SD protocol

5.2.1 Example of parameters

If we consider, $q = 2, n = 22, m = 23$ and $k = 9$ one obtains a minimal distance of 8 by the rank Gilbert-Varshamov bound [24], hence we can take $r = 7$ for the rank weight of the secret. The security of the protocol relies then on the security of the hash function and on a general random instance of the RSD problem defined in Section 2. In that case with these parameters, the best practical known attacks lead to a complexity of at least 2^{80} operations (these parameters are updated from [17] after the recent improvements on generic attacks of [18]). The fact that one can take a rank weight r close to the rank Gilbert-Varshamov bound permits to greatly decrease the size of the parameters.

Public matrix H : $(n - k) \times k \times m = 2691$ bits

Public key i : $(n - k)m = 299$ bits

Secret key s : $r(m + n) = 360$ bits

Average number of bits exchanged in one round: 2 hash + one word of $\mathrm{GF}(q^m) \sim 820$ bits.

Overall the protocol is more efficient than the Stern SD scheme and can probably be optimized as in [1]. For instance by cyclicity: if one considers a double-circulant matrix, the size of the public key decreases to only a few hundred bits: a $[22, 11]$ double circulant code, $q = 2$, $m = 29$, $r = 7$ leads to a public key of 319 bits. A security of 2^{100} can be reached with $k = 9, n = 27, q = 2, m = 24$ and $r = 10$.

6 Signature with rank metric

6.1 High level overview

The traditional approach for decoding random syndromes, that is used by the CFS scheme [8] in Hamming distance for instance, consists in taking advantage of the decoding properties of a code (e.g. a Goppa code) and considering parameters for which the proportion of decodable vectors – the decodable density – is not too low. For the Hamming metric, this approach leads to very flat dual matrices, i.e, codes with high rate and very low Hamming distance. In the rank metric case, this approach leads to very small decodable densities and does not work in practice. However, it is possible to proceed otherwise. It turns out that the decoding algorithm of LRPC codes can be adapted so that it is possible to decode not only errors but also (generalized) erasures. This new decoding algorithm allows us to decode more rank errors since the support is then partially known. In that case since the size of the balls depends directly on the dimension of the support, it leads to a dramatic increase of the size of the decodable balls. Semantically, what happens is that the signer can fix an erasure space, which relaxes the condition for finding a preimage. This approach works because in the particular case of our algorithm, it is possible to consider the erasure space at no cost in terms of error correction: to put it differently, the situation for LRPC codes is different from traditional Hamming metric codes for which "an error equals two erasures".

In practice it is possible to find parameters (not flat at all) for which it is possible to decode a random syndrome with the constraint that its support contains a fixed random subspace. Fixing part of the rank-support of the error, (the generalized erasure) allows us more rank-errors. For suitable parameters, the approach works then as follows: for a given random syndrome-space element s, one chooses a random subspace T of fixed dimension t (a generalized erasure of Definition 2), and the algorithm returns a small rank-weight word, whose rank-support E contains T, and whose syndrome is the given element s. Of course, there is no unicity of the error e since different choices of T lead to different errors e, which implies that the rank of the returned error is above the GVR bound (the Rank Gilbert-Varshamov bound): it is however only just above the GVR bound for the right choice of parameters.

n	n-k	m	q	d	t	r'	pk	sign	sec
16	8	18	2^{40}	2	2	4	57600	8640	**130**
16	8	18	2^{8}	2	2	4	11520	1728	**80**
16	8	18	2^{16}	2	2	4	23040	3456	**120**
20	10	24	2^{8}	2	3	5	24960	3008	**104**
27	9	20	2^{6}	3	2	3	23328	1470	**120**

Table 1: Examples of parameters for the RankSign signature scheme

6.2 The RankSign signature scheme

In the following Figure 4 describes the RankSign signature scheme based on the previous high level overview and Table 1 gives some possible parameters, all the details of the algorithm can be found in [16]. Notice that there is a security reduction which details the unleakibility of the signature depending on parameters of the scheme.

Parameters: *Public key size:* $(k+t)(n-k)mLog_2(q)$ *Signature size:* $(m+n+t)rLog_2(q)$.

7 Advantages and limitations of rank-based cryptography

We saw in this paper that there existed a very efficient encryption algorithm for rank-based cryptography with very small public key comparable to RSA and that it was also possible to obtain efficient signature and zero-knowledge authentication algorithms.

7.1 A semantical advantage

Rank-based cryptography belongs to the larger class of post-quantum cryptosystems, which is an alternative class of cryptosystems which are *a priori* resistant to a putative quantum computer. Indeed for

1. **Secret key:** an augmented LRPC code over $GF(q^m)$ with parity-check matrix $(R|H)$ of size $(n-k) \times (n+t)$ which can decode r' errors and t generalized erasures: a randomly chosen $(n-k) \times (n-k)$ matrix A that is invertible in $GF(q^m)$ a randomly chosen $(n+t) \times (n+t)$ matrix P invertible in $GF(q)$.

2. **Public key:** the matrix $H' = A(R|H)P$, a small integer value l, a hash function $hash$.

3. **Signature of a message M:**

   a) *initialization*: seed $\leftarrow \{0,1\}^l$, pick t random independent elements $(e_1, \cdots, e_t)$ of $GF(q^m)$

   b) *syndrome*: $s \leftarrow hash(M||seed) \in GF(q^m)^{n-k}$

   c) decode by the LRPC matrix H, the syndrome $s' = A^{-1}.s^T - R.(e_1, \cdots, e_t)^T$ with erasure space $T = \langle e_1, \cdots, e_t \rangle$ and r' errors by Algorithm 1.

   d) if the decoding algorithm works and returns a word $(e_{t+1}, \cdots, e_{n+t})$ of weight $r = t + r'$, signature=$((e_1, \cdots, e_{n+t}).(P^T)^{-1}, seed)$, else return to a).

4. **Verification**: Verify that $Rank(e) = r = t + r'$ and $H'.e^T = s = hash(M||seed)$.

Figure 4: The *RankSign* signature algorithm

post-quantum cryptography the security of the cryptosystems is usually related to a NP-hard problem. Of course the notion of NP-hardness being a worst-case reduction, it is a crude fit for cryptography, but for confidence in the intrinsic hardness of a problem it is better than no theoretical result at all as the recent breakthrough on the discrete log problem in small characteristic shows [2]. The particular appeal of rank metric based problems compared to lattices or (Hamming) codes based problems is that the practical complexity of the best known attacks for rank-based problems [18] grows very quickly when compared to their Hamming counterpart [3]. As mentioned earlier in this paper, such attacks have a quadratic term (related to parameters of the rank code) in their exponential coefficient, while for Hamming distance problems (and somehow also for heuristic LLL attacks for problems based on Euclidean distance in lattice-based cryptography), the best practical attacks have only an exponential term whose exponent is linear in the code parameters. This translates into rank codes having a decoding complexity that behaves as $exp(\Omega(N^{2/3}))$ rather than $exp(\Omega(N^{1/2}))$ for Hamming codes, where N is the input size, i.e. the number of q-ary symbols needed to describe the code. In practice it means that it is possible to obtain secure practical parameters (e.g. key sizes) of only a few thousand bits for cryptographic schemes based on the difficulty of decoding generic rank metric codes when at least a hundred thousand bits are needed for Hamming distance codes or for lattices.

This situation is indeed a very important feature of rank based cryptography since for lattice-based or (Hamming) code-based cryptography, in order to have practical cryptosystems one has to add structure anyway, since the smaller size of parameters corresponding to hard practical instances of general instances of a problem (like finding the minimum norm of a lattice or the minimum weight of a code) are of size at least several thousand bits. In practice the difficulty of attacking a particular cryptosystem is related to attacking a structured matrix, like for instance an ideal lattice or a quasi-cyclic code, but in that case nothing is really known on the theoretical difficulty of the problem. For instance the possibility does exist that there would be a quantum algorithm which could address these problems in polynomial time (even a non quantum algorithm is an option), when in the case of completely random instances, since the problems are NP-complete, it is widely believed that finding such an algorithm is unlikely.

In particular is possible to construct cryptographic schemes, like for instance the zero-knowledge authentication scheme of [17], whose security relies on general random instances of the Rank Syndrome Decoding problem.

In terms Hamming and rank distance we can do an analogy with the case of discrete logarithm: best known attacks for solving the discrete logarithm prob-

lem on a field are sub-exponential when best known attacks on a more complex structure like elliptic curves makes the best attacks exponential. In the case of the decoding problem, going from Hamming distance to rank distance increases the complexity from a linear exponential complexity to a quadratic exponential complexity. As explained earlier this fact results from the combinatorial analogy between Newton and Gaussian binomials.

7.2 Practical cryptosystems

In the previous section we described a semantical advantage of rank metric compared to Euclidean distance or Hamming distance, what about the practucality of it ?

As we saw in Section 4 it is possible to construct cryptographic schemes, like for instance the zero-knowledge authentication scheme of [17], whose security relies on general random instances of the Rank Syndrome Decoding problem. But at the difference of lattice-based cryptography there is not yet an equivalent of the LWE algorithm, so that the security of an encryption scheme relies on attacking random instances of a difficult problem, but still if it was to exist it would be a very strong result. It is possible to construct encryption scheme analogous to NTRU like we saw in section 3, but there is an added structure.

In practice though rank-based cryptography proposes very efficent scheme for encryption (like LRPC encryption scheme) relying on trapdoor with low structure (like the NTRU scheme), and efficient schemes for signature and authentication. Moreover in the case of authentication the security relies on attacking random general instances of the Rank Decoding problem.

7.3 Potential limitations

If one compares rank-based cryptography with (Hamming) code based cryptography in term of cryptosystems efficiency the situation is very similar in term of type of available primitives but for rank distance the key sizes seems though a little smaller in general. Of course general attacks for the Hamming case have been studied for a longer time, but since best known combinatorial attacks for rank metric are in the spirit the best known attacks for Hamming distance (like for instance the Information Set Decoding attacks), it would be a surprise that it would be possible to do clearly better that what is done for equivalent attacks for the Hamming case. Indeed in the last 50 years, the improvement on the general decoding problem of random codes for the Hamming case have resulted in an improvement of 15% in the constant of the exponential factor ([3]) (in the asymptotic case moreover). Therefore even if some work still need to be done the situation seems rather positive.

Regarding the comparison with lattices, there are clearly less available primitives, in the present situation it is not possible to obtain IBE or FHE schemes for instance with rank distance. Moreover in term of reduction there is not, for the moment, equivalent notion of the LWE scheme which permits to have reduction to general instance of Euclidean distance problems, neither is there the notion of worst-case to average-case reduction, even if one can question the real interest of this notion since it gives an average case reduction to a clearly weaker problem (an approximation with a polynomial factor), not to mention the case of structured lattices for which no reduction is known besides the worst-case to average-case reduction.

8 Conclusion

In this presentation paper we tried to present rank-based cryptography and especially recent results regarding cryptosystems and computational complexity results. These results show that rank-based cryptography has a strong potential in terms of size of keys because of the inherent difficulty of the RSD problem and its links with the Gaussian binomial which counts subspaces of given dimension. We highlighted the LRPC cryptosystem which benefits from a very low public key of less than $2,000$ bits and is moreover very fast. We also highlighted the analogy between the NTRU cryptosystem, the MDPC cryptosystem and the present LRPC cryptosystem which are based on similar ideas. We also presented the rank-SD authentication scheme and eventually gave the general ideas

on which relies the RankSign signature scheme which has also relatively small public keys.

Overall even if more scrutiny is needed from the communauty, all these recent results propose new promising direction for rank-based cryptography and indicate that rank-based cryptography is a credible candidate for post-quantum cryptography, with still many areas to improve but with also already efficient schemes like for instance the encryption LRPC scheme. In the post-quantum perspective there are not so many alternative and rank-based cryptography seems to be one the very few possible candidates.

References

[1] Carlos Aguilar, Philippe Gaborit and Julien Schrek: A new zero-knowledge code based identification scheme with reduced communication in Information Theory Workshop (ITW), 2011 IEEE, p. 648 - 652

[2] Razvan Barbulescu and Pierrick Gaudry and Antoine Joux and Emmanuel Thomé, "A quasi-polynomial algorithm for discrete logarithm in finite fields of small characteristic", eprint iacr 2013/400

[3] Anja Becker, Antoine Joux, Alexander May, Alexander Meurer: Decoding Random Binary Linear Codes in 2 n/20: How 1 + 1 = 0 Improves Information Set Decoding. EUROCRYPT 2012: 520-536

[4] Thierry P. Berger, Pierre-Louis Cayrel, Philippe Gaborit, Ayoub Otmani: Reducing Key Length of the McEliece Cryptosystem. AFRICACRYPT 2009: 77-97

[5] Thierry P. Berger, Pierre Loidreau: Designing an Efficient and Secure Public-Key Cryptosystem Based on Reducible Rank Codes. INDOCRYPT 2004: 218-229

[6] Florent Chabaud, Jacques Stern: The Cryptographic Security of the Syndrome in Decoding Problem for Rank Distance Codes. ASIACRYPT 1996: 368-381

[7] Kefei Chen: A New Identification Algorithm-Cryptography in book Policy and Algorithms, p. 244-249 (1995).

[8] N. Courtois, M. Finiaz and N. Sendrier : How to achieve a Mc-Eliece-based digital signature scheme. Proceedings of ASIACRYPT'01, LNCS vol 2248, pp. 157-174, 2001.

[9] P. Delsarte, "Bilinear forms overe a finite fied with applications to coding theory", Journal of Comb. Theory A, vol. 25, pp. 226-241, (1978).

[10] J.-C. Faugère, F. Levy-dit-Vehel, L. Perret. Cryptanalysis of MinRank. In *CRYPTO 2008*, LNCS 5157, pages 280–296. Springer Verlag, 2008.

[11] Jean-Charles Faugère, Mohab Safey El Din, Pierre-Jean Spaenlehauer: Computing loci of rank defects of linear matrices using Grbner bases and applications to cryptology. ISSAC 2010: 257-264

[12] Cédric Faure, Pierre Loidreau: A New Public-Key Cryptosystem Based on the Problem of Reconstructing p-Polynomials. WCC 2005: 304-315

[13] Ernst M. Gabidulin, Theory of Codes with Maximum Rank Distance, Probl. Peredachi Inf, (21), pp. 3-16 (1985).

[14] Ernst M. Gabidulin, A. V. Paramonov, O. V. Tretjakov: Ideals over a Non-Commutative Ring and their Applications in Cryptology. EUROCRYPT 1991: 482-489

[15] Philippe Gaborit, Gaétan Murat, Olivier Ruatta and Gilles Zémor: Low Rank Parity Check Codes and their application in cryptography, in the preproceedings of Workshop on Coding and Cryptography (WCC) 2013, Borgen, Norway, p. 167-179.

[16] Philippe Gaborit, Olivier Ruatta, Julien Schrek, Gilles Zmor: RankSign: An Efficient Signature Algorithm Based on the Rank Metric. PQCrypto 2014: 88-107

[17] Philippe Gaborit, Julien Schrek, Gilles Zmor: Full Cryptanalysis of the Chen Identification Protocol. PQCrypto 2011: 35-50

[18] Philippe Gaborit, Olivier Ruatta and Julien Schrek: On the complexity of the rank syndrome decoding problem, eprint (submitted to IEEE trans. Information Theory)

[19] Philippe Gaborit and Gilles Zémor, "On the hardness of the decoding and the minimum distance problems for rank codes", submitted to CCC, available at http://arxiv.org/abs/1404.3482

[20] von zur Gathen, Joachim and Gerhard, Jurgen, "Modern computer algebra", Cambridge University Press, 2003.

[21] Jeffrey Hoffstein, Jill Pipher, Joseph H. Silverman: NTRU: A Ring-Based Public Key Cryptosystem. ANTS 1998: 267-288

[22] F. Levy-dit-Vehel and L. Perret, Algebraic decoding of rank metric codes, proceedings of YACC06.

[23] Pierre Loidreau: Designing a Rank Metric Based McEliece Cryptosystem. PQCrypto 2010: 142-152

[24] P. Loidreau, Properties of codes in rank metric, http://arxiv.org/abs/cs/0610057

[25] H. Loo-Kengn, "A theorem on matrices over a field and its applications" ", Chinese mathematical society, Vol. 1, No. 2, pp. 109-163 (1951).

[26] Rafael Misoczki and Jean-Pierre Tillich and Nicolas Sendrier and Paulo S. L. M. Barreto, MDPC-McEliece: New McEliece Variants from Moderate Density Parity-Check Codes Cryptology ePrint Archive: Report 2012/409

[27] O. Ore, On a special class of polynomials, Trans. American Math. Soc. (1933)

[28] Ourivski, A. V. and Johansson, T., New Technique for Decoding Codes in the Rank Metric and Its Cryptography Applications,Probl. Inf. Transm.(38), 237–246 (2002)

[29] Raphael Overbeck: Structural Attacks for Public Key Cryptosystems based on Gabidulin Codes. J. Cryptology 21(2): 280-301 (2008)

[30] Stern J.: A new paradigm for public key identification *IEEE Transactions on Information THeory* 42 (6) 1996, 2757–2768