

A New Algorithm for Solving the Rank Syndrome Decoding Problem

Nicolas Aragon, Philippe Gaborit, Adrien Hauteville, Jean-Pierre Tillich

▶ **To cite this version:**

Nicolas Aragon, Philippe Gaborit, Adrien Hauteville, Jean-Pierre Tillich. A New Algorithm for Solving the Rank Syndrome Decoding Problem. ISIT 2018 - IEEE International Symposium on Information Theory, Jun 2018, Vail, United States. pp.2421-2425, 10.1109/ISIT.2018.8437464 . hal-01957179

HAL Id: hal-01957179
https://inria.hal.science/hal-01957179

Submitted on 17 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A New Algorithm for Solving the Rank Syndrome Decoding Problem

Nicolas ARAGON[1*], Philippe GABORIT[2*], Adrien HAUTEVILLE[3*], Jean-Pierre TILLICH[4#]
[*]XLIM-DMI, University of Limoges, 123 Avenue Albert Thomas, 87060 Limoges Cedex, France
[#]Inria, 2 rue Simone Iff, 75012 Paris, France
[1]nicolas.aragon@unilim.fr, [2]philippe.gaborit@unilim.fr, [3]adrien.hauteville@unilim.fr,
[4]jean-pierre.tillich@inria.fr

***Abstract*—In this paper, we propose an improvement of the attack on the Rank Syndrome Decoding (RSD) problem found in [1], usually the best attack considered for evaluating the security of rank based cryptosystems. For $\boldsymbol{H}$ a full-rank $(n-k)\times n$ matrix over $\mathbb{F}_{q^m}$ and $\boldsymbol{e}\in\mathbb{F}_{q^m}^n$ of small norm r, the RSD problem consists in recovering $\boldsymbol{e}$ from $s=\boldsymbol{H}\boldsymbol{e}^T$. In our case, the norm of a vector over $\mathbb{F}_{q^m}$ is defined by the dimension of the $\mathbb{F}_q$-subspace generated by its coordinates. This problem is very similar to the Syndrome Decoding problem in the Hamming metric (only the metric and the field of the coefficients are different) and the security of several cryptosystems relies on its hardness, like McEliece-based PKE [2], [3] or IBE [4]. Our attack is in $\mathcal{O}\big((n-k)^3m^3q^{w\left\lceil\frac{(k+1)m}{n}\right\rceil-m}\big)$ operations in $\mathbb{F}_q$ whereas the previous best attacks are in $\mathcal{O}\big((n-k)^3m^3q^{(w-1)\min\left(\left\lceil\frac{(k+1)m}{n}\right\rceil,k+1\right)}\big)$ [1], [5]. In particular in the case $m\leq n$, our attack permits to obtain an exponential gain in $q^{m(1-R)}$ for $R=k/n$ the rate of the code. We give examples of broken parameters for recently proposed cryptosystems based on LRPC codes or Gabidulin codes. Our attack does not fully break these cryptosystems but implies larger parameters for the same security levels.**

I. INTRODUCTION

The hardness of the problem of decoding a linear code makes its use very attractive for post-quantum cryptography. Indeed it has been proven to be NP-complete for the Hamming metric [6]. One of the first cryptosystem partly based on this problem is the McEliece cryptosystem, whose instances using the family of Goppa codes or the family of MDPC codes still remain unbroken today. One of the main drawback of these cryptosystems are the relatively large public key sizes. A solution to decrease their size is to change the metric used for the code.

The rank metric $\mathbb{F}_{q^m}$-linear codes has been introduced in [7] and their first use for the cryptosystem of McEliece was in 1991 [8]. This cryptosystem used the family of Gabidulin codes and a structural attack has been found by Overbeck in [9]. However another family, the Low Rank Parity Check (LRPC) codes, was proposed for the McEliece cryptosystem in [2]. These codes are similar to the MDPC codes in the Hamming metric [10] or the NTRU cryptosystem in the euclidean metric [11] and have a poor algebraic structure, which ensures a good resilience against structural attacks.

Since the security of the McEliece cryptosystem relies directly on the hardness of decoding a random linear code, it is very important to evaluate the complexity of generic algorithms to solve it. One of the first algorithm [5] has a complexity of $\mathcal{O}\big((wm)^3q^{(w-1)(k+1)+2}\big)$ for an error of weight w in a code of length n and of dimension k over the field $\mathbb{F}_{q^m}$. As we can see, this complexity does not take into account the length of the code. The GRS algorithm [1] gives a complexity of $\mathcal{O}\big((n-k)^3m^3q^{(w-1)\left\lceil\frac{(k+1)m}{n}\right\rceil}\big)$ in the case $n\leqslant m$ and this article introduces another algorithm based on the q-polynomials of complexity $\mathcal{O}\big((wm)^3q^{r\left\lceil w\frac{(k+1)(w+1)-n-1}{w}\right\rceil}\big)$. In this paper, we improve the complexity of the GRS algorithm and we obtain a complexity of $\mathcal{O}\big((n-k)^3m^3q^{w\left\lceil\frac{(k+1)m}{n}\right\rceil-m}\big)$.

II. BACKGROUND ON THE RANK METRIC

A. Definitions

Let us begin with the definitions of matrix codes and rank metric.

Definition II.1 (Linear matrix codes)**.** *A linear matrix code $\mathcal{C}$ of length $m\times n$ and dimension K over $\mathbb{F}_q$ is a subspace of dimension K of $\mathcal{M}_{m\times n}(\mathbb{F}_q)$. We say $\mathcal{C}$ is an $[m\times n,K]_q$ linear matrix code, or simply an $[m\times n,K]$ code if there is no ambiguity.*

We define the rank metric over matrix codes as such:

- the distance between two words $\boldsymbol{A}$ and $\boldsymbol{B}$ is $d_R(\boldsymbol{A},\boldsymbol{B})\stackrel{\text{def}}{=}\mathrm{Rank}(\boldsymbol{A}-\boldsymbol{B})$.
- the weight of a word $\boldsymbol{A}$ is $|\boldsymbol{A}|_r\stackrel{\text{def}}{=}\mathrm{Rank}(\boldsymbol{A})=d_R(\boldsymbol{A},\boldsymbol{0})$.

An important family of matrix code are the $\mathbb{F}_{q^m}$-linear codes.

Definition II.2 ($\mathbb{F}_{q^m}$-linear codes)**.** *A $\mathbb{F}_{q^m}$-linear code $\mathcal{C}$ of length n and dimension k is a subspace of dimension k of $\mathbb{F}_{q^m}^n$. We say $\mathcal{C}$ is an $[n,k]_{q^m}$ linear code, or simply an $[n,k]$ code if there is no ambiguity.*

A natural way to define the rank metric over $\mathbb{F}_{q^m}$-linear codes is to consider the matrix associated to a word of $\mathbb{F}_{q^m}^n$.

Definition II.3 (Associated matrix)**.** *Let $\boldsymbol{x}=(x_1,\ldots,x_n)\in\mathbb{F}_{q^m}^n$ and $\boldsymbol{\beta}=(\beta_1,dots,\beta_m)$ a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$.*
$\boldsymbol{x}$ can be represented by an $m\times n$ matrix $\boldsymbol{M_x}$ such that its i^{th} column represents the coordinates of x_i in the basis $\boldsymbol{\beta}$.

$$\boldsymbol{x}\leftrightarrow\begin{pmatrix}x_{11}&\cdots&x_{1n}\\ \vdots&&\vdots\\ x_{m1}&\cdots&x_{mn}\end{pmatrix}$$

with $x_j = \sum_{i=1}^{m} x_{ij}\beta_i$ *for all* $j \in [1..n]$.

By definition, $|\boldsymbol{x}|_r \stackrel{\text{def}}{=} |\boldsymbol{M}_{\boldsymbol{x}}|_r$ and $d_R(\boldsymbol{x},\boldsymbol{y}) = d_R(\boldsymbol{M}_{\boldsymbol{x}},\boldsymbol{M}_{\boldsymbol{y}})$. These definitions do not depend on the choice of the basis.

Definition II.4 (Matrix code associated to an $\mathbb{F}_{q^m}$-linear code)**.** *Let* $\mathcal{C}$ *be an* $[n,k]_{q^m}$ *linear code. The matrix code* $\mathcal{C}^M$ *associated to* $\mathcal{C}$ *(with respect to a basis of* $\mathbb{F}_{q^m}$ *over* $\mathbb{F}_q$*) is the code* $\{\boldsymbol{M}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{C}\}$.

It is an $[m \times n, km]_q$ *matrix code.*

An $[n,k]_{q^m}$ linear code $\mathcal{C}$ can be represented by a generator matrix $\boldsymbol{G} \in \mathbb{F}_{q^m}^{k\times n}$ where each row $\boldsymbol{g}_i$ of $\boldsymbol{G}$ is an element of a basis of $\mathcal{C}$ or by a parity-check matrix $\boldsymbol{H} \in \mathbb{F}_{q^m}^{(n-k)\times n}$ such that $\boldsymbol{c} \in \mathcal{C} \iff \boldsymbol{c}\boldsymbol{H}^T = \boldsymbol{0}$. In the first case, $\mathcal{C}$ is viewed as the image of a morphism and in the second case as the kernel of a morphism. If $\boldsymbol{G}$ (respectively $\boldsymbol{H}$) is of the form $(\boldsymbol{I}_k|\boldsymbol{A})$ (respectively $(\boldsymbol{B}|\boldsymbol{I}_{n-k})$), we say that it is under systematic form.

The advantage of $\mathbb{F}_{q^m}$-linear codes with respect to matrix codes of the same parameters is that they have a much more compact representation. Indeed, under their systematic form, an $[n,k]_{q^m}$ $\mathbb{F}_{q^m}$-linear can be represented by $(n-k)k$ coefficients in $\mathbb{F}_{q^m}$ which is $(n-k)km$ coefficients in $\mathbb{F}_q$ whereas an $[m\times n, km]_q$ matrix code can be represented by $(nm-km)km = (n-k)m^2$ coefficients in $\mathbb{F}_q$.

Definition II.5 (Support of a word)**.** *Let* $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$*. The support of* $\boldsymbol{x}$ *is the* $\mathbb{F}_q$*-subspace of* $\mathbb{F}_{q^m}$ *generated by the coordinates of* $\boldsymbol{x}$*. It is denoted* $\mathrm{Supp}(\boldsymbol{x})$.

$$\mathrm{Supp}(\boldsymbol{x}) = \langle x_1,\dots,x_n\rangle_{\mathbb{F}_q}$$

The weight of a word is equal to the dimension of its support.

The number of supports of dimension w in $\mathbb{F}_{q^m}$ is denoted by the Gaussian coefficient $\begin{bmatrix} m \\ w \end{bmatrix}_q$ and is equal to:

$$\begin{bmatrix} m \\ w \end{bmatrix}_q = \prod_{i=0}^{w-1} \frac{q^m - q^i}{q^w - q^i}$$

In practice Gaussian coefficients are approximated by $q^{w(m-w)}$. In the case of matrix codes, we can consider the subspace of $\mathbb{F}_q^m$ generated by the columns of the matrix, called the columns support, or the subspace of $\mathbb{F}_q^n$ generated by its rows, called the row support.

B. Hard problem in rank metric

Rank-based cryptography relies on difficult problems in rank metric. These problems are the same as in the Hamming metric, but with the rank metric replacing the Hamming metric. In this subsection, we only consider $\mathbb{F}_{q^m}$-linear codes but all the problems are defined exactly the same way for linear matrix codes.

The first one corresponds to the decoding problem by syndromes.

Definition II.6 (Rank Syndrome Decoding (RSD) problem)**.** *Let* $\boldsymbol{H}$ *be the parity-check matrix of an* $[n,k]$ *code,* $\boldsymbol{s} \in \mathbb{F}_{q^m}^{n-k}$ *and* w *an integer. The RSD problem consists in finding* $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ *such that:*

$$\begin{cases} \boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s}^T \\ |\boldsymbol{e}|_r = w \end{cases} \tag{1}$$

The second problem is the search for small-weight codewords.

Definition II.7 (Small-weight codeword problem)**.** *Let* $\mathcal{C}$ *be an* $[n,k]$ *code and* w *an integer. The problem consists in finding a codeword* $\boldsymbol{c} \in \mathcal{C}$ *such that* $|\boldsymbol{c}|_r = w$.

Remark: If $\boldsymbol{H}$ is a parity-check matrix of $\mathcal{C}$, the small-weight codeword problem corresponds to an instance of the RSD problem with $\boldsymbol{s} = \boldsymbol{0}$.

III. Description of the Attacks on the RSD Problem

Our algorithm is an improvement of the GRS algorithm [1]. We first recall the description of this algorithm in the first subsection then we present our improvement in the second one.

A. The GRS algorithm

The general idea to solve the RSD problem is to find a subspace F such that $\mathrm{Supp}(\boldsymbol{e}) \subset F$. Then we can express the coordinates of $\boldsymbol{x}$ in a basis of F and solve the linear system given by the parity-check equations. There are two possible cases we will describe more precisely.

- **first case:** $n \geqslant m$**.**
  Let F be a subspace of $\mathbb{F}_{q^m}$ of dimension r and $(F_1,\dots,F_r)$ a basis of F. Let us assume that $\mathrm{Supp}(\boldsymbol{e}) \subset F$.
  $$\Rightarrow \forall i \in [1..n], e_i = \sum_{j=1}^{r} \lambda_{ij} F_j$$
  This gives us nr unknowns over $\mathbb{F}_q$.
  We can now rewrite the parity-check equations in these unknowns:
  $$\boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s} \tag{2}$$
  $$\Leftrightarrow \begin{cases} H_{11}e_1 & +\cdots+ & H_{1n}e_n & = & s_1 \\ \vdots & & \vdots & & \vdots \\ H_{n-k,1}e_1 & +\cdots+ & H_{n-k,n}e_n & = & s_{n-k} \end{cases}$$
  $$\Leftrightarrow \begin{cases} \sum_{l=1}^{n} H_{1l} & \sum_{j=1}^{r} \lambda_{lj}F_j & = & s_1 \\ \vdots & & & \vdots \\ \sum_{l=1}^{n} H_{n-k,l} & \sum_{j=1}^{r} \lambda_{lj}F_j & = & s_{n-k} \end{cases} \tag{3}$$
  Now, we need to embed these $n-k$ equations over $\mathbb{F}_{q^m}$ into $(n-k)m$ equations over $\mathbb{F}_q$.
  Let φ_i the i^{th} canonical projection from $\mathbb{F}_{q^m}$ on $\mathbb{F}_q$:
  $$\begin{aligned} \varphi_i : \quad \mathbb{F}_{q^m} &\to \mathbb{F}_q \\ \sum_{i=1}^{m} x_i\beta_i &\mapsto x_i \end{aligned}$$

We have:

$$\begin{aligned}&\boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s}\\ &\Leftrightarrow \forall i \in [1..m],\\ &\begin{cases} \sum_{l=1}^{n}\sum_{j=1}^{r}\lambda_{1j}\varphi_i(H_{1l}F_j) &= \varphi_i(s_1)\\ \vdots &\vdots\\ \sum_{l=1}^{n}\sum_{j=1}^{r}\lambda_{1j}\varphi_i(H_{n-k,l}F_j) &= \varphi_i(s_{n-k}) \end{cases}\end{aligned} \tag{4}$$

Since we assume $\mathrm{Supp}(\boldsymbol{e}) \subset F$, this system has at least one solution. We want more equations than unknowns in order to have only one solution with overwhelming probability. So we have the condition:

$$(n-k)m \geqslant nr \Leftrightarrow r \leqslant m - \left\lceil \frac{km}{n} \right\rceil$$

If the system has a solution, we check if its rank is equal to w. Otherwise, we choose another F and restart at the beginning.

The average complexity of the algorithm is $\mathcal{O}\big(\frac{(n-k)^3m^3}{p}\big)$ where p is the probability that $\mathrm{Supp}(\boldsymbol{e}) \subset F$.

The probability p is equal to the number of subspaces of dimension w in a subspace of dimension r divided by the number of subspaces of dimension w in $\mathbb{F}_{q^m}$.

$$p = \frac{\begin{bmatrix} r \\ w \end{bmatrix}_q}{\begin{bmatrix} m \\ w \end{bmatrix}_q} \approx q^{-w(m-r)}$$

By taking $r = m - \left\lceil \frac{km}{n} \right\rceil$ we obtain a complexity of $\mathcal{O}\big((n-k)^3m^3q^{w\left\lceil \frac{km}{n} \right\rceil}\big)$

- **second case:** $m > n$. In this case we consider the matrix code associated to the $\mathbb{F}_{q^m}$-linear code. Instead of searching for a subspace which contains the columns support of the matrix of the error, we search for a subspace F of $\mathbb{F}_q^n$ which contains the rows support of the error. The rest of the algorithm is the same, the only differences are:
  - the number of unknowns is mr, which implies $r \leqslant n-k$.
  - probability to choose F correctly is $\dfrac{\begin{bmatrix} r \\ w \end{bmatrix}_q}{\begin{bmatrix} n \\ w \end{bmatrix}_q} \approx q^{-w(n-r)}$.

  Thus, the complexity in this case is $\mathcal{O}\big((n-k)^3m^3q^{wk}\big)$.

This algorithm does not use the $\mathbb{F}_{q^m}$-linearity of the code. To exploit this structure, the main idea is to consider the code $\mathcal{C}' = \mathcal{C} + \mathbb{F}_{q^m}\boldsymbol{e}$, where $\mathcal{C}$ is the code with parity-check matrix $\boldsymbol{H}$.

The problem is reduced to the search for a codeword of weight w in $\mathcal{C}'$. If $\boldsymbol{e}$ is the only solution of the system 1, then the only codewords of $\mathcal{C}'$ of weight w are of the form $\alpha\boldsymbol{e}, \alpha \in \mathbb{F}_{q^m}^*$. These codewords are solutions of the system

$$\begin{cases} \boldsymbol{H}'\boldsymbol{e}'^T = \boldsymbol{0} \\ |\boldsymbol{e}'|_r = w \end{cases} \tag{5}$$

where $\boldsymbol{H}'$ is a parity-check matrix of $\mathcal{C}'$.

Thus, instead of looking for the support E of $\boldsymbol{e}$, we can look for any multiple αE of the support. In [1], the authors specialize one element of the support by searching for small weight codeword $\boldsymbol{c}$ such that $1 \in \mathrm{Supp}(\boldsymbol{c}) = E$ (but one can choose any non zero element of $\mathbb{F}_{q^m}^*$). One can use this information to improve the probability that $F \supset \mathrm{Supp}(\boldsymbol{c})$ by choosing F such that $1 \in F$. Let φ be the projection from the $\mathbb{F}_q$-subspaces of $\mathbb{F}_{q^m}$ which contains 1 to the subspaces of the quotient-space $\mathbb{F}_{q^m}/\mathbb{F}_q$ such that $\varphi(V) = V/\mathbb{F}_q$. We have $\dim\varphi(F) = \dim F - 1$ and $\dim\varphi(E) = \dim E - 1$. Moreover, $E \subset F$ if and only if $\varphi(E) \subset \varphi(F)$.

Thus, the probability is equal to $\dfrac{\begin{bmatrix} w-1 \\ r-1 \end{bmatrix}_q}{\begin{bmatrix} w-1 \\ m-1 \end{bmatrix}_q} \approx q^{(w-1)(m-r)}$.

We use the same techniques as before to obtain the system

$$\begin{cases} \sum_{l=1}^{n}\sum_{j=1}^{r}\lambda_{1j}\varphi_i(H'_{1l}F_j) &= 0\\ \vdots &\vdots\\ \sum_{l=1}^{n}\sum_{j=1}^{r}\lambda_{1j}\varphi_i(H'_{n-k-1,l}F_j) &= 0 \end{cases} \tag{6}$$

We take $r = \left\lfloor \frac{m(n-k-1)}{n} \right\rfloor = m - \left\lceil \frac{m(k+1)}{n} \right\rceil$ in order to have more equations than unknowns. If $F \supset E$, then the kernel of system 6 is of positive dimension and we can compute a non-zero vector $\boldsymbol{e}'$ solution of $\boldsymbol{H}'\boldsymbol{e}'^T = \boldsymbol{0}$. Then if it is of rank w, we solve the equation $\boldsymbol{H}\boldsymbol{e}'^T = \alpha\boldsymbol{H}\boldsymbol{e}^T = \alpha\boldsymbol{s}^T$ of unknown α to recover $\boldsymbol{e}$. Finally, we get an average complexity of $\mathcal{O}\big((n-k)^3m^3q^{(w-1)\frac{(k+1)m}{n}}\big)$.

In the case $m > n$, we cannot use the attack of Ourivski and Johansson to obtain an average complexity of $\mathcal{O}\big((n-k)^3m^3q^{(w-1)(k+1)}\big)$ [5].

B. Improvement of this algorithm

The idea of our improvement is still to search for a codeword of weight w in the code $\mathcal{C}'$. However, instead of choosing F such that $1 \in F$, we choose a random F. If it contains a multiple αE of E, then we can compute the codeword $\alpha\boldsymbol{e}$ of $\mathcal{C}'$ as before. The probability that $F \supset \alpha E$ depends on the number of different subspaces of the form $\alpha E, \alpha \in \mathbb{F}_{q^m}^*$.

Proposition III.1. *Let E be an $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^m}$ of dimension w. There are at most $\frac{q^m-1}{q-1}$ subspaces of the form $\alpha E, \alpha \in \mathbb{F}_{q^m}^*$, with equality if w and m are coprime.*

Proof. Let S be the cardinality of the set of the subspaces of this form. Let α and $\beta \in \mathbb{F}_{q^m}$ such that $\alpha\beta^{-1} \in \mathbb{F}_q^*$ By linearity, it is obvious that $\alpha E = \beta E$, hence $S \leqslant \frac{q^m-1}{q-1}$.

To prove the equality, let $\alpha \in \mathbb{F}_{q^m}^*$ be such that $E = \alpha E$. By induction, we have $\mathbb{F}_q(\alpha)E = E$ thus E is an $\mathbb{F}_q(\alpha)$-subspace of $\mathbb{F}_{q^m}$. Let $m' = [\mathbb{F}_q(\alpha) : \mathbb{F}_q]$. We have $w = \dim_{\mathbb{F}_q} E = m' \dim_{\mathbb{F}_q(\alpha)} E$ hence m' divides w and m. If $GCD(m, w) = 1$ then $\alpha \in \mathbb{F}_q^*$, which concludes the proof. □

This also proves that if the number S of different subspaces of the form αE is less than $\frac{q^m-1}{q-1}$ then E is an $\mathbb{F}_{q^{m'}}$-subspace of dimension $\frac{w}{m'}$ of $\mathbb{F}_{q^m}$. The probability of such an event is negligible, so we can assume $S = \frac{q^m-1}{q-1}$.

The probability p that F of dimension $r = m - \left\lceil \frac{m(k+1)}{n} \right\rceil$ contains a subspace of the form αE, $\alpha \in \mathbb{F}_{q^m}^*$ can be approximated by the product of S by the probability that F contains a fixed subspace of dimension w. This approximation is correct if

$$\frac{q^m - 1}{q - 1} \begin{bmatrix} r \\ w \end{bmatrix}_q \ll \begin{bmatrix} m \\ w \end{bmatrix}_q \Leftrightarrow q^{m+w(r-w)} \ll q^{w(m-w)}.$$

Notice that the bound $\frac{q^m-1}{q-1}$ on S is an upper bound, the previous discussion shows that in practice the practical value of S can be approximated by this bound, which is used for evaluating the complexity of the attack. Now a value of S lower than the bound may permit to improve a little bit the attack but is unlikely to happen. The likely value of S associated with the probability that F contains contains a fixed subspace of dimension w leads us to the following theorem:

Theorem III.2. *Let $\boldsymbol{H} \in \mathbb{F}_{q^m}^{(n-k)\times n}, \boldsymbol{s} \in \mathbb{F}_{q^m}^{n-k}$ and w an integer. Our algorithm solves the RSD problem II.6 with average complexity of*

$$\mathcal{O}\big((n-k)^3 m^3 q^{w\left\lceil \frac{(k+1)m}{n} \right\rceil - m}\big)$$

operations in $\mathbb{F}_q$.

An implementation of the attack showed experimental values very close to the theoretic ones, which confirms our approximation. In the case $m \leqslant n$, our algorithm is always better than [1]. The gain in the exponent is equal to $m - \left\lceil \frac{(k+1)m}{n} \right\rceil = m - \lceil mR' \rceil \approx m(1 - R')$ where R' is the rate of $\mathcal{C}'$. In the case $m > n$ our algorithm may also be faster for some parameters.

IV. Examples of broken parameters

In this section, we present two public key encryption cryptosystems: the Loidreau cryptosystem [3] (based on Gabidulin codes masked by an LRPC matrix) and the LRPC cryptosystem [2]. Some of their parameters are broken by our attack. They are both based on the McEliece cryptosystem. To encrypt a message $\boldsymbol{m}$, one computes the cipher text $\boldsymbol{c} = \boldsymbol{m}\boldsymbol{G}_{pub} + \boldsymbol{e}$ where $\boldsymbol{G}_{pub}$ is the public generator matrix of the masked code ans $\boldsymbol{e}$ is an error of wight w. To decrypt, the receiver removes the mask and decodes the error.

Under the assumption that the masked code is indistinguishable from a random code, an attacker has to solve an instance of the RSD problem of weight w to recover the message.

The following tables give some parameters which are broken by our attack:

- Parameters from Loidreau's cryptosystem [3]

n	k	m	w	claimed security	complexity of our attack
82	41	41	4	2^{80}	2^{75}
106	53	53	5	2^{128}	2^{116}

- Parameters from LRPC cryptosystem [12]

n	k	m	w	claimed security	complexity of our attack
50	32	50	3	2^{81}	2^{75}
112	80	112	4	2^{259}	2^{247}

For all these cryptosystems our attack improves on the best attack on which their security is based, the gain is a little less than $m(1 - R)$ since the authors probably took a security margin on their parameters, but our attack still permits to break these parameters. Notice that some parameters of these cryptosystems are not affected by our attack since it is not necessarily the best attack if $m > n$.

Our attack also breaks the parameters of the IBE in [4] but they have been updated in the eprint version of the paper.

In the process of the standardization of post-quantum cryptosystems engaged by the NIST, the cryptosystems LAKE, LOCKER, Ouroboros-R, RankSign and RQC are based on rank metric codes and take account on this attack for the choice of their security parameters. See https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-1-Submissions for the documentation on these submissions.

References

[1] P. Gaborit, O. Ruatta, and J. Schrek, "On the complexity of the rank syndrome decoding problem," *CoRR*, vol. abs/1301.1026, 2013. [Online]. Available: http://arxiv.org/abs/1301.1026

[2] P. Gaborit, O. Ruatta, J. Schrek, and G. Zémor, "New results for rank-based cryptography," in *Progress in Cryptology - AFRICACRYPT 2014*, ser. LNCS, vol. 8469, 2014, pp. 1–12.

[3] P. Loidreau, "A new rank metric codes based encryption scheme," in *International Workshop on Post-Quantum Cryptography*. Springer, 2017, pp. 3–17.

[4] P. Gaborit, A. Hauteville, D. H. Phan, and J. Tillich, "Identity-based encryption from rank metric," in *Advances in Cryptology - CRYPTO2017*, ser. LNCS, vol. 10403. Springer, Aug. 2017, pp. 194–226.

[5] A. V. Ourivski and T. Johansson, "New technique for decoding codes in the rank metric and its cryptography applications," *Problems of Information Transmission*, vol. 38, no. 3, pp. 237–246, 2002.

[6] E. Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 384–386, May 1978.

[7] E. M. Gabidulin, "Theory of codes with maximum rank distance," *Problemy Peredachi Informatsii*, vol. 21, no. 1, pp. 3–16, 1985.

[8] E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov, "Ideals over a non-commutative ring and their applications to cryptography," in *Advances in Cryptology - EUROCRYPT'91*, ser. LNCS, no. 547, Brighton, Apr. 1991, pp. 482–489.

[9] R. Overbeck, "A new structural attack for GPT and variants," in *Mycrypt*, ser. LNCS, vol. 3715, 2005, pp. 50–63.

[10] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. S. L. M. Barreto, "MDPC-McEliece: New McEliece variants from moderate density parity-check codes," in *Proc. IEEE Int. Symposium Inf. Theory - ISIT*, 2013, pp. 2069–2073.

[11] J. Hoffstein, J. Pipher, and J. H. Silverman, "NTRU: A ring-based public key cryptosystem," in *Algorithmic Number Theory, Third International Symposium, ANTS-III, Portland, Oregon, USA, June 21-25, 1998, Proceedings*, ser. LNCS, J. Buhler, Ed., vol. 1423. Springer, 1998, pp. 267–288.

[12] P. Gaborit, O. Ruatta, J. Schrek, and G. Zémor, "Ranksign: An efficient signature algorithm based on the rank metric (extended version on arxiv)," in *Post-Quantum Cryptography 2014*, ser. LNCS, vol. 8772. Springer, 2014, pp. 88–107. [Online]. Available: https://arxiv.org/pdf/1606.00629.pdf