Two new algorithms for error support recovery of low rank parity check codes

Ermes Franch, Chunlei Li, *Senior Member, IEEE*
Department of Informatics, University of Bergen, Norway
Emails: {ermes.franch, chunlei.li}@uib.no

***Abstract*—Due to their weak algebraic structure, low rank parity check (LRPC) codes have been employed in several post-quantum cryptographic schemes. In this paper we propose new improved decoding algorithms for $[n,k]_{q^m}$ LRPC codes of dual rank weight d. The proposed algorithms can efficiently decode LRPC codes with the parameters satisfying $n-k=rd-c$, where r is the dimension of the error support and $c \leq d-2$. They outperform the original decoding algorithm of LRPC codes when $d>2$ and allow for decoding LRPC codes with a higher code rate and smaller values m.**

I. Introduction

Rank-metric codes, which are embedded in a rank metric space, have applications in network coding [1], space-time codes [2], distributed storage [3], and cryptography [4]–[11]. Rank-based cryptography relies on the difficulty of the rank syndrome decoding (RSD) problem. So far the best-known method of solving the RSD problem has an exponential complexity which is quadratic in the parameter size [12], [13]. This nice feature allows for smaller sizes of keys in rank-based cryptosystems to achieve the same level of security provided by those cryptosystems based on Hamming-metric codes. Existing rank-based cryptographic schemes mainly used two types of rank-metric codes: Gabidulin codes [14] and low rank parity-check (LRPC) codes [6] and their variants, for which efficient decodings have been extensively studied [15]–[17]. Due to the strong algebraic structure of Gabidulin codes, the GPT cryptosystem and its variants are subject to the structural attack by Overbeck [18]. LRPC codes can be seen as the equivalent of LDPC codes in rank metric and have a very weak algebraic structure. These codes could be masked more easily in cryptosystems. Consequently, different schemes based on LRPC codes were proposed in recent years, RankSign [7], ROLLO [11], Durandal [10], etc. On the other hand, without a certain algebraic structure, LRPC codes can only be decoded in a probabilistic manner. The original decoding algorithm of LPRC codes in [6] works only for the cases where the support of the syndrome is exactly the product space of the parity-check support and the error support. When decoding LRPC codes, the error support recovery step contributes a dominating factor to decoding failures. In the extended paper [19] on LRPC codes and their cryptographic applications, the authors further considered the cases where the syndrome support has a dimension $rd-c$ with $c<r$, where r is the dimension of the error support, d is the dual rank weight of the LRPC codes and rd is the dimension of the product space between the two supports. By applying two expansion functions on the syndrome support, they proposed new decoders that can correct errors with higher rank weights and decrease the decoding failure rate.

In this paper we consider an alternative approach to decoding LRPC codes for the cases where the syndrome support has a dimension $rd-c$ with $c<d$. The proposed decoders rely on a crucial observation that employs all the elements (instead of only the basis elements) in the parity-check support, which enables us to significantly loosen the restriction on m as required in [19]. The paper is organized as follows: Section II introduces notations, basics on rank metric codes, the problems of rank syndrome decoding and error support recovery. Section III recalls the LRPC codes and their decoding approach. In Section IV we start with some theoretical analysis and then proposed two new algorithms for LRPC codes where the syndrome support has dimension $rd-c$ with $c \leq d-2$; and Section V discusses the decoding failure rate of the proposed algorithms and their connections to the improved decoding algorithms in [19].

II. Preliminaries

We denote by $\mathbb{F}_{q^m}$ the finite field of q^m elements where q is a prime power. Vectors will be indicated by bold lower-case letters, and the i-th component of a vector will be indicated by the same letter in normal font, for example, $\mathbf{v}=(v_1,\ldots,v_n)$. The notation $\mathbb{F}_q^n$ denotes the vector space of all the vectors of length n over $\mathbb{F}_q$. A matrix M will be indicated by upper-case letter, and the (i,j)-th entry of M will be indicated by $m_{i,j}$. The notation $\mathbb{F}_q^{m\times n}$ indicates all the $m\times n$ matrices over $\mathbb{F}_q$. The notation $[n]$ indicates the interval $\{1,\ldots,n\}\subset\mathbb{N}$. Given

This work was supported by the Research Council of Norway under Grant No. 311646/O70

a set $S \subseteq \mathbb{F}_{q^m}$ and an element $a \in \mathbb{F}_{q^m}$ the notation Sa corresponds to the set $\{sa \mid s \in S\}$.

The field $\mathbb{F}_{q^m}$ can be regarded as a vector space of dimension m over the field $\mathbb{F}_q$. We will denote $\mathbb{F}_q$-linear subspaces $\mathcal{S}$ of $\mathbb{F}_{q^m}$ by upper case calligraphic letters. Given a set $S \subseteq \mathbb{F}_{q^m}$, we denote by $\langle S \rangle_{\mathbb{F}_q}$ the $\mathbb{F}_q$-linear subspace generated by the elements of this space. For a given vector $\mathbf{v} \in \mathbb{F}_{q^m}^n$, we denote as $\langle \mathbf{v} \rangle_{\mathbb{F}_q} = \langle v_1, \ldots, v_n \rangle_{\mathbb{F}_q}$ the subspace generated by the components of $\mathbf{v}$, this will be called the support of $\mathbf{v}$. Similarly, for a matrix $M \in \mathbb{F}_{q^m}^{k \times n}$, we will denote the subspace generated by all its entries as $\langle M \rangle_{\mathbb{F}_q} = \langle m_{i,j} \mid (i,j) \in [k] \times [n] \rangle_{\mathbb{F}_q}$. The notations $\mathcal{A}^n$ and $\mathcal{A}^{m \times n}$ stand for the set of all the vectors of length n having support $\mathcal{A}$ and the set of all the $m \times n$ matrices with support $\mathcal{A}$, respectively.

Using the notion of support, we can define the rank distance over $\mathbb{F}_{q^m}^n$. Consider $\mathbf{u}, \mathbf{v} \in \mathbb{F}_{q^m}^n$, the **rank-weight** of $\mathbf{u}$ is given by $\mathrm{w_R}(\mathbf{u}) = \dim(\langle \mathbf{u} \rangle_{\mathbb{F}_q})$, and the **rank distance** between two vectors is defined as $\mathrm{d_R}(\mathbf{u}, \mathbf{v}) = \mathrm{w_R}(\mathbf{u} - \mathbf{v})$. It can be proved that the function $\mathrm{d_R}$ is a distance in the mathematical sense.

Definition 1. A **rank metric code** is a subset $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ equipped with the rank distance. The **minimum rank distance** of the code $\mathcal{C}$ is given by the minimum distance between any two different elements of the code, i.e., $\mathrm{d_R}(\mathcal{C}) = \min(\{\mathrm{d_R}(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \neq \mathbf{v} \in \mathcal{C}\})$. If $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ is an $\mathbb{F}_{q^m}$-linear subspace of $\mathbb{F}_{q^m}$ we call $\mathcal{C}$ an $\mathbb{F}_{q^m}$**-linear rank metric code**.

An $\mathbb{F}_{q^m}$-linear rank metric code can be defined by the use of a generator matrix or the use of a parity-check matrix.

Definition 2. Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}$ be a vector subspace of $\mathbb{F}_{q^m}^n$ of dimension k. A **parity check matrix** of $\mathcal{C}$ is a matrix $H \in \mathbb{F}_{q^m}^{(n-k) \times n}$ of rank $n-k$ such that all the elements $\mathbf{c} \in \mathcal{C}$ satisfy $\mathbf{c}H^\intercal = \mathbf{0}$. For a generic $\mathbf{v} \in \mathbb{F}_{q^m}^n$, we will have that $\mathbf{v}H^\intercal = \mathbf{s} \in \mathbb{F}_{q^m}^{n-k}$, which is known as the **syndrome** of $\mathbf{v}$.

One interesting problem in coding theory is the syndrome decoding problem. Below we recall this problem in the context of rank metric codes and a closely-related problem.

Definition 3. Given $H \in \mathbb{F}_{q^m}^{n-k \times n}$ a parity check matrix of a $\mathbb{F}_{q^m}$-linear rank metric code $\mathcal{C}$, a syndrome $\mathbf{y}H^\intercal = \mathbf{s} \in \mathbb{F}_{q^m}^{n-k}$ and a small integer r. The **Rank Syndrome Decoding (RSD) problem** consists in finding $\mathbf{e} \in \mathbb{F}_{q^m}^n$ such that

$$\mathbf{e}H^\intercal = \mathbf{s}, \quad \mathrm{w_R}(\mathbf{e}) \leq r.$$

Definition 4. Given $H \in \mathbb{F}_{q^m}^{n-k \times n}$ a parity check matrix of an $\mathbb{F}_{q^m}$-linear rank metric code $\mathcal{C}$, a syndrome $\mathbf{y}H^\intercal = \mathbf{s} \in \mathbb{F}_{q^m}^{n-k}$ and a small integer r. The **Rank Support Recovery problem** is to find a subspace $\mathcal{E} \subseteq \mathbb{F}_{q^m}$ of dimension $\leq r$ such that there exists $\mathbf{e} \in \mathcal{E}^n$ and $\mathbf{e}H^\intercal = \mathbf{s}$.

A vector $\mathbf{e} \in \mathbb{F}_{q^m}$ of rank r will have a support $\mathcal{E} = \langle \mathbf{e} \rangle_{\mathbb{F}_q} = \langle \beta_1, \ldots, \beta_r \rangle_{\mathbb{F}_q}$ where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)$ is a basis of $\mathcal{E}$. The vector $\mathbf{e}$ can then be represented as $\mathbf{e} = \boldsymbol{\beta} C_e$ where $C_e \in \mathbb{F}_q^{r \times n}$ is a matrix over $\mathbb{F}_q$ which is the matrix of the coordinates of $\mathbf{e}$ with respect to the basis $\boldsymbol{\beta}$.

Many decoding algorithms for rank metric codes consist of two major steps: the first step is to find the error support and recover one basis of the support, and the second is to use the basis to recover the matrix of the coordinates of the error. Usually, once the error support is known, finding the matrix of the coordinates reduces to solving a linear system in nr variables corresponding to the nr entries of C_e. In this paper, we will focus on the Rank Support Recovery problem for the LRPC codes.

III. LRPC CODES AND THEIR DECODING

LRPC codes were introduced in 2013 by Gaborit, Murat, Ruatta and Zémor [6].

Definition 5 (LRPC code)**.** Let $\mathcal{A} \subseteq \mathbb{F}_{q^m}$ be an $\mathbb{F}_q$-linear subspace of $\mathbb{F}_{q^m}$ of dimension d and $H \in \mathcal{A}^{(n-k) \times n}$ a matrix of rank $n-k$. The code $\mathcal{C}$ having H as a parity check matrix is called an $[n,k]_{q^m}$ **LRPC code** of dual rank weight d.

Due to their lack of a strong algebraic structure, these codes were proposed for several cryptographic applications [7], [10], [11]. LRPC codes have a polynomial time decoding algorithm divided into two steps. The first step aims to recover the error support; the second step uses the error support acquired in the first step to find the exact coordinates of the error. Due to the page limit, in this paper we will focus only on the error support recovery part of the algorithm.

Consider $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be an LRPC code as in Definition 5. Suppose we receive a message $\mathbf{y} = \mathbf{x} + \mathbf{e}$, where $\mathcal{E} = \langle \mathbf{e} \rangle_{\mathbb{F}_q}$. If $\dim(\mathcal{E}) = r$ then $\mathcal{E}$ will have a basis $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r) \in \mathbb{F}_{q^m}^r$. Consider now the syndrome $\mathbf{s} = \mathbf{e}H^\intercal$, where $H \in \mathcal{A}^{(n-k) \times n}$ for some space $\mathcal{A}$ of dimension d. Each component $s_i = \sum_{j=1}^n e_j h_{i,j}$ of the syndrome is obtained as the sum of elements in the set $\mathcal{A}\mathcal{E} = \{eh \mid e \in \mathcal{E}, h \in \mathcal{A}\}$. A key observation is that, if we denote with $\mathcal{A}.\mathcal{E} = \langle \mathcal{A}\mathcal{E} \rangle_{\mathbb{F}_q}$ the smallest $\mathbb{F}_q$-linear vector subspace containing the set $\mathcal{A}\mathcal{E}$, we automatically have that $\mathbf{s} \in (\mathcal{A}.\mathcal{E})^{n-k}$.

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ be a basis of $\mathcal{A}$ and $\mathbf{b} = (\beta_1, \ldots, \beta_r)$ be a basis of $\mathcal{E}$. Observe that for any $c = ae \in \mathcal{A}.\mathcal{E}$, we have

$$c = \Big(\sum_{i=1}^{d} a_i \alpha_i \Big) \Big(\sum_{j=1}^{r} e_j \beta_j \Big) = \sum_{(i,j) \in [d] \times [r]} a_i e_j (\alpha_i \beta_i).$$

As a consequence $\mathcal{A}.\mathcal{E} = \langle \boldsymbol{\alpha} \otimes \boldsymbol{\beta} \rangle_{\mathbb{F}_q}$, where $\boldsymbol{\alpha} \otimes \boldsymbol{\beta} = (\alpha_1 \beta_1, \ldots, \alpha_d \beta_r)$ is a vector of length rd. This means that $\dim(\mathcal{A}.\mathcal{E}) \leq \min(rd, m)$. In [6] the authors proved that the equality holds with a high probability.

After this observation, we are ready to present the error support recovery algorithm. Since $\mathbf{s} \in (\mathcal{A}.\mathcal{E})^{n-k}$, we have that $\mathcal{S} = \langle \mathbf{s} \rangle_{\mathbb{F}_q} \subseteq \mathcal{A}.\mathcal{E}$. For $n - k \geq rd$, if we consider the

$n-k$ components of $\mathbf{s}$ as elements randomly extracted with a uniform distribution from $\mathcal{A}.\mathcal{E}$, with a good probability, we have $\mathcal{S} = \mathcal{A}.\mathcal{E}$. For the success of this algorithm it is crucial that $\mathcal{S} = \mathcal{A}.\mathcal{E}$, therefore we need the condition $n-k \geq rd$. The probability that $\mathcal{S} = \mathcal{A}.\mathcal{E}$ is estimated to be at least $1 - q^{rd-(n-k)}$ [11]. This probability can be made arbitrarily small by choosing $n-k$ significantly larger than rd. Notice that, if $\dim(\mathcal{A}.\mathcal{E}) = rd-t$, then we might require just $n-k \geq rd-t$. This means that in the case $\dim(\mathcal{A}.\mathcal{E}) < rd$, using the same number $n-k \geq rd$ of parity check equations, we will have an even better probability that $\mathcal{S} = \mathcal{A}.\mathcal{E}$. Therefore, considering $\dim(\mathcal{A}.\mathcal{E}) = rd$ can then be regarded as a worst-case scenario.

Suppose that $\mathcal{S} = \mathcal{A}.\mathcal{E}$, notice that, for any element α_i of the basis $\boldsymbol{\alpha}$, we have that $\mathcal{E} \subset \mathcal{S}\alpha^{-1}$. If we intersect all these sets, we have that $\mathcal{E} \subseteq \mathcal{S}\alpha_1^{-1} \cap \ldots \cap \mathcal{S}\alpha_d^{-1}$ where the equality holds with an estimated probability of at least $1 - q^{-(d-1)(m-rd-r)}$ [11, Prop. 2.4.2]. For large values of m, this probability becomes quickly negligible.

Both of the failure probabilities reduce exponentially in q. The first probability of failure is harder to reduce than the second when $d > 2$ since, increasing $n-k$ by one, improves the probability by a factor q^{-1} while, increasing m by one, improves the second probability by a factor $q^{-(d-1)}$.

In [19] Aragon, Gaborit, Hauteville, Ruatta, and Zémor gave two improved versions of the decoding algorithm that make use of two different expanding functions to be able to decode when $\mathcal{S} \subsetneq \mathcal{A}.\mathcal{E}$. In particular, they were able to decode when $\dim(S) = rd - c$, with $c < r$. The main drawback of these two new algorithms is the need of a larger m which has to be in the order of $3rd-2$ in the first algorithm, and $2rd-r$ in the second. Their first algorithm was able to decode LRPC codes with $n-k > (d-1)r$. The second improved considerably the success probability while keeping $n-k \geq rd$.

The algorithms we propose tackle the same problem. They offer similar improvements while keeping m in the order of rd. We are able to decode when $c \leq d-2$ while asking $n-k \geq (r-1)d+2$.

IV. Improved Error Support Recovery Algorithms

Consider an LRPC code defined over $\mathbb{F}_{q^m}^n$ with parity check matrix $H \in \mathcal{A}^{(n-k)\times n}$ where $\dim(\mathcal{A}) = d$. Suppose we receive $\mathbf{y} = \mathbf{x}+\mathbf{e}$ where $\mathbf{x} \in \mathcal{C}$ and $\mathbf{e} \in \mathbb{F}_{q^m}^n$ is an error of rank r. Then the support $\mathcal{E} = \langle \mathbf{e} \rangle_{\mathbb{F}_q}$ has dimension $\dim(\mathcal{E}) = r$.

We already showed $\mathbf{s} = \mathbf{e}H^\intercal \in (\mathcal{A}.\mathcal{E})^{n-k}$. For the rest of this section let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ be a basis of $\mathcal{A}$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)$ be a basis of $\mathcal{E}$. We will assume $\dim(\mathcal{A}.\mathcal{E}) = rd$ which is both the most common and the worst case,

Suppose all of the $n-k$ elements of the syndrome are linearly independent and that $n-k = rd-c$. The vector space $\mathcal{S} = \langle \mathbf{s} \rangle_{\mathbb{F}_q}$ will be a proper subspace of $\mathcal{A}.\mathcal{E}$ of co-dimension c. In [19] it was shown how the code can still be decoded when $c < r$. The reason behind this condition was given by the following lemma [19].

Lemma 1. *Let $\mathcal{A}, \mathcal{E}$ be two subspaces of $\mathbb{F}_{q^m}$ of dimension d and r, respectively, and $\mathcal{S}$ be a subspace of $\mathcal{A}.\mathcal{E}$ with $\dim(\mathcal{S}) = \dim(\mathcal{A}.\mathcal{E}) - c$. For any nonzero elements $a \in \mathcal{A}$, $b \in \mathcal{E}$, we have $\dim(\mathcal{S}a^{-1} \cap \mathcal{E}) \geq r-c$ and $\dim(\mathcal{S}b^{-1} \cap \mathcal{A}) \geq d-c$.*

Proof: Note that $\mathcal{S}a^{-1} + \mathcal{E} = (\mathcal{S} + \mathcal{E}a)a^{-1} \subseteq (\mathcal{A}.\mathcal{E})a^{-1}$. Thus $\dim(\mathcal{S}a^{-1} + \mathcal{E}) \leq \dim(\mathcal{A}.\mathcal{E})$. From this fact, we have the following inequality

$$\begin{aligned}\dim(\mathcal{S}a^{-1} \cap \mathcal{E}) &= \dim(\mathcal{S}a^{-1}) + \dim(\mathcal{E}) - \dim(\mathcal{S}a^{-1} + \mathcal{E}) \\ &\geq \dim(\mathcal{A}.\mathcal{E}) - c + r - \dim(\mathcal{A}.\mathcal{E}) = r - c.\end{aligned}$$

Due to the symmetrical role of $\mathcal{A}$ and $\mathcal{E}$, the second statement follows similarly. ∎

Lemma 1 shows that, when $c < r$, many elements of $\mathcal{E}$ are contained in $\mathcal{S}a^{-1}$ for every nonzero element a in $\mathcal{A}$.

The original algorithm considers only the d sets of the form $\mathcal{S}\alpha_i^{-1}$ where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ is a basis of $\mathcal{A}$. In our case, since different subsets of $\mathcal{E}$ might be contained in different sets, we want to use as many different sets as we can. To better understand why considering the elements of a basis might not be enough, consider the following example. Let $s_1 = \alpha_1\beta_1 + \alpha_2\beta_2$ and $s_2 = \alpha_1\beta_2 + \alpha_2\beta_1$. If we consider the set $\mathcal{S} = \langle s_1, s_2 \rangle_{\mathbb{F}_q}$ we have that $s_1 + s_2 = (\alpha_1 + \alpha_2)(\beta_1 + \beta_2) \in \mathcal{S}$ then $(\beta_1 + \beta_2) \in \mathcal{S}(\alpha_1 + \alpha_2)^{-1} \cap \mathcal{E}$ while $\mathcal{S}\alpha_1^{-1} \cap \mathcal{E}$ and $\mathcal{S}\alpha_2^{-1} \cap \mathcal{E}$ are both empty.

Intuitively, if we consider all the sets of the form Sa^{-1} for some $a \in \mathcal{A}^*$, the elements of $\mathcal{E} \cap Sa^{-1}$ will appear in many of the other sets with the same form, while the elements of $\mathcal{S}a^{-1} \setminus \mathcal{E}$, will occur with significantly less frequency in the other sets.

The following theorem gives a simple way to count how many of the sets in the form $\mathcal{S}a^{-1}$ for $a \in \mathcal{A}^*$ contain an element $x \in \mathbb{F}_{q^m}$.

Theorem 1. *Let $\mathcal{S}$ be a subspace of $\mathcal{A}.\mathcal{E}$, where $\mathcal{A}, \mathcal{E} \subseteq \mathbb{F}_{q^m}$ are two subspaces of $\mathbb{F}_{q^m}$ of dimension d and r. Define a multi-set Z as a union of $\mathcal{S}a^{-1}$ for all $a \in \mathcal{A}^*$, i.e.,*

$$Z = \biguplus_{a \in A^*} \mathcal{S}a^{-1}. \tag{1}$$

The multiplicity of $x \in \mathbb{F}_{q^m}^*$ in Z is given by

$$\mathrm{Mul}(x, Z) = |\mathcal{S}x^{-1} \cap \mathcal{A}^*| = q^{\dim(\mathcal{S}x^{-1} \cap \mathcal{A})} - 1 \tag{2}$$

and $\mathrm{Mul}(x = 0, Z) = q^d - 1$.

Proof: Consider the set

$$\alpha(x) = \{a \in \mathcal{A}^* \mid \exists s \in \mathcal{S} : x = sa^{-1}\}.$$

By definition of Z, we have that $|\alpha(x)| = \mathrm{Mul}(x, Z)$. For $x = 0$, since $0 \in \mathcal{S}$, we have that $0 = 0a^{-1}$ for all possible

$a \in \mathcal{A}^*$ therefore $\alpha(0) = \mathcal{A}^*$ and $\mathrm{Mul}(0, Z) = |\mathcal{A}^*| = q^d - 1$. To complete the proof of the theorem it is sufficient to show that $\alpha(x) = \mathcal{S}x^{-1} \cap \mathcal{A}^*$. Notice that, for $a \in \mathcal{A}^*, 0 \neq x \in \mathcal{S}$ we have

$$x = sa^{-1} \iff a = sx^{-1}.$$

By the definition of $\alpha(x)$ and $\mathcal{S}x^{-1}$, the first equation is equivalent to $a \in \alpha(x)$ while the second equation is equivalent to $a \in \mathcal{S}x^{-1} \cap \mathcal{A}^*$. This shows that $\alpha(x) = \mathcal{S}x^{-1} \cap \mathcal{A}^*$. We know that $|\mathcal{S}x^{-1} \cap \mathcal{A}| = q^{\dim(\mathcal{S}x^{-1} \cap \mathcal{A})}$. The desired conclusion thus follows. ∎

The following corollaries characterize the multiplicities of different elements in Z.

Corollary 1. *Let $\mathcal{A}, \mathcal{E}, \mathcal{S}$ and Z be as in Theorem 1, where $\mathcal{S}$ has dimension $rd - c$. Let $w = \max\{1, rd - c + d - m\}$. Then for any $x \in Z \setminus \mathcal{E}$, its multiplicity* $\mathrm{Mul}(x, Z) \geq q^w - 1$.

Proof: Since $x \in Z$, there exists $a \in \mathcal{A}$ such that $x = sa^{-1}$ for some $s \in \mathcal{S}$, implying $a = sx^{-1} \in \mathcal{S}x^{-1}$. Thus $\dim(\mathcal{S}x^{-1} \cap \mathcal{A}) \geq 1$. In addition, we know that

$$\begin{aligned}\dim(\mathcal{S}x^{-1} \cap \mathcal{A}) &= \dim(\mathcal{S}x^{-1}) + \dim(\mathcal{A}) - \dim(\mathcal{S}x^{-1} + \mathcal{A}) \\ &= \dim(\mathcal{S}) + \dim(\mathcal{A}) - \dim(\mathcal{S}x^{-1} + \mathcal{A}) \\ &\geq rd - c + d - m\end{aligned}$$

since $\mathcal{S}x^{-1} + \mathcal{A}$ has dimension at most m. ∎

Corollary 2. *Let $\mathcal{A}, \mathcal{E}, \mathcal{S}$ and Z be as in Theorem 1, where $\mathcal{S}$ has dimension $rd - c$. Then* $\mathrm{Mul}(b, Z) \geq q^{d-c} - 1$, $\forall b \in \mathcal{E}$.

Proof: It is clear that $Mul(0, Z) = q^d - 1 \geq q^{d-c} - 1$. For $b \in \mathcal{E}^*$, it follows from Theorem 1 that

$$\mathrm{Mul}(b, Z) = |\mathcal{S}b^{-1} \cap \mathcal{A}^*| = |\mathcal{S}b^{-1} \cap \mathcal{A}| - 1. \tag{3}$$

By Lemma 1 we know that $|\mathcal{S}b^{-1} \cap \mathcal{A}| \geq q^{d-c}$. This together with the above equality leads to the desired statement $\mathrm{Mul}(b, Z) \geq q^{d-c} - 1$. ∎

Our main goal is to recover the error support $\mathcal{E}$ when the support $\mathcal{S}$ is a proper subspace of $\mathcal{A}.\mathcal{E}$. From the above analysis, we can create the multi-set Z and focus only on the elements with multiplicity greater than or equal to $q^{d-c} - 1$. Consider the set

$$\tilde{\mathcal{E}} = \{x \in Z \mid \mathrm{Mul}(x, Z) \geq q^{d-c} - 1\}. \tag{4}$$

Thanks to Corollary 2, we know that $\mathcal{E} \subseteq \tilde{\mathcal{E}}$. From Corollary IV, it is better to choose $m \geq rd - c + d - 1$ such that the generic element of Z can have multiplicity as low as $q - 1$ while the elements of $\mathcal{E}$ will always have multiplicity at least $q^{d-c} - 1$. It is clear that $d - c \geq 2$ is a minimal condition to distinguish a generic element in Z from the elements of $\mathcal{E}$ when $q = 2$.

It is possible that some elements of Z have large multiplicity even if they are not elements of $\mathcal{E}$, when that happens we have $|\tilde{\mathcal{E}}| > |\mathcal{E}| = q^r$. Assume that $\tilde{\mathcal{E}} = \mathcal{E} \cup X \subseteq \mathbb{F}_{q^m}$ where X is a set of small cardinality $|X| < q^r$. Note that for any $x \in \mathcal{E}$, $\mathcal{E} \subset (x + \tilde{\mathcal{E}}) \cap \tilde{\mathcal{E}}$. With this fact, we can quickly obtain $\mathcal{E}$ from $\tilde{\mathcal{E}}$ in the following way: take a random $x \in \tilde{\mathcal{E}}$, if $|\tilde{\mathcal{E}} \cap (\tilde{\mathcal{E}} + x)| > q^r$, then take $\tilde{\mathcal{E}} = \tilde{\mathcal{E}} \cap (\tilde{\mathcal{E}} + x)$, and continue this process until $|\tilde{\mathcal{E}}| = |\mathcal{E}| = q^r$. Such a process of filtering works well when the number of outliers is small compared to the size of $\mathcal{E}$.

The above process of selecting elements in Z with multiplicites at least $q^{d-c} - 1$ is summarized in **Alg. 1**. Empirically, setting a large enough m, we will directly obtain $\tilde{\mathcal{E}} = \mathcal{E}$. Diminishing the value of m, the set $\tilde{\mathcal{E}} \setminus \mathcal{E}$ will progressively grow until it will not be possible anymore to retrieve the correct $\mathcal{E}$.

Algorithm 1: Error support recovery of LRPC codes (Theoretical Version)

Input: A parity check matrix $H \in \mathcal{A}^{(n-k)\times n}$ where $\dim(\mathcal{A}) = d$ and $\mathcal{S} = \langle \mathbf{y}H^\intercal \rangle_{\mathbb{F}_q}$ of dimension $rd - c$ where $\mathbf{y} = \mathbf{x} + \mathbf{e} \in \mathbb{F}_{q^m}^n$, $\mathbf{x} \in \mathcal{C}$ and $\mathbf{e}$ is an error of rank r.

Output: The support $\mathcal{E} = \langle \mathbf{e} \rangle_{\mathbb{F}_q}$ of dimension r.

```
// Assumption: dim(A.E) = rd
if rd − dim(S) < d − 2 then
    c = rd − dim(S);
    Z = {∗ ∗} ;          // Create an empty multi-set
    for a ∈ A* do
        for s ∈ S do
            Z = Z ⊎ {sa^{-1}} ;          // Compute Z
        end
    end
    E = {} ;
    for z ∈ Z do
        if Mul(z, Z) ≥ q^{d−c} − 1 then
            E = E ∪ {z} ;
        end
    end
    Return E;
// The dimension of S is too low
else
    Error Support Recovery Failure;
end
```

In **Alg. 1**, the generation of the multi-set $Z = \biguplus_{a \in \mathcal{A}^*} \mathcal{S}a^{-1}$ has a high time and space complexity. In order to address this drawback, we propose a more practical algorithm, which select elements from the intersection of $t > 2$ subspaces $\mathcal{S}a^{-1}$ for some rounds and generates the error support generated by those elements. We first give some theoretical analysis before presenting the second algorithm.

Proposition 1. *Let $\mathcal{S} \subset \mathcal{A}.\mathcal{E}$ be given as in Theorem 1 with dimension $rd - c$. Take nonzero elements $a_1, \ldots, a_t$ from $\mathcal{A}$. Then* $\dim(\mathcal{S}a_1^{-1} \cap \cdots \cap \mathcal{S}a_t^{-1} \cap \mathcal{E}) \geq r - tc$.

Proof: Take $\mathcal{T}_i = \mathcal{S}a_1^{-1} \cap E$ of dimension $\geq r - c$. Then

$$\begin{aligned}
&\dim\left(\mathcal{T}_1 \cap \cdots \cap \mathcal{T}_t\right) \\
&= \dim(\mathcal{T}_2 \cap \cdots \cap \mathcal{T}_t) + \dim\left(\mathcal{T}_1\right) - \dim(\mathcal{T}_1 + (\mathcal{T}_2 \cap \cdots \cap \mathcal{T}_t)) \\
&\geq \dim(\mathcal{T}_2 \cap \cdots \cap \mathcal{T}_t) + (r-c) - r = \dim(\mathcal{T}_2 \cap \cdots \cap \mathcal{T}_t) - c.
\end{aligned}$$

Iterating this process on t leads to the desired inequality. ■

By Proposition 1, the intersection of t subspace $\mathcal{S}a_i^{-1}$ may contribute to $r - tc$ independent elements in $\mathcal{E}$. This implies that we can recover the error support $\mathcal{E}$ by accumulating elements from such intersections. The following result is obtained by applying [11, Prop. 2.4.2].

Proposition 2. *Consider $\mathcal{E}$ a subspace of dimension r. Let Ω be the set of all subspaces of dimension $rd - c$ having intersection of dimension at least $r-c$ with $\mathcal{E}$. The probability that the intersection of t subspaces $\mathcal{S}_1, \ldots, \mathcal{S}_t$ independently chosen uniformly at random in Ω contains some elements not in $\mathcal{E}$ is approximated by*

$$Prob((\mathcal{S}_1 \cap \cdots \cap \mathcal{S}_t) \subset \mathcal{E})) \approx 1 - 1/q^{(t-1)m + r - trd}.$$

Proof: Following the proof of [11, Prop. 2.4.2]. The subspaces $\mathcal{S}_1, \ldots, \mathcal{S}_t$ are not independent since each contains part of $\mathcal{E}$. If consider the quotient space $V = \mathbb{F}_{q^m}/\mathcal{E}$ of dimension $m - r$. The sets $\mathcal{R}_i = \mathcal{S}_i/\mathcal{E}$ are now independent and will have dimension $\dim(\mathcal{R}_i) = \dim(\mathcal{S}_i) - \dim(\mathcal{S}_i \cap \mathcal{E}) \leq rd - c - r + c = rd - r$. Fix $0 \neq y \in \mathcal{R}_1$, if we consider $\mathcal{R}_2$ independent from $\mathcal{R}_1$ the probability that $y \in \mathcal{R}_2$ is $(|\mathcal{R}_2| - 1)/(|\mathbb{F}_{q^m}/\mathcal{E}| - 1) \leq (q^{rd-r} - 1)/(q^{m-r} - 1)$. The same will be true for the other $\mathcal{R}_i$ which gives a probability of $((q^{rd-r} - 1)/(q^{m-r} - 1))^{t-1}$. If we consider each y as independent we have to multiply this probability by the number of y which is $|\mathcal{R}_1| - 1 = q^{rd-r} - 1$ This gives us

$$\begin{aligned}
\text{Prob}(\dim(\cap_{i\in[t]}\mathcal{R}_i) > 0) &\leq (q^{rd-r} - 1)\left(\frac{q^{rd-r} - 1}{q^{m-r} - 1}\right)^{(t-1)} \\
&\approx q^{-t(m-rd)+m-r}
\end{aligned}$$

Notice that $\text{Prob}(\dim(\mathcal{R}_1 \cap \ldots \cap \mathcal{R}_t) > 0)$ is equal to $\text{Prob}((\mathcal{S}_1 \cap \ldots \cap \mathcal{S}_t) \setminus \mathcal{E} \neq \{0\})$. This ends the proof. ■

With the statements in Propositions 1 and 2, we propose a more practical decoding of LRPC codes in **Alg. 2**. Even though the subspaces $\mathcal{S}a^{-1}$ in **Alg. 2** cannot be considered as uniformly independently chosen, heuristically they seem to follow Proposition 2.

V. Discussion

For the two improved decoding algorithms in this paper, **Alg. 1** utilizes the observation that all nonzero elements in the parity-check support $\mathcal{A}$ can contribute to the process of error support recovery, yet it can only work well for small parameters due to its high complexity in memory. **Alg. 2** derives elements in the error support by repeatedly intersecting $t > 2$ subspaces $\mathcal{S}a_i^{-1}$. In this way, the algorithm is pretty efficient for different choices of parameters m, r, d, c, and t.

Algorithm 2: Error support recovery of LRPC codes

Input: A parity check matrix $H \in \mathcal{A}^{(n-k)\times n}$ where $\dim(\mathcal{A}) = d$ and $\mathcal{S} = \langle \mathbf{y}H^\intercal \rangle_{\mathbb{F}_q}$ of dimension $rd - c$ where $\mathbf{y} = \mathbf{x} + \mathbf{e} \in \mathbb{F}_{q^m}^n$, $\mathbf{x} \in \mathcal{C}$ and $\mathbf{e}$ is an error of rank r.

Output: The support $\mathcal{E} = \langle \mathbf{e} \rangle_{\mathbb{F}_q}$ of dimension r.

```
// Assumption: dim(A.E) = rd
if rd − dim(S) < d then
    S = {s_1, ..., s_u} = Basis(S);
    t = q^⌈log_q(r/c)⌉;
    E = { };
    while dim(⟨E⟩_{F_q}) < r do
        // Generate t random elements from A*
        Y = {a_1, ..., a_t} =Random(A*, t);
        for a ∈ Y do
            Generate Sa^{-1} = ⟨{a^{-1}s_1, ..., a^{-1}s_u}⟩_{F_q};
        end
        E = E + ⋂_{1≤i≤t} Sa_i^{-1};
    end
    Return E ;
else
    Return "Support recovery failure" ;
end
```

c	r, d	t	m	Success	r, d	t	m	Success
1	5, 5	4	40	99.9%	5, 6	4	46	99.4%
	5, 5	4	41	100%	5, 6	4	47	100%
2	5, 5	4	42	99.9%	5, 6	4	48	99.9%
	5, 5	4	43	100%	5, 6	4	49	100%

Table 1: Success rate of **Alg. 2** with $n - k = rd - c$, $c = 1, 2$

Our algorithms and the algorithms in [19] both improved the original decoding of LRPC codes for the cases where the syndrome support has dimension $rd - c$ for $c > 0$. While the algorithms in [19] require $m \geq 3rd - 2$ and $m \geq 2rd - r$, respectively, our new algorithms have significantly loosened the requirements on m, namely, **Alg. 1** requires $m \geq rd - 2(d - c)$ and **Alg. 2** requires $m \geq \frac{t}{t-1}rd$. Increasing the value of m by 1 will improve the failure rate by $q^{-(t-1)}$.

In Table 1 we provide some experimental results for **Alg. 2**, where we run a series of 1000 experiments for parameters $n - k = rd - c$ for $c = 1, 2$, $(r, d) = (5, 5), (5, 6)$, $t = 4$ for different m, and report the lower bound of m that gives nearly 100% success rates of error support recovery. The parameters that rule the probability of success are c and m. In our experiments we used $k = 1$ and $n = rd - c + k$. We observe that repeating the same experiments for different values of k did not affect the results reported in Table 1.

REFERENCES

[1] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3579–3591, 2008.

[2] P. Lusina, E. Gabidulin, and M. Bossert, "Maximum rank distance codes as space-time codes," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2757–2760, 2003.

[3] E. Gabidulin, *Rank Codes*. TUM.University Press, 2021.

[4] H. Bartz, L. Holzbaur, H. Liu, S. Puchinger, J. Renner, and A. Wachter-Zeh, "Rank-metric codes and their applications," 2022.

[5] E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov, "Ideals over a non-commutative ring and their application in cryptology," in *Advances in Cryptology – EUROCRYPT'91* (D. W. Davies, ed.), pp. 482–489, Springer, 1991.

[6] Philippe Gaborit, Gaétan Murat, Olivier Ruatta, and Gilles Zémor, "Low rank parity check codes and their application to cryptography. in proceedings of the workshop on coding and cryptography WCC'2013 Bergen Norway 2013. available on www.selmer.uib.no/wcc2013/pdfs/gaborit.pdf."

[7] P. Gaborit, O. Ruatta, J. Schrek, and G. Zémor, "Ranksign: an efficient signature algorithm based on the rank metric," in *Post-Quantum Cryptography* (M. Mosca, ed.), pp. 88–107, Springer International Publishing, 2014.

[8] P. Gaborit, A. Hauteville, D. H. Phan, and J.-P. Tillich, "Identity-based encryption from codes with rank metric," in *Advances in Cryptology – CRYPTO 2017* (J. Katz and H. Shacham, eds.), (Cham), pp. 194–224, Springer International Publishing, 2017.

[9] P. Gaborit, O. Ruatta, J. Schrek, and G. Zémor, "New results for rank-based cryptography," in *Progress in Cryptology – AFRICACRYPT 2014* (D. Pointcheval and D. Vergnaud, eds.), (Cham), pp. 1–12, Springer International Publishing, 2014.

[10] N. Durante and A. Siciliano, "Non-linear maximum rank distance codes in the cyclic model for the field reduction of finite geometries," *The Electronic Journal of Combinatorics*, vol. 24, no. 2.33, pp. 1–18, 2017.

[11] C. A. Melchor, N. Aragon, M. Bardet, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, A. Hauteville, A. Otmani, O. Ruatta, J.-P. Tillich, and G. Zémor, "ROLLO (merger of Rank-Ouroboros, LAKE and LOCKER)," in *Second round submission to the NIST post-quantum cryptography call*, April, 2020.

[12] P. Gaborit and G. Zémor, "On the hardness of the decoding and the minimum distance problems for rank codes," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7245–7252, 2016.

[13] P. Gaborit, O. Ruatta, and J. Schrek, "On the complexity of the rank syndrome decoding problem," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 1006–1019, 2016.

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

[15] P. Loidreau, "A Welch–Berlekamp like algorithm for decoding Gabidulin codes," in *International Workshop on Coding and Cryptography (WCC)* (Ø. Ytrehus, ed.), (Berlin, Heidelberg), pp. 36–45, Springer, 2006.

[16] A. Wachter-Zeh, V. Afanassiev, and V. Sidorenko, "Fast decoding of Gabidulin codes," *Designs, Codes and Cryptography*, vol. 66, no. 1-3, pp. 57–73, 2013.

[17] W. K. Kadir and C. Li, "On decoding additive generalized twisted Gabidulin codes," *Cryptography and Communications*, vol. 12, pp. 987 – 1009, 2020.

[18] R. Overbeck, "Structural attacks for public key cryptosystems based on Gabidulin codes," *Journal of Cryptology*, vol. 21, pp. 280–301, Apr 2008.

[19] N. Aragon, P. Gaborit, A. Hauteville, O. Ruatta, and G. Zémor, "Low rank parity check codes: New decoding algorithms and applications to cryptography," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 7697–7717, 2019.