Generalized low rank parity check codes

Ermes Franch
University of Bergen
Bergen, Norway
ermes.franch@uib.no

Philippe Gaborit
University of Limoges
Limoges, France
gaborit@unilim.fr

Chunlei Li, *Senior Member, IEEE*
University of Bergen
Bergen, Norway
chunlei.li@uib.no

***Abstract*—In this work we propose a family of $\mathbb{F}_q$-linear low-rank parity check (LRPC) codes based on a bilinear product over $\mathbb{F}_q^m$ defined by a generic 3-tensor over $\mathbb{F}_q$. A particular choice of this tensor corresponds to the classical $\mathbb{F}_{q^m}$-linear LRPC codes; and other tensors yield $\mathbb{F}_q$-linear codes, which, with some caveats, can be efficiently decoded with the same idea of decoding LRPC codes. The proposed codes contribute to the diversity of rank metric codes for cryptographic applications, particularly for the cases where attacks utilize $\mathbb{F}_{q^m}$-linearity to reduce decoding complexity.**

***Index Terms*—Algebraic coding theory, rank metric codes, network coding, cryptography**

I. Introduction

Rank metric codes play an important role in coding theory [1], [2] and have found a variety of applications in networking [3] and cryptography [4]–[7]. One desirable property of rank metric codes for cryptography is the hardness of the syndrome decoding problem. The decoding of a random $\mathbb{F}_q$-linear rank metric code is proven to be NP-complete [8], and for the $\mathbb{F}_{q^m}$-linear case, the syndrome decoding can be probabilistically reduced to an NP-complete problem [9]. So far the best-known attacks for solving the rank syndrome decoding problem have an exponential complexity which is quadratic in the parameters. This allows for significantly smaller key sizes in cryptographic schemes based on rank metric codes, when compared to those with codes in the Hamming metric. Recent years have seen a resurgence of interest in rank-based cryptography. Researchers have proposed various rank-based cryptographic schemes, including RankSign [5], identity-based encryption [10], ROLLO [6], the signature scheme Durandal [7], etc. On the other hand, recent developments of cryptanalytic attacks against rank-based cryptography also challenged parameters of schemes based on $\mathbb{F}_{q^m}$-linear rank metric codes. The early GPT cryptosystem [4] and its variants based on Gabidulin codes are vulnerable to algebraic attacks by Overbeck [11]. For the decoding problem with $\mathbb{F}_{q^m}$-linear rank metric codes, Ourivsky and Johansson [12] exploited the $\mathbb{F}_{q^m}$-linear structure to reduce the decoding complexity. Very recently, refined attacks using the same model were proposed in a series of papers [13]–[15], which challenged the security parameters of several schemes based on low-rank parity check (LRPC) codes [16] without a significant structure.

In the theory of rank metric codes, it is of great interest to study codes with efficient decoding, which is vital for their applications. As for cryptographic applications, another requirement is that the codes should not exhibit a significant algebraic structure, which is difficult to mask securely. Motivated by recent developments of algebraic attacks [13]–[15] on the decoding problem for $\mathbb{F}_{q^m}$-linear rank metric codes and relevant cryptographic schemes, in this paper we propose a family of $\mathbb{F}_q$-linear rank metric codes, which have no significant algebraic structure and can be efficiently decoded. For efficient decoding of the proposed codes, we introduce a bilinear product over $\mathbb{F}_q^m$, which is based on a generic 3-tensor T over $\mathbb{F}_q$. The bilinear product has the property that the product of elements in two subspaces of small dimensions r, d lies in a subspace with dimension upper bounded by rd. This property allows for an efficient probabilistic decoding algorithm similar to that of LRPC codes. It can be shown that the LRPC codes [16] correspond to a particular choice of the tensor T, and that there exist other choices of T that produce $\mathbb{F}_q$-linear codes and allow for a compact public key as well. Due to the space limitation, in this work we omit proofs and examples, which are included in the full version of the paper.

II. Preliminaries

In this section we will introduce basic notations and auxiliary results for subsequent sections.

To avoid heavy notation we use $[n]$ to indicate the set $\{1, \ldots, n\}$. We denote by $\mathbb{F}_q$ the finite field with q elements, where q is a power of a prime number. The vector space $\mathbb{F}_q^n$ is the set of all the n-tuples over $\mathbb{F}_q$ while $\mathbb{F}_q^{m \times n}$ is the set of all the $m \times n$ matrices over the same field. Vectors will be indicated by lower bold case. Given a vector $\mathbf{v}$, its i-th component will be indicated as v_i. Matrices will be indicated by uppercase letters. Given a matrix A its i, j-th entry will be denoted by $a_{i,j}$. For a given set S in $\mathbb{F}_q^m$ or in $\mathbb{F}_{q^m}$ we call the $\mathbb{F}_q$-linear space generated by the elements of S the support of

The work of C. Li is supported by the Research Council of Norway under Grant No. 311646/O7

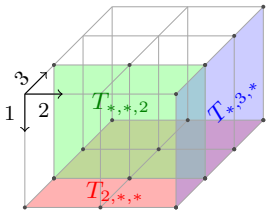

Figure 1: Slices of 3-tensors

S, denoted by $\langle S\rangle_{\mathbb{F}_q}$. Similarly, for a vector $\mathbf{v}\in\mathbb{F}_{q^m}$ and a matrix M the notation $\langle\mathbf{v}\rangle_{\mathbb{F}_q}$ and $\langle M\rangle_{\mathbb{F}_q}$ will denote the $\mathbb{F}_q$-linear space generated by the entries of $\mathbf{v}$ and M. We will call these two subspaces the support of $\mathbf{v}$ and the support of M.

A. 3-tensors over $\mathbb{F}_q$

In this paper elements in $\mathbb{F}_q^{n_1\times n_2\times n_3}$ will be termed *3-tensors* and will be also denoted by upper-case letters. Given a 3-tensor $T\in\mathbb{F}_q^{n_1\times n_2\times n_3}$ we will indicate with $t_{i,j,k}$ its i,j,k-th entry. Algebraically a 3-tensor T can be expressed as a vector of n_3 matrices T_i of size $n_1\times n_2$, i.e., $T=(T_1,\ldots,T_{n_3})$. A 3-tensor can be visualized as a rectangular cuboid (a closed box with six rectangular faces) of size $n_1\times n_2\times n_3$ in a system of three coordinates as displayed in Figure 1, where the first index indicates the vertical axis, the second indicates the horizontal axis and the third indicates the axis perpendicular to the paper. Given a 3-tensor T, one obtains a matrix of size $n_2\times n_3$ when fixing the 1st index of T to a certain value i for $1\le i\le n_1$. Likewise, one obtains a matrix of size $n_1\times n_3$ when fixing the 2nd index and a matrix of size $n_1\times n_2$ when fixing the 3rd index. We will denote by $T_{i,*,*},T_{*,j,*},T_{*,*,k}$ the matrices derived by fixing the 1st, 2nd and 3rd index of T to i,j and k, respectively. In Figure 1 we show 3 examples of this notation over a tensor $T\in\mathbb{F}_q^{2\times3\times4}$. More concretely, $T_{2,*,*}$ is the matrix by fixing the 1st index of T to 2, $T_{*,3,*}$ is the matrix by fixing the 2nd index of T to 3 and $T_{*,*,2}$ is the matrix by fixing the 2rd index of T to 2.

Multiplications over $\mathbb{F}_q^m$ associated with 3-tensors will be a core feature in the proposed generalized LRPC codes. Below we shall introduce multiplications between 3-tensors and vectors with respect to indices 1, 2 and 3, which, in a visualized manner, can be interpreted as directional multiplications. Given a 3-tensor $T\in\mathbb{F}_q^{n_1\times n_2\times n_3}$, vectors $\mathbf{x}\in\mathbb{F}_q^{n_1}$, $\mathbf{y}\in\mathbb{F}_q^{n_2}$, $\mathbf{z}\in\mathbb{F}_q^{n_3}$, we define the vertical multiplication between T and $\mathbf{x}$, denoted by, $T_{\mathbf{x},*,*}$, as the linear combination of T w.r.t $\mathbf{x}$ along the vertical direction, that is $T_{\mathbf{x},*,*}=\sum_{i=1}^{n_1}x_iT_{i,*,*}$ is a $n_2\times n_3$ matrix, where the j,k-th entry of $T_{\mathbf{x},*,*}$ is given by $\sum_{i=1}^{n_1}x_it_{i,j,k}$. Similarly, the horizontal multiplication between T and $\mathbf{y}$ will be defined as $T_{*,\mathbf{y},*}=\sum_{j=1}^{n_1}y_jT_{*,j,*}$ and the perpendicular multiplication between T and $\mathbf{z}$ will be defined as $T_{*,*,\mathbf{z}}=\sum_{k=1}^{n_1}z_kT_{*,*,k}$. Let $\mathbf{e}_i\in\mathbb{F}_q^{n_1}$ denote the i-th element of the standard base of $\mathbb{F}_q^{n_1}$. (i.e. the vector of length n_1 which is 1 in its j-th position and 0 elsewhere). The vertical multiplication between T and $\mathbf{e}_i$ is $T_{\mathbf{e}_i,*,*}=T_{i,*,*}$. Similarly the horizontal and the perpendicular multiplication with the standard vectors $\mathbf{e}_j$ of length n_2 and $\mathbf{e}_k$ of length n_3 is $T_{*,\mathbf{e}_j,*}=T_{*,j,*}$ and $T_{*,*,\mathbf{e}_k}=T_{*,*,k}$.

Suppose we want to multiply the matrix $T_{*,\mathbf{y},*}$ with the vector $\mathbf{x}$ along its first index. We can extend the notation introduced above indicating as $T_{\mathbf{x},\mathbf{y},*}=\mathbf{x}T_{*,\mathbf{y},*}$ Similarly $T_{*,\mathbf{y},\mathbf{z}}=T_{*,\mathbf{y},*}\mathbf{z}^\intercal$ and $T_{\mathbf{x},\mathbf{y},\mathbf{z}}=(\mathbf{x}T_{*,\mathbf{y},*})\mathbf{z}^\intercal$. While $T_{\mathbf{x},\mathbf{y},*}$ and $T_{*,\mathbf{y},\mathbf{z}}$ are vectors respectively in $\mathbb{F}_q^{n_3}$ and $\mathbb{F}_q^{n_1}$, we have that $T_{\mathbf{x},\mathbf{y},\mathbf{z}}$ is just an element of $\mathbb{F}_q$.

B. Matrix rank metric codes

The **column support** of a matrix $M\in\mathbb{F}_q^{m\times n}$ is the vector space $\mathrm{Colsp}(M)\subseteq\mathbb{F}_q^m$ generated by the columns of M. The **rank** of a matrix can be equivalently interpreted as the dimension of its column support and will be denoted by $\mathrm{Rank}(M)=\dim(\mathrm{Colsp}(M))$. The **rank distance** between two matrices $A,B\in\mathbb{F}_q^{m\times n}$ is then given by $d_R(A,B)=\mathrm{Rank}(A-B)$. Equipped with this distance, a subset $\mathcal{C}\subseteq\mathbb{F}_q^{m\times n}$ is called a **matrix (rank metric) code** if it is an $\mathbb{F}_q$-linear subspace of $\mathbb{F}_q^{m\times n}$. The **minimum rank distance** of a matrix code $\mathcal{C}$ is given by $d_R(\mathcal{C})=\min\{\mathrm{Rank}(U)\mid\mathbf{0}\neq U\in\mathcal{C}\}$. In the context of rank metric codes, for $m\ge n$ the **Singleton bound** is given as $|\mathcal{C}|\le q^{m(n-d_R+1)}$ and a code achieving the Singleton bound is said to be **maximum rank distance code (MRD).** Let $\mathcal{C}$ be an $\mathbb{F}_q$-linear matrix rank metric code of dimension k, namely, there are $G_1,\ldots,G_k\in\mathbb{F}_q^{m\times n}$ linearly independent matrices that generate this code. We can define a 3-tensor generator $G\in\mathbb{F}_q^{m\times n\times k}$ given by $G=(G_1,\ldots,G_k)$. Using the notation we introduced for 3-tensors, the code $\mathcal{C}$ can be expressed as $\mathcal{C}=\{G_{*,*,\mathbf{x}}\mid\mathbf{x}\in\mathbb{F}_q^k\}$.

Given two matrices $A,B\in\mathbb{F}_q^{m\times n}$ the trace inner product of A,B is defined as $\mathrm{Tr}(AB^\intercal)=\sum_{i,j}a_{i,j}b_{i,j}\in\mathbb{F}_q$. With this notion of inner product, we can define the dual of a k-dimensional linear code $\mathcal{C}\subseteq\mathbb{F}_q^{m\times n}$ as $\mathcal{C}^\perp=\{X\in\mathbb{F}_q^{m\times n}|\,\mathrm{Tr}(XC^\intercal)=0,\forall C\in\mathcal{C}\}$. The code $\mathcal{C}^\perp$ can be seen as the solution of a system of k linear equations in mn unknowns, therefore its dimension is lower bounded by $mn-k$. The 3-tensor generator of $\mathcal{C}^\perp$ will be a 3-tensor $H\in\mathbb{F}_q^{m\times n\times(mn-k)}$ and will be called a **parity-check tensor** of $\mathcal{C}$. The party-check tensor H of a code $\mathcal{C}$ gives a tool to quickly determine whether a matrix $A\in\mathbb{F}_q^{m\times n}$ belongs to $\mathcal{C}$. Indeed we have that $A\in\mathcal{C}$ iff $\mathrm{Tr}(H_{*,*,i}A^\intercal)=0$ for all $i\in[mn-k]$. The vector $\mathbf{s}\in\mathbb{F}_q^{mn-k}$ such that $s_i=\mathrm{Tr}(H_{*,*,i}A^\intercal)$ is called the **syndrome** of A. With this notion of syndrome we can express the rank syndrome decoding (RSD) problem for the matrix codes as follows.

Definition 1 (RSD Problem)**.** *Given a 3-tensor parity-check $H\in\mathbb{F}_q^{m\times n\times(nm-k)}$ of an $\mathbb{F}_q$-linear matrix code $\mathcal{C}\subseteq\mathbb{F}_q^{m\times n}$,*

a syndrome $\mathbf{s} \in \mathbb{F}_q^{nm-k}$ *and a small integer* r*, find a matrix* $E \in \mathbb{F}_q^{m\times n}$ *such that* $(\mathrm{Tr}(EH_1^\intercal), \ldots, \mathrm{Tr}(EH_{nm-k}^\intercal)) = \mathbf{s}$ *and* $\mathrm{Rank}(E) \leq r$*.*

Suppose a code $\mathcal{C}$ has minimum distance $d_R(\mathcal{C}) = d$. If we receive a matrix $Y = X + E$ where $X \in \mathcal{C}$ and E is an error of small rank-weight $< \frac{d}{2}$, one can recover X by solving the RSD problem. More concertely, one can compute the syndrome of Y, since $(\mathrm{Tr}(YH_1^\intercal), \ldots, \mathrm{Tr}(YH_{nm-k}^\intercal)) = 0+(\mathrm{Tr}(EH_1^\intercal), \ldots, \mathrm{Tr}(EH_{nm-k}^\intercal))$. Solving the RSD problem gives us E, from which we get the correct matrix $X = Y-E$.

III. Generalized LRPC codes

In this section we will propose a family of $\mathbb{F}_q$-linear matrix codes that generalizes the LRPC codes introduced in [16], [17]. It will be shown that the proposed codes include both the classical $\mathbb{F}_{q^m}$-linear LRPC codes and a large number of matrix codes that are only $\mathbb{F}_q$-linear. For efficiently decoding the proposed generalized LRPC codes, we will introduce a T-product over $\mathbb{F}_q^m$ which plays the same role as the standard product over $\mathbb{F}_{q^m}$.

A. T-product over $\mathbb{F}_q^m$

The product over $\mathbb{F}_{q^m}$ has two properties that are exploited in the decoding algorithm of the LRPC codes. The first property is that, for two given subspaces $\mathcal{H}, \mathcal{E} \subseteq \mathbb{F}_{q^m}$ of dimension d and r, the set $\mathcal{HE} = \{he \mid h \in \mathcal{H}, e \in \mathcal{E}\}$ is contained in a space of dimension upper bounded by rd. In particular if $\mathcal{H} = \langle h_1, \ldots, h_d\rangle_{\mathbb{F}_q}, \mathcal{E} = \langle e_1, \ldots, e_r\rangle_{\mathbb{F}_q}$, then $\mathcal{HE} \subseteq \langle h_i e_j \mid (i,j) \in [d] \times [r]\rangle_{\mathbb{F}_q} = \langle \mathcal{EH}\rangle_{\mathbb{F}_q}$. If the support of a parity check matrix and the support of the error are contained in two small subspaces $\mathcal{H}$ and $\mathcal{E}$ then, the support of the syndrome will be contained in $\langle \mathcal{HE}\rangle_{\mathbb{F}_q}$ of dimension upper bounded by rd. The second property of the standard product of $\mathbb{F}_{q^m}$ is the multiplicative inverse. Observe that for any $0 \neq h \in \mathcal{H}$ the subspace $\mathcal{E} \subseteq h^{-1}(\mathcal{H}.\mathcal{E})$. Intersecting such spaces, with a good probability, it is possible to recover the error space $\mathcal{E}$.

The vector space $\mathbb{F}_q^m$ does not have a product by default. For generalizing the LRPC codes, we will introduce a product over $\mathbb{F}_q^m$ satisfying the two properties discussed above.

Definition 2 (invertible bilinear product)**.** *The binary operation* $\star : \mathbb{F}_q^m \times \mathbb{F}_q^m \to \mathbb{F}_q^m$ *is an* ***invertible bilinear product*** *if it satisfies the following two properties. For* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{F}_q^m$*,*

- *(bilinear)* $(\mu\mathbf{a} + \nu\mathbf{b}) \star \mathbf{c} = \mu(\mathbf{a} \star \mathbf{c}) + \nu(\mathbf{b} \star \mathbf{c})$ *and* $\mathbf{a} \star (\mu\mathbf{b} + \nu\mathbf{c}) = \mu(\mathbf{a} \star \mathbf{b}) + \nu(\mathbf{a} \star \mathbf{c})$ *for any* $\mu, \nu \in \mathbb{F}_q$*.*
- *(invertible) Given* $\mathbf{a} \star \mathbf{b} = \mathbf{c}$*, for all* $\mathbf{b} \neq 0$ *the value of* $\mathbf{a}$ *such that* $\mathbf{a} \star \mathbf{b} = \mathbf{c}$ *is unique. So, the function* $\mathrm{inv}_{\mathbf{b}}$ *such that* $\mathrm{inv}_{\mathbf{b}}(\mathbf{c}) = \mathbf{a}$ *is well defined.*

Let $B = (\beta_1, \ldots, \beta_m)$ be a base of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. For $a \in \mathbb{F}_{q^m}$ we define $\phi_B(a) \in \mathbb{F}_q^m$ the vector of the coordinates of a with respect to B, i.e., $a = (\beta_1, \ldots, \beta_m)\phi_B(a)^\intercal$. The function ϕ_B induces an isomorphism between $\mathbb{F}_{q^m}$ and $\mathbb{F}_q^m$. An example of an invertible bilinear product is obtained by the composition of the standard product over $\mathbb{F}_{q^m}$ with ϕ_B. Explicitly, the product is defined as $\mathbf{a} \star \mathbf{b} = \phi_B(ab)$ where $a = \phi_B^{-1}(\mathbf{a}), b = \phi_B^{-1}(\mathbf{b}) \in \mathbb{F}_{q^m}$. The first property comes from the bi-linearity of the product in a field. For the second property, the function $\mathrm{inv}_{\mathbf{b}}$ we are looking for is just the $\star$-product by $\phi_B(b^{-1})$. That is $(\mathbf{a} \star \mathbf{b}) \star \phi_B(b^{-1}) = \mathbf{a}$. While in this case the product $\star$ inherits the commutativity by the field structure of $\mathbb{F}_{q^m}$, in general it is not always possible to express the function $\mathrm{inv}_{\mathbf{b}}$ as a right or left $\star$-multiplication by an element of $\mathbb{F}_{q^m}$. Below we discuss some properties of the bilinear product which will be used for decoding.

Proposition 1. *Let* $\star$ *be an invertible bilinear product. Let* $\mathcal{A} = \langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r\rangle_{\mathbb{F}_q}, \mathcal{B} = \langle \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_d\rangle_{\mathbb{F}_q} \subseteq \mathbb{F}_q^m$ *be two linear subspaces of dimension* r *and* d *such that* $rd \leq m$*. Consider* $\mathcal{A} \star \mathcal{B} = \{\mathbf{a} \star \mathbf{b} \mid \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}$*, then* $\dim(\langle \mathcal{A} \star \mathcal{B}\rangle_{\mathbb{F}_q}) \leq rd$*. Moreover, for* $0 \neq \mathbf{b} \in \mathcal{B}$ *we have*

$$\mathcal{A} \subseteq \mathrm{inv}_{\mathbf{b}}(\langle \mathcal{A} \star \mathcal{B}\rangle_{\mathbb{F}_q}) = \{\mathrm{inv}_{\mathbf{b}}(\mathbf{c}) \mid \mathbf{c} \in \langle \mathcal{A} \star \mathcal{B}\rangle_{\mathbb{F}_q}\}.$$

Notice that invertibility of the product $\star$ is not necessary for the statement $\dim(\langle \mathcal{A} \star \mathcal{B}\rangle_{\mathbb{F}_q}) \leq \dim(\mathcal{A})\dim(\mathcal{B})$.

We have seen that the composition of the standard product over $\mathbb{F}_{q^m}$ and an isomorphism ϕ_B is an invertible bilinear product, which is essentially the product used in classical LRPC codes. Below we introduce a more generalized product based on a generic 3-tensor T.

Definition 3 (T-product)**.** *Let* T *be a* 3*-tensor in* $\mathbb{F}_q^{m\times m\times m}$*. For* $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^m$ *we define the* T*-product between* $\mathbf{a}$ *and* $\mathbf{b}$ *as*

$$\mathbf{a} \cdot_T \mathbf{b} := \mathbf{a}T_{*,\mathbf{b},*} = T_{\mathbf{a},*,*}\mathbf{b}^\intercal = T_{\mathbf{a},\mathbf{b},*}.$$

The k*-th component of* $\mathbf{c} = \mathbf{a} \cdot_T \mathbf{b} = T_{\mathbf{a},\mathbf{b},*}$ *is given by*

$$c_k = \sum_{i=1}^{m}\sum_{j=1}^{m} t_{i,j,k}a_i b_j.$$

While the T-product is bilinear for any tensor T in $\mathbb{F}_q^{m\times m\times m}$, being invertible, in general, is not granted.

We already discussed how we can interpret a 3-tensor as the generator of a matrix linear code. Studying the code generated by the tensor T will give us a necessary and sufficient condition to establish if, for a given tensor T, its associated T-product is invertible or not. In particular all known finite presemifields can be used to generate invertible T-products. This connection, in a similar context, was explored in [18, Theorem 3] and previously in [19, Theorem 4.4.1].

Theorem 1. *A tensor* $T \in \mathbb{F}_q^{m\times m\times m}$ *defines an invertible product iff* $\{T_{*,i,*} \in \mathbb{F}_q^{m\times m}\}$ *is a base of an MRD code of dimension* m*. Equivalently iff*

$$\mathrm{Rank}(T_{*,\mathbf{b},*}) = m, \ \forall \mathbf{b} \in \mathbb{F}_q^m \setminus \{\mathbf{0}\}.$$

The T-product can be used also to define an inner product between matrices in $\mathbb{F}_q^{m\times n}$.

Definition 4 (T-inner product)**.** *Let $A, B \in \mathbb{F}_q^{m\times n}$ two matrices, we denote by $\mathbf{a}_i, \mathbf{b}_i$ the i-th column of A and B. For a 3-tensor $T \in \mathbb{F}_q^{m\times m\times m}$, the T-**inner product** of A and B is defined as*

$$A \cdot_T B = \sum_{i\in[n]} \mathbf{a}_i \cdot_T \mathbf{b}_i \in \mathbb{F}_q^m.$$

The vector $A \cdot_T B \in \mathbb{F}_q^m$ can be rewritten using the trace function component by component as

$$(A \cdot_T B)_k = \mathrm{Tr}(A^\intercal T_{*,*,k} B) = \mathrm{Tr}(T_{*,*,k} B A^\intercal).$$

It can be shown that the T-inner product between two matrices in $\mathbb{F}_q^{m\times n}$, for a certain tensor, coincides with the standard inner product between two vectors in $\mathbb{F}_{q^m}^n$.

B. Generalized LRPC codes

The subspace generated by k $\mathbb{F}_{q^m}$-linearly independent vectors in $\mathbb{F}_{q^m}^n$ has dimension k over $\mathbb{F}_{q^m}$ and km over $\mathbb{F}_q$. Similarly, using the T-product we can expand a code generated by k linearly independent matrices in $\mathbb{F}_q^{m\times n}$ to a code of dimension upper-bounded by km as follows.

Definition 5 (T-expanded code)**.** *Let $\mathcal{C} = \langle G_j \mid j \in [k]\rangle_{\mathbb{F}_q} \subseteq \mathbb{F}_q^{m\times n}$ be a matrix code of dimension k and T be a 3-tensor in $\mathbb{F}_q^{m\times m\times m}$. The T-**expanded code** of $\mathcal{C}$ is defined as*

$$\mathcal{C}_T = \langle T_{*,*,i} G_j \mid (i,j) \in [m]\times[k]\rangle_{\mathbb{F}_q}.$$

The dimension of $\mathcal{C}_T$ will be at most km.

We are ready to introduce the main topic of this paper.

Definition 6 (Generalized LRPC codes)**.** *Let T be a 3-tensor in $\mathbb{F}_q^{m\times m\times m}$ and let the matrices $H_1, \dots, H_{n-k} \in \mathbb{F}_q^{m\times n}$ satisfy the following two properties:*

1. *there exists a subspace $\mathcal{B} \subseteq \mathbb{F}_q^m$ such that $\mathrm{Colsp}(H_i) \subseteq \mathcal{B}, \forall i \in [n-k]$ and $\dim(\mathcal{B}) = d < m$.*
2. *there exists a base $\mathbf{b}_1, \dots, \mathbf{b}_d \in \mathbb{F}_q^m$ of $\mathcal{B}$ such that $T_{*,\mathbf{b}_i,*}$ is of full rank for all $i \in [d]$.*

The following matrix code

$$\mathcal{C} = \{C \in \mathbb{F}_q^{m\times n} \mid \mathrm{Tr}(T_{*,*,i} H_j C^\intercal) = 0, \forall (i,j) \in [m]\times[n-k]\}$$

is called a generalized LRPC code. Equivalently, let $\mathcal{H} = \langle H_i \mid i \in [n-k]\rangle_{\mathbb{F}_q}$, the code $\mathcal{C}$ is the dual of $\mathcal{H}_T$ given by

$$\mathcal{H}_T = \langle T_{*,*,i} H_j \mid (i,j) \in [m]\times[n-k]\rangle_{\mathbb{F}_q}.$$

Note that the right product $\cdot_T$ in general is not necessarily invertible for all the elements of $\mathbb{F}_q^m$. Hence the second condition needs to be verified. For the special product $\cdot_T$ given in Definition 2, which satisfy Proposition 1, the second condition is trivially satisfied by any base of any subspace $\mathcal{B}$.

Starting from the basis B of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, we can build a tensor T such that its associated T-product is $\mathbf{a}\cdot_T\mathbf{b} = \phi_B(ab)$. This specific T yields the classical LRPC codes. Meanwhile, there are a number of tensors T in $\mathbb{F}_q^{m\times m\times m}$ which result in codes inequivalent to the classical LRPC codes. It can be also shown that, starting from matrices $H_1, \dots, H_{n-k}$ as in Definition 6, when two tensors T and U satisfy certain relation, the resulting generalized LRPC codes $\mathcal{H}_T^\perp$ and $\mathcal{H}_U^\perp$ might be identical or isomorphic.

IV. Decoding of generalized LRPC codes

The decoding algorithm used for classical LRPC codes can be easily adapted to the generalized version. As in the case of classical LRPC codes the decoding algorithm will be probabilistic.

Let $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ be a generalized LRPC code of dimension mk. From Definition 6, there exists a code $\mathcal{H} = \langle H_1, \dots, H_{n-k}\rangle_{\mathbb{F}_q}$, where $\mathrm{Colsp}(H_i) \subseteq \mathcal{B} = \langle \mathbf{b}_1, \dots, \mathbf{b}_d\rangle_{\mathbb{F}_q}, \forall i \in [n-k]$ and a 3-tensor $T \in \mathbb{F}_q^{m\times m\times m}$ such that $(\mathcal{H}_T)^\perp = \mathcal{C}$. Moreover Definition 6 ensures us that $T_{*,\mathbf{b}_i,*} \in \mathbb{F}_q^{m\times m}$ is invertible for all $\mathbf{b}_i$.

A. Decoding Procedure

Suppose we receive the message $Y = C + E$ where $C \in \mathcal{C}$ and $E \in \mathbb{F}_q^{m\times n}$ is a matrix of low rank r. We can divide the decoding process into two steps. In the first step we will recover the column support of E. In this way we will be able to write $E = FX$ where $F = (\mathbf{f}_1, \dots, \mathbf{f}_r) \in \mathbb{F}_q^{m\times r}$ such that $\mathrm{Colsp}(F) = \mathrm{Colsp}(E) = \langle \mathbf{f}_1, \dots, \mathbf{f}_r\rangle_{\mathbb{F}_q}$ and $X \in \mathbb{F}_q^{r\times n}$ is a matrix of nr unknowns. In the second step we will solve a linear system in these nr unknowns.

Step 1. For $C \in \mathcal{C}$, from the definition we have that $C \cdot_T H_i = \mathbf{0}, \forall i \in [n-k]$. Therefore $Y \cdot_T H_i = (C+E) \cdot_T H_i = E \cdot_T H_i = \mathbf{s}_i$. Each column of E belongs to a subspace $\mathcal{E} = \langle \mathbf{f}_1, \dots, \mathbf{f}_r\rangle_{\mathbb{F}_q}$. Each column $\mathbf{h}_{i,j}$ of H_i belongs to $\mathcal{B} = \langle \mathbf{b}_1, \dots, \mathbf{b}_d\rangle_{\mathbb{F}_q}$. We denote by $\mathbf{e}_j$ the j-th column of E, since the T-product is bilinear, from Proposition 1 we have

$$\mathbf{s}_i = E \cdot_T H_i = \sum_{j\in[n]} \mathbf{e}_j \cdot_T \mathbf{h}_{i,j} \in \langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q}, \forall i \in [n-k] \quad (1)$$

where $\dim(\langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q}) \leq rd$. Letting $S = (\mathbf{s}_1, \dots, \mathbf{s}_{n-k})$ we will have $\mathrm{Colsp}(S) \subseteq \langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q}$, where the equality holds with a a good probability if $(n-k) \geq rd$.

To recover $\mathcal{E} = \mathrm{Colsp}(E)$ we can exploit the knowledge of a base of $\mathcal{B}$ over which the T-product is invertible. The space $\langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q}$ can be expressed as

$$\langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q} = \langle \mathbf{f}_i \cdot_T \mathbf{b}_j\rangle_{\mathbb{F}_q} = \langle \mathbf{f}_i T_{*,\mathbf{b}_j,*} \mid (i,j) \in [r]\times[d]\rangle_{\mathbb{F}_q}.$$

Notice that $\mathcal{E} \subseteq \langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q} (T_{*,\mathbf{b}_j,*})^{-1}$ for all $\mathbf{b}_j$. Therefore we have

$$\mathcal{E} \subseteq \bigcap_{j\in[d]} \langle \mathcal{E} \cdot_T \mathcal{B}\rangle_{\mathbb{F}_q} (T_{*,\mathbf{b}_j,*})^{-1}. \quad (2)$$

With a high probability the equality will hold and we will be able to recover $\mathcal{E}$.

Step 2 Assuming that the first step was successful, we obtained $\mathbf{f}_1, \ldots, \mathbf{f}_r$ which generates $\mathcal{E}$. We can collect them in a matrix $F = (\mathbf{f}_1, \ldots, \mathbf{f}_r) \in \mathbb{F}_q^{m \times r}$ and express the error as $E = FX$, where $X \in \mathbb{F}_q^{r \times n}$. Consider $\mathbf{s}_i = E \cdot_T H_i$, from Definition 4 its j-th component $\mathbf{s}_{i,j} = \mathrm{Tr}(T_{*,*,j} H_i E^\intercal) = \mathrm{Tr}(T_{*,*,j} H_i X^\intercal F^\intercal)$. For each $\mathbf{s}_{i,j}$ we get a linear equation in the nr variables contained in X. In total we will have $(n-k)m$ such linear equations in nr variables. It turns out that these equations are not linearly independent. We can get at most $(n-k)rd$ linearly independent equations. The space $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$ is generated by the rd vectors $\mathbf{z}_{k,l} = \mathbf{f}_k \cdot_T \mathbf{b}_l = \mathbf{f}_k T_{*,\mathbf{b}_l,*} \in \mathbb{F}_q^m$, let $Z = \{\mathbf{z}_{k,l} \mid (k,l) \in [r] \times [d]\}$ denote this set of generators. Each vector $\mathbf{s}_i = E \cdot_T H_i \in \langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$ can be expanded as

$$\mathbf{s}_i = \sum_{k=1}^{r} \sum_{l=1}^{d} \eta_{i,k,l} \mathbf{z}_{k,l}, \tag{3}$$

where $\eta_{i,k,l} \in \mathbb{F}_q$ are the coordinates of $\mathbf{s}_i$ with respect to the set of generator Z. Another way to express $\mathbf{s}_i$ is given by

$$\mathbf{s}_i = E \cdot_T H_i = \sum_{j=1}^{n} \mathbf{e}_j \cdot_T \mathbf{h}_{i,j} = \sum_{j=1}^{n} \mathbf{e}_j T_{*,\mathbf{h}_{i,j},*}, \tag{4}$$

where $\mathbf{e}_j$ is the j-th column of E and $\mathbf{h}_{i,j}$ is the j-th column of the matrix H_i. We have that $\mathbf{e}_j = \sum_{k=1}^{r} x_{k,j} \mathbf{f}_k$ and $\mathbf{h}_{i,j} = \sum_{l=1}^{d} \mu_{i,j,l} \mathbf{b}_l$, notice that $T_{*,\mathbf{h}_{i,j},*} = \sum_{l=1}^{d} \mu_{i,j,l} T_{*,\mathbf{b}_l,*}$. Substituting in (4) we obtain

$$\mathbf{s}_i = \sum_{j,k,l=1}^{n,r,d} x_{k,j} \mu_{i,j,l} (\mathbf{f}_k T_{*,\mathbf{b}_l,*}) = \sum_{j,k,l=1}^{n,r,d} x_{k,j} \mu_{i,j,l} \mathbf{z}_{k,l}. \tag{5}$$

From (3) and (5) we get the system of $(n-k)rd$ equations

$$\sum_{j=1}^{n} x_{k,j} \mu_{i,j,l} = \eta_{i,k,l}, \quad (i,k,l) \in [n-k] \times [r] \times [d]. \tag{6}$$

Finally, as in the case of classical LRPC codes, we have nr unknowns and $(n-k)rd$ equations, if $n \leq (n-k)d$ and at least nr of the equations in (6) are linearly independent, the system has a unique solution. If the system (6) has only $nr-a$ linearly independent equations the algorithm will give a list of q^a possible solutions.

B. Success probability

Similarly to the classical LRPC codes the algorithm for decoding generalized LRPC codes is not deterministic. In **Step 1** we have that $\mathrm{Colsp}(S) \subseteq \langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$. The space $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$ has dimension upper-bounded by rd. It could happen that, even if $n - k \geq rd$, the space $\mathrm{Colsp}(S)$ is strictly contained in $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$. Heuristically we can assume that the columns of S are vectors uniformly sampled from $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$. Under this assumption, the probability that a set of size $n-k \geq rd$ whose elements are extracted uniformly form a space $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$ of dimension rd spans the whole space $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}$ is given by [16]

$$P(\mathrm{Colsp}(S) = \langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}) = 1 - q^{rd-(n-k)}.$$

Notice that, in the case $\dim(\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}) = s < rd$, this probability improves to $1 - q^{s-(n-k)}$. The assumption that $\dim(\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q}) = rd$ is a worst case scenario.

Similarly to the classical LRPC codes, the second reason of failure in **Step 1** is given by the probability that the intersection of $\langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q} (T_{*,\mathbf{b}_i,*})^{-1}$ is not equal to $\mathcal{E}$. This probability can be approximated by the probability that d subspaces $\mathcal{R}_1, \ldots, \mathcal{R}_d \subseteq \mathbb{F}_q^m$ of dimension rd, each containing the same subspace $\mathcal{E}$ of dimension r, intersect in something bigger than $\mathcal{E}$. Assuming $\mathcal{R}_1, \ldots, \mathcal{R}_d$ are independently randomly chosen, the probability of their intersection to be bigger than $\mathcal{E}$ is given by $q^{-(d-1)(m-rd-r)}$ [6]. Considering these two possible reasons of failure, the success probability for **Step 1** will be lower bounded by $1 - (q^{rd-(n-k)} + q^{-(d-1)(m-rd-r)})$. Notice that, in the case $\mathcal{E} \subsetneq \bigcap_{i \in [d]} \langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q} (T_{*,\mathbf{b}_i,*})^{-1}$, it could still be possible to correct uniquely the error in some cases. Suppose $r < r' = \dim(\bigcap_{i \in [d]} \langle \mathcal{E} \cdot_T \mathcal{B} \rangle_{\mathbb{F}_q} (T_{*,\mathbf{b}_i,*})^{-1}$, the linear system in **Step 2** will have nr' unknowns and $(n-k)rd$ equations. If $nr' \leq (n-k)rd$ it will be still possible to uniquely recover the correct error.

$\langle q^m, n, k, d, r \rangle$	Product Space Recovery	Error Recovery
$\langle 2^{24}, 24, 5, 4, 3 \rangle$	99	99
$\langle 2^{20}, 20, 5, 4, 3 \rangle$	86	85
$\langle 2^{19}, 19, 5, 4, 3 \rangle$	76	75
$\langle 2^{17}, 24, 5, 4, 3 \rangle$	100	93
$\langle 2^{12}, 15, 5, 2, 3 \rangle$	93	83
$\langle 13^{13}, 15, 5, 2, 4 \rangle$	100	98

Table 1: Success rate of decoding Generalized LRPC codes

We run a test in Magma for 100 randomized trials for different sets of parameters, the success rate for each set of parameters are listed in Table 1. The numerical results in Table 1 appear to be in line with our heuristic analysis.

V. Conclusion

In this work we extend the important class of LRPC codes which have been used in different cryptographic schemes in recent years. Thanks to the bilinear product defined based on a 3-tensor over $\mathbb{F}_q$, the proposed $\mathbb{F}_q$-linear matrix codes allow for efficient decoding, which opens for potential applications of the codes. In addition, starting from the same set of matrices $H_1, \ldots H_{n-k}$ having all of their columns in a small subspace, it is possible to define many different codes depending on the choice of the tensor T. It is also interesting to investigate the relation between the codes defined by different tensors.

References

[1] E. Gabidulin, *Rank Codes*. TUM.University Press, 2021. [Online]. Available: https://mediatum.ub.tum.de/doc/1601193/1601193.pdf

[2] H. Bartz, L. Holzbaur, H. Liu, S. Puchinger, J. Renner, and A. Wachter-Zeh, "Rank-metric codes and their applications," 2022. [Online]. Available: https://arxiv.org/abs/2203.12384

[3] 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.

[4] 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. Springer, 1991, pp. 482–489.

[5] 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. Springer International Publishing, 2014, pp. 88–107.

[6] 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.

[7] N. Aragon, O. Blazy, P. Gaborit, A. Hauteville, and G. Zémor, *Durandal: A Rank Metric Based Signature Scheme*. Springer, 04 2019, pp. 728–758.

[8] N. T. Courtois, "Efficient zero-knowledge authentication based on a linear algebra problem minrank," 2001, https://eprint.iacr.org/2001/058. [Online]. Available: https://eprint.iacr.org/2001/058

[9] 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.

[10] 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. Springer International Publishing, 2017, pp. 194–224.

[11] R. Overbeck, "Structural attacks for public key cryptosystems based on gabidulin codes," *Journal of Cryptology*, vol. 21, pp. 280–301, 04 2008.

[12] A. V. Ourivski and T. Johansson, "New technique for decoding codes in the rank metric and its cryptography applications," *Probl. Inf. Transm.*, vol. 38, no. 3, pp. 237–246, 2002.

[13] M. Bardet, P. Briaud, M. Bros, P. Gaborit, V. Neiger, O. Ruatta, and J.-P. Tillich, "An algebraic attack on rank metric code-based cryptosystems," 10 2019.

[14] M. Bardet, M. Bros, D. Cabarcas, P. Gaborit, R. Perlner, D. Smith-Tone, J.-P. Tillich, and J. Verbel, "Algebraic attacks for solving the rank decoding and minrank problems without gröbner basis," *arXiv preprint arXiv:2002.08322*, 2020.

[15] M. Bardet, M. Bros, D. Cabarcas, P. Gaborit, R. Perlner, D. Smith, J.-P. Tillich, and J. Verbel, *Improvements of Algebraic Attacks for Solving the Rank Decoding and MinRank Problems*. Springer, 12 2020, pp. 507–536.

[16] 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."

[17] 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.

[18] J. Cruz, M. Kiermaier, A. Wassermann, and W. Willems, "Algebraic structures of mrd codes," *Advances in Mathematics of Communications*, vol. 10, 01 2015.

[19] D. E. Knuth, "Finite semifields and projective planes," *Journal of Algebra*, vol. 2, no. 2, pp. 182–217, 1965. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0021869365900189