

# Two Public-Key Cryptosystems Based on Expanded Gabidulin Codes

Wenshuo Guo<sup>\*</sup> and Fang-Wei Fu<sup>†</sup>

## Abstract

This paper presents two public-key cryptosystems based on the so-called expanded Gabidulin codes, which are constructed by expanding Gabidulin codes over the base field. Exploiting the fast decoder of Gabidulin codes, we propose an efficient algorithm to decode these new codes when the noise vector satisfies a certain condition. Additionally, these new codes have an excellent error-correcting capability because of the optimality of their parent Gabidulin codes. Based on different masking techniques, we give two encryption schemes by using expanded Gabidulin codes in the McEliece setting. According to our analysis, these two cryptosystems can both resist the existing structural attacks. Furthermore, our proposals also have an obvious advantage in public-key representation without using the cyclic or quasi-cyclic structure compared to some other code-based cryptosystems. To achieve the security of 256 bits, for instance, a public-key size of 37005 bytes is enough for our first proposal, while around 1044992 bytes are needed for Classic McEliece selected as a candidate of the third round of the NIST PQC project.

**Keywords** Post-quantum cryptography · Code-based cryptography · Rank metric codes · Expanded Gabidulin codes

## 1 Introduction

Over the past decades, cryptosystems based on coding theory have been drawing more and more attention due to the rapid development of quantum computers. The first code-based cryptosystem, known as McEliece cryptosystem [1] based on Goppa codes, was proposed by McEliece in 1978. The principle for McEliece's proposal is to first encode the plaintext with a random generator matrix of the distorted Goppa code and then add some random errors. Since then various studies [2–6] have been made to investigate the security of McEliece cryptosystem.

In addition to potential resistance against quantum computer attacks, McEliece cryptosystem has pretty fast encryption and decryption procedures. However, this system has never been used in practice due to its large public key size. To overcome this problem, some variants were proposed one

---

<sup>\*</sup>Wenshuo Guo is with the Chern Institute of Mathematics and LPMC, Nankai University, Tianjin 300071, China.  
E-mail: ws\_guo@mail.nankai.edu.cn

<sup>†</sup>Fang-Wei Fu is with the Chern Institute of Mathematics and LPMC, Nankai University, Tianjin 300071, China.  
E-mail: fwfu@nankai.edu.cn

after another. For instance, the authors in [7] proposed to use the automorphism groups of Goppa codes to build decodable error patterns of larger weight, which greatly enhances the system against decoding attacks. By doing this, smaller codes are allowed in the design of encryption schemes to reduce the public-key size. Unfortunately, this variant was shown to be vulnerable against the chosen-plaintext attacks proposed in [8]. In [9], the authors proposed the family of quasi-dyadic Goppa codes, which admit a very compact representation of parity-check or generator matrix, for efficiently designing syndrome-based cryptosystems. However, the authors in [10] mounted an efficient key-recovery attack against this variant for almost all the proposed parameters.

Besides endowing Goppa codes with some special structures, replacing Goppa codes with other families of codes is another approach to shorten the public keys. For instance, Niederreiter [11] introduced a knapsack-type cryptosystem based on generalized Reed-Solomon (GRS) codes. In Niederreiter’s proposal, the message sender first converts the plaintext into a vector of fixed weight and then multiplies it with a parity-check matrix of the public code. The advantage of GRS codes consists in their optimal error-correcting capability, which enables us to reduce the public-key size by exploiting codes with smaller parameters. However, this variant was proved to be insecure by Sidelnikov and Shestakov in [12] for the reason that GRS codes are highly structured. But if we use Goppa codes in the Niederreiter setting, it was proved to be equivalent to McEliece cryptosystem in terms of security [13]. To strengthen resistance against structural attacks, the authors in [14] performed a column-mixing transformation instead of a simple permutation to the underlying GRS code. According to their analysis, this variant could prevent some well-known attacks, such as Sidelnikov-Shestakov attack [12] and Wieschebrink’s attack [15]. However, in [16] the authors presented a polynomial key-recovery attack in some cases. Although one can adjust the parameters to prevent such an attack, it would introduce some other problems, such as the decryption complexity increasing dramatically and a higher request of error-correcting capability for the underlying code.

In [19] Gabidulin introduced a new family of rank metric codes, known as Gabidulin codes, which can be seen as an analogue of GRS codes in the rank metric. The particular appeal of rank metric codes is that the general decoding problem is much more difficult than that of Hamming metric codes [20, 21]. This inspires us to obtain a much smaller public-key size by building cryptosystems in the rank metric. In [22] the authors proposed the GPT cryptosystem by using Gabidulin codes in the McEliece setting, which requires a public-key size of only a few thousand bits for the security of 100 bits. Similar to the cryptosystems based on GRS codes, the GPT cryptosystem and some of its variants [23–26] have been subjected to many structural attacks [27–30]. Faure and Loidreau proposed another cryptosystem [31] that is quite different from the GPT proposal. The security of this scheme closely relates to the intractability of reconstructing linearized polynomials. Until the work in [32], the Faure-Loidreau scheme had never been severely attacked. In [33] Loidreau designed another rank-based cryptosystem in the McEliece setting, where a column-mixing transformation was imposed to the underlying code with the inverse of an invertible matrix whose entries are taken from an  $\mathbb{F}_q$ -subspace of  $\mathbb{F}_{q^m}$  of dimension  $\lambda$ . Loidreau claimed that his proposal could prevent all the existing structural attacks. However, this claim was proved to be invalidated by the authors in [34] for the case of  $\lambda = 2$  and the code rate being greater than  $1/2$ . Not long after this, the author in [35] generalized this attack to the case where  $\lambda > 2$  and the code rate is greater than  $1 - \frac{1}{\lambda}$ .

In [17], the authors introduced the concept of expanded GRS codes and designed an encryption scheme by using these codes in the Niederreiter setting. Our work in the present paper is inspired by this variant but uses the so-called expanded Gabidulin codes as the underlying codes. Benefitting

from the optimality of their parent Gabidulin codes, these new codes have excellent capability of correcting Hamming errors. This enables us to reduce the public-key size by exploiting smaller codes. Meanwhile, our proposals could resist the existing structural attacks such as Overbeck’s attack, and some potential attack according to our analysis.

The remaining part of this paper is arranged as follows. In Section 2, notations and some basic concepts about rank metric codes and Gabidulin codes will be given. In Section 3, we shall introduce the so-called expanded Gabidulin codes, investigate some of their algebraic properties and propose an efficient decoding algorithm for these new codes. Section 4 is devoted to the description of our two proposals constructed by using expanded Gabidulin codes in the McEliece setting. Section 5 will present the security analysis of our proposals, including structural attacks and generic attacks. In Section 6, we give some suggested parameters for different security levels and make a comparison on public-key sizes with some other code-based cryptosystems. Following this, we will make a few concluding remarks in Section 7.

## 2 Preliminaries

In this section we first introduce some notations used throughout this paper, and recall some basic concepts about linear codes and rank metric codes. After that, we will introduce the definition of Gabidulin codes and some related results.

### 2.1 Notations and basic concepts

Let  $q$  be a prime power. Denote by  $\mathbb{F}_q$  a finite field with  $q$  elements, and by  $\mathbb{F}_{q^m}$  an extension field of  $\mathbb{F}_q$  of degree  $m$ . For two positive integers  $k$  and  $n$ , let  $\mathcal{M}_{k,n}(\mathbb{F}_q)$  denote the space of all  $k \times n$  matrices over  $\mathbb{F}_q$ , and  $GL_n(\mathbb{F}_q)$  denote the general linear group formed by all invertible matrices in  $\mathcal{M}_{n,n}(\mathbb{F}_q)$ . For a matrix  $M \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ , let  $\langle M \rangle_{\mathbb{F}_{q^m}}$  be the vector space spanned by the rows of  $M$  over  $\mathbb{F}_{q^m}$ .

A linear code  $\mathcal{C}$  of length  $n$  and dimension  $k$  over  $\mathbb{F}_{q^m}$  is a  $k$ -dimensional subspace of  $\mathbb{F}_{q^m}^n$ . The dual code of  $\mathcal{C}$ , denoted by  $\mathcal{C}^\perp$ , is the orthogonal space of  $\mathcal{C}$  under the usual inner product over  $\mathbb{F}_{q^m}$ . A matrix  $G$  is called a generator matrix of  $\mathcal{C}$  if its row vectors form a basis of  $\mathcal{C}$ . A generator matrix  $H$  of  $\mathcal{C}^\perp$  is called a parity-check matrix of  $\mathcal{C}$ . For a codeword  $c \in \mathcal{C}$ , the Hamming weight of  $c$ , denoted by  $w_H(c)$ , is the number of nonzero components of  $c$ . The minimum Hamming distance of  $\mathcal{C}$ , denoted by  $d_H(\mathcal{C})$ , is defined as the minimum Hamming weight of nonzero codewords in  $\mathcal{C}$ . The minimum Hamming distance of  $\mathcal{C}$  has  $n - k + 1$  as an upper bound, and we call  $\mathcal{C}$  Maximum Distance Separable (MDS) when its minimum Hamming distance reaches to this bound.

### 2.2 Rank metric codes

Now we recall some basic concepts about rank metric codes.

*Definition 1.* Let  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{F}_{q^m}^n$ , the rank support of  $\mathbf{x}$ , denoted by  $\text{Supp}_R(\mathbf{x})$ , is defined to be the linear space spanned by components of  $\mathbf{x}$  over  $\mathbb{F}_q$ . Formally we have

$$\text{Supp}_R(\mathbf{x}) = \left\{ \sum_{i=1}^n \lambda_i x_i : \lambda_i \in \mathbb{F}_q, 1 \leq i \leq n \right\}.$$

*Definition 2.* For a vector  $\mathbf{x} \in \mathbb{F}_{q^m}^n$ , the rank weight of  $\mathbf{x}$ , denoted by  $w_R(\mathbf{x})$ , is defined to be the dimension of  $\text{Supp}_R(\mathbf{x})$  over  $\mathbb{F}_q$ . Formally we have

$$w_R(\mathbf{x}) = \dim_q (\text{Supp}_R(\mathbf{x})).$$

*Definition 3.* For two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{F}_{q^m}^n$ , the rank distance between  $\mathbf{x}$  and  $\mathbf{y}$ , denoted by  $d_R(\mathbf{x}, \mathbf{y})$ , is defined to be the rank weight of  $\mathbf{x} - \mathbf{y}$ . Formally we have

$$d_R(\mathbf{x}, \mathbf{y}) = w_R(\mathbf{x} - \mathbf{y}).$$

*Definition 4.* For a linear code  $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ , the minimum rank distance of  $\mathcal{C}$ , denoted by  $d_R(\mathcal{C})$ , is defined to be the minimum rank weight of nonzero codewords in  $\mathcal{C}$ . Formally we have

$$d_R(\mathcal{C}) = \min\{w_R(\mathbf{c}) : \mathbf{c} \in \mathcal{C} \text{ and } \mathbf{c} \neq \mathbf{0}\}.$$

A linear code endowed with the rank metric is called a rank metric code. Similar to Hamming metric codes, the minimum rank distance of a rank metric code is bounded from above by the Singleton-type bound defined as follows.

*Definition 5 (Singleton-type bound).* For positive integers  $k \leq n \leq m$ , let  $\mathcal{C}$  be an  $[n, k]$  rank metric code over  $\mathbb{F}_{q^m}$ , then the minimum rank distance of  $\mathcal{C}$  with respect to  $\mathbb{F}_q$  satisfies the following inequality

$$d_R(\mathcal{C}) \leq n - k + 1.$$

*Remark 1.* A rank metric code attaining the Singleton-type bound is called a Maximum Rank Distance (MRD) code. Let  $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$  be an  $[n, k]$  MRD code and  $\mathbf{c} \in \mathcal{C}$  be a nonzero codeword having the minimum Hamming weight. It is easy to see that  $n - k + 1 \leq w_R(\mathbf{c}) \leq w_H(\mathbf{c}) \leq n - k + 1$ . Hence we have  $w_H(\mathbf{c}) = n - k + 1$ , which enables us to conclude that an MRD code is MDS in the Hamming metric.

## 2.3 Gabidulin codes

Before giving the definition of Gabidulin codes, we shall introduce the concept of linearized polynomials. A linearized polynomial  $f(x) \in \mathbb{F}_{q^m}[x]$  is a polynomial of the form

$$f(x) = \sum_{i=0}^s a_i x^{q^i}, \text{ where } a_i \in \mathbb{F}_{q^m}.$$

The  $q$ -degree of  $f(x)$ , denoted by  $\deg_q(f)$ , is the largest  $i$  such that  $a_i \neq 0$ .

Let  $\mathcal{L}$  be the set of all linearized polynomials over  $\mathbb{F}_{q^m}$ . For a positive integer  $s$ , let  $\mathcal{L}_{<s}$  be the set of linearized polynomials of  $q$ -degree less than  $s$ , namely we have

$$\mathcal{L}_{<s} = \{f(x) \in \mathcal{L} : \deg_q(f) < s\}.$$

*Definition 6 (Gabidulin codes).* For positive integers  $k \leq n \leq m$  and a vector  $\mathbf{g} = (g_1, \dots, g_n) \in \mathbb{F}_{q^m}^n$  with  $w_R(\mathbf{g}) = n$ , the Gabidulin code  $\text{Gab}_{n,k}(\mathbf{g})$  generated by  $\mathbf{g}$  with length  $n$  and dimension  $k$  is defined to be

$$\text{Gab}_{n,k}(\mathbf{g}) = \{(f(g_1), \dots, f(g_n)) : f(x) \in \mathcal{L}_{<k}\}.$$

Equivalently, the Gabidulin code  $\text{Gab}_{n,k}(\mathbf{g})$  is defined to be a linear code having a generator matrix of the form

$$G = \begin{pmatrix} g_1 & g_2 & \cdots & g_n \\ g_1^q & g_2^q & \cdots & g_n^q \\ \vdots & \vdots & & \vdots \\ g_1^{q^{k-1}} & g_2^{q^{k-1}} & \cdots & g_n^{q^{k-1}} \end{pmatrix}.$$

Similar to GRS codes in the Hamming metric, Gabidulin codes admit an excellent error-correcting capability and simple algebraic structure. The following two theorems describe some properties of Gabidulin codes.

**Lemma 7.** [36] *A Gabidulin code is an MRD code. In other words, the minimum rank weight of  $\text{Gab}_{n,k}(\mathbf{g})$  attains the Singleton-type bound for rank metric codes.*

This implies that the Gabidulin code  $\text{Gab}_{n,k}(\mathbf{g})$  can theoretically correct up to  $\lfloor \frac{n-k}{2} \rfloor$  rank errors, which is an important reason for Gabidulin codes being widely used in the design of cryptosystems.

**Lemma 8.** [32] *The dual of a Gabidulin code is also a Gabidulin code. Particularly, we have  $\text{Gab}_{n,k}(\mathbf{g})^\perp = \text{Gab}_{n,n-k}(\mathbf{h}^{q^{-(n-k-1)}})$  for some  $\mathbf{h} \in \text{Gab}_{n,n-1}(\mathbf{g})^\perp$  with  $w_R(\mathbf{h}) = n$ , where  $\mathbf{h}^j$  denotes the  $j$ -th component-wise power of  $\mathbf{h}$ .*

### 3 Expanded Gabidulin codes

In this section, we first introduce the definition of expanded Gabidulin codes and then investigate some of their algebraic properties. After that, we will propose an efficient algorithm to decode these new codes.

#### 3.1 Expanded Gabidulin codes

Note that  $\mathbb{F}_{q^m}$  can be regarded as an  $\mathbb{F}_q$ -linear space of dimension  $m$ . Suppose that a subset of  $\mathbb{F}_{q^m}$ , say  $\mathcal{B} = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ , forms a basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ . For any  $\alpha \in \mathbb{F}_{q^m}$ , there exists  $(a_1, a_2, \dots, a_m) \in \mathbb{F}_q^m$  such that  $\alpha = \sum_{i=1}^m a_i \alpha_i$ . Based on this observation, we define an  $\mathbb{F}_q$ -linear isomorphism from  $\mathbb{F}_{q^m}$  to  $\mathbb{F}_q^m$  with respect to  $\mathcal{B}$  as follows

$$\begin{aligned} \phi_{\mathcal{B}} : \mathbb{F}_{q^m} &\mapsto \mathbb{F}_q^m, \\ \phi_{\mathcal{B}}(\alpha) &= (a_1, a_2, \dots, a_m). \end{aligned}$$

As for a vector  $\mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{F}_{q^m}^n$ , we define  $\phi_{\mathcal{B}}(\mathbf{v})$  to be the vector formed by performing the map  $\phi_{\mathcal{B}}$  to each component of  $\mathbf{v}$ , namely we have

$$\phi_{\mathcal{B}}(\mathbf{v}) = (\phi_{\mathcal{B}}(v_1), \phi_{\mathcal{B}}(v_2), \dots, \phi_{\mathcal{B}}(v_n)) \in \mathbb{F}_q^{mn}.$$

Theoretically we can always compute  $\phi_{\mathcal{B}}(\alpha)$  for any  $\alpha \in \mathbb{F}_{q^m}$ , but the authors in [17] did not specify how to do this in practice. Now we give an effective method to perform this operation, which is based on the following lemma.

**Lemma 9.** [18] Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_m\}$  be a basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ , then there exists a basis  $\mathcal{B}^* = \{\alpha_1^*, \dots, \alpha_m^*\}$  of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$  such that for  $1 \leq i, j \leq m$  we have

$$\mathrm{Tr}(\alpha_i \alpha_j^*) = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j, \end{cases} \quad (1)$$

where  $\mathrm{Tr}(\cdot)$  denotes the trace function  $\mathrm{Tr}(x) = \sum_{i=0}^{m-1} x^{q^i}$ .

*Remark 2.* A basis  $\mathcal{B}^*$  satisfying the condition (1) is called a dual basis of  $\mathcal{B}$ , and furthermore,  $\mathcal{B}^*$  is uniquely determined by  $\mathcal{B}$ . For any  $\alpha \in \mathbb{F}_{q^m}$ , assume that  $a_1, \dots, a_m \in \mathbb{F}_q$  satisfy  $\alpha = \sum_{i=1}^m a_i \alpha_i$ . Then we can obtain  $a_j$  for  $1 \leq j \leq m$  by computing

$$\mathrm{Tr}(\alpha \alpha_j^*) = \mathrm{Tr}\left(\sum_{i=1}^m a_i \alpha_i \alpha_j^*\right) = \sum_{i=1}^m a_i \mathrm{Tr}(\alpha_i \alpha_j^*) = a_j.$$

Finally we have  $\phi_{\mathcal{B}}(\alpha) = (\mathrm{Tr}(\alpha \alpha_1^*), \dots, \mathrm{Tr}(\alpha \alpha_m^*))$ .

Now we formally introduce the concept of expanded Gabidulin codes.

*Definition 10* (Expanded Gabidulin codes). Let  $\mathcal{G}$  be an  $[n, k]$  Gabidulin code over  $\mathbb{F}_{q^m}$ . For a basis  $\mathcal{B}$  of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ , let  $\phi_{\mathcal{B}}$  be the  $\mathbb{F}_q$ -linear isomorphism from  $\mathbb{F}_{q^m}$  to  $\mathbb{F}_q^m$  with respect to  $\mathcal{B}$ . The expanded code of  $\mathcal{G}$  induced by  $\phi_{\mathcal{B}}$  is defined as

$$\bar{\mathcal{G}} = \{\phi_{\mathcal{B}}(\mathbf{c}) : \mathbf{c} \in \mathcal{G}\}.$$

We call  $\mathcal{G}$  the parent Gabidulin code of  $\bar{\mathcal{G}}$ .

The following theorem gives a method of constructing a generator (parity-check) matrix of an expanded Gabidulin code when we already know a generator (parity-check) matrix of its parent Gabidulin code.

**Lemma 11.** Let  $\mathcal{G} \subseteq \mathbb{F}_{q^m}^n$  be an  $[n, k]$  Gabidulin code and  $\mathcal{B} = \{\alpha_1, \dots, \alpha_m\}$  be a basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ . The expanded code  $\bar{\mathcal{G}}$  of  $\mathcal{G}$  with respect to  $\phi_{\mathcal{B}}$  is a linear code of length  $mn$  and dimension  $mk$ . Furthermore, we have the following two conclusions.

(1) Suppose  $G = [\mathbf{g}_1, \dots, \mathbf{g}_k]^T$  is a generator matrix of  $\mathcal{G}$ , then  $\bar{\mathcal{G}}$  has an  $mk \times mn$  generator matrix of the form

$$\bar{G} = [\phi_{\mathcal{B}}(\alpha_1 \mathbf{g}_1), \dots, \phi_{\mathcal{B}}(\alpha_m \mathbf{g}_1), \dots, \phi_{\mathcal{B}}(\alpha_1 \mathbf{g}_k), \dots, \phi_{\mathcal{B}}(\alpha_m \mathbf{g}_k)]^T.$$

(2) Let  $H = [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_n^T]$  be a parity-check matrix of  $\mathcal{G}$ , then

$$\bar{H} = [\phi_{\mathcal{B}}(\alpha_1 \mathbf{h}_1)^T, \dots, \phi_{\mathcal{B}}(\alpha_m \mathbf{h}_1)^T, \dots, \phi_{\mathcal{B}}(\alpha_1 \mathbf{h}_n)^T, \dots, \phi_{\mathcal{B}}(\alpha_m \mathbf{h}_n)^T] \quad (2)$$

forms a parity-check matrix of  $\bar{\mathcal{G}}$ .

*Proof.* (1) Firstly, it is easy to verify that  $\bar{\mathcal{G}}$  forms a linear code over  $\mathbb{F}_q$ . Together with  $|\bar{\mathcal{G}}| = |\mathcal{G}| = q^{mk}$ , we have  $\dim_q(\bar{\mathcal{G}}) = \log_q(|\bar{\mathcal{G}}|) = mk$ . Apparently each row vector  $\phi_{\mathcal{B}}(\alpha_i \mathbf{g}_j)$  of  $\bar{G}$  ( $1 \leq i \leq m, 1 \leq j \leq k$ ) is contained in  $\bar{\mathcal{G}}$ . In the remaining part, it suffices to prove that  $\bar{G}$  is of full rank.

Suppose that there exists a vector

$$\mathbf{x} = (x_{11}, \dots, x_{m1}, \dots, x_{1k}, \dots, x_{mk}) \in \mathbb{F}_q^{mk}$$

such that  $\mathbf{x}\bar{G} = \mathbf{0}$ , then we have

$$\mathbf{x}\bar{G} = \sum_{j=1}^k \sum_{i=1}^m x_{ij} \phi_{\mathcal{B}}(\alpha_i \mathbf{g}_j) = \phi_{\mathcal{B}}\left(\sum_{j=1}^k \sum_{i=1}^m x_{ij} \alpha_i \mathbf{g}_j\right) = \phi_{\mathcal{B}}\left(\sum_{j=1}^k \beta_j \mathbf{g}_j\right) = \mathbf{0},$$

where  $\beta_j = \sum_{i=1}^m x_{ij} \alpha_i$  ( $1 \leq j \leq k$ ).

By performing the inverse of  $\phi_{\mathcal{B}}$  to both sides of the last equation, we can obtain  $\sum_{j=1}^k \beta_j \mathbf{g}_j = (\beta_1, \dots, \beta_k)G = \mathbf{0}$ . This implies that  $\beta_j = 0$  for  $1 \leq j \leq k$ . Further more, for  $1 \leq j \leq k$  we have  $x_{ij} = 0$  ( $1 \leq i \leq m$ ) because of  $\alpha_1, \dots, \alpha_m$  being a basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ . Finally, we come to the conclusion that  $\mathbf{x}\bar{G} = \mathbf{0}$  holds if and only if  $\mathbf{x} = \mathbf{0}$ , which means that  $\bar{G}$  is of full rank.

(2) We first show that each row vector of  $\bar{H}$  is contained in the dual code of  $\bar{G}$ . For any  $\bar{\mathbf{c}} = (c_{11}, \dots, c_{m1}, \dots, c_{1n}, \dots, c_{mn}) \in \bar{\mathcal{G}}$ , there exists  $\mathbf{c} = (c_1, \dots, c_n) \in \mathcal{G}$  such that  $\bar{\mathbf{c}} = \phi_{\mathcal{B}}(\mathbf{c})$ . By  $\mathbf{c}H^T = \mathbf{0}$ , we have

$$\bar{\mathbf{c}}\bar{H}^T = \sum_{j=1}^n \sum_{i=1}^m c_{ij} \phi_{\mathcal{B}}(\alpha_i \mathbf{h}_j) = \phi_{\mathcal{B}}\left(\sum_{j=1}^n \sum_{i=1}^m c_{ij} \alpha_i \mathbf{h}_j\right) = \phi_{\mathcal{B}}\left(\sum_{j=1}^n c_j \mathbf{h}_j\right) = \phi_{\mathcal{B}}(\mathbf{c}H^T) = \mathbf{0}.$$

This implies that each row vector of  $\bar{H}$  is contained in  $\bar{\mathcal{G}}^\perp$ .

It remains to prove that  $\bar{H}$  is of full rank. To do this, it suffices to prove that there exist  $m(n-k)$  columns of  $\bar{H}$  linearly independent over  $\mathbb{F}_q$ . Without loss of generality, we consider the first  $m(n-k)$  columns of  $\bar{H}$  and let  $\bar{H}_s$  be a submatrix of  $\bar{H}$  formed by these columns. Suppose that

$$\mathbf{x} = (x_{11}, \dots, x_{m1}, \dots, x_{1r}, \dots, x_{mr}) \in \mathbb{F}_q^{mr}$$

satisfies  $\mathbf{x}\bar{H}_s^T = \mathbf{0}$  where  $r = n - k$ , then we have

$$\mathbf{x}\bar{H}_s^T = \sum_{j=1}^r \sum_{i=1}^m x_{ij} \phi_{\mathcal{B}}(\alpha_i \mathbf{h}_j) = \phi_{\mathcal{B}}\left(\sum_{j=1}^r \sum_{i=1}^m x_{ij} \alpha_i \mathbf{h}_j\right) = \phi_{\mathcal{B}}\left(\sum_{j=1}^r \beta_j \mathbf{h}_j\right) = \mathbf{0},$$

where  $\beta_j = \sum_{i=1}^m x_{ij} \alpha_i$  ( $1 \leq j \leq r$ ).

By performing the inverse of  $\phi_{\mathcal{B}}$  to both sides of the last equation, we can obtain  $\sum_{j=1}^r \beta_j \mathbf{h}_j = (\beta_1, \dots, \beta_r)H_s^T = \mathbf{0}$ , where  $H_s$  is the first  $r$  columns of  $H$ . Since Gabidulin codes are MDS in the Hamming metric, any  $r$  columns of  $H$  are linearly independent over  $\mathbb{F}_{q^m}$ . Hence  $\beta_j = 0$  ( $1 \leq j \leq r$ ) and furthermore we have  $\mathbf{x} = \mathbf{0}$  because of  $\alpha_1, \dots, \alpha_m$  being linearly independent over  $\mathbb{F}_q$ . This implies that  $\bar{H}_s$  is of full rank and hence the conclusion is proved.  $\square$

We already know that Gabidulin codes are optimal in both the Hamming metric and rank metric. However, expanded Gabidulin codes are far from optimal in the Hamming metric according to our analysis. Specially, we have the following proposition.

**Proposition 1.** *Let  $\mathcal{G}$  be an  $[n, k]$  Gabidulin code over  $\mathbb{F}_{q^m}$ . For a given basis  $\mathcal{B}$  of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ , denote by  $\bar{\mathcal{G}}$  the expanded code of  $\mathcal{G}$  under the  $\mathbb{F}_q$ -linear isomorphism  $\phi_{\mathcal{B}}$ . Then the minimum Hamming distance of  $\bar{\mathcal{G}}$  satisfies the following inequality*

$$n - k + 1 \leq d_H(\bar{\mathcal{G}}) \leq m(n - k) + 1.$$

*In particular, with a proper choice of  $\mathcal{B}$ , the minimum Hamming distance of  $\bar{\mathcal{G}}$  can reach to  $n - k + 1$ .*

*Proof.* For any  $\bar{\mathbf{u}} \in \bar{\mathcal{G}}$ , there exists  $\mathbf{u} = (u_1, \dots, u_n) \in \mathcal{G}$  such that  $\bar{\mathbf{u}} = \phi_{\mathcal{B}}(\mathbf{u})$ . Since  $\mathcal{G}$  is MDS in the Hamming metric, there must be  $w_H(\mathbf{u}) \geq n - k + 1$  for a nonzero  $\mathbf{u}$ . Let  $I = \{1 \leq i \leq n : u_i \neq 0\}$ , then  $|I| \geq n - k + 1$ . Apparently  $w_H(\bar{\mathbf{u}}) = \sum_{i \in I} w_H(\phi_{\mathcal{B}}(u_i)) \geq n - k + 1$  because of  $w_H(\phi_{\mathcal{B}}(\alpha)) \geq 1$  for any  $\alpha \in \mathbb{F}_{q^m}^*$ . Hence we have  $d_H(\bar{\mathcal{G}}) \geq n - k + 1$ . On the other hand, by the Singleton bound for Hamming metric codes, it is easy to see that  $d_H(\bar{\mathcal{G}}) \leq m(n - k) + 1$ .

Let  $\mathbf{v} = (v_1, \dots, v_n) \in \mathcal{G}$  with  $w_H(\mathbf{v}) = n - k + 1$ , and let  $S = \{v_i \neq 0 : 1 \leq i \leq n\}$ . If  $S \subseteq \mathcal{B}$ , then  $w_H(\phi_{\mathcal{B}}(\mathbf{v})) = \sum_{v_i \neq 0} w_H(\phi_{\mathcal{B}}(v_i)) = n - k + 1$  because of  $w_H(\phi_{\mathcal{B}}(\alpha)) = 1$  for any  $\alpha \in \mathcal{B}$ . This implies that  $d_H(\bar{\mathcal{G}}) = n - k + 1$ .  $\square$

### 3.2 Decoding expanded Gabidulin codes

As for Gabidulin codes, several efficient decoding algorithms already exist [19, 37, 38]. In this part, we mainly study the decoding problem of expanded Gabidulin codes. Our analysis shows that when the noise vector satisfies a certain condition, decoding an expanded Gabidulin code can be converted into decoding its parent Gabidulin code.

Let  $\mathcal{G} \subseteq \mathbb{F}_{q^m}^n$  be an  $[n, k]$  Gabidulin code having  $H$  as a parity-check matrix. Denote by  $\bar{\mathcal{G}}$  an expanded code of  $\mathcal{G}$  induced by some  $\mathbb{F}_q$ -linear isomorphism  $\phi_{\mathcal{B}}$ . Suppose  $\mathbf{y} = \mathbf{c} + \mathbf{e}$  is the received word, where  $\mathbf{c} \in \bar{\mathcal{G}}$  and  $\mathbf{e} = (e_1, \dots, e_n) \in \mathbb{F}_q^{mn}$  is the noise vector with  $e_j = (e_{1j}, \dots, e_{mj}) \in \mathbb{F}_q^m$  ( $1 \leq j \leq n$ ). Let  $E$  be an  $n \times m$  matrix whose  $j$ -th row vector is  $e_j$ , called the error matrix corresponding to  $\mathbf{e}$ . If the following inequality holds

$$\text{Rank}(E) \leq \lfloor \frac{n-k}{2} \rfloor,$$

then we say  $\mathbf{e}$  satisfies the decodable condition. In this situation, we can obtain a fast decoder  $\mathcal{D}_{\bar{\mathcal{G}}}$  for  $\bar{\mathcal{G}}$  to decode  $\mathbf{y}$  by exploiting the syndrome decoding procedure of  $\mathcal{G}$ .

Denote by  $\bar{H}$  a parity-check matrix of  $\bar{\mathcal{G}}$ . It is easy to see that

$$\mathbf{y}\bar{H}^T = \mathbf{e}\bar{H}^T = \sum_{j=1}^n \sum_{i=1}^m e_{ij} \phi_{\mathcal{B}}(\alpha_i \mathbf{h}_j) = \phi_{\mathcal{B}}\left(\sum_{j=1}^n \sum_{i=1}^m e_{ij} \alpha_i \mathbf{h}_j\right) = \phi_{\mathcal{B}}\left(\sum_{j=1}^n e_j^* \mathbf{h}_j\right) = \phi_{\mathcal{B}}(\mathbf{e}^* H^T),$$

where  $\mathbf{e}^* = (e_1^*, \dots, e_n^*)$  with  $e_j^* = \sum_{i=1}^m e_{ij} \alpha_i$  ( $1 \leq j \leq n$ ). Since  $w_R(\mathbf{e}^*) = \text{Rank}(E) \leq \lfloor \frac{n-k}{2} \rfloor$ , applying the decoding procedure of  $\mathcal{G}$  to  $\phi_{\mathcal{B}}^{-1}(\mathbf{y}\bar{H}^T) = \mathbf{e}^* H^T$  will lead to  $\mathbf{e}^*$ . Hence we can recover  $\mathbf{e}$  by computing  $\phi_{\mathcal{B}}(\mathbf{e}^*)$  and then the codeword  $\mathbf{c}$  can be computed as  $\mathbf{y} - \mathbf{e}$ .

Apparently four steps are needed to decode expanded Gabidulin codes. Firstly, we shall compute the syndrome of the received word  $\mathbf{y}$ , which requires an operation of multiplying  $\mathbf{y}$  and  $\bar{H}^T$  together with a complexity of  $\mathcal{O}(m^2 n(n-k))$  in  $\mathbb{F}_q$ . Secondly, we shall perform the inverse transformation of  $\phi_{\mathcal{B}}$  to the syndrome obtained in the first step, requiring a complexity of  $\mathcal{O}(mn)$  in  $\mathbb{F}_{q^m}$ . The third step shall call the fast decoder of the parent Gabidulin code to obtain an error vector  $\mathbf{e}^*$  with  $w_R(\mathbf{e}^*) \leq \lfloor \frac{n-k}{2} \rfloor$ , which requires a complexity of  $\mathcal{O}(\frac{5}{2}n^2 - \frac{3}{2}k^2)$  in  $\mathbb{F}_{q^m}$  [37]. In the last step, we shall compute  $\phi_{\mathcal{B}}(\mathbf{e}^*)$  through the method described in Remark 2 with a complexity of  $\mathcal{O}((m-1)(q-1) + 1)mn$  in  $\mathbb{F}_{q^m}$ . Finally the total complexity of decoding expanded Gabidulin codes is  $\mathcal{O}(m^2 n(q-1) + mn(3-q) + \frac{5}{2}n^2 - \frac{3}{2}k^2)$  in  $\mathbb{F}_{q^m}$  plus  $\mathcal{O}(m^2 n(n-k))$  in  $\mathbb{F}_q$ .



## 4 Description of our proposals

Now we give a formal description of our two cryptosystems.

### 4.1 Proposal I

Let  $\mathcal{G} \subseteq \mathbb{F}_q^n$  be an  $[n, k]$  Gabidulin code, correcting up to  $t = \lfloor \frac{n-k}{2} \rfloor$  rank errors. Given a basis  $\mathcal{B} = \{\alpha_1, \dots, \alpha_m\}$  of  $\mathbb{F}_q^m$  over  $\mathbb{F}_q$ , let  $\phi_{\mathcal{B}}$  be the  $\mathbb{F}_q$ -linear isomorphism with respect to  $\mathcal{B}$ . Let  $\bar{\mathcal{G}}$  be the expanded code of  $\mathcal{G}$  induced by  $\phi_{\mathcal{B}}$ . Denote by  $\bar{H}$  a parity-check matrix of  $\bar{\mathcal{G}}$  of the form (2).

For a positive integer  $\lambda$  with  $\frac{m(n-k)}{n} < \lambda < m$ , we define  $I_j = \{mj + 1, \dots, mj + \lambda\}$  for  $0 \leq j \leq n-1$ . Let  $S = \cup_{j=0}^{n-1} I_j$  and denote by  $\bar{H}_S$  the submatrix of  $\bar{H}$  being restricted to  $S$ . Let  $\bar{\mathcal{G}}_S$  be a linear code that has  $\bar{H}_S$  as a parity-check matrix and denote by  $\bar{G}_S$  a generator matrix of  $\bar{\mathcal{G}}_S$ . It is easy to see that  $\bar{\mathcal{G}}_S$  has length  $N = \lambda n$ . As for the dimension of  $\bar{\mathcal{G}}_S$ , we have  $K = \lambda n - m(n-k)$  holds with high probability according to our experiments on Magma.

- Key generation

Randomly choose a matrix  $A \in GL_{\lambda}(\mathbb{F}_q)$ . Denote by  $I_n$  the identity matrix of order  $n$ , and let  $T$  be the tensor product  $I_n \otimes A$ , namely  $T$  is a quasi-diagonal matrix of order  $n\lambda$  that has  $A$  as the diagonal block. Randomly choose a matrix  $M \in GL_K(\mathbb{F}_q)$  such that  $G_{pub} = M\bar{G}_S T^{-1}$  is of systematic form. We publish  $(G_{pub}, t)$  as the public key, and keep  $(\bar{H}_S, A, \mathcal{D}_{\bar{\mathcal{G}}})$  as the secret key.

- Encryption

For a plaintext  $\mathbf{x} \in \mathbb{F}_q^K$ , randomly choose a matrix  $E \in \mathcal{M}_{n,\lambda}(\mathbb{F}_q)$  with  $\text{Rank}(E) = t$ . Let  $\mathbf{e} = (e_1, \dots, e_n) \in \mathbb{F}_q^N$ , where  $e_i$  is the  $i$ -th row vector of  $E$ . The ciphertext corresponding to  $\mathbf{x}$  is computed as  $\mathbf{y} = \mathbf{x}G_{pub} + \mathbf{e}$ .

- Decryption

For a ciphertext  $\mathbf{y} \in \mathbb{F}_q^N$ , let  $\mathbf{e}' = \mathbf{e}T$  and compute

$$\mathbf{s} = \mathbf{y}T\bar{H}_S^T = \mathbf{x}M\bar{G}_S T^{-1}T\bar{H}_S^T + \mathbf{e}'T\bar{H}_S^T = \mathbf{e}'\bar{H}_S^T.$$

Applying the fast decoder  $\mathcal{D}_{\bar{\mathcal{G}}}$  of  $\bar{\mathcal{G}}$  to  $\mathbf{s}$  will lead to a vector  $\mathbf{e}'' \in \mathbb{F}_q^{mn}$ . The restriction of  $\mathbf{e}''$  to  $S$  will be  $\mathbf{e}'$ , then we can recover  $\mathbf{e}$  by computing  $\mathbf{e}'T^{-1}$ . The plaintext will be the restriction of  $\mathbf{y} - \mathbf{e}$  to the first  $K$  coordinates.

**Correctness of Decryption.** Let  $\mathbf{e}' = (e'_1, \dots, e'_n)$ , where  $e'_i = e_i A$ . Define  $E' \in \mathcal{M}_{n,\lambda}(\mathbb{F}_q)$  to be a matrix whose  $i$ -th row vector is  $e'_i$ . Let  $\mathbf{e}'' = (e''_1, \dots, e''_n)$ , where  $e''_i = (e'_i, \mathbf{0})$  and  $\mathbf{0}$  denotes the zero vector of length  $m - \lambda$ . Define  $E'' \in \mathcal{M}_{n,m}(\mathbb{F}_q)$  to be a matrix whose  $i$ -th row vector is  $e''_i$ . It is easy to see that

$$E'' = [E'|O] = [EA|O],$$

where  $O$  denotes the  $n \times (m - \lambda)$  zero matrix. Apparently we have

$$\text{Rank}(E'') = \text{Rank}(E') = \text{Rank}(E) = t,$$

which implies that  $\mathbf{e}''$  satisfies the decodable condition described in Section 3.2. Applying the fast decoder of  $\bar{\mathcal{G}}$  to  $\mathbf{s} = \mathbf{e}'\bar{H}_S^T = \mathbf{e}''\bar{H}^T$  will lead to  $\mathbf{e}''$ , then the restriction of  $\mathbf{e}''$  to  $S$  will be  $\mathbf{e}'$ .

## 4.2 Proposal II

Let  $\mathcal{G} \subseteq \mathbb{F}_q^n$  be an  $[n, k]$  Gabidulin code, correcting up to  $\lfloor \frac{n-k}{2} \rfloor$  rank errors. For an  $\mathbb{F}_q$ -linear isomorphism  $\phi_{\mathcal{B}}$  from  $\mathbb{F}_q^m$  to  $\mathbb{F}_q^m$ , let  $\bar{\mathcal{G}}$  denote the expanded code of  $\mathcal{G}$  induced by  $\phi_{\mathcal{B}}$ . According to Lemma 11,  $\bar{\mathcal{G}}$  is a linear code of length  $N = mn$  and dimension  $K = mk$ . Denote by  $\bar{H}$  a parity-check matrix of  $\bar{\mathcal{G}}$ , and by  $\bar{G}$  a generator matrix respectively. For a positive integer  $\lambda \ll n$ , let  $u_f = \lfloor \frac{n}{\lambda} \rfloor$ ,  $u_c = \lceil \frac{n}{\lambda} \rceil$  and  $v = n - \lambda u_f$ .

- **Key generation**

Randomly choose an invertible matrix  $A \in GL_{m\lambda}(\mathbb{F}_q)$  such that the  $mv \times mv$  submatrix  $A_{sub}$  in the top left corner of  $A$  is also invertible. Set

$$T = \begin{pmatrix} A_{ten} & \\ & A_{sub} \end{pmatrix} \in GL_N(\mathbb{F}_q),$$

where  $A_{ten}$  is the tensor product  $I_{u_f} \otimes A$ . Randomly choose an invertible matrix  $M \in GL_K(\mathbb{F}_q)$  such that  $G_{pub} = M\bar{G}T^{-1}$  is of systematic form. Then we publish  $(G_{pub}, t)$  as the public key where  $t = \lfloor \frac{n-k}{2\lambda} \rfloor$ , and keep  $(\bar{H}, A, \mathcal{D}_{\bar{\mathcal{G}}})$  as the secret key.

- **Encryption**

For a plaintext  $\mathbf{x} \in \mathbb{F}_q^K$ , randomly choose a matrix  $E \in \mathcal{M}_{u_c, m\lambda}(\mathbb{F}_q)$  of the form

$$E = \begin{pmatrix} \mathbf{e}_1 & \cdots & \cdots & \mathbf{e}_{\lambda-1} & \mathbf{e}_{\lambda} \\ \mathbf{e}_{1+\lambda} & \cdots & \cdots & \mathbf{e}_{2\lambda-1} & \mathbf{e}_{2\lambda} \\ \vdots & & & \vdots & \vdots \\ \mathbf{e}_{1+(u_c-1)\lambda} & \cdots & \mathbf{e}_n & \cdots & \mathbf{0} \end{pmatrix} \quad (3)$$

with  $\text{Rank}(E) = t$ , where  $\mathbf{e}_i \in \mathbb{F}_q^m$  ( $1 \leq i \leq n$ ). Let  $\mathbf{e} = (\mathbf{e}_1, \dots, \mathbf{e}_n) \in \mathbb{F}_q^N$ , then the ciphertext corresponding to  $\mathbf{x}$  is computed as  $\mathbf{y} = \mathbf{x}G_{pub} + \mathbf{e}$ .

- **Decryption**

For a ciphertext  $\mathbf{y} \in \mathbb{F}_q^N$ , compute  $\mathbf{s} = \mathbf{y}T\bar{H}^T = \mathbf{e}T\bar{H}^T$ . Applying the syndrome decoding procedure  $\mathcal{D}_{\bar{\mathcal{G}}}$  of  $\bar{\mathcal{G}}$  to  $\mathbf{s}$  will lead to  $\mathbf{e}' = \mathbf{e}T$ , then we can recover  $\mathbf{e}$  by computing  $\mathbf{e}'T^{-1}$ . The restriction of  $\mathbf{y} - \mathbf{e}$  to the first  $K$  coordinates will be the plaintext.

**Correctness of Decryption.** Before proving the correctness, we first introduce the following proposition.

**Proposition 2.** Given  $\lambda$  matrices  $M_1, M_2, \dots, M_{\lambda} \in \mathcal{M}_{u,v}(\mathbb{F}_q)$ , let

$$M = [M_1, M_2, \dots, M_{\lambda}] \text{ and } M' = \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_{\lambda} \end{pmatrix}.$$

Suppose  $\text{Rank}(M) = t$ , then there must be  $\text{Rank}(M') \leq \lambda t$ .

*Proof.* Since  $\text{Rank}(M) = t$ , there must be  $\text{Rank}(M_i) \leq t$  for  $1 \leq i \leq \lambda$ . Hence we have  $\text{Rank}(M') \leq \sum_{i=1}^{\lambda} \text{Rank}(M_i) \leq \lambda t$ .  $\square$

Now we return to the proof. From the decrypting process of Proposal II, it suffices to prove that  $e'$  satisfies the decodable condition described in Section 3.2. Let  $e' = (e'_1, e'_2, \dots, e'_n)$  where  $e'_i \in \mathbb{F}_q^m$  ( $1 \leq i \leq n$ ), then there exist  $e'_{n+1}, \dots, e'_{u_c \lambda} \in \mathbb{F}_q^m$  such that

$$E' = \begin{pmatrix} e'_1 & \cdots & \cdots & e'_{\lambda-1} & e'_\lambda \\ e'_{1+\lambda} & \cdots & \cdots & e'_{2\lambda-1} & e'_{2\lambda} \\ \vdots & & & \vdots & \vdots \\ e'_{1+(u_c-1)\lambda} & \cdots & e'_n & \cdots & e'_{u_c \lambda} \end{pmatrix} = EA = [E'_1 \ E'_2 \ \cdots \ E'_\lambda],$$

where  $E'_i \in \mathcal{M}_{u_c, m}(\mathbb{F}_q)$  ( $1 \leq i \leq \lambda$ ). Apparently we have  $\text{Rank}(E') = \text{Rank}(E) = t$ . Let

$$F = \begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_n \end{pmatrix} \text{ and } F' = \begin{pmatrix} E'_1 \\ E'_2 \\ \vdots \\ E'_\lambda \end{pmatrix},$$

then  $\text{Rank}(F) \leq \text{Rank}(F') \leq \lambda t \leq \lfloor \frac{n-k}{2} \rfloor$  because of Proposition 2.

*Remark 3.* The cryptosystem presented above aims at the general situation where  $\lambda$  does not divide  $n$ , or equivalently  $u_f \neq u_c$ . As for the case of  $u_f = u_c$ , just a few changes are needed in the key generation procedure. To generate the column scrambling matrix  $T^{-1}$ , any non-singular matrix  $A \in GL_{m\lambda}(\mathbb{F}_q)$  is feasible for computing  $T = I_{u_f} \otimes A$ .

## 5 Security analysis

This section mainly discusses the security of the proposed cryptosystems. Attacks on code-based cryptosystems can be divided into two categories, namely the structural attack and generic attack. A structural attack aims to recover the structure of the underlying code from the published information, which amounts to recovering the secret key or its equivalent form that can be used to decrypt any valid ciphertext in polynomial time. Therefore, a code-based cryptosystem is said to be completely broken if there exists a polynomial-time structural attack on it. A generic attack is to recover the plaintext directly without knowing the secret key, which implies that one has to deal with the underlying computational problem. This type of attack usually requires exponential time, which reveals the security level of a code-based cryptosystem under current conditions.

### 5.1 A distinguisher for Gabidulin codes

Before giving the analysis, we introduce the so-called Frobenius transformation and some algebraic properties of Gabidulin codes under this transformation that will be useful to explain why our proposals can prevent the related structural attacks.

For a non-negative integer  $i$ , denote by  $[i]$  the  $i$ -th Frobenius power  $q^i$ , namely  $[i] = q^i$ . Under this notation, the element  $\alpha^{q^i}$  can be simply written as  $\alpha^{[i]}$  for any  $\alpha \in \mathbb{F}_{q^m}$ . Generalizing this transformation to a vector  $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{F}_{q^m}^n$ , there will be  $\mathbf{v}^{[i]} = (v_1^{[i]}, \dots, v_n^{[i]})$ . As for a linear code  $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ , the  $i$ -th Frobenius power of  $\mathcal{C}$  is defined as  $\mathcal{C}^{[i]} = \{\mathbf{c}^{[i]} : \mathbf{c} \in \mathcal{C}\}$ .

Now we introduce the following propositions without proof. These propositions provide us with a method of distinguishing Gabidulin codes from general ones.

**Proposition 3.** [32] Let  $\mathcal{G} \subseteq \mathbb{F}_{q^m}^n$  be an  $[n, k]$  Gabidulin code. In terms of the intersection of  $\mathcal{G}$  and its Frobenius power  $\mathcal{G}^{[1]}$ , we have

$$\dim(\mathcal{G} \cap \mathcal{G}^{[1]}) = k - 1.$$

**Proposition 4.** [32] Let  $\mathcal{G} \subseteq \mathbb{F}_{q^m}^n$  be an  $[n, k]$  Gabidulin code. For any positive integer  $i$ , the following equality holds

$$\dim(\mathcal{G} + \mathcal{G}^{[1]} + \dots + \mathcal{G}^{[i]}) = \min\{n, k + i\}.$$

**Proposition 5.** [32] Let  $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$  be an  $[n, k]$  random linear code. For any positive integer  $i$ , the following equality holds with high probability

$$\dim(\mathcal{C} + \mathcal{C}^{[1]} + \dots + \mathcal{C}^{[i]}) = \min\{n, k(i + 1)\}.$$

## 5.2 Structural attacks

**Existing attacks.** Most cryptosystems based on Gabidulin codes have been proved to be insecure due to their vulnerability against structural attacks, such as Overbeck's attack [28], Coggia-Couvreur attack [34] and the attack proposed in [32]. Although these attacks were designed to cryptanalyze different variants, most of them rely on the fact that one can distinguish Gabidulin codes from general ones by observing how their dimensions behave under the Frobenius mapping according to Propositions 3, 4 and 5. However, this observation is no longer valid when considering our proposals. Since our proposals are built over the base field  $\mathbb{F}_q$ , apparently we have  $\bar{\mathcal{G}}^{[i]} = \bar{\mathcal{G}}$  for any positive integer  $i$ . In this situation, Gabidulin codes will behave the same as general linear codes in terms of dimensions. Hence it is reasonable to conclude that all these attacks do not work on our proposals.

**A potential attack.** Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_m\} \subset \mathbb{F}_{q^m}$  be a basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ . Denote by  $\phi_{\mathcal{B}}$  the  $\mathbb{F}_q$ -linear isomorphism from  $\mathbb{F}_{q^m}^m$  to  $\mathbb{F}_{q^m}^m$  with respect to  $\mathcal{B}$ . For an  $[n, k]$  Gabidulin code  $\mathcal{G}$  over  $\mathbb{F}_{q^m}$ , denote by  $\bar{\mathcal{G}}$  the expanded code of  $\mathcal{G}$  induced by  $\phi_{\mathcal{B}}$ .

Let  $\mathcal{B}^* = \{\alpha_1^*, \dots, \alpha_m^*\} \subset \mathbb{F}_{q^m}$  be another basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ . Define  $\hat{\phi}_{\mathcal{B}^*}$  to be an  $\mathbb{F}_q$ -linear isomorphism from  $\mathbb{F}_q^m$  to  $\mathbb{F}_{q^m}$  as follows

$$\begin{aligned} \hat{\phi}_{\mathcal{B}^*} : \mathbb{F}_q^m &\mapsto \mathbb{F}_{q^m}, \\ \hat{\phi}_{\mathcal{B}^*}(\mathbf{a}) &= \sum_{i=1}^m a_i \alpha_i^*, \end{aligned}$$

where  $\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{F}_q^m$ . For a vector  $\mathbf{v} = (\mathbf{a}_1, \dots, \mathbf{a}_n) \in \mathbb{F}_q^{mn}$  where  $\mathbf{a}_i \in \mathbb{F}_q^m$  ( $1 \leq i \leq n$ ), we define  $\hat{\phi}_{\mathcal{B}^*}(\mathbf{v}) = (\hat{\phi}_{\mathcal{B}^*}(\mathbf{a}_1), \dots, \hat{\phi}_{\mathcal{B}^*}(\mathbf{a}_n))$ . For the expanded Gabidulin code  $\bar{\mathcal{G}}$ , we define  $\hat{\phi}_{\mathcal{B}^*}(\bar{\mathcal{G}}) = \{\hat{\phi}_{\mathcal{B}^*}(\mathbf{c}) : \mathbf{c} \in \bar{\mathcal{G}}\}$ . It is easy to verify that  $\hat{\phi}_{\mathcal{B}^*}(\bar{\mathcal{G}}) \subseteq \mathbb{F}_{q^m}^n$  forms an  $\mathbb{F}_q$ -linear space of dimension  $mk$ .

A potential adversary may randomly choose a basis  $\mathcal{B}^*$  and generate an  $\mathbb{F}_q$ -linear code by computing  $\hat{\phi}_{\mathcal{B}^*}(\bar{\mathcal{G}})$ . Theoretically  $\hat{\phi}_{\mathcal{B}^*}(\bar{\mathcal{G}})$  could be the parent Gabidulin code of  $\bar{\mathcal{G}}$  or some of its equivalent codes. If such an extreme case happens, then the encryption system will be completely broken. According to our analysis, however, this is merely a small probability event.

Let  $\Phi = \hat{\phi}_{\mathcal{B}^*} \circ \phi_{\mathcal{B}}$  be a composite mapping. It is easy to verify that  $\Phi$  forms an  $\mathbb{F}_q$ -linear automorphism of  $\mathbb{F}_{q^m}^n$ . If  $\Phi$  is a stretching transformation, namely there exists  $\gamma \in \mathbb{F}_{q^m}^*$  such that

$\Phi(\alpha) = \gamma\alpha$  for any  $\alpha \in \mathbb{F}_{q^m}$ , then there must be  $\Phi(\mathcal{G}) = \mathcal{G}$ . For any  $\gamma \in \mathbb{F}_{q^m}^*$ , there exists a stretching transformation  $\Phi$  induced by  $\gamma$ . Therefore, there are at least  $q^m - 1$   $\mathbb{F}_q$ -linear automorphisms  $\Phi$  such that  $\Phi(\mathcal{G})$  is also a Gabidulin code. In general cases, however,  $\Phi(\mathcal{G})$  is no longer a Gabidulin code, not even an  $\mathbb{F}_{q^m}$ -linear code. We will illustrate this point by Example 1.

On the other hand, the total number of  $\mathbb{F}_q$ -linear automorphisms of  $\mathbb{F}_{q^m}$  can be evaluated as  $\prod_{i=0}^{m-1} (q^m - q^i)$ . Among all these automorphisms, transformations that behave like a stretching transformation take a small proportion according to our experiments on Magma. For instance, we construct a  $[15, 7]$  Gabidulin code over  $\mathbb{F}_{3^{30}}$  and then perform one million random  $\mathbb{F}_3$ -linear automorphisms to this code, but none of these converted codes are  $\mathbb{F}_{3^{30}}$ -linear.

According to the analysis above, we conclude that this potential attack against expanded Gabidulin codes is infeasible. Similar to the masking techniques employed in [17], in Proposal I we first divide the underlying expanded Gabidulin code into  $n$  blocks, and then perform a column-mixing operation to each of these blocks by multiplying the punctured generator matrix with an invertible block diagonal matrix. In Proposal II, we disguise the underlying code by mixing  $\lambda$  adjacent blocks instead of mixing columns inside each block. We believe that all these techniques further strengthen our two proposals against structural attacks.

*Example 1.* Let  $f(x) = x^3 + x + 1$  be an irreducible polynomial over  $\mathbb{F}_2$ . Denote by  $\mathbb{F}_8$  the extension field of  $\mathbb{F}_2$  derived from  $f(x)$ . Let  $\alpha \in \mathbb{F}_8^*$  such that  $f(\alpha) = 0$ , then  $1, \alpha, \alpha^2$  form a basis of  $\mathbb{F}_8$  over  $\mathbb{F}_2$ . We define an  $\mathbb{F}_2$ -linear automorphism  $\Phi$  over  $\mathbb{F}_8$  as follows

$$\Phi(1) = \alpha^4, \Phi(\alpha) = \alpha^2, \Phi(\alpha^2) = \alpha^3.$$

Let  $\mathbf{g} = (\alpha, \alpha^2, \alpha^3) \in \mathbb{F}_8^3$  and define a  $[3, 2]$  Gabidulin code  $\mathcal{G} \subseteq \mathbb{F}_8^3$  generated by  $\mathbf{g}$ . It is easy to verify that  $\mathcal{G}$  has a generator matrix of the form

$$G = \begin{pmatrix} 1 & 0 & \alpha^3 \\ 0 & 1 & \alpha^4 \end{pmatrix}.$$

If there exists a word of  $\langle \Phi(\mathbf{g}), \Phi(\mathbf{g}^2) \rangle_{\mathbb{F}_8}$  that is not contained in  $\Phi(\mathcal{G})$ , we conclude that  $\Phi(\mathcal{G})$  is not  $\mathbb{F}_8$ -linear. Let  $\mathbf{a} = \alpha\Phi(\mathbf{g})$ , then  $\mathbf{a} \notin \Phi(\mathcal{G})$  holds if and only if  $\Phi^{-1}(\mathbf{a}) \notin \mathcal{G}$ . Since  $\Phi^{-1}(\mathbf{a}) = (\alpha^2, 1, \alpha)$ , apparently we have  $\Phi^{-1}(\mathbf{a}) \notin \mathcal{G}$ . Hence the previous claim is true.

### 5.3 Generic attacks

Note that in Hamming metric codes based cryptography, the best known generic attack is the information set decoding (ISD) attack [39]. Since expanded Gabidulin codes are far from optimal in the Hamming metric according to Proposition 1, the Hamming weight of the intended error vector is much greater than the error-correcting capability of the public code with high probability. Hence the ISD attack is not applicable to our cases.

In what follows, we introduce two hard problems closely related to the security of our proposals, namely the rank syndrome decoding (RSD) problem and MinRank problem. Generally speaking, approaches to solve these two problems are mainly divided into two categories, namely the combinatorial approach and algebraic approach. In this paper, we consider the combinatorial attack on the RSD problem and the algebraic attack on the MinRank problem to evaluate the security level of our proposals.

### 5.3.1 Combinatorial attacks

*Definition 12* (RSD problem). Let  $H \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$  be a matrix of full rank,  $\mathbf{s} \in \mathbb{F}_{q^m}^{n-k}$  and  $t$  be a positive integer. An RSD instance  $\mathcal{R}(q, m, n, k, t)$  is to solve  $\mathbf{s} = \mathbf{e}H^T$  for  $\mathbf{e} \in \mathbb{F}_{q^m}^n$  such that  $w_R(\mathbf{e}) \leq t$ .

The RSD problem plays a crucial role in rank-based cryptography. Although this problem is not known to be NP-complete, it is believed to be hard by the community. In the paper [40], the authors proved that if there is a probabilistic polynomial-time algorithm for solving the RSD problem, then a probabilistic polynomial-time algorithm can be obtained to solve the syndrome decoding problem in the Hamming metric, which has been proved to be NP-complete in [41]. Generally speaking, attacks on the RSD problem can be divided into two categories, namely the combinatorial attack and algebraic attack. Up to now, the best known combinatorial attacks can be found in [20, 42–44].

Now we first recall the principle of the combinatorial attack proposed in [43]. Although there are some improvements [44] for this attack, they are not applicable to our proposals. For an RSD instance  $\mathcal{R}(q, m, n, k, t)$ , we consider the following two cases to solve the problem.

**Case 1:**  $n \geq m$ . Let  $\mathcal{V}$  be an  $\mathbb{F}_q$ -subspace of  $\mathbb{F}_{q^m}$  of dimension  $t' \geq t$ . If  $\text{Supp}_R(\mathbf{e}) \subseteq \mathcal{V}$ , then each component of  $\mathbf{e}$  can be expressed as an  $\mathbb{F}_q$ -linear combination of a basis of  $\mathcal{V}$ . Computing  $\mathbf{s} = \mathbf{e}H^T$  will lead to a linear system of  $n-k$  equations with  $nt'$  variables in  $\mathbb{F}_q$ . By expressing each monomial of these parity-check equations as a vector in  $\mathbb{F}_q^m$  under some basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ , we eventually obtain a system of  $m(n-k)$  equations over  $\mathbb{F}_q$  with  $nt'$  variables. Apparently this system has at least one solution when the condition  $\text{Supp}_R(\mathbf{e}) \subseteq \mathcal{V}$  is satisfied. To have only one solution with overwhelming probability, we should make sure that  $m(n-k) \geq nt'$  and then  $t' \leq m - \lceil \frac{km}{n} \rceil$ . By solving this system, we can finally recover the error vector  $\mathbf{e}$ . The total complexity of this algorithm is  $\mathcal{O}(m^3(n-k)^3/p)$ , where  $\mathcal{O}(m^3(n-k)^3)$  represents the average complexity required for solving the linear system and  $p$  is the probability that a random  $\mathbb{F}_q$ -subspace of  $\mathbb{F}_{q^m}$  of dimension  $t'$  happens to have  $\text{Supp}_R(\mathbf{e})$  as a subspace. By the Gaussian binomial coefficient, we have

$$p = \frac{\begin{bmatrix} m-t \\ t'-t \end{bmatrix}_q}{\begin{bmatrix} m \\ t' \end{bmatrix}_q} = \frac{\prod_{i=0}^{t'-t-1} \frac{q^{m-t} - q^i}{q^{t'-t} - q^i}}{\prod_{i=0}^{t'-1} \frac{q^m - q^i}{q^{t'} - q^i}} \approx \frac{1}{q^{t(m-t')}}.$$

By taking  $t' = m - \lceil \frac{km}{n} \rceil$ , we get a complexity of  $\mathcal{O}(m^3(n-k)^3 q^{t \lceil \frac{km}{n} \rceil})$  in  $\mathbb{F}_q$ .

**Case 2:**  $n < m$ . Note that  $\mathbb{F}_{q^m}$  can be viewed as an  $\mathbb{F}_q$ -linear space of dimension  $m$ . For a basis of  $\mathbb{F}_{q^m}$  over  $\mathbb{F}_q$ , by expressing each component of  $\mathbf{e}$  as a column vector with respect to this basis, we can obtain a  $m \times n$  matrix  $E$  over  $\mathbb{F}_q$ . Let  $\mathcal{E}$  be a linear space spanned by rows of  $E$  over  $\mathbb{F}_q$ . If we can find a linear space  $\mathcal{V} \subseteq \mathbb{F}_q^n$  of dimension  $t'$  such that  $\mathcal{E} \subseteq \mathcal{V}$ , then each row of  $E$  can be expressed as an  $\mathbb{F}_q$ -linear combination of a basis of  $\mathcal{V}$ . With a similar analysis as Case 1, we eventually get a linear system of  $m(n-k)$  equations over  $\mathbb{F}_q$  with  $mt'$  variables. Let  $mt' \leq m(n-k)$ , then  $t' \leq n-k$ . On the other hand, the probability that a random subspace of  $\mathbb{F}_q^n$  of dimension  $t'$  happens to contain  $\mathcal{E}$  can be evaluated as

$$p = \frac{\begin{bmatrix} n-t \\ t'-t \end{bmatrix}_q}{\begin{bmatrix} n \\ t' \end{bmatrix}_q} = \frac{\prod_{i=0}^{t'-t-1} \frac{q^{n-t} - q^i}{q^{t'-t} - q^i}}{\prod_{i=0}^{t'-1} \frac{q^n - q^i}{q^{t'} - q^i}} \approx \frac{1}{q^{t(n-t')}}.$$

By taking  $t' = n - k$ , we get a complexity of  $\mathcal{O}(m^3(n-k)^3 q^{tk})$ .

Having introduced the general idea of the combinatorial attack on the RSD problem, we now apply it to the case of our two proposals.

**Proposai I.** In the case of  $n < \lambda$ , let  $\mathcal{E}$  be the linear space spanned by columns of the error matrix  $E$ . Let  $\mathcal{V}$  be a random subspace of  $\mathbb{F}_q^n$  of dimension  $t' \geq t$ . If  $\mathcal{E} \subseteq \mathcal{V}$ , then each column of  $E$  can be expressed as a linear combination of a basis of  $\mathcal{V}$  with  $t'$  undetermined coefficients in  $\mathbb{F}_q$ . With such a matrix  $E$  of this form, we can construct the error vector  $e$ . Let  $H_{pub}$  be a parity-check matrix of the public code, and  $\mathbf{y} = \mathbf{x}G_{pub} + e$  be the received ciphertext. Then computing  $\mathbf{y}H_{pub}^T = eH_{pub}^T$  will result in a linear system of  $m(n - k)$  equations with  $\lambda t'$  variables in  $\mathbb{F}_q$ . To have only one solution, there should be  $m(n - k) \geq \lambda t'$  and hence we have  $t' \leq \lfloor \frac{m(n-k)}{\lambda} \rfloor$ . With a similar analysis as Case 2, the probability that a randomly chosen  $\mathcal{V}$  happens to contain  $\mathcal{E}$  can be evaluated as  $p \approx q^{-t(n-t')}$ . By taking  $t' = \lfloor \frac{m(n-k)}{\lambda} \rfloor$ , we get a complexity of  $\mathcal{O}\left(m^3(n - k)^3 q^{t(n - \lfloor \frac{m(n-k)}{\lambda} \rfloor)}\right)$ .

In the case of  $n \geq \lambda$ , let  $\mathcal{E}$  be the linear space spanned by rows of  $E$ . Let  $\mathcal{V}$  be a random subspace of  $\mathbb{F}_q^\lambda$  of dimension  $t' \geq t$  such that  $\mathcal{E} \subseteq \mathcal{V}$ . With a similar analysis as above, we can obtain a linear system of  $m(n - k)$  equations with  $nt'$  variables in  $\mathbb{F}_q$ . Let  $m(n - k) \geq nt'$ , then  $t' \leq m - \lceil \frac{km}{n} \rceil$ . The probability that a randomly chosen  $\mathcal{V}$  happens to contain  $\mathcal{E}$  can be evaluated as  $p \approx q^{-t(\lambda-t')}$ . By taking  $t' = m - \lceil \frac{km}{n} \rceil$ , we get a complexity of  $\mathcal{O}\left(m^3(n - k)^3 q^{t(\lambda - m + \lceil \frac{km}{n} \rceil)}\right)$ .

**Proposai II.** On the one hand. According to the description of Proposal II, the linear space  $\mathcal{E} = \langle e_1, \dots, e_n \rangle_{\mathbb{F}_q}$  has dimension at most  $\lambda t$ . With a similar analysis as Case 1, we get a complexity of  $\mathcal{O}\left(m^3(n - k)^3 q^{\lambda t \lceil \frac{km}{n} \rceil}\right)$  for  $n \geq m$ . With a similar analysis as Case 2, we get a complexity of  $\mathcal{O}\left(m^3(n - k)^3 q^{\lambda tk}\right)$  for  $n < m$ .

On the other hand. Note that  $e_i$  ( $1 \leq i \leq n$ ) is obtained from a matrix  $E \in \mathcal{M}_{u_c, m\lambda}(\mathbb{F}_q)$  with  $\text{Rank}(E) = t$ . For a nontrivial  $\lambda \geq 2$ , apparently we have  $u_c = \lceil \frac{n}{\lambda} \rceil < n < m\lambda$ . Denote by  $\mathcal{E}$  the linear space spanned by columns of  $E$  over  $\mathbb{F}_q$ . Let  $\mathcal{V}$  be a random subspace of  $\mathbb{F}_q^{u_c}$  of dimension  $t' \geq t$ . If  $\mathcal{E} \subseteq \mathcal{V}$ , then each column of  $E$  can be expressed as a linear combination of a basis of  $\mathcal{V}$ . This implies that we can represent the error vector  $e$  with  $m\lambda t'$  undetermined coefficients in  $\mathbb{F}_q$ . Let  $m(n - k) \geq m\lambda t'$ , then  $t' \leq \lfloor \frac{n-k}{\lambda} \rfloor$ . The probability that a random  $t'$ -dimensional subspace of  $\mathbb{F}_q^{u_c}$  happens to contain  $\mathcal{E}$  can be evaluated as

$$p = \frac{\binom{u_c - t}{t' - t}}{\binom{u_c}{t'}} \approx \frac{1}{q^{t(u_c - t')}}.$$

By taking  $t' = \lfloor \frac{n-k}{\lambda} \rfloor$ , we get a complexity of  $\mathcal{O}\left(m^3(n - k)^3 q^{t(u_c - \lfloor \frac{n-k}{\lambda} \rfloor)}\right)$ .

### 5.3.2 Algebraic attacks

*Definition 13 (MinRank problem).* For positive integers  $n_1, n_2, m$  and  $t$ , let  $M_1, M_2, \dots, M_m \in \mathcal{M}_{n_1, n_2}(\mathbb{F}_q)$ . A MinRank instance of parameters  $(q, n_1, n_2, m, t)$  is to search for  $m$  coefficients  $x_1, x_2, \dots, x_m \in \mathbb{F}_q$  such that  $\text{Rank}\left(\sum_{i=1}^m x_i M_i\right) \leq t$ .

The MinRank problem was originally proposed by Buss et al. [45] as a natural question in linear algebra, and proved to be NP-complete at the same time. This problem plays a central role in both multivariate cryptography [47] and rank-based cryptography [43]. The rank decoding (RD) problem, a dual version of the RSD problem  $\mathcal{R}(q, m, n, k, t)$ , can be reduced to a structured MinRank instance of parameters  $(q, m, n, mk + 1, t)$  [48]. There are mainly three approaches to solve

the MinRank problem, namely the Kipnis-Shamir modeling [49], minors modeling [50] and linear algebra search [51].

The authors in [50, 52] investigated the “square” case where  $n_1 = n_2 = n$ , and gave an upper bound for the complexity of solving the MinRank problem in this situation. For a MinRank instance of parameters  $(q, n, n, m, t)$ , the complexity is

$$\mathcal{O} \left( \binom{m + t(n - t) + \min\{m, t(n - t)\} + 1}{\min\{m, t(n - t)\} + 1} \right)^\omega,$$

where  $\omega = 2.8$  represents the linear algebra constant. In the paper [48], Faugère et al. managed to solve a MinRank instance of parameters  $(q, n, n, m, t)$  with a complexity of  $\mathcal{O} \left( \log(q) n^{3(n-t)^2} \right)$  when  $m = (n - t)^2 + 1$  holds.

For the “nonsquare” case, Bardet et al. [46] proposed to solve a MinRank instance of parameters  $(q, n_1, n_2, m, t)$  with a complexity of

$$\mathcal{O} \left( m(t + 1) \binom{n_2}{t} \binom{m + b - 1}{b} \right),$$

where  $b$  is the smallest positive integer such that  $b < \min\{q, t + 2\}$  satisfying the following condition

$$\binom{n_2}{t} \binom{m + b - 1}{b} - 1 \leq \sum_{i=1}^b (-1)^{i+1} \binom{n_2}{t+i} \binom{n_1 + i - 1}{i} \binom{m + b - i - 1}{b - i}.$$

In what follows, we will show that decrypting any valid ciphertext in our proposals can be converted into solving a MinRank instance. For the convenience of our statement, we first introduce an  $\mathbb{F}_q$ -linear isomorphism  $\sigma_n$  from  $\mathbb{F}_q^{ns}$  to  $\mathcal{M}_{n,s}(\mathbb{F}_q)$  for some positive integer  $s$ . For a vector  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{F}_q^{ns}$  with  $\mathbf{x}_i \in \mathbb{F}_q^s$  ( $1 \leq i \leq n$ ), we define  $\sigma_n(\mathbf{x})$  as follows

$$\sigma_n(\mathbf{x}) = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T \in \mathcal{M}_{n,s}(\mathbb{F}_q).$$

Generalizing this transformation to a set  $\mathcal{X} \subseteq \mathbb{F}_q^{ns}$ , we define  $\sigma_n(\mathcal{X}) = \{\sigma_n(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ . For any  $\mathbf{x} \in \mathbb{F}_q^{ns}$ , by  $w_R(\mathbf{x})$  we mean the rank of  $\sigma_n(\mathbf{x})$  hereafter when no ambiguity arises.

**Proposal I.** With the concept above, we now introduce the following proposition with respect to our first proposal.

**Proposition 6.** Denote by  $\mathbf{m}_i$  the  $i$ -th row vector of the public matrix  $G_{pub}$  ( $1 \leq i \leq K$ ), and by  $\mathcal{G}_{pub}$  the public code generated by  $G_{pub}$ . Let  $\mathbf{y} = \mathbf{c} + \mathbf{e}$  be the received word, where  $\mathbf{c} \in \mathcal{G}_{pub}$  and  $\mathbf{e} \in \mathbb{F}_q^N$  with  $w_R(\mathbf{e}) = t$ . If there exist  $a_0, a_1, \dots, a_K \in \mathbb{F}_q$  such that  $0 < w_R(a_0 \mathbf{y} + \sum_{i=1}^K a_i \mathbf{m}_i) \leq t$ , then there must be  $a_0 \mathbf{y} + \sum_{i=1}^K a_i \mathbf{m}_i = a_0 \mathbf{e}$  and hence  $\mathbf{e} = \frac{1}{a_0} (a_0 \mathbf{y} + \sum_{i=1}^K a_i \mathbf{m}_i)$ .

*Proof.* It is easy to see that  $\bar{\mathcal{G}}_S$  is obtained by shortening  $\bar{\mathcal{G}}$  at  $S$ , then there must be

$$d(\bar{\mathcal{G}}_S) \geq d(\bar{\mathcal{G}}) \geq n - k + 1 \geq 2t + 1,$$

where the minimum distance  $d(\cdot)$  is defined with respect to the mapping  $\sigma_n(\cdot)$ . On the other hand, apparently we have  $\mathcal{G}_{pub} = \langle \bar{\mathcal{G}}_S T^{-1} \rangle_{\mathbb{F}_q} = \bar{\mathcal{G}}_S T^{-1}$ . For any  $\mathbf{u} \in \mathcal{G}_{pub}$ , there exists  $\mathbf{v} \in \bar{\mathcal{G}}_S$  such that  $\mathbf{u} = \mathbf{v} T^{-1}$ . Following this we have  $\sigma_n(\mathbf{u}) = \sigma_n(\mathbf{v}) A^{-1}$  and then

$$w_R(\mathbf{u}) = \text{Rank}(\sigma_n(\mathbf{u})) = \text{Rank}(\sigma_n(\mathbf{v})) = w_R(\mathbf{v}) \geq 2t + 1,$$



which implies that  $d(\mathcal{G}_{pub}) \geq 2t + 1$ . Hence if  $a_0\mathbf{y} + \sum_{i=1}^K a_i\mathbf{m}_i = (a_0\mathbf{c} + \sum_{i=1}^K a_i\mathbf{m}_i) + a_0\mathbf{e}$  has rank weight at most  $t$ , then there must be  $a_0\mathbf{c} + \sum_{i=1}^K a_i\mathbf{m}_i = \mathbf{0}$  and  $a_0 \neq 0$ . Otherwise, there will be  $w_R(a_0\mathbf{y} + \sum_{i=1}^K a_i\mathbf{m}_i) \geq t + 1$ . This completes the proof.  $\square$

Based on the analysis above, decrypting a valid ciphertext in Proposal I can be reduced to solving a MinRank instance of parameter  $(q, n, \lambda, K + 1, t)$ , formally we have the following proposition.

**Proposition 7.** *Suppose  $\mathbf{y} = \mathbf{c} + \mathbf{e}$  is a valid ciphertext in Proposal I, where  $\mathbf{c} \in \mathcal{G}_{pub}$  and  $\mathbf{e} \in \mathbb{F}_q^N$  with  $w_R(\mathbf{e}) = t$ . Let  $M_0 = \sigma_n(\mathbf{y})$  and  $M_i = \sigma_n(\mathbf{m}_i)$  for  $1 \leq i \leq K$ . Then recovering  $\mathbf{e}$  can be reduced to a MinRank instance of searching for  $a_0, a_1, \dots, a_K \in \mathbb{F}_q$  such that  $\text{Rank}(\sum_{i=0}^K a_i M_i) \leq t$ .*

**Proposal II.** Similar to Proposition 7, Proposal II can be reduced to a MinRank instance of parameter  $(q, n, m, K + 1, \lambda t)$ , formally we introduce the following proposition.

**Proposition 8.** *Suppose  $\mathbf{y} = \mathbf{c} + \mathbf{e}$  is a valid ciphertext in Proposal II, where  $\mathbf{c} \in \mathcal{G}_{pub}$  and  $\mathbf{e} \in \mathbb{F}_q^N$  is the intended error vector. Let  $M_0 = \sigma_n(\mathbf{y})$  and  $M_i = \sigma_n(\mathbf{m}_i)$  for  $1 \leq i \leq K$ . Then recovering  $\mathbf{e}$  can be reduced to a MinRank instance of searching for  $a_0, a_1, \dots, a_K \in \mathbb{F}_q$  such that  $\text{Rank}(\sum_{i=0}^K a_i M_i) \leq \lambda t$ .*

Apparently there exists at least one solution for this reduced MinRank instance. For a nontrivial  $\lambda \geq 2$ , the right column scrambler  $T^{-1}$  does not preserve the rank metric with respect to the mapping  $\sigma_n(\cdot)$  in general, thus we cannot decide as in Proposition 6 whether or not a solution reveals the intended error vector  $\mathbf{e}$  or some of its multiples. But we can still give an answer to this question according to the following proposition.

**Proposition 9.** *Suppose  $a_0, a_1, \dots, a_K \in \mathbb{F}_q$  form a solution of the MinRank instance described in Proposition 8, then let  $\mathbf{e}' = \sigma_n^{-1}(\sum_{i=0}^K a_i M_i) = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$  with  $\mathbf{e}'_j \in \mathbb{F}_q^m$  for  $1 \leq j \leq n$ . Define  $E' \in \mathcal{M}_{u_c, m\lambda}(\mathbb{F}_q)$  to be the associated matrix of  $\mathbf{e}'$  of the form (3). If  $\text{Rank}(E') \leq t$ , then there must be  $\mathbf{e}' = a_0\mathbf{e}$ .*

*Proof.* Apparently  $\mathbf{e}' = \sigma_n^{-1}(\sum_{i=0}^K a_i M_i) = a_0\mathbf{y} + \sum_{i=1}^K a_i\mathbf{m}_i = a_0\mathbf{e} + (a_0\mathbf{c} + \sum_{i=1}^K a_i\mathbf{m}_i)$ . On the one hand,  $\mathbf{e}'$  can be viewed as a valid ciphertext obtained by encrypting  $\mathbf{0} \in \mathbb{F}_q^K$  because of  $\text{Rank}(E') \leq t$ . Hence decrypting  $\mathbf{e}'$  will lead to  $\mathbf{e}'$  itself. On the other hand,  $\mathbf{e}'$  can also be viewed as a valid ciphertext that has  $a_0\mathbf{e}$  as the intended error vector because of  $a_0\mathbf{c} + \sum_{i=1}^K a_i\mathbf{m}_i \in \mathcal{G}_{pub}$ . Then decrypting  $\mathbf{e}'$  will lead to  $a_0\mathbf{e}$ . By the uniqueness of decryption, there must be  $\mathbf{e}' = a_0\mathbf{e}$ . This completes the proof.  $\square$

## 6 Parameter choice and public-key size

In this section, we compute the public-key sizes and information rates of the proposed cryptosystems for the security of 128 bits, 192 bits and 256 bits against the generic attacks described in Section 5.3. After that we will make a comparison on public-key sizes with some other code-based cryptosystems for different security levels.

In Proposal I, the public key is a systematic generator matrix of the public code with length  $n\lambda$  and dimension  $n\lambda - mr$  where  $r = n - k$ , resulting in a public-key size of  $mr(n\lambda - mr) \cdot \log_2(q)$

Instance	Parameters	Key Size	Rate	Security
Proposal I	m=25, n=25, k=17, $\lambda=24$	15850	0.67	128
	m=32, n=32, k=22, $\lambda=31$	42604	0.68	192
	m=37, n=37, k=25, $\lambda=36$	78114	0.67	256
Proposal II	m=43, n=43, k=19, $\lambda=2$	167044	0.44	128
	m=57, n=57, k=33, $\lambda=2$	509805	0.58	192
	m=67, n=67, k=39, $\lambda=2$	971184	0.58	256

Table 1: Public-key sizes (in bytes) and information rates of the proposed cryptosystems in the case of  $q = 3$ .

Instance	Parameters	Key Size	Rate	Security
Proposal I	m=20, n=20, k=14, $\lambda=19$	10949	0.68	128
	m=24, n=24, k=16, $\lambda=23$	24256	0.65	192
	m=28, n=28, k=18, $\lambda=27$	46771	0.63	256
Proposal II	m=35, n=35, k=23, $\lambda=2$	118646	0.66	128
	m=45, n=45, k=29, $\lambda=2$	329724	0.64	192
	m=51, n=51, k=31, $\lambda=2$	565900	0.61	256

Table 2: Public-key sizes (in bytes) and information rates of the proposed cryptosystems in the case of  $q = 7$ .

Instance	Parameters	Key Size	Rate	Security
Proposal I	m=17, n=17, k=11, $\lambda=16$	8021	0.63	128
	m=22, n=22, k=16, $\lambda=21$	20149	0.71	192
	m=25, n=25, k=17, $\lambda=24$	37005	0.67	256
Proposal II	m=29, n=29, k=17, $\lambda=2$	79358	0.59	128
	m=37, n=37, k=21, $\lambda=2$	212768	0.57	192
	m=43, n=43, k=23, $\lambda=2$	393422	0.53	256

Table 3: Public-key sizes (in bytes) and information rates of the proposed cryptosystems in the case of  $q = 13$ .

bits. In Proposal II, the public key is a systematic generator matrix of the public code with length  $mn$  and dimension  $mk$ , resulting in a public-key size of  $rk m^2 \cdot \log_2(q)$  bits. As for the information rates, this value is evaluated as  $(n\lambda - mr)/n\lambda$  for Proposal I, and  $k/n$  for Proposal II respectively.

In Table 1 we consider the case of  $q = 3$ , in Table 2 we consider the case of  $q = 7$ , and in Table 3 we consider  $q = 13$ . For Proposal I, we suggest the parameter set ( $q = 13, m = 17, n = 17, k = 11, \lambda = 16$ ) for the security of 128 bits, the parameter set ( $q = 13, m = 22, n = 22, k = 16, \lambda = 21$ ) for the security of 192 bits, and the parameter set ( $q = 13, m = 25, n = 25, k = 17, \lambda = 24$ ) for the security of 256 bits.

For Proposal II, we suggest the parameter set ( $q = 13, m = 29, n = 29, k = 17, \lambda = 2$ ) for the security of 128 bits, the parameter set ( $q = 13, m = 37, n = 37, k = 21, \lambda = 2$ ) for the security of 192 bits, and the parameter set ( $q = 13, m = 43, n = 43, k = 23, \lambda = 2$ ) for the security of 256 bits.

In Table 4, we make a comparison on public-key sizes with some other code-based cryptosystems. The first four instances, namely BIKE [53], HQC [54], Classic McEliece [55] and NTS-

Instance	128 bits	192 bits	256 bits
BIKE	1540	3082	5121
HQC	2249	4522	7245
Classic McEliece	261120	524160	1044992
NTS-KEM	319488	929760	1419704
KRW			578025
Proposal I	8021	20149	37005
Proposal II	79358	212768	393422

Table 4: Comparison on public-key sizes (in bytes) with some other cryptosystems.

KEM [56], have been selected to move on to the third round of the NIST PQC Standardization Process. Note that Classic McEliece published by the NIST PQC team is a merged version of NTS-KEM and the original Classic McEliece for their specifications being very similar. The KRW cryptosystem proposed in [17] is based on the so-called expanded GRS codes, which inspires us to exploit expanded Gabidulin codes in the design of encryption schemes. Compared to the KRW cryptosystem, our proposals admit a more compact representation of public keys for the reason that the general decoding problem in the rank metric is much more difficult than that in the Hamming metric.

## 7 Conclusion

This paper presented two cryptosystems by using the so-called expanded Gabidulin codes in the McEliece setting. In our proposals, the underlying expanded Gabidulin code is divided into  $n$  blocks, with each block having size  $m$ . By the definition of expanded Gabidulin codes, each block corresponds to one coordinate of the parent Gabidulin code. To destroy this correspondence, in Proposal I we first shortened the expanded Gabidulin code and then performed a column-mixing transformation to each block. In Proposal II, we adopted a rather different column-mixing transformation to the underlying code by mixing  $\lambda$  neighbouring blocks. According to our analysis in Section 5.2, these two variants can both resist the existing structural attacks. Furthermore, our proposals have great advantage in public-key representation compared to some other code-based cryptosystems. For instance, we have reduced the public-key size by around 96% compared to Classic McEliece entering the third round of the NIST PQC project.

**Acknowledgements** This research is supported by the National Key Research and Development Program of China (Grant No. 2018YFA0704703), the National Natural Science Foundation of China (Grant No. 61971243), the Natural Science Foundation of Tianjin (20JCZDJC00610), and the Fundamental Research Funds for the Central Universities of China (Nankai University).

## References

- [1] McEliece, R.J.: A public-key cryptosystem based on algebraic coding theory. Jet Propuls. Lab. DSN Progr. Rep. 42-44, 114–116 (1978).

- [2] Lee, P.J., Brickell, E.F.: An observation on the security of McEliece's public-key cryptosystem. In: Guenther, C.G. (Ed.): Proceedings of Advances in Cryptology-EUROCRYPT'88, LNCS, vol. 330, pp. 275–280. Springer (1988).
- [3] Canteaut, A., Sendrier, N.: Cryptanalysis of the original McEliece cryptosystem. In: Ohta, K., Pei, D. (Eds.): Proceedings of ASIACRYPT'98, LNCS, vol. 1514, pp. 187–199. Springer (2000).
- [4] Loidreau, P., Sendrier, N.: Weak keys in the McEliece public-key cryptosystem. *IEEE Trans. Inf. Theory* 47(3), 1207–1211 (2001).
- [5] Faugère, J.-C., Gauthier-Umana, V., Otmani, A., Perret, L., Tillich, J.-P.: A distinguisher for high-rate McEliece cryptosystems. *IEEE Trans. Inf. Theory* 59(10), 6830–6844 (2013).
- [6] Couvreur, A., Otmani, A., Tillich, J.-P.: Polynomial time attack on wild McEliece over quadratic extensions. *IEEE Trans. Inf. Theory* 63(1), 404–427 (2016).
- [7] Loidreau, P.: Strengthening McEliece cryptosystem. In: Okamoto, T. (Ed.): Proceedings of ASIACRYPT 2000, LNCS, vol. 1976, pp. 585–598. Springer (2000).
- [8] Kobara, K., Imai, H.: On the one-wayness against chosen-plaintext attacks of the Loidreau's modified McEliece PKC. *IEEE Trans. Inf. Theory* 49(12), 3160–3168 (2003).
- [9] Misoczki, R., Barreto, P.S.: Compact McEliece keys from Goppa codes. In: M.J. Jacobson Jr., Rijmen V., Safavi-Naini R. (Eds.): Proceedings of SAC 2009, LNCS, vol. 5867, pp. 376–392. Springer (2009).
- [10] Faugère, J.-C., Otmani, A., Perret, L., Tillich J.-P.: Algebraic cryptanalysis of McEliece variants with compact keys. In: Gilbert, H. (Ed.): Proceedings of EUROCRYPT 2010, LNCS, vol. 6110, pp. 279–298. Springer (2010).
- [11] Niederreiter, H.: Knapsack-type cryptosystems and algebraic coding theory. *Prob. Contr. Inform. Theory* 15(2), 159–166 (1986).
- [12] Sidelnikov, V.M., Shestakov, S.O.: On insecurity of cryptosystems based on generalized Reed-Solomon codes. *Discret. Math. Appl.* 2(4), 439–444 (1992).
- [13] Li, Y.X., Deng, R.H., Wang, X.M.: On the equivalence of McEliece's and Niederreiter's public-key cryptosystems. *IEEE Trans. Inf. Theory* 40(1), 271–273 (1994).
- [14] Baldi, M., Bianchi, M., Chiaraluce, F., Rosenthal J., Schipani, D.: Enhanced public key security for the McEliece cryptosystem. *J. Cryptology* 29(1), 1–27 (2016).
- [15] Wieschebrink, C.: Cryptanalysis of the Niederreiter public key scheme based on GRS subcodes. In: Sendrier, N. (Ed.): Proceedings of PQCrypto 2010, LNCS, vol. 6061, pp. 61–72. Springer (2010).
- [16] Couvreur, A., Gaborit, P., Otmani, A., Tillich, J.-P.: Distinguisher-based attacks on public-key cryptosystems using Reed-Solomon codes. *Des. Codes Cryptogr.* 73(2), 641–666 (2014).

- [17] Khathuria, K., Rosenthal, J., Weger, V.: Encryption scheme based on expanded Reed-Solomon codes. *Adv. Math. Commun.* 15(2), 207–218 (2021).
- [18] Mullen, G.L., Panario, D.: *Handbook of Finite Fields*. CRC Press (2013).
- [19] Gabidulin, E.M.: Theory of codes with maximum rank distance. *Prob. Peredachi Inf.* 21(1), 3–16 (1985).
- [20] Chabaud, F., Stern, J.: The cryptographic security of the syndrome decoding problem for rank distance codes. In: *Proceedings of International Conference on the Theory and Application of Cryptology and Information Security*, pp. 368–381. Springer (1996).
- [21] Ourivski, A.V., Johansson, T.: New technique for decoding codes in the rank metric and its cryptography applications. *Probl. Inf. Transm.* 38(3), 237–246 (2002).
- [22] Gabidulin, E.M., Paramonov, A.V., Tretjakov, O.V.: Ideals over a non-commutative ring and their application in cryptology. In: Davies, D.W. (Ed.): *Proceedings of Advances in Cryptology-EUROCRYPT'91*, LNCS, vol. 547, pp. 482–489. Springer (1991).
- [23] Gabidulin, E.M.: Attacks and counter-attacks on the GPT public key cryptosystem. *Des. Codes Cryptogr.* 48(2), 171–177 (2008).
- [24] Gabidulin, E.M., Rashwan, H., Honary, B.: On improving security of GPT cryptosystems. In: *Proceedings of 2009 IEEE International Symposium on Information Theory*, pp. 1110–1114. IEEE (2009).
- [25] Loidreau, P.: Designing a rank metric based McEliece cryptosystem. In: Sendrier, N. (Ed.): *Proceedings of PQCrypto 2010*, LNCS, vol. 6061, pp. 142–152. Springer (2010).
- [26] Rashwan, H., Gabidulin, E.M., Honary, B.: A smart approach for GPT cryptosystem based on rank codes. In: *Proceedings of 2010 IEEE International Symposium on Information Theory*, pp. 2463–2467. IEEE (2010).
- [27] Gibson, K.: The security of the Gabidulin public key cryptosystem. In: *Proceedings of Advances in Cryptology-EUROCRYPT'96*, LNCS, vol. 1070, pp. 212–223. Springer (1996).
- [28] Overbeck, R.: Structural attacks for public key cryptosystems based on Gabidulin codes. *J. Cryptology* 21(2), 280–301 (2008).
- [29] Horlemann-Trautmann, A.-L., Marshall, K., Rosenthal, J.: Extension of Overbeck's attack for Gabidulin-based cryptosystems. *Des. Codes Cryptogr.* 86(2), 319–340 (2018).
- [30] Otmani, A., Kalachi, H.T., Ndjeya, S.: Improved cryptanalysis of rank metric schemes based on Gabidulin codes. *Des. Codes Cryptogr.* 86(9), 1983–1996 (2018).
- [31] Faure, C., Loidreau, P.: A new public-key cryptosystem based on the problem of reconstructing  $p$ -polynomials. In: Ytrehus, Ø. (Ed.): *Proceedings of WCC 2005*, LNCS, vol. 3969, pp. 304–315. Springer (2005).

- [32] Gaborit, P., Otmani, A., Kalachi, H.T.: Polynomial-time key recovery attack on the Faure–Loidreau scheme based on Gabidulin codes. *Des. Codes Cryptogr.* 86(7),1391–1403 (2018).
- [33] Loidreau, P.: A new rank metric codes based encryption scheme. In: Lange, T., Takagi, T. (Eds.): *Proceedings of PQCrypto 2017*, LNCS, vol. 10346, pp. 3–17. Springer (2017).
- [34] Coggia, D., Couvreur, A.: On the security of a Loidreau rank metric code based encryption scheme. *Des. Codes Cryptogr.* 88(9), 1941–1957 (2020).
- [35] Ghatak, A. Extending Coggia-Couvreur attack on Loidreau’s rank-metric cryptosystem. [arXiv:2007.07354 \[cs.IT\]](https://arxiv.org/abs/2007.07354) (2020).
- [36] Horlemann-Trautmann, A.-L., Marshall, K.: New criteria for MRD and Gabidulin codes and some rank-metric code constructions. [arXiv:1507.08641 \[cs.IT\]](https://arxiv.org/abs/1507.08641) (2015).
- [37] Loidreau, P.: A Welch-Berlekamp like algorithm for decoding Gabidulin codes. In: Ytrehus, Ø. (Ed.): *Proceedings of WCC 2005*, LNCS, vol. 3969, pp. 36–45. Springer (2005).
- [38] Richter, G., Plass, S.: Error and erasure decoding of rank-codes with a modified Berlekamp-Massey algorithm. *ITG FACHBERICHT*, pp. 203–210 (2004).
- [39] Peters, C.: Information-set decoding for linear codes over  $\mathbb{F}_q$ . In: Sendrier, N. (Ed.): *Proceedings of PQCrypto 2010*, LNCS, vol. 6061, pp. 81–94. Springer (2010).
- [40] Gaborit, P., Zémor, G.: On the hardness of the decoding and the minimum distance problems for rank codes. *IEEE Trans. Inf. Theory* 62(12), 7245–7252 (2016).
- [41] Berlekamp, E.R., McEliece, R.J., van Tilborg, H.: On the inherent intractability of certain coding problems. *IEEE Trans. Inf. Theory* 24(3), 384–386 (1978).
- [42] Ourivski, A.V., Johansson, T.: New technique for decoding codes in the rank metric and its cryptography applications. *Problems Inform. Transm.* 38(3), 237–246 (2002).
- [43] Gaborit, P., Ruatta, O., Schrek, J.: On the complexity of the rank syndrome decoding problem. *IEEE Trans. Inf. Theory* 62(2), 1006–1019 (2016).
- [44] Aragon, N., Gaborit, P., Hauteville, A., Tillich, J.-P.: A new algorithm for solving the rank syndrome decoding problem. In: *Proceedings of 2018 IEEE International Symposium on Information Theory (ISIT)*, pp. 2421–2425. IEEE (2018).
- [45] Buss, J.F., Frandsen, G.S., Shallit, J.O.: The computational complexity of some problems of linear algebra. *J. Comput. Syst. Sci.* 58(3), 572–596 (1999).
- [46] Bardet, M., Bros, M., Cabarcas, D., et al.: Improvements of algebraic attacks for solving the rank decoding and MinRank problems. In: Moriai, S., Wang, H. (Eds.): *Proceedings of ASIACRYPT 2020*, LNCS, vol. 12491, pp. 507–536. IACR (2020).
- [47] Cabarcas, D., Smith-Tone, D., Verbel, J.A.: Key recovery attack for ZHFE. In: Lange, T., Takagi, T. (Eds.): *Proceedings of PQCrypto 2017*, LNCS, vol. 10346, pp. 289–308. Springer (2017).

- [48] Faugère, J.-C., Levy-dit-Vehel, F., Perret, L.: Cryptanalysis of MinRank. In: Wagner, D. (Ed.): Proceedings of Advances in Cryptology-CRYPTO 2008, LNCS, vol. 5157, pp. 280–296. Springer (2008).
- [49] Kipnis, A., Shamir, A.: Cryptanalysis of the HFE public key cryptosystem by relinearization. In: Wiener, M. (Ed.): Proceedings of CRYPTO 1999, LNCS, vol. 1666, pp. 19–30. Springer (1999).
- [50] Faugère, J.C., El Din, M.S., Spaenlehauer, P.J.: Computing loci of rank defects of linear matrices using Gröbner bases and applications to cryptography. In: Proceedings of Symbolic and Algebraic Computation, International Symposium, ISSAC 2010, pp. 257–264. ACM (2010).
- [51] Goubin, L., Courtois, N.T.: Cryptanalysis of the TTM cryptosystem. In: Okamoto, T. (Ed.): Proceedings of ASIACRYPT 2000, LNCS, vol. 1976, pp. 44–57. Springer (2000).
- [52] Faugère, J.-C., El Din, M.S., Spaenlehauer, P.-J.: Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree  $(1, 1)$ : algorithms and complexity. *J. Symb. Comput.* 46(4), 406–437 (2011).
- [53] Aragon, N., Barreto, P.S., et al.: BIKE: bit flipping key encapsulation. [https://bikesuite.org/files/v4.1/BIKE\\_Spec.2020.10.22.1.pdf](https://bikesuite.org/files/v4.1/BIKE_Spec.2020.10.22.1.pdf). Accessed October 10, 2020.
- [54] Melchor, C.A., Aragon, N., et al.: Hamming quasi-cyclic (HQC). [http://pqc-hqc.org/doc/hqc-specification\\_2020-10-01.pdf](http://pqc-hqc.org/doc/hqc-specification_2020-10-01.pdf). Accessed October 10, 2020.
- [55] Albrecht, M.R., Bernstein, D.J., et al.: Classic McEliece: conservative code-based cryptography. <https://classic.mceliece.org/nist/mceliece-20201010.pdf>. Accessed October 10, 2020.
- [56] Albrecht, M., Cid, C., Paterson, K.G., et al.: NTS-KEM. <https://drive.google.com/file/d/1N3rv4HKCt9yU4xn6wuepsBUrfQW8cuFy/view>. Accessed November 29, 2019.