Noname manuscript No.
(will be inserted by the editor)

Two modifications for Loidreau's code-based cryptosystem

Wenshuo Guo · Fang-Wei Fu

Received: date / Accepted: date

Abstract This paper presents two modifications for Loidreau's code-based cryptosystem. Loidreau's cryptosystem is a rank metric code-based cryptosystem constructed by using Gabidulin codes in the McEliece setting. Recently a polynomial-time key recovery attack was proposed to break Loidreau's cryptosystem in some cases. To prevent this attack, we propose the use of subcodes to disguise the secret codes in Modification I. In Modification II, we choose a random matrix of low column rank over $\mathbb{F}_q$ to mix with the secret matrix. According to our analysis, these two modifications can both resist the existing structural attacks. Additionally, we adopt the systematic generator matrix of the public code to make a reduction in the public-key size.

Keywords Code-based cryptography · Rank metric codes · Gabidulin codes · Loidreau's cryptosystem

Mathematics Subject Classification (2010) 94B05 · 81Q99

1 Introduction

In 1978, McEliece proposed the first code-based public-key cryptosystem, namely the well-known McEliece cryptosystem based on Goppa codes [1]. Since then cryptologists have made extensive study on its security [2–5]. Apart from some weak keys [6], the McEliece cryptosystem still remains secure in general cases. The main drawback of this cryptosystem lies in its large public-key size, which makes it unpractical in many situations. To overcome this problem, many variants have been proposed. In 1986, Niederreiter [7] introduced a knapsack-type cryptosystem using GRS codes, which was shown to be insecure by Sidelnikov in [8]. But if we

✉ Wenshuo Guo
Chern Institute of Mathematics and LPMC, Nankai University, Tianjin 300071, China
E-mail: ws_guo@mail.nankai.edu.cn

Fang-Wei Fu
Chern Institute of Mathematics and LPMC, Nankai University, Tianjin 300071, China
E-mail: fwfu@nankai.edu.cn

use Goppa codes in the Niederreiter setting, it was proved to be equivalent to the McEliece cryptosystem in terms of security [9]. GRS codes allow us to reduce the public-key size due to their optimal error-correcting capability. Many variants based on GRS codes were proposed after Niederreiter's work. However, nearly all of these variants were broken one after another because of GRS codes being highly structured. In the variant [10], the BBCRS cryptosystem, the authors proposed the use of a dense matrix rather than a permutation matrix to disguise the structure of the underlying GRS code. In this proposal, the column scrambler is a matrix of the form $(R+T)^{-1}$, where T is a sparse matrix and R is a dense matrix of low rank. With this approach, the public code seems quite different from GRS codes. This variant therefore can resist some known structural attacks, such as the Sidelnikov-Shestakov attack [8]. However, in [11] the authors presented a polynomial-time key recovery attack against this variant in some cases. Although we can adjust the parameters to prevent such an attack, it would bring some other problems such as the decryption complexity increasing exponentially and a higher request of error-correcting capability for the underlying code.

In 1985 Gabiduin [12] introduced a new family of rank metric codes, known as the Gabidulin codes. Since the complexity of decoding general rank metric codes is much higher than that of decoding Hamming metric codes [15,16], it is feasible to obtain much smaller public-key sizes by building cryptosystems in the rank metric. In [17] the authors proposed to use Gabidulin codes in the McEliece setting and introduced the GPT cryptosystem. Unfortunately, several structural attacks were put forward to completely break this system [23–25].To prevent these attacks, variants based on different masking skills for Gabidulin codes were proposed [18–22]. But in [28] the authors declare the failure of all the previous masking techniques for Gabidulin codes. In [26] Faure and Loidreau proposed a cryptosystem also relying on the Gabidulin codes but not in the McEliece setting. Until the work in [27], the Faure-Loidreau system had never been severely attacked. Recently, in [29] Loidreau proposed a cryptosystem constructed by using Gabidulin codes in the McEliece setting. Different from the original GPT cryptosystem, the isometric matrix is replaced with a matrix whose inverse is taken in an $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^m}$ of dimension λ. By doing this, the public code seems quite random. Loidreau claimed that his proposal could prevent the existing structural attacks. However, this claim was proved to be invalid by the authors in [30] when $\lambda = 2$ and the code rate is greater than $1/2$. Soon after this, the author in [31] generalized this attack to the case of $\lambda > 2$ and the code rate greater than $1 - \frac{1}{\lambda}$. However, it is feasible to prevent this attack even when the secret code rate is greater than $1 - \frac{1}{\lambda}$ according to our analysis.

The rest of this paper is organised as follows. In Section 2 notations and some concepts about rank metric codes used throughout this paper are given. Section 3 is devoted to a simple descripton of Loidreau's cryptosystem. In Section 4 we shall introduce part of the Coggia-Couvreur attack (please refer to [30] for more details). Following this, our two modifications for Loidreau's cryptosystem will be introduced in Section 5, then security analysis of our modifications will be given in Section 6. In Section 7, we will give some suggested parameters for different security levels and make a comparison with Loidreau's original scheme in Table 1 and with some NIST-PQC submissions in Table 2. Section 8 is our conclusion.

2 Preliminaries

2.1 Notations and basic concepts

Let q be a prime power. Denote by $\mathbb{F}_q$ the finite field with q elements, and $\mathbb{F}_{q^m}$ an extension field of $\mathbb{F}_q$ of degree m. For two positive integers k and n, denote by $\mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ the set of all $k \times n$ matrices over $\mathbb{F}_{q^m}$, and by $GL_n(\mathbb{F}_{q^m})$ the set of all $n \times n$ invertible matrices over $\mathbb{F}_{q^m}$. For a matrix $M \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$, the column rank of M with respect to $\mathbb{F}_q$, denoted by $\mathrm{Clr}_q(M)$, is the largest number of columns of M linearly independent over $\mathbb{F}_q$. Denote by $\langle M \rangle$ the vector space spanned by rows of M over $\mathbb{F}_{q^m}$.

An $[n,k]$ linear code $\mathcal{C}$ 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 Euclidean inner product over $\mathbb{F}_{q^m}$. A $k \times n$ full-rank matrix $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ is called a generator matrix of $\mathcal{C}$ if the vector space $\langle G \rangle$ is exactly the code $\mathcal{C}$. A generator matrix of $\mathcal{C}^\perp$ is called a parity-check matrix of $\mathcal{C}$.

2.2 Rank metric codes

Now we recall some basic concepts for rank metric and rank metric codes.

Definition 1 For a vector $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{F}_{q^m}^n$, the support of $\boldsymbol{x}$ denoted by $\mathrm{Supp}(\boldsymbol{x})$, is defined to be the linear space spanned by coordinates of $\boldsymbol{x}$ over $\mathbb{F}_q$. Formally we have

$$\mathrm{Supp}(\boldsymbol{x}) = \left\{ \sum_{i=1}^{n} \lambda_i x_i : \lambda_i \in \mathbb{F}_q, 1 \leqslant i \leqslant n \right\}.$$

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

Given two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_{q^m}^n$, the rank distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, denoted by $d_R(\boldsymbol{x}, \boldsymbol{y})$, is defined to be the rank weight of $\boldsymbol{x} - \boldsymbol{y}$. It is easy to verify that the function $d_R(\cdot, \cdot)$ defines a proper metric on $\mathbb{F}_{q^m}^n$. A code endowed with the rank metric is called a rank metric code, and in this paper by rank metric codes we always mean linear rank metric codes.

Definition 3 For a rank metric code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$, the minimum rank distance of $\mathcal{C}$, denoted by $d(\mathcal{C})$, is defined as

$$d(\mathcal{C}) = \min\{d_R(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x}, \boldsymbol{y} \in \mathcal{C} \text{ and } \boldsymbol{x} \neq \boldsymbol{y}\}.$$

It is easy to verify that the minimum rank (Hamming) distance of a linear code is equal to its minimum rank (Hamming) weight. In the context of Hamming metric codes, the minimum distance d of an $[n,k]$ linear code satisfies the Singleton bound $d \leqslant n - k + 1$ [32]. Similarly, the minimum rank distance of a rank metric code $\mathcal{C}$ satisfies the following Singleton-style bound.

Theorem 1 (Singleton-style bound) *[33] Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be an $[n, k]$ rank metric code, then the minimum rank distance of $\mathcal{C}$ with respect to $\mathbb{F}_q$ satisfies the following inequality*

$$d(\mathcal{C}) \leqslant n - k + 1.$$

Remark 1 A linear code attaining the Singleton-style bound is called a Maximum Rank Distance (MRD) code. Apparently an $[n, k]$ MRD code can correct up to $\lfloor \frac{n-k}{2} \rfloor$ rank errors.

The following proposition implies that the maximum rank weight of a rank metric code is bounded from above by the column rank of its generator matrix.

Proposition 1 *For a matrix $M \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ with $\mathrm{Clr}_q(M) = r$, the maximum rank weight of the code $\langle M \rangle$ is bounded by r from above.*

Proof It suffices to prove that for any $\boldsymbol{v} \in \langle M \rangle$, we have $w_R(\boldsymbol{v}) \leqslant r$. Since $\mathrm{Clr}_q(M) = r$, there exists $Q \in GL_n(\mathbb{F}_q)$ such that $MQ = [M'|O]$, where $M' \in \mathcal{M}_{k,r}(\mathbb{F}_{q^m})$ with $\mathrm{Clr}_q(M') = r$ and O is a zero matrix. For any $\boldsymbol{v} \in \langle M \rangle$, there exists $\boldsymbol{x} \in \mathbb{F}_{q^m}^k$ such that $\boldsymbol{v} = \boldsymbol{x}M$ and

$$\boldsymbol{v}Q = \boldsymbol{x}MQ = \boldsymbol{x}[M'|O] = (\boldsymbol{x}', \boldsymbol{0}),$$

where $\boldsymbol{x}' \in \mathbb{F}_{q^m}^r$ and $\boldsymbol{0}$ is a zero vector. Hence we have $w_R(\boldsymbol{v}) = w_R(\boldsymbol{v}Q) \leqslant r$. This concludes the proof.

2.3 Gabidulin codes

Gabidulin codes can be viewed as an analogue of GRS codes in the rank metric setting, and these two types of codes resemble each other closely in the construction principle. GRS codes admit generator matrices with the Vandermonde structure, while Gabidulin codes can be described by Moore matrices defined as follows.

Definition 4 For an integer s, denote by $[s]$ the s-th Frobenius power q^s. A matrix $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ is called a Moore matrix generated by $\boldsymbol{a} = (a_1, \cdots, a_n) \in \mathbb{F}_{q^m}^n$ if the s-th row of G equals the coordinate-wise Frobenius power $\boldsymbol{a}^{[s-1]} = (a_1^{[s-1]}, \cdots, a_n^{[s-1]})$ for each $1 \leqslant s \leqslant k$. Formally we have

$$G = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_1^{[1]} & a_2^{[1]} & \cdots & a_n^{[1]} \\ \vdots & \vdots & & \vdots \\ a_1^{[k-1]} & a_2^{[k-1]} & \cdots & a_n^{[k-1]} \end{pmatrix}. \tag{1}$$

For a matrix $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$, we define $G^{[s]} = (G_{ij}^{[s]})$. For a set $S \subseteq \mathbb{F}_{q^m}^n$, we define $S^{[s]} = \{\boldsymbol{x}^{[s]} : \boldsymbol{x} \in S\}$. For a linear code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$, it is easy to verify that $\mathcal{C}^{[s]}$ is also an $\mathbb{F}_{q^m}$-linear code.

Definition 5 (Gabidulin codes) For a vector $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{a}) = n \leqslant m$, let G be the $k \times n$ Moore matrix generated by $\boldsymbol{a}$. The $[n, k]$ Gabidulin code $\mathcal{G}_{n,k}(\boldsymbol{a})$ over $\mathbb{F}_{q^m}$ generated by $\boldsymbol{a}$ is defined to be the linear space $\langle G \rangle$, namely we have $\mathcal{G}_{n,k}(\boldsymbol{a}) = \langle G \rangle$.

A major reason for Gabidulin codes being widely used in the design of cryptosystems consists in their remarkable error-correcting capability and simple algebraic structure. Now we recall some properties of Gabidulin codes through the following two theorems without proving.

Theorem 2 *[34] The Gabidulin code $\mathcal{G}_{n,k}(\boldsymbol{a})$ is an MRD code. In other words, $\mathcal{G}_{n,k}(\boldsymbol{a})$ attains the Singleton-style bound for rank metric codes.*

According to Theorem 2, the minimum rank distance of $\mathcal{G}_{n,k}(\boldsymbol{a})$ is $n-k+1$. This implies that any $\lfloor \frac{n-k}{2} \rfloor$ rank errors can be corrected. In fact, several efficient docoding algorithms for Gabidulin codes already exist (for instance [12–14]).

Theorem 3 *[27] The dual code of $\mathcal{G}_{n,k}(\boldsymbol{a})$ is the Gabidulin code $\mathcal{G}_{n,n-k}(\boldsymbol{b}^{[k-n+1)]})$ for some $\boldsymbol{b} \in \mathcal{G}_{n,n-1}(\boldsymbol{a})^{\perp}$ with $w_R(\boldsymbol{b}) = n$.*

3 Loidreau's scheme

For a vector $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{a}) = n$, denote by G a generator matrix of $\mathcal{G}_{n,k}(\boldsymbol{a})$. For a positive integer $\lambda \ll m$, let $\mathcal{V} \subseteq \mathbb{F}_{q^m}$ be an $\mathbb{F}_q$-linear space of dimension λ. Now we give a simple description of Loidreau's scheme through the following three algorithms.

- Key Generation
  Randomly choose $P \in GL_n(\mathbb{F}_{q^m})$ whose entries are taken from $\mathcal{V}$ and compute $G_{pub} = GP^{-1}$. We publish (G_{pub}, t) as the public key where $t = \lfloor \frac{n-k}{2\lambda} \rfloor$, and keep $(\boldsymbol{a}, P)$ as the secret key.
- Encryption
  For a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^k$, randomly choose a vector $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{e}) = t$. The ciphertext corresponding to $\boldsymbol{m}$ is computed as $\boldsymbol{c} = \boldsymbol{m}G_{pub} + \boldsymbol{e}$.
- Decryption
  Compute $\boldsymbol{c}' = \boldsymbol{c}P = \boldsymbol{m}G + \boldsymbol{e}P$. Since $w_R(\boldsymbol{e}P) \leqslant w_R(\boldsymbol{e}) \cdot \dim_q(\mathcal{V}) \leqslant \lfloor \frac{n-k}{2} \rfloor$, decoding $\boldsymbol{c}'$ will lead to the plaintext $\boldsymbol{m}$.

4 The Coggia-Couvreur attack

Before describing the Coggia-Couvreur attack, we first introduce a distinguisher for Gabidulin codes. This distinguisher provides us with a method of distinguishing Gabidulin codes from general ones.

4.1 The distinguisher for Gabidulin codes

Most of cryptosystems based on Gabidulin codes have been proved to be insecure against structural attacks. Although these attacks were proposed to cryptanalyze different variants of the GPT cryptosystem, the principle for their work is based on the same observation that one can distinguish Gabidulin codes from general ones by performing a simple operation on these codes.

Given a random linear code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ of dimension $k \leqslant n/2$, the expected dimension of the code $\mathcal{C} + \mathcal{C}^{[1]}$ equals $2k$, or equivalently $\mathcal{C} \cap \mathcal{C}^{[1]} = \{\boldsymbol{0}\}$ holds with high probability. But for a Gabidulin code $\mathcal{G}_{n,k}(\boldsymbol{a})$, we have $\mathcal{G}_{n,k}(\boldsymbol{a})+\mathcal{G}_{n,k}(\boldsymbol{a})^{[1]} = \mathcal{G}_{n,k+1}(\boldsymbol{a})$, namely the dimension of $\mathcal{G}_{n,k}(\boldsymbol{a}) + \mathcal{G}_{n,k}(\boldsymbol{a})^{[1]}$ is $k+1$. More generally, we have the following two propositions.

Proposition 2 *[30] Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be a random linear code of length n and dimension k. For a non-negative integer l and a positive integer $s < k$, we have*

$$\Pr\left(\dim(\mathcal{C} + \mathcal{C}^{[1]} + \cdots + \mathcal{C}^{[s]}) \leqslant \min\{n, (s+1)k\} - l\right) = O(q^{-ml}).$$

Proposition 3 *[30] Let $k \leqslant n$ and s be a positive integer, then for any $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{a}) = n$, we have*

$$\mathcal{G}_{n,k}(\boldsymbol{a}) \cap \mathcal{G}_{n,k}(\boldsymbol{a})^{[1]} = \mathcal{G}_{n,k-1}(\boldsymbol{a}^{[1]});$$
$$\mathcal{G}_{n,k}(\boldsymbol{a}) + \mathcal{G}_{n,k}(\boldsymbol{a})^{[1]} + \cdots + \mathcal{G}_{n,k}(\boldsymbol{a})^{[s]} = \mathcal{G}_{n,k+s}(\boldsymbol{a}).$$

4.2 Description of the Coggia-Couvreur attack

In this part we investigate the structural vulnerability of Loidreau's cryptosystem in the case of $\lambda = 2$ and the dimension of the public code $\mathcal{C}_{pub} = \langle G_{pub}\rangle$ being greater than $n/2$. The principle for the Coggia-Couvreur attack lies in Propositions 2 and 3. Instead of directly operating the public code, the authors in [30] consider the dual of the public code because of the following lemma.

Lemma 1 *[30] Any parity-check matrix H_{pub} of $\mathcal{C}_{pub}$ can be expressed as*

$$H_{pub} = H_{sec}P^T,$$

where H_{sec} is a parity-check matrix of the secret Gabidulin code $\mathcal{G}_{n,k}(\boldsymbol{a})$.

The authors considered the case of $\lambda = 2$, namely the linear space $\mathcal{V} \subseteq \mathbb{F}_{q^m}$ has dimension 2 over $\mathbb{F}_q$. Suppose $\mathcal{V}$ is spanned by $\alpha, \beta \in \mathbb{F}_{q^m}^*$ over $\mathbb{F}_q$, namely $\mathcal{V} = \langle \alpha, \beta\rangle_{\mathbb{F}_q}$. Let $H'_{sec} = \alpha H_{sec}$ and $P' = \alpha^{-1}P$, apparently we have $H_{pub} = H'_{sec}P'^T$. It is easy to see that H'_{sec} spans the same code as H_{sec} and entries of P' are contained in $\mathcal{V}' = \langle 1, \alpha^{-1}\beta\rangle_{\mathbb{F}_q}$. Hence it is reasonable to suppose that $\mathcal{V} = \langle 1, \gamma\rangle_{\mathbb{F}_q}$ for some $\gamma \in \mathbb{F}_{q^m}^*$. In this situation, we can express P^T in the form of

$$P^T = P_0 + \gamma P_1,$$

where $P_0, P_1 \in \mathcal{M}_{n,n}(\mathbb{F}_q)$.

According to Theorem 3, there exists some $\boldsymbol{b} \in \mathcal{G}_{n,n-1}(\boldsymbol{a})^{\perp}$ with $w_R(\boldsymbol{b}) = n$ such that $\mathcal{G}_{n,k}(\boldsymbol{a})^{\perp} = \mathcal{G}_{n,n-k}(\boldsymbol{b})$. We define

$$\boldsymbol{g} = \boldsymbol{b}P_0, \boldsymbol{h} = \boldsymbol{b}P_1.$$

As for the triple $(\gamma, \boldsymbol{g}, \boldsymbol{h})$, the authors made the following two assumptions:

(1) $\mathcal{G}_{n,n-k+2}(\boldsymbol{g}) \cap \mathcal{G}_{n,n-k+2}(\boldsymbol{h}) = \{\boldsymbol{0}\}$ and $w_R(\boldsymbol{g}), w_R(\boldsymbol{h}) \geq n-k+2$;
(2) $m > 2$ and γ is not contained in any proper subfield of $\mathbb{F}_{q^m}$.

The rationality for these two assumptions can be explained as follows. According to the authors' experiments on Magma, Assumption (1) holds with an extremely high probability. Apparently $m > 2$ is reasonable because of $m \geqslant n$. On the other hand, if γ is contained in some proper subfield of $\mathbb{F}_{q^m}$, then the adversary can find γ through the exhausting method for the reason that even the union of all proper subfields of $\mathbb{F}_{q^m}$ contains much less elements than $\mathbb{F}_{q^m}$. Hence γ cannot be contained in any proper subfield of $\mathbb{F}_{q^m}$.

The core of the Coggia-Couvreur attack is to find the triple $(\gamma, \boldsymbol{g}, \boldsymbol{h})$ or one of its equivalent forms (see [30] for more details). With the knowledge of the triple $(\gamma, \boldsymbol{g}, \boldsymbol{h})$ or one of its equivalent forms, one can decrypt any valid ciphertext in polynomial time and hence completely break Loidreau's cryptosystem.

What follows are two lemmas that will be useful for analysing the security of our modifications. For the remaining part of the Coggia-Couvreur attack, interested readers can refer to [30] for more details. Now we introduce these two lemmas without proving.

Lemma 2 *[30] The code $\mathcal{C}_{pub}^{\perp}$ is spanned by*

$$\boldsymbol{g} + \gamma\boldsymbol{h}, \boldsymbol{g}^{[1]} + \gamma\boldsymbol{h}^{[1]}, \cdots, \boldsymbol{g}^{[n-k-1]} + \gamma\boldsymbol{h}^{[n-k-1]}. \tag{2}$$

Lemma 3 *[30] Under Assumption (1), we have that $\mathcal{C}_{pub}^{\perp} + {\mathcal{C}_{pub}^{\perp}}^{[1]}$ is spanned by*

$$\boldsymbol{g} + \gamma\boldsymbol{h} \text{ and } \boldsymbol{g}^{[1]}, \boldsymbol{h}^{[1]}, \cdots, \boldsymbol{g}^{[n-k-1]}, \boldsymbol{h}^{[n-k-1]} \text{ and } \boldsymbol{g}^{[n-k]} + \gamma^{[1]}\boldsymbol{h}^{[n-k]},$$

and

$$(\mathcal{C}_{pub}^{\perp} + {\mathcal{C}_{pub}^{\perp}}^{[1]}) \cap ({\mathcal{C}_{pub}^{\perp}}^{[1]} + {\mathcal{C}_{pub}^{\perp}}^{[2]})$$

is spanned by

$$\boldsymbol{g}^{[1]} + \gamma^{[1]}\boldsymbol{h}^{[1]} \text{ and } \boldsymbol{g}^{[2]}, \boldsymbol{h}^{[2]}, \cdots, \boldsymbol{g}^{[n-k-1]}, \boldsymbol{h}^{[n-k-1]} \text{ and } \boldsymbol{g}^{[n-k]} + \gamma^{[1]}\boldsymbol{h}^{[n-k]}.$$

Remark 2 Similar to Lemma 3, it is easy to verify that

$$(\mathcal{C}_{pub}^{\perp} + {\mathcal{C}_{pub}^{\perp}}^{[1]}) \cap ({\mathcal{C}_{pub}^{\perp}}^{[1]} + {\mathcal{C}_{pub}^{\perp}}^{[2]}) \cap \cdots \cap ({\mathcal{C}_{pub}^{\perp}}^{[n-k-1]} + {\mathcal{C}_{pub}^{\perp}}^{[n-k]}) \tag{3}$$

yields a code spanned by

$$\boldsymbol{g}^{[n-k-1]} + \gamma^{[n-k-1]}\boldsymbol{h}^{[n-k-1]} \text{ and } \boldsymbol{g}^{[n-k]} + \gamma^{[1]}\boldsymbol{h}^{[n-k]}. \tag{4}$$

The key point for the Coggia-Couvreur attack is that one can obtain (4) by computing (3). But if ${\mathcal{C}_{pub}^{\perp}}^{[i]} + {\mathcal{C}_{pub}^{\perp}}^{[i+1]} (0 \leqslant i \leqslant n-k-1)$ happens to be the whole space $\mathbb{F}_{q^m}^n$, computing (4) will lead to nothing but the whole space itself, which means that the Coggia-Couvreur attack will fail in this situation. Our first modification for Loidreau's cryptosystem is inspired by this observation. On the other hand, if $\mathcal{C}_{pub}^{\perp}$ does not contain the full code spanned by (2), then one cannot obtain (4) from (3) either even if ${\mathcal{C}_{pub}^{\perp}}^{[i]} + {\mathcal{C}_{pub}^{\perp}}^{[i+1]} (0 \leqslant i \leqslant n-k-1)$ is not the whole space. Modification II is based on this observation and this is really true according to our analysis in Section 6.

5 Our modifications

In code-based cryptography, randomness is widely used in both the key generation and encryption procedures. In terms of the intersection of a given linear code and a randomly chosen linear space, we have the following proposition.

Proposition 4 *Let n, k, l be positive integers with $k + l < n$. Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be a linear code of dimension k, and $\mathcal{V}$ be a random linear subspace of $\mathbb{F}_{q^m}^n$ of dimension l. In terms of the intersection $\mathcal{C} \cap \mathcal{V}$, we have the following inequality*

$$\Pr\{\mathcal{C} \cap \mathcal{V} = \{\mathbf{0}\}\} \geqslant 1 - O\big(q^{-ms}\big),$$

where $s \geqslant 2$ is a positive integer.

Proof Exploiting the Gaussian coefficient, the number of l-dimensional subspaces of $\mathbb{F}_{q^m}^n$ linearly independent of $\mathcal{C}$ can be computed as

$$N_1 = \prod_{i=0}^{l-1} \frac{(q^m)^n - (q^m)^{k+i}}{(q^m)^l - (q^m)^i} = \prod_{i=0}^{l-1} \frac{q^{mn} - q^{m(k+i)}}{q^{ml} - q^{mi}}.$$

Similarly, the number of all l-dimensional subspaces of $\mathbb{F}_{q^m}^n$ can be computed as

$$N_2 = \prod_{i=0}^{l-1} \frac{(q^m)^n - (q^m)^i}{(q^m)^l - (q^m)^i} = \prod_{i=0}^{l-1} \frac{q^{mn} - q^{mi}}{q^{ml} - q^{mi}}.$$

Then the target probability $\Pr\{\mathcal{C} \cap \mathcal{V} = \{\mathbf{0}\}\}$ can be computed as

$$\begin{aligned} \frac{N_1}{N_2} &= \prod_{i=0}^{l-1} \frac{q^{mn} - q^{m(k+i)}}{q^{mn} - q^{mi}} \\ &= \prod_{i=0}^{l-1} \frac{q^{mn} - q^{mi} - q^{mk+mi} + q^{mi}}{q^{mn} - q^{mi}} \\ &= \prod_{i=0}^{l-1} \Big(1 - \frac{q^{mk} - 1}{q^{m(n-i)} - 1}\Big) \\ &\geqslant \Big(1 - \frac{q^{mk} - 1}{q^{m(n-l+1)} - 1}\Big)^l. \end{aligned} \tag{5}$$

By Taylor expansion, the right hand side of (5) can be expressed as

$$\begin{aligned} \Big(1 - \frac{q^{mk} - 1}{q^{m(n-l+1)} - 1}\Big)^l &= 1 - l \cdot \frac{q^{mk} - 1}{q^{m(n-l+1)} - 1} + o\Big(\frac{q^{mk} - 1}{q^{m(n-l+1)} - 1}\Big) \\ &= 1 - O\Big(q^{-m(n-k-l+1)}\Big). \end{aligned}$$

Let $s = n - k - l + 1$, apparently $s \geqslant 2$ because of $k + l < n$. Finally we have $\Pr\{\mathcal{C} \cap \mathcal{V} = \{\mathbf{0}\}\} \geqslant 1 - O\big(q^{-ms}\big)$. This completes the proof.

Remark 3 Proposition 4 states a fact that for a linear code $\mathcal{C}$ and a randomly chosen linear space $\mathcal{V}$, we have that $\mathcal{C} \cap \mathcal{V} = \{\mathbf{0}\}$ holds with high probability. Meanwhile, it is reasonable to conclude that for a $k \times n$ full-rank matrix H and a randomly chosen $l \times n$ full-rank matrix A with $k + l < n$, the block matrix $\begin{pmatrix} A \\ H \end{pmatrix}$ is of full rank with high probability.

5.1 Description of Modification I

Let $\mathcal{G}$ be an $[n,k]$ Gabidulin code generated by $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{a}) = n$. Denote by H a parity-check matrix of $\mathcal{G}$. For a positive integer $l \geqslant k - \frac{n}{2}$, randomly choose an $l \times n$ full-rank matrix A over $\mathbb{F}_{q^m}$ and set $H_{sub} = \begin{pmatrix} A \\ H \end{pmatrix}$. Let G_{sub} be a generator matrix of $\langle H_{sub} \rangle^{\perp}$. By Proposition 4, H_{sub} has rank $k + l$ with high probability. It would be well if we assume that H_{sub} is of full rank, otherwise we rechoose the matrix A. Apparently G_{sub} spans a subcode of $\mathcal{G}$ of dimension $k' = k - l$. For a positive integer $\lambda \ll m$, let $\mathcal{V} \subseteq \mathbb{F}_{q^m}$ be an $\mathbb{F}_q$-linear space of dimension λ.

- Key generation
  Let $P \in GL_n(\mathbb{F}_{q^m})$ with entries contained in $\mathcal{V}$. Without loss of generality, we assume that the submatrix of $G_{sub}P^{-1}$ formed by the first k' columns is invertible. Choose a matrix $S \in GL_{k'}(\mathbb{F}_{q^m})$ to change $G_{pub} = SG_{sub}P^{-1}$ into systematic form. We publish (G_{pub}, t) as the public key where $t = \lfloor \frac{n-k}{2\lambda} \rfloor$, and keep $(\boldsymbol{a}, P)$ as the secret key.
- Encryption
  For a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^{k'}$, randomly choose $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{e}) = t$. Then the ciphertext corresponding to $\boldsymbol{m}$ is computed as $\boldsymbol{c} = \boldsymbol{m}G_{pub} + \boldsymbol{e}$.
- Decryption
  For a ciphertext $\boldsymbol{c}$, compute $\boldsymbol{c}' = \boldsymbol{c}P = \boldsymbol{m}SG_{sub} + \boldsymbol{e}P$. Since $w_R(\boldsymbol{e}P) \leqslant w_R(\boldsymbol{e}) \cdot \lambda \leqslant \lfloor \frac{n-k}{2} \rfloor$. Applying the decoding procedure of $\mathcal{G}$ to $\boldsymbol{c}'$ will lead to $\boldsymbol{e}' = \boldsymbol{e}P$, then we have $\boldsymbol{e} = \boldsymbol{e}'P^{-1}$. The restriction of $\boldsymbol{c} - \boldsymbol{e}$ to the first k' coordinates will be the plaintext $\boldsymbol{m}$.

Remark 4 According to the analysis in Section 4.2, we can always assume that $1 \in \mathcal{V}$. If $\lambda = 1$, there will be $\mathcal{V} = \mathbb{F}_q$ and $P^{-1} \in GL_n(\mathbb{F}_q)$. In this situation, G_{pub} spans a subcode of $\mathcal{G}$. Then one can exploit the r-Frobenius weak attack [35] to completely break this modification. To prevent this attack, we should make sure that $\lambda \geqslant 2$ in Modification I.

5.2 Description of Modification II

Let $\mathcal{G}$ be an $[n,k]$ Gabidulin code generated by $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{a}) = n$. Denote by G a generator matrix of $\mathcal{G}$. For a positive integer $l \ll \min\{k, n-k\}$, randomly choose $M \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ with $\mathrm{Clr}_q(M) = l$ and let $G_M = G + M$. It is easy to see that G_M is of full rank. Indeed, if there exists $\boldsymbol{x} \in \mathbb{F}_{q^m}^k$ such that $\boldsymbol{x}G_M = \boldsymbol{0}$, then we have $\boldsymbol{x}G \in \langle M \rangle$. By Proposition 1, the maximum rank weight of $\langle M \rangle$ does not exceed l. Together with $d(\mathcal{G}) = n - k + 1 \gg l$, we have $\boldsymbol{x}G = \boldsymbol{0}$ and hence $\boldsymbol{x} = \boldsymbol{0}$. For a positive integer $\lambda \ll m$, let $\mathcal{V} \subseteq \mathbb{F}_{q^m}$ be an $\mathbb{F}_q$-linear space of dimension λ.

- Key generation
  Let $P \in GL_n(\mathbb{F}_{q^m})$ with entries contained in $\mathcal{V}$. Without loss of generality, we assume that the submatrix of G_MP^{-1} formed by the first k columns is invertible. Choose a matrix $S \in GL_k(\mathbb{F}_{q^m})$ to change $G_{pub} = SG_MP^{-1}$ into systematic form. We publish (G_{pub}, t) as the public key where $t = \lfloor \frac{n-k-2l}{2\lambda} \rfloor$, and keep (S, G, P) as the secret key.
- Encryption

For a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^k$, randomly choose a vector $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ with $w_R(\boldsymbol{e}) = t$. Then the ciphertext corresponding to $\boldsymbol{m}$ is computed as $\boldsymbol{c} = \boldsymbol{m}G_{pub} + \boldsymbol{e}$.

- Decryption

  For a ciphertext $\boldsymbol{c}$, compute $\boldsymbol{c}' = \boldsymbol{c}P = \boldsymbol{m}SG + \boldsymbol{m}SM + \boldsymbol{e}P$. Since

$$w_R(\boldsymbol{m}SM + \boldsymbol{e}P) \leqslant w_R(\boldsymbol{m}SM) + w_R(\boldsymbol{e}P) \leqslant l + \lambda t \leqslant \lfloor \frac{n-k}{2} \rfloor,$$

  applying the decoding procedure of $\mathcal{G}$ to $\boldsymbol{c}'$ will lead to $\boldsymbol{m}SG$. Then the plaintext $\boldsymbol{m}$ can be recovered by solving a linear system with a complexity of $O(n^3)$.

Remark 5 Similar to the analysis in Remark 4, we should make sure that $\lambda \geqslant 2$ in this modification. Otherwise, Modification II can be reduced to the GPT cryptosystem that has been completely broken.

6 Security analysis

In general, there are two types of attacks on code-based cryptosystems, namely the generic attacks and structural attacks.

Generic attacks. These attacks aim to recover the plaintext directly from the ciphertext when nothing but the public key is known. In the context of code-based cryptography, generic attacks involve the problem of decoding general linear codes or equivalently the syndrome decoding problem, which has been proved to be NP-complete by Berlekamp et al. in [36]. However, the general decoding problem in the rank metric, or equivalently the rank syndrome decoding (RSD) problem, is not known to be NP-complete. In the paper [37], the authors proved that a randomized reduction exists from the RSD problem to the general decoding problem in the Hamming metric. There are mainly two types of approaches to solve the RSD problem, one is the combinatorial method and the other is the algebraic method. Up to now, the best combinatorial attack on the RSD problem is the one proposed in [38], and the best algebraic attacks are those proposed in [39, 40].

Suppose $\mathcal{C}$ is an $[n, k]$ rank metric code over $\mathbb{F}_{q^m}$, correcting up to t rank errors. Let $\boldsymbol{y} = \boldsymbol{c} + \boldsymbol{e}$ be the received word, where $\boldsymbol{c}$ is a codeword in $\mathcal{C}$ and $\boldsymbol{e}$ is a random vector with $w_R(\boldsymbol{e}) = t$. To recover the error vector $\boldsymbol{e}$, the combinatorial attack proposed in [38] requires

$$O\left(m^2 n^2 q^{t\lceil \frac{m(k+1)}{n} \rceil - m}\right)$$

operations in $\mathbb{F}_q$.

The algebraic attack proposed in [39] is divided into two categories, namely the overdetermined case and the underdetermined case. An RSD instance with the parameters (m, n, k, t) is called overdetermined if the following condition is satisfied

$$m\binom{n-k-1}{t} \geqslant \binom{n}{t} - 1,$$

otherwise we call it underdetermined.

For the overdetermined case, solving the RSD problem requires

$$O\left(m\binom{n-p-k-1}{t}\binom{n-p}{t}^{\omega-1}\right)$$

operations in $\mathbb{F}_q$, where $p = \max\{1 \leqslant i \leqslant n : m\binom{n-i-k-1}{t} \geqslant \binom{n-i}{t} - 1\}$ and $\omega = 2.81$ represents the constant of linear algebra.

For the underdetermined case, the RSD instance can either be reduced to the overdetermined case by introducing another parameter a or be kept underdetermined. In the former situation, the complexity of solving the RSD problem is

$$O\left(q^{at}m\binom{n-k-1}{t}\binom{n-a}{t}^{\omega-1}\right),$$

where $a \geqslant 0$ is the smallest integer such that $m\binom{n-k-1}{t} \geqslant \binom{n-a}{t} - 1$. In this paper we consider the base field $\mathbb{F}_3$, then the complexity in the latter situation can be expressed as

$$O\left(K(t+1)\binom{n}{t}^2\binom{K+b-1}{b}^2\right),$$

where $K = km + 1$ and b is the smallest positive integer such that $b < 3$ satisfies the following condition

$$\binom{n}{t}\binom{K+b-1}{b} - 1 \leqslant \sum_{i=1}^{b}(-1)^{i+1}\binom{n}{t+i}\binom{m+i-1}{i}\binom{K+b-i-1}{b-i}.$$

In the paper [40], the complexity for the overdetermined case described above is $O\left(\left(\frac{((m+n)t)^t}{t!}\right)^{\omega}\right)$ in $\mathbb{F}_q$, and $O\left(\left(\frac{((m+n)t)^{t+1}}{(t+1)!}\right)^{\omega}\right)$ for the underdetermined case respectively.

Structural attacks. These attacks aim to recover the structure of the secret code from a random-looking public matrix. In fact, recovering the structure amounts to obtaining the secret key in some sense, which means that the cryptosystem will be completely broken in this situation. In [29], Loidreau argued that his cryptosystem could resist the invariant subspace attack, also known as Overbeck's attack. Since our modifications exploit the same masking technique to disguise the structure of the secret code, naturally we believe that our modifications can also prevent Overbeck's attack. Therefore, in the remaining part of this section we only consider the security against the Coggia-Couvreur attack.

6.1 Analysis of Modification I

Before giving the analysis, we shall introduce the following theorem. This theorem states a simple fact that if $\mathcal{C} \subseteq \mathbb{F}_{q^m}^m$ is a linear code with a generator matrix G, then its s-th Frobenius power $\mathcal{C}^{[s]}$ is also a linear code over $\mathbb{F}_{q^m}$ and has $G^{[s]}$ as a generator matrix.

Theorem 4 *Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be an $[n, k]$ linear code that has G as a generator matrix. For any integer s, $\mathcal{C}^{[s]}$ is also an $[n, k]$ linear code over $\mathbb{F}_{q^m}$ and has $G^{[s]}$ as a generator matrix.*

Proof On the one hand. For any $\boldsymbol{u} \in \mathcal{C}^{[s]}$, there exists $\boldsymbol{x} \in \mathbb{F}_{q^m}^k$ such that $\boldsymbol{u} = (\boldsymbol{x}G)^{[s]} = \boldsymbol{x}^{[s]}G^{[s]} \in \langle G^{[s]}\rangle$, then we have

$$\mathcal{C}^{[s]} \subseteq \langle G^{[s]}\rangle.$$

On the other hand. For any $\boldsymbol{v} \in \langle G^{[s]} \rangle$, there exists $\boldsymbol{x} \in \mathbb{F}_{q^m}^k$ such that $\boldsymbol{v} = \boldsymbol{x}G^{[s]} = (\boldsymbol{x}^{[m-s]}G)^{[s]} \in \mathcal{C}^{[s]}$, then we have

$$\langle G^{[s]} \rangle \subseteq \mathcal{C}^{[s]}.$$

Hence we have $\mathcal{C}^{[s]} = \langle G^{[s]} \rangle$.

It remains to prove that $G^{[s]}$ is of full rank. Suppose there exists $\boldsymbol{x} \in \mathbb{F}_{q^m}^k$ such that $\boldsymbol{x}G^{[s]} = (\boldsymbol{x}^{[m-s]}G)^{[s]} = \boldsymbol{0}$, then we have $\boldsymbol{x}^{[m-s]}G = \boldsymbol{0}$ and consequently $\boldsymbol{x} = \boldsymbol{x}^{[m-s]} = \boldsymbol{0}$ because of G being of full rank. This concludes the proof.

Now we show that ${\mathcal{C}_{pub}^{\perp}}^{[i]} + {\mathcal{C}_{pub}^{\perp}}^{[i+1]}$ $(0 \leqslant i \leqslant n-k-1)$ is exactly the whole space $\mathbb{F}_{q^m}^n$, namely all these $n-k$ codes have dimension n. By Theorem 4, it suffices to consider the case of $\mathcal{C}_{pub}^{\perp} + {\mathcal{C}_{pub}^{\perp}}^{[1]}$.

Let H_{pub} be a parity-check matrix of $\mathcal{C}_{pub}$, then we have $H_{pub} = H_{sub}P^T$ and

$$\mathcal{C}_{pub}^{\perp} = \langle H_{sub}P^T \rangle = \langle HP^T \rangle + \langle AP^T \rangle.$$

Hence

$$\mathcal{C}_{pub}^{\perp} + {\mathcal{C}_{pub}^{\perp}}^{[1]} = \langle HP^T \rangle + \langle HP^T \rangle^{[1]} + \langle AP^T \rangle + \langle AP^T \rangle^{[1]}.$$

According to Lemma 3, $\langle HP^T \rangle + \langle HP^T \rangle^{[1]}$ is spanned by

$$\boldsymbol{g} + \gamma\boldsymbol{h} \text{ and } \boldsymbol{g}^{[1]}, \boldsymbol{h}^{[1]}, \cdots, \boldsymbol{g}^{[n-k-1]}, \boldsymbol{h}^{[n-k-1]} \text{ and } \boldsymbol{g}^{[n-k]} + \gamma^{[1]}\boldsymbol{h}^{[n-k]}, \qquad (6)$$

where $\gamma, \boldsymbol{g}$ and $\boldsymbol{h}$ are defined as in Section 4.

Note that these $2(n-k)$ vectors in (6) are linearly independent over $\mathbb{F}_{q^m}$. Indeed, if there exist $x_i, y_i \in \mathbb{F}_{q^m}$ $(0 \leqslant i \leqslant n-k-1)$ such that

$$x_0(\boldsymbol{g} + \gamma\boldsymbol{h}) + y_0(\boldsymbol{g}^{[n-k]} + \gamma^{[1]}\boldsymbol{h}^{[n-k]}) + \sum_{i=1}^{n-k-1} x_i\boldsymbol{g}^{[i]} + \sum_{i=1}^{n-k-1} y_i\boldsymbol{h}^{[i]} = \boldsymbol{0}.$$

Then we have

$$y_0\boldsymbol{g}^{[n-k]} + \sum_{i=0}^{n-k-1} x_i\boldsymbol{g}^{[i]} = -x_0\gamma\boldsymbol{h} - y_0\gamma^{[1]}\boldsymbol{h}^{[n-k]} - \sum_{i=1}^{n-k-1} y_i\boldsymbol{h}^{[i]}.$$

Apparently $y_0\boldsymbol{g}^{[n-k]} + \sum_{i=0}^{n-k-1} x_i\boldsymbol{g}^{[i]} \in \mathcal{G}_{n,n-k+2}(\boldsymbol{g})$ and $-x_0\gamma\boldsymbol{h} - y_0\gamma^{[1]}\boldsymbol{h}^{[n-k]} - \sum_{i=1}^{n-k-1} y_i\boldsymbol{h}^{[i]} \in \mathcal{G}_{n,n-k+2}(\boldsymbol{h})$. Hence $x_i = y_i = 0\,(0 \leqslant i \leqslant n-k-1)$ because of Assumption (1).

By Proposition 2, we have that $\dim(\langle AP^T \rangle + \langle AP^T \rangle^{[1]}) = 2l$ holds with extremely high probability. Together with Proposition 4, we have that $\dim(\mathcal{C}_{pub}^{\perp} + {\mathcal{C}_{pub}^{\perp}}^{[1]}) = n = \min\{2(n-k+l), n\}$. This means that by computing the intersection (3) the adversary can obtain nothing but the whole space and hence the Coggia-Couvreur attack will fail in this situation.

6.2 Analysis of Modification II

Since $\mathrm{Clr}_q(M) = l$, there must be $1 \leqslant \mathrm{Rank}(M) \leqslant l$. Assume that $\mathrm{Rank}(M) = l'$, apparently we have $\dim(\langle M \rangle) = l' \leqslant l$. By Proposition 1, we have $w_R(\boldsymbol{v}) \leqslant l$ for any $\boldsymbol{v} \in \langle M \rangle$. Together with $d(\mathcal{G}) = n - k + 1 \gg l$, we have $\langle M \rangle \cap \mathcal{G} = \{\mathbf{0}\}$.

Let $\mathcal{C}_{pub} = \langle G_{pub} \rangle = \langle S G_M P^{-1} \rangle$, then a parity-check matrix for $\mathcal{C}_{pub}$ can be written as $H_{pub} = H_M P^T$, where H_M is an $(n-k) \times n$ full-rank matrix such that $S G_M H_M^T = O$. It is easy to see that $\langle H_M \rangle$ contains a subcode of $\mathcal{G}^\perp$ of dimension $n - k - l'$. Hence $\mathcal{C}_{pub}^\perp$ contains a subcode of $\mathcal{C}_1$ of dimension $n - k - l'$, where $\mathcal{C}_1$ is spanned by

$$\boldsymbol{g} + \gamma \boldsymbol{h}, \boldsymbol{g}^{[1]} + \gamma \boldsymbol{h}^{[1]}, \cdots, \boldsymbol{g}^{[r]} + \gamma \boldsymbol{h}^{[r]}, \text{ where } r = n - k - 1.$$

Similarly ${\mathcal{C}_{pub}^\perp}^{[1]}$ contains a subcode of $\mathcal{C}_2$ of dimension $n - k - l'$, where $\mathcal{C}_2$ is spanned by

$$\boldsymbol{g}^{[1]} + \gamma^{[1]} \boldsymbol{h}^{[1]}, \boldsymbol{g}^{[2]} + \gamma^{[1]} \boldsymbol{h}^{[2]}, \cdots, \boldsymbol{g}^{[r+1]} + \gamma^{[1]} \boldsymbol{h}^{[r+1]}.$$

Finally we have that $\mathcal{C}_{pub}^\perp + {\mathcal{C}_{pub}^\perp}^{[1]}$ contains a subcode of $\mathcal{C} = \mathcal{C}_1 + \mathcal{C}_2$ of dimension at most $2(n - k - l')$, where $\mathcal{C}$ is spanned by

$$\boldsymbol{g} + \gamma \boldsymbol{h} \text{ and } \boldsymbol{g}^{[1]}, \boldsymbol{h}^{[1]}, \cdots, \boldsymbol{g}^{[r]}, \boldsymbol{h}^{[r]} \text{ and } \boldsymbol{g}^{[r+1]} + \gamma^{[1]} \boldsymbol{h}^{[r+1]}.$$

In the Coggia-Couvreur attack, the adversary can obtain (4) by computing (3). Our analysis shows that the adversary cannot perform the same operation on Modification II to obtain (4). Here we demonstrate this point with the method of reduction to absurdity.

Suppose that

$$\langle \boldsymbol{g}^{[r]} + \gamma^{[r]} \boldsymbol{h}^{[r]}, \boldsymbol{g}^{[r+1]} + \gamma^{[1]} \boldsymbol{h}^{[r+1]} \rangle \subseteq \bigcap_{i=0}^{r} ({\mathcal{C}_{pub}^\perp}^{[i]} + {\mathcal{C}_{pub}^\perp}^{[i+1]}). \tag{7}$$

Then for any $0 \leqslant i \leqslant r$, we have

$$\boldsymbol{g}^{[r]} + \gamma^{[r]} \boldsymbol{h}^{[r]}, \boldsymbol{g}^{[r+1]} + \gamma^{[1]} \boldsymbol{h}^{[r+1]} \in {\mathcal{C}_{pub}^\perp}^{[i]} + {\mathcal{C}_{pub}^\perp}^{[i+1]}. \tag{8}$$

Applying the inverse of the i-th Frobenius map to both sides of (8), there will be

$$\boldsymbol{g}^{[r-i]} + \gamma^{[r-i]} \boldsymbol{h}^{[r-i]}, \boldsymbol{g}^{[r-i+1]} + \gamma^{[1-i]} \boldsymbol{h}^{[r-i+1]} \in \mathcal{C}_{pub}^\perp + {\mathcal{C}_{pub}^\perp}^{[1]},$$

or equivalently

$$\boldsymbol{g} + \gamma \boldsymbol{h} \text{ and } \boldsymbol{g}^{[1]}, \boldsymbol{h}^{[1]}, \cdots, \boldsymbol{g}^{[r]}, \boldsymbol{h}^{[r]} \text{ and } \boldsymbol{g}^{[r+1]} + \gamma^{[1]} \boldsymbol{h}^{[r+1]} \in \mathcal{C}_{pub}^\perp + {\mathcal{C}_{pub}^\perp}^{[1]}.$$

This implies that $\mathcal{C} \subseteq \mathcal{C}_{pub}^\perp + {\mathcal{C}_{pub}^\perp}^{[1]}$, which conflicts with the previous conclusion that $\mathcal{C}_{pub}^\perp + {\mathcal{C}_{pub}^\perp}^{[1]}$ contains a subcode of $\mathcal{C}$ of dimension at most $2(n - k - l')$.

Hence the assumption (7) cannot be true and the adversary cannot recover (4) from (3) as the Coggia-Couvreur attack on Loidreau's cryptosystem. Therefore the Coggia-Couvreur attack does not work on Modification II.

7 Parameters and key size

In Table 1 we give some parameters suggested for different security levels, and make a comparison on public-key size with Loidreau's original scheme in the case of $k \leqslant \frac{n}{2}$. When considering the parameters, we exploit the complexity assessment of generic attacks described in Section 6.

In Modification I, the public key is a systematic generator matrix of an $[n, k-l]$ rank metric code, resulting in a public-key size of $(k-l)(n-k+l) \cdot m \cdot \log_2(q)$ bits. In Modification II, the public key is a systematic generator matrix of an $[n, k]$ rank metric code, resulting in a public-key size of $k(n-k) \cdot m \cdot \log_2(q)$ bits. For the concrete instances, we consider the case where $q = 3$ and $\lambda = 2$. It is not difficult to see from Table 1 that our modifications have obvious advantage over Loidreau's original scheme in public-key representation.

Instance	Parameters	Public-key Size	Security
Loidreau's system	m=50, n=50, k=25	12383	128
	m=60, n=60, k=27	19258	192
	m=75, n=75, k=34	37891	256
Modification I	m=50, n=50, k=26, l=1	6192	128
	m=60, n=60, k=28, l=1	10592	192
	m=75, n=75, k=35, l=1	20714	256
Modification II	m=52, n=52, k=26, l=1	6965	128
	m=62, n=62, k=28, l=1	11694	192
	m=77, n=77, k=35, l=1	22426	256

Table 1: Comparison on public-key sizes (in bytes) with Loidreau's original scheme for different security levels.

Instance	128 bits	192 bits	256 bits
HQC	2249	4522	7245
BIKE	1540	3082	5121
Classic McEliece	261120	524160	1044992
NTS-KEM	319488	929760	1419704
Modification I	6192	10592	20714
Modification II	6965	11694	22426

Table 2: Comparison on public-key sizes (in bytes) with some other cryptosystems.

In Table 2, we make a comparison on public-key sizes with some other code-based cryptosystems that have been selected as the third round candidates of the NIST PQC Standardization Process. These candidates are HQC [41], BIKE [42], NTS-KEM [43] and Classic McEliece [44]. Note that the Classic McEliece published in the third round of the NIST PQC project is a merged version of NTS-KEM and the original Classic McEliece for their specifications being very similar.

From Table 2 we can see that our modifications behave pretty well without using codes endowed with special algebraic structures.

8 Conclusion

In this paper, we propose two modifications for Loidreau's cryptosystem. According to our analysis, both of these two modifications can resist the existing structural attacks designed for cryptosystems based on Gabidulin codes, including Overbeck's attack and the Coggia-Couvreur attack. In our modifications, we adopt a systematic generator matrix of the public code to reduce the public-key size. Note that this method of describing the public code may reveal some information about the plaintext because of the sparsity of the intended errors in the Hamming metric [45], which means a security flaw to the cryptosystem. In the rank metric, however, the intended errors may happen in all coordinates of the error vector with high probability. Particularly, if we generate the error vector by randomly and uniformly choosing n elements from an $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^m}$ of dimension t, then the expected Hamming weight of the subvector of length k is $k(1-\frac{1}{q^t}) \sim k$, while in Hamming metric this value is kt/n. Therefore there is no need to worry about this problem in our modifications.

Acknowledgements We are grateful to Terry Lau Shue Chien for his valuable suggestion that we should consider the algebraic attacks described in Section 6 and rechoose the paremeter sets. 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. 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).
3. 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).
4. Kobara, K., Imai, H.: Semantically secure McEliece public-key cryptosystems-conversions for McEliece PKC. Kim, K. (Ed.): Proceedings of PKC 2001, LNCS, vol. 1992, pp. 19–35. Springer (2001).
5. Faugère, J.-C., Otmani, A., Perret, L., Portzamparc, F.de., Tillich, J.-P.: Structural cryptanalysis of McEliece schemes with compact keys. Des. Codes Cryptogr. 79(1), 87–112 (2016).
6. Loidreau, P., Sendrier, N.: Weak keys in the McEliece public-key cryptosystem. IEEE Trans. Inf. Theory 47(3), 1207–1211 (2001).
7. Niederreiter, H.: Knapsack-type cryptosystems and algebraic coding theory. Prob. Control and Inf. Theory 15(2), 159–166 (1986).
8. Sidelnikov, V.M., Shestakov, S.O.: On insecurity of cryptosystems based on generalized Reed-Solomon codes. Discret. Math. Appl. 2(4), 439–444 (1992).
9. 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).
10. 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).

11. 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).
12. Gabidulin, E.M.: Theory of codes with maximum rank distance. Prob. Peredachi Inf. 21(1), 3–16 (1985).
13. Loidreau, P.: A Welch-Berlekamp like algorithm for decoding Gabidulin codes. In: Ytrehus, ∅. (Ed.): Proceedings of WCC 2005, LNCS, vol. 3969, pp. 36–45. Springer (2005).
14. Richter, G., Plass, S.: Error and erasure decoding of rank-codes with a modified Berlekamp-Massey algorithm. ITG FACHBERICHT, pp. 203–210 (2004).
15. 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).
16. 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).
17. 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).
18. Rashwan, H., Gabidulin, E.M., Honary, B.: Security of the GPT cryptosystem and its applications to cryptography. Secur. Commun. Netw. 4(8), 937–946 (2011).
19. Gabidulin, E.M.: Attacks and counter-attacks on the GPT public key cryptosystem. Des. Codes Cryptogr. 48(2), 171–177 (2008).
20. 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).
21. 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).
22. 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).
23. 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).
24. Overbeck, R.: Structural attacks for public key cryptosystems based on Gabidulin codes. J. Cryptology 21(2), 280–301 (2008).
25. 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).
26. Faure, C., Loidreau, P.: A new public-key cryptosystem based on the problem of reconstructing p-polynomials. In: Ytrehus, ∅. (Ed.): Proceedings of WCC 2005, LNCS, vol. 3969, pp. 304–315. Springer (2005).
27. 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).
28. 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).
29. 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).
30. Coggia, D., Couvreur, A.: On the security of a Loidreau rank metric code based encryption scheme. Des. Codes Cryptogr. 88(9), 1941–1957 (2020).
31. Ghatak, A.: Extending Coggia-Couvreur attack on Loidreau's rank-metric cryptosystem. arXiv:2007.07354 [cs.IT] (2020).
32. Ling, S., Xing, C.: Coding theory: A First Course. Cambridge University Press (2004).
33. Gabidulin, E.M., Ourivski, A.V., Honary, B., Ammar, B.: Reducible rank codes and their applications to cryptography. IEEE Trans. Inf. Theory 49(12), 3289–3293 (2003).
34. Horlemann-Trautmann, A.-L., Marshall, K.: New criteria for MRD and Gabidulin codes and some rank-metric code constructions. arXiv:1507.08641 [cs.IT] (2015).
35. Horlemann-Trautmann, A.-L., Marshall, K., Rosenthal, J.: Considerations for rank-based cryptosystems. In: Proceedings of 2016 IEEE International Symposium on Information Theory, pp. 2544–2548. IEEE (2016).
36. 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).
37. 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).

38. 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, pp. 2421–2425. IEEE (2018).
39. Bardet, M., Bros, M., Cabarcas, D., et al.: Improvements of algebraic attacks for solving the rank decoding and MinRank problems. In: Proceedings of ASIACRYPT 2020, LNCS, vol. 12491, pp. 507–536. IACR (2020).
40. Bardet, M., Briaud, P., Bros, M., et al.: An algebraic attack on rank metric code-based cryptosystems. In: Proceedings of EUROCRYPT 2020, LNCS, vol. 12107, pp. 64–93. IACR (2020).
41. Melchor, C.A., Aragon, N., et al.: Hamming quasi-cyclic (HQC). http://pqc-hqc.org/doc/hqc-specification_2020-10-01.pdf. Accessed October 10, 2020.
42. 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. Accessed October 10, 2020.
43. Albrecht, M., Cid, C., Paterson, K.G., et al.: NTS-KEM. https://drive.google.com/file/d/1N3rv4HKCt9yU4xn6wuepsBUrfQW8cuFy/view. Accessed November 29, 2019.
44. 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.
45. Canteaut, A., Chabaud, F.: Improvements of the attacks on cryptosystems based on error-correcting codes. Rapport interne du Departement Mathematiques et Informatique, LIENS-95-21 (1995).