Cryptanalysis and Repair of a Gabidulin Code Based Cryptosystem from ACISP 2018

Wenshuo Guo(✉) and Fang-Wei Fu

Chern Institute of Mathematics and LPMC, Nankai University, Tianjin, China
ws_guo@mail.nankai.edu.cn, fwfu@nankai.edu.cn

Abstract. This paper presents a key recovery attack on a rank metric based cryptosystem proposed by Lau and Tan at ACISP 2018, which uses Gabidulin codes as the underlying decodable code. This attack is shown to cost polynomial time and therefore completely breaks the cryptosystem. Specifically, we convert the problem of recovering the private key into solving a multivariate linear system over the base field. We then present a simple repair for this scheme, which is shown to require exponential complexity for the proposed attack. Additionally, we apply this attack to cryptanalyze another Gabidulin code based cryptosystem proposed by Loidreau at PQCrypto 2017, and improve Loidreau's result in a talk at CBCrypto 2021.

Keywords: Post-quantum cryptography · Code-based cryptography · Gabidulin codes · Key recovery attack

1 Introduction

In post-quantum era, public key cryptosystems based on number theoretic problems will suffer serious security threat due to Shor's algorithm [37]. To prevent attacks from quantum computers, people have paid much attention to seeking alternatives for future use. Among these alternatives, code-based cryptography is one of the most promising candidates, whose security depends on the NP-completeness of decoding general linear codes [8]. The first cryptosystem of this type was proposed by McEliece [30] in 1978 using Goppa codes as the underlying linear code, which is now known as the McEliece cryptosystem. Although this scheme remains secure, it has never been used in practical situations due to the drawback of large key size. To tackle this problem, various improvements have been proposed one after another. In general, these variants can be divided into

This research was 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).

K. Nguyen et al. (Eds.): ACISP 2022, LNCS 13494, pp. 189–205, 2022.
https://doi.org/10.1007/978-3-031-22301-3_10

two categories: one is to replace Goppa codes with other Hamming metric codes [1,4,23,31], the other is to use codes endowed with other metric [2,22].

In 1991, Gabidulin et al. [15] proposed an encryption scheme based on rank metric codes, namely the GPT cryptosystem based on Gabidulin codes. The greatest advantage of rank metric based cryptosystems consists in their compact representation of public keys. Some representative variants based on Gabidulin codes can be found in [7,12,14,24,27,35]. Unfortunately, most of these variants have been completely or partially broken due to the inherent structural weakness of Gabidulin codes [9,11,16,20,32,34].

In [25], Lau and Tan proposed a public key cryptosystem based on Gabidulin codes, which was later published in [26] with an extended version. In this proposal, the public key consists of two parts, namely a generator matrix of the disturbed Gabidulin code by a random code that has maximum rank weight n and a vector of rank weight n. This technique of masking the structure of Gabidulin codes, as claimed by Lau and Tan, can prevent some existing attacks such as Frobenius weak attack [19], reduction attack [32], and Overbeck's attack [34]. Additionally, the recent Coggia-Couvreur attack [11] and Ghatak's attack [18] designed for Loidreau's cryptosystem [27] do not work on this scheme either.

Our Contributions. Firstly, we show that all the generating vectors of a Gabidulin code, together with the zero vector, form a 1-dimensional linear space. In other words, for a fixed generating vector $\boldsymbol{g}$ of a Gabidulin code $\mathcal{G} \subseteq \mathbb{F}_{q^m}^n$, any other generating vector must be of the form $\gamma \boldsymbol{g}$ for some $\gamma \in \mathbb{F}_{q^m}^*$. This suggests that there are totally $q^m - 1$ generating vectors for a Gabidulin code over $\mathbb{F}_{q^m}$. Secondly, we introduce a different approach from the one in [20] to compute the generating vector of Gabidulin codes from an arbitrary generator matrix. Thirdly, this paper presents a simple yet efficient key recovery attack on the Lau-Tan cryptosystem. Fourthly, we give a simple but effective repair for this system, which is shown to be secure against the existing structural attacks and have larger information transfer rate. Lastly, when applying this attack to analyze Loidreau's cryptosystem, we get a reduction in the complexity of recovering an equivalent private key.

The rest of this paper is organized as follows. Section 2 introduces basic notions used throughout this paper, as well as the concept of Moore matrices and Gabidulin codes. Section 3 gives a simple description of the Lau-Tan cryptosystem. Section 4 mainly describes the principle of our attack. Specifically, we first present some further results about Gabidulin codes that will be helpful for explaining why our attack works. Then a detailed description of this attack will be given in two steps. Lastly, we give a complexity analysis of this attack and some experimental results. In Sect. 5, we propose a modification for this scheme, investigate its security and give some practical parameters. In Sect. 6, we apply this attack to cryptanalyze Loidreau's cryptosystem. Section 7 concludes this paper.

2 Preliminaries

In this section, we first introduce some notations and basic concepts in coding theory. After that, we recall the concept of Gabidulin codes, and present some related results in the meanwhile.

2.1 Notations and Basic Concepts

For a prime power q, we 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. Note that $\mathbb{F}_{q^m}$ can be seen as a linear space over $\mathbb{F}_q$ of dimension m. A vector $\boldsymbol{a} \in \mathbb{F}_{q^m}^m$ is called a basis vector if the components of $\boldsymbol{a}$ form a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. Particularly, we call $\boldsymbol{a}$ a normal basis vector if $\boldsymbol{a}$ has the form $(\alpha^{q^{m-1}}, \alpha^{q^{m-2}}, \ldots, \alpha)$ for some $\alpha \in \mathbb{F}_{q^m}^* = \mathbb{F}_{q^m} \backslash \{0\}$. For two positive integers k and n, denote by $\mathcal{M}_{k,n}(\mathbb{F}_q)$ the space of all $k \times n$ matrices over $\mathbb{F}_q$, and by $\mathrm{GL}_n(\mathbb{F}_q)$ the set of all invertible matrices in $\mathcal{M}_{n,n}(\mathbb{F}_q)$. For a matrix $M \in \mathcal{M}_{k,n}(\mathbb{F}_q)$, denote by $\langle M \rangle_q$ the linear space spanned by the rows of M over $\mathbb{F}_q$.

An $[n, k]$ linear code $\mathcal{C}$ over $\mathbb{F}_{q^m}$ is a k-dimensional subspace of $\mathbb{F}_{q^m}^n$, and any element in $\mathcal{C}$ is called a codeword of $\mathcal{C}$. The dual code of $\mathcal{C}$, denoted by $\mathcal{C}^\perp$, is the orthogonal space of $\mathcal{C}$ under the usual inner product over $\mathbb{F}_{q^m}^n$. A $k \times n$ matrix G is called a generator matrix of $\mathcal{C}$ if its row vectors form a basis of $\mathcal{C}$ over $\mathbb{F}_{q^m}$. A generator matrix H of $\mathcal{C}^\perp$ is called a parity-check matrix of $\mathcal{C}$. For a codeword $\boldsymbol{c} \in \mathcal{C}$, the rank support of $\boldsymbol{c}$, denoted by $\mathrm{Supp}(\boldsymbol{c})$, is the linear space spanned by the components of $\boldsymbol{c}$ over $\mathbb{F}_q$. The rank weight of $\boldsymbol{c}$ with respect to $\mathbb{F}_q$, denoted by $\mathrm{rk}(\boldsymbol{c})$, is defined to be the dimension of $\mathrm{Supp}(\boldsymbol{c})$ over $\mathbb{F}_q$. The minimum rank distance of $\mathcal{C}$, denoted by $\mathrm{rk}(\mathcal{C})$, is defined to be the minimum rank weight of all nonzero codewords in $\mathcal{C}$. For a matrix $M \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$, the rank support of M, denoted by $\mathrm{Supp}(M)$, is defined to be the linear space spanned by entries of M over $\mathbb{F}_q$. The rank weight of M with respect to $\mathbb{F}_q$, denoted by $\mathrm{rk}(M)$, is defined as the dimension of $\mathrm{Supp}(M)$ over $\mathbb{F}_q$.

2.2 Gabidulin Codes

This section recalls the concept of Gabidulin codes. Before doing this, we first introduce the definition of Moore matrices and some related results.

Definition 1 (Moore matrices). *For an integer i and $\alpha \in \mathbb{F}_{q^m}$, we define $\alpha^{[i]} = \alpha^{q^i}$ to be the i-th Frobenius power of α. For a vector $\boldsymbol{a} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}_{q^m}^n$, we define $\boldsymbol{a}^{[i]} = (\alpha_1^{[i]}, \alpha_2^{[i]}, \ldots, \alpha_n^{[i]})$ to be the i-th Frobenius power of $\boldsymbol{a}$. For positive integers $k \leqslant n$, a $k \times n$ Moore matrix generated by $\boldsymbol{a}$ is defined as*

$$\mathrm{Mr}_k(\boldsymbol{a}) = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^{[1]} & \alpha_2^{[1]} & \cdots & \alpha_n^{[1]} \\ \vdots & \vdots & & \vdots \\ \alpha_1^{[k-1]} & \alpha_2^{[k-1]} & \cdots & \alpha_n^{[k-1]} \end{pmatrix}.$$

For a positive integer l and a matrix $M = (M_{ij}) \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$, we denote by $M^{[l]} = (M_{ij}^{[l]})$ the l-th Frobenius power of M. For a set $\mathcal{V} \subseteq \mathbb{F}_{q^m}^n$, we denote by $\mathcal{V}^{[l]} = \{\boldsymbol{v}^{[l]} : \boldsymbol{v} \in \mathcal{V}\}$ the l-th Frobenius power of $\mathcal{V}$. Particularly, for a linear code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$, it is easy to verify that $\mathcal{C}^{[l]}$ is also a linear code over $\mathbb{F}_{q^m}$.

The following proposition presents simple properties of Moore matrices.

Proposition 1. *(1) For two $k \times n$ Moore matrices $A, B \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$, the sum $A + B$ is also a $k \times n$ Moore matrix.*

(2) For a Moore matrix $M \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ and a matrix $Q \in \mathcal{M}_{n,l}(\mathbb{F}_q)$, the product MQ forms a $k \times l$ Moore matrix.

(3) For a vector $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $\mathrm{rk}(\boldsymbol{a}) = l$, there exist $\boldsymbol{a}' \in \mathbb{F}_{q^m}^l$ with $\mathrm{rk}(\boldsymbol{a}') = l$ and $Q \in \mathrm{GL}_n(\mathbb{F}_q)$ such that $\boldsymbol{a} = (\boldsymbol{a}'||\boldsymbol{0})Q$. Furthermore, let $A = \mathrm{Mr}_k(\boldsymbol{a})$ and $A' = \mathrm{Mr}_k(\boldsymbol{a}')$, then $A = [A'|0]Q$.

(4) For positive integers $k \leqslant n \leqslant m$, let $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ be a vector such that $\mathrm{rk}(\boldsymbol{a}) = n$, then the Moore matrix $\mathrm{Mr}_k(\boldsymbol{a})$ has rank k.

Proof. Statements (1), (2) and (3) are trivial and the proof is omitted here.

(4) Let $\boldsymbol{a} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_{q^m}^n$. If $\mathrm{Rank}(\mathrm{Mr}_k(\boldsymbol{a})) < k$, then there exists $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_{k-1}) \in \mathbb{F}_{q^m}^k \backslash \{\boldsymbol{0}\}$ such that $\boldsymbol{\lambda}\mathrm{Mr}_k(\boldsymbol{a}) = \boldsymbol{0}$. Let $f(x) = \sum_{j=0}^{k-1} \lambda_j x^{[j]} \in \mathbb{F}_{q^m}[x]$, then $f(\alpha_i) = 0$ holds for any $1 \leqslant i \leqslant n$. It follows that $f(\alpha) = 0$ for any $\alpha \in \langle \alpha_1, \ldots, \alpha_n \rangle_q$, which conflicts with the fact that $f(x) = 0$ admits at most q^{k-1} roots.

In particular, we have the following proposition, which was once exploited by Loidreau in [28] to cryptanalyze an encryption scheme [27] based on Gabidulin codes.

Proposition 2 (Moore matrix decomposition). *Let $\boldsymbol{a}$ be a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. For a positive integer $k \leqslant m$, let $M = \mathrm{Mr}_k(\boldsymbol{a})$ be a Moore matrix generated by $\boldsymbol{a}$. Then for any $k \times n$ Moore matrix $M' \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$, there exists $Q \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ such that $M' = MQ$.*

Now we formally introduce the definition of Gabidulin codes.

Definition 2 (Gabidulin codes). *For positive integers $k < n \leqslant m$, let $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{a}) = n$. The $[n, k]$ Gabidulin code generated by $\boldsymbol{a}$, denoted by $\mathrm{Gab}_{n,k}(\boldsymbol{a})$, is defined as the linear space spanned by the rows of $\mathrm{Mr}_k(\boldsymbol{a})$ over $\mathbb{F}_{q^m}$. $\mathrm{Mr}_k(\boldsymbol{a})$ is called a canonical generator matrix of $\mathrm{Gab}_{n,k}(\boldsymbol{a})$, and $\boldsymbol{a}$ a generating vector respectively.*

Remark 1. Gabidulin codes can be seen as a rank metric counterpart of generalized Reed-Solomon (GRS) codes, both of which admit good algebraic properties. The dual of an $[n, k]$ Gabidulin code is an $[n, n-k]$ Gabidulin code [16]. An $[n, k]$ Gabidulin code has minimum rank distance $n-k+1$ [21] and can therefore correct up to $\left\lfloor \frac{n-k}{2} \right\rfloor$ rank errors in theory. Efficient decoding algorithms for Gabidulin codes can be found in [13,29,36].

To reduce the public key size, Lau and Tan exploited the so-called partial circulant matrix in the cryptosystem, which is defined as follows.

Definition 3 (Partial circulant matrices). *For* $\boldsymbol{a} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}_{q^m}^n$, *the circulant matrix generated by* $\boldsymbol{a}$, *denoted by* $\mathrm{Cir}_n(\boldsymbol{a})$, *is defined to be a matrix whose first row is* $\boldsymbol{a}$ *and* i*-th row is obtained by cyclically right shifting the* $i-1$*-th row for* $2 \leqslant i \leqslant n$. *The* $k \times n$ *partial circulant matrix generated by* $\boldsymbol{a}$, *denoted by* $\mathrm{Cir}_k(\boldsymbol{a})$, *is defined to be the first* k *rows of* $\mathrm{Cir}_n(\boldsymbol{a})$.

Remark 2. Let $\boldsymbol{a}$ be a normal basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, then it is easy to verify that the $k \times m$ partial circulant matrix generated by $\boldsymbol{a}$ is exactly the $k \times m$ Moore matrix generated by $\boldsymbol{a}$. In other words, we have $\mathrm{Cir}_k(\boldsymbol{a}) = \mathrm{Mr}_k(\boldsymbol{a})$.

3 Lau-Tan Cryptosystem

In this section, we mainly give a simple description of the Lau-Tan cryptosystem that uses Gabidulin codes as the underlying decodable code. For a given security level, choose positive integers $m > n > k > k'$ and r such that $k' = \lfloor \frac{k}{2} \rfloor$ and $r = \lfloor \frac{n-k}{2} \rfloor$. The Lau-Tan cryptosystem consists of the following three algorithms.

- Key Generation

  Let $\mathcal{G}$ be an $[n, k]$ Gabidulin code over $\mathbb{F}_{q^m}$, and $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ be a generator matrix of $\mathcal{G}$ of canonical form. Randomly choose matrices $S \in \mathrm{GL}_k(\mathbb{F}_{q^m})$ and $T \in \mathrm{GL}_n(\mathbb{F}_q)$. Randomly choose $\boldsymbol{u} \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{u}) = n$ and set $U = \mathrm{Cir}_k(\boldsymbol{u})$. Let $G_{pub} = SG + UT$, then we publish $(G_{pub}, \boldsymbol{u})$ as the public key, and keep (S, G, T) as the private key.
- Encryption

  For a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^{k'}$, randomly choose a vector $\boldsymbol{m}_s \in \mathbb{F}_{q^m}^{k-k'}$ such that $\mathrm{rk}((\boldsymbol{m}||\boldsymbol{m}_s)U) > \lceil \frac{3}{4}(n-k) \rceil$. Randomly choose $\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{e}_1) \leqslant \frac{r}{2}$ and $\mathrm{rk}(\boldsymbol{e}_2) \leqslant \frac{r}{2}$. Compute $\boldsymbol{c}_1 = (\boldsymbol{m}||\boldsymbol{m}_s)U + \boldsymbol{e}_1$ and $\boldsymbol{c}_2 = (\boldsymbol{m}||\boldsymbol{m}_s)G_{pub} + \boldsymbol{e}_2$. Then the ciphertext is $\boldsymbol{c} = (\boldsymbol{c}_1, \boldsymbol{c}_2)$.
- Decryption

  For a ciphertext $\boldsymbol{c} = (\boldsymbol{c}_1, \boldsymbol{c}_2) \in \mathbb{F}_{q^m}^{2n}$, compute $\boldsymbol{c}' = \boldsymbol{c}_2 - \boldsymbol{c}_1 T = (\boldsymbol{m}||\boldsymbol{m}_s)SG + \boldsymbol{e}_2 - \boldsymbol{e}_1 T$. Note that $\mathrm{rk}(\boldsymbol{e}_2 - \boldsymbol{e}_1 T) \leqslant \mathrm{rk}(\boldsymbol{e}_2) + \mathrm{rk}(\boldsymbol{e}_1 T) \leqslant r$, decoding $\boldsymbol{c}'$ with the fast decoder of $\mathcal{G}$ will lead to $\boldsymbol{m}' = (\boldsymbol{m}||\boldsymbol{m}_s)S$, then by computing $\boldsymbol{m}' S^{-1}$ one can recover the plaintext $\boldsymbol{m}$.

4 Key Recovery Attack

This section discusses how to efficiently recover an equivalent private key of the Lau-Tan cryptosystem. We point out that the knowledge of T is of great importance for the security of the whole cryptosystem. Specifically, if one can find the private T, then one is able to recover everything needed to decrypt an arbitrary ciphertext in polynomial time. Before describing this attack, we first introduce some further results about Gabidulin codes.

4.1 Further Results About Gabidulin Codes

Similar to GRS codes in the Hamming metric, Gabidulin codes also have good algebraic structure. For instance, if $\mathcal{G}$ is a Gabidulin code over $\mathbb{F}_{q^m}$, then its l-th Frobenius power is also a Gabidulin code. Formally, we introduce the following proposition.

Proposition 3. *Let $\mathcal{G}$ be an $[n,k]$ Gabidulin code over $\mathbb{F}_{q^m}$, with $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ as a generator matrix. For any positive integer l, $\mathcal{G}^{[l]}$ is also an $[n,k]$ Gabidulin code and has $G^{[l]}$ as a generator matrix.*

Proof. Trivial from a straightforward verification.

For a proper positive integer l, the intersection of a Gabidulin code and its l-th Frobenius power is still a Gabidulin code, as described in the following proposition.

Proposition 4. *For an $[n,k]$ Gabidulin code $\mathcal{G}$ over $\mathbb{F}_{q^m}$, let $\boldsymbol{g} \in \mathbb{F}_{q^m}^n$ be a generating vector of $\mathcal{G}$. For a positive integer $l \leqslant \min\{k-1, n-k\}$, the intersection of $\mathcal{G}$ and its l-th Frobenius power is an $[n, k-l]$ Gabidulin code with $\boldsymbol{g}^{[l]}$ as a generating vector. In other words, we have the following equality*

$$\mathcal{G} \cap \mathcal{G}^{[l]} = \mathrm{Gab}_{n,k-l}(\boldsymbol{g}^{[l]}).$$

Proof. By Definition 2, $\mathcal{G}$ is an $\mathbb{F}_{q^m}$-span of $\boldsymbol{g}, \ldots, \boldsymbol{g}^{[k-1]}$, i.e. $\mathcal{G} = \langle \boldsymbol{g}, \ldots, \boldsymbol{g}^{[k-1]} \rangle_{q^m}$. By Proposition 3, we have $\mathcal{G}^{[l]} = \langle \boldsymbol{g}^{[l]}, \ldots, \boldsymbol{g}^{[k+l-1]} \rangle_{q^m}$. Note that $l \leqslant \min\{k-1, n-k\}$, then $k + l \leqslant n$ and $\boldsymbol{g}, \ldots, \boldsymbol{g}^{[k+l-1]}$ are linearly independent over $\mathbb{F}_{q^m}$. It follows that $\mathcal{G} \cap \mathcal{G}^{[l]} = \langle \boldsymbol{g}^{[l]}, \ldots, \boldsymbol{g}^{[k-1]} \rangle_{q^m}$ forms an $[n, k-l]$ Gabidulin code, having $\boldsymbol{g}^{[l]}$ as a generating vector. This completes the proof.

Proposition 5. *For positive integers $k < n \leqslant m$, let $\mathcal{G} \subset \mathbb{F}_{q^m}^n$ be an $[n,k]$ Gabidulin code, and $A \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ a nonzero Moore matrix. If all the row vectors of A are codewords in $\mathcal{G}$, then A forms a generator matrix of $\mathcal{G}$.*

Proof. It suffices to prove $\mathrm{Rank}(A) = k$. Suppose that A is generated by $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$, i.e. $A = \mathrm{Mr}_k(\boldsymbol{a})$. Let $l = \mathrm{rk}(\boldsymbol{a})$, then there exist $\boldsymbol{a}' \in \mathbb{F}_{q^m}^l$ with $\mathrm{rk}(\boldsymbol{a}') = l$ and $Q \in \mathrm{GL}_n(\mathbb{F}_q)$ such that $\boldsymbol{a} = (\boldsymbol{a}' || \boldsymbol{0})Q$. Let $A' \in \mathcal{M}_{k,l}(\mathbb{F}_{q^m})$ be a Moore matrix generated by $\boldsymbol{a}'$, then it follows immediately that $A = [A'|0]Q$. If $l > k$, then $\mathrm{Rank}(A) = \mathrm{Rank}(A') = k$ due to Proposition 1 and therefore the conclusion is proved. Otherwise, there will be $\langle A' \rangle_{q^m} = \mathbb{F}_{q^m}^l$. From this we can deduce that the minimum rank distance of $\mathcal{G}$ will be 1, which conflicts with the fact that $\mathrm{rk}(\mathcal{G}) = n - k + 1 \geqslant 2$. Hence $l > k$ and $\mathrm{Rank}(A) = k$. This completes the proof.

By Definition 2, a Gabidulin code is uniquely determined by its generating vector. Naturally, it is important to make clear what all these vectors look like and how many generating vectors there exist for a Gabidulin code.

Proposition 6. *Let $\mathcal{G}$ be an $[n,k]$ Gabidulin code over $\mathbb{F}_{q^m}$, with $\boldsymbol{g} \in \mathbb{F}_{q^m}^n$ as a generating vector. Let $\boldsymbol{g}' \in \mathbb{F}_{q^m}^n$ be a codeword in $\mathcal{G}$, then $\boldsymbol{g}'$ forms a generating vector if and only if there exists $\gamma \in \mathbb{F}_{q^m}^*$ such that $\boldsymbol{g}' = \gamma \boldsymbol{g}$.*

Proof. Assume that $\boldsymbol{g} = (\alpha_1, \ldots, \alpha_n)$ and $\boldsymbol{g}' = (\alpha'_1, \ldots, \alpha'_n)$, let $G = \mathrm{Mr}_k(\boldsymbol{g})$ and $G' = \mathrm{Mr}_k(\boldsymbol{g}')$. The conclusion is trivial if $\boldsymbol{g} = \boldsymbol{g}'$. Otherwise, without loss of generality we assume that $\alpha'_1 \neq \alpha_1$, then there exists $\gamma \in \mathbb{F}_{q^m}^* \backslash \{1\}$ such that $\alpha'_1 = \gamma\alpha_1$. Let

$$S = \begin{pmatrix} \gamma & 0 & \cdots & 0 \\ 0 & \gamma^{[1]} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \gamma^{[k-1]} \end{pmatrix},$$

then $SG = \mathrm{Mr}_k(\gamma\boldsymbol{g})$. Let $\boldsymbol{g}^* = \gamma\boldsymbol{g} - \boldsymbol{g}' = (0, \gamma\alpha_2 - \alpha'_2, \ldots, \gamma\alpha_n - \alpha'_n)$ and $G^* = \mathrm{Mr}_k(\boldsymbol{g}^*)$, then $G^* = SG - G'$. Apparently all the row vectors of G^* are codewords in $\mathcal{G}$. If $\boldsymbol{g}^* \neq \boldsymbol{0}$, then G^* forms a generator matrix of $\mathcal{G}$ of canonical form due to Proposition 5. Together with $\mathrm{rk}(\boldsymbol{g}^*) \leqslant n-1$, easily we can deduce that $\mathrm{rk}(\boldsymbol{c}) \leqslant n-1$ for any $\boldsymbol{c} \in \mathcal{G}$, which clearly contradicts the fact that $\mathrm{rk}(\boldsymbol{g}) = n$. Therefore there must be $\boldsymbol{g}^* = \boldsymbol{0}$, or equivalently $\boldsymbol{g}' = \gamma\boldsymbol{g}$. The opposite is obvious from a straightforward verification.

The following corollary is drawn immediately from Proposition 6.

Corollary 1. *An $[n, k]$ Gabidulin code over $\mathbb{F}_{q^m}$ admits $q^m - 1$ generator matrices of canonical form, or equivalently $q^m - 1$ generating vectors.*

Remark 3. Let $\mathcal{G} \subseteq \mathbb{F}_{q^m}^n$ be an $[n, k]$ Gabidulin code, and $M \in \mathcal{M}_{k,m}(\mathbb{F}_{q^m})$ a Moore matrix generated by a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. By Proposition 2, for any canonical generator matrix G, there exists a unique $Q \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ such that $G = MQ$. For a fixed M, there exist $q^m - 1$ Q's in $\mathcal{M}_{m,n}(\mathbb{F}_q)$ such that MQ forms a canonical generator matrix of $\mathcal{G}$. Furthermore, all these Q's together with the zero matrix form an $\mathbb{F}_q$-linear space of dimension m.

4.2 Recovering the Private T

This section mainly describes an efficient algorithm for recovering the private T. The technique we adopt here is to convert the problem of recovering T into solving a multivariate linear system, which clearly costs polynomial time. Before doing this, we first introduce the so-called subfield expanding transform.

Subfield Expanding Transform. For $\beta_1, \ldots, \beta_n \in \mathbb{F}_{q^m}$, we construct an equation as

$$\sum_{j=1}^{n} x_j \beta_j = 0, \tag{1}$$

where x_j's are underdetermined variables in $\mathbb{F}_q$. Let $\boldsymbol{a}$ be a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. For each $1 \leqslant j \leqslant n$, there exists $\boldsymbol{b}_j \in \mathbb{F}_q^m$ such that $\beta_j = \boldsymbol{b}_j \boldsymbol{a}^T$. It follows that $\sum_{j=1}^n x_j \beta_j = \sum_{j=1}^n x_j (\boldsymbol{b}_j \boldsymbol{a}^T) = (\sum_{j=1}^n x_j \boldsymbol{b}_j)\boldsymbol{a}^T$, and moreover, (1) holds if and only if

$$\sum_{j=1}^{n} x_j \boldsymbol{b}_j = \boldsymbol{0}. \tag{2}$$

Obviously, the linear systems (1) and (2) share the same solution space. A transform that derives (2) from (1) is called a subfield expanding transform (SET for short).

In the Lau-Tan cryptosystem, let $H \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$ be a parity-check matrix of $\mathcal{G}$ of canonical form. Let $M \in \mathcal{M}_{n-k,m}(\mathbb{F}_{q^m})$ be a Moore matrix generated by a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, then there exists an underdetermined matrix $X \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ such that $H = MX$. On the other hand, there exists another underdetermined matrix $T^* \in \mathrm{GL}_n(\mathbb{F}_q)$ such that $G_{pub} - UT^*$ forms a generator matrix of $\mathcal{G}$. This leads to a parity-check matrix equation as follows

$$(G_{pub} - UT^*)(MX)^T = G_{pub}X^TM^T - UT^*X^TM^T = 0. \tag{3}$$

We therefore obtain a system of $k(n-k)$ multivariate quadratic equations, with $n(m+n)$ variables in $\mathbb{F}_q$. This system admits at least q^m solutions. Specifically, we introduce the following proposition.

Proposition 7. *The linear system* (3) *has at least* q^m *solutions.*

Proof. If $T^* = T$, then we can deduce from (3) that

$$(G_{pub} - UT^*)(MX)^T = (SG + UT - UT^*)(MX)^T = SG(MX)^T = 0.$$

Note that $SG \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ forms a generator matrix of $\mathcal{G}$. By $SG(MX)^T = 0$, all the row vectors of MX are contained in $\mathcal{G}^\perp$, which is an $[n, n-k]$ Gabidulin code. On the other hand, it is clear that MX forms an $(n-k) \times n$ Moore matrix. By Proposition 5, MX forms a canonical generator matrix of $\mathcal{G}^\perp$ for a nonzero X. Then the conclusion is immediately proved from Corollary 1. Furthermore, we have that X is an $m \times n$ matrix of full rank.

Note that solving a multivariate quadratic system generally requires exponential time. Instead of solving the system (3) directly, the technique we exploit here is to consider each entry of T^*X^T as a new variable in $\mathbb{F}_q$ and set $Y = XT^{*T}$. In other words, we rewrite (3) into a matrix equation as follows

$$G_{pub}X^TM^T - UY^TM^T = 0. \tag{4}$$

This leads to a linear system of $k(n-k)$ equations, with coefficients in $\mathbb{F}_{q^m}$ and $2mn$ variables in $\mathbb{F}_q$. To solve the system (4), we usually convert this problem into an instance over the base field $\mathbb{F}_q$. Applying SET to (4) leads to a linear system of $mk(n-k)$ equations over $\mathbb{F}_q$, with $2mn$ variables to be determined. For cryptographic use, generally we have $mk(n-k) \geqslant 2mn$.

Remark 4. With each solution (X, T^*) of (3), one can obtain a solution of (4) by computing $Y = XT^{*T}$, which implies that (4) also has at least q^m solutions. Conversely, if (4) has exactly q^m solutions, then these solutions must correspond to those of (3) where $T^* = T$. In this situation, solving (4) for any nonzero solution (X, Y) enables us to recover the private T by solving the matrix equation $Y = XT^{*T}$.

As for whether or not the system (4) has other types of solutions, we make an **Assumption** that the answer is negative. According to our experimental results in MAGMA [10], this assumption holds with high probability. To make it easier, a simplified version of this problem is considered. Let G be an arbitrary generator matrix of an $[n,k]$ Gabidulin code and $\boldsymbol{u} \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{u}) = n$. We then construct a matrix equation as

$$GX^TM^T + \mathrm{Cir}_k(\boldsymbol{u})Y^TM^T = 0,$$

where $M \in \mathcal{M}_{n-k,m}(\mathbb{F}_{q^m})$ is a Moore matrix generated by a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ and $X, Y \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ are two underdetermined matrices. By applying SET to this system above, we obtain a new system over $\mathbb{F}_q$. By Remark 4, if this newly obtained system admits a solution space of dimension m, then there must be $Y = 0$. Finally, we ran 1000 random tests for $q = 2, m = 25, n = 23, k = 10$, and for $q = 3, m = 18, n = 15, k = 7$ respectively. It turns out that this assumption holds in all of these random instances.

Algorithm 1 : T-Recovering Algorithm

Input: (G_{pub}, U)
Output: T

1: Let $\boldsymbol{a}$ be a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ and set $M = \mathrm{Mr}_{n-k}(\boldsymbol{a})$
2: Let $X, Y \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ be two underdetermined matrices and set

$$G_{pub}X^TM^T - UY^TM^T = 0 \tag{5}$$

3: Apply SET to (5) to obtain a linear system over $\mathbb{F}_q$
4: Solve the system from Step 3 for any nonzero (X, Y)
5: Solve the matrix equation $Y = XT^{*T}$ for T^*
6: **return** $T = T^*$

4.3 Finding an Equivalent (S', G')

In Sect. 4.2, we have discussed how to efficiently recover T from (G_{pub}, U). With the knowledge of T, one can recover SG by computing $SG = G_{pub} - UT$, which forms a generator matrix of $\mathcal{G}$. To decrypt a ciphertext as the legitimate receiver does, one needs to recover a generator matrix G' of $\mathcal{G}$ of canonical form and an invertible matrix S' such that $S'G' = SG$, where (S', G') is called an equivalent form of (S, G). Once such a G' is obtained, then one can recover S' by solving a matrix equation.

Now we investigate how to derive a canonical generator matrix of a Gabidulin code, or equivalently a generating vector, from an arbitrary generator matrix. In [20] the authors presented an iterative method of computing the generating vector. Here in this paper we present a different approach to do this.

An Approach to Compute the Generating Vector. For an $[n,k]$ Gabidulin code $\mathcal{G}$ over $\mathbb{F}_{q^m}$, let $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ be an arbitrary generator matrix of $\mathcal{G}$. We

first compute a parity-check matrix of $\mathcal{G}$ from G, say H. Let $M \in \mathcal{M}_{k,m}(\mathbb{F}_{q^m})$ be a Moore matrix generated by a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, then there exists an underdetermined matrix $X \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ such that MX forms a canonical generator matrix of $\mathcal{G}$. By setting $(MX)H^T = 0$, we obtain a linear system of $k(n-k)$ equations, with coefficients in $\mathbb{F}_{q^m}$ and mn variables in $\mathbb{F}_q$. Applying SET to this system leads to a new linear system over the base field $\mathbb{F}_q$, with $mk(n-k)$ equations and mn variables. For cryptographic use, generally we have $mk(n-k) \geqslant mn$. By Corollary 1, this newly obtained system admits $q^m - 1$ nonzero solutions. And for any nonzero solution, say X, the first row of MX will be a generating vector of $\mathcal{G}$.

Algorithm 2 : (S', G')-Recovering Algorithm

Input: (G_{pub}, U, T)
Output: (S', G')

1: Let $\boldsymbol{a}$ be a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ and set $M = \mathrm{Mr}_k(\boldsymbol{a})$
2: Compute $SG = G_{pub} - UT$ and let $\mathcal{G} = \langle SG \rangle_{q^m}$
3: Let $H \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$ be a parity-check matrix of $\mathcal{G}$
4: Let $X \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ be an underdetermined matrix and set

$$(MX)H^T = 0 \tag{6}$$

5: Apply SET to (5) to obtain a linear system over $\mathbb{F}_q$
6: Solve the system from Step 5 for any nonzero X
7: Compute $G' = MX$
8: Compute $S' \in \mathrm{GL}_k(\mathbb{F}_{q^m})$ such that $S'G' = SG$
9: **return** (S', G')

4.4 Complexity of the Attack

Our attack consists of two phases: firstly, we manage to recover the private T from the published information, as described in Algorithm 1; secondly, with the knowledge of T and the public key, we compute a canonical generator matrix G' of the secret Gabidulin code and an invertible matrix S', as described in Algorithm 2. Hence the complexity analysis is done in the following two aspects.

Complexity of Algorithm 1. In Step 1 we construct a Moore matrix $M \in \mathcal{M}_{n-k,m}(\mathbb{F}_{q^m})$ whose first row forms a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. To avoid executing the Frobenius operation, here we choose $\boldsymbol{a}$ to be a normal basis vector, then we set $M = \mathrm{Cir}_{n-k}(\boldsymbol{a})$. In Step 2 we construct a multivariate linear system by performing matrix multiplication, requiring $\mathcal{O}(mn^3)$ operations in $\mathbb{F}_{q^m}$. The subfield expanding transform performed to (5) requires $\mathcal{O}(m^3n^3)$ operations in $\mathbb{F}_{q^m}$. Step 4 requires $\mathcal{O}(m^3n^3)$ operations to solve the linear system over $\mathbb{F}_q$ and Step 5 requires $\mathcal{O}(n^3)$ operations in $\mathbb{F}_q$. The total complexity of Algorithm 1 consists of $\mathcal{O}(m^3n^3 + mn^3)$ operations in $\mathbb{F}_{q^m}$ and $\mathcal{O}(m^3n^3 + n^3)$ operations in $\mathbb{F}_q$.

Complexity of Algorithm 2. In Step 1 we still choose a normal basis vector to construct M. To compute SG, we perform matrix addition and multiplication with $\mathcal{O}(n^3)$ operations in $\mathbb{F}_{q^m}$. Step 3 computes a parity-check H of $\mathcal{G}$ from SG, requiring $\mathcal{O}(n^3)$ operations in $\mathbb{F}_{q^m}$. Then we construct a linear system in Step 4, which costs $\mathcal{O}(mn^3)$ operations in $\mathbb{F}_{q^m}$. In Step 5 we apply SET to (6) to obtain a new system over $\mathbb{F}_q$, requiring $\mathcal{O}(m^3n^3)$ operations in $\mathbb{F}_{q^m}$. Solving this new system in Step 6 costs $\mathcal{O}(m^3n^3)$ operations in $\mathbb{F}_q$, and computing $G' = MX$ in Step 7 requires $\mathcal{O}(mn^2)$ operations in $\mathbb{F}_{q^m}$. In Step 8, we shall compute S' from $S'G'$ with $\mathcal{O}(n^3)$ operations. The total complexity of Algorithm 2 consists of $\mathcal{O}(m^3n^3 + mn^3 + n^3)$ operations in $\mathbb{F}_{q^m}$ and $\mathcal{O}(m^3n^3)$ operations in $\mathbb{F}_q$.

Finally, the total complexity of the attack is $\mathcal{O}(m^3n^3 + mn^3 + n^3)$ in $\mathbb{F}_{q^m}$ plus $\mathcal{O}(m^3n^3 + n^3)$ in $\mathbb{F}_q$.

4.5 Implementation

This attack has been implemented in MAGMA and permits to recover the private T. We tested this attack on a personal computer and succeeded for parameters as illustrated in Table 1. For each parameter set, this attack has been run 100 times and the last column gives the average timing (in seconds). Our implementation is just a proof of feasibility of this attack and does not consider the proposed parameters in [25,26].

Table 1. These tests were performed using MAGMA V2.11-1 on 11th Gen Intel$^{\text{R}}$ Core$^{\text{TM}}$ i7-11700 @ 2.5 GHz processor with 16 GB of memory.

q	m	n	k	t
2	22	18	9	8.6
2	28	22	9	40.7
2	35	26	12	173.2

5 A Repair

To prevent the proposed attack, we give a simple repair for the Lau-Tan cryptosystem in this section. Then we explain why this repair can resist the existing structural attacks, as well as the key recovery attack described in Sect. 4. After that, practical security of this repair against generic attacks is investigated. Following this, we suggest parameters for the security of at least 128 bits, 192 bits, and 256 bits. Public key sizes under these parameters are also given.

5.1 Description of the Repair

For a given security level, choose a field $\mathbb{F}_q$ and positive integers $m, n, k, \lambda, r_1, r_2$ such that $r = \lfloor \frac{n-k}{2} \rfloor$ and $r_1 + \lambda r_2 \leqslant r$. Our repair consists of the following three procedures.

– Key Generation
  Let $\mathcal{G}$ be an $[n,k]$ Gabidulin code over $\mathbb{F}_{q^m}$, and $G \in \mathcal{M}_{k,n}(\mathbb{F}_{q^m})$ a generator matrix of $\mathcal{G}$ of canonical form. Randomly choose matrices $S \in \mathrm{GL}_k(\mathbb{F}_{q^m})$ and $T \in \mathrm{GL}_n(\mathcal{V})$, where $\mathcal{V} \subseteq \mathbb{F}_{q^m}$ is a randomly chosen $\mathbb{F}_q$-linear space of dimension λ. Randomly choose $\boldsymbol{u} \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{u}) = n$ and set $U = \mathrm{Cir}_k(\boldsymbol{u})$. Let $G_{pub} = (SG + U)T^{-1}$, then we publish $(G_{pub}, \boldsymbol{u})$ as the public key, and keep (S, G, T) as the private key.
– Encryption
  For a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^k$, randomly choose $\boldsymbol{e}_1 \in \mathbb{F}_{q^m}^n$ with $\mathrm{rk}(\boldsymbol{e}_1) = r_1$ and $\boldsymbol{e}_2 \in \mathbb{F}_{q^m}^n$ with $\mathrm{rk}(\boldsymbol{e}_2) = r_2$. Compute $\boldsymbol{c}_1 = \boldsymbol{m}U + \boldsymbol{e}_1$ and $\boldsymbol{c}_2 = \boldsymbol{m}G_{pub} + \boldsymbol{e}_2$, then the ciphertext is $\boldsymbol{c} = (\boldsymbol{c}_1, \boldsymbol{c}_2)$.
– Decryption
  For a ciphertext $\boldsymbol{c} = (\boldsymbol{c}_1, \boldsymbol{c}_2) \in \mathbb{F}_{q^m}^{2n}$, compute $\boldsymbol{c}' = \boldsymbol{c}_2T - \boldsymbol{c}_1 = \boldsymbol{m}SG + \boldsymbol{e}_2T - \boldsymbol{e}_1$. Note that $\mathrm{rk}(\boldsymbol{e}_2T - \boldsymbol{e}_1) \leqslant \mathrm{rk}(\boldsymbol{e}_2T) + \mathrm{rk}(\boldsymbol{e}_1) \leqslant \lambda r_2 + r_1 \leqslant r$. Decoding $\boldsymbol{c}'$ with the decoder of $\mathcal{G}$ leads to $\boldsymbol{m}' = \boldsymbol{m}S$, then by computing $\boldsymbol{m}'S^{-1}$ one can recover the plaintext $\boldsymbol{m}$.

Remark 5. It is clear that the public key will degenerate into an instance of the original system if $\lambda = 1$ and $\mathcal{V} = \mathbb{F}_q$, which has been completely broken in the present paper. To achieve the IND-CPA security, the original encryption procedure chooses an extra vector $\boldsymbol{m}_s$ to concatenate $\boldsymbol{m}$, which greatly reduces the information transfer rate. To avoid this defect, we remove the use of $\boldsymbol{m}_s$ in the encrypting process. Consequently, a problem arises that the repaired scheme only satisfies the security notion of One-Wayness. However, we can follow the approach in [24] to convert this repair into an IND-CCA2 secured encryption scheme.

5.2 Security Analysis

Now we investigate the security of this repair in the following three aspects.

Structural Attacks. Resistance of our repair against the existing structural attacks [11,18,19,32,34] is apparent. In what follows, therefore, we only consider the key recovery attack presented in Sect. 4. With a similar analysis, we construct a matrix equation as follows

$$G_{pub}Y^TM^T - UX^TM^T = 0. \tag{7}$$

What differs from (4) is that $X \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ and $Y \in \mathcal{M}_{m,n}(\mathbb{F}_{q^m})$ is taken in an $\mathbb{F}_q$-linear space of dimension λ. Applying SET to (7) will lead to a linear system over $\mathbb{F}_q$, with $k(n-k)m$ equations and $(\lambda+1)mn$ variables. Solving this system generally requires $\mathcal{O}\big((\lambda+1)^3m^3n^3\big)$ operations. Note that we cannot presuppose $\mathbb{F}_q \subseteq \mathcal{V}$ because of the additive structure, which suggests that one has to enumerate λ-dimensional $\mathbb{F}_q$-subspaces of $\mathbb{F}_{q^m}$ with a complexity of $\mathcal{O}\big(q^{\lambda(m-\lambda)}\big)$. Finally the whole complexity of our attack on this repair can be evaluated as $\mathcal{O}\big((\lambda+1)^3m^3n^3q^{\lambda(m-\lambda)}\big)$. It is easy to see that this repair can easily reach the desired security for parameters of proper size.

Generic Attacks. We first introduce the so-called rank syndrome decoding (RSD) problem on which the security of most code-based cryptosystems relies. An RSD problem with parameters (q, m, n, k, t) is to search for a vector $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{e}) = t$ and $\boldsymbol{s} = \boldsymbol{e}H^T$, where $H \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$ is a matrix of full rank and $\boldsymbol{s} \in \mathbb{F}_{q^m}^{n-k}$. Generic attacks on the RSD problem can be divided into two categories, namely the combinatorial attacks as listed in Table 2 and the algebraic attacks as listed in Table 3. The security of a code-based cryptosystem under these attacks only relate to the practical parameters, and does not rely on the algebraic structure of the underlying code.

Table 2. Best known combinatorial attacks on the RSD problem.

Attack	Complexity
[33]	$\mathcal{O}\left(\min\left\{m^3t^3q^{(t-1)(k+1)}, (k+t)^3t^3q^{(t-1)(m-t)}\right\}\right)$
[17]	$\mathcal{O}\left((n-k)^3m^3q^{\min\left\{t\left\lceil\frac{mk}{n}\right\rceil,(t-1)\left\lceil\frac{m(k+1)}{n}\right\rceil\right\}}\right)$
[3]	$\mathcal{O}\left((n-k)^3m^3q^{t\left\lceil\frac{m(k+1)}{n}\right\rceil-m}\right)$

Table 3. Best known algebraic attacks on the RSD problem.

Attack	Condition	Complexity
[17]	$\left\lceil\frac{(t+1)(k+1)-(n+1)}{t}\right\rceil \leqslant k$	$\mathcal{O}\left(k^3t^3q^{t\left\lceil\frac{(t+1)(k+1)-(n+1)}{t}\right\rceil}\right)$
[6]	$m\binom{n-k-1}{t} \geqslant \binom{n}{t} - 1$	$\mathcal{O}\left(m\binom{n-p-k-1}{t}\binom{n-p}{t}^{\omega-1}\right)$, where $\omega = 2.81$ and $p = \min\{1 \leqslant i \leqslant n : m\binom{n-i-k-1}{t} \geqslant \binom{n-i}{t} - 1\}$
[5]		$\mathcal{O}\left(\left(\frac{((m+n)t)^t}{t!}\right)^{\omega}\right)$
[6]	$m\binom{n-k-1}{t} < \binom{n}{t} - 1$	$\mathcal{O}\left(q^{at}m\binom{n-k-1}{t}\binom{n-a}{t}^{\omega-1}\right)$, where $a = \min\{1 \leqslant i \leqslant n : m\binom{n-k-1}{t} \geqslant \binom{n-i}{t} - 1\}$
[5]		$\mathcal{O}\left(\left(\frac{((m+n)t)^{t+1}}{(t+1)!}\right)^{\omega}\right)$

Proposed Parameters. Now we consider the practical security of this repair and propose some parameters for the security of at least 128 bits, 192 bits, and 256 bits. As illustrated in Table 4, we consider $m = n$ and $r_1 = r_2 = t = \lfloor\frac{n-k}{2(\lambda+1)}\rfloor$. The ciphertext of our repair consists of the following two parts

$$\boldsymbol{c}_1 = \boldsymbol{m}U + \boldsymbol{e}_1, \boldsymbol{c}_2 = \boldsymbol{m}G_{pub} + \boldsymbol{e}_2,$$

which lead to an RSD instance of parameters (q, m, n, k, t). Meanwhile, it is easy to see that

$$(\boldsymbol{c}_1||\boldsymbol{c}_2) = \boldsymbol{m}[U|G_{pub}] + (\boldsymbol{e}_1||\boldsymbol{e}_2),$$

and this results in another RSD instance of parameters $(q, m, 2n, k, 2t)$. Additionally, we also consider the proposed key recovery attack described above, which requires $\mathcal{O}\big((\lambda+1)^3 m^3 n^3 q^{\lambda(m-\lambda)}\big)$ operations in $\mathbb{F}_q$. Finally we give some suggested parameters in Table 4, as well as the corresponding public-key sizes.

Table 4. Parameters and public-key size (in bytes).

Parameters							Public-Key Size	Security
q	m	n	k	λ	r_1	r_2		
2	79	79	37	2	7	7	29645	128
2	91	91	43	2	8	8	45546	193
2	110	110	50	2	10	10	77138	265

6 Cryptanalysis of Loidreau's Cryptosystem

The success of our attack on the Lau-Tan cryptosystem relies on four points. One is the fact of Moore matrix decomposition as described in Proposition 2, the second is to construct a system of equations from the parity-check matrix equation, the third is to reduce the problem of solving a multivariate quadratic system into solving a multivariate linear system, and the last is any nonzero solution of this linear system leads to an equivalent private key.

Based on Points 1, 2, and 4 described above, we provide another perspective on the security of Loidreau's cryptosystem [27], which has been completely broken for specific parameters [11,18]. Firstly, we give a simple description for the principle of Loidreau's cryptosystem. The public key in this cryptosystem is published as $G_{pub} = GP^{-1}$, where G is a generator matrix of an $[n, k]$ Gabidulin code $\mathcal{G} \subseteq \mathbb{F}_{q^m}$ and $P \in \mathrm{GL}_n(\mathbb{F}_{q^m})$ with entries contained in a small λ-dimensional $\mathbb{F}_q$-linear space $\mathcal{V} \subseteq \mathbb{F}_{q^m}$. To encrypt a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^k$, one first encodes $\boldsymbol{m}$ by computing $\boldsymbol{m}G_{pub}$, then disguises this codeword by adding an error vector $\boldsymbol{e} \in \mathbb{F}_{q^m}^n$ with $\mathrm{rk}(\boldsymbol{e}) = \lfloor \frac{n-k}{2\lambda} \rfloor$. To decrypt a ciphertext $\boldsymbol{c} = \boldsymbol{m}G_{pub} + \boldsymbol{e}$, one first computes $\boldsymbol{c}' = \boldsymbol{c}P$, then decodes $\boldsymbol{c}'$ with the decoder of $\mathcal{G}$ to recover $\boldsymbol{e}P$ due to $\mathrm{rk}(\boldsymbol{e}P) \leqslant \lfloor \frac{n-k}{2} \rfloor$. Then one can obtain $\boldsymbol{m}$ by solving the linear system $\boldsymbol{m}G = \boldsymbol{c}' - \boldsymbol{e}P$.

In a talk [28] at CBCrypto 2021, Loidreau proposed an attack to recover a polynomial-time decoder of the public code with a complexity of $\mathcal{O}\big((\lambda n + (n-k)^2)^3 m^3 q^{(\lambda-1)m}\big)$ for $q = 2$, which can be easily generalized to any field $\mathbb{F}_q$. Loidreau's attack manages to recover $Y \in \mathcal{M}_{m,n}(\mathcal{V})$ such that $G_{pub}(MY)^T = 0$, where $M \in \mathcal{M}_{n-k,m}(\mathbb{F}_{q^m})$ is a Moore matrix generated by a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. Then with the knowledge of Y, one can decrypt any ciphertext in polynomial time. Specifically, let $\boldsymbol{c} = \boldsymbol{m}G_{pub} + \boldsymbol{e}$ be the received ciphertext, then one computes $\boldsymbol{s} = \boldsymbol{c}(MY)^T = \boldsymbol{e}Y^T M^T$. Note that $\mathrm{rk}(\boldsymbol{e}Y^T) \leqslant \lfloor \frac{n-k}{2} \rfloor$, then one can recover $\boldsymbol{e}' = \boldsymbol{e}Y^T$ by using the syndrome decoder of an $[m, m-n+k]$ Gabidulin code that has M as a parity-check matrix. After that, one can recover $\boldsymbol{e}$ by solving the linear system $\boldsymbol{e}' = \boldsymbol{e}Y^T$.

Now we apply the proposed attack to Loidreau's cryptosystem. Let $H \in \mathcal{M}_{n-k,n}(\mathbb{F}_{q^m})$ be a canonical parity-check matrix of $\mathcal{G}$, and $\mathcal{G}_{pub} = \langle G_{pub} \rangle_{q^m}$ the public code. It is clear that $H_{pub} = HP^T$ forms a parity-check matrix of $\mathcal{G}_{pub}$. Let $M \in \mathcal{M}_{n-k,m}(\mathbb{F}_{q^m})$ be a Moore matrix generated by a basis vector of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, then there exists $Y^* \in \mathcal{M}_{m,n}(\mathbb{F}_q)$ such that $H = MY^*$. Let $Y = Y^*P^T$, then one can construct a linear system from the parity-check matrix equation $G_{pub}H_{pub}^T = 0$, which is equivalent to

$$G_{pub}(MY)^T = G_{pub}Y^T M^T = 0. \tag{8}$$

While in Loidreau's attack, the corresponding linear system is constructed from $SH_{pub} = MY$, which introduces extra variables from an underdetermined matrix $S \in \mathrm{GL}_{n-k}(\mathbb{F}_{q^m})$. Applying SET to (8) leads to a linear system over $\mathbb{F}_q$ of $k(n-k)m$ equations and λmn variables. For cryptographic use, $k(n-k)m \geqslant \lambda mn$ always holds in practical situations. And this system admits q^m solutions when $\mathcal{V}$ is correctly guessed, which has been validated through numerous experiments. Solving this system for any nonzero solution permits us to obtain $Y' \in \mathcal{M}_{m,n}(\mathcal{V})$ such that $G_{pub}(MY')^T = 0$. On the other hand, one can always presuppose $1 \in \mathcal{V}$ since $G_{pub} = \alpha^{-1}G(\alpha^{-1}P)^{-1}$ for any nonzero $\alpha \in \mathcal{V}$. Finally this attack requires a complexity of $\mathcal{O}\big(\lambda^3 n^3 m^3 q^{(\lambda-1)m}\big)$ in $\mathbb{F}_q$, which is clearly lower than Loidreau's attack.

7 Conclusion

Our attack has revealed the structural weakness of the Lau-Tan cryptosystem. Although the first part of the public key hides the structure of Gabidulin codes nicely, the second part reveals important information that can be used to design a key recovery attack. Specifically, we convert the problem of recovering the private key into solving a multivariate linear system over the base field. Extensive experiments have been performed and the results accord with our theoretical expectations. To prevent this attack, we give a simple but effective repair for this cryptosystem, which is shown to be secure against all the existing structural attacks. Furthermore, when applying this attack to analyze Loidreau's cryptosystem, we reduce the complexity of recovering a polynomial-time decoder of the public code.

References

1. Aguilar-Melchor, C., Blazy, O., Deneuville, J.-C., Gaborit, P., Zémor, G.: Efficient encryption from random quasi-cyclic codes. IEEE Trans. Inform. Theory **64**(5), 3927–3943 (2018)
2. Aragon, N., Gaborit, P., Hauteville, A., Ruatta, O., Zémor, G.: Low rank parity check codes: new decoding algorithms and applications to cryptography. IEEE Trans. Inform. Theory **65**(12), 7697–7717 (2019)

3. 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 2018), pp. 2421–2425. IEEE (2018)
4. Baldi, M., Chiaraluce, F., Garello, R.: On the usage of quasi-cyclic low-density parity-check codes in the McEliece cryptosystem. In: Proceedings of 2007 IEEE International Conference on Communications (ICC 2007), pp. 951–956. IEEE (2007)
5. Bardet, M., Briaud, P., Bros, M., Gaborit, P., Neiger, V., Ruatta, O., Tillich, J.-P.: An algebraic attack on rank metric code-based cryptosystems. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020. LNCS, vol. 12107, pp. 64–93. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45727-3_3
6. Bardet, M., Bros, M., Cabarcas, D., Gaborit, P., Perlner, R., Smith-Tone, D., Tillich, J.-P., Verbel, J.: Improvements of algebraic attacks for solving the rank decoding and MinRank problems. In: Moriai, S., Wang, H. (eds.) ASIACRYPT 2020. LNCS, vol. 12491, pp. 507–536. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64837-4_17
7. Berger, T., Loidreau, P.: Designing an efficient and secure public-key cryptosystem based on reducible rank codes. In: Canteaut, A., Viswanathan, K. (eds.) INDOCRYPT 2004. LNCS, vol. 3348, pp. 218–229. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30556-9_18
8. Berlekamp, E.R., McEliece, R.J., Van Tilborg, H.: On the inherent intractability of certain coding problems. IEEE Trans. Inform. Theory **24**(3), 384–386 (1978)
9. Bombar, M., Couvreur, A.: Decoding supercodes of gabidulin codes and applications to cryptanalysis. In: Cheon, J.H., Tillich, J.-P. (eds.) PQCrypto 2021 2021. LNCS, vol. 12841, pp. 3–22. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-81293-5_1
10. Bosma, W., Cannon, J., Playoust, C.: The MAGMA algebra system I: the user language. J. Symbolic Comput. **24**(3–4), 235–265 (1997)
11. Coggia, D., Couvreur, A.: On the security of a Loidreau rank metric code based encryption scheme. Des. Codes Crypt. **88**(9), 1941–1957 (2020). https://doi.org/10.1007/s10623-020-00781-4
12. Faure, C., Loidreau, P.: A new public-key cryptosystem based on the problem of reconstructing p–polynomials. In: Ytrehus, Ø. (ed.) WCC 2005. LNCS, vol. 3969, pp. 304–315. Springer, Heidelberg (2006). https://doi.org/10.1007/11779360_24
13. Gabidulin, E.M.: Theory of codes with maximum rank distance. Prob. Peredachi Inf. **21**(1), 3–16 (1985)
14. Gabidulin, E.M., Ourivski, A.V., Honary, B., Ammar, B.: Reducible rank codes and their applications to cryptography. IEEE Trans. Inform. Theory **49**(12), 3289–3293 (2003)
15. 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.) EUROCRYPT 1991. LNCS, vol. 547, pp. 482–489. Springer, Heidelberg (1991). https://doi.org/10.1007/3-540-46416-6_41
16. 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)
17. Gaborit, P., Ruatta, O., Schrek, J.: On the complexity of the rank syndrome decoding problem. IEEE Trans. Inf. Theory **62**(2), 1006–1019 (2016)
18. Ghatak, A.: Extending Coggia-Couvreur attack on Loidreau's rank-metric cryptosystem. Des. Codes Cryptogr. **90**, 215–238 (2022)

19. Horlemann-Trautmann, A.-L., Marshall, K., Rosenthal, J.: Considerations for rank-based cryptosystems. In: Proceedings of 2016 IEEE International Symposium on Information Theory (ISIT 2016), pp. 2544–2548. IEEE (2016)
20. 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)
21. 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)
22. Horlemann-Trautmann, A.-L., Werger, V.: Information set decoding in the Lee metric with applications to cryptography. Adv. Math. Commun. **15**(4), 677–699 (2021)
23. Janwa, H., Moreno, O.: McEliece public key cryptosystems using algebraic-geometric codes. Des. Codes Cryptogr. **8**(3), 293–307 (1996)
24. Lau, T.S.C., Tan, C.H.: New rank codes based encryption scheme using partial circulant matrices. Des. Codes Crypt. **87**(12), 2979–2999 (2019). https://doi.org/10.1007/s10623-019-00659-0
25. Lau, T.S.C., Tan, C.H.: A new encryption scheme based on rank metric codes. In: Susilo, W., Yang, G. (eds.) ACISP 2018. LNCS, vol. 10946, pp. 750–758. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93638-3_43
26. Lau, T.S.C., Tan, C.H.: A new technique in rank metric code-based encryption. Cryptography **2**(4), 32 (2018)
27. Loidreau, P.: A new rank metric codes based encryption scheme. In: Lange, T., Takagi, T. (eds.) PQCrypto 2017. LNCS, vol. 10346, pp. 3–17. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59879-6_1
28. Loidreau, P.: Analysis of a rank metric codes based encryption scheme. https://drive.google.com/file/d/1FuMgqm0NfGMJOxaZyrIrI1OWn0UICwPo/view. Accessed 1 July 2021
29. Loidreau, P.: A welch–berlekamp like algorithm for decoding gabidulin codes. In: Ytrehus, Ø. (ed.) WCC 2005. LNCS, vol. 3969, pp. 36–45. Springer, Heidelberg (2006). https://doi.org/10.1007/11779360_4
30. McEliece, R.J.: A public-key cryptosystem based on algebraic coding theory. Jet Propuls. Lab. DSN Progr. Rep. **42–44**, 114–116 (1978)
31. Niederreiter, H.: Knapsack-type cryptosystems and algebraic coding theory. Prob. Contr. Inform. Theory **15**(2), 157–166 (1986)
32. 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)
33. 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)
34. Overbeck, R.: Structural attacks for public key cryptosystems based on Gabidulin codes. J. Cryptology **21**(2), 280–301 (2008)
35. Renner, J., Puchinger, S., Wachter-Zeh, A.: LIGA: a cryptosystem based on the hardness of rank-metric list and interleaved decoding. Des. Codes Cryptogr. **89**(6), 1279–1319 (2021). Springer
36. Richter, G., Plass, S.: Error and erasure decoding of rank-codes with a modified Berlekamp-Massey algorithm. ITG FACHBERICHT, pp. 203–210 (2004)
37. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Rev. **41**(2), 303–332 (1994)