RQC Revisited and More Cryptanalysis for Rank-Based Cryptography

Loïc Bidoux, Pierre Briaud, Maxime Bros, and Philippe Gaborit

Abstract—In this paper, we revisit the Rank Quasi-Cyclic (RQC) (Melchor et al., IEEE IT, 2018) encryption scheme by proposing three possible variations for its design. Our first improvement relies on the introduction of Augmented Gabidulin codes, a new family of decodable codes exploiting the concept of support erasure for the rank metric. Following the work of Melchor et al. (PQCrypto, 2022), our second improvement uses multiple syndromes to increase the weight of the error to be decoded. As pioneered in Melchor et al. (NIST PQC, 2020), our third variation considers non-homogeneous error weights in order to decrease the parameters. These improvements can be combined together to design schemes offering various trade-offs in terms of security and size. Our Multi-UR-AG (multiple syndromes, unstructured, augmented Gabidulin) scheme achieves a size of 11kB (public key + ciphertext) for 128 bits of security while featuring a conservative design as it relies on pure random instances without any ideal structure. Besides, our NH-Multi-RQC-AG (non-homogeneous error, multiple syndromes, ideal structure, augmented Gabidulin) achieves a size of 2.7 kB for 128 bits of security, namely a 50 % improvement with respect to classical RQC. Our second and third variations respectively rely on the security of the RSL and NHRSD problems (or NHRSRL when considered together). In this paper, we also provide new security analysis and attacks for these problems. While these results are important for our new schemes, they are of independent interest as well. Our security analysis for the RSL problem provides an improvement on the recent algebraic attacks for some instances. In addition, we show that the RSL problem can be solved in polynomial time when \( N \geq (k+1) \frac{m}{m-1} \), this improves the best known combinatorial attack (Gaborit et al., Crypto, 2017). We also propose the first combinatorial attack against the NRSDL problem along with a precise complexity analysis of the algebraic attack described Melchor et al. (NIST PQC, 2020). At last, we combine these analysis to provide an attack against the NRSL problem.

Index Terms—Public key cryptosystem, coding and information theory.

I. INTRODUCTION

A. Background on Rank Metric Code-Based Cryptography

In the last decade, rank Metric code-based cryptography has evolved to become a real alternative to traditional code-based cryptography based on the Hamming metric. The original scheme based on rank metric was the GPT cryptosystem [5], an adaptation of the McEliece scheme in a rank metric context where Gabidulin codes [6], a rank metric analog of Reed-Solomon codes, were the masked codes. However, the strong algebraic structure of these codes was successfully exploited for attacking the original GPT cryptosystem and its variants with the Overbeck attack [7] (see [8] for the latest developments). This situation is similar to the Hamming metric where most of McEliece cryptosystems based on variants of Reed-Solomon codes have been broken.

Besides the McEliece scheme where a secret code is masked through using permutation, it is possible to generalize the approach by considering public key matrices with trapdoor. Examples of such an approach are NTRU [9] or MDPC [10] cryptosystems where the masking consists in knowing a very small weight vector of the given public matrix. Such an approach was adapted to the rank metric through the introduction of LRPC codes [11].

The security of such type of cryptosystems relies on the generic rank decoding problem together with the computational indistinguishability of the public key which is a matrix. The fact that the public matrix is used both for encryption and decryption, permits to obtain very efficient schemes, at the cost of an inversion. It is worth noticing that Loidreau’s scheme, which uses homogeneous LRRC matrices in a McEliece context, seems to resist to structural attacks with an homogeneous matrix of sufficiently high rank [12].

B. The RQC Scheme

Another approach, proposed by Aleknovich in [13], permits to rely solely on random instances of the Syndrome Decoding problem without any masking of a public key. However, such an approach is strongly inefficient in practice. A few years later, a more optimized approach was proposed with the HQC scheme [14], relying on Quasi-Cyclic codes. It has been generalized to rank metric with the RQC scheme [3]. For these schemes, two types of codes are used: a first random double circulant code permits to ensure the security of the scheme when a second public code permits to decode/decrypt the ciphertext. In RQC, Gabidulin codes are

 Manuscript received 10 January 2023; revised 29 August 2023; accepted 2 November 2023. Date of publication 1 December 2023; date of current version 16 February 2024. The work of Philippe Gaborit was supported by the French ANR France 2030, under Project ANR-22-PETQ-0008.

(Corresponding author: Maxime Bros.)

Loïc Bidoux was with the Cryptography Research Center, Technology Innovation Institute, Abu Dhabi, United Arab Emirates.
Pierre Briaud was with Inria, 75012 Paris, France, and also with UPMC University, Sorbonne Universités, 75006 Paris, France.
Maxime Bros was with the National Institute of Standards and Technology, Gaithersburg, MD 20899 USA. (e-mail: maxime.bros@nist.gov).
Philippe Gaborit was with CNRS, XLIM, University of Limoges, 87000 Limoges, France.

Communicated by A. Wachter-Zeh, Associate Editor for Coding and Decoding.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2023.3337945.

Digital Object Identifier 10.1109/TIT.2023.3337945

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.

See https://www.ieee.org/publications/rights/index.html for more information.

© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

0018-9448 © 2023 IEEE.
used as public decryption codes. Besides RQC, some other variations were proposed in [4], [15], and [16]. The main advantage of RQC compared to LRPC-based cryptosystems is the fact that its security reduction is done to random decoding instances whereas the LRPC approach requires an additional indistinguishability assumption; however, this advantage comes at a price since parameters are larger for RQC than for LRPC.

The RQC scheme was proposed to the NIST Standardization Process with very competitive parameters, however the algebraic attacks of [17] and [18], which were published during the standardization process, had a dreadful impact on RQC parameters so that, in order not to increase too much RQC parameters, the introduction of non-homogeneous errors [3] permitted to limit the impact of these algebraic attacks.

The idea of non-homogeneous errors is to consider errors in three parts of length $n$ such that the error weight is the same for the first two sets, but larger for the third one. Such an approach permits to limit the impact of the security reduction of RQC to decoding random $[3n, n]$ codes rather than $[2n, n]$ codes in LRPC cryptosystems. The notion of non-homogeneous error led to the introduction of the Non Homogeneous RSD problem (NHRSD) and was a first approach to decrease the size of RQC parameters. At this point, it is meaningful to notice that for LRPC and RQC systems, the weight of the error to attack is structurally $O(\sqrt{n})$ (where $n$ is the length of the code), a type of parameters for which algebraic attacks are very efficient.

Besides the RSD problem, the RSL problem which consists in having $N$ syndromes whose associated errors share the same support, was introduced in [4] to construct the RankPKE scheme and later in [15]. This problem which generalizes RSD is meaningful to give more margin in building cryptosystems; it has been recently used to improve on the LRPC schemes [2]. It permits, in particular, to increase the weight of the error to decode from $O(\sqrt{n})$ to a weight closer to the Rank Gilbert-Varshamov (RGV) bound; this is important since for that type of parameters, i.e. close to the RGV bound, algebraic attacks become relatively less efficient than combinatorial attacks.

### C. Attacks and Problems in Rank Metric

There are two types of attacks in rank metric. Combinatorial attacks which were the first to be introduced in the late 1990's then algebraic attacks ten years later. At first, combinatorial attacks were the most efficient ones, but recently, and especially for parameters where the error has weight $O(\sqrt{n})$, the seminal approaches of [17] and [18] permitted to have a strong impact on such parameters.

Besides the RSD problem, the RSL problem was studied in [4] and [19], and more recently algebraic attacks were considered in [20]. In particular, it was shown in [4] that any RSL instance with more than $nr$ syndromes could be solved in polynomial time. Moreover, the Non Homogeneous RSD (NHRSD) problem was introduced in [3] where a first algebraic attack was proposed.

### D. Contributions

As mentioned above, before this present paper some new problems in rank metric emerged (NHRSD, RSL); they permitted to improve the parameters both for RQC and LRPC-based cryptosystems. In this paper, our contributions are twofold: first, we propose new variations on the RQC scheme in order to improve its parameters; second, we study in details the new problems on which are based these approaches. The NHRSD and RSL problems appear as natural variations of RSD and are bound to be the future problems on which will be relying rank metric-based cryptosystems.

1) New Schemes: The new schemes we propose are based on three types of improvements.

Our first and main improvement is the introduction of a new family of decodable codes, namely Augmented Gabidulin codes. These codes exploit the concept of support erasure in a rank-metric context. The introduction of known support erasure permits to decrease the value $m$ down to a value close to the weight of the error, whereas $m$ had to be at least twice bigger with classical Gabidulin codes. This comes at the cost of a probabilistic decoding; however the decryption failure rate (DFR) can be controlled very easily as it is done with LRPC in [2] and [11].

Second, as for the recent LRPC-based cryptosystems improvement [2], we consider the use of multiple syndromes in RQC. This approach permits to greatly improve the decoding capacity of the RQC scheme by increasing the information available for the decryption at a lower cost than directly increasing all parameters. This variation implies that the scheme relies on RSL [4], [15], [21] rather than RSD. In particular, as for the LRPC approach [2], it permits to increase the weight of the error to decode, thus getting closer to the rank Gilbert-Varshamov bound, and no longer in the classical $O(\sqrt{n})$ area.

Third, like pioneered in [3], we use a variant of the RSD problem by considering an error with non-homogeneous rank weights $(w_1, w_2)$ which is the Non-Homogeneous RSD problem (NHRSD). In short, this error contains a part of weight $w_1$ and a part of weight $w_1 + w_2$. This optimization allows one extra degree of freedom while choosing the target error weights, and this has a strong impact on decreasing the parameters.

Overall, we propose two types of scheme with very competitive sizes. First, Multi-RQC-AG has parameters similar to MS-LRPC [2], i.e., around 4.5 KBytes for the public key together with the ciphertext, and its security relies on ideal codes. Second, Multi-UR-AG has parameters a little bit larger than MS-LRPC, around 11 KBytes total, this time without any structure; more precisely, it relies only on pure random instances of the RSL problem. This is the most conservative security one could expect. For both of the aforementioned schemes, one could add a non-homogeneous structure in order to shorten the sizes down by 30%, this corresponds to our schemes: NH-Multi-RQC-AG and NH-Multi-UR-AG.

The scheme we propose without any ideal structure with small parameters of 11KBytes is meaningful; indeed, since it is not proven that any ideal structure cannot be used
to get more efficient attacks with a quantum computer, schemes without any ideal structure may hence provide a better security. Naturally, in that case, the size of parameters increases a lot; however, with the rank metric it remains small compared to Hamming metric for which having no additional structure implies very large public key (see for instance McEliece schemes). Moreover, our scheme does not require any additional indistinguishability assumption.

2) New Attacks and Analysis: As previously mentioned, our new improvements on RQC rely on recent problems for rank metric, namely the NHRSD et RSL problems (or on NHRSL, a combination of the two latter). Although these problems have begun to be considered, we go deeper in their security evaluation by proposing new attacks and adaptation of known attacks. The motivation comes both from the general interest of these problems for rank-based cryptography, and for the attacks. The motivation comes both from the general interest of these problems for rank-based cryptography, and for the attacks.

More precisely, recall that an RSL \((m, n, k, r, N)\) instance is like a rank syndrome decoding instance of parameters \((m, n, k, r)\) where \(N\) instead of 1 syndromes are given, and all their associated errors share the same support. The security of RSL is inherent to the value of \(N\), and it is known since [4] that the problem can be solved in polynomial time as long as \(N > nr\). Our contributions are then the following:

- With our new combinatorial attack against RSL, first we improve on the most recent algebraic attack for some instances; most importantly, we improve the aforementioned bound, unchanged since 2017 [4], showing that RSL can be solved in polynomial time as long as

\[ N \geq (k + 1)r \frac{m}{m - r}. \]

- We also propose the first combinatorial attack against NHRSD, together with a precise complexity analysis of the algebraic attack described in [3].

- At last, we propose an attack against NHRSL. That it is to say that we were able to take advantage of two structure in the same attack: the fact that the error is non-homogeneous and that one is given several syndromes.

II. PRELIMINARIES

A. Coding Theory and Rank Metric

Let \(q\) be a prime power, let \(m\) a positive integer, let \(\mathbb{F}_q^m\) an extension of degree \(m\) of \(\mathbb{F}_q\) and let \(\beta := (\beta_1, \ldots, \beta_m)\) be an \(\mathbb{F}_q\)-basis of \(\mathbb{F}_q^m\). Any vector in \(\mathbb{F}_q^m\) can naturally be viewed as a matrix in \(\mathbb{F}_q^{m \times n}\) by expressing its coordinates in \(\beta\).

**Definition 1 (Rank Weight):** Let \(x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n\) be a vector. The rank weight of \(x\) denoted \(\|x\|\) is defined as the rank of the matrix \(\text{Mat}(x) := (x_{ij})_{i,j} \in \mathbb{F}_q^{m \times n}\) where \(x_{ij} = \beta_1 x_{i1} + \cdots + \beta_m x_{mj}\) for \(j \in \{1, n\}\). The set of vectors of weight \(w\) in \(\mathbb{F}_q^n\) is denoted \(S^w(\mathbb{F}_q^n)\).

**Definition 2 (Support):** The support of \(x \in \mathbb{F}_q^n\) is the \(\mathbb{F}_q\)-linear space generated by the coordinates of \(x\), i.e., \(\text{Supp}(x) := (x_1, \ldots, x_n)_{\mathbb{F}_q}\). It follows from the definition that \(\|x\| = \dim_{\mathbb{F}_q}(\text{Supp}(x))\).

These notions can be extended to matrices. The support of a matrix \(M \in \mathbb{F}_q^{m \times c}\) denoted \(\text{Supp}(M)\) is the \(\mathbb{F}_q\)-vector space spanned by all its \(r \times c\) entries, and the rank weight \(\|M\|\) is defined as the dimension of this support.

**Definition 3 (\(\mathbb{F}_q^m\)-linear code):** An \(\mathbb{F}_q^m\)-linear code \(C\) of length \(n\) and dimension \(k\) is an \(\mathbb{F}_q^m\)-linear subspace of \(\mathbb{F}_q^n\) of dimension \(k\). We say that it has parameters \([n, k]_{\mathbb{F}_q^m}\).

A generator matrix for \(C\) is a full-rank matrix \(G \in \mathbb{F}_q^{k \times n}\) such that \(C = \{mG, m \in \mathbb{F}_q^k\}\). A parity-check matrix is a full-rank matrix \(H \in \mathbb{F}_q^{(n-k) \times n}\) such that \(C = \{x \in \mathbb{F}_q^n, \; xH^T = 0\}\). Finally, the rowspace of \(H\) is a basis of the dual code \(C^\perp\).

The use of \(\mathbb{F}_q^m\)-linear codes instead of standard \(\mathbb{F}_q\)-linear codes permits to obtain a more compact description for the public key in code-based cryptosystems. Another classical way to reduce the keysize is to use some type of cyclic structure, which leads to the notion of ideal codes. Let \(P \in \mathbb{F}_q[X]\) denote a polynomial of degree \(n\). The linear map \(u := (0_1, \ldots, u_{n-1}) \mapsto u(X) := \sum_{i=0}^{n-1} u_i X^i\) is a vector space isomorphism between \(\mathbb{F}_q^n\) and \(\mathbb{F}_q[X]/(P)\), and we use it to define a product between two elements \(u, v \in \mathbb{F}_q^n\) via \(u \cdot v := u(X)v(X) \mod P\). Note that we have

\[ u \cdot v = \left(\sum_{i=0}^{n-1} u_i X^i\right) \cdot \left(\sum_{i=0}^{n-1} v_i X^i\right) \mod P = \sum_{i=0}^{n-1} u_i \left(X^i v(X) \mod P\right), \]

so that the product by \(v \in \mathbb{F}_q^n\) can be seen as a matrix-vector product by the so-called ideal matrix generated by \(x\) and \(P\).

**Definition 4 (Ideal Matrix):** Let \(P \in \mathbb{F}_q[X]\) a polynomial of degree \(n\) and let \(v \in \mathbb{F}_q^n\). The ideal matrix generated by \(v\) and \(P\), noted \(\mathcal{I}_M(P,v)\), is the \(n \times n\) matrix, with entries in \(\mathbb{F}_q^n\), and whose rows are the following: \(v(X) \mod P, \; Xv(X) \mod P, \ldots, \; X^{n-2}v(X) \mod P\). For conciseness, we use the notation \(\mathcal{I}_M(v)\) since there will be no ambiguity in the choice of \(P\) in the paper.

One can see that \(u \cdot v = u \mathcal{I}_M(v) = v \mathcal{I}_M(u) = v \cdot u\).

An ideal code \(C\) of parameters \([sn, tn]\) is an \(\mathbb{F}_q^n\)-linear code which admits a generating matrix made of \(s \times t\) ideal matrix blocks in \(\mathbb{F}_q^{sn \times tn}\). A crucial point regarding the choice of the modulus \(P\) (see [3, Lemma 1]) is that if \(P \in \mathbb{F}_q[X]\) is irreducible of degree \(n\) and if \(n\) and \(m\) are prime, then such a code \(C\) always admits a systematic generator matrix made of ideal blocks. Hereafter, we only consider \(t = 1\).

**Definition 5 (Ideal Codes):** Let \(P \in \mathbb{F}_q[X]\) a polynomial of degree \(n\). An \([ns, nt]\) \(\mathbb{F}_q^n\)-code \(C\) is an ideal code if it has a generator matrix of the form \(G = (I_n \mathcal{I}_M(g_i) \ldots \mathcal{I}_M(g_{s-1})) \in \mathbb{F}_q^{n \times ns}\), where \(g_i \in \mathbb{F}_q^n\) for \(1 \leq i \leq s - 1\). Similarly, \(C\) is an ideal code if it admits a parity-check matrix of the form

\[ H = \begin{pmatrix} I_{n(s-1)} & \mathcal{I}_M(h_1)^T \\ \vdots & \vdots \\ \mathcal{I}_M(h_{s-1})^T \end{pmatrix} \in \mathbb{F}_q^{(n(s-1)) \times ns}. \]

B. Gabidulin Codes

Gabidulin codes were introduced by Gabidulin in 1985 [6]. These codes can be seen as the rank metric analog of...
Reed-Solomon codes [22], where standard polynomials are replaced by \(q\)-polynomials (also called Ore polynomials or linearized polynomials).

Definition 6 (\(q\)-Polynomials): The set of \(q\)-polynomials over \(\mathbb{F}_{q^m}\) is the set of polynomials with the following shape:

\[
P(X) = \sum_{i=0}^{r} p_i X^i, \quad \text{with } p_i \in \mathbb{F}_{q^m} \text{ and } p_r \neq 0.
\]

The \(q\)-degree of a \(q\)-polynomial \(P\) is defined as \(\text{deg}_q(P) = r\).

Definition 7 (Ring Structure): The set of \(q\)-polynomials over \(\mathbb{F}_{q^m}\) is a non-commutative ring for \((+ , \circ)\), where \(\circ\) is the composition of \(\mathbb{F}_q\)-linear endomorphisms.

Due to their structure, the \(q\)-polynomials are inherently related to decoding problems in the rank metric as stated by the following propositions.

Theorem 1 ([23]): Any \(\mathbb{F}_{q}\)-subspace of \(\mathbb{F}_{q^m}\) of dimension \(r\) is the set of the roots of a unique monic \(q\)-polynomial \(P\) such that \(\text{deg}_q(P) = r\).

Corollary 1: Let \(x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_{q^m}^n\) and \(V\) be the monic \(q\)-polynomial of smallest \(q\)-degree such that \(V(x_i) = 0\) for \(1 \leq i \leq n\), then \(\|x\| = r\) if and only if \(\text{deg}_q(V) = r\).

Finally, Gabidulin codes can be seen as the evaluation of \(\mathbb{F}_q\)-polynomials of bounded degree on the coordinates of a fixed vector over \(\mathbb{F}_{q^m}\).

Definition 8 (Gabidulin Codes): Let \(k, n, m \in \mathbb{N}\) such that \(k \leq n \leq m\) and let \(g = (g_1, \ldots, g_n)\) be an \(\mathbb{F}_q\) linearly independent family of elements of \(\mathbb{F}_{q^m}\). The Gabidulin code \(G_{g}(m, k)\) is the code of parameters \([n, k]_{q^m}\) defined by

\[
G_{g}(m, k) := \{ P(g) \mid \text{deg}_q(P) < k \},
\]

where \(P(g) := (P(g_1), \ldots, P(g_n))\).

C. Hard Problems in Rank-Based Cryptography

As in the Hamming metric, the main source of computational hardness for rank-based cryptosystems is a decoding problem. More precisely, it is the decoding problem in the rank metric setting restricted to \(\mathbb{F}_{q^m}\)-linear codes which is called the Rank Syndrome Decoding Problem (RSD).

Problem 1 (RSD Problem, Search): Given \(H \in \mathbb{F}_{q^m}^{(n-k) \times n}\), a full rank parity-check matrix for a random \(\mathbb{F}_{q^m}\)-linear code \(C\), an integer \(w \in \mathbb{N}\) and a syndrome \(s \in \mathbb{F}_{q^m}^{n-k}\), the Rank Syndrome Decoding problem \(\text{RSD}(m, n, k, w)\) asks to find \(e \in \mathbb{F}_{q^m}\) such that \(\|e\| = w\) and \(He^T = s^T\).

The decision version is denoted by \(\text{DRSD}\). Even if RSD is not known to be NP-complete, there exists a randomized reduction from RSD to an NP-complete problem, namely to decoding in the Hamming metric [24]. The average number of solutions for a fixed weight \(w\) is given by the following Gilbert-Varshamov bound for the rank metric.

Definition 9 (Gilbert-Varshamov Bound): The Gilbert-Varshamov bound \(w_{GV}(q, m, n, k)\) for \(\mathbb{F}_{q^m}\)-linear codes of length \(n\) and dimension \(k\) in the rank metric is defined as the smallest positive integer \(t\) such that

\[q^{m(n-k)} \leq B_t, \quad \text{where } B_t := \sum_{j=0}^{t} \left(\prod_{i=0}^{j-1} (q^n - q^i)\right) (\begin{array}{c} m \end{array})\]

is the size of the ball of radius \(t\) in the rank metric.

In other words, it means that, with overwhelming probability, as long as \(w \leq w_{GV}(q, m, n, k)\), a random RSD instance will have at most a unique solution. In this paper, we also focus on a slightly less standard assumption which is the NHRSRSD problem. This RSD variant was proposed in the Second Round update of RQC [3] in order to mitigate the impact of the recent algebraic RSD attacks [17, 18] on the choice of the parameters. In NHRSRSD, the error \(e\) is no longer a random low weight vector but instead a vector with a \(\text{non-homogeneous weight}\).

Problem 2 (NHRSRSD Problem, Search): Given \(H \in \mathbb{F}_{q^m}^{(n+n_1) \times (2n+n_1)}\), a full rank parity-check matrix of a random \(\mathbb{F}_{q^m}\)-linear code \(C\) of parameters \([2n + n_1, n_1]_{q^m}\), integers \((w_1, w_2)\in \mathbb{N}^2\), and a syndrome \(s = (2n+n_1)\), the Non-Homogeneous Rank Syndrome Decoding problem \(\text{NHRSRSD}(m, n, n_1, w_1, w_2)\) asks to find a vector \(e = (e_1, e_2, e_3) \in \mathbb{F}_{q^m}^{2n+n_1}\) such that \(\|e_1, e_3\| = w_1, \|e_2\| = w_1 + w_2, \text{Supp}(e_1, e_3) \subset \text{Supp}(e_2)\), and such that \(He^T = s^T\).

We denote by \(\text{DNHRSRSD}\) the corresponding decisional version. Note that Definition 2 is slightly more general than the one of [3] where it is assumed that \(n = n_1\). Finally, recall that one of our improvements on the RQC scheme uses multiple syndromes which are correlated since they correspond to errors which share the same support. This formulation exactly corresponds to the definition of the Rank Support Learning problem (RSL and DRSL for the decision version). This problem can be seen as the rank metric analog of the Support-Learning problem in the Hamming metric [25, 26].

Problem 3 (RSL Problem, Search): Given \((H, HE^T)\), where \(H \in \mathbb{F}_{q^m}^{(n+n_1) \times n}\) is of full-rank, and \(E \in \mathbb{F}_{q^m}^{N \times n}\) has all its entries lying in a subspace \(V \subset \mathbb{F}_{q^m}\) of dimension \(w \in \mathbb{N}\), the Rank Support Learning problem \(\text{RSL}(m, n, k, w, N)\) asks to find the secret subspace \(V\).

RSL may enable the construction of more advanced cryptographic primitives in the rank metric. It was introduced in [4], and is at the core of the Durandal signature scheme [21], and the recent Multi-LRPC proposal [2]. Naturally, it is possible to somehow combine the error distributions from Problems 2 and 3.

Problem 4 (NHRSRSL Problem, Search): Given \((H, HE^T)\), where \(H = (n + n_1) \times (2n + n_1)\) is of full rank, and \(E \in \mathbb{F}_{q^m}^{N \times (2n+n_1)}\) such that \(e_i = E_i, s = (e_{i,1}, e_{i,2}, e_{i,3}) \in \mathbb{F}_{q^m}^{2n+n_1}, \|e_{i,1}, e_{i,3}\| = w_1, \|e_{i,2}\| = w_1 + w_2, \text{and such that the supports } V := \text{Supp}(e_{i,1}, e_{i,3}) \subset W := \text{Supp}(e_{i,2})\), and such that the supports \(V, W\) are independent of \(i\); the Non-Homogeneous Rank Support Learning problem \(\text{NHRSRSL}(m, n, n_1, w_1, w_2, N)\) asks to find the secret subspaces \(V\) and \(W\).

Finally, both classic RQC and our Multi-RQC-AG proposal involve ideal codes, so that we have to consider the ideal versions of Problems 1, 2, 3, and 4 (denoted by, respectively, IRSD, NHIRSD, IRSL, and NHIRSL). For the sake of conciseness, we do not give a formal definition of these ideal variants.
SetUp(1^n): Generates and outputs param = (n, k, d, w, w_1, w_2, P) where P ∈ ℤ_q[X] is an irreducible polynomial of degree n.

KeyGen(param): Samples h ← ℤ_q^n, g ← ℤ_q^n, and (x, y) ← ℤ_q^{n_1} × ℤ_q^{n_2}, computes the generator matrix G ∈ ℤ_q^{n×m} of a code C, sets pk = (g, h, s = x + h · y mod P) and sk = (x, y), returns (pk, sk).

Encrypt(pk, m, θ): Uses randomness θ to generate (r_1, e, r_2) ← ℤ_q^{n_1} × ℤ_q^n × ℤ_q^n, sets u = r_1 + h · r_2 mod P and v = mG + s · r_2 + e mod P, returns c = (u, v).

Decrypt(sk, c): Returns C.Decode(v - u · y mod P).

Fig. 1. Description the RQC PKE scheme.

D. RQC Scheme

On Figure 1 we briefly recall the classical RQC scheme [3], for which one needs the following notation:

\[ S_{w,1}^{n}(ℤ_q^n) = \{ x ∈ ℤ_q^n \mid \| x \| = w, 1 ∈ \text{Supp}(x) \}, \]

\[ S_{w,2}^{n}(ℤ_q^n) = \{ x = (x_1, x_2, x_3) ∈ ℤ_q^{3n} \mid \| (x_1, x_3) \| = w_1, \]

\[ \| x_2 \| = w_1 + w_2, \]

\[ \text{Supp}(x_1, x_3) ⊆ \text{Supp}(x_2) \}. \]

III. AUGMENTED GABIDULIN CODE: A NEW FAMILY OF EFFICIENTLY DECODABLE CODES FOR CRYPTOGRAPHY

In what follows, we introduce a new family of efficiently decodable codes, namely Augmented Gabidulin codes. The main idea behind these codes is to add a sequence of zeros at the end of the Gabidulin codes; by doing this, one directly gets elements of the support of the error, which correspond to support erasure in a rank metric context. More precisely, support erasures are defined as a subspace of the vector space spanned by the coordinates of the error, i.e. the support of the error. The decoding of this code corresponds to the decoding of a classical Gabidulin code to which support erasures are added. In practice, this approach permits to decrease the size of m at the cost of having a probabilistic decoding. The probability of decoding failure can then be easily controlled at the cost of sacrificing only a few support erasures; indeed, for these codes, the decoding failure probability decreases exponentially fast, with a quadratic exponent, see Equation (1).

This approach is then especially suitable in the case where many errors have to be corrected which exactly corresponds to our case where we want to decode has a very low rate. In what follows, we give a definition of augmented Gabidulin codes, and for didactic purpose, in Proposition 1, we recall a simple and natural way to decode Gabidulin code with support erasures and we give their decoding failure rate. For other or more efficient approaches, the reader may refer to [27], [28], and [29].

Notice that this type of approach (adding zeros) is not relevant in Hamming metric since the errors are independent in a classical noisy channel, whereas in rank metric, errors located on different coordinates are linked since they share the same support.

Definition 10 (Augmented Gabidulin codes): Let (k, n, n', m) ∈ ℤ^3 such that k ≤ n' < m < n. Let g = (g_1, . . . , g_m) be an ℤ_q^n-linearly independent family of n' elements of ℤ_q^n and let ℤ_q^n be the vector of length n which is equal to g padded with n - n' extra zeros on the right. The Augmented Gabidulin code \( G^+_ℤ_q(n, n', k, m) \) is the code of parameters \([n, k, m]^q \) defined by

\[ G^+_ℤ_q(n, n', k, m) := \{ P(ℤ_q), \deg_q(P) < k \}, \]

where \( P(ℤ_q) := (P(g_1), . . . , P(g_m), 0, . . . , 0) \).

Proposition 1 (Decoding capacity of Augmented Gabidulin Codes): Let \( G^+_ℤ_q(n, n', k, m) \) be an augmented Gabidulin code, and let

\[ ε ∈ \{1, 2, . . . , \min(n-n', n'-k)\} \]

be the dimension of the vector space generated by the support erasures. Then, \( G^+_ℤ_q(n, n', k, m) \) can uniquely decode an error of rank weight up to

\[ t ::= \left\lfloor \frac{n'-k+ε}{2} \right\rfloor. \]

Proof: The minimal distance of \( G^+_ℤ_q(n, n', k, m) \) is clearly \( d = n'-k+1 \) since it is made of a Gabidulin code augmented with zeros. Let \( x = c_1 + e_1 \) be a noisy codeword where \( c_1 ∈ G^+_ℤ_q \) and \( \| e_1 \| ≤ t \).

Let us assume that \( x \) is not uniquely decodable to find a contradiction. If \( x \) is not uniquely decodable, it means that there exists \( e_2 \) in \( G^+_ℤ_q \) such that \( e_2 ≠ c_1 \) and \( x = c_2 + e_2 \) where \( \| e_2 \| ≤ t \).

Recall that we assume that one knows support erasures which span a vector space of dimension \( ε \). These support erasures come from the \( n-n' \) last coordinates of the code \( G^+_ℤ_q \), thus these support elements are common to \( e_1 \) and \( e_2 \). Since \( \text{Supp}(e_1) \) and \( \text{Supp}(e_2) \) share \( ε \) elements, one has that

\[ d(e_1, e_2) ≤ 2(t-ε) + ε = 2t - ε ≤ n'-k. \]

Since \( x = c_1 + e_1 = c_2 + e_2 \), one clearly has that \( d(c_1, c_2) = d(e_1, e_2) \), thus \( d(c_1, c_2) ≤ n'-k \), which is a contradiction.

Thus, \( G^+_ℤ_q \) can uniquely decode errors of rank weight up to

\[ t ::= \left\lfloor \frac{n'-k+ε}{2} \right\rfloor. \]

Finally, the condition 1 ≤ ε ≤ \( \min(n-n', n'-k) \) comes from the fact that the dimension of the vector space spanned by support erasures can not exceed the maximum rank weight of the error nor the number of zero coordinates of the augmented Gabidulin code; in other words, on one hand \( ε \) is clearly smaller than \( n-n' \), and on the other hand

\[ \frac{n'-k+ε}{2} \Rightarrow 2ε \leq n'-k + ε \Rightarrow ε ≤ n'-k. \]
Proposition 2 (Decoding Algorithm for Augmented Gabidulin Codes): Let $G^\perp_{\mathbb{F}_q}(n, n', k, m)$ be an augmented Gabidulin code, and let
$$
\varepsilon \in \{1, 2, \ldots, \min(n-n', n'-k)\}
$$
be the dimension of the vector space generated by the support erasures.

This code benefits from an efficient decoding algorithm correcting errors of rank weight up to $\delta := \left\lfloor \frac{n'-k+k+1}{2} \right\rfloor$ with a decryption failure rate (DFR) of
$$
1 - q^{\delta(n'-n)} \sum_{i=\varepsilon}^{\delta} \prod_{j=0}^{i-1} (q^\delta - q^j) (q^{n-n'} - q^i). \tag{1}
$$

Proof: The proof gives the decoding algorithm. One is given a noisy encoded word $y = c + \varepsilon \in \mathbb{F}_q^n$, where $c := xG$ belongs to $G^\perp_{\mathbb{F}_q}(n, n', k, m)$ and $\|c\| \leq \varepsilon$. This yields to the probability given by Equation (1).

A. Step 1: Recovering a Part of the Error Support

By construction we have $c = (x, 0, \ldots, 0)$, so that the last $n-n'$ coordinates of $y$ are exactly the last coefficients of $\varepsilon$. Thus, one may use these coefficients to recover $\varepsilon$ elements in $E := \text{Supp}(\varepsilon)$. This will be doable as long as these $n-n'$ coefficients contain at least $\varepsilon$ linearly independent ones. The converse probability is the probability that a random $\delta \times (n-n')$ matrix with coefficients in $\mathbb{F}_q$ has rank less than $\varepsilon$. This yields to the probability given by Equation (1).

B. Step 2: Recovering $c$

Assume now that $\varepsilon$ elements in the support of $c$ are known and let $E_2$ be the vector space spanned by these elements. In what follows, we focus on the first $n'$ coordinates of $\tilde{y}, \tilde{c}$, and $\tilde{e}$ which are denoted by $y, c$ and $e$ respectively. By definition of $G^\perp_{\mathbb{F}_q}(n, n', k, m)$, there exists a $q$-polynomial $P$ of degree at most $k-1$ such that for $1 \leq i \leq n'$:
$$
y_i = P(g_i) + e_i. \tag{2}
$$

Let also $V$ and $V_2$ be the unique monic $q$-polynomials of $q$-degree $\delta$ and $\varepsilon$ which vanish on the vector spaces $E$ and $E_2$ respectively. The ring of $q$-polynomials being left Euclidean, there exists a unique monic $q$-polynomial $W$ of degree $\delta - \varepsilon$ such that $V = W \circ V_2$. As $E_2$ is known, one can easily build the $q$-polynomial $V_2$, for instance using the iterative process described in [23] and [30]. Evaluating $V$ at both sides of Equation (2), one gets $V(y_i) = (V \circ P)(g_i) + V(e_i) = V \circ P(g_i)$. This secret polynomial $P$ can be written symbolically using $\delta - \varepsilon$ unknowns in $\mathbb{F}_q^m$, and similarly we view $R := V \circ P$ as a $q$-polynomial of $q$-degree $k-1+\delta$ with unknown coefficients. Thus, we can derive a linear equation containing $k+2\delta-\varepsilon$ unknowns in $\mathbb{F}_q^m$ from
$$
V(y_i) = R(g_i), \tag{3}
$$
and the same goes for any $i \in \{1, 2, \ldots, n'\}$. Overall, this gives a linear system with $n'$ equations in $k+2\delta-\varepsilon$ variables. This linear system has more equations than unknowns as long as $\delta \leq \left\lfloor \frac{n'-k+k+1}{2} \right\rfloor$, which is the case by assumption. Moreover, this system has a unique solution by Proposition 1. This means that exactly $k+2\delta-\varepsilon$ equations are linearly independent, thus one can solve the system to recover $V$ and $R$, so one finally gets $P$.

IV. NEW RANK-BASED ENCRYPTION SCHEMES

A. Multi-RQC-AG Scheme

Our new encryption scheme denoted Multi-RQC-AG stands for RQC with multiple syndromes. Indeed, it uses several syndromes $U$ and $V$ which differ from the original RQC proposal that relies on unique syndromes $u$ and $v$. As a consequence, our new scheme is based on the IRSL problem which can be seen as a generalization of the IRSD problem used by RQC.

Notation: We start by introducing several sets and operators required to define the Multi-RQC-AG scheme. Let $S_{w_1}^{2n}(\mathbb{F}_q^m)$ and $S_{(w_1, w_2)}^{2n \times 3n}(\mathbb{F}_q^m)$ be defined as:
$$
S_{w_1}^{2n}(\mathbb{F}_q^m) = \{ x = (x_1, x_2) \in \mathbb{F}_q^{2n} | \|x\| = w_1 \in \text{Supp}(x) \},
$$
$$
S_{(w_1, w_2)}^{2n \times 3n}(\mathbb{F}_q^m) = \{ X = (X_1, X_2, X_3) \in \mathbb{F}_q^{2n \times 3n} | \langle X_1, X_3 \rangle = w_1, \langle X_2 \rangle = w_1 + w_2, \text{Supp}(X_1, X_3) \subset \text{Supp}(X_2) \}.
$$

Let $n_1, n_2$ be positive integers such that $n = n_1 \times n_2$, for a vector $v \in \mathbb{F}_q^{n_2 \times n_1}$ and a matrix $M \in \mathbb{F}_q^{n \times n_1}$ whose columns are labelled $M_1, \ldots, M_{n_1}$, we extend the aforementioned dot product such that:
$$
v \cdot M = ((v \cdot M_1^\top \mod P)^\top, \ldots, (v \cdot M_{n_1}^\top \mod P)^\top)
$$
where $v \cdot M \in \mathbb{F}_q^{n_2 \times n_1}$. Let $v = (v_1, \ldots, v_{n_1}) \in \mathbb{F}_q^{n_1}$ with $v_i \in \mathbb{F}_q^m$. If $i \in \{1, \ldots, n_1\}$, the Fold() procedure turns the vector $v$ into a $n_1 \times n_1$ matrix $\text{Fold}(v) = (v_1^\top, \ldots, v_{n_1}^\top) \in \mathbb{F}_q^{n_1 \times n_1}$. The procedure $\text{Unfold}()$ is naturally defined as the converse of $\text{Fold}()$.

Protocol: The Multi-RQC-AG is described on Figure 2. It relies on two codes namely an augmented Gabidulin code $G^\perp_{\mathbb{F}_q}(n, n', k, m)$ that can correct up to $\delta := \left\lfloor \frac{n'-k+k+1}{2} \right\rfloor$ errors using the efficient decoding algorithm $G^\perp_{\mathbb{F}_q}(n, n', k, m)$ as well as a random ideal $[2n_2, n_2]_{\mathbb{F}_q^m}$-code with parity check matrix (I, $\mathbb{F}_q^m$). The correctness of the protocol follows from:
$$
V \cdot y \cdot U = \text{Fold}(mG) + (x + h \cdot y) \cdot R_2 + E
$$
$$
- y \cdot (R_1 + h \cdot R_2)
$$
$$
= \text{Fold}(mG) + x \cdot R_2 - y \cdot R_1 + E.
$$

As a consequence, $\text{Unfold}(V \cdot y \cdot U) = mG + \text{Unfold}(x \cdot R_2 - y \cdot R_1 + E) \in \mathbb{F}_q^m$ which means that $G^\perp_{\mathbb{F}_q}(n, n', k, m)$ as long as:
$$
\| \text{Unfold}(x \cdot R_2 - y \cdot R_1 + E) \| \leq \delta.
$$

Theorem 2: The Multi-RQC-AG scheme depicted in Figure 2 is IND-CPA under the DIRSD and the DHNIRSL assumptions.
\[
\begin{align*}
\text{Setup}(1^\lambda) & : \\
& \text{Generate and output the parameters } \text{param} = (n', n_1, n_2, k, \epsilon, \delta, w, w_1, w_2, P) \text{ where } P \in \mathbb{F}_q[X] \text{ is an irreducible polynomial of degree } n_2. \\
\text{KeyGen}(\text{param}): & \\
& \text{Sample } g \leftarrow S_n^\omega(\mathbb{F}_q), h \leftarrow \mathbb{F}_q^n \text{ and } (x, y) \leftarrow S_{n_1}^\omega(\mathbb{F}_q). \\
& \text{Compute } s = x + h \cdot y \mod P. \\
& \text{Output } pk = (g, h, s) \text{ and } sk = (x, y). \\
\text{Encrypt}(pk, m, \theta): & \\
& \text{Compute } \mathcal{g} = (g | 0 \ldots 0) \in \mathbb{F}_q^{n_1 n_2} \\
& \text{Compute the generator matrix } G \in \mathbb{F}_q^{n \times n_2} \text{ of } \mathcal{g}^+_{n_1, n_2}(n_1, n_2, k, m) \text{ using randomness } \theta. \\
& \text{Sample } (R_1, E, R_2) \leftarrow S_{n_2}^\omega \times \mathbb{F}_q^{n_1} \text{ using randomness } \theta. \\
& \text{Compute } U = R_1 + h \cdot R_2 \text{ and } V = \text{Fold}(mG) + s \cdot R_2 + E. \\
& \text{Output } C = (U, V). \\
\text{Decrypt}(pk, sk, C): & \\
& \text{Output } m = \mathcal{g}^+_{n_2} \cdot \text{Decode}(\text{Unfold}(V - y \cdot U)). \\
\end{align*}
\]

Fig. 2. Multi-RQC-AG encryption scheme.

\textbf{Proof:} The proof of the Multi-RQC-AG scheme is similar to the proof from [3] with an IRSD \((m, 2n_2, n_2, \omega)\) instance defined from a \([2n_2, n_2]\) code and an NHIRSL \((m, n_2, n_2, \omega_1, \omega_2, n_1)\) instance defined from a \([3n_2, n_2]\) code. These instances are defined by the following products:

\[
\begin{align*}
(I_{n_2} & \quad IM(h)) \times (x, y)^T = s^T, \\
(I_{n_2} & \quad 0) \quad IM(h) \times \begin{pmatrix} R_1 \\ E \\ R_2 \end{pmatrix} = \begin{pmatrix} U \\ V - \text{Fold}(mG) \end{pmatrix}.
\end{align*}
\]

B. Multi-UR-AG Scheme

Our new encryption scheme denoted Multi-UR-AG stands for Multiple syndromes Unstructured Rank with Augmented Gabidulin codes encryption scheme. It is particularly interesting security wise as it does not use structured codes contrarily to existing constructions such as ROLLO, RQC or our new proposal Multi-RQC-AG. Indeed, it only relies on the security of the RSL problem. Multi-UR-AG leverages multiple syndromes and augmented Gabidulin codes. In addition, it features two variants as it can be instantiated with either homogeneous or non-homogeneous errors.

\textbf{Notation:} Hereafter, \text{Fold} and \text{Unfold} refer to the procedure introduced in Section IV-A. Let \(S_{n_1}^{n \times 2n_1}(\mathbb{F}_q^m)\) and \(S_{n_2 \times (n+n_1+n^2)}(\mathbb{F}_q^m)\) be defined as:

\[
S_{n_1}^{n \times 2n_1}(\mathbb{F}_q^m) = \{ X = (X_1, X_2, X_3) \in \mathbb{F}_q^{n \times 2n_1}, |X| = w, 1 \in \text{Supp}(X) \},
\]

\[
S_{n_2 \times (n+n_1+n^2)}(\mathbb{F}_q^m) = \{ X = (X_1, X_2, X_3) \in \mathbb{F}_q^{n \times (n+n_1+n^2)}, |X| = w, 1 \in \text{Supp}(X) \}.
\]

\textbf{Setup}(1^\lambda) & : \\
& \text{Generate and output the parameters } \text{param} = (n, n', n_1, n_2, k, \epsilon, \delta, w, w_1, w_2). \\
\text{KeyGen}(\text{param}): & \\
& \text{Sample } g \leftarrow S_n^\omega(\mathbb{F}_q), H \leftarrow \mathbb{F}_q^{n \times n}, (X, Y) \leftarrow S_{n \times n^2}(\mathbb{F}_q^m). \\
& \text{Compute } S = X + HY. \\
& \text{Output } pk = (g, H, S) \text{ and } sk = (X, Y). \\
\text{Encrypt}(pk, m, \theta): & \\
& \text{Compute } g = (g | 0 \ldots 0) \in \mathbb{F}_q^{n_1 n_2} \\
& \text{Compute the generator matrix } G \in \mathbb{F}_q^{n \times n_2} \text{ of } \mathcal{g}^+_{n_1 n_2}(n_1, n_2, n', k, m) \text{ using randomness } \theta. \\
& \text{Sample } (R_1, E, R_2) \leftarrow S_{n_2}^\omega \times \mathbb{F}_q^{n_1} \text{ using randomness } \theta. \\
& \text{Compute } U = R_1 + R_2 H \text{ and } V = \text{Fold}(mG) + R_2 S + E. \\
& \text{Output } C = (U, V). \\
\text{Decrypt}(pk, sk, C): & \\
& \text{Output } m = \mathcal{g}^+_{n} \cdot \text{Decode}(\text{Unfold}(V - UY)). \\
\end{align*}
\]

Fig. 3. Multi-UR-AG (with non-homogeneous errors) encryption scheme.

\textbf{Protocol:} The Multi-UR-AG is described on Figure 3. It relies on two codes namely an augmented Gabidulin code \(\mathcal{g}^+_{n}(n', n_1, k, m)\) that can correct up to \(\delta := \frac{n'-k+\epsilon}{2}\) errors using the efficient decoding algorithm \(\mathcal{g}^+_{n} \cdot \text{Decode}()\) as well as a random \([2n, n]\) code with parity check matrix \((I \ H)\). The correctness of the protocol follows from:

\[
V - UY = \text{Fold}(mG) + R_2(X + HY) + E \\
- (R_1 + R_2 H)Y \\
= \text{Fold}(mG) + R_2 X - R_1 Y + E.
\]

As a consequence, \text{Unfold}(V - YU) = mG + \text{Unfold}(XR_2 - YR_1 + E) \in \mathbb{F}_q^m \text{ which means that } \mathcal{g}^+_{n} \cdot \text{Decode}(\text{Unfold}(V - YU)) = m \text{ as long as:}

\[
\| \text{Unfold}(XR_2 - YR_1 + E) \| \leq \delta.
\]

\textbf{Theorem 3:} The Multi-UR-AG scheme is IND-CPA under the DRSL and the DNHRLS assumptions.

\textbf{Proof:} The proof of the Multi-UR-AG scheme is similar to the proof from [3] with an \(\text{RSL}(m, 2n, n, \omega, n_1)\) instance defined from a \([2n, n]\) code and an \(\text{NHIRSL}(m, n, n_1, \omega_1, \omega_2, n_2)\) instance defined from a \([2n + n_1, n]\) code.
These instances are defined by the following products:

\[(I_n \times H) \times (X \times Y) = S,\]

\[(R_1 \times E \times R_2) \times \begin{pmatrix} I_n & 0 \\ 0 & I_{n_1} \\ H & S \end{pmatrix} = (U, V - \text{Fold}(mG)).\]

V. Security Analysis

In this section, we provide the complexity to solve some hard problems in rank-based cryptography.

A. Attacks on the RSD Problem [18], [31], [32]

There are two general classes of attacks to RSD, based on combinatorial or algebraic techniques. On the one hand, combinatorial attacks can be seen as the equivalent of ISD-type attacks in the rank metric setting. Relying on [31] and [32], we estimate that the complexity of the best combinatorial attack is in

\[
\min\left(2^{(w-1)\left\lceil \frac{(k+1)m}{w} \right\rceil}, 2^w\left\lceil \frac{(k+1)m}{w} \right\rceil - m\right)
\]

(4)

\[F_q\)-operations. On the other hand, algebraic attacks on the RSD problem are by modeling the decoding instance into a system of polynomial equations, and the overall cost is reduced to the one of solving this system. To design our parameters, we take into account the most recent algebraic attack, namely the MaxMinors attack [18]. Its complexity in \(F_q\) operations is estimated to be

\[
\mathcal{O}\left(q^a m^{(n-k-1)}(n-w)\omega^{-1}\right),
\]

(5)

where \(a \geq 0\) the smallest integer such that \(m^{(n-k-1)} \geq (n-w)^a - 1\) and where \(\omega\) is a linear algebra constant.

B. Attacks on the NHRS RSD Problem

This section is dedicated to the first cryptanalysis of the NHRS RSD problem by proposing two attacks which exploit the inhomogeneous structure of the error.

1) A New Combinatorial Attack: Given an error vector \(e = (e_2, e_1, e_3)\), let \(S_1 := \text{Supp}(e_1, e_3)\) and let \(S_2 := \text{Supp}(e_2)\), which means that \(S_1\) (resp. \(S_2\)) is a vector space of dimension \(w_1\) (resp. \(w_1 + w_2\)). In this section, we may assume for clarity that the \(n_1\) leftmost coordinates of \(e\) correspond to the part of weight \(w_1 + w_2\), namely \(e = (e_2, e_1, e_3)\), and we also adopt a systematic form for the parity-check matrix \(H_e = (I_{n+n_1-1} - e)\) of the public code \(C_e := C \oplus \langle e \rangle\). The parity-check equations for this code which are traditionally used in this type of attack are as follows:

1) those associated to the \(n\) first rows of \(H_e\) provide \(n\) linear relations over \(F_q^m\) which can be mapped into \(nm\) relations over \(F_q\) between unknowns coming from \(e_2, e_1\) and \(e_3\).

2) those associated to the \(n_1 - 1\) last rows of \(H_e\) give \((n_1 - 1)m\) equations over \(F_q\) in unknowns coming from the components of \(e_1\) and \(e_3\) only.

Before describing our attack, let us recall how [31], [32] would solve a non-structured RSD\((m, 2n, n, w_1)\) instance to recover \((e_1, e_3)\). The most enhanced version of [32] consists in guessing a subspace \(V\) of dimension \(r_1 \geq w_1\) such that \(\alpha S_1 \subset V\) for some element \(\alpha \in F_q^m\) instead of simply \(S_1 \subset V\) as it provides a better success probability. Then, it aims at solving the linear system given by the parity-check equations from 2. The largest value of \(r_1\) for which one may expect a unique solution is given by

\[r_1 := \left\lceil \frac{m(n-1)}{2n} \right\rceil = m - \left\lceil \frac{m(n+1)}{2n} \right\rceil.
\]

(6)

The classical cost given in [32] is then roughly

\[\tilde{O}\left(q^{w_1(m-r_1)-m}\right) = \tilde{O}\left(q^{w_1(m+n+1) - m}\right),
\]

(7)

where \(\tilde{O}\) hides a polynomial factor corresponding to solving this linear system. To benefit from the inhomogeneous structure of \(e\) from NHRS RSD, our approach follows the natural path of making a guess on a random subspace \(V\) of dimension \(r \geq w_1\) such that \(S_1 \subset V\) and a random subspace \(Z \subset F_q^m/V\) of dimension \(\rho \in \{w_2..m-r\}\) such that \(S_2 \subset V \oplus Z\).

Theorem 4: Our proposed combinatorial algorithm runs in time

\[\tilde{O}\left(q^{w_1+w_2(m-r)-w_2\rho-m}\right).
\]

(8)

The complexity given by Equation (8) is of the same shape as Equation (7) since the rest of our attack is totally similar to [31] and [32]: expressing the coordinates of \((e_1, e_3)\) in a fixed basis of \(V\) yields \(2nr\) variables over \(F_q\), while we get \(n_1(r+\rho)\) variables over \(F_q\) by writing the coordinates of \(e_2\) in a fixed basis of \(V \oplus Z\). For the linear algebra step, \(n_1(r+\rho)\) random equations from 1. are used in order to express all the variables from \(e_2\) in terms of the other variables, and we are left with a linear system of \((n+n_1-1)m-n_1(r+\rho)\) equations over \(F_q\) in only \(2nr\) variables. This leads to the condition

\[2nr \leq m(n+n_1-1) - n_1(r+\rho)
\]

in order to expect at most one solution. Overall, the main task to prove Theorem 4 is to compute the success probability

\[\Pr_{V, Z} \left[3a \in F_q^m, \alpha S_1 \subset V, \alpha S_2 \subset V \oplus Z\right] \approx \frac{q^m-1}{q-1},
\]

(9)

and in case of success decoding the word \(\alpha e\) instead of \(e\). For clarity Appendix A presents the plain version of the attack, but as this trick is compatible with our analysis the corresponding \(q^{-m}\) factor appears in Equation (8). Finally, one has to consider the couple \((r, \rho)\) which leads to the best exponent in Equation (8). In other words, the goal will be to maximize the quantity \((w_1+w_2)r+w_2\rho\) under the constraints \((2n+n_1)r+n_1\rho \leq m(n+n_1-1), w_1 \leq r, w_2 \leq \rho, r+\rho \leq m-1\), where \(r, \rho \in \mathbb{N}\).
This is an example of integer linear program (ILP), and to solve this instance we have used dedicated tools.

2) Adaptation of the Algebraic Attack of [18] Against NHRSR: A first approach of this attack was proposed in [3], we build upon this work and give a thorough analysis of the complexity of this attack.

**Theorem 5:** Let \( a \geq 0 \) the smallest integer such that

\[
N_{\mathcal{F}, q} \geq (2n + n - 1 - a) - M_1 - \nu_{\mathcal{F}, q} - 1,
\]

where

\[
N_{\mathcal{F}, q} = m \sum_{i=1}^{w_1+w_2} (n_i-1)(n_i) - \nu_{\mathcal{F}, q} = m^{(n_i-1)} \nu_{\mathcal{F}, q} - M_a - \nu_{\mathcal{F}, q} - 1.
\]

The hybrid MaxMinors attack adapted to NHRSR costs

\[
O \left( q^{w_1+w_2}(2n + n - 1 - a) - M_1 - \nu_{\mathcal{F}, q} - 1 \right)
\]

operations in \( \mathbb{F}_q \), where \( \omega \) is a linear algebra constant.

a) MaxMinors linear system [18]: The MaxMinors system is a system of equations over \( \mathbb{F}_q \), which vanish on the solutions to the RSD instance. Let \( y = c + e \in \mathbb{F}_q^{w_1+n_1} \) be the noisy codeword to be decoded in a random \( \mathbb{F}_q \)-linear code \( C \) of length \( 2n + n_1 \) and dimension \( n \) with generator matrix \( G \in \mathbb{F}_q^{n \times (2n + n_1)} \). The extended code \( C_1 = C_2 \) is generated by the matrix \( G_y \in \mathbb{F}_q^{w_1+n_1} \) a full-rank parity-check matrix for this code. We clearly have

\[
0 = eH_y^T = \beta \text{Mat}(e)H_y^T = \beta SC\beta H_y^T,
\]

so that the matrix \( \beta \text{Mat}(e)H_y^T \) contains a non-zero vector \( \beta \text{S} \) in its left kernel and cannot be full-rank. In particular, the MaxMinors system is the system of maximal minors \( \mathcal{P} := \{ P_J \}_{J} \) such that \( P_J := CH_y^T_{e,J} \) for each subset \( J \subset \{ 1..n + n_1 - 1 \} \), \#\( J \) \( = w_1 + w_2 \). The crux is that these equations are actually linear in the minor variables \( c_T := (C_1^e)_{J,T} \in \mathbb{F}_q \) by using the Cauchy-Binet formula for the determinant of a product of rectangular matrices, see [17], [18]. In this section, the \( c_T \)'s will be sorted with respect to the following ordering on the \( T \)-s: we consider that \( T = \{ t_1 < \cdots < t_p \} < T' = \{ t_1' < \cdots < t_p' \} \) if \( t_j = t_j' \) for \( j < j_0 \) and \( t_j < t_j' \) \( \Rightarrow \) assuming that \( 1 < 2 \cdots < n \). We will further assume that \( H_y := (*) I_{n + n_1 - 1} \) and from that assumption [18] derive the fundamental Lemma 1 on the shape of the MaxMinors equations:

**Lemma 1 (Prop. 2, [18]):**

\[
P_J = c_{J+n+1} + \sum_{T \subset \{ 1..n+1 \}, \#T = w_1 + w_2} c_T |H_y|_{J,T}.
\]

A direct consequence of Lemma 1 is that the equations of \( \mathcal{P} \) are linearly independent over \( \mathbb{F}_q \) as their leading terms are distinct.

b) Removing variables corresponding to zero minors: The very same MaxMinors system can be employed to attack NHRSR. A main difference in this case is that if one wants to decrease the number of minor variables by relying on the special structure of \( e \) as shown in [3] and [18], then linear relations between the equations after removing these variables also occur and must be taken into account in the analysis.

Recall from [3, 6.2.2] that the row support of \( \text{Mat}(e) \in \mathbb{F}_q^{n \times (n + n_1)} \) can be written as

\[
C = \begin{pmatrix} C_1 & C_2 \\ 0 & 0 \end{pmatrix} \in \mathbb{F}_q^{w_1+w_2 \times (n + n_1 + n)},
\]

where \( C_1, C_2 \in \mathbb{F}_q^{w_1 \times n}, C_2 \in \mathbb{F}_q^{n \times n_1} \) and \( C_3 \in \mathbb{F}_q^{n_1 \times n_1} \). From Equation (11), it has been noted that the minors \( |C|_{n \times 2} \) such that \( T \cap \{ n+1..n + n_1 \} \leq w_2 - 1 \) are always zero. This means that the

\[
M := \sum_{i=0}^{w_2-1} \left( \begin{array}{c} n_1 \\ i \end{array} \right) \left( \begin{array}{c} 2n \\ w_1 + w_2 - i \end{array} \right)
\]

variables from the set

\[
\zeta := \{ c_{T}, T \subset \{ 1..(n+1) \}, \#T = w_1 + w_2, T \cap \{ n+1..n + n_1 \} \leq w_2 - 1 \}
\]

can be set to zero in the MaxMinors system. It is then relevant to separate the initial \( P_J \) equations into several subsets in function of the presence or the absence of these \( c_T \) variables. We consider the partition \( \mathcal{P} := \mathcal{P}_{\text{lost}} \sqcup \mathcal{P}_{\text{residual}} \sqcup \mathcal{P}_{\text{independent}} \), where

\[
\mathcal{P}_{\text{lost}} := \{ P_J : \#J \leq w_1 + w_2, \#(J \cap \{ 1..(n+1) \}) \leq w_2 - 2 \},
\]

\[
\mathcal{P}_{\text{residual}} := \{ P_J : \#J = w_1 + w_2, \#(J \cap \{ 1..(n+1) \}) = w_2 - 1 \},
\]

\[
\mathcal{P}_{\text{independent}} := \{ P_J : \#J = w_1 + w_2, \#(J \cap \{ 1..(n+1) \}) \geq w_2 \}.
\]

Using Lemma 1, it is easy to grasp the shape of the equations from \( \mathcal{P}_{\text{lost}} \) and \( \mathcal{P}_{\text{independent}} \) after removing the minor variables belonging to \( \zeta \).

**Proposition 3:** After setting the minor variables from \( \zeta \) to zero in the MaxMinors system \( \mathcal{P} \), we have the following properties:

1) The equations in \( \mathcal{P}_{\text{lost}} \) all become zero.
2) The equations in \( \mathcal{P}_{\text{independent}} \) keep the same leading terms and therefore they are still linearly independent.

We have

\[
\dim_{\mathbb{F}_q}(\mathcal{P}_{\text{independent}}) = \#\mathcal{P}_{\text{independent}} = \sum_{i=0}^{w_1+w_2} \left( \begin{array}{c} n_1 - 1 \\ i \end{array} \right) \left( \begin{array}{c} n \\ w_1 + w_2 - i \end{array} \right).
\]

Finally, the system \( \mathcal{P}_{\text{independent}} \) contains at most \( \left( \begin{array}{c} 2n + n_1 + n \\ n_1 \end{array} \right) - M \) variables.

**Proof:** See Appendix C-A.
the $M$ minor variables to zero. More precisely, by Lemma 1, an equation $P_J \in \mathcal{P}_{\text{rest}}$ becomes
\[
P_J = \sum_{T^- \subset \{1, \ldots, n\}, T^+ \subset (J+n+1), \#(T^- \cap (n+2, \ldots, n+1)) = w_2-1} c_T \mathbf{H}_y|_{J,T}
\]
\[
= \sum_{T^- \subset \{1, \ldots, n\}, T^+ \subset (J+n+1), \#(T^- \cap (n+2, \ldots, n+1)) = w_2-1} c_T \mathbf{H}_y|_{J,T}. \tag{13}
\]

For clarity, we still denote the resulting system by $\mathcal{P}_{\text{rest}}$. We analyze it in the following Proposition 4.

**Proposition 4:** After setting the minor variables from $\zeta$ to zero in $\mathcal{P}_{\text{rest}}$, one obtains a system of rank $(n_1-1)(n_1-1)$ and whose equations are also independent from $\mathcal{P}_{\text{indep}}$. Finally, these equations contain at most $(n_2-1)(2n)$ variables.

The first part of Proposition 4 is obvious. Using Equation (13), the leading term of $P_J \in \mathcal{P}_{\text{rest}}$ is a $c_T$ variable such that $n + 1 \in T$, whereas the leading term of any $P_{J'} \in \mathcal{P}_{\text{indep}}$ is $c_{J' + n + 1}$ and $c_{J' + n + 1} > c_T$ for any such $T$. Thus, what is left to prove in Proposition 4 is that $\dim_{\mathbb{F}_q'(n_1-1)(n_1-1)}$ and that the number of variables is $(n_2-1)(2n)$. For this we rely on the following lemma, whose proofs can be found in Appendix C-B.

**Lemma 2:** For $A \subset \{n+2, \ldots, n+n_1\}$, $\#A = w_1-1$, let $\mathcal{P}_{\text{rest}, A} := \{P_J \in \mathcal{P}_{\text{rest}} : J \cap \{1, \ldots, n_1-1\} = A \}$, so that $\mathcal{P}_{\text{rest}, A}$ is a partition of $\mathcal{P}_{\text{rest}}$. We have $\langle \mathcal{P}_{\text{rest}, A} \rangle = \oplus A (\mathcal{P}_{\text{rest}, A})$.

**Lemma 3:** For $A \subset \{n+2, \ldots, n+n_1\}$, $\#A = w_2-1$, let $\mathcal{P}_{\text{rest}, A}$ as defined in Lemma 2. With very high probability, we have $\dim_{\mathbb{F}_q'(n_1-1)}$.

c) **Finishing the attack by projecting over $\mathbb{F}_q'$.** The last step of the initial MaxMinors attack on RSD is by solving the “projected” linear system $\mathcal{P}_{\text{rest}} := \{P_{J,j}, j, J\}$ obtained by expressing the coefficients of the $P_J$’s in a fixed basis of $\mathbb{F}_q^m$ over $\mathbb{F}_q$ and taking each component, yielding $m$ more equations. We proceed in a very similar way as in [18] and due to space constraints we do not recall all the details of this step. Our final complexity estimate relies on the following assumption.

**Assumption 1:** Let $\mathcal{P}_{\text{indep,}\mathbb{F}_q}$ (resp. $\mathcal{P}_{\text{rest,}\mathbb{F}_q}$) be the system over $\mathbb{F}_q$ obtained by projecting $\mathcal{P}_{\text{indep}}$ (resp. $\mathcal{P}_{\text{rest}}$) where the variables in $\mathbb{z}$ had already been removed, let $N_{\mathbb{F}_q} := \dim_{\mathbb{F}_q'}(\mathcal{P}_{\text{indep,}\mathbb{F}_q})$, let $\nu_{\mathbb{F}_q} := \dim_{\mathbb{F}_q'}(\mathcal{P}_{\text{rest,}\mathbb{F}_q})$ and let $M$ as defined in Equation (12). We assume that
\[
N_{\mathbb{F}_q} = m \dim_{\mathbb{F}_q'}(\mathcal{P}_{\text{indep}}) = m \sum_{w_1+w_2} (n_1-1) \binom{n}{w_1+w_2} - M - 1, \tag{14}
\]
when this value is $\leq \left(\frac{2^{n_1}+1}{2^{n_1}+2}\right) - M$ and $N_{\mathbb{F}_q} = \left(\frac{2^{n_1}+1}{2^{n_1}+2}\right) - M$ otherwise, and
\[
\nu_{\mathbb{F}_q} = m \dim_{\mathbb{F}_q'}(\mathcal{P}_{\text{rest}}) = m \left(\frac{n_1-1}{n_1-2}\right) \binom{n}{w_1}, \tag{15}
\]
provided that this value is $\leq \left(\frac{n_1-1}{n_2-2}\right) \binom{2n}{w_1}$.

To solve the final system, one can start by performing linear algebra on $\mathcal{P}_{\text{rest,}\mathbb{F}_q}$ and then substitute $\nu_{\mathbb{F}_q}$ variables corresponding to an echelonized basis of $\langle \mathcal{P}_{\text{rest,}\mathbb{F}_q} \rangle$ in the system $\mathcal{P}_{\text{indep,}\mathbb{F}_q}$ to get a new system $\mathcal{P}'_{\text{indep,}\mathbb{F}_q}$. The final step is then to solve the linear system $\mathcal{P}'_{\text{indep,}\mathbb{F}_q}$ in $(\frac{2^{n_1}+1}{2^{n_1}+2}) - M$ variables.

**Corollary 2 (Same Notation as in Assumption 1):** Let $\mathcal{P}_{\text{indep,}\mathbb{F}_q}$ and let $\mathcal{P}_{\text{rest,}\mathbb{F}_q}$ denote the projected systems from Assumption 1. We consider $\mathcal{P}'_{\text{indep,}\mathbb{F}_q}$, the linear system obtained from $\mathcal{P}_{\text{indep,}\mathbb{F}_q}$ after plugging $\nu_{\mathbb{F}_q}$ equations from the echelon form of $\mathcal{P}_{\text{rest,}\mathbb{F}_q}$ to substitute variables. Assuming that the system $\mathcal{P}'_{\text{indep,}\mathbb{F}_q}$ can be solved, namely $N_{\mathbb{F}_q} \leq \left(\frac{2^{n_1}+1}{2^{n_1}+2}\right) - M - \nu_{\mathbb{F}_q} - 1$, the complexity of solving the system is
\[
O\left(N_{\mathbb{F}_q} \left(\frac{2^{n_1}+1}{2^{n_1}+2}\right) - M - \nu_{\mathbb{F}_q} - 1\right)
\]
operations in $\mathbb{F}_q$, where $\omega$ is a linear algebra constant.

However, the projected linear system cannot be solved directly when there are not enough equations compared to the number of minor variables, i.e. $N_{\mathbb{F}_q} < \left(\frac{2^{n_1}}{2^{n_1}+2}\right) - M - \nu_{\mathbb{F}_q} - 1$. In this case, a method suggested in [18] is an hybrid approach by adding linear constraints on these minor variables which are obtained by fixing the entries of $a \geq 0$ columns in the matrix $C$. Here, like it was done in [3], it is possible to take advantage of the particular structure of $C$ given in Equation (11) by fixing columns containing only $w_1$ non-zero coordinates, which leads to a smaller exponential factor of $q^{\omega w_1}$ in the final cost instead of the naive $q^\omega (w_1+w_2)$. The cost claimed in Theorem 5 follows.

### C. Attacks on the RSL Problem

In this section, we consider an RSL $(m, n, k, r, N)$-instance, say $N$ distinct RSD instances whose errors share the same support of dimension $r$. This number $N$ is a crucial parameter to estimate the hardness of RSL and in particular to compare it to RSD. For instance, this problem can be solved in polynomial time when $N \geq nr$ due to [4]. A more powerful attack was later found in [19] and it suggests that secure RSL instances must satisfy a stronger condition: $N < kr$.

In what follows, we give a new combinatorial attack against RSL, it is more efficient than the previous combinatorial attacks, plus it enables us to decrease the threshold where the RSL problem starts to be solvable in polynomial time. In addition to this, we give more explicit formulas to clarify the recent algebraic attack of [20].

1) **New Combinatorial Attack on RSL**

**Theorem 6 (Combinatorial Attack on RSL):** There exists a combinatorial attack on RSL $(m, n, k, r, N)$ with complexity
\[
O\left(q^\omega (m-\frac{n(n-k)-N}{n-k})\right)
\]
on operations in $\mathbb{F}_q$.

**Proof:** Let $s_i \in \mathbb{F}_q^m, 1 \leq i \leq N$ denote the $N$ syndromes from the RSL instance. By definition there exist $e_i \in \mathbb{F}_q^m, \|e_i\| = r, H e_i^T = s_i^T$, where $H \in \mathbb{F}_q^{(m-k) \times n}$ is a parity-check matrix and where $\text{Supp}(e_i)$ does not depend on $i$. Similarly to [20] and [33], this last property enables us to use the fact that there exists an $\mathbb{F}_q$-linear combination $(e_a | \overline{e}) \in \mathbb{F}_q^m$ of the $e_i$’s which is all-zero on its first
a := \left\lceil \frac{N}{r} \right\rceil \text{ coordinates. This error corresponds to a secret linear combination of the syndromes, more precisely}
\exists \lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{F}_q, \quad H(0_a | \bar{e})^T = \sum_{i=1}^{N} \lambda_i s_i^T.

By setting \( \bar{H} := H_{*,(a+1\ldots n)} \), this is equivalent to
\[ \bar{H} e^T = \sum_{i=1}^{N} \lambda_i s_i^T. \] (16)

Equation (16) can be seen as a \( n - k \) parity-check equations which may be exploited by the classical combinatorial technique, see [31], [32] or the discussion above Equation (6).

The main difference here is that the right hand this equation also contains \( N \) unknowns \( \lambda_i \in \mathbb{F}_q \). Still, we can pick a vector space \( V \) of dimension \( r_1 \geq r \) and hoping that \( \text{Supp}(\bar{e}) \subset V \).

If this is the case, one can derive from (16) a linear system of \( (n - k)m \) equations over \( \mathbb{F}_q \) in \( N + (n - a)r_1 \) variables, where the first \( N \) variables merely correspond to the \( \lambda_i \)'s.

The final cost is then obtained by looking at the optimal value of \( r_1 \) which allows to solve this linear system, namely \( r_1 = \min \left( \frac{m(n-k)-N}{n-a}, m \right) \) (as a value above \( m \) would not make sense).

It is clear from the proof of Theorem 6 that the attack is polynomial if and only \( \min \left( \frac{m(n-k)-N}{n-a}, m \right) = m \).

By recalling that \( a = \left\lceil \frac{N}{r} \right\rceil \), this is equivalent to
\[ \frac{m(n-k)-N}{n-a} \geq m \iff m(a-k) \geq N \iff \left\lceil \frac{N}{r} \right\rceil - N \geq m \geq k. \] (17)

It readily implies \( \frac{N}{r} - \frac{N}{m} \geq k \), hence \( N \geq kr \frac{m-r}{m-r} \). However, the converse is not true as \( N \geq kr \frac{m-r}{m-r} \) does not always imply Equation (17). A sufficient condition for it to hold is \( \frac{N}{r} - 1 - \frac{N}{m} \geq k \), hence \( N \geq (k+1)r \frac{m-r}{m-r} \).

Proposition 5: The proposed combinatorial technique on an RSL instance with parameters \((m, n, k, r, N)\) such that \( N \geq (k+1)r \frac{m-r}{m-r} \) is expected to take polynomial time.

2) Algebraic Attack of [20]: This attack consists in solving a bilinear system at some bi-degree \((b, 1)\) for \( b \geq 1 \) by using an XL approach similar to [18]. The two cases \( \delta = 0 \) and \( \delta > 0 \) presented below correspond to two different specializations of this bilinear system which lead to different costs. Here, we provide explicit formulas to compute these two complexities (for the binary field \( \mathbb{F}_2 \)). In particular, we also include the values of \( \alpha_R \) and \( \alpha_\lambda \) which correspond to the hybrid approach mentioned in [20]. Finally, note that these formulas are valid only when \( N > n - k - r \).

a) First case: \( \delta = 0 \): Let \( a \) be the unique integer such that \( ar \leq N \leq (a+1)r \), and let \( N' := ar + 1 \). For \( 1 \leq b \leq r + 1 \), the number of variables for linearization is
\[ M^{\delta = 0}_{\leq b} := \sum_{i=1}^{b} \left( n - a - \alpha_R \right) \left( N' - \alpha_\lambda \right), \] (18)
where \( 0 \leq \alpha_R < n - a - r \), and \( 0 \leq \alpha_\lambda < N' - b \), and the number of linearly independent equations at hand is equal to

\[ mN^{\delta = 0}_{\leq b} \]
where
\[ N^{\delta = 0}_{\leq b} := \sum_{i=1}^{b} \sum_{d=1}^{n-k} \sum_{j=1}^{r} \left( d - 1 \right) \left( r - d + 1 \right) \left( N' - \alpha_\lambda - j \right). \] (19)

The complexity is given by
\[ O \left( \min \left( 2^{\alpha_R + \alpha_\lambda} mN^{\delta = 0}_{\leq b} \left( M^{\delta = 0}_{\leq b} \right)^{-1}, \right. \right) \]
\[ \left. 2^{\alpha_R + \alpha_\lambda} \left( N' - \alpha_\lambda \right) \left( k - a + 1 + r \right) \left( M^{\delta = 0}_{\leq b} \right)^{1/2} \right) \] (20)
provided that \( mN^{\delta = 0}_{\leq b} \geq M^{\delta = 0}_{\leq b} - 1 \), and where the values of \( b, \alpha_R, \) and \( \alpha_\lambda \) are chosen to minimize the complexity.

b) \( \delta > 0 \) case: Let \( \delta \) be a positive integer such that \( N \geq \delta (n - r + \delta) \), let \( a \) be the greatest integer such that \( N \geq (k+1)r \frac{m-r}{m-r} \) and let \( N' := \delta (n - r + \delta) + a (r - \delta) \) and let \( N' := \delta (n - r + \delta) + a (r - \delta) \). To find the complexity of this attack, one replaces \( r \) by \( r - \delta \) in the expressions of \( M^{\delta = 0}_{\leq b} \) and \( N^{\delta = 0}_{\leq b} \) from Equations (18) and (19). The complexity is finally obtained with Equation (20) and its minimal value now depends on \( \delta > 0 \) as well as \( b, \alpha_R, \) and \( \alpha_\lambda \) as above.

3) Visualization of the Attacks Against RSL: Last but not least, thanks to our analysis of the complexity to solve RSL with different attacks, we were able to draw a graph, see Figure 4, of the complexity to solve an RSD instance as a function of the number of given syndromes \( N \).

The instance parameters are \([m, n, k, r] = [61, 100, 50, 7]\), this is precisely the instance corresponding to attacking our scheme NH-Multi-RQC-AG-128 (see Table I). The complexity to solve this RSD instance using the algebraic attack MaxMinors (see Section V-A) is 196 bits; it corresponds to
the horizontal black thick line. Starting with 44 syndromes; recall that it is the threshold for the algebraic attack against RSL (see Section V-C), one sees that it beats the RSD attack. It is worth noticing that with approximately 225 syndromes, our new combinatorial attack against RSL (see Theorem 6), starts to beat the algebraic attack of [20]. And finally, one notices that, with a lot of syndromes, all the aforementioned RSL attacks complexities drop down, which is quite logical.

D. Combinatorial Attack on NHRSL

In this section, we adapt the combinatorial attack against RSL, given in the proof of Theorem 6, to the case of non-homogeneous error, i.e. to the NHRSL problem (see Problem 4).

For the sake of simplicity, and since it is the case for all cryptographic parameters studied in this paper, we focus only on NHRSL instances where \( n_1 < n \).

Theorem 7 (Combinatorial Attack Against NHRSL):

There exists a combinatorial attack against an NHRSL instance with parameters \((m, n, n_1, w_1, w_2)\) whose complexity, in terms of elementary operations in \( \mathbb{F}_q \), is given by

\[
\tilde{O} \left( q^{(w_1+w_2)(m-r)-w_2^2} \right),
\]

where \( r, \rho \) are integers chosen to maximize the quantity \((w_1+w_2)r + w_2^2 \rho \) under the following constraints: \( N_1, N_2, r, \rho \in \mathbb{N} \), \( N_1 + N_2 = N \), \( w_1 \leq r \), \( w_2 \leq \rho \), \( r + \rho \leq m - 1 \), \( a := \left[ \frac{N_2}{w_1+w_2} \right] \leq n_1 \), \( b := \left[ \frac{N_2}{w_1+w_2} \right] \leq 2n \), \( m(n + n_1) \geq (n_1 - b)(r + \rho) + (2n - a)r + N \).

Proof: Straightforward adaptation of the attack in the proof of Theorem 6, combined with the probability results given in Appendix A.

VI. SECURITY AND PARAMETERS OF OUR SCHEMES

A. Security Comparison for Our Schemes

According to Theorem 2, the security of Multi-RQC-AG relies on the Decisional Ideal Rank Syndrome Decoding problem (DIRSD) and on the Decisional Ideal Non-Homogeneous Rank Support Learning problem (DNIHRSL). So far, there is no known attack to solve the decisional versions of these problems without solving the associated search instances. In addition to this, there is currently no attack that takes advantage of the ideal structure; thus, studying the security of Multi-RQC-AG comes down to evaluating the complexity of RSD and NHRSL. Unlike Multi-RQC-AG, our new scheme Multi-UR-AG does not use ideal structure. Despite the aforementioned absence of attack that exploits ideal structure, it might induce a weakness in a scheme. This is why Multi-UR-AG, which does not use any structure like its name suggests it, is more secure. To study its complexity, according to Theorem 3, one has to study RSL and NHRSL. However, for an even better security, one could use Multi-UR-AG with homogeneous weight, making its security relying solely on RSL (see for instance the parameters sets Multi-UR-AG -128 and Multi-UR-AG -192 in Section VI).

B. Examples of Parameters

Parameters proposed (see Table I) for combinatorial attacks are compliant with NIST security levels 1 and 3 of 143 and 207 classical bit security. The quantum complexity of combinatorial attacks has been considered in ([34] and [35]), in terms of algebraic attacks no quantum improvement is known to the best of our knowledge. At the end of this section, we recall what these parameters correspond to, we do so by describing two sets in details together with their associated attacks.

Among the different codes that can be attacked for each of our schemes (see proofs of Theorems 2 and 3), there is not a weaker one which enables us to fix all of our parameters. More precisely, sometimes attacking the public key, i.e. a code \([2n, n]\), gives the lowest complexity, but for another set of parameters, it will be the \([2n_1 + n, n, w_1, w_2]\)-code instead. However, there seems to be an invariant: no matter the length \( n \) of the code or the dimension \( m \) of the extension, it looks like the closer to GV bound the target rank \( r \) is, the better the combinatorial attacks are, and the worse are the algebraic attacks. In other words, for a given \([m, n, k]\)-code, there seems to always be a value of \( r \) such that all the combinatorial attacks will beat the algebraic ones. This seems to be the case both for homogeneous and non-homogeneous versions of the aforementioned problems, and with or without multiple syndromes.

Similarly to [3], we use the fact that \( 1 \in \text{Supp}(x, y) \) to set \( \delta := ww_1 \) in the homogeneous case and \( \delta := ww_1 + w_2 \) in the

| Instance          | Struct. | \( m \) | \( n' \) | \( n \) | \( n_1 \) | \( n_2 \) | \( k \) | \( w \) | \( w_1 \) | \( w_2 \) | DFR |
|-------------------|---------|--------|--------|--------|--------|--------|------|------|--------|--------|------|
| Loong-128 [15]    | Random  | 191    | 182    | 35     | 13     | 16     | 6    | 0    | 8      | 11     | 0    |
| Multi-RQC-AG-128  | Ideal   | 83     | 82     | -      | 5      | 38     | 2    | 7    | 0      | 11     | -138 |
| NH-Multi-RQC-AG-128 | Ideal | 61     | 60     | -      | 3      | 50     | 3    | 51    | 7      | 7     | -158 |
| Multi-RQC-AG-192  | Ideal   | 113    | 112    | -      | 4      | 60     | 2    | 98    | 8      | 13    | 0    | -215 |
| NH-Multi-RQC-AG-192 | Ideal | 79     | 78     | -      | 2      | 95     | 5    | 65    | 8      | 8     | -238 |
| Multi-UR-AG-128   | Random  | 97     | 96     | 24     | 14     | 15     | 3    | 83    | 9      | 11    | -190 |
| NH-Multi-UR-AG-128 | Random | 73     | 72     | 22     | 13     | 14     | 2    | 66    | 8      | 8     | -133 |
| Multi-UR-AG-192   | Random  | 127    | 126    | 35     | 15     | 16     | 3    | 93    | 12     | 9     | -350 |
| NH-Multi-UR-AG-192 | Random | 97     | 96     | 30     | 14     | 14     | 3    | 77    | 9      | 9     | -214 |

| Instance          | Sizes in KB | \( pk \) | \( ct \) | Total |
|-------------------|-------------|---------|---------|-------|
| Loong-128 [15]    | 10.9        | 16.0    | 26.9    |       |
| Multi-RQC-AG-128  | 0.4         | 3.9     | 4.4     |       |
| NH-Multi-RQC-AG-128 | 0.4        | 2.3     | 2.7     |       |
| Multi-RQC-AG-192  | 0.9         | 6.8     | 7.7     |       |
| NH-Multi-RQC-AG-192 | 0.9       | 3.8     | 4.7     |       |
| Multi-UR-AG-128   | 4.1         | 6.9     | 11.0    |       |
| NH-Multi-UR-AG-128 | 2.7       | 4.5     | 7.1     |       |
| Multi-UR-AG-192   | 8.4         | 12.7    | 21.1    |       |
| NH-Multi-UR-AG-192 | 5.1       | 7.5     | 12.6    |       |
non-homogeneous case. Recall that this quantity corresponds to the weight of the error decoded by the public Augmented Gabidulin code. For all our protocols (where “NH” denote non-homogeneous errors), both 128 and 192 bits security level are considered. As a comparison, we also updated the parameters of the code-based KEM Loong [15]. Note that this scheme does not use augmented Gabidulin codes nor non-homogeneous error but it does uses multiple syndromes.

The parameters sets given in Table I come with the sizes of the associated public key \( pk \) and ciphertext \( ct \) expressed in kilo-bytes (KB). For Multi-RQC-AG, \( |pk| = 40 + \left\lfloor \frac{n + m}{8} \right\rfloor \) and \( |ct| = \left\lfloor \frac{2n + m}{8} \right\rfloor \). For Multi-UR-AG, \( |pk| = 40 + \left\lfloor \frac{n + m}{8} \right\rfloor \) and \( |ct| = \left\lfloor \frac{m}{8(n + n_2)} \right\rfloor \). The term 40 represents the length of a seed used to generate \((g, h)\), recall that the public key consists in \((g, h, s)\) and the ciphertext in the couple \((u, v)\). Note that the size of the secret key is not relevant since it is only a seed, thus it always has size 40 bytes.

Our most competitive parameter set in terms of sizes is NH-Multi-RQC-AG-128 which relies on non-homogeneous errors and on an ideal structure. Otherwise, our best non-structured parameter set whose security solely depends on RSL is Multi-UR-AG-128. Overall, Table III enables one to compare all our sizes to the ones of other KEMs based on ideal or non-structured matrices. For structured lattice-based schemes, we chose to focus on CRYSTALS-KYBER [39] which will be considered for standardization by NIST. Outside of this comparison, we may also add the isogeny-based SIKE [41] achieving shorter sizes (676 and 948 bytes for the same \( |pk| + |ct| \) metric). Still, its security assumption is radically different and also a clear bottleneck is that it has slow running times compared to other post-quantum schemes.

Last but not least, the vertical green line at \( N = 150 \) on Figure 4 shows the number of syndromes available for an adversary trying to attack a ciphertext of our scheme NH-MRQC-AG-128. It is worth noticing that, even though the blue squares are below the black line (complexity of the plain RSD attack), they are still way above the security level of 128 bits, and even given 150 syndromes, an attacker could not break our scheme. Note that it is far away from the area where the complexities of the different RSL attacks start to drop. More generally, we picked all our parameters that way, not only to resist to these attacks, but to be sure not to be targeted by any minor improvements.

1) Understanding Parameters From Tables I and II: First, let us consider the parameters set for NH-Multi-RQC-AG-128 in Table I: \((m = 61, n' = 60, n_1 = 3, n_2 = 50, k = 3, \varepsilon = 51, w = 7, w_1 = 7, w_2 = 5)\). With this set of parameters, the public Augmented Gabidulin code will have length \( n_1n_2 = 150 \) and dimension \( k = 3 \), its zero block will have size

\[ n_1n_2 - n' = 90. \]

One expects to recover \( \varepsilon = 51 \) support erasures in the decryption process; this will not happen with probability \( 2^{-138} \), hence the DFR.

These values enables one to compute the correction capacity of this code:

\[ \delta = \left\lfloor \frac{n' - k + \varepsilon}{2} \right\rfloor = \left\lfloor \frac{60 - 3 + 51}{2} \right\rfloor = 54 \]

which corresponds precisely to the rank weight to be decoded, namely

\[ wu_1 + w_2 = 54. \]

To evaluate the security of this set of parameters according to Theorem 2, one has to consider attacking the two following instances:

- an IRSD instance with parameters \((m, 2n_2, w, n_2)\) which yields an RSL instance with parameters

\[ (m, 2n_2, n_2, w, n_2) = (61, 100, 50, 7, 50) \]

where the last \( n_2 \) is the number of syndromes \( N \) one gets from the ideal structure.

- an NHIRSL instance with parameters \((m, n_2, n_2, w_1, w_2, n_1)\) which yields an NHIRSL instance with parameters

\[ (m, n_2, n_2, w_1, w_2, n_1n_2) = (61, 50, 50, 7, 5, 150) \]

where \( n_1n_2 \) is the number of syndromes \( N \) one gets from the ideal structure. Recall that this involves a \([n_2, n_2]-\)code, i.e. a \([150, 50]-\)code.

Second, let us consider the parameters set for NH-Multi-UR-AG-128 in Table I: \((m = 73, n' = 72, n = 22, n_1 = 13, n_2 = 14, k = 2, \varepsilon = 66, w = 8, w_1 = 8, w_2 = 4)\). With this set of parameters, the public Augmented Gabidulin code will have length \( n_1n_2 = 182 \) and dimension \( k = 2 \), its zero block will have size

\[ n_1n_2 - n' = 110. \]

One expects to recover \( \varepsilon = 66 \) support erasures in the decryption process; this will not happen with probability \( 2^{-133} \), hence the DFR.

| Instance           | 128 bits | 192 bits |
|--------------------|----------|----------|
| NH-Multi-UR-AG     | 7,122    | 12,602   |
| LRPC-MS [2]        | 7,205    | 14,279   |
| Multi-UR-AG        | 11,026   | 21,075   |
| FrodoKEM [36]      | 19,336   | 31,376   |
| Loong-128 [15]     | 26,948   | -        |
| Leindeau [37]      | 36,300   | -        |
| Classic McEliece [38] | 264,244 | 524,348  |

TABLE III

Comparison of Sizes for Unstructured (Random) and Structured (Ideal) KEMs. The Sizes Represent the Sum of the Public Key and the Ciphertext, Expressed in Bytes
These values enables one to compute the correction capacity of this code:
\[
\delta = \left\lfloor \frac{n' - k + \ell}{2} \right\rfloor = \left\lfloor \frac{72 - 2 + 66}{2} \right\rfloor = 68
\]
which corresponds precisely to the rank weight to be decoded, namely
\[w_1 + w_2 = 68.\]

To evaluate the security of this set of parameters according to Theorem 3, one has to consider attacking the two following instances:

- an RSL instance with parameters
  \[(m, 2n, n, w_1, w_2, n_1) = (73, 44, 22, 8, 13)\]
  where the last \(n_1\) is the number of syndromes \(N\).
- an \(\text{NRHSD}\) instance with parameters
  \[(m, n, n_1, w_1, w_2, n_2) = (73, 22, 13, 8, 4, 14)\]
  where \(n_2\) is the number of syndromes \(N\). Recall that this involves a \([2n + n_1, n]-\text{code}\), i.e. a \([57, 22]-\text{code}\).

**VII. CONCLUSION**

In this paper, we introduce new variations on the RQC scheme, and more specifically, we introduce the Augmented Gabidulin codes which are very well suited to RQC. These new codes, together with the multiple syndromes and the non-homogeneous approaches, lead to very small parameters which compare very well with other existing code-based schemes. In addition to this, we propose a meaningful scheme only relying on pure random decoding instances, without any ideal structure and with small parameters, around 11KBytes.

We also study more deeply the security of the rank-based problems used in our new schemes. Because of their properties, problems like \(\text{NRHSD}\) or \(\text{RSL}\), are probably bound to be used in many future schemes based on the rank metric.

**APPENDIX A**

**COMPUTATION OF THE SUCCESS PROBABILITY II**

To compute \(\Pi := \Pr_{V,Z}[S_1 \in V, S_2 \subset V \oplus Z]\) we use

**Lemma 4**: Let \(\Pi := \Pr_{V,Z}[S_1 \subset V, S_2 \subset V \oplus Z]\), where the randomness comes from the choice of a random subspace \(V \subset F_q^m\) and a random complementary subspace \(Z\) (hence isomorphic to a subspace of \(F_q^m\)). We have

\[
\Pi = \Pr \left[ S_1 \subset V, S_2/S_1 \subset (V \oplus Z) / S_1 \right] / \Pi_{\text{cond}}
\]

where \(\Pi_{\text{cond}} := \Pr_{V,Z}[S_2/S_1 \subset (V \oplus Z) / S_1 \mid S_1 \subset V]\).

**Proof**: The only non-trivial equality is the first one. For \(\leq\), this is clear by taking the quotient by \(S_1\). For \(\geq\), let \(\pi_{S_1}\) denote the quotient map \(F_q^m \to F_q^m / S_1\). The event at the right-hand side can be seen as \(S_1 \subset V, \pi_{S_1}(S_2) \subset \pi_{S_1}(V \oplus Z) \subset S_1\), and by considering the inverse image by \(\pi_{S_1}\), this event is included in \(S_1 \subset V, \pi_{S_1}^{-1}(\pi_{S_1}(S_2)) \subset \pi_{S_1}^{-1}(\pi_{S_1}(V \oplus Z) \subset S_1))\). This gives \(S_1 \subset V\) and \(S_2 + \ker(\pi_{S_1}) = S_1 + S_2 = S_2\). This space is included in \(V \oplus Z \oplus S_1 + \ker(\pi_{S_1}) = V \oplus Z \oplus S_1 = V \oplus Z \oplus S_1\), hence \(S_1 \subset V\) and \(S_2 \subset V \oplus Z\).

We now focus on the \(\Pi_{\text{cond}}\) factor. Note that we have the decomposition

\[
\{S_2/S_1 \subset (V \oplus Z) / S_1 \mid S_1 \subset V\} = \left\{ S_2/S_1 \subset (V \oplus Z) / S_1 \mid \ell \right\}_{\ell = 0}^{w_2} \left\{ \dim_{F_q}(S_2/S_1 \cap V/S_1) = \ell \right\}
\]

where \(A_\ell := \dim_{F_q}(S_2/S_1 \cap V/S_1) = \ell\) and \(B := \dim_{F_q}(S_2/S_1 \cap V/S_1) = \ell'\) and \(\ell' \leq \ell \leq w_2\), let \(p_\ell := \Pr[A_\ell \cap B]\), let \(s_\ell := \Pr[A_\ell]\) and let \(t_\ell := \Pr[B \setminus A_\ell]\) so that \(p_\ell = s_\ell t_\ell\) and \(\Pi_{\text{cond}} = \sum_{\ell = 0}^{w_2} p_\ell\). To compute \(s_\ell\), we rely on

**Lemma 5** ([89,3.2,p. 269, [42]]): Let \(F\) be an \(F_q\)-linear space of dimension \(n\).

1) If \(X \) is a \(d\)-dimensional subspace of \(F\), then there are \(q^d \binom{n-d}{d}\) \(d\)-dimensional subspaces \(Y\) such that \(X \cap Y = 0\).

2) If \(X\) is a \(d\)-dimensional subspace of \(F\), then there are \(q^{d(n-d)} \binom{n-d}{d}\) \(d\)-dimensional subspaces \(Y\) such that \(X \cap Y\) has dimension \(\ell\).

More precisely, we use Lemma 5.2, with \(F := F_q^m / S_1\), fixed \(X := S_2 / S_1 \subset F_q^m / S_1\) of dimension \(j := w_2\) and random \(Y := V/S_1 \subset F_q^m / S_1\) of dimension \(i := r - w_1\). We obtain

\[
s_\ell = q^{-(r-w_1)}q^{-(r_1-w_1)} \frac{(w_2 \ell - 1)}{q} \frac{(w_2 \ell)}{q}.
\]

To compute \(t_\ell\), note that conditioned on \(\dim_{F_q}(S_2/S_1 \cap V/S_1) = \ell\) the probability that \(S_2/S_1 \subset (V \oplus Z) / S_1\) is the probability that a random subspace of dimension \(\rho\) contains a fixed subspace of dimension \(w_2 - \ell\) in the ambient space \(F_q^m / S_1 \cong F_q^m / V\). From there we obtain \(t_\ell = \frac{(w_2 \ell)}{q}\), and finally by combining this with Equation (21):

\[
p_\ell = q^{(r-w_1)}q^{(r_1-w_1)} \frac{(w_2 \ell - 1)}{q} \frac{(w_2 \ell)}{q} \frac{(\rho)}{q}.
\]

**APPENDIX B**

**FINISHING THE PROOF OF THEOREM 4**

Obviously \(p_0 < \Pi_{\text{cond}}\) and one can also easily show that \(\Pi_{\text{cond}} = \Theta(p_0)\). By including the \(q^{-m}\) factor from [32], the number of \(F_q\)-operations in the attack is

\[
\mathcal{K} = O \left( L \times \Pi^{-1} \times q^{-m} \right) = \tilde{O} \left( \Pr[C]^{-1} p_0^{-1} q^{-m} \right),
\]

where \(\Pr[C] := \Pr[V \mid S_1 \subset V]\) and where \(L\) is the polynomial factor coming from the linear algebra step whose exact
formula is not relevant for the discussion. Using the classical \( \binom{n}{k} \theta = \Theta\left(q^{\binom{n-k}{2}}\right) \) when \( \max(a, b) \to +\infty \) together with
\[
p_0 = q^{(r-w_1)w_2} \frac{\binom{m-w_1-w_2}{r-w_1}}{\binom{m-w_1-w_2}{r-w_1}} \frac{p}{w_2} q,
\]
we obtain \( p_0 = \Theta\left(q^{(r-w_1)w_2} \times q^{-(r-w_1)w_2} \times q^{-(w_2(m-r) - n)}\right) = \Theta\left(q^{w_2(m-r)}\right) \), and similarly \( \Pr[C] = \Theta\left(q^{w_1(m-r) - w_2n + m}\right) \), which is the statement of Theorem 4.

APPENDIX C

PROOFS FOR THE MAXMINORS ATTACK ON NRHSD

A. Proof of Proposition 3

For item 1., let \( J \in \{1, n + n + 1\} \), \#J = w_1 + w_2 such that \( P_J \in \mathcal{P}_\text{lost} \). By definition of \( \mathcal{P}_\text{lost} \) the set \( J + n + 1 \) has intersection \( \leq w_2 - 2 \) with \( \{n+2-n+n\} \), hence any subset \( T = T^* \cup T^{*+} \subseteq \{1, n + 1\} \). \( T^{*+} \subseteq \{J + n + 1\} \) satisfies \#(\{J \cap \{n + n + n + 1\} \} \leq w_2 - 1 \) since \( T^{*+} \) might also contain \( n + 1 \). This means that the corresponding minor variable \( c_T \in \zeta \) belongs to \( \zeta \) and can be set to zero in \( P_J \). Using the shape depicted in Equation (10), this implies that the whole \( P_J \) equation becomes zero. For item 2., recall that the leading term of \( P_J \in \mathcal{P}_\text{inde} \) is \( c_J + n + 1 \). Moreover we have \#(\{J \cap \{n + n + 1\} \} \# J + n + 1 \cap \{n + 2-n+n\} \geq w_2 \), which means \( c_J + n + 1 \not\in \zeta \). In particular, all the equations from \( \mathcal{P}_\text{inde} \) keep the same leading terms after fixing the \( M \) minor variables to zero and therefore they remain linearly independent. The last statement on the number of variables in obvious.

B. Lemma to prove Proposition 4

1) Proof of Lemma 2: Using Equation (13), one has that the equations in \( \mathcal{P}_{\text{rest}, A} \) all have their monomials in \( \mu_A := \{c_T, T \subseteq \{1, 2n+n+1\}, \#T = w_1 + w_2, n + 1 \in T, T \cap \{n+2-n+n+1\} = A\} \), and this set has size \((2n)^{w_1}\). Finally, note that \( \mu_A \) and \( \mu'_{\text{A}} \) are disjoint when \( A \not= A' \), which concludes the proof.

2) Proof of Lemma 3, Under Assumptions: Using Equation (13), it is readily verified that the set of leading terms of all equations in \( \mathcal{P}_{\text{rest}, A} \) is
\[
\tau_A := \{c_{n+1} \cup U \mid U \subseteq \{n + n + 2 \ldots (2n + n)\}, \#U = w_1\},
\]
and for instance note that the equation \( P_U \) with \( J_U, n + 1 = \sum_{U} A \cup (n + n + 1) \cup U \) has leading term \( c_{n+1} \cup U \in \tau_A \). This already shows that \( \text{dim}_{\mathbb{F}_q}(\mathcal{P}_{\text{rest}, A}) \geq \#\tau_A = (2n)^{w_1} \). For the converse inequality, we need to rely on some assumption on the randomness of the entries of the \( \text{P}_J \)'s in \( \mathbb{F}_q \) to argue that we cannot construct an element in \( \mathcal{P}_{\text{rest}, A} \) whose leading term does not belong to \( \tau_A \) with very high probability. First, note that the variables from \( P_J \in \mathcal{P}_{\text{rest}, A} \) with \( J + n + 1 = A \cup V_J \) where \( V_J = \{v_{(1)}^J, \ldots, v_{(w_1)}^J\} \) which belong to \( \tau_A \) are the \( c_{n+1} \cup U \cup V_J \) for \( 1 \leq J \leq w_2 + 1 \). To kill the leading term of \( P_J \), one would then consider an equation with the same leading term, namely a \( P_J \) with \( J' \not= J \), \( J' + n + 1 = A \cup V_{J'} \) and such that \( V_{J'}(v_{(1)}^J) = V_{J'}(v_{(w_1)}^J) = B \) for some \( B \). In this case, one can check that the only monomial from \( \tau_A \) present in both \( P_J \) and \( P_{J'} \) is \( c_{n+1} \cup U \cup B \), so that \( P_J + \lambda_{J'} P_{J'} \) contains at least \( w_2n \) monomials from \( \tau_A \). Similarly, by using a third \( J'' \) one could kill at most one extra monomial in \( P_J \) and in the worst case one in \( P_{J'} \) as well. This means that a linear combination of the form \( P_J + \lambda J P_{J'} + \lambda_{J''} P_{J''} \) contains at least \( 2(w_1 + 1) + (w_1 + 1 + 2) = 3w_1 + 1 \) monomials from \( \tau_A \), and the lower bound is reached if and only if those monomials in \( P_J \) and \( P_{J'} \) are killed at the same time by \( J'' \). This is extremely unlikely if the coefficients of the MaxMinors equations are random elements in \( \mathbb{F}_q \), so that we assume instead that \( P_J + \lambda J P_{J'} + \lambda_{J''} P_{J''} \) contains at least \( (w_1 + 1) + (w_1 + 1 + 1) = 3w_1 + 1 \) monomials in \( \tau_A \). Relying on the same type of assumption, one can proceed by induction on the numbers of terms to show that the non-zero linear combination in \( \mathcal{P}_{\text{rest}, A} \) always has a monomial in \( \tau_A \).

ACKNOWLEDGMENT

The author Maxime Bomar would like to thank Maxime Bomar for helpful discussion.

REFERENCES

[1] C. Aguilar-Melchor, O. Blazy, J.-C. Deneuville, P. Gaborit, and G. Zémor, “Efficient encryption from random quasi-cyclic codes,” IEEE Trans. Inf. Theory, vol. 64, no. 5, pp. 3927–3943, May 2018.
[2] C. A. Melchor, N. Aragon, V. Dyserly, P. Gaborit, and G. Zémor, “LRPC codes with multiple syndromes: Near ideal-size KEMs without ideals,” in Post-Quantum Cryptography, J. H. Cheon and T. Johansson, Eds., 2022.
[3] C. A. Melchor et al., “Rank quasi cyclic (RQC),” Second Round submission to NIST Post-Quantum Cryptography Call, Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep., Apr. 2020.
[4] P. Gaborit, A. Hauteville, D. H. Phan, and J. Tillich, “Identity-based encryption from rank metric,” in Advances in Cryptology—CRYPTO (Lecture Notes in Computer Science), vol. 10403, Santa Barbara, CA, USA: Springer, Aug. 2017, pp. 194–226.
[5] E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov, “Ideals over a non-commutative ring and their applications to cryptography,” in Advances in Cryptology—EUROCRYPT (Lecture Notes in Computer Science), vol. 547, D. W. Davies, Ed. Brighton, U.K., Apr. 1991, pp. 482–489.
[6] E. M. Gabidulin, “Theory of codes with maximum rank distance,” Problemy Peredachi Informatsii, vol. 21, no. 1, pp. 3–16, 1985.
[7] R. Overbeck, “A new structural attack for GPT and variants,” in Progress in Cryptology (Lecture Notes in Computer Science), vol. 3715, E. Dawson and S. Vaudenay, Eds. Kuala Lumpur, Malaysia, 2005, pp. 50–63.
[8] A. Otmani, H. T. Kalachi, and S. Ndjeya, “Improved cryptanalysis of rank metric schemes based on Gabidulin codes,” Des., Codes Cryptogr., vol. 86, no. 9, pp. 1983–1996, Sep. 2018.
[9] J. Hoffstein, J. Pipher, and J. H. Silverman, “NTU: A ring-based public key cryptosystem,” in Proc. 3rd Int. Symp. Algorithmic Number Theory (ANTS-III) (Lecture Notes in Computer Science), vol. 1423, J. Buhler, Ed. Portland, OR, USA: Springer, Jun. 1998, pp. 267–288.
[10] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. S. L. M. Barreto, “MDPC-McEliece: New McEliece variants from moderate density parity-check codes,” in Proc. IEEE Int. Symp. Inf. Theory, Jul. 2013, pp. 2069–2073.
[11] P. Gaborit, G. Murat, O. Ruatta, and G. Zémor, “Low rank parity check codes and their application to cryptography,” in Proc. Workshop Coding Cryptography, Bergen, Norway, 2013, pp. 1–14.
[12] P. Loideau, “A new rank metric codes based encryption scheme,” in Post-Quantum Cryptography (Lecture Notes in Computer Science), vol. 10346, Berlin, Germany: Springer, 2017, pp. 3–17.
[13] M. Alekhnovich, “More on average case vs approximation complexity,” in Proc. 44th Annu. IEEE Symp. Found. Comput. Sci., Oct. 2003, pp. 298–307.
M. Bardet et al., “An algebraic attack on rank metric code-based cryptosystems,” in Advances in Cryptology—EUROCRYPT, A. Canteaut and Y. Ishai, Eds. Zagreb, Croatia, 2020.

M. Bardet et al., “Improvements of algebraic attacks for solving the rank decoding and minrank problems,” in Advances in Cryptology—ASIACRYPT, S. Moriai and H. Wang, Eds., 2020, pp. 507–536.

T. Debris-Alazard and J.-P. Tillich, “Two attacks on rank metric code-based schemes: RankSign and an identity-based-encryption scheme,” in Advances in Cryptology—ASIACRYPT (Lecture Notes in Computer Science), vol. 11722. Brisbane, QLD, Australia: Springer, Dec. 2018, pp. 62–92.

M. Bardet and P. Briaud, “An algebraic approach to the rank support learning problem,” in Post-Quantum Cryptography (Lecture Notes in Computer Science), J. H. Cheon and J.-P. Tillich, Eds. Berlin, Germany: Springer, 2021.

N. Aragon, O. Blazy, P. Gaborit, A. Hauteville, and G. Zemor, “Durandal: A rank metric based signature scheme,” in Advances in Cryptology—EUROCRYPT (Lecture Notes in Computer Science), vol. 11478. Darmstadt, Germany: Springer, 2019, pp. 728–758.

I. S. Reed and G. Solomon, “Polynomial codes over certain finite fields,” in J. Soc. Ind. Appl. Math., vol. 8, no. 2, pp. 300–304, Jun. 1960.

O. Ore, “On a special class of polynomials,” Trans. Amer. Math. Soc., vol. 35, no. 3, pp. 559–584, 1933.

P. Gaborit and G. Zémor, “On the hardness of the decoding and the minimum distance problems for rank codes,” IEEE Trans. Inf. Theory, vol. 62, no. 12, pp. 7245–7252, Dec. 2016.

G. Kabatianskii, E. Krouk, and B. M. Smeets, “A digital signature scheme based on random error-correcting codes,” in Proc. IMA Int. Conf. (Lecture Notes in Computer Science), vol. 1355. Cham, Switzerland: Springer, 1997, pp. 161–167.

A. Otmani and J.-P. Tillich, “An efficient attack on all concrete KKS proposals,” in Post-Quantum Cryptography (Lecture Notes in Computer Science), vol. 7071. B. Y. Yang, Eds. Taipei, Taiwan, 2011, pp. 98–116.

D. Augot, P. Loidreau, and G. Robert, “Generalized Gabidulin codes over fields of any characteristic,” Des., Codes Cryptogr., vol. 86, no. 8, pp. 1807–1848, Aug. 2018.

A. Couvreur and M. Bombard, “Right-hand side decoding of Gabidulin codes and applications,” in Proc. WCC, 2022, pp. 1–11.

E. M. Gabidulin and N. I. Pilipchuk, “Error and erasure correcting algorithms for rank codes,” Des., Codes Cryptogr., vol. 49, nos. 1–3, pp. 105–122, Dec. 2008.

P. Loidreau, “Properties of codes in rank metric,” 2006, arXiv:0610057.

P. Gaborit, O. Ruatta, and J. Schrek, “On the complexity of the rank syndrome decoding problem,” IEEE Trans. Inf. Theory, vol. 62, no. 2, pp. 1006–1019, Feb. 2016.

N. Aragon, P. Gaborit, A. Hauteville, and J.-P. Tillich, “A new algorithm for solving the rank syndrome decoding problem,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Vail, CO, USA, Jun. 2018, pp. 2421–2425.

P. Gaborit, O. Ruatta, J. Schrek, and G. Zemor, “New results for rank-based cryptography,” in Progress in Cryptology—AFRICACRYPT (Lecture Notes in Computer Science), vol. 8469. D. Pointcheval and D. Vergnaud, Eds. Marrakesh, Morocco, 2014, pp. 1–12.

P. Gaborit, A. Hauteville, and J.-P. Tillich, “RankSynd a PRNG based on rank metric,” in Post-Quantum Cryptography. Fukuoka, Japan: Springer, Feb. 2016, pp. 18–28.

A. Wakasugi and M. Tada, “A proposal for quantum GRS algorithm and the cryptanalysis for ROLLO and RQC,” Cryptol. ePrint Arch., 2023. [Online]. Available: https://eprint.iacr.org/2023/0904