Study and design of new encryption primitives based on rank metric error correcting codes

Ba Duc Pham

▶ **To cite this version:**

Ba Duc Pham. Study and design of new encryption primitives based on rank metric error correcting codes. Cryptography and Security [cs.CR]. Université de Rennes 1, 2021. English. NNT : . tel-03516724

HAL Id: tel-03516724
https://hal.science/tel-03516724

Submitted on 7 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : « *Mathématiques et leurs Interactions* »

Par
PHAM Ba Duc

Étude et conception de nouvelles primitives de chiffrement fondées sur les codes correcteurs d'erreurs en métrique rang

Thèse présentée et soutenue à Rennes, le 10 Decembre 2021
Unité de recherche : IRMAR, UMR CNRS 6625 Institut de Recherche Mathématiques de Rennes

Rapporteurs avant soutenance :

Alain Couvreur	Directeur de recherche, INRIA Saclay
Philippe Gaborit	Professeur des Universités, Université de Limoges

Composition du Jury :

Président :	Ayoub Otmani	Professeur, Université de Rouen
Examinateurs :	Delphine Boucher	Maître de conférences, l'IRMAR Université de Rennes 1
	Antonia Wachter-Zeh	Professeur, Technical University of Munich
	Ayoub Otmani	Professeur, Université de Rouen
	Alain Couvreur	Directeur de recherche, INRIA Saclay
	Philippe Gaborit	Professeur des Universités, Université de Limoges
Dir. de thèse :	Pierre Loidreau	Chercheur, DGA, l'IRMAR Université de Rennes 1

Acknowledgement

First of all, I want to give my deepest gratitude to my supervisor Pierre Loidreau. You not only gave me the opportunity to do this thesis but also always stood by me for the past 3 years with all of your sincerity and dedication. It is no exaggeration to say that I could not have completed my thesis without your help. You have given me the instructions, corrections whenever I were wrong, and pointers to steer me in the right direction. Every Friday, you always spend hours for listening to me and answering my questions. It is your enthusiasm, erudition and carefulness that inspires me to confidently complete this thesis. My thesis passed through the pandemic Covid-19, but during this hard time, you have always encouraged and motivated me. I still want to say a lot of things to thank you, but I don't want the words of thanks to be cliched. Sincerely, it's great and honor for me to be able to do thesis under your guidance.

I also want to express my thank to professor Delphine Boucher. You are my first teacher in my first class in France. When I first came to France, I thought everything was difficult since I had to study in French but I wasn't really good at it. But at that time, you always helped me, tried to communicate with me, answered all of my questions in English. Your help gives me a lot of confidence, and it's like a solid fulcrum for me to complete my Master well. You are also the one who first introduced me to the beautiful world of "error-correcting code" and made me love it. I would also like to thank professor Sylvain Duquesne. You are not only the director of Master cryptography where you always took care all of student including me, but also the one who introduced me to professor Pierre Loidreau and helped me to have a chance to do this thesis.

I also want to thank the Centre Henri Lebesgue that give me the scholarship funding my Master. I truly thank the IRMAR laboratory and ED MATHSTIC for your hospitable environment. On the other hand, I would like to show my gratitude to all the secretaries in IRMAR (Madame Xhensila Lachambre, Chantal Halet, Nelly Loton), MATHSTIC (Madame Elodie Cottrel), who helped me a lot with all the confusing and complicated administrative procedures.

For my rapporteurs, professor Alain Couveur and Philippe Gaborit, thank you for spending time to read my manuscript. I am also sincerely grateful to team of my juries:

professors Delphine Boucher, Antonia Wachter-Zeh, Ayoub Otmani, Alain Couvreur, and Philippe Gaborit.

My thanks to all of my colleagues who support me a lot throughout my thesis. Special thanks to Julien, who co-worked with me in the two first year of my thesis. You are a very friendly and humorous people in daily life and a serious and careful person in work. You have taught me a lot about how to write an academic paper and enthusiastically answered my questions even when you were busy with preparation for the qualification of MdC. I always greatly appreciate your help. I also thank Elie and Fabien, who are my co-offices. Although we don't have a lot of time to talk because of work but my time of thesis can't be complete without you.

I am indebted to all of my friends. Andy, who is my best friend in France. From the time of Master, you have been always helped me to translate all the French lessons into English and spent time to discuss with me about the TD and TP where I was not very good at programming. Without you, I could not finish my Master that well. Although during the thesis, we didn't have as much time to talk as before but I won't forget all of your enthusiasm and sincerity.

For my family, I sincerely thank my parents, my sister who give me their caring, sympathy and love, that encourage me to finish my study in foreign country. Finally, special thanks to my girlfriend, who always support me during 5 years in France. Thank you for sharing with me all the moments not only the happiness but also the sorrow, it makes my life colorful.

Abstract

In the 1990s, Gabidulin, Paramonov and Tretjakov proposed a McEliece type encryption system based on the difficulty of decoding a linear metric rank code [1]. Since the complexity of decoding in rank-metric is exponentially more expensive with fixed parameters than Hamming metric decoding, the use of this metric allows to design encryption systems with more compact keys. The underlying code family used is the Gabidulin code family.

One of the objectives of this thesis is to study and design new primitives encryption using rank metric codes. To do this, we will study the structure of a new family of codes, derived from the Gabidulin codes. In 2005, Faure and Loidreau proposed a new rank-metric cryptosystem [2] inspired from the Hamming metric scheme of Augot-Finiasz in 2003. In 2018, it was broken by the attack of Gaborit, Otmani and Kalachi [3]. Recently, there are some attempts of repairing the Faure-Loidreau scheme, for example the work of Renner, Puchinger and Wachter–Zeh which is called LIGA [4]. In this thesis, we also introduce a new cryptosystem so-called RAMESSES [5] which is another repairing of Faure-Loidreau scheme.

Besides, we also study about the recent attack of Coggia and Couveur [6] in the Loidreau's cryptosystem (2017) [7] . Although they only propose an idea for a special case of the dimension of secret subspace, this attack can be generalized. In this thesis, we propose an analysis of Coggia-Couvreur attack on Loidreau's rank-metric public-key encryption scheme in the general case.

The last part is a study about the decoding of the sum of Gabidulin codes which is inspired from the work of Loidreau in 2005 "Welch-Berlekamp Like Algorithm for Decoding Gabidulin Codes" [8]. This work is also an attempt to repair the Loidreau's cryptosystem (2017) to avoid the Coggia-Couveur's attack.

RÉSUMÉ FRANCAIS

Cryptographie et système de cryptographie à clé publique

Tout au long de l'histoire du développement humain, l'échange d'informations a été un élément indispensable de la société humaine. Avec la nécessité d'échanger des lettres sur de longues distances, l'authentification, la confidentialité et l'intégrité des informations deviennent partie intégrante des règles incontournables de la sécurité dans la communication. Dans l'histoire du développement du système de communication, il existe deux types de cryptosystèmes. Le premier apparaît dès le départ, le cryptosystème qui repose sur l'utilisation d'un secret partagé entre les utilisateurs. Ce type de systèmes est ce qu'on appelle la cryptographie à clé secrète. Dans ce système, l'algorithme de chiffrement et de déchiffrement ont la même clé et il est appelé chiffrement à clé symétrique, par exemple des algorithmes DES, AES. La seconde apparaît après les recherches de Diffie et Hellman [9], appelée cryptographie à clé publique, dans laquelle chaque utilisateur dispose de deux clés : clé secrète et clé publique :

— La clé secrète est privée et utilisée pour le décodage

— La clé publique est publique et utilisée pour le encodage

La méthode de chiffrement et de déchiffrement est la suivante : l'expéditeur veut envoyer le message au destinataire, il utilise la clé publique du destinataire pour chiffrer le message par l'algorithme de chiffrement. Le récepteur reçoit le texte chiffré, il utilise sa clé secrète dans l'algorithme de déchiffrement pour le transférer en texte clair. Étant donné que la clé publique et la clé secrète de chaque personne sont généralement différentes (dans certains cas, la clé publique est générée à partir de la clé privée par une fonction unidirectionnelle), ce type de système est appelé cryptosystème asymétrique. L'un des avantages les plus importants du système de chiffrement à clé publique est qu'il réduit le nombre de clés de stockage utilisées pour un groupe d'utilisateurs. De nos jours, le cryptosystème à clé publique a de nombreuses applications dans les domaines bancaire, des télécommunications, Internet, etc.

Code correcteur d'erreurs et cryptographie basée sur le code

Ce n'est qu'au cours de la Seconde Guerre mondiale, avec l'avènement des technologies de l'information et des modèles de communication créés par Shannon [10], que la cryptographie a vraiment fait un grand pas en avant et a attiré plus d'attention. Le codage de l'information apparaît comme une partie de la théorie de l'information du grand domaine. L'échange de code via un canal bruyant implique au cœur du code correcteur d'erreurs, qui a été introduit par Shannon dans l'article [11]. Dans cet article, basé sur l'idée de Harry Nyquist et Ralph Hartley, il a prouvé que nous pouvons échanger des informations sans erreur via un canal bruyant sur la capacité du canal. Ce théorème est appelé par la suite théorème de codage de canal bruité.

Bien que la publication de Shannon nous montre l'existence du code qui atteint la limite de Shannon, elle n'a pas montré comment construire de tels types de code. En 1950, Hamming a introduit une métrique [12], qui est devenue une métrique très connue de nos jours : la métrique de Hamming. Sur la base de cette métrique, jusqu'à nos jours, de nombreux types de code ont été construits tels que le code Reed-Solomon, BCH, etc,.... Ces familles de code bien construites nous apportent leur application intéressante non seulement en théorie mais aussi dans la vie quotidienne comme les disques compacts [13], la connexion internet ADSL [14], etc. En 1978, McEliece propose un cryptosystème à clé publique à code [15], qui est à la base de cryptographie basée sur le code.

Codes en métrique rang et problématique

La sécurité du cryptosystème à clé publique basé sur le code dans [15] que propose McEliece, repose sur la difficulté de résoudre le problème de décodage borné pour un code linéaire en métrique de Hamming, qui a été prouvé comme un problème NP-difficile [16]. Cela signifie qu'en général, il n'existe pas d'algorithme efficace pour résoudre ce problème en temps polynomial. Cependant, le décodage par ensemble d'informations, introduit par Prange en 1962 [17], nous montre une attaque pour les petits paramètres. Par conséquent, pour garder un tel système toujours sécurisé, nous avons besoin d'une taille de clés suffisamment grande.

La même année, Delsarte [18] et sept ans plus tard, en 1985, Gabidulin [19], ont introduit et développé une nouvelle métrique dite métrique rang en [20]. Par la suite,

Gabidulin, Paramonov et Tretjakov ont développé un nouveau cryptosystème de type McEliece basé sur cette métrique, qui est appelé par la suite GPT-cryptosystem [21]. Ce cryptosystème utilise l'algorithme de encodage et de décodage des codes Gabidulin. Pour comparer avec l'autre cryptosystème de type McEliece basé sur la métrique de Hamming, celui-ci a besoin d'une taille de clé plus petite. Théoriquement, pour le même ensemble de paramètres, cela pourrait nous donner un niveau de sécurité plus élevé.

Gabidulin codes

En 1962, Singleton a introduit une limite pour la distance minimale d'un code [22]. Le type de codes dont la distance minimale atteint la limite de Singleton attirent beaucoup d'intérêt et d'attention. Dans son propre article [19], Gabidulin a comparé les propriétés de la métrique de Hamming et de la métrique de rang et a constaté que dans le cas de la métrique rang, nous avons également la borne de Singleton pour la distance minimale, la même que la métrique de Hamming. En conséquence, cela conduit à la recherche des codes dont la distance minimale atteint la borne de Singleton (dans le cas de la métrique de Hamming, ce sont les codes dits MDS et pour la métrique rang, il s'agit du code MRD). Dans la section 1.2.1, Gabidulin a montré un moyen de créer une « équivalence » du code Reed Solomon en métrique rang, appelée code de Gabidulin. Ces codes sont des MRD avec un algorithme efficace pour le décodage jusqu'à sa capacité d'erreur (par exemple, l'algorithme de Loidreau dans [8], que j'ai implémenté dans MAGMA, voir `https://github.com/BaDucPham/RAMESSES/blob/main/DecGab.mgm`). Ces codes jouent un rôle important dans le développement de la cryptographie basée sur le code métrique rang, en particulier le cryptosystème de type GPT. Cependant, de nombreux systèmes de type GPT étaient en panne. La principale faiblesse de ces systèmes repose sur le fait que les codes de Gabidulin sont vulnérables à l'attaque invariante puisqu'ils contiennent un immense espace vectoriel invariant par l'action de l'automorphisme de Frobenius.

Cryptographie post-quantique

Le cryptosystème à clé publique dont la sécurité repose sur l'un des trois problèmes difficiles : la factorisation en nombres entiers, le logarithme discret dans les corps finis ou le logarithme discret à courbe elliptique, peut être brisé sur un ordinateur quantique suffisamment puissant. Il répond aux besoins de la cryptographie Post-quantique, qui a beaucoup retenu l'attention (conférence PQCrypto depuis 2006 par exemple). Pour la

cryptographie basée sur le code, l'un des candidats les plus connus pour la cryptographie post-quantique est le cryptosystème McEliece.

Avec la probabilité croissante de l'existence d'un ordinateur quantique dans un futur proche, il est devenu important de proposer des alternatives aux cryptosystèmes à clé publique et aux protocoles d'échange de clés existants basés sur la théorie des nombres. Le récent processus de normalisation de la cryptographie post-quantique du NIST motive des propositions dans ce sens. Avec la cryptographie basée sur lattice, la cryptographie basée sur le code est la plus représentée parmi les propositions de cryptosystèmes ou de mécanismes d'encapsulation de clé (KEM). Les soumissions basées sur le code reposent de manière générique sur la dureté des problèmes de décodage, soit dans la métrique de Hamming, soit dans la métrique rang. Les problèmes de décodage métrique de Hamming bénéficient d'une étude de longue date et de peu d'améliorations pratiques depuis plus de cinquante ans, ce qui atteste de leur sécurité. A l'inverse, les problèmes de décodage de métriques rang sont étudiés depuis moins de vingt ans [23], et leur complexité de résolution n'est pas encore totalement stabilisée (voir les résultats récents de [24]). Néanmoins, ils bénéficient de clés beaucoup plus courtes et semblent très attrayants pour une mise en œuvre pratique, aboutissant à des soumissions pour le processus de normalisation NIST [25, 26]. Afin de réduire encore les tailles de clés, les concepteurs utilisent souvent des structures spécifiques comme quasi-cyclicité (équivalent du Module-LWE pour les treillis) qui pourraient être suspectées d'introduire des faiblesses supplémentaires [27].

Organisation de ma thèse

Dans ma thèse, je m'intéresse au cryptosystème de type GPT. Il a été motivé par l'étude de certaines recherches de mon directeur, professeur Pierre Loidreau, dans son cryptosystème, le cryptosystème Faure-Loidreau dans [2] et le récent dans [7]. Malheureusement, ces systèmes ont été récemment attaqués, les premiers ont été attaqués par Gaborit, Otmani et Kalachi dans [3] et les seconds dans l'article de Coggia et Couvreur [6]. Toutes ces recherches m'ont inspiré à étudier davantage la structure du code Gabidulin et à trouver un moyen de réparer ou de modifier le système pour résister à l'attaque.

Cette thèse est organisée comme suit :

— chapter 1 Le premier chapitre porte sur l'extension de l'attaque Coggia et Couvreur sur le cryptosystème de Loidreau. Cette attaque est basée sur l'idée de la faib-

lesse du code Gabidulin par le distingueur d'Overbeck, que nous pouvons distinguer les codes Gabidulin des codes aléatoires. De plus, comme beaucoup d'autres cryptosystèmes de type GPT, le cryptosystème de Loidreau montre toujours la faiblesse de la structure algébrique. Bien qu'il puisse parfaitement éviter l'attaque directe du distinguateur dans le code public, Coggia et Couvreur ont tout de même exploité un autre distinguateur pour le dual code du code public pour le paramètre secret $\lambda = 2$. Récemment, il a été étendu par Ghatak dans [28] pour le cas de $\lambda = 3$ mais il était incomplet. Ce chapitre concerne mon travail de généralisation de cette attaque pour le cas de tout λ et la considération sur la complexité de cette attaque et le comblement du vide pour le cas $\lambda = 3$ pour montrer l'efficacité de l'attaque dans ce cas. L'implémentation dans MAGMA peut être vue dans `https://github.com/BaDucPham/Coggia-and-Couvreur-attack`

— chapter 2 Le deuxième chapitre porte sur un nouveau cryptosystème qui s'inspire du cryptosystème Faure-Loidreau, dans lequel, nous considérons le texte clair comme l'espace des lignes d'une erreur. Dans cette section, dans un premier temps, nous présenterons un cryptosystème de type Augot-Finiasz et ensuite, étudierons sa sécurité en considérant la complexité de certaines attaques connues sur celui-ci. Récemment, ce cryptosystème a été brisé lors de l'attaque Bombar-Couveur dans [29]. Par conséquent, à la fin, nous présenterons l'idée de Bombar et Couvreur pour casser ce système.

— chapter 3 Le dernier chapitre est mon travail sur la structure des codes de Gabidulin. De l'idée d'utiliser des q-polynômes reconstruits pour le décodage des codes de Gabiduline [8], notre travail porte sur le décodage de la somme des codes de Gabiduline. Ce travail est également issu d'une tentative de modification du cryptosystème Loidreau [7] pour résister à l'attaque Coggia et Couvereur [6]. Bien qu'il n'y ait aucune preuve de la résistance de cette modification à cette attaque, elle a encore quelques applications. Dans ce chapitre, je présenterai le décodeur en temps polynomial pour la somme des codes de Gabiduline avec une probabilité de défaillance exponentiellement faible et ensuite, j'introduire certaines de ses applications dont une idée sur le remplacement du code de Gabiduline par la somme des codes de Gabiduline dans le Loireau cryptosystem et ma propre supposition sur la résistance de cette modification contre l'attaque Coggia et Couvereur.

Le résultat du premier chapitre est un article en cours et présenté à CBCrypt 2021, le deuxième est un article pré-imprimé et le troisième est publié dans la conférence ISIT 2021 (International Symposium on Information Theory).

Table of Contents

NOTATIONS

Finite Fields	
q	Power of a prime
$\mathbb{F}_q$	Finite field of order q
$\mathbb{F}_{q^m}$	Finite field extension of degree m of $\mathbb{F}_q$
$\mathbb{F}_q^{m\times n}$	Set of all $m \times n$ matrices over $\mathbb{F}_q$
$\mathbb{F}_{q^m}^n$	Set of all row vectors of length n over $\mathbb{F}_{q^m}$
$\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m) \in \mathbb{F}_{q^m}^m$	A basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$
$[i] = q^i$	The power i of q
Subspaces	
$\mathrm{Gr}(t, \mathbb{F}_2^n)$	Set of subspaces of $\mathbb{F}_2^n$ of dimension t
Matrices	
$\boldsymbol{A} = (a_{i,j})_{i=0,j=0}^{m,n}$	$m \times n$ matrix
$\mathrm{rk}_q(\boldsymbol{A})$	Rank of A over $\mathbb{F}_q$
$\mathrm{rk}_{q^m}(\boldsymbol{A})$	Rank of A over $\mathbb{F}_{q^m}$
$\ker(\boldsymbol{A})$	Nullspace of matrix $\boldsymbol{A}$
$\mathrm{RowSp}_{\boldsymbol{\beta}}(\boldsymbol{a})$	Row space of vector $\boldsymbol{a}$
$\mathrm{ColSp}_{\boldsymbol{\beta}}(\boldsymbol{a})$	Column space of vector $\boldsymbol{a}$
Codes	
$\mathcal{G}_k(\boldsymbol{g})$	Gabidulin code of dimension k, support $\boldsymbol{g}$
$\mathcal{C}_{\mathrm{pub}}$	The public code
$\mathcal{C}^\perp$	The dual code of the code $\mathcal{C}$
$\boldsymbol{G}$	The generator matrix of a code
$\boldsymbol{H}$	The parity-check matrix of a code
Linearized polynomials	
θ	Generator of a Galois group, Frobenius map
$\mathbb{F}_{q^m}[X;\theta]$	The skew polynomial ring over $\mathbb{F}_{q^m}$

Acronyms

RREF	Reduced row echelon form
MDS	Maximum distance separable
MRD	Maximum rank distance
KEM	Key-encapsulation mechanisms
PKE	Public key encryption

INTRODUCTION

Cryptography and public key cryptosystem

Throughout the history of human's development, the exchange of information has been an indispensable part of human society. With the need of exchanging letters back and forth over long distances, the authentication, confidentiality and information integrity become the integral parts of the inevitable rules of the security in communication. In the history of the development of the communication system, there are two kinds of cryptosystems. The first one appears from the beginning, the cryptosystem which relies on using a shared secret between users. This kind of systems is so-called secret key cryptography. In this system, the algorithm of encryption and decryption have same key and it is called symmetric key cipher, for example DES, AES algorithms. The second one appears after the research of Diffie and Hellman [9], called public key cryptography, in which each user has a couple of keys: secret key and public key:

— The secret key is private and used for the decryption

— The public key is public and used for the encryption

The method of encryption and decryption is as follows: the sender wants to send message to the receiver, he uses the receiver's public key to encrypt the message by the encryption algorithm. The receiver receives the cipher text, he uses his secret key in the decryption algorithm to transfer into plaintext. Since the public key and secret key for each person is generally different (in some case, the public key is generated from the private key by an one-way function), this kind of system is called asymmetric cryptosystems. One of the most important advantages of the public key cryptosystem is that it reduces the number of used keys storage for a group of users. Nowadays, public key cryptosystem has a lot of applications in banking, telecommunication, internet, etc.

Error-correcting code and code based cryptography

It was not until World War II, with the advent of the information technology and the communication models created by Shannon [10], that the cryptography really received a big move and earned more attention. The information coding appears as a part of the large domain information theory. The exchanging code through a noisy channel implies to the core of error-correcting code, which was introduced by Shannon in the paper [11]. In this paper, based on the idea of Harry Nyquist and Ralph Hartley, he proved that we can error-freely exchange information through noisy channel upon the channel capacity. This theorem is after called the noisy channel coding theorem.

Although the publication of Shannon shows us the existence of the code that reach the Shannon's limit, it did not show how to construct such kinds of code. In 1950, Hamming introduced a metric [12], which has become a very well known metric nowadays: Hamming metric. Based on this metric, until nowadays, a lot of types of code have been constructed such as Reed-Solomon code, BCH, etc,.... These well-constructed families of code therefrom bring to us their interesting application not only in theory but also in daily life such as compact disks [13], internet connection ADSL [14], etc. In 1978, McEliece proposes a code-based public-key cryptosystem [15], which is the foundation of code-based cryptography.

Codes in rank metric and problematic

The security of the code-based public-key cryptosystem in [15] which McEliece proposes, relies on the hardness of solving the bounded decoding problem for a linear code in Hamming metric, which was proven as NP- hard problem [16]. It means that, in general, there does not exist an efficient algorithm to solve this problem in polynomial time. However, the information set decoding, which was introduced by Prange in 1962 [17], shows us an attack for the small parameters. Hence, to keep such system still secure, we need a sufficiently large size of keys.

In the same year, Delsarte [18] and seven years later, in 1985, Gabidulin [19], introduced and developed a new metric so-called a rank metric in [20]. Afterwards, Gabidulin, Paramonov and Tretjakov developed a new McEliece-like cryptosystem based on this metric, which is after called GPT-cryptosystem [21]. This cryptosystem uses the encryption and decryption algorithm of the Gabidulin codes. To compare with the other McEliece-like

cryptosystem based on Hamming metric, this one needs a smaller key size. Theorically, for the same set of parameters, it might give us a higher level of security.

Gabidulin codes

In 1962, Singleton introduced a bound for the minimum distance of a code [22]. The kind of codes whose minimum rank distance reaches Singleton bound attract a lot of interest and attention. In his own paper [19], Gabidulin compared the properties of Hamming metric and rank metric and found that in the case of rank metric we also have the Singleton bound for the minimum distance, the same as Hamming metric. As consequence, it leads to the research about the codes whose the minimum rank distance reaches Singleton bound (in case of Hamming metric, it is the so-called MDS codes and for rank metric, it is MRD code). In Section 1.2.1, Gabidulin showed a way to created an "equivalence" of Reed Solomon code in rank metric, called Gabidulin code. These codes are MRD with an efficient algorithm for the decoding up to its error capacity (for example, the Loidreau's algorithm in [8], which I have implemented in MAGMA, see `https://github.com/BaDucPham/RAMESSES/blob/main/DecGab.mgm`). These codes play an important role in the development of rank metric code based cryptography, especially the GPT-like cryptosystem. However, many of GPT-like systems were broken. The main weakness of these systems relies on the fact that Gabidulin codes are vulnerable against the invariant attack since they contain a huge vector space invariant by the action of the Frobenius automorphism.

Post quantum cryptography

The public key cryptosystem whose security relies on one of three hard problems: the integer factorization, the discrete logarithm in finite fields or the elliptic-curve discrete logarithm, can be broken on a sufficiently powerful quantum computer. It tends to the needs of the Post-quantum cryptography, which has attracted a lot of attention (PQCrypto conference since 2006 for example). For code based cryptography, one of the most well-known candidates for the post-quantum cryptography is McEliece cryptosystem.

With the growing probability of the existence of a near-future quantum computer, it has become important to propose alternatives to existing public-key encryption schemes and key exchange protocols based on number theory. The recent NIST Post-Quantum Cryptography Standardization process motivates proposals in this sense. Along with

lattice-based cryptography, code-based cryptography is the most represented among proposals for encryption schemes or key-encapsulation mechanisms (KEMs). Code-based submissions generically rely on the hardness of decoding problems, either in the Hamming metric or in the rank metric. Hamming metric decoding problems enjoy a long-standing study and few practical improvements for more than fifty years, which ascertain their security. On the opposite, rank metric decoding problems have been studied for less than twenty years [23], and their solving complexity is not yet fully stabilized (see the recent results of [24]). Nevertheless, they benefit from much shorter keys and seem very attractive for practical implementation, culminating in submissions for the NIST standardization process [25, 26]. So as to further reduce the key sizes, designers often use specific structures as quasi-cyclicity (equivalent of Module-LWE for lattices) which could be suspected to introduce additional weaknesses [27].

Organization of my thesis

In my thesis, I am interested in GPT-like cryptosystem. It was motivated from the study about some researche of my supervisor, professor Pierre Loidreau, in his cryptosystem, the Faure-Loidreau cryptosystem in [2] and the recent one in [7]. Unfortunately, these systems have been recently attacked, the former were attacked by Gaborit, Otmani and Kalachi in [3] and the latter in the paper of Coggia and Couveur [6]. All of these researches inspired me to study more about the structure of Gabidulin code and find a way to repair or modify the system to resist against the attack.

This thesis is organized as follows:

— chapter 1 The first chapter is about the extension of the Coggia and Couveur attack on Loidreau's cryptosystem. This attack is based on the idea about the weakness of Gabidulin code by the Overbeck's distinguisher, that we can distinguish Gabidulin codes from random ones. Moreover, the same as many other GPT-like cryptosystem, Loidreau's cryptosystem still shows the weakness in algebraic structure. Although it can perfectly avoid the direct distinguisher attack into the public code, Coggia and Couveur still exploited another distinguisher for the dual code of the public code for the secret parameter $\lambda = 2$. Recently, it was extended by Ghatak in [28] for the case of $\lambda = 3$ but it was incomplete. This chapter is about my work of the generalisation of this attack for the case of any λ and the consideration about the complexity of this attack and filling the gap for the case $\lambda = 3$ to show the

efficiency of the attack in this case. The implementation in MAGMA can be seen in `https://github.com/BaDucPham/Coggia-and-Couvreur-attack`

— chapter 2 The second chapter is about a new cryptosystem which is inspired from the Faure-Loidreau cryptosystem, in which, we consider the plaintext as the row space of an error. In this section, firstly, we will introduce a Augot-Finiasz-type cryptosytem and afterwards, study about its security by considering the complexity of some known attacks on it. Recently, this cryptosystem was broken under the Bombar-Couveur attack in [29]. Therefore, in the end, we will introduce the idea of Bombar and Couveur to break this system.

— chapter 3 The last chapter is my work about the structure of Gabidulin codes. From the idea of using reconstructing q-polynomials for the decoding of Gabidulin codes [8], our work is about the decryption of the sum of Gabidulin codes. This work also comes from an attempt to modify the Loidreau cryptosystem [7] to resist against the Coggia and Couveur attack [6]. Although there is no proof about the resistance of this modification from this attack but it still has some applications. In this chapter, I will introduce the polynomial time decoder for the sum of Gabidulin codes with an exponentially small failure probability and afterwards, I introduce some of its applications including an idea about the replacement of Gabidulin code by the sum of Gabidulin codes in the Loireau's cryptosystem and my own supposition about the resistance of this modification against the Coggia and Couveur attack.

The result of first chapter is a paper which is ongoing work and presented at CBCrypt 2021, the second one is a pre-print paper and the third one is published in the conference ISIT 2021 (International Symposium on Information Theory).

Chapter 1

An analysis of Coggia-Couvreur Attack on Loidreau's Rank-metric public-key encryption scheme in the general case

---

Since the use of $\mathbb{F}_{q^m}$-linear rank metric permits to design a short public key encryption scheme, one of the directions of code based cryptography consists in instantiating McEliece encryption scheme [30] with codes in rank metric, [1, 31].

Because of the structure of Gabidulin codes, any cryptosystem instantiated with codes containing Gabidulin codes not sufficiently scrambled was attacked [32]. In 2017, Loidreau proposed a scheme based on Gabidulin codes masked with a small dimensional vector space [7]. If the dimension of the vector space is two small, then there exists a very simple polynomial-time distinguishing algorithm.

The question was to know if distinguishing is enough to break. Coggia and Couvreur [6] showed that in the case where the dimension of the masking space is 2, a decryption procedure can be recovered in polynomial-time. This attack exploited a distinguisher on the dual of the public code. Their approach gives the possibility for cryptanalysis of rank-metric schemes for any λ.

In this work we show that this can be extended to any dimension. The attack is not necessary polynomial, but we include the previous results. Moreover we are able to prove rigorously under some assumptions the efficiency of the attack.

Besides, recently, the extended Coggia-Couveur's attack [28] claimed the attack for the case where the dimension of the masking space is 3 and in this paper, the author uses only one reduced polynomial to determine the secret subspace but their assumption is not clear in practice. In our way, instead of one polynomial equation, we use a system of polynomial equations. This approach gives us a proof for the equivalence between the set

of roots and the orbit of one root under the action of $\mathbf{PGL}(3, \mathbb{F}_q)$. From this, we complete the polynomial time key recovery attack in case where the dimension of the masking space is 3.

In this chapter, the first section gives some preliminaries and notations. Section 2 outlines Loidreau's scheme and describes the distinguisher between the dual of the public code and random codes in 3.1 and afterwards, the key recovery attack in 3.2. In 3.3, we analyze the complexity of the attack for the specific case $\lambda = 3$.

1 Preliminaries and notations

Let q be a power of a prime and let $\mathbb{F}_q$ denote the finite field of order q. We consider the finite field extension of degree m: $\mathbb{F}_{q^m}/\mathbb{F}_q$. We use $\mathbb{F}_q^{m\times n}$ to denote the set of all $m \times n$ matrices over $\mathbb{F}_q$ and $\mathbb{F}_{q^m}^n$ for the set of all row vectors of length n over $\mathbb{F}_{q^m}$.

Let θ be a generator of the Galois group $\mathrm{Gal}(\mathbb{F}_{q^m}/\mathbb{F}_q)$. For instance, it could be the mapping $x \mapsto x^q$, but everything we write stays true for any generator of the Galois group. Moreover, for simplicity, we denote by $x^{[i]}$ the value $\theta^i(x)$.

In this setting, we define the skew polynomial ring or Ore ring [33] denoted by $\mathbb{F}_{q^m}[X;\theta]$ by defining the usual operations

— Addition is classical addition ;

— $X \cdot a = \theta(a) \cdot X$

With these operations, this ring is left and right Euclidean. We denote by $P\langle X\rangle = \sum_{i=0}^{\ell} p_i X^i$ any element P of $\mathbb{F}_{q^m}[X;\theta]$ of degree ℓ to distinguish it from the usual polynomial ring.

There are several ways to define an evaluation map on this ring [34]. Here, we choose the so-called operator evaluation, meaning that for any α in some finite field where the action θ is meaningful (for instance any finite field with the same characteristic as $\mathbb{F}_q$), we have

$$\forall P \in \mathbb{F}_{q^m}[X;\theta], \quad P\langle\alpha\rangle \stackrel{def}{=} \sum_{i=0}^{\ell} p_i \theta^i(\alpha)$$

If θ corresponds to the Frobenius automorphism, then this evaluation corresponds to the evaluation of so-called ring of linearized polynomials defined in [34]. We naturally extend the notion of evaluation to a vector :

$$\forall \boldsymbol{y} = (y_1, \ldots, y_n), \quad P\langle\boldsymbol{y}\rangle = (P\langle y_1\rangle, \ldots, P\langle y_n\rangle)$$

Let $\boldsymbol{M}$ be a matrix over $\mathbb{F}_{q^m}$, we denote by $\mathrm{rk}_q(\boldsymbol{M})$ and $\mathrm{rk}_{q^m}(\boldsymbol{M})$ its rank over $\mathbb{F}_q$ and $\mathbb{F}_{q^m}$, which is the dimension of the vector space generated by the columns of $\boldsymbol{M}$ over $\mathbb{F}_q$ ($\mathbb{F}_{q^m}$ respectively)

Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m) \in \mathbb{F}_{q^m}^m$ be a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. We define the extension map

$$\begin{array}{rccc} \mathrm{Ext}_{\boldsymbol{\beta}} : & \mathbb{F}_{q^m}^n & \to & \mathbb{F}_q^{m\times n} \\ & \boldsymbol{a} = (a_1, \ldots, a_n) & \mapsto & \boldsymbol{A} = (\boldsymbol{\alpha}_1^\top, \ldots, \boldsymbol{\alpha}_n^\top) \end{array}$$

where, for all $1 \le j \le n$, the vector $\boldsymbol{\alpha}_j \in \mathbb{F}_q^m$ consists of coordinates of $a_j \in \mathbb{F}_{q^m}$ in the basis $\boldsymbol{\beta}$, *i.e.* $a_j = \sum_{i=1}^n \beta_i A_{i,j}$. In particular, for every $\boldsymbol{A} \in \mathbb{F}_q^{n\times n}$, we have $\mathrm{Ext}_{\boldsymbol{\beta}}(\boldsymbol{\beta}\boldsymbol{A}) = \boldsymbol{A}$.

The *rank* of $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$, denoted $\mathrm{rk}(\boldsymbol{a})$, is defined as $\mathrm{rk}_q(\mathrm{Ext}_{\boldsymbol{\beta}}(\boldsymbol{a}))$. Notice that $\mathrm{rk}(\boldsymbol{a})$ does not depend on the choice of the basis $\boldsymbol{\beta}$

In this setting, Gabidulin codes are defined as evaluation codes of skew polynomials over linearly independent elements.

Definition 1 ([35, 36])**.** *Let $\boldsymbol{g} = (g_1, \ldots, g_n) \in \mathbb{F}_{q^m}^n$, formed with $\mathbb{F}_q$-linearly independent elements. The Gabidulin code of dimension k and of support $\boldsymbol{g}$ denoted by $\mathcal{G}_k(\boldsymbol{g})$ is defined by*

$$\mathcal{G}_k(\boldsymbol{g}) = \left\{ f(\boldsymbol{g}), \quad \begin{array}{c} f \in \mathbb{F}_{q^m}[X;\theta] \\ \deg(f) \le k-1 \end{array} \right\}$$

In the following, and since we are in finite fields we will simply call them Gabidulin codes rather than Generalized Gabidulin codes.

2 The encryption scheme

2.1 Generalities

Let $\boldsymbol{G}$ a random generator matrix of a Gabidulin code $\mathcal{G}_k(\boldsymbol{g})$. Fix an integer $\lambda \le m$ and an $\mathbb{F}_q$-vector subspace $\mathcal{V}$ of $\mathbb{F}_{q^m}$ of dimension λ. Let $\boldsymbol{P} \in \boldsymbol{GL}(n, \mathbb{F}_{q^m})$ whose entries are all in $\mathcal{V}$. Then, let

$$\boldsymbol{G}_{\mathrm{pub}} = \boldsymbol{G}\boldsymbol{P}^{-1}$$

— KeyGen: Public key $(\boldsymbol{G}_{\mathrm{pub}}, t)$ where $t = \left\lfloor \frac{n-k}{2\lambda} \right\rfloor$
Secret key $(\boldsymbol{g}, \boldsymbol{P})$

— Encryption: Given a plaintext $\boldsymbol{m} \in \mathbb{F}_{q^m}^k$, choose $e \in \mathbb{F}_{q^m}^n$ of rank weight t. The ciphertext is:

$$\boldsymbol{c} = \boldsymbol{m}\boldsymbol{G}_{\text{pub}} + \boldsymbol{e}$$

— Decryption:

— Compute $\boldsymbol{c}\boldsymbol{P} = \boldsymbol{m}\boldsymbol{G} + \boldsymbol{e}\boldsymbol{P}$.

— Decode in $\mathcal{G}_k(\boldsymbol{g})$ and $\text{rk}(\boldsymbol{e}\boldsymbol{P}) \leq t\lambda \leq \frac{n-k}{2}$

Let us denote by $\mathcal{C}_{\text{pub}}$ the code generated by $\boldsymbol{G}_{\text{pub}}$ and by $\mathcal{C}_{\text{pub}}^{\perp}$, the dual code. Let $\boldsymbol{H}_{\text{pub}}$ be a generator matrix of $\mathcal{C}_{\text{pub}}^{\perp}$. It is immediate that

$$\boldsymbol{H}_{\text{pub}} = \boldsymbol{H}_{\text{sec}}\boldsymbol{P}^T$$

where $\boldsymbol{H}_{\text{sec}}$ is a parity-check matrix of $\mathcal{G}_k(\boldsymbol{g})$.

2.2 Goal of a reconstructing attack and solution set

Our main goal is to design a reconstructing attack from the knowledge of $\mathcal{C}_{\text{pub}}^{\perp}$ and under some particular sets of parameters.

W.l.o.g, one can suppose that $1 \in \mathcal{V}$. Suppose that $\mathcal{V} = \langle 1, \beta_1, \ldots, \beta_{\lambda-1} \rangle_{\mathbb{F}_q}$ for some $\{\beta_i\}_{i=1}^{\lambda-1} \in \mathbb{F}_{q^m} \backslash \mathbb{F}_q$. Therefore, $\boldsymbol{P}^T$ can be decomposed into

$$\boldsymbol{P}^T = \boldsymbol{P}_0 + \sum_{i=1}^{\lambda-1} \beta_i \boldsymbol{P}_i$$

where $\boldsymbol{P}_i$ are $n \times n$ matrices with entries in $\mathbb{F}_q$ not necessarily invertible.

Let $\mathcal{C}_{\text{sec}}^{\perp}$ the dual code of $\mathcal{G}_k(\boldsymbol{g})$. Thus, $\mathcal{C}_{\text{sec}}^{\perp} = \mathcal{G}_{n-k}(\boldsymbol{a})$ for some $\boldsymbol{a} \in \mathbb{F}_{q^m}^n$ with $\text{rk}(\boldsymbol{a}) = n$. We define

$$\boldsymbol{h}_0 = \boldsymbol{a}\boldsymbol{P}_0, \boldsymbol{h}_1 = \boldsymbol{a}\boldsymbol{P}_1, \ldots, \boldsymbol{h}_{\lambda-1} = \boldsymbol{a}\boldsymbol{P}_{\lambda-1}$$

Lemma 1. *The code $\mathcal{C}_{pub}^{\perp}$ is spanned by $\boldsymbol{h}_0^{[i]} + \sum_{j=1}^{\lambda-1} \beta_j \boldsymbol{h}_j^{[i]}$ for $i = 0, \ldots, n-k-1$*

Proof. For any $\boldsymbol{c} \in \mathcal{C}_{\text{pub}}^{\perp}$, there exists $P \in \mathbb{F}_{q^m}[X; \theta]$ of degree smaller than $n-k$ such that

$$\boldsymbol{c} = P\langle \boldsymbol{a} \rangle \boldsymbol{P}^T = P\langle \boldsymbol{a} \rangle \boldsymbol{P}_0 + \sum_{i=1}^{\lambda-1} \beta_i P\langle \boldsymbol{a} \rangle \boldsymbol{P}_i = P\langle \boldsymbol{h}_0 \rangle + \sum_{i=1}^{\lambda-1} \beta_i P\langle \boldsymbol{h}_i \rangle$$

□

Let us define the so-called solution set of the encryption scheme

Definition 2 (Solution set)**.** *The set $\mathcal{S}$ of all $(\boldsymbol{h}, \beta) \in (\mathbb{F}_{q^m}^n)^\lambda \times \mathbb{F}_{q^m}^{\lambda-1}$ such that*

$$\mathcal{C}_{pub}^{\perp} = \left\langle \boldsymbol{h}_0^{[i]} + \sum_{j=1}^{\lambda-1} \beta_j \boldsymbol{h}_j^{[i]},\ i = 0, \ldots, n-k-1 \right\rangle \tag{1.1}$$

where $\forall j = 0, \ldots, \lambda$, $\boldsymbol{h}_j$ has rank n and $\langle 1, \beta_1, \ldots, \beta_{\lambda-1} \rangle_{\mathbb{F}_q}$ has dimension λ is called solution set of the encryption scheme.

It is obvious that finding an element of the solution set $\mathcal{S}$ implies the ability to design a polynomial-time decryption algorithm. What we call a reconstructing attack corresponds to finding an element in $\mathcal{S}$. The solution set $\mathcal{S}$ has the following properties.

Proposition 1. *Let $(\boldsymbol{h}, \beta) \in (\mathbb{F}_{q^m}^n)^\lambda \times \mathbb{F}_{q^m}^{\lambda-1}$. Let $\boldsymbol{A} = (a_{j,i})_{j,i=0}^{\lambda-1} \in GL_\lambda(\mathbb{F}_q)$. Let us define the following group action on $(\mathbb{F}_{q^m}^n)^\lambda \times \mathbb{F}_{q^m}^{\lambda-1}$ by $\boldsymbol{A} \cdot (\boldsymbol{h}, \beta) = (\boldsymbol{h}', \beta')$ where*

$$\begin{cases} \boldsymbol{h}_j = \dfrac{a_{j,0}\boldsymbol{h}'_0 + \sum_{i=1}^{\lambda-1} a_{j,i}\boldsymbol{h}'_i}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i},\ j = 0, \ldots, \lambda-1 \\ \beta'_j = \dfrac{a_{0,j} + \sum_{i=1}^{\lambda-1} a_{i,j}\beta_i}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i},\ j = 1, \ldots, \lambda-1 \end{cases}$$

1. *Then if $(\boldsymbol{h}, \beta) \in \mathcal{S}$ we have $\boldsymbol{A} \cdot (\boldsymbol{h}, \beta) \in \mathcal{S}$*
2. *Moreover let $\overline{\boldsymbol{A}} = \{\boldsymbol{B} \in GL_\lambda(\mathbb{F}_q) | \exists c \in \mathbb{F}_q^*, \boldsymbol{B} = c\boldsymbol{A}\}$. Then, for any $\boldsymbol{B} \in \overline{\boldsymbol{A}}$, and for any $(\boldsymbol{h}, \beta) \in (\mathbb{F}_{q^m}^n)^\lambda \times \mathbb{F}_{q^m}^{\lambda-1}$ we have*

$$\boldsymbol{A} \cdot (\boldsymbol{h}, \beta) = \boldsymbol{B} \cdot (\boldsymbol{h}, \beta)$$

Proof. Let $(\boldsymbol{h}, \beta) \in \mathcal{S}$. Since from the definition of $\mathcal{S}$ the elements $1, \beta_1, \ldots, \beta_{\lambda-1}$ are $\mathbb{F}_q$-linearly independent, and since $\boldsymbol{A}$ is non singular, $(a_{0,0}, \ldots, a_{0,\lambda-1}) \neq 0$ this implies that $a_{0,0} + \sum_{i=1}^{\lambda-1} a_{0,j}\beta_i \neq 0$. Therefore, the elements $\beta'_1, \ldots, \beta'_{\lambda-1}$ are well defined and for all $0 \leq \ell \leq n - k - 1$,

$$
\begin{aligned}
&\boldsymbol{h}'^{[\ell]}_0 + \sum_{j=1}^{\lambda-1} \beta'_j \boldsymbol{h}'^{[\ell]}_j \\
=&\frac{1}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i} \left(\left(a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i \right) \boldsymbol{h}'^{[\ell]}_0 + \sum_{j=1}^{\lambda-1} \left(a_{0,j} + \sum_{i=1}^{\lambda-1} a_{i,j}\beta_i \right) \boldsymbol{h}'^{[\ell]}_j \right) \\
=&\frac{1}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i} \left(\left(a_{0,0}\boldsymbol{h}'_0 + \sum_{j=1}^{\lambda-1} a_{0,j}\boldsymbol{h}'_j \right)^{[\ell]} + \sum_{i=1}^{\lambda-1} \beta_i \left(a_{i,0}\boldsymbol{h}'_0 + \sum_{j=1}^{\lambda-1} a_{i,j}\boldsymbol{h}'_j \right)^{[\ell]} \right) \\
=&\left(a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i \right)^{[\ell]-1} \left(\boldsymbol{h}_0^{[\ell]} + \sum_{i=1}^{\lambda-1} \beta_i \boldsymbol{h}_i^{[\ell]} \right)
\end{aligned}
$$

Therefore

$$
\mathcal{C}^{\perp}_{\text{pub}} \overset{def}{=} \left\langle \boldsymbol{h}_0^{[\ell]} + \sum_{i=1}^{\lambda-1} \beta_i \boldsymbol{h}_i^{[\ell]},\ 0 \le \ell \le n-k-1 \right\rangle = \left\langle \boldsymbol{h}'^{[\ell]}_0 + \sum_{i=1}^{\lambda-1} \beta'_i \boldsymbol{h}'^{[\ell]}_i,\ 0 \le \ell \le n-k-1 \right\rangle
$$

Thus, $(\boldsymbol{h}', \beta') \in \mathcal{S}$.

For the second point of the proposition : Let $\boldsymbol{B} \in \overline{\boldsymbol{A}}$. It means that there exists $c \in \mathbb{F}_q^*$ such that $\boldsymbol{B}\boldsymbol{A}^{-1} = c\boldsymbol{I}$ or equivalently $\boldsymbol{B} = c\boldsymbol{A}$.

Let $(\boldsymbol{h}'', \beta'') \overset{def}{=} \boldsymbol{B} \cdot (\boldsymbol{h}', \beta')$, where $(\boldsymbol{h}', \beta') = \boldsymbol{A} \cdot (\boldsymbol{h}, \beta) \in (\mathbb{F}_{q^m}^n)^\lambda \times \mathbb{F}_{q^m}^{\lambda-1}$. It means that

$$
\begin{cases}
\boldsymbol{h}'_j = \dfrac{b_{j,0}\boldsymbol{h}''_0 + \sum_{i=1}^{\lambda-1} b_{j,i}\boldsymbol{h}''_i}{b_{0,0} + \sum_{i=1}^{\lambda-1} b_{i,0}\beta_i},\ j = 0, \dots, \lambda-1 \\
\beta''_i = \dfrac{b_{0,j} + \sum_{i=1}^{\lambda-1} b_{i,j}\beta_i}{b_{0,0} + \sum_{i=1}^{\lambda-1} b_{i,0}\beta_i},\ j = 1, \dots, \lambda-1
\end{cases}
$$

It is obvious that the actions of 2 matrix $\boldsymbol{A}, \boldsymbol{B}$ give us same image, which is $(\boldsymbol{h}'', \beta'') = (\boldsymbol{h}', \beta')$.

□

3 Attacks on the system

Now the section is organised as follows: In a first part we make a brief summary of the attack. Since it is technical, this section highlights the different principles. In first subsection, we introduce the distinguisher between the dual of the public code and the random codes. Afterwards, we make some assumptions in the beginning and under these assumptions, we exploit the attack based on the distinguisher introduced in 3.1. We also consider the special case of $\lambda = 3$ and analyze its complexity in the section 3.3.

We will also need to introduce a special setting to simplify the technicalities of the proofs. For any $(\boldsymbol{h}, \beta) \in \mathcal{S}$, we denote

$$\boldsymbol{y}_{(\boldsymbol{h},\beta)}^{[u,j]} = \boldsymbol{h}_0^{[j]} + \sum_{i=1}^{\lambda-1} \beta_i^{[u]} \boldsymbol{h}_i^{[j]}$$

for any integers (u, j). For a given $(\boldsymbol{h}, \beta)$, to simplify, we denote it by $\boldsymbol{y}^{[u,j]}$. Moreover, we can denote $\boldsymbol{y}^{[M]} = \{\boldsymbol{y}^{[u,j]}, (u, j) \in M \subset \mathbb{Z} \times \mathbb{Z}\}$. The codes that we will consider will be generated by the $\boldsymbol{y}^{[u,j]}$, where $(u, j) \in \mathbb{Z} \times \mathbb{Z}$. Let $I \subset \mathbb{Z} \times \mathbb{Z}$. We denote by

$$\mathcal{C}_I \overset{def}{=} \left\langle \boldsymbol{y}^{[u,j]}, (u, j) \in I \right\rangle,$$

if I is non-empty and $\mathcal{C}_\varnothing \overset{def}{=} \{0\}$. From the expression of $\mathcal{C}_{\text{pub}}^\perp$ in Definition 2, we have

$$\mathcal{C}_{\text{pub}}^\perp = \mathcal{C}_{\{0\} \times [0,\dots,n-k-1]} \tag{1.2}$$

Now we introduce a very fundamental theorem which will support all of our future proofs

Theorem 1 (CodeSet theorem)**.** *We have*

$$\forall I,\ J \subset \mathbb{Z} \times \mathbb{Z},\ \begin{cases} \mathcal{C}_{I \cup J} = \mathcal{C}_I + \mathcal{C}_J \\ \mathcal{C}_{I \cap J} \subset \mathcal{C}_I \cap \mathcal{C}_J \end{cases}$$

If moreover $M \subset \mathbb{Z} \times \mathbb{Z}$, where $\boldsymbol{y}^{[u,j]},\ (u, j) \in M$ are $\mathbb{F}_q^m$ linearly independent, then

- *for all $I \subset M$, $\dim(\mathcal{C}_I) = |I|$*
- *$\forall I,\ J \subset M,\ \mathcal{C}_I \cap \mathcal{C}_J = \mathcal{C}_{I \cap J}$*
- *$\forall I,\ J \subset M,\ \mathcal{C}_{I \sqcup J} = \mathcal{C}_I \oplus \mathcal{C}_J$, where $\sqcup$ means that the sets do not intersect.*

Proof. The code generating set for $\mathcal{C}_I + \mathcal{C}_J$ is the union of generating sets for $\mathcal{C}_I$ and $\mathcal{C}_J$, since $\boldsymbol{y}^{[I]}$ and $\boldsymbol{y}^{[J]}$ are generating sets respectively for $\mathcal{C}_I$ and $\mathcal{C}_J$ then $\boldsymbol{y}^{[I\cup J]}$ is a generating set for $\mathcal{C}_I + \mathcal{C}_J$. Hence $\mathcal{C}_{I\cup J} = \mathcal{C}_I + \mathcal{C}_J$. Now the generating set $\boldsymbol{y}^{[I\cap J]}$ of $\mathcal{C}_{I\cap J}$ is included in the generating set of $\mathcal{C}_I$ and of $\mathcal{C}_J$. Therefore $\mathcal{C}_{I\cap J} \subset \mathcal{C}_I \cap \mathcal{C}_J$.

Let us consider now M such that $\boldsymbol{y}^{[M]}$ is formed with linearly independent vectors. It is immediate that for any $I \subset M$, a basis of $\mathcal{C}_I$ is $\boldsymbol{y}^{[I]}$, therefore the dimension of $\mathcal{C}_I$ is exactly equal to $|I|$. Let $\boldsymbol{c} \in \mathcal{C}_I \cap \mathcal{C}_J$. We have

$$\boldsymbol{c} = \sum_{(u,j)\in I} c_{u,j}\boldsymbol{y}^{[u,j]} = \sum_{(u,j)\in I\setminus J} c_{u,j}\boldsymbol{y}^{[u,j]} + \sum_{(u,j)\in I\cap J} c_{u,j}\boldsymbol{y}^{[u,j]}$$

and similarly

$$\boldsymbol{c} = \sum_{(u,j)\in J} c'_{u,j}\boldsymbol{y}^{[u,j]} = \sum_{(u,j)\in J\setminus I} c'_{u,j}\boldsymbol{y}^{[u,j]} + \sum_{(u,j)\in I\cap J} c'_{u,j}\boldsymbol{y}^{[u,j]}$$

Since by hypotheses on M, the $\boldsymbol{y}^{[M]}$ are linearly independent, this implies the equality of the coefficients on this bases and thus that $c_{u,j} = 0$, for $(u,j) \in I \setminus J$. Therefore, $\boldsymbol{c} \in \mathcal{C}_{I\cap J}$.

The last item comes from the fact that is I and J do not intersect then $I \cap J = \varnothing$, thus $\mathcal{C}_I \cap \mathcal{C}_J = \{0\}$.

□

3.1 A distinguishing attack in the general case

We show that if n, k, λ satisfy $k > \frac{(\lambda-1)n}{\lambda} + 1$, then one can distinguish the public-code from a random code in polynomial time. First we prove the following theorem.

Theorem 2. $\dim_{\mathbb{F}_{q^m}} \left(\mathcal{C}_{pub}^{\perp} + \mathcal{C}_{pub}^{\perp\,[1]} + \cdots + \mathcal{C}_{pub}^{\perp\,[\lambda]}\right) \leq \lambda(n-k) + \lambda$

Proof. Let $\mathcal{S}_0 \overset{def}{=} \sum_{i=0}^{\lambda-1} \mathcal{C}_{\text{pub}}^{\perp\,[i]}$. We want to show that $\dim\left(\mathcal{S}_0 + \mathcal{C}_{\text{pub}}^{\perp\,[\lambda]}\right) \leq \lambda(n-k)+\lambda$. For any $(\boldsymbol{h}, \beta) \in \mathcal{S}$ from the expression of $\mathcal{C}_{\text{pub}}^{\perp}$ under the form (1.2) and from the CodeSet theorem we obtain

$$\mathcal{S}_0 = \mathcal{C}_{S_0}, \text{ where } S_0 = \bigsqcup_{u=0}^{\lambda-1} \{u\} \times [u, n-k+u-1]$$

and

$$\mathcal{C}_{\text{pub}}^{\perp\,[\lambda]} = \mathcal{C}_{\{\lambda\}\times[\lambda,n-k+\lambda-1]}$$

Let

$$I = \bigsqcup_{u=0}^{\lambda-1} \{u\} \times [\lambda - 1, n - k - 1] = [0, \lambda - 1] \times [\lambda - 1, n - k - 1]$$

We have clearly $I \subset S_0$, implying $\mathcal{C}_I \subset \mathcal{S}_0$.

By the hypotheses on $\mathcal{S}$ the $(1, \beta_1, \ldots, \beta_{\lambda-1})$ are linearly independent over $\mathbb{F}_q$. Thus,

$$\det \begin{bmatrix} 1 & \beta_1 & \ldots & \beta_{\lambda-1} \\ 1 & \beta_1^{[1]} & \ldots & \beta_{\lambda-1}^{[1]} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \beta_1^{[\lambda-1]} & \ldots & \beta_{\lambda-1}^{[\lambda-1]} \end{bmatrix} \neq 0,$$

it implies that for any $j \in \lambda - 1, \ldots, n - k - 1$, $\mathcal{C}_{[0,\lambda-1]\times\{j\}} = \langle \boldsymbol{h}_i^{[j]}, 0 \leq i \leq \lambda - 1 \rangle$. Hence,

$$\mathcal{C}_I = \left\langle \boldsymbol{h}_i^{[j]}, \begin{array}{l} 0 \leq i \leq \lambda - 1 \\ \lambda - 1 \leq j \leq n - k - 1 \end{array} \right\rangle$$

In particular from the structure of $\boldsymbol{y}^{[u,j]}$ for any $J \subset * \times [\lambda - 1, n - k - 1]$, we have $\mathcal{C}_J \subset \mathcal{C}_I \subset \mathcal{S}_0$.

Thus, $\mathcal{C}_{\{\lambda\}\times[\lambda,n-k-1]} \subset \mathcal{S}_0 \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda]}$. From its construction, ${\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda]}$ has dimension $n - k$. The vectors $\boldsymbol{y}^{[\lambda,j]}, j \in [\lambda, n - k + \lambda - 1]$ are linearly independent and from the CodeSet theorem, $\mathcal{C}_{\{\lambda\}\times[\lambda,n-k-1]}$ has dimension $n - k - \lambda$. Therefore, $\dim(\mathcal{S}_0 \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda]}) \geq n - k - \lambda$. Conversely,

$$\begin{aligned} \dim\left(\mathcal{S}_0 + {\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda]}\right) &= \dim(\mathcal{S}_0) + \dim({\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda]}) - \dim(\mathcal{S}_0 \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda]}) \\ &\leq \lambda(n - k) + (n - k) - (n - k - \lambda) = \lambda(n - k) + \lambda \end{aligned}$$

□

Now the distinguishing attack comes from this proposition

Proposition 2 ([6] Proposition 2)**.** *If $\mathcal{C}_{rand}$ is a random code of length n and dimension k, then for a non-negative integer a and a positive $\lambda < k$, we have*

$$\mathbb{P}\left(\dim_{\mathbb{F}_{q^m}}\left(\mathcal{C}_{rand} + \mathcal{C}_{rand}^{[1]} + \cdots + \mathcal{C}_{rand}^{[\lambda]}\right) \leq \min(n, (\lambda + 1)k) - a\right) = O(q^{-ma}).$$

Now whenever $k > \frac{(\lambda-1)n}{\lambda} + 1$, the dimension of $\mathcal{C}_{rand} + \mathcal{C}_{rand}^{[1]} + \cdots + \mathcal{C}_{rand}^{[\lambda]}$ is very probably equal to $(\lambda+1)(n-k)$ whereas the dimension of $\mathcal{C}_{\text{pub}}^{\perp} + \mathcal{C}_{\text{pub}}^{\perp\,[1]} + \cdots + \mathcal{C}_{\text{pub}}^{\perp\,[\lambda]}$ is probably equal to $\lambda(n-k+1)$ (since $\lambda(n-k+1) < n$) , which is strictly less than $(\lambda+1)(n-k)$.

3.2 Reconstructing attack

We suppose that the public code has rate larger than $(\lambda-1)/\lambda$, so that the distinguisher introduced in Section 3.1 works on it.

Although the attack we describe should work heuristically, to have rigorous proofs of work we need the following assumptions, which are not very contraining

(1) There exists an element $(\boldsymbol{h}, \beta) \in \mathcal{S}$ such that $\forall i_1, \ldots, i_\lambda \in \{1, \ldots, n-k-1\}$ distinct.

$$\det \begin{bmatrix} 1 & \beta_1^{[i_1]} & \beta_2^{[i_1]} & \ldots & \beta_{\lambda-1}^{[i_1]} \\ 1 & \beta_1^{[i_2]} & \beta_2^{[i_2]} & \ldots & \beta_{\lambda-1}^{[i_2]} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \beta_1^{[i_\lambda]} & \beta_2^{[i_\lambda]} & \ldots & \beta_{\lambda-1}^{[i_\lambda]} \end{bmatrix} \neq 0,$$

(2) $\dim_{\mathbb{F}_{q^m}} \mathcal{C}_{\text{pub}}^{\perp} + \mathcal{C}_{\text{pub}}^{\perp\,[1]} + \mathcal{C}_{\text{pub}}^{\perp\,[2]} + \cdots + \mathcal{C}_{\text{pub}}^{\perp\,[\lambda]} = \lambda(n-k) + \lambda$

(3) There is no $\overline{\boldsymbol{A}} \in \mathbf{PGL}(\lambda, \mathbb{F}_q) \backslash \overline{\boldsymbol{I}_\lambda}$ and $\boldsymbol{A} = (a_{ij})_{i,j=1}^{\lambda}$ that satisfies

$$\beta_j = \frac{a_{0,j} + \sum_{i=1}^{\lambda-1} a_{i,j}\beta_i}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i}, \ \forall j = 1, \ldots, \lambda-1$$

The first step of the attack is dedicated to finding one dimensional vector-spaces $\mathcal{A}_i$ for $i = 1, \ldots, n-k-1$, such that any element $(\boldsymbol{h}, \beta) \in \mathcal{S}$ satisfies:

$$\mathcal{A}_i = \left\langle \boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-i]} \boldsymbol{h}_j \right\rangle$$

From the $\mathcal{A}_i$'s, one obtains a system of $\lambda-1$ multivariate polynomials which are of degree $q^{\lambda+1} - q^i$ for $i = 1, \ldots, \lambda-1$ satisfied by all the vectors β such that $(\boldsymbol{h}, \beta) \in \mathcal{S}$.

Under the above assumptions, we can also prove the stabilization of the set of solution $\mathcal{S}$ under the action of $\mathbf{PGL}(\lambda, \mathbb{F}_q)$ in the end of 3.2.2.

The complexities of the steps (by operations over $\mathbb{F}_{q^m}$):

— Step 1. It costs $O(n^3 \log q)$ operations for computing ${\mathcal{C}_{\text{pub}}^{\perp}}^{[i]}$ and $O(n^{\omega+1})$ for taking the intersection.

— Step 2. The principal complexity of this is finding the roots of the system of polynomials. In case of $\lambda = 3$, it can be done in polynomial time where the complexity is $\tilde{O}(\tilde{d}^2 n \log q)$ for $\tilde{d} = (q^4 - q)(q^4 - q^2)$.

Once such a root is found, the remaining of step 2 needs a finite number of linear systems solving which costs $O(n^{\omega})$.

3.2.1 First step: Recovering one-dimensional vector spaces

We now suppose that the three assumptions in section 2.2 are true we have the following theorem:

Theorem 3. *Let $d := n - k - \lambda + 1$. Under Assumptions (1), (2), (3), Algorithm (1)*

Algorithm 1: Recovering 1-dimensional vector spaces

Input: $\mathcal{C}_{\text{pub}}^{\perp}$, $\lambda \leq \dfrac{n}{n-k+1}$

Output: $\mathcal{A}_i$ for $i = 0, \ldots, n-k-1$

$\mathcal{S}_0 \leftarrow {\mathcal{C}_{\text{pub}}^{\perp}}^{[0]} + {\mathcal{C}_{\text{pub}}^{\perp}}^{[1]} + \cdots + {\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda-1]}$

$\mathcal{A} \leftarrow \left(\bigcap_{i=0}^{d} \mathcal{S}_0^{[i]} \right)^{[-d]}$

$\mathcal{D}_{\lambda-1} \leftarrow \mathcal{A}^{[\lambda-2]} \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[\lambda-1-d]}$ and $\mathcal{B}_0 \leftarrow \mathcal{A} + \mathcal{D}_{\lambda-1}^{[1-\lambda]}$

$\mathcal{D}_0 \leftarrow \mathcal{B}_0 \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[-1]}$

for $\ell \in 1, \ldots, \lambda - 2$ **do**

  $\mathcal{B}_\ell \leftarrow \mathcal{A} + \sum_{j=0}^{\ell-1} \mathcal{D}_j^{[\ell-j]}$;

  $\mathcal{D}_\ell \leftarrow \mathcal{B}_\ell \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[-1]}$

$\mathcal{H} \leftarrow \sum_{j=0}^{\lambda-1} \mathcal{C}_j^{[2-j-\lambda]}$

for $i \in 0, \ldots, n-k-1$ **do**

  Return $\mathcal{A}_i \leftarrow \mathcal{H} \cap {\mathcal{C}_{\text{pub}}^{\perp}}^{[-i]}$

returns the 1*-dimensional vector spaces*

$$\mathcal{A}_i = \left\langle \boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-i]} \boldsymbol{h}_j \right\rangle, \ i = 0, \ldots, n-k-1$$

for any $(\boldsymbol{h}, \beta) \in \mathcal{S}$.

Proof. For the proof we will thus make intensive use of the CodeSet theorem. First from assumption (2) and theorem 2, the set

$$M = \bigsqcup_{u=0}^{\lambda} \{u\} \times [u, n-k+u-1] = S_0 \sqcup \{\lambda\} \times [\lambda, n-k+\lambda-1]$$

with cardinality $\lambda(n-k)+\lambda$, is such that $\boldsymbol{y}^{[M]}$ is formed of linearly independent vectors and $\mathcal{C}_M = \sum_{i=0}^{\lambda} \mathcal{C}_{\text{pub}}^{\perp\ [i]}$. This point is very important since this is the crucial point of the proof of the theorem.

Line 1. From theorem 2 we have

$$\mathcal{S}_0 = \mathcal{C}_{S_0}, \text{ where } S_0 = \bigsqcup_{u=0}^{\lambda-1} \{u\} \times [u, n-k+u-1]$$

We can write S_0 under the form

$$S_0 = \underbrace{\left(\bigsqcup_{u=0}^{\lambda-2} \{u\} \times [u, \lambda-2] \right)}_{I_1} \sqcup \underbrace{[0, \lambda-1] \times [\lambda-1, n-k-1]}_{I_2} \sqcup \underbrace{\left(\bigsqcup_{u=0}^{\lambda-1} \{u\} \times [n-k, n-k+u-1] \right)}_{I_3}$$

Let $I_4 = I_3 \sqcup \{\lambda\} \times [n-k, n-k+\lambda-1]$. With these notations, we have

$S_0 = I_1 \sqcup I_2 \sqcup I_3 \subset I_1 \sqcup I_2 \sqcup I_4 = M$. Since their cardinalities satisfy

$|I_1| = \frac{\lambda(\lambda-1)}{2}$, $|I_2| = \lambda(n-k-\lambda+1)$, $|I_3| = \frac{\lambda(\lambda-1)}{2}$ and $|I_4| = \lambda$,

from theorem 2 the dimension of $\mathcal{S}_0$ and $\mathcal{C}_M$ is exactly $\lambda(n-k)$ and $\lambda(n-k)+\lambda$ respectively, and additionally under the CodeSet theorem,

$$\mathcal{S}_0 = \mathcal{C}_{I_1} \oplus \mathcal{C}_{I_2} \oplus \mathcal{C}_{I_3}$$
$$\mathcal{C}_M = \mathcal{C}_{I_1} \oplus \mathcal{C}_{I_2} \oplus \mathcal{C}_{I_4}$$

The set I_2 corresponds to the set denoted by I in the proof of theorem 2. We have

$$\mathcal{C}_{I_2} = \left\langle \boldsymbol{h}_i^{[j]}, \begin{array}{l} 0 \le i \le \lambda - 1 \\ \lambda - 1 \le j \le n - k - 1 \end{array} \right\rangle$$

This property gives us the flexibility for the modification of the set M to obtain several sets of indexes M' such that $\boldsymbol{y}^{[M']}$ is formed of linearly independent vectors. It can be done by the replacement of the set $[0, \lambda - 1]$ by the set A_j of λ elements corresponding to any j. We can see it precisely as the following lemma:

Lemma 2. *For every set* $I_2' = \bigsqcup_{j=\lambda-1}^{n-k-1} A_j \times \{j\}$ *where* $|A_j| = \lambda$, *then* $M' = I_1 \sqcup I_2' \sqcup I_4$ *satisfies*

— $\mathcal{C}_M = \mathcal{C}_{M'}$.

— $\boldsymbol{y}^{[M']}$ *is formed of linearly independent vectors.*

Proof. $\mathcal{C}_{I_2'} = \left\langle \boldsymbol{h}_i^{[j]}, \begin{array}{l} 0 \le i \le \lambda - 1 \\ \lambda - 1 \le j \le n - k - 1 \end{array} \right\rangle = \mathcal{C}_{I_2}$ (Assumption (1)). Moreover, $|I_2'| = |I_2| = \lambda(n - k - \lambda + 1)$. Hence $\mathcal{C}_{M'} = \mathcal{C}_{I_1} \oplus \mathcal{C}_{I_2'} \oplus \mathcal{C}_{I_4} = \mathcal{C}_M$ and $\boldsymbol{y}^{[M']}$ is formed of linearly independent vectors. □

Through this section, to apply the CodeSet theorem, in the beginning of each step, we will define its set of indexes M' such that $\boldsymbol{y}^{[M']}$ are linearly independent vectors and it contains the set of indexes of subspace that we want to compute the intersection. To be convenient, we will use some images where the red dot • are indexes of some transformation of I_1, the blue square ▪ are indexes of some transformation of I_4, the green × are indexes of some transformation of I_2 and the black diamond ◆ are indexes of $\mathcal{C}_{\text{pub}}^{\perp}$. On the other hand, the integer points which are inside the blue figures are indexes of linearly independent vectors. The left triangular covers all the points of the set of indexes I_1, the right triangular covers all the points of the set of indexes I_4 and the rectangular covers all the points of the set of indexes which is flexible modification of I_2.

To be convenient, for $I \subset \mathbb{Z}\times\mathbb{Z}$ and $a \in \mathbb{Z}$, we denote $I + a = \{(u+a, j+a),\ (u,j) \in I\}$. In the figures, we can consider $I+a$ as the translation of I by the vector (a, a). For example, in the figure 1.1, the set of red points shows I_1 in the first image and $I_1 + d + 1$ in the second ones.

Line 2. We show that $\mathcal{A} = \mathcal{C}_{I_1 \sqcup I_3 + (\lambda - (n-k) - 1)}$.

Lemma 3. *Let* $\mathcal{S}_i \overset{def}{=} \mathcal{C}_{pub}^{\perp\ [i]} + \mathcal{C}_{pub}^{\perp\ [i+1]} + \cdots + \mathcal{C}_{pub}^{\perp\ [i+\lambda-1]}$. *For any set* $*$ *of* λ *distinct integers modulo* m *we have*

$$\forall 0 \leq d \leq n-k-\lambda+1,\ \bigcap_{i=0}^{d} \mathcal{S}_i = \mathcal{C}_{(I_1+d)\sqcup *\times[\lambda-1+d,n-k-1]\sqcup I_3}$$

Proof. We prove the theorem by induction. This lemma is true for $d = 0$. We suppose that it is true until $0 \leq d \leq n-k-\lambda$, then we need to prove that it must be true for $d+1$.

Indeed,

$$\begin{aligned}\bigcap_{i=0}^{d+1} \mathcal{S}_i &= \mathcal{S}_0 \cap \left(\bigcap_{i=0}^{d} \mathcal{S}_i\right)^{[1]} \\ &= \mathcal{C}_{S_0} \cap \mathcal{C}_{(I_1+d+1)\cup *\times[\lambda+d,n-k]\cup(I_3+1)}\end{aligned}$$

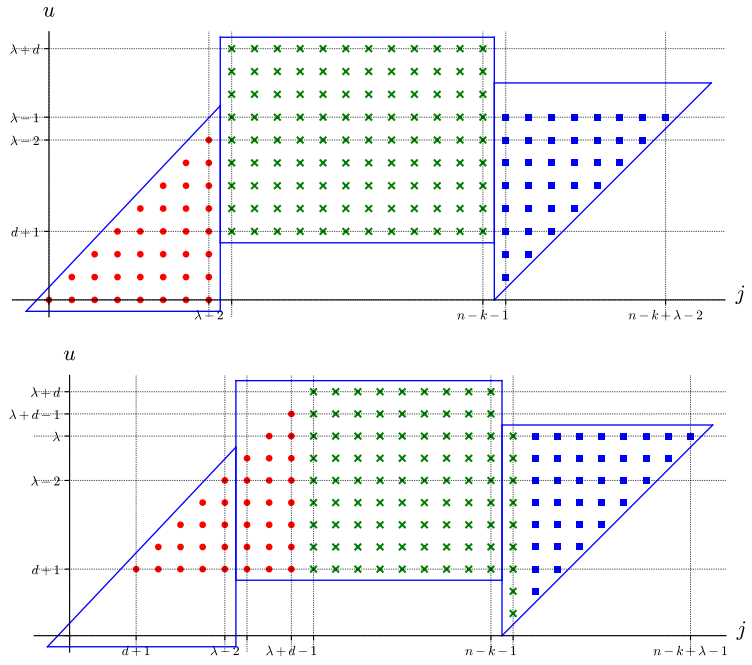

Figure 1.1 – Points of $\mathcal{S}_0$ (above) and $(I_1 + d + 1) \cup * \times [\lambda + d, n-k] \cup (I_3 + 1)$

Let $M_1 = I_1 \sqcup I_4 \sqcup [d+1, \lambda+d] \times [\lambda-1, n-k-1]$

Concerning $S_0 \subset M_1$ and $J \overset{def}{=} (I_1 + d + 1) \cup * \times [\lambda + d, n-k] \cup (I_3 + 1) \subset M_1$. We can apply the CodeSet theorem

$$\mathcal{C}_{S_0} \cap \mathcal{C}_J = \mathcal{C}_{S_0 \cap J}$$

And by a slightly fastidious computation on the sets intersections, we see that

$$S_0 \cap J = (I_1 + d + 1) \cup * \times [\lambda + d, n - k - 1] \cup I_3$$

It is not very difficult to check that the sets do not intersect which gives the result.

□

In the rest of the proof we will suppose that $d = n - k - \lambda + 1$. If we instantiate the lemma with d and elevate to the power $[-d]$ we obtain the following corollary:

Corollary 1. $\mathcal{A} = \mathcal{C}_{I_1 \sqcup I_3 - d}$

Proof. We have $\bigcap_{i=0}^{d} \mathcal{S}_i = \mathcal{C}_{(I_1+d)\sqcup I_3}$, with I_1 and I_3 subsets of M. Thus, from CodeSet theorem we have

$$\mathcal{A}^{[d]} = \mathcal{C}_{I_1+d} \oplus \mathcal{C}_{I_3}$$

Implying that $\mathcal{A} = \mathcal{C}_{I_1} \oplus \mathcal{C}_{I_3 - d}$.

□

Line 3. Since $\mathcal{C}_{\text{pub}}^{\perp} = \mathcal{C}_{\{0\}\times[0,n-k-1]}$, we have

$$\mathcal{A}^{[d-1]} \cap \mathcal{C}_{\text{pub}}^{\perp} = \mathcal{C}_{(I_1+(d-1))\sqcup(I_3-1)} \cap \mathcal{C}_{\{0\}\times[0,n-k-1]}$$

with

$$I_1 + (d-1) = \{(u, j) : d - 1 \le u \le j \le n - k - 2\}$$
$$I_3 - 1 = \{(u, j) : 0 \le u \le \lambda - 2,\ n - k - 1 \le j \le n - k + u - 1\}$$

Let

$$I_2' = (\{0\} \sqcup [n - k - \lambda, n - k - 2]) \times [\lambda - 1, n - k - 2]) \sqcup [0, \lambda - 1] \times \{n - k - 1\}$$

and $M_2 = I_1 \sqcup I_4 \sqcup I_2'$

We can prove that $I_1 + (d-1), I_3 - 1$ and $\{0\} \times [0, n-k-1]$ are all in M_2. Now since $\lambda \le (n-k)/2$, we have $d \ge \lambda - 1$, it implies that

$$((I_1 + (d-1) \sqcup (I_3 - 1)) \cap \{0\} \times [0, n-k-1] = \{(0, n-k-1)\}$$

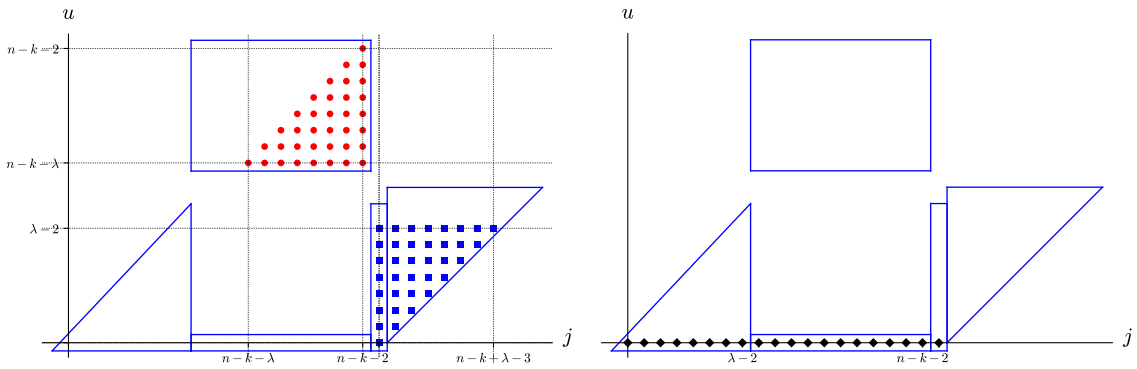

Figure 1.2 – Points of $(I_1 + (d-1)) \sqcup (I_3 - 1)$ and $\{0\} \times [0, n-k-1]$

Since the intersecting sets are all subsets of the set M_2, we can apply the CodeSet theorem, and we obtain

$$\mathcal{A}^{[d-1]} \cap \mathcal{C}_{\text{pub}}^{\perp} = \mathcal{C}_{\{(0,n-k-1)\}}$$

which by elevating to the power $[\lambda - d - 1]$ gives

$$\mathcal{D}_{\lambda-1} = \mathcal{A}^{[\lambda-2]} \cap \mathcal{C}_{\text{pub}}^{\perp\ [\lambda-d-1]} = \mathcal{C}_{\{(\lambda-d-1,2\lambda-3)\}}$$

Note that elevating to the power $[m]$ the scalars corresponds to the identity operator, we also have $\mathcal{D}_{\lambda-1} = \mathcal{C}_{\{(m+\lambda-d-1,2\lambda-3)\}}$. This will be of use in the proof of the algorithm. Now since $\mathcal{B}_0 = \mathcal{A} + \mathcal{D}_{\lambda-1}^{[1-\lambda]}$, we deduce that

$$\mathcal{B}_0 = \mathcal{C}_{I_1 \sqcup I_3 - d} + \mathcal{C}_{\{(-d,\lambda-2)\}}$$

Line 4. Compute $\mathcal{D}_0 = \mathcal{B}_0 \cap \mathcal{C}_{\text{pub}}^{\perp\ [-1]}$

We compute

$$\begin{aligned}
\mathcal{D}_0^{[1]} &= \mathcal{B}_0^{[1]} \cap \mathcal{C}_{\text{pub}}^{\perp} \\
&= \left(\mathcal{C}_{(I_1+1)\sqcup I_3-(d-1)} + \mathcal{C}_{\{(-d+1,\lambda-1)\}}\right) \cap \mathcal{C}_{\{0\}\times[0,(n-k)-1]} \\
&= \left(\mathcal{C}_{(I_1+1)\setminus[1,\lambda-1]\times\{\lambda-1\}} \oplus \mathcal{C}_{*\times\{\lambda-1\}} \oplus \mathcal{C}_{I_3-(d-1)}\right) \cap \mathcal{C}_{\{0\}\times[0,(n-k)-1]}
\end{aligned}$$

where * is instantiated for the set $[0, \lambda-2] \sqcup \{-d+1\}$ which contains λ different integers.

Let $I_2' = ([0, \lambda-1] \times \{\lambda-1\}) \sqcup ([2-d, \lambda-d] \times [\lambda, n-k-1])$ and $M_3 = I_1 \sqcup I_4 \sqcup I_2'$

Now $(I_1 + 1) \sqcup I_3 - (d-1) \sqcup (-d+1, \lambda-1) \subset M_3$. Moreover,

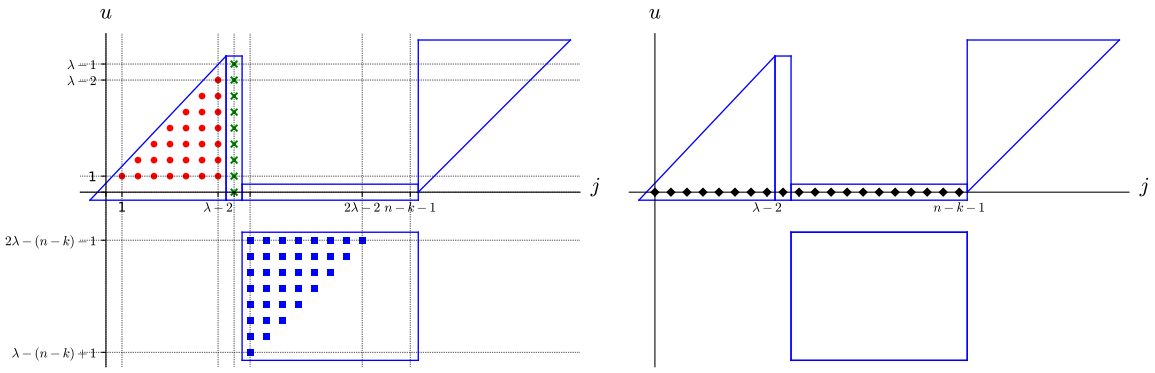

Figure 1.3 – Points of $((I_1 + 1)\backslash[1, \lambda - 1] \times \{\lambda - 1\}) \sqcup * \times \{\lambda - 1\} \sqcup I_3 - (d - 1)$ and $\{0\} \times [0, (n - k) - 1]$

$$((I_1 + 1) \sqcup I_3 - (d - 1)) \cap \{0\} \times [0, (n - k) - 1] = \varnothing$$

and $* \times \{\lambda - 1\} \cap \{0\} \times [0, n - k - 1] = \{(0, \lambda - 1)\}$

Hence, by the CodeSet theorem,

$$\mathcal{B}_0^{[1]} \cap \mathcal{C}_{\text{pub}}^{\perp} = \mathcal{C}_{\{(0,\lambda-1)\}}$$
$$\mathcal{B}_0 \cap \mathcal{C}_{\text{pub}}^{\perp\ [-1]} = \mathcal{C}_{\{(m-1,\lambda-2)\}}$$

Line 5,6,7 For $1 \leq i \leq \lambda - 2$, we compute

$$\mathcal{B}_i = \mathcal{A} + \sum_{j=0}^{i-1} \mathcal{C}_j^{[i-j]}$$
$$\mathcal{D}_i = \mathcal{B}_i \cap \mathcal{C}_{\text{pub}}^{\perp\ [-1]} = \mathcal{C}_{\{(m-1,\lambda+i-2)\}}$$

We prove by induction, suppose that $\mathcal{D}_i = \mathcal{C}_{\{(m-1,\lambda+i-2)\}}$ for all $1 \leq i \leq \ell \leq \lambda - 2$. we prove this for $i = \ell + 1$

$$\mathcal{B}_{\ell+1} = \mathcal{A} + \sum_{j=0}^{\ell} \mathcal{C}_j^{[\ell+1-j]}$$
$$\mathcal{D}_{\ell+1} = \mathcal{B}_{\ell+1} \cap \mathcal{C}_{\text{pub}}^{\perp\ [-1]} = \mathcal{C}_{\{(m-1,\lambda+\ell-1)\}}$$

Indeed,

$$\begin{aligned}\mathcal{B}_{\ell+1} &= \mathcal{A} + \sum_{j=0}^{\ell} \left(\mathcal{C}_{\{(m-1,\lambda+j-2)\}}\right)^{[\ell+1-j]} \\ &= \mathcal{A} + \sum_{j=0}^{\ell} \mathcal{C}_{\{(\ell-j,\lambda+\ell-1)\}} \\ &= \mathcal{C}_{I_1 \sqcup I_3 - d} + \mathcal{C}_{[0,\ell]\times\{\lambda+\ell-1\}} \\ &= \mathcal{C}_{I_1 \sqcup (I_3 - d \setminus [\ell-d+1,\lambda-d-1]\times\{\lambda+\ell-1\})} + \mathcal{C}_{*\times\{\lambda+\ell-1\}}\end{aligned}$$

where $*$ is instantiated for the set $[\ell - d + 1, \lambda - d - 1] \sqcup [0, d]$ of λ distinct integers.

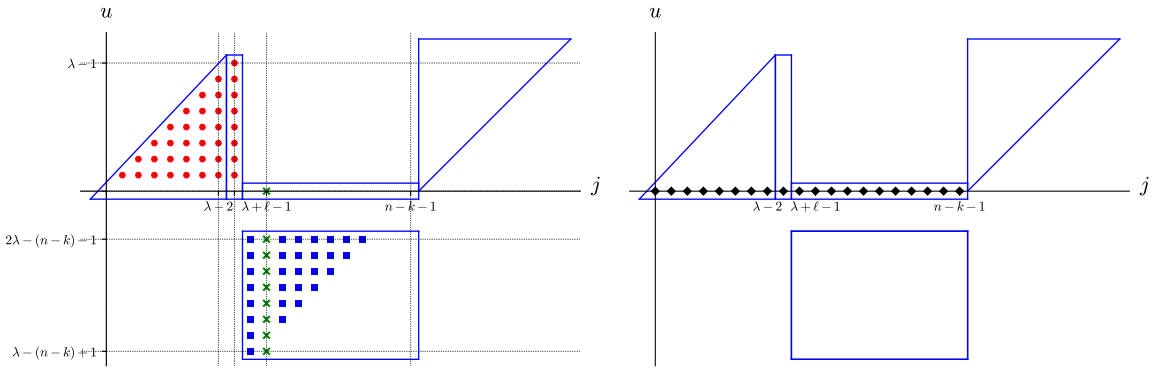

Figure 1.4 – Points of $I_1 \sqcup (I_3 - d \setminus [\ell - d + 1, \lambda - d - 1] \times \{\lambda + \ell - 1\}) \sqcup * \times \{\lambda + \ell - 1\}$ and $\{0\} \times [0, (n-k) - 1]$

We compute $\mathcal{D}_{\ell+1}^{[1]} = \mathcal{B}_{\ell+1}^{[1]} \cap \mathcal{C}_{\text{pub}}^{\perp}$. Let

$$I_2' = ([0, \lambda - 1] \times \{\lambda - 1\}) \sqcup (\{0\} \sqcup [2 - d, \lambda - d]) \times [\lambda, n - k - 1]$$

and $M_4 = I_1 \sqcup I_4 \sqcup I_2'$. Then, $I_1 + 1, I_3 - d + 1$ and $* \times \{\lambda + \ell + 1\} \subset M_4$.

Moreover,

$$\begin{aligned} I_3 - d + 1 \cap \{0\} \times [0, n - k - 1] &= \varnothing \\ I_1 + 1 \cap \{0\} \times [0, n - k - 1] &= \varnothing \end{aligned}$$

Hence, by CodeSet theorem,

$$\begin{aligned}\mathcal{B}_{\ell+1}^{[1]} \cap \mathcal{C}_{\text{pub}}^{\perp} &= \mathcal{C}_{*\times[\lambda+\ell]\cap\{0\}\times[0,n-k-1]} \\ &= \mathcal{C}_{\{(0,\lambda+\ell)\}}\end{aligned}$$

Therefore, $\mathcal{D}_{\ell+1} = \mathcal{B}_{\ell+1} \cap \mathcal{C}_{\text{pub}}^{\perp\ [-1]} = \mathcal{C}_{\{(m-1,\lambda+\ell-1)\}}$

Line 8. Compute $\mathcal{H} = \sum_{j=0}^{\lambda-1} \mathcal{C}_j^{[2-j-\lambda]}$.

We consider the sum of subspace:

$$\begin{aligned}&\sum_{j=0}^{\lambda-1} \mathcal{C}_j^{[2-j-\lambda]} = \sum_{j=0}^{\lambda-2} \mathcal{C}_{\{(m+1-\lambda-j,0)\}} + \mathcal{C}_{\{(m+1-(n-k),0)\}} \\ &= \mathcal{C}_{([m+3-2\lambda,m+1-\lambda]\sqcup\{m+1-(n-k)\})\times\{0\}} = \mathcal{C}_{*\times\{0\}} =: \mathcal{H}\end{aligned}$$

where * is instantiated for the set $[m+2-2\lambda, m+1-\lambda] \sqcup \{m+1-(n-k)\}$ of λ distinct integers.

Line 9-10 Next, for any $i \in \{0, \dots n-k-1\}$, one can compute

$$\mathcal{C}_{\text{pub}}^{\perp\ [-i]} \cap \mathcal{H} = \langle \boldsymbol{h}_0 + \sum_{i=1}^{\lambda-1} \beta_i^{[-i]} \boldsymbol{h}_i \rangle$$

— For $\lambda - 1 \le i \le n-k-1$, $* \times \{i\} \subset M$

$$\{(0,i)\} = (\{0\} \times [0, n-k-1]) \cap (* \times \{i\})$$

Hence, by the CodeSet theorem, $\mathcal{C}_{(0,i)} = \mathcal{C}_{\text{pub}}^{\perp} \cap \mathcal{H}^{[i]}$

— For $0 \le i \le \lambda-2$, $\mathcal{C}_{\text{pub}}^{\perp\ [\lambda-1]} = \mathcal{C}_{\{\lambda-1\}\times[\lambda-1,n-k+\lambda-2]}$ and $\mathcal{H}^{[\lambda+i-1]} = \mathcal{C}_{*\times\{\lambda+i-1\}}$. Moreover, $* \times \{\lambda+i-1\} \subset * \times [\lambda-1, 2\lambda-3] \subset M$ and

$$\{(\lambda-1, \lambda+i-1)\} = (\{\lambda-1\} \times [\lambda-1, n-k+\lambda-2]) \cap (* \times \{\lambda+i-1\}))$$

Hence, by the CodeSet theorem,

$$\begin{aligned}\mathcal{C}_{\text{pub}}^{\perp\ [\lambda-1]} \cap \mathcal{H}^{[\lambda+i-1]} &= \mathcal{C}_{\{(\lambda-1,\lambda+i-1)\}} \\ \mathcal{C}_{\text{pub}}^{\perp} \cap \mathcal{H}^{[i]} &= \mathcal{C}_{\{(0,i)\}}\end{aligned}$$

Therefore, for any $i \in \{0, \dots n-k-1\}$, one can compute

$$\mathcal{C}_{\text{pub}}^{\perp\ [-i]} \cap \mathcal{H} = \mathcal{C}_{\{(-i,0)\}} = \langle \boldsymbol{h}_0 + \sum_{i=1}^{\lambda-1} \beta_i^{[-i]} \boldsymbol{h}_i \rangle$$

□

Note that this specialization of one element of $\mathcal{S}$ should be true for any element in $\mathcal{S}$. Indeed, for 2 elements $(\boldsymbol{h}', \beta')$ and $(\boldsymbol{h}, \beta)$, if there exists $\boldsymbol{A} \in GL(\lambda, \mathbb{F}_q)$ such that $(\boldsymbol{h}', \beta') = \boldsymbol{A} \cdot (\boldsymbol{h}, \beta)$, then $\langle \boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-i]} \boldsymbol{h}_j \rangle = \langle \boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} {\beta'}_j^{[-i]} \boldsymbol{h}'_j \rangle$ since $\boldsymbol{h}'_0 + \sum_{j=1}^{\lambda-1} {\beta'}_j^{[-\ell]} \boldsymbol{h}'_j = \left(a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i\right)^{1-[-\ell]} \left(\boldsymbol{h}_0 + \sum_{i=1}^{\lambda-1} \beta_i^{[-\ell]} \boldsymbol{h}_i\right)$

3.2.2 Second step: Recovering the vector space

From step 1, we recovered the 1-dimensional vector-spaces

$$\forall i = 0, \dots, n-k-1,\ \mathcal{A}_i = \left\langle \boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-i]} \boldsymbol{h}_j \right\rangle$$

The vector spaces $\mathcal{A}_i$ do not depend on $(\boldsymbol{h}, \beta) \in \mathcal{S}$. We introduce the following lemma.

Lemma 4. *For any $\boldsymbol{u}_0 \in \mathcal{A}_0$, and for any set $\mathcal{I} = \{i_1, \dots, i_\lambda\} \subset \{1, \dots, n-k-1\}$ of λ distinct elements, there exists a unique λ-tuple $\boldsymbol{u}_\mathcal{I} \overset{def}{=} (\boldsymbol{u}_{i_1}, \boldsymbol{u}_{i_2}, \dots, \boldsymbol{u}_{i_\lambda}) \in \bigtimes_{j=1}^{\lambda} \mathcal{A}_{i_j}$ such that*

$$\sum_{i_j \in \mathcal{I}} \boldsymbol{u}_{i_j} = \boldsymbol{u}_0$$

Proof. We observe that, from assumption (1) we have

$$\mathcal{A}_{i_1} \oplus \cdots \oplus \mathcal{A}_{i_\lambda} = \langle \boldsymbol{h}_0, \dots, \boldsymbol{h}_{\lambda-1} \rangle.$$

Since $\mathcal{A}_0 \subset \langle \boldsymbol{h}_0, \dots, \boldsymbol{h}_{\lambda-1} \rangle$, this completes the proof. □

We denote $k_{i_\ell} \in \mathbb{F}_{q^m}$ such that $\forall i_\ell \in \mathcal{I}, \boldsymbol{u}_{i_\ell} = k_{i_\ell}\left(\boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-i_\ell]}\boldsymbol{h}_j\right)$. A vector $\boldsymbol{u}_0 \in \mathcal{A}_0$ can be written under the form

$$\boldsymbol{u}_0 = \alpha_{\boldsymbol{h},\beta}(\boldsymbol{h}_0 + \sum_{j=1}^{\lambda} \beta_j \boldsymbol{h}_j)$$

From the structure of the solution space $\mathcal{S}$, there exists an $(\boldsymbol{h}, \beta) \in \mathcal{S}$ such that $\alpha_{\boldsymbol{h},\beta} = 1$. It means that we can fix $\boldsymbol{u}_0 := \boldsymbol{h}_0 + \sum_{j=1}^{\lambda} \beta_j \boldsymbol{h}_j$ as a known vector. For this element and for any $\mathcal{I} = \{i_1, \ldots, i_\lambda\}$ from Lemma 4 we have

$$\begin{aligned}\sum_{i_\ell \in \mathcal{I}} k_{i_\ell}^{\mathcal{I}}\left(\boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-i_\ell]}\boldsymbol{h}_j\right) &= \left(\sum_{i_\ell \in \mathcal{I}} k_{i_\ell}^{\mathcal{I}}\right)\boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1}\left(\sum_{i_\ell \in \mathcal{I}} k_{i_\ell}^{\mathcal{I}}\beta_j^{[-i_\ell]}\right)\boldsymbol{h}_j \\ &= \boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j \boldsymbol{h}_j\end{aligned}$$

Since the $\boldsymbol{h}_j$ are linearly independent we obtain the following system

$$(k_{i_1}^{\mathcal{I}}, k_{i_2}^{\mathcal{I}}, \ldots, k_{i_\lambda}^{\mathcal{I}})\begin{bmatrix} 1 & \beta_1^{[-i_1]} & \beta_2^{[-i_1]} & \cdots & \beta_{\lambda-1}^{[-i_1]} \\ 1 & \beta_1^{[-i_2]} & \beta_2^{[-i_2]} & \cdots & \beta_{\lambda-1}^{[-i_2]} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \beta_1^{[-i_\lambda]} & \beta_2^{[-i_\lambda]} & \cdots & \beta_{\lambda-1}^{[-i_\lambda]} \end{bmatrix} = (1, \beta_1, \beta_2, \ldots, \beta_{\lambda-1})$$

in the unknowns $k_i^{\mathcal{I}}$ and β_i. From assumption (1), knowing the β_i's, the solution is unique. To solve the system, let us consider the associate matrix

$$Mat^{\mathcal{I}}(\mathbf{X}) := \begin{bmatrix} 1 & X_1^{[i_1]} & X_2^{[i_1]} & \cdots & X_{\lambda-1}^{[i_1]} \\ 1 & X_1^{[i_2]} & X_2^{[i_2]} & \cdots & X_{\lambda-1}^{[i_2]} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_1^{[i_\lambda]} & X_2^{[i_\lambda]} & \cdots & X_{\lambda-1}^{[i_\lambda]} \end{bmatrix}$$

where $\mathbf{X} = (X_1, X_2, \ldots, X_{\lambda-1})$ is formed with the unknowns. We define the multivariate polynomial

$$f^{\mathcal{I}}(\mathbf{X}) \stackrel{def}{=} \det(Mat^{\mathcal{I}}(\mathbf{X}))$$

Since $f^{\mathcal{I}} \in \mathbb{F}_q[\mathbf{X}]$ we have

Lemma 5. *$f^{\mathcal{I}}$ has degree $\sum_{j\in\mathcal{I}}[j]$ and for all $u \in \mathbb{Z}$, $f^{\mathcal{I}+u}(\mathbf{X}) = f^{\mathcal{I}}(\mathbf{X})^{[u]}$.*

By Cramer's rule, for any $j = 1, \ldots, \lambda$ we have

$$k_{i_j}^{\mathcal{I}} = \frac{f^{-(\mathcal{I}\setminus\{i_j\})\cup\{0\}}(\beta)}{f^{-\mathcal{I}}(\beta)}, \tag{1.3}$$

where $\beta = (\beta_1, \ldots, \beta_\lambda)$. Let us define $\mathcal{J}_s = (\{1, \ldots, \lambda+1\}) \setminus \{s+1\}$, for all $s = 1, \ldots, \lambda$. From (1.3), we have

$$\forall s \in \{1, \ldots, \lambda\},\ k_1^{\mathcal{J}_s} = \frac{f^{-(\mathcal{J}_s\setminus\{1\})\cup\{0\}}(\beta)}{f^{-\mathcal{J}_s}(\beta)}$$

By elevating the equation to the power $[\lambda+1]$, from Lemma 5 we have

$$\forall s \in \{1, \ldots, \lambda\},\ (k_1^{\mathcal{J}_s})^{[\lambda+1]} = \frac{f^{(\lambda+1)-(\mathcal{J}_s\setminus\{1\})\cup\{0\}}(\beta)}{f^{(\lambda+1)-\mathcal{J}_s}(\beta)}$$

Now since we know only the vector space $\mathcal{A}_1$ and not the exact vectors $\boldsymbol{h}_0 + \sum_{j=1}^{\lambda-1} \beta_j^{[-1]}\boldsymbol{h}_j$, we do not know $k_1^{\mathcal{J}_s}$. However, we can compute the quantity $k_1^{\mathcal{J}_\lambda}/k_1^{\mathcal{J}_s}$ for $s \in \{1, \ldots, \lambda-1\}$ thank to Algorithm 2 and Lemma 4.

Algorithm 2: Determining quotient $k_1^{\mathcal{J}_\lambda}/k_1^{\mathcal{J}_s}$

Input: $\{\mathcal{A}_i\}_{i=1}^{n-k-1}$, $\{\mathcal{J}_s\}_{s=1}^{\lambda}$ and the vector $\boldsymbol{u}_0 \in \mathcal{A}_0$

Output: $\alpha_s = k_1^{\mathcal{J}_\lambda}/k_1^{\mathcal{J}_s}$ for $s \in \{1, \ldots, \lambda-1\}$

1 For $i = 1, \ldots, n-k-1$, fix $\boldsymbol{u}_i$ arbitrarily in $\mathcal{A}_i$

2 For $s = 1, \ldots, \lambda$, find $a_j^{\mathcal{J}_s}$ such that of $\sum_{j\in\mathcal{J}_s} a_j^{\mathcal{J}_s}\boldsymbol{u}_j = \boldsymbol{u}_0$

3 Return $\dfrac{a_1^{\mathcal{J}_\lambda}}{a_1^{\mathcal{J}_s}}$, for $s = 1, \ldots, \lambda-1$

Now let us define by $\alpha_s = (k_1^{\mathcal{J}_\lambda}/k_1^{\mathcal{J}_s})^{[\lambda+1]}$, for $s = 1, \ldots, \lambda-1$. To simplify notations, we also define

$$\forall s \in \{1, \ldots, \lambda\} \begin{cases} \mathcal{L}_s = (\lambda+1) - (\mathcal{J}_s \setminus \{1\} \cup \{0\}) \\ \mathcal{M}_s = (\lambda+1) - \mathcal{J}_s \end{cases}$$

We obtain the set of equations

$$\forall s \in \{1, \ldots, \lambda - 1\}, \quad f^{\mathcal{L}_\lambda}(\beta) f^{\mathcal{M}_s}(\beta) - \alpha_s f^{\mathcal{M}_\lambda}(\beta) f^{\mathcal{L}_s}(\beta) = 0$$

Let

$$\mathcal{F}_s(\mathbf{X}) \stackrel{def}{=} f^{\mathcal{L}_\lambda}(\mathbf{X}) f^{\mathcal{M}_s}(\mathbf{X}) - \alpha_s f^{\mathcal{M}_\lambda}(\mathbf{X}) f^{\mathcal{L}_s}(\mathbf{X}) \in \mathbb{F}_{q^m}[\mathbf{X}].$$

The polynomial $\mathcal{F}_s$ has degree $q^{\lambda+1} + q^\lambda + 2\sum_{j=1}^{\lambda-1} q^j + 1 - q^{\lambda-s}$

This gives us a multivariate polynomial system over $\mathbb{F}_{2^m}$ for which β is a solution. However, from our hypotheses we can do better and even reduce the degrees of the polynomials.

Since $\beta_1, \ldots, \beta_\lambda$ are linearly independent they cannot be roots of linear factors over $\mathbb{F}_q$ of $\mathcal{F}_s$. Therefore we can reduce for all s the polynomial $\mathcal{F}_s(\mathbf{X})$ by its $\mathbb{F}_q$-linear factors.

Lemma 6. *Let us define*

$$f_0(\mathbf{X}) = \prod_{a \in \mathbb{F}_q} (X_1 + a) \prod_{i=2}^{\lambda-1} \left(\prod_{a_0, \ldots a_{i-1} \in \mathbb{F}_q} (X_i + \sum_{j=1}^{i-1} a_j X_j + a_0) \right)$$

For any set $\mathcal{I}$ of cardinality λ, $f^{\mathcal{I}}(\mathbf{X})$ is divisible by $f_0(\mathbf{X})$

Proof. Let β be a root of $X_i + \sum_{j=1}^{i-1} a_j X_j + a_0$ then they are $\mathbb{F}_q$ co-linear. Hence, for all set of cardinality λ, $\mathcal{I}$, the corresponding columns of $Mat^{\mathcal{I}}(\beta)$ are co-linear. Therefore, $f^{\mathcal{I}}(\beta) = \det(Mat^{\mathcal{I}}(\beta)) = 0$ □

We have the following two corollaries

Corollary 2. *We have $f^{\mathcal{J}_\lambda - 1}(\mathbf{X}) = f_0(\mathbf{X})$*

Proof. Both polynomials are monic. Since $\mathcal{J}_\lambda - 1 = \{0, \ldots, \lambda - 1\}$, they also have the same degree $\sum_{i=0}^{\lambda-1} q^i$ □

Corollary 3. *For all $\mathcal{I} = \{i_1, \ldots, i_\lambda\}$, we have $(f_0(\mathbf{X}))^{[i_1]} \mid f^{\mathcal{I}}(\mathbf{X})$*

We have

— From the lemma: $f_0(\mathbf{X})$ divides $f^{\mathcal{M}_s}(\mathbf{X})$ and $f^{\mathcal{L}_s}(\mathbf{X})$ for all $s \in \{1, \ldots, \lambda - 1\}$.

— From corollary 3 : $f_0(\mathbf{X})^{[1]}$ divides $f^{\mathcal{M}_\lambda}(\mathbf{X})$ and $f^{\mathcal{L}_\lambda}(\mathbf{X})$, since the minimum index of the sets is equal to 1.

Therefore, for all $s \in \{1, \dots, \lambda - 1\}$, $\mathcal{F}_s(\mathbf{X})$ can be divided by $f_0(\mathbf{X})^{q+1}$. We now consider the reduced polynomials

$$\forall s \in \{1, \dots, \lambda - 1\}, \quad \mathcal{P}_s(\mathbf{X}) \stackrel{def}{=} \frac{\mathcal{F}_s(\mathbf{X})}{(f_0(\mathbf{X}))^{q+1}} = \frac{\mathcal{F}_s(\mathbf{X})}{f^{\mathcal{J}_\lambda - 1}(\mathbf{X}) f^{\mathcal{J}_\lambda}(\mathbf{X})}$$

This gives us a new polynomial system for which β is also a solution, but the degree is reduced.

Lemma 7. *Let* $\mathbf{A} = (a_{i,j})_{i=0,j=0}^{\lambda-1,\lambda-1} \in \boldsymbol{PGL}(\lambda; \mathbb{F}_q)$. *Consider the tranformation on* $f^{\mathcal{I}}(\mathbf{X})$ *defined on* $\mathbf{X} = (X_1, \dots, X_{\lambda-1})$ *by*

$$\forall j \in \{1, \dots, \lambda - 1\}, \quad X_j \longmapsto \frac{a_{0,j} + \sum_{i=1}^{\lambda-1} a_{i,j} X_i}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0} X_i}$$

then the polynomial $f^{\mathcal{I}}(\mathbf{X})$ *is transformed into*

$$f^{\mathcal{I}}(\mathbf{X}) \longmapsto \mathbf{A}.f^{\mathcal{I}}(\mathbf{X}) \stackrel{def}{=} \frac{\Delta_A}{(a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0} X_i)^{\deg(f^{\mathcal{I}})}} f^{\mathcal{I}}(\mathbf{X})$$

where Δ_A *is the determinant of* $\boldsymbol{A}$.

Proof. Let $D = a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0} X_i$. Thus, for $j = 1, \dots, \lambda$, the jth row of $Mat^{\mathcal{I}}(\mathbf{X})$ denoted by $Row_j(Mat^{\mathcal{I}}(\mathbf{X}))$ becomes

$$Row_j(Mat^{\mathcal{I}}(\mathbf{X})) \longmapsto \frac{Row_j\left(Mat^{\mathcal{I}}(\mathbf{X}) \cdot \boldsymbol{A}\right)}{D^{[i_j]}}.$$

Therefore, since $\deg(f^{\mathcal{I}}) = \sum_{j \in \mathcal{I}} [j]$, from lemma 5, we obtain

$$\det(Mat^{\mathcal{I}}(\mathbf{X})) \longmapsto \frac{\det \mathbf{A}}{D^{\deg(f^{\mathcal{I}})}} \det(Mat^{\mathcal{I}}(\mathbf{X}))$$
$$f^{\mathcal{I}}(\mathbf{X}) \longmapsto \frac{\Delta_A}{D^{\deg(f^{\mathcal{I}})}} f^{\mathcal{I}}(\mathbf{X})$$

□

Apply the lemma, we have

$$\mathcal{F}_s(\mathbf{X}) \longmapsto \frac{\Delta_A^2}{D^{q^{\lambda+1}+q^{\lambda}+2\sum_{j=1}^{\lambda-1} q^j+1-q^{\lambda-s}}}\mathcal{F}_s(\mathbf{X})$$
$$f^{\mathcal{J}_\lambda-1}(\mathbf{X}) \longmapsto \frac{\Delta_A}{D^{\sum_{j=0}^{\lambda-1} q^j}} f^{\mathcal{J}_\lambda-1}(\mathbf{X})$$
$$f^{\mathcal{J}_\lambda}(\mathbf{X}) \longmapsto \frac{\Delta_A}{D^{\sum_{j=1}^{\lambda} q^j}} f^{\mathcal{J}_\lambda}(\mathbf{X})$$

Hence,

$$\mathcal{P}_s(\mathbf{X}) \mapsto \frac{1}{D^{q^{\lambda+1}-q^{\lambda-s}}}\mathcal{P}_s(\mathbf{X})$$

We therefore have

Proposition 3. *If there isn't any common factor between the polynomials $\mathcal{P}_s(X)$, then the set of root of the polynomial system*

$$\forall i = 1, \ldots, \lambda - 1, \quad \mathcal{P}_i(\mathbf{X}) = 0 \tag{1.4}$$

equals the orbit of any root under the group action of $\boldsymbol{PGL}(\lambda, \mathbb{F}_q)$

Proof. If there isn't any common factor between the polynomials $\mathcal{P}_s(X)$ then the number of roots is at bounded by $\prod_{j=1}^{\lambda-1}(q^{\lambda+1} - q^j) = |\mathbf{PGL}(\lambda, q)|$ (Bezout bound [37]). Moreover, any element in the orbit of a solution β under the group action of $\mathbf{PGL}(\lambda, \mathbb{F}_q)$ is again root of the system. From Assumption (3) the orbit of β under $\mathbf{PGL}(\lambda, \mathbb{F}_q)$ has cardinality $= |\mathbf{PGL}(\lambda, \mathbb{F}_q)|$ which means that the stabilization of β with respect to this group action is trivial. In that case any root of the system (1.4) corresponds to an element of $\mathcal{S}$. □

For instance, when $q = 2$ and $\lambda = 3$ the system of equation below taking (β_1, β_2) as solution:

$$\begin{cases} Pr_1(X, Y) = 0 \\ Pr_2(X, Y) = 0 \end{cases}$$

This is a system of 2 polynomial equation in 2 variables. In practice, by using MAGMA, we can see that there isn't any common factor between $Pr_1(X)$ and $Pr_2(X)$. Therefore,

the number of roots has Bezout 's upper bound by the product of the degrees of $Pr_1(X,Y)$ and $Pr_2(X,Y)$.

Therefore, the number of roots are at most $(q^4-q)(q^4-q^2) = |\mathbf{PGL}(3,q)|$. Thus, all the roots are in the orbit of a root under an action of $\mathbf{PGL}(3,q)$

The remaining problem is finding a root of the system of equation above. It can be done by the following steps:

1. Calculating $Res(Pr_1, Pr_2, Y)$ the resultant of Pr_1 and Pr_2 in the variable Y. We obtain a univariate polynomial of degree 168 in variable X. Finding one root x_0 of this polynomial.
2. Calculating $\gcd(Pr_1(x_0,Y), Pr_2(x_0,Y))$ which is a polynomial of degree 4 in variable Y. Taking one root y_0 and verify it is a root of the system of equation.

In general, the problem of finding one root of a system of polynomial equation is a hard question as well as finding all roots of a system of polynomial equation.

Polynomial System Solving over Finite Fields Let $\mathbb{F}$ is a finite field. **Input:** $f_1(x_1,...,x_n),...,f_m(x_1,...,x_n) \in \mathbb{F}[x_1,...,x_n]$.

Goal: Find a vector $\alpha = (\alpha_1,...,\alpha_n) \in \mathbb{F}^n$ s.t: $f_1(\alpha) = \cdots = f_m(\alpha) = 0$.

Theoretically it is a NP-hard problem (problem AN9 p.251 in Appendix: A list of NP-complete problem [38]). For the special cases $\lambda = 2$ and $\lambda = 3$ finding a solution can be done in polynomial-time (by using properties of resultants for $\lambda = 3$).

However, in the case of no common factor, the number of roots is bounded by Bezout bound. To check that $Pr_i(X)_{i=1}^{\lambda-1}$ don't have common factor, we can check whether $Res(Pr_i(X), Pr_j(X), X_1) \neq 0,\ \forall 1 \leq i < j \leq \lambda-1$. (Prop. 1, Ch. 3, [39]). It costs $O(d^3)$ where $d = \prod_{j=1}^{\lambda-1}(q^{\lambda+1}-q^j)$ arithmetic operations over $\mathbb{F}_{q^m}[X_2,\ldots,X_{\lambda-1}]$.

We can see the importance of the Assumption(3) in the Proposition 3. In the case where this assumption does not satisfy, i.e there exists $\overline{\boldsymbol{A}} \in \mathbf{PGL}(\lambda,\mathbb{F}_q)\backslash\overline{\boldsymbol{I}_\lambda}$ and $\boldsymbol{A} = (a_{ij})_{i,j=1}^{\lambda}$ such that

$$\beta_j = \frac{a_{0,j} + \sum_{i=1}^{\lambda-1} a_{i,j}\beta_i}{a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}\beta_i}$$

Thus, β is a root of a system of $\lambda - 1$ polynomial equations of degree 2:

$$(a_{0,0} + \sum_{i=1}^{\lambda-1} a_{i,0}X_i)X_j - (a_{0,j} + \sum_{i=1}^{\lambda-1} a_{i,j}X_i) = 0,\ j = 1, \ldots, \lambda - 1$$

This polynomial is different from 0. Indeed, if it was, $a_{0,0} = a_{j,j}$ for $j = 1, \ldots, \lambda - 1$ and $a_i, j = 0$ for $i \neq j$, which means $\boldsymbol{A} \in \overline{I_\lambda}$.

This system is multivariate quadratic (MQ)-system, the associated problem to decide if this system is solvable or not, also known as MQ-problem, is proven to be NP-complete [38]. Some algorithms used to solve this system is reviewed in the paper [40]. In case of $\lambda = 3$, this can be solved easily by Resultant. Therefore, when the Assumption (3) does not satisfy, we can exploit some information about β by solving a multivariate quadratic system.

3.2.3 Final step:

Now from a solution β to (1.4), we aim at finding the corresponding vector $\mathbf{h} \in (\mathbb{F}_{q^m}^n)^\lambda$ such that $(\mathbf{h}, \beta) \in \mathcal{S}$.

We point out the key steps in the Coggia-Couvreur attack for λ as follows. To be convenient, we denote known elements by blue color and unknown elements by red color. Given $\beta'_1, \ldots, \beta'_{\lambda-1}$, recover $(\boldsymbol{h}'_0, \ldots, \boldsymbol{h}'_{\lambda-1}, \beta'_1, \ldots, \beta'_{\lambda-1})$ corresponding.

1. For $\mathcal{I} = \{1, \ldots, \lambda\}$, since β' is known, $k_i = \dfrac{f^{-(\mathcal{I}\setminus\{i\})\cup\{0\}}(\beta')}{f^{-\mathcal{I}}(\beta')}, i = 1, \ldots, \lambda$ can be computed. Moreover, from the Lemma 4, there exists a unique λ-tuple $\boldsymbol{u}_\mathcal{I} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_\lambda) \in \bigtimes_{j=1}^{\lambda} \mathcal{A}_j$ such that $\sum_{i=1}^{\lambda} \boldsymbol{u}_i = \boldsymbol{u}_0$, so we can compute

$$\boldsymbol{h}'_0 + \sum_{j=1}^{\lambda-1} \beta'^{[-i]}_j \boldsymbol{h}'_j = \frac{\boldsymbol{u}_i}{k_i}, i = 1, \ldots, \lambda.$$

$$(\boldsymbol{h}'_0, \ldots, \boldsymbol{h}'_{\lambda-1}) \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \beta_1^{[-1]} & \beta_1^{[-2]} & \ldots & \beta_1^{[-\lambda]} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{\lambda-1}^{[-1]} & \beta_{\lambda-1}^{[-2]} & \ldots & \beta_{\lambda-1}^{[-\lambda]} \end{bmatrix} = \left(\frac{\boldsymbol{u}_1}{k_1}, \frac{\boldsymbol{u}_2}{k_2}, \ldots, \frac{\boldsymbol{u}_\lambda}{k_\lambda} \right)$$

It implies to a linear system of λ equation and λ unknowns which are vectors $\boldsymbol{h}'_0, \ldots, \boldsymbol{h}'_{\lambda-1}$ and the determinant of the matrix of coefficients is non-zero.

2. After recovering an alternate key of the form $(\boldsymbol{h}'_0, \ldots, \boldsymbol{h}'_{\lambda-1}, \beta'_1, \ldots, \beta'_{\lambda-1})$, we can compute the dual code $\mathcal{C}^{\perp}_{\text{pub}}$ and hence decrypt the ciphertext.

3.3 Complexity of the case $\lambda = 3$

This part shows the complexity of the attack by giving the number of operation in $\mathbb{F}_{q^m}$. Let ω be the exponent of the complexity of linear algebra operations. The Frobenius map costs $O(\log q)$ operations.

Step 1.

— Computation of dual code $\mathcal{C}^{\perp}_{\text{pub}}$ cots $O(n^{\omega})$ operations.

— Computation of ${\mathcal{C}^{\perp}_{\text{pub}}}^{[i]}, \forall i = 1, \ldots, n-k+1$ costs $O(n^2 \log q)$ operations.

— Computation $S_j = \sum_{i=j}^{j+\lambda-1} {\mathcal{C}^{\perp}_{\text{pub}}}^{[i]}$ uses Gaussian elimination, so it costs $O(n^{\omega})$. Thus, computation $\bigcap_{i=0}^{n-k-\lambda+1} S_j$ costs $O(n^{\omega+1})$.
Overall step 1 costs $O(n^3 \log q + n^{\omega+1})$ operations.

Step 2.

— Computation $(u_1^{\mathcal{I}}, \ldots, u_{\lambda}^{\mathcal{I}})$ represents the resolution of a linear system λ unknowns and n equations costs $O(n)$ operations. This computation performed $O(n)$ times, so it costs $O(n^2)$ operations.

— Complexity of finding a root of a polynomial of degree $\tilde{d}$ by Cantor–Zassenhaus algorithm ([41]) costs $\tilde{O}(\tilde{d}^2 m \log q)$ operations in $\mathbb{F}_{q^m}$ for $\tilde{d} = (q^4 - q)(q^4 - q^2)$.

— Computation of resultant of bivariate polynomials $Res(P_1, P_2, X)$ which P_1, P_2 of degree d, e by Lickteig–Roy subresultant algorithm costs $O(d^2 e)$ ([42]).

— A finite number of linear systems solving costs $O(n^{\omega})$

Summary. For $m = O(n)$, overall cost of $O(n^3 \log q + n^{\omega+1}) + \tilde{O}(d^2 n \log q)$ for $d = (q^4 - q)(q^4 - q^2)$.

Chapter 2

Cryptosystems based on the polynomial reconstruction RAMESSES

This chapter comes from a work in the pre-publication [5]. In this chapter we aim at designing a new one-way encryption scheme featuring very compact keys, based on rank metric decoding problems. The long-standing idea finds origins in [2] which was an extended idea of a proposal in Hamming metric [43]. The original rank metric encryption scheme was broken in [44], and a recent repair was proposed in [45]. However it implies to choose a specific code and a syndrome coming from a structured vector of moderate rank, which we want to avoid here.

In a first section we introduce necessary notation and definitions. Then we describe the encryption scheme and we propose sets of parameters such that keys and ciphertext sizes are not larger than few hundreds of bytes. In the next section, we prove the consistency of the encryption scheme (this consistency also can be seen in the implementation in MAGMA `https://github.com/BaDucPham/RAMESSES/blob/main/RAMESSES.mgm`) and we analyze its security by showing to which problems the security can be reduced, and by giving the complexity of algorithms solving these problems. Recently, this cryptosystem was broken by Bombar-Couveur attack in [29], so we also introduce this attack in the end of this chapter.

1 Preliminaries

1.1 Notation and definitions

Although the whole analysis can be done even if q is not a power of 2, throughout the chapter, we set $q = 2^n$ for some integer $n \geq 1$, and we let $\mathbb{F}_q$ denote the finite field with q

elements. The field $\mathbb{F}_q$ can also be viewed as a vector space of dimension n over $\mathbb{F}_2$. The Frobenius automorphism $\theta : \mathbb{F}_q \to \mathbb{F}_q$, $x \mapsto x^2$, is $\mathbb{F}_2$-linear. Its inverse is the $(n-1)$-fold composition $\theta^{n-1} = \theta \circ \cdots \circ \theta$. For convenience, we sometimes write $x^{[i]} \coloneqq \theta^i(x)$, for $i \in [0, n-1] \coloneqq \{0, \ldots, n-1\}$.

Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{F}_q^n$ be a basis of $\mathbb{F}_q$ over $\mathbb{F}_2$. We recall the extension map

$$\begin{array}{rccc} \mathrm{Ext}_{\boldsymbol{\beta}} : & \mathbb{F}_q^n & \to & \mathbb{F}_2^{n\times n} \\ & \boldsymbol{a} = (a_1, \ldots, a_n) & \mapsto & \boldsymbol{A} = (\boldsymbol{\alpha}_1^\top, \ldots, \boldsymbol{\alpha}_n^\top) \end{array}$$

where, for all $1 \le j \le n$, the vector $\boldsymbol{\alpha}_j \in \mathbb{F}_2^n$ consists of coordinates of $a_j \in \mathbb{F}_q$ in the basis $\boldsymbol{\beta}$, *i.e.* $a_j = \sum_{i=1}^n \beta_i A_{i,j}$. In particular, for every $\boldsymbol{A} \in \mathbb{F}_2^{n\times n}$, we have $\mathrm{Ext}_{\boldsymbol{\beta}}(\boldsymbol{\beta}\boldsymbol{A}) = \boldsymbol{A}$.

We also define the row space of $\boldsymbol{a} \in \mathbb{F}_q^n$ with respect to $\boldsymbol{\beta}$ as

$$\mathrm{RowSp}_{\boldsymbol{\beta}}(\boldsymbol{a}) \coloneqq \{\boldsymbol{x}\,\mathrm{Ext}_{\boldsymbol{\beta}}(\boldsymbol{a}), \boldsymbol{x} \in \mathbb{F}_2^n\} \subseteq \mathbb{F}_2^n\,.$$

Similarly, the column space of $\boldsymbol{a} \in \mathbb{F}_q^n$ is $\mathrm{ColSp}_{\boldsymbol{\beta}}(\boldsymbol{a}) \coloneqq \{\sum_{i=1}^n x_i a_i \mid \boldsymbol{x} \in \mathbb{F}_2^n\} \subseteq \mathbb{F}_q$.

We let $\mathrm{Gr}(t, \mathbb{F}_2^n)$ denote the set of subspaces of $\mathbb{F}_2^n$ of dimension t, which contains $\begin{bmatrix} n \\ t \end{bmatrix}_2 \coloneqq \frac{(2^n-1)(2^{n-1}-1)\cdots(2^{n-t+1}-1)}{(2^t-1)(2^{t-1}-1)\cdots(2^1-1)}$ elements. Each subspace $\mathcal{V} \in \mathrm{Gr}(t, \mathbb{F}_2^n)$ can be represented by the unique reduced row echelon form (RREF) of any matrix $\boldsymbol{V} \in \mathbb{F}_2^{n\times n}$ whose row space generates $\mathcal{V}$. We know from [46, 47] that this representation can be computed efficiently (in time $\tilde{O}(nt(n-t))$). Recall that a matrix is in reduced row echelon form if the following holds:

- the index of the pivot (*i.e.* the first non-zero coefficient) of row i is strictly larger than the index of the pivot of row $i-1$;
- all pivots are ones;
- each pivot is the only non-zero entry in its column.

We finally define $\mathcal{P}_{t,n} \coloneqq \{\boldsymbol{P} \in \mathbb{F}_2^{n\times n} \mid \mathrm{rk}(\boldsymbol{P}) = t, \boldsymbol{P} \text{ is in RREF}\}$.

Definition 3 ($\boldsymbol{g}$-degree)**.** *Let $\boldsymbol{x} \in \mathbb{F}_q^n$ and $\boldsymbol{X} = \mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{x})$. The* $\boldsymbol{g}$-degree of $\boldsymbol{x}$*, denoted* $\deg_{\boldsymbol{g}}(\boldsymbol{x})$*, is the unique integer $\ell \in [0, n-1]$ such that $\boldsymbol{x} \in \mathcal{G}_{\ell+1}(\boldsymbol{g}) \setminus \mathcal{G}_\ell(\boldsymbol{g})$. Similarly, one defines the $\boldsymbol{g}$-degree of $\boldsymbol{X}$ as $\deg_{\boldsymbol{g}}(\boldsymbol{X}) = \deg_{\boldsymbol{g}}(\boldsymbol{x})$.*

In other words, a vector $\boldsymbol{x} \in \mathbb{F}_q^n$ of $\boldsymbol{g}$-degree ℓ can be written

$$\boldsymbol{x} = \lambda_\ell \boldsymbol{g}^{[\ell]} + \sum_{j=0}^{\ell-1} \lambda_j \boldsymbol{g}^{[j]}$$

for some non-zero $\lambda_\ell \in \mathbb{F}_q \setminus \{0\}$ and some ℓ-tuple $(\lambda_{\ell-1}, \ldots, \lambda_0) \in \mathbb{F}_q^\ell$.

2 The encryption scheme

This implementation in MAGMA can be seen in `https://github.com/BaDucPham/RAMESSES/blob/main/RAMESSES.mgm`

System parameters. Integers $1 \leq w, k, \ell, t \leq n$ are public parameters and specified according to the desired security level (see Section 3). We set $q = 2^n$, and we also make public a basis $\boldsymbol{g}$ of $\mathbb{F}_q/\mathbb{F}_2$. We let $\boldsymbol{H}$ denote a *fixed* parity-check matrix of $\mathcal{G}_k(\boldsymbol{g})$.

Key generation. Alice picks uniformly at random a vector $\boldsymbol{k}_{\text{priv}} \in \mathbb{F}_q^n$ of rank w. As explained in Algorithm 3, the public key is the syndrome of $\boldsymbol{k}_{\text{priv}}$ with respect to the parity-check matrix $\boldsymbol{H}$ of $\mathcal{G}_k(\boldsymbol{g})$, and the private key is $\boldsymbol{k}_{\text{priv}}$.

Algorithm 3: $\mathsf{KeyGen}(1^\lambda)$

Input:
Output: a pair of public/private keys $(\boldsymbol{k}_{\text{pub}}, \boldsymbol{k}_{\text{priv}})$

1 Pick $\boldsymbol{k}_{\text{priv}} \xleftarrow{\$} \{\boldsymbol{x} \in \mathbb{F}_q^n, \mathrm{rk}(\boldsymbol{x}) = w\}$
2 Compute $\boldsymbol{k}_{\text{pub}} \in \mathbb{F}_q^{n-k}$ such that ${\boldsymbol{k}_{\text{pub}}}^\top = \boldsymbol{H}{\boldsymbol{k}_{\text{priv}}}^\top$
3 Output $(\boldsymbol{k}_{\text{pub}}, \boldsymbol{k}_{\text{priv}}) \in \mathbb{F}_q^{n-k} \times \mathbb{F}_q^n$

Encryption. The set of plaintexts is $\mathcal{P}_{t,n}$, as defined in Section 1.1. Encryption is presented in Algorithm 4. Notice that in steps 3-4, the computation of $\boldsymbol{p}'$ should be understood as a the generation of a uniform random vector such that $\mathrm{RowSp}_{\boldsymbol{g}}(\boldsymbol{p}')$ is the rowspan of $\boldsymbol{P}$.

Decryption. We present in Algorithm 5 a decryption algorithm which may fail with negligible probability. The failure rate is devoted to be cryptographically small, and is bounded in Section 4.2. We also make use of an $\mathbb{F}_2$-linear map $V_{\boldsymbol{k}_{\text{priv}}} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ such that $V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{k}_{\text{priv}} \rangle = \boldsymbol{0}$. This map can be efficiently computed from the knowledge of the private key $\boldsymbol{k}_{\text{priv}}$. Mathematical properties of this map are given in Section 4.1

In Algorithm 5, one needs to decode Gabidulin codes up to half their minimum distance, *i.e.* to decode errors of rank less than $\lfloor \frac{n - \dim \mathrm{Gab}}{2} \rfloor$. Many such algorithms can be

Algorithm 4: $\mathsf{Encrypt}(\boldsymbol{k}_{\text{pub}}, \boldsymbol{P})$

Input: public key $\boldsymbol{k}_{\text{pub}} \in \mathbb{F}_q^{n-k}$, plaintext $\boldsymbol{P} \in \mathcal{P}_{t,n}$
Output: ciphertext $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$

1 Compute any $\boldsymbol{y} \in \mathbb{F}_q^n$ such that $\boldsymbol{H}\boldsymbol{y}^\top = \boldsymbol{k}_{\text{pub}}{}^\top$
2 Pick $\boldsymbol{T} \xleftarrow{\$} \{\boldsymbol{M} \in \mathbb{F}_2^{n\times n}, \deg_{\boldsymbol{g}}(\boldsymbol{M}) = \ell\}$
3 Pick $\boldsymbol{S} \xleftarrow{\$} \{\boldsymbol{M} \in \mathbb{F}_2^{n\times n}, \mathrm{rk}(\boldsymbol{M}) = n\}$
4 Compute $\boldsymbol{p}' = \boldsymbol{g}\boldsymbol{S}\boldsymbol{P} \in \mathbb{F}_q^n$
5 Output $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$ such that $\boldsymbol{u}^\top = \boldsymbol{H}(\boldsymbol{y}\boldsymbol{T} + \boldsymbol{p}')^\top$

Algorithm 5: $\mathsf{Decrypt}(\boldsymbol{k}_{\text{priv}}, \boldsymbol{u})$

Input: private key $\boldsymbol{k}_{\text{priv}} \in \mathbb{F}_q^n$, ciphertext $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$
Output: plaintext $\hat{\boldsymbol{P}} \in \mathcal{P}_{t,n}$, or failure

1 Compute a solution $\boldsymbol{x} \in \mathbb{F}_q^n$ to the linear system $\boldsymbol{H}\boldsymbol{x}^\top = \boldsymbol{u}^\top$.
2 Compute $\boldsymbol{z} = V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{x} \rangle \in \mathbb{F}_q^n$.
3 Decode $\boldsymbol{z}$ as a corrupted $\mathcal{G}_{k+\ell+w}(\boldsymbol{g})$-codeword. If success, one gets an error vector $\boldsymbol{a} \in \mathbb{F}_q^n$ of rank $\leq t$.
4 **If** $\mathrm{rk}(\boldsymbol{a}) < t$, output failure.
5 **Otherwise**, output $\hat{\boldsymbol{P}} = \mathtt{RREF}(\mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{a}))$.

found in the literature since the seminal work of Gabidulin [19] such as [8, 48, 49, 50, 51, 52].

3 Parameters

Public key size. The public key consists in a vector $\boldsymbol{k}_{\text{pub}} \in \mathbb{F}_q^{n-k}$. Thus, its size is $(n-k)n$ bits, or $\frac{(n-k)n}{8}$ bytes.

Private key size. For the private key $\boldsymbol{k}_{\text{priv}} \in \mathbb{F}_q^n$, Alice actually needs to store only the map $V_{\boldsymbol{k}_{\text{priv}}}$. From Section 4.1, this map is a monic polynomial over $\mathbb{F}_q$ of degree w. Hence only w coefficients over $\mathbb{F}_q$ actually need to be stored, the size of the private key is thus wn bits, or $\frac{wn}{8}$ bytes.

Ciphertext size. The ciphertext is a vector $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$, hence its size is $(n-k)n$ bits, *i.e.* $\frac{(n-k)n}{8}$ bytes.

4 Analysis

4.1 Mathematical background

For $\boldsymbol{x} \in \mathbb{F}_q^n$, the polynomial $P(X) \in \mathbb{F}_q[X;\theta]$ of minimum degree such that $P\langle \boldsymbol{g} \rangle = \boldsymbol{x}$ is the *$\boldsymbol{g}$-interpolating polynomial of $\boldsymbol{x}$* and is denoted $L_{\boldsymbol{x}}(X)$. By definition $\deg(L_{\boldsymbol{x}}) = \deg_{\boldsymbol{g}}(\boldsymbol{x})$.

Finally, given $\boldsymbol{e} \in \mathbb{F}_q^n$, the set of polynomials $P \in \mathbb{F}_q[X;\theta]$ satisfying $P(\boldsymbol{e}) = \boldsymbol{0}$ is a left-ideal $I_{\boldsymbol{e}}$ of $\mathbb{F}_q[X;\theta]$. Since skew polynomial rings are principal ideal domains, we can define the *minimum vanishing polynomial* $V_{\boldsymbol{e}}(X) \in \mathbb{F}_q[X;\theta]$ of $\boldsymbol{e}$ as the unique monic skew polynomial which generates $I_{\boldsymbol{e}}$. Notice that $\deg(V_{\boldsymbol{e}}) = \mathrm{rk}(\boldsymbol{e}) \geq n - \deg_{\boldsymbol{g}}(\boldsymbol{e})$.

Theorem 4 ([53])**.** *Let $\boldsymbol{e} = (e_1, \ldots, e_n) \in \mathbb{F}_q^n$, then if $\mathrm{rk}(\boldsymbol{e}) = t$, there exist a unique $V_{\boldsymbol{e}} \in \mathbb{F}_q[X;\theta]$, monic of degree t such that $V_{\boldsymbol{e}}\langle \boldsymbol{e} \rangle = 0$.*

The following lemma will be helpful for the analysis of the scheme consistency.

Lemma 8. *Let $P(X) \in \mathbb{F}_q[X;\theta]$ and $\boldsymbol{a} \in \mathbb{F}_q^n$. Then we have $\mathrm{RowSp}_{\boldsymbol{g}}(P(\boldsymbol{a})) \subseteq \mathrm{RowSp}_{\boldsymbol{g}}(\boldsymbol{a})$. Moreover, if $\mathrm{RowSp}_{\boldsymbol{g}}(P(\boldsymbol{a})) \neq \mathrm{RowSp}_{\boldsymbol{g}}(\boldsymbol{a})$, then there exists a non-zero $x = \sum_{i=1}^n \lambda_i a_i \in \mathrm{ColSp}(\boldsymbol{a})$ such that $P(x) = 0$.*

Proof. Let $\boldsymbol{B} \in \mathbb{F}_2^{n \times n}$ satisfy $\mathrm{RowSp}_{\boldsymbol{g}}(\boldsymbol{a}) = \{\boldsymbol{x} \in \mathbb{F}_2^n, \boldsymbol{x}\boldsymbol{B} = \boldsymbol{0}\}$. In particular, one can see that $\boldsymbol{a}\boldsymbol{B} = \boldsymbol{0}$. Hence, by $\mathbb{F}_2$-linearity $P(\boldsymbol{a})\boldsymbol{B} = P(\boldsymbol{a}\boldsymbol{B}) = \boldsymbol{0}$. Thus, every $\boldsymbol{y} \in \mathrm{RowSp}_{\boldsymbol{g}}(P(\boldsymbol{a}))$ satisfies $\boldsymbol{y}\boldsymbol{B} = \boldsymbol{0}$, leading to $\mathrm{RowSp}_{\boldsymbol{g}}(P(\boldsymbol{a})) \subseteq \mathrm{RowSp}_{\boldsymbol{g}}(\boldsymbol{a})$.

Assume now that $\mathrm{RowSp}_{\boldsymbol{g}}(P(\boldsymbol{a})) \neq \mathrm{RowSp}_{\boldsymbol{g}}(\boldsymbol{a})$. It implies that $\dim \mathrm{ColSp}(P(\boldsymbol{a})) < \dim \mathrm{ColSp}(\boldsymbol{a})$. Let $(a_{i_j})_{1 \leq j \leq k} \subset \mathbb{F}_q$ be an ordered basis of $\mathrm{ColSp}(\boldsymbol{a}) \subseteq \mathbb{F}_q$ over $\mathbb{F}_2$. Then there must exist a non-zero $(\lambda_j) \in \mathbb{F}_2^k$ such that $\sum_{j=1}^k \lambda_j P(a_{i_j}) = 0$, otherwise we would have $\dim \mathrm{ColSp}(P(\boldsymbol{a})) = k$. If we set $x = \sum_j \lambda_j a_{i_j} \in \mathbb{F}_q \setminus \{0\}$, then we get $P(x) = 0$ by $\mathbb{F}_2$-linearity. □

4.2 Consistency

In this section we characterize the output of algorithm Decrypt described in Section 2. As input, Decrypt receives a vector $\boldsymbol{k}_{\mathrm{priv}} \in \mathbb{F}_q^n$ of rank w and a vector $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$ such that $\boldsymbol{u} = \boldsymbol{H}(\boldsymbol{y}\boldsymbol{T} + \boldsymbol{p}')^\top$, where

- vector $\boldsymbol{y} \in \mathbb{F}_q^n$ satisfies $\boldsymbol{H}\boldsymbol{y}^\top = \boldsymbol{H}\boldsymbol{k}_{\mathrm{priv}}{}^\top$,
- matrix $\boldsymbol{T} \in \mathbb{F}_2^{n \times n}$ has $\boldsymbol{g}$-degree ℓ,

– vector $\boldsymbol{p}' = \boldsymbol{gSP} \in \mathbb{F}_q^n$ has rank $t := \lfloor \frac{n-k-\ell-w}{2} \rfloor$.

First, notice that $\boldsymbol{y} = \boldsymbol{k}_{\text{priv}} + \boldsymbol{c}$ for some $\boldsymbol{c} \in \mathcal{G}_k(\boldsymbol{g})$. In the first step of Algorithm 5, a vector $\boldsymbol{x} \in \mathbb{F}_q^n$ solution to $\boldsymbol{Hx}^\top = \boldsymbol{u}^\top$ is computed. One can see that the set S of such solutions is

$$S = \{\boldsymbol{yT} + \boldsymbol{p}' + \boldsymbol{c}' \mid \boldsymbol{c}' \in \mathcal{G}_k(\boldsymbol{g})\} \subseteq \mathbb{F}_q^n .$$

Therefore, in step 2 of Algorithm 5, we have

$$\begin{aligned} \boldsymbol{z} = V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{x} \rangle &= V_{\boldsymbol{k}_{\text{priv}}}\langle (\boldsymbol{c} + \boldsymbol{k}_{\text{priv}})\boldsymbol{T} + \boldsymbol{p}' + \boldsymbol{c}' \rangle \\ &= V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{c}' + \boldsymbol{c}T + \boldsymbol{k}_{\text{priv}}\boldsymbol{T} + \boldsymbol{p}' \rangle \\ &= V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{c}' + \boldsymbol{cT} \rangle + \underbrace{V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{k}_{\text{priv}} \rangle \boldsymbol{T}}_{\mathbf{0}} + V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{p}' \rangle . \end{aligned}$$

We notably used the $\mathbb{F}_2$-linearity of $V_{\boldsymbol{k}_{\text{priv}}}$. Also recall that, for any $\boldsymbol{a} \in \mathbb{F}_q^n$, $L_{\boldsymbol{a}}(X)$ denotes the $\boldsymbol{g}$-interpolating polynomial of $\boldsymbol{a}$. Then we get:

$$\boldsymbol{z} = (V_{\boldsymbol{k}_{\text{priv}}} \cdot (L_{\boldsymbol{c}'} + L_{\boldsymbol{cT}}))\langle \boldsymbol{g} \rangle + V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{p}' \rangle .$$

Moreover, $L_{\boldsymbol{cT}} = L_{\boldsymbol{c}} \cdot L_{\boldsymbol{gT}}$ yields $\deg(L_{\boldsymbol{cT}}) \leq k - 1 + \ell$ since $\deg_{\boldsymbol{g}}(\boldsymbol{T}) = \ell$. Therefore, the polynomial $V_{\boldsymbol{k}_{\text{priv}}} \cdot (L_{\boldsymbol{c}'} + L_{\boldsymbol{cT}})$ has degree at most $\deg(V_{\boldsymbol{k}_{\text{priv}}}) + \max\{\deg(L_{\boldsymbol{c}'}), \deg(L_{\boldsymbol{cT}})\} \leq w + k - 1 + \ell$.

We also know that $\text{rk}(V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{p}' \rangle) \leq \text{rk}(\boldsymbol{p}') = \text{rk}(\boldsymbol{P}) = t = \lfloor \frac{n-k-\ell-w}{2} \rfloor$. Hence, in third step of Algorithm 5, any decoding algorithm for $\mathcal{G}_{k+w+\ell}(\boldsymbol{g})$ that decodes errors of rank at most t will retrieve $V_{\boldsymbol{k}_{\text{priv}}}(\boldsymbol{p}')$ from $\boldsymbol{z}$. Finally, Algorithm 5 outputs a matrix $\hat{\boldsymbol{P}} \in \mathcal{P}_{t,n}$ such that $\text{RowSp}(\hat{\boldsymbol{P}}) = \text{RowSp}_{\boldsymbol{g}}(V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{p}' \rangle)$.

As a consequence, decryption fails whenever $\text{RowSp}_{\boldsymbol{g}}(V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{p}' \rangle) \neq \text{RowSp}(\boldsymbol{P})$, where $\boldsymbol{P}$ is the original plaintext. First notice that $\text{RowSp}(\boldsymbol{P}) = \text{RowSp}_{\boldsymbol{g}}(\boldsymbol{p}')$. Then, Lemma 8 shows that if decryption fails, then there exists a non-zero $x \in \text{ColSp}(\boldsymbol{p}')$ such that $V_{\boldsymbol{k}_{\text{priv}}}\langle \boldsymbol{x} \rangle = 0$. Let us now recall that the set of zeroes of $V_{\boldsymbol{k}_{\text{priv}}}$ is exactly $\text{ColSp}(\boldsymbol{k}_{\text{priv}})$. Hence we get the following result.

Lemma 9. *Let $\boldsymbol{P} \in \mathcal{P}_{t,n}$. If, on input $(\boldsymbol{k}_{\text{priv}}, \mathsf{Encrypt}(\boldsymbol{k}_{\text{pub}}, \boldsymbol{P}))$ where $(\boldsymbol{k}_{\text{pub}}, \boldsymbol{k}_{\text{priv}}) \leftarrow \mathsf{KeyGen}$, algorithm $\mathsf{Decrypt}$ does not output $\boldsymbol{P}$, then the matrix $\boldsymbol{S}$ chosen at step 4 satisfies $\text{ColSp}(\boldsymbol{k}_{\text{priv}}) \cap \text{ColSp}(\boldsymbol{SP}) \neq \{0\}$.*

One can now estimate the probability of failure of $\mathsf{Decrypt}$.

Lemma 10. *Let* $(\boldsymbol{k}_{\text{pub}}, \boldsymbol{k}_{\text{priv}}) \leftarrow \mathsf{KeyGen}$ *be any pair of keys generated by* KeyGen*, on public parameters* n, w, t*. Then, for every* $\boldsymbol{P} \in \mathcal{P}_{t,n}$*,*

$$\mathbb{P}_{\boldsymbol{S},\boldsymbol{T},\boldsymbol{y}}\left(\hat{\boldsymbol{P}} \neq \boldsymbol{P} \;\middle|\; \begin{array}{l} \boldsymbol{u} \leftarrow \mathsf{Encrypt}(\boldsymbol{k}_{\text{pub}}, \boldsymbol{P}) \\ \hat{\boldsymbol{P}} \leftarrow \mathsf{Decrypt}(\boldsymbol{k}_{\text{priv}}, \boldsymbol{u}) \end{array}\right) \leq 2^{-(n-t-w)}\,.$$

Proof. Using Lemma 9, we have

$$\begin{aligned}\mathbb{P}_{\boldsymbol{S},\boldsymbol{T},\boldsymbol{y}}&\left(\hat{\boldsymbol{P}} \neq \boldsymbol{P} \;\middle|\; \begin{array}{l} \boldsymbol{u} \leftarrow \mathsf{Encrypt}(\boldsymbol{k}_{\text{pub}}, \boldsymbol{P}) \\ \hat{\boldsymbol{P}} \leftarrow \mathsf{Decrypt}(\boldsymbol{k}_{\text{priv}}, \boldsymbol{u}) \end{array}\right) \\ &= \mathbb{P}_{\boldsymbol{S}}\big(\,\mathrm{ColSp}(\boldsymbol{k}_{\text{priv}}) \cap \mathrm{ColSp}(\boldsymbol{SP}) \neq \{0\}\big)\,.\end{aligned}$$

It is easy to check that the probability that a t-dimensional random subspace of $\mathbb{F}_2^n$ intersects non-trivially a fixed subspace of dimension w is bounded by $\frac{(2^t-1)(2^w-1)}{2^n-1} \leq 2^{t+w-n}$. This concludes the proof. □

4.3 Existing attacks

In this section, we will consider known attacks on the system at the time of the publication of this preprint paper. In the following, we denote by λ the desired security parameter, *i.e.*, any attack against the cryptosystem must cost at least 2^λ operations over $\mathbb{F}_2$.

Exhaustive search attacks. In order to avoid attacks by exhaustive search, one has the following constraints on the parameters.

1. $|\mathcal{P}_{t,n}| = \begin{bmatrix} n \\ t \end{bmatrix}_2 \geq 2^\lambda$, satisfied when $t(n-t) \geq \lambda$.
2. $|\{\boldsymbol{k}_{\text{priv}}\}| \geq \begin{bmatrix} n \\ w \end{bmatrix}_2 \geq 2^\lambda$, satisfied when $w(n-w) \geq \lambda$.
3. $|\mathcal{M}_\ell| \geq 2^\lambda$, satisfied when $(\ell+1)n \geq \lambda$.

where $\mathcal{M}_\ell = \{\boldsymbol{M} \in \mathbb{F}_2^{n\times n}, \deg_{\boldsymbol{g}}(\boldsymbol{M}) = \ell\}$.

Attack by decoding beyond the unique decoding radius of Gabidulin codes. Let $\boldsymbol{e}' \in \mathbb{F}_q^n$ be *any* solution of $\boldsymbol{H}\boldsymbol{e}'^\top = \boldsymbol{k}_{\text{pub}}{}^\top$ of rank $\leq w$. From the consistency analysis one can see that $\boldsymbol{e}'$ can be used as an alternate private key in the Decrypt algorithm. The computation of such a vector $\boldsymbol{e}'$ actually corresponds to the search version of Gab-SD problem.

This problem is easy for $w \leq \lfloor \frac{n-k}{2} \rfloor$ (it corresponds to half-minimum-distance decoding) and for $w \geq n-k$ (equivalent to interpolation for linearized polynomials). For our concern, we have $\lfloor \frac{n-k}{2} \rfloor < w < n-k$, and we believe that the search version of Gab-SD is hard in this range of parameters.

An approach is proposed in [54] in order to tackle this problem. The solution consists in enumerating vector spaces of dimension slightly higher than w, checking whether they guessed correctly a large part of the solution space, and in such case, interpolating the solution. Roughly speaking, the number of valid choices for the subspace is large, but the complexity of finding one remains exponential in the code length. Precisely, in our settings ($m = n$, and $n-k$ even) the number of vector spaces to test before finding one solution is on average

$$\mathcal{N}_{Class-\text{Gab-SD}} \approx 0.3 \cdot 2^{\delta(n+k-2\delta)},$$

where $\delta \coloneqq w - \lfloor \frac{n-k}{2} \rfloor > 0$. This quantity is used as a bound for the complexity of solving Gab-SD. By using a straightforward Grover algorithm, we obtain that the number of iterations to be completed on a quantum computer is roughly

$$\mathcal{N}_{Quant-\text{Gab-SD}} \approx 0.55 \cdot 2^{\frac{\delta}{2}(n+k-2\delta)}.$$

Attack *via* a reduction to a MinRank instance. The recovery of a representative $\boldsymbol{p}' = \boldsymbol{gSP} \in \mathbb{F}_q^n$ of the plaintext $\boldsymbol{P}$, given only a ciphertext $\boldsymbol{u}$ and $\boldsymbol{k}_{\text{priv}}$, can be modeled as follows. First, one computes (i) any solution $\boldsymbol{x} \in \mathbb{F}_q^n$ of $\boldsymbol{Hx}^\top = \boldsymbol{u}^\top$, and (ii) any solution $\boldsymbol{y} \in \mathbb{F}_q^n$ to $\boldsymbol{Hy}^\top = \boldsymbol{k}_{\text{pub}}{}^\top$. Due to the form of the ciphertext, this leads us to

$$\boldsymbol{x} - \boldsymbol{yT} - \boldsymbol{c} = \boldsymbol{p}', \tag{2.1}$$

where $\boldsymbol{c} \in \mathcal{G}_k(\boldsymbol{g})$ and $\boldsymbol{T} \in \mathbb{F}_2^{n \times n}$ are unknown to the attacker. Notice that $\boldsymbol{T}$ lies in a $\mathbb{F}_2$-vector space of dimension $(\ell+1)n$, since $\boldsymbol{gT} \in \mathcal{G}_{\ell+1}(\boldsymbol{g})$. Two kinds of attacks can then be mounted to solve (2.1).

First, Equation (2.1) can be written $\boldsymbol{x} = (\boldsymbol{c} + \boldsymbol{yT}) + \boldsymbol{p}'$, which means that the problem can be rephrased as decoding an error $\boldsymbol{p}'$ of rank t in the underlying code

$$\mathcal{D} \coloneqq \mathcal{G}_k(\boldsymbol{g}) + \langle \{\boldsymbol{yT} \mid \boldsymbol{T} \in \mathcal{M}_\ell\} \rangle_{\mathbb{F}_2}$$

Notice that $\mathcal{D} \subseteq \mathbb{F}_q^n$ is an $\mathbb{F}_2$-linear code of $\mathbb{F}_2$-dimension at most $(k+\ell+1)n$. One can then write $\boldsymbol{yT} = L_{\boldsymbol{y}}\langle \boldsymbol{gT} \rangle$, which yields $\mathcal{D} = \mathcal{G}_k(\boldsymbol{g}) + L_{\boldsymbol{y}}\langle \mathcal{G}_{\ell+1}(\boldsymbol{g}) \rangle$. One could try to decode

in the smallest $\mathbb{F}_q$-linear code containing $\mathcal{D}$, and use the additional structure provided by the $\mathbb{F}_q$-linearity. This structure has been widely employed in the recent improvements, see [24].

Second, one can see Equation (2.1) as an instance of MinRank, a problem formally introduced by Courtois in [55] after the cryptanalysis of HFE [56].

Problem 1 (MinRank search problem)**.** *Let* $\mathbb{K}$ *be a field.*

— **Input:** $\boldsymbol{M}_0, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K \in \mathbb{K}^{N\times n}$ *and an integer* t*.*

— **Goal:** *Find* $(x_1, \ldots, x_K) \in \mathbb{K}^K$ *such that* $\mathrm{rk}_{\mathbb{K}}(\boldsymbol{M}_0 - \sum_{i=1}^{K} x_i \boldsymbol{M}_i) \leq t$*.*

Let us denote by $\{\boldsymbol{T}_1, \ldots, \boldsymbol{T}_{n(\ell+1)}\} \subseteq \mathbb{F}_2^{n\times n}$ an $\mathbb{F}_2$-basis of $\mathcal{M}_\ell$. Similarly, $\mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{c})$ can be written in some basis $\{\boldsymbol{C}_1, \ldots, \boldsymbol{C}_{nk}\} \subseteq \mathbb{F}_2^{n\times n}$ of the $\mathbb{F}_2$-vector space of dimension nk representing $\mathcal{G}_{k+\ell}(\boldsymbol{g})$. Applying $\mathrm{Ext}_{\boldsymbol{g}}$ to Equation (2.1), we get:

$$\boldsymbol{X} - \sum_{i=1}^{n(\ell+1)} t_i \boldsymbol{Y}\boldsymbol{T}_i - \sum_{i=1}^{nk} c_i \boldsymbol{C}_i = \boldsymbol{P}',$$

where $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{P}') = (\mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{x}), \mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{y}), \mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{p}'))$. Since $\mathrm{rk}(\boldsymbol{P}') = t$, one gets an instance of the MinRank problem, with one « base matrix » $\boldsymbol{X} \in \mathbb{F}_2^{n\times n}$ and $K := n(k + \ell + 1)$ « summand matrices » $\{\boldsymbol{Y}\boldsymbol{T}_1, \ldots, \boldsymbol{Y}\boldsymbol{T}_{n(\ell+1)}, \boldsymbol{C}_1, \ldots, \boldsymbol{C}_{n(k+\ell)}\}$.

There exist several approaches to solve the MinRank problem. In [57], Goubin and Courtois gave an algorithm which finds a solution in expected time $\mathcal{O}(K^3 2^{t\lceil K/n\rceil})$. In 1999, Kipnis and Shamir [56] proposed a multivariate formulation of MinRank which can be solved by computing Groebner bases. Such computations can be run in time $\mathcal{O}(\binom{m+d-1}{d}^\omega)$, where $2 \leq \omega < 3$ is the linear algebra constant, $m = t(n-t) + K$ and d is the *degree of regularity* of the system [58]. Faugère, Levy-dit-Vehel and Perret [59] proved that, in the Kapnis-Shamir formalism, any instance can be reduced to a simpler one if $\Delta := K - (n-t)^2 > 0$. In our case, setting $w \geq \ell + 1$ ensures that $\Delta \leq 0$. Moreover, the authors proved that the degree of regularity is lower than what is expected for random systems, and it seems to be upper bounded by $t + 2$ heuristically. This heuristic was confirmed by Verbel *et al.* [60] for superdetermined instances, and by Bardet *et al.* [24] in the context of decoding low rank errors in random codes. Finally, the latter work also presents instances for which the solving degree decreases to $d = t$. We choose to consider this conservative setting; the running time for the computation of the associated Groebner

basis is thus in

$$\mathcal{O}\left(\binom{t(n-t)+n(k+\ell+1)+t-1}{t}^{\omega}\right).$$

To sum up, the reduction to MinRank leads us to the following bounds on the parameters:

$$w \geq \ell+1, \qquad \omega \cdot \log\binom{n(k+\ell+t+1)-t^2+t-1}{t} \geq \lambda, \qquad n(k+\ell+1) \geq \lambda\,.$$

4.4 Bombar - Couveur attack

Shortly after the publication of our preprint paper, Bombar and Couvreur proposed an attack on it [29]. This is an attack on the ciphertext $\boldsymbol{y}$ by considering all vectors as block matrices in $\mathbb{F}_2^{n\times n}$ and the linearized polynomials as the $\mathbb{F}_2$-endomorphisms.

By the definition of Gabidulin code, we also can denote $\mathcal{G}_k(\boldsymbol{g}) = \{\boldsymbol{M}_P\boldsymbol{G}\}$ where $\boldsymbol{M}_P$ is representation matrix of a linearized polynomial $P \in \mathbb{F}_q[X;\theta]$ of degree $< k$ and $\boldsymbol{G} = \mathrm{Ext}_\beta(\boldsymbol{g})$.

Remark that for a matrix of $\boldsymbol{g}$-degree ℓ, there exists a q-polynomial $L_{\boldsymbol{T}}(X) \in \mathbb{F}_q[X;\theta]$ of degree ℓ such that $\boldsymbol{gT} = L_{\boldsymbol{T}}(\boldsymbol{g})$. Let $\boldsymbol{M}_T$ be the representation of a linearized polynomial $L_{\boldsymbol{T}}(X)$ then $\boldsymbol{GT} = \boldsymbol{M}_T\boldsymbol{G}$, or equivalently, $\boldsymbol{GT} = \boldsymbol{G}^{-1}M_T\boldsymbol{G}$

The ciphertext equation:

$$\boldsymbol{Y} = \boldsymbol{C} + \boldsymbol{KT} + \boldsymbol{E}$$

where $\boldsymbol{Y}, \boldsymbol{C}, \boldsymbol{K}, \boldsymbol{E} \in \mathbb{F}_2^{n\times n}$ are the representations of $\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{k}, \boldsymbol{e}$ in $\mathbb{F}_2$

Since $\mathrm{rk}(\boldsymbol{e}) = t$, there exists $\boldsymbol{R} = \boldsymbol{G}^{-1}\boldsymbol{M}_R\boldsymbol{G} \in \mathbb{F}_2^{n\times n}$ of $\boldsymbol{g}$-degree t such that $\boldsymbol{ER} = 0$. (see Proposition 2, [29]). Thus, $\boldsymbol{YR} = \boldsymbol{M}_C\boldsymbol{M}_R\boldsymbol{G} + \boldsymbol{K}\boldsymbol{G}^{-1}\boldsymbol{M}_T\boldsymbol{M}_R\boldsymbol{G}$.

The right side is an element in $\mathcal{G}_{k+\ell+t}(\boldsymbol{g}) + \boldsymbol{K}\boldsymbol{G}^{-1}\mathcal{G}_{\ell+t+1}$ which is well computed by $\mathcal{G}_k(\boldsymbol{g}) \oplus \boldsymbol{K}$.

We need to solve the following system of equations:

$$\boldsymbol{Y}\boldsymbol{G}^{-1}\boldsymbol{M}_R\boldsymbol{G} = \boldsymbol{N}. \tag{2.2}$$

where $\boldsymbol{N} \in \mathcal{G}_{k+\ell+t}(\boldsymbol{g}) + \boldsymbol{K}\boldsymbol{G}^{-1}\mathcal{G}_{\ell+t+1}(\boldsymbol{g})$ and $\deg(\boldsymbol{R}) \leq t$. This is an linear system of equation of $k+2\ell+3t+2$ unknowns ($k+2\ell+t+1$ of $\boldsymbol{N}$ and $t+1$ of $\boldsymbol{M}_R$) and n equations.

In the case of $k+2\ell+3t+2 \leq n$, then solving this system of equations will very likely gives us pairs of the form $(\boldsymbol{M}_R, \boldsymbol{N})$ where $\boldsymbol{ER} = 0$. The cryptanalysis is as follows:

1. Solve System 2.2
2. Take the nonzero solution of the system and obtain M_R. Compute $\boldsymbol{R} = \boldsymbol{G}^{-1}\boldsymbol{M}_R\boldsymbol{G}$
3. Compute the left kernel of $\boldsymbol{R}$. This kernel contains the rowspace $\mathcal{V}$ of $\boldsymbol{E}$.
4. Recover the RREF of any matrix V whose rowspace generates $\mathcal{V}$ and it is the plaintext.

In the nutshell, we summarize the parameters corresponding to the security against the existing attack in section 2.4.3. However, all of these parameters are vulnerable for Bombar- Couvreur attack as the following table

n	k	w	ℓ	t	Security (bits)	Bombar and Couvreur attack
64	32	19	3	5	141	$O(2^{17})$
80	40	23	3	7	202	$O(2^{18})$
96	48	27	3	9	265	$O(2^{19})$
164	116	27	3	9	256	$O(2^{21})$

Table 2.1 – 3 first rows are sets of parameters for RAMESSES as a KEM and the last one as a PKE, with different levels of security.

Chapter 3

A new family of codes : Sum of Gabidulin codes and their decoding

There are very few families of decodable codes in rank metric. Namely, the family of trivial codes [61], the family of Gabidulin codes [36], and the family of LRPC codes [62]. Apart from that, there are codes derived from Gabidulin codes that are used in [63]. These codes are masked versions of Gabidulin codes, enabling to design public-key encryption schemes.

The work in this chapter comes from a published paper in the conference ISIT 2021 [64]. In this chapter, we investigate the problem of decoding the sum of Gabidulin codes. Interestingly enough, this is a problem which appears when one analyses cryptosystems based on the problem of reconstructing linearized polynomials, [2, 5].

We show that the formulation of this problem can give some insight in decoding problems and give rise to further understanding on how to design public-key cryptosystems.

In the first section, we introduce notations and especially the notion of skew polynomial rings which are an elegant and simple manner to deal algebraically with rank metric and Gabidulin codes. Under this setting, Gabidulin codes are just the evaluation codes of bounded degree skew polynomials using the operator evaluation.

Then, we state the problem of decoding the sum of Gabidulin codes and show that there is a simple probabilistic polynomial-time decoder up to some bound. Under some assumptions, we show that the failure probability of this algorithm is exponentially small.

Finally, we present some potential applications. We show that by considering a random k-dimensional code as the sum of k 1-dimensional Gabidulin codes, we recover the result of [65] for the decoding of random codes. We also show that investigating properties of the sum of Gabidulin codes could be of interest in designing and analysing rank metric based public-key cryptography based on algebraic decoding.

1 Decoding of the sum of Gabidulin codes

We recall here the definition of Gabidulin code which is the definition 1 in chapter 1. In this setting, Gabidulin codes are defined as evaluation codes of skew polynomials over linearly independent elements (see p.20-21).

Definition 4. *Let $\boldsymbol{g} = (g_1, \ldots, g_n) \in \mathbb{F}_{q^m}^n$, formed with $\mathbb{F}_q$-linearly independent elements. The Gabidulin code of dimension k and of support $\boldsymbol{g}$ denoted by $\mathcal{G}_k(\boldsymbol{g})$ is defined by*

$$\mathcal{G}_k(\boldsymbol{g}) = \left\{ f\langle \boldsymbol{g} \rangle, \quad \begin{array}{l} f \in \mathbb{F}_{q^m}[X;\theta] \\ \deg(f) \leq k-1 \end{array} \right\}$$

Definition 5 (Sum of Gabidulin codes)**.** *Let $\{\mathcal{G}_{k_j}\langle \boldsymbol{g}_j \rangle \subset \mathbb{F}_{q^m}^n\}_{j=1}^{\ell}$ be a set of k_j-dimensional Gabidulin codes with support vectors $\boldsymbol{g}_j$. We define by*

$$\mathcal{C} = \sum_{j=1}^{\ell} \mathcal{G}_{k_j}(\boldsymbol{g}_j) = \left\{ \sum_{j=1}^{\ell} f_j\langle \boldsymbol{g}_j \rangle, \quad \begin{array}{l} f_j \in \mathbb{F}_{q^m}[X;\theta] \\ \deg(f_j) \leq k_j - 1 \end{array} \right\}$$

the code formed with the sum of Gabidulin codes.

Our goal is to study in which case we can decode it and up to which bound in the rank metric.

1.1 Problem

The decoding problem we address is the following: Let

$$\boldsymbol{y} = \mathbf{c} + \mathbf{e}$$

where $\mathbf{c} \in \mathcal{C}$ and $\mathbf{e}$ has rank t. This implies that

— There exists skew polynomials $f_j \in \mathbb{F}_{q^m}[X;\theta]$ with degree $\leq k_j - 1$, such that

$$\mathbf{c} = \sum_{j=1}^{\ell} f_j\langle \boldsymbol{g}_j \rangle$$

— There exists a unique skew polynomial $A_{\mathbf{e}} \in \mathbb{F}_{q^m}[X;\theta]$, monic and of degree t such that $A_{\mathbf{e}}\langle \mathbf{e} \rangle = 0$

From the $\mathbb{F}_q$-linearity of the evaluation of skew polynomials, we can rewrite the decoding problem under the following form

$$A_{\mathbf{e}}\langle \boldsymbol{y} \rangle = \sum_{j=1}^{\ell} (A_{\mathbf{e}} \cdot f_j)\langle \boldsymbol{g}_j \rangle \tag{3.1}$$

The unknowns of the system are the coefficients of the skew polynomials. Hence, we obtain a non homogenous bivariate system with $t + \sum_{j=1}^{\ell} k_j + 1$ unknowns and n equations.

A way to decode would then be to homogeneize the system and solve it by using Gröbner bases, but this is not the direction we investigate. As in [66, 67], we prefer to linearize the system and understand when the solution space is 1-dimensional to relate it directly to the decoding of $\mathcal{C}$.

1.2 Linearizing the problem

We now consider the following system:

$$A\langle \boldsymbol{y} \rangle = \sum_{j=1}^{\ell} N_j \langle \boldsymbol{g}_i \rangle \tag{3.2}$$

where A has degree t and for $j \in \{1, \dots, \ell\}$, $\deg(N_j) \leq t + k_j - 1$. The number of equations is equal to n and the number of variables can be counted as such:

— $t + 1$ variables to characterize the skew polynomial A;

— $\ell t + \sum_{j=1}^{\ell} k_j$ to characterize the polynomials N_j.

Hence, the number of variables is $(\ell + 1)t + \sum_{j=1}^{\ell} k_j + 1$. In case $(\ell + 1)t + \sum_{j=1}^{\ell} k_j < n$, this implies that the matrix of the system is degenerate, and the solution space often is of small dimension, typically 1.

We can relate the solutions of system (3.1) and (3.2) by the following immediate theorem.

Theorem 5. *Let $(f_1, \dots, f_\ell, V)$ be a solution of (3.1) then $(V \cdot f_1, \dots, V \cdot f_\ell, V)$ is a solution to (3.2)*

This theorem is a straightforward generalisation of the systems written for the Welch-Berlekamp decoding algorithm [66, 67]. More precisely, we can prove the following theorem

Theorem 6. *Let* $\boldsymbol{y} \in \mathbb{F}_{q^m}^n$. *Let*

$$\mathcal{L}_{\boldsymbol{y},t}(\mathcal{C}) = \{(\boldsymbol{c}_i, \boldsymbol{e}_i) \mid \boldsymbol{y} = \boldsymbol{c}_i + \boldsymbol{e}_i,\ \boldsymbol{c}_i \in \mathcal{C}, \mathrm{rk}_q(\boldsymbol{e}_i) \leq t\}$$

If the solution space of (3.2) is 1*-dimensional, then there is at most one element in* $\mathcal{L}_{\boldsymbol{y},t}(\mathcal{C})$. *Moreover any non zero solution* $(A, N_1, \ldots, N_\ell)$ *of the system provides the same solution* $(A, A\backslash N_1, \ldots, A\backslash N_\ell)$ *to (3.1), where* $A\backslash N$ *denotes the left Euclidean division of* N *by* A *in* $\mathbb{F}_{q^m}[X;\theta]$.

This gives a natural decoding algorithm consisting in enumerating the solution space of system (3.2). Let d be the dimension of this solution space this gives a list decoding algorithm recovering $\mathcal{L}_{\boldsymbol{y},t}$ with complexity

$$\mathcal{O}\left(P(n,m)q^{m(d-1)}\right),$$

where P is a polynomial of degree at most 3. The exponent is $d-1$ and not d since we only need to enumerate the 1-dimensional vector spaces and not all the elements in the solution space. Namely, we need to enumerate the solution space of the linear system, and then complete the left Euclidean division in $\mathbb{F}_{q^m}[X;\theta]$ corresponding to the linear algebra operations.

Corollary 4. *A necessary condition for the dimension of system (3.1) to be* ≤ 1 *is*

$$(\ell+1)t + \sum_{j=1}^{\ell} k_j < n$$

Proof. Suppose that $|\mathcal{L}_{\boldsymbol{y},t}(\mathcal{C})| \geq 2$. Let $(\boldsymbol{c}_1, \boldsymbol{e}_1)$ and $(\boldsymbol{c}_2, \boldsymbol{e}_2)$ be two distinct elements of $\mathcal{L}_{\boldsymbol{y},t}(\mathcal{C})$. Then they respectively correspond to solutions $(A_{\boldsymbol{e}_1}, f_1, \ldots, f_\ell)$ and $(A_{\boldsymbol{e}_2}, h_1, \ldots, h_\ell)$ of (3.1). Therefore $(A_{\boldsymbol{e}_1}, A_{\boldsymbol{e}_1}\cdot f_1, \ldots, A_{\boldsymbol{e}_1}\cdot f_\ell)$ and $(A_{\boldsymbol{e}_2}, A_{\boldsymbol{e}_2}\cdot h_1, \ldots, A_{\boldsymbol{e}_2}\cdot h_\ell)$ are solutions of (3.2) . From the hypothesis that the solution vector space is 1-dimensional, and the fact that $A_{\boldsymbol{e}_1}$ and $A_{\boldsymbol{e}_2}$ are monic, this implies that $A_{\boldsymbol{e}_1} = A_{\boldsymbol{e}_2}$, and that any solution has the form

$$\alpha \cdot (A_{\boldsymbol{e}_1}, A_{\boldsymbol{e}_1} \cdot f_1, \ldots, A_{\boldsymbol{e}_1} \cdot f_\ell),\ \alpha \in \mathbb{F}_{q^m}$$

□

1.3 Discussion on the failure probability

In this section, we investigate the failure probability, that is we consider that we are in the conditions of Corollary 4, where $(\ell+1)t+\sum_{j=1}^{\ell} k_j < n$ and where the solution space of (3.2) has dimension $d \geq 2$. This corresponds to the case where the decoding cannot be completed in polynomial-time.

As in [68], we define the operator λ_t which maps a matrix $\boldsymbol{M} = (m_{ij}) \in \mathbb{F}_{q^m}^{s\times n}$ to a block matrix:

$$\begin{array}{rccc} \lambda_t : & \mathbb{F}_{q^m}^{s\times n} & \to & \mathbb{F}_{q^m}^{s(t+1)\times n} \\ & \boldsymbol{M} & \mapsto & \begin{pmatrix} \boldsymbol{M} \\ \vdots \\ \boldsymbol{M}^{[t]} \end{pmatrix}, \end{array}$$

where $\boldsymbol{M}^{[u]} := (\theta^u(m_{ij}))$. Let $(A, N_1, \ldots, N_\ell) \in \mathbb{F}_{q^m}[X;\theta]$ be a solution to the linear system (3.2). We now identify any polynomial in $\mathbb{F}_{q^m}[X;\theta]$ with the vector formed by its coefficients. Then solving (3.2) is equivalent to solving the following linear system

$$\boldsymbol{M} \cdot \begin{pmatrix} -A \\ N_1 \\ \vdots \\ N_\ell \end{pmatrix} = 0$$

where $\boldsymbol{M} = (\lambda_t(\boldsymbol{y})^\top, \lambda_{t+k_1-1}(\boldsymbol{g}_1)^\top, \cdots, \lambda_{t+k_\ell-1}(\boldsymbol{g}_\ell)^\top)$.

The matrix $\boldsymbol{M}$ is a $n \times ((\ell+1)t + \sum_{j=1}^{\ell} k_j + 1)$ - matrix in $\mathbb{F}_{q^m}$. Its kernel $\ker(\boldsymbol{M})$ is the solution space of (3.2). From our hypotheses, the dimension of $\ker(\boldsymbol{M})$ is at least 1. A sufficient condition to be able to decode in polynomial-time is that the kernel of $\boldsymbol{M}$ is exactly 1.

We need to compute the probability of non-unique decoding $\mathcal{P}(|\mathcal{L}_{\boldsymbol{y},t}(\mathcal{C})| > 1) = \mathcal{P}(\dim\ker(\boldsymbol{M}) > 1)$. By the rank-nullity theorem, $\dim\ker(\boldsymbol{M}) + \mathrm{rk}_{q^m}(\boldsymbol{M}) = (\ell+1)t + \sum_{j=1}^{\ell} k_j + 1$. If the dimension of the solution space is greater than 1, then $\mathrm{rk}_{q^m}(\boldsymbol{M}) < (\ell+1)t + \sum_{j=1}^{\ell} k_j$. Let $\mathcal{M}$ be the set of all $n \times ((\ell+1)t + \sum_{j=1}^{\ell} k_j + 1)$ - matrix in $\mathbb{F}_{q^m}$ of the form

$$\left(\lambda_t(\boldsymbol{g}_0)^\top, \lambda_{t+k_1-1}(\boldsymbol{g}_1)^\top, \cdots, \lambda_{t+k_\ell-1}(\boldsymbol{g}_\ell)^\top\right).$$

Let us also define $\mathcal{A}$ as the set of all $n \times (\ell+1)$ - matrices in $\mathbb{F}_{q^m}$ of the form

$$\left(\boldsymbol{g}_0^\top, \boldsymbol{g}_1^\top, \ldots, \boldsymbol{g}_\ell^\top\right)$$

such that $\boldsymbol{g}_0, \boldsymbol{g}_1, \ldots, \boldsymbol{g}_\ell \in \mathbb{F}_{q^m}^n$ and $\forall j \in \{1, \ldots, \ell\}, \mathrm{rk}(\boldsymbol{g}_j) = n$. Given a set of integers $(k_0 = 1, k_1, \ldots, k_\ell)$, we define the following bijection

$$\begin{array}{rccc} \varphi: & \mathcal{A} & \to & \mathcal{M} \\ & \left(\boldsymbol{g}_0^\top, \ldots, \boldsymbol{g}_\ell^\top\right) & \mapsto & \left(\lambda_{t+k_0-1}(\boldsymbol{g}_0)^\top, \ldots, \lambda_{t+k_\ell-1}(\boldsymbol{g}_\ell)^\top\right) \end{array}$$

Theorem 7. *For $\boldsymbol{A} \xleftarrow{\$} \mathcal{A}$, and $\boldsymbol{M} = \varphi(\boldsymbol{A})$. Let $r := \mathrm{rk}_{q^m}(\boldsymbol{M})$. Then*

$$\mathcal{P}_{\mathcal{A}}\left[r < (\ell+1)t + \sum_{j=1}^{\ell} k_j\right] \leq \frac{\binom{n}{r+1}4^\ell}{q^{2m}} \leq \frac{1}{q^m}$$

Proof. For $\boldsymbol{A}$ is chosen uniformly from $\mathcal{A}$ and $\boldsymbol{M} = \varphi(\boldsymbol{A})$, of rank $r < (\ell+1)t + \sum_{j=1}^{\ell} k_j$, let

$$\mathcal{S} := \{\boldsymbol{h} \in \mathbb{F}_{q^m}^n \,|\, \boldsymbol{h} \text{ has } n-(r+1) \text{ coordinates } 0\}$$

Since $r < (\ell+1)t + \sum_{j=1}^{\ell} k_j$, there exists $\boldsymbol{h} \in \mathcal{S}$ such that $\boldsymbol{h}\boldsymbol{M} = 0$. Thus, for $j \in \{0, \ldots, \ell\}$,

$$\lambda_{t+k_j-1}(\boldsymbol{g}_j)\boldsymbol{h}^\top = 0 \tag{3.3}$$

This in particular implies that $\boldsymbol{h} \in \bigcap_{j=1}^{\ell} \mathcal{G}^{\perp}_{t+k_j-1}(\boldsymbol{g}_j)$. Therefore $\mathrm{rk}_q(\boldsymbol{h}) \geq t + k_{\max}$ where $k_{\max} := \max\{k_j\}$. Let

$$\mathcal{A}_{\boldsymbol{h}} = \{\boldsymbol{A} \in \mathcal{A} \mid \boldsymbol{h}\varphi(\boldsymbol{A}) = 0\}$$

We determine the probability that for a fixed $\boldsymbol{h} \in \mathcal{S}$ with $\mathrm{rk}_q(\boldsymbol{h}) \geq t + k_{\max}$, there exists $\boldsymbol{A} \in \mathcal{A}_{\boldsymbol{h}}$. This probability will be $\dfrac{|\mathcal{A}_{\boldsymbol{h}}|}{|\mathcal{A}|}$. Then,

$$\mathcal{P}_{\mathcal{A}}\left[r < (\ell+1)t + \sum_{j=1}^{\ell} k_j\right] \leq \frac{1}{q^m - 1} \sum_{\boldsymbol{h} \in \mathcal{S}, \mathrm{rk}_q(\boldsymbol{h}) \geq t + k_{\max}} \frac{|\mathcal{A}_{\boldsymbol{h}}|}{|\mathcal{A}|} \tag{3.4}$$

The term $1/(q^m - 1)$ coming from the fact that for any vector $\boldsymbol{h} \in \mathcal{S}$, and for any $\alpha \in \mathbb{F}_{q^m} \setminus \{0\}$, we have $\mathcal{A}_{\boldsymbol{h}} = \mathcal{A}_{\alpha\boldsymbol{h}}$. For a given $\boldsymbol{h} \in \mathcal{S}$ we now look at the cardinality of $\mathcal{A}_{\boldsymbol{h}}$. Now let $\boldsymbol{A} \in \mathcal{A}_{\boldsymbol{h}}$, this implies that $\lambda_{t+k_j-1}(\boldsymbol{g}_j)\boldsymbol{h}^\top = 0$, for $j \in \{0, \dots, \ell\}$. In particular this implies

$$\forall i \in \{0, \dots, t + k_j - 1\}, \quad \boldsymbol{g}_j^{[i]}\boldsymbol{h}^\top = 0.$$

Therefore, by applying the inverse of θ a sufficient number of times, we obtain $\forall i \in \{0, \dots, t+k_j-1\}$, $\boldsymbol{g}_j(\boldsymbol{h}^{[-i]})^\top = 0$. Now let $\boldsymbol{h}_j := \boldsymbol{h}^{[-(t+k_j-1)]}$. Then $\forall i \in \{0, \dots, t+k_j-1\}$, $\boldsymbol{g}_j(\boldsymbol{h}_j^{[i]})^\top = 0$. It implies that $\lambda_{t+k_j-1}(\boldsymbol{h}_j)\boldsymbol{g}_j^\top = 0$.

We need the following lemma:

Lemma 11 (Lemma 3.51 [69])**.** *Given $\boldsymbol{g} \in \mathbb{F}_{q^m}^n$ then $\mathrm{rk}_{q^m}(\lambda_k(\boldsymbol{g})) = \min\{k+1, \mathrm{rk}_q(\boldsymbol{g})\}$.*

Since $\mathrm{rk}_q(\boldsymbol{h}) \geq t + k_{\max}$, Lemma 11 implies that for $j \in \{0, \dots, \ell\}$, $\mathrm{rk}_{q^m}(\lambda_{t+k_j-1}(\boldsymbol{h}_j)) = t + k_j$. Moreover,

$$\dim \ker(\lambda_{t+k_j-1}(\boldsymbol{h}_j)) + \mathrm{rk}_{q^m}(\lambda_{t+k_j-1}(\boldsymbol{h}_j)) = n$$

Hence, $\dim \ker(\lambda_{t+k_j-1}(\boldsymbol{h}_j)) = n - (t + k_j)$. It implies that the number of possible vectors $\boldsymbol{g}_j$ is at most $(q^m)^{n-(t+k_j)}$. Therefore,

$$|\mathcal{A}_{\boldsymbol{h}}| \leq \prod_{j=0}^{\ell} (q^m)^{n-(t+k_j)} = (q^m)^{(\ell+1)(n-t)-\sum_{j=1}^{\ell} k_j - 1}$$

To complete the proof, we also need the following lemma

Lemma 12 (Lemma 3.13 [68])**.** *Given $n \leq m$, the number of matrixes $\boldsymbol{A} \in \mathbb{F}_q^{m\times n}$ such that $\mathrm{rk}_q(\boldsymbol{A}) = n$ is larger than $\dfrac{q^{mn}}{4}$. As a consequence, it is also the lower bound for the number of vectors $\boldsymbol{u} \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}_q(\boldsymbol{u}) = n$.*

Since $\forall j \in \{1, \dots, \ell\}$, $\mathrm{rk}_q(\boldsymbol{g}_j) = n$, from Lemma 12, the number of possible vectors $\boldsymbol{g}_j$ is $\dfrac{q^{mn}}{4}$. Moreover $\boldsymbol{g}_0$ can be chosen completely arbitrarily, thus adding a factor of q^{mn}. Hence, the number of possible matrices $\boldsymbol{A} \in \mathcal{A}$ is greater than $\left(\dfrac{q^{mn}}{4}\right)^{\ell} q^{mn}$. Thus,

$$|\mathcal{A}| \geq \frac{(q^m)^{n(\ell+1)}}{4^\ell} \text{ and } \frac{|\mathcal{A}_{\boldsymbol{h}}|}{|\mathcal{A}|} \leq \frac{4^\ell}{(q^m)^{\sum_{j=1}^{\ell} k_j + (\ell+1)t + 1}}$$

Finally we have

$$\frac{|\mathcal{S}|}{q^m-1} = \frac{\binom{n}{r+1}(q^m-1)^{r+1}}{q^m-1} \approx \binom{n}{r+1} q^{mr}$$

From the inequality (3.4), we obtain that

$$\mathcal{P}_{\mathcal{A}}\left[r < (\ell+1)t + \sum_{j=1}^{\ell} k_j\right] \leq \frac{\binom{n}{r+1}4^{\ell} q^{mr}}{q^{m\cdot(\sum_{j=1}^{\ell} k_j + (\ell+1)t+1)}}$$

Now since $1 \leq \sum_{j=1}^{\ell} k_j + (\ell+1)t - r$, we have

$$\mathcal{P}_{\mathcal{A}}\left[r < (\ell+1)t + \sum_{j=1}^{\ell} k_j\right] \leq \frac{\binom{n}{r+1}4^{\ell}}{q^{2m}} \leq \frac{1}{q^m}$$

□

Now we can sum up and establish our main result

Theorem 8 (Main theorem)**.** *Let $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_\ell$ be a randomly chosen set of vectors of rank n in $(\mathbb{F}_{q^m})^n$. Let*

$$\mathcal{C} = \sum_{j=1}^{\ell} \mathcal{G}_{k_j}(\boldsymbol{g}_j)$$

then $\mathcal{C}$ can be decoded up to t with a failure probability upper-bounded by q^{-m} with a polynomial-time complexity, under the condition that $(\ell+1)t + \sum_{j=1}^{\ell} k_j < n$.

2 Applications

In this section we give examples where the previous theorem has some applications. We do not claim to have obtained extraordinary new results, but we emphasize that this new point of view in decoding could have interesting cryptographic applications.

2.1 Decoding of Interleaved code

As in [70], we consider the following model of channel: The error positions are all taken in the same q-ary vector space $\mathcal{E}$, of dimension t, i.e, every error vector $\boldsymbol{e} = (e_1, \ldots, e_n)$ of length n such that for all $i \in \{1, \ldots, n\}$, $e_i \in \mathcal{E}$. Let be A the unique monic linearized

polynomial of degree t such that for all $e \in \mathcal{E}, A\langle e\rangle = 0$. Suppose that through this channel, one receives u messages $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(u)}$, such that

$$\forall i \in \{1, \ldots, u\},\ \boldsymbol{y}^{(i)} = \boldsymbol{c}_i + \boldsymbol{e}_i,$$

where $\boldsymbol{c}_i \in \mathcal{C}$. Thus, for all $i \in \{1, \ldots, u\}$, $\boldsymbol{y}^{(i)} = \sum_{j=1}^{\ell} f_j^{(i)} \langle \boldsymbol{g}_j \rangle + \boldsymbol{e}_i$ and

$$A\langle \boldsymbol{y}^{(i)} \rangle = \sum_{j=1}^{\ell} (A \cdot f_j^{(i)}) \langle \boldsymbol{g}_j \rangle \tag{3.5}$$

As in the normal case of interleaving, this implies that

$$A\langle \boldsymbol{y}^{(i)} \rangle = \sum_{j=1}^{\ell} N_j^{(i)} \langle \boldsymbol{g}_j \rangle \tag{3.6}$$

where, for $i \in \{1, \ldots, u\}$, $N_j^{(i)} \in \mathbb{F}_q[X; \theta]$ has degree $\leq t + k_j - 1$. The system (3.6) is linear in $t(u\ell + 1) + u(\sum_{j=1}^{\ell} k_j) + 1$ unknowns (the coefficients of polynomials) and nu equations.

Therefore we can hope to decode up to errors of rank

$$t \leq \left\lfloor u \left(n - \sum_{j=1}^{\ell} k_j \right) / (u\ell + 1) \right\rfloor$$

We denote $\boldsymbol{G} = \left[(\lambda_{t+k_1}(\boldsymbol{g}_1))^\top, \cdots, (\lambda_{t+k_\ell}(\boldsymbol{g}_\ell))^\top \right]$, $\boldsymbol{Y}_i = (\lambda_t(\boldsymbol{y}^{(i)}))^\top$ and

$$\boldsymbol{M} = \left(\begin{array}{cccc|c} \boldsymbol{G} & 0 & \cdots & 0 & \boldsymbol{Y}_1 \\ 0 & \boldsymbol{G} & \cdots & 0 & \boldsymbol{Y}_2 \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & \cdots & \boldsymbol{G} & \boldsymbol{Y}_u \end{array} \right)$$

Then,

$$(N_1^{(1)}, \cdots, N_\ell^{(1)}, \cdots, N_1^{(u)}, \cdots, N_\ell^{(u)}, -A)\boldsymbol{M}^\top = 0$$

$\boldsymbol{M}$ is a $nu \times t(u\ell + 1) + u(\sum_{j=1}^{\ell} k_j) + 1$ matrix with entries are in $\mathbb{F}_{q^m}$

Since $\dim\ker(\boldsymbol{M}) + \mathrm{rk}_{q^m}(\boldsymbol{M}) = t(u\ell + 1) + u(\sum_{j=1}^{\ell} k_j) + 1$, the case of non-unique decoding happens if and only if $\mathrm{rk}_{q^m}(\boldsymbol{M}) < t(u\ell + 1) + u(\sum_{j=1}^{\ell} k_j)$. It means that, for all $i \in 1, \ldots, u$, $\mathrm{rk}_{q^m}(\boldsymbol{G}|\boldsymbol{Y}_i) < (\ell + 1) + \sum_{j=1}^{\ell} k_j + t$.

For $i = 1, \ldots, u$, similarly to the Theorem 7 the probability that $\mathrm{rk}_{q^m}(\boldsymbol{G}|\boldsymbol{Y}_i) < (\ell+1)+\sum_{j=1}^{\ell} k_j + t$ is at most $\frac{1}{q^m}$. Therefore, the probability that non-unique decoding happens is at most $\frac{1}{q^{mu}}$.

2.2 On McEliece type rank-metric based cryptosystem

In GPT-type cryptosystem, we could expect to replace the family of Gabidulin codes with the family of sums of Gabidulin codes. However, by studying the effect of Overbeck's distinguisher, we show that it cannot be replaced directly. More recently in [63], a new technique was introduced to scramble Gabidulin codes. If the parameters are not carefully chosen, there exists a simple distinguisher leading to an efficient key recovery attack [6]. We investigate the effect of this attack if the family of Gabidulin codes is replaced by a sum of Gabidulin codes. We show that the attack cannot be easily adapted.

2.2.1 Overbeck 's distinguisher

Let $\mathcal{C} = \sum_{j=1}^{\ell} \mathcal{G}_{k_j}(\boldsymbol{g}_j)$. The idea of Overbeck's distinguisher is to use the automorphism θ to distinguish $\mathcal{C}$ from a random code of same dimension. Let $k = \sum_{j=1}^{\ell} k_j$ and $\mathcal{C}_{rand}$ a random code of dimension k. For the random code $\mathcal{C}_{rand}$, we expect that $\dim_{\mathbb{F}_{q^m}} \mathcal{C}_{rand} + \mathcal{C}_{rand}^{[1]} = \min(n, 2k)$ with high probability, since the usual hypothesis in that case is to suppose that $\mathcal{C}_{rand}$ and $\mathcal{C}_{rand}^{[1]}$ behave like two k-dimensional vector spaces randomly and uniformly chosen. By studying the dimension of $\mathcal{C} + \mathcal{C}^{[1]}$, we can show that it is at most $k + \ell$. For $\ell < k < n/2$, this implies a distinguisher between this code and the random ones. This indicates that substituting Gabidulin codes by sum of Gabidulin codes as such is probably not a good idea.

2.2.2 Loidreau-like encryption scheme

The security of the scheme is supported by two hypotheses

— The public code is indistinguishable from a random code

— Bounded distance decoding in rank metric is a computationally difficult problem

The second point is beyond the scope of this thesis. We are interested in the first point. So let us recall the procedure for generating a public-key/private key pair.

— The private key is $\mathcal{C} = \sum_{j=1}^{\ell} \mathcal{G}_{k_j}(\boldsymbol{g}_j)$

— The public-key is a randomly chosen generator matrix of $\mathcal{C}\boldsymbol{P}^{-1}$ where $\boldsymbol{P} \in \mathbf{GL}(\mathcal{V})$ where $\mathcal{V}$ is a random λ-dimensional $\mathbb{F}_q$- linear subspace of $\mathbb{F}_{q^m}$.

We observe the attack by distinguisher.

1. Distinguishing $\mathcal{C}\boldsymbol{P}^{-1}$ from random codes: If we raise a public-key $\boldsymbol{G}_{\text{pub}}^{[i]}$ to the i-th power of θ we have
$$\boldsymbol{G}_{\text{pub}}^{[i]} = \boldsymbol{S}^{[i]}\boldsymbol{G}^{[i]}(\boldsymbol{P}^{-1})^{[i]}$$
The matrix $\boldsymbol{P}$ has entries in $\mathcal{V}$ but the matrix $\boldsymbol{P}^{-1}$ has no reason to belong to some strict subspace of $\mathbb{F}_{q^m}$. Thus we avoid the invariant subspace attack ([71],[72]).
2. Distinguishing $\mathcal{C}^{\perp}\boldsymbol{P}^{\top} := \mathcal{C}_{\text{pub}}^{\perp}$ from random codes. A generator matrix of $\mathcal{C}_{\text{pub}}^{\perp}$ is $\boldsymbol{H}_{\text{pub}} = \boldsymbol{H}\boldsymbol{P}^{\top}$, where $\boldsymbol{H}$ is a parity-check matrix of $\mathcal{C}$. The invariant subspace attack requires computing $\dim_{\mathbb{F}_{q^m}}(\mathcal{C}_{\text{pub}}^{\perp} + \cdots + \mathcal{C}_{\text{pub}}^{\perp\,[i]})$ but we may not have enough information for $\mathcal{C}^{\perp}$.

Lemma 13. *The dual code of $\mathcal{C}_{pub}$ is*
$$\mathcal{C}_{pub}^{\perp} = \bigcap_{j=1}^{\ell} \mathcal{G}_{n-k_j}(\boldsymbol{h}_j)\boldsymbol{P}^{\top}$$
for some $\boldsymbol{h}_j \in \mathbb{F}_{q^m}^n$ such that $\mathrm{rk}(\boldsymbol{h}_j) = n$.

Proof. This lemma is straightforward from the fact that $\boldsymbol{H}_{\text{pub}} = \boldsymbol{H}\boldsymbol{P}^{\top}$ and $\mathcal{G}_k^{\perp}(\boldsymbol{g}) = \mathcal{G}_{n-k}^{\perp}(\boldsymbol{h})$ for some $\boldsymbol{h} \in \mathbb{F}_{q^m}^n$ ([36]) □

The attack of Alain Couveur and Coggia ([6]) needs the construction of $\mathcal{C}^{\perp}$ to compute $\dim_{\mathbb{F}_{q^m}}(\mathcal{C}_{\text{pub}}^{\perp} + \cdots + \mathcal{C}_{\text{pub}}^{\perp\,[i]})$, so it requires a representation for the basis of $\mathcal{C}_{\text{pub}}^{\perp}$. However, from the lemma, it is only a $n-k$-dimensional subspace of $\mathcal{G}_{n-k_j}(\boldsymbol{h}_j)\boldsymbol{P}^{\top}$. Thus, this approach cannot directly lead to the recovery of the private key.

2.3 Probabilistic polynomial-time decoding of random codes

A direct consequence of Theorem 6, is just a reformulation of a result in [65] showing that it is possible to have a probabilistic polynomial-time decoder for random codes up to a certain dimension. Namely, a k-dimensional random code is the direct sum of k 1-dimensional random codes.

Namely, suppose that $\mathcal{C}$ is a random code with generator matrix

$$G = \begin{pmatrix} \boldsymbol{g}_1 \\ \vdots \\ \boldsymbol{g}_k \end{pmatrix}$$

where $\dim(Span\langle \boldsymbol{g}_j \rangle) = k$. Then,

$$\mathcal{C} = \sum_{j=1}^{k} \mathcal{G}_1(\boldsymbol{g}_j)$$

Therefore, we have the immediate following corollary of theorem 6

Corollary 5. *Let $\mathcal{C}$ be a $[n,k]_r$ linear code over $\mathbb{F}_{q^m}$, then there is a probabilistic polynomial time decoder for $\mathcal{C}$ up to errors of rank*

$$t \leq \left\lfloor \frac{n-k}{k+1} \right\rfloor$$

The sum of k 1-dimensional codes is a random code of dimension k. With this approach, we recover the decoding of a random rank-metric code [65].

Conclusions and perspectives

We end this thesis by summarizing the main contributions and some perspectives.

Chapter 1

In Chapter 1, we generalize the Coggia and Couveur attack in case that the secret subspace is of any dimension λ. This generalization comes from the idea of the Coggia and Couveur attack [6] on Loidreau cryptosystem [7] and fills the gap from the work of Ghatak [28]. In the section 1.4.1, we prove the general version of the distinguisher for any λ

$$\dim_{\mathbb{F}_{q^m}} \left(\mathcal{C}_{\text{pub}}^{\perp} + \mathcal{C}_{\text{pub}}^{\perp\,[1]} + \cdots + \mathcal{C}_{\text{pub}}^{\perp\,[\lambda]} \right) \leq \lambda(n-k) + \lambda$$
$$\mathbb{P}(\dim_{\mathbb{F}_{q^m}} \left(\mathcal{C}_{rand} + \mathcal{C}_{rand}^{[1]} + \cdots + \mathcal{C}_{rand}^{[\lambda]} \right) \leq \min(n, (\lambda+1)k) - a) = O(q^{-ma})$$

where $\mathcal{C}_{rand}$ is a random code of length n and dimension k, a nonnegative integer a and a positive $\lambda < k$. This is a distinguisher for the Loidreau's scheme for any λ and the public code has rate $R_{pub} \geq 1 - 1/\lambda$.

Afterwards, in the section 1.4.2, we exploit the distinguisher for a key-recovery attack. For the case of $\lambda = 3$, we also see the incomplete work in [28] in case $\lambda = 3$ and we give a proof for the claim that the extension of the Coggia and Couvreur attack can be done in polynomial time.

The parameters of (k, n), which $R_{pub} \geq 1-1/\lambda$ should be avoided in Loidreau's scheme. In the future, it will be worthwhile to attempt a modification of the attack to work for lower rate codes $R_{pub} < 1 - 1/\lambda$ as well. Additionally, the algebraic attack in [73] gives us the new set of parameters in [7]:

$m = n$	k	λ	t	PK size	CT size	Decoding	K. Rec
128	20	3	18	34.5kB	1.8kB	2^{180}	2^{311}
128	44	3	14	58kB	1.3kB	2^{275}	2^{308}

In the future, we can continue study the case that $R_{pub} < 1 - 1/\lambda$. In this case, $\dim_{\mathbb{F}_{q^m}}(\mathcal{S})$ probably equals to n where $\mathcal{S} = \left(\mathcal{C}_{\text{pub}}^{\perp} + \mathcal{C}_{\text{pub}}^{\perp\,[1]} + \cdots + \mathcal{C}_{\text{pub}}^{\perp\,[\lambda]}\right)$. Although, we do not have the structure of $\mathcal{S}$, a natural idea is considering the subspace of $\mathcal{S}$ which generated by vectors $\boldsymbol{h}_0^{[j]} + \sum_{i=1}^{\lambda-1} \beta_i^{[u]} \boldsymbol{h}_i^{[j]}$. This subspace can be exploited some information, similarly the Chapter 1.

On the other hand, another direction of research is modifying the code to avoid this kind of attack. It is unsure but one of the ideas is using the sum of Gabidulin codes instead of Gabidulin code in Chapter 3.

Chapter 2

In Chapter 2, based on the idea of the Finiasz-Augot's cryptosystem [74] and Faure-Loidreau cryptosystem [2], we propose a new cryptosystem RAMESSES. In this cryptosystem, the plaintext is a t-dimensional subspace of $\mathbb{F}_2^n$. After giving the proof for the consistency and the estimation of the failure probability, in the section 2.5.3, we considered some known attacks on this cryptosystem such as exhaustive search attack, the reduction to a quadratic system over $\mathbb{F}_2$ and the reduction to a MinRank instance. Recently, it was broken by the Bombar and Couvreur attack [29], which is also introduced in section 2.5.4. Until now, we haven't had a proper way of modifying this system to resist against this attack.

In the future, we will try to repair the RAMESSES cryptosystem to avoid Bombar-Couveur attack. For example, using the multi-key

Algorithm 6: $\mathsf{KeyGen}(1^\lambda)$

Input:

Output: a pair of public/private keys $(\boldsymbol{k}_{\text{pub}}, \boldsymbol{k}_{\text{priv}})$

1 Pick $\mathcal{V} \subset \mathbb{F}_q$ $\mathbb{F}_2$-vector space of dimension $w > \dfrac{n-k}{2}$

2 Pick $\boldsymbol{k}_{\text{priv}} := (\boldsymbol{k}_{\text{priv}}{}^{(1)}, \boldsymbol{k}_{\text{priv}}{}^{(2)}) \xleftarrow{\$} \{\boldsymbol{x} \in \mathcal{V}^n \times \mathcal{V}^n, \text{rk}(\boldsymbol{x}_1) = \text{rk}(\boldsymbol{x}_2) = w\}$

3 Compute $\boldsymbol{k}_{\text{pub}} = (\boldsymbol{k}_{\text{pub}}{}^{(1)}, \boldsymbol{k}_{\text{pub}}{}^{(2)}) \in \mathbb{F}_q^{n-k} \times \mathbb{F}_q^{n-k}$ such that $\boldsymbol{k}_{\text{pub}}{}^{(i)\top} = \boldsymbol{H}\boldsymbol{k}_{\text{priv}}{}^{(i)\top}$

4 Output $(\boldsymbol{k}_{\text{pub}}, \boldsymbol{k}_{\text{priv}})$

Follow the Bombar-Couveur attack, we have a linear system of equation which has $k + 3\ell + 4t + 2$ unknowns and n equations. We can choose the parameter such that

Algorithm 7: $\mathsf{Encrypt}(\boldsymbol{k}_{\text{pub}}, \boldsymbol{P})$

Input: public key $\boldsymbol{k}_{\text{pub}} \in \mathbb{F}_q^{n-k} \times \mathbb{F}_q^{n-k}$, plaintext $\boldsymbol{P} \in \mathcal{P}_{t,n}$
Output: ciphertext $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$

1 Compute any $\boldsymbol{y}_i \in \mathbb{F}_q^n$ such that $\boldsymbol{H}\boldsymbol{y}_i^\top = {\boldsymbol{k}_{\text{pub}}}^{(i)\top}$
2 Pick $\boldsymbol{T}_1, \boldsymbol{T}_2 \xleftarrow{\$} \{\boldsymbol{M} \in \mathbb{F}_2^{n\times n}, \deg_{\boldsymbol{g}}(\boldsymbol{M}) = \ell\}$
3 Pick $\boldsymbol{S} \xleftarrow{\$} \{\boldsymbol{M} \in \mathbb{F}_2^{n\times n}, \mathrm{rk}(\boldsymbol{M}) = n\}$
4 Compute $\boldsymbol{p}' = \boldsymbol{g}\boldsymbol{S}\boldsymbol{P} \in \mathbb{F}_q^n$
5 Output $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$ such that $\boldsymbol{u}^\top = \boldsymbol{H}(\boldsymbol{y}_1\boldsymbol{T}_1 + \boldsymbol{y}_2 T_2 + \boldsymbol{p}')^\top$

Algorithm 8: $\mathsf{Decrypt}(\boldsymbol{k}_{\text{priv}}, \boldsymbol{u})$

Input: private key $\boldsymbol{k}_{\text{priv}} \in \mathbb{F}_q^n$, ciphertext $\boldsymbol{u} \in \mathbb{F}_q^{n-k}$
Output: plaintext $\hat{\boldsymbol{P}} \in \mathcal{P}_{t,n}$, or failure

1 Compute a solution $\boldsymbol{x} \in \mathbb{F}_q^n$ to the linear system $\boldsymbol{H}\boldsymbol{x}^\top = \boldsymbol{u}^\top$.
2 Compute $\boldsymbol{z} = V_{\boldsymbol{k}_{\text{priv}}}(\boldsymbol{x}) \in \mathbb{F}_q^n$.
3 Decode $\boldsymbol{z}$ as a corrupted $\mathrm{Gab}_{k+\ell+w}(\boldsymbol{g})$-codeword. If success, one gets an error vector $\boldsymbol{a} \in \mathbb{F}_q^n$ of rank $\leq t$.
4 **If** $\mathrm{rk}(\boldsymbol{a}) < t$, output failure.
5 **Otherwise**, output $\hat{\boldsymbol{P}} = \mathtt{RREF}(\mathrm{Ext}_{\boldsymbol{g}}(\boldsymbol{a}))$.

$k + 3\ell + 4t + 2 > n$ and this system can avoid the Bombar-Couveur attack. Unfortunately, this system can be attacked by the GOT attack [3].

Chapter 3

In this chapter, based on the idea of linearizing the decoding problem to decode Gabidulin codes in [8], we consider the problem of decoding the sum of Gabidulin codes. For a sum of Gabidulin codes $\mathcal{C} = \sum_{j=1}^{\ell} \mathcal{G}_{k_j}(\boldsymbol{g}_j)$ where $\boldsymbol{g}_j \xleftarrow{\$} \mathbb{F}_{q^m}^n$, it can be decoded with exponentially small failure probability in polynomial-time under the condition that $(\ell+1)t + \sum_{j=1}^{\ell} k_j < n$. In the section 3.3.2, we also consider a replacement of the sum of Gabidulin code in the Loidreau's cryptosystem [7] and we hope that this modified system can resist the Coggia and Couvreur attack. It means that we need to consider the distinguisher of the dual of the public code from the random one. In the future, we can find more applications for this decoding.

BIBLIOGRAPHY

[1] E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov, « Ideals over a Non-Commutative Ring and their Application in Cryptology », *in*: *Advances in Cryptology — EUROCRYPT '91*, ed. by D. W. Davies, Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, pp. 482–489, ISBN: 978-3-540-46416-7.

[2] C. Faure and P. Loidreau, « A New Public-Key Cryptosystem Based on the Problem of Reconstructing p-Polynomials », *in*: *Coding and Cryptography, International Workshop, WCC 2005, Bergen, Norway, March 14-18, 2005. Revised Selected Papers*, 2005, pp. 304–315.

[3] P. Gaborit, A. Otmani, and H. Talé-Kalachi, « Polynomial-time key recovery attack on the Faure-Loidreau scheme based on Gabidulin codes », *in*: *Des. Codes Cryptogr.* 86.7 (2018), pp. 1391–1403, DOI: 10.1007/s10623-017-0402-0, URL: https://doi.org/10.1007/s10623-017-0402-0.

[4] J. Renner, S. Puchinger, and A. Wachter-Zeh, English, *in*: *Designs, Codes and Cryptography* (2021), pp. 1279–1319, ISSN: 0925-1022, DOI: 10.1007/s10623-021-00861-z.

[5] J. Lavauzelle, P. Loidreau, and P. B. Duc, « RAMESSES, a Rank Metric Encryption Scheme with Short Keys », *in*: (2019), URL: http://arxiv.org/abs/1911.13119.

[6] D. Coggia and A. Couvreur, « On the security of a Loidreau's rank metric code based encryption scheme », *in*: *WCC 2019 - Workshop on Coding Theory and Cryptography*, Saint Jacut de la mer, France, Mar. 2019, URL: https://hal.archives-ouvertes.fr/hal-02064465.

[7] P. Loidreau, « A new rank metric codes based encryption scheme », *in*: *PQCrypto 2017*, ed. by T. Lange and T. Takagi, vol. 10346, Lecture Notes in Computer Science, Utrecht, Netherlands: Springer, June 2017, pp. 3–17, URL: https://hal.archives-ouvertes.fr/hal-01673462.

[8] P. Loidreau, « A Welch-Berlekamp Like Algorithm for Decoding Gabidulin Codes », *in*: *Coding and Cryptography, International Workshop, WCC 2005, Bergen, Norway, March 14-18, 2005. Revised Selected Papers*, ed. by Ø. Ytrehus, vol. 3969, Lecture Notes in Computer Science, Springer, 2005, pp. 36–45, DOI: 10.1007/11779360_4, URL: https://doi.org/10.1007/11779360%5C_4.

[9] W. Diffie and M. Hellman, « New directions in cryptography », *in*: *IEEE transactions on Information Theory* 22.*6* (1976), pp. 644–654.

[10] C. E. Shannon, « Communication theory of secrecy systems », *in*: *The Bell System Technical Journal* 28.*4* (1949), pp. 656–715, DOI: 10.1002/j.1538-7305.1949.tb00928.x.

[11] C. E. Shannon, « A mathematical theory of communication », *in*: *The Bell System Technical Journal* 27.*3* (1948), pp. 379–423, DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[12] R. W. Hamming, « Error detecting and error correcting codes », *in*: *The Bell System Technical Journal* 29.*2* (1950), pp. 147–160, DOI: 10.1002/j.1538-7305.1950.tb00463.x.

[13] K. S. Immink, « Reed-Solomon codes and the compact disc », *in*: Oct. 1999, ISBN: 978-0-7803-5391-6.

[14] V. Stylianakis and S. Toptchiyski, « A Reed-Solomon coding/decoding structure for an ADSL modem », *in*: *ICECS'99. Proceedings of ICECS '99. 6th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.99EX357)*, vol. 1, 1999, 473–476 vol.1, DOI: 10.1109/ICECS.1999.812325.

[15] R. J. McEliece, « A Public-Key System Based on Algebraic Coding Theory », *in*: DSN Progress Report 44, Jet Propulsion Lab, 1978, pp. 114–116.

[16] A. Vardy, « The Intractability of Computing the Minimum Distance of a Code », *in*: *IEEE Trans. Inform. Theory* 43.*6* (Nov. 1997), pp. 1757–1766.

[17] E. Prange, « The use of information sets in decoding cyclic codes », *in*: *IRE Transactions on Information Theory* 8.*5* (1962), pp. 5–9, DOI: 10.1109/TIT.1962.1057777, URL: http://dx.doi.org/10.1109/TIT.1962.1057777.

[18] P. Delsarte, « Bilinear Forms over a Finite Field, with Applications to Coding Theory », *in*: *J. Comb. Theory, Ser. A* 25.*3* (1978), pp. 226–241, URL: https://doi.org/10.1016/0097-3165(78)90015-8.

[19] E. Gabidulin, « Theory of codes with maximum rank distance », *in*: *Problems of Information Transmission* 21.*1* (1985), pp. 1–12.

[20] E. Gabidulin, « Theory of codes with maximum rank distance (translation) », *in*: *Problems of Information Transmission* 21 (Jan. 1985), pp. 1–12.

[21] E. Gabidulin, A. V. Paramonov, and O. V. Tretjakov, « Ideals over a non-commutative ring and their applications to cryptography », *in*: *Advances in Cryptology - EUROCRYPT'91*, LNCS 547, Brighton, Apr. 1991, pp. 482–489.

[22] R. Singleton, « Maximum distance q-nary codes », *in*: *IEEE Transactions on Information Theory* 10.*2* (1964), pp. 116–118, DOI: 10.1109/TIT.1964.1053661.

[23] F. Chabaud and J. Stern, « The Cryptographic Security of the Syndrome Decoding Problem for Rank Distance Codes », *in*: *Advances in Cryptology - ASIACRYPT 1996*, vol. 1163, LNCS, Kyongju, Korea: Springer, Nov. 1996, pp. 368–381.

[24] M. Bardet et al., « An Algebraic Attack on Rank Metric Code-Based Cryptosystems », *in*: *CoRR* abs/1910.00810 (2019), arXiv: 1910.00810, URL: http://arxiv.org/abs/1910.00810.

[25] C. A. Melchor et al., « ROLLO: Rank-Ouroboros, LAKE and LOCKER », *in*: *Submission to the NIST Post-Quantum Standardization project* (2017), URL: pqc-rollo.org.

[26] C. A. Melchor et al., « RQC: Rank Quasi-Cyclic », *in*: *Submission to the NIST Post-Quantum Standardization project* (2017), URL: pqc-rqc.org.

[27] P. Loidreau, « On Cellular codes and their cryptographic applications », *in*: *Proceedings of ACCT 2014, Fourteenth International Workshop on Algebraic and Combinatorial Coding Theory, Zvenigorod, 2014.* 2014.

[28] A. Ghatak, « Extending Coggia-Couvreur Attack on Loidreau's Rank-metric Cryptosystem », *in*: (2020), arXiv: 2007.07354 [cs.IT].

[29] M. Bombar and A. Couvreur, « Decoding Supercodes of Gabidulin Codes and Applications to Cryptanalysis », *in*: *Post-Quantum Cryptography*, ed. by J. H. Cheon and J.-P. Tillich, Cham: Springer International Publishing, 2021, pp. 3–22, ISBN: 978-3-030-81293-5.

[30] R. J. McEliece, « A Public-Key Cryptosystem Based On Algebraic Coding Theory », *in*: *Deep Space Network Progress Report* 44 (Jan. 1978), pp. 114–116.

[31] P. Gaborit et al., « Low Rank Parity Check codes and their application to cryptography », *in*: *The International Workshop on Coding and Cryptography (WCC 13)*, ed. by L. Budaghyan, T. Helleseth, and M. G. Parker, ISBN 978-82-308-2269-2, Bergen, Norway, Apr. 2013, 13 p. URL: https://hal.archives-ouvertes.fr/hal-00913719.

[32] R. Overbeck, « Structural Attacks for Public Key Cryptosystems based on Gabidulin Codes », *in*: *J. Cryptology* 21 (2008), pp. 280–301, DOI: 10.1007/s00145-007-9003-9.

[33] O. Ore, « On a special class of polynomials », *in*: *Transactions of the American Mathematical Society* 35.*3* (1933), pp. 559–584.

[34] D. Boucher and F. Ulmer, « Linear codes using skew polynomials with automorphisms and derivations », *in*: *Des. Codes Cryptogr.* 70.*3* (2014), pp. 405–431.

[35] P. Delsarte, « Bilinear Forms over a Finite Field, with Applications to Coding Theory », *in*: *J. Comb. Theory, Ser. A* 25.*3* (1978), pp. 226–241.

[36] E. M. Gabidulin, « Theory of codes with maximum rank distance », *in*: *Problemy Peredachi Informatsii* 21.*1* (1985), pp. 3–16.

[37] J.-C. Faugère et al., « Polynomial Systems Solving by Fast Linear Algebra », 27 pages, Apr. 2013, URL: https://hal.archives-ouvertes.fr/hal-00816724.

[38] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*, USA: W. H. Freeman Co., 1990, ISBN: 0716710455.

[39] D. A. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*, Berlin, Heidelberg: Springer-Verlag, 2007, ISBN: 0387356509.

[40] E. Thomae and C. Wolf, *Solving Systems of Multivariate Quadratic Equations over Finite Fields or: From Relinearization to MutantXL*, 2010.

[41] A. Bostan et al., *Algorithmes Efficaces en Calcul Formel*, French, Palaiseau: Frédéric Chyzak (auto-édit.), Sept. 2017, ISBN: 979-10-699-0947-2, URL: https://hal.archives-ouvertes.fr/AECF/.

[42] G. Lecerf, « On the complexity of the Lickteig-Roy subresultant algorithm », working paper or preprint, Jan. 2017, URL: https://hal.archives-ouvertes.fr/hal-01450869.

[43] D. Augot and M. Finiasz, « A Public Key Encryption Scheme Based on the Polynomial Reconstruction Problem », *in*: *Advanceds in Cryptology - EUROCRYPT 2003* 2656 (May 2003), ed. by E. Biham, pp. 229–240.

[44] P. Gaborit, A. Otmani, and H. T. Kalachi, « Polynomial-time key recovery attack on the Faure-Loidreau scheme based on Gabidulin codes », *in*: *Des. Codes Cryptogr.* 86.7 (2018), pp. 1391–1403, DOI: 10.1007/s10623-017-0402-0.

[45] A. Wachter-Zeh, S. Puchinger, and J. Renner, « Repairing the Faure–Loidreau Public-Key Cryptosystem », *in*: *IEEE Int. Symp. Inf. Theory (ISIT)*, 2018.

[46] N. Silberstein and T. Etzion, « Enumerative Coding for Grassmannian Space », *in*: *IEEE Trans. Information Theory* 57.1 (2011), pp. 365–374, DOI: 10.1109/TIT.2010.2090252, URL: https://doi.org/10.1109/TIT.2010.2090252.

[47] Y. Medvedeva, « Fast enumeration for Grassmannian space », *in*: *2012 XIII International Symposium on Problems of Redundancy in Information and Control Systems*, 2012, pp. 48–52, DOI: 10.1109/RED.2012.6338406.

[48] R. M. Roth, « Maximum-rank array codes and their application to crisscross error correction », *in*: *IEEE Trans. Information Theory* 37.2 (1991), pp. 328–336, DOI: 10.1109/18.75248, URL: https://doi.org/10.1109/18.75248.

[49] A. Paramonov and O. Tretjakov, « An analogue of Berlekamp-Massey algorithm for decoding codes in rank metric », *in*: *Proceedings of Moscow Inst. Physics and Technology (MIPT)* (1991).

[50] E. M. Gabidulin, « A Fast Matrix Decoding Algorithm for Rank-Error-Correcting Codes », *in*: *Algebraic Coding, First French-Soviet Workshop, Paris, France, July 22-24, 1991, Proceedings*, ed. by G. D. Cohen et al., vol. 573, Lecture Notes in Computer Science, Springer, 1991, pp. 126–133, DOI: 10.1007/BFb0034349, URL: https://doi.org/10.1007/BFb0034349.

[51] G. Richter and S. Plass, « Fast decoding of rank-codes with rank errors and column erasures », *in*: *Proceedings of the 2004 IEEE International Symposium on Information Theory, ISIT 2004, Chicago Downtown Marriott, Chicago, Illinois, USA, June 27 - July 2, 2004*, IEEE, 2004, p. 398, DOI: 10.1109/ISIT.2004.1365435, URL: https://doi.org/10.1109/ISIT.2004.1365435.

[52] A. Wachter-Zeh, V. B. Afanassiev, and V. Sidorenko, « Fast decoding of Gabidulin codes », *in*: *Des. Codes Cryptogr.* 66.*1-3* (2013), pp. 57–73, DOI: 10.1007/s10623-012-9659-5, URL: https://doi.org/10.1007/s10623-012-9659-5.

[53] G. Schmidt, V. R. Sidorenko, and M. Bossert, « Error and Erasure Correction of Interleaved Reed–Solomon Codes », *in*: *Coding and Cryptography*, ed. by Ø. Ytrehus, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 22–35.

[54] J. Renner et al., « Randomized Decoding of Gabidulin Codes Beyond the Unique Decoding Radius », *in*: *CoRR* abs/1911.13193 (2019), arXiv: 1911.13193, URL: http://arxiv.org/abs/1911.13193.

[55] N. Courtois, « Efficient Zero-Knowledge Authentication Based on a Linear Algebra Problem MinRank », *in*: *Advances in Cryptology - ASIACRYPT 2001, 7th International Conference on the Theory and Application of Cryptology and Information Security, Gold Coast, Australia, December 9-13, 2001, Proceedings*, ed. by C. Boyd, vol. 2248, Lecture Notes in Computer Science, Springer, 2001, pp. 402–421, DOI: 10.1007/3-540-45682-1_24.

[56] A. Kipnis and A. Shamir, « Cryptanalysis of the HFE Public Key Cryptosystem by Relinearization », *in*: *Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings*, ed. by M. J. Wiener, vol. 1666, Lecture Notes in Computer Science, Springer, 1999, pp. 19–30, DOI: 10.1007/3-540-48405-1_2.

[57] L. Goubin and N. Courtois, « Cryptanalysis of the TTM Cryptosystem », *in*: *Advances in Cryptology - ASIACRYPT 2000, 6th International Conference on the Theory and Application of Cryptology and Information Security, Kyoto, Japan, December 3-7, 2000, Proceedings*, ed. by T. Okamoto, vol. 1976, Lecture Notes in Computer Science, Springer, 2000, pp. 44–57, DOI: 10.1007/3-540-44448-3_4.

[58] D. Lazard, « Gröbner-Bases, Gaussian elimination and resolution of systems of algebraic equations », *in*: *Computer Algebra, EUROCAL '83, European Computer Algebra Conference, London, England, March 28-30, 1983, Proceedings*, ed. by J. A. van Hulzen, vol. 162, Lecture Notes in Computer Science, Springer, 1983, pp. 146–156, DOI: 10.1007/3-540-12868-9_99.

[59] J.-C. Faugère, F. Levy-dit-Vehel, and L. Perret, « Cryptanalysis of MinRank », *in*: *Advances in Cryptology - CRYPTO 2008, 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings*, ed. by D. A. Wagner, vol. 5157, Lecture Notes in Computer Science, Springer, 2008, pp. 280–296, DOI: 10.1007/978-3-540-85174-5_16.

[60] J. A. Verbel et al., « On the Complexity of "Superdetermined" Minrank Instances », *in*: *Post-Quantum Cryptography - 10th International Conference, PQCrypto 2019, Chongqing, China, May 8-10, 2019 Revised Selected Papers*, ed. by J. Ding and R. Steinwandt, vol. 11505, Lecture Notes in Computer Science, Springer, 2019, pp. 167–186, DOI: 10.1007/978-3-030-25510-7_10.

[61] D. Silva, F. R. Kschischang, and R. Kötter, « Communication over finite-field matrix channels », *in*: *IEEE Trans. Information Theory* 56.*3* (2010), pp. 1296–1305.

[62] P. Gaborit et al., « Low Rank Parity Check codes and their application to cryptography », *in*: *Proceedings of the Workshop on Coding and Cryptography WCC'2013*, Bergen, Norway, 2013, URL: www.selmer.uib.no/WCC2013/pdfs/Gaborit.pdf.

[63] P. Loidreau, « A new rank metric codes based encryption scheme », *in*: *Post-Quantum Cryptography 2017*, vol. 10346, LNCS, Springer, 2017, pp. 3–17.

[64] P. B. Duc and P. Loidreau, « On the decoding of the sum of Gabidulin codes », *in*: *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 166–171, DOI: 10.1109/ISIT45174.2021.9517869.

[65] N. Aragon et al., « A new algorithm for solving the rank syndrome decoding problem », *in*: *2018 IEEE International Symposium on Information Theory, ISIT 2018, Vail, CO, USA, June 17-22, 2018*, IEEE, 2018, pp. 2421–2425, DOI: 10.1109/ISIT.2018.8437464.

[66] P. Loidreau, « A Welch–Berlekamp Like Algorithm for Decoding Gabidulin Codes », *in*: *Coding and Cryptography*, ed. by Ø. Ytrehus, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 36–45.

[67] D. Augot, P. Loidreau, and G. Robert, « Generalized Gabidulin codes over fields of any characteristic », *in*: *Des. Codes Cryptogr.* 86.*8* (2018), pp. 1807–1848.

[68] R. Overbeck, « Public Key Cryptography based on Coding Theory », PhD thesis, Darmstadt: Technische Universität, June 2007, URL: http://tuprints.ulb.tu-darmstadt.de/823/.

[69] R. Lidl and H. Niederreiter, *Finite fields*, Second, vol. 20, Encyclopedia of Mathematics and its Applications, With a foreword by P. M. Cohn, Cambridge University Press, Cambridge, 1997, pp. xiv+755, ISBN: 0-521-39231-4.

[70] J. Renner, S. Puchinger, and A. Wachter-Zeh, « Interleaving Loidreau's Rank-Metric Cryptosystem », *in*: *CoRR* abs/1901.10413 (2019), arXiv: 1901.10413, URL: http://arxiv.org/abs/1901.10413.

[71] A. Kshevetskiy, « Security of GPT-like public-key cryptosystems based on linear rank codes », *in*: *2007 3rd International Workshop on Signal Design and Its Applications in Communications*, 2007, pp. 143–147, DOI: 10.1109/IWSDA.2007.4408344.

[72] A. Otmani, H. T. Kalachi, and S. Ndjeya, « Improved Cryptanalysis of Rank Metric Schemes Based on Gabidulin Codes », *in*: *CoRR* abs/1602.08549 (2016), arXiv: 1602.08549, URL: http://arxiv.org/abs/1602.08549.

[73] M. Bardet et al., « Improvements of Algebraic Attacks for Solving the Rank Decoding and MinRank Problems », *in*: *Advances in Cryptology – ASIACRYPT 2020*, ed. by S. Moriai and H. Wang, Cham: Springer International Publishing, 2020, pp. 507–536, ISBN: 978-3-030-64837-4.

[74] D. Augot and M. Finiasz, « A Public Key Encryption Scheme Based on the Polynomial Reconstruction Problem », *in*: *Advances in Cryptology - EUROCRYPT 2003*, vol. 2656, LNCS, Springer, 2003, pp. 229–240.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

Titre : Étude et conception de nouvelles primitives de chiffrement fondées sur les codes correcteurs d'erreurs en métrique rang

Mot clés : Métrique rang, Codes de Gabidulin, Cryptosystème de McEliece

Résumé : En 2005, Faure et Loidreau ont proposé un nouveau métrique rang cryptosystème inspiré du schéma métrique de Hamming d'Augot-Finiasz en 2003. En 2018, il a été attaqué par Gaborit, Otmani et Kalachi. Récemment, il y a eu quelques tentatives de réparation du schéma Faure-Loidreau, par exemple les travaux de Renner, Puchinger et Wachter-Zeh qui s'appelle LIGA. Dans cette thèse, on introduit également un nouveau cryptosystème appelé RAMESSES qui est une autre réparation du schéma de Faure-Loidreau.

Par ailleurs, on étude également l'attaque récente de Coggia et Couveur sur le cryptosystème de Loidreau (2017). Bien qu'ils ne proposent qu'une idée pour un cas particulier de la dimension du sous-espace secret, cette attaque peut être généralisée. Dans cette thèse, on propose une analyse de l'attaque Coggia-Couvreur sur le schéma de chiffrement à clé publique en métrique rang de Loidreau dans le cas général.

La dernière partie est une étude sur le décodage de la somme des codes de Gabidulin qui s'inspire du "Welch-Berlekamp Like Algorithm for Decoding Gabidulin Codes" de Loidreau en 2005. Ce travail est également une tentative de réparation du cryptosystème de Loidreau (2017) pour éviter l'attaque de Coggia-Couveur.

Title: Study and design of new encryption primitives based on rank metric error correcting codes

Keywords: Rank metric, Gabidulin codes, McEliece cryptosystem

Abstract: In 2005, Faure and Loidreau proposed a new rank-metric cryptosystem inspired from the Hamming metric scheme of Augot-Finiasz in 2003. In 2018, it was broken by the attack of Gaborit, Otmani and Kalachi. Recently, there are some attempts of repairing the Faure-Loidreau scheme, for example the work of Renner, Puchinger and Wachter–Zeh which is called LIGA. In this thesis, we also introduce a new cryptosystem so-called RAMESSES which is another repairing of Faure-Loidreau scheme.

Besides, we also study about the recent attack of Coggia and Couveur in the Loidreau's cryptosystem (2017). Although they only propose an idea for a special case of the dimension of secret subspace, this attack can be generalized. In this thesis, we propose an analysis of Coggia-Couvreur attack on Loidreau's rank-metric public-key encryption scheme in the general case.

The last part is a study about the decoding of the sum of Gabidulin codes which is inspired from the work of Loidreau in 2005 "Welch-Berlekamp Like Algorithm for Decoding Gabidulin Codes". This work is also an attempt to repair the Loidreau's cryptosystem (2017) to avoid the Coggia-Couveur's attack.