A New Rank Metric Codes based Identification Scheme

Peter Arnaud Kidoudou
Faculty of Science and Technology, Marien NGOUABI University

Regis F. Babindamana
Faculty of Science and Technology, Marien NGOUABI University

Benjamin Mampassi
Faculty of Science and Technology, Marien NGOUABI University

ABSTRACT

In this paper, we propose a version of the identification rank metric code-based schemes. This protocol is an improvement of rankbased Véron protocol.It provide smaller public and private key sizes.With the same complexity as the Véron scheme, we make the scheme more secure by hiding the structure of the code used.

General Terms

MSC Classification 2020: 68P30, 94B05, 94A60

Keywords

Codes in rank metric, Gabidulin codes, error-correcting codes, Véron Identification Scheme.

1. INTRODUCTION

Most of the cryptographic protocols used today are based on traditional number theory problems such as traditional number theory problems such as factoring a large number into the product of two numbers into a product of two prime integers, the calculation of the discrete logarithm in group. These protocols include RSA, DSA, ECDSA, ECC, etc.

However, with the arrival of quantum computers, the shor [16] shor shows that these protocols are vulnerable to polynomial-time attacks.

Cryptography based on code theory is one of the few supposedly secure alternatives in post-quantum cryptography.

The best-known cryptosystems in code theory are Mc Eliece [13] and Niederreiter [11].These cryptosystems use the Goppa code [11].

The main advantage of these two public-key cryptosystems is the speed of encryption and decryption. But theexcessive length of the size of the cl ́es makes them impracticable.

To solve the storage problem, the GPT cryptosystem based on rank-metric codes was proposed by Gabidulin and al (1991).

Identification is a cryptographic mechanism that verifies the identities of correspondents.

Several identification schemes based on the theory of error-correcting codes have been proposed. this is the case of stern [07] based on the decoding syndrome problem andVeron(1995) based on a search of a low weight problem [15].

Contribution

In this paper,we are inspired by Veron [07] identification schema, we propose a similar scheme using Gabidulin codes. We introduce the distortion matrix proposed by Loidreau [11] to guarantee the indistinguishability of the Gabidulin codes. We get the same performance as the Veron scheme. However, the key size is smaller and lead to faster identification algorithms.

Organization of the paper:

This paper is organized as follows: In section II we present the basic concepts of code-based cryptography. In section III we present the Veron scheme. In section IV we describe the new scheme with its properties. In section V we show a comparison between the key size of our scheme and that of Veron scheme and finally, we conclude.

2. BACKGROUND OF CODING THEORY

Let$F$be a finite field with q elements s is denoted by $\mathrm{F_q}$where q is a power of a prime number. For any subfield$\mathrm{U} \subseteq \mathrm{F}$ of a field Fsuch that $k \leq n$ with positive integers k and n. The vector U **s**pace spanned by $\beta_1, \ldots, \beta_n$ with$\beta_i \in F^n$is denoted by $\sum_{i=1}^{k} U\beta_1$. The set of matrices with m rowsand n columns and entries in$F$is denoted by$\mathrm{M_{m,n}}$.$\mathrm{GL_n(F)}$isdenoted group of invertiblematrices of size n over$\mathrm{F}$.

Definition 1

Let A be a matrix from$M_{m,n}(F)$he rank weight of A denoted by |A| is the rank of A.The rank distance between two matrices A and B from$M_{m,n}(F)$is defined as |A-B|.

The rank distance on $M_{m,n}(F)$has the properties of a metric. this rank distance is rather defined for vectors $x \in F_q^n$. Considering the field F_{q^m} as an F_q^nvector space and hence any vector $x \in F_{q^m}$as a matrix fromon $M_{m,n}(F)$bydecomposing each entry $x_i \in F_{q^m}$.The rank weight of x also denoted by |x| is then its rank viewed as a matrix of $M_{m,n}(F_q)$.

Definition 2

Letthe finite field extension $\mathrm{F_{q^m}/F_q}$The rank weightof a vector $\mathrm{x} = (\mathrm{x_1}, \dots, \mathrm{x_n}) \in \mathrm{F_{q^m}}$denoted by $|\mathrm{x}|$is the dimension of the $\mathrm{F_q}$vector spacegenerated by $(\mathrm{x_1}, \dots, \mathrm{x_n})$

$$|x| = dim \sum_{i=1}^{n} F_q \, x_i$$

The column rank over $F_q$of a matrix A matrix from$M_{m,n}(F)$is denoted by$|A|$.defined with the dimension of $\sum_{i=1}^{n} \mathrm{o} F_q \, A_i$ where $(A_1, \dots., A_n)$are the columns of A.

Proposition 1

Let $\mathrm{A} \in \mathrm{M_{k,n}}(\mathrm{F_{q^m}})$and set $\mathrm{s} = |\mathrm{A}|$ with$\mathrm{s} < n$.There exist then $\mathrm{A}^* \in \mathrm{M_{k,s}}(\mathrm{F_{q^m}})$ with $|\mathrm{A}^*| = \mathrm{s}$ and $\mathrm{B} \in \mathrm{GL_n}(\mathrm{F_q})$such that: $\mathrm{AB} = (\mathrm{A}^*|0)$.

For any $\mathrm{x} \in (\backslash \mathrm{F_{q^m}^n})$such that s=$|\mathrm{x}|$there exists then $\mathrm{B} \in \mathrm{GL_n}(\mathrm{F_q})$for witchxB=$(\mathrm{x}^*|0)$where $\mathrm{x}^* \in \mathrm{F_{q^m}^s}$ and $\mathrm{s} = |\mathrm{A}^*|$.

2.1 Gabidulin code

Definition 3

Let $1 \leq \mathrm{k} < n \leq m$ be integers,and $(\mathrm{g}_1, \dots, \mathrm{g}_\mathrm{n}) \in \mathrm{F}_\mathrm{q}^\mathrm{m}$ be linearly independent elements over F_q. An(n, k)Gabidulin code [8]) over $\mathrm{F}_{\mathrm{q}^\mathrm{m}}$ defined a points $(\mathrm{g}_1, \dots, \mathrm{g}_\mathrm{n})$is the set of code words, each of which is defined as $\left(\mathrm{p(g1)} \dots \mathrm{p(gn)}\right)$,for a distinct linearized polynomial p over $\mathrm{F}_{\mathrm{q}^\mathrm{m}}$ of degree less thanq^k.

Gab(n,k)=$\left\{\left(p(g1)\dots p(gn)\right): p(x) \in L_q^m, deg\left(p(x)\right) < q^k\right\}$

Let g$\in \mathrm{F}_\mathrm{q}^\mathrm{m}$such that |g| =n. The Gab(k,n) g) is the code of length n and dimension k generated by the matrix [1]

$$G = \begin{pmatrix} g_1 & . & . & . & g_n. \\ . & . & . & . & . \\ . & . & . & . & . \\ g_1^{[k-1]} & . & . & . & g_n^{[k-1]} \end{pmatrix}$$

Gabidulin codes possess a fastdecoding algorithm that can decode errors of weight t provided that $\mathrm{t} \leq 1/2$

Remark

Gabidulin codes are rare rank metric codes with a polynomial time decoding algorithm.

Example

let f: $\mathrm{x}^5 + \mathrm{x}^2 + 1$ n=5, k=4, q=2, t=1

$$F_{2^5} = \{0,1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \dots . \alpha^{30}\}$$

$$\alpha^5 \doteq \alpha^2 + 1\beta = \{1, \alpha, \alpha^2, \alpha^3, \alpha^4\}$$

$$G_{gab_{(4,5)}} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ 1 & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^2+1 \\ 1 & \alpha^3 & \alpha_4 & \alpha^2+1 & \alpha^3+\alpha \\ 1 & \alpha^4 & \alpha^2+1 & \alpha^3+\alpha & \alpha^4+\alpha \end{pmatrix}$$

2.2 Distinguishing Properties of Gabidulin Codes

We recallthat Gabidulin codes [09]can be easily distinguished from random linear codes.

Definition 4

For any integer $i \geq 0$.

$\theta_i: M_{k,n}\left(\mathrm{F}_{q^m}\right) \to M_{(i+1)k,n}\left(\mathrm{F}_{q^m}\right)$operator that maps any A from$M_{(i+1)k,n}\left(F_{q^m}\right)to\theta_i(M)$where definition:

$$\theta_i = \begin{pmatrix} A^{[0]} \\ .A^{[n]} \end{pmatrix}$$

For any code Gab generated by a matrix G we denote by θ_i(Gab)the code generated by θ_i(G).when one compares the dimension of the code spanned by θ_i for a randomly drawn matrix G and the dimension obtained when G generates a Gabidulin code.We note the importance of θ_i

Proposition 2

- let G$\in M_{k,n}(\mathrm{F}_{q^m})$be a generator matrix of the Gabidulin code Gab(n,k)
- Pick S$\in \mathrm{GL}_k\left(\mathrm{F}_{q^m}\right)$, X$\in M_{k,l}(\mathrm{F}_{q^m})$,P$\in \mathrm{GL}_{n+l}\left(\mathrm{F}_\mathrm{q}\right)$
- Compute$\mathrm{G}_\mathrm{pub} = \mathrm{S(G|X)P}$
- Return $pk = \mathrm{G}_{\mathrm{pub},\mathrm{t}}$and $\mathrm{sk} = (\mathrm{S}, \mathrm{P})$

Let $\mathrm{g} \in \mathrm{F}_{\mathrm{q}^\mathrm{m}}^\mathrm{n}$ with$|\mathrm{g}| = \mathrm{n}$, $\mathrm{n} \leq \mathrm{m}$.
For any integers k and i we have $\mathrm{k} \leq \mathrm{n}$; $\mathrm{i} \leq \mathrm{n} - \mathrm{k} - 1$
$\theta_\mathrm{i}\left(\mathrm{Gab(k, n)}\right)$ =Gab(k+i,n)

Proposition 3

Let *l*, k and n be positive integers with$l< n$; $k \leq k < n$.

We have G generator matrix of a Gabidulin code$\in M_{k,n}(\mathrm{F}_{q^m})$

and X be a randomly drawn$\in M_{k,l}(\mathrm{F}_{q^m})$Denote β as the code defined by the generator matrix (X|G).

$\mathrm{k} + \mathrm{i} \leq \dim \wedge_{\{\mathrm{i}\}} (\beta) \leq \mathrm{k} + \mathrm{i} + \mathrm{d}$ with $\mathrm{i} \geq 0$

where d=min$\{(\mathrm{i} + 1)\mathrm{k}, l\}$Note that by construction
$l \leq$ n and in Overbeck's attack,with i=n-k-1 so that we have both d =*l* with high probability.
$\dim \wedge_{\mathrm{n-k-1}} (\beta) = \mathrm{k} + (n - k - 1) + d = n + l - 1$

This implies that the dimension of $\dim \wedge_{\{\mathrm{i}\}} (\beta)^\perp$=l. Thisfact is then harnessed by to recover an equivalent Gabidulin code which enables to decrypt any ciphertext.

Proposition 4

Let public matrix $\mathrm{G}_\mathrm{pub} = \mathrm{S(G|X)P}$ with$\in M_{k,l}(\mathrm{F}_{q^m})$ P $\in \mathrm{GL}_{n+l}\left(\mathrm{F}_\mathrm{q}\right)$, S$\in \mathrm{GL}_k\left(\mathrm{F}_\mathrm{q}\right)$, and G generating a Gabidulin code Gab(k,n).If $\dim \wedge_{\{\mathrm{n-k-1}\}} (\beta)^\perp = \mathrm{l}$ hen it is possible to recover(with$\vartheta((n + l)^3)$field operations)matrices $\mathrm{X}^* \in M_{k,l}(\mathrm{F}_{q^m})$,$\mathrm{P}^* \in \mathrm{GL}_{n+l}\left(\mathrm{F}_\mathrm{q}\right)$ and G^*generating a Gabidulin code Gab(k,n). such that:$\mathrm{G}_\mathrm{pub} = \mathrm{S(G^*|X^*)P^*}$.

Overbeck's[14] attack uses crucially two important facts.The column scrambler matrix$\in \mathrm{F}_\mathrm{q}$and the codimension of$\dim \wedge_{\mathrm{n-k-1}} (\beta) = 1$.

Several works propose to resist to Overbeck's attack by taking special distortion matrix so that the second property is not true.

Remark

the introduction of the distortion matrix guarantees the distinguishability property of the codes.

2.3 The GPT cryptosystem

This is an example of the use of Gabidulin codes in an encryption mechanism.

The work in Gabidulin, Paramonov, and Tretjakov (1991) proposed a cryptosystem, named GPT in honor of its authors, based on McElieces but this time using the so-called Gabidulin codes.

The key generation algorithm of the general GPT cryptosystem takes as input the integers k, l, n and m with$\mathrm{k} < n \leq m$and$1 \ll n$and outputs the public key and private key pair.

The GPT cryptosystem is composed by a triple of probabilistic polynomial-time algorithms(KeyGen,Encrypt, Decrypt).

KeyGen$(\mathrm{n, m, k, l, q}) = (\mathrm{pk, sk})$

To encrypt a message x$\in (F_{q^m}^k)$

Encrypt(x,pk) =y

1. e ← $\$\mathrm{F}^\mathrm{n}{}_{q^m}$
2. return y=x$\mathrm{G}_\mathrm{pub} + \mathrm{e}$

Decrypt(y,sk) =x'

1. Compute yP^{-1}==mS(G|X) +eP^{-1}
2. Extract the last n components y' of yP^{-1}
3. Apply a fast decoding algorithm of Gab(n,k) to y' to obtainX^*
4. Return X^*S^{-1}

we have z'=mSG +e' such that e' is sub-vector of eP^{-1},hence $|e'|_q \le$ t.

The output$X^*$of the decoding algorithm for Gab(n,k) then satisfiesX^*=XS.

3. VERON IDENTIFICATION SCHEME

In this section, we present the identification scheme based on error-correcting codes. The security of this scheme is based on the decoding syndrome problem [10].

This scheme was introduced by Véron[04]uses a generating matrix G of a random binary linear code as a public key..

Data:G(n, k) generator matrix andhash

private key:n∈ F_{q^m} and e ∈ $F_{q^m}^n$

public key:x'=mG+e, r=wt(e)

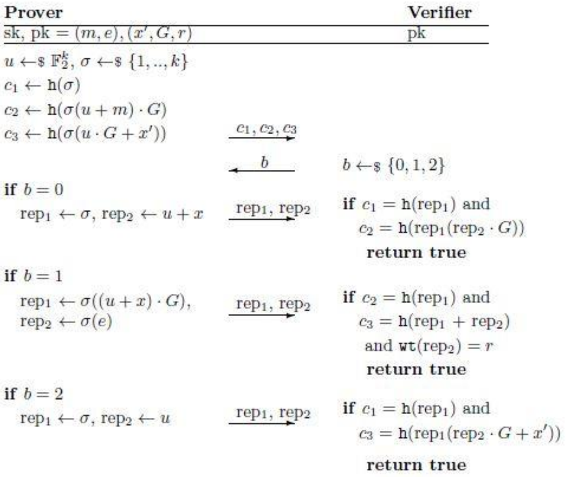

Prover
Verifier
sk, pk = $(m, e), (x', G, r)$
pk
$u \leftarrow_\$ \mathbb{F}_2^k$, $\sigma \leftarrow_\$ \{1, .., k\}$
$c_1 \leftarrow h(\sigma)$
$c_2 \leftarrow h(\sigma(u + m) \cdot G)$
$c_3 \leftarrow h(\sigma(u \cdot G + x'))$
c_1, c_2, c_3
b
$b \leftarrow_\$ \{0, 1, 2\}$
if $b = 0$
$rep_1 \leftarrow \sigma$, $rep_2 \leftarrow u + x$
rep_1, rep_2
if $c_1 = h(rep_1)$ and
$c_2 = h(rep_1(rep_2 \cdot G))$
return true
if $b = 1$
$rep_1 \leftarrow \sigma((u + x) \cdot G)$,
$rep_2 \leftarrow \sigma(e)$
rep_1, rep_2
if $c_2 = h(rep_1)$ and
$c_3 = h(rep_1 + rep_2)$
and $wt(rep_2) = r$
return true
if $b = 2$
$rep_1 \leftarrow \sigma$, $rep_2 \leftarrow u$
rep_1, rep_2
if $c_1 = h(rep_1)$ and
$c_3 = h(rep_1(rep_2 \cdot G + x'))$
return true

Fig 1: code-based Veron protocol [5]

4. New identification schemebased rank metric codes

4.1 Protocole

To use the advantages of the veron identification scheme [4][2].In this section we present a new identification scheme based rank metric codes.

Data:G(n, k) generator matrix andhash

h hash function, **S**(k, k) non-singular matrix**P**(n,n),non-singular matrix, **X**(n,k) distortion matrix. $\boldsymbol{g_{pub}}$= **S(G|X)P**

private key:*l*∈ F_{q^m} and *e*∈ $F_{q^m}^n$

public key:*x'= l* $\boldsymbol{g_{pub}}$ *+e, r=wt(e)*

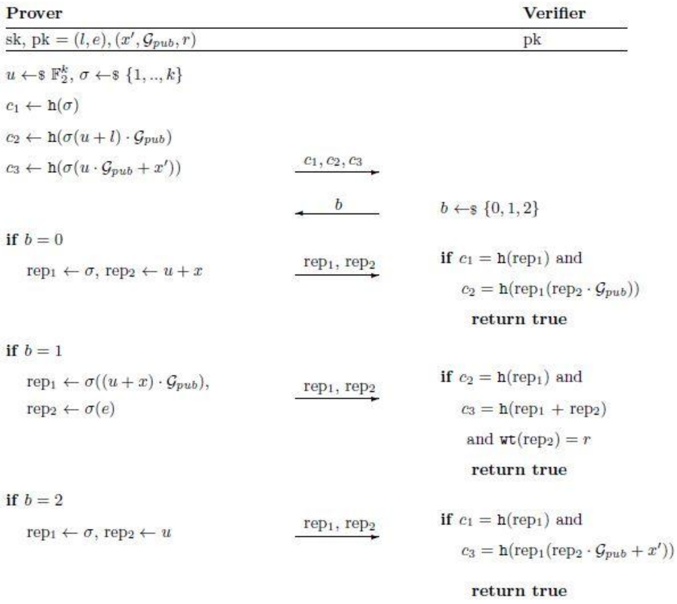

Prover
Verifier
sk, pk = $(l, e), (x', \mathcal{G}_{pub}, r)$
pk
$u \leftarrow_\$ \mathbb{F}_2^k$, $\sigma \leftarrow_\$ \{1, .., k\}$
$c_1 \leftarrow h(\sigma)$
$c_2 \leftarrow h(\sigma(u + l) \cdot \mathcal{G}_{pub})$
$c_3 \leftarrow h(\sigma(u \cdot \mathcal{G}_{pub} + x'))$
c_1, c_2, c_3
b
$b \leftarrow_\$ \{0, 1, 2\}$
if $b = 0$
$rep_1 \leftarrow \sigma$, $rep_2 \leftarrow u + x$
rep_1, rep_2
if $c_1 = h(rep_1)$ and
$c_2 = h(rep_1(rep_2 \cdot \mathcal{G}_{pub}))$
return true
if $b = 1$
$rep_1 \leftarrow \sigma((u + x) \cdot \mathcal{G}_{pub})$,
$rep_2 \leftarrow \sigma(e)$
rep_1, rep_2
if $c_2 = h(rep_1)$ and
$c_3 = h(rep_1 + rep_2)$
and $wt(rep_2) = r$
return true
if $b = 2$
$rep_1 \leftarrow \sigma$, $rep_2 \leftarrow u$
rep_1, rep_2
if $c_1 = h(rep_1)$ and
$c_3 = h(rep_1(rep_2 \cdot \mathcal{G}_{pub} + x'))$
return true

Fig 2: Veron protocol based on Gabidulin code

4.2.Security and Paramaters

The main strength of this scheme is that the private key is chosen randomly with entries in a secret λ-dimensional vector space. Thus, it prevents all types of attacks attempting to use algebraic properties to break the scheme.

In our zero-knowledge [7][12][5]. Veron of scheme (with several rounds) the probability of cheating is 2/3 for the security of 2^{80},we need 150 rounds. The number of rounds decreases the probability of impersonation according to our needs.

In general, to achieve a security levelwith a probability of impersonation, the number of roundsδ=$Log_q(\frac{1}{2^l})$ is determined.

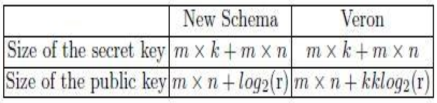
The ISO/IEC-9798-5 standard specifies two probabilities: $2^{-16}$and 2^{-32} or 28 and 56 rounds.

The use of an X-distortion matrix is crucial.This must prevent any structural attack consisting in recovering a decoder of the public code.

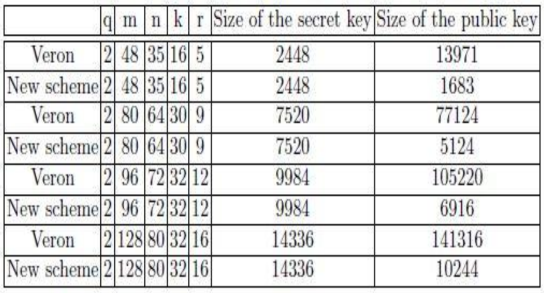
5. COMPARISON TABLE OF KEYS

Our thanks to the experts who have contributed towards development of the template.

Table1: comparison table of keys



	New Schema	Veron
Size of the secret key	$m \times k + m \times n$	$m \times k + m \times n$
Size of the public key	$m \times n + log_2(r)$	$m \times n + kklog_2(r)$



	q	m	n	k	r	Size of the secret key	Size of the public key
Veron	2	48	35	16	5	2448	13971
New scheme	2	48	35	16	5	2448	1683
Veron	2	80	64	30	9	7520	77124
New scheme	2	80	64	30	9	7520	5124
Veron	2	96	72	32	12	9984	105220
New scheme	2	96	72	32	12	9984	6916
Veron	2	128	80	32	16	14336	141316
New scheme	2	128	80	32	16	14336	10244

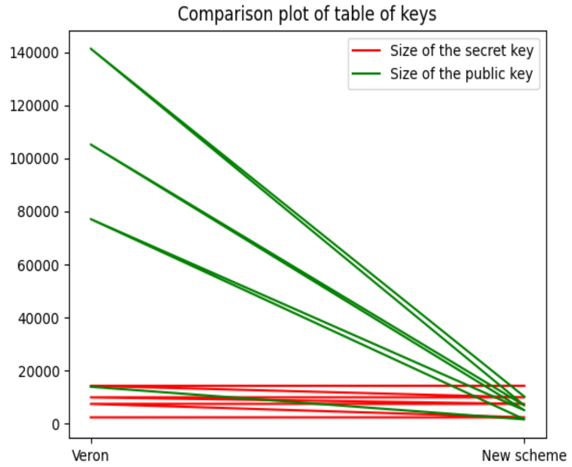


Fig 3: comparison plot of table of keys

6. CONCLUSION

In this new scheme, we use Gabidulin codes while hiding their structure. Gabidulin codes are among the few rank metric codes with a polynomial time decoding algorithm. The introduction of the distortion matrix guarantees the indistinguishability of Gabidulin codes.
This new identification schema veron-type elaborate decreases the probability of impersonation and generates small-size keys easy to store.

In the future, we are considering the protocol using another code family in rank metric.

7. REFERENCES

[1] Aragon, N: Cryptographie à base de codes correcteurs d'erreurs en métriquerang et application. Theses, Université de Limoges (Dec 2020), https://tel.archives-ouvertes.fr/tel-03115370.

[2] Aragon, N., Blazy, O., Gaborit, P., Hauteville, A., Z´emor, G.: Durandal: a rank metric based signature scheme. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 728–758. Springer (2019.

[3] Baldi, M., Battaglioni, M., Chiaraluce, F., Horlemann-Trautmann, A.L., Persichetti, E., Santini, P., Weger, V.: A new path to code-based signatures via identification schemes with restricted errors. arXiv preprint arXiv:2008.06403 (2020).

[4] Bellini, E., Caullery, F., Hasikos, A., Manzano, M., Mateu, V.: Code-based signature schemes from identification protocols in the rank metric. In: International Conference on Cryptology and Network Security. pp. 277–298. Springer (2018).

[5] Bellini, E., Gaborit, P., Hasikos, A., Mateu, V.: Enhancing code basedzeroknowledgeproofsusingrankmetric. IACRCryptol. ePrintArch. 2020, 1472(2020), https://eprint.iacr.org/2020/1472.

[6] Cayrel, P.L., Alaoui, S.: Dual construction of stern-based signature scheme 63,98–103 (03 2010).

[7] Cayrel, P.L., Véron, P., El YousfiAlaoui, S.M.: A zero-knowledge identification scheme based on the q-ary syndrome decoding problem. In: International Workshop on Selected Areas in Cryptography. pp. 171–186. Springer (2010).

[8] Gabidulin, E.M.: Rank-metric codes and applications. MoscowInst. Phys.Technol., StateUniv., Dolgoprudny, Russia.[Online].Available:http://iitp.ru/upload/content/839/Gabidulin. pdf

[9] Gabidulin, E.M.: Theory of codes with maximum rank distance. ProblemyPeredachiInformatsii 21(1), 3–16 (1985)

[10] Gaborit, P., Ruatta, O., Schrek, J.: On the complexity of the rank syndrome decoding problem. IEEE Transactions on Information Theory 62(2), 1006–1019 (2015)

[11] Loidreau, P.: Etude et optimisation de cryptosystèmes à clé publique fondés sur lathéorie des codes correcteurs. Ph.D. thesis (5 2001)

[12] Melchor, C.A., Gaborit, P., Schrek, J.: A new zero-knowledge code based identification scheme with reduced communication. 2011 IEEE Information Theory Workshop pp. 648–652 (2011)

[13] Moufek, H.: Les codes correcteurs pour la cryptographie. Ph.D. thesis, Faculté d eMathématiques (2017)

[14] Overbeck, R., Sendrier, N.: Code-based cryptography. In: Post-quantum cryptography, pp. 95–145. Springer (2009)

[15] Richmond, T.: Implantation sécurisée de protocoles cryptographiques basés sur les codes correcteurs d'erreurs. (secure implementation of cryptographic protocols based on error-correcting codes) (2016).

[16] Shor, P.W.: Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings 35th annual symposium on foundations of computer science.pp. 124–134. Ieee (1994)