# A Code-based Group Signature Scheme with Shorter Public Key Length

Hafsa Assidi, Edoukou Berenger Ayebie and El Mamoun Souidi
*Mohammed V University in Rabat, Faculty of Sciences, Laboratory of Mathematics, Computer Science and Applications, BP 1014 RP, Rabat, Morocco*

Keywords: Code-based Group Signature, Quasi-cyclic Moderate Density Parity-Check (QC-MDPC) Codes, McEliece Cryptosystem, Syndrome Decoding.

Abstract: Group signatures allow members to sign on behalf of a group while maintaining signer's identity anonymous. In this paper, we show that it is possible to reduce the public key length of the first provably secure group signature scheme from code-based assumptions without losing the security properties. More precisely, the public key can be 466 times shorter than the original scheme, typically for a group of 16 users when the public key length is 1.34 kilo-bytes, while the size is 625 kilo-bytes in the original scheme (Ezerman et al., 2015). Our technic consist in using a Quasi-cyclic Moderate Density Parity-Check McEliece variant for encrypting user identity and a random double circulant matrix for the Underlying Zero Knowledge Argument System.

## 1 INTRODUCTION

A fundamental cryptographic primitive which allows group users to anonymously sign documents on behalf of the whole group is called group signature, but in case of abuse, an administrator can revoke the anonymity of the signer. Group signature was introduced in (Chaum and van Heyst, 1991). Since then, many works have been proposed in this area [(Ateniese et al., 2000), (Boneh et al., 2004), (Camenisch et al., 2012), (Gordon et al., 2010)]. For anonymity and traceability properties of these schemes, group signatures are highly useful in various real-life scenarios such as controlled anonymous printing services, digital right management systems, e-bidding and e-voting schemes. The majority of group signature schemes are based on number theory assumptions [(Libert et al., 2012), (Camenisch and Stadler, 1997), (Boyen and Waters, 2006),(Boneh et al., 2004)], but number-theoretic based cryptography will not resist to the quantum computing. Recently, the research for post-quantum group signatures is quite active as shows these publications (Laguillaumie et al., 2013), (Langlois et al., 2014), (Ling et al., 2015), (Alamélou et al., ), (Gordon et al., 2010). The majority of these works are based on lattice assumptions, while in code-based cryptography we denote two group signature schemes: the first one is based on BSZ model (Bellare et al., 2005) and it is presented in (Alamélou et al., ). The second one is based on BMW model (Bellare et al., 2003) and is presented in (Ezerman et al., 2015), this scheme satisfies the CPA-anonymity and traceability requirements in the random oracle model. But the size of the signature and the public key makes this scheme impractical.

In this article, we propose a new provably secure group signature scheme based on code assumptions presented in (Ezerman et al., 2015). In this construction we replace the original McEliece cryptosystem used to encrypt the signer identity by the Quasi-cyclic Moderate Density Parity-Check (QC-MDPC) version of McEliece. The advantage of QC-MDPC codes is that for the same security level, the QC-MDPC version of McEliece has very short public key size than the original version. For example, for 80 bits security level, the QC-MDPC public key is 4801 bits unlike the original McEliece public key which is around 500000 bits. The second improvement consist in using a random double circulant matrix for the Underlying Zero Knowledge Argument System as proposed in (Gaborit and Girault, 2007), we suggest to replace a random matrix used by Stern (Stern, 1996) in the Zero-Knowledge identification scheme by a random double circulant matrix. This construction allows to have a very short public key, only 349 bits to obtain $2^{83}$ of security level in the Zero-Knowledge identification scheme. These two improvements reduce dramatically the public key size of the group signature scheme without impacting the security level. For instance, in the case of 16 users we achieve a public key size of 1.34 KB, while



Assidi, H., Ayebie, E. and Souidi, E.
A Code-based Group Signature Scheme with Shorter Public Key Length.
DOI: 10.5220/0005969204320439
In *Proceedings of the 13th International Joint Conference on e-Business and Telecommunications (ICETE 2016) - Volume 4: SECRYPT*, pages 432-439
ISBN: 978-989-758-196-0

the size is around 625 KB in the original scheme (Ezerman et al., 2015). Moreover, the structure of these matrices makes the implementation easier and efficient, since the product of a circulant matrix and a vector can be obtained by multiplying only the first row of the circulant matrix with shifts of the vector.

The organization of this paper is as follows: in Section 2 we recall the main tools of code-based cryptography, then we explain in Section 3 how to decrease the size of the public key by using the QC-MDPC version of McEliece cryptosystem and double circulant matrices. In Section 4, we consider the security of this new construction, while Section 5 compares the performances of the new provably secure group signature with the previous one. Section 6 is devoted for conclusion.

# 2 PRELIMINARIES

In this section, we first provide the notations that will be used all along this work, secondly we give background in code-based cryptography and finally we define the group signature.

## 2.1 Notations

Let $\lambda$ denotes the security parameter and $negl(\lambda)$ denotes a negligible function in $\lambda$. We denote by $a \xleftarrow{\$} S$ if $a$ is chosen uniformly at random from the finite set $S$. The symmetric group of all permutations of $k$ elements is denoted by $S_k$. We use bold capital letters, (e.g., $\mathbf{A}$), to denote matrices, and bold lowercase letters, (e.g., $\mathbf{x}$), to denote row vectors. We use $\mathbf{x}^\top$ to denote the transpose of $\mathbf{x}$ and $wt(\mathbf{x})$ to denote the Hamming weight of $\mathbf{x}$. We denote by $B(m,\omega)$ the set of all vectors $\mathbf{x} \in \mathbb{F}_2^m$ ($\mathbf{x}$ is a vector of $m$ bits) of hamming weight $\omega$: $wt(\mathbf{x}) = \omega$. Throughout the paper, we define a function $I2B$ which takes a non-negative integer $a$ as an input, and outputs the binary representation $(a_0, \cdots, a_{\ell-1}) \in \{0,1\}^\ell$ of $a$ such that $a = \sum_{i=0}^{\ell-1} a_i 2^{\ell-1-i}$ , and a function $B2I$ which takes as an input the binary representation $(a_0, \cdots, a_{\ell-1}) \in \{0,1\}^\ell$ of $a$, and outputs $a$. All logarithms are of base 2. By $1^\lambda$ we denote the string of $N$ ones. We denote by $\oplus$ the bitwise XOR operator.

## 2.2 Code-based Cryptography Background

Now we give some necessary notions in code-based cryptography for the well understanding of our work.

**Double Circulant Matrices.**
We say that $\mathbf{H}$ is a double circulant matrix if $\mathbf{H} = [\mathbf{I}_p|\mathbf{A}]$ where $\mathbf{I}_p$ is the identity matrix of size $p$ and $\mathbf{A}$ is a circulant matrix of length $p$, which means a $p \times p$ matrix generated from its first row $\mathbf{a} = (a_0, ..., a_{p-1})$

$$\mathbf{A} = \begin{pmatrix} a_0 & a_1 & . & . & . & a_{p-1} \\ a_{p-1} & a_0 & . & . & . & a_{p-2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_1 & a_2 & . & . & . & a_0 \end{pmatrix} \quad (1)$$

**$(n,q,w)$-QC-MDPC Code Construction.**
We are specially interested in $(n,q,w)$-QC-MDPC codes where $n = n_0 q$. This means that the parity-check matrix has the form $\mathbf{H} = [\mathbf{H}_0|\mathbf{H}_1|...|\mathbf{H}_{n_0-1}]$, where $\mathbf{H}_i$ is a $q \times q$ circulant block.
We define the first row of $\mathbf{H}$ picking a random vector of length $n = n_0 q$ and weight $w$. The other $q-1$ rows are obtained from the $q-1$ quasi-cyclic shifts of this first row. Each block $\mathbf{H}_i$ will have a row weight $w_i$, such that $w = \sum_{i=0}^{n_0-1} w_i$.

A generator matrix $\mathbf{G}$ in row reduced echelon form can be easily derived from the $\mathbf{H}_i$'s blocks. Assuming the rightmost block $\mathbf{H}_{n_0-1}$ is non-singular (which particularly implies $w_{n_0-1}$ is odd, otherwise the rows of $\mathbf{H}_{n_0-1}$ would sum up to 0), we construct a generator-matrix as follows.

$$\mathbf{G} = \left(\begin{array}{c|c} & (\mathbf{H}_{n_0-1}^{-1}.\mathbf{H}_0)^T \\ & (\mathbf{H}_{n_0-1}^{-1}.\mathbf{H}_1)^T \\ \mathbf{I}_{n-q} & . \\ & . \\ & . \\ & (\mathbf{H}_{n_0-1}^{-1}.\mathbf{H}_{n_0-2})^T \end{array}\right) \quad (2)$$

**Remark 2.1.** *Since we work in $\mathbb{F}_2$, we notice that generating a double circulant matrix $\mathbf{H} = [\mathbf{I}_p|\mathbf{A}]$ of length p require only p bits corresponding to the first row $\mathbf{a} = (a_0, a_1, ..., a_{p-1})$ of $\mathbf{A}$ (1). On the other hand, the generator matrix of a $(n,q,w)$ QC-MDPC code can be deduced only from a $n-q$ bits corresponding to the first row of each block $\mathbf{H}_i$.*

**Syndrome Decoding Problem.**
The Syndrome Decoding problem is a problem based on coding theory proved to be NP-complete in (Berlekamp et al., 1978) .

**Definition 2.1.** *The $SD(m,k,\omega)$ problem is formulated as follows: given a uniformly random matrix $\mathbf{H} \in \mathbb{F}_2^{k \times m}$ and a uniformly random syndrome $\mathbf{y} \in \mathbb{F}_2^k$, find a vector $\mathbf{s} \in B(m,\omega)$ such that $\mathbf{H} \cdot \mathbf{s}^\top = \mathbf{y}^\top$.*
*When m, k and $\omega$ are chosen based on $\lambda$, we say that*

*the $SD(m,k,\omega)$ problem is hard, if the success probability of any probabilistic polynomial time (PPT) algorithm in solving the problem is at most $negl(\lambda)$.*

**Syndrome Decoding Problem for Random Double Circulant Codes.**
We define a new problem, which adapts the syndrome decoding problem to the case of random double circulant codes:

**Definition 2.2.** *The $SD(2p,p,\omega)$ problem is formulated as follows:*
***Instance:*** *given a random double circulant matrix $\mathbf{H} \in \mathbb{F}_2^{p\times 2p}$ and a vector $\mathbf{y} \in \mathbb{F}_2^p$.*
***Question:*** *find a vector $\mathbf{s} \in B(2p,\omega)$ such that $\mathbf{H}\cdot\mathbf{s}^\top = \mathbf{y}^\top$.*

It is not known whether this problem is NP-complete, but the problem is probably as hard as syndrome decoding problem, and on practical point of view (see (Gaborit and Girault, 2007)) the practical security is almost the same for the best known attack that syndrome decoding for random matrices.

**The Randomized QC-MDPC McEliece Encryption Scheme**.
We derive from the QC-MDPC McEliece encryption scheme (Misoczki et al., 2013) a randomized variant as it is suggested in (Nojima et al., 2008) for the original version (McEliece, 1978) with Goppa codes, where a uniformly random vector is concatenated to the plaintext. We describe the scheme as follows:

- Let $\mathcal{S}(1^\lambda)$ be the setup algorithm, it selects parameters $n$, $q$, $t$ and $w$ which are chosen based on $\lambda$ for a binary $t$-error-correcting $(n,q,w)$-QC-MDPC code. Choose integers $q_1$, $q_2$ such that $n-q = q_1+q_2$. Set the plaintext space as $\mathbb{F}_2^{q_2}$.
- Let $\mathcal{K}(n,q,w,t)$ be the algorithm that generate the keys, it performs in two steps:
  1. Generate a parity-check matrix $\mathbf{H} \in \mathbb{F}_2^{q\times n}$ of a $t$-error-correcting $(n,q,w)$-QC-MDPC code.
  2. Generate its corresponding generator matrix $\mathbf{G} \in \mathbb{F}_2^{(n-q)\times n}$ in row reduced echelon form.

  Output encrypting key $pk_{ME} = \mathbf{G}$ and decrypting key $sk_{ME} = \mathbf{H}$.
- Let $\mathcal{E}(pk_{ME},\mathbf{m})$ be the encryption algorithm, it encrypts a message $\mathbf{m} \in \mathbb{F}_2^{q_2}$, sample $\mathbf{u} \xleftarrow{\$} \mathbb{F}_2^{q_1}$ and $\mathbf{e} \xleftarrow{\$} B(n,t)$, then output the ciphertext $\mathbf{c} = (\mathbf{u} \,\|\, \mathbf{m})\cdot\mathbf{G} \oplus \mathbf{e} \in \mathbb{F}_2^n$.
- Let $\mathcal{D}(sk_{ME},\mathbf{c})$ be the decryption algorithm and let $\mathcal{R}_\mathbf{H}$ be a $t$-error correcting Low-Density Parity Check (LDPC) decoding algorithm equipped with the knowledge of $\mathbf{H}$. To decrypt $\mathbf{c} \in \mathbb{F}_2^n$ :
  1. compute $\mathbf{y} = \mathcal{R}_\mathbf{H}(\mathbf{c})$
  2. get the plaintext $\mathbf{m}$ as follow: extract $(\mathbf{u}\|\mathbf{m})$ from the first $(n-q)$ positions of $\mathbf{y}$.

**Definition 2.3** (**Codeword Finding Problem**). *The $CwF(n,k,w)$ problem is as follows: given a matrix $\mathbf{H} \in \mathbb{F}_2^{k\times n}$ and an integer $w$, the problem is to find a codeword of weight at most $w$ in the code of generator-matrix $\mathbf{H}$.*

**Definition 2.4** (**The Decisional McEliece Problem**). *The $DMcE(n,k,t)$ problem is as follows: given a matrix $\mathbf{G} \in \mathbb{F}_2^{k\times n}$, distinguish whether $\mathbf{G}$ is a uniformly random matrix over $\mathbb{F}_2^{k\times n}$ or it is generated by algorithm $\mathcal{K}(n,q,w,t)$ described above.*
*When $n$, $q$, $t$ are chosen based on $\lambda$ and $k = n-q$, we say that the $DMcE(n,k,t)$ problem is hard, if the success probability of any probabilistic polynomial time (PPT) distinguisher is at most $\frac{1}{2} + negl(\lambda)$.*

**Definition 2.5** (**The Decisional Learning Parity with (Fixed-weight) Noise Problem**). *The $DLPN(k,n,B(n,t))$ problem is as follows: given a pair $(\mathbf{A},\mathbf{v}) \in \mathbb{F}_2^{k\times n} \times \mathbb{F}_2^n$ , distinguish whether $(\mathbf{A},\mathbf{v})$ is a uniformly random pair over $\mathbb{F}_2^{k\times n} \times \mathbb{F}_2^n$ or it is obtained by choosing $\mathbf{A} \xleftarrow{\$} \mathbb{F}_2^{k\times n}$ , $\mathbf{u} \xleftarrow{\$} \mathbb{F}_2^k$ , $\mathbf{e} \xleftarrow{\$} B(n,t)$ and outputting $(\mathbf{A}, \mathbf{u}\cdot\boldsymbol{A} \oplus \mathbf{e})$. When $n$, $k$, $t$ are chosen based on $\lambda$, we say that the $DLPN(k,n,B(n,t))$ problem is hard, if the success probability of any PPT distinguisher is at most $\frac{1}{2} + negl(\lambda)$.*

We recall the following proposition from (Misoczki et al., 2013).

**Proposition 2.1.** *Given the hardness of Codeword finding problem from MDPC and QC-MDPC codes:*

- *Breaking the MDPC variant of McEliece or Niederreiter is not easier than solving the syndrome decoding problem for a random code.*
- *Breaking the QC-MDPC variant of McEliece or Niederreiter is not easier than solving the syndrome decoding problem for a random quasi-cyclic linear code.*

*Proof.* The proof is given in (Misoczki et al., 2013). □

**Remark 2.2.** *The Proposition 2.1 shows that DMcE and DLPN problems from $(n,q,w)$-QC-MDPC code are both very hard as the random case of these problems.*

In the standard model assuming the hardness of the $DMcE(n,k,t)$ problem and the $DLPN(q_1,n,B(n,t))$ problem described in [(Döttling, 2014), (Nojima et al., 2008)] and assuming the Remark 2.2, the scheme described above is CPA-Secure.

**The Underlying Zero-knowledge Argument System.**
We recall the Underlying Zero Knowledge Argument System used by Ezerman and al in (Ezerman et al., 2015) which is based on Stern Zero Knowledge Protocol (Stern, 1996) using random double circulant matrix instead of random matrix .

Let $n, q, t, w, p, \omega, \ell$ be positive integers, and $N = 2^{\ell}$. In this protocol, the public input consists of matrices $\mathbf{G} \in \mathbb{F}_2^{(n-q)\times n}$ , $\mathbf{H} \in \mathbb{F}_2^{p\times 2p}$ ; $N$ syndromes $\mathbf{y}_0, \cdots, \mathbf{y}_{N-1} \in \mathbb{F}_2^p$ ; and a vector $\mathbf{c} \in \mathbb{F}_2^n$. It allows prover $\mathcal{P}$ to simultaneously convince verifier $\mathcal{V}$ in Zero Knowledge that $\mathcal{P}$ possesses a vector $\mathbf{s} \in B(2p, \omega)$ corresponding to certain syndrome $\mathbf{y}_j \in \{\mathbf{y}_0, \cdots, \mathbf{y}_{N-1}\}$ with hidden index $j$, and that $\mathbf{c}$ is a correct encryption of $I2B(j)$ via the randomized QC-MDPC McEliece encryption. Specifically, the secret witness of $\mathcal{P}$ is a tuple $(j, \mathbf{s}, \mathbf{u}, \mathbf{e}) \in [0, N-1] \times \mathbb{F}_2^{2p} \times \mathbb{F}_2^{n-q-\ell} \times \mathbb{F}_2^n$ such that:
$\mathbf{H} \cdot \mathbf{s}^{\top} = \mathbf{y}_j^{\top}$ and $(\mathbf{u} \parallel I2B(j)) \cdot \mathbf{G} \oplus \mathbf{e} = \mathbf{c}$ where $\mathbf{s} \in B(2p, \omega)$ and $\mathbf{e} \in B(n, t)$.

## 2.3 Group Signature Scheme

In this paper, we are interested in the static group signature case. For that, we will follow the definitions presented in the BMW model (Bellare et al., 2003).

**Definition 2.6.** *A group signature scheme GS is a tuple of four polynomial time algorithms $(KeyGen, Sign, Verify, Open)$. The description of these algorithms is as follows:*

1. *KeyGen : this randomized algorithm takes as input $(1^{\lambda}; 1^N)$, where $N \in \mathbb{N}^{\star}$ is the number of group users, and outputs (gpk, gmsk, gsk), where gpk is the group public key, gmsk is the group manager's secret key and $gsk = \{gsk[j]\}_{j\in[0,N-1]}$ with $gsk[j]$ being the secret key for the group user of index j.*
2. *Sign : it is a randomized algorithm, it takes as input a secret signing key gsk[ j] for some $j \in [0, N-1]$ and a message M and returns a group signature $\Sigma$ on M.*
3. *Verify : it is a deterministic algorithm that takes as input the group public key gpk, a message M, a signature $\Sigma$ on M, and returns either 1 (Accept) or 0 (Reject).*
4. *Open : this deterministic algorithm takes as input the group manager's secret key gmsk, a message M, a signature $\Sigma$ on M and returns an index $j \in [0, N-1]$ associated with a particular user, or False indicating failure.*

A group signature scheme as described in BMW model (Bellare et al., 2003) must verify three security requirements:

- *Correctness*: for all integers $\lambda$ and $N$, all $(gpk, gmsk, gsk)$ obtained from *KeyGen* algorithm with $(1^{\lambda}; 1^N)$ as input, all $j \in \{0, .., N-1\}$ and $M \in \{0,1\}^*$ $Verify(gpk, M, Sign(gpk, gsk[j], M)) = 1$ and
$$Open(gpk, gmsk, M, Sign(gpk, gsk[j], M)) = j$$
- *Traceability*: requires that all signatures, even those produced by a coalition of group users and the group manager, can be traced back to a member of the coalition.
- *Anonymity*: requires that, signatures generated by two distinct group users are computationally indistinguishable to an adversary knowing all the user secret keys.
  Formal definitions of these properties are described in the following definitions:

**Definition 2.7** (CPA anonymity)**.** *We say that a group signature $GS = (KeyGen, Sign, Verify, Open)$ is CPA-anonymous if for all polynomial $N(\cdot)$ and any PPT adversaries $\mathcal{A}$, the advantage of $\mathcal{A}$ in the following experiment is negligible in $\lambda$:*

1. *Run $(gpk, gmsk, gsk) \leftarrow KeyGen(1^{\lambda}, 1^N)$ and send $(gpk, gsk)$ to $\mathcal{A}$.*
2. *$\mathcal{A}$ outputs two identities $j_0$, $j_1 \in [0, N-1]$ with a message M. Choose a random bit b and give $Sign(gsk[j_b], M)$ to $\mathcal{A}$. Then, $\mathcal{A}$ outputs a bit $b_0$. $\mathcal{A}$ succeeds if $b_0 = b$, and the advantage of $\mathcal{A}$ is defined to $|Pr[\mathcal{A}\ succeeds] - \frac{1}{2}|$.*

**Definition 2.8** (Traceability)**.** *We say that a group signature $GS = (KeyGen, Sign, Verify, Open)$ is traceable if for all polynomial $N(\cdot)$ and any PPT adversaries $\mathcal{A}$, the advantage of $\mathcal{A}$ in the following experiment is negligible in $\lambda$:*

1. *Run $(gpk, gmsk, gsk) \leftarrow KeyGen(1^{\lambda}, 1^N)$ and send $(gpk, gmsk)$ to $\mathcal{A}$.*
2. *$\mathcal{A}$ may query the following oracles adaptively and in any order:*
   - *A $O^{Corrupt}$ oracle that on input $j \in [0, N-1]$, outputs $gsk[j]$.*
   - *A $O^{Sign}$ oracle that on input j, a message M, returns $Sign(gsk[j], M)$.*

   *Let CU be the set of identities queried to $O^{Corrupt}$.*
3. *Finally, $\mathcal{A}$ outputs a message $M^*$ and a signature $\Sigma^*$.*

*$\mathcal{A}$ succeeds if (1) $Verify(gpk, M^*, \Sigma^*) = 1$ and (2) $Sign(gsk[j], M^*)$ was never queried for $j \notin CU$, yet (3) $Open(gmsk, M^*, \Sigma^*) \notin CU$.*

# 3 OUR PROPOSED GROUP SIGNATURE SCHEME

In this section, we present our improved group signature scheme of (Ezerman et al., 2015). The improvement concern the *KeyGen* algorithm which is responsible to generate group public key $gpk$, group manager secret key $gmsk$ and group secret key $gsk$. We aim to reduce the group signature public key's length, for that, we proceed in two steps: first we use a generator matrix $\mathbf{G}$ of QC-MDPC code, as described in (Misoczki et al., 2013), to run the randomised QC-MDPC McEliece cryptosystem variant in order to have a reduced length of public key for the same security level achieved using Goppa codes. Second, we act at the choice of the matrix $\mathbf{H}$ used in the Zero Knowledge interactive protocol by using random double circulant matrix than a random one as presented in (Gaborit and Girault, 2007). The algorithms *Sign*, $Verify$ and *Open* remain unchanged.

Our scheme is as follows:

**KeyGen**$(1^{\lambda}, 1^{N})$ : takes as input a security parameter $\lambda$ and the number of group users $N = 2^{\ell}$. Outputs: the group public key $gpk$, group manager secret key $gmsk$ and the group secret keys $gsk$.
Then, we select:
- Parameters: $n = n(\lambda)$, $q = q(\lambda)$, $w = w(t)$ for the $(n,q,w)$-QC-MDPC code.
- Parameters: $p = p(\lambda)$, $\omega = \omega(\lambda)$ for the syndrome decoding problem.
- A collision-resistant hash function $\mathcal{H} : \{0,1\}^{*} \rightarrow \{1,2,3\}^{k}$ (where $k = \omega(\log(\lambda))$) used in the Fiat Schamir paradigm (Fiat and Shamir, 1987).

Let $\mathbf{G}$ be the public key of QC-MDPC variant of McEliece cryptosystem, $\mathbf{G} \in \mathbb{F}_2^{(n-q)\times n}$ is the generator matrix of a $(n,q,w)$ QC-MDPC code. As described in Remark 2.1, to generate the matrix $\mathbf{G}$ we need only $(n-q)$ bits. Unlike the original scheme where $\mathbf{G} \in \mathbb{F}_2^{(n-q)\times n}$ (of length $(n-q)\times n$) is a generator matrix of $(n,k,t)$ Goppa code.
Let $\mathbf{H} = [\mathbf{I}_p \mid \mathbf{A}] \in \mathbb{F}_2^{p\times 2p}$ be a random double circulant matrix, where $\mathbf{I}_p$ is the identity matrix and $\mathbf{A}$ is a $p \times p$ circulant matrix of length $p$ . Therefore, we need only $p$ bits (corresponding to the first row of $\mathbf{A}$) to generate $\mathbf{H}$ which is not the case in the original scheme where $\mathbf{H}$ is chosen uniformly random. The *KeyGen* algorithm proceeds as follows:

1. Run $\mathcal{K}(n,q,w,t)$ to obtain a key pair $(pk_{ME} = \mathbf{G} \in \mathbb{F}_2^{(n-q)\times n}; sk_{ME})$ for the QC-MDPC variant of McEliece encryption scheme.
2. Choose a random double circulant matrix $\mathbf{H} = [\mathbf{I}_p \mid \mathbf{A}]$.
3. For each $j \in [0, N-1]$, pick $\mathbf{s}_j \xleftarrow{\$} B(2p,\omega)$ , and let $\mathbf{y}_j \in \mathbb{F}_2^p$ be its syndrome, i.e., $\mathbf{y}_j^{\top} = \mathbf{H}\cdot\mathbf{s}_j^{\top}$.
4. The keys are as follows: $gpk = (\mathbf{G}, \mathbf{H}, \mathbf{y}_0, \cdots, \mathbf{y}_{N-1})$, $gmsk = sk_{ME}$, $gsk = (\mathbf{s}_0, \cdots, \mathbf{s}_{N-1})$.

**Sign**$(gsk[j], M)$ : the user $j$ sign a message $M \in \{0,1\}^{*}$ under $gpk$ using his secret key $s = gsk[j]$ by following these steps:

1. Encrypt the binary representation of $j$, i.e., vector $I2B(j) \in \mathbb{F}_2^{\ell}$ using the QC-MDPC randomised McEliece public key $\mathbf{G}$. The cipher text is
$$\mathbf{c} = (\mathbf{u} \parallel I2B(j))\cdot\mathbf{G} \oplus \mathbf{e} \in \mathbb{F}_2^{n}$$
where $\mathbf{u} \xleftarrow{\$} \mathbb{F}_2^{n-q-\ell}$ and $\mathbf{e} \xleftarrow{\$} B(n,t)$
2. Generate the signature $\Pi$ by running the Zero Knowledge variant of Stern in order to prove the possession of $\mathbf{s} \xleftarrow{\$} B(2p,\omega)$ corresponding to a certain syndrome $\mathbf{y}_j \in \{\mathbf{y}_0, \cdots, \mathbf{y}_{N-1}\}$ with hidden index $j$, and that $\mathbf{c}$ is a correct QC-MDPC McEliece encryption of $I2B(j)$. This is done by employing the Underlying Interactive Zero Knowledge Protocol with public input $(\mathbf{G}, \mathbf{H}, \mathbf{y}_0, \cdots, \mathbf{y}_{N-1}, \mathbf{c})$ and prover's witness $(j, \mathbf{s}, \mathbf{u}, \mathbf{e})$ that satisfies:
$$\begin{cases} \mathbf{H}\cdot\mathbf{s}^{\top} = \mathbf{y}_j \wedge \mathbf{s} \in B(2p,\omega) \\ (\mathbf{u} \parallel I2B(j))\cdot\mathbf{G} \oplus \mathbf{e} = \mathbf{c} \wedge \mathbf{e} \in B(n,t) \end{cases} \quad (3)$$
3. The protocol is repeated $k = \omega(log(\lambda))$ times to achieve negligible soundness error, and then apply the Fiat-schamir heuristic (Fiat and Shamir, 1987) to transform it in a signature.
Let $\Pi$ be a non-interactive zero-knowledge arguments of knowledge
$\Pi = (CMT^{(1)}, \cdots, CMT^{(k)}; (Ch^{(1)}, \cdots, Ch^{(k)}); RSP^{(1)}, \cdots, RSP^{(k)})$ where $CMT^{i}$ and $RSP^{i}$ are respectively the commitment and the response for the $i^{th}$ iteration in the interactive protocol and $(Ch^{(1)} \cdots, Ch^{(k)}) = \mathcal{H}(M; CMT^{(1)}, \cdots, CMT^{(k)}; gpk, \mathbf{c}) \in \{1,2,3\}^{k}$.
4. The signature is $\Sigma = (\mathbf{c}, \Pi)$.

**Verify**$(gpk, M, \Sigma)$ : since $\Sigma = (\mathbf{c}, \Pi)$ and $\Pi = (CMT^{(1)}, \cdots, CMT^{(k)}; (Ch^{(1)}, \cdots, Ch^{(k)}); RSP^{(1)}, \cdots, RSP^{(k)})$, the verification occur in two steps:

1. If
$$(Ch^{(1)}, \cdots, Ch^{(k)}) \neq \mathcal{H}(M; CMT^{(1)}, \cdots, CMT^{(k)}; gpk, \mathbf{c}),$$ *then return* 0.
2. For $i = 1$ to $k$, run the verification step of the interactive protocol in Section 2 with public input $(\mathbf{G}, \mathbf{H}, \mathbf{y}_0, \cdots, \mathbf{y}_{N-1}, \mathbf{c})$ to check the validity of

$RSP^{(i)}$ with respect to $CMT^{(i)}$ and $Ch^{(i)}$. If any of the verification conditions does not hold, then return 0 else return 1.

**Open**$(gmsk, M, \Sigma)$ : Parse $\Sigma$ as $(\mathbf{c}, \Pi)$ and run $\mathcal{D}(gmsk, \mathbf{c})$ to decrypt $\mathbf{c}$. If decryption fails, then return False. If decryption outputs $\mathbf{g} \in \mathbb{F}_2^{\ell}$ then return $j = B2I(\mathbf{g}) \in [0, N-1]$.

**Theorem 3.1.** *The group signature scheme is correct, CPA anonymous, traceable and the public key has size of* $((n-q)+(N+1)p)$ *bits.*

A sketch of proof of this theorem is in Section 4.

# 4 SECURITY ANALYSIS

Our code-based group signature scheme satisfies all the security properties listed in Subsection 2.3.

It is clear that the size of the public key is $((n-q)+(N+1)p)$ bits, since $\mathbf{G}$ and $\mathbf{H}$ have size of $(n-q)$ bits and $p$ bits respectively and for all $j \in [0, N-1]$ $\mathbf{y}_j$ is of length $p$ bits.

- **Correctness**
  Let $j \in [0, N-1]$ be a honest user. Consequently, there exists a tuple $(j, \mathbf{s}, \mathbf{u}, \mathbf{e})$ that satisfy Equation 3. $\Pi$ is a valid signature thanks to the perfect completeness of the underlying Zero Knowledge Protocol, then the algorithm Verify$(gpk, M, \Sigma)$ always outputs 1 for all M $\in \{0,1\}^*$. On the other hand, let $\mathbf{c}$ be a ciphertext of the form $\mathbf{c} = (\mathbf{u} \parallel I2B(j)) \cdot \mathbf{G} \oplus \mathbf{e}$. If we run the algorithm $\mathcal{D}(gmsk, \mathbf{c})$, the output will be $I2B(j)$ by the correctness of the randomised QC-MDPC variant of McEliece. Finally, $Verify$ and $Open$ are correct which implies the correctness of the signature scheme.
- **Anonymity**
  Since we use a randomized version of QC-MDPC McEliece and given the hardness of $DMcE(n,k,t)$ problem and the $DLPN(q_1, n, B(n,t))$ probleme in the QC-MDPC case (Remark 2.2), the CPA anonymity proof will be the same as (Ezerman et al., 2015).
- **Traceability**
  We consider $\mathcal{A}$ a PPT adversary against the traceability of the scheme with success probability equal to $\varepsilon$. We construct a PPT adversary $\mathcal{F}$ that is able to solve a $SD(2p, p, \omega)$ problem with success probability polynomially related to $\varepsilon$. Algorithm $\mathcal{F}$ receives a challenge $SD(2p, p.\omega)$ instance $(\widehat{\mathbf{H}}, \widehat{\mathbf{y}}) \in \mathbb{F}_2^{p \times 2p} \times \mathbb{F}_2^p$ with $\widehat{\mathbf{H}}$ a random double circulant matrix.

  The goal of $\mathcal{F}$ is to find a vector $s \in B(2p, \omega)$ such that $\widehat{\mathbf{H}}\mathbf{s}^T = \mathbf{y}^T$. It then proceeds as follows:

  1. Pick a guess $j^* \in [0, N-1]$ and set $\mathbf{y}_{j^*} = \widehat{\mathbf{y}}$.
  2. Set $\mathbf{H} = \widehat{\mathbf{H}}$. For each $j \in [0, N-1]$ such that $j \neq j^*$, sample $\mathbf{s}_j \xleftarrow{\$} B(2p, \omega)$ and set $\mathbf{y}_j \in \mathbb{F}_2^p$ be its syndrome, i.e., $\mathbf{y}_j^T = \mathbf{H}\mathbf{s}_j^T$ .
  3. Run $\mathcal{K}(n, q, w, t)$ to obtain a key pair $(pk_{ME} = \mathbf{G} \in \mathbb{F}_2^{(n-q) \times n}, sk_{ME})$.
  4. Sent $gpk = (\mathbf{G}, \mathbf{H}, \mathbf{y}_0, ..., \mathbf{y}_{N-1})$, $gmsk = sk_{ME}$ to $\mathcal{A}$.

  The forger $\mathcal{F}$ then initializes a set $CU = \varnothing$ and handles the queries from $\mathcal{A}$ as follows:

  – Queries to the random oracle $H$ are handled by consistently returning uniformly random values in $\{1,2,3\}^k$
  – $O^{Corrupt(j)}$, for any $j \in [0, N-1]$ : if $j \neq j^*$, then $\mathcal{F}$ sets $CU = CU \cup j$ and gives $s_j$ to $\mathcal{A}$; If $j = j^*$, then $\mathcal{F}$ aborts.
  – $O^{Sign(j,M)}$, for any $j \in [0, N-1]$ and any message $M$ : if $j \neq j^*$, then $\mathcal{F}$ honestly computes a signature since it has $s_j$; If $j = j^*$, then $\mathcal{F}$ returns a simulated signature $\Sigma^*$.

  The adversary $\mathcal{A}$ give a forged signature $\Sigma^*$ on a message $M^*$. The requirements of the traceability experiment implies that:
  $Verify(gpk, M^*, \Sigma^*) = 1$ and
  $Open(gmsk, M^*, \Sigma^*) = j^*$.

  By applying the same technics used in the traceability proof (Ezerman et al., 2015), $\mathcal{F}$ is able to solve the $SD(2p, p, \omega)$ with non negligible probability which contradicted the hardness of $SD(2p, p, \omega)$ problem.

# 5 RESULTS

In Table 1, we compare our results with the original scheme one (Ezerman et al., 2015) for different size of group users and for 80 bits security level.

We use a $(n, q, w, t)$ QC-MDPC code of parameters $n = 9602$, $q = 4801$, $w = 90$, $t = 84$ the size of the matrix $\mathbf{G}$ is $n - q = 4801$ bits (parameters for 80 bits security). It is shown in (Misoczki et al., 2013) that for this parameters the QC-MDPC McEliece cryptosystem is secure against the best attacks (Key distinguishing attack, Key recovery attack and Decoding attack).

We also use a double circulant matrix $\mathbf{H}$ of size $p = 350$ bits to achieve the same security level (80 bits) as it was detailed in (Gaborit and Girault, 2007).

Table 1: Public keys in the original scheme case and in our new scheme case.

| N | old key size | Our key size | factor reduction |
|---|---|---|---|
| $N=2^4$ | 625 KB | 1.34 KB | 466 |
| $N=2^8$ | 642 KB | 11.84 KB | 54 |
| $N=2^{12}$ | 906 KB | 179.84 KB | 5 |
| $N=2^{16}$ | 5.13 MB | 2.86 MB | 1.79 |
| $N=2^{20}$ | 72.8 MB | 45.87 MB | 1.58 |
| $N=2^{24}$ | 1.16 GB | 0.73 GB | 1.58 |
| $N=2^{28}$ | 18.45 GB | 11.74 GB | 1.57 |

We notice that our results, for the public key size, are better than ones obtained in the original scheme especially for groups with small size. For instance, when we have a group of $N = 2^4$ the size of our public key is 466 times shorter than the previous scheme. However when $N \geq 2^{16}$, the reduction factor tends to 1.57.

# 6 CONCLUSION

In this paper, we have proposed a provably secure code based group signature scheme with reduced public key length, the public keys can be 466 times shorter than the original scheme, typically for a group of 16 users when the public key is 1.34 kB, while the size is 625 kB in the original scheme. The proposed scheme also satisfies the correctness, CPA anonimity and traceability properties which are the security requirements for a static group signature scheme. In a future work, we will try to reduce the signature size. We can also construct a scheme achieving CCA-anonymity and supporting revocation mechanism.

# REFERENCES


Alamélou, Q., Blazy, O., Cauchie, S., and Gaborit, P. A code-based group signature scheme.

Ateniese, G., Camenisch, J., Joye, M., and Tsudik, G. (2000). A practical and provably secure coalition-resistant group signature scheme. In *CRYPTO 2000, Santa Barbara, California, USA, August 20-24, 2000, Proceedings*, volume 1880 of *LNCS*, pages 255–270. Springer.

Bellare, M., Micciancio, D., and Warinschi, B. (2003). Foundations of group signatures: Formal definitions, simplified requirements, and a construction based on general assumptions. In *EUROCRYPT 2003, Warsaw, Poland, May 4-8, 2003, Proceedings*, volume 2656 of *LNCS*, pages 614–629. Springer.

Bellare, M., Shi, H., and Zhang, C. (2005). Foundations of group signatures: The case of dynamic groups. In *CT-RSA 2005, San Francisco, CA, USA, February 14-18, 2005, Proceedings*, volume 3376 of *LNCS*, pages 136–153. Springer.

Berlekamp, E. R., McEliece, R. J., and Van Tilborg, H. C. (1978). On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24(3):384–386.

Boneh, D., Boyen, X., and Shacham, H. (2004). Short group signatures. In *CRYPTO 2004, California, USA, August 15-19, 2004, Proceedings*, volume 3152 of *LNCS*, pages 41–55. Springer.

Boyen, X. and Waters, B. (2006). Compact group signatures without random oracles. In *EUROCRYPT 2006, St. Petersburg, Russia, May 28 - June 1, 2006. Proceedings*, volume 4004 of *LNCS*, pages 427–444. Springer.

Camenisch, J., Neven, G., and Rückert, M. (2012). Fully anonymous attribute tokens from lattices. In *SCN 2012, Amalfi, Italy, September 5-7, 2012. Proceedings*, volume 7485 of *LNCS*, pages 57–75. Springer.

Camenisch, J. and Stadler, M. (1997). In *CRYPTO 97, California, USA August 17–21, 1997 Proceedings*, volume 1294 of *LNCS*, pages 410–424. Springer.

Chaum, D. and van Heyst, E. (1991). Group signatures. In *EUROCRYPT '91, Brighton, UK, April 8-11, 1991, Proceedings*, volume 547 of *LNCS*, pages 257–265. Springer.

Döttling, N. M. (2014). *Cryptography based on the Hardness of Decoding*. PhD thesis, Karlsruhe Institute of Technology.

Ezerman, M. F., Lee, H. T., Ling, S., Nguyen, K., and Wang, H. (2015). A provably secure group signature scheme from code-based assumptions. In *ASIACRYPT 2015,Auckland, New Zealand, November 29 – December 3, 2015, Proceedings, Part I*, volume 9452 of *LNCS*, pages 260–285. Springer.

Fiat, A. and Shamir, A. (1987). How to prove yourself: Practical solutions to identification and signature problems. In *CRYPTO '86, Santa Barbara, California, USA, 1986, Proceedings*, volume 263 of *LNCS*, pages 186–194. Springer.

Gaborit, P. and Girault, M. (2007). Lightweight code-based identification and signature. In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 191–195. IEEE.

Gordon, S. D., Katz, J., and Vaikuntanathan, V. (2010). A group signature scheme from lattice assumptions. In *ASIACRYPT 2010, Singapore, December 5-9, 2010. Proceedings*, volume 6477 of *LNCS*, pages 395–412. Springer.

Laguillaumie, F., Langlois, A., Libert, B., and Stehlé, D. (2013). Lattice-based group signatures with logarithmic signature size. In *ASIACRYPT 2013, Bengaluru, India, December 1-5, 2013, Proceedings, Part II*, volume 8270 of *LNCS*, pages 41–61. Springer.

Langlois, A., Ling, S., Nguyen, K., and Wang, H. (2014). Lattice-based group signature scheme with verifier-local revocation. In *PKC 2014, Buenos Aires, Ar-*

*gentina, March 26-28, 2014. Proceedings*, volume 8383 of *LNCS*, pages 345–361. Springer.

Libert, B., Peters, T., and Yung, M. (2012). Scalable group signatures with revocation. In *EUROCRYPT 2012, Cambridge, UK, April 15-19, 2012. Proceedings*, volume 7237 of *LNCS*, pages 609–627. Springer.

Ling, S., Nguyen, K., and Wang, H. (2015). PKC 2015, gaithersburg, md, usa, march 30 – april 1, 2015, proceedings. volume 9020 of *LNCS*, pages 427–449. Springer.

McEliece, R. (1978). A public-key cryptosystem based on algebraic. *Coding Thv*, 4244:114–116.

Misoczki, R., Tillich, J.-P., Sendrier, N., and Barreto, P. S. (2013). Mdpc-mceliece: New mceliece variants from moderate density parity-check codes. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pages 2069–2073. IEEE.

Nojima, R., Imai, H., Kobara, K., and Morozov, K. (2008). Semantic security for the mceliece cryptosystem without random oracles. *Designs, Codes and Cryptography*, 49(1):289–305.

Stern, J. (1996). A new paradigm for public key identification. *Information Theory, IEEE Transactions on*, 42(6):1757–1768.