High Weight Code-based Signature Scheme from QC-LDPC Codes

Chik How Tan[1][0000−0001−7550−3890] and Theo Fanuela Prabowo[1][0000−0002−4225−9112]

Temasek Laboratories, National University of Singapore, Singapore, Singapore
{tsltch, tsltfp}@nus.edu.sg

Abstract. We propose a new Hamming metric code-based signature scheme (called HWQCS) based on quasi-cyclic low density parity-check (QC-LDPC) codes. We propose the use of high error on QC-LDPC codes for constructing this signature and analyse its complexity. We show that HWQCS signature scheme achieves EUF-CMA security in the classical random oracle model, assuming the hardness of the syndrome decoding problem and the codeword finding problem for QC-LDPC codes. Furthermore, we also give a detailed security analysis of the HWQCS signature scheme. Based on the complexities of solving the underlying problems, the public key size and signature size of the HWQCS signature scheme are 1568 bytes and 4759 bytes respectively at 128-bit security level.

Keywords: code-based cryptography · signature · QC-LDPC codes

1 Introduction

Code-based cryptography is based on the problem of decoding random linear codes, which is referred to as the syndrome decoding problem and is known to be NP-hard [11]. The most common code-based cryptosystems are the McEliece cryptosystem [30] and the Niederreiter cryptosystem [33], which are equivalent in terms of their security. Solving the NP-hard syndrome decoding problem is believed to be hard even for quantum computers. Over the years, a number of code-based cryptographic schemes have been proposed. These include some promising key encapsulation mechanisms called BIKE [4], Classic McEliece[12] and HQC [1], which become fourth-round candidates in the NIST call for post-quantum cryptography standardization.

Unlike encryption and key encapsulation mechanisms, the construction of code-based digital signature schemes seems to be more challenging. This is indicated by the absence of code-based signature scheme in the second round onwards of the NIST PQC standardization. The most common techniques to construct signatures are based on two generic frameworks, which are, hash-and-sign constructions and Fiat-Shamir framework [23] constructions. The hash-and-sign construction requires some trapdoor functions, such as CFS [17] and Wave [19]. On the other hand, Fiat-Shamir framework construction does not necessarily use trapdoor functions in general, such as Stern [41], CVA [15], MPT [31], CVE [8], cRVDC [9], etc. However, most of them are inefficient or have large key or signature sizes. Furthermore, some of the proposed code-based signatures were even found to be insecure. For example, the KKS [24], RZW [37], CVE [8], SHMWW [39] and MPT [31] are shown to be insecure in [34], [18], [25], [5] and [35] respectively.

Recently, there is a new technique to construct signature schemes, which is called MPC (multiparty computation) in the head paradigm. This approach combines secret key sharing scheme and identification scheme in the multi-party computations setting, for example, CCJ signature [14], FJR signature [22], etc. The purpose of this construction is to reduce the signature size. But most of the signature size is still around eight thousand bytes. Therefore, it is still a challenge to construct signature schemes with practical signature size and public key size.

In this paper, we proposed a new signature scheme (called HWQCS) based on quasi-cyclic low density parity-check (QC-LDPC) codes. The proposed signature scheme is based on the Fiat-Shamir transformation and introduces high weight error on QC-LDPC codes. HWQCS signature scheme resists Prabowo-Tan's attack [35] on MPT-like signature scheme [31]. This is achieved by signing a message depending on a new ephemeral secret key for each signature rather than relying only on a fixed secret key. So, each signature can be viewed as a one-time signature. Furthermore, this signature is also secure against Bit-Flipping algorithm attack and statistical attack.

The organization of this paper is as follows. In Section 2, we provide a brief review of the properties of linear codes, quasi-cyclic codes and also define the syndrome decoding problem, etc. In Section 3, we propose a new high weight signature scheme (called HWQCS) which is based on 2-quasi-cyclic codes. We also provide security proof of the proposed HWQCS signature scheme under the random oracle model. In Section 4, we give a detailed analysis of various possible attacks on the proposed signature scheme HWQCS. In Section 5, we examine the public/secret key size and signature size for various security levels. Finally, the paper is concluded in Section 6.

2 Preliminaries

In this paper, let n, k be integers, denote by $\mathbb{F}_2$ the finite field of two elements, let $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{F}_2^n$ be a vector in $\mathbb{F}_2^n$.

2.1 Linear Codes

Definition 1 *Let* $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{F}_2^n$. *The* support *of* $\mathbf{a}$ *is the set consisting of all indices* $i \in \{1, \ldots, n\}$ *such that* $a_i \neq 0$. *The* Hamming weight *of* $\mathbf{a}$, *denoted by* $\mathrm{wt}(\mathbf{a})$ *is the cardinality of its support. The* Hamming distance *between* $\mathbf{a}$ *and* $\mathbf{b}$, *denoted by* $\mathrm{d}(\mathbf{a}, \mathbf{b})$ *is defined as* $\mathrm{wt}(\mathbf{a} - \mathbf{b})$, *i.e., the number of coordinates* $\mathbf{a}$ *and* $\mathbf{b}$ *differs on.*

Definition 2 *Let k and n be two positive integers with $k \leq n$. An $[n, k]$-linear code $\mathcal{C}$ of length n and dimension k is a linear subspace of dimension k of the vector space $\mathbb{F}_2^n$. The rate of the code $\mathcal{C}$ is $R = \frac{k}{n}$.*

Definition 3 *Let $\mathcal{C}$ be an $[n, k]$-linear code of length n and dimension k. We call its* minimum distance *δ the minimum Hamming weight of a non-zero codeword in $\mathcal{C}$, i.e.,*

$$\begin{aligned}\delta &= \min\{\mathrm{wt}(\mathbf{a}) \mid \mathbf{a} \in \mathcal{C}, \mathbf{a} \neq \mathbf{0}\} \\ &= \min\{\mathrm{wt}(\mathbf{a} - \mathbf{b}) \mid \mathbf{a}, \mathbf{b} \in \mathcal{C}, \mathbf{a} \neq \mathbf{b}\}.\end{aligned}$$

We sometimes refer to $\mathcal{C}$ as an $[n, k, \delta]$-code if δ is known.

Definition 4 *A matrix $G \in \mathbb{F}_2^{k \times n}$ is said to be a generator matrix of an $[n, k]$-linear code $\mathcal{C}$ if its rows form a basis of $\mathcal{C}$. Then, $\mathcal{C} = \{\mathbf{u}G \mid \mathbf{u} \in \mathbb{F}_2^k\}$. The parity-check matrix of $\mathcal{C}$ is $H \in \mathbb{F}_2^{(n-k) \times n}$ such that $GH^T = 0$ or $\mathbf{c}H^T = 0$ for all $\mathbf{c} \in \mathcal{C}$. Furthermore, G and H are said to be in systematic form if they are written as*

$$G = [I_k \quad A] \quad \text{resp.} \quad H = [I_{n-k} \quad B],$$

for some $A \in \mathbb{F}_2^{k \times (n-k)}$ and $B \in \mathbb{F}_2^{(n-k) \times k}$.

Problem 1 (Syndrome Decoding Problem (SDP)). *Given a matrix $H \in \mathbb{F}_2^{(n-k) \times n}$, a vector $\mathbf{s} \in \mathbb{F}_2^{n-k}$ and an integer $w > 0$ as input. The Syndrome Decoding problem is to determine a vector $\mathbf{e} \in \mathbb{F}_2^n$ such that $\mathrm{wt}(\mathbf{e}) \leq w$ and $\mathbf{s} = \mathbf{e}H^T$.*

Problem 2 (Codeword Finding Problem (CFP)). *Given a matrix $H \in \mathbb{F}_2^{(n-k) \times n}$, and an integer $w > 0$ as input. The Codeword Finding problem is to determine a vector $\mathbf{e} \in \mathbb{F}_2^n$ such that $\mathrm{wt}(\mathbf{e}) = w$ and $\mathbf{e}H^T = 0$.*

The SDP problem and CFP problem are well known and was proved to be NP-complete by Berlekamp, McEliece and van Tilborg in [11]. Moreover, it is proved that there exists a unique solution to SDP if the weight w is below the so-called GV Distance.

Definition 5 *Let $\mathcal{C}$ be an $[n, k]$ linear code over $\mathbb{F}_2$. The Gilbert–Varshamov (GV) Distance is the largest integer d such that*

$$\sum_{i=0}^{d-1} \binom{n}{i} \leq 2^{n-k}.$$

The first generic decoding method to solve SDP is called the Information Set Decoding (ISD) method, introduced by Prange [36] (denoted as Pra62) in 1962. It is the best known algorithm for decoding a general linear code. Since then, several improvements of the ISD method have been proposed for codes over the binary field, such as LB88 [26], Leon88 [27], Stern88 [40], Dum91 [20], and more recently by BLP11 [13], MMT11 [28], BJMM12 [7], MO15 [29]. The computational complexity of solving the syndrome decoding problem is quantified by the work factor $\mathcal{WF}_{\mathcal{A}}(n, k, w)$, which is defined as the average cost in binary operations of algorithm $\mathcal{A}$ to solve it. The work factor of Pra62 is given as follows.

$$\mathcal{WF}_{\mathsf{Pra62}}(n, k, w) = \frac{\min\{\binom{n}{w}, 2^{n-k}\}}{\binom{n-k}{w}}.$$

When $w = o(n)$, then $\mathcal{WF}_{\mathsf{Pra62}}(n, k, w) = \frac{\binom{n}{w}}{\binom{n-k}{w}}$ and $\frac{1}{w}\log_2 \frac{\binom{n}{w}}{\binom{n-k}{w}} \approx c$, where $c := -\log_2(1 - \frac{k}{n})$. Therefore, we have $\mathcal{WF}_{\mathsf{Pra62}}(n, k, w) \approx 2^{cw(1+o(1))}$.

Among the variants of solving algorithms for the syndrome decoding problem, the following result from [42] shows that their work factors are asymptotically the same.

Proposition 1 *[42] Let k and w be two functions of n such that* $\lim_{n\to\infty} \frac{k}{n} = R$, $0 < R < 1$, *and* $\lim_{n\to\infty} \frac{w}{n} = 0$. *For any algorithm $\mathcal{A}$ among the variants of* Pra62, Stern88, Dum91, MMT11, BJMM12 and MO15*, their work factors are asymptotically the same as*

$$\mathcal{WF}_{\mathcal{A}}(n, k, w) = 2^{cw(1+o(1))}, \quad \textit{where } c = -\log_2(1 - R)$$

when n tends to infinity.

2.2 Quasi-Cyclic Linear Codes

Let $\mathbb{F}_2$ be the finite field of two elements and let $\mathcal{R} := \mathbb{F}_2[x]/(x^k - 1)$ be the quotient ring of polynomials over $\mathbb{F}_2$ of degree less than k. Given $a = a_0 + a_1x + \ldots + a_{k-1}x^{k-1} \in \mathcal{R}$, we denote $\mathbf{a} := (a_0, a_1, \ldots, a_{k-1}) \in \mathbb{F}_2^k$. Let $\mathcal{R}^* = \{a \in \mathcal{R} \mid a \,\text{is invertible in}\, \mathcal{R}\}$. Let $\mathcal{V}$ be a vector space of dimension k over $\mathbb{F}_2$. Denote $\mathcal{V}_{k,w} := \{a \in \mathcal{R} = \mathbb{F}_2[x]/(x^k-1) \mid \text{wt}(\mathbf{a}) = w\}$. We sometimes abuse the notation by interchanging $\mathbf{a}$ with $a \in \mathcal{R}$.

Definition 6 (Circulant Matrix) *Let* $\mathbf{v} = (v_0, \cdots, v_{k-1}) \in \mathcal{V}$*, a circulant matrix defined by* $\mathbf{v}$ *is*

$$V := \begin{bmatrix} v_0 & v_1 & \ldots & v_{k-1} \\ v_{k-1} & v_0 & \ldots & v_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ v_1 & v_2 & \ldots & v_0 \end{bmatrix} \in \mathbb{F}_2^{k\times k}.$$

For $\mathbf{u}, \mathbf{v} \in \mathcal{R}$, the product $\mathbf{w} = \mathbf{uv}$ can be computed as $\mathbf{w} = \mathbf{u}V = \mathbf{v}U$, and $w_l = \sum_{i+j=l \bmod k} u_i v_j$ for $l = 0, \cdots, k-1$, where $\mathbf{w} = (w_0, \cdots, w_{k-1})$. To find the weight of $\mathbf{uv}$, we first compute the probability that $w_i = 1$, say p', then $\text{wt}(\mathbf{w}) = p' * k$. Now, we compute the probability that $w_i = 1$ as follows.

Lemma 1 [35] *Let* $\mathbf{u} \in \mathcal{V}_{k,\omega_u}$, $\mathbf{v} \in \mathcal{V}_{k,\omega_v}$ *and* $\mathbf{w} = \mathbf{uv} = (w_0, \cdots, w_{k-1})$. *Denote the probability that* $w_i = 1$, *for* $i \in \{0, \cdots, k-1\}$, *as* $P(k, \omega_u, \omega_v)$. *Then*

$$P(k, \omega_u, \omega_v) = \frac{1}{\binom{k}{\omega_v}} \sum_{\substack{1 \leq l \leq \min(\omega_u, \omega_v) \\ l \text{ odd}}} \binom{\omega_u}{l} \binom{k - \omega_u}{\omega_v - l}.$$

Definition 7 (Quasi-Cyclic Codes) *A linear block code* $\mathcal{C}$ *of length* lk *over* $\mathbb{F}_2$ *is called a quasi-cyclic code of index* l *if for any* $\mathbf{c} = (\mathbf{c}_0, \cdots, \mathbf{c}_{l-1}) \in \mathcal{C}$, *the vector obtained after applying a simultaneous circular shift to every block* $\mathbf{c}_0, \cdots, \mathbf{c}_{l-1}$ *is also a codeword.*

Definition 8 (Systematic 2-Quasi-Cyclic Codes, 2-QC Codes) *A systematic 2-quasi-cyclic* $[2k, k]$*-code has generator matrix of the form* $[H \ I_k] \in \mathbb{F}_2^{k \times 2k}$ *and parity check matrix* $[I_k \ H^T] \in \mathbb{F}_2^{k \times 2k}$.

Due to the quasi-cyclic structure of a code, any blockwise circular shift of a codeword is also a codeword. So, any circular shift of a syndrome will correspond to a blockwise circular shift of the error pattern. It has been shown in [38] that the work factor of the ISD algorithm for solving the syndrome decoding problem and the codeword finding problem for 2-quasi-cyclic codes for $n = 2k$ are

$$\mathcal{WF}_{\mathcal{A},\mathsf{2QCSD}}(n, k, w) := \frac{\mathcal{WF}_{\mathcal{A}}(n, k, w)}{\sqrt{n-k}} = 2^{c[1/2 + w(1+o(1))] - (\log_2 n)/2}$$

and

$$\mathcal{WF}_{\mathcal{A},\mathsf{2QCCF}}(n, k, w) := \frac{\mathcal{WF}_{\mathcal{A}}(n, k, w)}{n-k} = 2^{c[1 + w(1+o(1))] - \log_2 n}$$

respectively. Since the methods and the work factors for solving the syndrome decoding problem and the codeword finding problem for 2-quasi-cyclic codes require exponential time, therefore, we assume that the syndrome decoding problem and the codeword finding problem on quasi-cyclic codes are hard problems. We define the decisional codeword finding problem for 2-quasi-cyclic codes as follows.

Problem 3 (Decisional Codeword Finding Problem for 2-Quasi-Cyclic Codes (2QC-DCFP)). *Given a matrix* $[I_k \ \mathbf{h}] \in \mathbf{F}_2^{2k \times k}$, *and an even integer* $w > 0$ *as input, decide if there exists* $\mathbf{h}_0, \mathbf{h}_1 \in \mathcal{R}$ *such that* $\mathrm{wt}(\mathbf{h}_0) = \mathrm{wt}(\mathbf{h}_1) = w/2$ *and* $(\mathbf{h}_0, \mathbf{h}_1) \begin{bmatrix} \mathbf{I}_k \\ \mathbf{h} \end{bmatrix} = 0$.

In the special case of 2-quasi-cyclic codes with parity check matrix $H = [\mathbf{h}_0 \ \mathbf{h}_1] \in \mathbb{F}_2^{k \times 2k}$, where $(\mathbf{h}_0, \mathbf{h}_1)$ and $\mathbf{e}$ are of low weight approximate to $\sqrt{2k}$, we have what is called the quasi-cyclic low density parity check (QC-LDPC) codes. These codes are commonly used in the construction of key encapsulation mechanisms and signatures, such as BIKE [4] and HQC [1]. The Bit-Flipping algorithm [43] is used to decode an error $\mathbf{e}$ in BIKE.

On the other hand, for our signature (proposed in Section 3), we have $n = 2k$, $H = [\mathbf{I}_k\ \mathbf{c}]$ and $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2)$ is of high weight such that $\mathrm{wt}(\mathbf{e}) \gg \sqrt{n}$, $\frac{\mathtt{wt}(\mathbf{e})}{n} < \frac{1}{2}$, $\frac{\mathtt{wt}(\mathbf{e}_1)}{k} + \frac{\mathtt{wt}(\mathbf{ce}_2)}{k} > \frac{1}{2}$ and $\mathrm{wt}(\mathbf{c}) < \sqrt{k}$. Experimental results show that the Bit-Flipping algorithm [43] is unable to obtain $\mathbf{e}$ correctly in this case (many bits are decoded incorrectly). Up to our knowledge, there is no efficient decoding algorithm for high weight error. Therefore, we define the following problem and assume that it is a hard problem.

Problem 4 (Syndrome Decoding Problem for High Weight on QC-LDPC Codes (`HWQC-LDPC-SDP`)) *Let ω be integer, $n = 2k$, $H = [\mathbf{I}_k \quad \mathbf{c}]$ and $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2)$ is of high weight such that ω=$\mathrm{wt}(\mathbf{e}) \gg \sqrt{n}$, $\frac{\mathtt{wt}(\mathbf{e})}{n} < \frac{1}{2}$, $\frac{\mathtt{wt}(\mathbf{e}_1)}{k} + \frac{\mathtt{wt}(\mathbf{ce}_2)}{k} > \frac{1}{2}$ and $\mathrm{wt}(\mathbf{c}) < \sqrt{k}$. Given $H \in \mathbb{F}_2^{k\times 2k}$, $s \in \mathbb{F}_2^k$ and ω as input. The syndrome decoding problem for high weight on QC-LDPC code is to determine $\mathbf{e}$ such that $\mathrm{wt}(\mathbf{e}) = \omega$ and $\mathbf{s} = \mathbf{e}H^T$.*

3 HWQCS Signature Scheme

In this section, we present the Hamming-metric code-based digital signature scheme from QC-LDPC codes with high weight errors, which we call the HWQCS signature scheme. The HWQCS signature scheme is based on the hardness of the syndrome decoding problem and the codeword finding problem on quasi-cyclic codes. Furthermore, the HWQCS signature scheme is different from the MPT signature scheme [31] and is resistant to Prabowo-Tan's attack [35] as each signature can be thought of as a one-time signature with a new ephemeral secret key, while the MPT signature is based on a fixed secret key.

A signature scheme consists of three algorithms: KeyGen, Sign and Verify.

- KeyGen: Given a security parameter λ, the key generation algorithm returns a key pair (pk, sk) where pk and sk are the public key and the secret key respectively.
- Sign: The algorithm, on input a message m and the secret key sk, returns a signature σ.
- Verify: Given a message m, a public key pk and a signature σ as input, the algorithm returns either 0 or 1 depending on whether the signature σ is valid or not.

Before we describe a HWQCS signature scheme, we first define the required parameters. Let $k, \omega_f, \omega_u, \omega_e, \omega_c, \omega_s, \omega_t$ be integers as public parameters. The HWQCS signature scheme is described as follows.

Algorithm 1: Key Generation of HWQCS Signature Scheme

```
  Input  : $k, \omega_f$, security parameter $\lambda$
  Output: $pk = (\mathbf{h})$
1 Choose random $\mathbf{f}_1, \mathbf{f}_2 \in \mathcal{V}_{k,\omega_f}$ and both are invertible
2 Compute $\mathbf{h} := \mathbf{f}_1^{-1}\mathbf{f}_2$ in $\mathcal{R}^*$
3 The public key is $pk = (\mathbf{h})$ and the secret key is $sk = (\mathbf{f}_1, \mathbf{f}_2)$
```

Algorithm 2: Signing of HWQCS Signature Scheme

```
  Input  : $k, \omega_f, \omega_u, \omega_e, \omega_c, \omega_s, \omega_t$, message $m$, $pk = (\mathbf{h})$ and $sk = (\mathbf{f}_1, \mathbf{f}_2)$
  Output: signature $\sigma$
1 Choose random $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{V}_{k,\omega_e}$ and $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{V}_{k,\omega_u}$
2 Compute $\mathbf{b} := (\mathbf{e}_1, \mathbf{e}_2)\begin{bmatrix}\mathbf{h}\\ \mathbf{h}^{-1}\end{bmatrix}$ in $\mathcal{R}$
3 Compute $\mathbf{c} := \mathcal{H}(m\|\mathbf{b}\|(\mathbf{u}_1\mathbf{f}_2 + \mathbf{u}_2\mathbf{f}_1)\|pk) \in \mathcal{V}_{k,\omega_c}$
4 Compute $\mathbf{s}_i := \mathbf{u}_i\mathbf{f}_i + \mathbf{c}\mathbf{e}_i$ in $\mathcal{R}$ for $i = 1, 2$
5 if $\mathrm{wt}(\mathbf{s}_1) > \omega_s$ or $\mathrm{wt}(\mathbf{s}_2) > \omega_s$ or $\mathrm{wt}(\mathbf{u}_1\mathbf{f}_2 + \mathbf{u}_2\mathbf{f}_1) > \omega_t$ then
6     repeat from Step 1
7 else
8     the signature is $\sigma = (\mathbf{c}, \mathbf{b}, \mathbf{s}_1, \mathbf{s}_2)$
9 end if
```

Algorithm 3: Verification of HWQCS Signature Scheme

```
  Input  : message $m$, $pk$, signature $\sigma = (\mathbf{c}, \mathbf{b}, \mathbf{s}_1, \mathbf{s}_2)$
  Output: validity of the signature
1 Compute $\mathbf{t} := (\mathbf{s}_1, \mathbf{s}_2)\begin{bmatrix}\mathbf{h}\\ \mathbf{h}^{-1}\end{bmatrix} - \mathbf{cb}$ in $\mathcal{R}$
2 Compute $\mathbf{c}' := \mathcal{H}(m\|\mathbf{b}\|\mathbf{t}\|pk) \in \mathcal{V}_{k,\omega_c}$
3 if $\mathbf{c}' = \mathbf{c}$ and $\mathrm{wt}(\mathbf{t}) \leq \omega_t$ and $\mathbf{t} \neq 0$ $in$ $\mathcal{R}$ then
4     the signature is valid
5 else
6     the signature is invalid
7 end if
```

Correctness:

$$\begin{aligned}\mathbf{t} &= (\mathbf{s}_1, \mathbf{s}_2)\begin{bmatrix}\mathbf{h}\\ \mathbf{h}^{-1}\end{bmatrix} - \mathbf{cb}\\ &= (\mathbf{u}_1\mathbf{f}_2 + \mathbf{ce}_1\mathbf{h}) + (\mathbf{u}_2\mathbf{f}_1 + \mathbf{ce}_2\mathbf{h}^{-1}) - \mathbf{c}(\mathbf{e}_1\mathbf{h} + \mathbf{e}_2\mathbf{h}^{-1})\\ &= \mathbf{u}_1\mathbf{f}_2 + \mathbf{u}_2\mathbf{f}_1\end{aligned}$$

We define the notion of existential unforgeability under adaptive chosen message attack as follows.

Definition 9 (**EUF-CMA Security**) *A signature scheme is existential unforgeable under adaptive chosen message attack* (EUF-CMA) *if given a public key* pk *to any polynomial-time adversary* $\mathcal{A}$ *who can access the signing oracle* $\mathsf{Sign}(\mathsf{sk}, \cdot)$ *and query a number of signatures, then the adversary* $\mathcal{A}$ *can produce a valid signature* σ *for a message* m *which has not been previously queried to the signing oracle only with negligible success probability (the success probability is denoted as* $\Pr[\texttt{Forge}]$*).*

The advantage Adv of an adversary $\mathcal{A}$ in successfully solving a problem is defined as follows.

Definition 10 *The advantage of an adversary* $\mathcal{A}$ *in solving a problem* B *denoted as* $\mathsf{Adv}(\mathsf{B})$ *is defined as the probability that* $\mathcal{A}$ *successfully solves problem* B.

We define the following assumptions which are used to prove the security of the proposed signature scheme.

Assumption 1 (Syndrome Decoding for 2-Quasi-Cyclic Code (`2QC-SDP`) Assumption) *The syndrome decoding for 2-quasi-cyclic code assumption is the assumption that the advantage of an adversary* $\mathcal{A}$ *in solving* `2QC-SDP` *is negligible, i.e.* $\texttt{Adv}(\texttt{2QC-SDP}) < \epsilon_{\texttt{2QC-SDP}}$.

Assumption 2 (Codeword Finding for 2-Quasi-Cyclic Codes (`2QC-CFP`) Assumption) *The codeword finding for quasi-cyclic codes assumption is the assumption that the advantage of an adversary* $\mathcal{A}$ *in solving* `2QC-CFP` *is negligible, i.e.* $\texttt{Adv}(\texttt{2QC-CFP}) < \epsilon_{\texttt{2QC-CFP}}$.

Assumption 3 (Decisional Codeword Finding for 2-Quasi-Cyclic Codes (`2QC-DCFP`) Assumption) *The decisional codeword finding for 2-quasi-cyclic codes assumption is the assumption that the advantage of an adversary* $\mathcal{A}$ *in solving* `2QC-DCFP` *is negligible, i.e.* $\texttt{Adv}(\texttt{2QC-DCFP}) < \epsilon_{\texttt{2QC-DCFP}}$.

Assumption 4 (Syndrome Decoding for High Weight on QC-LDPC Codes (`HWQC-LDPC-SDP`) Assumption) *The syndrome decoding for high weight of QC-LDPC codes assumption is the assumption that the advantage of an adversary* $\mathcal{A}$ *in solving* `HWQC-LDPC-SDP` *is negligible, i.e.* $\texttt{Adv}(\texttt{HWQC-LDPC-SDP}) < \epsilon_{\texttt{HWQC-LDPC-SDP}}$.

Theorem 1 *Under the* 2QC-SDP, 2QC-DCFP, 2QC-CFP, HWQC-LDPC-SDP *assumptions, the* HWQCS *signature scheme with parameters* $(k, \omega_f, \omega_u, \omega_e, \omega_c, \omega_s, \omega_t)$ *is secure under the* EUF-CMA *model in the classical random oracle model.*

Proof. We consider a chosen-message EUF adversary $\mathcal{A}$ against the HWQCS signature scheme. To prove the security, adversary $\mathcal{A}$ interacts with the real signature scheme and makes a sequence of experiments. The adversary $\mathcal{A}$ is first given a public key $\mathbf{h}$. $\mathcal{A}$ made q_s signing queries and $q_{\mathcal{H}}$ hash $(\mathcal{H})$ queries. Finally, $\mathcal{A}$ outputs a message/signature pair such that the message has not been queried previously to the signing oracle. Let $\mathrm{Pr}_i[\mathsf{Forge}]$ be the probability of an event in experiment i that $\mathcal{A}$ obtains a valid signature of a message that has not been queried previously to the signing oracle. Let $\mathrm{Pr}_0[\mathsf{Forge}]$ be the success probability of an adversary $\mathcal{A}$ at the beginning (Experiment 0). Our goal is to give an upper-bound of $\mathrm{Pr}_0[\mathsf{Forge}]$.

Experiment 1. During the course of the experiment, if there is a collision in $\mathcal{H}$, then we abort the experiment. The number of queries to the hash oracle or the signing oracle throughout the experiment is at most $q_s + q_{\mathcal{H}}$. Thus,

$$|\ \mathrm{Pr}_0[\mathsf{Forge}] - \mathrm{Pr}_1[\mathsf{Forge}]\ | \le \frac{q_s + q_{\mathcal{H}}}{\binom{k}{\omega_c}}.$$

Experiment 2. During the course of the experiment, $\mathcal{A}$ received a number of signatures $\sigma_j = (\mathbf{c}, \mathbf{b}, \mathbf{s}_1, \mathbf{s}_2)_j$ for $j = 1, \cdots, q_s$. If $\mathcal{A}$ could solve for $(\mathbf{e}_1, \mathbf{e}_2)_j$ from $\mathbf{b}_j = (\mathbf{e}_1, \mathbf{e}_2)_j \begin{bmatrix} \mathbf{h} \\ \mathbf{h}^{-1} \end{bmatrix}$ for some j, then $\mathcal{A}$ could forge a new signature. But, the probability that $\mathcal{A}$ could solve it is bounded by $\epsilon_{\texttt{2QC-SDP}}$ and we abort the experiment in this case. Thus,

$$|\ \mathrm{Pr}_1[\mathsf{Forge}] - \mathrm{Pr}_2[\mathsf{Forge}]\ | \le \epsilon_{\texttt{2QC-SDP}}.$$

Experiment 3. During the course of the experiment, $\mathcal{A}$ received a number of signatures $\sigma_j = (\mathbf{c}, \mathbf{b}, \mathbf{s}_1, \mathbf{s}_2)_j$ for $j = 1, \cdots, q_s$. If $\mathcal{A}$ could solve for $(\mathbf{u}_i\mathbf{f}_i, \mathbf{e}_i)$ from $(\mathbf{s}_i)_j = (\mathbf{u}_i\mathbf{f}_i, \mathbf{e}_i)_j \begin{bmatrix} \mathbf{I}_k \\ \mathbf{c}_j \end{bmatrix}$ for $i = 1, 2$ for some j, then $\mathcal{A}$ could forge a new signature. But, the probability that $\mathcal{A}$ could solve it is bounded by $\epsilon_{\texttt{HWQC-LDPC-SDP}}$ and we abort the experiment in this case. Thus,

$$|\ \mathrm{Pr}_2[\mathsf{Forge}] - \mathrm{Pr}_3[\mathsf{Forge}]\ | \le 2\epsilon_{\texttt{HWQC-LDPC-SDP}}.$$

Experiment 4. In this experiment, a public key $\mathbf{h}$ is replaced by a random $\mathbf{h}' \in \mathcal{R}^*$. To distinguish Experiment 4 from Experiment 3, the adversary must in fact distinguish a well-formed public key $\mathbf{h} = \mathbf{f}_1^{-1}\mathbf{f}_2$ from a random invertible element of $\mathcal{R}$. Thus, we have

$$|\ \mathrm{Pr}_3[\mathsf{Forge}] - \mathrm{Pr}_4[\mathsf{Forge}]\ | \le \epsilon_{\texttt{2QC-DCFP}}.$$

Furthermore, in this experiment, an adversary $\mathcal{A}$ has no signature information on $\mathbf{h}'$ and needs to solve a codeword finding problem for 2-quasi-cyclic codes in order to forge a signature. Thus,

$$| \Pr_4[\mathsf{Forge}] \,| \leq \epsilon_{\mathtt{2QC\text{-}CFP}}.$$

Combining the above experiments, the success probability of the adversary $\mathcal{A}$ is

$$\begin{aligned} | \Pr_0[\mathsf{Forge}] \,| &\leq \sum_{i=0}^{3} | \Pr_i[\mathsf{Forge}] - \Pr_{i+1}[\mathsf{Forge}] \,| + | \Pr_4[\mathsf{Forge}] \,| \\ &\leq \epsilon_{\mathtt{2QC\text{-}CFP}} + \epsilon_{\mathtt{2QC\text{-}DCFP}} + 2\epsilon_{\mathtt{HWQC\text{-}LDPC\text{-}SDP}} + \epsilon_{\mathtt{2QC\text{-}SDP}} + \frac{q_s + q_{\mathcal{H}}}{\binom{k}{\omega_c}}. \end{aligned}$$

4 Security Analysis

Let λ be the security level. For the security analysis, we consider two common types of attacks, namely, key recovery attacks and signature forgeries.

4.1 Key Recovery Attack

Finding the secret key $(\mathbf{f}_1, \mathbf{f}_2)$ from the public key $\mathbf{h} = \mathbf{f}_1^{-1}\mathbf{f}_2$ is equivalent to finding the codeword $(\mathbf{f}_1, \mathbf{f}_2)$ with parity check matrix $[\mathbf{h} \quad \mathbf{I}_k]$ such that $(\mathbf{f}_1, \mathbf{f}_2)\begin{bmatrix}\mathbf{h}\\ \mathbf{I}_k\end{bmatrix} = \mathbf{0}$. The work factor of solving the codeword finding problem for quasi-cyclic parity-check codes is

$$\mathcal{WF}_{\mathcal{A},\mathsf{2QCCF}}(2k, k, 2\omega_f) = 2^{c[1+2\omega_f(1+o(1))]-\log_2 2k}, \qquad \text{where } c = 1.$$

Therefore, we can prevent key recovery attack by choosing the parameters such that $1 + 2\omega_f(1 + o(1)) - \log_2 2k \geq \lambda$, where λ is the security level.

Another method to find the secret key $(\mathbf{f}_1, \mathbf{f}_2)$ is by performing exhaustive search for $\mathbf{f}_1$ and checking whether $\mathbf{f}_1\mathbf{h}$ is of small Hamming weight w_f. The complexity of performing this exhaustive search is $\binom{k}{\omega_f}$. So, we must choose the parameters such that $\log_2 \binom{k}{\omega_f} \geq \lambda$, where λ is the security level.

Based on the above, we choose the parameters such that

$$\min\left\{ \log_2 \binom{k}{\omega_f},\ 1 + 2\omega_f(1 + o(1)) - \log_2 2k \right\} \geq \lambda.$$

This ensures that the scheme is resistant against key recovery attacks.

4.2 Signature Forgery

4.2.1 Collision

For a signature scheme based on the Schnorr scheme, it is important to address the issue of collisions between different messages. In order to prevent collisions, one way is to use a collision-free hash function. Another way is to use a secure hash function such that the collision is minimal, that is, satisfying $\log_2 \binom{k}{\omega_c} \geq 2\lambda$, where λ is the security level.

4.2.2 Forgery From Known Signature

We consider the following methods to forge a signature.

4.2.2.1 Forgery via Syndrome Decoding Algorithm

From a given signature, we have $\mathbf{b} = \mathbf{e}_1\mathbf{h} + \mathbf{e}_2\mathbf{h}^{-1}$, $\mathbf{s}_i = \mathbf{u}_i\mathbf{f}_i + \mathbf{c}\mathbf{e}_i$, where $i = 1, 2$. Equivalently, $\mathbf{b} = (\mathbf{e}_1, \mathbf{e}_2) \begin{bmatrix} \mathbf{h} \\ \mathbf{h}^{-1} \end{bmatrix}$, $\mathbf{s}_i = (\mathbf{u}_i\mathbf{f}_i, \mathbf{e}_i) \begin{bmatrix} \mathbf{I}_k \\ \mathbf{c} \end{bmatrix}$ for $i = 1, 2$.

(1) One may use syndrome decoding algorithms to recover $(\mathbf{e}_1, \mathbf{e}_2)$ from $\mathbf{b} = (\mathbf{e}_1, \mathbf{e}_2) \begin{bmatrix} \mathbf{h} \\ \mathbf{h}^{-1} \end{bmatrix}$. The work factor is

$$\mathcal{WF}_{\mathcal{A},\mathsf{2QCSD}}(2k, k, 2\omega_e) = 2^{c[1/2+2\omega_e(1+o(1))]-(\log_2 2k)/2}, \quad \text{where } c = 1.$$

In order to prevent this attack, we choose k, ω_e such that

$$1/2 + 2\omega_e(1 + o(1)) - (\log_2 2k)/2 \geq \lambda,$$

where λ is the security level.

(2) One may also use syndrome decoding algorithms to recover $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{u}_1\mathbf{f}_1, \mathbf{u}_2\mathbf{f}_2)$ from

$$\begin{bmatrix} \mathbf{b} \\ \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h} & \mathbf{h}^{-1} & \mathbf{0}_k & \mathbf{0}_k \\ \mathbf{c} & \mathbf{0}_k & \mathbf{I}_k & \mathbf{0}_k \\ \mathbf{0}_k & \mathbf{c} & \mathbf{0}_k & \mathbf{I}_k \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{u}_1\mathbf{f}_1 \\ \mathbf{u}_2\mathbf{f}_2 \end{bmatrix}$$

Note that the weight of $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{u}_1\mathbf{f}_1, \mathbf{u}_2\mathbf{f}_2)$ is $\omega = 2(\omega_e + \mathrm{wt}(\mathbf{u}_1\mathbf{f}_1))$. So, the work factor is

$$\mathcal{WF}_{\mathcal{A},\mathsf{4QCSD}}(4k, k, \omega) = \frac{\min\{\binom{4k}{\omega}, 2^{4k-k}\}}{\binom{4k-k}{\omega}\sqrt{4k-k}} = \frac{\min\{\binom{4k}{\omega}, 2^{3k}\}}{\binom{3k}{\omega}\sqrt{3k}}.$$

(3) Another method to find the ephemeral secret $(\mathbf{e}_1, \mathbf{e}_2)$ is by performing exhaustive search on $\mathbf{e}_1$ and checking whether $\mathbf{e}_2 = \mathbf{b}\mathbf{h} + \mathbf{e}_1\mathbf{h}^2$ is of small Hamming weight w_e. The complexity of performing this method is $\binom{k}{\omega_e}$. In order to prevent this attack, we choose k, ω_e such that $\log_2 \binom{k}{\omega_e} \geq \lambda$, where λ is the security level.

Suppose an adversary can recover $(\mathbf{e}_1, \mathbf{e}_2)$ using any of the above methods. Then, the adversary obtains $\mathbf{u}_i\mathbf{f}_i = \mathbf{s}_i - \mathbf{c}\mathbf{e}_i$ for $i = 1, 2$. Afterwards, he can forge a new signature by generating new $\mathbf{b}' = \mathbf{e}'_1\mathbf{h}+\mathbf{e}'_2\mathbf{h}^{-1}$ and setting $\mathbf{s}'_i = \mathbf{u}_i\mathbf{f}_i+\mathbf{c}'\mathbf{e}'_i$, for $i = 1, 2$.

Based on the above analysis, in order to resist forgery attacks with security level λ, we choose the parameters k, ω, ω_e satisfying the following conditions:

$$\min\left\{\log_2\binom{k}{\omega_e},\quad 1/2+2\omega_e(1+o(1))-(\log_2 2k)/2, \log_2\frac{\min\{\binom{4k}{\omega},2^{3k}\}}{\binom{3k}{\omega}\sqrt{3k}}\right\}\geq\lambda.$$

4.2.2.2 Forgery via Bit-Flipping Algorithm

Given a signature, we have $\mathbf{s}_i = \mathbf{u}_i\mathbf{f}_i + \mathbf{c}\mathbf{e}_i$, where $i = 1, 2$. One may try to apply the bit-flipping algorithm on $\mathbf{s}_i = (\mathbf{u}_i\mathbf{f}_i, \mathbf{e}_i)\begin{bmatrix}\mathbf{I}_k\\ \mathbf{c}\end{bmatrix}$ for $i = 1, 2$ to recover $\mathbf{e}_i$.

In this case, $n = 2k$, $H = \begin{bmatrix}\mathbf{I}_k\\ \mathbf{c}\end{bmatrix}$ and the threshold $\tau = \lfloor\rho\cdot\omega_c\rfloor$, where ρ is the probability that $(\mathbf{e}_i)_j = (\mathbf{s}_i)_j = 1$ for $j \in \{0, \cdots, k-1\}$ and will be given in the following proposition.

Proposition 2 *If* $\mathbf{c}_0 = 1$ *and* $(\mathbf{s}_i)_j = 1$, *then* $\rho = Prob[\,(\mathbf{e}_i)_j = (\mathbf{s}_i)_j = 1\,]$ *is equal to*

$$(1-P(k,\omega_u,\omega_f))*(1-P(k,\omega_c-1,\omega_e-1))+P(k,\omega_u,\omega_f)*P(k,\omega_c-1,\omega_e-1).$$

Proof. If $\mathbf{c}_0 = 1$, then

$$(\mathbf{s}_i)_j = (\mathbf{u}_i\mathbf{f}_i+\mathbf{c}\mathbf{e}_i)_j = (\mathbf{e}_i)_j + \sum_{l=0}^{k-1}(\mathbf{u}_i)_l(\mathbf{f}_i)_{j-l \bmod k} + \sum_{\substack{0\leq l\leq k-1\\ l\neq j}}\mathbf{c}_l(\mathbf{e}_i)_{j-l \bmod k}.$$

Note that the probability that $(\mathbf{u}_i\mathbf{f}_i)_j = 1$ and $\sum_{l\neq j}(\mathbf{c}_i)_l(\mathbf{e}_i)_{j-l \bmod k} = 1$ are $P(k, \omega_u, \omega_f)$ and $P(k, \omega_c - 1, \omega_e - 1)$ respectively. Hence, the probability that $(\mathbf{e}_i)_j = (\mathbf{s}_i)_j = 1$ is

$$(1-P(k,\omega_u,\omega_f))*(1-P(k,\omega_c-1,\omega_e-1))+P(k,\omega_u,\omega_f)*P(k,\omega_c-1,\omega_e-1).$$

As in Problem 4, we choose the parameters such that $\mathrm{wt}(\mathbf{u}_i\mathbf{f}_i)+\omega_e \gg \sqrt{2k}$, $\frac{\mathrm{wt}(\mathbf{u}_i\mathbf{f}_i)+\mathrm{wt}(\mathbf{c}\mathbf{e}_i)}{k} > \frac{1}{2}$ and $\omega_c \ll \sqrt{k}$. With this choice of parameters, the bit-flipping algorithm will not be able to decode correctly to obtain $\mathbf{e}_i$ for $i = 1, 2$. Hence, one cannot obtain $\mathbf{u}_i\mathbf{f}_i$ and forge a new signature.

4.2.3 Forgery Without Knowing Any Signature

Note that an adversary can generate $\mathbf{b} = \mathbf{e}_1\mathbf{h}+\mathbf{e}_2\mathbf{h}^{-1}$. To forge a signature, the adversary has to produce $\mathbf{s}_i$ of low weight. As the adversary needs to produce $\mathbf{u}_i\mathbf{f}_i$ of low Hamming weight and $\mathbf{u}_1\mathbf{f}_1\mathbf{h}$ such that $\mathbf{u}_2\mathbf{f}_2\mathbf{h}^{-1}$ are also of low Hamming weight, therefore $\mathrm{wt}(\mathbf{u}_i\mathbf{f}_i)$ must be set to low. In order to ensure this, we need to define the normal distribution and present the following lemma and corollary.

Let $\mathcal{N}(0, \sigma^2)$ be the normal distribution with mean 0 and standard deviation σ. Its density function is $\rho_\sigma(x) = (\frac{1}{\sqrt{2\pi\sigma^2}})e^{-\frac{x^2}{2\sigma^2}}$ for $x \in \mathbb{R}$.

Lemma 2 [16] *For $k > 2$, $Z \sim \mathcal{N}(0, \sigma^2)$, then*

$$\Pr[\,|z| > k\sigma \mid z \leftarrow Z\,] \leq \frac{1}{2}(e^{-k^2} + e^{-\frac{k^2}{2}}).$$

Corollary 1 (1) *For $\kappa > 2$, $Y \sim \mathcal{N}(\mu, \sigma^2)$, then* $\Pr[\,|y - \mu| > \kappa\sigma \mid y \leftarrow Y\,] \leq \frac{1}{2}(e^{-\kappa^2} + e^{-\frac{\kappa^2}{2}})$.
(2) *Let n be a large positive integer and $0 < p < 1$. If Y is a binomial distribution with parameters n and p (denoted* $\mathrm{Bin}(n, p)$*), then Y approximates to $\mathcal{N}(\mu, \sigma^2)$, where $\mu = np$ and $\sigma = \sqrt{np(1-p)}$.*
(3) *In (2), if $0 < l < p < 1$ and $\kappa = \frac{(p-l)\sqrt{n}}{\sqrt{p(1-p)}}$, then*

$$\Pr[\,|y - np| > (p - l)n \mid y \leftarrow Y\,] \leq \frac{1}{2}(e^{-\kappa^2} + e^{-\frac{\kappa^2}{2}}) < e^{-\kappa^2/2}.$$

Setting $n = k$, $p = \frac{1}{2}$ and $l < \frac{1}{2}$ in Corollary 1 (3), we have $\Pr[|y - \frac{k}{2}| > (\frac{1}{2} - l)k \mid y \leftarrow \mathrm{Bin}(k, p)] < e^{-\kappa^2/2}$. To ensure that the probability is negligible, we should choose κ such that $\kappa = (1 - 2l))\sqrt{k}$ and $\frac{1}{2}(e^{-\kappa^2} + e^{-\kappa^2/2}) < e^{-\kappa^2/2} < 2^{-\lambda}$, that is,

$$\frac{\kappa^2}{2}\log_2 e > \lambda \implies \kappa > \sqrt{\frac{2\lambda}{\log_2 e}}.$$

Letting $\kappa_0 = \sqrt{\frac{2\lambda}{\log_2 e}}$, we have

λ	128	192	256
κ_0	13.320	16.314	18.838

This means that if an adversary randomly picks an element $\mathbf{a}$ in place of $\mathbf{u}_i\mathbf{f}_i$ for $i = 1, 2$, then the probability that $|\mathrm{wt}(\mathbf{a}) - \frac{k}{2}| \leq \kappa\sqrt{k/4}$ is more than $1 - 2^{-\lambda}$. Hence, by selecting appropriate l, k such that $(1 - 2l)\sqrt{k} \geq \kappa_0$, we can ensure that the adversary cannot find $\mathbf{a}$ of weight less than lk. Therefore, it is not possible to forge a signature with probability more than $2^{-\lambda}$.

5 Parameters Selections

Based on the above security analysis, the parameters $(k, \omega_f, \omega_u, \omega_e, \omega_c, \omega_s)$ of the signature scheme must be chosen properly in order to achieve λ-bit computational security. The following conditions are to be fulfilled:

$$\min\left\{\log_2\binom{k}{\omega_f},\ 1+2\omega_f(1+o(1))-\log_2 2k\right\} \geq \lambda,$$

$$\log_2\binom{k}{\omega_c} \geq 2\lambda,$$

$$\min\left\{\log_2\binom{k}{\omega_e},\ \frac{1}{2}+2\omega_e(1+o(1))-\frac{\log_2 2k}{2},\ \log_2\frac{\min\{\binom{4k}{\omega},2^{3k}\}}{\binom{3k}{\omega}\sqrt{3k}}\right\} \geq \lambda,$$

$$(1-2l)\sqrt{k} > \sqrt{\frac{2\lambda}{\log_2 e}},$$

$$\mathrm{wt}(\mathbf{u}_i\mathbf{f}_i)+\omega_e \gg \sqrt{2k},$$

$$\frac{\mathrm{wt}(\mathbf{u}_i\mathbf{f}_i)+\mathrm{wt}(\mathbf{c}\mathbf{e}_i)}{k} > \frac{1}{2}.$$

The parameters for various security levels are given in the following Table 1.

Table 1. The parameters of the HWQCS signature

Name	λ	k	ω_f	ω_u	ω_e	ω_c	$\frac{\mathrm{wt}(\mathbf{s})}{k}$	$\frac{\mathrm{wt}(\mathbf{uf})}{k}$	$\frac{\mathrm{wt}(\mathbf{t})}{k}$
Para-1	128	12539	145	33	141	31	0.3863	0.2694	0.3937
Para-2	192	18917	185	41	177	39	0.3938	0.2779	0.4013
Para-3	256	25417	201	51	191	51	0.3978	0.2786	0.4019

To compute the size of HWQCS signature scheme, the public key size is $\lceil k/8 \rceil$ bytes, the secret key size is $2 * \lceil \lceil \log_2 k \rceil * \omega_f/8 \rceil$ bytes and the signature size is $3 * \lceil k/8 \rceil + \lceil \lceil \log_2 k \rceil * \omega_c/8 \rceil$ bytes. We list their sizes for various security levels in Table 2.

Table 2. Size of Signature Schemes (at certain classical security levels)

Scheme	Security	Size (in Bytes)		
		PK	SK	Sg
HWQCS-I	128	1,568	508	4,759
HWQCS-II	192	2,365	694	7,169
HWQCS-III	256	3,178	754	9,630

As listed in Table 2, the public key size, secret key size and signature size of the proposed signature scheme HWQCS-I are 1568 bytes, 508 bytes and 4759 bytes respectively for 128-bit classical security level.

We provide comparison of the key sizes and signature size for various code-based signature schemes in Table 3.

Table 3. Comparison of Various Code-based Signature Schemes (at certain classical security levels)

Scheme	PK size	SK size	Sg size	C.Sec
HWQCS-I	1.568 KB	508 B	4.759 KB	128
Durandal-I19 [3]	15.25 KB	2.565 KB	4.060 KB	128
WAVE23 [32]	3.60 MB	2.27 MB	737 B	128
CCJ23 [14]	90 B	231 B	12.52 KB	128
SDitH23 [2]	120 B	404 B	8.26 KB	128
BG23 [10]	1 KB	2 KB	13.5 KB	128
cRVDC19 [9]	0.152 KB	0.151 KB	22.480 KB	125
CVE18 [8]	7.638 KB	0.210 KB	436.600 KB	80

In Table 3, it can be observed that the signature size of the proposed signature scheme HWQCS-I is smaller than the other signature schemes except for the WAVE23 signature scheme [32] and the Durandal-I19 signature scheme [3]. However, it should be noted that the public key sizes for both the WAVE23 and Durandal-I19 signature schemes exceed ten thousand bytes. These are larger than that of the signature scheme HWQCS-I. Moreover, recently there is an attack on Durandal-I19 [6] which requires it to increase its parameter sizes.

Although the public key size of the CCJ23 signature scheme [14] and the SDitH23 signature scheme [2] are relatively small, but their signature sizes are more than eight thousand bytes. Overall, the proposed signature scheme HWQCS-I has shorter combined key and signature sizes than other signature schemes.

6 Conclusion

In this paper, we constructed a new Hamming metric code-based signature scheme (called HWQCS signature scheme). The security of HWQCS signature is based on the hardness of the syndrome decoding problem and the codeword finding problem on 2-quasi-cyclic codes, as well as on high error for quasi-cyclic low parity-check codes respectively. We provided security proof of the HWQCS signature scheme under the random oracle model and gave detailed analysis on the security of the HWQCS signature scheme against Bit-Flipping attack and statistical attack. Furthermore, we also provided concrete parameter choices for the HWQCS signature scheme and compared its key sizes and signature size to other existing signature schemes. The signature scheme HWQCS-I outperforms other code-based signature schemes with a public key size of 1568 bytes, secret key size of 508 bytes and signature size of 4759 bytes at 128-bit security level.

References

1. Aguilar-Melchor, C., et al.: Hamming Quasi-Cyclic (HQC). Submission to the NIST post quantum standardization process (2017). https://www.pqc-hqc.org/doc/hqc-specification_2021-06-06.pdf
2. Aguilar-Melchor, C., et al.: The Syndrome Decoding in the Head (SD-in-the-Head) Signature Scheme. Submission to the NIST call for additional post-quantum signatures (2023). https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-1/spec-files/SDitH-spec-web.pdf
3. Aragon, N., Blazy, O., Gaborit, P., Hauteville, A., Zémor, G.: Durandal: a rank metric based signature scheme. In: Ishai, Y., Rijmen, V. (eds.) EUROCRYPT 2019, Part III. LNCS, vol. 11478, pp. 728-758. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17659-4_25
4. Aragon, N., et al.: BIKE: Bit Flipping Key Encapsulation. Submission to the NIST post quantum standardization process (2017). https://bikesuite.org/files/v5.0/BIKE_Spec.2022.10.10.1.pdf
5. Aragon, N., Baldi, M., Deneuville, J.C., Khathuria, K., Persichetti, E., Santini, P. : Cryptanalysis of a code-based full-time signature. Designs, Codes and Cryptography 89 (9) (2021) 2097–2112. https://doi.org/10.1007/s10623-021-00902-7
6. Aragon, N., Dyseryn, V., Gaborit, P.: Analysis of the Security of the PSSI Problem and Cryptanalysis of the Durandal Signature Scheme. In: Handschuh, H., Lysyanskaya, A. (eds) CRYPTO 2023. LNCS, vol. 14083. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-38548-3_5
7. Becker, A., Joux, A., May, A., Meurer, A.: Decoding Random Binary Linear Codes in $2^{n/20}$: How 1+1=0 improves information set decoding. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 520–536. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-29011-4_31
8. Bellini, E., Caullery, F., Hasikos, A., Manzano, M., Mateu, V.: Code-based signature schemes from identification protocols in the rank metric. In: Camenisch, J., Papadimitratos, P. (eds.) CANS 2018. LNCS, vol. 11124, pp. 277-298. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00434-7_14
9. Bellini, E., Caullery, F., Gaborit, P., Manzano, M., Mateu, V.: Improved Veron identification and signature schemes in the rank metric. In: IEEE International Symposium on Information Theory, pp. 1872–1876 (2019). https://doi.org/10.1109/ISIT.2019.8849585
10. Bidoux, L., Gaborit, P.: Compact Post-quantum signatures from proofs of knowledge leveraging structure for the PKP, SD and RSD Problems. In: Hajji,S. El., et al. (Eds.): C2SI 2023, LNCS, vol. 13874, pp. 10-42. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33017-9_2

11. Berlekamp, E., McEliece, R., van Tilborg, H.: On the inherent intractability of certain coding problems (corresp.). IEEE Trans. Inf. Theory 24(3), 384–386 (1978). https://doi.org/10.1109/TIT.1978.1055873
12. Bernstein, D.J., et al.: Classic McEliece: conservative code-based cryptography. Submission to the NIST post quantum standardization process (2017). https://classic.mceliece.org/mceliece-rationale-20221023.pdf
13. Bernstein, D.J., Lange, T., Peters, C.: Smaller decoding exponents: ball-collision decoding. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 743-760. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22792-9_42
14. Carozza, E., Couteau, G. Joux, A. : Short Signatures from Regular Syndrome Decoding in the Head. In: Hazay, C., Stam, M. (eds) EUROCRYPT 2023. LNCS, vol. 14008, pp. 532-563. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-30589-4_19
15. Cayrel P.L., Véron P., El Yousfi Alaoui S.M.: A zero-knowledge identification scheme based on the q-ary syndrome decoding problem. In: Biryukov, A., Gong, G., Stinson, D.R. (eds) SAC 2010. LNCS, vol. 6544, pp. 171-186. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-19574-7_12
16. Chiani, M., Dardari, D., Simon, M.K.: New exponential bounds and approximations for the computation of error probability in fading channels. IEEE Trans. Wireless Commun. **2**(4), 840-845 (2003)
17. Courtois, N.T., Finiasz, M., Sendrier, N.: How to achieve a McEliece-based digital signature scheme. In: Boyd, C. (ed.) ASIACRYPT 2001. LNCS, vol. 2248, pp. 157–174. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45682-1_10
18. D'Alconzo, G., Meneghetti, A., Piasenti, P.: Security issues of CFS-like digital signature algorithms. arXiv preprint arXiv:2112.00429, 2021. https://arxiv.org/abs/2112.00429
19. Debris-Alazard, T., Sendrier, N., Tillich, J. P.: Wave: a new family of trapdoor one-way preimage sampleable functions based on codes. In: Galbraith, S.D., Moriai, S. (eds.) ASIACRYPT 2019. LNCS, vol. 11921, pp. 21–51. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-34578-5_2
20. Dumer, I.: On minimum distance decoding of linear codes. In: Proceedings of the 5th Joint Soviet-Swedish International Workshop Information Theory, pp. 50-52 (1991)
21. Esser, A., Bellini, E.: Syndrome decoding estimator. In: Hanaoka, G., Shikata, J., Watanabe, Y. (eds.) PKC 2022. LNCS, vol. 13177, pp. 112-141. Springer, Heidelberg (2022). https://doi.org/10.1007/978-3-030-97121-2_5
22. Feneuil, T., Joux, A., Rivain, M.: Syndrome decoding in the head: shorter signatures from zero-knowledge proofs. In: Dodis, Y., Shrimpton, T. (eds) CRYPTO 2022. LNCS, vol. 13508, pp. 541-572. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15979-4_19
23. Fiat, A., Shamir, A.: How to prove yourself: practical solutions to identification and signature problems. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 186–194. Springer, Heidelberg (1987). https://doi.org/10.1007/3-540-47721-7_12
24. Kabatianskii, G., Krouk, E., Smeets, B.: A digital signature scheme based on random error-correcting codes. In: Darnell, M. (ed.) Cryptography and Coding 1997. LNCS, vol. 1355, pp. 161-167. Springer, Heidelberg (1997). https://doi.org/10.1007/BFb0024461
25. Lau, T.S.C., Tan, C.H., Prabowo, T.F.: Key recovery attacks on some rank metric code-based signatures. In: Albrecht, M. (ed.) IMACC 2019. LNCS, vol. 11929, pp. 215-235. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-35199-1_11

26. Lee, P.J., Brickell, E.F.: An observation on the security of McEliece's public-key cryptosystem. In: Barstow, D., et al. (eds.) EUROCRYPT 1988. LNCS, vol. 330, pp. 275-280. Springer, Heidelberg (1988). https://doi.org/10.1007/3-540-45961-8_25
27. Leon, J.: A probabilistic algorithm for computing minimum weights of large error-correcting codes. In: IEEE Trans. Inform. Theory 34.5 (1988), pp. 1354–1359
28. May, A., Meurer, A., Thomae, E.: Decoding random linear codes in $\tilde{O}(2^{0.054n})$. In: Lee, D.H., Wang, X. (eds.) ASIACRYPT 2011. LNCS, vol. 7073, pp. 107-124. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25385-0_6
29. May, A., Ozerov, I.: On computing nearest neighbors with applications to decoding of binary linear codes. In: Oswald, E., Fischlin, Marc (eds.) EUROCRYPT 2015. LNCS, vol. 9056, pp. 203-228. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46800-5_9
30. McEliece, R.: A public-key cryptosystem based on algebraic coding theory. Jet Propulsion Lab., Pasadena, CA, USA, DSN Progress Rep. 42–44, pp. 114–116, 1978.
31. Meneghetti, A., Picozzi, C., Tognolini, G.: A Post-Quantum Digital Signature Scheme from QC-LDPC Codes. IACR Cryptology ePrint Archive 2022/1477 (2022). https://eprint.iacr.org/2022/1477
32. Sendrier, N.: Wave Parameter Selection. IACR Cryptology ePrint Archive 2023/588 (2023). https://eprint.iacr.org/2023/588
33. Niederreiter, H.: Knapsack-type cryptosystems and algebraic coding theory. Probl. Control Inf. Theory, vol. 15, no. 2, pp. 159–166, 1986.
34. Otmani, A., Tillich, J. P.: An efficient attack on all concrete KKS proposals. In: Yang, B.-Y. (ed.) PQCrypto 2011. LNCS, vol. 7071, pp. 98-116. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25405-5_7
35. Prabowo, T.F., Tan, C.H.: Attack on a Code-Based Signature Scheme from QC-LDPC Codes. In: El Hajji, S., Mesnager, S., Souidi, E.M. (eds.) Codes, Cryptology and Information Security (C2SI 2023). LNCS, vol. 13874, pp. 136-149. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33017-9_9
36. Prange, E.: The use of information sets in decoding cyclic codes, IRE Transactions on Information Theory, Vol. 8, No. 5, pp. 5-9, 1962.
37. Ren, F., Zheng, D., Wang, W.: An efficient code based digital signature algorithm. Int. J. Netw. Secur., 19(6):1072-1079 (2017). https://doi.org/10.6633/IJNS.201711.19(6).24
38. Sendrier, N.: Decoding one out of many. In: Yang, B.-Y. (ed.) PQCrypto 2011. LNCS, vol. 7071, pp. 51-67. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25405-5_4
39. Song, Y., Huang, X., Mu, Y., Wu, W., Wang, H.: A code-based signature scheme from the lyubashevsky framework, Theoretical Computer Science 835 (2020) 15–30. https://doi.org/10.1016/j.tcs.2020.05.011
40. Stern, J.: A method for finding codewords of small weight. In: Cohen, G., Wolfmann, J. (eds.) Coding Theory 1988. LNCS, vol. 388, pp. 106-113. Springer, Heidelberg (1989). https://doi.org/10.1007/BFb0019850
41. Stern, J.: A new identification scheme based on syndrome decoding. In: Stinson, D.R. (ed.) CRYPTO 1993. LNCS, vol. 773, pp. 13–21. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-48329-2_2
42. Canto Torres, R., Sendrier, N.: Analysis of information set decoding for a sub-linear error weight. In: Takagi, T. (ed.) PQCrypto 2016. LNCS, vol. 9606, pp. 144-161. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-29360-8_10
43. Vasseur, V.: Post-quantum cryptography: a study of the decoding of QC-MDPC codes. Ph.D. thesis, Université de Paris (2021)