# Compact-LWE-MQ ${ }^{\text {H }}$ : Public Key Encryption without Hardness Assumptions 

Dongxi Liu ${ }^{\star} \quad$ Surya Nepal<br>CSIRO Data61, Australia


#### Abstract

Modern public key encryption relies on various hardness assumptions for its security. Hardness assumptions may cause security uncertainty, for instance, when a hardness problem is no longer hard or the best solution to a hard problem might not be publicly released. In this paper, we propose a public key encryption scheme Compact-LWE-MQ ${ }^{\text {H }} 1$ to demonstrate the feasibility of constructing public key encryption without relying on hardness assumptions. Instead, its security is based on problems that are called factually hard.The two factually hard problems we propose in this work are stratified system of linear and quadratic equations, and layered learning with relatively big errors. The factually hard problems are characterized by their layered structures, which ensure that the secrets at a lower layer can only be recovered after the secrets in a upper layer have been found correctly (i.e., leading to consistent lower layer secrets, not necessarily the original upper layer ones). On the other hand, without knowing the secrets in the lower layer, the upper layer subproblem can only be solved by exhaustive search. Based on the structure of factually hard problems, we prove that without bruteforce search the adversary cannot recover plaintexts or correct private key, and then discuss CPA-security and CCA-security of Compact-LWE-MQ ${ }^{\mathrm{H}}$. We have implemented Compact-LWE-MQ ${ }^{\mathrm{H}}$ with a number of lines of SageMath code. Simplicity of Compact-LWE-MQ ${ }^{\mathrm{H}}$ makes it easy for understanding, cryptanalysis, and implementation. In our configuration for 128 -bit security, the dimensional parameter is $n=4$ ( $n$ has the same meaning as in LWE). For such a tiny parameter, the current analysis tools like LLL lattice reduction algorithm are already efficient enough to perform attacks if the security claim of Compact-LWE-MQ ${ }^{H}$ does not hold. That is, the security of Compact-LWE-MQ ${ }^{\mathrm{H}}$ is not assumed with the capability of cryptanalysis tools. SageMath code of verifying Compact-LWE-MQ ${ }^{\mathrm{H}}$ security is also included in Appendix.


## 1 Introduction

Public key encryption is usually constructed around various assumptions of computationally hard problems as its security foundation. The well-known hardness assumptions include the problems of factoring the product of big prime numbers, computing

[^0]discrete logarithms over finite fields, multivariate quadratic problems [9], and the problems in lattice, such as short integer solution problem and shortest vector problem [1].

When there are new algorithms that can improve the efficiency of solving hard problems, the relevant public key encryption schemes need to increase their key length as a countermeasure. On the other hand, if a new computing paradigm (e.g., quantum computing) makes hardness assumptions no longer hold, new public key encryption schemes have to be designed, as being done by NIST Post-Quantum Cryptography Standardization project.

In this paper, we demonstrate the feasibility of designing public key encryption based on hardness facts, rather than assumptions. To this end, we design the public key encryption scheme Compact-LWE-MQ ${ }^{\mathrm{H}}$, which has its security based on the following factually hard problems.

### 1.1 Factual Hardness Problems

The factually hard problems described below have layered structures, with secrets contained in each layer. Any attempt to find secret values in a lower layer must be preceded by necessary operations to correctly find secret values in a upper layer. The correct condition means that the candidate secret values from the upper layer must lead to consistent secret values in a lower layer.

Without knowing the secrets in the lower layer, the upper layer subproblem can only be solved by exhaustively search in factually hard problems. Even configured with very small parameters, a factually hard problem still needs to be solved with brute-force search.

To define the problems below, we let $h, h^{\prime}$, and $p$ be three prime numbers, and $m$, $n$, a_max three integers. We require $q>m * p *\left(h^{\prime}+h\right), h>m * p * p, h^{\prime}>m * p * p$, $m>2 * n$, and $p>a \_m a x$.
1.1.1 Layered Learning with Relatively Big Errors Let $s$ and $\mathbf{s}^{\prime}$ be uniformly sampled at random from $\mathbb{Z}_{q}^{n}, \mathbf{z}$ be uniformly sampled at random from $\mathbb{Z}_{h}^{n}$, and $\mathbf{z}^{\prime}$ be uniformly sampled at random from $\mathbb{Z}_{h^{\prime}}^{n}$. For $0 \leq i \leq m-1$, we can have $m$ samples:

$$
\left(\mathbf{a}_{i}, \mathbf{a}_{i}^{\prime}, b_{i}=\left\langle\mathbf{s}, \mathbf{a}_{i}\right\rangle+r_{i} \bmod q, b_{i}^{\prime}=\left\langle\mathbf{s}^{\prime}, \mathbf{a}_{i}^{\prime}\right\rangle+r_{i}^{\prime} \bmod q\right)_{i}
$$

where $\mathbf{a}_{i}$ and $\mathbf{a}^{\prime}{ }_{i}$ are uniformly sampled from $\mathbb{Z}_{a_{-} \max }^{n}$, and $r_{i} \in \mathbb{Z}_{h}$ and $r_{i}^{\prime} \in \mathbb{Z}_{h^{\prime}}$ are defined as,

$$
r_{i}=u_{i}+\left\langle\mathbf{z}, \mathbf{a}_{i}\right\rangle \bmod h, r_{i}^{\prime}=u_{i}+\left\langle\mathbf{z}^{\prime}, \mathbf{a}_{i}^{\prime}\right\rangle \bmod h^{\prime},
$$

where $u_{i}$ is uniformly sampled from $\mathbb{Z}_{p}$. In the samples, $\mathbf{a}_{i}, \mathbf{a}^{\prime}{ }_{i}, b_{i}$, and $b_{i}^{\prime}$ are public values. Finding the candidate values of $\mathbf{z}$ and $\mathbf{z}^{\prime}$ must be preceded by the operations to find the values of $\mathbf{s}, \mathbf{s}^{\prime}, h$ and $h^{\prime}$.

Compared with Learning with Errors (LWE) [8], the elements in $\mathbf{a}_{i}$ and $\mathbf{a}^{\prime}{ }_{i}$ here can be much smaller than $r_{i}$ and $r_{i}^{\prime}$. Thus, given proper parameters ( $m, p$, and $a \_m a x$ ), the solutions to the noised equations below (i.e., $s$ and $s^{\prime}$ ) will not necessarily be unique.

$$
b_{i} \approx\left\langle\mathbf{s}, \mathbf{a}_{i}\right\rangle \bmod q, b_{i}^{\prime} \approx\left\langle\mathbf{s}^{\prime}, \mathbf{a}^{\prime}{ }_{i}\right\rangle \bmod q
$$

This fact can be verified with experiments. Intuitively, since $\mathbf{a}_{i}[0]$ is small, $r_{i}+\mathbf{a}_{i}[0]$ can be a legitimate error term (i.e., its value is in $\mathbb{Z}_{h}$ ) for $0 \leq i \leq m-1$ with a high chance; thus $\tilde{\mathbf{s}}$ is a valid solution, where $\tilde{\mathbf{s}}[0]=\mathbf{s}[0]-1$ and $\tilde{\mathbf{s}}[j]=\mathbf{s}[j]$ for $1 \leq j \leq n-1$.

Correct values of $\mathbf{z}$ and $\mathbf{z}^{\prime}$ are required to keep $u_{i}$ the same in $r_{i}$ and $r_{i}^{\prime}$ in each sample. Note that we do not require the secret values must be the same as the original ones for correctness. To find a candidate value for $\mathbf{z}$ and $\mathbf{z}^{\prime}$, the adversary must first have a solution to $\mathbf{s}$ and $\mathbf{s}^{\prime}$, which leads to some values for $r_{i}$ and $r_{i}^{\prime}$ (not necessarily the original ones). And then, $h$ and $h^{\prime}$ have to be guessed to proceed. The experiment code in Appendix confirms that an arbitrary guess of $h$ and $h^{\prime}$ cannot fulfill the correctness requirement of $\mathbf{z}$ and $\mathbf{z}^{\prime}$; so exhaustive search of moduli $h$ and $h^{\prime}$ is necessary.

Even worse for the adversary, there might be no $h$ and $h^{\prime}$ that can generate correct $\mathbf{z}$ and $\mathbf{z}^{\prime}$ for some solutions to $\mathbf{s}$ and $\mathbf{s}^{\prime}$. For instance, this is the case, where $\mathbf{s}$ is $\tilde{\mathbf{s}}$ as defined above and $\mathbf{s}^{\prime}$ is the original one.
1.1.2 Stratified System of Linear and Quadratic Equations Let $x_{i}(0 \leq i \leq$ $m-1$ ) and $v$ all be random numbers uniformly sampled from $\mathbb{Z}_{p}$, and let $b_{i}$ and $b_{i}^{\prime}$ be defined as above. Then, we define the following linear equations, where $c b, c b^{\prime}, b_{i}$, and $b_{i}^{\prime}$ are public values,

$$
\begin{aligned}
c b & =\sum_{i=0}^{m-1} x_{i} * b_{i} \bmod q \\
c b^{\prime} & =\sum_{i=0}^{m-1} x^{\prime}{ }_{i} * b_{i}^{\prime}{ }_{i} \bmod q
\end{aligned}
$$

and $m$ multivariate quadratic (MQ) equations,

$$
x^{\prime}{ }_{i}=v * x_{i} \bmod p .
$$

In this problem, $x_{i}$ and $x_{i}^{\prime}(0 \leq i \leq m-1)$ are secret values in the upper layer, and $v$ is the secret value in the lower layer.

The adversary aims to recover $v$ from the above equations. Note that $v$ only appears in the $m$ quadratic equations in the lower layer. To recover $v$ without guessing it, $x_{i}$ and $x_{i}^{\prime}(0 \leq i \leq m-1)$ must be determined and their values must lead to consistent $v$, since each pair of $x_{i}$ and $x_{i}^{\prime}$ can calculate a value for $v$.

We call the system with above equations stratified, because the values of $x_{i}$ and $x_{i}^{\prime}$ (or at least one of them) can only be obtained by solving the corresponding linear equations, and then these values are used to check whether a correct $v$ can be determined from the quadratic equations. The method of solving a linear system for the values of some variables and then determining the values of other variables in quadratic terms is also taken by all existing algorithms to solve multivariate quadratic (MQ) equations, such as [3, 4].

The linear equations in our stratified system are underdetermined, thus allowing a solution space for $x_{i}$ and $x_{i}^{\prime}$. As a result, recovering $v$ in this stratified system is factually hard, since each solution to $x_{i}$ and $x_{i}^{\prime}$ in the solution space has to be found and checked. In Compact-LWE-MQ ${ }^{\mathrm{H}}, v$ is generalized into a vector.

### 1.2 Overview of Compact-LWE-MQ ${ }^{\mathrm{H}}$

Compact-LWE-MQ ${ }^{\mathrm{H}}$ is constructed around the above hardness facts, with further security enhancement. Briefly, the first hardness fact ensures a correct private key cannot be recovered from the public key, and the second prevents efficient attacks of recovering plaintexts from ciphertexts and the public key.

In Compact-LWE-MQ ${ }^{\mathrm{H}}$, the private key includes a number of extra secret components, in addition to those listed in the first factually hard problem. First, another lower layer of samples are added into $r_{i}$ and $r_{i^{\prime}}$, as to be detailed in the construction later. Second, $r_{i}$ and $r_{i}^{\prime}$ are multiplied with secret values in $\mathbb{Z}_{q}$ before they are added with the corresponding inner products.

The plaintext in Compact-LWE-MQ ${ }^{\mathrm{H}}$ is a $n^{\prime}$-dimensional vector $\mathbf{v} \in \mathbf{Z}_{p}^{n^{\prime}}$. The minimum of $n^{\prime}$ is determined by the expected security level and distribution of plaintexts. Compact-LWE-MQ ${ }^{\mathrm{H}}$ achieves CPA-security when $\mathbf{v}$ is random enough for the adversary. After being slightly revised, Compact-LWE-MQ ${ }^{\mathrm{H}}$ is also CCA-secure.

In the security proof of Compact-LWE-MQ ${ }^{\mathrm{H}}$, we identify the necessary operations (or steps) that have to be taken by the adversary to manipulate the public keys and ciphertexts. The factually hard problems underlying Compact-LWE-MQ ${ }^{H}$ ensure that the adversary has to do brute-force search at these necessary steps, enforcing by their layered structure.

The construction of Compact-LWE-MQ ${ }^{H}$ is conceptually simple. The following concepts used in existing post-quantum public key encryption algorithms are not needed to understand Compact-LWE-MQ ${ }^{\mathrm{H}}$ : polynomials, complex probability distributions or sampling techniques, complex coding/decoding, hashing. The basic knowledge of modular arithmetic, matrix and vector operations (i.e., inner product, inverse of matrix, multiplication of matrix and vector), and uniform distribution is sufficient to understand and analyze Compact-LWE-MQ ${ }^{\mathrm{H}}$ for a large cohort of people. The selection of parameters is also straightforward for a desired security level.

Compact-LWE-MQ ${ }^{\mathrm{H}}$ is straightforward to implement due to its simple construction. We have implemented Compact-LWE-MQ ${ }^{\mathrm{H}}$ in SageMath and the implementation is in Appendix. In a configuration for 128-bit security, we have parameters $m=24$ and $n=4$ (with the same meaning as in LWE), and $p$ with 128 bits. In the configuration, the size of the public key is 3708 -byte and a ciphertext is 574 -byte long. Note that performance is not a focus of Compact-LWE-MQ ${ }^{\mathrm{H}}$ in this paper.

Since $m$ and $n$ take tiny values, the existing cryptanalysis tools like LLL algorithm are already efficient enough to completely attack Compact-LWE-MQ ${ }^{H}$ in a short period of time if Compact-LWE-MQ ${ }^{\mathrm{H}}$ has a structure flaw. In addition, any suspicion to the security of Compact-LWE-MQ ${ }^{\mathrm{H}}$ can also be checked with the existing cryptanalysis tools. The SageMath code of verifying the security claims relevant to both factually hard problems is included in Appendix.

### 1.3 Notations

Given a positive integer $q, \mathbb{Z}_{q}$ refers to the set $\{0, \ldots, q-1\}$. The $n$-ary Cartesian product of $\mathbb{Z}_{q}$ is denoted by $\mathbb{Z}_{q}^{n}$. For a finite set, e.g. $S, x \leftarrow S$ means that $x$ is uniformly sampled from $S$ at random.

A lower-case boldface letter denotes a vector or a list (e.g., s), while an upper-case boldface letter (e.g., L) is used to represent a matrix (or a list of lists). Given two vectors with the same dimension, such as $\mathbf{a}$ and $\mathbf{s}$, their inner product is denoted by $\langle\mathbf{a}, \mathbf{s}\rangle$.

The $i$ th element of list $\mathbf{a}$ is written as $\mathbf{a}[i]$ and $i$ starts from 0 . The sublist of list $\mathbf{a}$ from its $i$ th element to $(j-1)$ th element, inclusively, is denoted by $\mathbf{a}[i: j]$. If $j \leq i$, then $\mathbf{a}[i]$ returns an empty list [] . Two vectors $\mathbf{a}$ and $\mathbf{b}$ are added by $\mathbf{a}+\mathbf{b}$. The concatenation of two lists $\mathbf{a}$ and $\mathbf{b}$ is denoted by $\mathbf{a} \| \mathbf{b}$. An element $x$ is appended to the tail of list $\mathbf{a}$ by $\mathbf{a}+x$. The length of list $\mathbf{a}$ is denoted by len(a). The prime number just after an integer $n$ is denoted by $\mathrm{np}(n)$. 0 means a zero vector, with its dimension derived from its usage context.

## 2 Construction of Compact-LWE-MQ ${ }^{H}$

As a public-key encryption scheme, Compact-LWE-MQ ${ }^{\mathrm{H}}$ is defined with three algorithms: key generation, encryption, and decryption. Compact-LWE-MQ ${ }^{H}$ partially shares its name with Compact-LWE [7], which has been broken. However, they are completely different on each algorithm in the schemes. Particularly, Compact-LWE (and other hardness assumption) does not have layered structure in their problems. The only similarity with Compact-LWE is that a noise in a sample can be bigger than elements in $\mathbf{a}_{i}$.

Compact-LWE-MQ ${ }^{\mathrm{H}}$ is parameterized with four integers: $n, m, p$, and $a \_m a x$. We require $p$ be a prime, $a \_\max <p, 2 * n<m<p$. Another integer parameter $q$ will be generated by the key generation algorithm. These public parameters are represented by $p p$. They are taken as implicit inputs by the encryption and decryption algorithms.

### 2.1 Key Generation

The key generation algorithm generates the private key $\mathbf{S K}$ and the public key $\mathbf{P K}$, denoted by $\operatorname{gen}(p p)=(\mathbf{S K}, \mathbf{P K})$. The key generation algorithm is comprised of private key generation and public key generation.
2.1.1 Private Key A private key $\mathbf{S K}$ is a tuple ( $\mathbf{s}, \mathbf{k}, \mathbf{t}, \mathbf{z}, s, h, k, \mathbf{s}^{\prime}, \mathbf{k}^{\prime}, \mathbf{t}^{\prime}, \mathbf{z}^{\prime}, s^{\prime}, h^{\prime}, k^{\prime}$ ), generated in the steps below.

- $h=\mathrm{np}(m * p * p+r)$, where $r \leftarrow \mathbb{Z}_{p}$.
- $h^{\prime}=\operatorname{np}\left(m * p * p+r^{\prime}\right)$, where $r^{\prime} \leftarrow \mathbb{Z}_{p}$.
- $q=m * p *\left(h+h^{\prime}\right)+q^{\prime}$, where $q^{\prime} \leftarrow \mathbb{Z}_{p}$.
$-\mathrm{s} \leftarrow \mathbb{Z}_{q}^{n}, \mathbf{k} \leftarrow \mathbb{Z}_{p}^{n}, \mathbf{t} \leftarrow \mathbb{Z}_{p}^{n}, \mathrm{~s}^{\prime} \leftarrow \mathbb{Z}_{q}^{n}, \mathbf{k}^{\prime} \leftarrow \mathbb{Z}_{p}^{n}$, and $\mathbf{t}^{\prime} \leftarrow \mathbb{Z}_{p}^{n}$.
$-\mathbf{z} \leftarrow \mathbb{Z}_{h}^{n}$, and $\mathbf{z}^{\prime} \leftarrow \mathbb{Z}_{h^{\prime}}^{n}$.
$-s \leftarrow \mathbb{Z}_{q}$ and $s^{\prime} \leftarrow \mathbb{Z}_{q}$, ensuring
- $s$ and $s$ are co-prime with $q$.
$-k \leftarrow \mathbb{Z}_{p}, k^{\prime} \leftarrow \mathbb{Z}_{p}$, ensuring
- $k \neq 0$, and $k^{\prime} \neq 0$.

Note that $q$ generated above is a public parameter and it is randomized with a number sampled from $\mathbb{Z}_{p}$; additionally, we require $n * \log _{2}\left(a \_m a x\right)+\log _{2}(p * p)>\log _{2}(q)$. All secret values in $\mathbf{S K}$ and random values used below are randomly sampled from uniform distributions. Given SK, we write $\widehat{\mathbf{S K}}$ for $\left(\mathbf{s}^{\prime}, \mathbf{k}^{\prime}, \mathbf{t}^{\prime}, \mathbf{z}^{\prime}, s^{\prime}, h^{\prime}, k^{\prime}, \mathbf{s}, \mathbf{k}, \mathbf{t}, \mathbf{z}, s, h, k\right)$.
2.1.2 Public Key Based on the private key $\mathbf{S K}$, the algorithm gen generates the corresponding public key $\mathbf{P K}$, which is comprised of $m$ samples. The $i$ th sample, denoted $\left(\mathbf{a}_{i}, b_{i}, \mathbf{a}_{i}^{\prime}, b_{i}^{\prime}\right)$, is defined as:

$$
\begin{aligned}
& -\mathbf{a}_{i} \leftarrow \mathbb{Z}_{a \_m a x}^{n} \text { and } \mathbf{a}_{i}^{\prime} \leftarrow \mathbb{Z}_{a \_m a x}^{n} \\
& -u_{i} \leftarrow \mathbb{Z}_{p} \\
& -b_{i}=\left\langle\mathbf{a}_{i}, \mathbf{s}\right\rangle+s * r_{i} \bmod q, \text { where }
\end{aligned}
$$

$$
\text { - } r_{i}=\left(\left\langle\mathbf{a}_{i}, \mathbf{k}\right\rangle+\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{t}\right\rangle+k * u_{i} \bmod p\right)+\left\langle\mathbf{a}_{i}, \mathbf{z}\right\rangle \bmod h
$$

$$
-b_{i}^{\prime}=\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{s}^{\prime}\right\rangle+s^{\prime} * r_{i}^{\prime} \bmod q, \text { where }
$$

$$
\text { - } r_{i}^{\prime}=\left(\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{k}^{\prime}\right\rangle+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle+k^{\prime} * u_{i} \bmod p\right)+\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{z}^{\prime}\right\rangle \bmod h^{\prime}
$$

As shown above, $\mathbf{P K}$ is defined by embedding one layer of Learning with Relatively Big Error samples into another; so it has a layered structure. The top layer is $\left(\mathbf{a}_{i}, \mathbf{a}^{\prime}{ }_{i}, b_{i}, b_{i}^{\prime}\right)_{i}$, and the bottom layer is $\left(\mathbf{a}_{i}, \mathbf{a}^{\prime}{ }_{i}, r_{i}, r_{i}^{\prime}\right)_{i}$. The bottom layer is a hidden layer, in terms that the moduli (i.e., $h$ and $h^{\prime}$ ), $r_{i}$, and $r_{i}^{\prime}$ are not known publicly. Without guessing the hidden values $\left(h, h^{\prime}, r_{i}\right.$, and $r_{i}^{\prime}$ ), the adversary does not have any way to recover $\mathbf{z}$ and $\mathbf{z}^{\prime}$. Other secrets $\mathbf{k}, \mathbf{k}^{\prime}, \mathbf{t}, \mathbf{t}^{\prime}, k, k^{\prime}$ are hidden even deeper in this structure.

The notation $\widehat{\mathbf{P K}}$ is used in in the encryption algorithm below. It has the same number of samples as $\mathbf{P K}$, and for each $\left(\mathbf{a}, b, \mathbf{a}^{\prime}, b^{\prime}\right) \in \mathbf{P K},\left(\mathbf{a}^{\prime}, b^{\prime}, \mathbf{a}, b\right) \in \widehat{\mathbf{P K}}$.

### 2.2 Encryption

A plaintext $\mathbf{v}$ in Compact-LWE-MQ ${ }^{\mathrm{H}}$ is an element in $\mathbb{Z}_{p}^{n^{\prime}}$, where $n^{\prime}$ is an odd integer when bigger than two. The minimum of $n^{\prime}$ is relevant to the expected security level and the size of $p$. There is no upper bound on $n^{\prime}$. That is, the size of $\mathbf{v}$ can be incremented for longer messages. We will discuss the choice of $n^{\prime}$ later when we analyze the security of Compact-LWE-MQ ${ }^{\mathrm{H}}$.

The encryption algorithm enc is defined in Algorithm 1. With the public key PK, this algorithm encrypts a message $\mathbf{v} \in \mathbb{Z}_{p}^{n^{\prime}}$ into a ciphertext $\mathbf{c}$, denoted $\mathbf{e n c}(\mathbf{P K}, \mathbf{v})=$ $\mathbf{c}$. The ciphertext $\mathbf{c}$ is comprised of $n^{\prime}$ components, each of which is generated with the algorithm encS in Algorithm 2.

In the enc algorithm, there are $n^{\prime}$ vectors each randomly sampled from $\mathbb{Z}_{p}^{m}$ and stored in $\mathbf{L}$, which is then passed as a parameter to encS. The algorithm encS is invoked $n^{\prime}$ times; at the $i$ th time $\left(0 \leq i \leq n^{\prime}-1\right)$, the $i$ th ciphertext component $c$ is generated and appended to the ciphertext $\mathbf{c}$.

In the encS algorithm, two vectors $\mathbf{l}$ and $\mathbf{l}^{\prime}$ are derived from $\mathbf{L}$ : $\mathbf{l}$ is simply $\mathbf{L}[i]$ and $\mathbf{1}^{\prime}$ is obtained by adding vectors $\mathbf{L}\left[(i+j) \bmod n^{\prime}\right] * \mathbf{v}[j]$ from $j=0$ to $j=$ $n^{\prime}-1$ over modulus $p$. Then, each sample in $\mathbf{P K}{ }^{\prime}$ is multiplied by the corresponding

```
Algorithm 1: enc for encryption
    input : PK, v
    output: c
    Let \(\mathbf{L}\) be a list with \(n^{\prime}\) entries.
    for \(i=0\) to \(n^{\prime}-1\) do
            \(\mathbf{L}[i] \leftarrow \mathbb{Z}_{p}^{m}\)
    end
    \(\mathbf{c}=[]\)
    for \(i=0\) to \(n^{\prime}-1\) do
        if \(i \bmod 2=0\) then
            \(c=\operatorname{encS}(\mathbf{P K}, \mathbf{L}, \mathbf{v}, i)\)
        end
        else
            \(c=\operatorname{encS}(\widehat{\mathbf{P K}}, \mathbf{L}, \mathbf{v}, i)\)
        end
        \(\mathbf{c}=\mathbf{c}+c\)
    end
    return \(\mathbf{c}\)
```

```
Algorithm 2: encS for encrypting ciphertext component
    input : \(\mathbf{P K}^{\prime}, \mathbf{L}, \mathbf{v}, i\)
    output: \(c\)
    \(\mathbf{l}=\mathbf{L}[i]\)
    \(\mathbf{l}^{\prime}=\mathbf{0}\)
    for \(j=0\) to \(n^{\prime}-1\) do
        \(\mathbf{1}^{\prime}=\mathbf{1}^{\prime}+\mathbf{v}[j] * \mathbf{L}\left[(i+j) \bmod n^{\prime}\right] \bmod p\)
    end
    \(\mathbf{c a _ { 1 }}=\mathbf{0}, c b_{1}=0, \mathbf{c a}_{2}=\mathbf{0}, c b_{2}=0\)
    for \(j=0\) to \(m-1\) do
        Let \(\left(\mathbf{a}_{1}, b_{1}, \mathbf{a}_{2}, b_{2}\right)=\mathbf{P K}^{\prime}[j]\).
        \(\mathbf{c} \mathbf{a}_{1}=\mathbf{c} \mathbf{a}_{1}+\mathbf{l}[j] * \mathbf{a}_{1}\)
        \(c b_{1}=c b_{1}+\mathbf{1}[j] * b_{1} \bmod q\)
        \(\mathbf{c a}_{2}=\mathbf{c a}_{2}+\mathbf{l}^{\prime}[j] * \mathbf{a}_{2}\)
        \(c b_{2}=c b_{2}+\mathbf{l}^{\prime}[j] * b_{2} \bmod q\)
    end
    4 4 \(=\left(\mathbf{c a}_{1}, c b_{1}, \mathbf{c a}_{2}, c b_{2}\right)\)
    15 return \(c\)
```

elements in $\mathbf{l}$ and $\mathbf{l}^{\prime}$, and all such samples are summed up as the ciphertext component $c=\left(\mathbf{c a}_{1}, c b_{1}, \mathbf{c a}_{2}, c b_{2}\right)$.

Depending on whether $n^{\prime}$ is even or odd, the public key passed to encS is different. Particularly, $\mathbf{P K}^{\prime}$ is $\mathbf{P K}$ when $n^{\prime}$ is even, and $\widehat{\mathbf{P K}}$ otherwise. Suppose $n^{\prime}=2$. Then, let the two ciphertext components in $\mathbf{c}$ be $c_{0}=\left(\mathbf{c a}_{1}, c b_{1}, \mathbf{c a}_{2}, c b_{2}\right)$ and $c_{1}=$ $\left(\mathbf{c a}_{1}^{\prime}, c b_{1}^{\prime}, \mathbf{c a}_{2}^{\prime}, c b_{2}^{\prime}\right)$. Since $\mathbf{P K}$ and $\widehat{\mathbf{P K}}$ are used for $c_{0}$ and $c_{1}$, respectively, $c b_{1}$ and $c b_{1}^{\prime}$ (or $c b_{2}$ and $c b_{2}^{\prime}$ ) cannot be homomorphically added together; otherwise, values from $\mathbf{Z}_{h}$ and $\mathbf{Z}_{h^{\prime}}$ are mixed.

```
Algorithm 3: dec for decryption
    input : SK, c
    output: \(\mathbf{v}\) or -1
    \(n^{\prime}=\operatorname{len}(\mathbf{c})\)
    \(\mathbf{g}=[], \mathbf{y}=[]\)
    for \(i=0\) to \(n^{\prime}-1\) do
        if \(i \bmod 2=0\) then
            \(g, y=\operatorname{decS}(\mathbf{S K}, \mathbf{c}[i])\)
        end
        else
            \(g, y=\operatorname{decS}(\widehat{\mathbf{S K}}, \mathbf{c}[i])\)
        end
        \(\mathbf{g}=\mathbf{g}+g\)
        \(\mathbf{y}=\mathbf{y}+y\)
    end
    Let \(\mathbf{G}\) be a \(n^{\prime} * n^{\prime}\) matrix.
    for \(i=0\) to \(n^{\prime}-1\) do
        \(\mathbf{G}[i]=\mathbf{g}\left[i: n^{\prime}\right] \| \mathbf{g}[0: i]\)
    end
    if \(\mathbf{G}\) is singular with repect to \(p\) then
        return -1
    end
    \(\mathbf{v}=\mathbf{G}^{-1} * \mathbf{y} \bmod p\)
    return \(\mathbf{v}\)
```

When $\mathbf{v}$ is a zero vector, each ciphertext component in $\mathbf{c}$ contains zero for the last two elements. For example, $\mathbf{v}=\mathbf{0}$ leads to $\mathbf{c a}_{2}=\mathbf{0}$ and $\mathbf{c b}_{2}=0$ in the above ciphertext component. Hence, before being passed to the encryption algorithm, v should be processed to ensure it is not a zero vector (e.g., by re-sampling some random elements or extending it to include more random elements).

### 2.3 Decryption

The decryption algorithm dec is defined in Algorithm 3. With the private key SK, the decryption algorithm recovers the value $\mathbf{v}$ from ciphertext $\mathbf{c}$ if $\mathbf{c}$ can be decrypted, denoted $\operatorname{dec}(\mathbf{S K}, \mathbf{c})=\mathbf{v}$; otherwise, the decryption algorithm returns -1 .

The ciphertext $\mathbf{c}$ is a list, consisting of ciphertext components. Let the length of $\mathbf{c}$ be $n^{\prime}$ (i.e., $n^{\prime}$ ciphertext components contained in $\mathbf{c}$ ). For each ciphertext component $\mathbf{c}[i]$, the dec algorithm invokes the algorithm decS, defined in Algorithm 4, to return a pair of integers $g$ and $y$ in $\mathbf{Z}_{p}$, which are then appended to two lists $\mathbf{g}$ and $\mathbf{y}$, respectively. Note that when $i$ is even, $\mathbf{S K}$ is passed to decS; otherwise, $\widehat{\mathbf{S K}}$ is used, corresponding to $\mathbf{P K}$ and $\widehat{\mathbf{P K}}$ in encryption.

After processing all ciphertext components, $\mathbf{g}$ and $\mathbf{y}$ each contains $n^{\prime}$ elements. From $\mathbf{g}$, the dec algorithm derives the $n^{\prime} * n^{\prime}$ matrix $\mathbf{G}$, which contains $\mathbf{g}$ as the first row and left rotations of $\mathbf{g}$ as the remaining rows, left rotating by one element each time for $n^{\prime}-1$ times.

If $\mathbf{G}$ is invertible with respect to $p$, the plaintext $\mathbf{v}$ is recovered by $\mathbf{G}^{-1} * \mathbf{y}$, where $\mathbf{G}^{-1}$ is the inverse of $\mathbf{G}$ with respect to modulus $p$. Otherwise, $\mathbf{c}$ cannot be decrypted and the decryption algorithm returns -1 .

```
Algorithm 4: decS for decrypting ciphertext component
    input : \(\mathbf{S K}^{\prime},\left(\mathbf{c a}_{1}, c b_{1}, \mathbf{c a}_{2}, c b_{2}\right)\)
    output: \(g, y\)
    Let \(\mathbf{S K}^{\prime}=\left(\mathbf{s}_{1}, \mathbf{k}_{1}, \mathbf{t}_{1}, \mathbf{z}_{1}, s_{1}, h_{1}, k_{1}, \mathbf{s}_{2}, \mathbf{k}_{2}, \mathbf{t}_{2}, \mathbf{z}_{2}, s_{2}, h_{2}, k_{2}\right)\).
    Let \(s_{1}^{-1} \in \mathbb{Z}_{q_{1}}\), such that \(s_{1}^{-1} * s_{1}=1 \bmod q\).
    Let \(s_{2}^{-1} \in \mathbb{Z}_{q_{2}}\), such that \(s_{2}^{-1} * s_{2}=1 \bmod q\).
    Let \(k_{1}^{-1} \in \mathbb{Z}_{p}\), such that \(k_{1}^{-1} * k_{1}=1 \bmod p\).
    Let \(k_{2}^{-1} \in \mathbb{Z}_{p}\), such that \(k_{2}^{-1} * k_{2}=1 \bmod p\).
    \(d=s_{1}^{-1} *\left(c b_{1}-\left\langle\mathbf{c a}_{1}, \mathbf{s}_{1}\right\rangle\right) \bmod q\)
    \(d=\left(d-\left\langle\mathbf{c a}_{1}, \mathbf{z}_{1}\right\rangle\right) \bmod h_{1}\)
    \(g=k_{1}^{-1} *\left(d-\left\langle\mathbf{c a}_{1}, \mathbf{k}_{1}\right\rangle\right)+\left\langle\mathbf{c a}_{1}, \mathbf{t}_{2} * k_{2}^{-1}\right\rangle \bmod p\)
    \(d^{\prime}=s_{2}^{-1} *\left(c b_{2}-\left\langle\mathbf{c} \mathbf{a}_{2}, \mathbf{s}_{2}\right\rangle\right) \bmod q\)
    \(d^{\prime}=\left(d^{\prime}-\left\langle\mathbf{c a}_{2}, \mathbf{z}_{2}\right\rangle\right) \bmod h_{2}\)
    \(y=k_{2}^{-1} *\left(d-\left\langle\mathbf{c a}_{2}, \mathbf{k}_{2}\right\rangle\right)+\left\langle\mathbf{c a}_{2}, \mathbf{t}_{1} * k_{1}^{-1}\right\rangle \bmod p\)
    return \(g, y\)
```

In Algorithm 4, we need the multiplicative inverses of $s_{1}, s_{2}, k_{1}$, and $k_{2}$. Their existence is guaranteed, since they are either co-prime with $q$ or required not being 0 in private key generation.

### 2.4 Correctness

Compact-LWE-MQ ${ }^{\mathrm{H}}$ is deterministically correct if the ciphertext can be decrypted. A cipherext cannot be decrypted when the determinant of $\mathbf{G}$ modulo $p$ is zero. Hence, the probability of failing to decrypt is $\frac{1}{p}$.

Theorem 1 (Correctness). Let $\operatorname{gen}(p p)=(\mathbf{S K}, \mathbf{P K}), \mathbf{v} \in \mathbb{Z}_{p}^{n^{\prime}}$ and $\mathbf{c}=\mathbf{e n c}(\mathbf{P K}, \mathbf{v})$. Then, $\operatorname{dec}(\mathbf{S K}, \mathbf{c})=\mathbf{v}$ with probability $\frac{p-1}{p}$, or $\operatorname{dec}(\mathbf{S K}, \mathbf{c})=-1$ with probability $\frac{1}{p}$.

Proof. For brevity, we prove the case where $n^{\prime}=2$. Let $\left(\mathbf{a}_{i}, b_{i}, \mathbf{a}_{i}^{\prime}, b_{i}^{\prime}\right)=\mathbf{P K}[i]$ for $0 \leq i \leq m-1$, and let $\mathbf{l}_{1}=\mathbf{L}[0], \mathbf{l}_{1}^{\prime}=\mathbf{v}[0] * \mathbf{L}[0]+\mathbf{v}[1] * \mathbf{L}[1] \bmod p, \mathbf{l}_{2}=\mathbf{L}[1]$, and $\mathbf{l}_{2}^{\prime}=\mathbf{v}[0] * \mathbf{L}[1]+\mathbf{v}[1] * \mathbf{L}[0] \bmod p$, where $\mathbf{L}$ is supposed to be sampled in encryption.

When $n^{\prime}=2$, $\mathbf{c}$ contains two ciphertext components: $\left(\mathbf{c a}_{1}, c b_{1}, \mathbf{c a}_{1}^{\prime}, c b_{1}^{\prime}\right)$, and $\left(\mathbf{c a}_{2}, c b_{2}, \mathbf{c a}_{2}^{\prime}, c b_{2}^{\prime}\right)$, where

$$
\begin{aligned}
\mathbf{c a}_{1} & =\sum_{i=0}^{m-1} \mathbf{l}_{1}[i] * \mathbf{a}_{i} \\
c b_{1} & =\sum_{i=0}^{m-1} \mathbf{l}_{1}[i] * b_{i} \bmod q \\
& =\sum_{i=0}^{m-1} \mathbf{l}_{1}[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{s}\right\rangle+s * r_{i}\right) \bmod q \\
\mathbf{c a}_{1}^{\prime} & =\sum_{i=0}^{m-1} \mathbf{l}_{1}{ }_{1}[i] * \mathbf{a}_{i}^{\prime} \\
c b^{\prime}{ }_{1} & =\sum_{i=0}^{m-1} \mathbf{l}^{\prime}[i] * b^{\prime}{ }_{i} \bmod q \\
& =\Sigma_{i=0}^{m-1} \mathbf{l}_{1}{ }_{1}[i] *\left(\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{s}^{\prime}\right\rangle+s^{\prime} * r^{\prime}{ }_{i}\right) \bmod q \\
\mathbf{c a}_{2} & =\Sigma_{i=0}^{m-1} \mathbf{l}_{2}[i] * \mathbf{a}_{i}^{\prime} \\
c b_{2} & =\sum_{i=0}^{m-1} \mathbf{l}_{2}[i] * b^{\prime}{ }_{i} \bmod q \\
& =\sum_{i=0}^{m-1} \mathbf{l}_{2}[i] *\left(\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{s}^{\prime}\right\rangle+s^{\prime} * r^{\prime}{ }_{i}\right) \bmod q \\
\mathbf{c a}^{\prime}{ }_{2} & =\sum_{i=0}^{m-1} \mathbf{l}_{2}[i] * \mathbf{a}_{i} \\
c b^{\prime}{ }_{2} & =\sum_{i=0}^{m-1} \mathbf{l}_{2}{ }_{2}[i] * b_{i} \bmod q \\
& =\sum_{i=0}^{m-1} \mathbf{l}_{2}^{\prime}[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{s}\right\rangle+s * r_{i}\right) \bmod q
\end{aligned}
$$

Due to $m * p *\left(h+h^{\prime}\right)<q, m * p * p<h$, and $m * p * p<h^{\prime}$, in Algorithm 4, we obtain

$$
\begin{aligned}
d_{1} & =\sum_{i=0}^{m-1} \mathbf{l}_{1}[i] * r_{i} \\
& =\sum_{i=0}^{m-1} \mathbf{l}_{1}[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{k}\right\rangle+\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{t}\right\rangle+k * u_{i}\right) \bmod p \\
d^{\prime}{ }_{1} & =\sum_{i=0}^{m-1} \mathbf{l}^{\prime}{ }_{1}[i] * r^{\prime}{ }_{i} \\
& =\sum_{i=0}^{m-1} \mathbf{1}^{\prime}{ }_{1}[i] *\left(\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{k}^{\prime}\right\rangle+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle+k^{\prime} * u_{i}\right) \bmod p \\
d_{2} & =\sum_{i=0}^{m-1} \mathbf{l}_{2}[i] * r^{\prime}{ }_{i} \\
& =\sum_{i=0}^{m-1} \mathbf{l}_{2}[i] *\left(\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{k}^{\prime}\right\rangle+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle+k^{\prime} * u_{i}\right) \bmod p \\
d^{\prime}{ }_{2} & =\sum_{i=0}^{m-1} \mathbf{1}^{\prime}{ }_{2}[i] * r_{i} \\
& =\sum_{i=0}^{m-1} \mathbf{l}^{\prime}{ }_{2}[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{k}\right\rangle+\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{t}\right\rangle+k * u_{i}\right) \bmod p
\end{aligned}
$$

Still in Algorithm 4, from $d_{1}, d_{1}^{\prime}, d_{2}$, and $d_{2}$, the following values can be recovered.

$$
\begin{aligned}
g_{1} & =\sum_{i=0}^{m-1} \mathbf{l}_{1}[i] *\left(\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{t}\right\rangle * k^{-1}+u_{i}+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{\prime-1}\right) \bmod p \\
y_{1} & =\sum_{i=0}^{m-1} \mathbf{1}^{\prime}{ }_{1}[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{-1}+u_{i}+\left\langle\mathbf{a}_{i}, \mathbf{t}\right\rangle * k^{-1}\right) \bmod p \\
& =\sum_{i=0}^{m-1}{ }^{\prime}\left(\mathbf{v}[0] * \mathbf{l}_{1}+\mathbf{v}[1] * \mathbf{l}_{2}\right)[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{\prime-1}+u_{i}+\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{t}\right\rangle * k^{-1}\right) \bmod p \\
g_{2} & =\sum_{i=0}^{m-1} \mathbf{l}_{2}[i] *\left(\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{\prime-1}+u_{i}+\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{t}\right\rangle * k^{-1}\right) \bmod p \\
y_{2} & =\sum_{i=0}^{m-1} \mathbf{1}^{\prime}{ }_{2}[i] *\left(\left\langle\mathbf{a}_{i}{ }_{i}, \mathbf{t}\right\rangle * k^{-1}+u_{i}+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{\prime-1}\right) \bmod p \\
& =\sum_{i=0}^{m-1}\left(\mathbf{v}[0] * \mathbf{l}_{2}+\mathbf{v}[1] * \mathbf{l}_{1}\right)[i] *\left(\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{t}\right\rangle * k^{-1}+u_{i}+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{\prime-1}\right) \bmod p
\end{aligned}
$$

Thus, we have $y_{1}=\mathbf{v}[0] * g_{1}+\mathbf{v}[1] * g_{2} \bmod p$, and $y_{2}=\mathbf{v}[0] * g_{2}+\mathbf{v}[1] * g_{1} \bmod p$. Hence, if the matrix $\mathbf{G}=\left[\begin{array}{ll}g_{1} & g_{2} \\ g_{2} & g_{1}\end{array}\right]$ is invertible with respect to $p$, the message $\mathbf{v}$ can be correctly recovered. The same proof applies when $n^{\prime}>2$; the only difference is that $\mathbf{G}$ becomes a bigger $n^{\prime} * n^{\prime}$ matrix.

## 3 Security Analysis

The security of Compact-LWE-MQ ${ }^{\mathrm{H}}$ is not based on computational hardness assumptions, so the conventional reduction-based proof methodology cannot be applied. Instead, we prove its security by analyzing the necessary steps the adversary has to take in order to derive plaintexts from ciphertexts or to derive secret key components from the public key. The necessary steps are enforced by the layered structures in our factually hard problems. Theorem 2 and Theorem 3 below will show that brute-force search is necessary in the identified necessary steps.

Note that the security claims below (Theorem 2 and Theorem 3) do not rely on the size of parameters ( $n, m, p$, and $\left.a \_m a x\right)$; small size of parameters just means a small space for brute-force search. Hence, by choosing small parameters, existing cryptanalysis tools can be used to verify these security claims. On the contrary, for existing schemes based on hardness assumption, their security cannot be verified when parameters are small, because at this case they can be solved without brute-force search (e.g., factoring the product of two small primes).

### 3.1 Attack to Ciphertexts

In this attack, the adversary tries to recover the plaintext $\mathbf{v}$ from the ciphertext $\mathbf{c}$ and the public key PK.

Theorem 2. Let $\operatorname{gen}(p p)=(\mathbf{S K}, \mathbf{P K})$, and $\mathbf{c}=\mathbf{e n c}(\mathbf{P K}, \mathbf{v})$, where $\mathbf{v} \in \mathbb{Z}_{p}^{n^{\prime}}$. Then, from $\mathbf{P K}$ and $\mathbf{c}$, the plaintext $\mathbf{v}$ can only be recovered by exhaustive search.

Proof. For brevity, we take $n^{\prime}=2$. When $n^{\prime}=2$, c contains two ciphertext components: $\left(\mathbf{c a}_{1}, c b_{1}, \mathbf{c a}_{1}^{\prime}, c b_{1}^{\prime}\right)$, and $\left(\mathbf{c a}_{2}, c b_{2}, \mathbf{c a}_{2}^{\prime}, c b_{2}^{\prime}\right)$. For $0 \leq i \leq m-1$, let $x_{1 i}, x^{\prime}{ }_{1 i}$, $x_{2 i}$, and $x^{\prime}{ }_{2 i}$ be $4 * m$ variables.

Then, from the ciphertext $\mathbf{c}$ and the pubic key PK, the adversary can obtain a stratified system, consisting of the following linear equations,

$$
\begin{gather*}
\mathbf{c a}_{1}=\Sigma_{i=0}^{m-1} x_{1 i} * \mathbf{a}_{i}  \tag{1}\\
c b_{1}=\Sigma_{i=0}^{m-1} x_{1 i} * b_{i} \bmod q  \tag{2}\\
\mathbf{c a}_{1}^{\prime}=\Sigma_{i=0}^{m-1} x^{\prime}{ }_{1 i} * \mathbf{a}^{\prime}{ }_{i}  \tag{3}\\
c b^{\prime}{ }_{1}=\Sigma_{i=0}^{m-1} x^{\prime}{ }_{1 i} * b^{\prime}{ }_{i} \bmod q  \tag{4}\\
\mathbf{c a}_{2}=\Sigma_{i=0}^{m-1} x_{2 i} * \mathbf{a}^{\prime}{ }_{i}  \tag{5}\\
c b_{2}=\Sigma_{i=0}^{m-1} x_{2 i} * b^{\prime}{ }_{i} \bmod q  \tag{6}\\
\mathbf{c a}_{2}=\Sigma_{i=0}^{m-1} x^{\prime}{ }_{2 i} * \mathbf{a}_{i}  \tag{7}\\
c b^{\prime}{ }_{2}=\Sigma_{i=0}^{m-1} x^{\prime}{ }_{2 i} * b_{i} \bmod q \tag{8}
\end{gather*}
$$

and the following quadratic equations $(0 \leq i \leq m-1)$, where $\mathbf{v}[0]$ and $\mathbf{v}[1]$ are also regarded as unknowns,

$$
\begin{equation*}
x^{\prime}{ }_{1 i}=\mathbf{v}[0] * x_{1 i}+\mathbf{v}[1] * x_{2 i} \bmod p \tag{9}
\end{equation*}
$$

$$
\begin{equation*}
x^{\prime}{ }_{2 i}=\mathbf{v}[0] * x_{2 i}+\mathbf{v}[1] * x_{1 i} \bmod p \tag{10}
\end{equation*}
$$

The variables $\mathbf{v}[0]$ and $\mathbf{v}[1]$ appear only in multivariate quadratic terms in equations (9) and (10). Hence, the values of $\mathbf{v}[0]$ and $\mathbf{v}[1]$ can only be determined from (9) and (10) by first knowing the values of $x_{1 i}, x^{\prime}{ }_{1 i}, x_{2 i}$, and $x^{\prime}{ }_{2 i}$. Moreover, for each $0 \leq$ $i \leq m-1, x_{1 i}, x^{\prime}{ }_{1 i}, x_{2 i}$, and $x^{\prime}{ }_{2 i}$ must lead to the same $\mathbf{v}[0]$ and the same $\mathbf{v}[1]$ for consistency.

The values of $x_{1 i}, x^{\prime}{ }_{1 i}, x_{2 i}$, and $x^{\prime}{ }_{2 i}$ can only be determined by solving linear equations (1)-(8). Based on the condition on public parameters, there can be a large number of solutions for equations (1)-(8). The adversary thus has to search the solution space of equations (1)-(8) for each $x_{1 i}, x^{\prime}{ }_{1 i}, x_{2 i}$, and $x^{\prime}{ }_{2 i}$ to find consistent $\mathbf{v}[0]$ and $\mathbf{v}[1]$.

On the other hand, for $0 \leq i \leq m-1$, by multiplying $\mathbf{a}^{\prime}{ }_{i}$ (or $\mathbf{a}_{i}$ ) at both sides of equations (9) (or equations (10)) and summing each side, the adversary can have $\sum_{i=0}^{m-1} x^{\prime}{ }_{1 i} * \mathbf{a}^{\prime}{ }_{i}=\mathbf{c a}^{\prime}{ }_{1}$ and $\sum_{i=0}^{m-1} x_{2 i} * \mathbf{a}^{\prime}{ }_{i}=\mathbf{c} \mathbf{a}_{2}\left(\right.$ or $\sum_{i=0}^{m-1} x^{\prime}{ }_{2 i} * \mathbf{a}_{i}=\mathbf{c a}^{\prime}{ }_{2}$ and $\sum_{i=0}^{m-1} x_{1 i} * \mathbf{a}_{i}=\mathbf{c a}_{1}$ ). Thus, equations (9) and (10) are reduced to:

$$
\begin{align*}
& \mathbf{c a}_{1}^{\prime}=\mathbf{v}[0] *\left(\Sigma_{i=0}^{m-1} x_{1 i} * \mathbf{a}_{i}^{\prime}\right)+\mathbf{v}[1] * \mathbf{c a}_{2} \bmod p  \tag{11}\\
& \mathbf{c a}^{\prime}{ }_{2}=\mathbf{v}[0] *\left(\Sigma_{i=0}^{m-1} x_{2 i} * \mathbf{a}_{i}\right)+\mathbf{v}[1] * \mathbf{c a}_{1} \bmod p \tag{12}
\end{align*}
$$

In the equations (11) and (12), the values of $\sum_{i=0}^{m-1} x_{1 i} * \mathbf{a}^{\prime}{ }_{i}$ and $\Sigma_{i=0}^{m-1} x_{2 i} * \mathbf{a}_{i}$ are not known. Hence, in order to to recover $\mathbf{v}_{0}$ and $\mathbf{v}_{1}$ from (11) and (12), the adversary still needs to search the solution spaces of equations (1)-(2) and (5)-(6) for possible values of $x_{1 i}$ and $x_{2 i}$.

At last, the adversary can search each pair of $\mathbf{v}[0]$ and $\mathbf{v}[1]$ from $\mathbb{Z}_{p}^{2}$, respectively. If the search space is small, for instance when $\mathbf{v}$ is not uniform, this search step is feasible. By guessing $\mathbf{v}[0]$ and $\mathbf{v}[1]$, equations (1) to (10) become a linear system. If this linear system allows $x_{1 i}, x^{\prime}{ }_{1 i}, x_{2 i}$, and $x^{\prime}{ }_{2 i}$ to have absolute values around $p$, then the adversary knows both $\mathbf{v}[0]$ and $\mathbf{v}[1]$ are correctly guessed.

With the implementation of Compact-LWE-MQ ${ }^{\mathrm{H}}$, we have done experiments to confirm the solution space of equations (1)-(2) (or other equations from (3) to (8)) and the attacks by guessing $\mathbf{v}[0]$ and $\mathbf{v}[1]$.

### 3.2 Attack to Public Keys

The aim of the adversary in this attack is to find a correct private key SK from the public key $\mathbf{P K}$. Note that a correct $\mathbf{S K}$ is not necessarily the original one; instead, it must guarantee the consistency of $u_{i}$ in all samples.

In the sample of $\mathbf{P K}$, the error term $r_{i}$ (or $r_{i}^{\prime}$ ) is multiplied by $s \in \mathbf{Z}_{q}$ (or $s^{\prime} \in \mathbf{Z}_{q}$ ). Hence, the adversary has to know $s$ and $s^{\prime}$ before performing any possible attacks to recover candidate solutions to $s$ and $\mathbf{s}^{\prime}$.

One way to achieve this purpose is that the adversary generates a short $m$-dimensional vector, denoted $\mathbf{l}$, such that $\sum_{i=1}^{m}\left(\mathbf{l}[i] * \mathbf{a}_{i}\right)=\mathbf{0}$ (and/or $\sum_{i=1}^{m}\left(\mathbf{l}[i] * \mathbf{a}_{i}{ }_{i}\right)=\mathbf{0}$ ). Then, the adversary can get $\sum_{i=1}^{m}\left(\mathbf{l}[i] * b_{i}\right)=s * \sum_{i=1}^{m}\left(\mathbf{l}[i] * r_{i}\right) \bmod q$. Then, the adversary can guess $\sum_{i=1}^{m}\left(\mathbf{l}[i] * r_{i}\right)$, which is much smaller than $s$.

In the parameter configuration proposed later, the size of $p$ is 128 bits. Thus, guessing $\sum_{i=1}^{m}\left(\mathbf{l}[i] * r_{i}\right)$ still needs to search in a big space. However, we will assume the adversary has somehow obtained the original $s$ and $s^{\prime}$. We then focus on proving the necessity of exhaustively searching $h$ and $h^{\prime}$, two moduli in the bottom layer.

Theorem 3. Let $\operatorname{gen}(p p)=(\mathbf{S K}, \mathbf{P K}), \mathbf{S K}=\left(\mathbf{s}, \mathbf{k}, \mathbf{t}, \mathbf{z}, s, h, k, \mathbf{s}^{\prime}, \mathbf{k}^{\prime}, \mathbf{t}^{\prime}, \mathbf{z}^{\prime}, s^{\prime}, h^{\prime}, k^{\prime}\right)$. Suppose $s$ and $s^{\prime}$ has been known by the adversary. Then, $h$ and $h^{\prime}$ must be exhaustively searched to check whether secret values (i.e., $\mathbf{k}, \mathbf{k}^{\prime}, \mathbf{t}, \mathbf{t}^{\prime}, \mathbf{z}, \mathbf{z}^{\prime}, k$, and $k^{\prime}$ ) can lead to consistent $u_{i}$ in each sample of $\mathbf{P K}$.

Proof. Let $i$ th sample of $\mathbf{P K}$ be $\left(\mathbf{a}_{i}, b_{i}, \mathbf{a}_{i}^{\prime}, b_{i}^{\prime}\right)$. Since $s$ and $s^{\prime}$ are assumed to be known by the adversary, we just let $s=1$ and $s^{\prime}=1$ in this proof. Then, from the pubic key, the adversary obtains the following equations:

$$
\begin{gather*}
b_{i}=\left\langle\mathbf{a}_{i}, \mathbf{s}\right\rangle+r_{i} \bmod q  \tag{13}\\
{b^{\prime}}_{i}=\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{s}^{\prime}\right\rangle+r_{i}^{\prime} \bmod q \tag{14}
\end{gather*}
$$

We know that $a \_\max <p<h$ and the number of public key samples $m$ is limited. Thus, in addition to the original values of $s$ and $s^{\prime}$ specified in the private key, there can be other values of $\mathbf{s}$ and $\mathbf{s}^{\prime}$, denoted by $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{s}^{\prime}}$, which satisfy the equations below, as confirmed by experiments in [7].

$$
\begin{align*}
b_{i} & =\left\langle\mathbf{a}_{i}, \tilde{\mathbf{s}}\right\rangle+\tilde{r}_{i} \bmod q  \tag{15}\\
b_{i}^{\prime} & =\left\langle\mathbf{a}_{i}^{\prime}, \tilde{\mathbf{s}^{\prime}}\right\rangle+\tilde{r}_{i}^{\prime} \bmod q \tag{16}
\end{align*}
$$

Thus, from each possible value of $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{s}^{\prime}}$, the adversary gets a pair of $\tilde{r}_{i}$ and $\tilde{r}_{i}^{\prime}$. At this point, the adversary has to guess the moduli $h$ and $h^{\prime}$ for the possibility of solving the equations below and checking the correctness of $\mathbf{k}, \mathbf{k}^{\prime}, \mathbf{t}, \mathbf{t}^{\prime}, \mathbf{z}, \mathbf{z}^{\prime}, k$, and $k^{\prime}$, in terms that $u_{i}$ in each sample is the same.

$$
\begin{gather*}
\tilde{r}_{i}=\left(\left\langle\mathbf{a}_{i}, \mathbf{k}\right\rangle+\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{t}\right\rangle+k * u_{i} \bmod p\right)+\left\langle\mathbf{a}_{i}, \mathbf{z}\right\rangle \bmod h  \tag{17}\\
\tilde{r}_{i}^{\prime}=\left(\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{k}^{\prime}\right\rangle+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle+k^{\prime} * u_{i} \bmod p\right)+\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{z}^{\prime}\right\rangle \bmod h^{\prime} \tag{18}
\end{gather*}
$$

Moreover, an arbitrary guess of $h$ and $h^{\prime}$ will not lead to correct secret values in the above equations, with the reason discussed below. Hence, an exhaustive search to $h$ and $h^{\prime}$ is necessary.

Suppose the adversary has guessed $h^{\prime}$ and gets $u_{i}$ from equations (18). Then, the adversary moves to solve equations (17), where $u_{i}$ becomes known values. With different guesses of $h$, the expression $\left\langle\mathbf{a}_{i}, \mathbf{k}\right\rangle+\left\langle\mathbf{a}^{\prime}{ }_{i}, \mathbf{t}\right\rangle+k * u_{i} \bmod p$ in equations (17) is constrained to different values. The equation system consisting of such expressions and their values is over-determined because $2 * n<m$. An arbitrary guess of $h$ cannot make this over-determined system have consistent solutions to $\mathbf{k}$, $\mathbf{t}$, and $k$. This is confirmed with an experiment in Appendix.

Note that for an arbitrary solution to $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{s}^{\prime}}$, there might not exist $h$ and $h^{\prime}$ that can lead to correct secret values at the bottom layer; this point will be exploited in the future work to make Compact-LWE-MQ ${ }^{\mathrm{H}}$ lightweight.

The search space of $h$ and $h^{\prime}$ is $\mathbf{Z}_{p}$ as defined by the key generation algorithm. Based on this theorem, the selection of parameters for a desired security level will be very straightforward.

### 3.3 CPA-Security and CCA-Security of Compact-LWE-MQ ${ }^{\mathrm{H}}$

Let $\mathbf{v} \in \mathbb{Z}_{p}^{n^{\prime}}$ be the input to the encryption algorithm. It includes the plaintext in some of its entries, while other entries must include random numbers in order to be CPAsecure. Otherwise, according to Theorem 2, in the standard CPA indistinguishability experiment [5], the adversary simply guesses each message sent to the challenger, and the correct guess will return shorter solutions to the linear equations derived from the ciphertext and the public key, as shown in the proof of Theorem 2. If in one of its entries $\mathbf{v}$ includes a uniform random number from $\mathbb{Z}_{p}$, the probability of correctly guessing $\mathbf{v}$ is a negligible function with respect to the size of $p$ in bits.

In practical applications, if $\mathbf{v}$ already contains uniformly sampled random values and is long enough, the encryption algorithm does not need to randomize $\mathbf{v}$ by adding extra random components. Theorem 3 ensures that the adversary cannot guess efficiently in the CPA indistinguishability experiment by efficiently recovering a correct private key (not necessarily the original one).

Note that currently Compact-LWE-MQ ${ }^{\mathrm{H}}$ is not CCA-secure. Let $x$ be a small integer and $\mathbf{c}$ be the ciphertext encrypting one of two messages from the adversary in the standard CCA indistinguishability experiment [5]. Then, $\mathbf{c}$ and $\mathbf{c} * x$ decrypt into the same message, because the decryption of $\mathbf{c} * x$ returns $\mathbf{g} * x$ (and hence $\mathbf{G} * x$ ) and $\mathbf{y} * x$, leading to

$$
(\mathbf{G} * x)^{-1} *(\mathbf{y} * x)=\mathbf{G}^{-1} * \mathbf{y} \bmod p
$$

Thus, in the CCA experiment, the adversary can send $\mathbf{c} * x$ to the decryption oracle for decryption. In addition, there is another chosen ciphertext attack. Suppose $\mathbf{c}$ and $\mathbf{c}^{\prime}$ encrypts the same message and $\mathbf{c}_{a}$ is a new ciphertext obtained from the component wise addition of $\mathbf{c}$ and $\mathbf{c}^{\prime}$. Then, by sending $\mathbf{c}_{a}$ to the decryption oracle, the adversary can recover the message.

However, we can revise Compact-LWE-MQ ${ }^{\mathrm{H}}$ slightly, as described below, to make it non-malleable in both attack cases.

- Add a random value $w \in \mathbb{Z}_{p}$ in the private key $\mathbf{S K}$ and also $\widehat{\mathbf{S K}}$.
- For $u_{i}$ randomly selected in the generation of PK, ensure that

$$
\sum_{i=0}^{m-1}\left(\left\langle\mathbf{a}_{i}^{\prime}, \mathbf{t}\right\rangle * k^{-1}+u_{i}+\left\langle\mathbf{a}_{i}, \mathbf{t}^{\prime}\right\rangle * k^{\prime-1}\right)=w \bmod p
$$

- In Algorithm 2 of generating ciphertext component, change the initial value of $\mathbf{1}^{\prime}$ (line 2) into

$$
\mathbf{l}^{\prime}=\mathbf{1} *\left(\sum_{i=0}^{n^{\prime}-1} \mathbf{v}[i] \bmod p\right)
$$

where 1 is a $m$-dimensional vector containing integer 1 for each entry.

- In Algorithm 3 of decryption, change the code in line 10 into

$$
\mathbf{g}=\mathbf{g}+(g+w)
$$

We analyze the non-malleability of revised Compact-LWE-MQ ${ }^{\mathrm{H}}$ briefly for the first attack case. With this updated Compact-LWE-MQ ${ }^{\mathrm{H}}$, from $\mathbf{c} * x$, the decryption algorithm gets $\mathbf{g} * x+\mathbf{1} * w$, instead of $\mathbf{g} * x$ as above. Let $\mathbf{G}^{\prime}$ be the matrix derived from $\mathbf{g} * x+\mathbf{1} * w$ in the decryption of $\mathbf{c} * x$ and $\mathbf{G}$ be the matrix derived from $\mathbf{g}+\mathbf{1} * w$ when decrypting $\mathbf{c}$. Then, we have

$$
\mathbf{G}^{\prime-1} *(\mathbf{y} * x) \neq \mathbf{G}^{-1} * \mathbf{y} \bmod p
$$

## 4 Implementation and Attack Analysis

A prototype of Compact-LWE-MQ ${ }^{\mathrm{H}}$ for CCA-security has been implemented with SageMath and is included in Appendix. In this section, we present a configuration of parameters, aimed at 128 -bit security. Then, we use experiments to do security evaluation. As to be seen below, for a desired security level, the parameters can be easily determined, without relying on the capability of any cryptanalysis algorithms like a lattice-reduction algorithm.

### 4.1 Configuration of Parameters

The configuration of parameters is shown in Table 1, where $p$ has 128 bits, and $q$ around 394 bits.

As stated in Theorem 3, both $h$ and $h^{\prime}$ must be searched in the range of $\mathbf{Z}_{p}$ as a necessary step for performing attacks to find a correct private key. Since $p$ is 128 bits, only $h$ or $h^{\prime}$ can guarantee the 128-bit security level for SK. Hence, we do not need to quantify the extra guarantee from other secret components in SK.

| $p$ | a_max | $n$ | $m$ | $q$ |
| :---: | :---: | :---: | :---: | :---: |
| $2^{128}+51$ | $2^{56}$ | 4 | 24 | 394 bits |

Table 1: Configuration of Parameters or Parameter Size

In this configuration, the size of public key in bytes is close to

$$
2 * m *\left(n * \log _{2}\left(a \_\max \right)+\log _{2}(q)\right) / 8=2 * 24 *(4 * 56+394) / 8=3708 .
$$

Suppose the plaintext to be encrypted is $v_{1} \in \mathbb{Z}_{p}$. Note that $v_{1}$ might be known by the adversary with very high probability, just like in the CPA indistinguishability experiment. According to Theorem 2, $v_{1}$ must be padded with extra random numbers to make it hard to guess, before it is passed to the encryption algorithm.

Hence, the minimum of $n^{\prime}$ must be 2 in this configuration and the vector $\mathbf{v}$ passed to the encryption algorithm can be prepared as:

$$
\mathbf{v}[0] \leftarrow \mathbb{Z}_{p} \text { and } \mathbf{v}[1]=\mathbf{v}[0] \oplus v_{1}
$$

Similarly, if there are two plaintexts $v_{1} \in \mathbb{Z}_{p}$ and $v_{2} \in \mathbb{Z}_{p}$, which are not uniformly distributed, then we let $n^{\prime}=3$ and the vector $\mathbf{v}$ can be:

$$
\mathbf{v}[0] \leftarrow \mathbb{Z}_{p}, \mathbf{v}[1]=\mathbf{v}[0] \oplus v_{1} \text { and } \mathbf{v}[2]=\mathbf{v}[0] \oplus v_{2}
$$

The size of ciphertexts depends on $n^{\prime}$. When $n^{\prime}=2$, the size of the ciphertext in bytes is around $n^{\prime} * 2 *\left(4 *\left(128+56+\log _{2}(24)\right)+394\right) / 8=574$ bytes. The decryption failure for this configuration has the probability $\frac{1}{2^{128}+51}$.

### 4.2 Experiments

We use experiments to estimate whether $m$ is big enough for the expected 128-bit security level, and conduct attacks to ciphertexts. The SageMath script for such attacks is also in Appendix.
4.2.1 Estimation of $\boldsymbol{m}$ If $m$ is too small, there can be a small number of solutions for equations (1) and (2), and other equations (3) and (4), (5) and (6), (7) and (8) as well.

We have used experiments to evaluate whether $m=24$ is big enough to allow more than $2^{128}$ solutions to those equations. The lattice developed in [2] is used to solve equations.

In equations (1) and (2), $x_{10}$ (or other variables $x_{1 i}$ for $1 \leq i \leq m-1$ ) can be an integer between 0 and $p-1$. If for each $x_{10} \in \mathbb{Z}_{p}$ there is a different solution for other variables $x_{1 i}$ in equations (1) and (2), then there must be more than $2^{128}$ solutions to (1) and (2), because $p$ is 128 bits in the proposed configuration.

In our experiment, we select three sets of sample values for $x_{10}$ to check the existence of distinct solutions for $x_{1 i}(1 \leq i \leq m-1)$. The first set is for $x_{10}$ from 0 to 1023 , the second set from $\lfloor p / 2\rfloor$ to $1023+\lfloor p / 2\rfloor$, and the third set from $p-1024$ to $p-1$. In the test of all three sets, the distinct solutions exist for all other variables $x_{1 i}$ ( $1 \leq i \leq m-1$ ).

In addition, when $m$ is as small as 8 , with other parameters not changed, there exists only one solution to equations (1) and (2). Hence, $m=24$ is big enough for $2^{128}$ distinct solutions for $x_{1 i}$.
4.2.2 Attack Experiments In this experiment, we confirm that an arbitrary solution of $x_{1 i}, x_{1 i}^{\prime}, x_{2 i}$, and $x_{2 i}^{\prime}$ obtained by solving equations (1)-(8) cannot make equations (9) and (10) hold; that is, an arbitrary solution of $x_{1 i}, x_{1 i}^{\prime}, x_{2 i}$, and $x_{2 i}^{\prime}$ cannot be used to recover $\mathbf{v}$.

In this experiment, we first get one solution for $x_{1 i}^{\prime}, x_{2 i}$, and $x_{2 i}^{\prime}$ by solving equations (3)-(8). Then, we generate different solutions for $x_{1 i}$ from equations (1) and (2). Our experiment confirmed that the solutions to equations (1)-(8) cannot satisfy overdetermined equations (9) and (10) for $0 \leq i \leq m-1$ and hence $\mathbf{v}$ cannot be recovered.

Another experiment is conducted to attack ciphertexts by exhaustively guessing $\mathbf{v}$ by choosing a small value for $p$, i.e. $p=257$. With guessed $\mathbf{v}$, equations (1)-(10) become a linear system. In this experiment, we solve this system by guessing different values of $\mathbf{v}$, including the correct one.

Our experiment shows that $x_{1 i}, x_{1 i}^{\prime}, x_{2 i}$, and $x_{2 i}^{\prime}$ have solution values roughly in the range of $\mathbb{Z}_{p}$, if $\mathbf{v}$ is correctly guessed; otherwise, there are no valid solutions for both versions of Compact-LWE-MQ ${ }^{\mathrm{H}}$, or the solutions contain values much bigger than $p$ in the CPA version of Compact-LWE-MQ ${ }^{H}$.

## 5 Related Works

Compact-LWE-MQ ${ }^{\mathrm{H}}$ demonstrates the feasibility of defining public encryption without assuming hard computational problems. There is no existing work sharing the same purpose.

The public key samples in Compact-LWE-MQ ${ }^{\mathrm{H}}$ is syntactically similar to LWE samples [8] at its top layer. However, Compact-LWE-MQ ${ }^{\mathrm{H}}$ allows multiple solutions to secret vectors at the top layer from the samples, while it is not the case for LWE. Thus, when $n$ is a very small integer (e.g., $n=4$ ) as configured for Compact-LWE-MQ ${ }^{\mathrm{H}}$ in our evaluation, LWE samples can be solved to find the secret vector. More importantly, the ciphertexts in Compact-LWE-MQ ${ }^{\mathrm{H}}$ can be represented as equations with plaintexts appeared in quadratic (MQ) terms, while the public encryption scheme in [8] (and other lattice-based schemes) treat plaintetxs as a linear term, which is vulnerable to latticebased attacks for small parameters.

There are algorithms to solve MQ problems, such as [3,4]. One major step in all exiting algorithms solving MQ problems is to derive an initial system of linear equations, so that MQ problems can be solved by starting with the linear equations and then recursively finding the values of variables in quadratic terms. This step does not help attack Compact-LWE-MQ ${ }^{\mathrm{H}}$, because its ciphertext is already a stratified system and its security is analyzed in Theorem 2. In addition, the ciphertext equations in Compact-LWE-MQ ${ }^{\mathrm{H}}$ are defined over two moduli at different layers ( $q$ and $p$ ), while equations in MQ problems are usually defined over one modulus for all equations (i.e., no layered structures in current MQ problems with all secrets at the same layer).

Compact-LWE-MQ ${ }^{\mathrm{H}}$ shares its name with Compact-LWE [7], which have been attacked in [2,6] by recovering plaintexts from ciphertexts. This is because ciphertexts are defined as a linear system in Compact-LWE; hence, with small dimension parameters, the current lattice-reduction tools can attack Compact-LWE efficiently. All attacks to Compact-LWE are not applicable to Compact-LWE-MQ ${ }^{\mathrm{H}}$, because ciphertexts in Compact-LWE-MQ ${ }^{\mathrm{H}}$ is a stratified system of linear and quadratic equations and the upper layer subproblem has to be solved by exhaustive search when the lower layer secret is not known.

## 6 Conclusion

In this paper, we have developed a public key encryption scheme Compact-LWE-MQ ${ }^{\mathrm{H}}$, with the aim to demonstrate the feasibility of using hardness facts, rather than hardness
assumptions, as a new principle of constructing public key encryption. The factually hard problems ensure the plaintexts and secret values can only be recovered by taking brute-force search as one necessary step. Compact-LWE-MQ ${ }^{\mathrm{H}}$ is simple to understand, facilitating the thorough analysis of its security from researchers with different backgrounds. With very small parameter configuration, the existing cryptanalysis tools are already efficient enough to attack Compact-LWE-MQ ${ }^{\mathrm{H}}$ if its security claim does not hold. In other words, the security of Compact-LWE-MQ ${ }^{\mathrm{H}}$ does not assume the capability of cryptanalysis tools or hardness of problems.

## References

1. Ajtai, M.: Generating hard instances of lattice problems (extended abstract). In: Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing. pp. 99-108. STOC '96 (1996)
2. Bootle, J., Tibouchi, M., Xagawa, K.: Cryptanalysis of compact-lwe. In: Topics in Cryptology - CT-RSA 2018 - The Cryptographers’ Track at the RSA Conference 2018, San Francisco, CA, USA, April 16-20, 2018, Proceedings. Lecture Notes in Computer Science, vol. 10808, pp. 80-97. Springer (2018)
3. Courtois, N., Klimov, A., Patarin, J., Shamir, A.: Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In: Preneel, B. (ed.) Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques, Bruges, Belgium, May 14-18, 2000, Proceeding. Lecture Notes in Computer Science, vol. 1807, pp. 392-407. Springer (2000)
4. Faugère, J.: A new efficient algorithm for computing gröbner bases (f4). Journal of Pure and Applied Algebra 139(1-3), 61-88 (1999)
5. Katz, J., Lindell, Y.: Introduction to Modern Cryptography. Chapman \& Hall/CRC (2014)
6. Li, H., Liu, R., Pan, Y., Xie, T.: Cryptanalysis of compact-lwe submitted to NIST PQC project. IACR Cryptol. ePrint Arch. 2018, 20 (2018)
7. Liu, D., Li, N., Kim, J., Nepal, S.: Compact-lwe: Enabling practically lightweight public key encryption for leveled iot device authentication. IACR Cryptol. ePrint Arch. 2017, 685 (2017), http://eprint.iacr.org/2017/685
8. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. In: Gabow, H.N., Fagin, R. (eds.) Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005. pp. 84-93. ACM (2005)
9. Yasuda, T., Dahan, X., Huang, Y., Takagi, T., Sakurai, K.: MQ challenge: Hardness evaluation of solving multivariate quadratic problems. IACR Cryptol. ePrint Arch. (2015), http://eprint.iacr.org/2015/275

## Appendix A Compact-LWE-MQ ${ }^{\mathrm{H}}$ Implementation

```
# ===============================
# Compact-LWE-MQ^{H}, CCA version
# ================================
set_random_seed(1)
p = next_prime(2^128)
n}=
m}=2
a_max = 2^56
# ========================
def prikey_gen():
    h = next-prime(m*p*p+randint(0, p-1))
    hp}=next_prime(m*p*p+randint (0, p-1)
    q}=\mathrm{ next-prime(m*p*(h+hp)+randint (0, p-1))
    S = vector(ZZ,[randint(0, q-1) for_ - in range(n)])
    Sp}=\operatorname{vector(ZZ,[randint(0, q-1) for_ in range(n)])
    K = vector(ZZ,[randint(0, p-1) for_ in range(n)])
    Kp = vector(ZZ,[randint(0, p-1) for _ in range(n)])
    T= vector(ZZ,[randint(0, p-1) for_ _ in range(n)])
    Tp = vector(ZZ,[randint(0, p-1) for _ in range(n)])
    Z = vector(ZZ,[randint(0, h-1) for_ in range(n)])
    Zp}=\mathrm{ vector(ZZ,[randint(0, hp-1) for_ in range(n)])
    k = randint(1, p-1)
    kp = randint(1, p-1)
    w = randint(0, p-1)
    s = randint(1, q-1)
    sp = randint(1, q-1)
    return S,Sp,s,sp,K,Kp,k,kp,T,Tp,Z,Zp,h,hp,w,q
def pubkey_gen(q,S,Sp,s,sp,K,Kp,k,kp,T,Tp,Z,Zp,h,hp,w):
    A1 = random_matrix (ZZ,m,n,x=0,y=1)
    A2 = random_matrix (ZZ,m, n, x=0,y=1) #A1
    b1 = vector(ZZ, [0 for _ in range(m)])
    b2 = vector (ZZ, [0 for _ in range(m)])
    kInv = inverse_mod (k, p)
    kpInv = inverse_mod (kp, p)
```

```
    for i in range(m):
        ui = randint(0,p-1)
        A1[i,:] = vector(ZZ, [randint(0, a_max -1) for - in range(n)])
        A2[i,:] = vector(ZZ, [randint(0, a_max -1) for - in range(n)])
        if (i<m-1):
            w = (w - (A2[i]*T*kInv+ui+A1[i]*Tp*kpInv))%p
        else:
            ui = (w - (A2[i] *T*kInv+A1[i]*Tp*kpInv))%p
        r = ((K*A1[i]+T*A2[i]+k*ui)%p + Z*A1[i])%h
        rp = ((Kp*A2[i]+Tp*A1[i]+kp*ui)%p + Zp*A2[i])%hp
        e = randint(0,p-1)
        ep = randint (0,p-1)
        b1[i] = (A1[i]*S + s*(r) )%q
        b2[i] = (A2[i]*Sp + sp*(rp))%q
    return A1, b1, A2, b2
def encS(A1,b1,A2, b2, L, v, index, q):
    np = len(v)
    11 = L[index]
    12 = vector(ZZ, [sum(v)%p for i in range(m)])
    for i in range(np):
        12 = 12 + L[(index+i)%np]*v[i]
        12 = 12%p
    ca1 = 11 * A1
    cb1 = (11 * b1) % q
    ca2 = 12 * A2
    cb2 = (12 * b2) % q
    c = (ca1, cb1, ca2, cb2)
    return c
def enc(q, A1,b1,A2, b2, v):
    np = len(v)
    L = random_matrix (ZZ, np,m,x=0,y=0)
    for i in range(np):
        l = vector(ZZ,[randint(0, p-1) for _ in range(m)])
        L[i,:] = 1
    C=[]
    for i in range(np):
        if (i % 2==0):
            c = encS(A1,b1,A2, b2, L, v, i, q)
        else:
```

```
            c = encS(A2,b2,A1, b1, L, v, i, q)
        C=C+[c]
    return C
def decS(S,Sp,c,s,sp,K,Kp,q,k,kp,T,Tp,Z,Zp,h,hp):
    sInv = inverse_mod(s, q)
    spInv = inverse_mod(sp, q)
    kInv = inverse_mod (k, p)
    kpInv = inverse_mod (kp, p)
    ca1,cb1,ca2, cb2 = c
    d = sInv*(cb1-S*ca1)%q
    d = (d-Z*ca1)%h
    g = (kInv*(d-K*ca1)+Tp*ca1*kpInv)%p
    dp = spInv*(cb2-Sp*ca2)%q
    dp = (dp-Zp*ca2)%hp
    y = (kpInv*(dp-Kp*ca2)+T*ca2*kInv)%p
    return g, y
def dec(q, S, Sp, C, s, sp, K, Kp, k, kp, T, Tp, Z, Zp, h, hp, w):
    y = []
    g = []
    np = len(C)
    for i in range(np):
        if (i % 2==0):
            g1, y1= decS(S,Sp,C[i],s,sp,K,Kp,q,k,kp,T,Tp,Z,Zp,h,hp)
        else:
            g1, y1= decS(Sp,S,C[i],sp,s,Kp,K,q,kp,k,Tp,T,Zp,Z,hp,h)
        y = y + [y1]
        g = g + [(g1+w)%p]
    Rp = Integers(p)
    G = random_matrix (Rp,np,np)
    for i in range(np):
        G[i,:] = vector(Rp, g[i:np]+g[0:i])
    try:
        v = G.inverse()*vector(Rp, y)
    except:
        return -1
    return list(v)
```

```
#correctness
correct = 0
failure = 0
np = 2 #3, 5, 7, 9, 11, 13, ..., 101, ...
S, Sp, s, sp, K, Kp, k, kp, T, Tp, Z, Zp, h, hp, w, q = prikey_gen()
A1,b1,A2,b2 = pubkey_gen(q,S,Sp,s,sp,K,Kp,k,kp,T,Tp,Z,Zp,h,hp,w)
for i in range(10):
    v =[randint(0,p-1) for - in range(np)]
    C = enc(q, A1,b1,A2, b2, v)
    dec_v = dec(q,S,Sp,C,s,sp,K,Kp,k,kp,T,Tp,Z,Zp,h,hp,w)
    if(v==dec_v):
        correct = correct+1
    if (dec_v==-1):
        failure = failure+1
print ("Correctness check: ", correct, failure)
```


## Appendix B Ciphertext Attack

```
# ===========================
# Ciphertext Attack Experiment
# with latex L from [2,7]
# =========================
def short_x(A1,b1,ca1,cb1, q):
    pa=q
    pa2=pa
    pa3=1
    L=block_matrix (ZZ, \
        [[1, 0, -pa*cal.row(), -pa2 * cb1],\
        [0, pa3*identity_matrix (m), pa*A1, pa2 * b1.column()],\
        [0, 0, 0, pa2*q]
        ])
    L=L.LLL()
    #index of first non-zero entry in the first column of L
    idx=next((i for i,x in enumerate(L.column(0). list()) if x!=0))
    x = vector(ZZ,L[idx][1:(m)+1]/pa3) if L[idx][0] == 1 \
                                    else vector(ZZ,-L[idx][1:(m)+1]/pa3)
    return x
def short_x1(A1,b1,ca1,cb1, x11, q):
    pa=q
    pa2=pa
    pa3=1
    calp = cal - x11*A1[0]
    cb1p = (cb1 - x 11*b1[0])%q
    A10 = A1[0,:]
    b10=b1[0]
    if (x11>0):
        A1[0,:] = vector(ZZ, [0 for _ in range(n)])
        b1[0] = 0
    L=block_matrix (ZZ, \
        [[1, 0, -pa*ca1p.row(), -pa2 * cb1p], \
        [0, pa3*identity_matrix(m), pa*A1, pa2 * b1.column()], \
        [0, 0, 0, pa2*q]
        ])
```

```
L=L.LLL()
#index of first non-zero entry in the first column of L
idx=next((i for i,x in enumerate(L.column(0). list()) if x!=0))
x = vector(ZZ,L[idx][1:(m)+1]/pa3) if L[idx][0] == 1 \
                                    else vector(ZZ,-L[idx][1:(m)+1]/pa3)
if (x11>0):
    x[0] = x11
A1[0,:] = A10
b1[0] = b10
return x
S, Sp, s, sp, K, Kp, k, kp, T, Tp, Z, Zp, h, hp, w, q = prikey_gen () A1, b1, A2, b2 = pubkey_gen (q, S, Sp, s, sp, K, Kp, k, kp, T, Tp, Z, Zp, h, hp,w)
v =[randint(0,p-1) for - in range(2)]
C = enc(A1,b1,A2, b2, v)
ca1,cb1,ca2, cb2 = C[0]
ca3,cb3,ca4, cb4 = C[1]
x1p = short_x (A2,b2,ca2,cb2, qp)
print ("Check x1p:", x1p*A2==ca2, (x1p*b2)%qp == cb2)
x2 = short-x (A2,b2,ca3,cb3, qp)
print ("Check x2:", x2*A2==ca3, (x2*b2)%qp == cb3)
x2p = short-x (A1,b1,ca4,cb4, q)
print ("Check x2p:", x2p*A1==ca4, (x2p*b1)%q == cb4)
for x11 in range(2):
    x1 = short_x (A1,b1,ca1,cb1, x11, q)
    print ("Check x1:", x1*A1==ca1, (x1*b1)%q == cb1)
    Rp = Integers(p)
    G = matrix(Rp, [[x1[0]+1, x2[0]+1], [x2[0]+1, x1[0]+1]])
    vp = G.inverse()*vector(Rp, [x1p[0], x2p[0]])
    consist = 0
    for j in range(m):
        if ((x1[j]*vp[0]+x2[j]*vp[1]+vp[0]+vp[1])%p==(x1p[j])%p and \
                (x1[j]*vp[1]+x2[j]*vp[0]+vp[0]+vp[1])%p==(x2p[j])%p):
                consist= consist + 1
    if(consist==m):
        print ("+++++ successful attack+++++")
    else:
        print ("----failed attack ----", consist)
```


## Appendix C Check Arbitrary Guess of $\boldsymbol{h}$

```
A1 = random_matrix (ZZ, 3, 3, x=0,y=7)
A2 = random_matrix (ZZ,3,3,x=0,y=7)
h = 23
Rp = Integers(h)
K = vector(ZZ,[randint(0, h) for - in range(3)])
A1a = A1.change_ring(Rp)
A2a = A2.change_ring(Rp)
y1 = (A1a*K).1ift()
y2 = (A2a*K).1ift()
K1a = Ala.inverse()*yl
K2a = A2a.inverse()*y2
print ("---consistent solution with the correct modulus----")
print (K1a)
print (K2a)
h = 31 #19, 23, 29, 31, 37, 91, 1019
Rpp = Integers(h)
A1b = A1.change_ring(Rpp)
A2b = A2.change_ring(Rpp)
K1b = Alb.inverse()*yl
K2b = A2b.inverse()*y2
print ("----inconsistent solution with an arbitrary modulus")
print (K1b)
print (K2b)
```


[^0]:    * Corresponding author: dongxi.liu@csiro.au
    ${ }^{1}$ The scheme is the same as in the previous version of the report (10 Aug 2020); only explanations and discussions are changed.

