A Study on Lattice-Based Public-Key Cryptosystems

05M37122 Keita Xagawa

Supervisor: Keisuke Tanaka

Department of Mathematical and Computing Sciences
Tokyo Institute of Technology

January 18, 2007

Contents

Chapter 1

Introduction

1.1 Background

Lattice Problems and Cryptography. The lattice-based cryptosystems have been well-studied since Ajtai's seminal result [Ajt96a] on a one-way function based on the worst-case hardness of lattice problems, which initiated the cryptographic use of lattice problems. Ajtai and Dwork first succeeded to construct public-key cryptosystems [AD97] based on the unique shortest vector problem (uSVP). After their results, a number of lattice-based cryptosystems have been proposed in the last decade by using cryptographic advantages of lattice problems [GGH97b, CC99, HPS98, Reg04, Ajt05, Reg05].

We can roughly classify the lattice-based cryptosystems into two types: (A) those who are efficient on the size of their keys and ciphertexts and the speed of encryption/decryption procedures, but have no security proofs based on the hardness of well-known lattice problems, and (B) those who have security proofs based on the lattice problems but are inefficient.

For example, the GGH cryptosystem [GGH97b], NTRU [HPS98] and their improvements [Mic01, PJH03, Ngu02, HGNP+03] belong to the type A. These are efficient multi-bit cryptosystems related to lattices, however it is unknown whether their security is based on the hardness of well-known lattice problems. Actually, a few papers reported security issues of cryptosystems in this type [Ngu99, Gen01].

On the other hand, those in the type B have security proofs based on well-known lattice problems such as uSVP, the shortest vector problem (SVP) and the shortest linearly independent vectors problem (SIVP) [AD97, Reg04, Reg05]. (See Section 2.2 for their definitions and computational complexity.) In particular, the security of these cryptosystems can be guaranteed by the worst-case hardness of the lattice problems, i.e., breaking the cryptosystems on average is at least as hard as solving the lattice problems in the worst case. This attractive property of

the average-case/worst-case connection has been also studied from a theoretical point of view and obtained the families of one-way functions or collision-resistant hash functions [Ajt96a, GGH96, CN97, MR04, Mic04b, LM05, PR06].

Aside from the interesting property, such cryptosystems generally have longer keys and ciphertexts than those of the cryptosystems in the type A. To set their size practically reasonable, their security parameters must be small, which possibly makes the cryptosystems insecure in a practical sense [NS98]. Therefore, it is important to improve their efficiency for secure lattice-based cryptosystems in the type B.

In recent years, several researchers actually considered more efficient lattice-based cryptosystems with security proofs. For example, Regev constructed an efficient lattice-based cryptosystem with shorter keys [Reg05]. The security is based on the worst-case quantum hardness of certain approximation versions of SVP and SIVP, that is, his cryptosystem is secure if we have no polynomial-time quantum algorithm that solves the lattice problems in the worst case. Ajtai also constructed an efficient lattice-based cryptosystem with shorter keys by using a compact representation of special instances of uSVP [Ajt05], whose security is based on a certain Diophantine approximation problem.

Other Applications. In addition to public-key cryptosystems and families of one-way functions or collision-resistant hash functions, there are many cryptographic primitives, such as digital signature, bit commitment, proof of knowledge, zero knowledge, and etc. There are a few works on each primitive based on lattice problems.

On digital signature, there exist lattice-based signature schemes, the GGH signature scheme [GGH97b], NSS [HPS01], and NTRUSign [HHGP+03]. One year later from NSS [HPS01] appearing, it was analyzed by two reports [GJSS01, GS02]. In 2003, Szydlo proposed an attack on the GGH signature scheme and NTRUSign-251 without perturbation [Szy03]. Recently, Nguyen and Regev proposed a practical attack on the GGH signature scheme and NTRUSign-251 without perturbation using learning algorithm [NR06]. On string commitment (rather than bit commitment), it is already known that the family of collision-resistant hash functions implies a statistically-hiding computationally-binding string commitment scheme [HM96, DPP97, DPP98].

There are also a few works on zero knowledge and proof of knowledge. Goldreich and Goldwasser showed $\mathrm{coGapCVP}_{\Omega(\sqrt{n/\log n})} \in \mathrm{AM}$ and proposed a statistical zero-knowledge proof for $\mathrm{coGapCVP}_{\Omega(\sqrt{n/\log n})}$ and $\mathrm{coGapSVP}_{\Omega(\sqrt{n/\log n})}$ [GG00]. In 2003, Micciancio and Vadhan introduced a statistical zero-knowledge proof for $\mathrm{GapCVP}_{\Omega(\sqrt{n/\log n})}$ and $\mathrm{GapSVP}_{\Omega(\sqrt{n/\log n})}$ [MV03]. Recently, Goldwasser and Kharchenko published a proof of plaintext knowledge for the Ajtai-Dwork cryptosystem [GK05].

1.2 Motivation

First, we remark progress of quantum computation. Using Shor's algorithm [Sho97], most number-theoretical cryptosystems are insecure against quantum adversary. Here, we must study cryptosystems that are secure against quantum adversary. Many researchers consider that combinational problems are hard even in quantum computation, and pay attention to lattice-based cryptosystems. Recently, International Workshop on Post-Quantum Cryptography (PQCrypto 2006) [Eur06] was held, which covered lattice-based cryptosystems, multivariate cryptosystems, and quantum algorithms.

Next, we revisit efficiency of public-key encryption schemes. Let n be a security parameter. In most number-theoretical public-key encryption schemes, such as RSA, ElGamal, Cramer-Shoup, etc., the size of public-key is $O(n)$, the size of plaintext is $O(n)$, the size of ciphertext is $O(n)$, and the time of encryption is $O(n^3)$. In a few number-theoretical public-key encryption schemes, such as Goldwasser-Micali, the size of public-key is $O(n)$, the size of plaintext is 1, the size of ciphertext is $O(n)$, and the time of encryption is $O(n^2)$. See Table 1.1 for the efficiency of cryptosystems.

	Number-Theoretical (1)		Number-Theoretical (2)	
cryptosystem	RSA [RSA78]	ElGamal [ElG85]	Goldwasser-Micali [GM84]	
security	Unknown	DDH	Factoring of RSA modules	
size of public key	$O(n)$	$O(n)$	$O(n)$	
size of private key	$O(n)$	$O(n)$	$O(n)$	
size of plaintext	$O(n)$	$O(n)$	1	
size of ciphertext	$O(n)$	$O(n)$	$O(n)$	
time of encryption	$O(n^3)$	$O(n^3)$	$O(n^2)$	
	Lattice-Based (A)		Lattice-Based (B)	
cryptosystem	GGH [GGH97b]	NTRU [HPS98]	ADGGH [GGH97a]	R05 [Reg05]
security	Unknown	Unknown	$O(n^{11})$-uSVP	$\mathrm{SVP}_{\tilde{O}(n^{1.5})}$
size of public key	$O(n^2 \log n)$	$O(n \log n)$	$O(n^5 \log n)$	$O(n^2 \log^2 n)$
size of private key	$O(n^2 \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$
size of plaintext	$O(n)$	$O(n)$	1	1
size of ciphertext	$O(n \log n)$	$O(n \log n)$	$O(n^2 \log n)$	$O(n \log n)$
time of encryption	$O(n^2 \log n)$	$O(n \log n)$	$O(n^2 \log n)$	$O(n^2 \log n)$

Table 1.1: summary.

Consequently, in number-theoretical cryptosystems, the sizes is small but the speed is slow. On the other hand, in lattice-based cryptosystems, the size is big but the speed is fast. We

expect increasing storage capacity will allow us to use lattice-based cryptosystems practically. However, the size of plaintext of the lattice-based cryptosystems in type B is only 1 bit and impede practical use of cryptosystems in type B. Therefore, increasing the size of plaintext in type B is one of important issues.

Applications of lattice-based cryptosystems are also important issues. There exist many applications, such that signature, identification, proof of knowledge, and etc, based on number-theoretic public-key cryptosystems. As seen in Section 1.1, there exist a few applications based on lattice-based cryptosystems. Construction of applications needs research of properties of cryptosystems and primitive tools, such as zero knowledge and proof of knowledge. Thus, we have to study zero knowledge and proof of knowledge for lattice-based cryptosystems and properties of lattice-based cryptosystems.

1.3 Our Contribution

In this thesis we study applications of lattice-based cryptosystems which belong to the type B. For simplicity, we call the cryptosystems proposed in [GGH97a, Reg04, Reg05, Ajt05] ADGGH, R04, R05, and A05, respectively.

1: Multi-bit Public-Key Cryptosystems Based on Lattice Problems and Their Pseudohomomorphism. The first is efficient lattice-based cryptosystems with security proofs based on well-known lattice problems or other secure cryptosystems. Specifically, we propose a universal technique which admits four lattice-based cryptosystems, ADGGH, R04, R05, and A05, to encrypt a multi-bit plaintext without changing the size of the ciphertext. Furthermore, we study their pseudohomomorphisms, the property of the sum of ciphertexts. For more details, see Chapter 3.

2: A Modified Regev'05 Cryptosystem, Proofs of Knowledge on Its Secret Key, and Signature Schemes. Secondly, we propose a modified R05 and a proof of knowledge on its secret key. Although there already exist public-key identification schemes based on lattice problems, it is not known that its public key can be used as a public key of an encryption scheme. We need modify the original R05 to obtain a public-key identification scheme. We also propose concrete lattice-based signature scheme, obtaining by using the Fiat-Shamir transformation [FS86]. The security in the random oracle model follows theorems in [PS96, OO98, AABN02]. See Chapter 4 for more details.

3: Proofs of Plaintext Knowledge for the Regev'04 and Regev'05 Cryptosystems. At last, we propose proofs of plaintext knowledge for R04 and R05 which are based on the proof of plaintext knowledge for the Ajtai-Dwork cryptosystem in [GK05]. In the construction, we use the result of (1), tradeoffs and pseudohomomorphisms. We also remark that Goldwasser and Kharchenko's technique can not apply to the original R04 and R05. Applying it to the original cryptosystems will need new techniques. See Chapter 5 for details.

1.4 Organization

The rest of this thesis is organized as follows: We first recall basic notions and notations, and briefly review tools in Chapter 2. Chapter 3 describes the multi-bit versions of four lattice-based cryptosystems and their pseudohomomorphism property. We show the modification of the Regev'05 cryptosystem, the proof of knowledge on its secret key, and the lattice-based signature scheme in Chapter 4. We discuss the proof of plaintext knowledge of the Regev'04 and Regev'05 cryptosystems in Chapter 5.

Chapter 2

Preliminaries

In this chapter, we denotes notions, notations and definitions. The organization of this chapter is follows: Section 2.1 denotes fundamental notions and notations. We review the definition and notions of lattices, and list up lattice problems and their complexities in Section 2.2. Section 2.3 describes Gaussian and other distributions which we use in this thesis. We review the definition and notions of codes in Section 2.4. Finally, we recall the definition of zero knowledge and proof of knowledge in Section 2.5.

2.1 Fundamental Notions and Notations

We define a negligible amount in n as an amount that is asymptotically smaller than n^{-c} for any constant $c > 0$. More formally, $f(n)$ is a negligible function in n if $lim_{n\to\infty} n^c f(n) = 0$ for any $c > 0$. Similarly, a non-negligible amount is one which is at least n^{-c} for some $c > 0$.

For m-bit string $r \in \{0,1\}^m$, r_i denotes i-th bit of r (i.e., $r = r_1 \dots r_m$). We define $\mathbf{I}_n$ as the n by n identity matrix. We also define $\mathbf{u}_i \in \mathbb{R}^n$ as an n-dimensional vector whose i-th coordinate is 1 and other coordinates are all 0. The length of a vector $\mathbf{x} = {}^t(x_1, \dots, x_n) \in \mathbb{R}^n$, denoted by $\|\mathbf{x}\|$, is $(\sum_{i=1}^n x_i^2)^{1/2}$. For any field K, the inner product of two vectors $\mathbf{x} = {}^t(x_1, \dots, x_n) \in K^n$ and $\mathbf{y} = {}^t(y_1, \dots, y_n) \in K^n$, denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$, is $\sum_{i=1}^n x_i y_i$. Let $w_H(\mathbf{x})$ denote the Hamming weight of $\mathbf{x}$, i.e., the number of nonzero elements in $\mathbf{x}$. For any vector $\mathbf{x} \in \mathbb{R}^n$ and a set $S \subseteq \mathbb{R}^n$ we define $\mathrm{Dist}(\mathbf{x}, S) = \inf_{\mathbf{y}\in S} \|\mathbf{y} - \mathbf{x}\|$. Let $B_n(\mathbf{c}, r)$ denote an n-dimensional hyperball whose center is $\mathbf{c} \in \mathbb{R}^n$ and radius is $r \geq 0$, that is, $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{c}\| \leq r\}$.

Let $\lfloor x \rceil$ be the closest integer to $x \in \mathbb{R}$ (if there are two such integers, we choose the smaller.) and $\mathrm{frc}(x) = |x - \lfloor x \rceil|$ for $x \in \mathbb{R}$, i.e., $\mathrm{frc}(x)$ is the distance from x to the closest integer. We define $x \bmod y$ as $x - \lfloor x/y \rfloor y$ for $x, y \in \mathbb{R}$. For an element $x \in \mathbb{Z}_q$ we define $|x|_q$ as the integer x if $x \in \{0, 1, \dots, \lfloor p/2 \rfloor\}$ and as the integer $q - x$ otherwise. In other words, $|x|_q$ represents the

distance of x from 0 in $\mathbb{Z}_q$.

We call probability p exponentially close to 1 if $p = 1 - 2^{-\Omega(n)}$. We represent a real number by rounding its fractional part. If the fractional part of $x \in \mathbb{R}$ is represented in m bits, the rounded number $\bar{x}$ has the precision of $1/2^m$, i.e., we have $|x - \bar{x}| \leq 1/2^m$. The security parameter n of lattice-based cryptosystems is equal to dimension of a lattice in the lattice problems on which security of the cryptosystems are based. We say that an algorithm distinguishes between two distributions if the gap between the acceptance probability for their samples is non-negligible.

2.2 Lattice Problems

An n-dimensional lattice in $\mathbb{R}^n$ is the set $L(\mathbf{b}_1, \ldots, \mathbf{b}_n) = \{\sum_{i=1}^n \alpha_i \mathbf{b}_i : \alpha_i \in \mathbb{Z}\}$ of all integral combinations of n linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$. The sequence of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is called a *basis* of the lattice L. For clarity of notations, we represent a basis by the matrix $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n) \in \mathbb{R}^{n \times n}$. For any basis $\mathbf{B}$, we define the *fundamental parallelepiped* $\mathcal{P}(\mathbf{B}) = \{\sum_{i=1}^n \alpha_i \mathbf{b}_i : 0 \leq \alpha_i < 1\}$. The vector $\mathbf{x} \in \mathbb{R}^n$ reduced modulo the parallelepiped $\mathcal{P}(\mathbf{B})$, denoted by $\mathbf{x} \bmod \mathcal{P}(\mathbf{B})$, is the unique vector $\mathbf{y} \in \mathcal{P}(\mathbf{B})$ such that $\mathbf{y} - \mathbf{x} \in L(\mathbf{B})$. The dual lattice L^* of a lattice L is the set $L^* = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z} \text{ for all } \mathbf{y} \in L\}$. If L is generated by basis $\mathbf{B}$, then $({}^t\mathbf{B})^{-1}$ is a basis for the dual lattice, where ${}^t\mathbf{B}$ is the transpose of $\mathbf{B}$. For more details on lattices, see the textbook by Micciancio and Goldwasser [MG02].

We list up well-known hard problems used for lattice-based cryptosystems. Recall that the length of vectors is defined by the l_2 norm in this thesis.

The shortest vector problem (SVP) and its approximation version (SVP_γ) have been deeply studied in the computer science.

Definition 2.2.1 (SVP)**.** Given a basis $\mathbf{B}$ of a lattice L, find a non-zero vector $\mathbf{v} \in L$ such that for any non-zero vector $\mathbf{x} \in L$, $\|\mathbf{v}\| \leq \|\mathbf{x}\|$.

Definition 2.2.2 (SVP_γ)**.** Given a basis $\mathbf{B}$ of a lattice L, find a non-zero vector $\mathbf{v} \in L$ such that for any non-zero vector $\mathbf{x} \in L$, $\|\mathbf{v}\| \leq \gamma \|\mathbf{x}\|$.

The NP-hardness of SVP was shown by Ajtai [Ajt98] under a randomized reduction in 1998. Recently, Khot [Kho04] proved that SVP_c is NP-hard under the assumption NP $\not\subseteq$ RP for any constant c. He also proved that $\mathrm{SVP}_{2^{O((\log n)^{1/2-\varepsilon})}}$ is NP-hard within under the assumption NP $\not\subseteq$ RTIME$(2^{\mathrm{poly}(\log n)})$.

Even within a polynomial approximation factor, it is unknown whether there exists a polynomial-time algorithm for the approximation version of SVP. The most well-known solu-

tion to this approximation problem is the so-called LLL algorithm proposed in [LLL82]. This algorithm can solve $\mathrm{SVP}_{2^{n/2}}$ in polynomial time.

On the other hand, there are several non-NP-hardness results on the approximation version of SVP with a polynomial approximation factor. Goldreich and Goldwasser [GG00] showed $\mathrm{SVP}_{\Omega(\sqrt{n/\log n})}$ is in NP ∩ coAM. Aharonov and Regev [AR05] showed that $\mathrm{SVP}_{\Omega(\sqrt{n})}$ is in NP ∩ coNP.

The unique shortest vector problem (uSVP) is also well known as a hard lattice problem applicable to cryptographic constructions. We say the shortest vector $\mathbf{v}$ of a lattice L is f-unique if for any non-zero vector $\mathbf{x} \in L$ which is not parallel to $\mathbf{v}$, $f\,\|\mathbf{v}\| \leq \|\mathbf{x}\|$. The definition of uSVP is given as follows.

Definition 2.2.3 (f-uSVP)**.** Given a basis $\mathbf{B}$ of a lattice L whose shortest vector is f-unique, find a non-zero vector $\mathbf{v} \in L$ such that for any non-zero vector $\mathbf{x} \in L$ which is not parallel to $\mathbf{v}$, $f\,\|\mathbf{v}\| \leq \|\mathbf{x}\|$.

Similarly to the case of SVP, the exact version of uSVP is shown to be in NP-hard by Kumar and Sivakumar [KS01]. Cai [Cai98] showed that $\Omega(n^{1/4})$-uSVP is in NP∩coAM. See Figure 2.1 for the hardness of SVP and uSVP.

In the computational complexity theory on lattice problems, the shortest linearly independent vectors problem (SIVP) and its approximation version SIVP_γ are also considered as a hard lattice problem.

Definition 2.2.4 (SIVP)**.** Given a basis $\mathbf{B}$ of a lattice L, find a sequence of n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in L$ such that for any sequence of n linearly independent vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in L$, $\max_{i=1,\ldots,n} \|\mathbf{v}_i\| \leq \max_{i=1,\ldots,n} \|\mathbf{x}_i\|$.

Definition 2.2.5 (SIVP_γ)**.** Given a basis $\mathbf{B}$ of a lattice L, find a sequence of n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in L$ such that for any sequence of n linearly independent vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in L$, $\max_{i=1,\ldots,n} \|\mathbf{v}_i\| \leq \gamma \max_{i=1,\ldots,n} \|\mathbf{x}_i\|$.

The closest vector problem (CVP) is also an important problem.

Definition 2.2.6 (CVP)**.** Given a basis $\mathbf{B}$ of a lattice L and a target vector $\mathbf{y}$, find a closest vector $\mathbf{v} \in L$ such that for any vector $\mathbf{x} \in L$, $\|\mathbf{y} - \mathbf{v}\| \leq \|\mathbf{y} - \mathbf{x}\|$.

Definition 2.2.7 (CVP_γ)**.** Given a basis $\mathbf{B}$ of a lattice L and a target vector $\mathbf{y}$, find a closest vector $\mathbf{v} \in L$ such that for any vector $\mathbf{x} \in L$, $\|\mathbf{y} - \mathbf{v}\| \leq \gamma\,\|\mathbf{y} - \mathbf{x}\|$.

We often consider its decisional promise problem.

Definition 2.2.8 (GapCVP$_\gamma$)**.** For $\gamma > 1$, instances of the promise closest vector problem GapCVP$_\gamma$ are tuples $(\mathbf{B}, \mathbf{y}, t)$ where $\mathbf{B}$ is a basis of a lattice L in $\mathbb{R}^n$, $t > 0$, and a vector $\mathbf{y} \in \mathbb{R}^n$. $(\mathbf{B}, \mathbf{y}, t)$ is a YES instance of the GapCVP$_\gamma$ if there exists a lattice vector $\mathbf{x} \in L$ such that $\|\mathbf{x} - \mathbf{y}\| \leq t$. $(\mathbf{B}, \mathbf{y}, t)$ is a NO instance of the GapCVP$_\gamma$ if there exists no lattice vector $\mathbf{x} \in L$ such that $\|\mathbf{x} - \mathbf{y}\| \geq \gamma t$.

Although the Diophantine Approximation (DA) was originally a number-theoretic problem, DA is deeply related to the lattice theory. (See, e.g., [GLS88].) The problem DA is defined as follows.

Definition 2.2.9 (DA)**.** Given n real numbers $r_1, \ldots, r_n$ and an integer M, find an integer $m \in [1, M^n]$ such that $\max_{i=1}^{n} \mathrm{frc}(mr_i) \leq 1/M$.

From a complexity-theoretical point of view, Lagarias [Lag85] showed that decisional version of DA is NP-complete. Trolin [Tro01] also showed a reduction from the decisional version of DA to a certain lattice problem. In the context of cryptography, Ajtai defined a variant of DA and constructed an efficient lattice-based cryptosystem based on the hardness of this variant [Ajt05]. We refer to this variant as DA′, defined as follows.

Definition 2.2.10 (DA′, [Ajt05])**.** Let $c_1, c_2 > 0$ be constants. Assume that $r_1, \ldots, r_n$ are samples from the uniform distribution on $(0, 1)$ with the condition that there exists an integer m such that

$$1 \leq m \leq n^{c_1 n} \text{ and } \mathrm{frc}(mr_i) \leq n^{-(c_1+c_2)} \text{ for } i = 1, \ldots, n.$$

Given n, $r_1, \ldots, r_n$, c_1, c_2, find such an integer m.

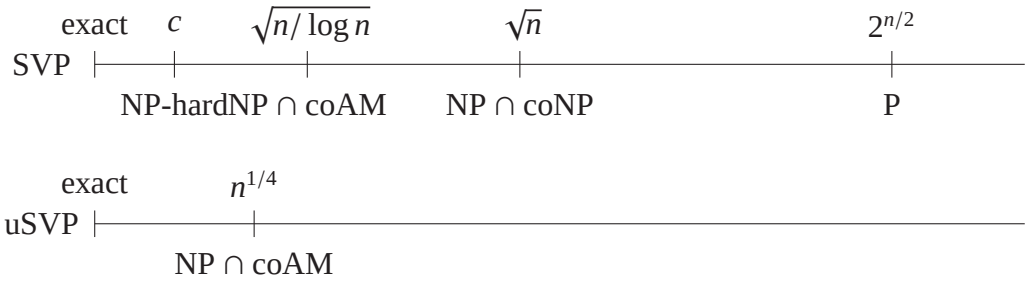

Figure 2.1: the complexity of SVP and uSVP.

2.3 Gaussian and Other Distributions

The normal distribution with mean 0 and variance σ^2 is the distribution on $\mathbb{R}$ given by the density function $\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right)$. For any distribution ϕ, we consider the distribution $\phi^{(n)}$ obtained as follows: (1) take n samples $x_1, \ldots, x_n$ from ϕ independently and (2) output ${}^t(x_1, \ldots, x_n)$. For a n-dimensional vector $\mathbf{x}$ and any $s > 0$, let $\rho_s^{(n)}(\mathbf{x}) = \exp(-\pi \|\mathbf{x}/s\|^2)$ be a Gaussian function scaled by a factor of s. Also, $\nu_s^{(n)} := \rho_s^{(n)}/s^n$ is an n-dimensional probability density function. For $\alpha \in \mathbb{R}^+$ the distribution Ψ_α is the distribution on $[0, 1)$ obtained by sampling from a normal variable with mean 0 and variance $\alpha^2/(2\pi)$ and reducing the result modulo 1:

$$\Psi_\alpha(r) := \sum_{k\in\mathbb{Z}} \frac{1}{\alpha}\exp\left(-\pi\left(\frac{r-k}{\alpha}\right)^2\right).$$

This distribution is obtained by "folding" a Gaussian distribution $N(0, \alpha^2/(2\pi))$ on $\mathbb{R}$ into the interval $[0, 1)$. Based on this distribution, the Regev'04 cryptosystem makes use of a periodic distribution $\Phi_{h,\alpha}$ defined by the density function $\Phi_{h,\alpha}(r) := \Psi_\alpha(rh \bmod 1)$. We can sample values according to this distribution by using samples from Ψ_α, as shown in [Reg04]: (1) We sample $x \in \{0, 1, \ldots, \lceil h\rceil\}$ uniformly at random and then (2) sample y according to Ψ_α. (3) If $0 \le (x + y)/h < 1$, we then take the value as a sample. Otherwise, we repeat. For an arbitrary probability distribution with density function $\phi : \mathbb{T} \to \mathbb{R}^+$ and some integer $q > 0$, we define its discretization $\bar{\phi} : \mathbb{Z}_q \to \mathbb{R}^+$ as the discrete probability distribution obtained by sampling from ϕ, multiplying by q, and rounding to the closest integer modulo q. More formally,

$$\bar{\phi}(i) := \int_{(i-1/2)q}^{(i+1/2)q} \phi(x)dx.$$

Given two probability density functions ϕ_1, ϕ_2 on $\mathbb{R}^n$, we define the statistical distance between them as $\Delta(\phi_1, \phi_2) := \frac{1}{2}\int_{\mathbb{R}^n} |\phi_1(\mathbf{x}) - \phi_2(\mathbf{x})|\, d\mathbf{x}$. A similar definition holds for discrete random variables. We sometimes abuse such notation, and use the same notation for two arbitrary functions. Note that the acceptance probability of any algorithm on inputs from X differs from its acceptance probability on inputs from Y by at most $\Delta(X, Y)$.

We use the following lemma in [Reg04] to bound the tail of Gaussian distribution.

Lemma 2.3.1 ([Reg04]). *The probability that the distance of a normal variable with variance σ^2 from its mean is more than t is at most $\sqrt{2/\pi}(\sigma/t)\exp(-(t/\sigma)^2/2)$. That is, $\Pr_{x\sim N(m,\sigma^2)}[|x - m| > t] \le \sqrt{2/\pi}(\sigma/t)\exp(-(t/\sigma)^2/2)$,*

We say that an algorithm $\mathcal{D}$ with oracle access is a distinguisher between two distributions if its acceptance probability when the oracle outputs samples of the first distribution and when the oracle outputs samples of the second distribution differ by a non-negligible amount.

2.4 Codes

Let $\mathbb{F}_q$ denote a field with q elements, where q is a prime power. A q-ary linear code C is a linear subspace of $\mathbb{F}_q^n$. If C has dimension k then C is called an $[n,k]_q$ code. A generator matrix $\mathbf{G}$ for a linear code C is a n by k matrix for which the columns are a basis of C. Note that $C := \{\mathbf{Gm} \mid \mathbf{m} \in \mathbb{F}_q^k\}$. We say that $\mathbf{G}$ is in standard form if $\mathbf{G} = \binom{\mathbf{I}_k}{\mathbf{P}}$. For an $[n,k]_q$ code C, we define the dual code $C^\perp$ by $C^\perp := \{\mathbf{y} \in \mathbb{F}_q^n \mid \text{ for any } \mathbf{x} \in C, \langle \mathbf{x}, \mathbf{y} \rangle = 0\}$. If $\mathbf{G} = \binom{\mathbf{I}_k}{\mathbf{P}}$ is a generator matrix in standard form of the code C, then $\mathbf{H} = \binom{-{}^t\mathbf{P}}{\mathbf{I}_{n-k}}$ is a generator matrix of the code $C^\perp$. This follows from the fact that $\mathbf{H}$ has the right size and rank and that ${}^t\mathbf{HG} = \mathbf{0}$, which implies every codeword $\mathbf{Gm}$ has inner product 0 with every column of $\mathbf{H}$. In other words, $\mathbf{x} \in C$ if and only if ${}^t\mathbf{Hx} = \mathbf{0}$. Thus, we call $\mathbf{H}$ a parity-check matrix. We note that, given any generator matrix $\mathbf{G}$ of the code C, we can efficiently compute C's generator matrix $\mathbf{G}'$ in standard form and C's parity-check matrix $\mathbf{H}$. If C is a linear code with a parity-check matrix $\mathbf{H}$ then for every $\mathbf{x} \in \mathbb{F}_q^n$ we call ${}^t\mathbf{Hx}$ the syndrome of $\mathbf{x}$.

It is well known that the question of finding the nearest codeword to a vector (Nearest Codeword Problem, NCP) is NP-hard even in approximation version [ABS97]. It is also difficult to find a word of a given weight from its syndrome [BMvT78].

Definition 2.4.1 (Symdrome Decoding Problem, SDP)**.** Given a parity-check matrix $\mathbf{H} \in \mathbb{Z}_2^{n \times m}$, a binary nonzero vector $\mathbf{y} \in \mathbb{Z}_2^m$, and a positive integer w, find a binary vector $\mathbf{x} \in \mathbb{Z}_2^n$ with no more than w 1's such that ${}^t\mathbf{Hx} = \mathbf{y}$.

2.5 Zero Knowledge and Proof of Knowledge

In this section, we recall the definitions and notations of zero knowledge and proof of knowledge.

Definition 2.5.1 (Auxiliary-Input Zero Knowledge)**.** An interactive proof system (P, V) for a language L is (perfect/statistical/computational) *auxiliary-input zero knowledge* if for every probabilistic polynomial-time machine V^* and polynomial $p(\cdot)$, there exists a probabilistic polynomial-time machine S such that the ensembles $\{(P, V^*(z))(x)\}$ and $\{S(x,z)\}$ are (perfectly/statistically/computationally) indistinguishable on the set $\{(x,z) : x \in L, |z| = p(|x|)\}$.

For a relation $R \subseteq \{0,1\}^* \times \{0,1\}^*$ and $x \in \{0,1\}^*$, we define a set of witness $R(x) := \{y : (x,y) \in R\}$.

Definition 2.5.2 (Proof of Knowlegde)**.** Let $\eta \in (0,1)$. An interactive protocol (P, V) with a prover P and a verifier V is a *proof of knowledge system with knowledge error κ for a relation R* if the following holds:

Completeness: For every common input x for which there exists y such that $(x, y) \in R$ the verifier V always accepts interacting with the prover P.

Validity with error η: There exists a polynomial-time interacting oracle Turing machine K and a constant $c > 0$ such that for every $x \in \{0,1\}^*$ such that $R(x) \neq \emptyset$ and for every prover P^* the following holds: $K^{P^*}(x) \in R(x) \cup \{\perp\}$ and $\Pr[K^{P^*}(x) \in R(x)] \geq (p - \kappa)^c$, where $p > \kappa$ is the probability that V accepts while interacting with P^* on common input x.

Chapter 3

Multi-bit Public-Key Cryptosystems Based on Lattice Problems and Their Pseudohomomorphism

3.1 Introduction

As seen in Chapter 1, the efficiency of lattice-based cryptosystems is an important problem. We continue to study efficient lattice-based cryptosystems with security proofs based on well-known lattice problems or other secure cryptosystems. In particular, we focus on the size of plaintexts encrypted by the cryptosystems in the type B. To the best of the authors' knowledge, all those in this type are single-bit cryptosystems. We therefore obtain more efficient lattice-based cryptosystems with security proofs if we succeed to construct their multi-bit versions without increase in the size of ciphertexts.

In this chapter, we consider multi-bit versions of the improved Ajtai-Dwork cryptosystem proposed by Goldreich, Goldwasser, and Halevi [GGH97a], the Regev cryptosystems given in [Reg04] and in [Reg05], and the Ajtai cryptosystem [Ajt05]. We develop a universal technique derived from a general structure behind them for constructing their multi-bit versions without increase in the size of ciphertexts.

Our technique requires precise evaluation of trade-offs between decryption errors and hardness of underlying lattice problems in the original lattice-based cryptosystems. We firstly give precise evaluation for the trade-offs to apply our technique to constructions of the multi-bit versions. This precise evaluation also clarifies a quantitative relationship between the security levels and the decryption errors in the lattice-based cryptosystems, which may be useful to improve the cryptosystems beyond our results.

	Ajtai-Dwork		Regev'04	
cryptosystem	ADGGH [GGH97a]	mbADGGH	R04 [Reg04]	mbR04
security	$O(n^{11})$-uSVP	$O(n^{11+\varepsilon})$-uSVP	$\tilde{O}(n^{1.5})$-uSVP	$\tilde{O}(n^{1.5+\varepsilon})$-uSVP
size of public key	$O(n^5 \log n)$	$O(n^5 \log n)$	$O(n^4)$	$O(n^4)$
size of private key	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^2)$
size of plaintext	1	$O(\log n)$	1	$O(\log n)$
size of ciphertext	$O(n^2 \log n)$	$O(n^2 \log n)$	$O(n^2)$	$O(n^2)$
rounding precision	2^{-n}	2^{-n}	2^{-8n^2}	2^{-8n^2}
	Regev'05		Ajtai	
cryptosystem	R05 [Reg05]	mbR05	A05 [Ajt05]	mbA05
security	$\mathrm{SVP}_{\tilde{O}(n^{1.5})}$	$\mathrm{SVP}_{\tilde{O}(n^{1.5+\varepsilon})}$	DA′	A05
size of public key	$O(n^2 \log^2 n)$	$O(n^2 \log^2 n)$	$O(n^2 \log n)$	$O(n^2 \log n)$
size of private key	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$
size of plaintext	1	$O(\log n)$	1	$O(\log n)$
size of ciphertext	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$	$O(n \log n)$
rounding precision	2^{-n}	2^{-n}	$1/n$	$1/n$

Table 3.1: summary. (ε is any positive constant and $\tilde{O}(f(n))$ means $O(f(n)\,\mathrm{poly}(\log n))$.)

Due to this evaluation of the cryptosystems, it is shown that our multi-bit versions encrypt $O(\log n)$-bit plaintexts into ciphertexts of the same length as the original ones with reasonable sacrifices of the hardness of the underlying lattice problems.

The ciphertexts of our multi-bit version are distributed in the same ciphertext space, theoretically represented with real numbers, as the original cryptosystem. To represent the real numbers in their ciphertexts, we have to round their fractional parts with certain precision. The size of ciphertexts then increases if we process the numbers with high precision. We stress that our technique does not need higher precision than that of the original cryptosystems, i.e., we take the same precision in our multi-bit versions as that of the original ones.

See Table 3.1 for the cryptosystems studied in this chapter. (The problems in the "security" fields are defined in Section 2.2.)

We call the cryptosystems proposed in [GGH97a, Reg04, Reg05, Ajt05] ADGGH, R04, R05, and A05, respectively. We also call the corresponding multi-bit versions mbADGGH, mbR04, mbR05, and mbA05.

We also focus on the algebraic property we call *pseudohomomorphism* of the lattice-based cryptosystems. The homomorphism of ciphertexts is quite useful for many cryptographic applications. (See, e.g., [Rap04].) In fact, the single-bit cryptosystems ADGGH, R04, R05 and

A05 implicitly have a similar property to the homomorphism. Let $E(x_1)$ and $E(x_2)$ be ciphertexts of x_1 and $x_2 \in \{0,1\}$, respectively. Then, $E(x_1) + E(x_2)$ becomes a variant of $E(x_1 \oplus x_2)$. More precisely, $E(x_1)+E(x_2)$ does not obey the distribution of the ciphertexts, but we can guarantee the same security level as that of the original cryptosystem and decrypt $E(x_1) + E(x_2)$ to $x_1 \oplus x_2$ by the original private key with a small decryption error. We refer to this property as the pseudohomomorphism. Goldwasser and Kharchenko actually made use of a similar property to construct the plaintext knowledge proof system for the Ajtai-Dwork cryptosystem [GK05] (See Chapter 5).

Unfortunately, it is only over $\mathbb{Z}_2$ (and direct product groups of $\mathbb{Z}_2$ by concatenating the ciphertexts) that we can operate the addition of the plaintexts in the single-bit cryptosystems. It is unlikely that we can naively simulate the addition over large cyclic groups by concatenating ciphertexts in such single-bit cryptosystems.

In this chapter, we present the pseudohomomorphic property of mbADGGH, mbR04, mbR05, and (a slightly modified version mbA05′ of) mbA05 over larger cyclic groups. We believe that this property extends the possibility of the cryptographic applications of the lattice-based cryptosystems.

Main Idea for Multi-Bit Constructions and Their Security. We can actually find the following general structure behind the single-bit cryptosystems ADGGH, R04, R05, and A05: Their ciphertexts of 0 are basically distributed according to a periodic Gaussian distribution and those of 1 are also distributed according to another periodic Gaussian distribution whose peaks are shifted to the middle of the period. We thus embed two periodic Gaussian distributions into the ciphertext space such that their peaks appear alternatively and regularly. (See the left side of Figure 3.1.)

Our technique is based on a generalization of this structure. More precisely, we regularly embed *multiple* periodic Gaussian distributions into the ciphertext space rather than only two ones. (See the right side of Figure 3.1.) Embedding p periodic Gaussian distributions as shown in this figure, the ciphertexts for a plaintext $i \in \{0, \ldots, p-1\}$ are distributed according the i-th periodic Gaussian distribution. This cyclic structure enables us not only to improve the efficiency of the cryptosystems but also to guarantee their security.

If we embed too many periodic Gaussian distributions, the decryption errors increase due to the overlaps of the distributions. We can then decrease the decryption errors by reducing their variance. However, it is known that smaller variance generally makes such cryptosystems less secure, as commented in [GGH97a]. We therefore have to evaluate the trade-offs in our multi-bit versions between the decryption errors and their security, which depend on their own structures of the cryptosystems.

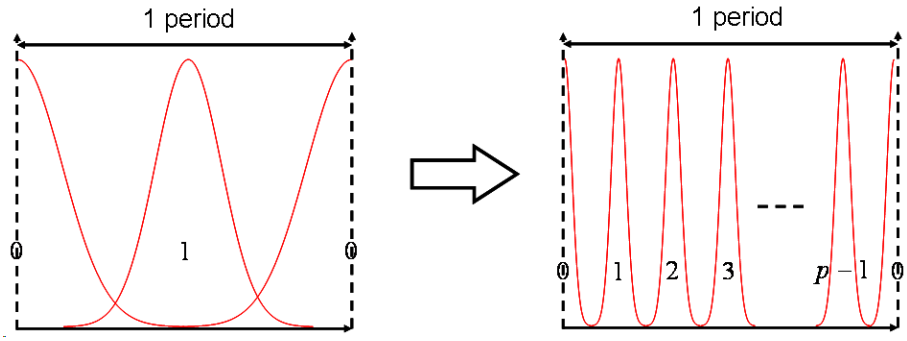

Figure 3.1: the embedding of periodic Gaussian distributions.

Once we evaluate their trade-offs, we can apply a general strategy based on the cyclic structure to the security proofs. The security of the original cryptosystems basically depends on the indistinguishability between a certain periodic Gaussian distribution Φ and a uniform distribution U since it is shown in their security proofs that we can construct an efficient algorithm for a certain hard lattice problem by employing an efficient distinguisher between Φ and U. The goal is thus to construct the distinguisher from an adversary against the multi-bit version.

We first assume that there exists an efficient adversary for distinguishing between two Gaussian distributions corresponding two kinds of ciphertexts in our multi-bit version with its public key. By the hybrid argument, the adversary can distinguish either between Φ_i and U or between Φ_j and U. We now suppose that it can distinguish between Φ_i and U. Note that we can slide Φ_i to Φ_0 corresponding to ciphertexts of 0 even if we do not know the private key by the cyclic property of the ciphertexts. Thus, we obtain an efficient distinguisher between Φ_0 and U. Φ_0 is in fact a variance-reduced version of the periodic Gaussian distribution Φ used in the original cryptosystem. We can guarantee the indistinguishability between such a version Φ_0 and U is based on the hardness of another lattice problem slightly easier than the original one. We can therefore guarantee the security of our multi-bit versions similarly to the original ones.

Encryption and Decryption in Multi-Bit Versions. We also exploit this cyclic structure for the correctness of encryption and decryption procedures. In the original cryptosystems except for R05, the private key is the period d of the periodic Gaussian distribution, and the public key consists of the information for generating the periodic Gaussian distribution corresponding to 0 and the information for shifting the distribution to the other distribution corresponding to 1. The latter information for the shift essentially is $k(d/2)$ for a random odd number k. Then, if we want to encrypt a plaintext 0, we generate the periodic Gaussian distribution corresponding

to 0. Also, if we want to encrypt 1, we generate the distribution corresponding to 0 and then shift it using the latter information.

The private and public keys in our multi-bit versions are slightly different from those of the original ones. The major difference is the information for shifting the distribution. If the size of the plaintext space is p, the information for the shift is essentially $k(d/p)$, where the number k must be a coprime to p for unique decryption. We then interpret the number k as a generator of the "group" of periodic Gaussian distributions. We adopt a prime as the size of the plaintext space p for efficient public key generation in our constructions. The private key also contains this number k other than the period d. Therefore, we can construct correct encryption and decryption procedures using this information k.

In the cases of R05 and mbR05, it is not necessary for keys to contain the information for the shift. We can actually obtain such information due to their own structures even if it is not given from the public key. Thus, p is not necessarily a prime in mbR05.

Pseudohomomorphism in Multi-Bit Versions. The regular embedding of the periodic Gaussian distributions also gives our multi-bit cryptosystems the algebraic property named *pseudohomomorphism*. Recall that a Gaussian distribution has the following reproducing property: For two random variables X_1 and X_2 according to $N(m_1, s_1^2)$ and $N(m_2, s_2^2)$, where $N(m, s^2)$ is a Gaussian distribution with mean m and standard deviation s, the distribution of $X_1 + X_2$ is equal to $N(m_1 + m_2, s_1^2 + s_2^2)$. This property implies that the sum of two ciphertexts (i.e., the sum of two periodic Gaussian distributions) becomes a variant of a ciphertext (i.e., a periodic Gaussian distribution with larger variance). This sum can be moreover decrypted into the sum of two plaintexts with the private key of the multi-bit version, and has the indistinguishability based on the security of the multi-bit version. By precise analysis of our multi-bit versions, we estimate the upper bound of the number of the ciphertexts which can be summed without the change of the security and the decryption errors.

3.2 A Multi-Bit Version of the Improved Ajtai-Dwork Cryptosystem

We discuss the improved Ajtai-Dwork cryptosystem ADGGH given by Goldreich, Goldwasser, and Halevi [GGH97a] in detail and apply our technique to construction of its multi-bit version mbADGGH in this section.

3.2.1 The Improved Ajtai-Dwork Cryptosystem and Its Multi-Bit Version

For understanding our construction intuitively, we first overview the protocol of ADGGH. Let $N = n^n = 2^{n \log n}$. We define an n-dimensional hypercube C and an n-dimensional ball B_r as $C = \{\mathbf{x} \in \mathbb{R}^n : 0 \leq x_i < N, i = 1, \ldots, n\}$ and $B_r = B_n(\mathbf{0}, n^{-r}/4) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq n^{-r}/4\}$ for any constant $r \geq 7$, respectively. For $\mathbf{u} \in \mathbb{R}^n$ and an integer i we define a hyperplane H_i as $H_i = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{u} \rangle = i\}$.

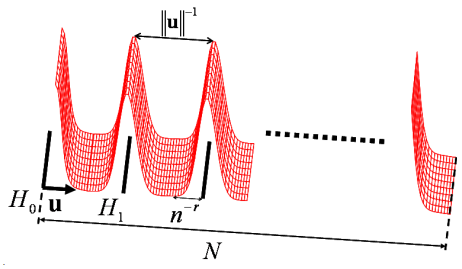

Figure 3.2: ciphertexts of 0 in ADGGH.

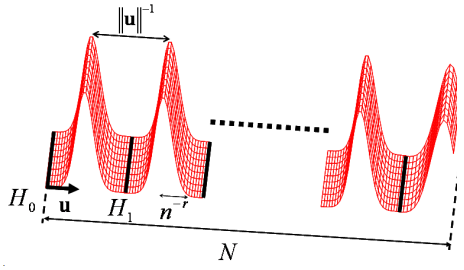

Figure 3.3: ciphertexts of 1 in ADGGH.

Roughly speaking, ADGGH encrypts 0 into a vector distributed closely around hidden $(n-1)$-dimensional parallel hyperplanes $H_0, H_1, H_2, \ldots$ for a normal vector $\mathbf{u}$ of H_0, and encrypts 1 into a vector distributed closely around their intermediate parallel hyperplanes $H_0 + \mathbf{u}/(2\|\mathbf{u}\|^2), H_1 + \mathbf{u}/(2\|\mathbf{u}\|^2), \ldots$. (See Figure 3.2 and Figure 3.3.) Then, the private key is the normal vector $\mathbf{u}$. These distributions of ciphertexts can be obtained from its public key, which consists of vectors on the hidden hyperplanes and information i_1 for shifting a vector on the hyperplanes to another vector on the intermediate hyperplanes. If we know the normal vector, we can reduce the n-dimensional distribution to on the 1-dimensional one along the normal vector. Then, we can easily find whether a ciphertext distributed around the hidden hyperplanes or the intermediate ones.

We now describe the protocol of ADGGH as follows. Our description slightly generalizes the original one by introducing a parameter r, which controls the variance of the distributions since we need to estimate a trade-off between the security and the size of plaintexts in our multi-bit version.

Cryptosystem 3.2.1 (ADGGH [AD97, GGH97a])**.** All the participants agree with the security parameter n, the variance-controlling parameter r, and the precision 2^{-n} for rounding real numbers.

Key Generation: We choose $\mathbf{u}$ uniformly at random from the n-dimensional unit ball. Let $m = n^3$. Repeating the following procedure m times, we sample m vectors

$\mathbf{v}_1, \ldots, \mathbf{v}_m$: (1) We choose $\mathbf{a}_i$ from $\{\mathbf{x} \in C : \langle \mathbf{x}, \mathbf{u} \rangle \in \mathbb{Z}\}$ uniformly at random, (2) choose $\mathbf{b}_1, \ldots, \mathbf{b}_n$ from B_r uniformly at random, (3) and output $\mathbf{v}_i = \mathbf{a}_i + \sum_{j=1}^{n} \mathbf{b}_j$ as a sample. We then take the minimum index i_0 satisfying that the width of $\mathcal{P}(\mathbf{v}_{i_0+1}, \ldots, \mathbf{v}_{i_0+n})$ is at least $n^{-2}N$, where width of a parallelepiped $\mathcal{P}(\mathbf{x}_1, \ldots \mathbf{x}_n)$ is defined as $\min_{i=1,\ldots,n} \mathrm{Dist}(\mathbf{x}_i, \mathrm{span}(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n))$ for a distance function $\mathrm{Dist}(\cdot, \cdot)$ between a vector and an $(n-1)$-dimensional hyperplane. Now let $\mathbf{w}_j = \mathbf{v}_{i_0+j}$ for every $j \in \{1, \ldots, n\}$, $V = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$, and $W = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$. We also choose an index i_1 uniformly at random from $\{i : \langle \mathbf{a}_i, \mathbf{u} \rangle \text{ is odd}\}$, where $\mathbf{a}_i$ is the vector appeared in the sampling procedure for $\mathbf{v}_i$. Note that there are such indices i_0 and i_1 with probability $1 - o(1)$. If such indices do not exist, we perform this procedure again. To guarantee the security, $\|\mathbf{u}\|$ should be in $[1/2, 1)$. The probability of this event is exponentially close to 1. If the condition is not satisfied, we sample the vector $\mathbf{u}$ again. Then, the private key is $\mathbf{u}$ and the public key is (V, W, i_1).

Encryption: Let S be a uniformly random subset of $\{1, 2, \ldots, m\}$. We encrypt a plaintext $\sigma \in \{0, 1\}$ to $\mathbf{x} = \frac{\sigma}{2}\mathbf{v}_{i_1} + \sum_{i \in S} \mathbf{v}_i \bmod \mathcal{P}(W)$.

Decryption: Let $\mathbf{x} \in \mathcal{P}(W)$ be a received ciphertext. We decrypt $\mathbf{x}$ to 0 if $\mathrm{frc}(\langle \mathbf{x}, \mathbf{u} \rangle) \leq 1/4$ and to 1 otherwise.

Carefully reading the results in [AD97, GGH97a], we obtain the following theorem on the cryptosystem ADGGH.

Theorem 3.2.2 ([GGH97a])**.** *The cryptosystem* ADGGH *encrypts a* 1*-bit plaintext into an* $n\lceil n(\log n + 1)\rceil$*-bit ciphertext with no decryption error. The security of* ADGGH *is based on the worst case of* $O(n^{r+5})$-uSVP *for* $r \geq 7$*. The size of the public key is* $O(n^5 \log n)$ *and the size of the private key is* $O(n^2)$*.*

As commented in [Cai03], we can actually improve the security of ADGGH by a result in [Cai03]. We give the precise proof in Section 3.2.5.

Theorem 3.2.3. *The security of* ADGGH *is based on the worst case of* $O(n^{r+4})$-uSVP *for* $r \geq 7$*.*

We next describe the multi-bit version mbADGGH of ADGGH. Let p be a prime such that $2 \leq p \leq n^{r-7}$, where the parameter r controls a trade-off between the size of the plaintext space and the hardness of underlying lattice problems. In mbADGGH, we can encrypt a plaintext of $\log p$ bits into a ciphertext of the same size as ADGGH. The strategy of our construction basically follows the argument in Section 3.1. Note that the parameter r is chosen to keep our version error-free.

Cryptosystem 3.2.4 (mbADGGH)**.** All the participants agree with the parameters n, r and the precision 2^{-n} similarly to ADGGH, and additionally the size p of the plaintext space.

Key Generation: The key generation procedure is almost the same as that of ADGGH. We choose an index i'_1 uniformly at random from $\{i : \langle \mathbf{a}_i, \mathbf{u} \rangle \not\equiv 0 \bmod p\}$ instead of i_1 in the original key generation procedure. We set decryption information $k \equiv \langle \mathbf{a}_{i'_1}, \mathbf{u} \rangle \bmod p$. Note that there is such a k with probability $1 - (1/p)^m = 1 - o(1)$. Then, the private key is $(\mathbf{u}, k)$ and the public key is (V, W, i'_1).

Encryption: Let S be a uniformly random subset of $\{0,1\}^m$. We encrypt $\sigma \in \{0, \ldots, p-1\}$ to $\mathbf{x} = \frac{\sigma}{p}\mathbf{v}_{i'_1} + \sum_{i \in S} \mathbf{v}_i \bmod \mathcal{P}(W)$.

Decryption: We decrypt a received ciphertext $\mathbf{x} \in \mathcal{P}(W)$ to $\lfloor p\langle \mathbf{x}, \mathbf{u} \rangle \rceil k^{-1} \bmod p$, where k^{-1} is the inverse of k in $\mathbb{Z}_p$.

Before evaluating the performance of mbADGGH precisely, we give the summary of the results as follows.

Theorem 3.2.5 (security and decryption errors)**.** *Let $r \geq 7$ be any constant and let $p(n)$ be a prime such that $2 \leq p(n) \leq n^{r-7}$. The cryptosystem* mbADGGH *encrypts a $\lfloor \log p(n) \rfloor$-bit plaintext into an $n\lceil n(\log n + 1)\rceil$-bit ciphertext without the decryption errors. The security of* mbADGGH *is based on the worst case of $O(n^{r+4})$-uSVP. The size of the public key is the same as that of the original one. The size of the private key is $\lceil \log p(n) \rceil$ plus that of the original one.*

Theorem 3.2.6 (pseudohomomorphism)**.** *Let $r \geq 7$ be any constant. Also, let p be a prime and let κ be an integer such that $\kappa p \leq n^{r-7}$. Let E_{m} be the encryption function of* mbADGGH*. For any κ plaintexts $\sigma_1, \ldots, \sigma_\kappa$ $(0 \leq \sigma_i \leq p-1)$, we can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i) \bmod \mathcal{P}(W)$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$ without decryption error. Moreover, if there exist two sequences of plaintexts $(\sigma_1, \ldots, \sigma_\kappa)$ and $(\sigma'_1, \ldots, \sigma'_\kappa)$, and a polynomial-time algorithm that distinguishes between $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i) \bmod \mathcal{P}(W)$ and $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i) \bmod \mathcal{P}(W)$ with its public key, then there exists a polynomial-time algorithm that solves $O(n^{r+4})$-uSVP in the worst case with non-negligible probability.*

3.2.2 Decryption Errors of mbADGGH

We first evaluate the decryption error probability in mbADGGH. The following theorem can be proven by a similar argument to the analysis of [AD97, GGH97a]. Since we generalize this theorem for analysis of the pseudohomomorphism in mbADGGH (Theorem 3.2.12), we here give a precise proof.

Theorem 3.2.7. *The cryptosystem* mbADGGH *makes no decryption errors.*

Proof. Since the decryption error probability for any ciphertext can be estimated by sliding the distribution to that of the ciphertext of 0, we first estimate the decryption error probability for the ciphertext of 0.

Let $H := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{u} \rangle \in \mathbb{Z}\}$. From the definition, $\mathrm{Dist}(\mathbf{v}_i, H) \leq n \cdot n^{-r}/4$ for $1 \leq i \leq m$. Thus, we can obtain $\mathrm{frc}(\langle \mathbf{v}_i, \mathbf{u} \rangle) \leq n^{1-r}/4$ and $\mathrm{frc}(\langle \sum_{i \in S} \mathbf{v}_i, \mathbf{u} \rangle) \leq n^{4-r}/4$. Next, we estimate an inner product between $\sum_{i \in S} \mathbf{v}_i \bmod \mathcal{P}(W)$ and $\mathbf{u}$. Let $\sum_{i \in S} \mathbf{v}_i = \mathbf{r} + \sum_{j=1}^{n} q_j \mathbf{w}_j$, where $\mathbf{r} \in \mathcal{P}(W)$. Since $\left\|\mathbf{w}_j\right\| \geq n^{-2}N$ and $p \leq n^{r-7}$, we have $\left|q_j\right| \leq n^5$ and

$$\mathrm{frc}(\langle \mathbf{r}, \mathbf{u} \rangle) \leq n \cdot n^5 \cdot \frac{1}{4} n^{1-r} + \frac{1}{4} n^{4-r} \leq \frac{5}{16} n^{7-r} \leq \frac{1}{2p}.$$

Therefore, we decrypt a ciphertext of 0 into 0 without decryption errors.

Now let ρ be a ciphertext of σ. Let $\mathbb{Z} \pm a := \{x \in \mathbb{R} : \mathrm{frc}(x) \leq a\}$ for $a \geq 0$ and $\mathbb{Z} + a \pm b := \{x \in \mathbb{R} : \mathrm{frc}(x - a) \leq b\}$ for $a, b \geq 0$. By a property of the key generation, we have $\langle \mathbf{v}_{i'_1}/p, \mathbf{u} \rangle \in \mathbb{Z} + k/p \pm n^{1-r}/4p$ and

$$\langle \rho, \mathbf{u} \rangle \in \mathbb{Z} + \frac{k}{p}\sigma \pm \frac{5}{16} n^{7-r} \pm \frac{1}{4p} n^{1-r}\sigma \pm \frac{1}{4} n^{4-r} \subset \mathbb{Z} + \frac{k}{p}\sigma \pm \frac{3}{8} n^{7-r}.$$

Therefore, we obtain $\langle \rho, \mathbf{u} \rangle \in \mathbb{Z} + k\sigma/p \pm 1/(2p)$ and decrypt ρ into σ without decryption errors. □

3.2.3 Security of mbADGGH

We next prove the security of mbADGGH. Let $U_{\mathcal{P}(W)}$ be a uniform distribution on $\mathcal{P}(W)$. We denote the encryption function of ADGGH by E defined as a random variable $E(\sigma, (V, W, i_1))$ for a plaintext σ and a public key (V, W, i_1). If the public key is obvious, we abbreviate $E(\sigma, (V, W, i_1))$ to $E(\sigma)$. Similarly, the encryption function E_{m} is defined for mbADGGH.

First, we show that the indistinguishability between two certain distributions is based on the worst-case hardness of uSVP. The following lemma can be obtained by combining Theorem 3.2.3 and the results in [AD97] and [GGH97a] with our generalization.

Lemma 3.2.8 ([AD97, GGH97a])**.** *If there exists a polynomial-time distinguisher between* $(E(0), (V, W, i_1))$ *and* $(U_{\mathcal{P}(W)}, (V, W, i_1))$*, there exists a polynomial-time algorithm for the worst case of* $O(n^{r+4})$-uSVP *for* $r \geq 7$.

We next present the indistinguishability between the ciphertexts of 0 in mbADGGH and $U_{\mathcal{P}(W)}$.

Lemma 3.2.9. *If there exists a polynomial-time algorithm* $\mathcal{D}_1$ *that distinguishes between* $(E_{\mathrm{m}}(0), (V, W, i'_1))$ *and* $(U_{\mathcal{P}(W)}, (V, W, i'_1))$*, there exists a polynomial-time algorithm* $\mathcal{D}_2$ *that distinguishes between* $(E(0), (V, W, i_1))$ *and* $(U_{\mathcal{P}(W)}, (V, W, i_1))$.

Proof. We denote by $\varepsilon(n)$ the non-negligible gap of the acceptance probability of $\mathcal{D}_1$ between $E_{\mathrm{m}}(0)$ and $U_{\mathcal{P}(W)}$ with its public key. We will construct the distinguisher $\mathcal{D}_2$ from the given algorithm $\mathcal{D}_1$. To run $\mathcal{D}_1$ correctly, we first find the index i'_1 by estimating the gap of acceptance probability between $E_{\mathrm{m}}(0)$ and $U_{\mathcal{P}(W)}$ with the public key. If we can find i'_1, we output the result of $\mathcal{D}_1$ using i'_1 with the public key. Otherwise, we output a uniformly random bit. For random inputs of ciphertexts and public keys, the above procedure can distinguish between them.

We now describe the details of $\mathcal{D}_2$ as follows. We denote by $\mathbf{x}$ and (V, W, i_1) a ciphertext and a public key of ADGGH given as an input for $\mathcal{D}_2$, respectively. Also, let $p_0 = \Pr[\mathcal{D}_1(E_{\mathrm{m}}(0), (V, W, j)) = 1]$ and $p_U = \Pr[\mathcal{D}_1(U_{\mathcal{P}(W)}, (V, W, j)) = 1]$, where the probability p_0 is taken over the inner random bits of the encryption procedure and p_U is taken over $U_{\mathcal{P}(W)}$.

(D1) For every $j \in \{1, \ldots, m\}$, we run $\mathcal{D}_1(E_{\mathrm{m}}(0), (V, W, j))$ and $\mathcal{D}_1(U_{\mathcal{P}(W)}, (V, W, j))$ $T = n/\varepsilon^2$ times. Let $x_0(j)$ and $x_U(j)$ be the number of 1 in the outputs of $\mathcal{D}_1$ for the ciphertexts of 0 and the uniform distribution with the index j, respectively.

(D2) If there exists the index j' such that $|x_0(j') - x_U(j')| / T > \varepsilon/2$, we take j' as the component of the public key.

(D3) We output $\mathcal{D}_1(\mathbf{x}, (V, W, j'))$ if we find j'. Otherwise, we output a uniformly random bit.

Note that we have $|p_0 - x_0(j')/T| \leq \varepsilon/4$ and $|p_U - x_U(j')/T| \leq \varepsilon/4$ with probability exponentially close to 1 by the Hoeffding bound [Hoe63]. Therefore, we succeed to choose the index j' with which $\mathcal{D}_1$ can distinguish between the target distributions with probability exponentially close to 1 if j' exists. By the above argument, $\mathcal{D}_1$ works correctly for a non-negligible fraction of all the inputs. □

The next lemma can be proven by the hybrid argument.

Lemma 3.2.10. *If there exist $\sigma_1, \sigma_2 \in \{0, \ldots, p-1\}$ and a polynomial-time algorithm $\mathcal{D}_3$ that distinguishes between $(E_{\mathrm{m}}(\sigma_1), (V, W, i'_1))$ and $(E_{\mathrm{m}}(\sigma_2), (V, W, i'_1))$, there exists a polynomial-time algorithm $\mathcal{D}_4$ that distinguishes between $(E_{\mathrm{m}}(0), (V, W, i'_1))$ and $(U_{\mathcal{P}(W)}, (V, W, i'_1))$.*

Proof. By the hybrid argument, the distinguisher $\mathcal{D}_3$ can distinguish between $E_{\mathrm{m}}(\sigma_1)$ and $U_{\mathcal{P}(W)}$ or between $E_{\mathrm{m}}(\sigma_2)$ and $U_{\mathcal{P}(W)}$ with its public key. Without loss of generality, we can assume that $\mathcal{D}_3$ can distinguish between $E_{\mathrm{m}}(\sigma_1)$ and $U_{\mathcal{P}(W)}$ with its public key. Note that we have $E_{\mathrm{m}}(\sigma_1, (V, W, i'_1)) = E_{\mathrm{m}}(0, (V, W, i'_1)) + \frac{\sigma_1}{p}\mathbf{v}_{i'_1} \bmod \mathcal{P}(W)$ by the definition of E_{m}. Then, we can transform a given $\mathbf{x}$ from $E_{\mathrm{m}}(0, (V, W, i'_1))$ to another sample $\mathbf{y}$ from $E_{\mathrm{m}}(\sigma_1, (V, W, i'_1))$. We can therefore obtain the polynomial-time algorithm $\mathcal{D}_4$ that distinguishes between $(E_{\mathrm{m}}(0), (V, W, i'_1))$ and $(U_{\mathcal{P}(W)}, (V, W, i'_1))$. □

By the above three lemmas, we obtain the security proof for our multi-bit version mbADGGH.

Theorem 3.2.11. *If there exist plaintexts $\sigma_1, \sigma_2 \in \{0, \ldots, p-1\}$ and a polynomial-time algorithm that distinguishes between the ciphertexts of σ_1 and σ_2 of* mbADGGH *with its public key, there exists a polynomial-time algorithm for the worst-case of $O(n^{r+4})$-*uSVP *for $r \geq 7$.*

3.2.4 Pseudohomomorphism of mbADGGH

As stated in Theorem 3.2.6, mbADGGH has the pseudohomomorphic property. To demonstrate this property, we have to evaluate the decryption errors for sum of ciphertexts and prove its security.

Decryption Errors for Sum of Ciphertexts. First, we evaluate the decryption errors when we apply the decryption procedure to the sum of ciphertexts in mbADGGH. Recall that $\mathbb{Z} \pm a := \{x \in \mathbb{R} : \mathrm{frc}(x) \leq a\}$ for $a \geq 0$ and $\mathbb{Z} + a \pm b := \{x \in \mathbb{R} : \mathrm{frc}(x-a) \leq b\}$ for $a, b \geq 0$.

Theorem 3.2.12. *Let $r \geq 7$ be any constant. Also let p be a prime and κ be an integer such that $\kappa p \leq n^{r-7}$. For any κ plaintexts $\sigma_1, \ldots, \sigma_\kappa$ $(0 \leq \sigma_i \leq p-1)$, we can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i) \bmod \mathcal{P}(W)$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$ without the decryption errors.*

Proof. We define $\rho_1, \ldots, \rho_\kappa$ as ciphertexts of $\sigma_1, \ldots, \sigma_\kappa$, respectively. We will show that we can decrypt $\rho := \sum_{i=1}^{\kappa} \rho_i \bmod \mathcal{P}(W)$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$. From the proof of Theorem 3.2.7, we have

$$\langle \rho_i, \mathbf{u} \rangle \in \mathbb{Z} + \frac{k}{p}\sigma_i \pm \frac{3}{8}n^{7-r}.$$

Hence, we obtain

$$\langle \sum_{i=1}^{\kappa} \rho_i, \mathbf{u} \rangle \in \mathbb{Z} + \frac{k}{p}\sum_{i=1}^{\kappa} \sigma_i \pm \frac{3}{8}\kappa n^{7-r}.$$

Combining with the fact $\rho_i \in \mathcal{P}(W)$ and $\kappa p \leq n^{r-7}$, we have

$$\langle \rho, \mathbf{u} \rangle \in \mathbb{Z} + \frac{k}{p}\sum_{i=1}^{\kappa} \sigma_i \pm \frac{3}{8}\kappa n^{7-r} \pm \frac{1}{4}\kappa n^{2-r} \subset \mathbb{Z} + \frac{k}{p}\sum_{i=1}^{\kappa} \sigma_i \pm \frac{1}{2}\kappa n^{7-r} \subset \mathbb{Z} + \frac{k}{p}\sum_{i=1}^{\kappa} \sigma_i \pm \frac{1}{2p}.$$

Therefore, we correctly decrypt ρ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$. □

Security for Sum of Ciphertexts. We can also give the security proof for the sum of ciphertexts in mbADGGH. The security proof obeys so general framework that we can apply the same argument to the security of sum of ciphertexts in the other multi-bit versions mbR04,

mbR05, and mbA05′. For convenience of the other multi-bit versions, we here present an abstract security proof for sum of ciphertexts. We denote the encryption function of our multi-bit cryptosystems by E_{m}, also regarded as a random variable $E_{\mathrm{m}}(\sigma, pk)$ for a plaintext σ and a public key pk. If the public key is obvious, we abbreviate $E_{\mathrm{m}}(\sigma, pk)$ to $E_{\mathrm{m}}(\sigma)$. Let C be the ciphertext space and U_C be the uniform distribution on C.

We first show that it is hard to distinguish between the sum of ciphertexts and the uniform distribution if it is hard to distinguish between κ samples from $E_{\mathrm{m}}(0)$ and those from U_C.

Lemma 3.2.13. *If there exist two sequences of plaintexts* $(\sigma_1, \ldots, \sigma_\kappa)$ *and* $(\sigma'_1, \ldots, \sigma'_\kappa)$ *and a polynomial-time algorithm* $\mathcal{D}_1$ *that distinguishes between* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i), pk)$ *and* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i), pk)$*, then there exists a polynomial-time algorithm* $\mathcal{D}_2$ *that distinguishes between* κ *ciphertexts and its public key* $(E_{\mathrm{m}}(0, pk), \ldots, E_{\mathrm{m}}(0, pk), pk)$ *and uniformly random* κ *ciphertexts and the public key* $(U_C, \ldots, U_C, pk)$.

Proof. By the hybrid argument, the distinguisher $\mathcal{D}_1$ can distinguish between $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i)$ and U_C or between $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i)$ and U_C with its public key. Without loss of generality, we can assume that $\mathcal{D}_1$ can distinguish between $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i), pk)$ and (U_C, pk). By $(\sigma_1, \ldots, \sigma_\kappa)$, we can transform $(E_{\mathrm{m}}(\sigma_1), \ldots, E_{\mathrm{m}}(\sigma_\kappa), pk)$ into $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i), pk)$. This shows the polynomial-time distinguisher $\mathcal{D}_2$. □

As already stated in Section 3.1 (and Lemma 3.2.9 in the case of ADGGH), the original security proofs of ADGGH, R04, R05 and A05 show that we have efficient algorithms for certain lattice problems if there is an efficient distinguisher between $E_m(0)$ and U_C with its public key. By the similar argument to that in original proofs, we also have such algorithms from efficient distinguisher $\mathcal{D}_2$ between $(E_m(0), \ldots, E_m(0), pk)$ and $(U_C, \ldots, U_C, pk)$. Thus, we obtain from $\mathcal{D}_2$ in Lemma 3.2.13 a probabilistic polynomial-time algorithm $\mathcal{A}$ that solve the worst case of $O(n^{r+4})$-uSVP in the case of mbADGGH.

By combining the above discussion with Lemma 3.2.13, we guarantee the security of the sum of ciphertexts in mbADGGH.

Theorem 3.2.14. *If there exist two sequences of plaintext* $(\sigma_1, \ldots, \sigma_\kappa)$ *and* $(\sigma'_1, \ldots, \sigma'_\kappa)$ *and a polynomial-time algorithm* $\mathcal{D}_1$ *that distinguishes between* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i), pk)$ *and* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i), pk)$*, then there exists a probabilistic polynomial-time algorithm* $\mathcal{A}$ *that solves the worst case of* $O(n^{r+4})$-uSVP *in the case of* mbADGGH.

3.2.5 Proof of Theorem 3.1.3

For the proof of Theorem 3.2.3, we first describe the transference theorems.

Transference theorems

Let $B(r)$ be an n-dimensional ball in $\mathbb{R}^n$ centered at $\mathbf{0}$ with radius r, i.e., $B(r) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq r\}$.

Definition 3.2.15 (Minkowski's successive minima)**.** For an n-dimensional lattice L in $\mathbb{R}^n$ the *i-th successive minima* $\lambda_i(L)$ is defined as follows:

$$\lambda_i(L) = \min_{\mathbf{v}_1,\ldots,\mathbf{v}_i \in L} \max_{1 \leq j \leq i} \left\|\mathbf{v}_j\right\|,$$

where the sequence of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_i \in L$ ranges over all i linearly independent lattice vectors.

It is not hard to show that

$$\lambda_i(L) = \min\{r : \max_{\mathbf{v}_1,\ldots,\mathbf{v}_i \in L \cap B(r)} \dim(\mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_i)) = i\}.$$

Banaszczyk showed the following transference theorem in [Ban93].

Theorem 3.2.16 ([Ban93])**.** *For every n-dimensional lattice L and every constant $c > 3/2\pi$,*

$$\lambda_i(L) \cdot \lambda_{n-i+1}(L^*) \leq cn,$$

for all sufficiently large n.

We say a sublattice $L' \subseteq L$ is a *saturated sublattice* if $L' = L \cap \mathrm{span}(L')$, where $\mathrm{span}(L')$ is the linear subspace of $\mathbb{R}^n$ spanned by the basis of L'. For $1 \leq i \leq n$, we define $g_i(L)$ to be the minimum r such that the sublattice generated by $L \cap B(r)$ contains an i-dimensional saturated sublattice L'. Clearly, $\lambda_i(L) \leq g_i(L)$ for $1 \leq i \leq n$.

Cai improved Theorem 3.2.16 as the following theorem.

Theorem 3.2.17 ([Cai03])**.** *For every an n-dimensional lattice L and for every constant $c > 3/2\pi$,*

$$\lambda_i(L) \cdot g_{n-i+1}(L^*) \leq cn,$$

for all sufficiently large n.

Main Proof

Now, we give the proof of Theorem 3.2.3.

Proof of Theorem 3.2.3. The proof is similar to the argument of [AD96, AD97]. Let $H_{\mathbf{u}}$ be the distribution of $\mathbf{v}_i$ in the key generation procedure of ADGGH. Ajtai and Dwork gave the following two lemmas.

Lemma 3.2.18 (Lemma 8.1, [AD97]). *If there exists a probabilistic polynomial-time algorithm $\mathcal{D}_1$ such that distinguishes between $E(0)$ and $U_{\mathcal{P}(W)}$ with (V, W), there exists a probabilistic polynomial-time algorithm $\mathcal{D}_2$ such that distinguishes between $H_{\mathbf{u}}$ and U_C, where U_C is an uniform distribution on C.*

Lemma 3.2.19 (Lemma 8.2, [AD97]). *If there exists a probabilistic polynomial-time algorithm $\mathcal{D}_2$ such that distinguishes between $H_{\mathbf{u}}$ and U_C, there exists a probabilistic polynomial-time algorithm $\mathcal{A}$ such that solve the worst case of $f(n)$-uSVP.*

We now evaluate the value of $f(n)$. Given an instance of $f(n)$-uSVP, we obtain a lattice L by certain linear transformations shown in [AD97] such that we can efficiently compute its shortest vector $\mathbf{u}$ if there exists an efficient attacking algorithm for ADGGH. Then, the dual lattice $J = L^*$ of L has a saturated sublattice J' on a hyperplane H_0 orthogonal to $\mathbf{u}$. Let l be the length of the smallest basis of J', where the length of the basis $\mathbf{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is defined as $\max_{i=1,\ldots,n} \|\mathbf{v}_i\|$.

It is also commented in [AD97] that the length l of the smallest basis of J' is approximately $O(n^2/f(n))$. It also holds that this upper bound is $O(n^{-r-3})$ by combining the argument in [AD97] with our generalization. Thus, we obtain $f(n) = O(n^{r+5})$.

On the other hand, we obtain $\lambda_2(L) \cdot g_{n-1}(L^*) \leq cn$ by Theorem 3.2.17 with $i = 2$, i.e., $\lambda_2(L) \cdot l \leq cn$ for some constant $c > 3/2\pi$. We can also see that $\lambda_2(L) \geq f(n)\|\mathbf{u}\|$ from the definition. Thus, we can obtain an upper bound $O(n/f(n))$ of l.

By the above argument, we obtain $f(n) = O(n^{r+4})$, which completes the proof of Theorem 3.2.3. □

3.3 A Multi-Bit Version of the Regev'04 Cryptosystem

3.3.1 The Regev'04 Cryptosystem and Its Multi-Bit Version

In this section we consider the Regev cryptosystem R04 proposed in [Reg04]. Roughly speaking, the ciphertexts of 0 and 1 approximately corresponds to two periodic Gaussian distributions in R04. (See Figure 3.4 and Figure 3.5.) We now denote the distributions of the ciphertexts of 0 and 1 as Φ_0 and Φ_1, respectively. Note that every peak in Φ_1 is regularly located in the middle of two peaks in Φ_0. A parameter h is approximately equal to the number of peaks in Φ_0, and a private key d, obtained from h, corresponds to length of the period. A public key is of the form $(a_1, \ldots, a_m, i_0)$, where $a_1, \ldots, a_m$ are samples from Φ_0 to make a ciphertext of 0 by summing up randomly chosen elements from the samples and a certain index $i_0 \in \{1, \ldots, m\}$ is used to shift a ciphertext of 0 to that of 1 by adding $a_{i_0}/2$ to a ciphertext of 0. One can easily see that we can

distinguish between Φ_0 and Φ_1 with d. It however seems hard to distinguish them only with polynomially many samples of Φ_0 and i_0. Actually, it is shown in [Reg04] that breaking R04 is at least as hard as the worst case of a certain uSVP.

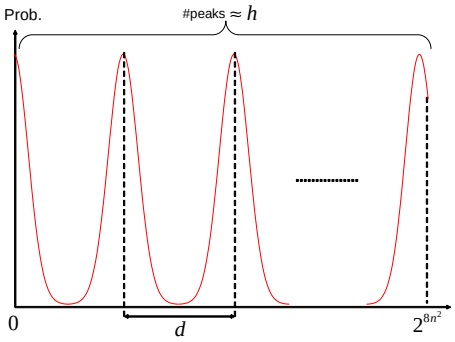

Figure 3.4: ciphertexts of 0 in R04.

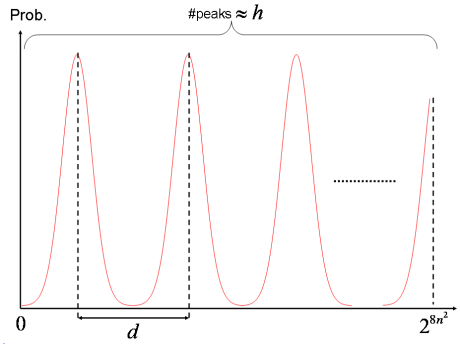

Figure 3.5: ciphertexts of 1 in R04.

In what follows, we precisely describe the original R04. First, we recall the definition of a folded Gaussian distribution Ψ_α whose density function is $\Psi_\alpha(l) = \sum_{k\in\mathbb{Z}}(1/\alpha)\exp(-\pi((l-k)/\alpha)^2)$. This distribution is obtained by "folding" a Gaussian distribution $N(0,\alpha^2/(2\pi))$ on $\mathbb{R}$ into the interval $[0,1)$. Note that this folded Gaussian distribution is equivalent with the fractional part of $N(0,\alpha^2/(2\pi))$. Based on this distribution, R04 makes use of a periodic distribution $\Phi_{h,\alpha}$ defined by the following density function: $\Phi_{h,\alpha}(l) = \Psi_\alpha(lh \bmod 1)$. We can sample values according to this distribution by using samples from Φ_α, as shown in [Reg04]: (1) We sample $x \in \{0,\ldots,\lceil h\rceil\}$ uniformly at random and then (2) sample y according to Ψ_α. (3) If $0 \le (x+y)/h < 1$, we then take the value as a sample. Otherwise, we repeat (1) and (2).

Let $N = 2^{8n^2}$, $m = c_0 n^2$ for a sufficiently large constant c_0, and $\gamma(n) = \omega(n\sqrt{\log n})$, specifying the size of the ciphertext space, the size of the public keys, and the variance of the folded Gaussian distribution, respectively. In this section, we require precision of $1/2^{8n^2} = 1/N$ for rounding real numbers.

Cryptosystem 3.3.1 (R04, [Reg04])**.** All the participants agree with the security parameter n and the precision 2^{-8n^2}.

Key Generation: Let $H = \{h \in [\sqrt{N}, 2\sqrt{N}) : \mathrm{frc}(h) < 1/(16m)\}$. We choose $h \in H$ uniformly at random and set $d = N/h$. The private key is the number d. Choosing $\alpha \in [2/\gamma(n), (2\sqrt{2})/\gamma(n))$, we sample m values $z_1,\ldots,z_m$ from the distribution $\Phi_{h,\alpha}$, where $z_i = (x_i + y_i)/h$ $(i = 1,\ldots,m)$ according to the above sampling procedure. Let $a_i = \lfloor Nz_i \rfloor$ for every $i \in \{1,\ldots,m\}$. Note that we have an index i_0 such that x_{i_0} is odd with a probability exponentially close to 1. Then, the public key is $(a_1,\ldots,a_m,i_0)$.

Encryption: We choose a uniformly random subset S of $\{1,\dots,m\}$. The ciphertext is $\sum_{i\in S} a_i \bmod N$ if the plaintext is 0, and $(\sum_{i\in S} a_i + \lfloor a_{i_0}/2 \rfloor) \bmod N$ if it is 1.

Decryption: We decrypt a received ciphertext $w \in \{0,\dots,N-1\}$ to 0 if $\mathrm{frc}(w/d) < 1/4$ and to 1 otherwise.

Summarizing the results in [Reg04] on the size of plaintexts, ciphertexts, and keys, the decryption errors, and the security of R04, Regev proved the following theorem.

Theorem 3.3.2 ([Reg04])**.** *The cryptosystem* R04 *encrypts a* 1*-bit plaintext into an* $8n^2$*-bit ciphertext with decryption error probability at most* $2^{-\Omega(\gamma^2(n)/m)} + 2^{-\Omega(n)}$*. The security of* R04 *is based on the worst case of* $O(\gamma(n)\sqrt{n})$-uSVP*. The size of the public key is* $O(n^4)$ *and the size of the private key is* $O(n^2)$*.*

We next propose a multi-bit version mbR04 of the cryptosystem R04. Let p be a prime such that $2 \le p \le n^r$ and $\delta(n) = \omega(n^{1+r}\sqrt{\log n})$ for any constant $r > 0$, where the parameter r controls the trade-off between the decryption errors (or the size of plaintext space) and the hardness of underlying lattice problems. Our cryptosystem mbR04 can encrypt one of p plaintexts in $\{0,\dots,p-1\}$ into a ciphertext of the same size as one of R04.

As mentioned above, R04 relates the ciphertexts to two periodic Gaussian distributions Φ_0 and Φ_1 such that each of them has one peak in a period of length d. Our construction follows the argument in Section 3.1. The idea of our cryptosystem is embedding of p periodic Gaussian distributions $\Phi_0,\dots,\Phi_{p-1}$ corresponding to the plaintexts $\{0,\dots,p-1\}$ into the same period of length d. We also adjust the parameter α, which affects the variance of the Gaussian distributions, to bound the decryption errors. Note that $\mathrm{frc}(h)$ also affects the decryption errors. Therefore, adjusting the set H simultaneously with α, we have to reduce the decryption errors by $\mathrm{frc}(h)$. Based on the above idea, we describe our cryptosystem mbR04 as follows.

Cryptosystem 3.3.3 (mbR04)**.** All the participants agree with the parameters n and r, the precision 2^{-8n^2}, and the size p of the plaintext space.

Key Generation: Let $H_r = \{h \in [\sqrt{N}, 2\sqrt{N}) : \mathrm{frc}(h) < 1/(8n^r m)\}$. We choose $h \in H_r$ uniformly at random and set $d = N/h$. Choosing $\alpha \in [2/\delta(n), (2\sqrt{2})/\delta(n))$, we sample m values $z_1,\dots,z_m$ from the distribution $\Phi_{h,\alpha}$, where $z_i = (x_i + y_i)/h$ $(i = 1,\dots,m)$ according to the above sampling procedure. Let $a_i = \lfloor Nz_i \rfloor$ for every $i \in \{1,\dots,m\}$. Additionally, we choose an index i_0' uniformly at random from $\{i : x_i \not\equiv 0 \bmod p\}$. Then, we compute $k \equiv x_{i_0'} \bmod p$. The private key is (d,k) and the public key is $(a_1,\dots,a_m,i_0')$.

Encryption: Let $\sigma \in \{0,\dots,p-1\}$ be a plaintext. We choose a uniformly random subset S of $\{1,\dots,m\}$. The ciphertext is $\left(\sum_{i\in S} a_i + \sigma\left\lfloor a_{i_0'}/p \right\rfloor\right) \bmod N$.

Decryption: For a received ciphertext $w \in \{0, \ldots, N-1\}$, we compute $\tau = w/d \bmod 1$. We decrypt the ciphertext w to $\lfloor p\tau \rceil k^{-1} \bmod p$, where k^{-1} is the inverse of k in $\mathbb{Z}_p$.

Before evaluating the performance of mbR04 precisely, we give the summary of the results as follows.

Theorem 3.3.4. *For any constant $r > 0$, let $\delta(n) = \omega(n^{1+r}\sqrt{\log n})$ and let $p(n)$ be a prime such that $2 \le p(n) \le n^r$. The cryptosystem* mbR04 *encrypts a $\lfloor \log p(n) \rfloor$-bit plaintext into an $8n^2$-bit ciphertext with decryption error probability at most $2^{-\Omega\left(\delta^2(n)/(n^{2r}m)\right)} + 2^{-\Omega(n)}$. The security of* mbR04 *is based on the worst case of $O(\delta(n)\sqrt{n})$-*uSVP*. The size of a public key is the same as that of the original one. The size of a private key is $\lceil \log p(n) \rceil$ plus that of the original one.*

For example, setting $\delta(n) = n^{1+r}\log n$ for any constant $r > 0$, we obtain an $\lfloor r \log n \rfloor$-bit cryptosystem with negligible decryption error, whose security is based on the worst-case of $O(n^{1.5+r}\log n)$-uSVP.

Theorem 3.3.5 (pseudohomomorphism)**.** *Let $\delta(n) = \omega(n^{1+r}\sqrt{\log n})$. Also let $p(n)$ be a prime and κ an integer such that $\kappa p \le n^r$ for any constant $r > 0$. Let E_{m} be the encryption function of* mbR04*. For any κ plaintexts $\sigma_1, \ldots, \sigma_\kappa$ $(0 \le \sigma_i \le p-1)$, we can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i) \bmod N$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega((\delta(n))^2/n^{2r}m)}$. Moreover, if there exist two sequences of plaintexts $(\sigma_1, \ldots, \sigma_\kappa)$ and $(\sigma'_1, \ldots, \sigma'_\kappa)$, and a polynomial-time algorithm that distinguishes between $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i) \bmod N$ and $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i) \bmod N$ with its public key, then there exists a polynomial-time algorithm that solves $O(\delta(n)\sqrt{n})$-*uSVP *in the worst case with non-negligible probability.*

In what follows, we demonstrate the performance of mbR04 stated in the above theorems.

3.3.2 Decryption Errors of mbR04

We give the analysis of the decryption errors without the proof since it can be done by a quite similar analysis to [Reg04] and we will prove the generalized theorem (Theorem 3.3.11) in Section 3.3.4.

Theorem 3.3.6. *The probability of the decryption errors in* mbR04 *is at most $2^{-\Omega\left(\delta^2(n)/(n^{2r}m)\right)} + 2^{-\Omega(n)}$.*

3.3.3 Security of mbR04

In what follows, we evaluate the security of our cryptosystem mbR04. We first mention the result in [Reg04] that the indistinguishability of two certain distributions is guaranteed by the

hardness of a certain uSVP. Let U_N and U_1 be the uniform distributions over $\{0, \ldots, N-1\}$ and $[0, 1)$, respectively.

Lemma 3.3.7 ([Reg04])**.** *If there exists a polynomial-time distinguisher between $\Phi_{h,\alpha}$ and U_1 over uniformly random choices of $h \in [\sqrt{N}, 2\sqrt{N})$ and $\alpha \in [2/\delta(n), 2\sqrt{2}/\delta(n))$, there exists a polynomial-time algorithm for the worst case of $O(\delta(n)\sqrt{n})$-uSVP.*

Thus, our task is to prove the security of our cryptosystem mbR04 from this indistinguishability. For convenience of the proof, we introduce a parameterized version R04′ of the cryptosystem R04. In the key generation procedure of R04′, we sample h from $H_r = \{h \in [\sqrt{N}, 2\sqrt{N}) : \mathrm{frc}(h) < 1/(8n^r m)\}$ and α from $[2/\delta, 2\sqrt{2}/\delta)$ uniformly at random. The other procedures in R04′ are the same as R04. Similarly to the case of R04, we can show that the indistinguishability between the ciphertexts of 0 in R04′ and U_N can be guaranteed by the indistinguishability between $\Phi_{h,\alpha}$ and U_N.

Lemma 3.3.8. *For any constant $r > 0$, let p be a prime such that $2 \le p \le n^r$ and $\delta(n) = \omega(n^{1+r}\sqrt{\log n})$. If there exists a polynomial-time algorithm that distinguishes between ciphertexts of 0 in* R04′ *and U_N with its public key, there exists a polynomial-time algorithm between $\Phi_{h,\alpha}$ and U_1 over uniformly random choices of $h \in [\sqrt{N}, 2\sqrt{N})$ and $\alpha \in [2/\delta(n), 2\sqrt{2}/\delta(n))$.*

This lemma can be proven by the same way as [Reg04] using the fact that $8n^r m \in \mathrm{poly}(n)$. By the same technique as the security proof of mbADGGH, we obtain the following lemma.

Lemma 3.3.9. *If there exist plaintexts $\sigma_1, \sigma_2 \in \{0, \ldots, p-1\}$ and a polynomial-time algorithm that distinguishes between the ciphertexts of σ_1 and σ_2 in* mbR04 *with its public key, there exists a polynomial-time algorithm that distinguishes between the ciphertexts of* 0 *in* R04′ *and U_N with its public key.*

By the above lemmas, we can show the security of mbR04 based on the hardness of uSVP.

Theorem 3.3.10. *If there exist plaintexts $\sigma_1, \sigma_2 \in \{0, \ldots, p-1\}$ and a polynomial-time algorithm that distinguishes between the ciphertexts of σ_1 and σ_2 in* mbR04 *with its public key, there exists a polynomial-time algorithm for the worst-case of $O(\delta(n)\sqrt{n})$-uSVP.*

3.3.4 Pseudohomomorphism of mbR04

Decryption Errors for Sum of Ciphertexts.

Theorem 3.3.11 (mbR04)**.** *Let $\delta(n) = \omega(n^{1+r}\sqrt{\log n})$. Also let $p(n)$ be a prime and κ be an integer such that $\kappa p \le n^r$ for any constant $r > 0$. For any κ plaintexts $\sigma_1, \ldots, \sigma_\kappa$ ($0 \le \sigma_i \le p-1$), we can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i) \bmod N$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega((\delta(n))^2/n^{2r}m)}$.*

Before the proof, recall Lemma 2.3.1 denoting a bound of the tails of Gaussian distributions. By Lemma 2.3.1, one can see easily that if $\sigma \leq 1/\sqrt{n}$, the probability $\Pr_{X\sim N(0,\sigma^2)}[|X| > 1/2]$ is exponentially small in n.

Proof. The proof is similar to the estimation of the decryption errors in [Reg04]. First, we show the case that we have κ ciphertexts of 0, $\rho_1, \ldots, \rho_\kappa$. The probabilities are taken over the choices of the private and public keys and the inner random bits of the encryption procedure. Let $S_1, \ldots, S_\kappa$ denote the subsets of indices used in the encryption procedure, i.e., $\rho_i = \sum_{j\in S_i} a_j \bmod N$. Let $\rho := \sum_{i=1}^{\kappa} \rho_i \bmod N$. Thus,

$$\left|\rho - \left(\sum_{i=1}^{\kappa}\left(\sum_{j\in S_i} a_j \bmod d\lfloor h\rceil\right) \bmod d\lfloor h\rceil\right)\right| \leq m\kappa|N - d\lfloor h\rceil| = m\kappa d \cdot \mathrm{frc}(h) < \frac{\kappa}{8n^r}d.$$

Similarly to the argument for evaluation of the decryption errors in [Reg04], we obtain

$$\begin{aligned}\mathrm{frc}\left(\frac{\rho}{d}\right) &< \frac{\kappa}{8n^r} + \mathrm{frc}\left(\frac{\sum_{i=1}^{\kappa}\left(\sum_{j\in S_i} a_i \bmod d\lfloor h\rceil\right) \bmod d\lfloor h\rceil}{d}\right)\\ &= \frac{\kappa}{8n^r} + \mathrm{frc}\left(\frac{\sum_{i=1}^{\kappa}\sum_{j\in S_i} a_j}{d}\right)\\ &< \frac{\kappa}{8n^r} + \frac{m\kappa}{d} + \mathrm{frc}\left(\frac{N}{d}\sum_{i=1}^{\kappa}\sum_{j\in S_i} z_j\right),\end{aligned}$$

where in the last inequality we use the fact that $a_j := \lfloor Nz_j \rfloor$. Since $z_j = (x_j + y_j)/h$ and $d = N/h$,

$$\mathrm{frc}\left(\frac{N}{d}\sum_{i=1}^{\kappa}\sum_{j\in S_i} z_j\right) = \mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i}(x_j + y_j)\right) = \mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} y_j\right).$$

Hence, we have

$$\mathrm{frc}\left(\frac{\rho}{d}\right) < \frac{\kappa}{8n^r} + \frac{m\kappa}{d} + \mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} y_j\right) < \frac{3\kappa}{16n^r} + \mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} y_j\right),$$

where we used the fact that $d = 2^{\Theta(4n^2)}$ is much larger than $m = c_0 n^2$. All x_i are strictly less than $\lceil h\rceil - 1$ with probability exponentially close to 1. Conditioned on that, $y_1, \ldots, y_m$ are distributed according to Ψ_α. Therefore, we have

$$\Pr\left[\mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} y_j\right) > \frac{1}{16p}\right] \leq \Pr\left[\mathrm{frc}\left(\sum_{j=1}^{m}\kappa y_j\right) > \frac{1}{16p}\right].$$

The distribution of $\sum_{j=1}^{m}\kappa y_j \bmod 1$ is $\Psi_{\sqrt{m\kappa}\alpha}$. Since $\sqrt{m\kappa}\alpha = O(\frac{\sqrt{m\kappa}}{\delta(n)})$, we obtain

$$\Pr\left[\mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} y_j\right) > \frac{1}{16p}\right] \leq 2^{-\Omega((\delta(n))^2/m\kappa p^2)} \leq 2^{-\Omega((\delta(n))^2/n^{2r}m)}$$

by Lemma 2.3.1. We thus obtain $\mathrm{frc}(\rho/d) < 1/(4p)$, which implies that we can decrypt ρ to 0 with decryption error probability at most $2^{-\Omega((\delta(n))^2/mn^{2r})}$.

Next, we consider κ ciphertexts $\rho'_1, \ldots, \rho'_\kappa$ of plaintexts $\sigma_1, \ldots, \sigma_\kappa$ respectively and set $\rho' := \sum_{i=1}^{\kappa} \rho'_i \bmod N$. From the encryption procedure, $\rho'_i = \rho_i + \sigma_i \left\lfloor a_{i'_0}/p \right\rceil \bmod N$. By using the fact that $k \equiv x_{i'_0} \bmod p$ and that $y_{i'_0} \in \mathbb{Z} \pm 1/(8n^r)$ with probability exponentially close to 1 , we get $\left\lfloor a_{i'_0}/p \right\rceil /d \in \mathbb{Z} + k/p \pm 1/(8pn^r) \pm 2/d$. Hence, we have $\sigma_i \left\lfloor a_{i'_0}/p \right\rceil /d \in \mathbb{Z} + \sigma_i k/p \pm 1/(8n^r) \pm 2p/d$. This implies that

$$\sum_{i=1}^{\kappa} \frac{\sigma_i \left\lfloor a_{i'_0}/p \right\rceil}{d} \in \mathbb{Z} + \frac{k}{p} \sum_{i=1}^{\kappa} \sigma_i \pm \frac{\kappa}{8n^r} \pm \frac{2\kappa p}{d}.$$

Since $\mathrm{frc}(\rho/d) < 1/(4p)$, we obtain

$$\frac{\rho'}{d} \in \mathbb{Z} + \frac{k}{p} \sum_{i=1}^{\kappa} \sigma_i \pm \frac{1}{4p} \pm \frac{\kappa}{8n^r} \pm \frac{\kappa+1}{8n^r m} \pm \frac{2\kappa p}{d} \subset \mathbb{Z} + \frac{k}{p} \sum_{i=1}^{\kappa} \sigma_i \pm \frac{1}{2p}$$

with the probability at most $2^{-\Omega((\delta(n))^2/mn^{2r})}$, which completes the proof. □

Security for Sum of Ciphertexts. By a similar argument in Section 3.2.4, we obtain the following theorem.

Theorem 3.3.12. *If there exist two sequences of plaintext* $(\sigma_1, \ldots, \sigma_\kappa)$ *and* $(\sigma'_1, \ldots, \sigma'_\kappa)$ *and a polynomial-time algorithm* $\mathcal{D}_1$ *that distinguishes between* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i), pk)$ *and* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i), pk)$*, then there exists a probabilistic polynomial-time algorithm* $\mathcal{A}$ *that solves the worst case of* $O(\delta(n)\sqrt{n})$-uSVP *in the case of* mbR04.

3.4 A Multi-Bit Version of the Regev'05 Cryptosystem

3.4.1 The Regev'05 Cryptosystem and Its Multi-Bit Version

The cryptosystem R05 proposed in 2005 [Reg05] is also constructed by using a variant of Gaussian distributions. A folded Gaussian distribution Ψ_α over $[0, 1)$ is given by a density function $\Psi_\alpha(l) = \sum_{k \in \mathbb{Z}} (1/\alpha) \exp(-\pi((l-k)/\alpha)^2)$. Let $m = 5(n+1)(2\log n + 1) = \Theta(n \log n)$ and $q(n) \in [n^2, 2n^2]$ be a prime. The parameter $\alpha = \alpha(n)$ satisfying conditions that $\alpha(n) = o(1/(\sqrt{n}\log n))$ and $\alpha(n)q(n) > 2\sqrt{n}$ is used to control the variance of the distribution Ψ_α. (In [Reg05], α is set to $1/(\sqrt{n}\log^2 n)$.) We also describe the discretized distribution on $\mathbb{Z}_q$ from Ψ_α. The Gaussian distribution $\bar{\Psi}_\alpha$ on $\mathbb{Z}_q$ is obtained by sampling from Ψ_α, multiplying q, and rounding the closest integer modulo q. The distribution can be formally defined as $\bar{\Psi}_\alpha(l) = \int_{(l-1/2)/q}^{(l+1/2)/q} \Psi_\alpha(x)dx$.

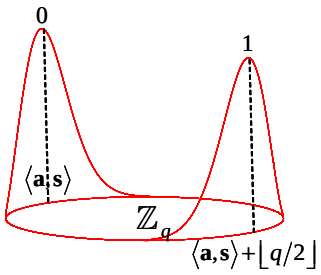

Figure 3.6: cryptosystem R05.

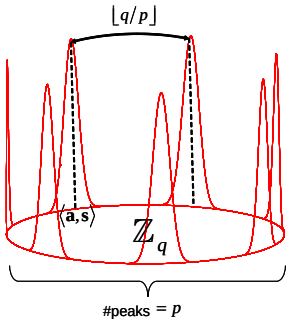

Figure 3.7: multi-bit version of R05.

In R05, the ciphertexts of 0 and 1 are vectors in $\mathbb{Z}_q^n$ obtained from some Gaussian distributions, which are specified by vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ shared among all the participants in the preparation procedure. Every coordinate i of the ciphertext of 0 corresponds to a Gaussian distribution on $\mathbb{Z}_q$ with mean $\langle \mathbf{a}_i, \mathbf{s} \rangle$ for the private key $\mathbf{s}$. On the other hand, the ciphertext of 1 corresponds to the "opposite" Gaussian distribution. (See Figure 3.6.)

Cryptosystem 3.4.1 (R05, [Reg05])**.** All the participants agree with the security parameter n, the variance-controlling parameter α, and the precision 2^{-n}. They also share m vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ chosen from $\mathbb{Z}_q^n$ uniformly at random.

Key Generation: The private key $\mathbf{s}$ is chosen uniformly at random from $\mathbb{Z}_q^n$. We also choose $e_1, \ldots, e_m$ according to the distribution $\bar{\Psi}_\alpha$. Let $b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i$ for every $i \in \{1, \ldots, m\}$. The public key is $\{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$.

Encryption: We choose a uniformly random subset S of $\{1, \ldots, m\}$. The ciphertext is $(\sum_{i \in S} \mathbf{a}_i, \sum_{i \in S} b_i)$ if the plaintext is 0, and $(\sum_{i \in S} \mathbf{a}_i, \lfloor q/2 \rfloor + \sum_{i \in S} b_i)$ if it is 1.

Decryption: We decrypt a received ciphertext $(\mathbf{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ into 0 if $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q < q/4$, and into 1 otherwise.

Note that the security reduction of R05 is done by a polynomial-time quantum algorithm. In other word, if R05 is insecure, there exists a polynomial-time quantum algorithm for certain lattice problems. As shown in [Reg05], the cryptosystem R05 has the following performance.

Theorem 3.4.2 ([Reg05])**.** *The cryptosystem* R05 *encrypts a* 1*-bit plaintext into an* $(n+1)\lceil \log q \rceil$*-bit ciphertext with decryption error probability at most* $2^{-\Omega\left(1/(m\alpha^2(n))\right)} + 2^{-\Omega(n)}$*. The security of* R05 *is based on the worst case of* $\mathrm{SVP}_{\tilde{O}(n/\alpha(n))}$ *and* $\mathrm{SIVP}_{\tilde{O}(n/\alpha(n))}$ *for polynomial-time quantum algorithms. The size of the public key is* $O(n \log^2 n)$ *and the size of the private key is* $O(n \log n)$*.*

We now give our cryptosystem mbR05 based on R05. (See Figure 3.7.) Let $r \in (0,1)$ be any constant, which controls the trade-off between the size of plaintext space and the hardness of underlying lattice problems, and p be an integer such that $p \le n^r = o(n)$, which is the size of the plaintext space in mbR05. mbR05 can encrypt a plaintext in $\{0,\dots,p-1\}$ into a ciphertext of the same size as R05. We use the same parameters m and q as R05 and introduce a parameter $\beta = \beta(n) = \alpha(n)/n^r = o(1/(n^{0.5+r}\log n))$ to control the distribution instead of α in R05. The parameter $\beta(n)$ must satisfy $\beta(n)q(n) > 2\sqrt{n}$.

Cryptosystem 3.4.3 (mbR05)**.** All the participants agree with the parameters n, β, the precision 2^{-n}, and the size p of the plaintext space. They also share m vectors $\mathbf{a}_1,\dots,\mathbf{a}_m$ chosen from $\mathbb{Z}_q^n$ uniformly at random.

Key Generation: This procedure is the same as R05 except that we sample $e_1,\dots,e_m$ from $\bar{\Psi}_\beta$.

Encryption: We choose a uniformly random subset S of $\{1,\dots,m\}$. For a plaintext $\sigma \in \{0,\dots,p-1\}$, the ciphertext is $(\sum_{i\in S}\mathbf{a}_i, \sigma\lfloor q/p\rceil + \sum_{i\in S} b_i)$.

Decryption: We decrypt a received ciphertext $(\mathbf{a}, b)$ to $\lfloor (b - \langle \mathbf{a}, \mathbf{s}\rangle)p/q\rceil \bmod p$.

Before evaluating the performance of mbR05 precisely, we give the summary of the results as follows.

Theorem 3.4.4. *Let $p = p(n)$ be an integer such that $p(n) \le n^r$ for any constant $0 < r < 1$. The cryptosystem* mbR05 *encrypts a $\lfloor \log p(n)\rfloor$-bit plaintext into an $(n+1)\lceil \log q\rceil$-bit ciphertext with decryption error probability at most $2^{-\Omega\left(1/(m\beta^2(n)n^{2r})\right)} + 2^{-\Omega(n)}$. The security of* mbR05 *is based on the worst case of $\mathrm{SVP}_{\tilde{O}(n/\beta(n))}$ and $\mathrm{SIVP}_{\tilde{O}(n/\beta(n))}$ for polynomial-time quantum algorithms. The size of the public key and private key is the same as that of the original one.*

For example, by setting $p(n) = n^r$ for a constant $0 < r < 1$ and $\beta(n) = 1/(n^{0.5+r}\log^2 n)$, we obtain a $\lfloor r \log n\rfloor$-bit cryptosystem with negligible decryption error whose security is based on $\mathrm{SVP}_{\tilde{O}(n^{1.5+r})}$ and $\mathrm{SIVP}_{\tilde{O}(n^{1.5+r})}$.

Theorem 3.4.5 (pseudohomomorphism)**.** *Let $p(n)$ be an integer and κ be an integer such that $\kappa p \le n^r$ for any constant $0 < r < 1$. Let E_{m} be the encryption function of* mbR05*. For any κ plaintexts $\sigma_1,\dots,\sigma_\kappa$ $(0 \le \sigma_i \le p-1)$, we can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i)$ into $\sum_{i=1}^{\kappa}\sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega(1/(m\beta^2(n)n^{2r}))}$, where the addition is defined over $\mathbb{Z}_q^n \times \mathbb{Z}_q$. Moreover, if there exist two sequences of plaintexts $(\sigma_1,\dots,\sigma_\kappa)$ and $(\sigma'_1,\dots,\sigma'_\kappa)$, and a polynomial-time algorithm that distinguishes between $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i)$ and $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i)$ with its public key, then there exist polynomial-time quantum algorithms that solve $\mathrm{SVP}_{\tilde{O}(n/\beta(n))}$ and $\mathrm{SIVP}_{\tilde{O}(n/\beta(n))}$ in the worst case with non-negligible probability.*

In what follows, we demonstrate the performance of mbR05 stated in the above theorems.

3.4.2 Decryption Errors of mbR05

We first estimate the decryption errors in our cryptosystem mbR05. By replacing the parameter α in R05 to the parameter β in mbR05, we immediately obtain the evaluation of the decryption errors from Theorem 3.4.2. The generalization of this theorem (Theorem 3.4.10) is also given in Section 3.4.4.

Theorem 3.4.6. *The probability of the decryption errors in* mbR05 *is at most* $2^{-\Omega\left(1/(m\beta^2(n)n^{2r})\right)} + 2^{-\Omega(n)}$.

3.4.3 Security of mbR05

We next discuss the security of our cryptosystem mbR05. Let U_{R05} be the uniform distribution over the ciphertext space $\mathbb{Z}_q^n \times \mathbb{Z}_q$ of R05 (and mbR05). The strategy of the security proof for mbR05 is similar to mbR04. We first mention the result in [Reg05] that the indistinguishability between the ciphertexts of 0 in R05 and U_{R05} is guaranteed by the worst-case hardness of certain lattice problems.

Lemma 3.4.7 ([Reg05])**.** *If there exists a polynomial-time algorithm that distinguishes between the ciphertexts of* 0 *in* R05 *and* U_{R05} *with its public key, there exists a polynomial-time quantum algorithm for the worst case of* $\mathrm{SVP}_{\tilde{O}(n/\alpha(n))}$ *and* $\mathrm{SIVP}_{\tilde{O}(n/\alpha(n))}$.

We now consider a slightly modified version R05′ with the distribution parameter $\beta(n) = \alpha(n)/n^r = o(1/(n^{0.5+r}\log n))$ instead of $\alpha(n)$ in R05. Since the trade-off between the decryption error and the security of R05′ is obtained by Theorem 3.4.2, we can show the following lemma by the same technique as the security proof of mbADGGH.

Lemma 3.4.8. *If there exist plaintexts* $\sigma_1, \sigma_2 \in \{0, \ldots, p-1\}$ *and a polynomial-time algorithm that distinguishes between the ciphertexts of* σ_1 *and* σ_2 *in* mbR05 *with its public key, there exists a polynomial-time algorithm that distinguishes between the ciphertexts of* 0 *in* R05′ *and* U_{R05} *with its public key.*

By these lemmas, we can obtain the security of our cryptosystem mbR05.

Theorem 3.4.9. *If there exist plaintexts* $\sigma_1, \sigma_2 \in \{0, \ldots, p-1\}$*, and a polynomial-time algorithm that distinguishes between the ciphertext of* σ_1 *and* σ_2 *in* mbR05 *with its public key, there exists a polynomial-time quantum algorithm for the worst-case of* $\mathrm{SVP}_{\tilde{O}(n/\beta(n))}$ *and* $\mathrm{SIVP}_{\tilde{O}(n/\beta(n))}$.

We omit the proof of the security since it is quite similar to mbADGGH.

3.4.4 Pseudohomomorphism of mbR05

Decryption Errors for Sum of Ciphertexts.

Theorem 3.4.10 (mbR05). *Let $\beta(n) = o(1/(n^{0.5+r}\log n))$. Also let $p(n)$ be an integer and κ be an integer such that $\kappa p \le n^r$ for any constant $0 < r < 1$. For any κ plaintexts $\sigma_1, \ldots, \sigma_\kappa$ $(0 \le \sigma_i \le p-1)$, we can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i)$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega(1/(m\beta^2(n)n^{2r}))}$, where the addition is defined over $\mathbb{Z}_q^n \times \mathbb{Z}_q$.*

Proof. The proof is similar to [Reg05]. First, we estimate the decryption errors for the sum of κ ciphertexts of 0, $(\boldsymbol{\rho}_1, \upsilon_1), \ldots, (\boldsymbol{\rho}_\kappa, \upsilon_\kappa)$. The probabilities are taken over the choices of the private and public keys and the randomness of the encryption procedure. Let $S_1, \ldots, S_\kappa$ denote the subsets of indices used in the encryption procedure, i.e., $(\boldsymbol{\rho}_i, \upsilon_i) = (\sum_{j\in S_i} \mathbf{a}_j, \sum_{j\in S_i} b_j)$. Let $(\boldsymbol{\rho}, \upsilon) = (\sum_{i=1}^{\kappa} \boldsymbol{\rho}_i, \sum_{i=1}^{\kappa} \upsilon_i)$. Recall that we obtain $\sum_{i=1}^{\kappa}\sum_{j\in S_i} e_j = \upsilon - \langle \boldsymbol{\rho}, \mathbf{s}\rangle$ in the key generation. We will show

$$\Pr\left[\left|\sum_{i=1}^{\kappa}\sum_{j\in S_i} e_i \bmod q\right|_q > \frac{\lfloor q/p \rfloor}{4}\right] < 2^{-\Omega(1/(m\beta^2 n^{2r}))}, \tag{3.1}$$

where $e_1, \ldots, e_\kappa$ are samples from the distribution $\bar{\Psi}_\beta$. A sample from $\bar{\Psi}_\beta$ can be obtained by sampling x_i from Ψ_β and outputting $\lfloor qx_i \rceil \bmod q$. Notice that $\sum_{i=1}^{\kappa}\sum_{j\in S_i} \lfloor qx_j \rceil \bmod q$ is at most $m\kappa < q/(16p)$ away from $\sum_{i=1}^{\kappa}\sum_{j\in S_i} qx_i \bmod q$ for sufficiently large n. Therefore, it is sufficient to show

$$\Pr\left[\left|\sum_{i=1}^{\kappa}\sum_{j\in S_i} qx_i\right|_q > \frac{q}{16p}\right] < 2^{-\Omega(1/(m\beta^2 n^{2r}))},$$

where $x_1, \ldots, x_\kappa$ are independently distributed according to Ψ_β. That is, it is sufficient to show

$$\Pr\left[\mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} x_i\right) > \frac{1}{16p}\right] < 2^{-\Omega(1/(m\beta^2 n^{2r}))}.$$

Similarly to the argument in Theorem 3.3.11, we obtain

$$\Pr\left[\mathrm{frc}\left(\sum_{i=1}^{\kappa}\sum_{j\in S_i} x_i\right) > \frac{1}{16p}\right] \le \Pr\left[\mathrm{frc}\left(\sum_{j=1}^{m} \kappa x_i\right) > \frac{1}{16p}\right] \le 2^{-\Omega(1/m\kappa p^2\beta^2)} \le 2^{-\Omega(1/m\beta^2 n^{2r})}.$$

It follows that we can decrypt $(\boldsymbol{\rho}, \upsilon)$ into 0 with decryption error probability at most $2^{-\Omega(1/(m\beta^2 n^{2r}))}$.

Next, we consider κ ciphertexts$(\boldsymbol{\rho}'_1, \upsilon'_1), \ldots, (\boldsymbol{\rho}'_\kappa, \upsilon'_\kappa)$ of plaintexts $\sigma_1, \ldots, \sigma_\kappa$ respectively. We now set $(\boldsymbol{\rho}', \upsilon') := (\sum_{i=1}^{\kappa} \boldsymbol{\rho}'_i, \sum_{i=1}^{\kappa} \upsilon'_i)$. By the encryption procedure, $\upsilon'_i = \upsilon_i + \sigma_i \lfloor q/p \rceil$. Therefore, we have $\upsilon' - \langle \boldsymbol{\rho}', \mathbf{s}\rangle = \sum_{i=1}^{\kappa}\sum_{j\in S_i} e_j + \sum_{i=1}^{\kappa} \sigma_i \lfloor q/p \rceil$. Combining the equation (3.1) and the fact that $\left|\sum_{i=1}^{\kappa} \sigma_i \lfloor q/p \rceil - \sum_{i=1}^{\kappa} \sigma_i q/p\right| \le \kappa p < \lfloor q/p \rfloor /4$, we decrypt $(\boldsymbol{\rho}', \upsilon')$ into $\sum_{i=1}^{\kappa} \sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega(1/(m\beta^2 n^{2r}))}$. □

Security for Sum of Ciphertexts. By a similar argument in Section 3.2.4, we obtain the following theorem.

Theorem 3.4.11. *If there exist two sequences of plaintext* $(\sigma_1, \ldots, \sigma_\kappa)$ *and* $(\sigma'_1, \ldots, \sigma'_\kappa)$ *and a polynomial-time algorithm* $\mathcal{D}_1$ *that distinguishes between* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma_i), pk)$ *and* $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}(\sigma'_i), pk)$*, then there exists a polynomial-time quantum algorithm for the worst case of* $\mathrm{SVP}_{\tilde{O}(n/\alpha(n))}$ *and* $\mathrm{SIVP}_{\tilde{O}(n/\alpha(n))}$ *in the case of* mbR05.

3.5 A Multi-Bit Version of the Ajtai Cryptosystem

3.5.1 The Ajtai Cryptosystem and Its Multi-Bit Version

Let b be a uniformly random string of $O(n^2 \log n)$ bits and t be a random string of $O(n \log n)$ bits specified later. We denote by $\nu_s^{(n)}$ a Gaussian distribution on an n-dimensional Euclidean space with mean $\mathbf{0}$ and standard deviation s. The density function is given by $\nu_s^{(n)}(\mathbf{x}) = s^{-n} \exp(-\pi \|\mathbf{x}/s\|^2)$.

Note that, given an orthonormal basis for $\mathbb{R}^n$, $\nu_s^{(n)}$ can be written as the sum of n orthogonal 1-dimensional Gaussian distributions along one of the basis vectors. For instance, given a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, $\nu_s^{(n)}(\mathbf{x}) = \prod_{i=1}^{n} (1/s) \exp(-\pi (x_i/s)^2)$ for any $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$.

Ajtai showed how to generate a certain class of efficiently representable lattices related to hard problems in [Ajt05]. He also succeeded to construct two lattice-based cryptosystems based on the original Ajtai-Dwork cryptosystem [AD97] and the improved Ajtai-Dwork cryptosystem [GGH97a]. The latter one reduces decryption error from the former one by the idea of [GGH97a]. In this section, we only describe the former one, which is related to security of our cryptosystem.

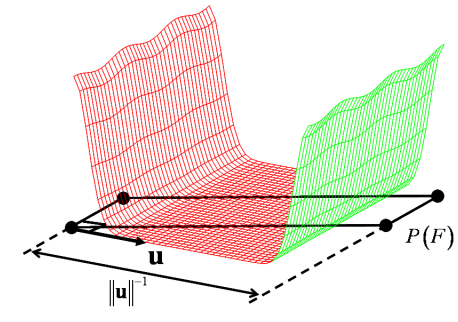

Figure 3.8: ciphertexts of 0 in A05.

In the Ajtai cryptosystem A05, we make use of a periodic Gaussian distribution on $\mathbb{R}^n$ such that its peaks are located on the points of the dual lattice spanned by a basis F of an instance $L(b,t)$ of uSVP obtained in the preparation procedure. Then, the periodic Gaussian distribution looks like a "wave" going along the shortest vector $\mathbf{u}$ of $L(b,t)$ since the dual lattice of $L(b,t)$, which is an instance of uSVP, has a much longer interval between two $(n-1)$-dimensional sublattices orthogonal to $\mathbf{u}$ than others. (See Figure 3.8.) Then, the ciphertexts of 0 correspond to the periodic Gaussian distribution modulo $\mathcal{P}(F)$ and those of 1 correspond to the uniform distribution on $\mathcal{P}(F)$ in the cryptosystem A05. Similarly to the previous cryptosystems, if we know $\mathbf{u}$, we can easily decrypt a received ciphertext by the inner product between the ciphertext and $\mathbf{u}$ with high probability.

We now describe the details of the Ajtai cryptosystem A05. All the participants share a probabilistic polynomial-time algorithm $\mathcal{D}$, a deterministic polynomial-time algorithm $\mathcal{B}$, and a uniformly random string b. In the preparation procedure, $\mathcal{D}$ generates a random string t and a vector $\mathbf{u}$ in a lattice $L(b,t)$ from b. Also, $\mathcal{B}$ generates a basis $B(b,t)$ of the lattice $L(b,t)$ if strings b and t are given. Then, the probability that $L(b,t)$ is an instance of $n^{1/2+r}$-uSVP and $\mathbf{u}$ is its unique shortest vector such that $n^{-r/2} \leq \|\mathbf{u}\| \leq n^{-r/3}$ is exponentially close to 1. Now let $F = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ be a basis of the dual lattice of $L(b,t)$. We also denote by $U_{\mathcal{P}(F)}$ the uniform distribution on $\mathcal{P}(F)$.

Cryptosystem 3.5.1 (A05, [Ajt05])**.** All the participants agree with the security parameter n, and share the algorithms $\mathcal{B}, \mathcal{D}$ and the random string b.

Key Generation: We give b to the procedure $\mathcal{D}$, and then obtain t and $\mathbf{u}$. Then, the private key is $\mathbf{u}$ and the public key is t.

Encryption: Let $\sigma \in \{0,1\}$ be an encrypted plaintext. If $\sigma = 0$, we choose $\mathbf{z}$ from a Gaussian distribution on the n-dimensional Euclidean space given by the density function $\nu^{(n)}(\mathbf{x}) = \exp(-\pi \|\mathbf{x}\|^2)$. We then set $\mathbf{y} = {}^t(y_1, \ldots, y_n) = \mathbf{z} \bmod \mathcal{P}(F)$. If $\sigma = 1$, we choose $\mathbf{y}$ from the uniform distribution $U_{\mathcal{P}(F)}$. These operations for real numbers are done with precision $2^{-n \log n}$. The ciphertext $\overline{\mathbf{y}} = {}^t(\overline{y}_1, \ldots, \overline{y}_n)$ is obtained by rounding $\mathbf{y}$ with precision of $1/n$, i.e., we have $|\overline{y}_i - y_i| \leq 1/n$ for every $i \in \{1, \ldots, n\}$.

Decryption: We decrypt a received ciphertext $\overline{\mathbf{y}}$ to 0 if $\mathrm{frc}\,(\langle \overline{\mathbf{y}}, \mathbf{u} \rangle) \leq \tilde{c} \sqrt{\log n}\, \|\mathbf{u}\|$ and to 1 otherwise, where $\tilde{c}$ is a constant given in [Ajt05]. This operation is also done with precision $2^{-n \log n}$.

Summarizing the results on A05, he mentioned the following theorem in [Ajt05]. Since the ciphertexts of A05 are rounded with precision of $1/n$ and use a compact representation of lattices, the ciphertexts and the keys can be represented by $O(n \log n)$ bits. For the definition of the underlying problem DA′, see Section 2.2.

Theorem 3.5.2 ([Ajt05])**.** *The cryptosystem* A05 *encrypts a* 1*-bit plaintext into an* $O(n \log n)$*-bit ciphertext with decryption error probability at most* $\tilde{O}(n^{-r/3})$*. The security of* A05 *is based on the average case of* DA′*. The size of the public key and the private key is* $O(n \log n)$*.*

We show the multi-bit cryptosystem mbA05 as follows. Let λ be the length of the unique non-zero shortest vector $\mathbf{u}$, i.e., $\lambda = \|\mathbf{u}\|$. We generalized the standard deviation of n-dimensional Gaussian distribution in encryption procedure for the sake of a discuss of a pseudohomomorphism. We use $\nu_s^{(n)}(\mathbf{x}) = s^{-n} \exp(-\pi \|\mathbf{x}/s\|^2)$ instead of $\nu^{(n)}$ in the original cryptosystem. If we set $s = 1$, the security of our cryptosystem is based on the security of the original one. We suppose that $\eta(n) = \omega(\sqrt{\log n})$ is a parameter to control a trade-off between decryption errors and size of plaintexts and $1/n$ is the precision of rounding in the encryption procedure as same as in the original. To guarantee the decryption errors, we suppose that $s > \sqrt{\lambda}/\eta(n)$. Let a prime p be the size of plaintext space such that $p < n^{r/6}/(4s\eta(n))$. Note that $p \leq 1/(4\sqrt{\lambda}s\eta(n))$.

Cryptosystem 3.5.3 (mbA05)**.** All the participants agree with the parameters n and s, and the size p of the plaintext space. They also share the algorithms $\mathcal{B}, \mathcal{D}$ and the random string b.

Key Generation: This procedure is the same as that of A05 except that we add an index i_1 chosen uniformly at random from $\{i : \langle \mathbf{f}_i, \mathbf{u} \rangle \not\equiv 0 \bmod p\}$ to the public key and $k \equiv \langle \mathbf{f}_{i_1}, \mathbf{u} \rangle \bmod p$ to the private key. Thus, the private key is $(\mathbf{u}, k)$ and the public key is (t, i_1).

Encryption: Let $\sigma \in \{0, \ldots, p-1\}$ be a plaintext. We choose $\mathbf{z}$ from the Gaussian distribution $\nu_s^{(n)}$. Then, the ciphertext $\overline{\mathbf{y}}$ is obtained by rounding $\mathbf{y} = \frac{\sigma}{p}\mathbf{f}_{i_1} + \mathbf{z} \bmod \mathcal{P}(F)$ with the precision of $1/n$, i.e., we have $|\overline{y}_i - y_i| \leq 1/n$ for every $i \in \{1, \ldots, n\}$.

Decryption: We decrypt a received ciphertext $\overline{\mathbf{y}}$ into $\lceil p\langle \overline{\mathbf{y}}, \mathbf{u} \rangle \rfloor k^{-1} \bmod p$, where k^{-1} is the inverse of k in $\mathbb{Z}_p$.

Before evaluating the performance of mbA05 precisely, we give the summary of the results as follows.

Theorem 3.5.4. *The cryptosystem* mbA05 *encrypts a* $\lfloor \log p(n) \rfloor$*-bit plaintext into an* $O(n \log n)$*-bit ciphertext with decryption error probability at most* $2^{-\Omega(\eta^2(n))}$*, where* $p < n^{r/6}/(4s\eta(n))$ *and* $s > \sqrt{\lambda}/\eta(n)$*. The security of* mbA05 *is based on the security of* A05*. The size of the public key is the same as that of the original one. The size of the private key is* $\lceil \log p \rceil$ *plus that of the original one.*

Setting $\eta(n) = \log n$, we obtain an $O(\log n)$-bit cryptosystem with negligible decryption errors.

Finally, we discuss the pseudohomomorphic property of mbA05. We consider a modified version mbA05′ of our multi-bit mbA05 is the same cryptosystem as mbA05 except that the precision is $2^{-n\log n}$ for its ciphertexts instead of $1/n$. This modified version mbA05′ actually has the pseudohomomorphism. We denote by E^s_{m} the encryption function of mbA05′ such that we use the Gaussian distribution with standard deviation s in the encryption procedure.

Theorem 3.5.5 (pseudohomomorphism)**.** *Let p be a prime and κ be an integer such that $\kappa p < n^{r/6}/(4\eta(n))$ for any constant $r > 0$. We can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E^1_{\mathrm{m}}(\sigma_i) \bmod \mathcal{P}(F)$ into $\sum_{i=1}^{\kappa}\sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega(\eta^2(n))}$. Moreover, if there exist two sequences of plaintexts $(\sigma_1,\ldots,\sigma_\kappa)$ and $(\sigma'_1,\ldots,\sigma'_\kappa)$, and a polynomial-time algorithm that distinguishes between $\sum_{i=1}^{\kappa} E^1_{\mathrm{m}}(\sigma_i) \bmod \mathcal{P}(F)$ and $\sum_{i=1}^{\kappa} E^1_{\mathrm{m}}(\sigma'_i) \bmod \mathcal{P}(F)$ with its public key, then there exists a polynomial-time algorithm that solves* DA′ *with non-negligible probability.*

In what follows, we demonstrate the performance of mbA05 and mbA05′ stated in the above theorems.

3.5.2 Decryption Errors of mbA05

We now give the decryption errors of our multi-bit version mbA05.

Theorem 3.5.6. *The probability of the decryption errors in* mbA05 *is at most* $2^{-\Omega(\eta^2(n))}$.

Proof. Let $\overline{\mathbf{y}}$ be a ciphertext of a plaintext σ. It is enough to show

$$\Pr\left[\mathrm{frc}\left(\langle\overline{\mathbf{y}},\mathbf{u}\rangle - \frac{k\sigma}{p}\right) > \frac{1}{2p}\right] \le 2^{-\Omega(\eta^2(n))}.$$

Since $p < 1/(4\sqrt{\lambda}s\eta(n))$ and $\sqrt{\lambda}s\eta(n) > \lambda$,

$$\begin{aligned}\Pr\left[\mathrm{frc}\left(\langle\overline{\mathbf{y}},\mathbf{u}\rangle - \frac{k\sigma}{p}\right) > \frac{1}{2p}\right] &\le \Pr\left[\mathrm{frc}\left(\langle\overline{\mathbf{y}},\mathbf{u}\rangle - \frac{k\sigma}{p}\right) > 2\sqrt{\lambda}s\eta(n)\right]\\ &\le \Pr\left[\mathrm{frc}\left(\langle\overline{\mathbf{y}},\mathbf{u}\rangle - \frac{k\sigma}{p}\right) > \sqrt{\lambda}s\eta(n) + \lambda\right].\end{aligned}$$

By the rounding precision of $1/n$, we also have $|\langle(\overline{\mathbf{y}} - \mathbf{y}),\mathbf{u}\rangle| \le \lambda$. Therefore, we have

$$\begin{aligned}\Pr\left[\mathrm{frc}\left(\langle\overline{\mathbf{y}},\mathbf{u}\rangle - \frac{k\sigma}{p}\right) > \sqrt{\lambda}s\eta(n) + \lambda\right] &\le \Pr\left[\mathrm{frc}\left(\langle\mathbf{y},\mathbf{u}\rangle - \frac{k\sigma}{p}\right) > \sqrt{\lambda}s\eta(n)\right]\\ &\le \Pr_{\mathbf{z}\sim\nu_s^{(n)}}\left[\mathrm{frc}\left(\langle\mathbf{z},\mathbf{u}\rangle\right) > \sqrt{\lambda}s\eta(n)\right] + 2^{-\Omega(n)}.\end{aligned}$$

(In the last inequality, we use the fact that $\mathbf{y} = \mathbf{z} + \frac{\sigma}{p}\mathbf{f}_{i'_0} \bmod \mathcal{P}(F)$ and $k \equiv \langle\mathbf{f}_{i'_0},\mathbf{u}\rangle \bmod p$.) Notice that the fractional part of $\langle\mathbf{z},\mathbf{u}\rangle$ then has a folded Gaussian distribution $\Psi_{\sqrt{\lambda}s}$. (Recall that its

density function Ψ_σ is of the form $\Psi_\sigma(l) = \sum_{k\in\mathbb{Z}}(1/\sigma)\exp\left(-\pi((l-k)/\sigma)^2\right).$) By Lemma 2.3.1, we have

$$\Pr_{\mathbf{z}\sim\nu_s^{(n)}}\left[\mathrm{frc}\left(\langle\mathbf{z},\mathbf{u}\rangle\right) > \sqrt{\lambda}s\eta(n)\right] \leq \frac{1}{\pi\eta(n)}\exp\left(-\pi\eta^2(n)\right).$$

This completes the proof. □

3.5.3 Security of mbA05

The security of our cryptosystem mbA05 can be also proven by a similar technique to mbADGGH.

Theorem 3.5.7. *If there exist plaintexts $\sigma_1, \sigma_2 \in \{0,\ldots,p-1\}$ and a polynomial-time algorithm that distinguishes between the ciphertext of σ_1 and σ_2 in* mbA05 *with its public key, there exists a polynomial-time algorithm that distinguishes between the ciphertexts of* 0 *and* 1 *in* A05 *with its public key.*

3.5.4 Pseudohomomorphism of mbA05′

Decryption Errors for Sum of Ciphertexts. Recall that we adopt the precision of $2^{-n\log n}$ for the ciphertexts in mbA05′. We denote by E_{m}^s the encryption function of mbA05′ such that we use the Gaussian distribution with standard deviation s in the encryption procedure.

Theorem 3.5.8 (mbA05′)**.** *Let $\eta(n) = \omega(\sqrt{\log n})$. Also let p be a prime and κ be an integer such that $\kappa p < n^{r/6}/(4\eta(n))$ for any constant $r > 0$. We can decrypt the sum of κ ciphertexts $\sum_{i=1}^{\kappa} E_{\mathrm{m}}^1(\sigma_i) \bmod \mathcal{P}(F)$ into $\sum_{i=1}^{\kappa}\sigma_i \bmod p$ with decryption error probability at most $2^{-\Omega(\eta^2(n))}$.*

Proof. Since the precision is $2^{-n\log n}$, we can consider $\sum_{i=1}^{\kappa} E_{\mathrm{m}}^1(\sigma_i) \bmod \mathcal{P}(F)$ as $E_{\mathrm{m}}^{\sqrt{\kappa}}(\sum_{i=1}^{\kappa}\sigma_i \bmod p)$. Replacing s and p by $\sqrt{\kappa}$ and κp respectively, we can evaluate the decryption errors with the same argument as the proof of Theorem 3.5.6 by the fact that $|\langle\overline{\mathbf{y}} - \mathbf{y}, \mathbf{u}\rangle| \leq n\lambda 2^{-n\log n} = 2^{-\Omega(n)}$. □

Security for Sum of Ciphertexts. Combining Lemma 3.2.13 with the security proof of A05 in [Ajt05], we guarantee the security of the sum of ciphertexts in mbA05′. Note that we can regard $\sum_{i=1}^{\kappa} E_{\mathrm{m}}^1(\sigma_i) \bmod \mathcal{P}(W)$ as $E_{\mathrm{m}}^{\sqrt{\kappa}}(\sum_{i=1}^{\kappa}\sigma_i \bmod p)$ in mbA05′ by replacing the precision $1/n$ of the ciphertexts to $2^{-n\log n}$.

Theorem 3.5.9. *If there exist two sequences of plaintexts $(\sigma_1,\ldots,\sigma_\kappa)$ and $(\sigma'_1,\ldots,\sigma'_\kappa)$ and a polynomial-time algorithm $\mathcal{D}_1$ that distinguishes between $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}^1(\sigma_i), pk)$ and $(\sum_{i=1}^{\kappa} E_{\mathrm{m}}^1(\sigma'_i), pk)$, then there exists a probabilistic polynomial-time algorithm $\mathcal{A}$ that solves* DA′.

3.6 Concluding Remarks

We have developed a universal technique for constructing multi-bit versions of lattice-based cryptosystems using periodic Gaussian distributions and revealed their pseudohomomorphism. In particular, we have showed the details of the multi-bit version of the improved Ajtai-Dwork cryptosystem in Section 3.2.

Although our technique achieved only logarithmic improvements on the length of plaintexts, we also obtained precise evaluation of the trade-offs between decryption errors and the hardness of underlying lattice problems in the single-bit cryptosystems. We believe that our evaluation is useful for further improvements of such single-bit cryptosystems.

Another direction of research on lattice-based cryptosystems is to find interesting cryptographic applications by their algebraic properties such as the pseudohomomorphism. Number-theoretic cryptosystems can provide a number of applications due to their algebraic structures, whereas lattice-based ones have few applications currently. For demonstration of the cryptographic advantages of lattice problems, it is important to develop the algebraic properties and their applications such as [GK05].

Chapter 4

A Modified Regev'05 Cryptosystem, Proofs of Knowledge on Its Secret Key, and Signature Schemes

4.1 Introduction

Summary. We propose a modified Regev'05 cryptosystem and introduce a proof of knowledge on its secret key in the common reference string (CRS) model. We consider the relation between the private key and the public key as that between the message and the codeword with the error in coding theory. To construct a proof of knowledge, we modify generation of the error. This modification admits a prover to prove the knowledge of the error and the message based on Stern [Ste96]. Thus, we obtain a proof of knowledge on a secret key of our cryptosystem. We also obtain a signature scheme via the Fiat-Shamir transformation [FS86, AABN02].

Related Results. There already exist public-key identification schemes based on lattice and coding problems. In 1989, Shamir showed an identification scheme based on permuted kernel problem [Sha89]. Stern proposed public-key identification based on syndrome decoding problem in 1996 [Ste96]. Micciancio and Vadhan introduced a zero-knowledge proof with efficient prover for GapCVP$_\gamma$ and discussed public-key identification schemes [MV03]. Recently, Hayashi and Tada showed public-key identification schemes based on binary non-negative exact length vector problem (or integer subset sum problem) [HT06]. Unfortunately, it is unknown whether their public keys can be used as a public key of cryptosystems or not. We stress that in our identification schemes, the information for identification is indeed a public key of cryptosystems.

Why can we not apply the MV protocol to R05? Before description of our idea, we briefly review the key generation of R05 and explain why the same approach with the Micciancio-Vadhan protocol [MV03] fails for our goal. (We abbreviate it to "the MV protocol".)

In R05, the secret key is $\mathbf{s} \in \mathbb{Z}_q^n$ and the public key is $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m] \in \mathbb{Z}_q^{n \times m}$ and $\mathbf{b} = {}^t\mathbf{A}\mathbf{s}+\mathbf{e}$, where $\mathbf{e} \in \mathbb{Z}_q^m$ and each coordinate of $\mathbf{e}$ is close to 0. From a coding-theoretical view, we can regard ${}^t\mathbf{A}$ as a generator matrix, $\mathbf{s}$ as a message, and $\mathbf{e}$ as an error. Remark that the length of $\mathbf{e}$ is short. Hence, one would think we can apply the MV protocol to proofs of knowledge for a secret key $\mathbf{s}$. However, we cannot apply it in a naive way. We explain more details.

We first review the intuition which is used in the MV protocol. (See Protocol 5.2.1 for more details.) Let $(\mathbf{B}, \mathbf{y}, t)$ be an instance of GapCVP$_\gamma$.[1] Let $B_m(\mathbf{c}, r)$ be an m-dimensional hyperball whose center is $\mathbf{c}$ and radius is r. In their protocol, the prover chooses a random bit c and a random vector $\mathbf{r}$ from $B_m(\mathbf{0}, \gamma t/2)$. The prover computes $\mathbf{m} = c\mathbf{y} + \mathbf{r} \bmod \mathbf{B}$ and sends $\mathbf{m}$ to the verifier. The verifier sends a challenge bit δ to the prover. Note that if $(\mathbf{B}, \mathbf{y}, t)$ is a YES instance then the ratio between the volume of $(B_m(\mathbf{0}, \gamma t/2) \bmod \mathbf{B}) \cap (B_m(\mathbf{y}, \gamma t/2) \bmod \mathbf{B})$ and that of $B(\mathbf{0}, \gamma t/2)$ is at least $1/\operatorname{poly}(n)$. If $\mathbf{m} \in (B_m(\mathbf{0}, \gamma t/2) \bmod \mathbf{B}) \cap (B_m(\mathbf{y}, \gamma t/2) \bmod \mathbf{B})$ the prover can flip a bit c. The prover sends the proof that $\mathbf{m}$ is chosen from $B_m(c\mathbf{y}, \gamma t/2)$. Note that if $(\mathbf{B}, \mathbf{y}, t)$ is a NO instance then $(B_m(\mathbf{0}, \gamma t/2) \bmod \mathbf{B}) \cap (B_m(\mathbf{y}, \gamma t/2) \bmod \mathbf{B}) = \emptyset$. Therefore the prover can not flip a bit c after a reception of the challenge bit.

Next, we consider applying their protocol to the Regev'05 cryptosystem, i.e., a proof of knowledge that, on input $(\mathbf{A}, \mathbf{b})$, the prover knows $\mathbf{s}$ such that $\mathbf{b} = {}^t\mathbf{A}\mathbf{s} + \mathbf{e}$, where $\mathbf{e} \in B_m(\mathbf{0}, t)$.[2] Note that a linear code is $\mathbb{Z}_q$-module in $\mathbb{Z}_q^m$ and a lattice is $\mathbb{Z}$-module in $\mathbb{R}^m$. Therefore, instead of reducing modulo $\mathbf{B}$, we multiply a parity-check matrix $\mathbf{H}$ of ${}^t\mathbf{A}$ to the vector in $\mathbb{Z}_q^m$. Suppose that $B_m(\mathbf{0}, \gamma t/2)$ and $B_m(\mathbf{b}, \gamma t/2)$ do not intersect. Unfortunately, we cannot ensure that $\mathbf{H}B_m(\mathbf{0}, \gamma t/2)$ and $\mathbf{H}B_m(\mathbf{b}, \gamma t/2)$ do not intersect because the dimension of $\mathbf{H}\mathbb{Z}_q^m$ is $m - n < m$. On such NO instance $(\mathbf{A}, \mathbf{b})$, the prover can cheat the verifier on which hyperball he chose $\mathbf{m}$ from. Hence the soundness of the protocol fails. Thus, we cannot apply their protocol to the Regev'05 cryptosystem in a straightforward way.

Main Idea. As seen in the above paragraphs, we cannot apply the protocol [MV03] to the Regev'05 cryptosystem straightforwardly. Let us reconsider multiplying a parity-check matrix $\mathbf{H}$. Let $\mathbf{s} \in \mathbb{Z}_q^n$ be a private key and let $(\mathbf{A}, \mathbf{b})$ be a public key, where $\mathbf{b} = {}^t\mathbf{A}\mathbf{s} + \mathbf{e}$. Multiplying a

[1] $(\mathbf{B}, \mathbf{y}, t)$ is a YES instance if there exists $\mathbf{w} \in \mathbb{Z}^n$ such that $\|\mathbf{B}\mathbf{w} - \mathbf{y}\| \leq t$. It is a NO instance if for any vector $\mathbf{w} \in \mathbb{Z}^n$, $\|\mathbf{B}\mathbf{w} - \mathbf{y}\| \geq \gamma t$. Although they consider only full-rank lattices in [MV03], we consider not only full-rank lattices. That is, an instance of GapCVP$_\gamma$ consists of $\mathbf{B}$, which is a basis of a lattice whose rank is n, $\mathbf{y} \in \mathbb{R}^m$, $\gamma \geq 1$.

[2] We abuse the notation $B_m(\cdot, \cdot)$.

parity-check matrix $\mathbf{H}$ to the equation $\mathbf{b} = {}^t\mathbf{A}\mathbf{s} + \mathbf{e}$, we obtain that $\mathbf{Hb} = \mathbf{He}$. The prover should prove the knowledge of $\mathbf{e}$ that satisfies the equation and each coordinate of $\mathbf{e}$ is in certain range. The difficulty to construct the protocol is to combine protocols that prove sufficiency of the equation and lying in the range.

Then, we modify a public key as follows: The secret key is $\mathbf{s} \in \mathbb{Z}_q^n$ and $\mathbf{s}' \in \{0,1\}^{m_1}$, whose Hamming weight is m_2. The public key is $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ and $\mathbf{E} \in \mathbb{Z}_q^{m \times m_1}$ and $\mathbf{b} = {}^t\mathbf{A}\mathbf{s} + \mathbf{E}\mathbf{s}'$. In this case, by multiplying a parity-check matrix $\mathbf{H}$, we have that $\mathbf{Hb} = \mathbf{HEs}'$. Translating a matrix $\mathbf{HE}$ as a parity-check matrix, we have an instance $(\mathbf{HE}, \mathbf{Hb}, m_2)$ and a witness $\mathbf{s}'$ of Syndrome Decoding Problem (SDP).[3] Since Stern proposed a proof of knowledge for SDP in 1996 [Ste96], we adopt it to prove knowledge of secret key $\mathbf{s}'$.

The proof of knowledge for SDP needs a statistically-hiding and computationally-binding commitment scheme. Fortunately, if $\mathbf{A}$ is chosen randomly then the function $f_{\mathbf{A}} : \{0,1\}^m \to \mathbb{Z}_q^n : \mathbf{m} \mapsto \mathbf{Am}$ is a collision-resistant function based on the approximation version of SVP [Ajt96b, GGH96, CN97, Mic04a, MR04]. Thus we employ that function to develop a statistically-hiding and computationally-binding string commitment scheme. Our construction of a string commitment is more straightforward than Damgård, Pedersen, and Pfizmann [DPP97, DPP98] and Halevi and Micali [HM96], which used the universal hash functions.

We also show the security of the modified R05, mR05. Unfortunately, we need a stronger assumption than the original one. The stronger assumption is the worst-case hardness of certain learning problem, which is based on well-known problem Learning With Error (LWE).

Organization. The rest of this chapter is organized as follows. We briefly note basic notions and notations in Section 4.2. We describe the Regev'05 cryptosystem and our modified cryptosystem in Section 4.3. Finally, we give our main results, a proof of knowledge on a secret key, in Section 4.4.

4.2 Preliminaries

For integers $m_1 \geq m_2 \geq 0$, we define $\mathrm{Set}_{m_1,m_2} := \{\mathbf{s}' \in \{0,1\}^{m_1} \mid w_H(\mathbf{s}') = m_2\}$. For any $\mathbf{s} \in \mathbb{Z}_q^m$, we define $A_{\mathbf{s}}$ obtained as follows: (1) Choose a random vector $\mathbf{a} \in \mathbb{Z}_q^m$. (2) Choose a random element $e \in \mathbb{Z}_q$ according to $\bar{\Psi}_\alpha$. (3) Outputs $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)$. For any $\mathbf{s} \in \mathbb{Z}_q^m$ and any $\mathbf{s}' \in \mathrm{Set}_{m_1,m_2}$, we define $A_{\mathbf{s},\mathbf{s}'}$ as the distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q^{m_1} \times \mathbb{Z}_q$ obtained as follows: (1) Choose a random vector $\mathbf{a} \in \mathbb{Z}_q^m$. (2) Choose a random vector $\mathbf{e} \in \mathbb{Z}_q^{m_1}$ according to $\bar{\Psi}_{\alpha/m_2}^{(m_1)}$. (3) Set

[3] Syndrome Decoding Problem: Given input $(\mathbf{H}, \mathbf{y}, m)$, where $\mathbf{H} \in \mathbb{Z}_2^{(n-k)\times n}$, $\mathbf{y} \in \mathbb{Z}_2^{n-k}$, $m \geq 0$, find $\mathbf{x} \in \mathbb{Z}_2^n$ such that $\mathbf{Hx} = \mathbf{y}$ and Hamming weight of $\mathbf{x}$ is exactly m.

$b := \langle \mathbf{a}, \mathbf{s} \rangle + \langle \mathbf{e}, \mathbf{s}' \rangle$ and output $(\mathbf{a}, \mathbf{e}, b)$. We also define U' as the distribution on $\mathbb{Z}_q^n \times \mathbb{Z}_q^{m_1} \times \mathbb{Z}_q$ obtained as follows: (1) Choose a random vector $\mathbf{a} \in \mathbb{Z}_q^m$. (2) Choose a random vector $\mathbf{e} \in \mathbb{Z}_q^{m_1}$ according to $\bar{\Psi}_{\alpha/m_2}^{(m_1)}$. (3) Choose a random elements $u \in \mathbb{Z}_q$ and output $(\mathbf{a}, \mathbf{e}, u)$.

We consider the following learning problems.

Definition 4.2.1 (Learning With Errors, $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$)**.** Given samples from $A_{\mathbf{s}}$, find $\mathbf{s}$.

Definition 4.2.2 (Learning With Known Errors, $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$)**.** Given samples from $A_{\mathbf{s},\mathbf{s}'}$, find $\mathbf{s}$.

We note that if there exists an adversary $\mathcal{A}$ that solves $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$ with non-negligible probability then there exists an adversary $\mathcal{A}'$ that solves $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ with non-negligible probability. If $\mathcal{A}$ needs $k = \mathrm{poly}(n)$ samples, then $\mathcal{A}'$ takes k samples $(\mathbf{a}_i, \mathbf{e}_i, b_i)$ from $A_{\mathbf{s},\mathbf{s}'}$. $\mathcal{A}'$ inputs $\{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,k}$ to $\mathcal{A}$ and obtains an output $\mathbf{s}$. $\mathcal{A}'$ outputs $\mathbf{s}$. Using the reproducibility of Gaussian distributions, we show that the sum of m_2 samples according to $\bar{\Psi}_{\alpha/m_2}$ is, in fact, distributed according to $\bar{\Psi}_\alpha$, and hence $\{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,k}$ which $\mathcal{A}'$ computes is indeed samples from $A_{\mathbf{s}}$.

4.2.1 String Commitment

We explain the notation for commitment schemes in the common reference string (CRS) model. Assume that there exists a trusted third party (TTP). Let $\mathrm{Com}_{(\cdot)}(\cdot;\cdot)$ be an indexed function which maps a pair of a message string and a random string to a commitment string. First, TTP on input 1^n outputs a random string a, which is the CRS and the index of the commitment function. To commit to a string s, the sender chooses a random string r, computes $c = \mathrm{Com}_a(s; r)$, and sends c to the receiver. To reveal commitment c, the sender sends s and r to the receiver. The receiver accepts if $c = \mathrm{Com}_a(s; r)$ or rejects otherwise.

Definition 4.2.3. We say a string commitment scheme $\mathrm{Com}_{(\cdot)}(\cdot;\cdot)$ is statistically hiding and computationally binding if it has the following properties:

Statistical Hiding: For any two strings s and s', the statistical distance between $(a, \mathrm{Com}_a(s; r))$ and $(a, \mathrm{Com}_a(s'; r'))$ is negligible, where a, r, r' are random and independent.

Computational Binding: For any probabilistic polynomial-time machine $\mathcal{A}$, if a is randomly chosen by TTP, then the probability that, given an input a, $\mathcal{A}$ outputs (s, r) and (s', r') such that $\mathrm{Com}_a(s; r) = \mathrm{Com}_a(s'; r')$ is negligible.

4.2.2 Subset-Sum Hash Functions and a String Commitment Scheme

As explained in Section 4.1 we need a string commitment scheme to construct a proof of knowledge of a secret key. We first argue the family of subset-sum hash functions and the string commitment scheme.

Let n be a security parameter (or a dimension of underlying lattice problems). For a prime $q = q(n) = n^{O(1)}$ and an integer $m = m(n) > n \log q(n)$, we define a family of hash functions, $\mathcal{H}_{q,m} = \{ f_{\mathbf{A}} : \{0,1\}^{m(n)} \to \mathbb{Z}_{q(n)}^n \mid \mathbf{A} \in \mathbb{Z}_{q(n)}^{n \times m(n)} \}$, where $f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \bmod q(n)$.

Originally, Ajtai [Ajt96a] showed $\mathcal{H}_{q,m}$ is a family of one-way functions under the assumption that SVP with some polynomial approximation factor is hard in the worst case for suitably chosen $q(n)$ and $m(n)$. It is known that $\mathcal{H}_{q,m}$ is indeed a family of collision-resistant hash functions for suitably chosen q and m by Goldreich, Goldwasser, and Halevi [GGH96], Cai and Nerurkar [CN97] and Micciancio [Mic04a]. Recently, Micciancio and Regev showed $\mathcal{H}_{q,m}$ is a family of collision-resistant hash functions under the assumption $\mathrm{SVP}_{\tilde{O}(n)}$ is hard in the worst case [MR04].

We construct a statistically-hiding and computationally-binding string commitment scheme based on the above hash functions. It is well known that if there exists a collision-resistant hash function then there exists a statistically hiding and computationally binding string commitment scheme [DPP97, DPP98, HM96]. Their construction used universal hash functions for the statistically hiding property. However, our construction do not use it, because if m is sufficiently large and a plaintext $\mathbf{s}$ is randomized, $\mathbf{As}$ is distributed statistically close to the uniform distribution. To prove the statistically-hiding property, we use Claim 4.2.5 below in [Reg05].

We describe how to achieve a string commitment scheme in the CRS model. We first split the domain $\{0,1\}^m$ into two domain $\{0,1\}^{m/2} \times \{0,1\}^{m/2}$. The first domain is used for randomization. The second domain is for message. We define $\mathrm{Com}_{\mathbf{A}}(s; r) := \mathbf{A}\mathbf{x}$, where $\mathbf{x} = {}^t(r_0, \ldots, r_{m/2}, s_1, \ldots, s_{m/2})$, $r = r_1 \ldots r_{m/2}$, and $s = s_1 \ldots s_{m/2}$.

Lemma 4.2.4. *For a prime $q = q(n) = n^{O(1)}$ and an integer $m = m(n) > 10n \log q$, if $\mathcal{H}_{q,m}$ is collision resistant and a trusted third party gives a random matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, then $\mathrm{Com}_{\mathbf{A}}$ is a statistically hiding and computationally binding string commitment scheme in the CRS model.*

Proof. The computationally-binding property immediately follows from the collision-resistant property. Next, we consider the statistically-hiding property. Using Claim 4.2.5 below, we have that with probability exponentially close to 1 the statistical distance between the distribution of $(\mathbf{A}, \mathrm{Com}_{\mathbf{A}}(0^{m/2}; r))$ and that of $(\mathbf{A}, \mathbf{u})$ is negligible in n, where r and $\mathbf{u}$ are random variables according to the uniform distribution on $\{0,1\}^{m/2}$ and $\mathbb{Z}_q^n$, respectively. Hence, for any two messages $m_1, m_2 \in \{0,1\}^{m/2}$, the statistical distance between the distribution of $(\mathbf{A}, \mathrm{Com}_{\mathbf{A}}(m_1; r_1))$ and that of $(\mathbf{A}, \mathrm{Com}_{\mathbf{A}}(m_2; r_2))$ is negligible in n with probability exponentially close to 1, where r_1 and r_2 are random variables according to the uniform distribution on $\{0,1\}^{m/2}$. This completes the proof. □

Claim 4.2.5 (Claim 5.3, [Reg05])**.** *Let G be a finite Abelian group and let $l = c \log |G|$. For $c \geq 5$, when choosing l elements $g_1, \ldots, g_l$ uniformly from G the probability that the statisti-*

cal distance between the uniform distribution on G and the distribution given by the sums of random subsets of $g_1, \ldots, g_l$ is more than $2/|G|$ is at most $1/|G|$.

4.3 The Regev'05 Cryptosystem and Its Modification

4.3.1 The Regev'05 Cryptosystem

Regev proposed a lattice-based cryptosystem in 2005 [Reg05]. Although we briefly review the Regev'05 cryptosystem, R05, in Section 3.4, we review it again.

Cryptosystem 4.3.1 (R05, [Reg05])**.** Let n be a security parameter (or a dimension of the underlying lattice problem). Let q be a prime and α be a parameter to define the variance of Gaussian distribution such that $\alpha q > 2\sqrt{n}$. Let m be an integer at least $5(n+1)\log q$.

Private Key: Choose $\mathbf{s} \in \mathbb{Z}_q^n$ randomly.

Public Key: Choose $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{Z}_q^n$ randomly. Choose $e_1, \ldots, e_m$ according to the distribution $\bar{\Psi}_\alpha$. Compute $b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i \bmod q$. The public key is $\{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$.

Encryption: A plaintext is $\sigma \in \{0,1\}$. Choose $S \subseteq_R \{1, \ldots, m\}$ randomly. The ciphertext is $(\sum_{i \in S} \mathbf{a}_i, \sigma \lfloor q/2 \rfloor + \sum_{i \in S} b_i)$.

Decryption: Let $(\mathbf{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ be a received ciphertext. If $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q \leq q/4$ then decrypt to 0. Otherwise decrypt to 1.

The size of a public key and a private key are $O(mn \log q) = O(n^2 \log^2 q)$ and $O(n \log q) = O(n \log n)$ respectively. If $\mathbf{a}_1, \ldots, \mathbf{a}_m$ is the CRS, this is the idea from Ajtai [Ajt05], the size of a public key is $O(m \log q) = O(n \log^2 q)$. We summarize the security and decryption errors of R05.

Theorem 4.3.2 (Thereom 3.1, Lemma 4.4, and Lemma 5.4, [Reg05])**.** *Let $\alpha = \alpha(n)$ be a real number on $(0,1)$ and $q = q(n)$ a prime such that $\alpha q > 2\sqrt{n}$. For $m \geq 5(n+1)\log q$, if there exists a polynomial time algorithm that distinguishes between encryptions of 0 and 1 then there exists a distinguisher that distinguishes between $A_{\mathbf{s}}$ and $U(\mathbb{Z}_q^n \times \mathbb{Z}_q)$ for a non-negligible fraction of all possible $\mathbf{s}$.*

Next, assume there exists a distinguisher that distinguishes $A_{\mathbf{s}}$ from $U(\mathbb{Z}_q^n \times \mathbb{Z}_q)$ for a non-negligible fraction of all possible $\mathbf{s}$. Then, there exists an efficient algorithm that solves $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$.

Finally, assume there exists an efficient (possibly quantum) algorithm that solves $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$. *Then there exists an efficient quantum algorithm for solving the worst-case of* $\mathrm{SVP}_{\tilde{O}(n/\alpha)}$ *and* $\mathrm{SIVP}_{\tilde{O}(n/\alpha)}$.

Lemma 4.3.3 (Lemma 5.1, [Reg05] (Correctness))**.** *The decryption error probability is at most* $2^{-\Omega(1/(m\alpha^2))} + 2^{-\Omega(n)}$.

Remark 4.3.4. The reduction in Theorem 4.3.2 is quantum. Therefore, the security of R05 depends on the worst-case hardness of $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$ in the classical sense.

4.3.2 A Modified Regev'05 Cryptosystem

We modify the Regev'05 cryptosystem to obtain a new cryptosystem mR05.

Cryptosystem 4.3.5 (mR05)**.** Let n be a security parameter (or a dimension of the underlying lattice problem). Let q be a prime and α be a parameter to define the variance of Gaussian distribution such that $\alpha q > 2\sqrt{n}$. Let t_α be a threshold such that $\Pr_{e\sim\bar{\Psi}_{\alpha/m_2}}[|e|_q \geq t_\alpha]$ is negligible in n (i.e., $t_\alpha = \omega(\alpha q \log n/m_2)$.) Let m be an integer at least $10(n+1)\log q$. Let m_1 and m_2 be integers such that $m_1, m_2 = \mathrm{poly}(n)$ and $\binom{m_1}{m_2}$ is exponential in n. Let $\mathrm{Set}_{m_1,m_2} := \{\mathbf{s}' \in \{0,1\}^{m_1} \mid w_H(\mathbf{s}') = m_2\}$. We need $4mm_2t_\alpha < q$ to ensure the correctness of the cryptosystem.

Private Key: Choose $\mathbf{s} \in \mathbb{Z}_q^n$ randomly. Choose $\mathbf{s}' \in \mathrm{Set}_{m_1,m_2}$ randomly.

Public Key: Choose $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{Z}_q^n$ randomly and $\mathbf{e}_1, \ldots, \mathbf{e}_{m_1}$ according to the distribution $\bar{\Psi}^{(m)}_{\alpha/m_2}$. Let $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m]$ and $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_{m_1}]$. Check for any i, $\mathbf{e}_i$'s coordinates are at most t_α in the sense of $|\cdot|_q$. Compute $\mathbf{e} := \mathbf{E}\mathbf{s}'$. Let $\mathbf{b} := {}^t\mathbf{A}\mathbf{s} + \mathbf{e} \in \mathbb{Z}_q^m$. The public key is $(\mathbf{A}, \mathbf{E}, \mathbf{b})$. The secret key is $\mathbf{s}, \mathbf{s}'$.

Encryption: A plaintext is $\sigma \in \{0,1\}$. Choose $S \subseteq_R \{1, \ldots, m\}$ randomly. The ciphertext is $(\sum_{i\in S} \mathbf{a}_i, \sigma\lfloor q/2\rfloor + \sum_{i\in S} b_i)$.

Decryption: Let $(\mathbf{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ be a received ciphertext. If $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq q/4$ then decrypt to 0. Otherwise decrypt to 1.

For example, we set $q = \Theta(n^3)$, $m = 10(n+1)\log q$, $\alpha = 1/m^2$, $t_\alpha = n/\log n$, $m_1 = m$, and $m_2 = \sqrt{m}$. Note that, with such parameters, we have that $4mm_2t_\alpha < q$.

The size of a public key and a private key are $O(mn\log q + m_1 n \log q) = O(n^2 \log^2 q)$ and $O(n\log q + m_1 \log q) = O(n\log^2 n)$ respectively. If $\mathbf{A}$ and $\mathbf{E}$ are the CRSs the size of a public key is $O(m\log q) = O(n\log^2 q)$. Note that, from a coding-theoretical view, ${}^t\mathbf{A}$ is a generator matrix and we can compute a parity check matrix $\mathbf{H}$ such that, for any $\mathbf{s} \in \mathbb{Z}_q^n$, $\mathbf{H}{}^t\mathbf{A}\mathbf{s} = \mathbf{0} \in \mathbb{Z}_q^{m-n}$.

First, we see the correctness of mR05.

Lemma 4.3.6 (Correctness)**.** *There exist no decryption errors in* mR05.

Proof. Suppose that $(\mathbf{a}, b)$ is a valid ciphertexts of 0, i.e., $(\mathbf{a}, \mathbf{b}) = (\sum_{i=1}^{m} r_i \mathbf{a}_i, \sum_{i=1}^{m} r_i b_i)$ for some $r \in \{0,1\}^m$. We have

$$\begin{aligned} |b - \langle \mathbf{a}, \mathbf{s} \rangle|_q &= \left| \sum_{i=1}^{m} r_i b_i - \langle \sum_{i=1}^{m} r_i \mathbf{a}_i, \mathbf{s} \rangle \right|_q \\ &= \left| \sum_{i=1}^{m} r_i e_i \right|_q \leq \left| \sum_{i=1}^{m} e_i \right|_q \leq m \, |e_i|_q \leq m m_2 t_\alpha, \end{aligned}$$

where e_i is i-th coordinate of $\mathbf{e} = \mathbf{E}\mathbf{s}'$. Since we set $4mm_2t_\alpha < q$, we obtain $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q < q/4$. Next we consider the case $(\mathbf{a}, b)$ is a valid ciphertexts of 1, i.e., $(\mathbf{a}, \mathbf{b}) = (\sum_{i=1}^{m} r_i \mathbf{a}_i, \lfloor q/2 \rfloor + \sum_{i=1}^{m} r_i b_i)$ for some $r \in \{0,1\}^m$. Similarly to the case of 0, we have

$$|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q \geq \lfloor q/2 \rfloor - m m_2 t_\alpha \geq q/4$$

and we can decrypt correctly. □

Combining Lemmas 4.3.8, 4.3.9, and 4.3.10 below, we obtain the following theorem on security of mR05.

Theorem 4.3.7 (Security)**.** *For $m \geq 10(n+1)\log q$, if there exists a polynomial-time algorithm $\mathcal{D}$ that distinguishes between encryptions of 0 and 1 with its public key, then there exists a polynomial-time algorithm $\mathcal{A}$ that solves $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ in the worst case.*

Lemma 4.3.8. *For $m \geq 5(n+1)\log q$, if there exists a polynomial time algorithm $\mathcal{D}$ that distinguishes between encryptions of 0 and 1 with its public key, then there exists a distinguisher $\mathcal{D}'$ that distinguishes between $A_{\mathbf{s},\mathbf{s}'}$ and U' for a non-negligible fraction of all possible $\mathbf{s}$ and $\mathbf{s}'$.*

We omit the proof, because the proof is quite similar to the security proof in [Reg05].

Lemma 4.3.9 (Average-case to Worst-case)**.** *Assume there exists a distinguisher $\mathcal{D}$ that distinguishes $A_{\mathbf{s},\mathbf{s}'}$ from U' for a non-negligible fraction of all possible $\mathbf{s}$ and $\mathbf{s}'$. Then there exists an algorithm $\mathcal{D}'$ that for all $\mathbf{s}$ and $\mathbf{s}'$ accepts with probability exponentially close to 1 on inputs from $A_{\mathbf{s},\mathbf{s}'}$ and rejects with probability exponentially close to 1 on inputs from U'.*

Proof. As similar to Regev's proof [Reg05], we prove the lemma based on the following transformation. For any $\mathbf{t} \in \mathbb{Z}_q^n$ and any permutation $\pi \in S_{m_1}$ consider the function $f_{\mathbf{t},\pi} : \mathbb{Z}_q^n \times \mathbb{Z}_q^{m_1} \times \mathbb{Z}_q \to \mathbb{Z}_q^n \times \mathbb{Z}_q^{m_1} \times \mathbb{Z}_q$ defined by

$$f_{\mathbf{t},\pi}(\mathbf{a}, \mathbf{e}, b) = (\mathbf{a}, \pi(\mathbf{e}), b + \langle \mathbf{a}, \mathbf{t} \rangle).$$

This function transforms the distribution $A_{\mathbf{s},\mathbf{s}'}$ into $A_{\mathbf{s}+\mathbf{t},\pi(\mathbf{s}')}$. Moreover, it transforms the distribution U' into itself.

Next we consider a random statistical test. Assume that for n^{-c_1} fraction of all possible $(\mathbf{s}, \mathbf{s}')$, the acceptance probability of W on inputs from $A_{\mathbf{s},\mathbf{s}'}$ and on inputs from U' differ by at least n^{-c_2}. We construct the distinguisher $\mathcal{D}'$ as follows. Let R denote the unknown input distribution. (0) Repeat the following n^{c_1+1} times. (1) Choose a vector $\mathbf{t} \in \mathbb{Z}_q^n$ and a permutation $\pi \in S_{m_1}$ uniformly at random. (2) Estimate p_R, the acceptance probability of $\mathcal{D}$ on $f_{\mathbf{t},\pi}(R)$, by calling $\mathcal{D}$ $T = n^{2c_2+1}$ times. Let x_R be the number of 1 in the outputs of $\mathcal{D}$. (3) Estimate p_U, the acceptance probability of $\mathcal{D}$ on U', by calling $\mathcal{D}$ T times. Let x_U be the number of 1 in the outputs of $\mathcal{D}$. (4) If $|x_U - x_R| / T \geq n^{-c_2}/2$ then stop and accept. Otherwise continue. (5) If the procedure ends without accepting, stop and reject.

When R is U', the probability that $|p_U - x_U/T| \geq n^{-c_2}/8$ is exponentially small by the Hoeffding bound. Since $f_{\mathbf{t},\pi}(U') = U'$, the probability that $|p_U - x_R/T| \geq n^{-c_2}/8$ is exponentially small. Therefore, the acceptance probability of $\mathcal{D}'$ is exponentially close to 0.

When R is $A_{\mathbf{s},\mathbf{s}'}$ for some $\mathbf{s}, \mathbf{s}'$. In each of the iterations, we are considering the distribution $f_{\mathbf{t},\pi}(A_{\mathbf{s},\mathbf{s}'}) = A_{\mathbf{s}+\mathbf{t},\pi(\mathbf{s}')}$ for some uniformly chosen $\mathbf{t}$ and π. Hence, with probability exponentially close to 1, in one of the n^{c_1+1} iterations, $(\mathbf{s} + \mathbf{t}, \pi(\mathbf{s}'))$ is such that the acceptance probability of $\mathcal{D}$ on inputs from $A_{\mathbf{s}+\mathbf{t},\pi(\mathbf{s}')}$ and on inputs from U' differ by at least n^{-c_2}. In this case, from the Hoeffding bound, the probability that $|p_U - x_U/T| \geq n^{-c_2}/8$ and $|p_R - x_R/T| \geq n^{-c_2}/8$ is exponentially small. Hence, $\mathcal{D}'$ accepts with probability exponentially close to 1. □

Lemma 4.3.10 (Decision to Search)**.** *Let $n \geq 1$ be some integer and $q \geq 2$ be a prime. Assume there exists an algorithm $\mathcal{D}$ that for all $\mathbf{s}, \mathbf{s}'$ accepts with probability exponentially close to 1 on inputs from $A_{\mathbf{s},\mathbf{s}'}$ and rejects with probability exponentially close to 1 on inputs from U'. Then, there exists an algorithm $\mathcal{D}'$ that, given samples from $A_{\mathbf{s},\mathbf{s}'}$ for some $\mathbf{s}$, outputs $\mathbf{s}$ with probability exponentially close to 1.*

Proof. We only show how $\mathcal{D}'$ find the first coordinate of $\mathbf{s}$ $s_1 \in \mathbb{Z}_q$. For any $k \in \mathbb{Z}_q$, consider the following transformation. Given a tuple $(\mathbf{a}, \mathbf{e}, b)$ we output the tuple $(\mathbf{a} + {}^t(l, 0, \ldots, 0), \mathbf{e}, b + lk)$ where $l \in \mathbb{Z}_q$ is chosen uniformly at random. This random transformation takes U' into itself. Moreover, if $k = s_1$ then this transformation also takes $A_{\mathbf{s},\mathbf{s}'}$ into itself. Finally, if $k \neq s_1$ then it transforms $A_{\mathbf{s},\mathbf{s}'}$ to U'. Therefore, using $\mathcal{D}$, we can test whether $k = s_1$ or not. Since there are only $q < \mathrm{poly}(n)$ possibilities for s_1, we can try all of them. □

Remark 4.3.11. The hardness of the worst case of $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ implies the hardness of the worst case of $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$. Note that it is unknown if the converse statement holds. We also note that, from Theorem 4.3.2, there exists a quantum reduction from $\mathrm{LWE}_{q,\bar{\Psi}_\alpha}$ to $\mathrm{SVP}_{\tilde{O}(n/\alpha)}$ and $\mathrm{SIVP}_{\tilde{O}(n/\alpha)}$.

4.4 Proofs of Knowledge on Its Secret Key

Recall that we can consider ${}^t\mathbf{A}$ as a generator matrix from a coding-theoretical view and a parity-check matrix $\mathbf{H}$ is easily computed. Informally, if Alice wants to prove that she has a secret key corresponding to a public key $\mathbf{b}$, it is sufficient that she proves that she has an error key $\mathbf{s}'$ such that $\mathbf{HEs}' = \mathbf{Hb}$.

Definition 4.4.1 (Relation R_{mR05})**.** Let $(\mathbf{A}, \mathbf{E}, \mathbf{b})$ be a public key of mR05, $\mathbf{H}$ a parity-check matrix of $\mathbf{A}$, $\mathbf{s}$ a vector in $\mathbb{Z}_q^n$, and $\mathbf{s}'$ a vector in $\mathbb{Z}_q^{m_1}$. We say that input $(\mathbf{A}, \mathbf{H}, \mathbf{E}, \mathbf{b})$ and witness $(\mathbf{s}, \mathbf{s}')$ are in R_{mR05} if $\mathbf{s}' \in \mathrm{Set}_{m_1,m_2}$, $\mathbf{As} + \mathbf{Es}' = \mathbf{b}$, and $\mathbf{HEs}' = \mathbf{Hb}$.

Next, we describe the protocol for a proof of knowledge for a secret key, which is mainly based on a proof of knowledge for SDP by Stern [Ste96].

Protocol 4.4.2 (Protocol PSK)**.** Let P and V be a prover and a verifier respectively. The CRS is $\mathbf{A}, \mathbf{E}$. The common input is $\mathbf{b}$. The auxiliary inputs to the prover are $\mathbf{s}$ and $\mathbf{s}'$ such that $\mathbf{b} = {}^t\mathbf{As} + \mathbf{Es}'$. Let $\mathrm{Com}(\cdot;\cdot) = \mathrm{Com}_{\mathbf{A}}(\cdot;\cdot)$.

Step P1 Choose a random permutation π for $\{1, \ldots, m_1\}$ and a random vector $\mathbf{y} \in \mathbb{Z}_q^{m_1}$. Compute $c_1 = \mathrm{Com}(\pi, \mathbf{HEy}; r_1)$, $c_2 = \mathrm{Com}(\pi(\mathbf{y}); r_2)$ and $c_3 = \mathrm{Com}(\pi(\mathbf{y} + \mathbf{s}'); r_3)$. Send c_1, c_2, c_3 to V.

Step V1 V sends a random challenge bit $\delta \in_R \{1, 2, 3\}$ to P.

Step P2 If $\delta = 1$, P opens c_1 and c_2 (i.e., sends $\pi, \mathbf{y}, r_1$, and r_2 to V). If $\delta = 2$, P opens c_1 and c_3 (i.e., sends $\pi, \mathbf{y} + \mathbf{s}', r_1$ and r_3 to V). If $\delta = 3$, P opens c_2 and c_3 (i.e., sends $\pi(\mathbf{s}'), \pi(\mathbf{y}), r_2$, and r_3 to V).

Step V2 If $\delta = 1$, received $\tilde{\pi}$, $\tilde{\mathbf{y}}$, $\tilde{r}_1$, and $\tilde{r}_2$, check the commitments c_1 and c_2 were correct (i.e., $c_1 = \mathrm{Com}(\tilde{\pi}, \mathbf{HE}\tilde{\mathbf{y}}; \tilde{r}_1)$ and $c_2 = \mathrm{Com}(\tilde{\pi}(\tilde{\mathbf{y}}); \tilde{r}_2)$). If $\delta = 2$, received $\tilde{\pi}$, $\tilde{\mathbf{x}}$, $\tilde{r}_1$, and $\tilde{r}_3$, check that the commitments c_1 and c_3 were correct (i.e., $c_1 = \mathrm{Com}(\tilde{\pi}, \mathbf{HE}\tilde{\mathbf{x}} - \mathbf{Hb}; \tilde{r}_1)$ and $c_3 = \mathrm{Com}(\tilde{\pi}(\tilde{\mathbf{x}}); \tilde{r}_3)$). If $\delta = 3$, received $\tilde{\mathbf{x}}_1$, $\tilde{\mathbf{x}}_2$, $\tilde{r}_2$, and $\tilde{r}_3$, check that the commitments c_2 and c_3 were correct (i.e., $c_2 = \mathrm{Com}(\tilde{\mathbf{x}}_1; \tilde{r}_2)$ and $c_3 = \mathrm{Com}(\tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_2; \tilde{r}_3)$) and that $w_H(\tilde{\mathbf{x}}_2) = m_2$.

Theorem 4.4.3. *An interactive protocol (P, V) is a proof of knowledge system with knowledge error $2/3$ for R_{mR05}. Moreover, the protocol (P, V) is a statistical zero-knowledge argument for R_{mR05} in CRS model under the assumption that the worst case of $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ and $\mathrm{SVP}_{O(\tilde{n})}$ is hard.*

Proof of completeness. We omit the proof since it is evident. □

We use Lemma 4.4.4 below in [Ste96] in the proof of knowledge error.

Lemma 4.4.4 (Theorem 1 and Lemma 1, [Ste96]). *Assume that some probabilistic polynomial-time adversary P^* is accepted with probability at least $(2/3)^r + \epsilon$, $\epsilon > 0$, after playing the identification protocol r times. Then there exists a polynomial-time probabilistic machine K such that outputs the witness $\mathbf{s}'$ from the common input or else finds collisions for the hash function with probability larger than $\epsilon^3/10$.*

The idea of Lemma 4.4.4 is follows: Assume that P^* can output response to all V's challenges correctly. Let P's response to V's challenge 1 be $\tilde{\pi}_1$, $\tilde{\mathbf{y}}$, $\tilde{r}_{1,1}$, and $\tilde{r}_{1,2}$. Let P's response to V's challenge 2 be $\tilde{\pi}_2$, $\tilde{\mathbf{x}}$, $\tilde{r}_{2,1}$, and $\tilde{r}_{2,3}$. Finally, let P's response to V's challenge 3 be $\tilde{\mathbf{x}}_1$, $\tilde{\mathbf{x}}_2$, $\tilde{r}_{3,2}$ and $\tilde{r}_{3,3}$. Since all response are correct, we obtain that

$$
\begin{aligned}
c_1 &= \mathrm{Com}(\tilde{\pi}_1, \mathbf{HE}\tilde{\mathbf{y}}; \tilde{r}_{1,1}) &&= \mathrm{Com}(\tilde{\pi}_2, \mathbf{HE}\tilde{\mathbf{x}} - \mathbf{Hb}; \tilde{r}_{2,1}) \\
c_2 &= \mathrm{Com}(\tilde{\pi}_1(\tilde{\mathbf{y}}); \tilde{r}_{1,2}) &&= \mathrm{Com}(\tilde{\mathbf{x}}_1; \tilde{r}_{3,2}) \\
c_3 &= \mathrm{Com}(\tilde{\pi}_2(\tilde{\mathbf{x}}); \tilde{r}_{2,3}) &&= \mathrm{Com}(\tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_2; \tilde{r}_{3,3})
\end{aligned}
$$

If there exists a distinct pair in the inputs of commitment, we find a collision. Then, we assume there exists no distinct pair in P^*'s responses. Since P^* is accepted, $w_H(\tilde{\mathbf{x}}_2) = m_2$. From c_1's equation, $\tilde{\pi}_1 = \tilde{\pi}_2$. Combining $\tilde{\pi}_1 = \tilde{\pi}_2$ and c_3's equations, we obtain $\tilde{\mathbf{x}} = \tilde{\pi}_2^{-1}(\tilde{\mathbf{x}}_1) + \tilde{\pi}_2^{-1}(\tilde{\mathbf{x}}_2)$. From c_2's equation, we have that $\tilde{\mathbf{y}} = \tilde{\pi}_2^{-1}(\tilde{\mathbf{x}}_1)$. Therefore, combining the above argument and c_1's equation, we obtain $\mathbf{Hb} = \mathbf{HE}(\tilde{\mathbf{x}} - \tilde{\mathbf{y}}) = \mathbf{HE}\tilde{\pi}_2^{-1}(\tilde{\mathbf{x}}_2)$ and a witness $\tilde{\pi}_2^{-1}(\tilde{\mathbf{x}}_2)$. Thus, we obtain a collision or a witness using P^*.

Proof of knowledge error with $2/3$. Assume that some probabilistic polynomial-time adversary P^* in Lemma 4.4.4. Using Lemma 4.4.4, we obtain K in the above. In Stern's proof, he consider binary linear codes. Although we play the protocol in q-ary linear codes, we can apply Stern's proof to q-ary codes. Note that, under the assumption that the worst case of $\mathrm{SVP}_{\tilde{O}(n)}$ is hard, finding collision is hard [MR04]. Therefore if assume that $\mathrm{SVP}_{\tilde{O}(n)}$ is hard in the worst case, we obtain a knowledge extractor K. □

Proof of zero knowledge. We construct the simulator as follows.

Step P1 Choose $\Delta \in \{1, 2, 3\}$ randomly. Choose a permutation π, a vector $\mathbf{y} \in \mathbb{Z}_q^{m_1}$, a vector $\mathbf{s}' \in \mathrm{Set}_{m_1,m_2}$ uniformly at random.

1. $\Delta = 1$: Compute $c_1 = \mathrm{Com}(\pi, \mathbf{HE}(\mathbf{y} + \mathbf{s}') - \mathbf{Hb}; r_1)$, $c_2 = \mathrm{Com}(\pi(\mathbf{y}); r_2)$, and $c_3 = \mathrm{Com}(\pi(\mathbf{y} + \mathbf{s}'); r_3)$. Sends c_1, c_2, and c_3 to V^*.
2. $\Delta = 2$: Compute $c_1 = \mathrm{Com}(\pi, \mathbf{HE}\mathbf{y}; r_1)$, $c_2 = \mathrm{Com}(\pi(\mathbf{y}); r_2)$, and $c_3 = \mathrm{Com}(\pi(\mathbf{y} + \mathbf{s}'); r_3)$. Sends c_1, c_2, and c_3 to V^*.

3. $\Delta = 3$: Compute $\mathbf{x} \in \mathbb{Z}_q^{m_1}$ such that $\mathbf{HEx} = \mathbf{HEy} + \mathbf{Hb}$. Compute $c_1 = \mathrm{Com}(\pi, \mathbf{HEy}; r_1)$, $c_2 = \mathrm{Com}(\pi(\mathbf{y}); r_2)$, and $c_3 = \mathrm{Com}(\pi(\mathbf{x}); r_3)$. Sends c_1, c_2, and c_3 to V^*.

Step V1 Receive a challenge $\delta \in \{1, 2, 3\}$.

Step P2 If $\Delta = \delta$ then output $\perp$ and halt. Else,

1. $(\Delta, \delta) = (1, 2)$: Send $\tilde{\pi} = \pi$, $\tilde{\mathbf{x}} = \pi(\mathbf{y} + \mathbf{s}')$, $\tilde{r}_1 = r_1$, and $\tilde{r}_3 = r_3$ to V^*.
2. $(\Delta, \delta) = (1, 3)$: Send $\tilde{\mathbf{x}}_1 = \pi(\mathbf{y})$, $\tilde{\mathbf{x}}_2 = \pi(\mathbf{s}')$, $\tilde{r}_2 = r_2$, and $\tilde{r}_3 = r_3$ to V^*.
3. $(\Delta, \delta) = (2, 1)$: Send $\tilde{\pi} = \pi$, $\tilde{\mathbf{y}} = \mathbf{y}$, $\tilde{r}_1 = r_1$, and $\tilde{r}_2 = r_2$ to V^*.
4. $(\Delta, \delta) = (2, 3)$: Send $\tilde{\mathbf{x}}_1 = \pi(\mathbf{y})$, $\tilde{\mathbf{x}}_2 = \pi(\mathbf{s}')$, $\tilde{r}_2 = r_2$, and $\tilde{r}_3 = r_3$ to V^*.
5. $(\Delta, \delta) = (3, 1)$: Send $\tilde{\pi} = \pi$, $\tilde{\mathbf{y}} = \mathbf{y}$, $\tilde{r}_1 = r_1$, and $\tilde{r}_2 = r_2$ to V^*.
6. $(\Delta, \delta) = (3, 2)$: Send $\tilde{\pi} = \pi$, $\tilde{\mathbf{x}} = \pi^{-1}(\mathbf{x})$, $\tilde{r}_1 = r_1$, and $\tilde{r}_3 = r_3$ to V^*.

Output the transcript and halt.

Since Com is statistically hiding, the simulator's outputs when the simulator did not output $\perp$ is statistically close to the real transcript. □

4.5 Signature Schemes

Background. In 1986, Fiat and Shamir proposed zero-knowledge proof of knowledge for the quadratic residue on $n = pq$ whose factorization is unknown. Firstly, Pointcheval and Stern [PS96] showed the securities of some signature schemes, the Fiat-Shamir signature and the ElGamal signature, in the random oracle model. Along this direction, Ohta and Okamoto [OO98] proved that a signature scheme from honest-verifier public-coin perfect zero-knowledge protocol via the Fiat-Shamir transformation is polynomially secure against chosen-message attacks in the random oracle model. Recently, Abdalla, An, Bellare, and Namprempre proved that a signature scheme from a polynomially-secure identification scheme via the Fiat-Shamir transformation is polynomially secure in the above sense [AABN02].

However, their proofs do not imply the security in real world. Indeed, Goldwasser and Tauman Kalai [GTK03] showed that existence of a signature scheme, obtained from a secure identification scheme via the Fiat-Shamir transformation, which is not secure in real world though secure in the random oracle model.

4.5.1 The Fiat-Shamir Transformation

We summary results in [AABN02]. We first review definitions and notations in [AABN02].

Canonical identification schemes. Let $\mathcal{ID} = (K, P, V, c)$ be an identification schemes; where K is the key generation algorithm which on input $n \in \mathbb{N}$ outputs $(\mathsf{sk}, \mathsf{pk})$, P is the prover algorithm taking input sk, V is the verifier algorithm taking input pk, and c is a function of n indicating the length of the verifier's challenge. We say $\mathcal{ID}$ is a canonical identification schemes if it is a public-coin 3-round protocol. See Figure 4.5.1 for details.

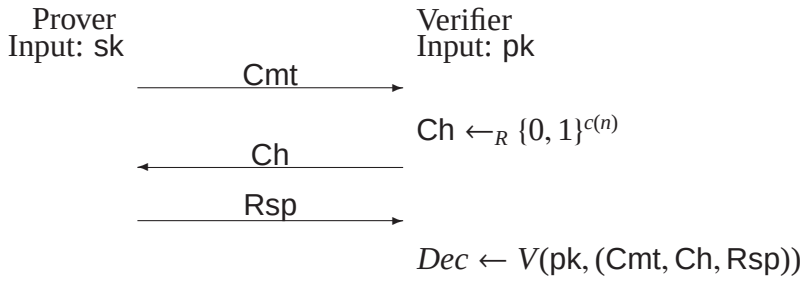

Figure 4.1: A canonical identification protocol.

Next, we define the security of an identification scheme. Let $\mathcal{TR}$ be a randomized transcript generation oracle which takes no inputs and returns a random transcript of an "honest" execution:

Oracle $\mathcal{TR}^{\mathcal{ID}}_{\mathsf{pk},\mathsf{sk},1^n}$**:**

Choose a random tape r of P

$\mathsf{Cmt} \leftarrow P(\mathsf{sk}; r)$; $\mathsf{Ch} \leftarrow_R \{0,1\}^{c(n)}$; $\mathsf{Rsp} \leftarrow P(\mathsf{sk}, (\mathsf{Cmt}, \mathsf{Ch}); r)$;

Return $(\mathsf{Cmt}, \mathsf{Ch}, \mathsf{Rsp})$

Definition 4.5.1 (Definition 2.1, [AABN02])**.** Let $\mathcal{ID} = (K, P, V, c)$ be an identification scheme, and let $\mathcal{I}$ be an impersonator, st be its state, and n be the security parameter. Define the *advantage of* $\mathcal{I}$ as

$$\mathbf{Adv}^{\mathrm{imp-pa}}_{\mathcal{ID},\mathcal{I}}(n) = \Pr\left[\mathbf{Exp}^{\mathrm{imp-pa}}_{\mathcal{ID},\mathcal{I}}(n) = 1\right],$$

where the experiment in question is

$\mathbf{Exp}^{\mathrm{imp-pa}}_{\mathcal{ID},\mathcal{I}}(n)$**:**

$(\mathsf{pk}, \mathsf{sk}) \leftarrow K(1^n)$; $(\mathsf{st}, \mathsf{Cmt}) \leftarrow \mathcal{I}^{\mathcal{TR}^{\mathcal{ID}}_{\mathsf{pk},\mathsf{sk},1^n}}(\mathsf{pk}, 1^n)$; $\mathsf{Ch} \leftarrow_R \{0,1\}^{c(n)}$

$\mathsf{Rsp} \leftarrow \mathcal{I}(\mathsf{st}, \mathsf{Ch})$; $\mathsf{Dec} \leftarrow V(\mathsf{pk}, (\mathsf{Cmt}, \mathsf{Ch}, \mathsf{Rsp}))$; Return Dec.

We say that $\mathcal{ID}$ is polynomially-secure against impersonation under passive attacks if $\mathbf{Adv}^{\mathrm{imp-pa}}_{\mathcal{ID},\mathcal{I}}(\cdot)$ is negligible for every probabilistic poly(n)-time impersonator $\mathcal{I}$.

Signature Schemes. Let $\mathcal{DS} = (K, S, Vf, c)$ be a digital signature scheme; where K is the key generation algorithm which on input $n \in \mathbb{N}$ outputs $(\mathsf{sk}, \mathsf{pk})$, S is the signing algorithm taking input sk and a message $M \in \{0,1\}^*$ and return a signature σ for M, Vf is the verification algorithm taking input pk, a message M, and a signature σ for M and returning a boolean decision. The signing and verifying algorithms have oracle access to a function $H : \{0,1\}^* \to \{0,1\}^{c(n)}$, which is the random oracle. The security of a signature scheme is defined as follows:

Definition 4.5.2 (Definition 2.2, [AABN02])**.** Let $\mathcal{DS} = (K, S, Vf, c)$ be a digital signature scheme, let $\mathcal{F}$ be a forger and $n \in \mathbb{N}$ the security parameter. Define the experiment

$\mathbf{Exp}^{\text{frg-cma}}_{\mathcal{DS},\mathcal{F}}(n)$:

$H \leftarrow_R \{f : \{0,1\}^* \to \{0,1\}^{c(n)}\}$;

$(\mathsf{pk}, \mathsf{sk}) \leftarrow K(1^n)$; $(M, \sigma) \leftarrow F^{S^H_{\mathsf{sk}}(\cdot), H(\cdot)}(1^n, \mathsf{pk})$; $\mathsf{Dec} \leftarrow Vf^H(\mathsf{pk}, (M, \sigma))$

If M was previously queried to $S^H_{\mathsf{sk}}(\cdot)$ then return 0 else return Dec.

Define the advantage of $\mathcal{F}$ as

$$\mathbf{Adv}^{\text{frg-cma}}_{\mathcal{DS},\mathcal{F}}(n) = \Pr\left[\mathbf{Exp}^{\text{frg-cma}}_{\mathcal{DS},\mathcal{F}}(n) = 1\right].$$

We say $\mathcal{DS}$ is polynomially-secure against chosen-message attacks is a is negligible for every probabilistic poly(n)-time forger $\mathcal{F}$.

The Fiat-Shamir Transformation. The idea of transformation is replacing a public coin Ch by the function H. We note formal construction.

Construction 4.5.3. Let $\mathcal{ID} = (K, P, V, c)$ be a canonical identification scheme. We associate to these a digital signature scheme $\mathcal{DS} = (K, S, Vf, c)$. It has the same key generation algorithm as the identification scheme, and the output length of the hash function equals to the challenge length of the identification scheme. (Let $H : \{0,1\}^* \to \{0,1\}^{c(n)}$ be a hash function.) The signing and verifying algorithms are defined as follows:

Algorithm $S^H(\mathsf{sk}, M)$

Choose a random tape r of P;

$\mathsf{Cmt} \leftarrow P(\mathsf{sk}; r)$; $\mathsf{Ch} \leftarrow H(\mathsf{Cmt}||M)$; $\mathsf{Rsp} \leftarrow P(\mathsf{sk}, (\mathsf{Cmt}, \mathsf{Ch}); r)$;

Return $(\mathsf{Cmt}, \mathsf{Rsp})$

Algorithm $Vf^H(\mathsf{pk}, M, \sigma)$

Parse σ as $(\mathsf{Cmt}, \mathsf{Ch})$;

$\mathsf{Ch} \leftarrow H(\mathsf{Cmt}||M)$; $\mathsf{Dec} \leftarrow V(\mathsf{pk}, (\mathsf{Cmt}, \mathsf{Ch}, \mathsf{Rsp}))$;

Return Dec

Finally, we review main theorem of [AABN02]. In [AABN02] they show the generalized Fiat-Shamir transformation and the security of obtained signature scheme in the random oracle model. Our strategy is using the basic Fiat-Shamir transformation. Thus, we quote only their theorem considering the basic transformation.

Theorem 4.5.4 (Therem 3.3, [AABN02]). *Let $\mathcal{ID} = (K, P, V, c)$ be a non-trivial and canonical identification scheme. Let $\mathcal{DS} = (K, S, Vf, c)$ be the associated signature scheme as per Construction 4.5.3. Then $\mathcal{DS}$ is polynomially-secure against chosen-message attacks in the random oracle model if and only if $\mathcal{ID}$ is polynomially-secure against impersonation under passive attacks.*

4.5.2 Concrete Signature Scheme

We first parallelize Protocol 4.4.2 and obtain a canonical identification scheme.

Protocol 4.5.5 (Identification Scheme). Let P and V be a prover and a verifier respectively. The CRS is $\mathbf{A}, \mathbf{E}$. The common input is $\mathbf{b}$. The auxiliary inputs to the prover are $\mathbf{s}$ and $\mathbf{s}'$ such that $\mathbf{b} = {}^t\mathbf{A}\mathbf{s} + \mathbf{E}\mathbf{s}'$ and $w_H(\mathbf{s}') = m_2$. Let $\mathrm{Com}(\cdot;\cdot) = \mathrm{Com}_{\mathbf{A}}(\cdot;\cdot)$. This protocol is obtained by parallelizing the Protocol 4.4.2 n times.

Step P1 Choose n random permutations π_i for $\{1, \ldots, m_1\}$ and n random vectors $\mathbf{y}_i \in \mathbb{Z}_q^{m_1}$. Compute $c_{i,1} = \mathrm{Com}(\pi_i, \mathbf{H}\mathbf{E}\mathbf{y}_i; r_{i,1})$, $c_{i,2} = \mathrm{Com}(\pi_i(\mathbf{y}_i); r_{i,2})$ and $c_{i,3} = \mathrm{Com}(\pi_i(\mathbf{y}_i + \mathbf{s}'); r_{i,3})$. Set $\mathsf{Cmt} := \{(c_{i,1}, c_{i,2}, c_{i,3})\}_{i=1,\ldots,n}$ and send Cmt to V.

Step V1 V sends n random challenge bits $\mathsf{Ch} := \delta_1 \| \ldots \| \delta_n \in_R \{1,2,3\}^n$ to P.

Step P2 Parse Ch as $\delta_1 \| \ldots \| \delta_n \in \{1,2,3\}^n$.

1. If $\delta_i = 1$, set $\mathsf{Rsp}_i := (\pi_i, \mathbf{y}_i, r_{i,1}, r_{i,2})$.
2. If $\delta_i = 2$, set $\mathsf{Rsp}_i := (\pi_i, \mathbf{y}_i + \mathbf{s}', r_{i,1}, r_{i,3})$.
3. If $\delta_i = 3$, set $\mathsf{Rsp}_i := (\pi_i(\mathbf{s}'), \pi_i(\mathbf{y}_i), r_{i,2}, r_{i,3})$.

Set $\mathsf{Rsp} := \{\mathsf{Rsp}_i\}_{i=1,\ldots,n}$ and send Rsp to V.

Step V2 Parse Rsp as $\{\mathsf{Rsp}_i\}_{i=1,\ldots,n}$.

1. If $\delta_i = 1$, received $\mathsf{Rsp}_i = (\tilde{\pi}_i, \tilde{\mathbf{y}}_i, \tilde{r}_{i,1}, \tilde{r}_{i,2})$, check the commitments $c_{i,1}$ and $c_{i,2}$ were correct (i.e., $c_{i,1} = \mathrm{Com}(\tilde{\pi}_i, \mathbf{H}\mathbf{E}\tilde{\mathbf{y}}_i; \tilde{r}_{i,1})$ and $c_2 = \mathrm{Com}(\tilde{\pi}_i(\tilde{\mathbf{y}}_i); \tilde{r}_{i,2})$). If correct, set $\mathsf{Dec}_i := 1$. Otherwise $\mathsf{Dec}_i := 0$.
2. If $\delta_i = 2$, received $\mathsf{Rsp}_i = (\tilde{\pi}_i, \tilde{\mathbf{x}}_i, \tilde{r}_{i,1}, \tilde{r}_{i,3})$, check that the commitments $c_{i,1}$ and $c_{i,3}$ were correct (i.e., $c_{i,1} = \mathrm{Com}(\tilde{\pi}_i, \mathbf{H}\mathbf{E}\tilde{\mathbf{x}}_i - \mathbf{H}\mathbf{b}; \tilde{r}_{i,1})$ and $c_{i,3} = \mathrm{Com}(\tilde{\pi}_i(\tilde{\mathbf{x}}_i); \tilde{r}_{i,3})$). If correct, set $\mathsf{Dec}_i := 1$. Otherwise $\mathsf{Dec}_i := 0$.

3. If $\delta_i = 3$, received $\mathsf{Rsp}_i = (\tilde{\mathbf{x}}_{i,1}, \tilde{\mathbf{x}}_{i,2}, \tilde{r}_{i,2}, \tilde{r}_{i,3})$, check that the commitments $c_{i,2}$ and $c_{i,3}$ were correct and the weight of witness is correct (i.e., $c_{i,2} = \mathrm{Com}(\tilde{\mathbf{x}}_{i,1}; \tilde{r}_{i,2})$ and $c_{i,3} = \mathrm{Com}(\tilde{\mathbf{x}}_{i,1} + \tilde{\mathbf{x}}_{i,2}; \tilde{r}_{i,3})$) and that $w_H(\tilde{\mathbf{x}}_{i,2}) = m_2$.

If all Dec_i is 1 then set $\mathsf{Dec} := 1$. Otherwise $\mathsf{Dec} := 0$. Output Dec.

Applying the Fiat-Shamir transformation to Protocol 4.5.5, we obtain the following signature scheme.

Signature Scheme 4.5.6. Let P and V be a prover and a verifier respectively. The CRS is $\mathbf{A}, \mathbf{E}$. Let $H : \{0,1\}^* \to \{0,1\}^n$ be a random oracle. This protocol is obtained by applying the Fiat-Shamir transformation.

Key Generation: Same as that of mbR05 (See Cryptosystem 4.3.5). The secret key is $(\mathbf{s}, \mathbf{s}')$. The public key is $(\mathbf{A}, \mathbf{E}, \mathbf{b})$.

Signing: Let a message be $M \in \{0,1\}^*$. First, compute Cmt as Step P1 in Protocol 4.5.5. Secondly, query to the random oracle H and obtain $\mathsf{Ch} := H(\mathsf{Cmt}||M)$. Finally, compute Rsp as Step P2 in Protocol 4.5.5. The signature is $\sigma := (\mathsf{Cmt}, \mathsf{Rsp})$.

Verifying: The input is the public key $(\mathbf{A}, \mathbf{E}, \mathbf{b})$, the message M, and the signature σ. First, parse σ as $(\mathsf{Cmt}, \mathsf{Rsp})$. Secondly, query to the random oracle H and obtain $\mathsf{Ch} := H(\mathsf{Cmt}||M)$. Finally, decide accept or reject as Step V2 in Protocol 4.5.5.

Let us show the security of the underlying identification scheme Protocol 4.5.5.

Lemma 4.5.7. *Assume that* $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ *and* $\mathrm{SVP}_{\tilde{O}(n)}$ *is hard in the worst case. Then, the underlying identification scheme, Protocol 4.5.5, is polynomially-secure against impersonation under passive attacks.*

Proof. The underlying identification scheme, Protocol 4.5.5, is obtained by parallelizing Protocol 4.4.2. We show if there exists an impersonator P^* which can impersonate Protocol 4.5.5 with non-negligible probability ϵ, then there exists an adversary $\mathcal{A}$ which can obtain the witness $\mathbf{s}'$ for $(\mathbf{A}, \mathbf{E}, \mathbf{b})$ using P^*.

As in Lemma 4.4.4, $\mathcal{A}$ finds a collision or obtain the witness $\mathbf{s}'$ for the common input $(\mathbf{A}, \mathbf{E}, \mathbf{b})$ with probability larger than $\epsilon^3/10$. From Theorem 4.3.7, if $\mathcal{A}$ exists then an adversary $\mathcal{A}'$ such that solves $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ in the worst case. □

Combining Theorem 4.5.4 and Lemma 4.5.7 below, we obtain the following theorem.

Theorem 4.5.8. *Assume that* $\mathrm{LWKE}_{q,\bar{\Psi}_\alpha}$ *and* $\mathrm{SVP}_{\tilde{O}(n)}$ *is hard in the worst case. Then, the above signature scheme (4.5.6) is polynomially-secure against chosen-message attacks in the random oracle model.*

We remark that if we use a random oracle as hash function for string commitment, we can omit the assumption that $\mathrm{SVP}_{\tilde{O}(n)}$ is hard in the worst case.

4.6 Concluding Remarks

In this chapter, we have proposed a modified Regev'05 cryptosystem (mR05) and introduced a proof of knowledge on its secret key. We stress that our signature scheme is based on the worst-case assumption.

At the end, we list up a few open problems: (1) A proof of knowledge on a secret key of the original Regev'05 cryptosystem (R05); mR05 needs stronger assumption than one which R05 needs. (2) Relation between LWE and LWKE; we have failed to show a reduction from LWE to LWKE. (3) Zero knowledge on coding problems; As seen in Section 4.1, the MV protocol can not apply to coding problems. Thus, we need a direct protocol for coding problems.

Chapter 5

Proofs of Plaintext Knowledge for the Regev'04 and Regev'05 Cryptosystems

5.1 Introduction

Proof of Plaintext Knowledge. Given an instance of a public-key cryptosystem with public key pk, a proof of plaintext knowledge (PPK) allows a prover to prove knowledge of the plaintext m of ciphertext $c \in E_{\mathsf{pk}}(m)$ to a verifier. If both the prover and the verifier are online, IND-CPA public-key cryptosystems with PPK protocol achieves interactive IND-CCA1 security [GHY85, Gol01]. It was known that efficient proofs of plaintext knowledge for the number-theoretic public-key cryptosystems, such that Rabin, RSA, ElGamal, and etc., using zero-knowledge public-coin proofs of knowledge protocols with 3 rounds (known as Σ-protocol). However, efficient proofs of plaintext knowledge for the lattice-based cryptosystems were unknown except that in [GK05].

Summary of Our Results. We construct PPK protocols for slightly modified versions of the Regev'04 cryptosystem (pR04) and the Regev'05 cryptosystem (pR05) based on the protocol in [GK05].

We show a relation between ciphertexts of cryptosystems, pR04 and pR05, and instances of GapCVP_γ. Although the cryptosystems are less secure than the original ones, we can show that their security are based on the worst-case of certain lattice problems as in Kawachi, Tanaka, and Xagawa [KTX06].

Our connection between the ciphertexts and GapCVP_γ implies that if we set a large factor for the underlying lattice problems, for small n, the LLL algorithm [LLL82] heuristically succeed to distinguish ciphertexts of 0 and 1. From the positive view, we can apply Micciancio

and Vadhan's zero-knowledge protocol for GapCVP$_\gamma$ [MV03] and obtain a verifiable encryption scheme. Based on the protocol in [GK05] and the above connection, we construct a proof of plaintext knowledge for pR04 and pR05.

Organization. In Section 5.2 we review tools for construction of proof of plaintext knowledge. In Section 5.3, we construct a proof of plaintext knowledge for the modified Regev'04 cryptosystem. We also construct it for the modified Regev'05 cryptosystem in Section 5.4. Finally we conclude in Section 5.5.

5.2 Tools for Proof of Plaintext Knowledge

5.2.1 The Ajtai-Dwork Cryptosystem and Nguyen and Stern's Embedding

The Ajtai-Dwork cryptosystem is a 1-bit lattice-based cryptosystem. Nguyen and Stern showed how to reduce distinguishing encryptions of 0 from one of 1 to GapCVP$_\gamma$ for some $\gamma > 1$. We briefly review error-free version of the Ajtai-Dwork cryptosystem, which is proposed by Goldreich, Goldwasser, and Halevi [GGH97a], and Nguyen and Stern's embedding techniques [NS98]. For more details of the embedding techniques, see [NS98, Section 4].

First, we briefly review the Ajtai-Dwork cryptosystem. For more details, see Section 3.2.1. The secret key of the Ajtai-Dwork cryptosystem is $\mathbf{u} \in \mathbb{R}^n$ whose length is 1. The public key is $m+n$ vectors in n-dimensional space and an index. We denote it as $(\mathbf{w}_1, \ldots, \mathbf{w}_n, \mathbf{v}_1, \ldots, \mathbf{v}_m, i_0)$. The vectors $\mathbf{w}_i, \mathbf{v}_i$ are chosen from hyperplanes $\{\mathbf{x} \in [0, n^n]^n \mid \langle \mathbf{x}, \mathbf{u} \rangle \in \mathbb{Z}\}$ and "blurred" by adding small noises. The index i_0 is chosen from $\{1, \ldots, m\}$ such that $\langle \mathbf{u}, \mathbf{v}_i \rangle$ is near by odd integers. Encryption of $\sigma \in \{0, 1\}$ is produced as follows: (1) Choose random string $r = r_1 \ldots r_m \in \{0,1\}^m$. (2) Compute $\mathbf{c} = (\sigma/2)\mathbf{v}_{i_0} + \sum_{i=1}^m r_i \mathbf{v}_i \bmod \mathcal{P}(\mathbf{w}_1, \ldots, \mathbf{w}_n)$. We decrypt a ciphertext $\mathbf{c} \in \mathcal{P}(\mathbf{w}_1, \ldots, \mathbf{w}_n)$ into 0 if $\mathrm{frc}(\langle \mathbf{c}, \mathbf{u} \rangle) \leq 1/4$ and into 1 if $\mathrm{frc}(\langle \mathbf{c}, \mathbf{u} \rangle) > 1/4$.

Nguyen and Stern showed the following embeddings [NS98]. For any public key pk of the

Ajtai-Dwork cryptosystem, let $\mathbf{B}_{\mathsf{pk}} \in \mathbb{R}^{(2n+m)\times(n+m)}$ be

$$\mathbf{B}_{\mathsf{pk}} = \begin{bmatrix} K_1\mathbf{w}_1 & \dots & K_1\mathbf{w}_n & K_1\mathbf{v}_1 & \dots & K_1\mathbf{v}_m \\ 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & K_2 & & \\ & & & & \ddots & \\ & & & & & K_2 \end{bmatrix},$$

where K_1 and K_2 are suitably chosen and all empty spaces are set by 0. For any ciphertext $\mathbf{c} \in \mathcal{P}(\mathbf{w}_1, \dots, \mathbf{w}_n)$, define $\mathbf{x}_{\mathbf{c}} = \binom{K_1\mathbf{c}}{\mathbf{0}} \in \mathbb{R}^{2n+m}$. Nguyen and Stern showed for suitably chosen K_1 and K_2, $\mathrm{Dist}(\mathbf{x}_{\mathbf{c}}, L(\mathbf{B}_{\mathsf{pk}}))$ is small if $\mathbf{c}$ is a legal ciphertext of 0 with pk and $\mathrm{Dist}(\mathbf{x}_{\mathbf{c}}, L(\mathbf{B}_{\mathsf{pk}}))$ is large if $\mathbf{c}$ decrypts into 1 with high probability.

5.2.2 Micciancio and Vadhan's Zero-Knowledge Protocol

In [MV03], Micciancio and Vadhan introduced a zero-knowledge protocol for GapCVP_{γ}. They use the following observation by Goldreich and Goldwasser [GG00]. Consider two n-dimensional unit hyperballs, one center locates the origin and the other center locates the point that distance is d, i.e., $B(\mathbf{0}, 1)$ and $B(\mathbf{y}, 1)$, where $\|\mathbf{y}\| = d$. If $d = \Omega(\sqrt{n/\log n})$, ratio between a volume of an intersection of two hyperballs and a volume of a hyperball is $1/\mathrm{poly}(n)$. Based on this observation, Goldreich and Goldwasser showed statistical zero-knowledge protocol for $\mathrm{coGapCVP}_{\Omega(\sqrt{n/\log n})}$ [GG00]. Micciancio and Vadhan also constructed honest-verifier statistical zero-knowledge proof system for $\mathrm{GapCVP}_{\Omega(\sqrt{n/\log n})}$ [MV03].

We refer Micciancio and Vadhan's protocol as the MV protocol.

Protocol 5.2.1 (The MV protocol, [MV03])**.** Let P_{MV} and V_{MV} denote the prover and the verifier, respectively. The common input is $(\mathbf{B}, \mathbf{y}, t)$. The auxiliary input to the prover is $\mathbf{w} \in \mathbb{Z}^n$ such that $\|\mathbf{B}\mathbf{w} - \mathbf{y}\| \leq t$.

Step P1 Choose k random bits $c_1, \dots, c_k \in \{0, 1\}$ independently. Also choose error vectors $\mathbf{r}_1, \dots, \mathbf{r}_k \in B(\mathbf{0}, \gamma t/2)$ independently and uniformly at random. Then, check if there exists an index i^* such that $\|\mathbf{r}_{i^*} + (2c_{i^*} - 1)\mathbf{u}\| \leq \gamma t/2$. If not, set $i^* = 1$ and redefine $c_{i^*} = 0$ and $\mathbf{r}_{i^*} = \mathbf{u}/2$, so that $\|\mathbf{r}_{i^*} + (2c_{i^*} - 1)\mathbf{u}\| \leq \gamma t/2$ is certainly satisfied. Finally, compute points $\mathbf{m}_i = c_i\mathbf{y} + \mathbf{r}_i \bmod \mathbf{B}$ for $i = 1, \dots, k$ and send them to V_{MV}.

Step V1 Send a random challenge bit $\delta \in \{0, 1\}$ to P_{MV}.

Step P2 Receive a challenge bit $\delta \in \{0,1\}$. If $\delta = \sum_{i=1}^{k} c_i \bmod 2$, then the prover completes the proof sending bits c_i and lattice vectors $\mathbf{Bv}_i = \mathbf{m}_i - (\mathbf{r}_i + c_i\mathbf{y})$ to V_{MV}. If $\delta \neq \sum_{i=1}^{k} c_i \bmod 2$, then the prover sends the same messages to V_{MV}, but with c_{i^*} and $\mathbf{Bv}_{i^*}$ replaced by $1 - c_{i^*}$ and $\mathbf{Bv}_{i^*} + (2c_{i^*} - 1)(\mathbf{y} - \mathbf{u})$.

Step V2 Receive k bits $c_1, \ldots, c_k$ and k lattice points $\mathbf{Bv}_1, \ldots, \mathbf{Bv}_k$ and check that they satisfy $\sum_{i=1}^{k} c_i = q \pmod 2$ and $\|\mathbf{m}_i - (\mathbf{Bv}_i + c_i\mathbf{y})\| \leq \gamma t/2$ for all $i = 1, \ldots, k$.

A completeness property is evident.

Theorem 5.2.2 (Zero Knowledge)**.** *$(P_{\mathrm{MV}}, V_{\mathrm{MV}})$ is a statistical zero-knowledge proof system with perfect completeness and soundness error $1/2$, provided one of the following conditions holds:*

- *$\gamma = \Omega(\sqrt{n/\log n})$ and $k = \mathrm{poly}(n)$ is a sufficiently large polynomial, or*
- *$\gamma = \Omega(\sqrt{n})$ and $k = \omega(\log n)$ is any superlogarithmic function of n, or*
- *$\gamma = n^{0.5+\Omega(1)}$ and $k = \omega(1)$ is any superconstant function of n.*

Theorem 5.2.3 (Proof of Knowledge)**.** *There is a probabilistic polynomial-time algorithm K_{MV} such that if a prover P^* makes V_{MV} accept with probability $1/2 + \epsilon$ on some instance $(\mathbf{B}, \mathbf{y}, t)$, then $K_{\mathrm{MV}}^{P^*}(\mathbf{B}, \mathbf{y}, t)$ outputs a vector $\mathbf{w} \in \mathbb{Z}^n$ satisfying $\|\mathbf{Bw} - \mathbf{y}\| \leq \gamma t$ with probability ϵ.*

5.2.3 A Proof of Plaintext Knowledge for the Ajtai-Dwork Cryptosystem

Goldwasser and Kharchenko [GK05] showed a interactive zero-knowledge proof of plaintext knowledge (PPK) for the Ajtai-Dwork cryptosystem using the above two results.

First, we immediately obtain a statistical zero-knowledge protocol for a statement that $\mathbf{c}$ is a legal ciphertext of 0 combining the above results. They also show a statistical zero-knowledge protocol for a statement that $\mathbf{c}$ is a legal ciphertext of 1 setting parameters carefully and using the fact that $\mathbf{c}_1 = \mathbf{v}_{i_0}/2 + \mathbf{c}_0 \bmod \mathcal{P}(\mathbf{w}_1, \ldots, \mathbf{w}_n)$ for some legal ciphertexts $\mathbf{c}_1$ of 1. Thus, in other words, they showed a verifiable encryption for a statement "the ciphertext $\mathbf{c}$ decrypts into σ."

They showed a proof of plaintext knowledge for the Ajtai-Dwork cryptosystem implicitly using pseudohomomorphism [KTX06] of the cryptosystem. We state informally their protocol: Let a common input be a pair $(\mathsf{pk}, \mathbf{c})$. The auxiliary inputs to the prover are a plaintext σ and a randomness that used in the ciphertext. In the first step, the prover makes a dummy ciphertext of a random bit σ'. The verifier sends a challenge bit δ. Suppose that $\delta = 0$. The prover sends the plaintext and the randomness that used in the dummy ciphertext. The verifier checks its consistency. Next, suppose that $\delta = 1$. The prover invokes a prover of the MV protocol with a

statement that the sum of input ciphertext and dummy one decrypts into $\sigma \oplus \sigma'$. The verifier invokes a verifier of the MV protocol. We note that the prover can not send the sum of plaintexts $\sigma \oplus \sigma'$ and the sum of randomness to the verifier since it leaks the part of the knowledge.

5.3 A Proof of Plaintext Knowledge for the Regev'04 Cryptosystem

5.3.1 The Regev'04 Cryptosystem

Instead of the original cryptosystem, we review the modified one in Section 3.3.1. Let $c \geq 0$ is a constant. The parameter of original one is $c = 0$.

Cryptosystem 5.3.1 (R04)**.** Let n be a security parameter, $N\ 2^{8n^2}$, and $m = c_m n^2$ where c_m is a constant. Let $\gamma(n) = \omega(n^{1+c}\sqrt{\log n})$. Let $H = \{h \in [\sqrt{N}, 2\sqrt{N}) \mid \mathrm{frc}(h) < 1/(8n^c m)\}$.

Private Key: Choose $h \in H$ uniformly at random. Let d denote N/h. The private key is the number h (or d).

Public Key: Choose $\alpha \in [2/\gamma(n), 2\sqrt{2}/\gamma(n))$ uniformly at random. We choose m values $z_1, \ldots, z_m$ from $\Phi_{h,\alpha}$ by choosing $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$, where each x_i is chosen from $\{0, 1, \ldots, \lceil h \rceil\}$ at random and each y_i is chosen according to Ψ_α. Let i_0 be an index such that x_{i_0} is odd. For $i \in \{1, \ldots, m\}$, let a_i be $\lfloor N z_i \rfloor$. The public key is $(a_1, \ldots, a_m, i_0)$.

Encryption: A plaintext is $\sigma \in \{0, 1\}$. Choose a random string $r = r_1 \ldots r_m \in \{0, 1\}^m$. The ciphertext is $\sigma \lfloor a_{i_0}/2 \rfloor + \sum_{i=1}^{m} r_i a_i \bmod N$.

Decryption: Let $w \in \{0, \ldots, N-1\}$ be a receiving ciphertext. We decrypt 0 if $\mathrm{frc}(w/d) < 1/4$ and 1 otherwise.

We summary the decryption errors and the security of R04 as follows.

Theorem 5.3.2. *The security of the Regev'04 cryptosystem is based on the worst case of* $O(\gamma(n)\sqrt{n})$-uSVP. *The decryption error probability is at most* $2^{-\Omega(\gamma^2(n)/n^{2c}m)}$.

We modify parameters and the key-generation algorithm as follows:

Cryptosystem 5.3.3 (pR04)**.**

Parameters: Let $c = 3$ and $t_\alpha = n^{-3.5}$. Let also $\gamma(n) = n^4 \log n$.

Private Key: Same as the original one.

Public Key: Choose $\alpha \in [2/\gamma(n), 2\sqrt{2}/\gamma(n))$ uniformly at random. We choose m values $z_1, \ldots, z_m$ from $\Phi_{h,\alpha}$ by choosing $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$. If $\mathrm{frc}(y_i) > t_\alpha$ we rechoose y_i. For $i \in \{1, \ldots, m\}$, let a_i be $\lfloor Nz_i \rfloor$. Let i_0 be an index such that x_{i_0} is odd and a_{i_0} is even. The public key is $(a_1, \ldots, a_m, i_0)$.

We refer this modified version as pR04.

Before summarizing security and correctness of pR04, we show a lemma to bound the tail of Gaussian distribution Ψ_α.

Lemma 5.3.4. *Let n be a security parameter. Let $\alpha > 0$ be a real number in $[2/\gamma(n), 2\sqrt{2}/\gamma(n))$. Let t_α be an integer that asymptotically larger than $2\sqrt{2}\log n/\gamma(n)$, i.e., $t_\alpha = \omega(\log n)/\gamma(n)$. Finally, let y be a random variable according to the distribution Ψ_α. Then, the probability that $\mathrm{frc}(y) \geq t_\alpha$ is negligible in n.*

Proof. By Lemma 2.3.1, we have that

$$\begin{aligned}
\Pr_{y\sim\Psi_\alpha}[\mathrm{frc}(y) \geq t_\alpha] &\leq \Pr_{y'\sim N(0,\alpha^2/(2\pi))}[|y'| \geq t_\alpha] \\
&\leq \sqrt{\frac{2}{\pi}}\frac{2\sqrt{2}/(\gamma(n)\sqrt{2\pi})}{t_\alpha}\exp\left(-\frac{t_\alpha^2}{2(2\sqrt{2}/(\gamma(n)\sqrt{2\pi}))^2}\right) \\
&\leq \frac{2\sqrt{2}}{\pi t_\alpha \gamma(n)}\exp\left(-\pi\frac{t_\alpha^2\gamma(n)^2}{8}\right).
\end{aligned}$$

Since we set $t_\alpha = \omega(\sqrt{\log n})/\gamma(n)$, we obtain $\exp(-\omega(\log n))$ as the upperbound of the probability. □

Let us argue the correctness of pR04.

Lemma 5.3.5 (Correctness)**.** *Let c_0 and c_1 be legal ciphertexts of 0 and 1 respectively. Then,*

$$\mathrm{frc}\left(\frac{c_0}{d}\right) \leq \frac{1}{4n^3} + mt_\alpha \leq \frac{2}{n} \text{ and } \mathrm{frc}\left(\frac{c_1}{d}\right) \geq \frac{1}{2} - \frac{1}{2n^3} + (m+1)t_\alpha \geq \frac{1}{2} - \frac{2}{n}.$$

I.e., there exist no decryption errors.

Proof. We first evaluate $\mathrm{frc}(c_0/d)$. Let $c_0 = \sum_{i=1}^m r_i a_i \bmod N$. Considering effects by modulo N at most m times, we have that

$$\left|c_0 - \left(\sum_{i=1}^m r_i a_i \bmod d\lfloor h \rceil\right)\right| \leq m|N - d\lfloor h \rceil| = md \cdot \mathrm{frc}(h) < \frac{1}{8n^3}d.$$

By the triangle inequality,

$$\begin{aligned}\mathrm{frc}\left(\frac{c_0}{d}\right) &\leq \frac{1}{8n^3} + \mathrm{frc}\left(\frac{\sum_{i=1}^{m} r_i a_i \bmod d\lfloor h \rceil}{d}\right) \\ &\leq \frac{1}{8n^3} + \mathrm{frc}\left(\frac{\sum_{i=1}^{m} a_i}{d}\right) \\ &\leq \frac{1}{8n^3} + \frac{m}{d} + \mathrm{frc}\left(\frac{N}{d}\sum_{i=1}^{m} z_i\right),\end{aligned}$$

where in the last inequality we use the fact $a_i = \lfloor N z_i \rfloor$. Since $z_i = (x_i + y_i)/h$ and $N = dh$,

$$\mathrm{frc}\left(\frac{N}{d}\sum_{i=1}^{m} z_i\right) = \mathrm{frc}\left(\sum_{i=1}^{m}(x_i + y_i)\right) = \mathrm{frc}\left(\sum_{i=1}^{m} y_i\right) \leq m t_\alpha.$$

Since d is much larger than m, $\frac{1}{8n^3} + \frac{m}{d} \leq \frac{1}{4n^3}$. Therefore, we obtain $\mathrm{frc}\,(c_0/d) \leq \frac{1}{4n^3} + m t_\alpha$.

We next evaluate $\mathrm{frc}\,(c_1/d)$. Note that for some legal ciphertext of 0 c_0, $c_1 = \lfloor a_{i_0}/2 \rfloor + c_0 \bmod N$. From the construction of a_{i_0},

$$\mathrm{frc}\left(\frac{\lfloor a_{i_0}/2 \rfloor}{d}\right) \geq \mathrm{frc}\left(\frac{a_{i_0}/2}{d}\right) - \frac{1}{d} \geq \mathrm{frc}\left(\frac{N z_{i_0}/2}{d}\right) - \frac{2}{d} \geq \mathrm{frc}\left(\frac{x_{i_0} + y_{i_0}}{2}\right) - \frac{2}{d} \geq \frac{1}{2} - \mathrm{frc}\left(\frac{y_{i_0}}{2}\right) - \frac{2}{d} \geq \frac{1}{2} - t_\alpha,$$

where in the last inequality we use the fact d is much larger than t_α. By the triangle inequality, we obtain that

$$\begin{aligned}\mathrm{frc}\left(\frac{c_1}{d}\right) &= \mathrm{frc}\left(\frac{\lfloor a_{i_0}/2 \rfloor + c_0 \bmod N}{d}\right) \\ &\geq \frac{1}{2} - t_\alpha - \left(\frac{1}{4n^3} + m t_\alpha\right) - \frac{1}{8n^3 m} \\ &\geq \frac{1}{2} - \frac{1}{2n^3} - (m+1) t_\alpha.\end{aligned}$$

□

We define the assumption IuSVP as follows:

Assumption 5.3.6 (Infeasibility of uSVP)**.** There exists no polynomial-time algorithm that solves $\tilde{O}(n^{4.5})$-uSVP with non-negligible probability.

5.3.2 Preliminaries for PPK

Let $\mathcal{E}(\mathsf{pk}, \sigma)$ be a set of legal ciphertexts of σ with a public key pk. We define a threshold of GapCVP as $t = \sqrt{m^2 + K_2^2 m}$ and an approximation factor of GapCVP as $\gamma = \sqrt{\frac{m+2}{\log(m+2)}}$.

Definition 5.3.7. Let $\mathsf{pk} = (a_1, \ldots, a_m, i_0)$ be a public key of pR04. Let c be an integer in $\{0, 1, \ldots, N-1\}$. Define a mapping $\mathcal{F}(\mathsf{pk}, c) = (\mathbf{B}_{\mathsf{pk}}, t, \mathbf{x}_c)$, where $\mathbf{x_c} = \binom{K_1 c}{\mathbf{0}} \in \mathbb{Z}^{m+2}$ and $\mathbf{B}_{\mathsf{pk}} \in \mathbb{Z}^{(m+2)\times(m+1)}$ is

$$\mathbf{B}_{\mathsf{pk}} = \begin{bmatrix} K_1 N & K_1 v_1 & \cdots & K_1 v_m \\ 1 & & & \\ & K_2 & & \\ & & \ddots & \\ & & & K_2 \end{bmatrix},$$

where $v_i = a_i$, $K_1 = n^4$, $K_2 = n^2$ and empty spaces are set by 0.

We remark that $K_1 > \gamma t$ and $\frac{1}{8n^3 m} + \frac{\sqrt{2mt_\alpha}}{K_2} \leq n^{-4}$.

5.3.3 From Ciphertexts to GapCVP (or Verifiable Encryption)

From Ciphertexts of 0 to Instances of GapCVP

We show that $\mathcal{F}(\cdot, \cdot)$ maps a valid ciphertext of 0 to a YES instance of GapCVP_γ and a ciphertext that decrypts to 1 to a NO instance of one. Hence, we have an interactive proof that c is a ciphertext of 0 using the MV protocol and this transformation.

Lemma 5.3.8.

1. *For* $(\mathsf{sk}, \mathsf{pk})$ *and* $c \in \mathcal{E}(\mathsf{pk}, 0)$, $\mathcal{F}(\mathsf{pk}, c)$ *is a YES instance of* GapCVP_γ.
2. *For any instance of* $(\mathsf{sk}, \mathsf{pk})$ *and* $c \in \{0, 1, \ldots, N-1\}$ *such that* $D(\mathsf{sk}, c) = 1$, $\mathcal{F}(\mathsf{pk}, c)$ *is a NO instance of* GapCVP_γ.

Proof. **(1)** Since $c \in \mathcal{E}(\mathsf{pk}, 0)$, there exists a string r such that $c = \sum_{i=1}^m r_i v_i \bmod N$. Thus, there exists a vector $\mathbf{w} = {}^t(\alpha_1, \beta_1, \ldots, \beta_m)$, where $\alpha_1 \in \{-m, \ldots, 0\}$ and $\beta_i \in \{0, 1\}$, such that $c = \alpha_1 N + \sum_{i=1}^m \beta_i v_i$. It is evident that $\mathbf{B}_{\mathsf{pk}}\mathbf{w} \in L_{\mathsf{pk}}$. Hence, we obtain that

$$\begin{aligned} \mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L_{\mathsf{pk}}\right) &\leq \mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, \mathbf{B}_{\mathsf{pk}}\mathbf{w}\right) \\ &= \sqrt{\alpha_1^2 + K_2^2 \sum_j^m \beta_j^2} \\ &\leq \sqrt{m^2 + K_2^2 m} = t. \end{aligned}$$

(2) Let $c \in \{0, 1, \ldots, N-1\}$ be any vector which decrypts to 1 and let $T = \gamma t$. From the remark, it follows that $T/n^4 \leq 1/4 \leq \mathrm{frc}\,(c/d)$. By Claim 5.3.9 $\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L_{\mathsf{pk}}\right) \leq T$ can not hold. Thus, $\mathcal{F}(\mathsf{pk}, c)$ is a NO instance. □

Claim 5.3.9. *Assume that $K_1 > T > 0$. Let* pk *be a public key of* pR04 *and $c \in \{0, 1, \ldots, N-1\}$. For sufficiently large n, If* $\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L_{\mathsf{pk}}\right) \leq T$ *then* $\mathrm{frc}\,(c/d) \leq T(\frac{1}{8n^3 m} + \frac{\sqrt{2mt_\alpha}}{K_2}) \leq T/n^4$.

Proof. From the assumption, there exists an integer vector $\mathbf{w} = {}^t(\alpha_1, \beta_1, \ldots, \beta_m)$ such that $\left\|\binom{K_1 c}{\mathbf{0}} - \mathbf{B}_{\mathsf{pk}}\mathbf{w}\right\| \leq T$. We define $e = K_1 c - K_1(\alpha_1 N + \sum_{i=1}^m \beta_i \mathbf{v}_i)$. From the construction of $\mathbf{B}_{\mathsf{pk}}$, we obtain that

$$\alpha_1^2 + K_2^2 \sum_{i=1}^m \beta_i^2 + e^2 \leq T^2.$$

From the fact $K_1 > T$ and $e \in K_1\mathbb{Z}$, e must be 0. Recall that $c = \alpha_1 N + \sum_{i=1}^m \beta_i \mathbf{v}_i + e/K_1$. Therefore,

$$\begin{aligned}\mathrm{frc}\,(c/d) &\leq |\alpha_1|\,\mathrm{frc}\,(N/d) + \sum_{i=1}^m |\beta_i|\,\mathrm{frc}\,(v_i/d) \\ &\leq T\,\mathrm{frc}\,(h) + \sum_{i=1}^m |\beta_i|\,(1/d + \mathrm{frc}\,(y_i))\end{aligned}$$

By the Cauchy-Schwartz inequality and the upper bound of $\sum \beta_i^2$, we have $\sum_{i=1}^m \beta_i(1/d + \mathrm{frc}\,(y_i)) \leq \sqrt{\sum_{i=1}^m \beta_i^2}\sqrt{\sum_{i=1}^m (1/d + \mathrm{frc}\,(y_i))^2} \leq \sqrt{\sum_{i=1}^m 2\mathrm{frc}\,(y_i)^2}\,T/K_2$. Moreover, from the key generation algorithm, we have $\sqrt{\sum_{i=1}^m 2\mathrm{frc}\,(y_i)^2} \leq \sqrt{2m}t_\alpha$. Hence, we obtain $\mathrm{frc}\,(c/d) \leq T(\frac{1}{8n^3 m} + \frac{\sqrt{2mt_\alpha}}{K_2})$ and complete the proof. □

Protocol 5.3.10 (Protocol$_0$: proving that a ciphertext decrypts to 0)**.** Let P_0 and V_0 denote the prover and the verifier, respectively. Let the common input be a pair (pk, c), where pk is a public key of pR04 and c is an element in $\{0, 1, \ldots, N-1\}$. The auxiliary input to the prover is $\beta_1, \ldots, \beta_m \in \{0, 1\}$ such that $c = \sum_{i=1}^m \beta_i v_i \bmod N$.

Prover P_0**:** Computes an integer α_1 such that $c = \alpha_1 N + \sum_{i=1}^m \beta_i v_i$. Invokes the prover P_{MV} to prove that input $\mathcal{F}(\mathsf{pk}, c)$ is a YES instance of GapCVP$_\gamma$ with an auxiliary input $\mathbf{B}_{\mathsf{pk}}\mathbf{w}$, where $\mathbf{w} = {}^t(\alpha_1, \beta_1, \ldots, \beta_m)$.

Verifier V_0**:** Invoke the verifier V_{MV} to verify that input $\mathcal{F}(\mathsf{pk}, c)$ is a YES instance of GapCVP$_\gamma$.

Hence we use the MV protocol, we obtain the following lemma.

Lemma 5.3.11. *Protocol (P_0, V_0) is a statistical zero-knowledge protocol.*

From Ciphertexts of 1 to Instances of GapCVP

If c is a valid ciphertext of 1 then $y := c - \lfloor v_{i_0}/2 \rfloor \bmod N$ is some valid ciphertext of 0. On the other hand, even if c be a ciphertext that decrypts to 0, there exists the case that y is *not*

a ciphertext that decrypts to 1 because $\mathrm{frc}(v_{i_0})$ is not 0 and there are effects by modulo N. However, we ensure $\mathcal{F}(\mathsf{pk}, y)$ is a NO instance of GapCVP_γ as follows.

Lemma 5.3.12. *Let* $y = c - \lfloor v_{i_0}/2 \rfloor \bmod N$.

1. *For* $(\mathsf{sk}, \mathsf{pk})$ *and* $c \in \mathcal{E}(\mathsf{pk}, 1)$, $\mathcal{F}(\mathsf{pk}, y)$ *is a YES instance of* GapCVP_γ.
2. *For any instance of* $(\mathsf{sk}, \mathsf{pk})$ *and* $c \in \{0, 1, \ldots, N-1\}$ *such that* $D(\mathsf{sk}, c) = 0$, $\mathcal{F}(\mathsf{pk}, y)$ *is a NO instance of* GapCVP_γ.

Proof. **(1)** Since c is a legal ciphertext of 1, we have y is a legal ciphertext of 0. Therefore, by Lemma 5.3.8, $\mathcal{F}(\mathsf{pk}, y)$ is a YES instance of GapCVP_γ.
(2) Let $c \in \{0, 1, \ldots, N-1\}$ be a ciphertext that decrypts into 0. By the triangle inequality,

$$\mathrm{frc}\left(\frac{c - \lfloor v_{i_0}/2 \rfloor \bmod N}{d}\right) \geq \mathrm{frc}\left(\frac{\lfloor v_{i_0}/2 \rfloor}{d}\right) - \mathrm{frc}\left(\frac{c}{d}\right) - \mathrm{frc}(h).$$

From the decryption algorithm, $\mathrm{frc}(c/d) \leq 1/4$. Therefore, we obtain

$$\mathrm{frc}\left(\frac{c - \lfloor v_{i_0}/2 \rfloor \bmod N}{d}\right) \geq \frac{1}{2} - t_\alpha - 1/4 - \frac{1}{8n^3 m} \geq \frac{1}{4} - \left(t_\alpha + \frac{1}{8n^3 m}\right).$$

Note that $\frac{\gamma t}{n^4} < \frac{1}{4} - \left(t_\alpha + \frac{1}{8n^3 m}\right)$. Thus, by Claim 5.3.9, $\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L_{\mathsf{pk}}\right) \leq \gamma t$ can not hold, and $\mathcal{F}(\mathsf{pk}, y)$ is a NO instance of GapCVP_γ. □

Protocol 5.3.13 ($\mathrm{Protocol}_1$: proving that a ciphertext decrypts to 1)**.** Let P_1 and V_1 denote the prover and the verifier, respectively. The common input is a pair (pk, c), where pk is a public key of pR04 and c is an integer in $\{0, 1, \ldots, N-1\}$. The auxiliary input to the prover is $\beta_1, \ldots, \beta_m \in \{0, 1\}$ such that $c = \lfloor v_{i_0}/2 \rfloor + \sum_{i=1}^m \beta_i v_i \bmod N$.

Prover P_1**:** Let $y = c - \lfloor v_{i_0}/2 \rfloor \bmod N$. Computes an integer α_1 such that $c = \alpha_1 N + \sum_{i=1}^m \beta_i v_i$. Invokes the prover P_{MV} to prove that input $\mathcal{F}(\mathsf{pk}, y)$ is a YES instance of GapCVP_γ with an auxiliary input $\mathbf{B}_{\mathsf{pk}}\mathbf{w}$, where $\mathbf{w} = {}^t(\alpha_1, \beta_1, \ldots, \beta_m)$.

Verifier V_1**:** Invoke the verifier V_{MV} to verify that input $\mathcal{F}(\mathsf{pk}, y)$ is a YES instance of GapCVP_γ.

Similar to the case of ciphertexts of 0, we obtain the following lemma.

Lemma 5.3.14. *Protocol* (P_1, V_1) *is a statistical zero-knowledge protocol.*

5.3.4 Definition of Relation

In this section, we consider the relation between the sum and the instance of GapCVP$_\gamma$. In the following section, we define $t' = 4t$.

Definition 5.3.15 (Relation for pR04)**.** Let $\mathsf{pk} = (a_1, \ldots, a_m, i_0)$ be a public key of pR04, c and c' elements from $\{0, 1, \ldots, N-1\}$, σ' and $\sigma'' \in \{0,1\}$, $r' \in \{0,1\}^m$, and $\mathbf{p}$ be a point from L_{pk}. We say that input (pk, c) and witness $(c', \sigma', r', \sigma'', \mathbf{p})$ are in R_{pR04} if:

- $c' = E_{\mathsf{pk}}(\sigma'; r')$
- $\mathrm{Dist}\left(\begin{pmatrix} K_1(c+c'-\sigma''\lfloor v_{i_0}/2 \rfloor \bmod N) \\ \mathbf{0} \end{pmatrix}, \mathbf{p}\right) \leq \gamma t'$ (i.e., $c + c' \bmod N$ decrypts to σ''.)

Theorem 5.3.16. *Let* $(\mathsf{pk}, \mathsf{sk})$ *be an instance of* pR04*. If* $((\mathsf{pk}, c), w) \in R_{\mathrm{pR04}}$ *for* $w = (c', \sigma', r', \sigma'', \mathbf{p})$*, then* $\sigma' \oplus \sigma'' = D(\mathsf{sk}, c)$*.*

Proof. We first consider the case $\sigma'' = 0$. In this case, we have that an inequality

$$\mathrm{Dist}\left(\begin{pmatrix} K_1(c + c' \bmod N) \\ \mathbf{0} \end{pmatrix}, \mathbf{p}\right) \leq \gamma t'.$$

Applying Claim 5.3.9, we obtain that $\mathrm{frc}\,((c + c' \bmod N)/d) \leq \gamma t'/n^4$. Suppose that $\sigma' = 0$. Since c' is a legal ciphertext, $\mathrm{frc}\,(c'/d) \leq 2/n$. It implies that $\mathrm{frc}\,(c/d) \leq \gamma t'/n^4 + 2/n + 1/8n^3m \leq 1/4$ and $D(\mathsf{sk}, c) = 0$. We also suppose that $\sigma' = 1$. Since c' is a legal ciphertexts, $\mathrm{frc}\,(c'/d) \geq 1/2 - 2/n$. Therefore, by triangle inequality $\mathrm{frc}\,(c/d) \geq 1/2 - 2/n - \gamma t'/n^4 - 1/8n^3m \geq 1/4$ and $D(\mathsf{sk}, c) = 1$.

Next, we consider the case $\sigma'' = 1$, i.e.,

$$\mathrm{Dist}\left(\begin{pmatrix} K_1(c + c' - \lfloor v_{i_0}/2 \rfloor \bmod N) \\ \mathbf{0} \end{pmatrix}, \mathbf{p}\right) \leq \gamma t'.$$

Applying Claim 5.3.9, we obtain that $\mathrm{frc}\,((c + c' - \lfloor v_{i_0}/2 \rfloor \bmod N)/d) \leq \gamma t'/n^4$. It implies that $\mathrm{frc}\,((c + c' \bmod N)/d) \geq 1/2 - (\mathrm{frc}\,(h) + 2t_\alpha) - \gamma t'/n^4 \geq 1/2 - 2/n$. Suppose that $\sigma' = 0$. Since $\mathbf{c}'$ is a legal ciphertext, $\mathrm{frc}\,(c'/d) \leq 2/n$. It implies that $\mathrm{frc}\,(c/d) \geq 1/2 - 2/n - 2/n - 1/8n^3m \geq 1/4$ and $D(\mathsf{sk}, \mathbf{c}) = 1$. Next, we suppose that $\sigma' = 1$. Since c' is a legal ciphertext, we have that $\mathrm{frc}\,(c'/d) \geq 1/2 - (2\mathrm{frc}\,(h) + 2mt_\alpha) \geq 1/2 - 2/n$. It implies that $\mathrm{frc}\,(c/d) \leq 2/n + 2/n + 1/8n^3m \leq 1/4$ and $D(\mathsf{sk}, \mathbf{c}) = 0$. We complete the proof. □

5.3.5 Main Protocol

Protocol 5.3.17 (Protocol PPK)**.** Let P and V denote a prover and a verifier, respectively. A common input is (pk, c). An auxiliary input to the prover is (σ, r) such that $c = E_{\mathsf{pk}}(\sigma; r)$.

Define a mapping $\mathcal{G}(\mathsf{pk}, c) = (\mathbf{B}_{\mathsf{pk}}, \mathbf{x}_c, t')$ where $t' = 2t$ and $\mathbf{B}_{\mathsf{pk}}$ and $\mathbf{x}_c$ are similar to $\mathcal{F}(\mathsf{pk}, c)$. Let $\text{Protocol}'_0$ (or $\text{Protocol}'_1$) be Protocol_0 (or Protocol_1) where $\mathcal{F}(\cdot,\cdot)$ is replaced by $\mathcal{G}(\cdot,\cdot)$ respectively.

Step P1 P selects $\sigma' \in \{0,1\}$ and $r' \in \{0,1\}^m$ randomly. Computes $c' = E_{\mathsf{pk}}(\sigma'; r')$ and sends c' to V.

Step V1 V sends a random challenge bit $\delta \in \{0,1\}$ to P.

Step P2 If $\delta = 0$, P sends pair (σ', r'). If $\delta = 1$, P computes $\sigma'' = \sigma + \sigma' \bmod 2$ and sends σ'' to V. Let $\overline{c} = (c + c') \bmod N$ and runs $\text{Protocol}'_{\sigma''}$ on input $(\mathsf{pk}, \overline{c})$ as prover.

Step V2 If $\delta = 0$. V accepts if $c' = E_{\mathsf{pk}}(\sigma'; r')$, else rejects. If $\delta = 1$. Run the $\text{Protocol}'_{\sigma''}$ on input $(\mathsf{pk}, \overline{c})$ as verifier.

Theorem 5.3.18 (Regev 04 PPK)**.** *An interactive protocol (P, V) is a proof of knowledge system with knowledge error $3/4$ for R_{pR04}. Moreover, the protocol (P, V) is a computational zero knowledge under the assumption IuSVP.*

The proofs of following Lemma 5.3.19 and Lemma 5.3.20 are in Section 5.3.6. We need the lemmas for larger protocol PPK.

Lemma 5.3.19. *For sufficiently large n,*

1. *If $(\mathsf{sk}, \mathsf{pk})$ is an instance of* pR04*, $c = c_1 + c_2 \bmod N$ such that $D(\mathsf{sk}, c) = 0$ and $c_1, c_2 \in \mathcal{E}(\mathsf{pk}, \cdot)$, $\mathcal{G}(\mathsf{pk}, c)$ is a YES instance of* GapCVP_γ.
2. *Let $(\mathsf{sk}, \mathsf{pk})$ be an instance of* pR04 *and $c \in \{0, 1, \ldots, N-1\}$. If $\text{frc}(c/d) > 1/8$, then $\mathcal{G}(\mathsf{pk}, c)$ is a NO instance of* GapCVP_γ.

Lemma 5.3.20. *For sufficiently large n,*

1. *If $(\mathsf{sk}, \mathsf{pk})$ is an instance of* pR04*, $c = c_1 + c_2 \bmod N$ such that $D(\mathsf{sk}, c) = 1$ and $c_1, c_2 \in \mathcal{E}(\mathsf{pk}, \cdot)$, $\mathcal{G}(\mathsf{pk}, y)$ is a YES instance of* GapCVP_γ*, where $y = c - \lfloor v_{i_0}/2 \rfloor \bmod N$.*
2. *Let $(\mathsf{sk}, \mathsf{pk})$ be an instance of* pR04 *and $c \in \{0, 1, \ldots, N-1\}$. If $\text{frc}(c/d) < 3/8$, then $\mathcal{G}(\mathsf{pk}, y)$ is a NO instance of* GapCVP_γ*, where $y = c - \lfloor v_{i_0}/2 \rfloor \bmod N$.*

Proof of Completeness. Since it is evident, we omit the proof. □

Proof of Validity with error $3/4$. Let $\mathsf{pk} = (a_1, \ldots, a_m, i_0)$ be a public key of pR04. and $c \in \{0, 1, \ldots, N-1\}$. Let P^* be an arbitrary prover that make V accept with probability $\epsilon + 3/4$ for $\epsilon > 0$ on common input (pk, c).

We construct a knowledge extractor K as follows. K's input is (pk, c). First, K choose a random tape of P^*. Let δ_1 denotes a challenge bit in $\text{Protocol}'_{\sigma''}$. K runs P^* three times, where

the challenge bit are 0, $(1,0)$ and $(1,1)$. K obtains three views T_0, T_1, and T_2. Views are in forms that $T_0 = (c', 0, \sigma', r')$, $T_1 = (c', 1, \sigma'', T_1')$, and $T_2 = (c', 1, \sigma'', T_2')$, where T_1' and T_2' are transcripts of Protocol$'_{\sigma''}$ that δ_1 are 0 and 1 respectively. If any one of three views is rejected, K outputs $\perp$ and halts. Otherwise, i.e., three views are accepted, K obtains a vector $\mathbf{p}$ that is witness of GapCVP$_\gamma$ using the extractor in Protocol$'_0$ or Protocol$'_1$. Outputs $(c', \sigma', r', \sigma'', \mathbf{p})$ and halts.

Note that the probability K does not output $\perp$ is at least ϵ. Therefore, K is indeed the knowledge extractor. □

Proof of Zero-knowledge of PPK. We construct a simulator S as follows: Let S_σ is a simulator for Protocol$'_\sigma$.

Step P1 Chooses $\Delta \in \{0,1\}$ randomly (Predictor of a challenge bit). If $\Delta = 0$, chooses σ', r' randomly and computes $c' = E_{\mathsf{pk}}(\sigma'; r')$. If $\Delta = 1$, chooses σ'', r'' randomly, computes $\bar{c} = E_{\mathsf{pk}}(\sigma''; r'')$, and sets $c' = \bar{c} - c \bmod N$. Sends c' to V^*.

Step V1 Receives a challenge bit δ from V^*.

Step P2, V2 If $\Delta \neq \delta$, outputs $\perp$ and halts. If $\Delta = \delta = 0$ outputs $(c', \delta, \sigma', r')$. If $\Delta = \delta = 1$, invoke $S_{\sigma''}$ with input $(\mathsf{pk}, \bar{c})$. Let $T = S_{\sigma''}(\mathsf{pk}, \bar{c})$. Outputs $(c', \delta, \sigma'', T)$ and halts.

We assume that ISVP holds, hence according to the security property of pR04 if $\Delta = 0$ then c' is computationally indistinguishable from the uniform distribution on $\{0, 1, \ldots, N-1\}$; if $\Delta = 0$ then $c' = \bar{c} - c \bmod N$ is also indistinguishable from the uniform distribution. Therefore, the generated transcripts is computationally indistinguishable from a real transcript. □

5.3.6 Proof of Lemmas

Proof of Lemma 5.3.19. **(1)** There are two cases that c can decrypts into 0: when both c_1 and c_2 are ciphertexts of 0 and when both are ciphertexts of 1.

Suppose that $c_1, c_2 \in \mathcal{E}(\mathsf{pk}, 0)$. From Lemma 5.3.8, $\mathrm{Dist}\left(\binom{K_1 c_i}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq t$ for $i = 1, 2$. By Lemma 5.3.21 below, Thus, for $c = c_1 + c_2 \bmod N$, we have that

$$\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + 1 \leq 4t = t'.$$

Next, suppose that $c_1, c_2 \in \mathcal{E}(\mathsf{pk}, 1)$. Thus, for $i = 1, 2$, $\bar{c}_i = c_i - v_{i_0}/2 \bmod N \in \mathcal{E}(\mathsf{pk}, 0)$. By Lemma 5.3.21 below, we have that for $\bar{c} = \bar{c}_1 + \bar{c}_2 \bmod N$, $\mathrm{Dist}\left(\binom{K_1 \bar{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + 1$. Consider the vector $c = \bar{c} + v_{i_0} \bmod N$. By Lemma 5.3.22, we have that

$$\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + 1 + \sqrt{K_2^2 + 1} \leq 4t = t'.$$

(2) Let $c \in \{0, 1, \ldots, N-1\}$ be any ciphertext such that $\mathrm{frc}(c/d) > 1/8$. Let $T = \gamma t'$. Note that $T/n^4 \leq 1/8 < \mathrm{frc}(c/d)$. Hence, by Claim 5.3.9 $\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq T$ can not hold. Thus, $\mathcal{G}(\mathsf{pk}, c)$ is a NO instance of the GapCVP$_\gamma$. □

Proof of Lemma 5.3.20. **(1)** Without a loss of generality, we suppose that $c_1 \in \mathcal{E}(\mathsf{pk}, 0)$ and $c_2 \in \mathcal{E}(\mathsf{pk}, 1)$. Since c_1 is a legal ciphertext of 0, from Lemma 5.3.8, for some $\mathbf{p}_1 \in L(\mathbf{B}_{\mathsf{pk}})$, $\mathrm{Dist}\left(\binom{K_1 c_1}{\mathbf{0}}, \mathbf{p}_1\right) \leq t$. Since c_1 is a legal ciphertext of 1, from Lemma 5.3.12, for some $\mathbf{p}_2 \in L(\mathbf{B}_{\mathsf{pk}})$, $\mathrm{Dist}\left(\binom{K_1(c_2 - v_{i_0}/2 \bmod N)}{\mathbf{0}}, \mathbf{p}_2\right) \leq t$. Hence, from $y = c_1 + c_2 - v_{i_0}/2 \bmod N$, we obtain

$$\mathrm{Dist}\left(\binom{K_1 y}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + 1 \leq 4t = t'$$

by Lemma 5.3.21.

(2) Let $c \in \{0, 1, \ldots, N-1\}$ be any ciphertext such that $\mathrm{frc}(c/d) < 3/8$. In this case, we obtain that $\mathrm{frc}(y/d) > 1/4$ in a similar way to the proof of Lemma 5.3.12. Let $T = \gamma t'$. Note that $T/n^4 \leq 1/4 < \mathrm{frc}(y/d)$. Hence, by Claim 5.3.9 $\mathrm{Dist}\left(\binom{K_1 y}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq T$ can not hold. Thus, $\mathcal{G}(\mathsf{pk}, y)$ is a NO instance of the GapCVP$_\gamma$. □

Lemma 5.3.21. *Let* pk *be a public key of* pR04, $\mathbf{p}_1$ *and* $\mathbf{p}_2$ *points from* $L(\mathbf{B}_{\mathsf{pk}})$. *If for* $c_1, c_2 \in \{0, 1, \ldots, N-1\}$, $\mathrm{Dist}\left(\binom{K_1 c_1}{\mathbf{0}}, \mathbf{p}_1\right) \leq d_1$ *and* $\mathrm{Dist}\left(\binom{K_1 c_2}{\mathbf{0}}, \mathbf{p}_2\right) \leq d_2$, *then* $\mathrm{Dist}\left(\binom{K_1(c_1 + c_2 \bmod N)}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq d_1 + d_2 + 1$.

Proof. Represent $K_1(c_1 + c_2 \bmod N) = K_1(c_1 + c_2 + \alpha_1 N)$. Since both vectors c_1 and c_2 belong to $\{0, 1, \ldots, N-1\}$, we can bound $|\alpha_1| \leq 1$. Consider a vector $\mathbf{p} = \mathbf{B}_{\mathsf{pk}}{}^t(\alpha_1, 0, \ldots, 0)$. Thus, we obtain that

$$\mathrm{Dist}\left(\binom{K_1 \alpha N}{\mathbf{0}}, \mathbf{p}\right) \leq 1.$$

By the triangle inequality, the lemma follows. □

Lemma 5.3.22. *Let* pk *be a public key of* pR04 *and* $\mathbf{p}$ *a point from* $L(\mathbf{B}_{\mathsf{pk}})$. *If for* $c \in \{0, 1, \ldots, N-1\}$, $\mathrm{Dist}\left(\binom{K_1 c}{\mathbf{0}}, \mathbf{p}\right) = d$ *then* $\mathrm{Dist}\left(\binom{K_1(c + v_{i_0} \bmod N)}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq d + \sqrt{K_2^2 + 1}$.

Proof. Represent $K_1(c + v_{i_0} \bmod N) = K_1(c + v_{i_0} + \alpha_1 N)$ for some $\alpha_1 \in \{-1, 0\}$. Consider a vector $\mathbf{p}'$ in $L(\mathbf{B}_{\mathsf{pk}})$ such that $\mathbf{p}' = L(\mathbf{B}_{\mathsf{pk}})^t(0, \ldots, 0, 1, 0, \ldots, 0)$ (with 1 at the (i_0+1)-th position). By the construction of $\mathbf{B}_{\mathsf{pk}}$, we have that $\mathrm{Dist}\left(\binom{K_1(v_{i_0} + \alpha_1 N)}{\mathbf{0}}, \mathbf{p}'\right) \leq \sqrt{K_2^2 + 1}$. By the triangle inequality, the lemma follows. □

5.4 A Proof of Plaintext Knowledge for the Regev'05 Cryptosystem

5.4.1 The Regev'05 Cryptosystem

Although we review the Regev'05 cryptosystem in Section 3.4.1, we briefly review the Regev'05 cryptosystem [Reg05] again.

Cryptosystem 5.4.1 (R05)**.** Let n be a security parameter (or a dimension of underlying lattice problems). Let q be a prime and $\alpha \in (0,1)$ a real such that $\alpha q > 2\sqrt{n}$. Let m be an integer larger than $5(n+1)\log q$.

Private Key: Choose $\mathbf{s} \in \mathbb{Z}_q^n$ uniformly at random. A private key is $\mathbf{s}$.

Public Key: Choose m vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{Z}_q^n$ independently at random. Choose $e_1, \ldots, e_m \in \mathbb{Z}_q$ independently according to $\bar{\Psi}_\alpha$. Compute $b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i \bmod q$. A public key is $\{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$.

Encryption: Choose a random string $r \in \{0,1\}^m$. Let $\sigma \in \{0,1\}$ be a plaintext. A ciphertext is $(\sum_{i=1}^m r_i \mathbf{a}_i \bmod q, \sigma \lfloor q/2 \rfloor + \sum_{i=1}^m r_i b_i \bmod q)$.

Decryption: Let $(\mathbf{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ be a received ciphertext. If $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q \leq q/4$ then decrypt into 0, otherwise into 1.

Regev recommended $q \in (n^2, 2n^2)$ and $\alpha = o(1/\sqrt{n}\log n)$ to tighten the approximation factor of underlying lattice problems.

Theorem 5.4.2 ([Reg05])**.** *The security of the Regev'05 cryptosystem is based on the worst case of* $\mathrm{SVP}_{\tilde{O}(n/\alpha(n))}$ *and* $\mathrm{SIVP}_{\tilde{O}(n/\alpha(n))}$ *for polynomial-time quantum algorithms. The decryption error probability is at most* $2^{-\Omega\left(1/(m\alpha^2(n))\right)} + 2^{-\Omega(n)}$.

We modify the key generation algorithm and parameters as follows:

Cryptosystem 5.4.3 (pR05)**.**

Parameter: Let $q = \Theta(n^4)$ be a prime and $m = 5(n+1)(\log q + 1)$. We also define $\alpha = 1/m^2$. Note that $q\alpha = \Theta(n^2/\log^2 n) > 2\sqrt{n}$ for sufficiently large n. Let $t_\alpha = n^2 \log n$. Note that $t_\alpha = \omega(q\alpha\sqrt{\log n})$.

Private Key: Same as the original one.

Public Key: Choose m vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{Z}_q^n$ independently at random. Choose m elements $e_1, \ldots, e_m \in \mathbb{Z}_q$ independently according to $\bar{\Psi}_\alpha$. If $|e_i|_q \leq t_\alpha$ for all i then compute $b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i \bmod q$, else re-choose $e_1, \ldots, e_m$. A public key is $\{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$.

We refer this modified version as pR05. Note that the probability that there exists i such that $|e_i|_q > t_\alpha$ is negligible in n from the following Lemma 5.4.4. We also note that there exist no decryption errors in pR05.

Lemma 5.4.4. *Let n be a security parameter. Let q be a prime and $\alpha > 0$ a real number such that $q\alpha > 2\sqrt{n}$. Let t_α be an integer that asymptotically larger than $q\alpha\sqrt{\log n}$, i.e., $t_\alpha = \omega(q\alpha\sqrt{\log n})$. Finally, let e be a random variable according to the distribution $\bar{\Psi}_\alpha$. Then, the probability that $|e|_q \geq t_\alpha$ is negligible in n.*

Proof. By Lemma 2.3.1, we have that

$$\begin{aligned}\Pr_{e\sim\bar{\Psi}_\alpha}[|e|_q \geq t_\alpha] &\leq \Pr_{e'\sim\Psi_\alpha}[|e'| \geq (t_\alpha - 1)/q] \\ &\leq \sqrt{\frac{2}{\pi}}\frac{\alpha/\sqrt{2\pi}}{(t_\alpha-1)/q}\exp\left(-\frac{(t_\alpha-1)^2/q^2}{2(\alpha/\sqrt{2\pi})^2}\right) \\ &\leq \frac{q\alpha}{\pi(t_\alpha-1)}\exp\left(-\pi\frac{(t_\alpha-1)^2}{q^2\alpha^2}\right).\end{aligned}$$

Since we set $t_\alpha = \omega(q\alpha\sqrt{\log n})$, we obtain $\exp(-\omega(\log n))$ as the upperbound of the probability. □

The security follows from Theorem 5.4.2. We summarize the property of pR05 as follows.

Theorem 5.4.5. *The security of* pR05 *is based on the worst case of* $\mathrm{SVP}_{\tilde{O}(n^3)}$ *and* $\mathrm{SIVP}_{\tilde{O}(n^3)}$ *for polynomial-time quantum algorithms. There exist no decryption errors.*

We define the assumption ISVP as follows:

Assumption 5.4.6 (Infeasibility of SVP)**.** There exists no quantum polynomial-time algorithm that solves $\mathrm{SVP}_{\tilde{O}(n^3)}$ and $\mathrm{SIVP}_{\tilde{O}(n^3)}$ with non-negligible probability.

5.4.2 Preliminaries for PPK

Let $\mathcal{E}(\mathsf{pk}, \sigma)$ be a set of legal ciphertexts of σ with a public key pk. We define a threshold of GapCVP as $t = \sqrt{(n+1)m^2 + K_2^2 m}$ and an approximation factor of GapCVP as $\gamma = \sqrt{\frac{2n+m+3}{\log(2n+m+3)}}$.

Definition 5.4.7. Let $\mathsf{pk} = \{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$ be a public key of pR05. Let $\mathbf{c}$ be a vector in $\mathbb{Z}_q^{n+1}$.

Define a mapping $\mathcal{F}(\mathsf{pk}, \mathbf{c}) = (\mathbf{B}_{\mathsf{pk}}, t, \mathbf{x}_{\mathbf{c}})$, where $\mathbf{x}_{\mathbf{c}} = \binom{K_1\mathbf{c}}{\mathbf{0}} \in \mathbb{Z}^{2n+m+3}$. $\mathbf{B}_{\mathsf{pk}} \in \mathbb{Z}^{(2n+m+3)\times(n+m+2)}$ is

$$\mathbf{B}_{\mathsf{pk}} = \begin{bmatrix} K_1 q\mathbf{I}_{n+1} & K_1(q-1)\mathbf{u}_{n+1} & K_1\mathbf{v}_1 & \dots & K_1\mathbf{v}_m \\ \mathbf{I}_{n+1} & & & & \\ & 1 & & & \\ & & K_2 & & \\ & & & \ddots & \\ & & & & K_2 \end{bmatrix},$$

where $\mathbf{v}_i = \binom{\mathbf{a}_i}{b_i} \in \mathbb{Z}_q^{n+1}$, $K_1 = n^4$, and $K_2 = n^2$.

From the definitions of t and γ, we have that $\gamma t = O(n^2 m)$. We remark that, for sufficiently large n, $4\gamma t = O(n^2 m) < K_1$ and $4\gamma t(1 + \sqrt{m}t_\alpha/K_2) = O(n^2 m)O(1 + \sqrt{m}\log n) < O(n^4) = q/8$ from the definitions of K_1, K_2, q, and t_α.

5.4.3 From Ciphertexts to Instances of GapCVP (or Verifiable Encryption)

From Ciphertexts of 0 to Instances of GapCVP

We show that $\mathcal{F}(\cdot,\cdot)$ maps a valid ciphertext of 0 to a YES instance of GapCVP_γ and a ciphertext that decrypts into 1 to a NO instance of one. Hence, we have an interactive proof that $\mathbf{c}$ is a ciphertext of 0 using the MV protocol and the transformation $\mathcal{F}(\cdot,\cdot)$.

Lemma 5.4.8.

1. *For* $(\mathsf{sk}, \mathsf{pk})$ *and* $\mathbf{c} \in \mathcal{E}(\mathsf{pk}, 0)$, $\mathcal{F}(\mathsf{pk}, \mathbf{c})$ *is a YES instance of* GapCVP_γ.
2. *For any instance of* $(\mathsf{sk}, \mathsf{pk})$ *and* $\mathbf{c} \in \mathbb{Z}_q^{n+1}$ *such that* $D(\mathsf{sk}, \mathbf{c}) = 1$, $\mathcal{F}(\mathsf{pk}, \mathbf{c})$ *is a NO instance of* GapCVP_γ.

Proof. **(1)** Since $\mathbf{c} \in \mathcal{E}(\mathsf{pk}, 0)$, there exists a string $r \in \{0,1\}^m$ such that $\mathbf{c} = \sum_{i=1}^m r_i\mathbf{v}_i \bmod q$. Thus, there exists an integer vector $\mathbf{w} = {}^t(\alpha_1, \dots, \alpha_{n+1}, 0, \beta_1, \dots, \beta_m)$, where $\alpha_i \in \{-m, \dots, 0\}$ and $\beta_i \in \{0,1\}$, such that $\mathbf{c} = \sum_{i=1}^{n+1} \alpha_i q\mathbf{u}_i + \sum_{j=1}^m \beta_j\mathbf{v}_j$. It is evident that $\mathbf{B}_{\mathsf{pk}}\mathbf{w} \in L(\mathbf{B}_{\mathsf{pk}})$. Hence, we obtain that

$$\begin{aligned} \mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) &\le \mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, \mathbf{B}_{\mathsf{pk}}\mathbf{w}\right) \\ &= \sqrt{\sum_i^{n+1} \alpha_i^2 + K_2^2 \sum_j^m \beta_j^2} \\ &\le \sqrt{(n+1)m^2 + K_2^2 m} = t. \end{aligned}$$

(2) Let $\mathbf{c} = \binom{\mathbf{a}}{b} \in \mathbb{Z}_q^{n+1}$ be any vector which decrypts into 1. Let $T = \gamma t$. From the remark, it follows that $T(1 + \sqrt{m}t_\alpha/K_2) \leq q/4 \leq |b - \langle \mathbf{a}, \mathbf{s}\rangle|_q$. By Claim 5.4.9 $\mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq T$ can not hold. Thus, $\mathcal{F}(\mathsf{pk}, \mathbf{c})$ is a NO instance. □

Claim 5.4.9. *Let $K_1 > T > 0$. Let* pk *be a public key of* pR05 *and $\mathbf{c} \in \mathbb{Z}_q^{n+1}$. For sufficiently large n, if* $\mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq T$ *then* $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq T(1 + \sqrt{m}t_\alpha/K_2)$.

Proof. From the assumption, there exists an integer vector $\mathbf{w} = {}^t(\alpha_1, \ldots, \alpha_{n+2}, \beta_1, \ldots, \beta_m)$ such that $\left\|\binom{K_1\mathbf{c}}{\mathbf{0}} - \mathbf{B}_{\mathsf{pk}}\mathbf{w}\right\| \leq T$. We define $\mathbf{e} = K_1\mathbf{c} - K_1(q\sum_i^{n+1}\alpha_i\mathbf{u}_i + (q-1)\alpha_{n+2}\mathbf{u}_{n+1} + \sum_{j=1}^m \beta_i\mathbf{v}_i)$. From the construction of $\mathbf{B}_{\mathsf{pk}}$, we obtain that

$$\sum_{i=1}^{n+2}\alpha_i^2 + K_2^2\sum_{j=1}^{m}\beta_i^2 + \|\mathbf{e}\|^2 \leq T^2.$$

From the fact $K_1 > T$ and $\mathbf{e} \in K_1\mathbb{Z}^{n+1}$, $\mathbf{e}$ must be $\mathbf{0}$. We note that $\alpha_{n+2}^2 \leq T^2$. Now, recall that $\mathbf{c} = \sum_{i=1}^{n+1}\alpha_i q\mathbf{u}_i + (q-1)\alpha_{n+2}\mathbf{u}_{n+1} + \sum_{j=1}^m \beta_i\mathbf{v}_i + \mathbf{e}/K_1$. Therefore,

$$b - \langle \mathbf{a}, \mathbf{s}\rangle \equiv (q-1)\alpha_{n+2} + \sum_{i=1}^{m}\beta_i b_i - \sum_{i=1}^{m}\beta_i\langle \mathbf{a}_i, \mathbf{s}\rangle \equiv -\alpha_{n+2} + \sum_{i=1}^{m}\beta_i e_i \pmod q.$$

By the Cauchy-Schwartz inequality and the upper bound of $\sum\beta_i^2$, we have $\left|\sum_{i=1}^m \beta_i e_i\right|_q \leq \sqrt{\sum_{i=1}^m\beta_i^2}\sqrt{\sum_{i=1}^m |e_i|_q^2} \leq \sqrt{\sum_{i=1}^m |e_i|_q^2}T/K_2$. Moreover, from the key generation algorithm, we have $\sqrt{\sum_{i=1}^m |e_i|_q^2} \leq \sqrt{m}t_\alpha$. Hence, by triangle inequality, we obtain $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq T(1 + \sqrt{m}t_\alpha/K_2)$ and complete the proof. □

Protocol 5.4.10 (Protocol$_0$: proving that a ciphertext decrypts into 0)**.** P_0 and V_0 denote the prover and the verifier, respectively. The common input is a pair $(\mathsf{pk}, \mathbf{c})$, where pk is a public key of pR05 and $\mathbf{c}$ is a vector in $\mathbb{Z}_q^{n+1}$. The prover's auxiliary input is $\beta_1, \ldots, \beta_m \in \{0, 1\}$ such that $\mathbf{c} = \sum_{i=1}^m \beta_i\mathbf{v}_i \bmod q$.

Prover P_0**:** Compute integers $\alpha_1, \ldots, \alpha_{n+1}$ such that $\mathbf{c} = \sum_{i=1}^m \beta_i\mathbf{v}_i + \sum_{j=1}^{n+1} q\alpha_i\mathbf{u}_i$. Invoke the prover P_{MV} to prove that the input $\mathcal{F}(\mathsf{pk}, \mathbf{c})$ is a YES instance of GapCVP_γ with an auxiliary input $\mathbf{B}_{\mathsf{pk}}\mathbf{w}$, where $\mathbf{w} = {}^t(\alpha_1, \ldots, \alpha_{n+1}, 0, \beta_1, \ldots, \beta_m)$.

Verifier V_0**:** Invoke the verifier V_{MV} to verify that the input $\mathcal{F}(\mathsf{pk}, \mathbf{c})$ is a YES instance of GapCVP_γ.

Hence we use the MV protocol, we obtain the lemma as follows.

Lemma 5.4.11. *The protocol (P_0, V_0) is a statistical zero-knowledge protocol.*

From Ciphertexts of 1 to Instances of GapCVP

Lemma 5.4.12. *Let* $\mathbf{y} = \mathbf{c} - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$.

1. *For* $(\mathsf{sk}, \mathsf{pk})$ *and* $\mathbf{c} \in \mathcal{E}(\mathsf{pk}, 1)$, $\mathcal{F}(\mathsf{pk}, \mathbf{y})$ *is a YES instance of* GapCVP_γ.
2. *For any instance of* $(\mathsf{sk}, \mathsf{pk})$ *and* $\mathbf{c} \in \mathbb{Z}_q^{n+1}$ *such that* $D(\mathsf{sk}, \mathbf{c}) = 0$, $\mathcal{F}(\mathsf{pk}, \mathbf{y})$ *is a NO instance of* GapCVP_γ.

Proof. **(1)** Since $\mathbf{c}$ is a legal ciphertext of 1, $\mathbf{y}$ is a legal ciphertext of 0. The proof is similar to that of Lemma 5.4.8.
(2) We consider $\mathbf{y} = \mathbf{c} - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$. In this case, $D(\mathsf{sk}, \mathbf{y}) = 1$. Therefore, we prove in a similar way to the proof of Lemma 5.4.8. □

Protocol 5.4.13 ($\mathrm{Protocol}_1$: proving that a ciphertext decrypts into 1)**.** P_1 and V_1 denote the prover and the verifier, respectively. The common input is a pair $(\mathsf{pk}, \mathbf{c})$, where pk is a public key of pR05 and $\mathbf{c}$ is a vector from $\mathbb{Z}_q^{n+1}$. The prover's auxiliary input is $\beta_1, \ldots, \beta_m \in \{0, 1\}$ such that $\mathbf{c} = \sum_{i=1}^{m} \beta_i \mathbf{v}_i \bmod q$.

Prover P_1**:** Let $\mathbf{y} = \mathbf{c} - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$. Compute integers $\alpha_1, \ldots, \alpha_{n+1}$ such that $\mathbf{c} = \lfloor q/2 \rfloor \mathbf{u}_{n+1} + \sum_{i=1}^{m} \beta_i \mathbf{v}_i + \sum_{j=1}^{n+1} q\alpha_i \mathbf{u}_i$. Invoke the prover P_{MV} to prove that input $\mathcal{F}(\mathsf{pk}, \mathbf{y})$ is a YES instance of GapCVP_γ with an auxiliary input $\mathbf{B}_{\mathsf{pk}}\mathbf{w}$, where $\mathbf{w} = {}^t(\alpha_1, \ldots, \alpha_{n+1}, 0, \beta_1, \ldots, \beta_m)$.

Verifier V_1**:** Invoke the verifier V_{MV} to verify that input $\mathcal{F}(\mathsf{pk}, \mathbf{y})$ is a YES instance of GapCVP_γ.

We obtain the following lemma in a similar way to the case of ciphertexts of 0.

Lemma 5.4.14. *The protocol* (P_1, V_1) *is a statistical zero-knowledge protocol.*

5.4.4 Definition of Relation

We define $t' = 4t$.

Definition 5.4.15 (Relation for pR05)**.** Let $\mathsf{pk} = \{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$ be a public key of pR05. Let $\mathbf{c}$ and $\mathbf{c}'$ be vectors from $\mathbb{Z}_q^{n+1}$. Let σ' and σ'' be bits, r' an m-bit string, and $\mathbf{p}$ a vector in $L(\mathbf{B}_{\mathsf{pk}})$. We say that input $(\mathsf{pk}, \mathbf{c})$ and witness $(\mathbf{c}', \sigma', r', \sigma'', \mathbf{p})$ are in R_{pR05} if:

- $\mathbf{c}' = E_{\mathsf{pk}}(\sigma'; r')$ and
- $\mathrm{Dist}\left(\begin{pmatrix} K_1(\mathbf{c}+\mathbf{c}'-\sigma''\lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q) \\ \mathbf{0} \end{pmatrix}, \mathbf{p}\right) \leq \gamma t'$ (i.e., $\mathbf{c} + \mathbf{c}' \bmod q$ decrypts into σ''.)

Theorem 5.4.16. *Let* $(\mathsf{pk}, \mathsf{sk})$ *be an instance of* pR05. *If* $((\mathsf{pk}, \mathbf{c}), w) \in R_{\mathrm{pR05}}$ *for* $w = (\mathbf{c}', \sigma', r', \sigma'', \mathbf{p})$, *then* $\sigma' \oplus \sigma'' = D(\mathsf{sk}, \mathbf{c})$.

Proof. Let $\mathsf{pk} = \{(\mathbf{a}_i, b_i)\}_{i=1,\ldots,m}$ be a public key of pR05.

We first consider the case $\sigma'' = 0$. In this case, we have that an inequality

$$\mathrm{Dist}\left(\begin{pmatrix} K_1(\mathbf{c} + \mathbf{c}' \bmod q) \\ \mathbf{0} \end{pmatrix}, \mathbf{p}\right) \leq \gamma t'.$$

Applying Claim 5.4.9, we obtain that $|b + b' - \langle \mathbf{a} + \mathbf{a}', \mathbf{s}\rangle|_q \leq \gamma t'(1 + \sqrt{m} t_\alpha / K_2)$. Suppose that $\sigma' = 0$. Since $\mathbf{c}'$ is a legal ciphertext, $|b' - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq m t_\alpha$. It implies that $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq m t_\alpha + \gamma t'(1 + \sqrt{m} t_\alpha / K_2) \leq q/4$ and $D(\mathsf{sk}, \mathbf{c}) = 0$. We also suppose that $\sigma' = 1$. Since $\mathbf{c}'$ is a legal ciphertext, $|b' - \langle \mathbf{a}, \mathbf{s}\rangle|_q \geq q/2 - m t_\alpha$. It implies that $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \geq q/2 - m t_\alpha - \gamma t'(1 + \sqrt{m} t_\alpha / K_2) \geq q/4$ and $D(\mathsf{sk}, \mathbf{c}) = 1$.

Next, we consider the case $\sigma'' = 1$, i.e.,

$$\mathrm{Dist}\left(\begin{pmatrix} K_1(\mathbf{c} + \mathbf{c}' - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q) \\ \mathbf{0} \end{pmatrix}, \mathbf{p}\right) \leq \gamma t'.$$

Applying Claim 5.4.9, we obtain that $|b + b' - \lfloor q/2 \rfloor - \langle \mathbf{a} + \mathbf{a}', \mathbf{s}\rangle|_q \leq \gamma t'(1 + \sqrt{m} t_\alpha / K_2)$. Hence we have $|b + b' - \langle \mathbf{a} + \mathbf{a}', \mathbf{s}\rangle|_q \geq q/2 - \gamma t'(1 + \sqrt{m} t_\alpha / K_2)$. Suppose that $\sigma' = 0$. Since $\mathbf{c}'$ is a legal ciphertext, $|b' - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq m t_\alpha$. It implies that $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \geq q/2 - m t_\alpha - \gamma t'(1 + \sqrt{m} t_\alpha / K_2) \geq q/4$ and $D(\mathsf{sk}, \mathbf{c}) = 1$. Next, we suppose that $\sigma' = 1$. Since $\mathbf{c}'$ is a legal ciphertext, we have that $|b' - \langle \mathbf{a}', \mathbf{s}\rangle|_q \geq q/2 - m t_\alpha$. It implies that $|b - \langle \mathbf{a}, \mathbf{s}\rangle|_q \leq \gamma t'(1 + \sqrt{m} t_\alpha / K_2) + m t_\alpha \leq q/4$ and $D(\mathsf{sk}, \mathbf{c}) = 0$. We complete the proof. □

5.4.5 Main Protocol

Protocol 5.4.17 (Protocol PPK)**.** Let P and V denote the prover and the verifier, respectively. The common input is a pair $(\mathsf{pk}, \mathbf{c})$. The auxiliary input is a pair (σ, r) such that $\mathbf{c} = E_{\mathsf{pk}}(\sigma; r)$.

Define a mapping $\mathcal{G}(\mathsf{pk}, \mathbf{c}) = (\mathbf{B}_{\mathsf{pk}}, \mathbf{x}_{\mathbf{c}}, t')$ where $t' = 4t$ and both $\mathbf{B}_{\mathsf{pk}}$ and $\mathbf{x}_{\mathbf{c}}$ are similar to $\mathcal{F}(\mathsf{pk}, \mathbf{c})$. Let $\mathrm{Protocol}'_0$ (or $\mathrm{Protocol}'_1$) be $\mathrm{Protocol}_0$ (or $\mathrm{Protocol}_1$) where $\mathcal{F}(\cdot, \cdot)$ is replaced by $\mathcal{G}(\cdot, \cdot)$ respectively.

Step P1 P selects $\sigma' \in \{0, 1\}$ and $r' \in \{0, 1\}^m$ randomly. P computes $\mathbf{c}' = E_{\mathsf{pk}}(\sigma'; r')$ and sends $\mathbf{c}'$ to V.

Step V1 V sends a random challenge bit $\delta \in \{0, 1\}$ to P.

Step P2 If $\delta = 0$, P sends the pair (σ', r'). If $\delta = 1$, P computes $\sigma'' = \sigma + \sigma' \bmod 2$ and sends σ'' to V. Let $\overline{\mathbf{c}} = (\mathbf{c} + \mathbf{c}') \bmod q$ and runs $\mathrm{Protocol}'_{\sigma''}$ on the input $(\mathsf{pk}, \overline{c})$ as the prover.

Step V2 If $\delta = 0$, V accepts if $\mathbf{c}' = E_{\mathsf{pk}}(\sigma'; r')$, else rejects. If $\delta = 1$, V runs the $\mathrm{Protocol}'_{\sigma''}$ on the input $(\mathsf{pk}, \overline{\mathbf{c}})$ as the verifier.

Theorem 5.4.18 (PPK for pR05)**.** *The interactive protocol* (P, V) *is a proof of knowledge system with knowledge error* $3/4$ *for* R_{pR05}*. Moreover, the protocol* (P, V) *is a computational zero knowledge under the assumption ISVP.*

Our proof is based on the proof of Goldwasser and Kharchenko [GK05]. Before describing the proof, we need lemmas that give the properties of the protocols.

Lemma 5.4.19. *For sufficiently large* n*,*

1. *If* $(\mathsf{sk}, \mathsf{pk})$ *be an instance of* pR05 *and* $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2 \bmod q$ *such that* $D(\mathsf{sk}, \mathbf{c}) = 0$ *and* $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{E}(\mathsf{pk}, \cdot)$*,* $\mathcal{G}(\mathsf{pk}, \mathbf{c})$ *is a YES instance of* GapCVP_γ*.*
2. *Let* $(\mathsf{sk}, \mathsf{pk})$ *be an instance of* pR05 *and* $\mathbf{c} = \binom{\mathbf{a}}{b} \in \mathbb{Z}_q^{n+1}$*. If* $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q > q/8$*, then* $\mathcal{G}(\mathsf{pk}, \mathbf{c})$ *is a NO instance of* GapCVP_γ*.*

Lemma 5.4.20. *For sufficiently large* n*,*

1. *If* $(\mathsf{sk}, \mathsf{pk})$ *be an instance of* pR05 *and* $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2 \bmod q$ *such that* $D(\mathsf{sk}, \mathbf{c}) = 1$ *and* $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{E}(\mathsf{pk}, \cdot)$*,* $\mathcal{G}(\mathsf{pk}, \mathbf{y})$ *is a YES instance of* GapCVP_γ*, where* $\mathbf{y} = \mathbf{c} - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$*.*
2. *Let* $(\mathsf{sk}, \mathsf{pk})$ *be an instance of* pR05 *and* $\mathbf{c} = \binom{\mathbf{a}}{b} \in \mathbb{Z}_q^{n+1}$*. If* $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q > 3q/8$*, then* $\mathcal{G}(\mathsf{pk}, \mathbf{y})$ *is a NO instance of* GapCVP_γ*, where* $\mathbf{y} = \mathbf{c} - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$*.*

The proofs of Lemma 5.4.19 and Lemma 5.4.20 are in Section 5.4.6. Let us prove Theorem 5.4.18.

Proof of completeness. Since it is evident, we omit the proof. □

Proof of validity with error $3/4$*.* Let $\mathsf{pk} = \{(\mathbf{a}_i, b_i)\}_{i=1,\dots,m}$ be a public key of pR05 and $\mathbf{c} = (\mathbf{a}, b) \in \mathbb{Z}_q^{n+1}$. Let P^* be an arbitrary prover that make V accept with probability $\epsilon + 3/4$ for $\epsilon > 0$ on the common input $(\mathsf{pk}, \mathbf{c})$.

We construct a knowledge extractor K as follows. K's input is $(\mathsf{pk}, \mathbf{c})$. First, K chooses a random tape of P^*. Let δ_1 denote a challenge bit in $\mathrm{Protocol}'_{\sigma''}$. K runs P^* three times, where the challenge bits are 0, $(1, 0)$ and $(1, 1)$. K obtains three views T_0, T_1, and T_2. Views are in forms that $T_0 = (\mathbf{c}', 0, \sigma', r')$, $T_1 = (\mathbf{c}', 1, \sigma'', T'_1)$, and $T_2 = (\mathbf{c}', 1, \sigma'', T'_2)$, where T'_1 and T'_2 are transcripts of $\mathrm{Protocol}'_{\sigma''}$ that δ_1 are 0 and 1 respectively. If any one of three views is rejected, K outputs $\perp$ and halts. Otherwise, i.e., three views are accepted, K obtains a vector $\mathbf{p}$ that is witness of GapCVP_γ using the extractor of $\mathrm{Protocol}'_0$ or $\mathrm{Protocol}'_1$. K outputs $(\mathbf{c}', \sigma', r', \sigma'', \mathbf{p})$ and halts.

Note that the probability K does not output $\perp$ is at least $\Theta(\epsilon)$. Therefore, K is indeed the knowledge extractor. □

Proof of zero-knowledge of PPK. Let $S_{\sigma''}$ be a simulator for Protocol$'_{\sigma''}$. We construct a simulator S as follows:

Step P1 Chooses $\Delta \in \{0,1\}$ randomly (a predictor of a challenge bit). If $\Delta = 0$, chooses σ', r' randomly and computes $\mathbf{c}' = E_{\mathsf{pk}}(\sigma'; r')$. If $\Delta = 1$, chooses σ'', r'' randomly, computes $\bar{\mathbf{c}} = E_{\mathsf{pk}}(\sigma''; r'')$, and sets $\mathbf{c}' = \bar{\mathbf{c}} - \mathbf{c} \bmod q$. Sends $\mathbf{c}'$ to V^*.

Step V1 Receives a challenge bit δ from V^*.

Step P2, V2 If $\Delta \neq \delta$, outputs $\perp$ and halts. If $\Delta = \delta = 0$ outputs $(\mathbf{c}', \delta, \sigma', r')$. If $\Delta = \delta = 1$, invoke $S_{\sigma''}$ with input $(\mathsf{pk}, \bar{\mathbf{c}})$. Let $T = S_{\sigma''}(\mathsf{pk}, \bar{\mathbf{c}})$. Outputs $(\mathbf{c}', \delta, \sigma'', T)$ and halts.

We assume that ISVP holds, hence according to the security property of pR05 if $\Delta = 0$ then $\mathbf{c}'$ is computationally indistinguishable from the uniform distribution on $\mathbb{Z}_q^{n+1}$; if $\Delta = 0$ then $\mathbf{c}' = \bar{\mathbf{c}} - \mathbf{c} \bmod q$ is also indistinguishable from the uniform distribution. Therefore, the generated transcripts is computationally indistinguishable from a real transcript. □

5.4.6 Proof of Lemmas

Proof of Lemma 5.4.19. **(1)** There are two cases that $\mathbf{c}$ can decrypts into 0: when both $\mathbf{c}_1$ and $\mathbf{c}_2$ are ciphertexts of 0 and when both are ciphertexts of 1.

Suppose that $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{E}(\mathsf{pk}, 0)$. From Lemma 5.4.8, $\mathrm{Dist}\left(\binom{K_1\mathbf{c}_i}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq t$ for $i = 1, 2$. By Lemma 5.4.21 below, Thus, for $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2 \bmod q$, we have that

$$\mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + \sqrt{n+1} \leq 4t = t'.$$

Next, suppose that $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{E}(\mathsf{pk}, 1)$. Thus, for $i = 1, 2$, $\bar{\mathbf{c}}_i = \mathbf{c}_i - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q \in \mathcal{E}(\mathsf{pk}, 0)$. By Lemma 5.4.21 below, we have that for $\bar{\mathbf{c}} = \bar{\mathbf{c}}_1 + \bar{\mathbf{c}}_2 \bmod q$, $\mathrm{Dist}\left(\binom{K_1\bar{\mathbf{c}}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + \sqrt{n+1}$. Consider the vector $\mathbf{c} = \bar{\mathbf{c}}_1 + \bar{\mathbf{c}}_2 + 2\lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$. Since q is a prime, we have $2\lfloor q/2 \rfloor = q - 1$. By Lemma 5.4.22 below, we have that $\mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + \sqrt{n+1} + 1 \leq 4t = t'$.
(2) Let $\mathbf{c} = \binom{\mathbf{a}}{b} \in \mathbb{Z}_q^{n+1}$ be any ciphertext such that $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q > q/8$. Let $T = \gamma t'$. Recall that $T(1 + \sqrt{m} t_\alpha / K_2) \leq q/8 < |b - \langle \mathbf{a}, \mathbf{s} \rangle|_q$. Hence, by Claim 5.4.9 $\mathrm{Dist}\left(\binom{K_1\mathbf{c}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq T$ can not hold. Thus, $\mathcal{G}(\mathsf{pk}, \mathbf{c})$ is a NO instance of the GapCVP$_\gamma$. □

Proof of Lemma 5.4.20. **(1)** Without loss of generality, we suppose that $\mathbf{c}_1 \in \mathcal{E}(\mathsf{pk}, 0)$ and $\mathbf{c}_2 \in \mathcal{E}(\mathsf{pk}, 1)$. From Lemma 5.4.8 and Lemma 5.4.12, for some $\mathbf{p}_1, \mathbf{p}_2 \in L(\mathbf{B}_{\mathsf{pk}})$ $\mathrm{Dist}\left(\binom{K_1\mathbf{c}_1}{\mathbf{0}}, \mathbf{p}_1\right) \leq t$ and $\mathrm{Dist}\left(\binom{K_1(\mathbf{c}_2 - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q)}{\mathbf{0}}, \mathbf{p}_2\right) \leq t$. Hence, from $\mathbf{y} = \mathbf{c}_1 + \mathbf{c}_2 - \lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q$, we obtain

$$\mathrm{Dist}\left(\binom{K_1\mathbf{y}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 2t + 1 \leq 4t = t'$$

by Lemma 5.4.21.

(2) Let $\mathbf{c} = \binom{\mathbf{a}}{b} \in \mathbb{Z}_q^{n+1}$ be any ciphertext such that $|b - \langle \mathbf{a}, \mathbf{s} \rangle|_q \leq 3q/8$. Let $\mathbf{y} = \binom{\mathbf{a}'}{b'}$. In this case, we obtain that $|b' - \langle \mathbf{a}', \mathbf{s} \rangle|_q \geq q/8$. Let $T = \gamma t'$. Note that $T(1 + \sqrt{m} t_\alpha / K_2) \leq q/8 < |b' - \langle \mathbf{a}', \mathbf{s} \rangle|_q$. Hence, by Claim 5.4.9 $\mathrm{Dist}\left(\binom{K_1 \mathbf{y}}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq T$ can not hold. Thus, $\mathcal{G}(\mathsf{pk}, \mathbf{y})$ is a NO instance of GapCVP_γ. □

Lemma 5.4.21. *Let* pk *be a public key of* pR05, $\mathbf{p}_1$ *and* $\mathbf{p}_2$ *points from* $L(\mathbf{B}_{\mathsf{pk}})$. *If for* $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{Z}_q^{n+1}$, $\mathrm{Dist}\left(\binom{K_1 \mathbf{c}_1}{\mathbf{0}}, \mathbf{p}_1\right) = d_1$ *and* $\mathrm{Dist}\left(\binom{K_1 \mathbf{c}_2}{\mathbf{0}}, \mathbf{p}_2\right) = d_2$, *then* $\mathrm{Dist}\left(\binom{K_1(\mathbf{c}_1 + \mathbf{c}_2 \bmod q)}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq d_1 + d_2 + \sqrt{n+1}$.

Proof. Represent $K_1(\mathbf{c}_1 + \mathbf{c}_2 \bmod q) = K_1(\mathbf{c}_1 + \mathbf{c}_2 + \sum_{i=1}^{n+1} \alpha_i q \mathbf{u}_i)$. Since both vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ belong to $\{0, 1, \ldots, q-1\}^{n+1}$, we can bound $|\alpha_i| \leq 1$ for all i. Consider a vector $\mathbf{p}_3 = \mathbf{B}_{\mathsf{pk}}{}^t(\alpha_1, \ldots, \alpha_{n+1}, 0, \ldots, 0)$. Thus, we obtain that

$$\mathrm{Dist}\left(\binom{K_1 \sum_{i=1}^{n+1} \alpha_i q \mathbf{u}_i}{\mathbf{0}}, \mathbf{p}_3\right) \leq \sqrt{\sum_{i=1}^{n+1} \alpha_i^2} \leq \sqrt{n+1}.$$

By the triangle inequality, the lemma follows. □

Lemma 5.4.22. *Let* pk *be a public key of* pR05 *and* $\mathbf{p}$ *a point from* $L(\mathbf{B}_{\mathsf{pk}})$. *If for* $\mathbf{c} \in \mathbb{Z}_q^{n+1}$, $\mathrm{Dist}\left(\binom{K_1 \mathbf{c}}{\mathbf{0}}, \mathbf{p}\right) = d$ *then* $\mathrm{Dist}\left(\binom{K_1(\mathbf{c} + 2\lfloor q/2 \rfloor \mathbf{u}_{n+1} \bmod q)}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq d + 1$.

Proof. Since q is an odd prime, we have that $2\lfloor q/2 \rfloor = q - 1$. Represent $K_1(\mathbf{c} + (q-1)\mathbf{u}_{n+1} \bmod q) = K_1(\mathbf{c} + (q-1)\mathbf{u}_{n+1} + \alpha_{n+2}(q-1)\mathbf{u}_{n+1})$ for some $\alpha \in \{-1, 0\}$. Consider a vector $\mathbf{p}'$ in $L(\mathbf{B}_{\mathsf{pk}})$ such that $\mathbf{p}' = L(\mathbf{B}_{\mathsf{pk}})^t(0, \ldots, 0, \alpha_{n+2}, 0, \ldots, 0)$ (with 1 at the $(n+2)$-th position). By the construction of $\mathbf{B}_{\mathsf{pk}}$, we have that $\mathrm{Dist}\left(\binom{K_1((q-1)\mathbf{u}_{n+1} + \alpha_{n+2}(q-1)\mathbf{u}_{n+1})}{\mathbf{0}}, L(\mathbf{B}_{\mathsf{pk}})\right) \leq 1$. By the triangle inequality, the lemma follows. □

5.5 Concluding Remarks

In this chapter we have constructed proofs of plaintext knowledge for pR04 and pR05.

We list up a few open problems: Verifiable decryption for the lattice-based cryptosystems and non-malleable proofs for plaintext knowledge for the lattice-based cryptosystems. The former has many applications. The latter are sources of interactive CCA2-secure cryptosystems.

Acknowledgement

I would like to thank my supervisor Keisuke Tanaka for insightful suggestions and creative comments. Without his guidance and wide knowledge, this thesis could not be done. I would also like to thank Akinori Kawachi for thoughtful discussions and supports. Finally, I would like to thank all members of Tanaka Laboratory for their advice, encouragement, and friendship.

References

[AABN02] Michel Abdalla, Jee Hea An, Mihir Bellare, and Chanathip Namprempre. From identification to signatures via the Fiat-Shamir transform: Minimizing assumptions for security and forward-security. In Lars Knudsen, editor, *Advances in Cryptology – EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 418–433, Amsterdam, The Netherlands, April 2002. Springer-Verlag.

[AR05] Dorit Aharonov and Oded Regev. Lattice problems in NP cap coNP. *Journal of the ACM*, 52(5):749–765, 2005.

[Ajt96a] Miklós Ajtai. Generating hard instances of lattice problems. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(007), 1996.

[Ajt96b] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *Proceedings on 28th Annual ACM Symposium on Theory of Computing (STOC '96)*, pages 99–108, Philadelphia, Pennsylvania, USA, May 1996. ACM. See also ECCC TR96-007.

[Ajt98] Miklós Ajtai. The shortest vector problem in L_2 is NP-hard for randomized reductions (extended abstract). In *Proceedings on 30th Annual ACM Symposium on Theory of Computing (STOC '98)*, pages 10–19, Dallas, Texas, USA, May 1998. ACM. See also ECCC TR97-047.

[Ajt05] Miklós Ajtai. Representing hard lattices with $O(n \log n)$ bits. In Harold N. Gabow and Ronald Fagin, editors, *Proceedings on the 37th Annual ACM Symposium on Theory of Computing (STOC 2005)*, pages 94–103, Baltimore, MD, USA, May 2005. ACM.

[AD96] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(065), 1996.

[AD97] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proceedings on 29th Annual ACM Symposium on Theory of Computing (STOC '97)*, pages 284–293, El Paso, Texas, USA, May 1997. ACM. See also ECCC TR96-065.

[ABS97] Sanjeev Arora, László Babai, and Z. Stern, Jacques Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *Journal of Computer and System Sciences*, 54(2):317–331, 1997.

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 4(296):625–635, 1993.

[BMvT78] Elwyn R. Berlekamp, Robert J. McEliece, and Henk C. A. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24(3):384–386, MAY 1978.

[Cai98] Jin-Yi Cai. A relation of primal-dual lattices and the complexity of shortest lattice vector problem. *Theoretical Computer Science*, 207(1):105–116, 1998.

[Cai03] Jin-Yi Cai. A new transference theorem in the geometry of numbers and new bounds for Ajtai's connection factor. *Discrete Applied Mathematics*, 126(1):9–31, 2003.

[CC99] Jin-Yi Cai and Thomas W. Cusick. A lattice-based public-key cryptosystem. *Information and Computation*, 151(1-2):17–31, 1999.

[CN97] Jin-Yi Cai and Ajay Nerurkar. An improved worst-case to average-case connection for lattice problems. In *38th Annual Symposium on Foundations of Computer Science (FOCS '97)*, pages 468–477, Miami Beach, Florida, USA, October 1997. IEEE Computer Society.

[DPP97] Ivan B. Damgård, Torben P. Pedersen, and Birgit Pfizmann. On the existence of statistically hiding bit commitment schemes and fail-stop signatures. *Journal of Cryptology*, 10(3):163–194, 1997. Preliminary version in *CRYPTO '93*, 1993.

[DPP98] Ivan B. Damgård, Torben P. Pedersen, and Birgit Pfizmann. Statistical secrecy and multibit commitments. *IEEE Transactions on Information Theory*, 44(3):1143–1151, MAY 1998.

[ElG85] Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transaction on Information Theory*, 31(4):469–472, 1985.

[Eur06] European Network of Excellence for Cryptology. *PQCrypto 2006: International Workshop on Post-Quantum Cryptography*, Katholieke Universiteit Leuven, Belgium, MAY 2006.

[FS86] Amos Fiat and Adi Shamir. How to prove yourself – practical solutions to identification and signature problems. In Andrew M. Odlyzko, editor, *Advances in Cryptology – CRYPTO '86*, volume 263 of *Lecture Notes in Computer Science*, pages 186–194, Santa Barbara, California, USA, August 1986. Springer-Verlag.

[GHY85] Zvi Galil, Stuart Haber, and Moti Yung. Symmetric public-key encryption. In Hugh C. Williams, editor, *Advances in Cryptology – CRYPTO '85*, volume 218 of *Lecture Notes in Computer Science*, pages 128–137, Santa Barbara, California, USA, August 1985. Springer-Verlag.

[Gen01] Craig Gentry. Key recovery and message attacks on NTRU-composite. In Birgit Pfitzmann, editor, *Advances in Cryptology – EUROCRYPT 2001*, volume 2045 of *Lecture Notes in Computer Science*, pages 182–194, Innsbruck, Austria, May 2001. Springer-Verlag.

[GJSS01] Craig Gentry, Jakob Jonsson, Jacques Stern, and Mike Szydlo. Cryptanalysis of the NTRU signature scheme (NSS) from Eurocyrpt 2001. In C. Boyd, editor, *Advances in Cryptology – ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 1–20, Gold Coast, Australia, December 2001. Springer-Verlag.

[GS02] Craig Gentry and Mike Szydlo. Cryptanalysis of the revised NTRU signature scheme. In Lars Knudsen, editor, *Advances in Cryptology – EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 299–320, Amsterdam, The Netherlands, April 2002. Springer-Verlag.

[Gol01] Oded Goldreich. *Foundations of Cryptography: Basic Tools*. Cambridge University Press, 2001.

[GG00] Oded Goldreich and Shafi Goldwasser. On the limits of nonapproximability of lattice problems. *Journal of Computer and System Sciences*, 60(3):540–563, 2000.

[GGH96] Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Collision-free hashing from lattice problems. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(42), 1996.

[GGH97a] Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Eliminating decryption errors in the Ajtai-Dwork cryptosystem. In Burton S. Kaliski, Jr., editor, *Advances in Cryptology – CRYPTO '97*, volume 1294 of *Lecture Notes in Computer Science*, pages 105–111, Santa Barbara, California, USA, August 1997. Springer-Verlag. See also ECCC TR97-018.

[GGH97b] Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Public-key cryptosystems from lattice reduction problems. In Burton S. Kaliski, Jr., editor, *Advances in Cryptology – CRYPTO '97*, volume 1294 of *Lecture Notes in Computer Science*, pages 112–131, Santa Barbara, California, USA, August 1997. Springer-Verlag.

[GK05] Shafi Goldwasser and Dmitriy Kharchenko. Proof of plaintext knowledge for the Ajtai-Dwork cryptosystem. In Joe Kilian, editor, *Theory of Cryptography, 2nd Theory of Cryptography Conference, TCC 2005*, volume 3378 of *Lecture Notes in Computer Science*, pages 529–555, Cambridge, MA, USA, February 2005. Springer-Verlag.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, 1984.

[GTK03] Shafi Goldwasser and Yael Tauman Kalai. On the (in)security of the Fiat-Shamir paradigm. *Electronic Colloquium on Computational Complexity (ECCC)*, 10(015), 2003.

[GLS88] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinational Optimization*. Springer-Verlag, 1988.

[HM96] Shai Halevi and Silvio Micali. Practical and provably-secure commitment scheme from collision-free hashing. In Neal Koblitz, editor, *Advances in Cryptology – CRYPTO '96*, volume 1109 of *Lecture Notes in Computer Science*, pages 201–215, Santa Barbara, California, USA, August 1996. Springer-Verlag.

[HT06] Shunichi Hayashi and Mitsuru Tada. A lattice-based public-key identification scheme. In *The 2006 International Symposium on Information Theory and its Applications (ISITA 2006)*, 2006.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[HHGP+03] Jeffrey Hoffstein, Nick Howgrave-Graham, Jill Pipher, Joseph H. Silverman, and William Whyte. NTRUsign: Digital signature using the NTRU lattice. In Marc Joye, editor, *Topics in Cryptology – CT-RSA 2003*, volume 2612 of *Lecture Notes in Computer Science*, pages 122–140, San Francisco, CA, USA, April 2003. Springer-Verlag.

[HPS98] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NTRU: A ring-based public key cryptosystem. In Joe Buhler, editor, *Algorithmic Number Theory, Third International Symposium, ANTS-III*, volume 1423 of *Lecture Notes in Computer Science*, pages 267–288, Portland, Oregon, USA, June 1998. Springer-Verlag.

[HPS01] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NSS: An NTRU lattice-based signature scheme. In Birgit Pfitzmann, editor, *Advances in Cryptology – EUROCRYPT 2001*, volume 2045 of *Lecture Notes in Computer Science*, pages 211–228, Innsbruck, Austria, May 2001. Springer-Verlag.

[HGNP+03] Nick Howgrave-Graham, Phong Q. Nguyen, David Pointcheval, John Proos, Joseph H. Silverman, Ari Singer, and William Whyte. The impact of decryption failures on the security of NTRU encryption. In Dan Boneh, editor, *Advances in Cryptology – CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 226–246, Santa Barbara, California, USA, August 2003. Springer-Verlag.

[KTX06] Akinori Kawachi, Keisuke Tanaka, and Keita Xagawa. Multi-bit cryptosystems based on lattice problems, 2006. Manuscript.

[Kho04] Subhash Khot. Hardness of approximating the shortest vector problem in lattices. In *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*, pages 126–135, Rome, Italy, October 2004. IEEE Computer Society.

[KS01] S.-Ravi Kumar and D. Sivakumar. On the unique shortest lattice vector problem. *Theoretical Computer Science*, 255(1-2):641–648, 2001.

[Lag85] Jeffrey C. Lagarias. The computational complexity of simultaneous Diophantine approximation problems. *SIAM Journal of Computing*, 14(1):196–209, 1985.

[LLL82] Arjen K. Lenstra, Hendrik W. Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):513–534, 1982.

[LM05] Vadim Lyubashevsky and Daniele Micciancio. Generalized compact knapsacks are collision resistant. *Electronic Colloquium on Computational Complexity (ECCC)*, 12(142), 2005.

[Mic01] Daniele Micciancio. Improving lattice based cryptosystems using the Hermite normal form. In Joseph H. Silverman, editor, *Cryptography and Lattices, International Conference, CaLC 2001*, volume 2146 of *Lecture Notes in Computer Science*, pages 126–145, Providence, RI, USA, March 2001. Springer-Verlag.

[Mic04a] Daniele Micciancio. Almost perfect lattices, the covering radius problem, and applications to Ajtai's connection factor. *SIAM Journal on Computing*, 34(1):118–169, 2004. Preliminary version in *STOC 2002*, 2002.

[Mic04b] Daniele Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions from worst-case complexity assumptions. *Electronic Colloquium on Computational Complexity (ECCC)*, 11(095), 2004.

[MG02] Daniele Micciancio and Shafi Goldwasser. *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, Massachusetts, March 2002.

[MR04] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*, pages 372–381, Rome, Italy, October 2004. IEEE Computer Society.

[MV03] Daniele Micciancio and Salil Vadhan. Statistical zero-knowledge proofs with efficient provers: Lattice problems and more. In Dan Boneh, editor, *Advances in Cryptology – CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 282–298, Santa Barbara, California, USA, August 2003. Springer-Verlag.

[Ngu99] Phong Q. Nguyen. Cryptanalysis of the Goldreich-Goldwasser-Halevi cryptosystem from Crypto '97. In Michael Wiener, editor, *Advances in Cryptology – CRYPTO '99*, volume 1666 of *Lecture Notes in Computer Science*, pages 288–304, Santa Barbara, California, USA, August 1999. Springer-Verlag.

[Ngu02] Phong Q. Nguyen. Analysis and improvements of NTRU encryption paddings. In Moti Yung, editor, *Advances in Cryptology – CRYPTO 2002*, volume 2442 of

Lecture Notes in Computer Science, pages 210–225, Santa Barbara, California, USA, August 2002. Springer-Verlag.

[NR06] Phong Q. Nguyen and Oded Regev. Learning a parallelepiped: Cryptanalysis of GGH and NTRU signatures. In Serge Vaudenay, editor, *Advances in Cryptology – EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pages 271–288, St. Peterburg, Russia, May/June 2006. Springer-Verlag.

[NS98] Phong Q. Nguyen and Jacques Stern. Cryptanalysis of the Ajtai-Dwork cryptosystem. In Hugo Krawczyk, editor, *Advances in Cryptology – CRYPTO '98*, volume 1462 of *Lecture Notes in Computer Science*, pages 223–242, Santa Barbara, California, USA, August 1998. Springer-Verlag.

[OO98] Kazuo Ohta and Tatsuaki Okamoto. On concrete security treatment of signatures derived from identification. In Hugo Krawczyk, editor, *Advances in Cryptology – CRYPTO '98*, volume 1462 of *Lecture Notes in Computer Science*, pages 354–369, Santa Barbara, California, USA, August 1998. Springer-Verlag.

[PJH03] Seong-Hun Paeng, Bae Eun Jung, and Kil-Chan Ha. A lattice based public key cryptosystem using polylnomial representations. In Yvo Desmedt, editor, *Public Key Cryptography – PKC 2003*, volume 2567 of *Lecture Notes in Computer Science*, pages 292–308, Miami, Florida, USA, January 2003. Springer-Verlag.

[PR06] Chris Peikert and Alon Rosen. Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography, 3rd Theory of Cryptography Conference, TCC 2006*, volume 3876 of *Lecture Notes in Computer Science*, pages 145–166, New York, NY, USA, March 2006. Springer-Verlag.

[PS96] David Pointcheval and Jacques Stern. Security proofs for signature schemes. In Ueli Maurer, editor, *Advances in Cryptology – EUROCRYPT '96*, volume 1070 of *Lecture Notes in Computer Science*, pages 387–398, Saragossa, Spain, May 1996. Springer-Verlag.

[Rap04] Dörte Rappe. *Homomorphic Cryptosystems and Their Applications*. PhD thesis, University of Dortmund, 2004. Also available at http://eprint.iacr.org/2006/001.

[Reg04] Oded Regev. New lattice-based cryptographic constructions. *Journal of the ACM*, 51(6):899–942, 2004. Preliminary version in *STOC 2003*, 2003.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *Proceedings on the 37th Annual ACM Symposium on Theory of Computing (STOC 2005)*, pages 84–93, Baltimore, MD, USA, May 2005. ACM.

[RSA78] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, Febrary 1978.

[Sha89] Adi Shamir. An efficient identification scheme based on permuted kernels (extended abstract). In Gilles Brassard, editor, *Advances in Cryptology – CRYPTO '89*, volume 435 of *Lecture Notes in Computer Science*, pages 606–609, Santa Barbara, California, USA, August 1989. Springer-Verlag.

[Sho97] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[Sil01] Joseph H. Silverman, editor. *Cryptography and Lattices, International Conference, CaLC 2001*, volume 2146 of *Lecture Notes in Computer Science*, Providence, RI, USA, March 2001. Springer-Verlag.

[Ste96] Jacques Stern. A new paradigm for public key identification. *IEEE Transactions on Information Theory*, 42(6):749–765, November 1996. Preliminary version in *CRYPTO '93*, 1993.

[Szy03] Michael Szydlo. Hypercubic lattice reduction and analysis of GGH and NTRU signatures. In Eli Biham, editor, *Advances in Cryptology – EUROCRYPT 2003*, volume 2656 of *Lecture Notes in Computer Science*, pages 433–448, Warsaw, Poland, May 2003. Springer-Verlag.

[Tro01] Mårten Trolin. The shortest vector problem in lattices with many cycles. In Joseph H. Silverman, editor, *Cryptography and Lattices, International Conference, CaLC 2001*, volume 2146 of *Lecture Notes in Computer Science*, pages 194–205, Providence, RI, USA, March 2001. Springer-Verlag.