SCIS 2020 2020 Symposium on
Cryptography and Information Security
Kochi, Japan, Jan. 28 – 31, 2020
The Institute of Electronics,
Information and Communication Engineers

On the hardness of LWE with Non-uniform Binary Error

Sun Chao * Mehdi Tibouchi † Masayuki Abe ‡

Abstract: Binary-Error LWE is the particular case of the learning with errors problem in which errors are chosen from $\{0, 1\}$. For uniform case, Arora and Ge showed that the problem can be solved in polynomial time given a number of samples quadratic in the dimension n, and it can also be solved in subexponential time given a superlinear number of samples, by applying Gröbner basis techniques to the system arising from Arora and Ge's approach. On the other hand, Micciancio and Peikert showed the Uniform Binary-Error LWE problem reduces to standard LWE (and thus is believed to be exponentially hard) when the number of samples is restricted to $n + O(n/\log n)$. In this paper, we examine more generally about the hardness of the non-uniform Binary-Error LWE and how it varies with the error rate and the number of available samples. In particular, we show that, when error rate is low, non-uniform Binary-Error LWE can be solved in polynomial time given $O(n)$ samples, and when the error rate is slightly higher, it can solved in subexponential time with $O(n)$ samples. Besides, we generalize the hardness proof by Micciancio and Peikert to the non-uniform case.

Keywords: LWE, Binary-Error LWE, Lossy Function Family, Arora-Ge algorithm, Time complexity, Reduction

1 Introduction

The learning with errors problem (LWE), introduced by Regev in 2005 [9], is one of the central problems of lattice-based cryptography. It can be seen as an average-case problem which, for suitable parameters, is as hard as worst-case lattice problems, and it is therefore very convenient to build secure lattice-based cryptographic schemes: it has been used to build various primitives from encryption and signatures all the way to fully-homomorphic encryption.

For efficiency reasons, constructions often rely on variants of LWE (such as its ring version Ring-LWE [6]) or instantiations in more aggressive ranges of parameters than those for which Regev's reduction to worst-case lattice problems holds. An important example is *Binary-Error LWE*, where the error term is sampled uniformly from $\{0, 1\}$ (instead of from a wider discrete Gaussian distribution). Binary-Error LWE is a particularly simple problem with various interesting cryptographic applications, such as Buchmann et al.'s efficient lattice-based encryption scheme for IoT and lightweight devices [3] (based on the ring version of Binary-Error LWE, with the additional constraint that the secret is binary as well).

However, the problem is not hard given arbitrarily many samples: in fact, an algebraic attack due to Arora and Ge [2] solves uniform Binary-Error LWE in polynomial time given around $n^2/2$ samples. The same approach can also be combined by Gröbner basis techniques to reduce the number of required samples [1]. On the other hand, Micciancio and Peikert [8] showed the Binary-Error LWE problem reduces to standard LWE (and thus is believed to be exponentially hard) when the number of samples is restricted to $n + O(n/\log n)$. Thus, the hardness of Binary-Error LWE crucially depends on the number of samples released to the adversary.

In this paper, we make a generalization of the uniform Binary-Error LWE to the non-uniform case, in which the error is chosen from $\{0, 1\}$ and the error is 1 with some probability p. We analyze this problem from two perspectives. On one hand, when the error rate is $p = O(1/n)$, it can be solved in polynomial time with $O(n)$ samples, and when the error rate is $p = O(1/n^\alpha)$ $(0 < \alpha < 1)$, it can be solved in subexponential time with $O(n)$ samples. On the other hand, we generalize the hardness proof given by Micciancio and Peikert. In particular, we show that non-uniform Binary-Error LWE with any constant error rate reduces to standard LWE, as long as the number of samples is restricted.

2 Preliminaries

2.1 Learning with Errors

The LWE problem asks to recover a secret $\mathbf{s} \in \mathbb{F}_q^n$, given a system of linear approximate equations. For

* Kyoto University, Kyoto, Japan (sun.chao.46s@st.kyoto-u.ac.jp)
† NTT Secure Platform Laboratories , Tokyo, Japan
‡ NTT Secure Platform Laboratories , Tokyo, Japan

instance,

$$\begin{aligned}
14s_1 + 15s_2 + 5s_3 + 2s_4 &\approx 8 \pmod{17} \\
13s_1 + 14s_2 + 14s_3 + 6s_4 &\approx 16 \pmod{17} \\
6s_1 + 10s_2 + 13s_3 + s_4 &\approx 3 \pmod{17} \\
10s_1 + 4s_2 + 12s_3 + 16s_4 &\approx 12 \pmod{17} \\
9s_1 + 5s_2 + 9s_3 + 6s_4 &\approx 9 \pmod{17} \\
3s_1 + 6s_2 + 4s_3 + 5s_4 &\approx 16 \pmod{17}
\end{aligned}$$

Each equation is satisfied up to some small error, sampled independently according to some known distribution (typically a discrete Gaussian distribution). The goal is to recover the secret $\mathbf{s}$. If the equation held without error, finding $\mathbf{s}$ would simply amount to solving a system of linear equations. We could therefore recover the secret $\mathbf{s}$ in polynomial time $O(n^\omega)$, where $2 \leq \omega \leq 3$ is the complexity exponent of linear algebra ($\omega \approx 2.37$ using the best known approach [5]). However, the errors introduced in LWE typically make the problem much harder. Formally, the LWE problem can be defined as follows.

Definition 2.1 (LWE)**.** The (search) LWE problem, defined with respect to a dimension n, a modulus q and an error distribution χ over $\mathbb{F}_q$, asks to recover a secret vector $\mathbf{s} \in \mathbb{F}_q^n$ given polynomially many samples of the form

$$\big(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e \bmod q\big) \in \mathbb{F}_q^n \times \mathbb{F}_q \tag{1}$$

where $\mathbf{a}$ is uniformly random in $\mathbb{F}_q^n$, and e is sampled according to χ. One can optionally specify the number $m = \text{poly}(n)$ of available samples as an additional parameter.

Remark 2.1. One can also similarly define a decision variant of the LWE problem, which asks to distinguish the distribution of the samples (1) above from the uniform distribution over $\mathbb{F}_q^n \times \mathbb{F}_q$.

The LWE problem given m samples has a simple expression in matrix form: it asks to recover $\mathbf{s}$ from the pair $(A, \mathbf{b})$ where $A \in \mathbb{F}_q^{m \times n}$ is a uniformly random matrix, and $\mathbf{b} = A\mathbf{s} + \mathbf{e} \bmod q$, where all the coefficients of $\mathbf{e} \in \mathbb{F}_q^m$ are sampled independently from χ.

2.2 Binary-Error LWE

The Binary-Error LWE is simply the special case of Definition 2.1 where χ is the uniform distribution over $\{0, 1\}$. In other words:

Definition 2.2 (Binary-Error LWE)**.** The Binary-Error LWE with parameters n, m and q asks to recover the vector $\mathbf{s} \in \mathbb{F}_q^n$ from m samples of the form:

$$\big(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e \bmod q\big) \in \mathbb{F}_q^n \times \mathbb{F}_q$$

where $\mathbf{a}$ is uniformly random in $\mathbb{F}_q^n$, and e is uniform in $\{0, 1\}$.

The dimension n is the main security parameter, and both m and q are typically chosen as polynomially bounded functions of n. In this paper, we assume that $q = n^{\Theta(1)}$.

2.3 One-Way Functions

A function family is a probability distribution $\mathcal{F}$ over a set of functions $\mathcal{F} \subset (X \to Y)$ with common domain X and range Y. Let $\mathcal{X}$ be a probability distribution over the domain X of a function family $\mathcal{F}$. We recall the following standard security notions:

- $(\mathcal{F}, \mathcal{X})$ is (t, ε)-one-way if for all probabilistic algorithms $\mathcal{A}$ running in time at most t,

  $$\Pr\left[f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X} : \mathcal{A}(f, f(x)) \in f^{-1}(f(x))\right] \leq \varepsilon$$

- $(\mathcal{F}, \mathcal{X})$ is (t, ε)-uninvertible if for all probabilistic algorithms $\mathcal{A}$ running in time at most t,

  $$\Pr[f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X} : \mathcal{A}(f, f(x)) = x] \leq \varepsilon$$

- $(\mathcal{F}, \mathcal{X})$ is (t, ε)-second preimage resistant if for all probabilistic algorithms $\mathcal{A}$ running in time at most t,

  $$\Pr\left[f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X}, x' \leftarrow \mathcal{A}(f, x) : f(x) = f\left(x'\right) \wedge x \neq x'\right] \leq \varepsilon$$

- $(\mathcal{F}, \mathcal{X})$ is (t, ε)-pseudorandom if the distributions $\{f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X} : (f, f(x))\}$ and $\{f \leftarrow \mathcal{F}, y \leftarrow \mathcal{U}(Y) : (f, y)\}$ are (t, ε)-indistinguishable.

2.4 Lossy Function Families

Definition 2.3 (Lossy Function Families [8])**.** Let ($\mathcal{L}$, $\mathcal{F}$) be two probability distributions(with possbily different supports) over the same set of (efficiently computable) functions $\mathcal{F} \subset X \to Y$, and let $\mathcal{X}$ be an efficiently sampleable distribution over the domain X. We say that ($\mathcal{L}$, $\mathcal{F}$, $\mathcal{X}$) is a lossy function family if the following properties are satisfied:

- the distributions $\mathcal{L}$ and $\mathcal{F}$ are indistinguishable.
- ($\mathcal{L}$, $\mathcal{X}$) is uninvertible.
- ($\mathcal{F}$, $\mathcal{X}$) is second preimage resistant.

Theorem (1). [8] Let $\mathcal{F}$ be a family of functions computable in time t'. If $(\mathcal{F}, \mathcal{X})$ is both (t, ε)-uninvertible and $(t + t', \varepsilon')$-second preimage resistant, then it is also $(t, \varepsilon + \varepsilon')$-one-way.

Theorem (2). [8] Let $\mathcal{F}$ and $\mathcal{F}'$ be any two indistinguishable, efficiently computable function families, and let $\mathcal{X}$ be an efficiently sampleable input distribution. Then if ($\mathcal{F}$, $\mathcal{X}$) is uninvertible(respectively, second-preimage resistant), then ($\mathcal{F}'$, $\mathcal{X}$) is also uninvertible(resp., second preimage resistant). In particular, if ($\mathcal{L}$, $\mathcal{F}$, $\mathcal{X}$) is a lossy function family, then ($\mathcal{L}$, $\mathcal{X}$) and ($\mathcal{F}$, $\mathcal{X}$) are both one-way.

2.5 SIS and LWE functions

The Short Integer Solution function family $\text{SIS}(m, n, q, X)$ is the set of all functions f_A indexed by $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ with domain $X \subseteq \mathbb{Z}^m$ and range $Y = \mathbb{Z}_q^n$ defined as $f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \bmod q$. The Learning With Error function family $\text{LWE}(m, n, q, X)$ is the set of all functions g_A indexed by $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ with domain $\mathbb{Z}_q^n \times X$ and range $Y = \mathbb{Z}_q^m$, defined as $g_{\mathbf{A}}(\mathbf{s}, \mathbf{x}) = \mathbf{A}^T\mathbf{s} + \mathbf{x} \bmod q$

Theorem (3). [7] For any $n, m \geq n + \omega(logn), q$, and distribution $\mathcal{X}$ over $\mathbb{Z}^m$, the LWE(m, n, q) function family is one-way (resp. pseudorandom, or uninvertible) with respect to input distribution $U(\mathbb{Z}_q^n) \times \mathcal{X}$ if and only if the SIS(m, m - n, q) function family is one-way(resp. pseudorandom , or uninvertible) with respect to the input distribution $\mathcal{X}$.

2.6 Uniqueness of solutions

Theorem (4). Suppose that the following condition is satisfied:

$$m \geq n \cdot \left(1 + \frac{c}{\log q}\right)$$

for some $c > \log 3$. Then, the Binary-Error LWE problem with parameters n, m, q has a unique solution with overwhelming probability.

Proof. Indeed, suppose that two solutions $\mathbf{s} \neq \mathbf{s}'$ exist to the Binary-Error LWE challenge $(A, \mathbf{b})$. This means that there exists binary error vectors $\mathbf{e}, \mathbf{e}'$ such that:

$$\mathbf{b} = A\mathbf{s} + \mathbf{e} = A\mathbf{s}' + \mathbf{e}'.$$

As a result, the vector $\mathbf{t} = \mathbf{s}' - \mathbf{s} \neq 0$ satisfies $A\mathbf{t} = \mathbf{e} - \mathbf{e}' \in \{-1, 0, 1\}^m$. It thus suffices to prove that for a random $A \in \mathbb{F}_q^{m \times n}$, such a vector $\mathbf{t}$ can only exist with negligible probability.

We can proceed as follows: fix $\mathbf{t} \in \mathbb{F}_q^n \setminus \{0\}$. For a uniformly random $A \in \mathbb{F}_q^{m \times n}$, the probability that $A\mathbf{t} \in \{-1, 0, 1\}^m$ is exactly $3^m/q^m$, since the product vector is uniformly distributed in $\mathbb{F}_q^m$. As a result, the union bound shows that:

$$\Pr_{A \xleftarrow{\$} \mathbb{F}_q^{m \times n}} \left[\exists \mathbf{t} \in \mathbb{F}_q^m \setminus \{0\}, A\mathbf{t} \in \{-1, 0, 1\}^m\right] \leq \left(\frac{3}{q}\right)^m \cdot q^n$$

since there are fewer than q^n possible vectors $\mathbf{t}$.

Therefore, assuming without loss of generality that $q > 3$, the probability ε that the challenge has at least two solutions is bounded as:

$$\begin{aligned}
\varepsilon &\leq \left(\frac{3}{q}\right)^m \cdot q^n \\
\log \varepsilon &\leq m \log\left(\frac{3}{q}\right) + n \log q \\
&\leq n\left(1 + \frac{c}{\log q}\right) \log\left(\frac{3}{q}\right) + n \log q \\
&= n\left(\log 3 - \log q + \frac{c \log 3}{\log q} - c + \log q\right) \\
&= n\big(\log 3 - c + o(1)\big)
\end{aligned}$$

and since $c > \log 3$, it follows that ε is negligible. □

2.7 Naive algorithm

From now on, we assume that the hypothesis of Theorem 4 is satisfied. It is easy to see that the matrix A is then of rank n with overwhelming probability (indeed, that probability is exactly $(1 - q^{-m})(1 - q^{1-m}) \cdots (1 - q^{n-1-m}) \geq 1 - q^{n-m}$, and this can be used to deduce a "naive" algorithm for Binary-Error LWE in time $O^*(2^n)$, essentially by guessing n coefficients of the error vector $\mathbf{e}$.

More precisely, since A is full rank, one can assume without loss of generality that its first n rows form an invertible square submatrix A_0. An algorithm for Binary-Error LWE is then as follows: guess the vector $\mathbf{e}_0 \in \{0, 1\}^n$ consisting of the first n coefficients of $\mathbf{e}$; then deduce the corresponding $\mathbf{s} = A_0^{-1}(\mathbf{b}_0 - \mathbf{e}_0)$, and check that $\mathbf{e} = \mathbf{b} - A\mathbf{s}$ is indeed in $\{0, 1\}^m$. The check is performed in poly(n) time, and by Theorem 4, there is with overwhelming probability a unique $\mathbf{e}_0 \in \{0, 1\}^n$ passing this check, which corresponds to the unique solution $\mathbf{s}$. Trying all possibilities yields an algorithm in $O^*(2^n)$ time.

2.8 Arora–Ge algorithm

In a paper published at ICALP 2011, Arora and Ge proposed an algebraic approach to the LWE problem, which essentially amounts to expressing LWE as a system of *polynomial* equations, and then solving that system by unique linearization techniques. In the case of Binary-Error LWE, the polynomial system is a system of multivariate *quadratic* equations, which can be solved in polynomial time by linearization when the number m of samples exceeds about $n^2/2$.

More precisely, solving an instance $(A, \mathbf{b})$ of the Binary-Error LWE problem amounts to finding a vector $\mathbf{s} \in \mathbb{F}_q^n$ (which we have seen is uniquely determined) such that for $i = 1, \ldots, m$, we have:

$$b_i - \langle \mathbf{a}_i, \mathbf{s} \rangle \in \{0, 1\},$$

where the vectors $\mathbf{a}_i$ are the rows of A, and the scalars b_i the coefficients of $\mathbf{b}$. The idea of Arora and Ge is to rewrite that condition as:

$$\big(b_i - \langle \mathbf{a}_i, \mathbf{s} \rangle\big) \cdot \big(b_i - \langle \mathbf{a}_i, \mathbf{s} \rangle - 1\big) = 0,$$

which is a quadratic equation in the coefficients $s_1, \ldots, s_n$ of $\mathbf{s}$.

In general, solving a multivariate quadratic system is hard. However, it becomes easy when many equations are available. Arora and Ge propose to solve this system using a simple linearization technique: replace all the monomials appearing in the system by a new variable.

There are $\binom{n+2}{2} = (n+2)(n+1)/2$ monomials of degree at most 2. Therefore, if the number of samples m is at least $(n+2)(n+1)/2$, linearizing the quadratic system should yield a full rank linear system with high probability, and the secret $\mathbf{s}$ can be recovered by solving this linear system. This takes time $O\Big(\binom{n+2}{2}^\omega\Big) = O(n^{2\omega})$, and therefore shows that Binary-Error LWE can be solved in polynomial time given $m \approx n^2/2$ samples.

However, many applications of LWE-like problems only give out much fewer than $\Theta(n^2)$ samples. For example, public-key encryption schemes based on LWE-like problems often have a public key consisting of $O(n \log q)$ samples (or in some cases, just $O(n)$ samples). It is

therefore interesting to analyze how the complexity of Binary-Error LWE varies as the number of available samples decreases. This is the goal of the present paper: more precisely, we describe an approach to extend the Arora–Ge attack to arbitrarily many samples at the cost of an increased attack complexity.

3 Attacks on LWE with Non-uniform Binary Error

In this section, we show the attacks on LWE with non-uniform binary error. In fact, we use a quite simple algorithm working as follows:

- Step 1: Get $3n$ samples from the LWE oracle.
- Step 2: Choose $2n$ samples randomly from the 3n samples got in step1.
- Step 3: By assuming the $2n$ samples are all correct, solve the linear equation system.
- Step 4: If failed, go back to step2.

Denote p as the error rate of the non-uniform binary error. We analyze two cases:

3.1 Case 1: $p = 1/n$

Theorem (5). By applying the above algorithm, non-uniform binary-error LWE with error rate $p = 1/n$ can be attacked in polynomial time with $O(n)$ samples.

Proof. The probability that 2n samples are all error free is

$$(1 - \frac{1}{n})^{2n}$$

Considering the asymptotic behaviour,

$$\lim_{n\to\infty}(1 - \frac{1}{n})^{2n} = \frac{1}{e^2}$$

which means that step2 and step3 succeeds in finding the secret vector with constant probability and the program is expected to end in polynomial time. □

3.2 Case 2: $p = 1/n^\alpha (0 < \alpha < 1)$

The proof for case 2 is similar.

Theorem (6). By applying the above algorithm, non-uniform binary-error LWE with error rate $p = 1/n^\alpha (0 < \alpha < 1)$ can be attacked in subexponential time with $O(n)$ samples.

Proof. The probability that 2n samples are all error free is

$$\begin{aligned}&\lim_{n\to\infty}(1 - \frac{1}{n^\alpha})^{2n}\\ &= \lim_{n\to\infty}((1 - \frac{1}{n^\alpha})^{n^\alpha})^{2\cdot n^{1-\alpha}}\\ &\approx (\frac{1}{e})^{2\cdot n^{1-\alpha}}\end{aligned}$$

which means that we are expected to try step2 and step3 subexponential times before getting the secret vector. □

4 Hardness of LWE with Non-uniform Binary Error

In this section we prove the hardness of Non-uniform Binary Error. We proceed similarly to [8], by using the standard LWE assumption to construct a lossy family of functions with respect to the non-uniform input distribution.

4.1 Construction of Lossy Function Family

The basic idea of [8] is as follows:

- Construct two indistinguishable function families $\mathcal{F} = \mathrm{SIS}(m, m-n, q)$ and $\mathcal{L} = \mathrm{SIS}(l, m-n, q) \circ \mathcal{I}(m, l, \mathcal{Y})$, where $\circ$ means the composition of two functions.
- Prove $(\mathcal{L}, \mathcal{X})$ is uninvertible with respect to input distribution $\mathcal{X}$.
- Prove $(\mathcal{F}, \mathcal{X})$ is second-preimage resistant with respect to input distribution $\mathcal{X}$.
- Use the above three properties to show that $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family.
- By using Theorem 1 to show that $(\mathcal{L}, \mathcal{X})$ and $(\mathcal{F}, \mathcal{X})$ are both one-way, so $\mathrm{SIS}(m, m-n, q)$ is one-way with respect to the input distribution $\mathcal{X}$.
- By using Theorem 3 to show that $\mathrm{LWE}(m, n, q)$ is one-way with respect to the input distribution $\mathcal{X}$.

In this construction, they first proved the one-wayness of $\mathrm{SIS}(m, m-n, q)$, and then use the equivalence of $\mathrm{LWE}(m, n, q)$ and $\mathrm{SIS}(m, m-n, q)$ to prove $\mathrm{LWE}(m, n, q)$ is also one-way. There are some other work(essentially the same) [4], which directly reduces Binary-Error LWE to standard LWE without using the notation of SIS. In this paper, we stick to the SIS notation.

4.2 Bounding Statistical Uninvertibility

Their proof of uninvertibility gives bound of statistical adversary. For uniform error, it works well. Suppose that f is a function and the domain, range of f is denoted as X, Y respectively. If $y \in Y$ has several preimages, since the input distribution is uniform, the adversary can not do better than random guessing one preimage, even with unbounded computation power. However, this is not the case for a non-uniform input distribution. Suppose that the domain of f is {0, 1, 2} with probability $\{\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\}$ respectively and they all map to the same image 0. Instead of randomly guessing, the adversary can get some advantage by guessing the preimage with the highest conditional probability, so the adversary can always output 0. If guessing randomly, the adversary only has $\frac{1}{3}$ probability of correctness, but if always guessing 0, the success probability becomes $\frac{2}{3}$. The following theorem is the uniform case for bounding uninvertibility.

Theorem (7). Let $\mathcal{L}$ be a family of functions on a common domain X, and let $\mathcal{X} = U(X)$ the uniform input distribution over X. Then $(\mathcal{L}, \mathcal{X})$ is ε-uninvertible(even statistically, with respect to computationally unbounded adversaries) for $\varepsilon = \mathbb{E}_{f \leftarrow \mathcal{L}}[|f(X)|]/|X|$.

Proof. Choose a function f and random input $x \in X$. Denote $y = f(x)$, suppose that y has t preimages in the domain X. Since the input distribution is uniform, the conditional distribution given y is also uniform, so the adversary can not do better than randomly guessing one preimage. The success probability of the adversary is bounded by 1/t. Thus the total success probability is bounded by

$$\sum_{y \in f(X)} \frac{\left|f^{-1}(y)\right|}{|X|} \cdot \frac{1}{|f^{-1}(y)|} = \frac{|f(X)|}{|X|}$$

□

However, if applying this proof to non-uniform input distribution, things become more complicated. Since the input distribution is not uniform, the adversary can always output the preimage with the maximum conditional probability. The success probability is bounded by

4.3 Second Preimage Resistance

Theorem (8). For any integers m, k, q, s and set $X \subset [s]^m$, the function family $\mathrm{SIS}(m, k, q)$ is (statistically) ε-second preimage resistant with respect to the non-uniform distribution $N(X)$ for $\varepsilon = |X| \cdot (s'/q)^k$, where s' is the largest factor of q smaller than s.

Proof. Let $x \leftarrow X$ and $A \leftarrow \mathrm{SIS(m, k, q)}$ be chosen at random. We want to evaluate the probability that there exists an $\mathbf{x}' \in X \backslash \{\mathbf{x}'\}$ such that $A\mathbf{x} = A\mathbf{x}'(\mathrm{mod}\ q)$, or, equivalently, $A(\mathbf{x} - \mathbf{x}') = 0(\mathrm{mod}\ q)$. For any two distinct vectors $\mathbf{x}, \mathbf{x}' \in X$ and let $\mathbf{z} = \mathbf{x} - \mathbf{x}'$. The vectors $A\mathbf{z}(\mathrm{mod}\ q)$ is distributed uniformly at random in $(d\mathbb{Z}/q\mathbb{Z}^k)$, where $d = gcd(q, z_1 \cdots, z_m)$. All coordinates of $\mathbf{z}$ are in the range $z_i \in \{-(s-1), \ldots, (s-1)\}$ and at least one of them is nonzero. Therefore, d is at most s' and $|dZ_q^k| = (q/d)^k \geq (q/s')^k$. By using union bound(over $\mathbf{x}' \in X \backslash \{mathbfx\}$) for any $\mathbf{x}$, the probability that there is a second preimage $\mathbf{x}'$ is at most $(|X| - 1)(x'/q)^k$. □

4.4 One-wayness

Theorem (9). Let q be a modulus and let $\mathcal{X}$, $\mathcal{Y}$ be two distributions over $\mathbb{Z}^m$ and $\mathbb{Z}^l$ respectively, where $l = m - n + k$ for some $0 < k \leq n \leq m$, such that

- $\mathcal{I}(m, l, \mathcal{Y})$ is uninvertible with respect to input distribution $\mathcal{X}$.
- $\mathrm{SIS}(l, m-n, q)$ is pseudorandom with respect to input distribution $\mathcal{Y}$.
- $\mathrm{SIS}(m, m-n, q)$ is second-preimage resistant with respect to input distribution $\mathcal{X}$.

Then $\mathcal{F} = \mathrm{SIS}(m, m-n, q)$ is one-way with respect to input distribution $\mathcal{X}$.

Proof. Here we construct a lossy function family where $\mathcal{F} = \mathrm{SIS}(m, m-n, q)$ and $\mathcal{L} = \mathrm{SIS}(l, m-n, q) \circ \mathcal{I}(m, l, \mathcal{Y})$. In order to prove $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family, we need to prove three things:

- $\mathcal{L}$ and $\mathcal{F}$ is indistinguishable.
- $(\mathcal{L}, \mathcal{X})$ is uninvertible.
- $(\mathcal{F}, \mathcal{X})$ is second-preimage resistant.

First, we prove the indistinguishability of $\mathcal{L}$ and $\mathcal{F}$. Suppose there is an adversary $\mathcal{A}$ that can distinguish $\mathcal{L}$ and $\mathcal{F}$ with advantage ε. Let $U(\mathbb{Z}^{m-n})$ denote the uniform distribution over $\mathbb{Z}^{m-n}$,

Secondly,

Thirdly, the second-preimage resistance of $(\mathcal{F}, \mathcal{X})$ is proved by Theorem 8. Thus, we prove that $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family. By using Theorem 1, we prove that $\mathcal{F} = \mathrm{SIS}(m, m-n, q)$ is one-way with respect to input distribution $\mathcal{X}$. □

4.5 Instantiation

5 Conclusion

In this paper, we analyze the hardness of LWE with non-uniform binary error. On one hand, we show some attacks against LWE when the error rate is low. On the other hand, when the error rate is a constant, we reduce it to standard LWE, as long as the number of samples is strongly restricted.

References

[1] Martin R Albrecht, Rachel Player, and Sam Scott. On the concrete hardness of learning with errors. *Journal of Mathematical Cryptology*, 9(3):169–203, 2015.

[2] Sanjeev Arora and Rong Ge. New algorithms for learning in presence of errors. In *International Colloquium on Automata, Languages, and Programming*, pages 403–415. Springer, 2011.

[3] Johannes Buchmann, Florian Göpfert, Tim Güneysu, Tobias Oder, and Thomas Pöppelmann. High-performance and lightweight lattice-based public-key encryption. In *Proceedings of the 2nd ACM International Workshop on IoT Privacy, Trust, and Security*, pages 2–9. ACM, 2016.

[4] Nico Döttling and Jörn Müller-Quade. Lossy codes and a new variant of the learning-with-errors problem. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 18–34. Springer, 2013.

[5] François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th international symposium on symbolic and algebraic computation*, pages 296–303. ACM, 2014.

[6] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 1–23. Springer, 2010.

[7] Daniele Micciancio and Petros Mol. Pseudorandom knapsacks and the sample complexity of lwe search-to-decision reductions. In *Annual Cryptology Conference*, pages 465–484. Springer, 2011.

[8] Daniele Micciancio and Chris Peikert. Hardness of sis and lwe with small parameters. In *Advances in Cryptology–CRYPTO 2013*, pages 21–39. Springer, 2013.

[9] Oded Regev. The learning with errors problem. *Invited survey in CCC*, 7, 2010.