Module-LWE versus Ring-LWE, Revisited

Yang Wang[0000−0001−9274−8195] and Mingqiang Wang✉[0000−0001−9221−4230]

School of Mathematics, Shandong University, Jinan Shandong 250100, China
wyang1114@mail.sdu.edu.cn
wangmingqiang@sdu.edu.cn

Abstract. Till now, the only reduction from the module learning with errors problem (MLWE) to the ring learning with errors problem (RLWE) is given by Albrecht *et al.* in ASIACRYPT 2017. Reductions from search MLWE to search RLWE were satisfactory over power-of-2 cyclotomic fields with relative small increase of errors. However, a direct reduction from decision MLWE to decision RLWE leads to a super-polynomial increase of errors and does not work even in the most special cases- -power-of-2 cyclotomic fields. Whether we could reduce decision MLWE to decision RLWE and whether similar reductions could also work for general fields are still open. In this paper, we give a reduction from decision MLWE with module rank d and computation modulus q in worst-case to decision RLWE with modulus q^d in average-case over any cyclotomic field. Our reduction increases the LWE error rate by a small polynomial factor. As a conclusion, we obtain an efficient reduction from decision MLWE with modulus $q \approx \tilde{O}(n^{5.75})$ and error rate $\alpha \approx \tilde{O}(n^{-4.25})$ in worst-case to decision RLWE with error rate $\Gamma \approx \tilde{O}(n^{-\frac{1}{2}})$ in average-case, hence, we get a reduction from worst-case module approximate shortest independent vectors problem (SIVP_γ) with approximation parameter $\gamma \approx \tilde{O}(n^5)$ to corresponding average-case decision RLWE problems. Meanwhile, our result shows that the search variant reductions of Albrecht *et al.* could work in arbitrary cyclotomic field as well. We also give an efficient self-reduction of RLWE problems and a converse reduction from decision MLWE to module SIVP_γ over any cyclotomic field as improvements of relative results showed by Rosca *et al.* in EUROCRYPT 2018 and Langlois *et al.* [DCC 15]. Our methods can also be applied to more general algebraic fields K, as long as we can find a good enough basis of the dual $R^\vee$ of the ring of integers of K.

Keywords: Lattice-based Cryptography · Security Reduction · Cyclotomic Fields · Ring-LWE · Module-LWE

1 Introduction

Cryptographic primitives based on hard lattice problems play a key role in the area of post-quantum cryptographic researches. In the round two submissions of post-quantum cryptography called by NIST, 12 out of 26 are lattice-based and most of which are based on the learning with errors problem (LWE) and its variants. Ever since introduced by Regev [33], LWE and its variants have become fundamental problems in lattice-based cryptography. A huge amount of cryptographic primitives based on LWE and its variants have been proposed, such as public-key encryption [22, 28], key exchange protocols [2, 7, 8], digital signatures [15, 16], identity-based encryption [17, 18], pseudo-random function families [6, 11, 14, 31], watermarking [19, 20], etc.

Regev established quantum reductions from worst-case lattice problems over Euclidean lattices (such as SIVP_γ) to LWE, making LWE a versatile and very attractive ingredient for post-quantum cryptography. Soon after, Peikert [28] gave a de-quantization by proposing a reduction from the decisional approximate shortest vector problem (GapSVP_γ) to plain LWE with exponential modulus. Combining the modulus-switch techniques, Brakerski *et al.* [10] showed the classical hardness of plain LWE with quite flexible choices of parameters. Cryptographic protocols relying on plain LWE therefore enjoy the property of being provably as secure as worst-case lattice problems which is strongly suspected of being extremely hard. However, cryptography primitives based on plain LWE suffer from large key sizes (or public data), hence, they are usually inherently inefficient. This drawback stimulates people to develop more efficient LWE variants, such as the Polynomial Ring Learning with Errors problem (PLWE)[36] and the Ring Learning with Errors problem (RLWE) [23]. It has been shown that RLWE is also at least as hard as worst-case lattice problems over special classes of ideal lattices [23, 30] and cryptographic applications of RLWE generally enjoy an increase in efficiency compared with those of plain LWE, especially in the power-of 2 cyclotomic rings. But, these ideal lattices received relatively less attention than their analogues on general Euclidean lattices. Most importantly, the de-quantization reductions could not applied to RLWE problem, since GapSVP_γ problems are actually easy on ideal lattices for the involved approximation factors γ as in [28]. Though a standard and well accepted conjecture is to assume that there is no probabilistic polynomial time (PPT) algorithm (even using quantum computer) to solve hard lattice problem (for example SIVP_γ) that achieves an approximation factor which is polynomial in the lattice dimension n [26], a series of works showed that finding short vectors in ideal lattices is potentially easier on a quantum computer than in Euclidean lattices [12, 13, 32]. The length of the short vectors found in quantum polynomial time is a sub-exponential multiple of the length of the shortest vectors in ideal lattices. While, it is not known how to efficiently find such vectors in Euclidean lattices.

As alluded to above, plain LWE is known to be at least as hard as standard worst-case problems on Euclidean lattices, whereas RLWE is only known to be as hard as their restrictions to special classes of ideal lattices. The Module Learning with Errors problem (MLWE) was proposed to address shortcomings in both plain LWE and RLWE by interpolating between two [9, 21]. Module lattices have more complicated algebraic structures than ideal lattices. While, compared with Euclidean lattices, they are more structured. Thus, MLWE might be able to offer a better level of security than RLWE and still have performance advantages over plain LWE. Furthermore, MLWE has been suggested as an interesting option to hedge against potential attacks exploiting the algebraic structure of RLWE [13]. Many submissions to NIST also provided constructions based on MLWE, such as KCL, CRYSTALS-KYBER, CRYSTALS-DILITHIUM, etc. In fact, it was posed as an open problem in [21] that whether there exists reductions from MLWE to RLWE. To the best of our knowledge, till now, the only reduction is given by Albrecht *et al.* in ASIACRYPT 2017. Their reduction is an application of the main result of Brakerski *et al.* [10] in the context of MLWE. Similar technique was also used by Langlois *et al.* [21] to give a self-reduction of decision MLWE problems.

In [1], they gave a very satisfactory reduction from search MLWE to search RLWE over power-of-2 cyclotomic fields. However, it turns out that for the decision variants, even in the special power-of-2 cyclotomic fields, one can't obtain a satisfactory bounds for the reduction to preserve non-negligible advantage unless one allows for super polynomial modulus q and absolute noise in addition to negligible noise rate, as pointed in [1]. The self-reduction of decision MLWE problems [21] suffers similar problem. This is just the point, since in applications, we usually use the decision variants of MLWE/RLWE. Moreover, as stressed in [24], "powers of 2 are sparsely distributed and the desired concrete security level for an application may call for a ring dimension much smaller than the next-largest power of 2. Restricting to powers of 2 could lead to key sizes and run-times that are at least twice as large as necessary." So, both in theory and applications, it's meaningful and instructive to investigate whether we could reduce decision MLWE problem to decision RLWE problem efficiently and whether we could get similar reductions over more general fields.

1.1 Our contributions

Our first result is a reduction from worst/average-case decision MLWE problems to average-case decision RLWE problems over any cyclotomic field. We reduce decision MLWE with module rank d and computation modulus q to decision RLWE with modulus q^d in average-case, deterioriting the LWE error rate by a small polynomial factor. As a result, for any cyclotomic field $K = \mathbb{Q}(\zeta_l)$ with ζ_l the primitive l-th root of unit, we deduce that if one could solve the decision RLWE problem with error rate $\Gamma \approx \tilde{O}(n^{-\frac{1}{2}})$ and modulus q^d in average-case over K, then he can also solve the worst-case decision MLWE problem with modulus $q \approx \tilde{O}(n^{7.25})$ and error rate $\alpha \approx \tilde{O}(n^{-4.75})$ over K^d. Combining the known reduction from module SIVP_γ to decision MLWE [21, 30], we conclude a reduction from worst-case module SIVP_γ with

$\gamma \approx \tilde{O}(n^5)$ to corresponding average-case decision RLWE problems. We must stress that we constrain our discussions in cyclotomic fields because we use the powerful basis of R and the decoding basis of $R^\vee$ [24], here R is the ring of integers of K. Our methods can be extended to general number field, as long as we could find a good basis of $R^\vee$.

We then use similar method to give a self-reduction of RLWE problems. This reduction can be regarded as a modulus switch of RLWE. Roughly speaking, we could reduce decision RLWE problem with error rate α and modulus q to decision RLWE problem with modulus p and error rate $\alpha' = \alpha \cdot \frac{q}{p} \cdot poly(n)$ for some small $poly(n)$. Then our reduction could be used to reduce a decision RLWE problem with arbitrary polynomially bounded modulus q to a decision RLWE problem with some split 'well' modulus p that is relatively closed to q, for example $p = 1 \bmod l$ and $\frac{q}{p} = poly(n)$. Since decision RLWE with such modulus p can be proved hard [23], we then can prove that decision RLWE is hard for large amount of modulus q's.

Finally, we give a converse reduction from decision MLWE problem to a special case of module SIVP$_\gamma$ problem over any cyclotomic field. We prove that if one could solve the module SIVP$_\gamma$ problem in lattice $A^\perp := \{\boldsymbol{z} \in R^m : A \cdot \boldsymbol{z} = \boldsymbol{0} \bmod qR^d\}$ for some $m > d$ and $A \hookleftarrow U(R_q^{d\times m})$ with non-negligible probability, he can also solve the average-case decision MLWE problem with error rate $\alpha \approx \tilde{O}(\frac{1}{\gamma \cdot m \cdot n^3 \cdot q^{\frac{d}{m}}})$. For the usual case $d = O(1)$, by taking $m = d \cdot \log q$, we obtain a reduction from decision MLWE with error rate $\alpha \approx \tilde{O}(\frac{1}{\gamma \cdot n^3})$ to average-case module SIVP$_\gamma$ over lattice $A^\perp$, with $A \hookleftarrow U(R_q^{d \times d \log q})$. As a corollary, we obtain a reduction from worst-case module SIVP$_{\tilde{O}(\gamma \cdot n^{3.75})}$ problem over K^d to average-case SIVP$_\gamma$ problem over lattice $A^\perp$ with $A \hookleftarrow U(R_q^{d\times d \log q})$.

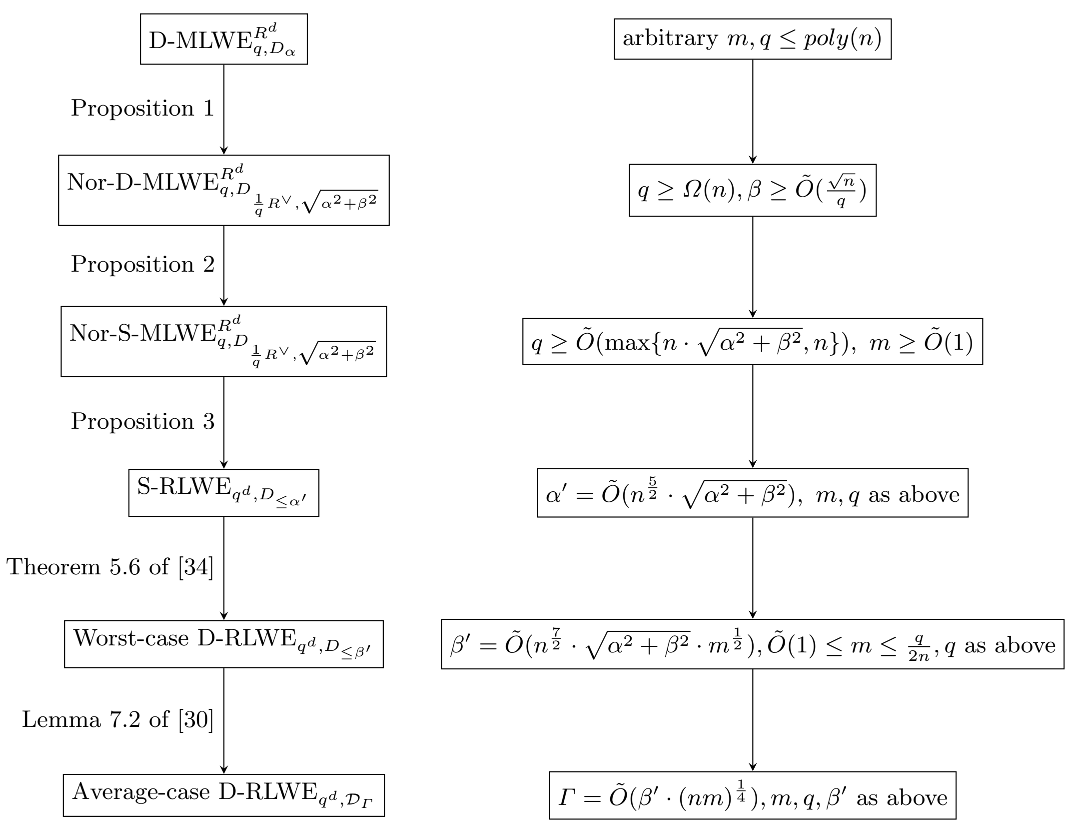

Fig. 1. Reduction road-map form decision MLWE to average-case decision RLWE

1.2 Reduction Road-map

Note that reductions from search MLWE to search RLWE in [1] are quite satisfactory. In order to get a reduction from decision MLWE to decision RLWE, a natural thought is to build some reduction from decision MLWE to search MLWE. Then, we could connect the decision MLWE and decision RLWE through the path: decision MLWE$\mapsto$search MLWE$\mapsto$search RLWE$\mapsto$decision RLWE. Many details need to be treated carefully and the reduction road-map are summarized in Figure 1.

For any cyclotomic field $K = \mathbb{Q}(\zeta_l)$, let R be the ring of integers of K, $n = \varphi(l)$ and $q \nmid l$ be some prime. We will denote D-MLWE$^{R^d}_{q,\psi}$ to be the decision MLWE problem with modulus q and error distribution ψ, denote D-RLWE$_{q,\psi}$ to be the decision RLWE problem with modulus q and error distribution ψ. Symbols for search variants are similar.

We start from D-MLWE$^{R^d}_{q,D_\alpha}$ for some continuous Gaussian distribution D_α without loss of generality [30] and reduce D-MLWE to D-RLWE step by step. If we change it to be the elliptic Gaussian distribution emerged in [21, 30], the same reduction also works with some slight modifications of error distributions. Denote m to be the number of samples we need and $D_{\leq\alpha} := \{D_{\boldsymbol{r}} : \boldsymbol{r}_k \leq \alpha$ and $\boldsymbol{r}_k = \boldsymbol{r}_{n+1-k}$ for all $k \in \{1, \cdots, \frac{n}{2}\}\}$. We first need to discretize the errors to a lattice in $K_{\mathbb{R}} = K \otimes_{\mathbb{Q}} \mathbb{R}$. We choose to discrete the errors to lattice $\frac{1}{q}R^\vee$. This can be done easily by using the fact, which is showed in [29], that for any $e \hookleftarrow D_\alpha$ and $f \hookleftarrow D_{\frac{1}{q}R^\vee - e,\beta}$ with $\beta \geq \eta_\varepsilon(\frac{1}{q}R^\vee)$, we have $e + f \overset{s}{\approx} D_{\frac{1}{q}R^\vee,\sqrt{\alpha^2+\beta^2}}$. Then we consider the normal form of corresponding MLWE problems (denoted by Nor-D-MLWE), i.e. the secret $\boldsymbol{s}$ and the error e obey the same distribution. Note that for a MLWE sample $(\boldsymbol{a}, b)$ with $b = \frac{1}{q}\boldsymbol{a}^T \cdot \boldsymbol{s} + e \bmod R^\vee$ for some secret $\boldsymbol{s} \in R_q^\vee$ and error $e \hookleftarrow D_\alpha$, we can represent it as the form $b = \boldsymbol{a}^T \cdot \boldsymbol{s'} + e$ with $\boldsymbol{s'} = \frac{1}{q}\boldsymbol{s} \in \frac{1}{q}R^\vee/R^\vee$. Hence, transformation used in [3] may also work if we could construct an invertible matrix $A \in R_q^{d\times d}$ which is consist of the $\boldsymbol{a}$ components, when given polynomial many samples. Fortunately, for $q = \Omega(n)$, we could construct such a matrix with very high probability by Lemma 9. Reduction from Nor-D-MLWE to Nor-S-MLWE is straight-forward. When given m samples, one only need to test if each component of $\boldsymbol{e'} = \boldsymbol{b} - A \cdot \boldsymbol{s} \bmod R^\vee$ has small norm, where $\boldsymbol{s}$ is the output of the Nor-S-MLWE oracle and $A \in R_q^{m\times d}$ is the matrix formed by the $\boldsymbol{a}$ components of given samples. In this reduction, we use some properties (inequality (4), which states that the smallest singular value of the matrix formed by the decoding basis is relatively large) of the decoding basis of $R^\vee$ [24] to estimate the probability $\Pr_{\boldsymbol{b}\hookleftarrow U((\frac{1}{q}R^\vee/R^\vee)^m)}[\exists \boldsymbol{s} \in (\frac{1}{q}R^\vee/R^\vee)^d : \max_{1\leq k\leq m} ||\boldsymbol{e}'_k|| < B,\ \boldsymbol{e'} = \boldsymbol{b} - A \cdot \boldsymbol{s} \bmod R^\vee]$, where B is some suitable upper-bound. For suitable B, if $\boldsymbol{b}$ is chosen uniformly at random, there will be at least one $\boldsymbol{e}'_k$ with norm lager than B. While, if $\boldsymbol{b}$ comes from MLWE distributions, the norm of all $\boldsymbol{e}'_k$ will be less than B with overwhelming probability. So, we could solve Nor-D-MLWE efficiently when given a Nor-S-MLWE oracle. As we have mentioned, reductions from search MLWE to search RLWE in [1] are acceptable, so we use similar method to reduce Nor-S-MLWE to S-RLWE. The main difference is that we need to add more error terms to amend the error distribution to elliptical Gaussian of corresponding S-RLWE problem. In this step, we need to bound the quantity $\max_{1\leq k\leq n} \frac{1}{|\sigma_k(s)|}$ for $s \hookleftarrow D_{\frac{1}{q}R^\vee,\sqrt{\alpha^2+\beta^2}}$. Our estimate also shows that the direct reduction of search variants in [1] also works for all cyclotomic fields. We then can use Theorem 5.6 of [34] to reduce S-RLWE to worst-case D-RLWE and use Lemma 7.2 of [30] to reduce worst-case D-RLWE to average-case D-RLWE with some spherical error distribution, as desired.

One may have noticed that the error parameter of D-RLWE in the above reduction is related heavily to m, meanwhile, m is bounded by $\frac{q}{2n}$. This is not very satisfactory. In applications, we may hope that m is polynomially bounded and should be independent of q. Meanwhile, the error rate should also be less dependent (or independent) of m. So, we provide a self-reduction of RLWE problem by using similar thoughts as above. More precisely, it is a modulus switch form q_1 to q_2 - - a reduction from S-RLWE$_{q_1,D_{\alpha''}}$ to S-RLWE$_{q_2,D_{\leq\alpha'''}}$ with $\alpha''' \approx \tilde{O}(\alpha'' \cdot n^{2.5} \cdot \frac{q_1}{q_2})$. In the above, we reduce D-MLWE$^{R^d}_{q,D_\alpha}$ to worst-case S-RLWE$_{q^d,D_{\leq\alpha'}}$. Then, though somewhat heuristically, for many choices of q and d, we can switch modulus q^d to some non-ramified prime p that splits 'well' (in the sense that the norm of the prime factors of pR are $poly(n)$ bounded) and $\frac{q}{p} \leq poly(n)$. Such p admits reductions from S-RLWE to D-RLWE by using the same method used in [23]. We can also reduce D-RLWE with modulus q to D-RLWE with modulus p by using similar procedure as reductions from D-MLWE to D-RLWE, too. We also remark that for many choices of q and d (for example $d = O(1)$ and $q = 1 \bmod l$), we could directly use reductions showed in [23] to reduce S-RLWE$_{q^d,D_{\leq\alpha'}}$ to average-case D-RLWE$_{q^d,D_\tau}$ for some small polynomially bounded $\tau \in \mathbb{R}$, hence reduce D-MLWE$^{R^d}_{q,D_\alpha}$ to average-case D-RLWE$_{q^d,D_\tau}$. These special cases have already covered all the usual applications, including the examples we give- -KCL, CRYSTALS-KYBER and CRYSTALS-DILITHIUM.

Reduction from D-MLWE to module SIVP$_\gamma$ is routine. It is well known that one of the classic way to solve LWE consists in solving an associated SIS instance [21, 26]. In the module context, the SIS problems over R^d (denoted by M-SIS$^{R^d}_{q,\beta}$) are defined as follows: given $A \hookleftarrow U(R_q^{m\times d})$, find a nonzero vector $\boldsymbol{z} \in R^m$ such that $\boldsymbol{z}^T \cdot A = \boldsymbol{0} \bmod qR^d$ and $||\boldsymbol{z}|| \leq \beta$ for some target norm β. We first reduce D-MLWE$^{R^d}_{q,D_\alpha}$ to M-SIS$^{R^d}_{q,\beta}$ with $\alpha \approx \tilde{O}(\frac{1}{\beta\cdot n})$. Essentially, when give a short vector $\boldsymbol{z}$ such that $\boldsymbol{z}^T \cdot A = \boldsymbol{0} \bmod qR^d$ and m sample $(A, \boldsymbol{b}) \in R_q^{m\times d} \times (K_\mathbb{R}/R^\vee)^m$, we can represent $\boldsymbol{z}^T \cdot \boldsymbol{b} \bmod R^\vee$ with respect to the decoding basis. Then, if $\boldsymbol{b}$ is distributed uniformly at random, the coefficients of $\boldsymbol{z}^T \cdot \boldsymbol{b} \bmod R^\vee$ will also be distributed randomly in the interval $(-\frac{1}{2}, \frac{1}{2}]$. On the other hand, if $\boldsymbol{b} = A\cdot \boldsymbol{s}+\boldsymbol{e}$ for some $\boldsymbol{e} \hookleftarrow D^m_\alpha$, with high probability, the coefficients of $\boldsymbol{z}^T \cdot \boldsymbol{b} \bmod R^\vee$ would be much closer to 0. Solving M-SIS$^{R^d}_{q,\beta}$ can be converted to solving module SIVP$_\gamma$ problem over the lattice $A^\perp := \{\boldsymbol{z} \in R^m : \boldsymbol{z}^T \cdot A = \boldsymbol{0} \bmod qR^d\}$ with $\beta \leq \gamma \cdot \lambda_n(A^\perp)$. By the transference theorem, $\lambda_n(A^\perp) \leq \frac{n\cdot d}{\lambda_1((A^\perp)^\vee)} \leq \frac{n\cdot d}{\lambda_\infty((A^\perp)^\vee)}$, where $(A^\perp)^\vee$ denotes the dual lattice of $A^\perp$. In fact, $(A^\perp)^\vee$ is equal to $\frac{1}{q}\{\boldsymbol{y} \in (R^\vee)^m : \exists\ \boldsymbol{s} \in (R_q^\vee)^d,\ A\cdot \boldsymbol{s} = \boldsymbol{y} \bmod q(R^\vee)^m\}$. We prove that for $A \hookleftarrow U(R_q^{m\times d})$, the lattice $(A^\perp)^\vee$ is extremely unlike to contain unusually short vectors under the infinity norm, which completes the reduction. Similar proof techniques are also used in [21, 34, 37] to obtain some kinds of ring-based leftover hash lemma and may be standard now.

We remark that we constrain our discussion in cyclotomic fields in order to use the powerful basis of R and decoding basis of $R^\vee$. Essentially, we use the property that the singular values of (one of) the basis matrix of lattice R of cyclotomic fields are well bounded [1]. We use this to discretize the errors in Subsection 3.2, to bound the probability (5) in Subsection 3.3 and to sample lattice Gaussians, whose parameter r is related to the singular values of the basis we use, in Subsection 3.4. For general algebraic field K, our reduction also works if we can find similar good basis of R. If our purpose is to get a (maybe very large) polynomially bounded reduction, a basis with a polynomially bounded singular values of R is sufficient.

1.3 Organization

We will introduce some useful definitions and results in Section 2. Reductions from D-MLWE to average-case D-RLWE are studied in Section 3. In Section 4, we will give the self-reduction of RLWE problems and some discussions. The converse reduction from D-MLWE to module SIVP$_\gamma$ is put in Section 5.

2 Preliminaries

In this section, we introduce some background results and notations.

2.1 Notations

Throughout this paper, we use $\mathbb{R}^+$ to denote the set of positive reals. Symbol $[n]$ represents the set $\{1, \cdots, n\}$ for any positive integer n. For any $M \in \mathbb{C}^{n\times n}$, we use $\mathfrak{s}_k(M)$ to denote the singular values of

[1] This also means that the singular values of (one of) the basis matrix of lattice $R^\vee$ of cyclotomic fields are well bounded.

M for $k \in [n]$. We will re-arrange singular values and assume $\mathfrak{s}_1(M) \geq \cdots \geq \mathfrak{s}_n(M)$. Matrix I_n denotes the matrix $\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}_{n\times n}$ and matrix J_n denotes the matrix $\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}_{n\times n}$. When we write $X \hookleftarrow \xi$, we mean the random variable X obeys to the distribution ξ. For a finite set S, we will use $|S|$ to denote its cardinality and $U(S)$ to denote the uniform distribution over S.

2.2 Cyclotomic Fields, Space H and Lattices

Through out this paper, we mainly consider cyclotomic fields for brevity. We now briefly introduce some basic facts about cyclotomic fields. For more details and similar results of general algebraic number fields, one can refer to [23, 34, 37].

For a cyclotomic field $K = \mathbb{Q}(\zeta)$ with $\zeta = \zeta_l$ the primitive l-th root of unity, its minimal polynomial is $\Phi_l(x) = \prod_{i|l}(x^i - 1)^{\mu(\frac{l}{i})} \in \mathbb{Z}[x]$ with degree $n = \varphi(l)$, where $\varphi(\cdot)$ denotes the Euler totient function. As usual, we set $R := \mathcal{O}_K = \mathbb{Z}[\zeta]$, which is the ring of integers of K. Then $[K : \mathbb{Q}] = n := 2\mathfrak{r}$, $K \cong \mathbb{Q}[x]/\Phi_l(x)$ and $R \cong \mathbb{Z}[x]/\Phi_l(x)$. K is Galois over $\mathbb{Q}$. We set $\mathrm{Gal}(K/\mathbb{Q}) = \{\sigma_i : \ i = 1, \cdots, n\}$ and use the canonical embedding σ on K, who maps $x \in K$ into a space $\{\sigma_i(x)\}_i \in H := \{(x_1, \cdots, x_n) \in \mathbb{C}^n : \ x_{n+1-i} = \overline{x_i}, \ \forall i \in [\mathfrak{r}]\}$ via embeddings in $\mathrm{Gal}(K/\mathbb{Q})$. H is isomorphic to $\mathbb{R}^n$ as an inner product space via the orthonormal basis $\boldsymbol{h}_{i\in[n]}$ defined as follows: for $1 \leq j \leq \mathfrak{r}$,

$$\begin{cases} \boldsymbol{h}_j = \frac{1}{\sqrt{2}}(\boldsymbol{e}_j + \boldsymbol{e}_{n+1-j}) \\ \boldsymbol{h}_{n+1-j} = \frac{\boldsymbol{i}}{\sqrt{2}}(\boldsymbol{e}_j - \boldsymbol{e}_{n+1-j}), \end{cases}$$

where $\boldsymbol{e}_j \in \mathbb{C}^n$ is the vector with 1 in its j-th coordinate and 0 elsewhere, $\boldsymbol{i}$ is the imaginary number such that $\boldsymbol{i}^2 = -1$.

The discriminant Δ_K of K is a measure of the geometry sparsity of its ring of integers. Let $\alpha_1, \cdots, \alpha_n$ represent a $\mathbb{Z}$ basis of R, we can define $\Delta_K = |(\sigma_i(\alpha_j))_{1\leq i,j\leq n}|^2$, where $|\cdot|$ represents the determinant of a matrix. In particular, the discriminant of the l-th cyclotomic number field is

$$\Delta_K = (-1)^{\frac{n}{2}} \cdot \left(\frac{l}{\prod_{p|l} p^{\frac{1}{p-1}}} \right)^n \leq n^n, \tag{1}$$

where p runs over all prime factors of l.

As in [23], we define a lattice as a discrete additive subgroup of H, which is equivalent to be a discrete additive subgroup of $\mathbb{R}^n$. The dual lattice of $\Lambda \subseteq H$ is defined as $\Lambda^\vee = \{\boldsymbol{y} \in H : \ \forall \ \boldsymbol{x} \in \Lambda, \ <\boldsymbol{x}, \overline{\boldsymbol{y}}> = \sum_{i=1}^n x_i \cdot y_i \in \mathbb{Z}\}$. One can check that this definition is actually the complex conjugate of the dual lattice as usually defined in $\mathbb{C}^n$. All of the properties of the dual lattice that we use also hold for the conjugate dual.

A fractional ideal I of K is an R-module such that $xI \subseteq R$ for some non-zero $x \in K$. So, any ideal in R (integral ideal) is also a fractional ideal. Any fractional ideal I of K is a free $\mathbb{Z}$ module of rank n. So, $\sigma(I)$ is a lattice of H, and we call $\sigma(I)$ an ideal lattice and identify I with this lattice and associate with I all the usual lattice quantities. Meanwhile, its dual is defined as $I^\vee = \{a \in K : \ \mathrm{Tr}(a \cdot I) \subseteq \mathbb{Z}\}$ [2]. Then, it is easy to verify that $(I^\vee)^\vee = I$, $I^\vee$ is a fractional ideal and $I^\vee$ embeds under σ as the dual lattice of I as defined above. Recall that we have $|\Delta_K| = \det(\sigma(R))^2$, the squared determinant of the lattice $\sigma(R)$. The algebraic norm of a non-zero integral ideal J is defined as $\mathrm{N}(J) = |R/J|$. Any fractional ideal can be represented as the quotient of two non-zero co-prime integral ideals. We can define the norm of a fractional ideal I as $\mathrm{N}(I) = \frac{\mathrm{N(J_1)}}{\mathrm{N}(J_2)}$ with $I = \frac{J_1}{J_2}$, $J_1, J_2 \subseteq R$ and $J_1 + J_2 = R$. We also have $\det(\sigma(I)) = \mathrm{N}(I) \cdot \sqrt{|\Delta_k|}$. The following lemma [29] gives upper and lower bounds on the minimum distance of an ideal lattice in l_2 norm and l_∞ norm.

[2] For any $a \in K$, we define $\mathrm{Tr}(a) = \sum_{i=1}^n \sigma_i(a)$ and $\mathrm{N}(a) = \prod_{i=1}^n \sigma_i(a)$.

Lemma 1. *For any fractional ideal I in a number field K of degree n, we have*

$$\sqrt{n} \cdot \mathrm{N}^{\frac{1}{n}}(I) \leq \lambda_1(I) \leq \sqrt{n} \cdot \mathrm{N}^{\frac{1}{n}}(I) \cdot |\Delta_K|^{\frac{1}{2n}}$$

and

$$\mathrm{N}^{\frac{1}{n}}(I) \leq \lambda_1^{\infty}(I) \leq \mathrm{N}^{\frac{1}{n}}(I) \cdot |\Delta_K|^{\frac{1}{2n}}.$$

2.3 Gaussian Distributions and Rényi Divergence

The Gaussian distribution is defined as usual. For any $s > 0$, $\boldsymbol{c} \in H$, which is taken to be $s = 1$ or $\boldsymbol{c} = 0$ when omitted, define the (spherical) Gaussian function $\rho_{s,\boldsymbol{c}} : H \to (0, 1]$ as $\rho_{s,\boldsymbol{c}}(\boldsymbol{x}) = e^{-\pi \frac{||\boldsymbol{x}-\boldsymbol{c}||^2}{s^2}}$. By normalizing this function, we obtain the continuous Gaussian probability distribution $D_{s,\boldsymbol{c}}$ of parameter s, whose density function is given by $s^{-n} \cdot \rho_{s,\boldsymbol{c}}(\boldsymbol{x})$. For a real vector $\boldsymbol{r} = (r_1, \cdots, r_n) \in (\mathbb{R}^+)^n$, we define the elliptical Gaussian distributions in the basis $\{\boldsymbol{h}_i\}_{i \leq n}$ as follows: a sample from $D_{\boldsymbol{r}}$ is given by $\sum_{i \in [n]} x_i \boldsymbol{h}_i$, where x_i is chosen independently from the Gaussian distribution D_{r_i} over $\mathbb{R}$. Note that, if we define a map $\varphi : H \to \mathbb{R}^n$ by $\varphi(\sum_{i \in [n]} x_i \boldsymbol{h}_i) = (x_1, \cdots, x_n)$, then $D_{\boldsymbol{r}}$ is also a (elliptical) Gaussian distribution over $\mathbb{R}^n$.

More generally, for some rank n matrix $B \in \mathbb{R}^{n \times n}$, we set $\Sigma = B \cdot B^T$ and say a random variable $\boldsymbol{x} \hookleftarrow D_{B,\boldsymbol{c}}$ (or $\boldsymbol{x} \hookleftarrow D_{\sqrt{\Sigma},\boldsymbol{c}}$) for some $\boldsymbol{c} \in \mathbb{R}^n$ if the density function of $\boldsymbol{x}$ is given by $\frac{1}{\sqrt{det(\Sigma)}} \rho_{B,\boldsymbol{c}}(\boldsymbol{x}) := \frac{1}{\sqrt{det(\Sigma)}} \cdot e^{-\pi(\boldsymbol{x}-\boldsymbol{c})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{c})}$. It is easy to check that if $B = \begin{pmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{pmatrix}$, then $D_B = D_{\boldsymbol{r}}$ with $\boldsymbol{r} = (r_1, \cdots, r_n) \in (\mathbb{R}^+)^n$. Distributions over H are sampled by choosing an element in $\mathbb{R}^n$ according to corresponding distributions and mapping back to H via the isomorphism $H \cong \mathbb{R}^n$. Moreover, if the element falls into the set $\sigma(K)$, we can map it back to K by using the inverse of canonical embedding efficiently.

A discrete Gaussian distribution over some n-dimensional lattice Λ and coset vector $\boldsymbol{c} \in \mathbb{R}^n$ with parameter s is denoted by $D_{\Lambda+\boldsymbol{c},s}$ with density function $\frac{\rho_s(\boldsymbol{x})}{\rho_s(\Lambda+\boldsymbol{c})}$, where $\rho_s(\Lambda + \boldsymbol{c}) = \sum_{\boldsymbol{x} \in \Lambda+\boldsymbol{c}} \rho_s(\boldsymbol{x})$. It was showed in [10] that we can sample a discrete Gaussian distribution efficiently.

Lemma 2. *There is a probabilistic polynomial-time algorithm that, given a basis B of an n-dimensional lattice $\Lambda = \mathcal{L}(B) \subseteq \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^n$ and a parameter $s \geq ||\tilde{B}|| \cdot \sqrt{\frac{\ln(2n+4)}{\pi}}$, outputs a sample distributed according to $D_{\Lambda+\boldsymbol{c},s}$.*

Here, $\tilde{B}$ is the Gram-Schmidt orthogonalization of basis $B = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ and $||\tilde{B}||$ is the length of the longest column vector in it. We will also use Rényi divergence in our reductions.

Definition 1. *For any distributions P and Q such that $Supp(P) \subseteq Supp(Q)$, the Rényi divergence of P and Q of order $a \in [1, \infty]$ is given by*

$$R_a(P||Q) = \begin{cases} e^{\sum_{x \in Supp(P)} P(x) \cdot \log \frac{P(x)}{Q(x)}} & \text{for } a = 1 \\ (\sum_{x \in Supp(P)} \frac{P(x)^a}{Q(x)^{a-1}})^{\frac{1}{a-1}} & \text{for } a \in (1, \infty) \\ \max_{x \in Supp(P)} \frac{P(x)}{Q(x)} & \text{for } a = \infty \end{cases}$$

For the case where P and Q are continuous distributions, we replace the sums by integrals and let $P(x)$ and $Q(x)$ denote probability density functions. We just give a collection of useful results of the Rényi divergence. For more details, one can refer to [1, 4].

Lemma 3. *Let $a \in [1, \infty]$ and let P, Q be distributions such that $Supp(P) \subseteq Supp(Q)$. Then we have*

- **Increasing Function of the Order**: *The function $a \mapsto R_a(P||Q)$ is nondecreasing, continuous and tends to $R_\infty(P||Q)$ as $a \mapsto \infty$.*
- **Log Positivity**: *$R_a(P||Q) \geq R_a(P||P) = 1$.*
- **Data Processing Inequality**: *$R_a(P^f||Q^f) \leq R_a(P||Q)$ for any function f, where P^f and Q^f denote the distributions induced by performing the function f on a sample from P and Q respectively.*
- **Multiplicativity**: *Let P and Q be distributions on a pair of random variables (Y_1, Y_2). Let $P_{2|1}(\cdot|y_1)$ and $Q_{2|1}(\cdot|y_1)$ denote the distributions of Y_2 under P and Q respectively given that $Y_1 = y_1$. Also, for $i \in \{1, 2\}$ denote the marginal distribution of Y_i under P resp. Q as P_i resp. Q_i. Then*
  - *$R_a(P||Q) = R_a(P_1||Q_1) \cdot R_a(P_2||Q_2)$ if Y_1 and Y_2 are independent for $a \in [1, \infty]$.*
  - *$R_a(P||Q) \leq R_\infty(P_1||Q_1) \cdot \max_{y_1 \in Supp(P_1)} R_a(P_{2|1}(\cdot|y_1)||Q_{2|1}(\cdot|y_1))$.*
- **Probability Preservation**: *Let $E \subseteq Supp(Q)$ be an arbitrary event. If $a \in (1, \infty)$, then $Q(E) \geq \frac{P(E)^{\frac{a}{a-1}}}{R_a(P||Q)}$. Furthermore, we have $Q(E) \geq \frac{P(E)}{R_\infty(P||Q)}$.*
- **Weak Triangle Inequality**: *Let P_1, P_2 and P_3 be three probability distributions such that $Supp(P_1) \subseteq Supp(P_2) \subseteq Supp(P_3)$. Then*

$$R_a(P_1||P_3) \leq \begin{cases} R_a(P_1||P_2) \cdot R_\infty(P_2||P_3) \\ R_\infty(P_1||P_2)^{\frac{a}{a-1}} \cdot R_a(P_2||P_3) & \text{if } a \in (1, \infty) \end{cases}$$

Recall that for a lattice Λ and positive real $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest real $s > 0$ such that $\rho_{\frac{1}{s}}(\Lambda^\vee \backslash \{0\}) \leq \varepsilon$. We will use the following lemmata from [5, 18, 21, 25, 27, 33].

Lemma 4. *For any real $\varepsilon > 0$ and n-dimensional lattice $\Lambda \subseteq \mathbb{R}^n$ with a set of basis B, we have $\sqrt{\frac{\ln(\frac{1}{\varepsilon})}{\pi}} \cdot \frac{1}{\lambda_1(\Lambda^\vee)} \leq \eta_\varepsilon(\Lambda) \leq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \max\{||\tilde{B}||, \lambda_n(\Lambda), \frac{1}{\lambda_1^\infty(\Lambda^\vee)}\}$.*

Lemma 5. *For any n-dimensional lattice Λ, $\sigma > 0$, $c > 0$ and $t \in (0, 1)$, we have $\frac{\rho_\sigma(\Lambda \backslash c\sqrt{n}B_n)}{\rho_\sigma(\Lambda)} \leq t^{-\frac{n}{2}} \cdot e^{-\pi \frac{(1-t)c^2 n}{\sigma^2}}$ [3]. In particular, we have $\Pr_{\boldsymbol{x} \hookleftarrow D_{\Lambda,\sigma}}[||\boldsymbol{x}|| \geq \sigma\sqrt{n}] \leq 2^{-2n}$.*

Lemma 6. *For any n-dimensional lattice Λ, $\varepsilon > 0$, $s \geq \eta_\varepsilon(\Lambda)$ and $\boldsymbol{c} \in \mathbb{R}^n$, we have*

$$\rho_s(\Lambda + \boldsymbol{c}) \in \left[\frac{1-\varepsilon}{1+\varepsilon}, 1\right] \cdot \rho_s(\Lambda).$$

Lemma 7. *Let Λ be an n-dimensional lattice, $\boldsymbol{u} \in \mathbb{R}^n$, $\boldsymbol{r} \in (\mathbb{R}^+)^n$, $\sigma > 0$ and $t_i = \sqrt{r_i^2 + \sigma^2}$ for all $i \in [n]$. Assume that $\min_i \frac{r_i \cdot \sigma}{t_i} \geq \eta_\varepsilon(\Lambda)$ for some $\varepsilon \in (0, \frac{1}{2})$. Consider the continuous distribution Y on $\mathbb{R}^n$ obtained by sampling from $D_{\Lambda+\boldsymbol{u},\boldsymbol{r}}$ and then adding a vector form $D_{\boldsymbol{\sigma}}$. Then we have $\Delta(Y, D_{\boldsymbol{t}}) \leq 4\varepsilon$ and $R_\infty(D_{\boldsymbol{t}}||Y) \leq \frac{1+\varepsilon}{1-\varepsilon}$.*

2.4 Ring-LWE and Module-LWE Problems

Let $K_\mathbb{R} := K \otimes_\mathbb{Q} \mathbb{R} \cong H$, $\mathbb{T}_{R^\vee} = K_\mathbb{R}/R^\vee$, $R_q = R/(qR)$ and $R_q^\vee = R^\vee/(qR^\vee)$ for some modulus $q \in \mathbb{Z}$. We define the Ring-LWE and Module-LWE distributions as follows.

Definition 2. *Let $M = R^d$ and ψ be some distribution over H,*

- *For $s \in R_q^\vee$, the Ring-LWE distribution $A_{q,s,\psi}$ over $R_q \times \mathbb{T}_{R^\vee}$ is (a, b) for some $a \hookleftarrow U(R_q)$ and $b = \frac{1}{q} a \cdot s + e \bmod R^\vee$ with $e \hookleftarrow \psi$.*
- *For $\boldsymbol{s} \in (R_q^\vee)^d$, the Module-LWE distribution $A_{q,\boldsymbol{s},\psi}^M$ over $R_q^d \times \mathbb{T}_{R^\vee}$ is $(\boldsymbol{a}, b)$ for some $\boldsymbol{a} \hookleftarrow U(R_q^d)$ and $b = \frac{1}{q} \sum_{k=1}^d a_k \cdot s_k + e \bmod R^\vee$ with $e \hookleftarrow \psi$.*

[3] Here, B_n denotes the unit open ball.

Now we can define the Search/Decision Ring-LWE and Module-LWE problems.

Definition 3. *Let $M = R^d$ and ψ be some distribution over H,*

- *The decision ring learning with errors problem D-RLWE$_{q,\psi}$ is to distinguish poly(n) many samples of $U(R_q \times \mathbb{T}_{R^\vee})$ from $A_{q,s,\psi}$, where $s \hookleftarrow U(R_q^\vee)$. The search variant S-RLWE$_{q,\psi}$ is to find the secret s with poly(n) many samples from $A_{q,s,\psi}$ for some arbitrary $s \in R_q^\vee$.*
- *The decision module learning with errors problem D-MLWE$^M_{q,\psi}$ is to distinguish poly(n) many samples of $U(R_q^d \times \mathbb{T}_{R^\vee})$ from $A^M_{q,\boldsymbol{s},\psi}$, where $\boldsymbol{s} \hookleftarrow U((R_q^\vee)^d)$. The search variant S-MLWE$^M_{q,\psi}$ is to find the secret $\boldsymbol{s}$ with poly(n) many samples from $A^M_{q,\boldsymbol{s},\psi}$ for some arbitrary $\boldsymbol{s} \in (R_q^\vee)^d$.*

Usually, the error distribution ψ may be chosen from a family of distributions Ψ over H. Let's take the Ring-LWE problem for an example. When the error distribution ψ is sampled from a family of distributions Ψ over $K_\mathbb{R}$, we call an algorithm solve the worst-case search (or decision) problems if it solves corresponding problems with probability ≈ 1 with the pair $(s, \psi) \in R_q^\vee \times \Psi$ arbitrary. Correspondingly, we call an algorithm solve the average-case problems if it solves corresponding problems with a non-negligible probability with the pair $(s, \psi) \hookleftarrow U(R_q^\vee) \times \mathcal{D}$ for some distribution $\mathcal{D}$ over Ψ[4]. The detailed definition of $\mathcal{D}$, which is denoted by Υ_α, in the worst-case to average-case reductions of corresponding LWE problems can be found in [21, 23, 30]. We just remark that ψ can be modified to be some spherical Gaussian distribution over $K_\mathbb{R}$ [21, 30]. Also, in the followings, we will reduce D-MLWE problems with spherical error distribution to average-case D-RLWE problems with some other spherical error distribution for brevity. So, we just use a single error distribution to define corresponding problems.

In the rest of this paper, we will use $D_{\alpha' \le \alpha}$ to denote the set of elliptical Gaussian distributions $D_{\boldsymbol{r}}$ with $\alpha' \le r_i \le \alpha$. We write $D_{\le \alpha}$ when $\alpha' = 0$. Meanwhile, we assume $\psi = D_\alpha$ without loss of generality, since we can reduce worst-case lattice problems to corresponding decision variant problems with some appropriate spherical error distribution [21, 23, 30]. We will also use the SIVP problems over rings and modules, so we give definition of SIVP prblem briefly.

Definition 4. *For an approximation factor $\gamma = \gamma(n) \ge 1$, the SIVP$_\gamma$ problem is: given a full-rank lattice Λ of dimension n, output n linearly independent lattice vectors of length at most $\gamma \cdot \lambda_n(\Lambda)$.*

2.5 Basis for R and $R^\vee$ in Cyclotomic Fields

In some of our reductions, we hope that the matrices whose columns are consisted of the basis of R or $R^\vee$ have smaller $\mathfrak{s}_1$ and larger $\mathfrak{s}_n$. In cyclotomic fields, there are good bases of R and $R^\vee$ with very nice magnitudes of singular values. So, we introduce the powerful basis and the decoding basis as in [24]. We set τ be the automorphism of K that maps ζ_l to $\zeta_l^{-1} = \zeta_l^{l-1}$, under the canonical embedding it corresponds to complex conjugation $\sigma(\tau(a)) = \overline{\sigma(a)}$.

Definition 5. *The Powerful basis $\overrightarrow{p}$ of $K = \mathbb{Q}(\zeta_l)$ and $R = \mathbb{Z}[\zeta_l]$ is defined as follows:*

- *For a prime power l, define $\overrightarrow{p}$ to be the power basis $(\zeta_l^j)_{(j \in \{0,1,\cdots,n-1\})}$, treated as a vector over $R \subseteq K$.*
- *For l having prime-power factorization $l = \prod l_k = \prod p_k^{\alpha_k}$, define $\overrightarrow{p} = \otimes_k \overrightarrow{p_k}$, the tensor product of the power basis $\overrightarrow{p_k}$ of each $K_k = \mathbb{Q}(\zeta_{l_k})$.*

The Decoding basis of $R^\vee$ is $\overrightarrow{d} = \tau(\overrightarrow{p})^\vee$, the dual of the conjugate of the powerful basis $\overrightarrow{p}$.

Also note that $\tau(\overrightarrow{p})$ is a $\mathbb{Z}$-basis of R. Different bases of R (or $R^\vee$) are connected by some unimodular matrice, hence the spectral norm (i.e. the $\mathfrak{s}_1$) may have different magnitudes. The following lemma comes from [24], which shows the estimates of $\mathfrak{s}_1(\sigma(\overrightarrow{p}))$ and $\mathfrak{s}_n(\sigma(\overrightarrow{p}))$. Define $rad(l) = \prod_{p|l} p$ and $\hat{l} = \begin{cases} l, & \text{if } l \text{ is odd,} \\ \frac{l}{2}, & \text{if } l \text{ is even.} \end{cases}$

[4] We also classify the cases, where the secret $s \hookleftarrow U(R_q^\vee)$ and the error distribution $\psi \in \Psi$ is arbitrary, into worst-case variants of corresponding problems.

Lemma 8. *We have* $\mathfrak{s}_1(\sigma(\overrightarrow{p})) = \sqrt{\hat{l}}$, $\mathfrak{s}_n(\sigma(\overrightarrow{p})) = \sqrt{\frac{l}{rad(l)}}$, $||\sigma(\overrightarrow{p})_i||_\infty = 1$ *and* $||\sigma(\overrightarrow{p})_i|| = \sqrt{n}$ *for all* $i = 1, \cdots, n$.

We also need the estimates of $\mathfrak{s}_1(\sigma(\overrightarrow{d}))$ and $\mathfrak{s}_n(\sigma(\overrightarrow{d}))$. Assume that $\sigma(\overrightarrow{p}) = T$, Lemma 8 shows that $\mathfrak{s}_1(T) = \sqrt{\hat{l}}$ and $\mathfrak{s}_n(T) = \sqrt{\frac{l}{rad(l)}}$. By the definitions of $\overrightarrow{d}$ and the dual ideal, an easy computation shows that $\sigma(\overrightarrow{d}) = (T^*)^{-1}$. Hence we have $\mathfrak{s}_n(\sigma(\overrightarrow{d})) = \frac{1}{\sqrt{\hat{l}}}$, $\mathfrak{s}_1(\sigma(\overrightarrow{d})) = \sqrt{\frac{rad(l)}{l}}$. Moreover, one can similarly deduce that $||\sigma(\overrightarrow{d})_i|| \le \sqrt{\frac{rad(l)}{l}}$ for all $i = 1, 2, \cdots, n$. The following definition is also useful.

Definition 6. *Given a basis B of a fractional ideal J, for any $x \in J$ with $x = x_1 b_1 + \cdots + x_n b_n$, the B-coefficient embedding of x is defined as the vector $(x_1, \cdots, x_n)$ and the B-coefficient embedding norm of x is defined as $||x||_B^c = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$.*

If we represent $x \in R$ (or $R^\vee$) with respect to the powerful basis (or decoding basis), we have

$$\sqrt{\frac{l}{rad(l)}} \cdot ||x||^c_{\sigma(\overrightarrow{p})} \le ||\sigma(x)|| \le \sqrt{\hat{l}} \cdot ||x||^c_{\sigma(\overrightarrow{p})}, \qquad for\ x \in\ R, \tag{2}$$

and

$$\frac{1}{\sqrt{\hat{l}}} \cdot ||x||^c_{\sigma(\overrightarrow{d})} \le ||\sigma(x)|| \le \sqrt{\frac{rad(l)}{l}} \cdot ||x||^c_{\sigma(\overrightarrow{d})}, \qquad for\ x \in\ R^\vee. \tag{3}$$

We will omit the subscript $\sigma(\overrightarrow{d})$ or $\sigma(\overrightarrow{p})$ in the following applications.

3 Reductions form D-MLWE to D-RLWE

In this section, we shall reduce D-MLWE problems to average-case D-RLWE problems step by step.

3.1 Actions of matrices on R^d

In this subsection, we shall introduce some facts of maps induced by matrices in $R^{d' \times d}$, which will be helpful for us to under the transformations in the following reductions.

Assume that matrix $G \in R^{d' \times d} : R^d \mapsto R^{d'}$ induces a map, we consider the corresponding map $G_H : \sigma(R^d) \mapsto \sigma(R^{d'})$, i.e. for any $\boldsymbol{x} = (x_1, \cdots, x_d)^T \in R^d$, we require that $\sigma(\boldsymbol{y}) = G_H \cdot \sigma(\boldsymbol{x}) \in H^{d'}$, where $\boldsymbol{y} = G \cdot \boldsymbol{x}$. If

$$G = \begin{pmatrix} g_{1,1} & \cdots & g_{1,d} \\ \vdots & & \vdots \\ g_{d',1} & \cdots & g_{d',d} \end{pmatrix},$$

we define

$$G_H = \begin{pmatrix} \sigma_1(g_{1,1}) & & & \ldots\ \sigma_1(g_{1,d}) & & \\ & \ddots & & & \ddots & \\ & & \sigma_n(g_{1,1}) & \ldots & & \sigma_n(g_{1,d}) \\ & \vdots & & & \vdots & \\ \sigma_1(g_{d',1}) & & & \ldots\ \sigma_1(g_{d',d}) & & \\ & \ddots & & & \ddots & \\ & & \sigma_n(g_{d',1}) & \ldots & & \sigma_n(g_{d',d}) \end{pmatrix}.$$

Then, it is easy to verify that $\sigma(\boldsymbol{y}) = G_H \cdot \sigma(\boldsymbol{x})$. The same calculation shows that the map $\sigma_H : R^{d\times d} \mapsto \mathbb{C}^{nd\times nd}$ given by $\sigma_H(A) = A_H$ defined as above is a ring homomorphism. In fact, for any $A \in R^{d_1\times d_2}$ and $B \in R^{d_2\times d_3}$ with $C = A \cdot B \in R^{d_1\times d_3}$, we have $A_H \cdot B_H = C_H$. Hence, $A \in R^{d\times d} \subseteq \mathbb{C}^{d\times d}$ is invertible if and only if $A_H \in \mathbb{C}^{nd\times nd}$ is invertible, since $I_H = I_{nd}$.

Assume further that $\varphi : K \mapsto \mathbb{R}^n$ is the composite of the canonical embedding σ and the isomorphism $H \cong \mathbb{R}^n$, we now decide the corresponding matrix $G_{\mathbb{R}}$ of G such that $\varphi(\boldsymbol{y}) = G_{\mathbb{R}} \cdot \varphi(\boldsymbol{x})$. For any element $x \in K$, we have $\varphi(x) = U \cdot \sigma(x)$ with $U = \begin{pmatrix} \frac{1}{\sqrt{2}} \cdot I_r & \frac{1}{\sqrt{2}} \cdot J_r \\ -\frac{\boldsymbol{i}}{\sqrt{2}} \cdot J_r & \frac{\boldsymbol{i}}{\sqrt{2}} \cdot I_r \end{pmatrix}$ (note that $U^{-1} = U^*$). Hence,

$$\varphi(\boldsymbol{y}) = \begin{pmatrix} U & & \\ & \ddots & \\ & & U \end{pmatrix} \cdot \sigma(\boldsymbol{y}) = \begin{pmatrix} U & & \\ & \ddots & \\ & & U \end{pmatrix} \cdot G_H \cdot \sigma(\boldsymbol{x}) = \begin{pmatrix} U & & \\ & \ddots & \\ & & U \end{pmatrix} \cdot G_H \cdot \begin{pmatrix} U^{-1} & & \\ & \ddots & \\ & & U^{-1} \end{pmatrix} \cdot \varphi(\boldsymbol{x}),$$

which implies that $G_{\mathbb{R}} = \begin{pmatrix} U & & \\ & \ddots & \\ & & U \end{pmatrix} \cdot G_H \cdot \begin{pmatrix} U^{-1} & & \\ & \ddots & \\ & & U^{-1} \end{pmatrix}$. Moreover, we also have $G \in R^{d\times d}$ is invertible if and only if $G_H \in \mathbb{C}^{nd\times nd}$ is invertible, and if and only if $G_{\mathbb{R}} \in \mathbb{R}^{nd\times nd}$ is invertible.

Addition and multiplication of field elements are carried out component-wise in space H, i.e. $\sigma(x\cdot y) = \sigma(x) \cdot \sigma(y)$ for any $x, y \in K$. While multiplication is not component-wise for φ in $\mathbb{R}^n$. In fact, we have $\varphi(x \cdot y) = x_{\mathbb{R}} \cdot \varphi(y) = y_{\mathbb{R}} \cdot \varphi(x)$, where $x_{\mathbb{R}} = U \cdot x_H \cdot U^{-1}$ and $x_H = \begin{pmatrix} \sigma_1(x) & & \\ & \ddots & \\ & & \sigma_n(x) \end{pmatrix}$. Note that $x_{\mathbb{R}} \cdot x_{\mathbb{R}}^T = x_{\mathbb{R}} \cdot x_{\mathbb{R}}^* = U \cdot x_H \cdot x_H^* \cdot U^{-1} = \begin{pmatrix} |\sigma_1(x)|^2 & & \\ & \ddots & \\ & & |\sigma_n(x)|^2 \end{pmatrix}$, the singular values of $x_{\mathbb{R}}$ are precisely given by $|\sigma_i(x)|$ for $i \in [n]$. Then, for any $s \in K$, if $x \hookleftarrow D_B$ for some nonsingular matrix B with $\Sigma = B \cdot B^T$, then $s \cdot x \hookleftarrow D_{\sqrt{\Sigma'}}$ with $\Sigma' = s_{\mathbb{R}} \cdot \Sigma \cdot s_{\mathbb{R}}^T$. In particular, if $B = \begin{pmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{pmatrix}$ with $r_k = r_{n+1-k}$ for all $k \in [\frac{n}{2}]$, then $s \cdot x \hookleftarrow D_{B'}$ with $B' = \begin{pmatrix} r_1 \cdot |\sigma_1(s)| & & \\ & \ddots & \\ & & r_n \cdot |\sigma_n(s)| \end{pmatrix}$.

Suppose q is a prime which does not ramify in R (equivalently, $q \nmid l$ in our settings), meanwhile, $qR = \mathfrak{q}_1 \cdots \mathfrak{q}_{\mathfrak{g}}$ with $\mathfrak{g} \cdot \mathfrak{f} = n$. We have $N(\mathfrak{q}_i) = q^{\mathfrak{f}}$ and $R_q \cong R/\mathfrak{q}_1 \times \cdots \times R/\mathfrak{q}_{\mathfrak{g}}$. The following lemma is useful for us to get some results about the normal form of module LWE problems where the secret distribution is a discretized version of the error distribution. Its proof is somewhat fundamental but fussy, so we put it in Appendix A. We call a set of vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\} \in R_q^d$ is R_q-linearly independent if $x_1 \cdot \boldsymbol{a}_1 + \cdots + x_k \cdot \boldsymbol{a}_k = 0 \bmod qR$ implies $x_1 = \cdots = x_k = 0$, where $x_i \in R_q$ for $i \in [k]$. Also, note that the determinant function of square matrices over the ring R_q is well defined.

Lemma 9. *For any $i \in \{0, \cdots, d-1\}$ and R_q-linearly independent vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i \in R_q^d$, the probability that sample a vector $\boldsymbol{b} \hookleftarrow U(R_q^d)$ such that $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i, \boldsymbol{b}$ are R_q-linearly independent is at least $1 - \frac{\mathfrak{g}}{q^{\mathfrak{f}}} \geq 1 - \frac{n}{q}$.*

Remark 1. Results in this Subsection can be easily modified to general algebraic number fields. The only difference is that in general fields, the $\{\mathfrak{f}, \mathfrak{g}\}$'s may not equal to each other. However, similar deduction implies that we still have the same lower bound $1 - \frac{n}{q}$ as in Lemma 9.

3.2 Hardness of Normal Form of Decision MLWE

In this subsection, we shall discuss the hardness of normal form of D-MLWE. However, in order to make multiplication well-defined in K, we need to discretize the errors. The discretized distribution would also be used in Subsection 3.3.

Given a lattice $\Lambda \subseteq H$ and a point $\boldsymbol{x} \in H$, we want to discretize $\boldsymbol{x}$ to a point $\lfloor \boldsymbol{x} \rceil_\Lambda \in \Lambda$. To do so, we can sample a point $\boldsymbol{f} \in \Lambda - \boldsymbol{x}$ and set $\lfloor \boldsymbol{x} \rceil_\Lambda = \boldsymbol{f} + \boldsymbol{x}$. The only requirement is that $\boldsymbol{f}$ can be chosen efficiently and dependent only on the coset $\Lambda - \boldsymbol{x}$. We call such a procedure valid discretization as in [24]. Then, it is easy to check that $\lfloor \boldsymbol{z} + \boldsymbol{x} \rceil_\Lambda = \boldsymbol{z} + \lfloor \boldsymbol{x} \rceil_\Lambda$ for any valid discretization and $\boldsymbol{z} \in \Lambda$.

Assume that D-MLWE$^{R^d}_{q,D_\alpha}$ is hard for some distribution D_α over $K_\mathbb{R}/R^\vee$, let $\phi = \lfloor D_\alpha \rceil_{\frac{1}{q}R^\vee}$ for some valid discretization $\lfloor \cdot \rceil_{\frac{1}{q}R^\vee}$. We can show that D-MLWE$^{R^d}_{q,\phi}$ is also hard by using the same method as in [24]. We just state the following lemma and its proof is put in Appendix B.

Lemma 10. *There is a transformation that given a pair $(\boldsymbol{a}', b') \in R_q^d \times K_\mathbb{R}/R^\vee$, outputs a pair $(\boldsymbol{a}, b) \in R_q^d \times \frac{1}{q}R^\vee/R^\vee$ with the following guarantees: if the input pair is uniformly distributed, then so is the output pair; and if the input pair is distributed according to the MLWE distribution $A^{R^d}_{q,\boldsymbol{s},D_\alpha}$, then the output pair is distributed according to $A^{R^d}_{q,\boldsymbol{s},\phi}$.*

Next, we show that D-MLWE is also hard when the secret $\boldsymbol{s}$ is distributed as the error e. We denote this kind of D-MLWE problem by Nor-D-MLWE (whose corresponding distribution is denoted by $A^{R^d*}_{q,\boldsymbol{s},\phi}$), i.e. a sample of $A^{R^d*}_{q,\boldsymbol{s},\phi}$ is of the form $(\boldsymbol{a}, b)$ with $\boldsymbol{a} \hookleftarrow U(R_q)$ and $b = \boldsymbol{a}^T \cdot \boldsymbol{s} + e \bmod R^\vee$, where $s_i, e \hookleftarrow \phi = \lfloor D_\alpha \rceil_{\frac{1}{q}R^\vee}$ for $i \in [d]$.

For $A \hookleftarrow U(R_q^{d\times d})$, Lemma 9 shows that with probability larger than $(1 - \frac{\mathfrak{g}}{q^\mathfrak{f}})^d$, A is invertible mod qR. When $q^\mathfrak{f} = O(d \cdot \mathfrak{g})$, this is a non-negligible probability. In fact, for any $q \geq 2n$, with polynomial many samples, we could find an invertible matrix A with probability ≈ 1. Assume we have d samples of the form $(A, \boldsymbol{b}) \in R_q^{d\times d} \times (\frac{1}{q}R^\vee/R^\vee)^d$, where A is invertible and $\boldsymbol{b} = A \cdot \boldsymbol{s}' + \boldsymbol{e}$ for some $\boldsymbol{s}', \boldsymbol{e} \in (\frac{1}{q}R^\vee/R^\vee)^d$ with $\boldsymbol{s}' = \frac{1}{q}\boldsymbol{s}$. Note that given A is equivalent to given A_H, the b-component of MLWE distribution is $\sigma(\boldsymbol{b}) = A_H \cdot \sigma(\boldsymbol{s}') + \sigma(\boldsymbol{e})$, i.e.

$$\sigma(\boldsymbol{b}) = \begin{bmatrix} \sigma_1(a_{1,1}) & & & \ldots\, \sigma_1(a_{1,d}) & & \\ & \ddots & & & \ddots & \\ & & \sigma_n(a_{1,1}) & \ldots & & \sigma_n(a_{1,d}) \\ & \vdots & & & \vdots & \\ \sigma_1(a_{d,1}) & & & \ldots\, \sigma_1(a_{d,d}) & & \\ & \ddots & & & \ddots & \\ & & \sigma_n(a_{d,1}) & \ldots & & \sigma_n(a_{d,d}) \end{bmatrix} \cdot \begin{pmatrix} \sigma_1(s_1') \\ \vdots \\ \sigma_n(s_1') \\ \vdots \\ \sigma_1(s_d') \\ \vdots \\ \sigma_n(s_d') \end{pmatrix} + \begin{pmatrix} \sigma_1(e_1) \\ \vdots \\ \sigma_n(e_1) \\ \vdots \\ \sigma_1(e_d) \\ \vdots \\ \sigma_n(e_d) \end{pmatrix}.$$

For another new sample $(\boldsymbol{a}, b) \hookleftarrow A^{R^d}_{q,\boldsymbol{s},\phi}$, we set $(\boldsymbol{a}', b')$ as $(\boldsymbol{a}')^T = -\boldsymbol{a}^T \cdot A^{-1} \bmod qR$ and $b' = b + (\boldsymbol{a}')^T \cdot \boldsymbol{b} \bmod R^\vee$. Then, we have

$$\begin{aligned} b' &= b + (\boldsymbol{a}')^T \cdot \boldsymbol{b} \\ &= \frac{1}{q}\boldsymbol{a}^T \cdot \boldsymbol{s} + e - \boldsymbol{a}^T \cdot A^{-1} \cdot (A \cdot \boldsymbol{s}' + \boldsymbol{e}) \\ &= (\boldsymbol{a}')^T \cdot \boldsymbol{e} + e, \end{aligned}$$

where the components of $\boldsymbol{e}$ and e obey the same distribution ϕ. Recall that $(A_H)^{-1} = (A^{-1})_H$, equivalently, we have

$$\begin{aligned}\sigma(b') &= \sigma(b) + \sigma(\boldsymbol{a}')^T \cdot \sigma(\boldsymbol{b}) \\ &= \sigma(\boldsymbol{a})^T \cdot \sigma(\boldsymbol{s}') + \sigma(e) - \sigma(\boldsymbol{a})^T \cdot A_H^{-1} \cdot (A_H \cdot \sigma(\boldsymbol{s}') + \sigma(\boldsymbol{e})) \\ &= \sigma(\boldsymbol{a}')^T \cdot \sigma(\boldsymbol{e}) + \sigma(e).\end{aligned}$$

It is easy to see that if $(\boldsymbol{a}, b) \hookleftarrow U(R_q^d \times \frac{1}{q}R^\vee/R^\vee)$, so is $(\boldsymbol{a}', b')$. Combining all above discussions, we get the following proposition.

Proposition 1. *There is a PPT reduction from* D-MLWE$^{R^d}_{q,D_\alpha}$ *to* Nor-D-MLWE$^{R^d}_{q,\phi}$ *for* $q \geq 2n$.

We mainly consider the following discretization in this paper: we use results showed in [29] to discretize $\boldsymbol{e}$ to a discrete Gaussian distribution. Note that $\varphi(\overrightarrow{d}) = U \cdot \sigma(\overrightarrow{d})$, so $\varphi(\overrightarrow{d})^T \cdot \varphi(\overrightarrow{d}) = \varphi(\overrightarrow{d})^* \cdot \varphi(\overrightarrow{d}) = \sigma(\overrightarrow{d})^* \cdot \sigma(\overrightarrow{d})$, which implies $\mathfrak{s}_1(\varphi(\overrightarrow{d})) = \mathfrak{s}_1(\sigma(\overrightarrow{d})) = \sqrt{\frac{rad(l)}{l}}$. Hence, if we set $\Lambda = \sigma(\frac{1}{q}R^\vee)$ and use the basis $\frac{1}{q}\overrightarrow{d}$, for any $\boldsymbol{c} \in H \cong \mathbb{R}^n$ and $\beta > \omega(\sqrt{\log n}) \cdot \frac{1}{q}\sqrt{\frac{rad(l)}{l}}$, we can use Algorithm 2 of [29] to output a vector $\boldsymbol{x}$ drawn from a distribution statistically close to $D_{\Lambda+\boldsymbol{c},\beta}$ in probabilistic polynomial time. We also have

$$\begin{aligned}
\eta_\varepsilon(\frac{1}{q}R^\vee) &\leq \frac{1}{q}\sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \lambda_n(R^\vee) && \text{(By Lemma 4)} \\
&= \frac{1}{q}\sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \lambda_1(R^\vee) && (\lambda_n = \lambda_1 \text{ in cyclotomic fields}) \\
&\leq \frac{1}{q}\sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \sqrt{n} \cdot N^{\frac{1}{n}}(R^\vee) \cdot |\Delta_K|^{\frac{1}{2n}} && \text{(By Lemma 1)} \\
&= \frac{1}{q}\sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \sqrt{n} \cdot |\Delta_K|^{-\frac{1}{2n}} && (N(R^\vee) = |\Delta_k|^{-1}) \\
&= \frac{1}{q}\sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \sqrt{n} \cdot (\frac{\prod_{p|l} p^{\frac{1}{p-1}}}{l})^{\frac{1}{2}} && \text{(By equation (1))} \\
&\leq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \frac{\sqrt{n}}{q}.
\end{aligned}$$

Note that the last inequality is rather loose. For any $\beta \geq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \frac{\sqrt{n}}{q}$, Theorem 3.1 of [29] shows that the distribution $\lfloor \boldsymbol{e} \rceil_\Lambda = \boldsymbol{e} + \boldsymbol{f}$ with $\boldsymbol{e} \hookleftarrow D_\alpha$ and $\boldsymbol{f} \hookleftarrow D_{\Lambda-\boldsymbol{e},\beta}$ is statistically close to $D_{\Lambda,\sqrt{\alpha^2+\beta^2}}$. In the rest of this paper, we will set $\phi = D_{\frac{1}{q}R^\vee,\sqrt{\alpha^2+\beta^2}}$ with $\beta \geq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \frac{\sqrt{2n}}{q} \geq \sqrt{2} \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$ [5], unless we specify it with other values.

Note that by the transference theorem [5] and Lemma 1, we have $\lambda_n(R^\vee) \leq \frac{n}{\lambda_1(R)} \leq \sqrt{n}$. So, we still have $\eta_\varepsilon(\frac{1}{q}R^\vee) \leq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \frac{\sqrt{n}}{q}$ for general number fields. We can save a factor $\approx \sqrt{n}$ in the above long inequalities for some special cyclotomic fields (e.g. $K = \mathbb{Q}(\zeta_l)$ with l a large prime power).

Results in this Subsection can also be extended to general number fields as long as we can find a good enough basis of $R^\vee$, since we use $\frac{1}{q}\overrightarrow{d}$ to sample a lattice Gaussian distribution. However, this constrain depends on the discretization we used and can be avoided by using other discretizations. For example, one

[5] The factor $\sqrt{2}$ is used in Subsection 3.4 for convenience.

can use the "coordinate-wise randomized rounding" or the simplest rounding [24] to obtain a Gaussian-like distribution. The adverse impact is that the error analysis in Subsection 3.4 would become much more complicated.

3.3 Reduction from Nor-D-MLWE to Nor-S-MLWE

We give a reduction from Nor-D-MLWE$^{R^d}_{q,\phi}$ to Nor-S-MLWE$^{R^d}_{q,\phi}$ in this subsection. Recall that, for cyclotomic field $K = \mathbb{Q}(\zeta_l)$, if we represent $x \in R^\vee$ with respect to the decoding basis, we have

$$\frac{1}{\sqrt{\hat{l}}} \cdot ||x||^c_{\sigma(\vec{d})} \leq ||\sigma(x)|| \leq \sqrt{\frac{rad(l)}{l}} \cdot ||x||^c_{\sigma(\vec{d})}, \qquad for\ x \in\ R^\vee. \tag{4}$$

Note that, by Lemma 5, ϕ is $(\sqrt{\alpha^2+\beta^2}\cdot\sqrt{n}, 2^{-2n})$ bound, i.e. $\Pr_{x\hookleftarrow\phi}[||x|| \geq \sqrt{\alpha^2+\beta^2}\cdot\sqrt{n}] \leq 2^{-2n}$. We also represent m Nor-D-MLWE samples as the form $(A, \boldsymbol{b})$, where $A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)^T \in R_q^{m\times d}$ and $\boldsymbol{b} \in (\frac{1}{q}R^\vee/R^\vee)^m$.

Now assume we have an oracle $\mathcal{O}$ for solving Nor-S-MLWE problem with advantage ε when given m samples. When we get m samples $(A, \boldsymbol{b}) \in R_q^{m\times d} \times (\frac{1}{q}R^\vee/R^\vee)^m$, we give it to the Nor-S-MLWE oracle $\mathcal{O}$ and get some $\boldsymbol{s} \in (\frac{1}{q}R^\vee/R^\vee)^d$ with probability ε. Then we compute $\boldsymbol{e} = \boldsymbol{b} - A\cdot\boldsymbol{s} \bmod R^\vee$ and $N = ||\boldsymbol{e}||^\infty$, where $||\boldsymbol{e}||^\infty = \max_{i\in[m]}||e_i||$. We output 1 if and only if $N < B := \sqrt{\alpha^2+\beta^2}\cdot\sqrt{n}$.

If $(A, \boldsymbol{b}) \hookleftarrow A^{R^d*}_{q,\boldsymbol{s},\phi}$, then the probability we output 1 is large than $\varepsilon - \Pr_{\boldsymbol{e}\hookleftarrow\phi^m}(||\boldsymbol{e}||^\infty \geq B)$. If $(A, \boldsymbol{b})$ is uniformly distributed, the probability we output 1 is less than

$$\Pr_{\boldsymbol{b}\hookleftarrow U((\frac{1}{q}R^\vee/R^\vee)^m)}[\exists \boldsymbol{s} \in (\frac{1}{q}R^\vee/R^\vee)^d : ||\boldsymbol{b} - A\cdot\boldsymbol{s}||^\infty < B]. \tag{5}$$

Hence, the distinguishing advantage we have is larger than $\varepsilon - \Pr_{\boldsymbol{e}\hookleftarrow\phi^m}(||\boldsymbol{e}||^\infty \geq B) - \Pr_{\boldsymbol{b}\hookleftarrow U((\frac{1}{q}R^\vee/R^\vee)^m)}[\exists \boldsymbol{s} \in (\frac{1}{q}R^\vee/R^\vee)^d : ||\boldsymbol{b} - A\cdot\boldsymbol{s}||^\infty < B]$.

Since ϕ is a (B, δ) bound distribution with $\delta = 2^{-2n}$, we have $\Pr_{\boldsymbol{e}\hookleftarrow\phi^m}(||\boldsymbol{e}||^\infty \geq B) \leq m\cdot\delta$. Also, note that $||x||^c_\infty \leq ||x||^c \leq \sqrt{\hat{l}}\cdot||x||$ for any $x \in \frac{1}{q}R^\vee/R^\vee$, we have

$$\Pr_{\boldsymbol{b}\hookleftarrow U((\frac{1}{q}R^\vee/R^\vee)^m)}[\exists \boldsymbol{s} \in (\frac{1}{q}R^\vee/R^\vee)^d : ||\boldsymbol{b} - A\cdot\boldsymbol{s}||^\infty < B] \leq q^{nd}\cdot\frac{(2\sqrt{\hat{l}}\cdot B)^{mn}}{q^{nm}}.$$

Note that $q^{nd}\cdot\frac{(2\sqrt{\hat{l}}\cdot B)^{mn}}{q^{nm}} = (2\sqrt{\hat{l}}\cdot B\cdot q^{\frac{d}{m}-1})^{mn}$. We now decide the conditions to bound $(2\sqrt{\hat{l}}\cdot B\cdot q^{\frac{d}{m}-1})^{mn} \leq \delta < 2^{-2n}$. For $q > 2\sqrt{\hat{l}}B$, this is equivalent to $(2\sqrt{\hat{l}}B\cdot q^{\frac{d}{m}-1})^m < 2^{-2}$. So, $m > \frac{d\log q+2}{\log q - \log(2\sqrt{\hat{l}}B)}$ and we get that the distinguishing advantage we have in the above reduction is larger than $\varepsilon - (m+1)\cdot 2^{-2n}$. Hence, we have the following proposition.

Proposition 2. *Assume that $q > 2\sqrt{\hat{l}}\cdot B$ with $B = \sqrt{\alpha^2+\beta^2}\cdot\sqrt{n}$, there is a reduction from* Nor-D-MLWE$^{R^d}_{q,\phi}$ *to* Nor-S-MLWE$^{R^d}_{q,\phi}$ *problems when given $m > \frac{d\log q+2}{\log q-\log(2\sqrt{\hat{l}}B)}$ samples.*

Remark 2. Note that in this section, we use (4) (a good basis of $R^\vee$ more precisely) to bound the probability (5).

3.4 Reduction from Nor-S-MLWE to S-RLWE

In this subsection, we use methods showed in [1] to reduce Nor-S-MLWE problems to the worst-case S-RLWE problems. We shall use the following lemma from [35] to bound some useful magnitude about the secret $\boldsymbol{s}$.

Lemma 11. *For any full rank lattice $\Lambda \subseteq H$, $c \in H$, $\varepsilon \in (0,1)$, $t \geq \sqrt{2\pi}$, unit vector $\boldsymbol{u} \in H$ and $\sigma \geq \frac{t}{\sqrt{2\pi}} \cdot \eta_\varepsilon(\Lambda)$, we have*

$$\Pr_{\boldsymbol{x} \hookleftarrow D_{\Lambda,\sigma,\boldsymbol{c}}}[|<\boldsymbol{x}-\boldsymbol{c},\boldsymbol{u}>| \leq \frac{\sigma}{t}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\sqrt{2\pi e}}{t}.$$

Similarly, if $\sigma \geq \eta_\varepsilon(\Lambda)$, we have

$$\Pr_{\boldsymbol{x} \hookleftarrow D_{\Lambda,\sigma,\boldsymbol{c}}}[|<\boldsymbol{x}-\boldsymbol{c},\boldsymbol{u}>| \geq t \cdot \sigma] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot t \cdot \sqrt{2\pi e} \cdot e^{-\pi t^2}.$$ [6]

We can deduce the following useful estimate, which will be used to bound the increase of errors, of $|\sigma_k(x)|$ for some $x \hookleftarrow \phi$ and any $k \in [n]$.

Lemma 12. *Let $\varepsilon \in (0,1)$, $t \geq \sqrt{2\pi}$, $\phi = D_{\frac{1}{q}R^\vee,\sqrt{\alpha^2+\beta^2}}$ with $\sqrt{\alpha^2+\beta^2} \geq \frac{t}{\sqrt{2\pi}} \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$, we have*

$$\Pr_{x \hookleftarrow \phi}[\max_{1\leq i\leq n} \frac{1}{|\sigma_i(x)|} \geq \frac{\sqrt{2}\cdot t}{\sqrt{\alpha^2+\beta^2}}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{n\cdot\sqrt{2\pi e}}{2t}$$

and

$$\Pr_{x \hookleftarrow \phi}[\max_{1\leq i\leq n} |\sigma_i(x)| \geq t \cdot \sqrt{\alpha^2+\beta^2}] \leq n \cdot \frac{1+\varepsilon}{1-\varepsilon} \cdot t \cdot \sqrt{2\pi e} \cdot e^{-\pi t^2}.$$

Proof. For any $x \hookleftarrow \phi$, by using Lemma 11 with $\boldsymbol{c} = \boldsymbol{0}$ and $\boldsymbol{u} = (\frac{1}{\sqrt{2}}, 0, \cdots, 0, \frac{1}{\sqrt{2}})$ or $\boldsymbol{u} = (\frac{\boldsymbol{i}}{\sqrt{2}}, 0, \cdots, 0, -\frac{\boldsymbol{i}}{\sqrt{2}})$, we have

$$\Pr[|\sqrt{2} \cdot \mathrm{Re}(\sigma_1(x))| \leq \frac{\sqrt{\alpha^2+\beta^2}}{t}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\sqrt{2\pi e}}{t},$$

or

$$\Pr[|\sqrt{2} \cdot \mathrm{Im}(\sigma_1(x))| \leq \frac{\sqrt{\alpha^2+\beta^2}}{t}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\sqrt{2\pi e}}{t}.$$

Since $|\sigma_1(x)| \geq \max\{|\mathrm{Re}(\sigma_1(x))|, |\mathrm{Im}(\sigma_1(x))|\}$, we get

$$\Pr[\sqrt{2} \cdot |\sigma_1(x)| \leq \frac{\sqrt{\alpha^2+\beta^2}}{t}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\sqrt{2\pi e}}{t},$$

which implies that

$$\Pr[\frac{1}{|\sigma_1(x)|} \geq \frac{\sqrt{2}\cdot t}{\sqrt{\alpha^2+\beta^2}}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\sqrt{2\pi e}}{t}.$$

For other $k \in [\mathfrak{r}]$, we can get the same result by using similar method. Then, by taking a union bound and noticing that $\sigma_k(x) = \overline{\sigma_{n+1-k}(x)}$, we conclude the first desired result. The second assertion can be obtained similarly.

Set $\Lambda = \frac{1}{q^{d-1}} \cdot \boldsymbol{g} \cdot R + qR^d$ with $\boldsymbol{g} = (1, q, q^2, \cdots, q^{d-1})^T \in R^d$ and denote B_Λ the basis of Λ, B_{s_iR} the basis of s_iR for some $s_i \in K$. For any basis B_R of R, it is easy to verify that

$$B_\Lambda = \begin{pmatrix} 1 & \frac{1}{q} & \cdots & \frac{1}{q^{d-1}} \\ & 1 & \cdots & \frac{1}{q^{d-2}} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix} \otimes B_R = \begin{pmatrix} B_R & \frac{1}{q}B_R & \cdots & \frac{1}{q^{d-1}}B_R \\ & B_R & \cdots & \frac{1}{q^{d-2}}B_R \\ & & \ddots & \vdots \\ & & & B_R \end{pmatrix}$$

[6] Here, $D_{\Lambda,\sigma,\boldsymbol{c}} = D_{\Lambda-\boldsymbol{c},\sigma}$ corresponds to the distribution $\frac{e^{-\pi\frac{||\boldsymbol{x}-\boldsymbol{c}||^2}{\sigma^2}}}{\sum_{\boldsymbol{y}\in\Lambda} e^{-\pi\frac{||\boldsymbol{y}-\boldsymbol{c}||^2}{\sigma^2}}}$.

is a basis of Λ. Moreover, $||\tilde{B}_\Lambda|| = ||\tilde{B}_R||$. We then take B_R to be the powerful basis of R, hence, $||\tilde{B}_\Lambda|| = ||\tilde{B}_R|| \leq ||B_R|| = \sqrt{n}$. Observe that for any $x \in \frac{1}{q}R^\vee$, $x \cdot R \subseteq \frac{1}{q}R^\vee$ is a fractional ideal of K with a set of basis $x \cdot B_R$, here B_R denotes the powerful basis of R. Moreover, we have $||\widetilde{x \cdot B_R}|| \leq ||x \cdot B_R|| \leq ||x||_\infty \cdot ||B_R|| \leq \sqrt{n} \cdot ||x||$. Now we can present the following lemma.[7]

Lemma 13. *Assume* $\boldsymbol{s} = (s_1, \cdots, s_d)^T \in (\frac{1}{q}R^\vee/R^\vee)^d$ *which satisfies* $\max_{1 \leq k \leq n} \frac{1}{|\sigma_k(s_i)|} < B_2$ *and* $\max_{1 \leq k \leq n} |\sigma_k(s_i)| \leq ||s_i|| < B_1$ *for all* $i \in [d]$. *Let* $r \geq \max\{\sqrt{n}, \sqrt{n} \cdot B_1 \cdot B_2\} \cdot \sqrt{\frac{2\ln(2nd(1+\frac{1}{\varepsilon}))}{\pi}}$. *There is a PPT transformation* $\mathcal{F} : R_q^d \times \frac{1}{q}R^\vee/R^\vee \mapsto R_{q^d} \times \mathbb{T}_{R^\vee}$ *such that*

$$R_\infty(A_{q^d,s',D_{\boldsymbol{\alpha}}} || \mathcal{F}(A^{R^d*}_{q,\boldsymbol{s},\phi})) \leq (\frac{1+\varepsilon}{1-\varepsilon})^{d+4},$$

where $s' \hookleftarrow U(R^\vee_{q^d})$ *and* $\boldsymbol{\alpha}_j = \sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot d \cdot B_1^2 + r^2 \cdot \sum_{k=1}^d |\sigma_j(s_k)|^2}$ *for* $j \in [n]$.

Proof. Suppose that we are given $(\boldsymbol{a}, b) \hookleftarrow A^{R^d*}_{q,\boldsymbol{s},\phi}$. Consider the following map $\mathcal{F} : R_q^d \times \frac{1}{q}R^\vee/R^\vee \mapsto R_{q^d} \times \mathbb{T}_{R^\vee}$:

1. Sample $\boldsymbol{f} \hookleftarrow D_{\Lambda - \boldsymbol{a}, r}$.
2. Let $\boldsymbol{v} = \boldsymbol{a} + \boldsymbol{f} \bmod qR^d$ and set $\tilde{a} = x \in R_{q^d}$, where $x \in R_{q^d}$ be a random solution of $\frac{1}{q^{d-1}}\boldsymbol{g} \cdot x = \boldsymbol{v} \bmod qR^d$.
3. Sample $\tilde{e} \hookleftarrow D_{r \cdot \gamma}$ with $\gamma = \sqrt{d} \cdot B_1$, $e' \hookleftarrow D_{\sqrt{\alpha^2+\beta^2}}$ and $y \hookleftarrow U(R^\vee_{q^d})$, set $\tilde{b} = b + \tilde{e} + e' + \frac{1}{q^d}\tilde{a} \cdot y \bmod R^\vee$.
4. Output $(\tilde{a}, \tilde{b})$.

Note that $\boldsymbol{a} \in R_q^d$, so the coset $\Lambda - \boldsymbol{a}$ is well defined. Meanwhile, $r \geq ||\tilde{B}_\Lambda|| \cdot \sqrt{\frac{\ln(2nd+4)}{\pi}}$, we can efficiently sample $\boldsymbol{f}$ by Lemma 2. Assume $\boldsymbol{a} = (a_1, \cdots, a_d)^T$, $\boldsymbol{s} = (s_1, \cdots, s_d)^T$, $\boldsymbol{f} = (f_1, \cdots, f_d)^T$ and $\tilde{s} = \boldsymbol{g}^T \cdot \boldsymbol{s} + \frac{1}{q}y$, we have

$$\begin{aligned} \tilde{b} - \frac{1}{q^{d-1}}\tilde{a} \cdot \tilde{s} \bmod R^\vee &= \boldsymbol{a}^T \cdot \boldsymbol{s} + e + e' + \tilde{e} - \frac{1}{q^{d-1}}\tilde{a} \cdot \boldsymbol{g}^T \cdot \boldsymbol{s} \bmod R^\vee \\ &= e + e' + \tilde{e} - \boldsymbol{f}^T \cdot \boldsymbol{s} \bmod R^\vee. \end{aligned} \tag{6}$$

Since we choose $\beta \geq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \frac{\sqrt{2n}}{q}$ in Subsection 3.2, by Lemma 7, we have $R_\infty(D_{\sqrt{2(\alpha^2+\beta^2)}} || e + e') \leq \frac{1+\varepsilon}{1-\varepsilon}$.

In the following, we denote $\mathcal{D}$ the distribution of the outputs of $\mathcal{F}$ and try to bound $R_\infty(A_{q^d,s',D_{\alpha'}} || \mathcal{D})$. Observe that $\Lambda \cong \frac{1}{q^{d-1}}\boldsymbol{g} \cdot R_{q^d} \bmod qR^d$, every $x \in R_{q^d}$ is a solution to the equation $\frac{1}{q^{d-1}}\boldsymbol{g} \cdot x = \boldsymbol{v} \bmod qR^d$ for some $\boldsymbol{v}$ and the number of solutions to this equation in R_{q^d} for different $\boldsymbol{v}$ is the same. For any $\bar{\boldsymbol{a}} \in R_q^d$ and $\bar{\boldsymbol{f}} \in \Lambda - \bar{\boldsymbol{a}}$, we have

$$\begin{aligned} \Pr[\boldsymbol{a} = \bar{\boldsymbol{a}} \wedge \boldsymbol{f} = \bar{\boldsymbol{f}}] &= \frac{1}{q^{nd}} \cdot \frac{\rho_r(\bar{\boldsymbol{f}})}{\rho_r(\Lambda - \bar{\boldsymbol{a}})} \\ &= C \cdot \frac{\rho_r(\Lambda)}{\rho_r(\Lambda - \bar{\boldsymbol{a}})} \cdot \rho_r(\bar{\boldsymbol{f}}) \\ &\in C \cdot \left[1, \frac{1+\varepsilon}{1-\varepsilon}\right] \cdot \rho_r(\bar{\boldsymbol{f}}), \end{aligned} \tag{7}$$

[7] In the following, we use powerful basis of R to implement Lemma 2. For general number field, we also need a good basis of R (or equivalently, a good basis of $R^\vee$) to efficient output discrete Gaussian samples.

where $C = \frac{q^{-nd}}{\rho_r(\Lambda)}$ and we have used Lemma 6 with $r \geq \eta_\varepsilon(\Lambda)$. Then, for any $\bar{\boldsymbol{v}} \in \Lambda \bmod qR^d$, by using Lemma 6 again, we get

$$\begin{aligned}\Pr[\boldsymbol{v} = \bar{\boldsymbol{v}}] &= \sum_{\boldsymbol{a} \in R_q^d} \Pr[\boldsymbol{a}] \cdot \Pr[\boldsymbol{f} = \bar{\boldsymbol{v}} - \boldsymbol{a} | \boldsymbol{a}] \\ &\in C \cdot \left[1, \frac{1+\varepsilon}{1-\varepsilon}\right] \sum_{\boldsymbol{a} \in R_q^d} \rho_r(\bar{\boldsymbol{v}} - \boldsymbol{a}) \\ &\in C \cdot \left[1, \frac{1+\varepsilon}{1-\varepsilon}\right] \cdot \rho_r(\bar{\boldsymbol{v}} - R^d) \\ &\in C' \cdot \left[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}\right],\end{aligned}$$

where $C' = C \cdot \rho_r(R^d)$, also we have used that $r \geq \eta_\varepsilon(R^d)$ and $\rho_r(-R^d) = \rho_r(R^d)$. Now, let $K_{\boldsymbol{v}}$ denote the number of $\boldsymbol{v}$ that has solutions in the equation $\frac{1}{q^{d-1}} \boldsymbol{g} \cdot x = \boldsymbol{v} \bmod qR^d$, we have

$$C' \cdot \frac{1-\varepsilon}{1+\varepsilon} \cdot K_{\boldsymbol{v}} \leq \sum_{\bar{\boldsymbol{v}}} \Pr[\boldsymbol{v} = \bar{\boldsymbol{v}}] = 1 \leq C' \cdot \frac{1+\varepsilon}{1-\varepsilon} \cdot K_{\boldsymbol{v}}.$$

So, for any $\bar{a} \in R_{q^d}$,

$$\begin{aligned}\Pr[\tilde{a} = \bar{a}] &= \sum_{\bar{\boldsymbol{v}}} \Pr[\tilde{a} = \bar{a} | \boldsymbol{v} = \bar{\boldsymbol{v}}] \cdot \Pr[\boldsymbol{v} = \bar{\boldsymbol{v}}] \\ &\in \frac{1}{q^{nd}} \cdot \left[(\frac{1-\varepsilon}{1+\varepsilon})^2, (\frac{1+\varepsilon}{1-\varepsilon})^2\right].\end{aligned}$$

Therefore, we have $R_\infty(U(R_{q^d}) || \tilde{a}) \leq (\frac{1+\varepsilon}{1-\varepsilon})^2$.

We now analyze the distribution of $-\boldsymbol{f}$ appeared in (6) condition on some fixed $\bar{a}$ (equivalently, condition on some fixed $\bar{\boldsymbol{v}}$). In this situation, $-\boldsymbol{f} \in R^d - \bar{\boldsymbol{v}}$ and fixing a value $\bar{\boldsymbol{f}}$ fixes $\boldsymbol{a} = \bar{\boldsymbol{v}} - \bar{\boldsymbol{f}} \bmod qR^d$. So, by (7), we have

$$\frac{1-\varepsilon}{1+\varepsilon} \cdot \frac{\rho_r(-\bar{\boldsymbol{f}})}{\rho_r(R^d - \bar{\boldsymbol{v}})} \leq \Pr[-\boldsymbol{f} = -\bar{\boldsymbol{f}} | \tilde{a} = \bar{a}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\rho_r(-\bar{\boldsymbol{f}})}{\rho_r(R^d - \bar{\boldsymbol{v}})}.$$

Hence, $R_\infty(D_{R^d - \bar{\boldsymbol{v}}, r} || -\boldsymbol{f}) \leq \frac{1+\varepsilon}{1-\varepsilon}$. This also implies that condition on some fixed $\bar{\boldsymbol{v}} = (v_1, \cdots, v_d)^T$, $\Delta(D_{R^d - \bar{\boldsymbol{v}}, r}, -\boldsymbol{f}) \leq 2\varepsilon$, i.e. $-f_i$ is almost distributed as $D_{R - v_i, r}$ for $i \in [d]$. It then follows that $-s_i \cdot f_i$ is almost distributed as $D_{s_i R - s_i \cdot v_i, \boldsymbol{r}_i}$ with $\boldsymbol{r}_i = (r \cdot |\sigma_1(s_i)| \cdots, r \cdot |\sigma_n(s_i)|)^T$ for $i \in [d]$. Note that $\tilde{e} \hookleftarrow D_{r \cdot \gamma}$ is equivalent to $\tilde{e} = \sum_{i=1}^d \tilde{e}_i$ with $\tilde{e}_i \hookleftarrow D_{r \cdot B_1}$. For $i \in [d]$, let $D^{(i)}$ denotes the distribution of $\varphi(-s_i \cdot f_i) + \tilde{e}_i$, $Y^{(i)}$ denotes the distribution obtained by sampling from $D_{s_i R - s_i \cdot v_i, \boldsymbol{r}_i}$ and then adding a vector sampled from $D_{r \cdot B_1}$, $\tilde{D}$ denotes the distribution of $-\sum_{i=1}^d s_i \cdot f_i + \tilde{e}$ in (6). By using the data-processing inequality of Rényi Divergence with the function $(-\boldsymbol{f}, \tilde{e}_1, \cdots, \tilde{e}_d) \mapsto (\varphi(-s_1 \cdot f_1) + \tilde{e}_1, \cdots, \varphi(-s_d \cdot f_d) + \tilde{e}_d)$, we obtain

$$\begin{aligned}R_\infty(Y^{(1)} \times \cdots \times Y^{(d)} || D^{(1)} \times \cdots \times D^{(d)}) &\leq R_\infty(D_{R^d - \bar{\boldsymbol{v}}, r} \times D_{r \cdot B_1}^d || -\boldsymbol{f} \times D_{r \cdot B_1}^d) \\ &\leq \frac{1+\varepsilon}{1-\varepsilon}.\end{aligned}$$

Then, noticing that by our choice of r, we can use Lemma 7 and conclude that

$$R_\infty(D_{\boldsymbol{t}_i} || Y^{(i)}) \leq \frac{1+\varepsilon}{1-\varepsilon}$$

for any $i \in [d]$, where $\boldsymbol{t}_i = (\sqrt{r^2 \cdot B_1^2 + r^2 \cdot |\sigma_1(s_i)|^2}, \cdots, \sqrt{r^2 \cdot B_1^2 + r^2 \cdot |\sigma_n(s_i)|^2})^T$. By first applying the data-processing inequality to the function that sums the samples and then considering the weak triangle inequality and independence, we have

$$\begin{aligned} R_\infty(D_{\boldsymbol{t}}||\tilde{D}) &\leq R_\infty(D_{\boldsymbol{t}_1} \times \cdots \times D_{\boldsymbol{t}_d}||Y^{(1)} \times \cdots \times Y^{(d)}) \cdot R_\infty(Y^{(1)} \times \cdots \times Y^{(d)}||D^{(1)} \times \cdots \times D^{(d)}) \\ &\leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \prod_{i=1}^{d} R_\infty(D_{\boldsymbol{t}_i}||Y^{(i)}) \leq (\frac{1+\varepsilon}{1-\varepsilon})^{d+1}, \end{aligned}$$

where $\boldsymbol{t} = (\sqrt{r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^{d} |\sigma_1(s_k)|^2}, \cdots, \sqrt{r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^{d} |\sigma_n(s_k)|^2})^T$.

Finally, note that $\frac{1}{q^{d-1}}\tilde{a} \cdot \tilde{s}$ for some $\tilde{s} \in \frac{1}{q}R^\vee/R^\vee$ is equivalent to $\frac{1}{q^d}\tilde{a} \cdot \tilde{s}'$ for $\tilde{s}' = q \cdot \tilde{s} \in R_q^\vee$. We obtain, by using data processing inequality and the multiplicativity of Rényi divergence,

$$R_\infty(A_{q^d,\tilde{s}',D_{\boldsymbol{\alpha}}}||\mathcal{D}) \leq (\frac{1+\varepsilon}{1-\varepsilon})^{d+4},$$

where $\boldsymbol{\alpha} = (\sqrt{2(\alpha^2+\beta^2) + r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^{d} |\sigma_1(s_k)|^2}, \cdots, \sqrt{2(\alpha^2+\beta^2) + r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^{d} |\sigma_n(s_k)|^2})^T$.

Combining Lemmata 12 and 13, we get the following proposition.

Proposition 3. *There is a reduction from* Nor-S-MLWE$_{q,\phi}^{R^d}$ *to the worst-case* S-RLWE$_{q^d,D_{\leq\alpha'}}$ *with m samples, where $\alpha' = \sqrt{2(\alpha^2+\beta^2)(1 + r^2 \cdot d \cdot n)}$ with $r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1+(d+4)m))}$ and $\sqrt{\alpha^2+\beta^2} \geq 2\sqrt{e} \cdot n \cdot d \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$.*

Proof. Recall that for Nor-S-MLWE$_{q,\phi}^{R^d}$ problem, the secret $\boldsymbol{s} \hookleftarrow \phi^d$. By using Lemma 5 and 12 with $\varepsilon = \frac{1}{m(d+4)}$ and $t = 2n \cdot d \cdot \sqrt{2\pi e}$, we have that with probability greater than $(1 - \frac{m(d+4)+1}{m(d+4)-1} \cdot \frac{1}{4d} - 2^{-2n})^d > (1 - \frac{1}{2d} - 2^{-2n})^d$, $\max_{1\leq k\leq n} \frac{1}{|\sigma_k(s_i)|} < B_2 := \frac{4n \cdot d \cdot \sqrt{\pi e}}{\sqrt{\alpha^2+\beta^2}}$ and $\max_{1\leq k\leq n} |\sigma_k(s_i)| \leq ||s_i|| < B_1 := \sqrt{n} \cdot \sqrt{\alpha^2+\beta^2}$ for all $i \in [d]$. So, $r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1+(d+4)m))}$ is sufficient to use Lemma 13. At the same time, the error distribution $D_{\boldsymbol{\alpha}}$ satisfies $\boldsymbol{\alpha}_i \leq \alpha'$.

Therefore, when given m samples, we can use the above settings and Lemma 13 to solve Nor-S-MLWE$_{q,\phi}^{R^d}$ problem with advantage greater than $(1 - 2^{-2n} - \frac{1+\varepsilon}{1-\varepsilon} \cdot n \cdot \frac{\sqrt{2\pi e}}{t})^d (\frac{1+\varepsilon}{1-\varepsilon})^{-(d+4)m} \geq \frac{1}{8}(1 - \frac{1}{2d} - 2^{-2n})^d > \frac{1}{16} - \frac{d}{2^{2n+3}}$, as desired.

Remark 3. The requirements of Proposition 3 can be released. One can see that we only need to assume that we can solve S-RLWE problem with $s' \hookleftarrow U(R_{q^d}^\vee)$ and non-negligible advantage δ. Then, we can solve the Nor-S-MLWE problem with non-negligible advantage $\delta' = \delta \cdot (\frac{1}{16} - \frac{d}{2^{2n+3}})$.

Now, we can collect the results of Propositions 1, 2 and 3 to conclude the following theorem.

Theorem 1. *Assume $\varepsilon \in (0, \frac{1}{2})$, $\alpha = \alpha(n) \in (0,1)$ and $\beta \geq \frac{\sqrt{2n}}{q} \cdot \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}}$ such that $\sqrt{\alpha^2+\beta^2} \geq 2\sqrt{e} \cdot n \cdot d \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$. Let $q > \max\{2n, 2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{\alpha^2+\beta^2}\}$ be a prime that does not ramify in R. When given $m > \frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{(\alpha^2+\beta^2)})}$ samples, there is a probabilistic polynomial-time reduction from* D-MLWE$_{q,D_\alpha}^{R^d}$ *in worst/average-case to* S-RLWE$_{q^d,D_{\leq\alpha'}}$ *in worst-case for arbitrary $d = poly(n)$, where $\alpha' = \sqrt{2(\alpha^2+\beta^2)(1 + d \cdot n \cdot r^2)}$ and $r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1+(d+4))m)}$.*

Remark 4. In many applications, for example, the NIST submissions KCL, CRYSTALS-KYBER and CRYSTALS-DILITHIUM, we usually set $d = O(1)$ and $q = 1 \bmod l$, then we can direct reduce corresponding S-RLWE to average-case D-RLWE by using the reductions showed in [23], hence reduce D-MLWE to average-case D-RLWE efficiently.

The term $2n$ in the inequality of q can be replaced by some $\Omega(n)$. As we will see later, we have to set q large than $\tilde{O}(n)$ usually. Till now, we obtain a reduction from D-MLWE to S-RLWE with polynomially bounded q and error parameters. For example, in order to obtain a meaningful reduction, we need to avoid the case $\alpha' \geq \eta_\varepsilon(R^\vee)$. Recall that, by Lemmata 1 and 4, for cyclotomic fields, we have

$$\sqrt{\frac{\ln(\frac{1}{\varepsilon})}{\pi}} \cdot n^{-\frac{1}{2}} \cdot (\frac{\prod_{p|l} p^{\frac{1}{p-1}}}{l})^{-\frac{1}{2}} \leq \eta_\varepsilon(R^\vee) \leq \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}} \cdot \sqrt{n} \cdot (\frac{\prod_{p|l} p^{\frac{1}{p-1}}}{l})^{\frac{1}{2}}.$$

The upper bound of $\eta_\varepsilon(R^\vee)$ can be as small as $\tilde{O}(1)$. Assume $d = \tilde{O}(n^{c_1})$ and $\alpha = \tilde{O}(n^{-c_2})$, we then can set $\beta \approx \alpha$, $q = \tilde{O}(n^{\frac{3}{2}+c_1+c_2})$ and $r = \tilde{O}(n^{2+c_1})$, which gives $\alpha' = \tilde{O}(n^{\frac{5+3c_1}{2}-c_2})$. So, $c_2 > \frac{5+3c_1}{2}$ is sufficient. In applications, we usually use very small $d = O(1)$, then we can set $\alpha \approx \beta = \tilde{O}(n^{-\frac{5}{2}})$ and $q = \tilde{O}(n^4)$ to obtain a very satisfactory result.

3.5 Reduction From S-RLWE to D-RLWE

We now need to reduce the worst-case S-RLWE problems to average-case D-RLWE problmes to finish our reduction. Note that the modulus in the S-RLWE problems we investigate is q^d, so we can't directly use the reduction showed in [23] even in the cyclotomic fields, unless we add more restricts on q and d, for example $d = 2, 3$ and $q = 1 \bmod l$. There are some results showed in [34], which state a reduction from S-RLWE problem to worst-case D-RLWE problem for arbitrary modulus q.

Theorem 2. *Let $\boldsymbol{r} \in (\mathbb{R}^+)^n$ be such that $\boldsymbol{r}_i = \boldsymbol{r}_{n+1-i}$ for all $i \in [\frac{n}{2}]$ and $\boldsymbol{r}_i \leq r$ for some r. Let $d' = n \cdot q^{\frac{1}{m}+\frac{1}{n}}$, and consider $\Sigma = \{\boldsymbol{r}' : \boldsymbol{r}'_i \leq \sqrt{d'^2 \cdot r^2 \cdot m + d'^2}\}$. Then, there exists a probabilistic polynomial-time reduction from S-RLWE$_{q,D_{\boldsymbol{r}}}$ with $m \leq \frac{q}{2n}$ input samples to worst-case D-RLWE$_{q,\Sigma}$.*

Collecting Theorem 1 and Theorem 2, we get the following theorem.

Theorem 3. *Assume $\varepsilon \in (0, \frac{1}{2})$, $\alpha = \alpha(n) \in (0, 1)$ and $\beta \geq \frac{\sqrt{2n}}{q} \cdot \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}}$ such that $\sqrt{\alpha^2+\beta^2} \geq 2\sqrt{e} \cdot n \cdot d \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$. Let $q > \max\{2n, 2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{\alpha^2+\beta^2}\}$ be a prime that does not ramify in R. When given $\frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{\alpha^2+\beta^2})} < m \leq \frac{q}{2n}$ samples, there is a probabilistic polynomial-time reduction from* D-MLWE$^{R^d}_{q,D_\alpha}$ *in worst/average-case to* D-RLWE$_{q^d, D_{\leq \beta'}}$ *in worst-case, where $\beta' = \sqrt{(n \cdot q^{\frac{d}{m}+\frac{d}{n}})^2 \cdot (1 + m \cdot \alpha'^2)}$, $\alpha' = \sqrt{2(\alpha^2+\beta^2)(1 + d \cdot n \cdot r^2)}$ and $r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1+(d+4))m)}$.*

Note that, the error parameter β' contains a term $q^{\frac{d}{m}+\frac{d}{n}}$. Assume $d = O(1)$ and $\alpha = \tilde{O}(n^{-c})$, we set $\beta \approx \alpha = \tilde{O}(n^{-c})$, $q = \tilde{O}(n^{c+\frac{3}{2}})$ and $r = \tilde{O}(n^2)$. Under this condition, we have $\alpha' = \tilde{O}(n^{\frac{5}{2}-c})$, since $m \geq \tilde{O}(1)$ implies $q^{\frac{d}{m}+\frac{d}{n}} = O(1)$. So, $\beta' = \tilde{O}(n^{\frac{7}{2}-c} \cdot m^{\frac{1}{2}})$. Meanwhile, $\frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{\alpha^2+\beta^2})} < m \leq \frac{q}{2n} = \tilde{O}(n^{c+\frac{1}{2}})$. We conclude that $c > \frac{7}{2}$ for $m = \tilde{O}(1)$ or $c > \frac{15}{2}$ for $m = \frac{q}{2n}$ is sufficient for us to obtain a meaningful reduction.

Next, we consider to reduce the worst-case D-RLWE to average-case D-RLWE. Variant solutions can be found in previous works. For example, one can use Lemma 2.14 of [34] to discuss the distribution $\mathcal{D}$ over the set of error distributions $D_{\leq \beta'}$. In this paper, we use the following lemma, which comes from [30], to reduce worst-case D-RLWE to average-case D-RLWE with a spherical error.

Lemma 14. *There is a randomized polynomial-time algorithm that given any $\beta' > 0$ and $m \geq 1$, as well as an oracle that solves D-RLWE$_{q,D_\xi}$ given only m samples for any modulus q, where $\xi = \beta' \cdot (\frac{nm}{\log(nm)})^{\frac{1}{4}}$, solves D-RLWE$_{q,D_{\leq \beta'}}$.*

Overall, we conclude the following theorem.

Theorem 4. *Assume* $\varepsilon \in (0, \frac{1}{2})$, $\alpha = \alpha(n) \in (0, 1)$ *and* $\beta \geq \frac{\sqrt{2n}}{q} \cdot \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}}$ *such that* $\sqrt{\alpha^2 + \beta^2} \geq 2\sqrt{e} \cdot n \cdot d \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$. *Let* $q > \max\{2n, 2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2}\}$ *be a prime that does not ramify in* R. *When given* $\frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\hat{l}} \cdot \sqrt{n} \cdot \sqrt{\alpha^2+\beta^2})} < m \leq \frac{q}{2n}$ *samples, there is a probabilistic polynomial-time reduction from* D-MLWE$^{R^d}_{q,D_\alpha}$ *to* D-RLWE$_{q^d,D_\Gamma}$ *in average-case, where* $\Gamma = \beta' \cdot (\frac{nm}{\log(nm)})^{\frac{1}{4}}$, $\beta' = \sqrt{(n \cdot q^{\frac{d}{m} + \frac{d}{n}})^2 \cdot (1 + m \cdot \alpha'^2)}$, $\alpha' = \sqrt{2(\alpha^2 + \beta^2)(1 + d \cdot n \cdot r^2)}$ *and* $r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1 + (d + 4))m)}$.

Usually, there are reductions from worst-case SIVP$_\gamma$ with $\gamma = \tilde{O}(\frac{n^{\frac{3}{4}}}{\alpha})$ over rings or modules to corresponding average-case D-LWE problem with error distribution D_α and $\alpha \leq \tilde{O}(n^{-\frac{1}{4}})$ [21, 30]. Hence, when $m = \tilde{O}(1)$ and $d = O(1)$, we obtain a reduction from worst-case SIVP$_\gamma$ to average-case D-RLWE$_{q^d,\mathcal{D}_\Gamma}$ with $q \leq \tilde{O}(n^{5.75})$, $\gamma \leq \tilde{O}(n^5)$ and $\Gamma \approx \tilde{O}(n^{-\frac{1}{2}})$.

4 Self-reductions of Ring-LWE Problems

Reductions from S-RLWE to D-RLWE in [34] restricts the number of samples. This increases requirements of capacities of the adversary. Meanwhile, the error rate is also related heavily to the number of samples. However, in applications, we may usually hope that the number of samples m should be independent of the modulus q and need only to be bounded by $poly(n)$. So is the error rate. In this section, we shall use similar method as in Section 3 to give a self-reduction of RLWE problems to offer an alternative solution to this problem.

We reset the values of α and β, and give a self-reduction of S-RLWE first. We begin with the problem S-RLWE$_{q,D_\alpha}$. It is easy to deduce that Nor-S-RLWE$_{q,\phi}$ (denote corresponding distribution $A^*_{q,s,\phi}$) is also hard for $\phi = D_{\frac{1}{q}R^\vee,\sqrt{\alpha^2+\beta^2}}$ with $\beta \geq \frac{\sqrt{2n}}{q} \cdot \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}}$. The proof of the following lemma is similar to that of Lemma 13.

Lemma 15. *Assume* $s \in \frac{1}{q}R^\vee/R^\vee$ *such that* $\frac{1}{|\sigma_k(s)|} \leq B_2$ *and* $|\sigma_k(s)| \leq ||s|| \leq B_1$ *for all* $k \in [n]$, *let* $r \geq \max\{\sqrt{n}, \frac{p}{q} \cdot \sqrt{n}, \sqrt{n} \cdot B_1 \cdot B_2 \cdot \sqrt{1 + \frac{p^2}{q^2}}\} \cdot \sqrt{\frac{\ln(2n(1+\frac{1}{\varepsilon}))}{\pi}}$, *there is a transformation* $\mathcal{F} : R_q \times \frac{1}{q}R^\vee/R^\vee \mapsto R_p \times \mathbb{T}_{R^\vee}$ *such that*

$$R_\infty(A_{p,\tilde{s},D_{\boldsymbol{t}}}||\mathcal{F}(A_{q,s,\phi})) \leq (\frac{1+\varepsilon}{1-\varepsilon})^5,$$

where $\tilde{s} \hookleftarrow U(R_p)$ *and* $\boldsymbol{t}_i = \sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot B_1^2 + \frac{q^2}{p^2} \cdot r^2 \cdot |\sigma_i(s)|^2}$ *for* $i \in [n]$.

Proof. We consider the following transformation with a given sample $(a, b) \in R_q \times \frac{1}{q}R^\vee/R^\vee$:

1. Sample $f \hookleftarrow D_{R - \frac{p}{q} \cdot a, r}$ and $s_1 \hookleftarrow U(R_p^\vee)$.
2. Set $\tilde{a} = f + \frac{p}{q} \cdot a \bmod pR$.
3. Set $\tilde{b} = b + \frac{1}{p}\tilde{a} \cdot s_1 + \tilde{e} + e' \bmod R^\vee$ with $\tilde{e} \hookleftarrow D_{r \cdot B_1}$ and $e' \hookleftarrow D_{\sqrt{\alpha^2+\beta^2}}$.
4. Output $(\tilde{a}, \tilde{b})$.

Since $a \in R_q$ and $r \geq ||\tilde{B}_R|| \cdot \sqrt{\frac{\ln(2n+4)}{\pi}}$, the coset $R - \frac{p}{q} \cdot a$ is well defined and f can be sampled efficiently. For any $\bar{a} \in R_q$ and $\bar{f} \in R - \frac{p}{q} \cdot \bar{a}$, we have

$$\Pr[a = \bar{a} \wedge f = \bar{f}] = q^{-n} \cdot \frac{\rho_r(\bar{f})}{\rho_r(R - \frac{p}{q} \cdot \bar{a})} \in C \cdot [1, \frac{1+\varepsilon}{1-\varepsilon}] \cdot \rho_r(\bar{f}),$$

where $C = \frac{q^{-n}}{\rho_r(R)}$. Hence, for any $a' \in R_p$,

$$\Pr[\tilde{a} = a'] = \sum_{\bar{a} \in R_q} \Pr[\bar{a}] \cdot \Pr[f = a' - \frac{p}{q} \cdot \bar{a} | a = \bar{a}] \in C' \cdot [\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}],$$

where $C' = C \cdot \rho_r(\frac{p}{q} \cdot R)$ and we have used $r \geq \eta_\varepsilon(\frac{p}{q} \cdot R)$. We conclude that $C' \in \frac{1}{p^n} \cdot [\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}]$ and $R_\infty(U(R_p)||\tilde{a}) \leq (\frac{1+\varepsilon}{1-\varepsilon})^2$.

If we set $\tilde{s} = q \cdot s + s_1 \bmod pR^\vee$, we have $\tilde{s} \hookleftarrow U(R_p^\vee)$ and $\tilde{b} - \frac{1}{p}\tilde{a} \cdot \tilde{s} = e + e' + \tilde{e} - \frac{q}{p} f \cdot s \bmod R^\vee$. Then, $R_\infty(D_{\sqrt{2(\alpha^2+\beta^2)}}||e + e') \leq \frac{1+\varepsilon}{1-\varepsilon}$. We now estimate the distribution of $-f$ condition on some fixed $\bar{a} \in R_p$. Similarly, in this situation, $-f \in \frac{p}{q}R - \bar{a}$ and we have

$$\frac{1-\varepsilon}{1+\varepsilon} \cdot \frac{\rho_r(-\bar{f})}{\rho_r(\frac{p}{q}R - \bar{a})} \leq \Pr[-f = -\bar{f} | \tilde{a} = \bar{a}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\rho_r(-\bar{f})}{\rho_r(\frac{p}{q}R - \bar{a})}.$$

Then, $R_\infty(D_{\frac{p}{q}R - \bar{a}, r}|| - f) \leq \frac{1+\varepsilon}{1-\varepsilon}$ and $\Delta(D_{\frac{p}{q}R - \bar{a}, r}, -f) \leq 2\varepsilon$. Meanwhile, by our choice of r and Lemma 7, we have $R_\infty(D_{\boldsymbol{t}'}||\tilde{e} - \frac{q}{p} f \cdot s) \leq \frac{1+\varepsilon}{1-\varepsilon}$, where $\boldsymbol{t}'_i = \sqrt{r^2 \cdot B_1^2 + \frac{q^2}{p^2} \cdot r^2 \cdot |\sigma_i(s)|^2}$ for $i \in [n]$. Therefore, we obtain

$$R_\infty(A_{p, \tilde{s}, D_{\boldsymbol{t}}}||\mathcal{F}(A_{q,s,\phi})) \leq (\frac{1+\varepsilon}{1-\varepsilon})^5,$$

with $\boldsymbol{t}_i = \sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot B_1^2 + \frac{q^2}{p^2} \cdot r^2 \cdot |\sigma_i(s)|^2}$, as desired.

Now, we can obtain the following proposition by combining Lemma 12 and 15.

Proposition 4. *There is a reduction from* Nor-S-RLWE$_{q,\phi}$ *to the worst-case* S-RLWE$_{p, D_{\leq \alpha'}}$ *with m samples, where $\alpha' = \sqrt{(\alpha^2 + \beta^2)(2 + r^2 \cdot n + r^2 \cdot n \cdot \frac{q^2}{p^2})}$ with $r \geq 4\sqrt{e} \cdot n^2 \cdot \sqrt{1 + \frac{p^2}{q^2}} \cdot \sqrt{\ln(2n(1+5m))}$ and $\sqrt{\alpha^2 + \beta^2} \geq 2\sqrt{e} \cdot n \cdot \eta_\varepsilon(\frac{1}{q}R^\vee)$.*

Proof. We set $\varepsilon = \frac{1}{5m}$ and $t = 2n \cdot \sqrt{2\pi e}$. Then, with probability $\geq 1 - \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{1}{4} - 2^{-2n} > \frac{1}{2} - 2^{-2n}$, $\max_{1 \leq k \leq n} \frac{1}{|\sigma_k(s)|} < B_2 := \frac{4n\sqrt{\pi e}}{\sqrt{\alpha^2 + \beta^2}}$ and $\max_{1 \leq k \leq n} ||\sigma_k(s)|| \leq ||s|| < \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2}$. So, $r \geq 4\sqrt{e} \cdot n^2 \cdot \sqrt{1 + \frac{p^2}{q^2}} \cdot \sqrt{\ln(2n(1+5m))}$ is sufficient to use Lemma 15. At the same time, the error distribution $D_{\boldsymbol{t}}$ satisfies $\boldsymbol{t}_i \leq \alpha'$.

Therefore, when given m samples, we can use the above settings and Lemma 15 to solve Nor-S-RLWE$_{q,\phi}$ problem with advantage greater than $(\frac{1}{2} - 2^{-2n}) \cdot (\frac{5m+1}{5m-1})^{-5m} \geq \frac{1}{16} - 2^{-2n+3}$, as desired.

Remark 5. A similar self-reduction (modulus switch) of D-MLWE was given in [21]. When applied to RLWE, it also gave a modulus-switch reduction of D-RLWE. But, we should note that in the case of decision variants, in order to remain a non-negligible advantage, the reduction will suffer the same problem as in [1] and make q to be at least super-polynomial. Since we usually set the error rate of D-MLWE to be constant or polynomial, the error rate of corresponding D-RLWE will deteriorate to be negligible. So, strictly speaking, we can't directly use their reductions.

Now, we can combine Theorem 1 and Proposition 4 to reduce D-MLWE$^{R^d}_{q,D_\alpha}$ in worst/average-case to S-RLWE$_{p,\Psi}$ in worst-case, where Ψ is some set of elliptical Gaussians. Recall that the search to decision reduction in [23] requires the modulus q to split 'well', so they assume $q = 1 \bmod l$. In fact, assume $R/qR \cong R/\mathfrak{q}_1 \times \cdots \times R/\mathfrak{q}_\mathfrak{g}$ with $\mathfrak{g} \cdot \mathfrak{f} = n$, if $|R/\mathfrak{q}_k| = q^\mathfrak{f} = poly(n)$ for $k \in [\mathfrak{g}]$, the reduction in [23] also works. This inspires us that if we can find a prime p that splits 'well', then for any q satisfies $\frac{q}{p} \leq poly(n)$, we can obtain a reduction from D-RLWE$_{q,D_\alpha}$ to D-RLWE$_{p,D_{\beta'}}$ for some $\beta' = poly(n)^{-1}$ by combining Proposition 4 and the search to decision reductions showed in [23] for arbitrary $m = poly(n)$ samples.

However, this process is of course somewhat heuristic. On the one hand, the Dirichlet's theorem on primes in arithmetic progressions tells us that there are infinite many primes in the arithmetic progression $h+k\cdot l$ for $k\in\mathbb{N}$ and $(h,l)=1$. So, we may have confidence that we could find a split 'well' prime in the $poly(n)$ interval in the asymptotic sense. On the other hand, the primes are very sparse, whether there are such primes in every desired interval and how to efficiently find such primes need to be considered carefully.

5 Reductions from D-MLWE to Module-SIVP

In this section, we give converse reductions from decision module LWE problems to module SIVP problems over cyclotomic fields. We will first reduce module LWE problems to module SIS problems, then reduce module SIS problems to corresponding module SIVP problems as in [21]. Combining techniques used in [37], we can conclude the above reductions in any cyclotomic field under canonical embedding. Recall that the definition of module SIS problems (denoted by M-SIS$^{R^d}_{q,\beta}$) is as follows: Given $A \hookleftarrow U(R_q^{d\times m})$, find $\boldsymbol{z}\in R^m\backslash\{\boldsymbol{0}\}$ such that $A\cdot\boldsymbol{z}=0 \bmod qR^d$ and $||\boldsymbol{z}||\le\beta$.

It's well known that one of the classical ways to solve LWE consists in solving an associated SIS instance [21, 26].

Lemma 16. *There is a PPT reduction from D-MLWE$^{R^d}_{q,D_\alpha}$ to M-SIS$^{R^d}_{q,\beta}$ with $\alpha<\frac{1}{\beta\cdot\omega(n\ln n\sqrt{\log\log n})}$.*

Proof. Given m samples $(A,\boldsymbol{b})\in R_q^{m\times d}\times\mathbb{T}^m_{R^\vee}$, we use the M-SIS oracle to obtain some $\boldsymbol{z}$ such that $A^T\cdot\boldsymbol{z}=0 \bmod qR^d$ and $||\boldsymbol{z}||\le\beta$. Then we compute $\boldsymbol{z}^T\cdot\boldsymbol{b} \bmod R^\vee=\sum_{k=1}^n x_k\cdot\varphi(\overrightarrow{d}_k)$ $(\mathbb{T}_{R^\vee}\cong\mathbb{R}^n/\varphi(R^\vee))$ with $x_k\in[-\frac{1}{2},\frac{1}{2})$. Note that, if $\boldsymbol{b}\hookleftarrow U(\mathbb{T}^m_{R^\vee})$, we have $\boldsymbol{z}^T\cdot\boldsymbol{b}\hookleftarrow U(\mathbb{T}_{R^\vee})$, so the coefficients $\{x_k\}$'s of $\boldsymbol{z}^T\cdot\boldsymbol{b}$ will be distributed uniformly in $[-\frac{1}{2},\frac{1}{2})$. If $\boldsymbol{b}=A\cdot\boldsymbol{s}+\boldsymbol{e} \bmod R^\vee$ for some $\boldsymbol{e}\hookleftarrow D_\alpha$, then $\boldsymbol{z}^T\cdot\boldsymbol{b}=\boldsymbol{z}^T\cdot\boldsymbol{e}=\sum_{j=1}^m z_j\cdot e_j\hookleftarrow D_{\boldsymbol{r}}$ with $\boldsymbol{r}_k=\sqrt{\alpha^2\cdot\sum_{1\le j\le m}|\sigma_k(z_j)|^2}$ for $k\in[n]$. By definition, in this situation, we have $\boldsymbol{z}^T\cdot\boldsymbol{e}=\sum_{k=1}^n x'_k\cdot\boldsymbol{h}_k$ with $x'_k\hookleftarrow D_{\boldsymbol{r}_k}$ for any $k\in[n]$, so $E[e^{\frac{\pi}{2\boldsymbol{r}_k^2}\cdot(x'_k)^2}]=\sqrt{2}$. By Markov's inequality, we have

$$\Pr[(x'_k)^2\ge\frac{2\boldsymbol{r}_k^2}{\pi}\cdot t^2]\le\sqrt{2}\cdot e^{-t^2}.$$

Setting $t=\omega(\ln n)$, we get $\Pr[|x'_k|<\sqrt{\frac{2}{\pi}}\cdot\boldsymbol{r}_k\cdot\omega(\ln n)]>1-n^{-\omega(\ln n)}$. Hence, by taking a union bound, we have $\Pr[||\boldsymbol{z}^T\cdot\boldsymbol{e}||<\sqrt{n}\cdot\alpha\cdot||\boldsymbol{z}||\cdot\omega(\ln n)]>1-n^{1-\omega(\ln n)}$. Therefore, $\Pr[\max_k|x_k|<\sqrt{\hat{l}\cdot n}\cdot\alpha\cdot\beta\cdot\omega(\ln n)]>1-n^{1-\omega(\ln n)}$. Since $\sqrt{\hat{l}}=O(\sqrt{n\cdot\log\log n})$, for $\alpha<\frac{1}{\beta\cdot\omega'(n\ln n\sqrt{\log\log n})}$, we have $x_k<\frac{1}{4}$ for all $k\in[n]$ with probability at least $1-n^{-\omega''(\ln n)}$ for some other functions $\omega'(\cdot)$ and $\omega''(\cdot)$. Thus, we can distinguish $A^{R^d}_{q,\boldsymbol{s},D_\alpha}$ and $U(\mathbb{T}_{R^\vee})$ efficiently by checking if $x_k<\frac{1}{4}$ for all $k\in[n]$.

The module SIS problems correspond to finding a short vector in the lattice

$$A^\perp=\{\boldsymbol{z}\in R^m: A\cdot\boldsymbol{z}=0 \bmod qR^d\}$$

for $A\hookleftarrow U(R_q^{d\times m})$. If we can solve Mod-SIVP$_\gamma$ in the lattice $A^\perp$ for $A\hookleftarrow U(R_q^{d\times m})$ with non-negligible probability, then, of course, we can solve M-SIS$^{R^d}_{q,\beta}$ with $\beta\le\gamma\cdot\lambda_N(\Lambda)$, here $N\le m\cdot n$ denotes the dimension of lattice $A^\perp$. Note that $\lambda_N(\Lambda)\le\frac{N}{\lambda_1(\Lambda^\vee)}\le\frac{N}{\lambda_1^\infty(\Lambda^\vee)}$ for any N-dimensional lattice Λ, we only need to estimate the lower bound of $\lambda_1^\infty((A^\perp)^\vee)$. Recall that the dual $M^\vee$ of a lattice $M\subseteq K^m$ is defined as the set of all $\boldsymbol{x}\in K^m$ such that $Tr(\boldsymbol{x}^T\cdot\boldsymbol{v})\in\mathbb{Z}$ for all $\boldsymbol{v}\in M$. It is easy to check $(A^\perp)^\vee=\frac{1}{q}L_q(A)$, where

$$L_q(A)=\{\boldsymbol{y}\in(R^\vee)^m,\exists\ \boldsymbol{s}\in(R_q^\vee)^d,\ A^T\cdot\boldsymbol{s}=\boldsymbol{y} \bmod q(R^\vee)^m\}.$$

Next, we give a probabilistic lower bound of $\lambda_1^\infty(L_q(A))$ for $A\hookleftarrow U(R_q^{d\times m})$, whose proof technique is an extension of methods used in [21, 34, 35, 37] and may be standard now.

Lemma 17. *Let q be a prime that does not ramify in R and $qR = \mathfrak{q}_1 \times \cdots \times \mathfrak{q}_{\mathfrak{g}}$ with $\mathfrak{g} \cdot \mathfrak{f} = n$, assume $m > d$ and $\varepsilon \in (0,1)$, then* $\Pr_{A \hookleftarrow U(R_q^{d\times m})}[\lambda_1^{\infty}(L_q(A)) < \frac{1}{n} \cdot q^{1-\frac{d}{m}-\varepsilon}] \leq 2^{2mn+\mathfrak{g}} \cdot q^{-mn\varepsilon}$.

Proof. By our assumption, we have $N(\mathfrak{q}_k) = q^{\mathfrak{f}}$ for all $k \in [\mathfrak{g}]$. By the union bound, the probability p that $L_q(A)$ contains a nonzero vector of infinity norm $< B := \frac{1}{n} \cdot q^{1-\frac{d}{m}-\varepsilon}$ is bounded from above by

$$\sum_{\substack{\boldsymbol{t} \in (R_q^{\vee})^m \\ 0 < ||\boldsymbol{t}||_\infty < B}} \sum_{\boldsymbol{s} \in (R_q^{\vee})^d} \Pr_{A \hookleftarrow U(R_q^{d\times m})}[A^T \cdot \boldsymbol{s} = \boldsymbol{t} \bmod q(R^{\vee})^m],$$

which is equal to

$$\sum_{\substack{\boldsymbol{t} \in (R_q^{\vee})^m \\ 0 < ||\boldsymbol{t}||_\infty < B}} \sum_{\boldsymbol{s} \in (R_q^{\vee})^d} \prod_{k=1}^{m} \Pr_{\boldsymbol{a} \hookleftarrow U(R_q^{d})}[\boldsymbol{a}^T \cdot \boldsymbol{s} = t_k \bmod qR^{\vee}].$$

By the CRT and Lemma 2.15 of [23], we have R-module isomorphisms $R_q^{\vee} \cong R^{\vee}/\mathfrak{q}_1 \cdot R^{\vee} \times \cdots \times R^{\vee}/\mathfrak{q}_{\mathfrak{g}} \cdot R^{\vee} \cong R/\mathfrak{q}_1 R \times \cdots \times R/\mathfrak{q}_{\mathfrak{g}} R \cong R_q \cong \mathbb{F}_{q^{\mathfrak{f}}}^{\mathfrak{g}}$. Now, $\boldsymbol{a}^T \cdot \boldsymbol{s} = t_k \bmod qR^{\vee}$ if and only if $\boldsymbol{a}^T \cdot \boldsymbol{s} = t_k \bmod \mathfrak{q}_j \cdot R^{\vee}$ for all $j \in [\mathfrak{g}]$. If $\boldsymbol{s} = \boldsymbol{0} \bmod \mathfrak{q}_j \cdot R^{\vee}$ for some $j \in [\mathfrak{g}]$, the probability $\prod_{k=1}^{m} \Pr_{\boldsymbol{a} \hookleftarrow U(R_q^{d})}[\boldsymbol{a}^T \cdot \boldsymbol{s} = t_k \bmod qR^{\vee}] \neq 0$ if and only if $\boldsymbol{t} = \boldsymbol{0} \bmod \mathfrak{q}_j \cdot R^{\vee}$ (denoted by $\mathfrak{q}_j \cdot R^{\vee} | \boldsymbol{t}$) for the same $j \in [\mathfrak{g}]$. We denote $S \subseteq [\mathfrak{g}]$ be the set of indices j such that $\boldsymbol{s} = \boldsymbol{0} \bmod \mathfrak{q}_j \cdot R^{\vee}$. Then, for any $j \in [\mathfrak{g}] \backslash S$, we have $\Pr_{\boldsymbol{a} \hookleftarrow U(R_q^{d})}[\boldsymbol{a}^T \cdot \boldsymbol{s} = t_k \bmod qR^{\vee}] \leq \frac{1}{q^{\mathfrak{f}}}$ for any $k \in [m]$. So,

$$\Pr_{\boldsymbol{a} \hookleftarrow U(R_q^{d})}[\boldsymbol{a}^T \cdot \boldsymbol{s} = t_k \bmod qR^{\vee}] \leq \prod_{i \in [\mathfrak{g}] \backslash S} \frac{1}{q^{\mathfrak{f}}} = (\frac{1}{q^{\mathfrak{f}}})^{\mathfrak{g}-|S|}.$$

Therefore, we have

$$p \leq \sum_{S \subseteq [\mathfrak{g}]} \sum_{\substack{\boldsymbol{s} \in (R_q^{\vee})^d \\ \forall i \in S,\ \mathfrak{q}_i R^{\vee} | \boldsymbol{s}}} \sum_{\substack{\boldsymbol{t} \in (R_q^{\vee})^m \\ 0 < ||\boldsymbol{t}||_\infty < B \\ \forall i \in S,\ \mathfrak{q}_i R^{\vee} | \boldsymbol{t}}} q^{m\mathfrak{f}|S| - mn}.$$

There are $((q^{\mathfrak{f}})^{\mathfrak{g}-|S|})^d$ elements in $(R_q^{\vee})^d$ satisfying $\mathfrak{q}_i R^{\vee} | \boldsymbol{s}$ for $i \in S$. Thus,

$$p \leq \sum_{S \subseteq [\mathfrak{g}]} \sum_{\substack{\boldsymbol{t} \in (R_q^{\vee})^m \\ 0 < ||\boldsymbol{t}||_\infty < B \\ \forall i \in S,\ \mathfrak{q}_i R^{\vee} | \boldsymbol{t}}} q^{(m-d)(\mathfrak{f} \cdot |S| - n)}.$$

Set $\mathfrak{h} = \prod_{i \in S} \mathfrak{q}_i R^{\vee}$ and denote $\mathfrak{B}(r, \boldsymbol{c})$ the open ball in H of center $\boldsymbol{c}$ and radius r under the infinity norm. We now estimate the number N of $\boldsymbol{t}$'s satisfying the conditions in the above sum. First note that, if we denote $\boldsymbol{t} = (t_1, \cdots, t_m)^T$, $t_i \in \mathfrak{h}$ for all $i \in [m]$, then, $||\boldsymbol{t}||_\infty = \max_{1 \leq i \leq m} ||t_i||_\infty \geq \frac{1}{\sqrt{n}} \max_{1 \leq i \leq m} ||t_i|| \geq \frac{1}{\sqrt{n}} \lambda_1(\mathfrak{h}) \geq N(\mathfrak{h})^{\frac{1}{n}} \geq \frac{1}{n} \cdot q^{\frac{|S|}{\mathfrak{g}}}$, since $N(R^{\vee}) = \Delta_K^{-1} \geq n^{-n}$. As a result, there is no such $\boldsymbol{t}$ when $|S| \geq (1 - \frac{d}{m} - \varepsilon) \cdot \mathfrak{g}$. For the case $|S| < (1 - \frac{d}{m} - \varepsilon) \cdot \mathfrak{g}$, we try to bound $|\mathfrak{B}(B, \boldsymbol{0}) \cap \mathfrak{h}|$. Let $\lambda = \frac{\lambda_1^{\infty}(\mathfrak{h})}{2}$, then for any two elements $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ of $\mathfrak{h}$, we have $\mathfrak{B}(\lambda, \boldsymbol{v}_1) \cap \mathfrak{B}(\lambda, \boldsymbol{v}_2) = \phi$. Meanwhile, for any $\boldsymbol{v} \in \mathfrak{B}(B, \boldsymbol{0})$, we have $\mathfrak{B}(\lambda, \boldsymbol{v}) \subseteq \mathfrak{B}(B + \lambda, \boldsymbol{0})$. Hence,

$$N \leq |\mathfrak{B}(B, \boldsymbol{0}) \cap \mathfrak{h}|^m \leq \left(\frac{\mathrm{Vol}(\mathfrak{B}(B + \lambda, \boldsymbol{0}))}{\mathrm{Vol}(\mathfrak{B}(\lambda, \boldsymbol{0}))} \right)^m \leq (\frac{B}{\lambda} + 1)^{mn} \leq 4^{mn} \cdot q^{mn(1 - \frac{d}{m} - \frac{|S|}{\mathfrak{g}} - \varepsilon)},$$

where we have used $\lambda_1^{\infty}(\mathfrak{h}) \geq \frac{1}{n} \cdot q^{\frac{|S|}{\mathfrak{g}}}$. Since there are $2^{\mathfrak{g}}$ subsets of $[\mathfrak{g}]$, we get

$$\begin{aligned} p &\leq 2^{\mathfrak{g}} \cdot \max_{\substack{S \subseteq [\mathfrak{g}] \\ |S| < (1 - \frac{d}{m} - \varepsilon)\mathfrak{g}}} 4^{mn} \cdot q^{mn(1 - \frac{d}{m} - \frac{|S|}{\mathfrak{g}} - \varepsilon)} \cdot q^{(m-d)(\mathfrak{f}|S| - n)} \\ &= 2^{\mathfrak{g}+2mn} \cdot \max_{\substack{S \subseteq [\mathfrak{g}] \\ |S| < (1 - \frac{d}{m} - \varepsilon)\mathfrak{g}}} q^{-mn\varepsilon - d|S|\mathfrak{f}} \leq 2^{2mn+\mathfrak{g}} \cdot q^{-mn\varepsilon}, \end{aligned}$$

as desired.

By Lemma 17, for any $\varepsilon \in (0,1)$, if we can solve Mod-SIVP$_\gamma$ problem over lattice $A^\perp$ for $A \hookleftarrow U(R_q^{d\times m})$ with advantage δ, then with advantage $\geq \delta \cdot (1 - 2^{(2m+1)n} \cdot q^{-mn\varepsilon})$, we can solve Mod-SIS$_{q,\beta}^{R^d}$ with $\beta \geq \gamma \cdot n^2 \cdot m \cdot q^{\frac{d}{m}+\varepsilon}$. Combining Lemmata 16 and 17, we get the following theorem.

Theorem 5. *Let $q \nmid l$ be a prime, $m > d$ and $\varepsilon \in (0,1)$ such that $q^\varepsilon \geq 8$, there is a PPT reduction from* D-MLWE$_{q,D_\alpha}^{R^d}$ *to* Mod-SIVP$_\gamma$ *over lattice $A^\perp$ with $A \hookleftarrow U(R_q^{d\times m})$, where $\alpha < \frac{1}{8\gamma\cdot m\cdot\omega(n^3 \ln n\sqrt{\log\log n})\cdot q^{\frac{d}{m}}}$.*

In particular, if we choose $m = d\cdot\log q$, we obtain a reduction from D-MLWE$_{q,D_\alpha}^{R^d}$ to Mod-SIVP$_\gamma$ over lattice $A^\perp$ with $A \hookleftarrow U(R_q^{d\times d\log q})$, with $\frac{1}{\alpha} \approx m\cdot\gamma\cdot\tilde{O}(n^3)$. So, for $d = O(1)$, we can obtain a reduction from worst-case module SIVP$_{\tilde{O}(\gamma\cdot n^{3.75})}$ problem over K^d to average-case SIVP$_\gamma$ problem over lattice $A^\perp$, with $A \hookleftarrow U(R_q^{d\times d\log q})$.

Acknowledgement: The authors are supported by National Cryptography Development Fund (Grant No. MMJJ20180210) and National Natural Science Foundation of China (Grant No. 61832012 and No. 61672019).

A Proof of Lemma 9

Suppose q is a prime which does not ramify in R. In the ring R_q, a non-zero element $x \notin R_q^\times$ [8] if and only if there is an element $y \in R_q$ such that $x\cdot y = 0$. In fact, assume $qR = \mathfrak{q}_1\cdots\mathfrak{q}_\mathfrak{g}$ with $\mathfrak{f}\cdot\mathfrak{g} = n$, then $0 \neq x \notin R_q^\times$ if and only if $x = 0 \bmod \mathfrak{q}_i$ for some $i \in S \subsetneq [\mathfrak{g}]$ and $x \neq 0 \bmod \mathfrak{q}_j$ for others $j \in [\mathfrak{g}]\backslash S$. Then, any element y such that $y = 0 \bmod \mathfrak{q}_j$ and $y \neq 0 \bmod \mathfrak{q}_i$ will satisfy $x\cdot y = 0 \bmod qR$.

A matrix $A = [\boldsymbol{a_1}, \cdots, \boldsymbol{a_k}]^T \in R_q^{k\times k}$ is invertible in R_q if and only if $det(A) \in R_q^\times$, since in this case, there is a matrix B such that $A\cdot B = I \bmod qR$, hence $det(A)\cdot det(B) = 1 \bmod qR$ (Note that the determinant function of square matrices in $R_q^{k\times k}$, which is a special staggered k-linear map such that $det(I_k) = 1$, over the ring R_q is well defined). We call a set of vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\} \in R_q^d$ is R_q-linearly independent if $x_1\cdot\boldsymbol{a}_1 + \cdots + x_k\cdot\boldsymbol{a}_k = 0 \bmod qR$ implies $x_1 = \cdots = x_k = 0$.

We have the following useful result.

Lemma 18. *For a matrix $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k]^T \in R_q^{k\times k}$, A is invertible modulo qR if and only if $\{\boldsymbol{a}_i\}$'s are R_q-linearly independent.*

Proof. Suppose that A is invertible. If $\{\boldsymbol{a}_i\}$'s are R_q-linearly dependent, then there exist $x_1, \cdots, x_k \in R_q$ such that $\{x_i\}$'s are not all zero and $x_1\boldsymbol{a}_1 + \cdots + x_k\boldsymbol{a}_k = 0 \bmod qR$. Hence, assume without loss of generality $x_k \neq 0$, $x_k\cdot det(A) = det([\boldsymbol{a}_1, \cdots, x_k\boldsymbol{a}_k]^T) = det([\boldsymbol{a}_1, \cdots, x_1\boldsymbol{a}_1 + \cdots + x_k\boldsymbol{a}_k]^T) = 0 \bmod qR$. This means that $det(A) \notin R_q^\times$, a contradiction.

On the other hand, if these $\{\boldsymbol{a}_i\}$'s are R_q-linearly independent, we want to show $det(A) \in R_q^\times$. We prove this fact by using induction on k. For $k = 1$, it is obvious, since an element in R_q is R_q-linear independent if and only if $a \in R_q^\times$. Assume this is true for $k-1$. In the case of k, we first claim that there exits an element $a_{i,j} \in R_q^\times$ for any i or j. Otherwise, for some $i \in [k]$, $\boldsymbol{a}_i^T = [a_{i,1}, \cdots, a_{i,k}] \in (R_q\backslash R_q^\times)$, we can set $b = b_{i,1}\cdots b_{i,k}$, where $a_{i,j}\cdot b_{i,j} = 0 \bmod qR$. Then $b\cdot\boldsymbol{a}_i^T = [0, \cdots, 0] \bmod qR$ and $b \neq 0 \bmod qR$, which implies that $\{\boldsymbol{a}_i\}$'s are R_q-linearly dependent and is contradicted to our assumption. Without loss of generality, we assume $a_{1,1} \in R_q^\times$, then $det(A) = a_{1,1}\cdot det(A')$, where $A' = [\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_k]^T \in R_q^{(k-1)\times(k-1)}$ with $\boldsymbol{a}'_i = \boldsymbol{a}_i - a_{1,1}^{-1}\cdot a_{i,1}\cdot\boldsymbol{a}_1$. Meanwhile, $\{\boldsymbol{a}'_i\}$'s are R_q-linearly independent. By induction assumption, $det(A') \in R_q^\times$. Hence, we have $det(A) \in R_q^\times$, as desired.

[8] Here, $R_q^\times$ denotes the set of invertible elements in R_q.

Note that Lemma 18 implies that any $d+1$ vectors of R_q^d are R_q-linearly dependent. Since we can consider $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_d, \boldsymbol{a}_{d+1}$, if $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_d$ are linearly dependent, we have done. Otherwise, for any $x_{d+1} \in R_q$, there exist unique $x_1, \cdots, x_d \in R_q$ such that $x_1 \cdot \boldsymbol{a}_1 + \cdots, x_d \cdot \boldsymbol{a}_d = -x_{d+1} \cdot \boldsymbol{a}_{d+1}$. It also means that for any matrix $A \in R_q^{k\times k}$, the row vectors are R_q-linearly independent if and only if the column vectors are R_q-linearly independent.

Lemma 19. *For any $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i]^T \in R_q^{i\times k}$ with $i \leq k$, $\{\boldsymbol{a}_j\}$'s are R_q-linearly independent for $j \in [i]$ if and only if there exist i columns of A such that the matrix they formed are invertible.*

Proof. If there exist i columns of A such that the matrix they formed are invertible, then by Lemma 18, it is obvious that $\{a_j\}$'s are R_q-linearly independent.

On the other hand, if $\{a_j\}$'s are R_q-linearly independent, we will prove the fact by using induction on i. When $i = 1$, $A = \boldsymbol{a}_1^T \in R_q^{1\times k}$ is R_q-linearly independent if and only if there exists some $a_{i,j} \in R_q^\times$. Assume the case $i = j-1$ is true, we consider $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_j]^T$. Since $\{\boldsymbol{a}_m\}$'s are R_q-linearly independent for $m \in [j]$, there exists at least one element of $\boldsymbol{a}_1$ that is in $R_q^\times$. Assume without loss of generality $a_{1,1} \in R_q^\times$, then vectors $\boldsymbol{a}_1, \boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_j$ with $\boldsymbol{a}'_m = \boldsymbol{a}_m - a_{1,1}^{-1} \cdot a_{m,1} \cdot \boldsymbol{a}_1$ and $m \in \{2, \cdots, j\}$ are also R_q-linearly independent. In particular, vectors $\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_j$ are R_q-linearly independent. Then, there exist $j-1$ columns of $[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_j]^T$ such that the matrix they formed are invertible. Assume without loss of generality that columns from 2 to j of $[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_j]^T$ are R_q-linearly independent, then it is obvious that the first j columns of $[\boldsymbol{a}_1, \boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_j]^T$ are R_q-linearly independent, since $a_{1,1} \in R_q^\times$. Hence the first j columns of A are R_q-linearly independent(the determinant of the first j columns of A is equal to $a_{1,1} \cdot det(B) \in R_q^\times$ with B the columns from 2 to j of $[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_j]^T$), as desired.

Noticing that $R_q \cong R/\mathfrak{q}_1 \times \cdots \times R/\mathfrak{q}_{\mathfrak{g}}$, we have $\Pr_{\boldsymbol{a}\hookleftarrow U(R_q^k)}(\boldsymbol{a}$ is R_q-linearly independent$) = 1-(1-(1-\frac{1}{q^{\mathfrak{f}}})^{\mathfrak{g}})^k \geq 1-(\frac{\mathfrak{g}}{q^{\mathfrak{f}}})^k \geq 1-\frac{\mathfrak{g}}{q^{\mathfrak{f}}} \geq 1-\frac{n}{q}$ [9] when $q \geq n$.

Now we can prove the lemma we need.

Lemma 20. *For any $i \in [k-1]$ and R_q-linearly independent vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i \in R_q^k$, the probability that sample a vector $\boldsymbol{b} \hookleftarrow U(R_q^k)$ such that $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i, \boldsymbol{b}$ are R_q-linearly independent is at least $\frac{(q^{\mathfrak{f}}-1)^{\mathfrak{g}}}{q^n} \geq 1-\frac{\mathfrak{g}}{q^{\mathfrak{f}}}$.*

Proof. Given R_q-linearly independent vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i$, by lemma 19, we can assume without loss of generality that the first i columns of $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i]^T$ are R_q-linearly independent. We consider the first $i+1$ columns of A together with the first $i+1$ elements of a vector $\boldsymbol{b} \in R_q^k$. Let $B = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,i} & a_{1,i+1} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,i} & a_{2,i+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,i} & a_{i,i+1} \\ b_1 & b_2 & \cdots & b_i & b_{i+1} \end{bmatrix}$ be the corresponding matrix. By assumption, there must be some $a_{1,j} \in R_q^\times$ for $j \in [i]$. Without loss of generality, set $j = 1$. Then we can get a new matrix $B' = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,i} & a_{1,i+1} \\ 0 & a'_{2,2} & \cdots & a'_{2,i} & a'_{2,i+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a'_{i,2} & \cdots & a'_{i,i} & a'_{i,i+1} \\ b_1 & b_2 & \cdots & b_i & b_{i+1} \end{bmatrix}$ as above, such that the first i rows of B are R_q-linearly independent if and only if the first i rows of B' are R_q-linearly independent. Also note that the first i rows of B' are R_q-linearly independent if and only if the rows from 2 to i are R_q-linearly independent, thanks to the special form of B' and $\{a'_{j,m}\}$'s with $j \in \{2, \cdots, i\}$ and

[9] For general number field K, we have $\Pr_{\boldsymbol{a}\hookleftarrow U(R_q^k)}(\boldsymbol{a}$ is R_q-linearly independent$) = 1-(1-\prod_{i=1}^{\mathfrak{g}}(1-\frac{1}{q^{\mathfrak{f}_i}}))^k \geq 1-(1-(1-\frac{1}{q^{\min_i \mathfrak{f}_i}})^{\mathfrak{g}})^k \geq 1-\frac{\mathfrak{g}}{q^{\min_i \mathfrak{f}_i}} \geq 1-\frac{n}{q}$, as well.

$m \in [i+1]$. Thus, repeating the above procedure, we can get a matrix $C = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,i} & a_{1,i+1} \\ 0 & c_{2,2} & \cdots & c_{2,i} & c_{2,i+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_{i,i} & c_{i,i+1} \\ b_1 & b_2 & \cdots & b_i & b_{i+1} \end{bmatrix}$ such that the first i rows of B are R_q-linearly independent if and only if the first i rows of C are R_q-linearly independent, which also means that B is invertible if and only if C is invertible.

By Lemma 18, C is invertible if and only if the columns of $C := [\boldsymbol{d}_1, \cdots, \boldsymbol{d}_{i+1}]$ are R_q-linearly independent, also by Lemma 19 and our construction, $\boldsymbol{d}_1, \cdots, \boldsymbol{d}_i$ are R_q-linearly independent. We then modify the matrix C to the following form $D = [\boldsymbol{d}_1, \cdots, \boldsymbol{d}_i, \boldsymbol{d}'_{i+1}] = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,i} & 0 \\ 0 & c_{2,2} & \cdots & c_{2,i} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_{i,i} & 0 \\ b_1 & b_2 & \cdots & b_i & b'_{i+1} \end{bmatrix}$. This can be done easily, since by construction, elements except b_{i+1} in the diagonal of matrix C are all in $R_q^\times$. So, $\boldsymbol{d}'_{i+1} = \boldsymbol{d}_{i+1} + y_i \cdot \boldsymbol{d}_i + \cdots + y_1 \cdot \boldsymbol{d}_1$ for some $y_1, \cdots, y_i \in R_q$. Note that $x_1 \cdot \boldsymbol{d}_1 + \cdots + x_i \cdot \boldsymbol{d}_i + x_{i+1} \cdot \boldsymbol{d}_{i+1} = (x_1 - x_{i+1} \cdot y_1) \cdot \boldsymbol{d}_1 + \cdots + (x_i - x_{i+1} \cdot y_i) \cdot \boldsymbol{d}_i + x_{i+1} \cdot \boldsymbol{d}'_{i+1}$ and when $b'_{i+1} \in R_q^\times$, D is invertible. Thus, we conclude that C is invertible if $b'_{i+1} \in R_q^\times$.

Finally, notice that $b'_{i+1} = b_{i+1} + y_i \cdot b_i + \cdots + y_1 \cdot b_1 \hookleftarrow U(R_q)$ since $\{b_j\}_{j=1}^{i+1}$ are sampled uniformly and independently from R_q. We get the conclusion as desired[10].

B Missing Proofs in Subsection 3.2

Proof of Lemma 10: Given $(\boldsymbol{a}', b')$, the transformation discretizes $b' \in K_{\mathbb{R}}/R^\vee$ to $\lfloor b' \rceil_{\frac{1}{q}R^\vee} \in \frac{1}{q}R^\vee + R^\vee$. It then sets $\boldsymbol{a} = \boldsymbol{a}' \bmod qR$ and $b = \lfloor b' \rceil_{\frac{1}{q}R^\vee} \bmod R^\vee$ and outputs $(\boldsymbol{a}, b)$.

If the distribution of $(\boldsymbol{a}', b')$ is $A_{q,s,\alpha}^{R^d}$, then $b' = \frac{1}{q}\sum_{i=1}^d a'_i \cdot s_i + e' \bmod R^\vee$ for $e' \hookleftarrow D_\alpha$. Since $\frac{1}{q}\sum_{i=1}^d a'_i \cdot s_i \bmod R^\vee \in \frac{1}{q}R^\vee/R^\vee$, by validity of this discretization, we have that $\lfloor b' \rceil_{\frac{1}{q}R^\vee}$ and $\frac{1}{q}\sum_{i=1}^d a'_i \cdot s_i + \lfloor e' \rceil_{\frac{1}{q}R^\vee}$ are identically distributed. Hence, we get $(\boldsymbol{a}, b) \hookleftarrow A_{q,s,\phi}^{R^d}$.

If $(\boldsymbol{a}', b')$ is uniformly random, then by validity so is the distribution of $(\boldsymbol{a}, b)$.

References

1. Albrecht, M.R., Deo, A.: Large modulus ring-lwe $\geq$ module-lwe. In: Takagi, T., Peyrin, T. (eds.) Advances in Cryptology – ASIACRYPT 2017. pp. 267–296. Springer International Publishing, Cham (2017)
2. Alkim, E., Ducas, L., Pöppelmann, T., Schwabe, P.: Post-quantum key exchange—a new hope. In: 25th USENIX Security Symposium (USENIX Security 16). pp. 327–343. USENIX Association, Austin, TX (2016), https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/alkim
3. Applebaum, B., Cash, D., Peikert, C., Sahai, A.: Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In: Halevi, S. (ed.) Advances in Cryptology - CRYPTO 2009. pp. 595–618. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
4. Bai, S., Lepoint, T., Roux-Langlois, A., Sakzad, A., Stehlé, D., Steinfeld, R.: Improved security proofs in lattice-based cryptography: Using the rényi divergence rather than the statistical distance. Journal of Cryptology **31**(2), 610–640 (Apr 2018). https://doi.org/10.1007/s00145-017-9265-9, https://doi.org/10.1007/s00145-017-9265-9
5. Banaszczyk, W.: New bounds in some transference theorems in the geometry of numbers. Mathematische Annalen **296**(1), 625–635 (Dec 1993). https://doi.org/10.1007/BF01445125, https://doi.org/10.1007/BF01445125

[10] For general number field K, the corresponding result is $\prod_{i=1}^{\mathfrak{g}}(1 - \frac{1}{q^{\mathfrak{f}_i}}) \geq 1 - \frac{\mathfrak{g}}{q^{\min_i \mathfrak{f}_i}} \geq 1 - \frac{n}{q}$.

6. Banerjee, A., Peikert, C., Rosen, A.: Pseudorandom functions and lattices. In: Pointcheval, D., Johansson, T. (eds.) Advances in Cryptology – EUROCRYPT 2012. pp. 719–737. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
7. Bos, J., Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schanck, J.M., Schwabe, P., Seiler, G., Stehle, D.: Crystals - kyber: A cca-secure module-lattice-based kem. In: 2018 IEEE European Symposium on Security and Privacy (EuroS P). pp. 353–367 (April 2018). https://doi.org/10.1109/EuroSP.2018.00032
8. Bos, J., Costello, C., Ducas, L., Mironov, I., Naehrig, M., Nikolaenko, V., Raghunathan, A., Stebila, D.: Frodo: Take off the ring! practical, quantum-secure key exchange from lwe. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 1006–1018. CCS '16, ACM, New York, NY, USA (2016). https://doi.org/10.1145/2976749.2978425, http://doi.acm.org/10.1145/2976749.2978425
9. Brakerski, Z., Gentry, C., Vaikuntanathan, V.: (leveled) fully homomorphic encryption without bootstrapping. In: Proceedings of the 3rd Innovations in Theoretical Computer Science Conference. pp. 309–325. ITCS '12, ACM, New York, NY, USA (2012). https://doi.org/10.1145/2090236.2090262, http://doi.acm.org/10.1145/2090236.2090262
10. Brakerski, Z., Langlois, A., Peikert, C., Regev, O., Stehlé, D.: Classical hardness of learning with errors. In: Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing. pp. 575–584. STOC '13, ACM, New York, NY, USA (2013). https://doi.org/10.1145/2488608.2488680, http://doi.acm.org/10.1145/2488608.2488680
11. Brakerski, Z., Tsabary, R., Vaikuntanathan, V., Wee, H.: Private constrained prfs (and more) from lwe. In: Kalai, Y., Reyzin, L. (eds.) Theory of Cryptography. pp. 264–302. Springer International Publishing, Cham (2017)
12. Cramer, R., Ducas, L., Peikert, C., Regev, O.: Recovering short generators of principal ideals in cyclotomic rings. In: Fischlin, M., Coron, J.S. (eds.) Advances in Cryptology – EUROCRYPT 2016. pp. 559–585. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)
13. Cramer, R., Ducas, L., Wesolowski, B.: Short stickelberger class relations and application to ideal-svp. In: Coron, J.S., Nielsen, J.B. (eds.) Advances in Cryptology – EUROCRYPT 2017. pp. 324–348. Springer International Publishing, Cham (2017)
14. Deshpande, A., Koppula, V., Waters, B.: Constrained pseudorandom functions for unconstrained inputs. In: Fischlin, M., Coron, J.S. (eds.) Advances in Cryptology – EUROCRYPT 2016. pp. 124–153. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)
15. Ducas, L., Durmus, A., Lepoint, T., Lyubashevsky, V.: Lattice signatures and bimodal gaussians. In: Canetti, R., Garay, J.A. (eds.) Advances in Cryptology – CRYPTO 2013. pp. 40–56. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
16. Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schwabe, P., Seiler, G., Stehl, D.: Crystals-dilithium: A lattice-based digital signature scheme (01 2018)
17. Ducas, L., Lyubashevsky, V., Prest, T.: Efficient identity-based encryption over ntru lattices. In: Sarkar, P., Iwata, T. (eds.) Advances in Cryptology – ASIACRYPT 2014. pp. 22–41. Springer Berlin Heidelberg, Berlin, Heidelberg (2014)
18. Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing. pp. 197–206. STOC '08, ACM, New York, NY, USA (2008). https://doi.org/10.1145/1374376.1374407, http://doi.acm.org/10.1145/1374376.1374407
19. Kim, S., Wu, D.J.: Watermarking cryptographic functionalities from standard lattice assumptions. In: Katz, J., Shacham, H. (eds.) Advances in Cryptology – CRYPTO 2017. pp. 503–536. Springer International Publishing, Cham (2017)
20. Kim, S., Wu, D.J.: Watermarking PRFs from lattices: Stronger security via extractable PRFs. In: CRYPTO (2019)
21. Langlois, A., Stehlé, D.: Worst-case to average-case reductions for module lattices. Designs, Codes and Cryptography **75**(3), 565–599 (Jun 2015). https://doi.org/10.1007/s10623-014-9938-4, https://doi.org/10.1007/s10623-014-9938-4
22. Lindner, R., Peikert, C.: Better key sizes (and attacks) for lwe-based encryption. In: Kiayias, A. (ed.) Topics in Cryptology – CT-RSA 2011. pp. 319–339. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)
23. Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. In: Gilbert, H. (ed.) Advances in Cryptology – EUROCRYPT 2010. pp. 1–23. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)

24. Lyubashevsky, V., Peikert, C., Regev, O.: A toolkit for ring-lwe cryptography. In: Johansson, T., Nguyen, P.Q. (eds.) Advances in Cryptology – EUROCRYPT 2013. pp. 35–54. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
25. Micciancio, D., Regev, O.: Worst-case to average-case reductions based on gaussian measures. SIAM J. Comput. **37**(1), 267–302 (Apr 2007). https://doi.org/10.1137/S0097539705447360, http://dx.doi.org/10.1137/S0097539705447360
26. Micciancio, D., Regev, O.: Lattice-based Cryptography, pp. 147–191. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
27. Peikert, C.: Limits on the hardness of lattice problems in l_p norms. In: Proceedings of the Twenty-Second Annual IEEE Conference on Computational Complexity. pp. 333–346. CCC '07, IEEE Computer Society, Washington, DC, USA (2007). https://doi.org/10.1109/CCC.2007.12, https://doi.org/10.1109/CCC.2007.12
28. Peikert, C.: Public-key cryptosystems from the worst-case shortest vector problem: Extended abstract. In: Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing. pp. 333–342. STOC '09, ACM, New York, NY, USA (2009). https://doi.org/10.1145/1536414.1536461, http://doi.acm.org/10.1145/1536414.1536461
29. Peikert, C.: An efficient and parallel gaussian sampler for lattices. In: Rabin, T. (ed.) Advances in Cryptology – CRYPTO 2010. pp. 80–97. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
30. Peikert, C., Regev, O., Stephens-Davidowitz, N.: Pseudorandomness of ring-lwe for any ring and modulus. In: Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. pp. 461–473. STOC 2017, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3055399.3055489, http://doi.acm.org/10.1145/3055399.3055489
31. Peikert, C., Shiehian, S.: Privately constraining and programming prfs, the lwe way. In: Abdalla, M., Dahab, R. (eds.) Public-Key Cryptography – PKC 2018. pp. 675–701. Springer International Publishing, Cham (2018)
32. Pellet-Mary, A., Hanrot, G., Stehlé, D.: Approx-svp in ideal lattices with pre-processing. In: Ishai, Y., Rijmen, V. (eds.) Advances in Cryptology – EUROCRYPT 2019. pp. 685–716. Springer International Publishing, Cham (2019)
33. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. J. ACM **56**(6), 34:1–34:40 (Sep 2009). https://doi.org/10.1145/1568318.1568324, http://doi.acm.org/10.1145/1568318.1568324
34. Rosca, M., Stehlé, D., Wallet, A.: On the ring-lwe and polynomial-lwe problems. In: Nielsen, J.B., Rijmen, V. (eds.) Advances in Cryptology – EUROCRYPT 2018. pp. 146–173. Springer International Publishing, Cham (2018)
35. Stehlé, D., Steinfeld, R.: Making ntruencrypt and ntrusign as secure as standard worst-case problems over ideal lattices. Cryptology ePrint Archive, Report 2013/004 (2013), https://eprint.iacr.org/2013/004
36. Stehlé, D., Steinfeld, R., Tanaka, K., Xagawa, K.: Efficient public key encryption based on ideal lattices. In: Matsui, M. (ed.) Advances in Cryptology – ASIACRYPT 2009. pp. 617–635. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
37. Wang, Y., Wang, M.: Provably secure ntruencrypt over any cyclotomic field. In: Cid, C., Jacobson Jr., M.J. (eds.) Selected Areas in Cryptography – SAC 2018. pp. 391–417. Springer International Publishing, Cham (2019)