Module-LWE versus Ring-LWE, Revisited

Yang Wang\textsuperscript{1,2} and Mingqiang Wang\textsuperscript{1,2}

\textsuperscript{1}School of Mathematics, Shandong University, Jinan, Shandong, 250100, P.R. China
\textsuperscript{2}Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education

\texttt{wyang1114@mail.sdu.edu.cn}
\texttt{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 \textit{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 $\Gamma \approx \tilde{O}(n^{-2})$ in average-case, hence, we get a reduction from worst-case module approximate shortest independent vectors problem (SIVP$_\gamma$) with approximation parameter $\gamma \approx \tilde{O}(n^7)$ to corresponding average-case decision RLWE problems. Meanwhile, our result shows that the search variant reductions of Albrecht \textit{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$_\gamma$ over any cyclotomic field as improvements of relative results showed by Rosca \textit{et al.} in EUROCRYPT 2018 and Langlois \textit{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$_\gamma$) 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$_\gamma$) to plain LWE with exponential modulus. Combining the modulus-switch techniques, Brakerski \textit{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, cryptographic 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, deteriorating the LWE error rate by a small polynomial factor. As a result, for any cyclotomic field $K = \mathbb{Q}(\zeta_l)$ with $\zeta_l$ 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 $q \approx \tilde{O}(n^{5.75})$ and error rate $\alpha \approx \tilde{O}(n^{-4.25})$ over $K^d$. Combining the known reduction from
module SIVP, to decision MLWE [21, 30], we conclude a reduction from worst-case module SIVP, with γ ≈ 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' [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'.

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 p to decision RLWE problem with modulus q and error rate α' = α \cdot \frac{p}{q} \cdot poly(n) for some small poly(n). Then our reduction could be used to reduce a decision RLWE problem with some split 'well' polynomially bounded modulus p to a decision RLWE problem with arbitrary polynomially bounded modulus q that is relatively close to p, for example p = 1 mod l and \frac{p}{q} = 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, problem over any cyclotomic field. We prove that if one could solve the module SIVP, problem in lattice \( A^k := \{ z \in \mathbb{R}^m : A \cdot z = 0 \mod qR^d \} \) for some \( m > d \) and \( A \leftarrow U(R_q^{d \times m}) \) with non-negligible probability, he can also solve the average-case decision MLWE problem with error rate \( \alpha \approx O(\frac{1}{\gamma \cdot m \cdot n \cdot q^m}) \). For the usual case \( d = O(1) \), by taking \( m = d \cdot \log q \), we obtain a reduction from decision MLWE with error rate \( \alpha \approx O(\frac{1}{\gamma \cdot n^3 \cdot q^m}) \) to average-case module SIVP, over lattice \( A^k \), with \( A \leftarrow U(R_q^{d \times d \log q}) \). As a corollary, we obtain a reduction from worst-case module SIVP, problem over \( K^d \) to average-case SIVP, problem over lattice \( A^k \) with \( A \leftarrow U(R_q^{d \times d \log q}) \).

### 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 decision MLWE and decision RLWE through the path: decision MLWE→search MLWE→search RLWE→decision RLWE. Many details need to be treated carefully and the reduction road-map are summarized in Figure [1].

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

We start from D-MLWE, for some continuous Gaussian distribution \( D_o \) 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 \( D_{\leq \alpha} := \{ D_r : \beta_k \leq \alpha \text{ and } \beta_k = \alpha_{n+1-k} \text{ for all } k \in \{1, \cdots, \frac{n}{2}\} \} \). We first need to discretize the errors to a lattice in \( K_R = K \otimes \mathbb{Q} \mathbb{R} \). We choose to displace the errors to lattice \( \frac{1}{q} R^{\sqrt{q}} \). This can be done easily by using the fact, which is showed in [29], that for any \( e \leftarrow D_{\alpha} \) and \( f \leftarrow D_{\frac{1}{q} R^{\sqrt{q}} - e_{\alpha}} \) with \( \beta \geq \eta_{\epsilon}(\frac{1}{q} R^{\sqrt{q}}) \), we have \( e + f \leq D_{\frac{1}{q} R^{\sqrt{q}} \cdot \sqrt{\alpha^2 + \beta^2}} \). Then we consider the normal form of corresponding MLWE problems (denoted by Nor-D-MLWE), i.e. the secret \( s \) and the error \( e \) obey the same distribution. Note that for a MLWE sample \((a, b)\) with \( b = b_{\alpha}^T \cdot s + e \mod R^{\sqrt{q}} \) for some secret \( s \in R_q^{\alpha} \) and error \( e \leftarrow D_{\alpha} \), we can represent it as the form \( b = a^T \cdot s' + \frac{1}{q} s \mod R^{\sqrt{q}}/R^{\sqrt{q}} \). Hence, transformation used in [3] may also work if we could construct an invertible matrix \( A \in R_q^{\alpha \times d} \) which is consist of the \( a \) components, when given polynomial many samples. Fortunately, for \( q = \Omega(n) \), we could construct this 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 \( e' = b - A \cdot s \mod R^{\sqrt{q}} \) has small norm, where \( s \) is the output of the Nor-S-MLWE oracle and \( A \in R_q^{\alpha \times d} \) is the matrix formed by the \( a \) components of given samples. In this reduction, we
Theorem 5.6 of [34]  

Lemma 7.2 of [30]  

Fig. 1. Reduction road-map from decision MLWE to average-case decision RLWE
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 \( R^\ast \) to estimate the probability \( \Pr_{\bf{b} \in U((l \cdot R^\ast) / \mathcal{R})^m)} |\exists \bf{s} \in (\frac{l}{q} \cdot R^\ast / \mathcal{R})^d : \max_{1 \leq k \leq m} ||e_k^t|| < B, e^t = \bf{b} - A \cdot \bf{s} \mod \mathcal{R}^\ast | \) where \( B \) is some suitable upper-bound. For suitable \( B \), if \( \bf{b} \) is chosen uniformly at random, there will be at least one \( e_k^t \) with norm larger than \( B \). While, if \( \bf{b} \) comes from MLWE distributions, the norm of all \( e_k^t \) 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 \( \max_{1 \leq k \leq m} ||e_k^t|| \) to\(^3\) amend the error distribution to elliptical Gaussian of corresponding S-RLWE problem. In this step, we need to bound the quantity \( \max_{1 \leq k \leq m} ||e_k^t|| \leq \frac{1}{\sqrt{l^2+q^2}} \) for \( s \in D_{\frac{1}{q} \cdot R^\ast / \mathcal{R}^\ast} \). Our estimate also shows that the direct reduction of search variants in [1] also works for all cyclotomic fields. We 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 \( \frac{2n}{\delta} \). 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\(_{q_1, D, n}^d \) to S-RLWE\(_{q_2, D, n}^d \), with \( \alpha'' \approx O(\alpha' \cdot n^{2.5} \cdot \frac{q_1}{q_2}) \). In the above, we reduce D-MLWE\(_{q_1, D, n}^d \) to worst-case S-RLWE\(_{q_1, D, n}^d \). 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 \( \text{poly}(n) \) bounded) and \( \frac{q}{p} \leq \text{poly}(n) \). Such \( p \) admits reductions from S-RLWE to D-RLWE by using the same method used in [23].

We can 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 \( q = 1 \mod l \)), we could directly use reductions showed in [23] to reduce S-RLWE\(_{q^d, D, n}^d \) to average-case D-RLWE\(_{q^d, D, n}^d \) for some small polynomially bounded \( \tau \in \mathbb{R} \), hence reduce D-MLWE\(_{q_1, D, n}^d \) to average-case D-RLWE\(_{q^d, D, n}^d \). These special cases have already covered many usual applications, including the examples we give:-KCL, CRYSTALS-KYBER and CRYSTALS-DILITHIUM.

Reduction from D-RLWE 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 \( \mathcal{R}^d \) (denoted by M-SIS\(_{q_1^d, \beta} \)) are defined as follows: given \( A 
Rightarrow U(\mathcal{R}_q^{n \times d}) \), find a nonzero vector \( \bf{z} \in \mathcal{R}^m \) such that \( \bf{z}^T \cdot A = 0 \mod q \mathcal{R}^d \) and \( ||\bf{z}|| \leq \beta \) for some target norm \( \beta \). We first reduce D-MLWE\(_{q_1^d, D, n}^d \) to M-SIS\(_{q_1^d, \beta} \) with \( \alpha \approx O(\frac{1}{\beta n}) \). Essentially, when give a short vector \( \bf{z} \) such that \( \bf{z}^T \cdot A = 0 \mod q \mathcal{R}^d \) and \( m \) sample \( (A, \bf{b}) \in \mathcal{R}^m_{q^d} \times (\mathcal{R}^d / \mathcal{R})^m \), we can represent \( \bf{z}^T \cdot \bf{b} \mod \mathcal{R}^d \) with respect to the decoding basis. Then, if \( \bf{b} \) is distributed uniformly at random, the coefficients of \( \bf{z}^T \cdot \bf{b} \mod \mathcal{R}^d \) will also be distributed randomly in the interval \( (-\frac{1}{2}, \frac{1}{2}) \). On the other hand, if \( \bf{b} = A \cdot \bf{s} + \bf{e} \) for some \( \bf{e} \sim D_{\frac{1}{q} \cdot \mathcal{R}^d} \), with high probability, the coefficients of \( \bf{z}^T \cdot \bf{b} \mod \mathcal{R}^d \) would be much closer to 0. Solving M-SIS\(_{q_1^d, \beta} \) can be converted to solving module SIVP\(_{\gamma} \) problem over the lattice \( A^\perp := \{ \bf{z} \in \mathcal{R}^d : \bf{z}^T \cdot A = 0 \mod q \mathcal{R}^d \} \) with \( \beta \leq \gamma \cdot \lambda_{n}(A^\perp) \). By the transference theorem, \( \lambda_{n}(A^\perp) \leq \frac{n \cdot d}{\lambda_{n}(A^\perp \gamma)} \leq \frac{n \cdot d}{\lambda_{n}(A^\perp \gamma)} \), where \( (A^\perp)^\Gamma \) denotes the dual lattice of \( A^\perp \). In fact, \( (A^\perp)^\Gamma \) is equal to \( \frac{1}{q} \{ \bf{y} \in (\mathcal{R}^d)^m \) : \exists \bf{s} \in (\mathcal{R}^d)^d, A \cdot \bf{s} = \bf{y} \mod q(\mathcal{R}^d)^m \} \).

We prove that for \( A \nRightarrow U(\mathcal{R}_q^{n \times d}) \), the lattice \( (A^\perp)^\Gamma \) 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, 38] 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 decoding basis of \( R^\ast \). Essentially, we use the property that the singular values of (one of) the basis matrix of lattice \( R \) of cyclotomic fields are well bounded \(^3\). We use this to discretize the errors in Subsection 3.2, to bound the

---

\(^3\) This also means that the singular values of (one of) the basis matrix of lattice \( R^\ast \) of cyclotomic fields are well bounded.
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, 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 $s_k(M)$ to denote the singular values of $M$ for $k \in [n]$. We will re-arrange singular values and assume $s_1(M) \geq \cdots \geq s_n(M)$. Matrix $I_n$ denotes the matrix $\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}_{n \times n}$ and matrix $J_n$ denotes the matrix $\begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 0 \end{pmatrix}_{n \times n}$. When we write $X \leftarrow \xi$, we mean the random variable $X$ obeys to the distribution $\xi$. 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

Throughout 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, 38].

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

$$h_j = \frac{1}{\sqrt{2}} (e_j + e_{n+1-j})$$

$$h_{n+1-j} = \frac{i}{\sqrt{2}} (e_j - e_{n+1-j}),$$

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

The discriminant $\Delta_K$ 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 $\mathbb{R}$, 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

$$\Delta_K = (-1)^{\frac{1}{2}} \cdot \left( \frac{l}{\prod_p ||p||^{\frac{1}{p}}} \right)^n \leq n^n, \quad (1)$$
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 \( \Lambda^\perp = \{ y \in H : \forall x \in \Lambda, \langle x, y \rangle = \sum_{i=1}^n x_i y_i \in \mathbb{Z} \} \). 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 \( I^\perp = \{ a \in K : \text{Tr}(a \cdot I) \subseteq \mathbb{Z} \} \). Then, it is easy to verify that \( (I^\perp)^\perp = I \), \( I^\perp \) is a fractional ideal and \( I^\perp \) embeds under \( \sigma \) as the dual lattice of \( I \) as defined above. Recall that we have \( |\Delta_K| = \text{det}(\sigma(R))^2 \), the squared determinant of the lattice \( \sigma(R) \).

The algebraic norm of a non-zero integral ideal \( J \) is defined as \( N(J) = |R/J| \). 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 \( N(I) = \frac{N(I_1)}{N(I_2)} \) with \( I_1 = J_1, J_2 \subseteq R \) and \( J_1 + J_2 = R \). We also have \( \text{det}(\sigma(I)) = N(I)^{1/\sqrt{|\Delta_K|}} \).

The following lemma [29] gives upper and lower bounds on the minimum distance of an ideal lattice in \( l_2 \) norm and \( l_{\infty} \) norm.

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

\[
\sqrt{n} \cdot N^{1/2}(I) \leq \lambda_1(I) \leq \sqrt{n} \cdot N^{1/2}(I) \cdot |\Delta_K|^{1/n}
\]

and

\[
N^{1/2}(I) \leq \lambda_1^\perp(I) \leq N^{1/2}(I) \cdot |\Delta_K|^{1/n}.
\]

### 2.3 Gaussian Distributions and Rényi Divergence

The Gaussian distribution is defined as usual. For any \( s > 0 \), \( c \in H \), which is taken to be \( s = 1 \) or \( c = 0 \) when omitted, define the (spherical) Gaussian function \( \rho_{s,c} : H \to (0, 1) \) as \( \rho_{s,c}(x) = e^{-\pi \|x-c\|^2/s^2} \). By normalizing this function, we obtain the continuous Gaussian probability distribution \( D_{s,c} \) of parameter \( s \), whose density function is given by \( s^{-n} \cdot \rho_{s,c}(x) \).

For a real vector \( r = (r_1, \cdots, r_n) \in (\mathbb{R}^+)^n \) which satisfies \( r_k = r_{n+1-k} \) for \( k \in [\frac{n}{2}] \), we define the elliptical Gaussian distributions in the basis \( \{ x_1, \cdots, x_n \} \) as follows: a sample from \( D_r \) is given by \( \sum_{i \in [n]} x_i h_i \), 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 \( \varphi(\sum_{i \in [n]} x_i h_i) = (x_1, \cdots, x_n) \), then \( D_r \) 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 \( x \sim D_{B,c} \) (or \( x \sim D_{\Sigma,c} \)) for some \( c \in \mathbb{R}^n \) if the density function of \( x \) is given by \( \frac{1}{\sqrt{\det(\Sigma)}} e^{-\pi \|x-c\|^2/\Sigma} \). It is easy to check that if \( B = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \), then \( D_{B,c} = D_r \) with \( r = (r_1, \cdots, r_n) \in (\mathbb{R}^+)^n \). 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 \( \Lambda \) and coset vector \( c \in \mathbb{R}^n \) with parameter \( s \) is denoted by \( D_{\Lambda + c, s} \) with density function \( \frac{\rho_s(x)}{\rho_s(\Lambda + c)} \), where \( \rho_s(\Lambda + c) = \sum_{x \in \Lambda + c} \rho_s(x) \). It was showed in [10] that we can sample a discrete Gaussian distribution efficiently.

\[4\text{ For any } a \in K, \text{ we define } \text{Tr}(a) = \sum_{i=1}^n \sigma_i(a) \text{ and } N(a) = \prod_{i=1}^n \sigma_i(a).\]
Lemma 2. There is a probabilistic polynomial-time algorithm that, given a basis $B$ of an $n$-dimensional lattice $\Lambda = L(B) \subseteq \mathbb{R}^n$, $c \in \mathbb{R}^n$ and a parameter $s \geq ||B|| \cdot \sqrt{\frac{\ln(2n+4)}{\pi}}$, outputs a sample distributed according to $D_{A+c,s}$.

Here, $\hat{B}$ is the Gram-Schmidt orthogonalization of basis $B = \{b_1, \cdots, b_n\}$ and $||\hat{B}||$ 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 $\text{Supp}(P) \subseteq \text{Supp}(Q)$, the Rényi divergence of $P$ and $Q$ of order $a \in [1, \infty]$ is given by

$$R_a(P||Q) = \begin{cases} \sum_{x \in \text{Supp}(P)} P(x)^{1-a} & \text{for } a = 1 \\ \frac{\left(\sum_{x \in \text{Supp}(P)} P(x)^{a} Q(x)^{1-a}\right)^{\frac{1}{a-1}}}{\max_{x \in \text{Supp}(P)} P(x)^{a}} & \text{for } a > 1 \end{cases}$$

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 $\text{Supp}(P) \subseteq \text{Supp}(Q)$. 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 $a \to \infty$.
- **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 $a \in [1, \infty]$.
  - $R_a(P||Q) \leq R_\infty(P_1||Q_1) \cdot \max_{y_1 \in \text{Supp}(P_1)} R_a(P_2|y_1)(||Q_{2|1}(\cdot|y_1)||)$.
- **Probability Preservation:** Let $E \subseteq \text{Supp}(Q)$ be an arbitrary event. If $a \in (1, \infty)$, then $Q(E) \geq \frac{P(E)^{\frac{1}{a}}}{R_a(P||Q)}$. Furthermore, we have $Q(E) \geq \frac{P(E)^{\frac{1}{a}}}{R_a(P||Q)}$.
- **Weak Triangle Inequality:** Let $P_1$, $P_2$ and $P_3$ be three probability distributions such that $\text{Supp}(P_1) \subseteq \text{Supp}(P_2) \subseteq \text{Supp}(P_3)$. Then

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

Recall that for a lattice $\Lambda$ and positive real $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest real $s > 0$ such that $\rho_\varepsilon(\Lambda^\vee \setminus \{0\}) \leq \varepsilon$. 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 $\frac{1}{\lambda(\Lambda^\vee)} \leq \eta_\varepsilon(\Lambda) \leq \sqrt{\frac{\ln(2n+4)}{\pi}} \cdot \max\{||\hat{B}||, \lambda_n(\Lambda), \frac{1}{\lambda^3(\Lambda^\vee)}\}$.

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

Here, $B_n$ denotes the unit open ball.
Lemma 6. For any $n$-dimensional lattice $\Lambda$, $\varepsilon > 0$, $s \geq \eta_\varepsilon(\Lambda)$ and $c \in \mathbb{R}^n$, we have
\[
\rho_s(A + c) \in \left[1 - \varepsilon, \frac{1}{1 + \varepsilon}\right] \cdot \rho_s(A).
\]

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

2.4 Ring-LWE and Module-LWE Problems

Let $K_R := K \otimes \mathbb{Z}/q \cong H$, $T_{R^\vee} = K_R/R^\vee$, $R_q = R/(qR)$ 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 $\psi$ be some distribution over $H$,

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

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

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

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

Usually, the error distribution $\psi$ may be chosen from a family of distributions $\Psi$ over $H$. Let’s take the Ring-LWE problem for an example. When the error distribution $\psi$ is sampled from a family of distributions $\Psi$ over $K_R$, we call an algorithm solve the worst-case search (or decision) problems if it solves corresponding problems with probability $\nu = 1$ 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 $(s, \psi) \leftarrow U(R_q^\vee \times \mathcal{D})$ for some distribution $\mathcal{D}$ over $\Psi^6$. The detailed definition of $\mathcal{D}$, which is denoted by $\mathcal{Y}_\alpha$, is the worst-case to average-case reductions of corresponding LWE problems can be found in [21, 23, 30]. We just remark that $\psi$ can be modified to be some spherical Gaussian distribution over $K_R$ [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 $D_{\alpha' \leq \alpha}$ to denote the set of elliptical Gaussian distributions $D_\alpha$ with $\alpha' \leq r_i \leq \alpha$. We write $D_{\leq \alpha}$ 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 problem briefly.

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

---

6 We also classify the cases, where the secret $s \leftarrow U(R_q^\vee)$ and the error distribution $\psi \in \Psi$ is arbitrary, into worst-case variants of corresponding problems.
2.5 Basis for $R$ and $R'$ in Cyclotomic Fields

In some of our reductions, we hope that the matrices whose columns are consisted of the basis of $R$ or $R'$ have smaller $s_1$ and larger $s_n$. In cyclotomic fields, there are good bases of $R$ and $R'$ with very nice magnitudes of singular values. So, we introduce the powerful basis and the decoding basis as in [24]. We set $\tau$ be the automorphism of $K$ that maps $\zeta_l$ to $\zeta_l^{-1} = \zeta_{l^{-1}}$, 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 $(\zeta_{l^j})_{j \in \{0,1,\ldots,n-1\}}$, treated as a vector over $R \subset K$.
- For $l$ having prime-power factorization $l = \prod l_k = \prod l_k^{a_k}$, 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 $R'$ 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 $R'$) are connected by some unimodular matrice, hence the spectral norm (i.e. the $\|\cdot\|$) may have different magnitudes. The following lemma comes from [24], which shows the estimates of $s_1(\sigma(\overrightarrow{p}))$ and $s_n(\sigma(\overrightarrow{p}))$. Define $\text{rad}(l) = \prod_{p \mid l} p$ and

$$
\hat{l} = \begin{cases} 
\lfloor \frac{l}{2} \rceil, & \text{if } l \text{ is odd,} \\
\frac{l}{2}, & \text{if } l \text{ is even.}
\end{cases}
$$

**Lemma 8.** We have $s_1(\sigma(\overrightarrow{p})) = \sqrt{l}$, $s_n(\sigma(\overrightarrow{p})) = \sqrt{\frac{l}{\text{rad}(l)}}$, $||\sigma(\overrightarrow{p})||_i = 1$ and $||\sigma(\overrightarrow{p})|| = \sqrt{n}$ for all $i = 1, \ldots, n$.

We also need the estimates of $s_1(\sigma(\overrightarrow{d}))$ and $s_n(\sigma(\overrightarrow{d}))$. Assume that $\sigma(\overrightarrow{p}) = T$, Lemma 8 shows that $s_1(T) = \sqrt{l}$ and $s_n(T) = \sqrt{\frac{l}{\text{rad}(l)}}$. By the definitions of $\overrightarrow{d}$ and the dual ideal, an easy computation shows that $\sigma(\overrightarrow{d}) = (T^*)^{-1}$. Hence we have $s_n(\sigma(\overrightarrow{d})) = \frac{1}{\sqrt{l}}$, $s_1(\sigma(\overrightarrow{d})) = \sqrt{\frac{\text{rad}(l)}{l}}$. Moreover, one can similarly deduce that $||\sigma(\overrightarrow{d})||_i \leq \sqrt{\frac{\text{rad}(l)}{l^2}}$ for all $i = 1, 2, \ldots, n$. 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 $R'$) with respect to the powerful basis (or decoding basis), we have

$$
\sqrt{\frac{l}{\text{rad}(l)}} \cdot ||x||_{\sigma(\overrightarrow{p})}^c \leq ||\sigma(x)|| \leq \sqrt{l} \cdot ||x||_{\sigma(\overrightarrow{p})}^c, \quad \text{for } x \in R,
$$

and

$$
\frac{1}{\sqrt{l}} \cdot ||x||_{\sigma(\overrightarrow{d})}^c \leq ||\sigma(x)|| \leq \sqrt{\frac{\text{rad}(l)}{l}} \cdot ||x||_{\sigma(\overrightarrow{d})}^c, \quad \text{for } x \in R'.
$$

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 understand the transformations in the following reductions.

Assume that matrix $G \in R^{d \times d} : R^d \rightarrow R^d$ induces a map, we consider the corresponding map $G_H : \sigma(R^d) \rightarrow \sigma(R^d)$, i.e. for any $x = (x_1, \cdots, x_d)^T \in R^d$, we require that $\sigma(y) = G_H \cdot \sigma(x) \in H^d$, where $y = G \cdot x$. If

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

we define

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

Then, it is easy to verify that $\sigma(y) = G_H \cdot \sigma(x)$. The same calculation shows that the map $\sigma_H : R^{d \times d} \rightarrow \mathbb{C}^{nd \times nd}$ 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 $\varphi : K \rightarrow \mathbb{R}^n$ is the composite of the canonical embedding $\sigma$ and the isomorphism $H \cong \mathbb{R}^n$, we now decide the corresponding matrix $G_\mathbb{R}$ of $G$ such that $\varphi(y) = G_\mathbb{R} \cdot \varphi(x)$. For any element $x \in K$, we have $\varphi(x) = U \cdot \sigma(x)$ with $U = \begin{pmatrix} \frac{1}{\sqrt{2}} \cdot I_r & \frac{1}{\sqrt{2}} \cdot J_r \\ \frac{-1}{\sqrt{2}} \cdot J_r & \frac{1}{\sqrt{2}} \cdot I_r \end{pmatrix}$ (note that $U^{-1} = U^*`). Hence,

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

which implies that $G_\mathbb{R} = \begin{pmatrix} U & \cdots & U \end{pmatrix} \cdot G_H \cdot \begin{pmatrix} U^{-1} & \cdots & U^{-1} \end{pmatrix}$. 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 $\varphi \in \mathbb{R}^n$. 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 $x_H = \begin{pmatrix} \sigma_1(x) \\ \vdots \\ \sigma_n(x) \end{pmatrix}$. Note that

$$x_\mathbb{R} \cdot x_\mathbb{R}^T = x_\mathbb{R} \cdot x_\mathbb{R}^T = U \cdot x_H \cdot x_H^T \cdot U^{-1} = \begin{pmatrix} |\sigma_1(x)|^2 & \cdots & |\sigma_n(x)|^2 \end{pmatrix},$$

given by $|\sigma_i(x)|$ for $i \in [n]$. Then, for any $s \in K$, if $x \leftrightarrow D_B$ for some nonsingular matrix $B$ with
\[ \Sigma = B \cdot B^T, \text{ then } s \cdot x \leftarrow D_{\sqrt{\Sigma}} \text{ with } \Sigma' = s_R \cdot \Sigma \cdot s_R^T. \text{ In particular, if } B = \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix} \text{ with } r_k = r_{n+1-k} \]

for all \( k \in \left[ \frac{n}{2} \right] \), then \( s \cdot x \leftarrow D_{\Sigma'} \) with \( B' = \begin{pmatrix} r_1 \cdot |\sigma_1(s)| & \cdots & r_n \cdot |\sigma_n(s)| \end{pmatrix} \).

Suppose \( q \) is a prime which does not ramify in \( R \) (equivalently, \( q \nmid l \) in our settings), meanwhile, \( qR = q_1 \cdots q_s \) with \( g \cdot f = n \). We have \( N(q_i) = q^i \) and \( R_q \cong R/q_1 \times \cdots \times R/q_s \). 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.

**Lemma 9.** For any positive integer \( d \) and non-ramified prime \( q \) with the prime ideal decomposition as above, we have \( \Pr_{A \leftarrow U(R_q^d \times \mathcal{A})}[A \text{ is invertible } \mod qR] \geq (1 - \frac{1}{q!})^d \geq (1 - \frac{2}{q})^d \geq (1 - \frac{n}{q})^d \).

**Remark 1.** Results in this subsection can be easily modified to general algebraic number fields. The only difference is that in general fields, the \{f, g\}'s may not equal to each other. However, similar deduction implies that we still have the same lower bound \((1 - \frac{2}{q})^d \) 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 \( R^\vee \), we need to discretize the errors. The discretized distribution would also be used in Subsection 3.3.

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

Assume that D-MLWE\(_{\mathcal{A},D_n}\) is hard for some distribution \( D_n \) over \( K_R/R^\vee \), let \( \phi = [D_n]_{\frac{1}{2}R^\vee} \) for some valid discretization \([\cdot]_{\frac{1}{2}R^\vee} \). We can show that D-MLWE\(_{\mathcal{A},\phi}\) 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 \((a', b') \in R_q^d \times K_R/R^\vee\), outputs a pair \((a, b) \in R_q^d \times \frac{1}{q}R^\vee/R^\vee\) 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 \( A_{\mathcal{A},s,D_n}^R \), then the output pair is distributed according to \( A_{\mathcal{A},s,\phi}^R \).

Next, we show that D-MLWE is also hard when the secret \( s \) is distributed as the error \( e \). We denote this kind of D-MLWE problem by Nor-D-MLWE (whose corresponding distribution is denoted by \( A_{\mathcal{A},s,\phi}^R \)), i.e. a sample of \( A_{\mathcal{A},s,\phi}^R \) is of the form \((a, b) \) with \( a \leftarrow U(R_q) \) and \( b = aT \cdot s + e \mod R^\vee \), where \( s_i, e \leftarrow \phi = [D_n]_{\frac{1}{2}R^\vee} \) for \( i \in [d] \).

For \( A \leftarrow U(R_q^d \times \mathcal{A}) \), Lemma 9 shows that with probability larger than \((1 - \frac{1}{q!})^d \), \( A \) is invertible \( \mod qR \). When \( q^l = O(g \cdot d) \), 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 \( \approx 1 \). Assume we have \( d \) samples of the form \((A, b) \in R_q^d \times \frac{1}{q}(\frac{1}{q}R^\vee/R^\vee)^d \), where \( A \) is invertible and \( b = A \cdot s' + e \) for some \( s', e \in (\frac{1}{q}R^\vee/R^\vee)^d \) with \( s' = \frac{1}{q}s \). Note that given \( A \) is equivalent to given \( A_H \), the \( b \)-component of MLWE distribution is
\[ \sigma(b) = A_H \cdot \sigma(s') + \sigma(e), \] i.e.

\[
\sigma(b) = \begin{bmatrix}
\sigma_1(a_{1,1}) & \cdots & \sigma_1(a_{1,d}) \\
\vdots & \ddots & \vdots \\
\sigma_n(a_{1,1}) & \cdots & \sigma_n(a_{1,d}) \\
\vdots & \cdots & \vdots \\
\sigma_1(a_{d,1}) & \cdots & \sigma_1(a_{d,d}) \\
\vdots & \cdots & \vdots \\
\sigma_n(a_{d,1}) & \cdots & \sigma_n(a_{d,d})
\end{bmatrix}
\begin{bmatrix}
\sigma_1(s'_1) \\
\vdots \\
\sigma_n(s'_1) \\
\vdots \\
\sigma_1(s'_d) \\
\vdots \\
\sigma_n(s'_d)
\end{bmatrix}
+ \begin{bmatrix}
\sigma_1(e_1) \\
\vdots \\
\sigma_n(e_1)
\end{bmatrix}.
\]

For another new sample \((a, b) \leftarrow A_{R^d_q, \phi}^R\), we set \((a', b')\) as \((a')^T = -a^T \cdot A^{-1} \mod qR\) and \(b' = b + (a')^T \cdot b \mod R^\nu\). Then, we have

\[
\begin{align*}
b' &= b + (a')^T \cdot b \\
&= \frac{1}{q} a^T \cdot s + e - a^T \cdot A^{-1} \cdot (A \cdot s' + e) \\
&= (a')^T \cdot e + e,
\end{align*}
\]

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

\[
\begin{align*}
\sigma(b') &= \sigma(b) + \sigma(a')^T \cdot \sigma(b) \\
&= \sigma(a')^T \cdot \sigma(s') + \sigma(e) - \sigma(a)^T \cdot A_H^{-1} \cdot (A_H \cdot \sigma(s') + \sigma(e)) \\
&= \sigma(a')^T \cdot \sigma(e) + \sigma(e).
\end{align*}
\]

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

**Proposition 1.** There is a PPT reduction from D-MLWE\(_{R^d_q, \nu}^R\) to Nor-D-MLWE\(_{q, \phi}^R\) for \(q \geq 2n\).

We mainly consider the following discretization in this paper: we use results showed in [29] to discretize \(e\) to a discrete Gaussian distribution. Note that \(\varphi(\vec{d}) = U \cdot \sigma(\vec{d})\), so \(\varphi(\vec{d})^T \cdot \varphi(\vec{d}) = \varphi(\vec{d})^* \cdot \varphi(\vec{d}) = \sigma(\vec{d})^* \cdot \sigma(\vec{d})\), which implies \(s_1(\varphi(\vec{d})) = s_1(\sigma(\vec{d})) = \sqrt{\frac{\text{rad}(\nu)}{q}}\). Hence, if we set \(\Lambda = \sigma(\frac{1}{q} R^\nu)\) and use the basis \(\frac{1}{q} \vec{d}\), for any \(c \in H \cong \mathbb{R}^n\) and \(\beta > \omega(\sqrt{\log n}) \cdot \frac{1}{q} \sqrt{\frac{\text{rad}(\nu)}}{n}}\), we can use Algorithm 2 of [29] to output a vector \(x\) drawn from a distribution statistically close to \(D_{A^\perp c, \beta}^\nu\) in probabilistic polynomial time. We
also have

$$\eta_{\varepsilon}(q R^\varepsilon) \leq \frac{1}{q} \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \lambda_n(R^\varepsilon)$$  \hspace{1cm} (By Lemma 4)

$$= \frac{1}{q} \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \lambda_1(R^\varepsilon)$$  \hspace{1cm} (\lambda_n = \lambda_1 in cyclotomic fields)

$$\leq \frac{1}{q} \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \sqrt{n} \cdot N_k(R^\varepsilon) \cdot |\Delta_K|^{\frac{1}{\pi}}$$  \hspace{1cm} (By Lemma 1)

$$= \frac{1}{q} \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \sqrt{n} \cdot |\Delta_K|^{-\frac{1}{\pi}}$$  \hspace{1cm} (By equation (1))

$$= \frac{1}{q} \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \sqrt{n} \cdot (\prod_{\nu \mid \nu R^\varepsilon})^{\frac{1}{\pi}}$$

Note that the last inequality is rather loose. For any \( \beta \geq \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \sqrt{\frac{n}{\pi}} \), Theorem 3.1 of [29] shows that the distribution \([e]_{\nu} = e + f\) with \( e \leftarrow D_{\nu} \) and \( f \leftarrow D_{\nu - e, \beta} \) is statistically close to \( D_{\nu, \sqrt{\alpha^2 + \beta^2}} \).

In the rest of this paper, we will set \( \phi = D_{\frac{1}{q} R^\varepsilon, \sqrt{\alpha^2 + \beta^2}} \) with \( \beta \geq \sqrt{\frac{\ln(2n(1 + \frac{1}{2}))}{\pi}} : \sqrt{\frac{n}{\pi}} \geq \sqrt{2} \cdot \eta_{\varepsilon}(\frac{1}{q} R^\varepsilon) \),

unless we specify it with other values.

Note that by the transference theorem [5] and Lemma 1, we have \( \lambda_n(R^\varepsilon) \geq q \cdot \eta_{\varepsilon}(\frac{1}{q} R^\varepsilon) \) 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 \( R^\varepsilon \), since we use \( \frac{1}{q} \sqrt{\frac{n}{\pi}} \) 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 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\(_{\nu, \phi}^{R^d,q}\) to Nor-S-MLWE\(_{\nu, \phi}^{R^d,q}\) in this subsection. Recall that, for cyclotomic field \( K = \mathbb{Q}(\zeta_l) \), if we represent \( x \in R^\varepsilon \) with respect to the decoding basis, we have

$$\frac{1}{\sqrt{l}} \cdot ||x||_{\sigma(d)}^c \leq ||\sigma(x)|| \leq \sqrt{\frac{\mathrm{rad}(l)}{l}} \cdot ||x||_{\sigma(d)}^c, \quad \text{for } x \in R^\varepsilon. \quad (4)$$

Note that, by Lemma 5, \( \phi \) is \( (\sqrt{\alpha^2 + \beta^2}, \sqrt{\alpha^2 + \beta^2} \cdot \sqrt{\beta^2}, 2^{-2n}) \) bound, i.e. \( \Pr_{x \leftarrow \phi} ||x|| \geq \sqrt{\alpha^2 + \beta^2} \cdot \sqrt{\beta^2} \cdot \sqrt{n} \leq 2^{-2n} \).

We also represent \( m \) Nor-D-MLWE samples as the form \((A, b)\), where \( A = (a_1, \cdots, a_m)^T \in R_{q \times d}^m \) and \( b \in (\frac{1}{q} R^\varepsilon / R^\varepsilon)^m \).

Now assume we have an oracle \( O \) for solving Nor-S-MLWE problem with advantage \( \varepsilon \) when given \( m \) samples. When we get \( m \) samples \((A, b) \in R_{q \times d}^m \times (\frac{1}{q} R^\varepsilon / R^\varepsilon)^m \), we give it to the Nor-S-MLWE oracle \( O \)

\[\text{The factor } \sqrt{2} \text{ is used in Subsection 3.4 for convenience.}\]
and get some \( s \in (\frac{1}{q} R^V/R^V)^d \) with probability \( \varepsilon \). Then we compute \( e = b - A \cdot s \) mod \( R^V \) and \( N = ||e||_\infty \), where \( ||e||_\infty = \max_{i \in [n]} ||e_i|| \). We output 1 if and only if \( N < B := \sqrt{\alpha^2 + \beta^2 \cdot \sqrt{n}} \).

If \( (A, b) \leftarrow R_{q, \phi}^d \), then the probability we output 1 is larger than \( \varepsilon - \Pr_{e \leftarrow \phi^m} (||e||_\infty \geq B) \). If \( (A, b) \) is uniformly distributed, the probability we output 1 is less than

\[
\Pr_{b \leftarrow U((\frac{1}{q} R^V/R^V)^d)} [\exists s \in (\frac{1}{q} R^V/R^V)^d : ||b - A \cdot s||_\infty < B].
\]

(5)

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

Since \( \phi \) is a \((B, \delta)\) bound distribution with \( \delta = 2^{-2n} \), we have \( \Pr_{e \leftarrow \phi^m}(||e||_\infty \geq B) \leq m \cdot \delta \). Also, note that \( ||x||_\infty \leq ||x||_c \leq \sqrt{t} \cdot ||x|| \) for any \( x \in \frac{1}{q} R^V/R^V \), we have

\[
\Pr_{b \leftarrow U((\frac{1}{q} R^V/R^V)^d)} [\exists s \in (\frac{1}{q} R^V/R^V)^d : ||b - A \cdot s||_\infty < B] \leq q^{nd} \cdot \frac{(2\sqrt{t} \cdot B)^{mn}}{q^{nm}}.
\]

Note that \( q^{nd} \cdot \frac{(2\sqrt{t} \cdot B)^{mn}}{q^{nm}} = (2\sqrt{t} \cdot B \cdot q^{d-1})^m \). We now decide the conditions to bound \( (2\sqrt{t} \cdot B \cdot q^{d-1})^m \leq \delta < 2^{-2n} \). For \( q > 2\sqrt{t} \), this is equivalent to \((2\sqrt{B} \cdot q^{d-1})^m < 2^{-2} \). So, \( m > \frac{d \log q + 2}{\log q - \log(2\sqrt{B})} \) 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 \( q > 2\sqrt{t} \cdot B \) with \( B = \sqrt{\alpha^2 + \beta^2 \cdot \sqrt{n}} \), there is a reduction from Nor-D-MLWE\(_q^d\) to Nor-S-MLWE\(_q^d\) problems when given \( m > \frac{d \log q + 2}{\log q - \log(2\sqrt{B})} \) samples.

**Remark 2.** Note that in this section, we use (4) (a good basis of \( R^V \) 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 \( s \).

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

\[
\Pr_{x \leftarrow D_{A, c, \varepsilon}} [||x - c, u|| \leq \frac{\sigma}{t}] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{\sqrt{2\pi} \varepsilon}{t}.
\]

Similarly, if \( \sigma \geq \eta_c(A) \), we have

\[
\Pr_{x \leftarrow D_{A, c, \varepsilon}} [||x - c, u|| \geq t \cdot \sigma] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot t \cdot \sqrt{2\pi} \varepsilon \cdot e^{-\pi t^2}.
\]

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

---

\( ^8 \) Here, \( D_{A, c, \varepsilon} = D_{A, c, \sigma} \) corresponds to the distribution \( \frac{e^{-\frac{||x-c||^2}{2t^2}}}{\sum_{y \in A} e^{-\frac{||y-c||^2}{2t^2}}} \).
Lemma 12. Let \( \varepsilon \in (0, 1) \), \( t \geq \sqrt{2\pi} \), \( \phi = D_{\frac{1}{q}R} \cdot \sqrt{\alpha^2 + \beta^2} \) with \( \sqrt{\alpha^2 + \beta^2} \geq \frac{1}{\sqrt{2\pi}} \cdot \eta_k(\frac{1}{q}R^\nu) \), we have
\[
\Pr_{x \leftarrow \phi} \left[ \max_{1 \leq i \leq n} \frac{1}{|\sigma_i(x)|} \geq \frac{\sqrt{2} \cdot t}{\sqrt{\alpha^2 + \beta^2}} \right] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{n \cdot \sqrt{2\pi} \varepsilon}{2t}
\]
and
\[
\Pr_{x \leftarrow \phi} \left[ \max_{1 \leq i \leq n} |\sigma_i(x)| \geq t \cdot \sqrt{\alpha^2 + \beta^2} \right] \leq n \cdot \frac{1 + \varepsilon}{1 - \varepsilon} \cdot t \cdot \sqrt{2\pi} \varepsilon \cdot e^{-\pi t^2}.
\]

Proof. For any \( x \leftarrow \phi \), by using Lemma 11 with \( c = 0 \) and \( u = (\frac{1}{\sqrt{2}}, 0, \ldots, 0, \frac{1}{\sqrt{2}}) \) or \( u = (\frac{1}{\sqrt{2}}, 0, \ldots, 0, -\frac{1}{\sqrt{2}}) \), we have
\[
\Pr[|\sqrt{2} \cdot \Re(\sigma_1(x))| \leq \frac{\sqrt{\alpha^2 + \beta^2} \cdot t}{2}] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{\sqrt{2\pi} \varepsilon}{t},
\]
or
\[
\Pr[|\sqrt{2} \cdot \Im(\sigma_1(x))| \leq \frac{\sqrt{\alpha^2 + \beta^2} \cdot t}{2}] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{\sqrt{2\pi} \varepsilon}{t}.
\]
Since \( |\sigma_1(x)| \geq \max\{|\Re(\sigma_1(x))|, |\Im(\sigma_1(x))|\} \), we get
\[
\Pr[|\sqrt{2} \cdot |\sigma_1(x)|| \leq \frac{\sqrt{\alpha^2 + \beta^2} \cdot t}{2}] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{\sqrt{2\pi} \varepsilon}{t},
\]
which implies that
\[
\Pr\left[ \frac{1}{|\sigma_1(x)|} \geq \frac{\sqrt{2} \cdot t}{\sqrt{\alpha^2 + \beta^2}} \right] \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \frac{\sqrt{2\pi} \varepsilon}{t}.
\]

For other \( k \in [r] \), we can get the same result by using similar method. Then, by taking a union bound and noticing that \( \sigma_k(x) = \sigma \) for some \( K \), we conclude the first desired result. The second assertion can be obtained similarly.

Set \( A = \frac{1}{\sqrt{q}} \cdot g \cdot R + qR^d \) with \( g = (1, q^2, \ldots, q^{d-1})^T \in R^d \) and denote \( B_A \) the basis of \( A \), \( B_{s_iR} \) the basis of \( s_iR \) for some \( s_i \in K \). For any basis \( B_R \) of \( R \), it is easy to verify that
\[
B_A = \begin{pmatrix}
1 & \frac{1}{q} & \ldots & \frac{1}{q^{d-1}} \\
1 & \frac{1}{q} & \ldots & \frac{1}{q^{d-1}} \\
\vdots & \ddots & \vdots & \vdots \\
1 & \frac{1}{q} & \ldots & 1
\end{pmatrix} \otimes B_R = \begin{pmatrix}
B_R & \frac{1}{q}B_R & \frac{1}{q^2}B_R \\
0 & B_R & \frac{1}{q}B_R \\
\vdots & \ddots & \ddots \\
0 & \ldots & 0 & B_R
\end{pmatrix}
\]
is a basis of \( A \). Moreover, \( ||B_A|| = ||B_R|| \). We then take \( B_R \) to be the powerful basis of \( R \), hence, \( ||B_A|| = ||B_R|| \leq ||B_R|| = \sqrt{n} \). Observe that for any \( x \in \frac{1}{q}R^\nu \), \( x \cdot R \subseteq \frac{1}{q}R^\nu \) 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 \( ||x \cdot B_R|| \leq ||x \cdot B_R|| \leq \sqrt{n} \cdot ||x|| \). Now we can present the following lemma.\(^9\)

Lemma 13. Assume \( s = (s_1, \ldots, s_d)^T \in (\frac{1}{q}R^\nu \cdot R^d) \) which satisfies \( \max_{1 \leq i \leq d} \frac{1}{|\sigma_{s_{(n,i)}]}(s_i)|} < B_2 \) and \( \max_{1 \leq k \leq n} |\sigma_k(s_1)| \leq ||s_i|| < B_1 \) for all \( i \in [d] \). Let \( r \geq \max\{\sqrt{n}, \sqrt{n} \cdot B_1 \cdot B_2\} \cdot \sqrt{\frac{2\ln(2n(1+1/i))}{\pi}} \).
There is a PPT transformation \( F : R^d \times \frac{1}{q}R^\nu \to R^d \times \mathbb{T}_{R^\nu} \) such that
\[
R_{\alpha} \cdot (A_{s, s', D_{\alpha}}^R) \cdot \mathcal{F}(A_{s, s', \phi}^R) \leq \left( \frac{1 + \varepsilon}{1 - \varepsilon} \right)^{d+1},
\]
where \( s' \leftarrow U(R^\nu) \) and \( \alpha_j = \sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot d \cdot B_1^2 + r^2 \cdot \sum_{k=1}^d |\sigma_j(s_k)|} \) for \( j \in [n] \).

\(^9\) 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 \( R^\nu \)) to efficient output discrete Gaussian samples.
Proof. Suppose that we are given \((a, b) \leftarrow A^{R^d}_{q, a, \phi}\). Consider the following map \(F : R^d_q \times \frac{1}{q}R^\vee / R^\vee \rightarrow R^d_q \times \mathbb{T}_{R^\vee}:

1. Sample \(f \leftarrow D_{A-a, r}\).
2. Let \(v = a + f \mod qR^d\) and set \(\tilde{a} = x \in R^d_q\), where \(x \in R^d_q\) be a random solution of \(\frac{1}{q}g \cdot x = v \mod qR^d\).
3. Sample \(\tilde{e} \leftarrow D_{r, \gamma}\) with \(\gamma = \sqrt{d} \cdot B_1\), \(e' \leftarrow D_{\frac{\ln(2n+4)}{\pi}}\) and \(y \leftarrow U(R^\vee_q)\), set \(\tilde{b} = b + e + e' + \frac{1}{q^d} \tilde{a} \cdot y \mod R^\vee\).
4. Output \((\tilde{a}, \tilde{b})\).

Note that \(a \in R^d_q\), so the coset \(A - a\) is well defined. Meanwhile, \(r \geq ||B_1|| \cdot \sqrt{\frac{\ln(2n+4)}{\pi}}\), we can efficiently sample \(f\) by Lemma 2. Assume \(a = (a_1, \cdots, a_d)^T\), \(s = (s_1, \cdots, s_d)^T\), \(f = (f_1, \cdots, f_d)^T\) and \(\tilde{s} = g^T \cdot s + \frac{1}{q} y\), we have

\[
\tilde{b} - \frac{1}{q^d} \tilde{a} \cdot \tilde{s} \mod R^\vee = a^T \cdot s + e + e' + \tilde{e} - \frac{1}{q^d} \tilde{a} \cdot g^T \cdot s \mod R^\vee = e + e' + \tilde{e} - f^T \cdot s \mod R^\vee.
\]

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

In the following, we denote \(D\) the distribution of the outputs of \(F\) and try to bound \(R^\infty(A_{q^d, s', D_{a'}} | | D\)).

Observe that \(A \cong \frac{1}{q^d} g \cdot R^d_q\) mod \(qR^d\), every \(x \in R^d_q\) is a solution to the equation \(\frac{1}{q^d} g \cdot x = v \mod qR^d\) for some \(v\) and the number of solutions to this equation in \(R^d_q\) for different \(v\) is the same. For any \(\bar{a} \in R^d_q\) and \(f \in A - \bar{a}\), we have

\[
\Pr[a = \bar{a} \land f = \bar{f}] = \frac{1}{q^{nd}} \cdot \frac{\rho_r(f)}{\rho_r(A - \bar{a})} = C \cdot \frac{\rho_r(A)}{\rho_r(A - \bar{a})} \cdot \rho_r(\bar{f})
\]

\[
\leq C \cdot \left[1, \frac{1 + \epsilon}{1 - \epsilon}\right] \cdot \rho_r(\bar{f}),
\]

where \(C = \frac{2^{-nd}}{\rho_r(A)}\) and we have used Lemma 6 with \(r \geq \eta_r(A)\). Then, for any \(\bar{v} \in A \mod qR^d\), by using Lemma 6 again, we get

\[
\Pr[v = \bar{v}] = \sum_{a \in R^d_q} \Pr[a] \cdot \Pr[f = \bar{v} - a|a] \leq C' \cdot \frac{1 + \epsilon}{1 - \epsilon} \cdot K_v \leq \sum_v \Pr[v = \bar{v}] = 1 \leq C' \cdot \frac{1 + \epsilon}{1 - \epsilon} \cdot K_v,
\]

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

\[
C' \cdot \frac{1 - \epsilon}{1 + \epsilon} \cdot K_v \leq \sum_v \Pr[v = \bar{v}] = 1 \leq C' \cdot \frac{1 + \epsilon}{1 - \epsilon} \cdot K_v.
\]
So, for any \( \bar{a} \in R_q \),
\[
\Pr[\bar{a} = a] = \sum_v \Pr[\bar{a} = \bar{a}|v = \bar{v}] \cdot \Pr[v = \bar{v}]
\]
\[
\geq \frac{1}{q^d} \cdot \left[\left(\frac{1 - \varepsilon}{1 + \varepsilon}\right)^2, \left(\frac{1 + \varepsilon}{1 - \varepsilon}\right)^2\right].
\]

Therefore, we have \( R_\infty(U(R_q)||\bar{a}) \leq (\frac{1 + \varepsilon}{1 - \varepsilon})^2 \).

We now analyze the distribution of \(-f\) appeared in (6) condition on some fixed \( \bar{a} \) (equivalently, condition on some fixed \( \bar{v} \)). In this situation, \(-f \in R^d - \bar{v} \) and fixing a value \( f \) fixes \( a = \bar{v} - f \mod qR^d \). So, by (7), we have
\[
1 - \varepsilon \cdot \frac{\rho_r(-\bar{f})}{\rho_r(R^d - \bar{v})} \leq \Pr[-f = -f|\bar{a} = \bar{a}] \leq 1 + \varepsilon \cdot \frac{\rho_r(-\bar{f})}{\rho_r(R^d - \bar{v})},
\]
Hence, \( R_\infty(D_{R^d - \bar{v}, r}||-f) \leq \frac{1 + \varepsilon}{1 - \varepsilon} \). This also implies that condition on some fixed \( \bar{v} = (v_1, \ldots, v_d)^T \), \( \Delta(D_{R^d - \bar{v}, r}||-f) \leq 2\varepsilon \), i.e. \(-f_i\) is almost distributed as \( D_{R^d - v_i, r} \) for \( i \in [d] \). It then follows that \(-s_i \cdot f_i\) is almost distributed as \( D_{s_i, R^d - v_i, r} \) with \( r_i = (r \cdot |s_1| \cdot \ldots, r \cdot |\sigma_1(s_i)|)^T \). 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 \( D_{s_i, R^d - v_i, r} \), and then adding a vector sampled from \( D_{r, B_1} \), \( \tilde{D} \) 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 \(-f, \tilde{e}_1, \ldots, \tilde{e}_d \mapsto (\varphi(-s_i \cdot f_i + \tilde{e}_i), \ldots, \varphi(-s_d \cdot f_d + \tilde{e}_d)\), we obtain
\[
R_\infty(Y^{(1)} \times \ldots \times Y^{(d)}||D^{(1)} \times \ldots \times D^{(d)}) \leq R_\infty(D_{R^d - v, r} \times D_{r, B_1}|| -f \times D_{r, B_1}) \leq \frac{1 + \varepsilon}{1 - \varepsilon}.
\]

Then, noticing that by our choice of \( r \), we can use Lemma 7 and conclude that
\[
R_\infty(D_{t_i}||Y^{(i)}) \leq \frac{1 + \varepsilon}{1 - \varepsilon}
\]
for any \( i \in [d] \), where \( t_i = (\sqrt{r^2 \cdot B_i^2 + r^2 \cdot |s_1|^2}, \ldots, \sqrt{r^2 \cdot B_i^2 + r^2 \cdot |s_1|^2}) \). 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
\[
R_\infty(D_t||\tilde{D}) \leq R_\infty(D_{t_i} \times \ldots \times D_{t_i}||Y^{(1)} \times \ldots \times Y^{(d)}) \cdot R_\infty(Y^{(1)} \times \ldots \times Y^{(d)}||D^{(1)} \times \ldots \times D^{(d)}) \leq \frac{1 + \varepsilon}{1 - \varepsilon} \cdot \prod_{i=1}^d R_\infty(D_{t_i}||Y^{(i)}) \leq \left(\frac{1 + \varepsilon}{1 - \varepsilon}\right)^{d+1},
\]
where \( t = (\sqrt{r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^d |s_k|^2}, \ldots, \sqrt{r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^d |s_k|^2}) \).

Finally, note that \( \frac{1}{q^d} \bar{a} \cdot \bar{s} \) for some \( \bar{s} \in \frac{1}{q} R^\gamma / R^\gamma \) is equivalent to \( \frac{1}{q^d} \bar{a} \cdot \bar{s}' \) for \( \bar{s}' = q \cdot \bar{s} \in R_q^\gamma \). We obtain, by using data processing inequality and the multiplicativity of Rényi divergence,
\[
R_\infty(A_{q^d, \bar{s}', D_{\alpha}}||\tilde{D}) \leq \left(\frac{1 + \varepsilon}{1 - \varepsilon}\right)^{d+4},
\]
where \( \alpha = (\sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^d |s_k|^2}, \ldots, \sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot \gamma^2 + r^2 \cdot \sum_{k=1}^d |s_k|^2}) \).
Proposition 3. There is a reduction from Nor-S-MLWE\textsubscript{q,φ} to the worst-case S-RLWE with m samples, where α′ = \sqrt{2(α^2 + β^2)(1 + r^2 \cdot d \cdot n)} \text{ with } r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1 + (d + 4)m))} \text{ and } \sqrt{α^2 + β^2} \geq 2\sqrt{e} \cdot n \cdot d \cdot \eta_e(\frac{1}{\sqrt{q}}) \text{.}

Proof. Recall that for Nor-S-MLWE\textsubscript{q,φ} problem, the secret s \leftarrow φ^d. By using Lemma 5 and 12 with ε = \frac{1}{\ln(d+4)} and t = 2n \cdot d \cdot \sqrt{2πe}, we have that with probability greater than (1 - \frac{m(d+4)+1}{\ln(d+4)+1} \cdot \frac{1}{2d} - 2^{-2n})^d > (1 - \frac{1}{2d} - 2^{-2n})^d, \text{ max}_{1 \leq k \leq n} \frac{1}{|σ_k(s)|} < B_2 := 4n \cdot d \cdot \sqrt{πe} \sqrt{α^2 + β^2} \text{ and } \text{ max}_{1 \leq k \leq n} |σ_k(s_i)| \leq ||s_i|| < B_1 := \sqrt{π} \cdot \sqrt{α^2 + β^2} \text{ for all } i \in [d]. \text{ So, } r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1 + (d + 4)m))} \text{ is sufficient to use Lemma 13. At the same time, the error distribution } D_α \text{ satisfies } α_i \leq α′.

Therefore, when given m samples, we can use the above settings and Lemma 13 to solve Nor-S-MLWE\textsubscript{q,φ} Problem with advantage greater than (1 - 2^{-2n} - \frac{1+ε}{2+ε} \cdot n \cdot \sqrt{\frac{α^2 + β^2}{π}})^{(1+ε)(d+4)m} \geq \frac{1}{8}(1 - \frac{1}{2d} - 2^{-2n})^d > \frac{1}{16} - \frac{d}{3πe^{48}} \text{, 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 α′ \leq α. Now, we can collect the results of Propositions 1, 2 and 3 to conclude the following theorem.

Theorem 1. Assume ε ∈ (0, \frac{1}{2}), α = α(n) ∈ (0, 1) and β ≥ \sqrt{\frac{2πe}{q}} \cdot \sqrt{\frac{\ln(2n(1+\frac{1}{2}))}{π}} \text{ such that } \sqrt{α^2 + β^2} ≥ 2\sqrt{e} \cdot n \cdot d \cdot \eta_e(\frac{1}{\sqrt{q}}) \text{.} \text{ Let } q > \max\{2n, 2\sqrt{1} \cdot \sqrt{n} \cdot \sqrt{α^2 + β^2}\} \text{ be a prime that does not ramify in } R. \text{ When given } m > \frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{1} \cdot \sqrt{n} \cdot \sqrt{α^2 + β^2})} \text{ samples, there is a probabilistic polynomial-time reduction from D-MLWE\textsubscript{q,φ} \text{ in worst/average-case to S-RLWE\textsubscript{q,φ}, } \text{ in worst-case for arbitrary } d = \text{ poly}(n), \text{ where } α′ = \sqrt{2(α^2 + β^2)(1 + d \cdot n \cdot r^2)} \text{ and } r \geq 4\sqrt{2e} \cdot n^2 \cdot d \cdot \sqrt{\ln(2nd(1 + (d + 4)m))} \text{.}

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 mod 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 Ω(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 α′ ≥ η_e(R′). Recall that, by Lemmata 1 and 4, for cyclotomic fields, we have

\[ \sqrt{\frac{\ln(3)}{π}} \cdot n^{-\frac{1}{2}} \cdot (\prod_{p \| l} \frac{p}{p-1})^{-\frac{1}{2}} \leq η_e(R′) \leq \sqrt{\frac{\ln(2n(1+\frac{1}{2}))}{π}} \cdot \sqrt{n} \cdot (\prod_{p \| l} \frac{p}{p-1})^{\frac{1}{2}}. \]

The upper bound of η_e(R′) can be as small as \tilde{O}(1). Assume d = \tilde{O}(n^{\epsilon_1}) and α = \tilde{O}(n^{-\epsilon_2}), we then can set β ≈ α, q = \tilde{O}(n^{\frac{3}{4} + \epsilon_1 + \epsilon_2}) and r = \tilde{O}(n^{2 + \epsilon_1}), which gives α′ = \tilde{O}(n^{\frac{3}{4} + \epsilon_1 + \epsilon_2}). So, c_2 > \frac{2+\epsilon_1}{\epsilon_2} is sufficient. In applications, we usually use very small d = O(1), then we can set α ≈ β = \tilde{O}(n^{\frac{3}{4}}) and q = \tilde{O}(n^{\frac{1}{2}}) 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 problems to finish our reduction. Note that the modulus in the S-RLWE problems we investigate is q\textsuperscript{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 mod l. There are some results showed in [34], which state a reduction from S-RLWE to worst-case S-RLWE problem to arbitrary modulus q.
Theorem 4. Let \( r \in (\mathbb{R}^+)^n \) be such that \( r_i = r_{n+1-i} \) for all \( i \in [\frac{n}{2}] \) and \( r_i \leq r \) for some \( r \). Let \( d' = n \cdot q^{-\frac{d}{2} + \frac{k}{2}} \), and consider \( \Sigma = \{ r' : r'_i \leq \sqrt{d' \cdot r^2 \cdot m + d'^2} \} \). Then, there exists a probabilistic polynomial-time reduction from S-RLWE\(_{q,D} \) with \( m \leq \frac{q}{\sqrt{m}} \) 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 \( \beta \geq \frac{\sqrt{2n}}{q} \cdot \sqrt{\ln(2n(1+\frac{1}{\varepsilon}))} \) such that \( \sqrt{\alpha^2 + \beta^2} \geq 2\sqrt{\frac{1}{2} \cdot \eta_{R^c}} \cdot \sqrt{(\frac{\ell}{d})^2} \). Let \( q > \max\{2n, 2\sqrt{\ell} \cdot \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2} \} \) be a prime that does not ramify in \( R \). When given \( \frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\ell} \cdot \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2})} < m \leq \frac{q}{\sqrt{m}} \) samples, there is a probabilistic polynomial-time reduction from D-MLWE\(_{q^d, D} \) in worst/average-case to D-RLWE\(_{q^d, D} \), \( m \) worst-case, where \( \beta' = \sqrt{(\frac{\sqrt{\alpha^2 + \beta^2}}{2}(1 + d \cdot n \cdot r^2))} \) and \( r \geq 4\sqrt{2}(\alpha^2 \cdot \sqrt{\frac{\alpha^2 + \beta^2}{2}})^2 \cdot (1 + m \cdot \alpha^2) \).

Note that, the error parameter \( \beta' \) contains a term \( q^{-\frac{d}{2} + \frac{k}{2}} \). Assume \( d = O(1) \) and \( \alpha = O(n^{-c}) \), we set \( \beta \approx \alpha = O(n^{-c}) \), \( q = O(n^{-\frac{1}{2}}) \) and \( r = O(n^2) \). Under this condition, we have \( \alpha' = O(n^{\frac{c}{2}-c}) \), since \( m \geq O(1) \) implies \( q^{-\frac{d}{2} + \frac{k}{2}} = O(1) \). So, \( \beta' = O(n^{\frac{c}{2}-c} \cdot m^\frac{1}{2}) \). Meanwhile, \( \frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\ell} \cdot \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2})} < m \leq \frac{q}{\sqrt{m}} \) 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 \( D \) 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\(_{q^d, D} \) given only \( m \) samples for any modulus \( q \), where \( \xi = \beta' \cdot \left( \frac{nm}{\log(n^m)} \right)^\frac{1}{2} \), solves D-RLWE\(_{q^d, D} \leq \beta' \).

Overall, we conclude the following theorem.

Theorem 4. Assume \( \varepsilon \in (0, \frac{1}{2}) \), \( \alpha = \alpha(n) \in (0, 1) \) and \( \beta \geq \frac{\sqrt{2n}}{q} \cdot \sqrt{\ln(2n(1+\frac{1}{\varepsilon}))} \) such that \( \sqrt{\alpha^2 + \beta^2} \geq 2\sqrt{\frac{1}{2} \cdot \eta_{R^c}} \cdot \sqrt{(\frac{\ell}{d})^2} \). Let \( q > \max\{2n, 2\sqrt{\ell} \cdot \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2} \} \) be a prime that does not ramify in \( R \). When given \( \frac{d \cdot \log q + 2}{\log q - \log(2\sqrt{\ell} \cdot \sqrt{n} \cdot \sqrt{\alpha^2 + \beta^2})} < m \leq \frac{q}{\sqrt{m}} \) samples, there is a probabilistic polynomial-time reduction from D-MLWE\(_{q^d, D} \), \( q \leq \hat{O}(n^{\gamma}) \), \( \gamma \leq \hat{O}(n^\frac{c}{2}) \) and \( r \geq 4\sqrt{2}(\alpha^2 \cdot \sqrt{\frac{\alpha^2 + \beta^2}{2}})^2 \cdot (1 + m \cdot \alpha^2) \).

Usually, there are reductions from worst-case SIVP\(_{\gamma} \) with \( \gamma = \hat{O}(\frac{\alpha^2}{\log} \) over rings or modules to corresponding average-case D-LWE problem with error distribution \( D_{\alpha} \) and \( \alpha \leq \hat{O}(n^{-\frac{1}{2}}) \) [21, 30]. Hence, when \( m = \hat{O}(1) \) and \( d = O(1) \), we obtain a reduction from worst-case SIVP\(_{\gamma} \) (over module lattices) to average-case D-RLWE\(_{q^d, D^r} \) with \( q \leq \hat{O}(n^{5.75}) \), \( \gamma \leq \hat{O}(n^\frac{c}{2}) \) and \( r \geq \hat{O}(n) \).

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 \( \text{poly}(n) \). 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 $\alpha$ and $\beta$, and give a self-reduction of S-RLWE first. We begin with the problem S-RLWE$_{q,d_0}$. It is easy to deduce that Nor-S-RLWE$_{q,\phi}$ (denote corresponding distribution $A_{q,s,\phi}^*$) is also hard for $\phi = D_{\frac{p}{q}}R' \cdot \sqrt{\alpha^2 + \beta^2}$ with $\beta \geq \frac{\sqrt{n}}{q} \cdot \sqrt{\frac{\ln(2n(1+a))}{\pi}}$. The proof of the following lemma is similar to that of Lemma 13.

Lemma 15. Assume $s \in \frac{1}{q}R'/R'$ 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 $r \geq \max\{\sqrt{n}, \frac{p}{q} \cdot \sqrt{n}, \sqrt{n} \cdot B_1 \cdot B_2 \cdot \sqrt{1 + \frac{p^2}{q^2}}, \sqrt{\frac{\ln(2n(1+a))}{\pi}}\}$, there is a transformation $F : R_q \times \frac{1}{q}R'/R' \rightarrow R_p \times T_{R'}$ such that

\[ R_\infty(A_{p,s,D_k}|| F(A_{q,s,\phi})) \leq (\frac{1 + \varepsilon}{1 - \varepsilon})^5, \]

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

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

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

Since $a \in R_q$ and $r \geq ||\tilde{B_R}|| \cdot \sqrt{\frac{\ln(2n(1+a))}{\pi}}$, the coset $R - \frac{p}{q} \cdot a$ is well defined and $f$ can be sampled efficiently. For any $\tilde{a} \in R_q$ and $f \in R - \frac{p}{q} \cdot \tilde{a}$, we have

\[ \Pr[a = \tilde{a} \land f = \tilde{f}] = q^{-n} \cdot \frac{\rho_r(f)}{\rho_r(R - \frac{p}{q} \cdot \tilde{a})} \in C \cdot \left[1, \frac{1 + \varepsilon}{1 - \varepsilon}\right] \cdot \rho_r(\tilde{f}), \]

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

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

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 $C' \in \frac{1}{p^2} \cdot \left[\frac{1 - \varepsilon}{1 + \varepsilon}, \frac{1 + \varepsilon}{1 - \varepsilon}\right]$ and

\[ R_\infty(U(R_p)|| \tilde{a}) \leq \left(\frac{1 + \varepsilon}{1 - \varepsilon}\right)^2. \]

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

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

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

\[ R_\infty(A_{p,s,D_k}|| F(A_{q,s,\phi})) \leq (\frac{1 + \varepsilon}{1 - \varepsilon})^5, \]

with $t_i = \sqrt{2(\alpha^2 + \beta^2) + r^2 \cdot B_1^2 + \frac{p^2}{q^2} \cdot r^2 \cdot |\sigma_i(s)|^2}$, 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\(_q,\varphi\) to the worst-case S-RLWE\(_{p,D_p}\), with \(m\) samples, where \(\alpha' = \sqrt{(\alpha^2 + \beta^2)(2 + r^2 \cdot n + r^2 \cdot n \cdot 2^p)}\) with \(r \geq 4\sqrt{e} \cdot n^2 \cdot \sqrt{1 + \frac{p^2}{q^2}} \cdot \sqrt{\ln(2n(1+5m))}\) and \(\sqrt{\alpha'^2 + \beta'^2} \geq 2\sqrt{e} \cdot n \cdot \eta_1(\frac{1}{q} R^\varphi)\).

**Proof.** We set \(\varepsilon = \frac{1}{5m}\) and \(t = 2n \cdot \sqrt{2\pi e}\). Then, with probability \(\geq 1 - \frac{1 + \frac{p^2}{q^2} \cdot \frac{1}{4} - 2^{-2n}}{\frac{1}{2} - 2^{-2n}}\), \(\max_{1 \leq k \leq n} \frac{1}{|\sigma_k(s)|} < B_2 := \frac{4n\sqrt{\pi e}}{\sqrt{\alpha'^2 + \beta'^2}}\) 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 \(D_t\) satisfies \(t_i \leq \alpha'\).

Therefore, when given \(m\) samples, we can use the above settings and Lemma 15 to solve Nor-S-RLWE\(_q,\varphi\) problem with advantage greater than \((\frac{1}{2} - 2^{-2n}) \cdot (\frac{5m+1}{5m})^{-5m} \geq \frac{1}{16} - 2^{-2n+3}\), 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 obtain a non-negligible advantage, the reduction will remain a non-provable advantage, and the reduction will remain the same problem as in [1] and make \(q\) to be at least super-polynomial. Since we usually set the error rate of D-RLWE 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\(_q,\varphi\) in worst/average-case to S-RLWE\(_{p,\Psi}\) in worst-case, where \(\Psi\) 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 \(q = 1 \mod l\). In fact, assume \(R/qR \cong R/q_1 \times \cdots \times R/q_k\) with \(q_1 \cdots q_k = n\), if \(|R/q_k| = q_1^{k'} = \text{poly}(n)\) for \(k \in [q]\), 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 \(\frac{q}{l} = \text{poly}(n)\), we can obtain a reduction from D-RLWE\(_{q,D_p}\) to D-RLWE\(_{p,D_p'}\) for some \(\beta' = \text{poly}(n)^{-1}\) by combining Proposition 4 and the search to decision reductions showed in [23] for arbitrary \(m = \text{poly}(n)\) 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 \(\text{poly}(n)\) interval in the asymptotic sense. However, 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, and then reduce module SIS problems to corresponding module SIVP problems as in [21]. Combining techniques used in [38], 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\(_{q,\beta}^{R_d}\)) is as follows: Given \(A \leftarrow U(R_d^{n \times d})\), find \(z \in R_m^{n} \setminus \{0\}\) such that \(A \cdot z = 0 \mod qR_d^{d} \) and \(||z|| \leq \beta\).

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\(_{q,D_p}\) to M-SIS\(_{q,\beta}^{R_d}\) with \(\alpha < \frac{1}{\beta \omega(n \ln n \sqrt{\log \log n})}\).

**Proof.** Given \(m\) samples \((A, b) \in R_m^{n \times d} \times \mathbb{T}^{m}_{R^\varphi}\), we use the M-SIS oracle to obtain some \(z\) such that \(A^T \cdot z = 0 \mod qR_d^{d}\) and \(||z|| \leq \beta\). Then we compute \(z^T \cdot b \mod R^\varphi = \sum_{k=1}^{n} x_k \cdot \varphi(\tilde{d}_k) (\mathbb{T}^{m}_{R^\varphi} \cong \mathbb{R}^{m} / \varphi(R^\varphi))\) with \(x_k \in [-\frac{1}{2}, \frac{1}{2})\). Note that, if \(b \leftarrow U(\mathbb{T}^{m}_{R^\varphi})\), we have \(z^T \cdot b \leftarrow U(\mathbb{T}^{m}_{R^\varphi})\), so the coefficients \(\{x_k\}\)'s
of $z^T \cdot b$ will be distributed uniformly in $[-\frac{1}{2}, \frac{1}{2}]$. If $b = A \cdot s + e \mod R^\gamma$ for some $e \leftrightarrow D_\alpha$, then $z^T \cdot b = z^T \cdot e = \sum_{j=1}^m z_j \cdot e_j \leftrightarrow D_\alpha$ with $r_k = \sqrt{\alpha^2 \cdot \sum_{1 \leq j \leq m} |\sigma_k(z_j)|^2}$ for $k \in [n]$. By definition, in this situation, we have $z^T \cdot e = \sum_{k=1}^n x'_k \cdot h_k$ with $x'_k \leftrightarrow D_{r_k}$ for any $k \in [n]$, so $E[e^{-\frac{1}{2}T} (x'_k)^2] = \sqrt{2}$. By Markov’s inequality, we have
\[
Pr[(x'_k)^2 \geq \frac{2r_k^2}{\pi} \cdot t^2] \leq \sqrt{2 \cdot e^{-t^2}}.
\]
Setting $t = \omega(\ln n)$, we get $Pr[|x'_k| < \sqrt{\frac{2}{n}} \cdot r_k \cdot \omega(\ln n)] > 1 - n^{-\omega(\ln n)}$. Hence, by taking a union bound, we have $Pr[||z^T \cdot e|| < \sqrt{n} \cdot \alpha \cdot ||z|| \cdot \omega(\ln n)] > 1 - n^{-\omega(\ln n)}$. Therefore, $Pr[\max_k |x_k| < \sqrt{\frac{1}{n \cdot \alpha \cdot \beta \cdot \omega(\ln n)}}, \alpha < \frac{1}{\beta \cdot \omega(\ln n) \log \log n}]$, 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 $\omega'(-)$ and $\omega''(-)$. Thus, we can distinguish $A_{q,s,D_\alpha}$ and $U(T_{R^\gamma})$ 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
\[
A^d = \{ z \in R^m : A \cdot z = 0 \mod qR^d \}
\]
for $A \leftarrow U(R_q^{d \times m})$. If we can solve Mod-SIVP$_\gamma$ in the lattice $A^\perp$ for $A \leftarrow U(R_q^{d \times m})$ with non-negligible probability, then, of course, we can solve M-SIS$_{q,\beta}$ with $\beta \leq \gamma \cdot \lambda_N(A)$, here $N \leq m \cdot n$ denotes the dimension of lattice $A^\perp$. Note that $\lambda_N(A) \leq \frac{N}{\lambda^\perp_{\infty}(A^\perp)} \leq \frac{N}{\lambda^\perp_{\infty}(A)}$ for any $N$-dimensional lattice $A$, we only need to estimate the lower bound of $\lambda^\infty_{\gamma}(A^\perp)^\gamma$. Recall that the dual $M^\perp$ of a lattice $M \subseteq K^m$ is defined as the set of all $x \in K^m$ such that $T_M(x^T \cdot v) \in \mathbb{Z}$ for all $v \in M$. It is easy to check $(A^\perp)^\perp = \frac{1}{q}L_q(A)$, where
\[
L_q(A) = \{ y \in (R_q^\gamma)^m, \exists s \in (R_q^\gamma)^d, A^T \cdot s = y \mod q(R^\gamma)^m \}.
\]
Next, we give a probabilistic lower bound of $\lambda^\infty_{\gamma}(L_q(A))$ for $A \leftarrow U(R_q^{d \times m})$, whose proof technique is an extension of methods used in [21, 34, 35, 38] and may be standard now.

**Lemma 17.** Let $q$ be a prime that does not ramify in $R$ and $qR = q_1 \times \cdots \times q_g$ with $g \cdot f = n$, assume $m > d$ and $\varepsilon \in (0, 1)$, then $Pr_{A \leftarrow U(R_q^{d \times m})}[\lambda^\infty_{\gamma}(L_q(A)) < \frac{1}{n} \cdot q^{1 - \frac{d}{m} - \varepsilon}] \leq 22^{mn + \delta \cdot q^{-mn}}$.

**Proof.** By our assumption, we have $N(q_k) = q^f$ for all $k \in [g]$. 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
\[
\sum_{\substack{t \in (R_q^\gamma)^m \\alpha < ||t||_\infty < B}} \sum_{s \in (R_q^\gamma)^d} Pr_{A \leftarrow U(R_q^{d \times m})}[A^T \cdot s = t \mod q(R^\gamma)^m],
\]
which is equal to
\[
\sum_{\substack{t \in (R_q^\gamma)^m \\alpha < ||t||_\infty < B}} \sum_{s \in (R_q^\gamma)^d} \prod_{k=1}^m Pr_{A \leftarrow U(R_q^\gamma)}[a^T \cdot s = t_k \mod qR^\gamma].
\]
By the CRT and Lemma 2.15 of [23], we have $R$-module isomorphisms $R_q^\gamma \cong R^\gamma/q_1 \cdot R^\gamma \times \cdots \times R^\gamma/q_g R \cong \prod_{i=1}^g q_i R$. Now, $a^T \cdot s = t_k \mod qR^\gamma$ if and only if $a^T \cdot s = t_k \mod q_j \cdot R^\gamma$ for all $j \in [g]$. If $s = 0$ mod $q_j \cdot R^\gamma$ for some $j \in [g]$, the probability $\prod_{i=1}^g Pr_{A \leftarrow U(R_q^\gamma)}[a^T \cdot s = t_k \mod qR^\gamma] \neq 0$ if and only if $t = 0$ mod $q_j \cdot R^\gamma$ (denoted by $q_j \cdot R^\gamma | t$) for the same $j \in [g]$. We denote $S \subseteq [g]$ be the set of indices $j$ such that $s = 0$ mod $q_j \cdot R^\gamma$. Then, for any $j \in [g] \setminus S$, we have $Pr_{A \leftarrow U(R_q^\gamma)}[a^T \cdot s = t_k \mod qR^\gamma] \leq \frac{1}{q_j}$ for any $k \in [m]$. So,
\[
Pr_{A \leftarrow U(R_q^\gamma)}[a^T \cdot s = t_k \mod qR^\gamma] \leq \prod_{i \in [g] \setminus S} \frac{1}{q_j} = \left(\frac{1}{q_j}\right)^{|S|},
\]
Therefore, we have
\[ p \leq \sum_{S \subseteq [d]} \sum_{\mathbf{t} \in (R_q^d)^m \atop 0 < \mathcal{R}(\mathbf{t}) \leq B} \sum_{\mathbf{v} \in S} q^{m||S|-m\mathbf{v}}. \]

There are \(((q^d)^{n-|S|})^d\) elements in \((R_q^d)^d\) satisfying \(q_i R^i \mathbf{s}\) for \(i \in S\). Thus,
\[ p \leq \sum_{S \subseteq [d]} \sum_{\mathbf{t} \in (R_q^d)^m \atop 0 < \mathcal{R}(\mathbf{t}) \leq B} q^{(m-d)(|S|-n)}. \]

Set \( \mathfrak{h} = \prod_{i \in S} q_i R^i \) and denote \( \mathfrak{B}(r, c) \) the open ball in \( H \) of center \( c \) and radius \( r \) under the infinity norm.
We now estimate the number of \( \mathbf{t} \)'s satisfying the conditions in the above sum. First note that, if we denote \( \mathbf{t} = (t_1, \ldots, t_m)^T \), \( t_i \in \mathfrak{h} \) for all \( i \in [m] \), then, \( \mathcal{R}(\mathbf{t}) = \max_{1 \leq i \leq m} ||t_i||_\infty \geq \frac{1}{\sqrt{m}} \max_{1 \leq i \leq m} ||t_i||_\infty \geq \frac{1}{\sqrt{m}} \lambda_1(\mathfrak{h}) \geq N(\mathfrak{h}) \frac{\sqrt{m}}{d} \geq \frac{1}{d} \cdot q^{|S|} \), since \( N(R^d) = \Delta_\mathfrak{h}^{-1} \geq n^{-\delta} \). As a result, there is no such \( \mathbf{t} \) when \( |S| \geq (1 - \frac{d}{m} - \varepsilon) \cdot g \). For the case \(|S| < (1 - \frac{d}{m} - \varepsilon) \cdot g \), we try to bound \( |\mathfrak{B}(B, 0) \cap \mathfrak{h}| \).
Let \( \lambda = \lambda_\mathfrak{h}(\mathfrak{h}) \), then for any two elements \( \mathfrak{v}_1 \) and \( \mathfrak{v}_2 \) of \( \mathfrak{h} \), we have \( \mathfrak{B}(\lambda, \mathfrak{v}_1) \cap \mathfrak{B}(\lambda, \mathfrak{v}_2) = \phi \). Meanwhile, for any \( \mathfrak{v} \in \mathfrak{B}(B, 0) \), we have \( \mathfrak{B}(\lambda, \mathfrak{v}) \subseteq \mathfrak{B}(B + \lambda, 0) \). Hence,
\[ N \leq |\mathfrak{B}(B, 0) \cap \mathfrak{h}|^m \leq \left( \frac{\text{Vol}(\mathfrak{B}(B + \lambda, 0))}{\text{Vol}(\mathfrak{B}(\lambda, 0))} \right)^m \leq B^{m} \leq 4^{mn} \cdot q^{mn(1 - \frac{d}{m} - \frac{|S|}{\theta} - \varepsilon)}, \]
where we have used \( \lambda_\mathfrak{h}(\mathfrak{h}) \geq \frac{1}{\sqrt{m}} \cdot q^{|S|} \). Since there are \( 2^g \) subsets of \([g]\), we get
\[ p \leq 2^g \cdot \max_{S \subseteq [d]} 4^{mn} \cdot q^{mn(1 - \frac{d}{m} - \frac{|S|}{\theta} - \varepsilon)} \cdot q^{(m-d)(|S|-n)} \]
\[ = 2^{g+2mn} \cdot \max_{S \subseteq [d]} q^{m-n\varepsilon-d|S|} \leq 2^{2m+g} \cdot q^{-m\varepsilon}, \]
as desired.

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

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

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

**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

Let’s first introduce some basic facts.

Suppose \( q \) is a prime which does not ramify in \( R \), then \( R_q \) is a principal ideal ring (this is true even when \( K = \mathbb{Q}(\alpha) \) is not a cyclotomic field, as long as \( q \not| \Delta_K \) and \( q \not| [R/\mathbb{Z}[\alpha]] \), see Section 3 of [37]). In the ring \( R_q \), a non-zero element \( x \not\in R_q^* \) if and only if there exists some non-zero element \( y \in R_q \) such that \( x \cdot y = 0 \). In fact, assume \( QR = q_1 \cdots q_\ell \) with \( \ell \cdot q = n \), then \( 0 \neq x \in R_q \) if and only if \( x = 0 \mod q_i \) for some \( i \in S \subseteq [\ell] \) and \( x \neq 0 \mod q_j \) for others \( j \in [\ell]\backslash S \). Then, any element \( y \) such that \( y = 0 \mod q_j \) and \( y \neq 0 \mod q_i \) will satisfy \( x \cdot y = 0 \mod qR \).

For any positive integer \( k \), the \( R_q \) modules \( R_q^k \) and \( R_q^{k \times k} \) are all free-\( R_q \) modules with rank \( k \) and \( k^2 \). A matrix \( A = [a_1, \ldots, a_k]^T \in R_q^{k \times k} \) is invertible in \( R_q \) if and only if \( \text{det}(A) \in R_q^* \), since in this case, there is a matrix \( B \) such that \( A \cdot B = I \mod qR \), hence \( \text{det}(A) \cdot \text{det}(B) = 1 \mod qR \). Note that the determinant function of square matrices in \( R_q^{k \times k} \), which is a special staggered \( k \)-linear map such that \( \text{det}(I_k) = 1 \), over the ring \( R_q \) is well defined). We call a set of vectors \( \{a_1, \ldots, a_k\} \subseteq R_q^k \) \( R_q \)-linearly independent if \( x_1 \cdot a_1 + \cdots + x_k \cdot a_k = 0 \mod qR \) implies \( x_1 = \cdots = x_k = 0 \). Note that for a square matrix \( A = (a_1, \ldots, a_k)^T \in R_q^{k \times k} \), under our definition, \( a_1, \ldots, a_k \) are \( R_q \)-linearly independent does not imply that \( A \) is invertible over \( R_q \). We consider the ring isomorphism \( \Psi : R_q^{k \times k} \simeq (R/q_1)^{k \times k} \times \cdots \times (R/q_\ell)^{k \times k} \) induced by the ring isomorphism \( R_q \simeq R/q_1 \cdots R/q_\ell \), i.e. for any \( A \in R_q^{k \times k} \), \( \Psi(A) = (A_1, \ldots, A_\ell) \) with \( A_i = A \mod q_i \) for \( i \in [\ell] \). Then, it is easy to verify that \( \Psi \leftarrow U(R_q^{k \times k}) \) if and only if \( A \leftarrow U((R/q_i)^{k \times k}) \) for all \( i \in [\ell] \), and \( A \) is invertible in \( R_q^{k \times k} \) if and only if \( A_i \) is invertible in \( (R/q_i)^{k \times k} \) for all \( i \in [\ell] \). So, we just need to estimate the probability that \( A_i \) is invertible for \( A \leftarrow U((R/q_i)^{k \times k}) \). Recall that for a matrix \( A = (a_1, \ldots, a_k)^T \in R_q^{k \times k} \), \( A \) is invertible if and only if vectors \( \{a_i\} \)’s are \( F \)-linearly independent.

We have the following useful result.

Lemma 18. For any \( i \in [k-1] \) and \( F \)-linearly independent vectors \( a_1, \ldots, a_i \in F^k \), the probability that sample a vector \( b \leftarrow U(F^k) \) such that \( a_1, \ldots, a_i, b \) are \( F \)-linearly independent is at least \( 1 - \frac{1}{q^i} \).

Proof. Given \( F \)-linearly independent vectors \( a_1, \ldots, a_i \), we can assume without loss of generality that the first \( i \) columns of \( A = [a_1, \ldots, a_i]^T \in F^{i \times k} \) are \( F \)-linearly independent. We consider the first \( i+1 \) columns of \( A \) together with the first \( i+1 \) elements of a vector \( b \in F^k \). 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 \( a_{1,j} \in F^\times \) for \( j \in [i] \). Without loss of generality, set \( j = 1 \). Then we can get a new matrix \( B' = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,i} & a_{1,i+1} \\ 0 & a_{2,2} & \cdots & a_{2,i} & a_{2,i+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{i,2} & \cdots & a_{i,i} & a_{i,i+1} \\ b_1 & b_2 & \cdots & b_i & b_{i+1} \end{bmatrix} \) such that the first \( i \) rows of \( B \) are \( F \)-linearly independent if and only if the first \( i \) rows of \( B' \) are \( F \)-linearly independent. Also note that the first \( i \) rows of \( B' \) are \( F \)-linearly independent if and only if the rows from \( 2 \) to \( i \) are \( F \)-linearly independent, thanks to the special form of \( B' \) and \( \{a'_{j,m}\} \)’s with \( j \in \{2, \cdots, i\}, m \in [i+1] \).

\[^{10}\text{Here, } R_q^* \text{ denotes the set of invertible elements under multiplication in } R_q.\]
Proof of Lemma 10: Given \( (a, b) \), the transformation discretizes \( b' \in \mathbb{K}_R / \mathbb{R}^d \) to \( [b']_{\frac{1}{q} R^d} \in \frac{1}{q} \mathbb{R}^d \) and outputs \( (a, b) \).

If the distribution of \( (a', b') \) is \( A_{\mathbb{K}_R / \mathbb{R}^d}^{R^d} \), then \( b' = \frac{1}{q} \sum_{i=1}^d a_i' \cdot s_i + e' \mod R^d \) for \( e' \sim D_{\mathbb{R}^d} \). Since \( \frac{1}{q} \sum_{i=1}^d a_i' \cdot s_i \mod R^d \in \frac{1}{q} \mathbb{R}^d / \mathbb{R}^d \), by validity of this discretization, we have that \( [b']_{\frac{1}{q} R^d} \) and \( \frac{1}{q} \sum_{i=1}^d a_i' \cdot s_i \) are identically distributed. Hence, we get \( (a, b) \sim A_{\mathbb{K}_R / \mathbb{R}^d}^{R^d} \).

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

### References


