Large Modulus Ring-LWE $\geq$ Module-LWE

Martin R. Albrecht[⋆] and Amit Deo[⋆⋆]

Information Security Group
Royal Holloway, University of London
martin.albrecht@royalholloway.ac.uk, amit.deo.2015@rhul.ac.uk

Abstract. We present a reduction from the module learning with errors problem (MLWE) in dimension d and with modulus q to the ring learning with errors problem (RLWE) with modulus q^d. Our reduction increases the LWE error rate α by a quadratic factor in the ring dimension n and a square root in the module rank d for power-of-two cyclotomics. Since, on the other hand, MLWE is at least as hard as RLWE, we conclude that the two problems are polynomial-time equivalent. As a corollary, we obtain that the RLWE instance described above is equivalent to solving lattice problems on *module* lattices. We also present a self reduction for RLWE in power-of-two cyclotomic rings that halves the dimension and squares the modulus while increasing the error rate by a similar factor as our MLWE to RLWE reduction. Our results suggest that when discussing hardness to drop the RLWE/MLWE distinction in favour of distinguishing problems by the module rank required to solve them.

Keywords: security reduction, learning with errors, lattice-based cryptography

1 Introduction

Lattice-based cryptography has emerged as a central area of research in the pursuit of designing quantum-safe primitives and advanced cryptographic constructions. For example, lattice-based schemes have been proposed for public-key encryption [Reg09, LP11], key exchange protocols [LP11, ADPS16, BCD$^+$16], digital signatures [BG14, DDLL13], identity-based encryption [GPV08, DLP14] and fully homomorphic encryption schemes [Gen09, BGV12, GSW13].

A fundamental problem in lattice-based cryptography is the Learning with Errors problem (LWE) [Reg05]. For a given dimension n, modulus q and and error distribution χ, the LWE *distribution* in normal-form is informally defined as $(\mathbf{a}, b = \frac{1}{q}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod 1)$, where $\mathbf{a} \in \mathbb{Z}_q^n$ is chosen uniformly at random and both the secret $\mathbf{s} \in \mathbb{Z}_q^n$ and e are drawn from the distribution χ. Distinguishing

[⋆] The research of this author was supported by EPSRC grant “Bit Security of Learning with Errors for Post-Quantum Cryptography and Fully Homomorphic Encryption” (EP/P009417/1).

[⋆⋆] The research of this author was supported by the EPSRC and the UK government as part of the Centre for Doctoral Training in Cyber Security at Royal Holloway, University of London (EP/K035584/1)

between the LWE and the uniform distribution is known as the decision LWE problem, whereas the search LWE problem asks to find the secret $\mathbf{s}$.

The seminal work of Regev [Reg05] establishes reductions from standard problems such as finding short vectors in general lattices to LWE, suggesting that LWE is indeed a difficult problem to solve. In particular, the ability to solve LWE in dimension n implies an efficient algorithm to find somewhat short vectors in *any* n-dimensional lattice. The concrete and asymptotic hardness of LWE has recently been surveyed in [APS15, HKM17].

Although LWE has proven to be a versatile ingredient for cryptography, it suffers from large key sizes (quadratic in the dimension) which motivated the development of more efficient LWE variants.

The Ring Learning with Errors problem (RLWE) was introduced in [LPR10]. RLWE can be seen as a specialisation of LWE where n-dimensional vectors are replaced by polynomials of degree smaller than n. Informally, for RLWE we first choose a ring R of dimension n, modulus q and error distribution χ over a related space of dimension n denoted $K_{\mathbb{R}}$. Then, to sample the RLWE distribution, we sample $a \in R/qR$ uniformly, a secret polynomial s in a suitable space and error e according to χ. We then output $(a, b = \frac{1}{q} a \cdot s + e \bmod R^{\vee})$ as the RLWE sample. Similar to the case of plain LWE, the decision problem is to distinguish the RLWE distribution from random and the search problem is to find the secret s. As alluded to above, the RLWE problem generally offers an increase in efficiency over plain LWE. Intuitively, this can be seen by considering each RLWE sample as a structured set of n LWE samples.

It has been shown that RLWE is at least as hard as standard lattice problems on *ideal* lattices [LPR10, PRSD17]. However, these ideal lattice problems have received much less attention than their analogues on general lattices. Furthermore, some problems that are presumed hard on general lattices such as GapSVP are actually easy on ideal lattices and a recent series of works [CGS14, CDPR16, CDW17] showed that finding short vectors in ideal lattices is potentially easier on a quantum computer than in the general case. More precisely, the length of the short vectors found in quantum polynomial time are a sub-exponential multiple of the length of the shortest vector in the lattice. Currently, it is not known how to efficiently find such vectors in general lattices. However, the vectors that can be found in quantum polynomial time are mainly of theoretical interest since they are still too long to affect current RLWE-based cryptography. Another important caveat to note is that if there was a way to find even shorter vectors in ideal lattices, RLWE could still prove to be a difficult problem. This is due to the fact that RLWE has not been proven to be *equivalent* to finding short vectors in ideal lattices, i.e. the problem might be *strictly* harder.

It is worth noting that the reductions from lattice problems to LWE resp. RLWE [Reg05, LPR10, PRSD17] mentioned above have no dependency on q apart from the fact that q must exceed some lower bound that depends on the dimension and error distribution. In these reductions, the class of lattices is simply defined by the dimension in plain LWE and the ring in the case of RLWE.

The approximation factors defining the lattice problems are also independent of q.

This theoretical conclusion is inconsistent with the current state-of-the-art cryptanalytic techniques for solving LWE. The cost of all known strategies scales with q [HKM17]. As an example, consider the simplest lattice attack, i.e. the dual attack on plain LWE using a set of samples $\{(\mathbf{a}_i, c_i) : i = 1, \ldots, m\}$. In order to perform this attack, a short vector $\mathbf{y}$ in the m-dimensional dual lattice to the lattice formed by the $\mathbf{a}_i$ must be found. This dual lattice has volume q^n whp. Using the Gaussian heuristic, the shortest vector in such a lattice is expected to have length $\approx q^{n/m}$. The attack proceeds by noticing that the inner-product $\langle \mathbf{y}, \mathbf{c} \rangle = \langle \mathbf{y}, \mathbf{e} \rangle$ should be small (modulo 1) in the case of LWE samples and uniform otherwise. In particular, the smaller $\langle \mathbf{y}, \mathbf{e} \rangle$, the more certain we are that the samples were in fact from an LWE distribution. Concretely, for a fixed error rate α, we have $\langle \mathbf{y}, \mathbf{c} \rangle \approx \alpha\, q^{n/m}$ in the case where the modulus is q. On the other hand, if the modulus is q^2, we expect $\langle \mathbf{y}, \mathbf{c} \rangle \approx \alpha\, q^{2n/m}$ which is larger for fixed α, n, m. Therefore, the performance of the dual attack diminishes for growing q.

This fact has only a marginal impact on the performance of LWE where we can choose n freely to increase security. However, in the case of RLWE the choice of ring R — and hence the dimension n — can lead to practical implementation advantages and a simpler interpretation of formally defined RLWE (see Section 3.1). Typically, a power-of-two cyclotomic ring is used, i.e. a ring isomorphic to $\mathbb{Z}[X]/\left\langle X^n + 1 \right\rangle$ with $n = 2^k$. In addition to its simplicity, this choice also improves performance due to its amenability to FFT-based algorithms. In fact, power-of-two cyclotomic rings have proven extremely popular in the literature and dominate the design space, e.g. [LMPR08, Gen10, BGV12, DDLL13, BCNS15, ADPS16]. However, as stressed in [LPR13], this choice severely limits the possible parameter space as powers-of-two are sparsely distributed. Indeed, the reductions mentioned above suggest that power-of-two cyclotomics may lead to bigger parameter sizes than strictly required, i.e. that for a given dimension n there is an upper bound on the security that can be achieved which then necessitates considering dimension $2n$ to increase security, cf. [LPR13]. Alternatively, if an implementation wishes to support intermediate field sizes, a new implementation of multiplication in the intermediate ring is required to achieve comparable performance.

The Module Learning with Errors problem (MLWE) [BGV12, LS15] was proposed to address shortcomings in both LWE and RLWE by interpolating between the two. It will be defined formally in Section 2. For now, one way to informally view the MLWE problem is to take the RLWE problem and replace the single ring elements (a and s) with module elements over the same ring. Using this intuition, RLWE can be seen as MLWE with module rank 1.

As expected, MLWE comes with hardness guarantees given by lattice problems based on a certain class of lattices. In this case, the lattices are generated by modules as opposed to ideals in the RLWE case and in contrast to RLWE, it has been shown that MLWE is *equivalent* to natural hard problems over these lattices. Indeed, solving the approximate shortest vector problem on module lattices for polynomial approximation factors would permit solving MLWE (and thus RLWE) efficiently. We note that this reduction, too, only has a mild dependency on q. Furthermore, MLWE has been suggested an interesting option to hedge against potential attacks exploiting the algebraic structure of RLWE [CDW17]. Thus, MLWE might be able to offer a better level of security than RLWE, while still offering performance advantages over plain LWE.

An example illustrating the flexibility of MLWE is given by the CRYSTALS suite [BBD$^+$17], where MLWE is used to build both key exchange and signature schemes. The advantage of using modules when implementing such systems is that the concrete-security/efficiency trade-off is highly tunable. Remembering that working in power-of-two dimensional rings enables efficient implementations, we can fix our ring and then change the rank of the module as desired. For example, suppose we were working in a module over a ring of dimension $n = 256$, then we can increase the effective dimension from 1024 to 1280 by simply increasing the rank of the module. This effective dimension would not be attainable using power-of-two dimensional rings in RLWE. Thus, MLWE promises to adjust the security level with much greater granularity than efficient RLWE instantiations and implementations for one security level can easily be extended to other security levels.

Contributions. After some preliminaries in Section 2, our main contribution is a reduction from MLWE in dimension d over some general ring R/qR to RLWE in R/q^dR. This was posed as an open problem in [LS15]. Our solution is given in Theorem 1 and Corollary 1. In Section 3.1, we carry out a tighter analysis of the reduction for power-of-two cyclotomic rings. It turns out that for the decision variants, we cannot obtain satisfactory bounds for our reduction to preserve non-negligible advantage unless we allow for super polynomial q and *absolute* noise in addition to negligible *noise rate*. We address this problem in Section 4 by considering the search variants. Specifically, we show that a probabilistic polynomial time (PPT) algorithm with non-negligible success probability for solving average-case search RLWE with error rate $\tilde{\mathcal{O}}(\alpha \cdot n^2\, d^{1/2})$ implies a PPT algorithm solving search MLWE in normal form with error rate α (Corollary 3). Here, the parameters n and d correspond to the ring dimension and module rank respectively. In essence, this says that RLWE with modulus q^d is at least as hard as MLWE with modulus q and module rank d in the same ring. More generally, Corollary 3 shows that there is a freedom to trade between the rank of module and the modulus as long as we hold $d \log q = d' \log q'$ fixed for cyclotomic power-of-two rings. This means that for any decrease in d, we can always balance this off by increasing q exponentially without loss of security.

Our reduction is an application of the main result of Brakerski et al. [BLP$^+$13] in the context of MLWE. In its simplest form, the reduction proceeds from the observation that for $\mathbf{a}, \mathbf{s} \in \mathbb{Z}_q^d$ with $\mathbf{s}$ small it holds that

$$q^{d-1} \cdot \langle \mathbf{a}, \mathbf{s} \rangle \approx \left(\sum_{i=0}^{d-1} q^i \cdot a_i \right) \cdot \left(\sum_{i=0}^{d-1} q^{d-i-1} \cdot s_i \right) \bmod q^d = \tilde{a} \cdot \tilde{s} \bmod q^d.$$

It should be noted that we incur an extra factor of $n^{3/2}\,d^{1/2}$ in error rate expansion when comparing our results to those in [BLP$^+$13]. The extra factor of $n^{3/2}$ arises since we need to drown an (unknown) discrete Gaussian over an (unknown) lattice determined by the secret of the input MLWE instance. Naturally, the factor of d accounts for summing Gaussians when compressing the MLWE sample in rank d into a RLWE sample.

The error distribution of the output in our reduction is an ellipsoidal Gaussian (with bounded widths) as opposed to a spherical one. This type of error distribution appears in the standard hardness result for RLWE [LPR10] and should not be considered unusual. However, we also describe how to perform a reduction from search MLWE to spherical error search RLWE using Rényi divergence arguments (see Section 4.1). This is a tool that has recently received attention in lattice-based cryptography because it allows to tighten security reductions for search (and some decisional) problems [LSS14, BLL$^+$15, BGM$^+$16, LLM$^+$16].

In Section 5, we present self-reductions from power-of-two RLWE in dimension n and modulus q to RLWE in dimension $n/2$ and modulus q^2 following the same strategy. Here, the error rate typically expands from α to $\tilde{\mathcal{O}}(\alpha \cdot n^{9/4})$ if we have access to $\mathcal{O}(1)$ samples and wish to preserve a non-negligible success probability.

Interpretation. Our reduction along with the standard hardness results for MLWE [LS15] implies that RLWE with modulus q^d and error rate α is at least as hard as solving the approximate lattice problem Module-SIVP over power-of-two cyclotomic rings. The approximation factor in this case is $\gamma = \tilde{\mathcal{O}}(n^{5/2}\,d^{1/2})$. As there are also *converse* reductions from RLWE to Module-SIVP e.g. the dual attack mentioned above which requires finding short vectors in a module lattice, these observations imply RLWE is equivalent to Module-SIVP. Previous hardness results only stated that RLWE is at least as hard as Ideal-SIVP [LPR10]. We note, though, that it is not known if Module-SIVP is strictly harder than Ideal-SIVP.

Our results suggest that the distinction between MLWE and RLWE does not yield a hardness hierarchy. There are two different interpretations of this implication. The first and perhaps suspicious conclusion is that MLWE should not be used to hedge against powerful algorithms solving RLWE for *any* modulus. However, as we show in Appendix B, an adversary solving our output RLWE instance with modulus q^d and any dimension n implies an adversary that can solve the standard LWE problem in dimension d and modulus q given n samples. While such an adversary cannot be ruled out in principle, it cannot be enabled by the algebraic structure of RLWE or ideal lattices. However, this line of argument is less powerful when restricting to small constant d.

On the other hand, assuming that such a powerful adversary does not exist, an alternative interpretation is that our results suggest that the difficulty of solving RLWE increases with the size of the modulus when keeping dimension n and noise rate α (roughly) constant. This interpretation is consistent with cryptanalytic results as the best, known algorithms for solving LWE depend on q [APS15, HKM17]. Indeed, our output RLWE instance in modulus q^d has noise of size at least $q^{d/2}$. Thus, as discussed above, our RLWE output instances *cannot* be solved by finding short vectors in lattices of module rank 2 using standard primal or dual attacks in contrast to typical RLWE instances used in the literature. Yet, this contrasts with standard reductions from LWE, RLWE resp. MLWE to SIVP, Ideal-SIVP resp. Module-SIVP [Reg05, LPR10, LS15] which do not suggest that the problem becomes harder with increasing q in their current form.

Finally, in Appendix A, we show how to achieve the same flexibility to MLWE-based constructions for public-key encryption by *explicitly* only considering RLWE elements but relying on a MLWE resp. large modulus RLWE assumption.

2 Preliminaries

An n-dimensional lattice is a discrete subgroup of $\mathbb{R}^n$. Any lattice Λ can be seen as the set of all integer linear combinations of a set of basis vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_j\}$. That is, $\Lambda := \left\{\sum_{i=1}^{j} z_i \mathbf{b}_i \colon z_i \in \mathbb{Z}^n \text{ for } i = 1, \ldots, j\right\}$. The lattices we will be considering will have full rank i.e. $j = n$. We use the matrix $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ to denote a basis. $\tilde{\mathbf{B}}$ is used to denote the Gram-Schmidt orthogonalisation of $\mathbf{B}$ (from left to right) and $\|\mathbf{B}\|$ is the length of the longest vector (in Euclidean norm) of the basis given by $\mathbf{B}$. Additionally, for any vector $\mathbf{x} \in \mathbb{R}^n$, we write $\|\mathbf{x}\|$ to denote the standard Euclidean norm of $\mathbf{x}$. The dual of a lattice Λ is defined as $\Lambda^* = \{\mathbf{x} \in \operatorname{span}(\Lambda) : \forall\ \mathbf{y} \in \Lambda, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}$.

Given a matrix $\mathbf{M} \in \mathbb{C}^{m \times n}$, the singular values of $\mathbf{M}$ are defined to be the positive square roots of the eigenvalues of $\mathbf{M}^\dagger \mathbf{M}$ where $\mathbf{M}^\dagger$ denotes the conjugate transpose of $\mathbf{M}$. The matrix $\mathbf{M}^\dagger \mathbf{M}$ takes a diagonal form in some orthonormal basis of $\mathbb{R}^n$ due to the fact that it is self-adjoint. We write $\sigma_i(\mathbf{M})$ for the ith singular value of $\mathbf{M}$ where $\sigma_1(\mathbf{M}) \geq \cdots \geq \sigma_n(\mathbf{M})$. We also denote the identity matrix in n dimensions using $\mathbb{I}_n$. In addition to the conjugate transpose denoted by $(\cdot)^\dagger$, the transpose of a matrix or vector will be denoted by $(\cdot)^T$. The complex conjugate of $z \in \mathbb{C}$ will be written as $\bar{z}$.

The uniform probability distribution over some finite set $\mathcal{S}$ will be denoted $U(\mathcal{S})$. If s is sampled from a distribution D, we write $s \leftarrow_\$ D$. Also, we let $\mathbf{s} = (s_0, \ldots, s_{d-1}) \leftarrow_\$ D^d$ denote the act of sampling each component s_i according to D independently. We also write $\operatorname{Supp}(D)$ to mean the support of the distribution D. Note that we use standard big-$\mathcal{O}$ notation where $\tilde{\mathcal{O}}$ hides logarithmic factors.

For an algebraic number field $K = \mathbb{Q}(\zeta)$ for algebraic number ζ, its *ring of integers* $\mathcal{O}_K$ is defined to be the ring of all integral elements in K. An element $x \in K$ is said to be integral if it is the solution to some monic polymonial with integer coefficients. We also denote isomorphisms via the symbol $\simeq$.

2.1 Coefficient Embeddings

Let $K := \mathbb{Q}(\zeta)$ be an algebraic number field of degree n where $\zeta \in \mathbb{C}$ is an algebraic number. Then for any $s \in K$, we can write $s = \sum_{i=0}^{n-1} s_i \cdot \zeta^i$ where $s_i \in \mathbb{Q}$. We can embed this field element into $\mathbb{R}^n$ by associating it with its vector of coefficients s_{vec}. Therefore, for any $s \in K$ we have $s_{vec} = (s_0, \ldots, s_{n-1})^T$.

We can also represent multiplication by $s \in K$ in this coefficient embedding. The appropriate matrix will be denoted by $\text{rot}(s) \in \mathbb{R}^{n \times n}$. In particular, for any $r, s, t \in K$ where $r = st$, we have that $r_{vec} = \text{rot}(s) \cdot t_{vec}$. Note that the matrix $\text{rot}(s)$ must be invertible with inverse $\text{rot}(s^{-1})$. The explicit form of $\text{rot}(s)$ depends on the particular field K. In the case where K is a cyclotomic power-of-two field, i.e. $K = \mathbb{Q}[X]/\left\langle X^n + 1 \right\rangle$ for power-of-two n, we have

$$\text{rot}(s) = \begin{bmatrix} s_0 & -s_{n-1} & -s_{n-2} & \ldots\ldots & -s_1 \\ s_1 & s_0 & -s_{n-1} & \ddots\;\ddots & -s_2 \\ \vdots & \vdots & \ddots & \ddots\;\ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots\;\ddots & \vdots \\ s_{n-1} & s_{n-2} & \ldots & \ldots\ldots & s_0 \end{bmatrix}. \tag{1}$$

2.2 Canonical Embeddings

We will often use canonical embeddings to endow field elements with a geometry. A number field $K(\zeta)$ has $n = r_1 + 2r_2$ field homomorphisms $\sigma_i : K \to \mathbb{C}$ fixing each element of $\mathbb{Q}$. Let $\sigma_1, \ldots, \sigma_{r_1}$ be the real embeddings and $\sigma_{r_1+1}, \ldots, \sigma_{r_1+2r_2}$ be complex. The complex embeddings come in conjugate pairs, so we have $\sigma_i = \overline{\sigma_{i+r_2}}$ for $i = r_1 + 1, \ldots, r_1 + r_2$. Define

$$H \coloneqq \{\mathbf{x} \in \mathbb{R}^{r_1} \times \mathbb{C}^{2r_2} : x_i = \overline{x_{i+r_2}}, i = r_1 + 1, \ldots, r_1 + r_2\}.$$

and let $(\mathbf{e}_i)_{i=1}^n$ be the (orthonormal) basis assumed in the above definition. We can easily change to the basis $(\mathbf{h}_i)_{i=1}^n$ defined by

- $\mathbf{h}_i = \mathbf{e}_i$ for $i = 1, \ldots, r_1$
- $\mathbf{h}_i = \frac{1}{\sqrt{2}}(\mathbf{e}_i + \mathbf{e}_{i+r_2})$ for $i = r_1 + 1, \ldots, r_1 + r_2$
- $\mathbf{h}_i = \frac{\sqrt{-1}}{2}(\mathbf{e}_i - \mathbf{e}_{i+r_2})$ for $i = r_1 + r_2 + 1, \ldots, r_1 + 2r_2$

to see that $H \simeq \mathbb{R}^n$ as an inner product space. The *canonical embedding* is defined as $\sigma_C : K \to \mathbb{R}^{r_1} \times \mathbb{C}^{2r_2}$ where

$$\sigma_C(x) \coloneqq (\sigma_1(x), \ldots, \sigma_n(x)).$$

The image of any field element under the canonical embedding lies in the space H, so we can always represent $\sigma_C(x)$ via the real vector $\sigma_H(x)$ through the change of basis described above. So for any $x \in K$, $\sigma_H(x) = U_H^\dagger \cdot \sigma_C(x)$ where the unitary matrix is given by

$$U_H = \begin{bmatrix} \mathbb{I}_{r_1} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}\mathbb{I}_{r_2} & \frac{i}{\sqrt{2}}\mathbb{I}_{r_2} \\ 0 & \frac{1}{\sqrt{2}}\mathbb{I}_{r_2} & \frac{-i}{\sqrt{2}}\mathbb{I}_{r_2} \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{2}$$

Addition and multiplication of field elements is carried out component-wise in the canonical embedding, i.e. for any $x, y \in K$, $\sigma_C(xy)_i = \sigma_C(x)_i \cdot \sigma_C(y)_i$ and $\sigma_C(x+y) = \sigma_C(x) + \sigma_C(y)$. Multiplication is not component-wise for σ_H. Specifically, in the basis $(\mathbf{e}_i)_{i=1}^n$, we have that multiplication by $x \in K$ can be written as left multiplication by the matrix $X_{ij} = \sigma_i(x)\delta_{ij}$ where δ_{ij} is the Kronecker delta. Therefore, in the basis $(\mathbf{h}_i)_{i=1}^n$, the corresponding matrix is $X_H = U_H^\dagger X U_H \in \mathbb{R}^{n\times n}$ which is not diagonal in general. However, for any X_H, we have $X_H \cdot X_H^T = X_H \cdot X_H^\dagger = U_H^\dagger X X^\dagger U_H$. Explicitly, $(X_H \cdot X_H^T)_{ij} = |\sigma_i(x)|^2 \delta_{ij}$ i.e. $X_H \cdot X_H^T$ is a diagonal matrix. Likewise for $X_H^T \cdot X_H$. Therefore, the singular values of X_H are precisely given by $|\sigma_i(x)|$ for $i = 1, \ldots, n$.

Remark 1. We use $\sigma_i(\cdot)$ to denote both singular values and embeddings of field elements. If the argument is a matrix, it should be assumed that we are referring to singular values. Otherwise, $\sigma_i(\cdot)$ denotes a field embedding.

For a ring R contained in field K, we define the canonical embedding of the module R^d into the space H^d in the obvious way, i.e. by embedding each component of R^d into H separately. Furthermore, if we have a matrix of ring elements $\mathbf{G} \in R^{d' \times d}$ for integers d and d', we denote the action of $\mathbf{G}$ on R^d in canonical space H^d as $\mathbf{G}_H \in \mathbb{R}^{nd' \times nd}$. It is well-known that the dimension of $\mathcal{O}_K$ as a $\mathbb{Z}$-module is equal to the degree of K over $\mathbb{Q}$, meaning that the lattice $\sigma_H(R)$ is of *full rank*.

2.3 Ring-LWE and Module-LWE

Let R be some ring with field of fractions K and dual $R^\vee := \{x \in K : \mathrm{Tr}(xR) \subseteq \mathbb{Z}\}$. Also let $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$ and define $\mathbb{T}_{R^\vee} := K_\mathbb{R}/R^\vee$. Note that distributions over $K_\mathbb{R}$ are sampled by choosing an element of the space H (as defined in Section 2.2) according to the distribution and mapping back to $K_\mathbb{R}$ via the isomorphism $H \simeq K_\mathbb{R}$. For example, sampling the Gaussian distribution D_α over $K_\mathbb{R}$ is done by sampling D_α over $H \simeq \mathbb{R}^n$ and then mapping back to $K_\mathbb{R}$. In all definitions below, let Ψ be a *family* of distributions over $K_\mathbb{R}$ and D be a distribution over $R_q^\vee$ where $R_q^\vee := R^\vee/(qR^\vee)$ and $R_q := R/(qR)$.

Definition 1 (RLWE Distribution). *For $s \in R_q^\vee$ and error distribution ψ over $K_\mathbb{R}$, we sample the ring learning with errors (RLWE) distribution $A_{q,s,\psi}^{(R)}$ over $R_q \times \mathbb{T}_{R^\vee}$ by outputting $(a, \frac{1}{q}(a \cdot s) + e \bmod R^\vee)$. where $a \leftarrow_\$ U(R_q)$ and $e \leftarrow_\$ \psi$.*

Definition 2 (Decision/Search RLWE problem). *The* decision *ring learning with errors problem $RLWE^{(R)}_{m,q,\Psi}(D)$ entails distinguishing m samples of $U(R_q \times \mathbb{T}_{R^\vee})$ from $A^{(R)}_{q,s,\psi}$ where $s \leftarrow_\$ D$ and ψ is an arbitrary distribution in Ψ.*

The search *variant s-$RLWE^{(R)}_{m,q,\Psi}(D)$ entails obtaining the secret $s \leftarrow_\$ D$.*

Definition 3 (MLWE Distribution). *Let $M := R^d$. For $\boldsymbol{s} \in (R_q^\vee)^d$ and an error distribution ψ over $K_\mathbb{R}$, we sample the module learning with error distribution $A^{(M)}_{d,q,\boldsymbol{s},\psi}$ over $(R_q)^d \times \mathbb{T}_{R^\vee}$ by outputting $(\boldsymbol{a}, \frac{1}{q}\langle \boldsymbol{a}, \boldsymbol{s}\rangle + e \bmod R^\vee)$ where $\boldsymbol{a} \leftarrow_\$ U((R_q)^d)$ and $e \leftarrow_\$ \psi$.*

Definition 4 (Decision/Search MLWE problem). *Let $M = R^d$. The* decision *module learning with errors problem $MLWE^{(M)}_{m,q,\Psi}(D)$ entails distinguishing m samples of $U((R_q)^d \times \mathbb{T}_{R^\vee})$ from $A^{(M)}_{q,\boldsymbol{s},\psi}$ where $\boldsymbol{s} \leftarrow_\$ D^d$ and ψ is an arbitrary distribution in Ψ.*

The search *variant s-$MLWE^{(M)}_{m,q,\Psi}(D)$ entails obtaining the secret element $\boldsymbol{s} \leftarrow_\$ D^d$.*

When $\Psi = \{\psi\}$, we replace Ψ by ψ in all of the definitions above. It can be shown that the *normal form* of the above problems where the secret distribution is a discretized version of the error distribution is at least as hard as the case where the secret is uniformly distributed. Therefore, it is customary to assume the normal form when discussing hardness.

2.4 Statistical Distance and Rényi Divergence

Definition 5 (Statistical Distance). *Let P and Q be distributions over some discrete domain X. The statistical distance between P and Q is defined as $\Delta(P,Q) := \sum_{i \in X} |P(i) - Q(i)|/2$. For continuous distributions, replace the sum by an appropriate integral.*

Claim. If P and Q are two probability distributions such that $P(i) \geq (1-\epsilon)Q(i)$ for all i, then $\Delta(P,Q) \leq \epsilon$.

We will also make use of the Rényi divergence as an alternative to the statistical distance as a measure of the similarity between two distributions.

Definition 6. *(Rényi Divergence) For any distributions P and Q such that* $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$*, the Rényi divergence of P and Q of order $a \in [1,\infty]$ is given by*

$$R_a(P||Q) = \begin{cases} \exp\left(\sum_{x \in Supp(P)} P(x) \log \frac{P(x)}{Q(x)}\right) & \textit{for } a = 1, \\ \left(\sum_{x \in Supp(P)} \frac{P(x)^a}{Q(x)^{a-1}}\right)^{\frac{1}{a-1}} & \textit{for } a \in (1,\infty), \\ \max_{x \in Supp(P)} \frac{P(x)}{Q(x)} & \textit{for } a = \infty. \end{cases}$$

For the case where P and Q are continuous distributions, we replace the sums by integrals and let $P(x)$ and $Q(x)$ denote probability densities. We also give a collection of well-known results on the Rényi divergence (cf. [LSS14]), many of which can be seen as multiplicative analogues of standard results for statistical distance. The proof of this lemma is given in [vEH14] and [LSS14].

Lemma 1 (Useful facts on Rényi divergence). *Let $a \in [1, +\infty]$. Also let P and Q be distributions such that $Supp(P) \subseteq Supp(Q)$. Then we have:*

- ***Increasing Function of the Order:*** *The function $a \mapsto R_a(P||Q)$ is non-decreasing, 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:*** *$R_a\left(P^f||Q^f\right) \leq R_a(P||Q)$ 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*
  - *$R_a(P||Q) = R_a(P_1||Q_1) \cdot R_a(P_2||Q_2)$.*
  - *$R_a(P||Q) = R_\infty(P_1||Q_1) \cdot \max_{y_1 \in Supp(P_1)} R_a\left(P_{2|1}(\cdot|y_1)||Q_{2|1}(\cdot|y_1)\right)$.*
- ***Probability Preservation:*** *Let $E \subseteq Supp(Q)$ be an arbitrary event. If $a \in (1, \infty)$, then $Q(E) \geq P(E)^{\frac{a}{a-1}} / R_a(P||Q)$. Furthermore, we have $Q(E) \geq P(E)/R_\infty(P||Q)$.*
- ***Weak Triangle Inequality:*** *Let P_1, P_2 and P_3 be three probability distributions such that $Supp(P_1) \subseteq Supp(P_2) \subseteq Supp(P_3)$. Then*

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

2.5 Gaussian Measures

Definition 7 (Continuous Gaussian distribution). *The Gaussian function of parameter r and centre c is defined as*

$$\rho_{r,c}(x) = \exp\left(-\pi(x-c)^2/r^2\right)$$

and the Gaussian distribution $D_{r,c}$ as the probability distribution whose probability density function is given by $\frac{1}{r}\rho_{r,c}$.

Definition 8 (Multivariate Gaussian distribution). *Let $\Sigma = S^T S$ for some rank-n matrix $S \in \mathbb{R}^{m \times n}$. The multivariate Gaussian function with covariance matrix Σ centred on $\boldsymbol{c} \in \mathbb{R}^n$ is defined as*

$$\rho_{S,\boldsymbol{c}}(\boldsymbol{x}) = \exp\left(-\pi(\boldsymbol{x}-\boldsymbol{c})^T \left(S^T S\right)^{-1} (\boldsymbol{x}-\boldsymbol{c})\right)$$

and the corresponding multivariate Gaussian distribution denoted $D_{S,\boldsymbol{c}}$ is defined by the density function $\frac{1}{\sqrt{\det(\Sigma)}}\rho_{S,\boldsymbol{c}}$.

Note that if the centre c is omitted, it should be assumed that $c = 0$. If the covariance matrix is diagonal, we describe it using the vector of its diagonal entries. For example, suppose that $(S^T S)_{ij} = (s_i)^2 \delta ij$ and let $\mathbf{s} = (s_1, \ldots s_n)^T$. Then we would write $D_{\mathbf{s}}$ to denote the centred Gaussian distribution D_S.

We are often interested in families of Gaussian distributions. For $\alpha > 0$, we write $\Psi_{\leq \alpha}$ to denote the set of Gaussian distributions with diagonal covariance matrix of parameter $\mathbf{r}$ satisfying $r_i \leq \alpha$ for all i.

We also have discrete Gaussian distributions i.e. normalised distributions defined over some discrete set. It is common to speak about discrete Gaussians over a lattice (or lattice cosets). The notation for a discrete Gaussian over some n-dimensional lattice Λ and coset vector $\mathbf{u} \in \mathbb{R}^n$ with parameter r is $D_{\Lambda+\mathbf{u},r}$. This distribution has probability mass function $\frac{1}{\rho_r(\Lambda+\mathbf{u})}\rho_r$ where $\rho_r(\Lambda + \mathbf{u}) = \sum_{\mathbf{x}\in\Lambda+\mathbf{u}} \rho_r(\mathbf{x})$. It was shown in [GPV08] that we can efficiently sample from a (not too narrow) discrete Gaussian over a lattice to within negligible statistical distance. It was further shown that we can actually sample the discrete Gaussian precisely in [BLP$^+$13]. This result is given below as Lemma 2.

Lemma 2 (Lemma 2.3 in [BLP$^+$13], Sampling discrete Gaussians). *There is a probabilistic polynomial-time algorithm that, given a basis $\boldsymbol{B}$ of an n-dimensional lattice $\Lambda = \mathcal{L}(\boldsymbol{B})$, $\boldsymbol{c} \in \mathbb{R}^n$ and parameter $r \geq \|\tilde{\boldsymbol{B}}\| \cdot \sqrt{\ln(2n+4)/\pi}$ outputs a sample distributed according to $D_{\Lambda+\mathbf{c},r}$.*

Next we define the smoothing parameter of a lattice followed by a collection of lemmas that we will make use of.

Definition 9 (Smoothing parameter). *For a lattice Λ and any $\epsilon > 0$, the* smoothing parameter *$\eta_\epsilon(\Lambda)$ is defined as the smallest $s > 0$ s.t. $\rho_{1/s}(\Lambda^* \backslash \{\boldsymbol{0}\}) \leq \epsilon$.*

Lemma 3 (Lemma 3.1 in [GPV08], Upper bound on smoothing parameter). *For any $\epsilon > 0$ and n-dimensional lattice Λ with basis $\boldsymbol{B}$,*

$$\eta_\epsilon(\Lambda) \leq \|\tilde{\boldsymbol{B}}\| \sqrt{\ln(2n(1+1/\epsilon))/\pi}.$$

Lemma 4 (Claim 3.8 in [Reg09], Sums of Gaussians over cosets). *For any n-dimensional lattice Λ, $\epsilon > 0$, $r \geq \eta_\epsilon(\Lambda)$ and $\boldsymbol{c} \in \mathbb{R}^n$, we have*

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

The claim $R_\infty(D_{\mathbf{t}}||Y) \leq \frac{1+\epsilon}{1-\epsilon}$ in the lemma below follows immediately from the proof given in [LS15].

Lemma 5 (Adapted from Lemma 7 in [LS15], Drowning ellipsoidal discrete Gaussians). *Let Λ be an n-dimensional lattice, $\boldsymbol{u} \in \mathbb{R}^n$, $\boldsymbol{r} \in (R^+)^n$, $\sigma > 0$ and $t_i = \sqrt{r_i^2 + \sigma^2}$ for all i. Assume that $r_i\sigma/t_i \geq \eta_\epsilon(\Lambda)$ for some $\epsilon \in (0, 1/2)$. Consider the continuous distribution Y on $\mathbb{R}^n$ obtained by sampling from $D_{\Lambda+\boldsymbol{u},\boldsymbol{r}}$ and then adding a vector from D_σ. Then we have $\Delta(Y, D_{\boldsymbol{t}}) \leq 4\epsilon$ and $R_\infty(D_{\boldsymbol{t}}||Y) \leq \frac{1+\epsilon}{1-\epsilon}$.*

In the lemma below, ring elements are sampled in the *coefficient* embedding.

Lemma 6 (Adapted from Lemma 4.1 in [SS13], Upper bound on least singular value). *Let n be a power of two and $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$. Then for any $\delta \in (0, 1)$, $t \geq \sqrt{2\pi}$ and $\sigma \geq \frac{t}{\sqrt{2\pi}} \cdot \eta_\delta(\mathbb{Z}^n)$, we have*

$$\Pr_{b \leftarrow\$ D_{\mathbb{Z}^n,\sigma}}\left[\frac{1}{\sigma_n(\mathrm{rot}(b))} \geq \frac{t\sqrt{2}}{\sigma\sqrt{n}}\right] \leq \frac{1+\delta}{1-\delta} \cdot \frac{n\sqrt{2\pi e}}{t}.$$

3 Reduction for General Rings

In this section, we show how to reduce an MLWE instance in module rank d and modulus q to an MLWE instance in rank d' and modulus q'. The particular case where $d' = 1$ yields a reduction from MLWE to RLWE. We start by describing the high-level intuition behind the reduction for the case $d' = 1$ and where the modulus goes from q to q^d. In this case, our strategy is to map $(\mathbf{a}, \mathbf{s}) \in (R_q)^d \times (R_q^\vee)^d$ to $(\tilde{a}, \tilde{s}) \in R_q \times R_{q'}^\vee$ aiming to satisfy the approximate equation

$$\frac{1}{q}\langle \mathbf{a}, \mathbf{s}\rangle \approx \frac{1}{q^d}(\tilde{a} \cdot \tilde{s}) \bmod R^\vee. \tag{3}$$

We then map from b to $\tilde{b} \approx b \bmod R^\vee$. For $q = \Omega(\mathrm{poly}(n))$, if we take $\tilde{s} = (q^{d-1}, \ldots, 1)^T \cdot \mathbf{s}$ and $\tilde{a} = (1, \ldots, q^{d-1})^T \cdot \mathbf{a}$, we obtain

$$\begin{aligned}\frac{1}{q^d}(\tilde{a} \cdot \tilde{s}) &= \frac{1}{q}\langle \mathbf{a}, \mathbf{s}\rangle + \frac{1}{q^2}(\ldots) + \frac{1}{q^3}(\ldots) + \ldots \bmod R \\ &\approx \frac{1}{q}\langle \mathbf{a}, \mathbf{s}\rangle \bmod R.\end{aligned} \tag{4}$$

This mapping satisfies the requirement but leads to a narrow, yet non-standard *discrete* error distribution. The reduction in Theorem 1 is a generalisation of the above idea. Specifically, take $\mathbf{G} \in (R)^{d' \times d}$ and $\tilde{\mathbf{s}} = \mathbf{G} \cdot \mathbf{s} \bmod (q'R)^{d'}$. Then we simply require that

$$\frac{1}{q'}\sum_{i=1}^{d'}\sum_{j=1}^{d}\tilde{a}_i g_{ij} s_j \approx \frac{1}{q}\sum_{j=1}^{d} a_j s_j \bmod R^\vee. \tag{5}$$

This requirement can be satisfied if we choose $\tilde{\mathbf{a}}$ such that

$$\frac{1}{q'}\sum_{i=1}^{d'}\tilde{a}_i g_{ij} \approx \frac{1}{q}a_j \bmod R \tag{6}$$

for $j = 1, \ldots, d$. To carry out this strategy, we will sample $\tilde{\mathbf{a}}$ over an appropriate lattice defined by $\mathbf{G}$ in the canonical embedding. The main challenge in applying this strategy is that we want the error in the new MLWE sample to follow a standard error distribution, i.e. a *continuous* Gaussian.

Theorem 1. *Let R be the ring of integers of some algebraic number field K of degree n, d, d', q, q' be integers, $\epsilon \in (0, 1/2)$, and $\boldsymbol{G} \in R^{d' \times d}$. Also, fix $\boldsymbol{s} = (s_1, \ldots, s_d) \in (R_q^{\vee})^d$. Further, let $\boldsymbol{B}_{\Lambda}$ be some known basis of the lattice $\Lambda = \frac{1}{q'} \boldsymbol{G}_H^T R^{d'} + R^d$ (in the canonical embedding), $\boldsymbol{B}_R$ be some known basis of R in H and*

$$r \geq \max \begin{cases} \|\tilde{\boldsymbol{B}}_{\Lambda}\| \ \cdot \sqrt{2\ln(2nd(1+1/\epsilon))/\pi} \\ \frac{1}{q} \ \|\tilde{\boldsymbol{B}}_R\| \cdot \sqrt{2\ln(2nd(1+1/\epsilon))/\pi} \\ \frac{1}{q}\|\tilde{\boldsymbol{B}}_{s_i R}\| \cdot \frac{1}{\min_k |\sigma_k(s_i)|} \cdot \sqrt{2\ln(2n(1+1/\epsilon))/\pi} \end{cases}$$

where $\boldsymbol{B}_{s_i R}$ is a basis of $s_i R$ in the canonical embedding. There exists an efficient probabilistic mapping $\mathcal{F} : (R_q)^d \times \mathbb{T}_{R^{\vee}} \longrightarrow (R_{q'})^{d'} \times \mathbb{T}_{R^{\vee}}$ such that:

1. *The output distribution given uniform input $\mathcal{F}(U((R_q)^d \times \mathbb{T}_{R^{\vee}}))$ is within statistical distance 4ϵ of the uniform distribution over $(R_{q'})^{d'} \times \mathbb{T}_{R^{\vee}}$.*
2. *Let $M = R^d$, $M' = R^{d'}$ and define $B := \max_{i,j} |\sigma_i(s_j)|$. The distribution of $\mathcal{F}(A^{(M)}_{q,\boldsymbol{s},D_{\alpha}})$ is within statistical distance $(4d+6)\epsilon$ of $A^{(M')}_{q',\boldsymbol{Gs},D_{\boldsymbol{\alpha}'}}$ where $(\boldsymbol{\alpha}')_i^2 = \alpha^2 + r^2(\beta^2 + \sum_{j=1}^d |\sigma_i(s_j)|^2)$ and β satisfies $\beta^2 \geq B^2 d$.*

Proof. We use the canonical embedding on each component of R^d individually, e.g. $\mathbf{a}_H = (\sigma_H(a_1), \ldots, \sigma_H(a_d)) \in H^d \simeq \mathbb{R}^{nd}$ and similarly for other module elements. We will also refer to the canonical embedding of R as simply R to ease notation. Suppose we are given $(\mathbf{a}, b) \in (R_q)^d \times \mathbb{T}_{R^{\vee}}$. The mapping $\mathcal{F}$ is performed as follows:

1. Sample $\mathbf{f} \leftarrow D_{\Lambda - \frac{1}{q}\mathbf{a}_H, r}$. Note that the parameter r is large enough so that we can sample the discrete Gaussian efficiently by Lemma 2.
2. Let $\mathbf{v} = \frac{1}{q}\mathbf{a}_H + \mathbf{f} \in \Lambda / R^d$ and set $\mathbf{x} \in R_{q'}^{d'}$ to be a random solution of $\frac{1}{q'}\mathbf{G}_H^T \mathbf{x} = \mathbf{v} \bmod R^d$. Then set $\tilde{\mathbf{a}} \in M'$ to be the unique element of M' such that $\tilde{\mathbf{a}}_H = \mathbf{x}$.
3. Sample $\tilde{e}$ from the distribution $D_{r\beta}$ over $K_{\mathbb{R}} \simeq H$ for some $\beta > B\sqrt{d}$ and set $\tilde{b} = b + \tilde{e}$.
4. Finally, output $(\tilde{\mathbf{a}}, \tilde{b}) \in (R_{q'})^{d'} \times \mathbb{T}_{R^{\vee}}$.

Distribution of $\tilde{a}$. Suppose that $\mathbf{a} \in (R_q)^d$ was drawn uniformly at random. Step 2 of the reduction can be performed by adding a random element of the basis of solutions to $\frac{1}{q'}\mathbf{G}_H^T \mathbf{y} = 0 \bmod R^d$ to a particular solution of $\frac{1}{q'}\mathbf{G}_H^T \mathbf{x} = \mathbf{v} \bmod R^d$. In order to show that $\tilde{\mathbf{a}}$ is *nearly* uniform random, we will show that the vector $\mathbf{x}$ is *nearly* uniform random over the set $(R_{q'})^{d'}$. Note that every $\mathbf{x} \in (R_{q'})^{d'}$ is a solution to $\frac{1}{q'}\mathbf{G}_H^T \mathbf{x} = \mathbf{v} \bmod R^d$ for some $\mathbf{v}$ and the number of solutions to this equation in $R_{q'}^{d'}$ for each $\mathbf{v}$ is the same. Thus, proving that $\mathbf{v}$ is *almost* uniform suffices. Observe that $r \geq \eta_{\epsilon}(\Lambda)$. Therefore, Lemma 4 tells us that for any particular $\bar{\mathbf{a}} \in (R_q)^d$ and $\bar{\mathbf{f}} \in \mathit{\Lambda} - \frac{1}{q}\bar{\mathbf{a}}_H$, we have

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

where $C := q^{-nd}/\rho_r(\mathit{\Lambda})$ is a constant. By summing this equation over appropriate values of $\bar{\mathbf{a}}$ and $\bar{\mathbf{f}}$, Lemma 4 tells us that for any coset $\bar{\mathbf{v}} \in \mathit{\Lambda}/R^d$,

$$\begin{aligned}\Pr[\mathbf{v} = \bar{\mathbf{v}}] &\in C \cdot \left[1, \frac{1+\epsilon}{1-\epsilon}\right] \cdot \rho_r(q^{-1}R^d + \bar{\mathbf{v}}) \\ &\in C \cdot \rho_r(q^{-1}R^d) \cdot \left[1, \frac{1+\epsilon}{1-\epsilon}\right] \cdot \frac{\rho_r(q^{-1}R^d + \bar{\mathbf{v}})}{\rho_r(q^{-1}R^d)} \\ &\in C' \cdot \left[\frac{1-\epsilon}{1+\epsilon}, \frac{1+\epsilon}{1-\epsilon}\right]\end{aligned} \tag{8}$$

where $C' := C\rho_r(q^{-1}R^d)$. Note that we may apply Lemma 4 here since we know that $r \geq \eta_\epsilon((q)^{-1}R^d)$ by Lemma 3. This allows us to conclude that the distribution of $\mathbf{v}$ is within statistical distance $1 - [(1-\epsilon)/(1+\epsilon)]^2 \leq 4\epsilon$ of the uniform distribution. This means that $\mathbf{x}$ is uniformly random over $(R_{q'})^{d'}$ to within statistical distance 4ϵ implying that $\tilde{\mathbf{a}}$ is uniform random over $(R_{q'})^{d'}$ to within statistical distance 4ϵ. It is also clear that $\tilde{b}$ is exactly uniform random given that b is uniform random. This proves the first claim (uniform-to-uniform).

Distribution of $-\boldsymbol{f}$. In our analysis of the resulting error, it will be useful to understand the distribution of the vector $-\mathbf{f}$ for fixed $\tilde{\mathbf{a}}$ (and thus fixed $\mathbf{v} = \bar{\mathbf{v}}$). Note that fixing a value $\mathbf{f} = \bar{\mathbf{f}}$ fixes $\frac{1}{q}\mathbf{a} = \bar{\mathbf{v}} - \bar{\mathbf{f}} \bmod R^d$. By summing over all appropriate values of $\mathbf{f}$ in Equation 7, one can show that the distribution of $-\mathbf{f}$ for any fixed $\tilde{\mathbf{a}}$ is within statistical distance $1 - (1-\epsilon)(1+\epsilon) \leq 2\epsilon$ of $D_{\frac{1}{q}R^d - \bar{\mathbf{v}}, r}$.

Distribution of the error. Suppose we are given the MLWE sample $(\mathbf{a}, b = \frac{1}{q}\langle \mathbf{a}, \mathbf{s}\rangle + e) \in (R_q)^d \times \mathbb{T}_{R^\vee}$ where $e \in K_\mathbb{R}$ is drawn from D_α. We have already shown that our map outputs $\tilde{\mathbf{a}} \in (R_{q'})^{d'}$ that is *almost* uniformly random. Now we condition on a fixed $\tilde{\mathbf{a}} = \bar{\tilde{\mathbf{a}}}$ and analyse the distribution of

$$(\tilde{b} - \frac{1}{q'}\langle \bar{\tilde{\mathbf{a}}} \cdot \tilde{\mathbf{s}}\rangle) \bmod R^\vee. \tag{9}$$

Let $\mathbf{f}_i \in \mathbb{R}^n$ be the vector consisting of the i^{th} block of n entries of $\mathbf{f} \in \mathbb{R}^{nd}$ for $i = 1, \ldots, d$. Using the fact that $\tilde{\mathbf{s}} = \mathbf{G}\mathbf{s}$ and that $R^\vee$ is closed under multiplication by elements of R, we can rewrite this as

$$(\tilde{b} - \frac{1}{q'} \left\langle \bar{\tilde{\mathbf{a}}} \cdot \tilde{\mathbf{s}} \right\rangle) = \sum_{i=1}^{d} s_i \cdot \sigma_H^{-1}(-\mathbf{f}_i) + \tilde{e} + e \bmod R^{\vee}. \qquad (10)$$

In fact, we want to analyse the RHS of the above equation in canonically embedded space. To do so, define the invertible matrix $S_{i,H} := U_H S_i U_H^{\dagger} \in \mathbb{R}^{n \times n}$ where U_H is given in Equation 2 and S_i is the diagonal matrix with the field embeddings of s_i along the diagonal i.e. $[S_i]_{jk} = \sigma_j(s_i)\delta_{jk}$. Note that $S_{i,H}$ is the matrix representing multiplication by s in the basis $(\mathbf{h}_i)_{i=1}^{n}$ of H. Therefore, in canonical space, the error is given by

$$\sum_{i=1}^{d} S_{i,H} \cdot (-\mathbf{f}_i) + \sigma_H(\tilde{e}) + \sigma_H(e) \bmod R^{\vee} \qquad (11)$$

where $\sigma_H(\tilde{e})$ and $\sigma_H(e)$ are distributed as $D_{r\beta}$ and D_α respectively. Also, letting $\bar{\mathbf{v}}_i$ denote the i^{th} block of n coordinates of $\bar{\mathbf{v}}$, we know that $-\mathbf{f}_i$ is *almost* distributed as $D_{\frac{1}{q}R - \bar{\mathbf{v}}_i, r}$. It then follows that $S_{i,H} \cdot (-\mathbf{f}_i)$ is close in distribution to $D_{\frac{1}{q}S_{i,H} \cdot R - S_{i,H} \cdot \bar{\mathbf{v}}_i, r(S_{i,H})^T}$ i.e. an ellipsoidal discrete Gaussian. In fact the covariance matrix $r^2 S_{i,H} S_{i,H}^T$ is diagonal with respect to our basis $(\mathbf{h}_i)_{i=1}^{n}$ of $\mathbb{R}^n$ (see Section 2.2) with eigenvalues given by $r^2 |\sigma_j(s_i)|^2$ for $j = 1, \ldots, n$. Note that we can conceptualise $\sigma_H(\tilde{e})$ as $\sum_{i=1}^{d} \tilde{e}^{(i)}$ where each $\tilde{e}^{(i)}$ is distributed as a continuous spherical Gaussian in $\mathbb{R}^n$ with parameter $\gamma_i \geq rB$. Recalling that $-\mathbf{f}$ is distributed as $D_{\frac{1}{q}R^d - \bar{\mathbf{v}}, r}$ to within statistical distance 2ϵ, we can now apply Lemma 5 d times to conclude that

$$\sum_{i=1}^{d} S_{i,H} \cdot (-\mathbf{f}_i) + \sigma_H(\tilde{e}) = \sum_{i=1}^{d} S_{i,H} \cdot (-\mathbf{f}_i) + \tilde{e}^{(i)} \qquad (12)$$

is distributed as the continuous Gaussian with a diagonal covariance matrix to within statistical distance $2\epsilon + 4d\epsilon$. In particular, the diagonal entries of the convariance matrix are given by $r^2 \left(\beta^2 + \sum_{j=1}^{d} |\sigma_i(s_j)|^2 \right)$ for $i = 1, \ldots, n$. Considering the original error term $\sigma_H(e)$ that follows the distribution D_α completes the proof. □

Remark 2. It is permissible to take $B := \min_{i,j} |\sigma_j(s_i)|$ in the above theorem. However, this will not save us any asymptotic factors in the output error distribution so we use $B := \max_{i,j} |\sigma_j(s_i)|$ to allow for cleaner looking bounds.

The following corollary specialises to a map from MLWE in module rank d to d/k and from modulus q to q^k for general rings. Taking $k = d$ constitutes a reduction from MLWE to RLWE. Note that the new secret distribution is nonstandard in general, but we can always use the usual re-randomizing process to obtain a uniform secret. We also highlight the fact that the lower bound on r is not particularly tight due to a loose upper bound on the quantities $\|\tilde{\mathbf{B}}_{s_i R}\|$. This issue is addressed for power-of-two cyclotomics in Section 3.1. In fact, for a general cyclotomic ring R, it holds that $\|\mathbf{B}_{s_i R}\| = \|\sigma_H(s_i)\|$.

Corollary 1. *Let R be a ring with basis $\boldsymbol{B}_R$ in the canonical embedding and χ be a distribution satisfying*

$$\Pr_{s \leftarrow\$ \chi}\left[\max_i |\sigma_i(s)| > B\right] \leq \delta \text{ and } \Pr_{s \leftarrow\$ \chi}\left[\max_{i,j} \frac{|\sigma_i(s)|}{|\sigma_j(s)|} > B'\right] \leq \delta'$$

for some (B, δ) and (B', δ'). Also let $\alpha > 0$ and take any $\epsilon \in (0, 1/2)$. For any $k > 1$ that divides d and

$$r \geq \max\begin{cases} \frac{1}{q} \ \|\tilde{\boldsymbol{B}}_R\| \ \cdot \sqrt{2\ln(2nd(1+1/\epsilon))/\pi} \\ \frac{1}{q} \ B' \ \|\tilde{\boldsymbol{B}}_R\| \ \cdot \sqrt{2\ln(2nd(1+1/\epsilon))/\pi} \end{cases},$$

there is an efficient reduction from $MLWE^{(R^d)}_{m,q,\Psi_{\leq \alpha}}(\chi^d)$ to $MLWE^{(R^{d/k})}_{m,q^k,\Psi_{\leq \alpha'}}(\boldsymbol{G} \cdot \chi^d)$ for $\boldsymbol{G} = \mathbb{I}_{d/k} \otimes (1, q, \ldots, q^{k-1}) \in R^{d/k \times d}$ and

$$(\alpha')^2 \geq \alpha^2 + 2r^2B^2d.$$

Moreover, this reduction reduces the advantage by at most $[1 - (1 - \delta - \delta')^d] + (4d + 10)\epsilon m$.

Proof. We run the reduction from Theorem 1, taking $q' = q^k$, $\beta^2 \geq B^2d$ and $\mathbf{G} \in R^{d/k \times d}$ as in the corollary statement. First, note that $\|\tilde{\mathbf{B}}_{s_iR}\| \leq \max_j |\sigma_j(s_i)| \cdot \|\tilde{\mathbf{B}}_R\|$ by considering multiplication in the canonical embedding and Lemma 2 from [ABB10]. In the *coefficient* embedding, we have that $\mathbf{G} = \mathbb{I}_{d/k} \otimes (1, q, \ldots, q^{k-1}) \otimes \mathbb{I}_n$ and the lattice of interest is $\frac{1}{q^k}\mathbf{G}^T\mathbb{Z}^{nd/k} + \mathbb{Z}^{nd}$ with basis $\mathbf{B} = \mathbb{I}_{d/k} \otimes \mathbf{Q} \otimes \mathbb{I}_n$ where

$$\mathbf{Q} = \begin{bmatrix} q^{-1} & q^{-2} & \ldots & q^{-k} \\ & q^{-1} & \ldots & q^{1-k} \\ & & \ddots & \vdots \\ & & & q^{-1} \end{bmatrix}.$$

To move from the coefficient embedding to the canonical embedding, we simply multiply by the matrix $\mathbf{B}_{R^d} := \mathbb{I}_d \otimes \mathbf{B}_R$. Therefore, in the canonical embedding, the basis is given by $\mathbf{B}_\Lambda = \mathbb{I}_{d/k} \otimes \mathbf{Q} \otimes \mathbf{B}_R$. Orthogonalising from left to right, we can see that $\|\tilde{\mathbf{B}}_\Lambda\|$ is precisely $\frac{1}{q}\|\tilde{\mathbf{B}}_R\|$.

Let E be the event that $\max_i |\sigma_i(s)| \leq B$ and F be the event $\max_{i,j} \frac{|\sigma_i(s)|}{|\sigma_j(s)|} \leq B'$ where $s \leftarrow\$ \chi$. The fact that $P(E \cap F) = P(E) + P(F) - P(E \cup F) \geq P(E) + P(F) - 1 \geq 1 - \delta - \delta'$ implies the result. □

3.1 Power-of-Two Cyclotomic Rings

We now give a more specific account of Theorem 1 in the case where R for power-of-two is a cyclotomic ring, i.e. $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ for power-of-two n. We will also be considering discrete Gaussian secret distributions and normal form MLWE. The corollary given in this section is almost identical to Corollary 1 apart from the definition of the pairs (B, δ) and (B', δ'). This change makes the corollary amenable to known results for discrete Gaussian secret s.

It can be shown that the map taking the canonical embedding to the coefficient embedding is a scaled isometry with scaling factor $1/\sqrt{n}$. In particular, the canonical to coefficient embedding map sends a spherical Gaussian r to $r/\sqrt{n}$. Furthermore, the dual ring is given by $R^\vee := \frac{1}{n} \cdot R$ and takes the simple form of $\frac{1}{n}\mathbb{Z}^n$ in the coefficient embedding.

Let $\tau > 0$. We will be considering the case where the secret s is drawn from $D_{R^\vee,\tau}$ (and then reduced modulo $qR^\vee$). In the coefficient embedding, this is equivalent to drawing the secret from the distribution $D_{\frac{1}{n}\mathbb{Z}^n, \tau/\sqrt{n}}$ i.e. $\frac{1}{n} \cdot D_{\mathbb{Z}^n, \tau\sqrt{n}}$.

Let S_H be the matrix of multiplication by s in the canonical embedding. For cyclotomic power-of-two rings, there is a simple relationship between components of the canonical embedding $\sigma_i(s)$ and the singular values of the matrix $\mathrm{rot}(s)$. Let $\mathbf{B}_R = \sqrt{n} \cdot U$ denote the scaled isometry mapping from coefficient space to canonical space where U is unitary. Then we have $S_H^T S_H = U^{-1} \cdot \mathrm{rot}(s)^T \mathrm{rot}(s) \cdot U$. Since $S_H^T S_H$ is diagonal with elements given by $|\sigma_i(s)|^2$, the eigenvalues of $\mathrm{rot}(s)^T \mathrm{rot}(s)$ are exactly these diagonal elements. This implies $|\sigma_i(s)|$ are exactly the singular values of $\mathrm{rot}(s)$. We will use this fact in the next claim.

Lemma 7. *Let $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ for some power-of-two n. Then for any $\delta \in (0,1)$, $t \geq \sqrt{2\pi}$ and $\tau \geq \frac{t}{\sqrt{2\pi n}} \cdot \eta_\delta(\mathbb{Z}^n)$, we have*

$$\mathrm{Pr}_{s \leftarrow\!\!\!\$\, D_{R^\vee,\tau}} \left[\frac{1}{\min_i |\sigma_i(s)|} \geq \frac{t\sqrt{2}}{\tau} \right] \leq \frac{1+\delta}{1-\delta} \cdot \frac{n\sqrt{2\pi e}}{t}.$$

Proof. Let $b = ns$. The distribution of b is $D_{\mathbb{Z}^n, \tau\sqrt{n}}$. Let $\sigma_n(\mathrm{rot}(b))$ denote the least singular value of $\mathrm{rot}(b)$. Now we can write

$$\begin{aligned}
\mathrm{Pr}_{s \leftarrow\!\!\!\$\, D_{R^\vee,\tau}} \left[\frac{1}{\min_i |\sigma_i(s)|} \geq \frac{t\sqrt{2}}{\tau} \right] &= \mathrm{Pr}_{s \leftarrow\!\!\!\$\, D_{\frac{1}{n}\mathbb{Z}^n, \frac{\tau}{\sqrt{n}}}} \left[\frac{1}{\sigma_n(\mathrm{rot}(s))} \geq \frac{t\sqrt{2}}{\tau} \right] \\
&= \mathrm{Pr}_{b \leftarrow\!\!\!\$\, D_{\mathbb{Z}^n, \tau\sqrt{n}}} \left[\frac{1}{\sigma_n(\mathrm{rot}(b))} \geq \frac{t\sqrt{2}}{(\tau\sqrt{n}) \cdot \sqrt{n}} \right] \\
&\leq \frac{1+\delta}{1-\delta} \cdot \frac{n\sqrt{2\pi e}}{t}
\end{aligned}$$

where the inequality comes from Lemma 6. □

In the proof of the following lemma, we will say that a distribution D over $\mathbb{Z}^n$ is (B, δ)-bounded for real numbers $B, \delta > 0$ if $\mathrm{Pr}_{\mathbf{x} \leftarrow\!\!\!\$\, D}[\|\mathbf{x}\| > B] \leq \delta$.

Lemma 8. *Let $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ for some power-of-two n. Then for any $\delta \in (0,1)$ and $\tau \geq 0$,*

$$\mathrm{Pr}_{s \leftarrow\!\!\!\$\, D_{R^\vee,\tau}} \left[\|\sigma_H(s)\| > C\tau\sqrt{n \log(n/\delta)} \right] \leq \delta$$

for some universal constant $C > 0$. We also have that

$$\Pr_{s \leftarrow\$ D_{R^\vee,\tau}}\left[\|\sigma_H(s)\| > \tau\sqrt{n}\right] \leq 2^{-n}.$$

Proof. Take $B > 0$ and let $b = ns$. We have

$$\begin{aligned}\Pr_{s \leftarrow\$ D_{R^\vee,\tau}}\left[\|\sigma_H(s)\| > B\right] &= \Pr_{s \leftarrow\$ D_{\frac{1}{n}\mathbb{Z}^n,\frac{\tau}{\sqrt{n}}}}\left[\|s_{vec}\| > B/\sqrt{n}\right] \\ &= \Pr_{b \leftarrow\$ D_{\mathbb{Z}^n,\tau\sqrt{n}}}\left[\|b\| > B\sqrt{n}\right].\end{aligned}$$

As mentioned in [BLP+13], we know that $D_{\mathbb{Z}^n,r}$ is $(Cr\sqrt{n\log(n/\delta)}, \delta)$-bounded for some universal constant $C > 0$ by taking a union bound over the n coordinates. Furthermore, an application of Lemma 1.5 in [Ban93] implies that $D_{\mathbb{Z}^n,r}$ is $(r\sqrt{n}, 2^{-n})$-bounded. Applying these results completes the proof. □

Corollary 2. *Let $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ for power-of-two n and χ be a distribution over $R^\vee$ satisfying*

$$\Pr_{s \leftarrow\$ \chi}\left[\|\sigma_H(s)\| > B_1\right] \leq \delta_1 \text{ and } \Pr_{s \leftarrow\$ \chi}\left[\max_j \frac{1}{|\sigma_j(s)|} \geq B_2\right] \leq \delta_2$$

for some (B_1, δ_1) and (B_2, δ_2). Also let $\alpha > 0$ and take any $\epsilon \in (0, 1/2)$. For any $k > 1$ that divides d,

$$r \geq \left(\frac{\max\{\sqrt{n}, B_1B_2\}}{q}\right) \cdot \sqrt{2\ln(2nd(1 + 1/\epsilon))/\pi},$$

there is an efficient reduction from $MLWE^{(R^d)}_{m,q,\Psi_{\leq\alpha}}(\chi^d)$ to $MLWE^{(R^{d/k})}_{m,q^k,\Psi_{\leq\alpha'}}(\boldsymbol{G}\cdot\chi^d)$ for $\boldsymbol{G} = \mathbb{I}_{d/k} \otimes (1, q, \ldots, q^{k-1}) \in R^{d/k\times d}$ and

$$(\alpha')^2 \geq \alpha^2 + 2r^2B_1^2d.$$

Moreover, this reduction reduces the advantage by at most $[1 - (1 - \delta_1 - \delta_2)^d] + (4d + 10)\epsilon m$.

Proof. We apply Theorem 1 taking $\beta^2 \geq B_1^2d$. For power-of-two cyclotomic rings, $\|\mathbf{B}_{s_iR}\| = \|\sigma_H(s)\|$. Furthermore, if $B_1 \geq \|\sigma_H(s)\|$, then it is guaranteed that $B_1 \geq \max_i |\sigma_i(s)|$. The rest of the proof is the same as in Corollary 1. □

To put the above corollary into context, we now discuss the pairs (B_1, δ_1) and (B_2, δ_2) when the secret distribution χ is $D_{R^\vee,\tau}$. From Lemma 8, for any $\delta_1 \in (0, 1)$, we have $B_1 = \mathcal{O}(\tau\sqrt{n\log(n/\delta_1)})$. Next, for any $\delta_2 \in (0, 1)$, we fix the parameter δ from Lemma 7 (e.g. $\delta = 1/2$) and take t from Lemma 7 proportional to n/δ_2. Then, as long as $\tau \geq \mathcal{O}(\sqrt{n\log(n)}/\delta_2)$, we can take $B_2 = \mathcal{O}(n/(\tau\delta_2))$. To summarize, we may take:

- $B_1 = \mathcal{O}(\tau\sqrt{n\log(n/\delta_1)})$ for arbitrary $\tau > 0$ and $\delta_1 \in (0, 1)$
- $B_2 = \tilde{\mathcal{O}}\left(\frac{n}{\tau\delta_2}\right)$ for $\tau \geq \mathcal{O}(\sqrt{n\log(n)}/\delta_2)$ and any $\delta_2 \in (0, 1)$

- $B_1 B_2 = \tilde{\mathcal{O}}\left(\frac{n\sqrt{n\log(n/\delta_1)}}{\delta_2}\right)$ for $\tau \geq \mathcal{O}(\sqrt{n\log(n)}/\delta_2)$ and any $\delta_1, \delta_2 \in (0,1)$.

In an ideal setting, we would like to conclude that a probabilistic polynomial-time algorithm that solves RLWE with non-negligible advantage implies a probabilistic polynomial-time algorithm capable of solving MLWE with non-negligible advantage. In order to achieve this, it is necessary that the loss in advantage incurred by any reduction should be negligible in the security parameter λ. Therefore, we would require that δ_1, δ_2 and ϵ all be negligible in the corollaries above. The requirement that δ_2 be negligible is particularly troublesome since this implies that B_1 and B_2 are super-polynomial in λ if we want to use the results above. This would mean that the resulting error in our reduction would also be super-polynomial. In particular, the case of normal form MLWE where $\tau = \alpha q$ $(= \mathsf{poly}(n))$ is not covered by the analysis given in the case that δ_2 is negligible. This issue will be addressed in Section 4 where we show that taking $\delta_2 = \mathcal{O}(1/d)$ suffices when considering *search* variants.

Yet, the analysis given so far remains relevant for sufficiently good algorithms for solving RLWE. For example, given access to an algorithm solving RLWE with advantage $1/\mathsf{poly}(\lambda)$, it would be adequate to consider δ_1, δ_2 and ϵ as $1/\mathsf{poly}(\lambda)$. These choices lead to a reduction from MLWE to RLWE (with polynomial noise) with $1/\mathsf{poly}(\lambda)$ loss in advantage which is acceptable given a sufficiently effective algorithm for solving RLWE.

4 Search Reductions Using Rényi Divergence

In this section, we apply the Rényi divergence. Given our analysis of the reduction explicited in Theorem 1, it is fairly straight-forward to obtain analogous results based on Rényi divergence. We will show that our reduction can be used to solve search MLWE with non-negligible probability given an algorithm for solving search RLWE with non-negligible success probability. Note that this result could be derived from statistical distance arguments, but we choose to use the Rényi divergence because it later allows us to reduce to a strictly spherical error distribution while increasing the width of the resulting error distribution only by small powers of n. In contrast, statistical distance arguments would require the drowning noise to increase by super-polynomial factors. This is because we require negligible statistical distances to target distributions whereas we only require that Rényi divergences are $\mathcal{O}(1)$ to obtain meaningful results.

Theorem 2. *Let R be the ring of integers of some algebraic number field K of degree n, d, d', q, q' be integers, $\epsilon \in (0, 1/2)$, and $\boldsymbol{G} \in R^{d' \times d}$. Also, fix $\boldsymbol{s} = (s_1, \ldots, s_d) \in (R_q^\vee)^d$. Further, let $\boldsymbol{B}_\Lambda$ be some known basis of the lattice $\Lambda = \frac{1}{q'}\boldsymbol{G}_H^T R^{d'} + R^d$ (in the canonical embedding), $\boldsymbol{B}_R$ be some known basis of*

R in H and

$$r \geq \max \begin{cases} \|\tilde{\boldsymbol{B}}_{\Lambda}\| \ \cdot \sqrt{2\ln(2nd(1+1/\epsilon))/\pi} \\ \frac{1}{q} \ \|\tilde{\boldsymbol{B}}_{R}\| \cdot \sqrt{2\ln(2nd(1+1/\epsilon))/\pi} \\ \frac{1}{q}\|\tilde{\boldsymbol{B}}_{s_iR}\| \cdot \frac{1}{\min_k |\sigma_k(s_i)|} \cdot \sqrt{2\ln(2n(1+1/\epsilon))/\pi} \end{cases}$$

where $\boldsymbol{B}_{s_iR}$ is a basis of s_iR in the canonical embedding. Let $M = R^d$, $M' = R^{d'}$ and define $B \coloneqq \max_{i,j} |\sigma_i(s_j)|$. There exists an efficient probabilistic mapping $\mathcal{F} : (R_q)^d \times \mathbb{T}_{R^\vee} \longrightarrow (R_{q'})^{d'} \times \mathbb{T}_{R^\vee}$ such that

$$R_\infty\left(A^{(M')}_{q',\boldsymbol{Gs},D_{\boldsymbol{\alpha}'}} || \mathcal{F}(A^{(M)}_{q,\boldsymbol{s},D_\alpha})\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^{d+3}$$

where $(\boldsymbol{\alpha}')_i^2 = \alpha^2 + r^2(\beta^2 + \sum_{j=1}^d |\sigma_i(s_j)|^2)$ and β satisfies $\beta^2 \geq B^2 d$.

Proof. We take the mapping $\mathcal{F}$ described in the proof of Theorem 1 and adopt the same notation. Recall that $(\tilde{\mathbf{a}}, \tilde{b})$ denotes the output of $\mathcal{F}$. Denote the distribution of interest $\mathcal{F}(A^{(M)}_{q,\mathbf{s},D_\alpha})$ as $\tilde{A}^{(M')}_{q',\mathbf{Gs},\tilde{D}}$ i.e. the distribution of $(\tilde{\mathbf{a}}, \tilde{b})$ given that $(\mathbf{a}, b)$ follows the distribution $A^{(M)}_{q,\mathbf{s},D_\alpha}$.

Distribution of $\tilde{\mathbf{a}}$. Let K_{sol} denote the number of solutions to the equation $\frac{1}{q'}\mathbf{G}_H^T\mathbf{x} = \mathbf{v} \bmod R^d$ and K_v the number of possible vectors $\mathbf{v}$. Recall that K_{sol} is constant in $\mathbf{v}$. For any $\bar{\tilde{\mathbf{a}}} \in R_{q'}^{d'}$, we have (from Equation 8) that

$$\begin{aligned} \Pr[\tilde{\mathbf{a}} = \bar{\tilde{\mathbf{a}}}] &= \sum_{\bar{\mathbf{v}} \in \Lambda/\mathbb{Z}^{nd}} \Pr[\tilde{\mathbf{a}} = \bar{\tilde{\mathbf{a}}} | \mathbf{v} = \bar{\mathbf{v}}] \cdot \Pr[\mathbf{v} = \bar{\mathbf{v}}] \\ &\geq C' \cdot \left(\frac{1-\epsilon}{1+\epsilon}\right) \frac{1}{K_{sol}} \geq \left(\frac{1-\epsilon}{1+\epsilon}\right)^2 \cdot \frac{1}{K_{sol}K_v}. \end{aligned}$$

Note that picking $\tilde{\mathbf{a}}$ at random is identical to choosing $\mathbf{v}$ at random followed by picking a uniformly random solution to $\frac{1}{q'}\mathbf{G}_H^T\mathbf{x} = \mathbf{v} \bmod R^d$. Therefore, the distribution of $\tilde{\mathbf{a}}$ which we denote by $D^{(\tilde{\mathbf{a}})}$ satisfies

$$R_\infty\left(U(R_{q'}^{d'}) || D^{(\tilde{\mathbf{a}})}\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2. \tag{13}$$

Distribution of $-\boldsymbol{f}$. Before, we concluded that the distribution of $-\mathbf{f}$ was close in statistical distance to $D_{\frac{1}{q}R^d - \bar{\mathbf{v}}, r}$ conditioned on some fixed $\tilde{\mathbf{a}}$. Once again, summing over appropriate values of $\mathbf{f}$ in Equation 7 tells us that

$$\Pr[-\mathbf{f} = \bar{\mathbf{f}} | \tilde{\mathbf{a}} = \bar{\tilde{\mathbf{a}}}] \geq C \cdot \rho_r(\bar{\mathbf{f}}) \geq \frac{1-\epsilon}{1+\epsilon} \cdot \frac{\rho_r(\bar{\mathbf{f}})}{\rho_r(\frac{1}{q}R^d - \bar{\mathbf{v}})}.$$

Therefore, writing $D^{(-\mathbf{f})}$ as the distribution of $-\mathbf{f}$, we see that

$$R_\infty\left(D_{\frac{1}{q}R^d - \bar{\mathbf{v}}, r} || D^{(-\mathbf{f})}\right) \leq \frac{1+\epsilon}{1-\epsilon}.$$

Distribution of the error term. We now analyse the distribution of the error term given in Equation 10. Let $\mathbf{f}_i$ denote the i^{th} block of n consecutive coordinates of $\mathbf{f} \in \mathbb{R}^{nd}$ Once again, we split the RHS of this error term and analyse it as $\sum_{i=1}^{d}\left(S_{i,H}^T \cdot (-\mathbf{f}_i) + \tilde{e}^{(i)}\right) + e$ where each $\tilde{e}^{(i)}$ is sampled independently from a continuous Gaussian on $\mathbb{R}^n$ with parameter $\gamma_i \geq rB$. Let $D^{(i)}$ denote the distribution of $\left(S_{i,H}^T \cdot (-\mathbf{f}_i) + \tilde{e}^{(i)}\right)$. We now use the data-processing inequality with the function $(-\mathbf{f}, \tilde{e}^{(1)}, \ldots, \tilde{e}^{(d)}) \longmapsto (S_{1,H}^T \cdot (-\mathbf{f}_1) + \tilde{e}^{(1)}, \ldots, S_{d,H}^T \cdot (-\mathbf{f}_d) + \tilde{e}^{(d)})$. For $i = 1, \ldots, d$, define $Y^{(i)}$ as the distribution obtained by sampling from $D_{\frac{1}{q}S_{i,H}R+S_{i,H}\cdot\bar{\mathbf{v}}_i,r(S_{i,H}^T)}$ and then adding a vector sampled from D_{γ_i}. Note that $Y^{(i)}$ is the distribution of $\mathbf{S}_i^T \cdot (-\mathbf{f}_i) + \tilde{e}^{(i)}$ in the case that the distribution of $-\mathbf{f}$ is *exactly* $D_{\frac{1}{q}R^d-\bar{\mathbf{v}},r}$. Let $D_\gamma = D_{\gamma_1} \times \ldots \times D_{\gamma_d}$. The data-processing inequality for Rényi divergence implies that

$$\begin{aligned} R_\infty\left(Y^{(1)} \times \cdots \times Y^{(d)}||D^{(1)} \times \cdots \times D^{(d)}\right) &\leq R_\infty\left(D_{\frac{1}{q}R^d-\bar{\mathbf{v}},r} \times D_\gamma||D^{(-\mathbf{f})} \times D_\gamma\right) \\ &\leq \frac{1+\epsilon}{1-\epsilon}. \end{aligned}$$

Now we apply Lemma 5 by recalling that the covariance matrix $S_{i,H}^T S_{i,H}$ is diagonal with elements $|\sigma_j(s_i)|$ for $j = 1, \ldots n$. This allows us to conclude that for $i = 1, \ldots, d$,

$$R_\infty\left(D_{(\gamma_i^2+r^2S_{i,H}^TS_{i,H})^{1/2}}||Y^{(i)}\right) \leq \frac{1+\epsilon}{1-\epsilon}.$$

By first applying the data-processing inequality to the function that sums the samples and then considering the triangle inequality and independence, the above equation implies that

$$\begin{aligned} R_\infty\left(D_{(\alpha^2+r^2\beta^2+r^2\sum_{i=1}^{d}S_{i,H}^TS_{i,H})^{1/2}}||\tilde{D}\right) &\leq \frac{1+\epsilon}{1-\epsilon} \cdot \prod_{i=1}^{d} R_\infty\left(D_{(\gamma^2+r^2S_{i,H}^TS_{i,H})^{1/2}}||Y^{(i)}\right) \\ &\leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^{d+1} \end{aligned} \tag{14}$$

where $\tilde{D}$ is the distribution of the RHS of Equation 10 (i.e. the sum of the distributions $D^{(i)}$).

Distribution of the reduction's output. We now complete the proof by combining the results above.

$$\begin{aligned} R_\infty\left(A^{(M')}_{q',\mathbf{Gs},D_{\boldsymbol{\alpha}'}}||\tilde{A}^{(M')}_{q',\mathbf{Gs},\tilde{D}}\right) &\leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \cdot R_\infty\left(D_{\boldsymbol{\alpha}'}||\tilde{D}\right) \\ &\leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \cdot \left(\frac{1+\epsilon}{1-\epsilon}\right)^{d+1} \end{aligned}$$

where the first inequality comes from the multiplicative property of Rényi divergence along with the inequality in (13) and the second comes from the weak triangle inequality along with (14). □

Corollary 3. *For power-of-two n, let $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$, m be a positive integer and χ be a distribution over $R^\vee$ satisfying*

$$\mathrm{Pr}_{s \leftarrow\$ \chi}\left[\|\sigma_H(s)\| > B_1\right] \leq \delta_1 \text{ and } \mathrm{Pr}_{s \leftarrow\$ \chi}\left[\max_j \frac{1}{|\sigma_j(s)|} \geq B_2\right] \leq \delta_2$$

for some (B_1, δ_1) and (B_2, δ_2). Also let $\alpha > 0$. For any $k > 1$ that divides $d > 1$ and

$$r \geq \left(\frac{\max\{\sqrt{n}, B_1 B_2\}}{q}\right) \cdot \sqrt{2\ln(2nd(1 + m(d+3)))/\pi},$$

there exists an efficient reduction from search $MLWE^{(R^d)}_{m,q,\Psi_{\leq\alpha}}(\chi^d)$ *to* search $MLWE^{(R^{d/k})}_{m,q^k,\Psi_{\leq\alpha'}}(U(R_q^\vee))$ *for*

$$(\alpha')^2 \geq \alpha^2 + 2r^2 B_1^2 d.$$

In particular, if there is an algorithm solving search $MLWE^{(R^{d/k})}_{m,q^k,\Psi_{\leq\alpha'}}(U(R_q^\vee))$ *with success probability p, then for* search $MLWE^{(R^d)}_{m,q,\Psi_{\leq\alpha}}(\chi^d)$ *an algorithm exists which succeeds with probability at least $[1 - (\delta_1 + \delta_2)]^d \cdot p/8$.*

Proof. We use the reduction and analysis from Theorem 2 with $\beta^2 \geq B_1^2 d$ and $\mathbf{G} = I_{d/k} \otimes (1, q, \ldots, q^{k-1}) \in R^{d/k \times d}$ followed by a standard re-randomization of the resulting secret. Since we sample d such ring elements, we are in the realm of Theorem 2 with probability at least $(1 - (\delta_1 + \delta_2))^d$. Since we have m samples, we must raise the Rényi divergence in Theorem 2 to the m^{th} power. Taking $\epsilon = \frac{1}{m(d+3)}$ ensures that $\left(\frac{1+\epsilon}{1-\epsilon}\right)^{(d+3)m} \leq 8$. The result now follows from the probability preservation property of the Rényi divergence and the fact that we can reverse the mapping between secrets. □

The results of this section are far more satisfying than the analysis given in the previous section when analysing a secret distribution of the form $D_{R^\vee,\tau}$. Let us assume that the probability of success p of an algorithm for solving RLWE is non-negligible. Then all we require is that $\delta_1, \delta_2 = O(1/d)$ in order to solve the search MLWE with non-negligible success probability. Therefore, we may take $B_1 = \tilde{\mathcal{O}}(\tau\sqrt{n})$ and $B_2 = \mathcal{O}(dn/\tau)$ for this secret distribution as long as $\tau \geq \tilde{\mathcal{O}}(d\sqrt{n})$. In this case, we have $\alpha' = \tilde{\mathcal{O}}(\tau n^2\sqrt{d}/q)$. This simplifies to $\alpha' = \tilde{\mathcal{O}}(\alpha n^2\sqrt{d})$ when considering the normal form of MLWE where $\tau = \alpha q$. Therefore, we see that even for typical error and secret distributions with polynomial standard deviations, search MLWE is not qualitatively harder than search RLWE, i.e. an efficient algorithm for the latter implies an efficient algorithm for the former.

4.1 Strictly Spherical Error Distributions

We will now present a lemma that allows us to reduce from MLWE to RLWE with a spherical error distribution.

Lemma 9. *For integers m, n, let $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ be a matrix with non-zero singular values σ_i for $i = 1, \ldots, n$ and take $\beta^2 \geq \sigma_1^2$. Then*

- $R_2\left(D_{r\beta} || D_{r(\beta^2\mathbb{I}+\boldsymbol{M}^T\boldsymbol{M})^{1/2}}\right) \leq \left(1 + \frac{\sigma_1^4}{\beta^4}\right)^{n/2}$,
- $R_\infty\left(D_{r\beta} || D_{r(\beta^2\mathbb{I}+\boldsymbol{M}^T\boldsymbol{M})^{1/2}}\right) \leq \left(1 + \frac{\sigma_1^2}{\beta^2}\right)^{n/2}$.

Proof. To prove this lemma, simply work in the orthogonal basis where the matrix $\mathbf{M}^T\mathbf{M}$ takes a diagonal form. For the first claim,

$$\begin{aligned}
&R_2\left(D_{r\beta} || D_{r(\beta^2\mathbb{I}+\mathbf{M}^T\mathbf{M})^{1/2}}\right) \\
&= \prod_{i=1}^{n} \frac{\sqrt{r^2(\beta^2+\sigma_i^2)}}{r^2\beta^2} \int_{\mathbb{R}} \exp\left[-\pi x_i^2 \left(\frac{2}{r^2\beta^2} - \frac{1}{r^2(\beta^2+\sigma_i^2)}\right)\right] dx_i \\
&= \prod_{i=1}^{n} \sqrt{\frac{\beta^2+\sigma_i^2}{r^2\beta^4}} \int_{\mathbb{R}} \exp\left[-\pi x_i^2 \left(\frac{\beta^2+2\sigma_i^2}{r^2\beta^2(\beta^2+\sigma_i^2)}\right)\right] dx_i \\
&= \prod_{i=1}^{n} \sqrt{\frac{\beta^2+\sigma_i^2}{r^2\beta^4}} \cdot \sqrt{\frac{r^2\beta^2(\beta^2+\sigma_i^2)}{\beta^2+2\sigma_i^2}} = \prod_{i=1}^{n} \sqrt{\frac{\left(\beta^2+\sigma_i^2\right)^2}{\beta^4+2\beta^2\sigma_i^2}} \\
&= \prod_{i=1}^{n} \sqrt{1 + \frac{\sigma_i^4}{\beta^4+2\beta^2\sigma_i^2}} \leq \left(1+\frac{\sigma_1^4}{\beta^4}\right)^{n/2}.
\end{aligned}$$

For the second claim, we have

$$\begin{aligned}
&R_\infty\left(D_{r\beta} || D_{r(\beta^2\mathbb{I}+\mathbf{M}^T\mathbf{M})^{1/2}}\right) \\
&= \max_{\mathbf{x}\in\mathbb{R}^n}\left(\prod_{i=1}^{n} \sqrt{\frac{\beta^2+\sigma_i^2}{\beta^2}} \cdot \exp\left[-\pi x_i^2\left(\frac{\sigma^2}{r^2\beta^2(\beta^2+\sigma_i^2)}\right)\right]\right) \\
&= \prod_{i=1}^{n} \sqrt{\frac{\beta^2+\sigma_i^2}{\beta^2}} \leq \left(1+\frac{\sigma_1^2}{\beta^2}\right)^{n/2}.
\end{aligned}$$

□

We can now extend Theorem 2 to get a spherical output error distribution by applying the above Lemma to the final result along with the triangle inequality. In particular, the Rényi divergences given in Theorem 2 increase by factors of $\left(1 + \frac{d^4 \max_{i,j}|\sigma_j(s_i)|^4}{\beta^4}\right)^{n/2}$ and $\left(1 + \frac{d^2 \max_{i,j}|\sigma_j(s_i)|^2}{\beta^2}\right)^{n/2}$ for orders 2 and ∞ respectively. Therefore, when applying the theorem to m MLWE samples, we require that β increase by factors of $(mn)^{1/4}$ for order 2 and $(mn)^{1/2}$ for infinite order to ensure $\mathcal{O}(1)$ Rényi divergences. These ideas will be concretised in the proof of Theorem 3 in the next section.

5 Reducing RLWE in (n, q) to $(n/2, q^2)$

Throughout this entire section, we assume that n is a power of two. The reduction strategy is to represent polynomial multiplications in ring dimension n using $n \times n$ matrices by working in the *coefficient* embedding. The reduction follows the same blueprint as in Section 3 apart from the fact that we are no longer working exclusively in the canonical embedding. Since we are considering power-of-two cyclotomic rings, polynomial multiplication is always represented by a matrix of the form given in Equation (1). Going from ring dimension n to $n/2$ just halves the dimension of these matrices. For clarity, we adopt the notation $R_{n,q} = \mathbb{Z}_q[X]/\langle X^n + 1\rangle$ and $R_n = \mathbb{Z}[X]/\langle X^n + 1\rangle$.

Our aim is to reduce RLWE in dimension and modulus (n, q) to RLWE in $(n/2, q^2)$ via some mapping:

- $a \in R_{n,q} \longmapsto \tilde{a} \in R_{n/2,q^2}$
- $b \in \mathbb{T}_{R_n{}^\vee} \longmapsto \tilde{b} \in \mathbb{T}_{R_{n/2}{}^\vee}$
- $s \in R_{n,q}{}^\vee \longmapsto \tilde{s} \in R_{n/2,q^2}{}^\vee$

We can start by defining a relationship between $\mathrm{rot}(s)$ and $\mathrm{rot}(\tilde{s})$. In order to make clear the distinction between the two rings, we denote $n \times n$ matrices associated with multiplications in $R_{n,q}$ by writing the subscript n, q. Given $\mathbf{G}$ and $\mathbf{H} \in \mathbb{Z}^{n/2\times n}$, the linear relationship will be defined via the equation

$$\mathrm{rot}(\tilde{s})_{n/2,q^2} = 2 \cdot \mathbf{H} \cdot \mathrm{rot}(s)_{n,q} \cdot \mathbf{G}^T. \tag{15}$$

Note that the factor of 2 is present to account for the fact that the new secret should be in the dual ring $R_{n/2,q^2}{}^\vee = \frac{2}{n}R$ and the matrix $\mathbf{H}$ ensures that we end up with a square matrix $\mathrm{rot}(\tilde{s})_{n/2,q^2}$. We also need to be careful that $\mathbf{G}$ and $\mathbf{H}$ are chosen so the matrix $\mathrm{rot}(\tilde{s})_{n/2,q^2}$ has the correct form. Define the map between b and $\tilde{b}$ (up to some Gaussian error) as

$$\tilde{b}_{vec} \approx 2\mathbf{H} \cdot b_{vec}. \tag{16}$$

In order for the reduction to work, we require that $\tilde{b} \approx \tilde{a} \cdot \tilde{s}/q^2 \bmod R_{n/2}{}^\vee$ i.e.

$$2 \cdot \mathbf{H} \cdot \mathrm{rot}(s)_{n,q} \cdot \frac{1}{q} a_{vec} \approx 2 \cdot \mathbf{H} \cdot \mathrm{rot}(s)_{n,q} \cdot \mathbf{G}^T \cdot \frac{1}{q^2} \tilde{a}_{vec} \bmod 2/n. \tag{17}$$

It is easy to see that we can satisfy this requirement by choosing $\tilde{a}$ such that

$$\frac{1}{q} a_{vec}^T = \frac{1}{q^2} \mathbf{G}^T \cdot \tilde{a}_{vec}^T \bmod 1. \tag{18}$$

Explicit forms for our choice of $\mathbf{G}$ and $\mathbf{H}$ are

$$\mathbf{G} = \mathbb{I}_{n/2} \otimes (1, q) \in \mathbb{Z}^{n/2\times n} \tag{19}$$

$$\mathbf{H} = \mathbb{I}_{n/2} \otimes (1, 0) \in \mathbb{Z}^{n/2\times n} \tag{20}$$

Claim. Take $\mathbf{G}$ and $\mathbf{H}$ as above. Then $\mathrm{rot}(\tilde{s})_{n/2,q^2}$ is of the correct form (i.e. represents multiplication by some polynomial in $(R_{n/2,q^2})$).

Proof. We can write simple explicit forms $(\mathbf{G}^T)_{kl} = \delta_{k,2l-1} + q\delta_{k,2l}$ and $(\mathbf{H})_{ij} = \delta_{2i-1,j}$.Then the matrix multiplication $\mathbf{H} \cdot \mathrm{rot}(s)_{n,q} \cdot \mathbf{G}^T$ yields $(\mathrm{rot}(\tilde{s})_{n/2,q^2})_{il} = (\mathrm{rot}(s)_{n,q})_{2i-1,2l-1} + (q\mathrm{rot}(s)_{n,q})_{2i-1,2l}$ which is of the correct form. □

Note that the mapping between secrets is

$$s = \sum_{i=0}^{n-1} s_i \cdot X^i \longmapsto \tilde{s} = (s_0 - qs_{n-1}) + \sum_{i=1}^{n/2-1} (s_{2i} + qs_{2i-1}) \cdot X^i. \tag{21}$$

Now the proof of correctness for this reduction is essentially the same as Theorem 2 with a few alterations. One of the more important changes is that we use Lemma 9 and target a spherical error. We do this to ensure that multiplication by $\mathbf{H}$ leads to a Gaussian with parameters that we can easily bound.

Theorem 3. *Let n be a power of two, q be an integer, fix $s \in {R_{n,q}}^\vee$ and*

$$r \geq \frac{1}{q} \cdot \max\left\{1, \frac{\|s_{vec}\|}{\sigma_n(\mathrm{rot}(s))}\right\} \cdot \sqrt{2\ln(2n(1+1/\epsilon))/\pi}.$$

Further, take $\boldsymbol{G} \in \mathbb{Z}^{n/2 \times n}$ and $\boldsymbol{H} \in \mathbb{Z}^{n \times n/2}$ as in Equations (19) and (20) respectively. Let $\sigma_1 \coloneqq \sigma_1(\mathrm{rot}(s))$ and $\beta \geq 2\sigma_1\sqrt{n}$.

For any $\alpha > 0$, there exists an efficient mapping $\mathcal{F} : R_{n,q} \times \mathbb{T}_{{R_{n,q}}^\vee} \to R_{n/2,q} \times \mathbb{T}_{{R_{n/2,q^2}}^\vee}$ such that

- $R_2\left(A^{R_{n/2}}_{q^2,\tilde{s},D_{\alpha'}} || \mathcal{F}(A^{R_n}_{q,s,D_\alpha})\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^4 \cdot \left(1 + \frac{16n^2\sigma_1^4}{\beta^4}\right)^{n/2}$,
- $R_\infty\left(A^{R_{n/2}}_{q^2,\tilde{s},D_{\alpha'}} || \mathcal{F}(A^{R_n}_{q,s,D_\alpha})\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^4 \cdot \left(1 + \frac{4n\sigma_1^2}{\beta^2}\right)^{n/2}$

where $\tilde{s}$ is given in Equation 21 and $(\alpha')^2 = 4\alpha^2 + r^2\beta^2$.

Proof. Suppose we are given $(a, b) \in R_{n,q} \times \mathbb{T}_{{R_{n,q}}^\vee}$. The mapping $\mathcal{F}$ is performed as follows:

1. Sample $\mathbf{f} \leftarrow D_{\Lambda - \frac{1}{q}a_{vec},r}$ over the lattice $\Lambda = \frac{1}{q^2}\mathbf{G}^T\mathbb{Z}^{n/2} + \mathbb{Z}^n$. Note that the parameter r is large enough so we can sample the discrete Gaussian efficiently by Lemma 2 since $\|\tilde{\mathbf{B}}_\Lambda\| = q^{-1}$.
2. Let $\mathbf{v} = \frac{1}{q}a_{vec} + \mathbf{f} \in \Lambda/\mathbb{Z}^n$ and set $\mathbf{x}$ to be a random solution of $\frac{1}{q^2}\mathbf{G}^T\mathbf{x} = \mathbf{v} \bmod 1$. Then set $\tilde{a} \in R_{n/2,q^2}$ to be the unique polynomial such that $\tilde{a}_{vec} = \mathbf{x}$.
3. Sample $\tilde{e}$ from the distribution $D_{r\beta}$ over $K_\mathbb{R} \simeq H \simeq \mathbb{R}^{n/2}$ and set $\tilde{b} = 2\mathbf{H} \cdot b + \tilde{e} \in \mathbb{T}_{{R_{n/2,q^2}}^\vee}$.
4. Finally, output $(\tilde{a}, \tilde{b}) \in (R_{n/2,q^2}) \times \mathbb{T}_{{R_{n/2,q^2}}^\vee}$.

Distribution of $\tilde{a}$: We can precisely repeat the argument given in the proof of Theorem 2 after noting that $r \geq \eta_\epsilon(\Lambda)$ and $r \geq \eta_\epsilon(q^{-1}\mathbb{Z}^n)$. The only conceptual difference is that we are now working in the coefficient embedding. Denoting the distribution of $\tilde{a}$ given that the distribution of a is uniform by $D^{(\tilde{a})}$, we find that

$$R_\infty\left(U(R_{n/2,q^2})||D^{(\tilde{a})}\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2. \tag{22}$$

Distribution of the error: We now condition on fixed $\tilde{a} = \bar{\tilde{a}}$ and set $\bar{\mathbf{v}} = \mathbf{G}^T\bar{\tilde{a}}_{vec}$. Denoting the distribution of $-\mathbf{f}$ as $D^{(-\mathbf{f})}$ we also have that

$$R_\infty\left(D_{\frac{1}{q}\mathbb{Z}^n-\bar{\mathbf{v}},r}||D^{(-\mathbf{f})}\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right).$$

All that remains is to analyse the distribution of

$$\begin{aligned}\left(\tilde{b} - \frac{1}{q^2}\tilde{a}\cdot\tilde{s}\right)_{vec} &= 2\mathbf{H}\cdot\mathrm{rot}(s)\cdot(-\mathbf{f}) + 2\mathbf{H}\cdot e_{vec} + \tilde{e}_{vec} \bmod 2/n \qquad (23)\\ &= 2\mathbf{H}\cdot(\mathrm{rot}(s)\cdot(-\mathbf{f}) + e_{vec} + \tilde{e}^*_{vec}) \bmod 2/n \qquad (24)\end{aligned}$$

where $\tilde{e}^*_{vec}$ (resp. e_{vec}) is drawn from the spherical distribution $D_{r\beta/(2\sqrt{n})}$ (resp. $D_{\alpha/\sqrt{n}}$). Note that the $\sqrt{n}$ factors take into account that we are working in the coefficient embedding.
The distribution of $\mathrm{rot}(s)\cdot D_{\frac{1}{q}\mathbb{Z}^n-\bar{\mathbf{v}},r}$ is $D_{\frac{1}{q}\mathrm{rot}(s)\mathbb{Z}^n-\mathrm{rot}(s)\bar{\mathbf{v}},r\cdot\mathrm{rot}(s)^T}$. By working in the orthogonal basis where the covariance matrix $\mathrm{rot}(s)^T\mathrm{rot}(s)$ is diagonal, we can apply Lemma 5. We also apply the data-processing inequality on $(-\mathbf{f},\tilde{e}^*_{vec}) \longmapsto -\mathrm{rot}(s)\cdot\mathbf{f} + \tilde{e}^*_{vec}$ along with the triangle inequality to obtain

$$R_\infty\left(D_{err}||D^{(-\mathrm{rot}(s)\cdot\mathbf{f}+\tilde{e}^*_{vec})}\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)\cdot\left(\frac{1+\epsilon}{1-\epsilon}\right), \tag{25}$$

where D_{err} is a continuous Gaussian distribution with covariance $\Sigma = r^2(\frac{\beta^2}{4n}\mathbb{I} + \mathrm{rot}(s)^T\mathrm{rot}(s))$ and $D^{(-\mathrm{rot}(s)\cdot\mathbf{f}+\tilde{e}^*_{vec})}$ is the *exact* distribution of $-\mathrm{rot}(s)\cdot\mathbf{f}+\tilde{e}^*_{vec}$.

Distance to spherical error: We now apply Lemma 9 to find that

$$R_2\left(D_{r\beta/(2\sqrt{n})}||D_{err}\right) \leq \left(1+\frac{16n^2\sigma_1^4}{\beta^4}\right)^{n/2},$$

$$R_\infty\left(D_{r\beta/(2\sqrt{n})}||D_{err}\right) \leq \left(1+\frac{4n\sigma_1^2}{\beta^2}\right)^{n/2}.$$

Finally, using the weak triangle inequality with intermediate distribution $2\mathbf{H}\cdot D_{err}$ and the data-processing inequality, we obtain

$$R_2\left(2\mathbf{H}\cdot D_{((r\beta)^2/(4n)+\alpha^2/n)^{1/2}}||D^{(RHS)}\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2\cdot\left(1+\frac{16n^2\sigma_1^4}{\beta^4}\right)^{n/2},$$

$$R_\infty\left(2\mathbf{H}\cdot D_{((r\beta)^2/(4n)+\alpha^2/n)^{1/2}}||D^{(RHS)}\right) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2\cdot\left(1+\frac{4n\sigma_1^2}{\beta^2}\right)^{n/2}$$

where $D^{(RHS)}$ is the distribution of Equation (23).

Distribution of the reduction output: We conclude by combining the above results in the same way as in the proof of Theorem 2. We must also scale up by a factor of $\sqrt{n}$ to account for the fact that we have been working in the coefficient embedding. □

Corollary 4. *Let n be a power of two and χ be a distribution over $R_n^\vee$ satisfying*

$$\Pr_{s \hookleftarrow\$ \chi}[\|\sigma_H(s)\| > B_1] \leq \delta_1 \text{ and } \Pr_{s \hookleftarrow\$ \chi}\left[\max_j \frac{1}{|\sigma_j(s)|} \geq B_2\right] \leq \delta_2$$

for some (B_1, δ_1) and (B_2, δ_2). Also, let $\alpha > 0$ and $\epsilon \in (0, 1/2)$. For any

$$r \geq \frac{1}{q} \cdot \max\{1, B_1 B_2\} \cdot \sqrt{2\ln(2n(4m+1))/\pi},$$

let $(\alpha'_c)^2 = 4\alpha^2 + 4r^2 B_1^2 (mn)^{2c}$. Suppose there exists an algorithm solving search $RLWE^{(R_{n/2})}_{m,q^2,D_{\alpha'_c}}(U(R^\vee_{n/2,q^2}))$ for $c = 1/4$ (resp. $c = 1/2$) with success probability $p_{1/4}$ (resp. $p_{1/2}$). Then there exists algorithms solving $RLWE^{(R_n)}_{m,q,D_\alpha}(\chi)$ with success probabilities at least $(1 - (\delta_1 + \delta_2))\frac{p_{1/4}^2}{8e^{1/2}}$ and $(1 - (\delta_1 + \delta_2))\frac{p_{1/2}}{8e^{1/2}}$.

Proof. We will be applying the reduction in Theorem 3 with $\epsilon = 1/(4m)$ along with a re-randomizing of the secret. We take $\beta = 2B_1(mn)^c$ in the theorem. Recall that for power-of-two cyclotomic rings, we have $\|\sigma_H(s)\| = \sqrt{n}\|s_{vec}\|$, $\min_j |\sigma_j(s)| = \sigma_n(\mathrm{rot}(s))$ and $\max_j |\sigma_j(s)| = \sigma_1(\mathrm{rot}(s))$. This means that we are able to apply the reduction and analysis of Theorem 3 with probability at least $1 - (\delta_1 + \delta_2)$. Since we have m samples, we need to raise the Rényi divergences to the m^{th} power. Therefore, in the case that $c = 1/4$ (resp. $c = 1/2$), we have that the Rényi divergence of order 2 (resp. order ∞) is upper bounded by $8 \cdot e^{1/2}$. Note that the reduction defines a reversible map between the secrets. Therefore, the result is obtained by running the reduction, re-randomizing the secret, solving the resulting search RLWE instance and then mapping back to the original secret. □

Typically, we would have access to $m = \mathcal{O}(1)$ RLWE samples. Considering the normal form of RLWE with secret distribution $D_{R^\vee,\alpha q}$, we can take the parameters B_1 and B_2 to be $\tilde{\mathcal{O}}(\alpha q \sqrt{n})$ and $\tilde{\mathcal{O}}(n/(\alpha q))$ respectively. Therefore, the above corollary says that if we can solve RLWE in dimension $n/2$, modulus q^2 and error rate $\alpha \cdot n^{9/4}$ with non-negligible probability in polynomial time, then we can also solve RLWE with dimension n, modulus q and error rate α is polynomial time with non-negligible probability.

Acknowledgements

We thank Adeline Roux-Langlois and Léo Ducas for helpful discussions on an earlier draft of this work.

References

ABB10. Shweta Agrawal, Dan Boneh, and Xavier Boyen. Efficient lattice (H)IBE in the standard model. In Gilbert [Gil10], pages 553–572.

ABD16. Martin R. Albrecht, Shi Bai, and Léo Ducas. A subfield lattice attack on overstretched NTRU assumptions - cryptanalysis of some FHE and graded encoding schemes. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology – CRYPTO 2016, Part I*, volume 9814 of *Lecture Notes in Computer Science*, pages 153–178. Springer, Heidelberg, August 2016.

ADPS16. Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Post-quantum key exchange - A new hope. In Thorsten Holz and Stefan Savage, editors, *25th USENIX Security Symposium, USENIX Security 16*, pages 327–343. USENIX Association, 2016.

APS15. Martin R. Albrecht, Rachel Player, and Sam Scott. On the concrete hardness of learning with errors. *Journal of Mathematical Cryptology*, 9(3):169–203, 2015.

Ban93. Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.

BBD$^+$17. Shi Bai, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, and Damien Stehlé. CRYptographic suiTe for Algebraic LatticeS. Contributed talk at Real World Crypto 2017, New York, USA, January 2017.

BCD$^+$16. Joppe W. Bos, Craig Costello, Léo Ducas, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Ananth Raghunathan, and Douglas Stebila. Frodo: Take off the ring! Practical, quantum-secure key exchange from LWE. In Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi, editors, *ACM CCS 16: 23rd Conference on Computer and Communications Security*, pages 1006–1018. ACM Press, October 2016.

BCNS15. Joppe W. Bos, Craig Costello, Michael Naehrig, and Douglas Stebila. Post-quantum key exchange for the TLS protocol from the ring learning with errors problem. In *2015 IEEE Symposium on Security and Privacy*, pages 553–570. IEEE Computer Society Press, May 2015.

BG14. Shi Bai and Steven D. Galbraith. An improved compression technique for signatures based on learning with errors. In Josh Benaloh, editor, *Topics in Cryptology – CT-RSA 2014*, volume 8366 of *Lecture Notes in Computer Science*, pages 28–47. Springer, Heidelberg, February 2014.

BGM$^+$16. Andrej Bogdanov, Siyao Guo, Daniel Masny, Silas Richelson, and Alon Rosen. On the hardness of learning with rounding over small modulus. In Eyal Kushilevitz and Tal Malkin, editors, *TCC 2016-A: 13th Theory of Cryptography Conference, Part I*, volume 9562 of *Lecture Notes in Computer Science*, pages 209–224. Springer, Heidelberg, January 2016.

BGV12. Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In Shafi Goldwasser, editor, *ITCS 2012: 3rd Innovations in Theoretical Computer Science*, pages 309–325. Association for Computing Machinery, January 2012.

BLL$^+$15. Shi Bai, Adeline Langlois, Tancrède Lepoint, Damien Stehlé, and Ron Steinfeld. Improved security proofs in lattice-based cryptography: Using the Rényi divergence rather than the statistical distance. In Tetsu Iwata and Jung Hee Cheon, editors, *Advances in Cryptology – ASIACRYPT 2015,* *Part I*, volume 9452 of *Lecture Notes in Computer Science*, pages 3–24. Springer, Heidelberg, November / December 2015.

BLP$^+$13. Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th Annual ACM Symposium on Theory of Computing*, pages 575–584. ACM Press, June 2013.

CDPR16. Ronald Cramer, Léo Ducas, Chris Peikert, and Oded Regev. Recovering short generators of principal ideals in cyclotomic rings. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology – EUROCRYPT 2016, Part II*, volume 9666 of *Lecture Notes in Computer Science*, pages 559–585. Springer, Heidelberg, May 2016.

CDW17. Ronald Cramer, Léo Ducas, and Benjamin Wesolowski. Short stickelberger class relations and application to ideal-svp. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology - EUROCRYPT 2017, Proceedings, Part I*, volume 10210 of *Lecture Notes in Computer Science*, pages 324–348, 2017.

CG13. Ran Canetti and Juan A. Garay, editors. *Advances in Cryptology – CRYPTO 2013, Part I*, volume 8042 of *Lecture Notes in Computer Science*. Springer, Heidelberg, August 2013.

CGS14. Peter Campbell, Michael Groves, and Dan Shepherd. Soliloquy: A cautionary tale. In *ETSI 2nd Quantum-Safe Crypto Workshop*, pages 1–9, 2014.

CJL16. Jung Hee Cheon, Jinhyuck Jeong, and Changmin Lee. An algorithm for ntru problems and cryptanalysis of the ggh multilinear map without a low-level encoding of zero. *LMS Journal of Computation and Mathematics*, 19(A):255–266, 2016.

DDLL13. Léo Ducas, Alain Durmus, Tancrède Lepoint, and Vadim Lyubashevsky. Lattice signatures and bimodal Gaussians. In Canetti and Garay [CG13], pages 40–56.

DLP14. Léo Ducas, Vadim Lyubashevsky, and Thomas Prest. Efficient identity-based encryption over NTRU lattices. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology – ASIACRYPT 2014, Part II*, volume 8874 of *Lecture Notes in Computer Science*, pages 22–41. Springer, Heidelberg, December 2014.

Gen09. Craig Gentry. Fully homomorphic encryption using ideal lattices. In Mitzenmacher [Mit09], pages 169–178.

Gen10. Craig Gentry. Toward basing fully homomorphic encryption on worst-case hardness. In Tal Rabin, editor, *Advances in Cryptology – CRYPTO 2010*, volume 6223 of *Lecture Notes in Computer Science*, pages 116–137. Springer, Heidelberg, August 2010.

Gil10. Henri Gilbert, editor. *Advances in Cryptology – EUROCRYPT 2010*, volume 6110 of *Lecture Notes in Computer Science*. Springer, Heidelberg, May 2010.

GPV08. Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In Richard E. Ladner and Cynthia Dwork, editors, *40th Annual ACM Symposium on Theory of Computing*, pages 197–206. ACM Press, May 2008.

GSW13. Craig Gentry, Amit Sahai, and Brent Waters. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In Canetti and Garay [CG13], pages 75–92.

HKM17. Gottfried Herold, Elena Kirshanova, and Alexander May. On the asymptotic complexity of solving lwe. *Designs, Codes and Cryptography*, pages 1–29, 2017.

KF17. Paul Kirchner and Pierre-Alain Fouque. Revisiting lattice attacks on overstretched NTRU parameters. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology - EUROCRYPT 2017, Proceedings, Part I*, volume 10210 of *Lecture Notes in Computer Science*, pages 3–26, 2017.

LLM$^+$16. Benoît Libert, San Ling, Fabrice Mouhartem, Khoa Nguyen, and Huaxiong Wang. Signature schemes with efficient protocols and dynamic group signatures from lattice assumptions. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *Advances in Cryptology – ASIACRYPT 2016, Part II*, volume 10032 of *Lecture Notes in Computer Science*, pages 373–403. Springer, Heidelberg, December 2016.

LMPR08. Vadim Lyubashevsky, Daniele Micciancio, Chris Peikert, and Alon Rosen. SWIFFT: A modest proposal for FFT hashing. In Kaisa Nyberg, editor, *Fast Software Encryption – FSE 2008*, volume 5086 of *Lecture Notes in Computer Science*, pages 54–72. Springer, Heidelberg, February 2008.

LP11. Richard Lindner and Chris Peikert. Better key sizes (and attacks) for LWE-based encryption. In Aggelos Kiayias, editor, *Topics in Cryptology – CT-RSA 2011*, volume 6558 of *Lecture Notes in Computer Science*, pages 319–339. Springer, Heidelberg, February 2011.

LPR10. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Gilbert [Gil10], pages 1–23.

LPR13. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-LWE cryptography. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology – EUROCRYPT 2013*, volume 7881 of *Lecture Notes in Computer Science*, pages 35–54. Springer, Heidelberg, May 2013.

LS15. Adeline Langlois and Damien Stehlé. Worst-case to average-case reductions for module lattices. *Des. Codes & Cryptography*, 75(3):565–599, 2015.

LSS14. Adeline Langlois, Damien Stehlé, and Ron Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology – EUROCRYPT 2014*, volume 8441 of *Lecture Notes in Computer Science*, pages 239–256. Springer, Heidelberg, May 2014.

Mit09. Michael Mitzenmacher, editor. *41st Annual ACM Symposium on Theory of Computing*. ACM Press, May / June 2009.

Pei09. Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In Mitzenmacher [Mit09], pages 333–342.

PRSD17. Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of Ring-LWE for any ring and modulus. In *STOC 2017*, 2017.

Reg05. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th Annual ACM Symposium on Theory of Computing*, pages 84–93. ACM Press, May 2005.

Reg09. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6), 2009.

SS13. Damien Stehlé and Ron Steinfeld. Making NTRUEncrypt and NTRUSign as secure as standard worst-case problems over ideal lattices. Cryptology ePrint Archive, Report 2013/004, 2013. `http://eprint.iacr.org/2013/004`.

vEH14. Tim van Erven and Peter Harremos. Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014.

A Design Space for RLWE Public-Key Encryption

Recall the simple public-key encryption scheme from [LPR10] which serves as the blueprint for many subsequent constructions. The scheme publishes a public-key $(a, b = a \cdot s + e)$, where both s and e are small elements from the ring of integers of a power-of-two cyclotomic field. Encryption of some polynomial m with $\{0,1\}$ coefficients is then performed by sampling short r, e_1, e_2 and outputting:

$$\big(u,\, v\big) = \big(a \cdot r + e_1, \quad b \cdot r + e_2 + \lfloor q/2 \rfloor \cdot m \bmod q\big).$$

The decryption algorithm computes

$$u \cdot s - v = (a \cdot r + e_1) \cdot s - (a \cdot s + e) \cdot r - e_2 - \lfloor q/2 \rfloor \cdot m.$$

Let σ be the norm of s, e, r, e_1, e_2. Clearly, the final message will have noise of norm $\geq \sigma^2$. Thus to ensure correct decryption, q has a quadratic dependency on σ. As a consequence, in this construction, increasing σ and q can only reduce security by increasing the gap between noise and modulus.

However, this issue can be avoided and is avoided in MLWE-based constructions by picking some $\sigma' < \sigma$ at the cost of publishing more RLWE samples in the public key. For example, if $d = 2$ the public key becomes

$$\big((a', b'),\, (a'', b'')\big) = \big((a', a' \cdot s + e'),\, (a'', a'' \cdot s + e'')\big),$$

where $s, e'e,''$ have norm σ. Encryption of some $\{0,1\}$ polynomial m is then performed by sampling short $r', r'', e_1,\, e_2$ with norm σ' and outputting

$$\big(u,\, v\big) = \big(a' \cdot r' + a'' \cdot r'' + e_1, \quad b' \cdot r' + b'' \cdot r'' + e_2 + \lfloor q/2 \rfloor \cdot m \bmod q\big).$$

The decryption algorithm computes

$$u \cdot s - v = (a' \cdot r' + a'' \cdot r'' + e_1) \cdot s - (a' \cdot s + e') \cdot r' - (a'' \cdot s + e'') \cdot r'' - e_2 - \lfloor q/2 \rfloor \cdot m.$$

The security of the public key reduces to the hardness of RLWE in dimension n with modulus q and noise size σ as before. The security of encryptions reduces to the hardness of MLWE in dimension $d = 2$ over ring dimension n, modulus q and noise size σ', i.e. the level of security is maintained for $\sigma' < \sigma$ by increasing the dimension. While we still require $q > \sigma \cdot \sigma'$, the size of σ' can be reduced at the cost of increasing d resp. by relying on RLWE with modulus q^d. Finally, note that we may think of Regev's original encryption scheme [Reg09] as one extreme corner of this design space (for LWE) with $d = 2\, n \log q$, where r', r'' are binary and where $e_1, e_2 = 0, 0$.

B Powerful Ring LWE Adversaries

We show that an adversary that is able to solve search $\mathrm{RLWE}^{(R)}_{m=1,q^d,D_\alpha}$ implies an adversary that can solve search $\mathrm{LWE}_{m=n,d,q,D_\alpha}$ where m denotes the number of samples, n is the ring dimension of R and d is the plain LWE dimension. As usual, q denotes the modulus and D_α the error distribution.

Suppose we start with n LWE samples in dimension d and modulus q. Then the main result in [BLP+13] says that we may transform these to samples of LWE in dimension 1 and modulus q^d while *slightly* increasing the error rate to $\alpha' > \alpha$. We now show how to further transform these into a single RLWE sample. Denote our m LWE samples of dimension 1 as

$$\left(a_i, b_i = \frac{1}{q^d} \cdot a_i s_0 + e_i\right) \in \mathbb{Z}_{q^d} \times \mathbb{T} \text{ for } i = 1, \dots, n \tag{26}$$

where s_0 is the uniform secret obtained having performed the reduction mentioned above.

Now we take the common example of a power-of-two cyclotomic ring $R \simeq \mathbb{Z}[X]/\langle X^n + 1\rangle$ for simplicity. In order to produce a RLWE sample, we choose random $s_1, \dots, s_{n-1} \leftarrow_{\$} \left\{\frac{0}{n}, \frac{1}{n}, \dots, \frac{q^d-1}{n}\right\}^{n-1}$ where the divisor n arises because secrets come from the dual ring. We now define

$$s := \frac{s_0}{n} + s_1 \cdot X + \dots + s_{n-1} \cdot X^{n-1} \in R_q^\vee, \tag{27}$$

$$a := a_0 + a_1 \cdot X + \dots + a_{n-1} \cdot X^{n-1} \in R_q. \tag{28}$$

When doing the multiplication $\frac{1}{q^d} a \cdot s \bmod R^\vee$, the only terms which we do not explicitly know are of the form $a_i \cdot s_0$. In particular, the only unknown term in the coefficient of X^i is $\frac{1}{q^d} a_i \cdot \frac{s_0}{n}$. However, we can simply replace this with $\frac{1}{n} b_i$ from Equation 26 to get an approximation of $\frac{1}{q^d} a \cdot s$. Following this strategy, we end up with a polynomial $\tilde{b}$ such that

$$\tilde{b} - \frac{1}{q^d} a \cdot s = \frac{1}{n}\left(e_0 + e_1 \cdot X + \dots e_{n-1} \cdot X^{n-1}\right) \bmod R^\vee. \tag{29}$$

Therefore, $(a, \tilde{b}) \in R_q \times \mathbb{T}_{R^\vee}$ is an RLWE sample with error distribution $D_{\alpha'}$ over $K_\mathbb{R}$. In the case of a general ring R, the same overall strategy works apart from the fact that the final error distribution may be skewed in the canonical embedding space. However, the error distribution can be made spherical by adding an appropriately skewed Gaussian if desired.