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 factor of $n^{c+1/2} \cdot \sqrt{d}$ for ring dimension n, module rank d and any constant $c > 0$ in the case of 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 power-of-two cyclotomic RLWE that reduces the ring dimension n by a power-of-two factor 2^i, while increasing the modulus by a power of 2^i and the error rate by a factor of $2^{i \cdot (1-c)} \cdot n^{c+1/2}$ for any constant $c > 0$. 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.

Changelog (01/2020): This updated version improves on our original work [AD17] (and the subsequent work of [WW19]) by reducing the error rate expansions incurred. For the normal form power-of-two cyclotomic MLWE to MLWE reduction, the error rate expansion is improved from a factor of $n^2\sqrt{d}$ to a factor of $n^{c+1/2}\sqrt{d}$ (for any constant $c > 0$). This improvement comes from an improved analysis. In addition, the RLWE self-reduction of our original work is generalised and improved by combining the recent results of Peikert and Pepin [PP19] along with our improved MLWE to MLWE reduction analysis. Notably, the RLWE to RLWE reduction in this version improves on previous results by covering *both* decisional and search variants of RLWE. In addition, the RLWE to RLWE dimension-halving, modulus-squaring reduction [AD17] was originally reported to have an error rate expansion factor of $n^{9/4}$ when ignoring small constant factors. However in this updated version, we show that a factor of $n^{c+1/2}$ (for any constant $c > 0$) suffices.

* The research of Albrecht 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 Deo 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)

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 [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 error distribution χ, samples of the LWE *distribution* in normal-form are constructed 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 all components of the secret $\mathbf{s} \in \mathbb{Z}_q^n$ and e are drawn from the distribution χ. Distinguishing the LWE distribution from uniform is known as the decision LWE problem, whereas finding the secret $\mathbf{s}$ is known as the search LWE problem.

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 where $R^\vee$ denotes the dual of the ring R. A complete and more precise definition is given in Section 2.3. Similar to the case of plain LWE, the decision problem is to distinguish the RLWE distribution from uniform 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, PRS17]. 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 find such vectors in general lattices efficiently. 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, PRS17] mentioned above have no dependency on q apart from the requirement 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 . Similarly, the approximation factors defining the lattice problems are also independent of q.

This interpretation of known hardness results 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.

Indeed, for LWE it is well-known [BLP$^+$13] that we can trade the size of the dimension n and the modulus q without affecting security, as long as $n \log q$ remains constant. Furthermore, in the case of plain LWE we can choose n freely, reducing our dependence on large q to increase security. However, in the case of RLWE the analogue reduction to [BLP$^+$13] is not known and the choice of ring R — and hence the dimension n — can lead to practical implementation advantages and a simpler interpretation of formally defined RLWE. Typically, a power-of-two cyclotomic ring is used, i.e. a ring isomorphic to $\mathbb{Z}[X] / \langle X^n + 1 \rangle$ 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], "powers of two are sparsely distributed, and the desired concrete security level for an application may call for a ring dimension much smaller than the next-largest power of two. So restricting to powers of two could lead to key sizes and runtimes that are at least twice as large as necessary." 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 as 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 [BDK$^+$17, DLL$^+$17], where MLWE is used to build both key encapsulation 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. These results are stated in terms of any ring R. In Section 3.4, we discuss an instantiation of our main results focusing on power-of-two cyclotomic rings. Informally, applying Corollary 1 for *power-of-two cyclotomic rings* and *normal-form* secret distributions (see Section 3.4) yields the following result:

Corollary. *There exists an efficient reduction from search* MLWE *in modulus q, rank d and error rate α to search* RLWE *in modulus q^d and error rate $\alpha \cdot n^{1/2+c} \cdot \sqrt{d}$ for any constant $c > 0$.*

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, Corollaries 1 and 2 show 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 $n^c \cdot d^{1/2}$ in error rate expansion when comparing our results to those in [BLP$^+$13]. In fact, we save a factor of $n^{1/2-c}$ as our reduction considers *infinity norms* of Gaussian secrets in the canonical embedding but lose a factor of $n^{1/2}$ because of the fact that the analysis of the errors takes place in a power-of-two cyclotomic ring interpreted as a lattice in canonical space rather than $\mathbb{Z}^n$. 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). 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 $n' := n/2^i$ and modulus q^{2^i} by combining a reduction from the work of Peikert and Pepin [PP19] and our MLWE to RLWE reduction. Here, the error rate expands by a factor of $\frac{n^{3/2}}{(n')^{1-c}}$ for any constant $c > 0$ if we have access to $\mathcal{O}(1)$ samples and wish to preserve a non-negligible success probability.

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

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 $\gamma = \tilde{\mathcal{O}}(n^{c+1} \cdot d^{3/2}/\alpha)$ for any constant $c > 0$. As there are also converse reductions from RLWE to

Module-SIVP e.g. the dual attack [MR09] which requires finding short vectors in a module lattice, these observations imply that RLWE is equivalent to Module-SIVP. Previous hardness results only stated that RLWE is at least as hard as Ideal-SIVP [LPR10].[1] 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, such an algorithm would essentially solve RLWE over any power-of-two cyclotomic field by our reduction in Section 5. Furthermore, as already mentioned in [BLP$^+$13], 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 (we give more details in Appendix B). While such an adversary cannot be ruled out in principle, it cannot be enabled by the algebraic structure of RLWE or ideal lattices. However, we note that 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] and the analogous result for LWE in [BLP$^+$13]. Indeed, our output RLWE instance in modulus q^d has noise of size at least $q^{d/2}$. Thus, 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. This augments standard reductions from RLWE resp. MLWE to Ideal-SIVP resp. Module-SIVP [Reg05, LPR10, LS15] which do not by themselves suggest that the problem becomes harder with increasing q.

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 columns in $\mathbf{B}$ (from left to right) and $\|\mathbf{B}\|$ is the length of the longest vector (in Euclidean norm) of the columns of $\mathbf{B}$. Additionally, for any $\mathbf{x} \in \mathbb{R}^n$, we write $\|\mathbf{x}\|$ to denote the standard Euclidean norm of $\mathbf{x}$. The dual of a lattice Λ is defined as $\Lambda^* = \{\mathbf{x} \in \operatorname{span}(\Lambda) : \forall\ \mathbf{y} \in \Lambda, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}$ where $\langle \cdot, \cdot \rangle$ is an inner product. We denote the identity matrix in n dimensions using $\mathbf{I}_n$. In addition to the conjugate

[1] Except for RLWE instances with modulus q^n which are known to be as hard as LWE in dimension n and modulus q [BLP$^+$13].

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 $\mathbf{s} = (s_0, \dots, s_{d-1}) \leftarrow D^d$ denote the act of sampling each component s_i according to D independently. We also write $\mathrm{Supp}(D)$ to mean the support of the distribution D. Note that we use standard big-$\mathcal{O}$ notation where $\tilde{\mathcal{O}}$ hides logarithmic factors.

Let K denote some algebraic number field. The degree of K is equal to the dimension of K as a vector space over $\mathbb{Q}$. The trace of a field element $x \in K$, denoted by $\mathrm{Tr}_{K/\mathbb{Q}}(x)$, is defined to be the trace of the linear map (acting on K when viewed as a vector space over $\mathbb{Q}$) corresponding to multiplication by x. Generalising this notion, if K'/K is a field extension, $\mathrm{Tr}_{K'/K}(x')$ for any $x' \in K'$ is defined to be the trace of the linear map (acting on K' when viewed as a vector space over K) corresponding to multiplication by x'. An element $x \in K$ is said to be integral if it is the root of a monic polynomial with integer coefficients. The set of all integral elements forms the ring of integers of K denoted by $\mathcal{O}_K$. An order $\mathcal{O}$ of an algebraic number field K of degree n is a subring containing 1 that is also a rank n $\mathbb{Z}$-module. The simplest example is the ring of integers $\mathcal{O}_K$ which corresponds to the *maximal* order i.e. for any order of K, $\mathcal{O}$, we have that $\mathcal{O} \subseteq \mathcal{O}_K$. Suppose that $\zeta \in \mathbb{C}$ is an algebraic number (i.e. the root of some polynomial with rational coefficients) and let $f(X)$ denote the minimal polynomial of ζ with coefficients in $\mathbb{Q}$. The field extension $\mathbb{Q}(\zeta) \supset \mathbb{Q}$ is isomorphic to $\mathbb{Q}(X)/\langle f(X) \rangle$, so we can view field elements in $\mathbb{Q}(\zeta)$ as polynomials with degree at most $\deg(f) - 1$, and consider operations as polynomial multiplication/addition modulo the polynomial $f(X)$. 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 $s_{vec} = (s_0, \dots, s_{n-1})^T$.

We can also represent multiplication by $s \in K$ in this coefficient embedding using matrices. The appropriate matrix will be denoted by $\mathrm{rot}(s) \in \mathbb{R}^{n \times n}$. In particular, for $r, s, t \in K$ with $r = st$, we have that $r_{vec} = \mathrm{rot}(s) \cdot t_{vec}$. Note that the matrix $\mathrm{rot}(s)$ must be invertible with inverse $\mathrm{rot}(s^{-1})$ for $s \neq 0$. The explicit form of $\mathrm{rot}(s)$ depends on the particular field K. In the case where K is a cyclotomic power-of-two field, i.e. $K = \mathbb{Q}[X]/\langle X^n + 1 \rangle$ for power-of-two n, we have

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

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 $\sigma_1, \dots, \sigma_{r_1}$ be the real embeddings and $\sigma_{r_1+1}, \dots, \sigma_{r_1+2r_2}$ be complex. The complex embeddings come in conjugate pairs, so we have $\sigma_i = \overline{\sigma_{i+r_2}}$ for $i = r_1 + 1, \dots, r_1 + r_2$ if we use an appropriate ordering of the embeddings. Define

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

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

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

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

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

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 $\sigma_H(x) \in \mathbb{R}^n$ through the change of basis described above. So for any $x \in K$, $\sigma_H(x) = \mathbf{U}_H^\dagger \cdot \sigma_C(x)$ where the unitary matrix is given by

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

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 $\mathbf{X}_H = \mathbf{U}_H^\dagger \mathbf{X} \mathbf{U}_H \in \mathbb{R}^{n \times n}$ which is not diagonal in general. However, for any $\mathbf{X}_H$, we have $\mathbf{X}_H \cdot \mathbf{X}_H^T = \mathbf{X}_H \cdot \mathbf{X}_H^\dagger = \mathbf{U}_H^\dagger \mathbf{X}\mathbf{X}^\dagger \mathbf{U}_H$. Explicitly, $(\mathbf{X}_H \cdot \mathbf{X}_H^T)_{ij} = |\sigma_i(x)|^2 \delta_{ij}$ i.e. $\mathbf{X}_H \cdot \mathbf{X}_H^T$ is a diagonal matrix. Likewise for $\mathbf{X}_H^T \cdot \mathbf{X}_H$.

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 with respect to the basis $(\mathbf{h}_i)_{i=1}^n$ 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*. We often refer to a ring R as a lattice. Whenever we do this, we are really referring to the lattice $\sigma_H(R)$.

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 a distribution D over $K_\mathbb{R}$ is done by sampling D 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 $\mathsf{RLWE}_{m,q,\Psi}^{(R)}(D)$ entails distinguishing m samples of $U(R_q \times \mathbb{T}_{R^\vee})$ from $A_{q,s,\psi}^{(R)}$ where $s \leftarrow D$ and ψ is an arbitrary distribution in Ψ. The* search *variant s-$\mathsf{RLWE}_{m,q,\Psi}^{(R)}(D)$ 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 $A_{d,q,\boldsymbol{s},\psi}^{(M)}$ 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 $\mathsf{MLWE}_{m,q,\Psi}^{(M)}(D)$ entails distinguishing m samples of $U((R_q)^d \times \mathbb{T}_{R^\vee})$ from $A_{q,\boldsymbol{s},\psi}^{(M)}$ where $\boldsymbol{s} \leftarrow D^d$ and ψ is an arbitrary distribution in Ψ. The* search *variant s-$\mathsf{MLWE}_{m,q,\Psi}^{(M)}(D)$ 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 1. *If P and Q are two probability distributions such that $P(i) \geq (1-\varepsilon)Q(i)$ for all i, then $\Delta(P,Q) \leq \varepsilon$.*

We will also make use of the Rényi divergence as an alternative to the statistical distance to measure 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*

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

For the case where P and Q are continuous distributions, we replace the sums by integrals and let $P(x)$ and $Q(x)$ denote probability 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 $a \mapsto R_a\left(P||Q\right)$ is non-decreasing, continuous and tends to $R_\infty\left(P||Q\right)$ as $a \to \infty$.*
- ***Log Positivity:*** *$R_a\left(P||Q\right) \geq R_a\left(P||P\right) = 1$.*
- ***Data Processing Inequality:*** *$R_a\left(P^f||Q^f\right) \leq R_a\left(P||Q\right)$ 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*
  - *$R_a\left(P||Q\right) = R_a\left(P_1||Q_1\right) \cdot R_a\left(P_2||Q_2\right)$.*
  - *$R_a\left(P||Q\right) = R_\infty\left(P_1||Q_1\right) \cdot \max_{y_1 \in Supp(P_1)} R_a\left(P_{2|1}(\cdot|y_1)||Q_{2|1}(\cdot|y_1)\right)$.*
- ***Probability Preservation:*** *Let $E \subseteq Supp(Q)$ be an arbitrary event. If $a \in (1,\infty)$, then $Q(E) \geq P(E)^{\frac{a}{a-1}}/R_a\left(P||Q\right)$. Furthermore, we have $Q(E) \geq P(E)/R_\infty\left(P||Q\right)$.*
- ***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*

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

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

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

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

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

Notice that the matrix $\boldsymbol{\Sigma}$ differs from the standard covariance matrix by a factor of 2π. However, for convenience, we refer to $\boldsymbol{\Sigma}$ as the covariance matrix throughout. Note that if the centre $\mathbf{c}$ is omitted, it should be assumed that $\mathbf{c} = \mathbf{0}$. If the covariance matrix is diagonal, we describe it using the vector of its diagonal entries. For example, suppose that $\left(\mathbf{S}^T\mathbf{S}\right)_{ij} = (s_i)^2\delta_{ij}$ and let $\mathbf{s} = (s_1, \ldots s_n)^T$. Then we would write $D_{\mathbf{s}}$ to denote the centred Gaussian distribution $D_{\mathbf{S}}$.

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 (typically lattices 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 $\rho_r(\Lambda+\mathbf{u}) = \sum_{\mathbf{x}\in\Lambda+\mathbf{u}}\rho_r(\mathbf{x})$. For a ring R contained in a number field K and any $x \in K$, we define $D_{R+x,r}$ to be the discrete Gaussian over the coset $R+x$ of the lattice R i.e. over the lattice coset $\sigma_H(R)+\sigma_H(x)$ of the lattice $\sigma_H(R)$. 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 [MR04]). *For a lattice Λ and any $\varepsilon > 0$, the* smoothing parameter *$\eta_\varepsilon(\Lambda)$ is defined as the smallest $s > 0$ s.t. $\rho_{1/s}(\Lambda^* \backslash \{\boldsymbol{0}\}) \leq \varepsilon$.*

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

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

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

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

Lemma 5 (Discrete Gaussian infinity norm bound[2]). *For any n-dimensional lattice Λ, $\sigma > 0$, if $\boldsymbol{x} \leftarrow D_{\Lambda,\sigma}$, then*

$$\Pr\left[\|\boldsymbol{x}\|_\infty \geq t\right] \leq 2n \exp\left(-\frac{\pi t^2}{\sigma^2}\right)$$

for some universal constant $c > 1$.

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 $\tilde{s} = (q^{d-1}, \ldots, 1) \cdot \mathbf{s}$ and $\tilde{a} = (1, \ldots, q^{d-1}) \cdot \mathbf{a}$, 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 error distribution. The reduction in Theorem 1 is a generalisation of the above

[2] For proof, see Corollary 10 of Daniele Micciancio's lecture notes (https://cseweb.ucsd.edu/classes/fa17/cse206A-a/LecGaussian.pdf)

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 $j = 1, \dots, d$. 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 narrow *continuous* Gaussian. In the next subsection, we state and prove some results to simplify the analysis of the reduction.

3.1 Drowning Lemma

We will use a drowning lemma stating that multiplying a discrete Gaussian by a field element and then adding a "wide-enough" continuous Gaussian results in a continuous Gaussian. This lemma is proven using the following claim and is formally stated in Lemma 6.

Claim 2. *For any $\boldsymbol{\tau} \in \mathbb{R}^n, r \in \mathbb{R}$, define $t_i = \sqrt{\tau_i^2 + r^2}$ for $i = 1, \dots, n$ and let Λ be an n-dimensional lattice. Let $X \sim D_{\Lambda+\mathbf{u},r}$ and $Y \sim D_{\boldsymbol{\tau}}$ and define the random variable $Z := X + Y$. Provided that $\tau_i r / t_i \geq \eta_\varepsilon(\Lambda)$ for all i, the distribution of Z is within statistical distance 2ε of $D_{\mathbf{t}}$.*

Proof. Throughout the proof, we denote the density function of Z at any point $\mathbf{z} \in \mathbb{R}^n$ as $p(\mathbf{z})$. Let $c_1 = \int_{\mathbb{R}^n} \rho_{\boldsymbol{\tau}}(\mathbf{x}) d\mathbf{x}$, let $\mathbf{T}$ be the diagonal matrix with $T_{i,i} = r\tau_i/t_i$ and let $\mathbf{u}_\mathbf{z}$ be the vector whose i^{th} entry is $\frac{r^2}{\tau_i^2 + r^2} z_i$. We have that

$$\begin{aligned} p(\mathbf{z}) &= \sum_{\mathbf{x} \in \Lambda + \mathbf{u}} \Pr[X = \mathbf{x}] \cdot \frac{\rho_{\boldsymbol{\tau}}(\mathbf{z} - \mathbf{x})}{c_1} \\ &= \frac{1}{c_1 \cdot \rho_r(\Lambda + \mathbf{u})} \cdot \left(\sum_{\mathbf{x} \in \Lambda + \mathbf{u}} \frac{\rho_r(\mathbf{x}) \cdot \rho_{\boldsymbol{\tau}}(\mathbf{z} - \mathbf{x})}{\rho_{\mathbf{t}}(\mathbf{z})} \right) \cdot \rho_{\mathbf{t}}(\mathbf{z}) \\ &= \underbrace{\frac{1}{c_1 \cdot \rho_r(\Lambda + \mathbf{u})}}_{(\clubsuit_1)} \cdot \underbrace{\left(\sum_{\mathbf{x} \in \Lambda + \mathbf{u}} \rho_{\mathbf{T}}(\mathbf{x} + \mathbf{u}_\mathbf{z}) \right)}_{(\clubsuit_2)} \cdot \rho_{\mathbf{t}}(\mathbf{z}). \end{aligned}$$

The term labelled $(\clubsuit_1)$ is a constant with respect to $\mathbf{z}$. Informally speaking, we will show that $(\clubsuit_2)$ is *almost* constant with respect to $\mathbf{z}$. This will imply that the mass function of Z is *almost* proportional to $\rho_{\mathbf{t}}$ as required. Let $\mathbf{v_z} := \mathbf{u} + \mathbf{u_z}$ and define $f(\mathbf{x}) := \rho(\mathbf{T}^{-1}(\mathbf{x} + \mathbf{v_z}))$. Note that[3] the Fourier transform of f is given by $\hat{f}(\mathbf{y}) = e^{2\pi i \langle \mathbf{y}, \mathbf{v_z} \rangle} \cdot \rho(\mathbf{T}\mathbf{y}) / \det(\mathbf{T}^{-1})$. Using this fact and the Poisson summation formula $\sum_{\mathbf{x} \in \Lambda} f(\mathbf{x}) = \det(\Lambda^*) \cdot \sum_{\mathbf{y} \in \Lambda^*} \hat{f}(\mathbf{y})$, we get

$$\begin{aligned}
\left| (\clubsuit_2) - \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \right| &= \left| \sum_{\mathbf{x} \in \Lambda} \rho_{\mathbf{T}}(\mathbf{x} + \mathbf{v_z}) - \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \right| \\
&= \left| \sum_{\mathbf{x} \in \Lambda} \rho(\mathbf{T}^{-1}(\mathbf{x} + \mathbf{v_z})) - \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \right| \\
&= \left| \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \sum_{\mathbf{y} \in \Lambda^*} e^{2\pi i \langle \mathbf{y}, \mathbf{v_z} \rangle} \cdot \rho(\mathbf{T}\mathbf{y}) - \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \right| \\
&= \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \cdot \left| \sum_{\mathbf{y} \in \Lambda^* \setminus \{\mathbf{0}\}} e^{2\pi i \langle \mathbf{y}, \mathbf{v_z} \rangle} \cdot \rho(\mathbf{T}\mathbf{y}) \right| \\
&\le \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \cdot \left| \sum_{\mathbf{y} \in \Lambda^* \setminus \{\mathbf{0}\}} \rho(\mathbf{T}\mathbf{y}) \right| \\
&\le \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})} \cdot \left| \sum_{\mathbf{y} \in \Lambda^* \setminus \{\mathbf{0}\}} \rho\left(\min_i \frac{\tau_i r}{t_i} \mathbf{y} \right) \right| \\
&\le \varepsilon \cdot \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})}
\end{aligned}$$

Therefore, we can conclude that $(\clubsuit_2) \in [1-\varepsilon, 1+\varepsilon] \cdot \frac{\det(\Lambda^*)}{\det(\mathbf{T}^{-1})}$ and

$$p(\mathbf{z}) \in [1-\varepsilon, 1+\varepsilon] \cdot C \cdot \rho_{\mathbf{t}}(\mathbf{z}) \tag{7}$$

for some constant C. Integrating over all $\mathbf{z}$, we get that $C \in [\frac{1}{1+\varepsilon}, \frac{1}{1-\varepsilon}] \cdot \frac{1}{c_1}$ implying that

$$p(\mathbf{z}) \in \left[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon} \right] \cdot \frac{\rho_{\mathbf{t}}(\mathbf{z})}{c_1}. \tag{8}$$

Using Claim 1 and the fact that $(1-\varepsilon)/(1+\varepsilon) \ge 1 - 2\varepsilon$ completes the proof. □

Lemma 6. *Let R be the ring of integers of a field K with degree n and take arbitrary positive $r, B \in \mathbb{R}$. For any non-zero $s \in K$, let $\boldsymbol{S} \in \mathbb{Q}^{n \times n}$ be the matrix corresponding to field multiplication by s in the space H and define $\sigma_i := \sigma_i(s)$ for $i = 1, \ldots, n$. Let $\boldsymbol{S}'$ be the diagonal matrix with $(i,i)^{th}$ entry $\sqrt{B^2 + |\sigma_i|^2}$. For any n-dimensional lattice Λ and $\mathbf{u} \in \mathbb{R}^n$, the random variable $Z = \boldsymbol{S} \cdot X + Y$ where $X \sim D_{\Lambda + \mathbf{u}, r}$ and $Y \sim D_{rB}$ is within statistical distance 2ε of $D_{r\boldsymbol{S}'}$ where provided that $\frac{rB}{\sqrt{B^2 + |\sigma_i|^2}} \ge \eta_\varepsilon(\Lambda)$ for all i.*

[3] using the convention that the Fourier transform of a function $f : \mathbb{R}^n \to \mathbb{C}$ is given by $\hat{f}(\mathbf{z}) = \int_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \cdot e^{-2\pi i \langle \mathbf{x}, \mathbf{z} \rangle} d\mathbf{x}$

Proof. Let $W = \mathbf{S}^{-1}Z = X + \mathbf{S}^{-1}Y$. The distribution of $\mathbf{S}^{-1}Y$ is a continuous Gaussian with covariance matrix $(rB)^2(\mathbf{S}^T \cdot \mathbf{S})^{-1}$ which is diagonal with i^{th} entry $(rB)^2/|\sigma_i|^2$. Therefore, we express the distribution of $\mathbf{S}^{-1}Y$ as $D_{\boldsymbol{\tau}}$ where $\tau_i = rB/|\sigma_i|$. We can now apply Claim 2 to the random variable W whenever $\frac{r^2B}{|\sigma_i|\sqrt{(rB/|\sigma_i|)^2+r^2}} = \frac{rB}{\sqrt{B^2+|\sigma_i|^2}} \geq \eta_\varepsilon(\Lambda)$. This allows us to conclude that the distribution of W is within statistical distance 2ε of $D_{\mathbf{t}}$ where $t_i = \sqrt{\tau_i^2 + r^2}$. In other words, the distribution of W is continuous Gaussian with diagonal covariance matrix $\mathbf{T}$ where $T_{i,i} = t_i^2$.

To complete the proof, we consider the distribution of $Z = \mathbf{S}W$. By using the data processing inequality for statistical distance, Z is at most a statistical distance 2ε away from a continuous Gaussian with covariance matrix $\mathbf{STS}^T$. Recall from the discussion in Section 2.2 that $\mathbf{S} = \mathbf{U}_H^\dagger \mathbf{D} \mathbf{U}_H$ for some diagonal matrix $\mathbf{D}$ and unitary $\mathbf{U}_H$ given in Equation (2). Let r_1 be the number of real field embeddings and r_2 the number of pairs of complex embeddings. This means that for $i = r_1 + 1, \ldots, r_1 + r_2$, $|\sigma_i| = |\sigma_{i+r_2}|$ and therefore that $t_i = t_{i+r_2}$. From these observations, we can see that $\mathbf{T}$ and $\mathbf{S}$ commute as they share a basis of eigenvectors given by the columns of

$$\mathbf{U}_H^\dagger = \begin{bmatrix} \mathbf{I}_{r_1} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}\mathbf{I}_{r_2} & \frac{1}{\sqrt{2}}\mathbf{I}_{r_2} \\ 0 & \frac{-i}{\sqrt{2}}\mathbf{I}_{r_2} & \frac{i}{\sqrt{2}}\mathbf{I}_{r_2} \end{bmatrix} \in \mathbb{C}^{n \times n}. \tag{9}$$

Therefore, we can write $\mathbf{STS}^T = \mathbf{T} \cdot \mathbf{SS}^T$. Since $\mathbf{T}$ and $\mathbf{SS}^T$ are both diagonal with i^{th} entry $|\sigma_i|^2$ and t_i^2 respectively, the covariance matrix associated with Z is diagonal with i^{th} entry $|\sigma_i|^2 t_i^2 = r^2(B^2 + |\sigma_i|^2)$. □

3.2 The Main Reduction

Theorem 1. *Let R be the ring of integers of some algebraic number field K of degree n, let d, d', q, q' be positive integers, $\varepsilon \in (0, 1/2)$, and $\boldsymbol{G} \in R^{d' \times d}$. Also, fix $\boldsymbol{s} = (s_1, \ldots, s_d) \in (R^\vee)^d$. 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 H^d), $\boldsymbol{B}_R$ be some known basis of R in H and*

$$r \geq \max\left(\|\tilde{\boldsymbol{B}}_\Lambda\|, \frac{1}{q}\|\tilde{\boldsymbol{B}}_R\|\right) \cdot \sqrt{2\ln(2nd(1 + 1/\varepsilon))/\pi}.$$

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 $(2d + 6)\varepsilon$ of $A^{(M')}_{q',\boldsymbol{Gs},D_{\boldsymbol{\alpha}'}}$ where $(\boldsymbol{\alpha}_i')^2 = \alpha^2 + r^2(\beta^2 + \sum_{j=1}^d |\sigma_i(s_j)|^2)$ for any β satisfying $\beta^2 \geq B^2 d$.*

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 $\mathbf{v} = \frac{1}{q}\mathbf{a}_H + \mathbf{f} \bmod R^d \in \Lambda / R^d$ 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. For some $\beta > B\sqrt{d}$ sample $\tilde{e}$ from the distribution $D_{r\beta}$ over $K_\mathbb{R} \simeq H$ 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 $\tilde{\mathbf{a}}$*.* 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 $r \geq \eta_\varepsilon(\Lambda)$. Therefore, Lemma 4 tells us that for any particular $\bar{\mathbf{a}} \in (R_q)^d$ and $\bar{\mathbf{f}} \in \Lambda - \frac{1}{q}\bar{\mathbf{a}}_H$, we have

$$\begin{aligned}
\Pr[\mathbf{a} = \bar{\mathbf{a}} \wedge \mathbf{f} = \bar{\mathbf{f}}] &= q^{-nd} \cdot \frac{\rho_r(\bar{\mathbf{f}})}{\rho_r\left(\Lambda - \frac{1}{q}\bar{\mathbf{a}}_H\right)} \\
&= \frac{q^{-nd}}{\rho_r(\Lambda)} \cdot \frac{\rho_r(\Lambda)}{\rho_r\left(\Lambda - \frac{1}{q}\bar{\mathbf{a}}_H\right)} \cdot \rho_r(\bar{\mathbf{f}}) \qquad (10) \\
&\in C \cdot \left[1, \frac{1+\varepsilon}{1-\varepsilon}\right] \cdot \rho_r(\bar{\mathbf{f}})
\end{aligned}$$

where $C := q^{-nd}/\rho_r(\Lambda)$ 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 $\bar{\mathbf{v}} \in \Lambda / R^d$,

$$\begin{aligned}
\Pr[\mathbf{v} = \bar{\mathbf{v}}] &\in C \cdot \left[1, \frac{1+\varepsilon}{1-\varepsilon}\right] \cdot \rho_r(q^{-1}R^d + \bar{\mathbf{v}}) \\
&\in C \cdot \rho_r(q^{-1}R^d) \cdot \left[1, \frac{1+\varepsilon}{1-\varepsilon}\right] \cdot \frac{\rho_r(q^{-1}R^d + \bar{\mathbf{v}})}{\rho_r(q^{-1}R^d)} \qquad (11) \\
&\in C' \cdot \left[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}\right]
\end{aligned}$$

where $C' := C\rho_r(q^{-1}R^d)$. Note that we may apply Lemma 4 here since we know that $r \geq \eta_\varepsilon(q^{-1}R^d)$ by Lemma 3. This allows us to conclude that the distribution of $\mathbf{v}$ is within statistical distance $1 - [(1-\varepsilon)/(1+\varepsilon)]^2 \leq 4\varepsilon$ 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 by the data processing inequality. It is also clear that if b is uniform random, then so is $\tilde{b}$. 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 $\bar{\mathbf{f}}$ in Equation (10), one can show that the distribution of $-\mathbf{f}$ for any fixed $\tilde{\mathbf{a}}$ is within statistical distance $1 - (1-\varepsilon)(1+\varepsilon) \leq 2\varepsilon$ of $D_{\frac{1}{q}R^d - \bar{\mathbf{v}}, r}$.

Distribution of the error. Suppose we are given the MLWE sample $(\mathbf{a}, b = \frac{1}{q}\langle \mathbf{a}, \mathbf{s}\rangle + e) \in (R_q)^d \times \mathbb{T}_{R^\vee}$ as input 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

$$(\tilde{b} - \frac{1}{q'}\langle \bar{\tilde{\mathbf{a}}} \cdot \tilde{\mathbf{s}}\rangle) \bmod R^\vee \tag{12}$$

where $\tilde{\mathbf{s}} = \mathbf{G}\mathbf{s}$. 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

$$(\tilde{b} - \frac{1}{q'}\langle \bar{\tilde{\mathbf{a}}}, \tilde{\mathbf{s}}\rangle) = \sum_{i=1}^{d} s_i \cdot \sigma_H^{-1}(-\mathbf{f}_i) + \tilde{e} + e \bmod R^\vee. \tag{13}$$

We want to show that the RHS of this equation is almost distributed as a Gaussian in canonically embedded space H. To do so, define the invertible matrix $\mathbf{S}_{i,H} \coloneqq \mathbf{U}_H \mathbf{S}_i \mathbf{U}_H^\dagger \in \mathbb{R}^{n\times n}$ where $\mathbf{U}_H$ is given in Equation (2) and $\mathbf{S}_i$ is the diagonal matrix with the field embeddings of s_i along the diagonal i.e. $[\mathbf{S}_i]_{jk} = \sigma_j(s_i)\delta_{jk}$. Note that $\mathbf{S}_{i,H}$ is the matrix representing field multiplication by s_i in the basis $(\mathbf{h}_i)_{i=1}^n$ of H. Therefore, in canonical space, the error is given by

$$\sum_{i=1}^{d} \mathbf{S}_{i,H} \cdot (-\mathbf{f}_i) + \sigma_H(\tilde{e}) + \sigma_H(e) \bmod R^\vee \tag{14}$$

where $\sigma_H(\tilde{e})$ and $\sigma_H(e)$ are distributed as $D_{r\beta}$ and D_α respectively. 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 $\gamma_i \geq rB$ provided that $\sum_{i=1}^d \gamma_i^2 = r^2\beta^2$. 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 $D_{\frac{1}{q}R - \bar{\mathbf{v}}_i, r}$. Therefore, writing

$$\sum_{i=1}^{d} \mathbf{S}_{i,H} \cdot (-\mathbf{f}_i) + \sigma_H(\tilde{e}) = \sum_{i=1}^{d} \mathbf{S}_{i,H} \cdot (-\mathbf{f}_i) + \tilde{e}^{(i)}, \tag{15}$$

and applying Lemma 6 to the summand on the RHS, we find (using the triangle and data processing inequalities for statistical distance) that the error term in Equation (14) is a statistical distance of at most $2\varepsilon + 2d\varepsilon$ away from $D_{\boldsymbol{\alpha}'}$. □

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 non-standard in general, but we can always use the usual re-randomizing process to obtain a uniform secret i.e. sample $s' \leftarrow U(R_q^\vee)$ and transform (a, b) to $(a, b + \frac{1}{q}(a \cdot s'))$.

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

$$\mathrm{Pr}_{s \leftarrow \chi}\left[\max_i |\sigma_i(s)| > B\right] \leq \delta$$

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

$$r \geq \frac{1}{q}\ \|\tilde{\boldsymbol{B}}_R\|\ \ \cdot \sqrt{2\ln(2nd(1 + 1/\varepsilon))/\pi},$$

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

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

Moreover, this reduction reduces the advantage by at most $[1 - (1 - \delta)^d] + (2d + 10)\varepsilon m$.

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. The main task is to show that the conditions on r in the theorem are satisfied by the choice of r in this corollary. In particular, for $\mathit{\Lambda} = \frac{1}{q'}\mathbf{G}_H^T R^{d'} + R^d$, we will attempt to express $\|\tilde{\mathbf{B}}_{\mathit{\Lambda}}\|$ in terms of $\|\tilde{\mathbf{B}}_R\|$. Define $\mathbf{g} := (1, q, \ldots, q^{k-1})$ so that $\mathbf{G} = \mathbf{I}_{d/k} \otimes \mathbf{g}$. In the canonical embedding, we may write the lattice from Theorem 1 as $\mathit{\Lambda} = \frac{1}{q^k}(\mathbf{I}_{d/k} \otimes \mathbf{g}^T \otimes \mathbf{I}_n) \cdot (\mathbf{I}_{d/k} \otimes \mathbf{B}_R) \cdot \mathbb{Z}^{nd/k} + (\mathbf{I}_d \otimes \mathbf{B}_R) \cdot \mathbb{Z}^{nd}$. Pre-multiplying by $\mathbf{I}_d \otimes \mathbf{B}_R^{-1}$, we get the related lattice

$$\begin{aligned}
\mathit{\Lambda}' &:= \left(\mathbf{I}_d \otimes \mathbf{B}_R^{-1}\right) \cdot \mathit{\Lambda} \\
&= \frac{1}{q^k}\left(\mathbf{I}_{d/k} \otimes \mathbf{g}^T \otimes \mathbf{B}_R^{-1}\right) \cdot \left(\mathbf{I}_{d/k} \otimes \mathbf{B}_R\right) \cdot \mathbb{Z}_q^{nd/k} + \mathbb{Z}^{nd} \\
&= \frac{1}{q^k}\left(\mathbf{I}_{d/k} \otimes \mathbf{g}^T \otimes \mathbf{I}_n\right) \cdot \mathbb{Z}^{nd/k} + \mathbb{Z}^{nd}
\end{aligned}$$

The lattice $\bar{\Lambda}$ can be shown to have basis $\mathbf{B}' = \mathbf{I}_{d/k} \otimes \mathbf{Q} \otimes \mathbf{I}_n$ where

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

Pre-multiplying $\mathbf{B}'$ by $\mathbf{I}_d \otimes \mathbf{B}_R$ gives us a basis $\mathbf{B}_\Lambda = \mathbf{I}_{d/k} \otimes \mathbf{Q} \otimes \mathbf{B}_R$ for Λ. Orthogonalising from left to right, we can see that $\|\tilde{\mathbf{B}}_\Lambda\|$ is precisely $\frac{1}{q}\|\tilde{\mathbf{B}}_R\|$.

Finally, the loss in advantage can be derived from the statistical distances in Theorem 1 and the fact that the reduction is only guaranteed to work when each of the d secret polynomials has embeddings of modulus at most B (which occurs with probability at least $(1-\delta)^d$). □

3.3 Normal Form Secrets

There are well-known results stating that an LWE problem where the secret is drawn from the error distribution is at least as hard as an LWE problem where the secret is uniform [ACPS09]. The connection is straight-forward for plain LWE, but for R/MLWE, we need to remember that the secrets lie in $(R^\vee)^d$ whereas the errors are sampled from $K \otimes \mathbb{R}$. However, this complication can be solved using discretisation techniques that transform continuous errors to discrete ones [LPR13]. Using discretization, Langlois et al. [LS15] showed that $\mathsf{MLWE}^M_{q,D_\alpha}(U)$ is at least as hard as $\mathsf{MLWE}^M_{q,D_{\frac{1}{q}R^\vee,\sqrt{2}\alpha}}\left(D_{(R^\vee)^d,\sqrt{2}q\alpha}\right)$. We can also add continuous Gaussian noise to $\mathsf{MLWE}^M_{q,D_{\frac{1}{q}R^\vee,\sqrt{2}\alpha}}\left(D_{(R^\vee)^d,\sqrt{2}q\alpha}\right)$ challenges in an attempt to reintroduce continuous noise distributions. By Claim 3.9 from [Reg05], adding the noise $D_{\sqrt{2}\alpha}$ results in challenges that are statistically close to those from $\mathsf{MLWE}^M_{q,D_{2\alpha}}\left(D_{(R^\vee)^d,\sqrt{2}q\alpha}\right)$ provided that $\alpha q \geq \|\tilde{\mathbf{B}}_{R^\vee}\| \cdot \tilde{O}(1)$. An informal summary of this is that we may use a discrete Gaussian secret with a continuous Gaussian error distribution roughly q times narrower without compromising hardness.

Now that we have established the significance of secret distributions of the form $D_{(R^\vee)^d,\sqrt{2}q\alpha}$, we next discuss valid choices of (B,δ) with respect to this secret distribution. The below lemma shows that we can choose $B = \sqrt{2}q\alpha n^c$ for any positive constant c along with $\delta = 2n\exp\left(-\pi n^{2c}\right)$ which is negligible assuming that n is polynomial in the security parameter.

Lemma 7. *For any algebraic number field K of degree n with ring of integers R, any $\sigma > 0$ and any constant $c > 0$*

$$\Pr_{s \leftarrow D_{R^\vee,\sigma}}\left[\max_i |\sigma_i(s)| > \sigma n^c\right] \leq 2n\exp\left(-\pi n^{2c}\right).$$

Proof. We first recall that $s \leftarrow D_{(R^\vee),\sqrt{2}q\alpha}$ means that $\sigma_H(s) \leftarrow D_{\sigma_H((R^\vee)),\sqrt{2}q\alpha}$. Assuming r_1 real embeddings and r_2 pairs of complex embeddings so that

$n = r_1 + 2r_2$, we have that

$$\sigma_H(s) = \begin{bmatrix} \sigma_1(s) \\ \vdots \\ \sigma_{r_1}(s) \\ \sqrt{2}\ \mathrm{Re}(\sigma_{r_1+1}(s)) \\ \vdots \\ \sqrt{2}\ \mathrm{Re}(\sigma_{r_1+r_2}(s)) \\ \sqrt{2}\ \mathrm{Im}(\sigma_{r_1+1}(s)) \\ \vdots \\ \sqrt{2}\ \mathrm{Im}(\sigma_{r_1+r_2}(s)) \end{bmatrix}.$$

It is clear that $|\sigma_i(s)| \le |\sigma_H(s)|_\infty$ for $i = 1, \dots, r_1$. For $i = r_1+1, \dots, r_1+r_2$, we have that $2|\sigma_i(s)|^2 = \sigma_H(s)_i^2 + \sigma_H(s)_{i+r_2}^2 \le 2|\sigma_H(s)|_\infty^2$. Therefore, we have that $\max_i |\sigma_i(s)| \le |\sigma_H(s)|_\infty$. Applying Lemma 5 to $\sigma_H(s)$ completes the proof. □

Corollary 2. *Let R be a ring with basis $\boldsymbol{B}_R$ in the canonical embedding, $c > 0$ be an arbitrary constant and χ denote the distribution $D_{(R^\vee)^d, \alpha q}$. Also take any $\alpha > 0$ and any $\varepsilon \in (0, 1/2)$. For any $k > 1$ that divides d and*

$$r \ge \frac{1}{q}\ \|\tilde{\boldsymbol{B}}_R\|\ \cdot \sqrt{2\ln(2nd(1+1/\varepsilon))/\pi},$$

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

$$(\alpha')^2 \ge \alpha^2 \left(1 + 2(rqn^c)^2 d\right).$$

Moreover, this reduction reduces the advantage by at most $[1 - (1-\delta)^d] + (2d + 10)\varepsilon m$ where $\delta = 2n\exp(-\pi n^2 c)$.

3.4 Instantiation: Power-of-Two Cyclotomic Rings

We now consider the case of cyclotomic rings with power-of-two dimension n. It can be shown that the map taking the coefficient embedding to the canonical embedding is a scaled isometry with scaling factor $\sqrt{n}$. In this case, we have $R = \mathbb{Z}(\xi)$ for $(2n)^{th}$ primitive root of unity ξ. Taking the "power basis" of R given by $1, \xi, \dots, \xi^{n-1}$, gives us an orthonormal lattice basis of R in the coefficient embedding. Applying the aforementioned scaled isometry, we find an *orthogonal* basis in the canonical embedding where each vector has length $\sqrt{n}$. Therefore, in the canonical embedding $\|\tilde{\mathbf{B}}_R\| = \sqrt{n}$ when using this basis.

In the following, we will informally take normal form MLWE to mean the case where the error distribution is D_α for some α and where the secret distribution is $D_{R^\vee, q\alpha}$ (i.e. we will ignore any small constant factor differences between the

Gaussian parameters of the secret and error distributions). We assume that m, n and d are $\mathsf{poly}(\lambda)$. We will also use $\delta = 2n \exp\left(-\pi n^{2c}\right)$, $\varepsilon = n^{-\log n}$ as negligible functions while ignoring constant/logarithmic factors. In this setting, the losses in advantage are negligible. Taking the above into account, Corollary 2 shows we can reduce normal form MLWE in modulus q, module rank d to MLWE in modulus q^k, module rank d/k with a uniform secret distribution (after re-randomising the secret). In particular, the bound on the width of the error distribution grows from α to roughly $n^{c+1/2}\sqrt{d}\alpha$ for any positive constant c. Since normal form MLWE is at least as hard as uniform secret MLWE, we also have that uniform secret MLWE in modulus q, rank d reduces to uniform secret MLWE in modulus q^k, rank d/k at the cost of roughly a factor $n^{c+1/2}\sqrt{d}$ blow-up in the bound on the error rate. Finally, taking $k = d$ gives us the result that RLWE in modulus q^d is at least as hard as MLWE with modulus q, rank d with error rate roughly $n^{c+1/2}\sqrt{d}$ smaller than the RLWE error rate bound.

4 Strictly Spherical Error Distributions

Note that the reduction presented in Theorem 1 results in a skewed, but bounded error distribution. We will now present a lemma that allows us to reduce from MLWE to MLWE with a *spherical* error distribution following the strategy laid out in [LPR10]. The price paid when targeting a strictly spherical error distribution is a larger (but still polynomial) blow-up in the error rate. Corollary 3 below shows that the extra blow-up factor incurred when targeting a strictly spherical distribution can be as low as $(mn)^{1/4}$ where m is the number of MLWE samples provided. It is important to note that we will be using the Renyi divergence to carry out this analysis. As a result, the analysis only applies to the *search variants* of the MLWE problem. Note that the reduction and analysis of Theorem 1 implicitly contains a reduction and analysis between *search* variants of MLWE. This is because the reduction takes the distribution $A^{(M)}_{q,\mathbf{s},D_\alpha}$ to a distribution statistically close to $A^{(M)}_{q'\mathbf{s},D_\alpha}$. It must be noted that the reduction maps a secret $\mathbf{s} \in (R^\vee)^d$ to $\tilde{\mathbf{s}} = \mathbf{G} \cdot \mathbf{s}$, so we only have a search to search reduction when this mapping is efficiently invertible between the spaces $(R_q^\vee)^d$ and $(R_{q^k}^\vee)^{d/k}$ as is the case for $\mathbf{G} = \mathbf{I}_{d/k} \otimes (1, q, \ldots, q^{k-1})$.

Lemma 8. *For integers m, n, let $\boldsymbol{M} \in \mathbb{R}^{m \times n}$ be a matrix with non-zero singular values σ_i for $i = 1, \ldots, n$ such that $\sigma_1^2 \geq \cdots \geq \sigma_n^2$ and take $\beta^2 \geq \sigma_1^2$. Then*

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

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,

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

For the second claim, we have

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

□

We now use the above lemma to show that there is a search MLWE to search MLWE reduction where the resulting error distribution is essentially spherical.

Corollary 3. *Let R be a ring with basis $\boldsymbol{B}_R$ in the canonical embedding and χ be a distribution satisfying*

$$\Pr_{s\leftarrow\chi}\left[\max_i|\sigma_i(s)|>B\right]\leq\delta$$

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

$$r\geq\frac{1}{q}\ \|\tilde{\boldsymbol{B}}_R\|\ \cdot\sqrt{2\ln(2nd(1+1/\varepsilon))/\pi}$$

and define $t:=\sqrt{\alpha^2+(rB(mn)^{1/4})^2}$. Suppose there exists a PPT algorithm solving s-$\mathsf{MLWE}^{(R^{d/k})}_{m,q^k,D_t}(\boldsymbol{G}\cdot\chi^d)$ where $\boldsymbol{G}=\boldsymbol{I}_{d/k}\otimes(1,q,\ldots,q^{k-1})\in R^{d/k\times d}$ with probability p. Then there is a PPT algorithm solving s-$\mathsf{MLWE}^{(R^d)}_{m,q,D_\alpha}(\chi^d)$ with probability at least $\frac{(1-\delta)^d p^2}{2}-\big(1-(1-\delta)^d+(2d+6)\varepsilon m\big)$. Alternatively, *if we define* $t := \sqrt{\alpha^2 + (rB(mn)^{1/2})^2}$*, there is a PPT algorithm solving* s-$\mathsf{MLWE}^{(R^d)}_{m,q,D_\alpha}(\chi^d)$ *with probability at least* $\frac{(1-\delta)^d p}{2} - \big(1 - (1-\delta)^d + (2d+6)\varepsilon m\big)$.

Proof. Consider the first value of t in the lemma statement. We run the reduction from Theorem 1 choosing $\beta = B(mn)^{1/4}\sqrt{d}$. The resulting error distribution has a diagonal covariance with i^{th} entry $(\boldsymbol{\alpha}')_i = \alpha^2 + r^2(\beta^2 + \sum_{j=1}^d |\sigma_i(s_j)|^2)$ where the s_j are the components of the original MLWE secret. In addition, the loss in success probability is at most $1 - (1-\delta)^d + (4d+6)\varepsilon$.

Next consider the loss in success probability when applying an algorithm solving s-$\mathsf{MLWE}^{(R^{d/k})}_{m,q^k,D_{\boldsymbol{\alpha}'}}(\mathbf{G}\cdot\chi^d)$ with probability p to the s-$\mathsf{MLWE}^{(R^{d/k})}_{m,q^k,D_t}(\mathbf{G}\cdot\chi^d)$ problem. We can apply Lemma 8 after setting $\mathbf{M}$ to be the diagonal matrix with $\sqrt{\sum_{j=1}^d |\sigma_i(s_j)|^2}$ in the i^{th} position. Conditioned on B being larger than $\max_i |\sigma_i(s_j)|$ for $j = 1, \ldots, d$ (which occurs with probability $(1-\delta)^d$), we find that

$$R_2\left((D_{r\beta})^m || (D_{r(\beta^2\mathbf{I}+\mathbf{M}^T\mathbf{M})^{1/2}})^m\right) \le \left(1 + \frac{\sigma_1^4}{\beta^4}\right)^{mn/2} \le \left(1 + \frac{1}{mn}\right)^{mn/2}.$$

This quantity is upper bounded by $\exp(1/2) \le 2$. By the data processing inequality, 2 is also an upper bound for the Renyi divergence between the MLWE distribution with error distributions D_t and $D_{\boldsymbol{\alpha}'}$. Therefore, a PPT algorithm solving s-$\mathsf{MLWE}^{(R^{d/k})}_{m,q^k,D_{\boldsymbol{\alpha}'}}(\mathbf{G}\cdot\chi^d)$ with probability p also solves s-$\mathsf{MLWE}^{(R^{d/k})}_{m,q^k,D_t}(\mathbf{G}\cdot\chi^d)$ with probability at least $(1-\delta)^d \cdot p^2/2$. Repeating the analysis for Renyi divergences of order infinity completes the proof. □

5 Power-of-Two Cyclotomic RLWE to RLWE Reductions

In this section we restrict our attention to the popular power-of-two cyclotomic rings. We use a recent result of Peikert and Pepin [PP19] along with our reduction from MLWE to RLWE to show that RLWE in power-of-two dimension n and modulus q is at least as hard as RLWE in power-of-two dimension $n' := n/2^i$ and modulus q^{2^i}. When performing the associated reduction, the error rate grows by a rough factor of $n^{3/2}/(n')^{1-c}$ for an arbitrary constant $c > 0$. Next, we state the particular result taken from [PP19] that is used to derive the aforementioned hardness result. Note that this result is associated to the order-LWE (OLWE) problem which is a direct generalisation of RLWE that replaces a ring of integers with an order of a number field.

Theorem 2 (Theorem 6.1 [PP19]). *Let* K'/K *be a number field extension;* $\mathcal{O}$ *be an order of* K*;* $\mathcal{O}'$ *be an order of* K' *that is a rank-*d *free* $\mathcal{O}$*-module with known basis* $\overrightarrow{b} = (b_1, \ldots, b_d) \in (K')^d$*;* ψ' *be a distribution over* $K_\mathbb{R}$*; and* q *be a positive integer. Then there is an efficient deterministic reduction preserving the number of samples from* $\mathsf{OLWE}^{\mathcal{O}'}_{q,\psi'}(\chi')$ *to* $\mathsf{OLWE}^{\mathcal{O}^d}_{q,\psi}(\chi)$ *where* $\psi = \mathrm{Tr}_{K'_\mathbb{R}/K_\mathbb{R}}(\psi')$ *and* $\chi = \mathrm{Tr}_{K'/K}(\chi' \cdot \overrightarrow{b})$*.*

Using previously established notation and the fact that a ring of integers of a field is an order of the field, we get the following corollary.

Corollary 4. *For any $i \in \{1, 2, \ldots, \log_2 n\}$ for power-of-two n, there exists an efficient reduction from $\mathsf{RLWE}^{R_n}_{m,q,D_\alpha}(D_{R_n^\vee,\alpha q})$ to $\mathsf{MLWE}^{R^{2^i}_{n/2^i}}_{m,q,D_{2^i\alpha}}(D_{R^\vee_{n/2},2^i\alpha q})$.*

Proof. We instantiate Theorem 2 by setting ξ to be a primitive $(2n)^{th}$ root of unity for power-of-two n. Then we set $K' = \mathbb{Q}(\xi)$ and $K = \mathbb{Q}(\xi^{2^i})$ to be cyclotomic fields. The rings of integers $R' := \mathbb{Z}(\xi) = R_n$ and $R := \mathbb{Z}(\xi^{2^i}) = R_{n/2^i}$ are cyclotomic rings of dimension n and $n/2^i$ respectively. Since rings of integers are orders, we may set $\mathcal{O}' = R'$ and $\mathcal{O} = R$. It is easy to see that R' is a rank 2^i R-module with basis $\overrightarrow{b} = (1, \xi, \ldots, \xi^{2^i-1})$. Finally, we show that $\mathrm{Tr}_{K'/K}(D_{R_n^\vee,\alpha q} \cdot \overrightarrow{b}) = D_{R^\vee_{n/2^i},2^i\alpha q}$ and $\mathrm{Tr}_{K'_\mathbb{R}/K_\mathbb{R}}(D_\alpha) = D_{2^i\alpha}$. As the simplest example, take $i = 1$. It is not hard to see that for any $e = \sum_{j=0}^{n-1} e_j\xi^j \in K'$, the linear map representing multiplication by e in the K-module basis $(1, \xi)$ is given by the 2×2 matrix

$$\begin{bmatrix} e_0 + e_2\xi^2 + \cdots + e_{n-2}\xi^{n-2} & e_1\xi^2 + e_3\xi^4 + \cdots + e_{n-1}\xi^n \\ e_1 + e_3\xi^2 + \cdots + e_{n-1}\xi^{n-1} & e_0 + e_2\xi^2 + \cdots + e_{n-2}\xi^{n-2} \end{bmatrix}.$$

Therefore $\mathrm{Tr}_{K'/K}(e) = 2 \cdot (e_0 + e_2\xi^2 + \cdots + e_{n-2}\xi^{n-2})$. In a similar fashion, one may see that $\mathrm{Tr}_{K'/K}(e\xi) = 2 \cdot (-e_{n-1} + e_1\xi^2 + \cdots + e_{n-3}\xi^{n-2})$. It is straightforward to show that $\mathrm{Tr}_{K'_\mathbb{R}/K_\mathbb{R}}$ takes exactly the same form. It can now be seen that if $\psi' = D_\alpha$, then $\psi = D_{2\alpha}$ and if $\chi' = D_{(R')^\vee,\alpha q}$, then $\chi = (D_{R^\vee,2\alpha q})^2$. This analysis can straight-forwardly be generalised to consider the remaining values of i. □

We can now compose Corollary 4 with Corollary 2 in the context of power-of-two cyclotomic rings and use the two-step reduction below:

$$\mathsf{RLWE}^{R_n}_{m,q,D_\alpha}(D_{R_n^\vee,\alpha q}) \xrightarrow{\text{Cor.4}} \mathsf{MLWE}^{R^{2^i}_{n/2^i}}_{m,q,D_{2^i\alpha}}(D_{R^\vee_{n/2^i},2^i\alpha q}) \xrightarrow{\text{Cor.2}} \mathsf{RLWE}^{R_{n/2^i}}_{m,q^{2^i},\Psi_{\leq 2^i\alpha'}}.$$

In terms of power-of-two cyclotomic rings, we get the following corollary.

Corollary 5. *For any $m = \mathsf{poly}(\lambda)$, R a cyclotomic ring with power-of-two dimension $n = \mathsf{poly}(\lambda)$, $\alpha > 0$, $i \in \{1, 2, 2^2, \ldots, n/2\}$ and constant $c > 0$, there exists an efficient reduction from $\mathsf{RLWE}^{R_n}_{m,q,D_\alpha}(D_{R_n^\vee,\alpha q})$ to $\mathsf{RLWE}^{R_{n/2^i}}_{m,q^{2^i},\Psi_{\leq\alpha''}}$ where $\alpha'' = 2^{i(1-c)}n^{c+1/2} \cdot \alpha$ that reduces the advantage by a negligible additive term.*

Importantly, the above corollary yields a decision to decision reduction improving over previous work [AD17] that only considered search to search reductions. In addition, reducing from dimension n to $n/2$ (i.e. taking $i = 1$) the error rate grows from α to $n^{c+1/2} \cdot \alpha$ for any constant $c > 0$ in this two step reduction (ignoring constant/logarithmic factors and choosing parameters such that losses in advantage are negligible). Another advantage of the two-step reduction over [AD17] is that we can reduce to any power of two dimension $n' < n$. Using normal form secrets, the error rate grows roughly by a factor of $\left(\frac{n}{n'}\right)^{3/2} \cdot (n')^{c+1/2} = \frac{n^{3/2}}{(n')^{1-c}}$. This shows the hardness of small power-of-two dimensional RLWE over cyclotomic rings, provided that the modulus is chosen to be sufficiently large.

Acknowledgements

We thank Katharina Boudgoust, Léo Ducas and Adeline Roux-Langlois for helpful discussions on earlier versions of this work. We also thank our anonymous referees for their feedback.

References

ACPS09. Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In Shai Halevi, editor, *CRYPTO 2009*, volume 5677 of *LNCS*, pages 595–618. Springer, Heidelberg, August 2009.

AD17. Martin R. Albrecht and Amit Deo. Large modulus ring-LWE $\geq$ module-LWE. In Tsuyoshi Takagi and Thomas Peyrin, editors, *ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 267–296. Springer, Heidelberg, December 2017.

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

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 2016*, 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.

BDK$^+$17. Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, and Damien Stehlé. CRYSTALS – Kyber: a CCA-secure module-lattice-based KEM. Cryptology ePrint Archive, Report 2017/634, 2017. http://eprint.iacr.org/2017/634.

BG14. Shi Bai and Steven D. Galbraith. An improved compression technique for signatures based on learning with errors. In Josh Benaloh, editor, *CT-RSA 2014*, volume 8366 of *LNCS*, 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, Part I*, volume 9562 of *LNCS*, 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*, pages 309–325. ACM, 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, *ASIACRYPT 2015, Part I*, volume 9452 of *LNCS*, 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 ACM STOC*, 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, *EUROCRYPT 2016, Part II*, volume 9666 of *LNCS*, 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, *EUROCRYPT 2017, Part I*, volume 10210 of *LNCS*, pages 324–348. Springer, Heidelberg, April / May 2017.

CG13. Ran Canetti and Juan A. Garay, editors. *CRYPTO 2013, Part I*, volume 8042 of *LNCS*. 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.

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

DLL$^+$17. Léo Ducas, Tancrède Lepoint, Vadim Lyubashevsky, Peter Schwabe, Gregor Seiler, and Damien Stehle. CRYSTALS – Dilithium: Digital signatures from module lattices. Cryptology ePrint Archive, Report 2017/633, 2017. `http://eprint.iacr.org/2017/633`.

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

Gen09. Craig Gentry. Fully homomorphic encryption using ideal lattices. In Michael Mitzenmacher, editor, *41st ACM STOC*, pages 169–178. ACM Press, May / June 2009.

Gen10. Craig Gentry. Toward basing fully homomorphic encryption on worst-case hardness. In Tal Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 116–137. Springer, Heidelberg, August 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 ACM STOC*, 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.

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, *ASIACRYPT 2016, Part II*, volume 10032 of *LNCS*, 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, *FSE 2008*, volume 5086 of *LNCS*, 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, *CT-RSA 2011*, volume 6558 of *LNCS*, 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 Henri Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 1–23. Springer, Heidelberg, May / June 2010.

LPR13. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-LWE cryptography. In Thomas Johansson and Phong Q. Nguyen, editors, *EUROCRYPT 2013*, volume 7881 of *LNCS*, 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, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 239–256. Springer, Heidelberg, May 2014.

MR04. Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th FOCS*, pages 372–381. IEEE Computer Society Press, October 2004.

MR09. Daniele Micciancio and Oded Regev. Lattice-based cryptography. In Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, editors, *Post-Quantum Cryptography*, pages 147–191. Springer, Heidelberg, Berlin, Heidelberg, New York, 2009.

PP19. Chris Peikert and Zachary Pepin. Algebraically structured LWE, revisited. In Dennis Hofheinz and Alon Rosen, editors, *TCC 2019, Part I*, volume 11891 of *LNCS*, pages 1–23. Springer, Heidelberg, December 2019.

PRS17. Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *49th ACM STOC*, pages 461–473. ACM Press, June 2017.

Reg05. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, 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.

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

WW19. Yang Wang and Mingqiang Wang. Module-lwe versus ring-lwe, revisited. Cryptology ePrint Archive, Report 2019/930, 2019. https://eprint.iacr.org/2019/930.

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 samples in the public key. For example, if $d = 2$ the public key becomes

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

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

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

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 $e_1, e_2 = 0, 0$. That is, in the construction above, we can replace the Module-LWE assumption by the leftover hash lemma if d is sufficiently big.

B Powerful Ring LWE Adversaries

We show that an adversary that is able to solve search $\mathsf{RLWE}^{(R)}_{m=1,q^d,D_\alpha}$ implies an adversary that can solve search $\mathsf{LWE}_{m=n,d,q,D_\alpha}$ 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. This result is already mention as a simple corollary in the introduction to [BLP$^+$13], but we give a slightly more detailed account here.

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

$$\left(a_i, b_i = \frac{1}{q^d} \cdot a_i s_0 + e_i\right) \in \mathbb{Z}_{q^d} \times \mathbb{T} \text{ for } i = 0, \dots, n-1 \tag{16}$$

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 $R \simeq \mathbb{Z}[X]/\left\langle X^n + 1\right\rangle$ 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

$$s := \frac{s_0}{n} + s_1 \cdot X + \dots + s_{n-1} \cdot X^{n-1} \in R^\vee_{q^d}, \tag{17}$$

$$a := a_0 + a_1 \cdot X + \dots + a_{n-1} \cdot X^{n-1} \in R_{q^d}. \tag{18}$$

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 (16) 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

$$\tilde{b} - \frac{1}{q^d} a \cdot s = \frac{1}{n}\left(e_0 + e_1 \cdot X + \dots + e_{n-1} \cdot X^{n-1}\right) \bmod R^\vee. \tag{19}$$

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.