Dimension-Preserving Reductions from LWE to LWR

Jacob Alperin-Sheriff[1] and Daniel Apon[1]

University of Maryland: jacobmas@umd.edu, dapon@cs.umd.edu

Abstract. The Learning with Rounding (LWR) problem was first introduced by Banerjee, Peikert, and Rosen (Eurocrypt 2012) as a *derandomized* form of the standard Learning with Errors (LWE) problem. The original motivation of LWR was as a building block for constructing efficient, low-depth pseudorandom functions on lattices. It has since been used to construct reusable computational extractors, lossy trapdoor functions, and deterministic encryption.

In this work we show two (incomparable) dimension-preserving reductions from LWE to LWR in the case of a *polynomial-size modulus*. Prior works either required a superpolynomial modulus q, or lost at least a factor $\log(q)$ in the dimension of the reduction. A direct consequence of our improved reductions is an improvement in parameters (i.e. security and efficiency) for each of the known applications of poly-modulus LWR.

Our results directly generalize to the ring setting. Indeed, our formal analysis is performed over "module lattices," as defined by Langlois and Stehlé (DCC 2015), which generalize both the general lattice setting of LWE and the ideal lattice setting of RLWE as the single notion M–LWE. We hope that taking this broader perspective will lead to further insights of independent interest.

1 Introduction

The Learning with Rounding (LWR) problem was first introduced by Banerjee, Peikert and Rosen [BPR12] as a derandomization of the standard Learning with Errors (LWE) problem [Reg09]. In dimension d with modulus q and a noise distribution ψ, an $\mathsf{LWE}_{d,q,\psi}$ sample for a secret $\mathbf{s}$ consists of a uniformly random $\mathbf{a} \in \mathbb{Z}_q^d$ and $b = \langle \mathbf{a}, \mathbf{s} \rangle + e \bmod q \in \mathbb{Z}_q$, where e is small, random "noise" sampled from ψ. By contrast, the corresponding $\mathsf{LWR}_{d,q,p}$ sample in dimension d with modulii $q > p$ is formed by simply rounding off the lower-order bits, instead setting $\mathbf{a} \in \mathbb{Z}_q^d$ uniformly random and $b = \lfloor (p/q)\langle \mathbf{a}, \mathbf{s} \rangle \rceil \in \mathbb{Z}_p$. The initial application for the LWR problem was as a building block in constructing efficient, low-depth pseudorandom functions, and there have been a number of further works in this area, cf. [BP14, BFP+15, BLMR13].

LWR can also be used to directly construct public key cryptosystems and other "Cryptomania" cryptographic primitives [AKPW13]. In particular, constructions based on LWR have an important implementation advantage over LWE. As mentioned, constructions based on LWE require sampling an error term

from a discrete (or rounded) Gaussian distribution. While this can be accomplished efficiently [GPV08, Pei10], it is not nearly as efficient as the rounding operation in LWR. More importantly, the known efficient algorithms for sampling are not naturally "constant-time" (although they can be modified to be made constant-time [PG13]) making an implementation potentially vulnerable to timing attacks. Moreover, it is significantly more difficult to correctly implement cryptosystems directly based on LWE (especially without a strong understanding of the underlying mathematics) than the simple, deterministic rounding operation in LWR. These implementation woes are only compounded when considering practical implementations, which *must* take place in the ring setting for efficiency reasons [LPR13b].

As a result of LWR's usefulness, there has been significant theoretical work devoted to obtaining good reductions from LWE (and hence from worst-case hard lattice problems) to LWR. As mentioned previously, [BPR12] first introduced LWR, and in particular gave a reduction from LWE to LWR when the ratio of the LWR modulii q/p is superpolynomial in the security parameter κ. In a little more detail, for any efficiently samplable B-bounded distribution ψ (where each noise sample $e \leftarrow \psi$ has magnitude at most B w.h.p.), any modulii q, p so that $q \geq p \cdot B \cdot \kappa^{\omega(1)}$, and any distribution over the secret $\mathbf{s}$, [BPR12] showed that distinguishing unboundedly-many $\mathsf{LWR}_{d,q,p}$ samples from uniform samples is at least as hard as distinguishing as many $\mathsf{LWE}_{d,q,\psi}$ samples from uniform.

Banerjee, Peikert and Rosen also show analogous results for the case of the Ring Learning with Errors (RLWE) problem [LPR13a] and Ring Learning with Rounding (RLWR), where for an appropriate choice of ring R and secret $s \in R$, $\mathsf{RLWE}_{R,q,\psi}$ samples $(a, b) \in R_q \times R_q$ are similarly obtained by sampling $a \leftarrow R_q$ uniformly and $e \leftarrow \psi$, then setting $b = a \cdot s + e \mod qR$; and where RLWR samples just round off the lower-order bits of the coefficients of ring elements, written as polynomials in the "power basis" representation for R.

In later work, Alwen et al. [AKPW13] gave the first reduction from LWE to LWR in the case when the modulus q is *a fixed polynomial* in the security parameter κ. However to do so, their techniques introduce a number of drawbacks not originally present in the [BPR12] reduction; namely: *(i)* restricting to only the case of LWE/LWR (rather than RLWE/RLWR as well), *(ii)* additionally *a-priori bounding the number of* LWE/LWR *samples, w,* for the proof to proceed, and *(iii)* losing samples during the course of the reduction. More precisely, they show the hardness of LWR with parameters d, w, q, p from the hardness of LWE with parameters d', w, q, ψ with ψ B-bounded, so long as $d > \log(q)/\log(\gamma) \cdot d'$ and $q \geq \gamma(dwBp)$ for some flexible parameter $\gamma \geq 1$., with an additional constraint that q must be prime or have its largest prime factor greater than or equal to $\gamma(dwBp)$.

Additionally, Alwen et al. demonstrate applications of "w-bounded-sample" LWR to constructing reusable computational extractors [DKL09], lossy trapdoor function families [PW11], and deterministic encryption [BBO07, BFOR08, BFO08, BS14, FOR15]. In particular, any improvements in their reduction from LWE to LWR immediately imply improvements to the security or efficiency of

these constructions, though – unfortunately – not to lattice-based pseudorandom function constructions, which (so far) only make use of unbounded-sample LWR for their security proofs.

More recently, Bogdanov et al. [BGM+16] extended this line of work in a number of ways, primarily through the novel introduction of Rényi divergence (rather than statistical distance) in order to fine-tune their statistical analyses. Of particular relevance to our work is their reduction to $\mathsf{LWR}_{d,w,q,p}$ from $\mathsf{LWE}_{d',w,q,\psi}$ with parameters d', w, q, ψ so long as $d > \log(q) \cdot d'$ and $q \geq 2wBp$, with no special conditions on the factorization of q, and it appears that if all of the prime factors of q are at least ρ, they only require that $d > \log_\rho(q) \cdot d'$, making it dimension preserving for the special case that q is prime. They are also able to give a reduction from the *search* version of RLWE to the *search* version of RLWR, but are unable to prove any search-to-decision reduction for RLWR. It is an open problem to recover all of the tighter parameters of [BGM+16] without losing in the dimension, perhaps by a applying Rényi divergence based analysis rather than our forthcoming use of statistical distance.

1.1 Our Contributions

Our main contribution is a dimension-preserving reduction from LWE to LWR with a polynomial-sized modulus. The two somewhat incomparable reductions described in the theorem stem from the two somewhat incomparable pre-existing reductions from LWE to Ext–LWE; see 3 for details.

Theorem 1.1. *Let κ be a security parameter on which all other parameters depend. Let ψ be a distribution over $\mathbb{Z}$ for some $B > 0$. Let $p, q = \mathrm{poly}(\kappa), w = \mathrm{poly}(\kappa), d \in \mathbb{N}$ such that $q \geq 4eBwp\kappa$. If $\exists$ a probabilistic polynomial-time $\mathcal{A}$ succeeding with advantage $\epsilon(\kappa) \geq (\kappa)^{-c}$ for some constant $c \geq 1$ in distinguishing $\mathsf{LWR}_{d,w,q,p}$ from uniform, then there exists a probabilistic polynomial-time algorithm $\mathcal{A}'$*

1. *succeeding with advantage $\epsilon(wB)^{-c}/4 \geq (\kappa wB)^{-c}/4$ in distinguishing $\mathsf{LWE}_{d,w,q,\psi}$ from uniform, as long as every prime factor of q is greater than B and ψ is B-bounded.*
2. *succeeding with advantage $\epsilon(w)^{-c}/4 \geq (\kappa w)^{-c}/4$ in distinguishing $\mathsf{LWE}_{d-c,w,q,\psi}$ from uniform, with no restrictions on the factors of q, as long as $\psi = D_\alpha$ and $D_{(a^2+\omega(\log\kappa))}$ is B-bounded.*

In Table 1, we give a comparison to the previous reductions, omitting constants, negligible amounts, and $O(\cdot)$ symbols for readability.

Extension to Ideal and Module Lattices. Although the above reduction is written in terms of LWE, as it turns out, we are able to prove essentially the same reduction in the ring setting. In particular, we are able to reduce Ext–RLWE to RLWR in a dimension-preserving manner, where here the hints for Ext–RLWE can be arbitrary $\mathbb{Q}$-linear functions of the coefficient vector. To avoid having to repeat what is essentially the same reduction for both settings, we instead prove

Work	**Unbounded Samples** (w)	**Modulus** (q)	**Advantage Change** ($\epsilon \to \epsilon'$)	**Dimension Change** ($d \to d'$)
[BPR12]	Yes	$Bp\kappa^{\omega(1)}$	$\epsilon - \mathrm{negl}(\kappa)$	d
[AKPW13]	No	$\gamma Bwp\kappa$	$\epsilon/(2dw)$	$d\log(\gamma)/\log q$
[BGM+16]	No	Bwp	$(\epsilon/qw)^2$	$d/\log_\rho q$
This work (1)	No	$Bwp\kappa$	$\epsilon(wB)^{-c}$	d
This work (2)	No	$Bwp\kappa$	$\epsilon(w)^{-c}$	$d-c$

Table 1. In this table, we have $c \leq -\ln(\epsilon)/\ln(\kappa)$, $\gamma \geq 1$, every prime factor of q is greater than ρ.

the reduction for the learning with errors problem over module lattices. However, we have written the reduction in such a way that someone who is only familiar with LWE over integer lattices can easily follow it, by simply replacing R with $\mathbb{Z}$.

We are unable to prove any reductions from RLWE to the form of Ext–RLWE described above, and leave this as a major open problem. We believe that it *should* be possible without too much difficulty to prove a reduction from RLWE to an alternative form of Ext–RLWE. Specifically, in this version one receives the error modulo some prime ideal of qR as the "hint," and the proof would work for those choices of q and R such that q factors over R into the product of prime ideals all having small norm (i.e. polynomial in the security parameter). Briefly, such a reduction would follow the reduction from LWE to Ext–LWE by Alperin-Sheriff and Peikert, [AP12] where the reduction guesses the value of the error term modulo the requested prime ideal, and then modifies the RLWE samples based on the guess so that a correct guess results in valid RLWE samples and an incorrect guess results in uniform samples. Some care would also have to be taken to show that the resulting samples are in fact uniform, using some of the techniques in the search-to-decision reduction for RLWE in [LPR13a].

However, such a reduction does not appear to be useful for our goal of getting a reduction for RLWR, at least in a manner similar to our reduction for regular LWR, so we omit a formal proof. The problem here is that the RLWE error will necessarily be statistically uniform modulo any small prime ideal, so we cannot simply round each coefficient of an LWE sample in the "CRT representation" and have any useful bound on the likelihood that the coefficient will round to an equivalent value in the errorless version of the sample. Conversely, in the coefficient representation and/or the canonical embedding, we will be able to obtain useful bounds on these likelihoods, but learning the error modulo a prime ideal will be entirely useless for dealing with the "bad" coefficients.

Instead, we define extended-LWE over the generalization of ideal lattices (and general lattices) that have been termed *module lattices* [LS15]. Over module lattices, we are able to show a reduction from M–LWE to Ext–M–LWE. This reduction is a fairly straightforward adaptation of the LWE to Ext–LWE reduction that can be found in [BLP+13], but requires some care since we are working over rings that in general fail to be principal ideal domains.

1.2 Our Approach

The connection given in each of the three previous works between LWE and LWR ultimately stems from the fact that, except for some "bad" cases, we have that

$$\lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p = \lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rceil_p \pmod{q},$$

when e is drawn from a B-bounded distribution and q is sufficiently larger than Bp. In Banerjee et al's original work [BPR12], they simply set q to be superpolynomially large than Bp, so that the above equation holds with all but negligible probability, allowing an unbounded polynomial number of samples. As in the subsequent two works [AKPW13, BGM+16], we provide a more sophisticated bound for the case that the number of LWR samples received w is an a priori bounded polynomial in the underlying security parameter κ.

The first insight into our reduction comes from noticing that we can tune the number of "bad" samples allowed to the advantage $\epsilon \geq \kappa^{-c}$ (where c is some constant independent of κ) of a given adversary $\mathcal{A}$ in attacking LWR. In this case, by setting $q \geq Cw\kappa Bp$ for a suitable small constant C, we can easily show that the probability that $\mathbf{b}$ contains more than c "bad" elements is at most $\frac{1}{2}\kappa^{-c}$. Consequently, if we could find a method to transform any LWE sample with at most c "bad" elements into an LWR sample, then we could construct $\mathcal{A}'$ that could use $\mathcal{A}$ to successfully attack LWE in the *same dimension*. Specifically, $\mathcal{A}'$ would work as follows:

1. $\mathcal{A}'$ queries its LWE oracle to receive back ($\mathbf{A} \in \mathbb{Z}_q^{d \times w}, \mathbf{b}$). If $\mathbf{b}$ contains c or more bad elements, $\mathcal{A}'$ aborts.
2. ??????
3. $\mathcal{A}'$ sends $\mathcal{A}$ a sample ($\mathbf{A} \in \mathbb{Z}_q^{d \times w}, \mathbf{b} \in \mathbb{Z}_p^w$) within statistical distance $\epsilon/2$ of LWR.
4. Profit(able reduction)!

Of course, this reduction has a missing step-how do we successfully transform an LWE sample with (at most) c bad elements into a valid LWR sample? Rounding the "good" elements properly is easy, as we know that for those elements, $\lfloor b_i \rceil_p = \lfloor \langle \mathbf{a}_i, \mathbf{s} \rangle \rceil_p$. But what about the "bad" elements, where it may be the case that $\lfloor b_i \rceil_p \neq \lfloor \langle \mathbf{a}_i, \mathbf{s} \rangle \rceil_p$. It turns out that there is a very easy solution: *we guess them*!

More formally, instead of show a reduction from LWE to LWR, we instead show a reduction from extended-LWE (Ext–LWE) (with c hints) to LWR, and then rely on the existing reductions from LWE to Ext–LWE to achieve a full reduction from LWE to LWR. The Ext–LWE$_{d,w,q,\alpha,c}$ problem allows one to query for the result of c linear functions of the error vector $\mathbf{e}$ to be received along with the LWE sample ($\mathbf{A}, \mathbf{b}$). In particular, we can query to learn arbitrary elements e_i of the error vector.

As a result, in order to round the "bad" elements properly, we simply make a uniform random guess (in advance) which c elements of $\mathbf{b}$ might be "bad." With probability at least $w^{-c} \geq 1/\operatorname{poly}(\kappa)$, we will guess correctly, in which case

we can simply subtract off the associated terms and then correctly round the result. If we guess incorrectly, we abort. Note that this step requires an artificial abort on correct guesses to ensure that the sample getting sent to $\mathcal{A}$ remains distributed within $\epsilon/2$ of actual LWR; see the body of the paper for details.

Organization. In Section 2 we recall various preliminary facts that we will need for the rest of the paper. In Section 3, we extend the definition of extended LWE to module lattices, and prove a reduction from M–LWE to Ext–LWE; this section is standalone and may be skipped for those unfamiliar with the ring setting. Finally, in Section 4, we give a dimension-preserving reduction from Ext–M–LWE to M–LWR. We stress that this section has been written to be accessible to someone only familiar with integer lattices, by simply viewing R as $\mathbb{Z}$ when reading it, and noting that this makes the ring dimension parameter $n = 1$.

2 Preliminaries

2.1 Notation

Throughout this work, by "ring" we mean a commutative ring with identity. We identify the elements in $\mathbb{Z}_q$ with their coset representatives in $[0, q)$. For an element $x \in Z_q$, we define

$$\lfloor x \rceil_p = \lfloor (p/q) \cdot x \rceil,$$

where the latter $\lfloor \cdot \rceil$ denotes standard rounding to the nearest integer. For a ring element $y \in R_q$, the rounding operation is performed coefficient-wise.

Note that we use d and w as matrix dimensional parameters instead of n and m used in most LWE papers, because we reserve the latter variables for the dimension and index of cyclotomic rings, respectively. We (sometimes implicitly) parameterize essentially all variables asymptotically in terms of κ instead of n, since we use n to denote the dimension of a ring and it may be small; in particular, for $\mathbb{Z}$, we have $n = 1$.

2.2 Algebraic Number Theory Background

Here we briefly review concepts from algebraic number theory necessary for (parts) of our work. For a more thorough background on algebraic number theory as it applies to cryptography based on ideal lattices, see [LPR13a, LPR13b].

The Space H. When working with cyclotomic fields under the canonical embedding, it is standard to work with the subspace $H \subseteq \mathbb{C}^n$, where $2s_2 = n$ is the dimension of the field (the $2s_2$ notation is to account for complex conjugates), defined as

$$H = \{(x_1, \ldots, x_n) \in \mathbb{C}^n : x_{s_2+j} = \bar{x_j}, \forall j \in [s_2]\} \subseteq \mathbb{C}^n.$$

Letting $\mathbf{e}_j \in C^n$ be the vector with 1 in its jth coordinate, 0 elsewhere, we have the following orthonormal basis for H. We let $\mathbf{h}_j = \frac{1}{\sqrt{2}}(\mathbf{e}_j + \mathbf{e}_{j+s_2})$ and $\mathbf{h}_{j+s_2} = \frac{\sqrt{-1}}{\sqrt{2}}(\mathbf{e}_j - \mathbf{e}_{j+s_2})$. This basis also shows that H is isomorphic to $\mathbb{R}^n$ as an inner product space.

Cyclotomic Fields and Rings. For a positive integer m called the *index*, the mth *cyclotomic number field* is $K = \mathbb{Q}(\zeta_m)$, where ζ_m is some fixed arbitrary primitive mth root of unity (for each m, we view it abstractly and not as any particular such root). We denote the ring of integers of K as $\mathcal{O}_m = \mathbb{Z}[\zeta_m]$, and refer to it as the mth cyclotomic ring. The minimal polynomial of ζ_m over $\mathbb{Q}$ is the mth *cyclotomic polynomial* $\Phi_m(X) = \prod_{i \in \mathbb{Z}_m^*}(X - \omega_m^i) \in \mathbb{Z}[X]$, where $\omega_m = \exp(2\pi\sqrt{-1}/m) \in \mathbb{C}$ is the principal mth complex root of unity, and the roots $\omega_m^i \in \mathbb{C}$ range over all the *primitive* complex mth roots of unity. Therefore, $\mathcal{O}_m$ is a ring extension of degree $n = \varphi(m)$ over $\mathbb{Z}$. (In particular, $\mathcal{O}_1 = \mathcal{O}_2 = \mathbb{Z}$.) Clearly, $\mathcal{O}_m$ is isomorphic to the polynomial ring $\mathbb{Z}[X]/\Phi_m(X)$ by identifying ζ_m with X, and has the "power basis" $\{1, \zeta_m, \ldots, \zeta_m^{n-1}\}$ as a $\mathbb{Z}$-basis. Note that for $m \in \{1, 2\}$, we have that $K = \mathbb{Q}$ and $R = \mathbb{Z}$.

Complex Embeddings. A cyclotomic number field $K = \mathbb{Q}(\zeta)$ degree n has exactly n ring embeddings $\sigma_i : K \to \mathbb{C}$, which can be defined by $\sigma_i(\zeta) = \zeta_i$ for $i \in \mathbb{Z}_m^*$. The *canonical embedding* is then defined as

$$\sigma(x) = (\sigma_1(x), \ldots, \sigma_{m-1}(x)) \in \mathbb{C}^n.$$

We also define the trace $\mathrm{Tr} : K \to \mathbb{Q}$ and the (field) norm $N : K \to \mathbb{Q}$ as $\mathrm{Tr}(x) = \sum \sigma_j(x)$, $N(x) = \prod_j \sigma_j(x)$, and recall that the trace is a "universal" $\mathbb{Q}$-linear function, in the sense that any linear function $L : K \to \mathbb{Q}$ can be expressed as $L(a) = \mathrm{Tr}(r \cdot a)$ for some fixed $r \in K$.

Ideals. An (integral) *ideal* I of R is an additive subgroup of R that is closed under multiplication by every element of R. We denote the smallest ideal of R containing the set S by (S), and by R/I the set of equivalence classes $x + I$ of R modulo I. The norm of a (nonzero) ideal is the number of elements in R/I.

For ideals I and J of a ring R, their sum $I + J = \{i + j : i \in I, j \in J\}$, and their product is $IJ = \{\sum_k i_k j_k : i_k \in I, j_k \in J\}$. A *prime ideal* $I \subseteq R$ is such that if $ab \in I$, then at least one of a and b is also in I. Over the ring of integers of any algebraic number field, any ideal of R can be represented uniquely as a product of prime ideals. A fractional ideal $I \subseteq K$ is a set such that $dI \subseteq R$ is an integral ideal for some non-zero $d \in R$. The fractional ideals of K form a group under multiplication, and in particular we have that $I \cdot I^{-1} = R$.

Duality. The *dual* of an ideal is defined as $I^\vee = \{x \in K : \mathrm{Tr}(xI) \subseteq \mathbb{Z}\}$, and we have that $I^\vee = I^{-1} \cdot R^\vee$. If R is the $m = 2^{e_0} p_1^{e_1} \ldots p_t^{e_t}$th cyclotomic ring, let $g = \prod_{i \in [t]}(1 - \zeta_{p_i}) \in R$ and let $\hat{m} = m/2$ is m is even, $\hat{m} = m$ otherwise. Then $R^\vee = \langle g/\hat{m} \rangle$. As a result, for an element $x \in R^\vee$, we have that $\hat{m}x \in R$, and that each coefficient of $\hat{m}x$ is scaled up by exactly $\hat{m}$. We could also scale by $\hat{m}/g$ to move it into R, but this would eliminate any nice guarantees on the change in the magnitudes of each coeficient.

Factorization of Ideals. Let $q = p^r \in \mathbb{Z}$ be a prime power. In the mth cyclotomic ring $R = \mathcal{O}_m = \mathbb{Z}[\zeta_m]$ (which has degree $n = \varphi(m)$ over $\mathbb{Z}$), the ideal pR

factors into prime ideals as follows. First write $m = \bar{m} \cdot p^k$ where $p \nmid \bar{m}$. Let $e = \varphi(p^k)$, and let d be the multiplicative order of p in $\mathbb{Z}^*_{\bar{m}}$. Note that d divides $\varphi(\bar{m}) = n/e$. The ideal qR then factors into the product of (re)th powers of $\varphi(\bar{m})/d = n/(de)$ distinct prime ideals $\mathfrak{p}_i$, i.e. $qR = \prod \mathfrak{p}_i^{re}$. Each prime ideal $\mathfrak{p}_i$ has norm $|R/\mathfrak{p}_i| = p^d$. The factorization of qR for general $q = p_1^{r_1} \dots p_t^{r_t}$ then follows by recalling the unique factorization of ideals for cyclotomic rings and that $qR = (p_1^{r_1} R) \dots (p_t^{r_t} R)$.

Modules. A subset $M \subseteq K^d$ is an R-module if it closed under addition and multiplication by elements of R. When K is a number field and R is its ring of integers, an R-module has a so-called *pseudo-basis* of linearly independent vectors $\mathbf{b}_i \in K^d$ such that $M = \sum_{i=1}^{d} I_i \cdot \mathbf{b}_i$ for some non-zero ideals I_k of R. The representation of element of M with respect to a pseudo-basis unique, but two pseudo-bases can generate the same module; however, these bases will always have the same rank.

2.3 Lattices, Ideal Lattices and Module Lattices

A lattice $\mathcal{L} \subseteq R^n$ is a discrete additive subgroup consisting of all integer linear combinations $\sum_{i \in m} x_i \mathbf{b}_i$, where $B = (\mathbf{b}_i)$ is called the basis.

Under the canonical embedding, a fractional ideal I with $\mathbb{Z}$-basis $U = (u_1, \dots, u_n)$ becomes a rank-n *ideal lattice* $\sigma(I)$ with basis $\{\sigma(u_1), \dots, \sigma(u_n)\} \subset H$. For convenience, we often identify the ideal I with its embedded lattice.

Module lattices were first implicitly used for cryptography in [BGV12], but were first formally investigated by Langlois and Stehlé [LS15]. For an n-dimensional number field, they can be defined via the embedding $\sigma_M : K^d \to \mathbb{R}^{nd}$ defined by $\sigma_M = (\sigma)_{i \in d}$, where σ is the canonical embedding. For a module lattice M, the set $\sigma_M(M)$ is a module lattice, and similarly to ideal lattices, we often identify the module M with its embedded lattice. One can see them as a generalization of both lattices and ideal lattices. In particular, if $K = \mathbb{Q}$, then it is easy to see that this defines an arbitrary integer lattice, while if we set $d = 1$, we end up with ideal lattices.

As mentioned by Langlois and Stehlé, for an $O(n)$-dimensional cyclotomic ring, $1 < d \ll n$, cryptography based on module lattices maintain much of the practical advantages over cryptography based on general lattices, while enjoying potentially stronger theoretical security guarantees. In particular, for $d = 1$ (corresponding to ideal lattices), the Minkowski upper bound on the length of the shortest vector is within a $\sqrt{n}$ factor of the actual length of the shortest vector in the lattice [LPR13a], which makes $\mathsf{GapSVP}_{\sqrt{n}}$ for cyclotomic rings into an easy problem. By contrast, for $d \geq 2$, it is an easy exercise to show that there exist module lattices with shortest vector shorter than Minkowski's upper bound by any arbitrarily large factor, and more generally, in this case no polynomial-time algorithm is known for approximating Mod–GapSVP (i.e. GapSVP over module lattices) to within any polynomial factor. As a result, we may apply Peikert's classical reduction from GapSVP to BDD in the module lattice setting [Pei09], using the oracle reducing Mod–BDD to M–LWE found

in [LS15] and gain potentially more confidence in the hardness of M–LWE (for $d \geq 2$) than we can in the hardness of RLWE, for which there only exists a *quantum* reduction from plausibly hard worst-case lattice problems.

2.4 Gaussian Measures

For $s > 0$, the n-dimensional Gaussian function ρ_s is defined as

$$\rho_s(\mathbf{x}) := \exp(-\pi \|\mathbf{x}\|^2 / s^2).$$

Normalizing this function gives the *continuous* Gaussian distribution D_s. More generally, we can define $D_{\mathbf{B}}$ as the distribution of $\mathbf{Bx}$ where $\mathbf{x}$ is sample from D_1. For an invertible $\mathbf{B}$, $D_{\mathbf{B}}$ is proportional to $\exp(-\pi \mathbf{x}^t \mathbf{B}^{-t} \mathbf{B}^{-1} \mathbf{x})$. For any $\mathbf{B}_1, \mathbf{B}_2$, the sum of a sample from $D_{\mathbf{B}_1}$ and $D_{\mathbf{B}_2}$ is distributed as $D_{(\mathbf{B}_1 \mathbf{B}_1^t + \mathbf{B}_2 \mathbf{B}_2^t)^{1/2}}$.

For an n-dimensional lattice Λ and a vector $\mathbf{u} \in \mathbb{R}^n$, we define the *discrete Gaussian distribution* $D_{\Lambda + \mathbf{u}, s}$ as the discrete distribution with support on the coset $\Lambda + \mathbf{u}$ whose probability mass function is proportional to ρ_s.

We require the following now standard facts about Gaussian distributions over lattices [MR07, GPV08, BLP+13].

Lemma 2.1. *Let $\mathcal{L}$ be a lattice with associated basis $\mathbf{B}$, and let $\|\tilde{\mathbf{B}}\|$ denote the length of the longest vector in the Gram-Schmidt orthogonalization of $\mathbf{B}$. Let $s \geq \|\tilde{\mathbf{B}}\| \sqrt{\ln(2n(1 + 1/\epsilon))/\pi}$ for some $0 < \epsilon \leq 1/2$. Then*

1. *There is a probabilistic polynomial-time algorithm that, given $\mathbf{c} \in \mathbb{R}^n$ outputs a sample distributed according to $D_{\Lambda + \mathbf{c}, s}$.*
2. *The distribution of $\mathbf{x} \bmod \Lambda$, where $\mathbf{x} \leftarrow D_s$, is within statistical distance $\epsilon/2$ of the uniform distribution over cosets of Λ.*
3. *Let $r > 0$. Then if we choose $\mathbf{x} \leftarrow D_r$ and then choose $\mathbf{y} \leftarrow D_{\Lambda - \mathbf{x}, s}$, we have that $\mathbf{x} + \mathbf{y}$ is within statistical distance 8ϵ of the discrete Gaussian $D_{\Lambda, (r^2 + s^2)^{1/2}}$.*

2.5 Learning with Errors

Here we formally define the learning with errors problem [Reg09] and its variants over rings and modules. As alluded to earlier, it is sufficient to simply define module learning with errors (M–LWE) and module learning with rounding (M–LWR), as the ring and g eneral variants fall out as special cases.

Definition 2.2 (M–LWE Problem). *Let K be some finitely generated field extension of $\mathbb{Q}$, and denote by n its dimension. Let $R = \mathcal{O}_K$ be its ring of integers, and let ψ be some distribution over R. Then we define the decision version of $\mathsf{M\text{–}LWE}_{R,d,w,q,\psi}$ as follows.*

The adversary gets $(\mathbf{A} \leftarrow R_q^{d \times w}, \mathbf{b} \in R_q^d)$, and must distinguish between the case where $\mathbf{b} = \mathbf{A}^t \mathbf{s} + \mathbf{e} \in R_q^w$, where $\mathbf{s} \leftarrow R_q^d$ uniformly at random and $\mathbf{e} \leftarrow \psi^w$ and the case where $\mathbf{b} \leftarrow R_q^w$ uniformly at random.

When ψ is a spherical Gaussian, we sometimes write $\mathsf{M\text{–}LWE}_{R,d,w,q,\alpha}$ to denote $\mathsf{M\text{–}LWE}_{R,d,w,q,D_\alpha}$, and similarly for LWE and RLWE.

Under this definition, it is easy to see that M–LWE is a generalization of both RLWE and LWE. When $d = 1$, $\mathsf{M\text{–}LWE}_{R,d,w,q,\psi}$ is equivalent to $\mathsf{RLWE}_{R,w,q,\psi}$, while when $R = \mathbb{Z}$, $\mathsf{M\text{–}LWE}_{R,d,w,q,\psi}$ is equivalent to $\mathsf{LWE}_{d,w,q,\psi}$.

Unlike the standard definition, the error and the secret are chosen from a distribution over the dual ideal of the ring $R_q^\vee$. However, as discussed above in Section 2.2, we can simply scale up $\mathbf{b}$ by $\hat{m}$ to move it into R, simplifying notation and making the paper more accessible.

Hardness Guarantees. We have quantum worst-case to average case reductions from SIVP for module lattices for $d \geq 1$, and we also have a (potentially meaningful) classical reduction from GapSVP for $d \geq 2$ [LS15]. There are also many different hardness reductions for the special cases of LWE and RLWE [Pei09, Reg09, LPR13a, BLP+13]; for details, consult the cited works.

Error Distributions. For the purposes of this paper, we will largely ignore the specific details of the error distribution, except for in Section 3.2 (which can be viewed in a standalone manner), where we use Gaussian distributions of the form described in Section 2.4. For our main result, it is sufficient that the error distribution is B-bounded.

Definition 2.3. *A distribution ψ over $\mathbb{Z}$ is said to be* B-bounded *if*

$$\Pr_{x \leftarrow \psi}[x \notin [-B, B]] \leq \mathrm{negl}(\kappa).$$

For completeness, we recall the following result of Banszcyzyk [Ban93], which allows us to upper-bound the magnitude of the error coefficients in an LWE sample.

Lemma 2.4 (Adaptation of [Ban93] Lemma 1.5(i)). *For any lattice Λ, $r > 0$, and let $\mathbf{x} \leftarrow D_{\Lambda,r}$ we have that except with probability 2^{-2n}, $\|\mathbf{x}\| \leq r\sqrt{n}$.*

We caution that this result cannot necessarily be used immediately as an upper-bound of (ring) coefficients in RLWE and M–LWE if one wishes to rely on the existing worst-case reductions from hard ideal/module lattice problems. This is because we need bounds on the magnitude of the individual coefficients of the error term, while the error is sampled to be short with respect to the canonical embedding. As in general (specifically, for cyclotomic polynomial rings with an index that is the product of many distinct small primes), the distortion between the coefficient embedding and the canonical embedding may be superpolynomially large in the index [Erd46], these reductions may be of less utility in such rings. However, in the rings of index 2^k (and more generally for rings of index $2^k p^\ell$, where p is a small odd prime), the distortion between the two embeddings will be very small, and so relying on the worst-case reductions becomes meaningful.

2.6 Learning With Rounding

The learning with rounding (LWR) problem was introduced by Banerjee, Peikert and Rosen as a "derandomization" of LWE [BPR12]. As with learning with errors, we will simply define the module version (M–LWR), and note that RLWR and LWR fall out as special cases.

Definition 2.5 (M–LWR Problem). *Let K be an algebraic number field of dimension n, $R = \mathcal{O}_K$ its ring of integers. Let $q \geq p \geq 1, d \geq 1 \in \mathbb{Z}$ be integers. Then we define the decision version of $\mathsf{M}\text{–}\mathsf{LWR}_{R,d,w,q,p}$ as follows: The adversary gets $(\mathbf{A} \leftarrow R_q^{d \times w}, \mathbf{b} \in R_p^w)$, and must distinguish between two cases. In the first case, which we refer to as the M–LWR distribution,*

$$\mathbf{b} = \lfloor \mathbf{A}^t \mathbf{s} \rceil_p = \lfloor (p/q)\mathbf{A}^t \mathbf{s} \rceil \in R_p^w,$$

where $\mathbf{s} \leftarrow R_q^d$ uniformly at random, while in the second case, which we refer to use at the uniform case, $\mathbf{b} = \lfloor \mathbf{u} \rceil_p$, where $\mathbf{u} \leftarrow R_q^w$ uniformly at random.

Note that if p divides q, the distribution of $\mathbf{b}$ is truly uniform, while if $p \not| q$, $\mathbf{b}$ will be slightly biased towards certain values modulo p.

3 Extended-LWE

In this section we define Ext–LWE and its variants, and give a security reduction from M–LWE to Ext–M–LWE for suitable parameters.

3.1 Definitions

The Extended-LWE (Ext–LWE) problem was first introduced formally by O'Neill et al. [OPW11], although it appeared implicitly in several earlier works. We provide a definition of the problem in the context of module lattices (Ext–M–LWE), following the definition of Brakerski et al. for regular Ext–LWE [BLP+13].

Definition 3.1. *For a number field K with ring of integers R, dimensional parameters $d, w \in \mathbb{Z}$, modulus $q \geq 1$, a set $\mathcal{Z} \subseteq K^d$, a parameter k denoting the number of hints received, and a distribution ψ over R, the $\mathsf{Ext}\text{–}\mathsf{M}\text{–}\mathsf{LWE}_{R,d,w,q,\psi,\mathcal{Z},k}$ problem is defined as follows: The algorithm gets to choose $\mathbf{z}_1, \ldots, \mathbf{z}_k \in \mathcal{Z}$ and receives back a tuple*

$$(\mathbf{A}, \mathbf{b}, (\mathrm{Tr}(\langle \mathbf{e}, \mathbf{z}_i \rangle))_{i \in [k]}) \in R_q^{d \times w} \times R_q^w \times \mathbb{Q}_q^k$$

The goal is to distinguish between two cases. In the first, which we refer to as the Ext–M–LWE distribution, $\mathbf{A}$ is chosen uniformly at random, $\mathbf{e} \leftarrow \psi^w$, and $\mathbf{b} = \mathbf{A}^t \mathbf{s} + \mathbf{e} \pmod q$ for $\mathbf{s}$ chosen uniformly at random. The second case, which we refer to as the uniform distirbution, is identical, except that $\mathbf{b}$ is chosen uniformly at random and independently of everything else. When $\psi = D_\alpha$, we sometimes write $\mathsf{Ext}\text{–}\mathsf{M}\text{–}\mathsf{LWE}_{R,d,w,q,\alpha,\mathcal{Z},k}$ to mean $\mathsf{Ext}\text{–}\mathsf{M}\text{–}\mathsf{LWE}_{R,d,w,q,D_\alpha,\mathcal{Z},k}$.

An explanation is in order for the use of the trace function in the definition of the hints, which is not found in previous definitions of Ext–LWE (and indeed has no meaning over general lattices, as $\mathrm{Tr}_{\mathbb{Q}/\mathbb{Q}}$ is the identity function). The purpose of including the trace function is to provide a method to request any $\mathbb{Q}$-linear function of the error vector, as viewed in the coefficient embedding, as the hints. The trace function, being a "universal" $\mathbb{Q}$-linear function, allows us to accomplish this.

In our proof below, we do not gain anything by limiting to extracting only the trace of the inner product instead of allowing extraction the entire inner product $\langle \mathbf{e}, \mathbf{z}_i \rangle$. However, the latter case clearly reveals strictly less information (given the hint alone) than the former, so there may exist an alternative reduction for the latter case that is less costly in parameters than the former. Moreover, using the full trace allows for a nicer reduction to M–LWR in the next section, as we end up only requesting the error coefficients we need instead of needing to request entire ring elements.

Following Brakerski, et al, we define the *quality* of a set $\mathcal{Z} \subseteq \mathbb{Z}^m$.

Definition 3.2. *For a real $\xi > 0$ and a set $\mathcal{Z} \subseteq K^w$, we say that $\mathcal{Z}$ is of* quality (ξ, t) *if given any $\mathbf{z}_1, \ldots, \mathbf{z}_t \in \mathcal{Z}$, we can efficiently construct a unimodular matrix $\mathbf{U} \in \mathbb{Z}^{w \times w}$ such that, if $\mathbf{U}' \in \mathbb{Z}^{w \times (w-t)}$ is the matrix obtained from $\mathbf{U}$ be removing its t leftmost columns, then all the columns of $\mathbf{U}'$ are orthogonal to span$(\mathbf{z}_1, \ldots, \mathbf{z}_t)$ and its largest singular value is at most ξ.*

For our reduction in Section 4, we will have $\mathcal{Z}$ be the set of (scaled down by the dimension of K) columns of the identity matrix $\mathbf{I}_w$. We can easily see that $\mathcal{Z}$ is of quality $(1, t)$ for all $t < w$ by letting $\mathbf{U}$ be a suitable permutation of the rows of the identity matrix.

For this particular setting of $\mathcal{Z}$, there exist two prior somewhat incomparable reductions for Ext–LWE (over general lattices), given by Alperin-Sheriff and Peikert [AP12] and by Brakersi et al., respectively. [BLP+13]. We state them both here:

Lemma 3.3. *There is a (transformation) reduction from $\mathsf{LWE}_{d,w,q,\alpha}$ to*

1. *[AP12]: $\mathsf{Ext\text{–}LWE}_{d,w,q,\alpha,k}$, that reduces the advantage by a multiplicative factor of at most q^k.*
2. *[BLP+13]: $\mathsf{Ext\text{–}LWE}_{d+k,w,q,(\alpha^2+r^2)^{1/2},k}$, where $r \geq \omega(\sqrt{\log w})$, that reduces the advantage by a negligible factor.*

3.2 Extended-LWE Over Module Lattices

Here we show that an analogous version of the Ext–LWE reduction found in [BLP+13] holds over module lattices as well. In order to prove this theorem, we follow the path of Brakerski et al. [BLP+13]. We first prove the reduction to the intermediate problem "first-is-errorless-M-LWE" and then proceed to prove the reduction from "first-is-errorless-M-LWE." to Ext–M–LWE.

First-Is-Errorless-LWE As in the previous work, we have the intermediate step of "first-is-errorless" M–LWE, in which the first inner product contains no error. The proof for this general case essentially follows that of Brakerski et al, with a few important differences stemming from the form of R will not in general be a Euclidean domain or even a principal ideal domain. This requires some careful proof restructuring.

Definition 3.4. *Let K be an algebraic number of field, $R = \mathcal{O}_K$ be its ring of integers. Let $q > 1, d \geq 2$ be integers, and let ψ be a distribution over R. Then the "first-is-errorless" version of* M–LWE *requires the adversary to distinguish between two cases. In the first case, the adversary receives $\mathbf{A} \in R_q^{d \times w}, \mathbf{b} = \mathbf{A}^t\mathbf{s} + \mathbf{e} \in R_q^w$, where $\mathbf{s} \leftarrow R_q^d$ uniformly at random and $\mathbf{e} \leftarrow (0, \mathbf{e}')$, where $\mathbf{e}' \leftarrow \psi^{w-1}$. In the second case, $\mathbf{b} \leftarrow R_q^m$ uniformly at random.*

We first prove a brief lemma relating to the probability that the greatest common divisor of qR and the principal ideals generated by d elements chosen uniformly at random modulo R_q is in fact all of R (making all of these ideals coprime).

Lemma 3.5. *Let $q = \prod_{i \in [\ell]} p_i^{e_i}$. Let $R = \mathcal{O}_m$ and write $m = \bar{m} \cdot p_i^{\kappa_i}$, where $p \nmid \bar{m}$. Let f_i be the multiplicative order of p_i modulo $\bar{m}$. Let $a_1, \ldots, a_d \leftarrow R_q$ uniformly at random (strictly speaking, elements in R from some canonical set of representatives of R_q). Then we have that $R = \gcd(qR, \langle a_1 \rangle, \ldots, \langle a_d \rangle) = qR + \langle a_1 \rangle + \ldots + \langle a_d \rangle$, except with probability at most*

$$\sum_{i \in \ell} \left(\frac{\varphi(\bar{m})}{f_i} p_i^{-d \cdot f_i} \right) \tag{3.1}$$

Proof. We have that the sum of all the ideals above will be equal to R unless $a_1, \ldots, a_d$ all lie in some prime ideal $\mathfrak{p}_i$ of qR. Recall from 2.2 that a prime ideal $\mathfrak{p}_{i,j}$ dividing $p_i R$ has norm $\|R/\mathfrak{p}_{i,j}\| = p_i^{f_i}$, the probability of this event is at most $p_i^{-d \cdot f_i}$. Equation 3.1 then follows from a simple union bound.

We now prove the reduction.

Lemma 3.6. *Let $d \geq 2$, w, $q > 1$, $R = \mathcal{O}_m$, ψ be an error distribution over R. There is an efficient (transformation) reduction from* M–LWE$_{R,d-1,w-1,q,\psi}$ *to the first-is-errorless variant of* M–LWE$_{R,d,w,q,\psi}$ *that reduces the advantage by at most $\sum_{i \in \ell} \left(\frac{\varphi(\bar{m})}{f_i} p_i^{-d \cdot f_i} \right)$, where f_i and $\bar{m}$ are as defined in Lemma 3.5.*

Proof. The reduction first chooses $\mathbf{a}' \leftarrow R_q^d$. If we have that $qR + \langle a_1' \rangle + \ldots + \langle a_d' \rangle \subset R$, the reduction aborts. Otherwise, it finds, via a generalization of the Euclidean algorithm [Coh96], a $d \times d$ invertible matrix $\mathbf{R} \in R_q^{d \times d}$ such that $\mathbf{R}\mathbf{a}' = (1, 0, \ldots, 0)^t$, and sets $\mathbf{U} = \mathbf{R}^{-1}$.

It also picks a uniform element $s_0 \in R_q$.It then output as the first sample $(\mathbf{a}', s_0)$. For each of the $w - 1$ remaining samples, we first sample $(\mathbf{a}, b)$ from the given M–LWE oracle, pick a fresh uniformly random $d \in R_q$, and output $(\mathbf{U}(d|\mathbf{a}), b + (s_0 \cdot d)$, where the vertical bar denotes concatenation.

Conditioned on not aborting, we will verify that the output is distributed according $\mathsf{M}\text{–}\mathsf{LWE}_{R,d,m,q,\psi}$. Combining this with the bound in Lemma 3.5 will give our desired result.

We now verify correctness, conditioned on not aborting. First, we consider the case that $(\mathbf{a}, b)$ samples are uniform and independent. In this case it is easy to see that the outputs remain uniform, since $(d|\mathbf{a})$ is uniform, $\mathbf{U}$ is invertible, and each b is uniform and independent of everything else. Secondly, if the samples are from $\mathsf{M}\text{–}\mathsf{LWE}$ with respect to a secret $\mathbf{s}$, it is easy to verify that the outputs are distributed according to the first-is-errorless variant $\mathsf{M}\text{–}\mathsf{LWE}_{R,d,w,q,\psi}$ with secret $\mathbf{s}' = \mathbf{R}^t(s_0|\mathbf{s}) \bmod q$. Since $\mathbf{U}$ is invertible modulo q and $(s_0|\mathbf{s})$ is uniform, so is $\mathbf{s}'$, which proves correctness.

By iterating this reduction and applying a union bound, we immediately have the following corollary which gives a reduction to a first-t-are-errorless variant of $\mathsf{M}\text{–}\mathsf{LWE}$. We use this corollary in the next section to reduce to $\mathsf{Ext}\text{–}\mathsf{M}\text{–}\mathsf{LWE}$ with t hints.

Corollary 3.7. *Let $d > t \geq 0$, w, $q > 1$, $R = \mathcal{O}_m$, ψ be an error distribution over R. There is an efficient (transformation) reduction from $\mathsf{M}\text{–}\mathsf{LWE}_{R,d-t,w-t,q,\psi}$ to the first-t-are-errorless variant of $\mathsf{M}\text{–}\mathsf{LWE}_{R,d,w,q,\psi}$ that reduces the advantage by at most $t\sum_{i\in\ell}\left(\frac{\varphi(\bar{m})}{f_i}p_i^{-(d-t)\cdot f_i}\right)$, where f_i and $\bar{m}$ are as defined in Lemma 3.5.*

In order for our reduction to "first-is-errorless" $\mathsf{M}\text{–}\mathsf{LWE}$ to be meaningful when $d-t$ is much smaller than linear in the underlying security parameter κ (as it will be in the case we are interested in for our reduction to $\mathsf{M}\text{–}\mathsf{LWR}$), we will require that all of the prime ideals of qR have norms superpolynomially large in the security parameter (which makes the reduction in advantage negligible). This is the case for many if not most practical choices of R and q. Similar conditions were required by Ducas and Micciancio [DM14] for the ring used in their signature scheme, although they were significantly further restricted because they needed the prime ideals to be exponentially large in the security parameter. As in their work, for a concrete example of choices of R and q for which this reduction is meaningful, we can take R to have index 2^k and let $q = 3^\ell$. In fact, for R of index 2^k, noting that $\mathbb{Z}_{2^k}^* \equiv \mathbb{Z}_2 \oplus \mathbb{Z}_{2^{k-2}}$, we have asymptotically by the prime number theorem for arithmetic progressions [Sop10] that $1 - 1/2^{k-1}$ of all primes (and their powers) are admissible as prime factors of the modulus q for which the reduction is meaningful.

First-Is-Errorless to Extended-LWE We now show a reduction from "first-t-are-errorless" $\mathsf{M}\text{–}\mathsf{LWE}$ to $\mathsf{Ext}\text{–}\mathsf{M}\text{–}\mathsf{LWE}$. The reduction primarily follows that of Brakerski et al, with slight differences to account for the multiple samples, as well as a correction of some small mistakes in the proof therein.

Lemma 3.8. *Let $\mathcal{Z} \subseteq K^d$ be of quality $(\xi > 0, t)$. Then for any $n, q > 1, \epsilon \in (0, 1/2)$, $\alpha, r \geq (\ln(2w(1 + 1/\epsilon))/\pi)^{1/2}/q$, $t < d$, there is a (transformation)*

reduction from the first-t-are-errorless variant of M–$\mathsf{LWE}_{R,d,w,q,\alpha}$ *to* Ext–M–$\mathsf{LWE}_{R,d,w,q,(\alpha^2\xi^2+r^2)^{1/2},\mathcal{Z},t}$ *that reduces the advantage by at most* 16ϵ.

Proof. The reduction proceeds as follows. Given $\mathbf{z}_1, \ldots, \mathbf{z}_t \in \mathcal{Z}$, we compute the unimodular $\mathbf{U} \in \mathbb{Z}^{d\times d}$ that can be efficiently computed as in Definition 3.2, and let $\mathbf{U}' \in \mathbb{Z}^{d\times(d-t)}$ be the matrix formed by removing the t leftmost columns of $\mathbf{U}$. We receive m samples $(\mathbf{A} \in \mathbb{R}_q^{d\times w}, \mathbf{b} \in \mathbb{R}_q^w)$ from our "first-t-errorless" M–LWE oracle. We then sample $\mathbf{f} \leftarrow D_{\alpha(\xi^2\mathbf{I}-\mathbf{U}\mathbf{U}^t)^{1/2}}$. The above distribution is well-defined since $\xi^2\mathbf{I} - \mathbf{U}'\mathbf{U}'^t$ is a positive semidefinite matrix according to our assumption on $\mathbf{U}$.

The reduction then sets $\bar{\mathbf{b}} = \mathbf{U}\mathbf{b}+\mathbf{f}$, and samples $\mathbf{c}$ from the discrete Gaussian distribution $D_{q^{-1}R^w-\bar{\mathbf{b}},r}$. It then outputs

$$(\mathbf{A}' = \mathbf{A}\mathbf{U}^t, \mathbf{b}' = \bar{\mathbf{b}} + \mathbf{c}, (\mathrm{Tr}(\langle \mathbf{z}_i, \mathbf{f}+\mathbf{c}\rangle))_{i\in[t]}) \in R_q^{d\times w} \times R_q^w \times \mathbb{Q}_q^t$$

We now analyze correctness of the reduction. First, we have that $\mathbf{U}$ is invertible, so that $\mathbf{A}' = \mathbf{A}\mathbf{U}^t$ is uniform solely over the choice of $\mathbf{A}$. For the remainder of the proof we condition on any fixed value of $\mathbf{A}'$ and focus on analyzing the distribution of the second and third components of the output.

First, consider the case that the samples received by our reduction came from the first-t-are-errorless variant of M–LWE. Then we have that

$$\bar{\mathbf{b}} = \mathbf{U}\mathbf{b} + \mathbf{f} = \mathbf{A}'^t\mathbf{s} + \mathbf{U}\mathbf{e} + \mathbf{f}$$

We have that $\mathbf{U}\mathbf{e}$ is distributed as a continuous Gaussian $D_{\alpha\mathbf{U}}$, so by additivity of Gaussians, we have that $\mathbf{U}\mathbf{e} + \mathbf{f}$ is distributed as a spherical continuous Gaussian $D_{\alpha\xi}$. Since $\mathbf{A}'^t\mathbf{s} \in R_q^w$, we have that $\mathbf{c}$ is in fact distributed according to $D_{q^{-1}R^w-\bar{\mathbf{b}},r}$. As a result, by Lemma 2.1, $\mathbf{U}\mathbf{e}+\mathbf{f}+\mathbf{c}$ is within statistical distance 8ϵ of $D_{q^{-1}R^w,(\alpha^2\xi^2+r^2)^{1/2}}$, which shows that $\mathbf{b}'$ is distributed correctly. For the third component, since the first t elements of $\mathbf{e}$ are zero and the last $d-t$ columns of $\mathbf{U}$ are orthogonal to each $\mathbf{z}_i$, we have that

$$\langle \mathbf{z}_i, \mathbf{f}+\mathbf{c}\rangle = \langle \mathbf{z}_i, \mathbf{U}\mathbf{e}+\mathbf{f}+\mathbf{c}\rangle,$$

so the third component gives the inner product of the noise with each vector $\mathbf{z}_i$, as desired.

Finally, we consider the case where the samples received by our reduction are truly uniform. Then since $\mathbf{b}$ is truly uniform and independent, we can view it as being distributed as $\mathbf{b} = \tilde{\mathbf{b}} + \mathbf{e}$, where $\tilde{\mathbf{b}} \in R_q^w$ is truly uniform and $\mathbf{e}$'s first t coordinates are 0, and the remaining coordinates are chosen independently from D_α. Then we have that $\mathbf{b}' = \mathbf{U}\tilde{\mathbf{b}} + \mathbf{U}\mathbf{e} + \mathbf{f} + \mathbf{c}$, where $\mathbf{U}\mathbf{e}+\mathbf{f}+\mathbf{c}$ is distributed exactly as in the case above. Since $\mathbf{U}$ is unimodular, we have that $\mathbf{U}(\tilde{\mathbf{b}}+\mathbf{e})$ is uniform and independent over R_q^w. Finally, since $\mathbf{U}\mathbf{e}+\mathbf{f}+\mathbf{c}$ is distributed exactly as above, we also have that the third component output is distributed correctly, as needed.

4 Reducing Extended-LWE to LWR

Theorem 4.1. *Let κ be a security parameter on which all other parameters depend. Let R be the ring of integers of an algebraic number field of dimension $n = O(\kappa)$. Let ψ be an arbitrary coordinate-wise B-bounded distribution over R for some $B > 0$. Let $p, q, w = \mathrm{poly}(\kappa), d \in \mathbb{N}$ such that $q \geq 4eBwn\kappa p$. Let $\mathcal{Z} \subseteq R^d$ be the set of vectors $\mathbf{z}$ such that $\mathrm{Tr}(\langle \mathbf{z}, \mathbf{e} \rangle)$ extracts the ith coefficient of the jth ring element of $\mathbf{e}$ for some $i \in [n], j \in [d]$.*

If $\exists$ a probabilistic polynomial-time $\mathcal{A}$ succeeding with advantage $\epsilon(\kappa) \geq (\kappa)^{-c}$ for some constant $c \geq 1$ in distinguishing M–$\mathsf{LWR}_{R,d,w,q,p}$ from uniform, then there exists a probabilistic polynomial-time algorithm $\mathcal{A}'$ succeeding with advantage $\epsilon(nw)^{-c}/4 \geq (\kappa nw)^{-c}/4$ in distinguishing Ext–M–$\mathsf{LWE}_{R,d,w,q,\psi,\mathcal{Z},c}$ from uniform.

4.1 The Set BAD

Before stating our reduction, we define and prove some basic results about the set $\mathsf{BAD}_{B,q,p}$. This set, following the notation of [BPR12], characterizes those elements in $\mathbb{Z}_q$ that are within distance B of an element that rounds (with $\lfloor \cdot \rceil_p$ as defined above) to a different value than it does.

Definition 4.2. *For $B, p, q \geq 1$, we define*

$$\mathsf{BAD}_{B,q,p} := \{x \in \mathbb{Z}_q : \lfloor x \rceil_p \neq \lfloor x \pm B \rceil_p\}$$

The reason for the terminology is that, letting one of these elements represent a coefficient with noise (from a B-bounded distribution) in an M–LWE sample, it may be the case that the noiseless version of these coefficients rounds to a different value than the noisy one, which is "bad" for our reduction.

We recall a bound on the probability that a uniform element in $\mathbb{Z}_q$ is in $\mathsf{BAD}_{B,q,p}$.

Lemma 4.3. *[AKPW13, Lemma 2.7]*

$$\Pr_{x \leftarrow \mathbb{Z}_q}[x \in \mathsf{BAD}_{B,q,p}] \leq 2Bp/q.$$

This next lemma uses a standard Chernoff bound to upper bound the probability any given number of elements in a uniformly $\mathbf{b} \in R_q^w$ will be in BAD.

Lemma 4.4. *Let B, q, p, κ, n, w satisfy the conditions in Theorem 4.1. Let $\mathbf{b} \leftarrow R_q^w$ uniformly at random, and let $S = \{b_{i,j} \in \mathsf{BAD}_{B,q,p}\}$. Then for all $c \geq 1$,*

$$\Pr[|S| \geq c] \leq \kappa^{-c}/2$$

Proof. Let X_i be the event that $b_i \in \mathsf{BAD}_{B,q,p}$. Let $X = \sum_{i \in [m]} X_i, \mu = E[X]$. First, by linearity of expectation and Lemma 4.3, we have that $\mu \leq 2Bnwp/q \leq 1/(2e\kappa)$.

Let $\delta = \frac{c}{\mu} - 1 > 0$. Then

$$\Pr[X > c] < \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu = \frac{e^{\frac{\delta c}{1+\delta}}}{(1+\delta)^c} \leq \left(\frac{e}{1+\delta}\right)^c = \kappa^{-c}/(2c)^c \leq \kappa^{-c}/2.$$

4.2 Reduction

We begin by choosing which c error elements we will receive as hints, by choosing uniformly at random c distinct elements

$$\mathsf{GUESSES} := \{\ell_1, \ell_2, \ldots, \ell_c \leftarrow [w] \times [n]\},$$

where the $\ell_\iota = (i_\iota, j_\iota)$ are tuples representing the jth coefficient of the ith element in the error vector $\mathbf{e} \in R_q^w$.

Since the trace is a "universal" $\mathbb{Q}$-linear function, there exist (efficiently computable) vectors $\mathbf{z}_{\ell_1}, \mathbf{z}_{\ell_2}, \ldots, \mathbf{z}_{\ell_c}$ such that $\mathrm{Tr}(\langle \mathbf{e}, \mathbf{z}_\iota \rangle) = \mathbf{e}_{\ell_\iota}$ for all i. We query the $\mathsf{Ext\text{–}M\text{–}LWE}_{R,d,w,q,\psi,\mathcal{Z},c}$ oracle with these vectors, receiving back ($\mathbf{A} \in R_q^{d \times w}, \mathbf{b} \in R_q^w, (e_{\ell_1}, \ldots, e_{\ell_c})$).

Let

$$T = \{(i,j) : b_{i,j} \in \mathsf{BAD}_{B,q,p}\},$$

where $b_{i,j}$ refers to the jth coefficient of the ith ring element. We have the following abort events.

E.1 If $|T| > c$, we abort.
E.2 Otherwise, if $T \not\subseteq \mathsf{GUESSES}$, we abort.
E.3 Otherwise, we abort with probability 1-1/$\binom{wn-|T|}{c-|T|}$.

Now, if we have not aborted, set $\tilde{\mathbf{e}}_{i,j} = e_{i,j}$ if $(i,j) \in \mathsf{GUESSES}$, $\tilde{\mathbf{e}}_{i,j} = 0$. Let $\mathbf{b}' := \lfloor \mathbf{b} - \tilde{\mathbf{e}} \rceil_p$, where the rounding is done element-wise, and let $y \leftarrow \mathcal{A}(\mathbf{A}, \mathbf{b}')$. $\mathcal{A}'$ outputs y as its guess (where 0 corresponds to $\mathsf{Ext\text{–}M\text{–}LWE}/\mathsf{M\text{–}LWR}$, respectively, 1 to uniform).

4.3 Analysis

Here we prove correctness of the above reduction. To do so, we need to prove two things. First, we need to show that we invoke $\mathcal{A}$ (i.e. we do not abort and output a random bit) with non-negligible probability. Second, conditioned on not aborting, we need to show that the distribution of the output sent to $\mathcal{A}$ is within statistical distance at most $\epsilon/2$ of the true $\mathsf{M\text{–}LWR}$ distribution if the oracle samples $\mathbf{b}$ from $\mathsf{Ext\text{–}M\text{–}LWE}$, and within distance at most $\epsilon/2$ of the uniform distribution if the oracles samples $\mathbf{b}$ uniformly and independently.

The following simple lemma shows that if the distribution of $|T|$ differed depending on whether the oracle gives samples from $\mathsf{Ext\text{–}M\text{–}LWE}$ or uniform, then we would have an (alternative standalone) attack on $\mathsf{Ext\text{–}M\text{–}LWE}$. As a result, we need only consider the distribution of $|T|$, when the oracle outputs a truly uniform $\mathbf{b}$, where it is simpler to analyze.

Lemma 4.5. *Under the* $\mathsf{Ext\text{–}M\text{–}LWE}_{R,d,w,q,\psi,\mathcal{Z},c}$ *assumption, we have that*

$$|\Pr[|T| > c \mid \textit{oracle is } \mathsf{Ext\text{–}M\text{–}LWE}] - \Pr[|T| > c \mid \textit{oracle is uniform}]| \leq \mathrm{negl}(n)$$

Proof. If there was a non-negligible difference in the two probabilities, we could simply query the oracle and compute $|T|$ to give an attack on $\mathsf{Ext\text{–}M\text{–}LWE}_{R,d,w,q,\psi,\mathcal{Z},c}$ succeeding with non-negligible probability, without having to rely on $\mathcal{A}$.

We can now upper bound the probability of abort event E.1.

Lemma 4.6.

$$\Pr[\mathbf{E.1} \mid \textit{oracle is } \mathsf{Ext\text{–}M\text{–}LWE}] \leq \mathrm{negl}(\kappa) + \Pr[\mathbf{E.1} \mid \textit{oracle is uniform}] \leq \epsilon/2$$

Proof. The first inequality is a direct consequence of Lemma 4.5, while the second follows immediately from 4.4.

The next lemma shows that the distribution conditioned only on **E.1** not happening is exactly the same as the distribution conditioned on none of the abort events happening, and that the probability of none of the abort events happening is at least $(1-\epsilon/2)/\binom{nw}{c}$.

Lemma 4.7.

$$\Pr[\neg(\mathbf{E.1} \vee \mathbf{E.2} \vee \mathbf{E.3})] \geq \frac{(1-\epsilon/2)}{\binom{nw}{c}}.$$

Moreover, for any $\mathbf{A}^* \in R_q^{d\times w}, \mathbf{b}^* \in R_q q^w \in \mathbb{Z}_q^c$*, we have*

$$\Pr[(\mathbf{A},\mathbf{b}) = (\mathbf{A}^*,\mathbf{b}^*) \mid \neg\mathbf{E.1}] = \Pr[(\mathbf{A},\mathbf{b}) = (\mathbf{A}^*,\mathbf{b}^*) \mid \neg(\mathbf{E.1} \vee \mathbf{E.2} \vee \mathbf{E.3})]$$

Proof. First, we consider **E.2**. Consider any fixed possible value t for the number of elements in BAD. For this fixed value of BAD, there are $\binom{nw-t}{c-t}$ possible distinct sets $\mathsf{GUESSES}$ of indices that cover all elements in BAD, while there are $\binom{nw}{c}$ total possible distinct sets $\mathsf{GUESSES}$ of indices.

Since the elements of $\mathsf{GUESSES}$ are chosen uniformly and independent of everything else, we have that

$$\Pr[\neg\mathbf{E.2} \mid \neg\mathbf{E.1} \wedge |\mathsf{BAD}| = t] = \binom{wn-t}{c-t}/\binom{wn}{c}$$

Next, by the definition of **E.3**,

$$\Pr[\neg\mathbf{E.3} \mid \neg(\mathbf{E.1} \vee \mathbf{E.2}) \wedge |\mathsf{BAD}| = t] = 1/\binom{wn-t}{c-t}$$

$$\Pr[\neg\mathbf{E.3} \wedge \neg\mathbf{E.2} \mid \neg\mathbf{E.1} \wedge |\mathsf{BAD}| = t] = 1/\binom{wn}{c}$$

independently of $|\mathsf{BAD}|$ and thus independently of the value of $\mathbf{A},\mathbf{b}$ received from the $\mathsf{Ext\text{–}M\text{–}LWE}$ oracle.

$$\Pr[\neg\mathbf{E.3} \wedge \neg\mathbf{E.2} \mid \neg\mathbf{E.1}] = 1/\binom{nw}{c}.$$

Combining the above result with Lemma 4.6 gives

$$\Pr[\neg(\mathbf{E.1} \vee \mathbf{E.2} \vee \mathbf{E1.C})] \geq \frac{1-\epsilon/2}{\binom{nw}{c}}$$

Finally, we show that conditioned on not aborting, the samples sent to $\mathcal{A}$ are distributed sufficiently close to the desired distributions.

Lemma 4.8. *The distribution of* $(\mathbf{A}, \mathbf{b}')$ *sent to* $\mathcal{A}$ *is within statistical distance at most* $\epsilon/2$ *of the true* M–LWR *distribution if the oracle sampled* $\mathbf{b}$ *from* Ext–M–LWE*, and uniform if it sampled it uniformly at random.*

Proof. Since the reduction only ever sends $(\mathbf{A}, \mathbf{b}')$ if none of the abort events (in particular, $\mathbf{E}.\mathbf{2}$) happen, we receive as our Ext–M–LWE hints the error coefficients for every $b_{i,j} \in \mathsf{BAD}_{B,q,p}$. Let $\mathbf{y} = \mathbf{b} - \tilde{\mathbf{e}} \in R_q^w$ (with $\tilde{\mathbf{e}}$ as defined above).

In the case the oracle samples $\mathbf{b} = \mathbf{A}^t\mathbf{s} + \mathbf{e}$ from Ext–M–LWE, we have that every coefficient of every ring element is either errorless or it is "good", meaning it is not in BAD. As a result, we have that

$$\lfloor \mathbf{y} \rceil_p = \lfloor \mathbf{A}^t\mathbf{s} \rceil_p,$$

so, combining with Lemmas 4.6 and 4.7 , we have that it is distributed within distance $\epsilon/2$ of M–LWR.

In the case the oracle samples $\mathbf{b}$ uniformly at random and independently, we have that $\tilde{\mathbf{x}}$ is independent of $\mathbf{b}$, and so $\mathbf{y}$ remains truly uniform, making $\lfloor \mathbf{y} \rceil_p$ have the identical distribution to the "uniform" distribution.

By Lemma 4.8 and our assumption on $\mathcal{A}$, conditioned on none of the abort events happening, $\mathcal{A}$ must succeed in distinguishing with advantage at least $\epsilon/2$.

Combining this with Lemma 4.7 gives

$$\epsilon' \geq \frac{\epsilon/2(1-\epsilon/2)}{\binom{nw}{c}} \geq \epsilon(nw)^{-c}/4,$$

as required.

This completes the reduction's correctness proof.

Acknowledgments

This work was performed under financial assistance award 70NANB15H328 from the U.S. Department of Commerce, National Institute of Standards and Technology.

References

[AKPW13] Joël Alwen, Stephan Krenn, Krzysztof Pietrzak, and Daniel Wichs. Learning with rounding, revisited - new reduction, properties and applications. In *Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, pages 57–74, 2013.

[AP12] Jacob Alperin-Sheriff and Chris Peikert. Circular and KDM security for identity-based encryption. In *Public Key Cryptography*, pages 334–352, 2012.

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993.

[BBO07] Mihir Bellare, Alexandra Boldyreva, and Adam O'Neill. Deterministic and efficiently searchable encryption. In *Advances in Cryptology - CRYPTO 2007, 27th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2007, Proceedings*, pages 535–552, 2007.

[BFO08] Alexandra Boldyreva, Serge Fehr, and Adam O'Neill. On notions of security for deterministic encryption, and efficient constructions without random oracles. In *Advances in Cryptology - CRYPTO 2008, 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings*, pages 335–359, 2008.

[BFOR08] Mihir Bellare, Marc Fischlin, Adam O'Neill, and Thomas Ristenpart. Deterministic encryption: Definitional equivalences and constructions without random oracles. In *Advances in Cryptology - CRYPTO 2008, 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings*, pages 360–378, 2008.

[BFP+15] Abhishek Banerjee, Georg Fuchsbauer, Chris Peikert, Krzysztof Pietrzak, and Sophie Stevens. Key-homomorphic constrained pseudorandom functions. In *Theory of Cryptography - 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part II*, pages 31–60, 2015.

[BGM+16] Andrej Bogdanov, Siyao Guo, Daniel Masny, Silas Richelson, and Alon Rosen. On the hardness of learning with rounding over small modulus. In *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, pages 209–224, 2016.

[BGV12] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In *ICTS*, pages 309–325, 2012.

[BLMR13] Dan Boneh, Kevin Lewi, Hart William Montgomery, and Ananth Raghunathan. Key homomorphic prfs and their applications. In *Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, pages 410–428, 2013.

[BLP+13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In *STOC*, pages 575–584, 2013.

[BP14] Abhishek Banerjee and Chris Peikert. New and improved key-homomorphic pseudorandom functions. In *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I*, pages 353–370, 2014.

[BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In *EUROCRYPT*, pages 719–737, 2012.

[BS14] Zvika Brakerski and Gil Segev. Better security for deterministic public-key encryption: The auxiliary-input setting. *J. Cryptology*, 27(2):210–247, 2014.

[Coh96] Henri Cohen. Hermite and smith normal form algorithms over dedekind domains. *Mathematics of Computation*, 65(216):1681–1699, 1996.

[DKL09] Yevgeniy Dodis, Yael Tauman Kalai, and Shachar Lovett. On cryptography with auxiliary input. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 621–630, 2009.

[DM14] Léo Ducas and Daniele Micciancio. Improved short lattice signatures in the standard model. In *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I*, pages 335–352, 2014.

[Erd46] Paul Erdős. On the coefficients of the cyclotomic polynomial. *Bulletin of the American Mathematical Society*, 52(2):179–184, 1946.

[FOR15] Benjamin Fuller, Adam O'Neill, and Leonid Reyzin. A unified approach to deterministic encryption: New constructions and a connection to computational entropy. *J. Cryptology*, 28(3):671–717, 2015.

[GPV08] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206, 2008.

[LPR13a] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. *J. ACM*, 2013. To appear. Preliminary version in Eurocrypt 2010.

[LPR13b] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-LWE cryptography. In *EUROCRYPT*, pages 35–54, 2013.

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

[MR07] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. Preliminary version in FOCS 2004.

[OPW11] Adam O'Neill, Chris Peikert, and Brent Waters. Bi-deniable public-key encryption. In *CRYPTO*, pages 525–542, 2011.

[Pei09] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem. In *STOC*, pages 333–342, 2009.

[Pei10] Chris Peikert. An efficient and parallel Gaussian sampler for lattices. In *CRYPTO*, pages 80–97, 2010.

[PG13] Thomas Pöppelmann and Tim Güneysu. Towards practical lattice-based public-key encryption on reconfigurable hardware. In *Selected Areas in Cryptography - SAC 2013 - 20th International Conference, Burnaby, BC, Canada, August 14-16, 2013, Revised Selected Papers*, pages 68–85, 2013.

[PW11] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. *SIAM J. Comput.*, 40(6):1803–1844, 2011.

[Reg09] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):1–40, 2009. Preliminary version in STOC 2005.

[Sop10] Ivan Soprounov. A short proof of the prime number theorem for arithmetic progressions. *preprint*, 2010.