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 \(\psi\), an LWE\(_{d,q,\psi}\) sample for a secret \(s\) consists of a uniformly random \(a \in \mathbb{Z}_q^d\) and \(b = \langle a, s \rangle + e \mod q \in \mathbb{Z}_q\), where \(e\) is small, random “noise” sampled from \(\psi\). By contrast, the corresponding LWR\(_{d,q,a}\) sample in dimension \(d\) with moduli \(q > p\) is formed by simply rounding off the lower-order bits, instead setting \(a \in \mathbb{Z}_q^d\) uniformly random and \(b = \lfloor (p/q)\langle a, s \rangle \rfloor \in \mathbb{Z}_p\). 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 \cite{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 \cite{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 \cite{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, \cite{BPR12} first introduced LWR, and in particular gave a reduction from LWE to LWR when the ratio of the LWR moduli \( q/p \) is superpolynomial in the security parameter \( \kappa \). In a little more detail, for any efficiently samplable \( B \)-bounded distribution \( \psi \) (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 \( s \), \cite{BPR12} showed that distinguishing unboundedly-many \( \text{LWR}_{d,q,p} \) samples from uniform samples is at least as hard as distinguishing as many \( \text{LWE}_{d,q,\psi} \) samples from uniform.

Banerjee, Peikert and Rosen also show analogous results for the case of the Ring Learning with Errors (RLWE) problem \cite{LPR13a} and Ring Learning with Rounding (RLWR), where for an appropriate choice of ring \( R \) and secret \( s \in R \), 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. \cite{AKPW13} gave the first reduction from LWE to LWR in the case when the modulus \( q \) is a fixed polynomial in the security parameter \( \kappa \). However to do so, their techniques introduce a number of drawbacks not originally present in the \cite{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 \( d', w, q, \psi \) \( B \)-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 \cite{DKL09}, lossy trapdoor function families \cite{PW11}, and deterministic encryption \cite{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 $LWR_{d,w,q,p}$ from $LWE_{d',w,q,\psi}$ with parameters $d',w,q,\psi$ 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 $\rho$, 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 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 $\kappa$ be a security parameter on which all other parameters depend. Let $\psi$ be a distribution over $\mathbb{Z}$ for some $B > 0$. Let $p,q = \text{poly}(\kappa), w = \text{poly}(\kappa), d \in \mathbb{N}$ such that $q \geq 4eBwp$. If $\exists$ a probabilistic polynomial-time $A$ succeeding with advantage $\epsilon(\kappa) \geq (\kappa)^{-c}$ for some constant $c \geq 1$ in distinguishing $LWR_{d,w,q,p}$ from uniform, then there exists a probabilistic polynomial-time algorithm $A'$

1. succeeding with advantage $\epsilon(wB)^{-c}/4 \geq (\kappa w B)^{-c}/4$ in distinguishing $LWE_{d,w,q,\psi}$ from uniform, as long as every prime factor of $q$ is greater than $B$ and $\psi$ is $B$-bounded.
2. succeeding with advantage $\epsilon(w)^{-c}/4 \geq (\kappa w)^{-c}/4$ in distinguishing $LWE_{d-c,w,q,\psi}$ from uniform, with no restrictions on the factors of $q$, as long as $\psi = D_a$ 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...
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 \textsc{LWE} over integer lattices can easily follow it, by simply replacing $R$ with $\mathbb{Z}$.

We are unable to prove any reductions from \textsc{RLWE} to the form of \textsc{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 \textsc{RLWE} to an alternative form of \textsc{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 \textsc{LWE} to \textsc{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 \textsc{RLWE} samples based on the guess so that a correct guess results in valid \textsc{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 \textsc{RLWE} in [LPR13a].

However, such a reduction does not appear to be useful for our goal of getting a reduction for \textsc{RLWR}, at least in a manner similar to our reduction for regular \textsc{LWR}, so we omit a formal proof. The problem here is that the \textsc{RLWE} error will necessarily be statistically uniform modulo any small prime ideal, so we cannot simply round each coefficient of an \textsc{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 \textit{module lattices} [LS15]. Over module lattices, we are able to show a reduction from \textsc{M-LWE} to \textsc{Ext-M-LWE}. This reduction is a fairly straightforward adaptation of the \textsc{LWE} to \textsc{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.

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

<table>
<thead>
<tr>
<th>Work</th>
<th>Unbounded Samples ((w))</th>
<th>Modulus ((q))</th>
<th>Advantage Change ((\epsilon \rightarrow \epsilon'))</th>
<th>Dimension Change ((d \rightarrow d'))</th>
</tr>
</thead>
<tbody>
<tr>
<td>[BPR12]</td>
<td>Yes</td>
<td>$B_{pk}^{\omega(t)}$</td>
<td>$\epsilon - \text{negl}(\kappa)$</td>
<td>$d$</td>
</tr>
<tr>
<td>[AKPW13]</td>
<td>No</td>
<td>$\gamma B_{wpk}$</td>
<td>$\epsilon/(2dw)$</td>
<td>$d \log(\gamma)/\log q$</td>
</tr>
<tr>
<td>[BGM+16]</td>
<td>No</td>
<td>$B_{wp}$</td>
<td>$(\epsilon/qw)^{\omega}$</td>
<td>$d/\log_q q$</td>
</tr>
<tr>
<td>This work (1)</td>
<td>No</td>
<td>$B_{wpk}$</td>
<td>$\epsilon(w)^{-\omega}$</td>
<td>$d$</td>
</tr>
<tr>
<td>This work (2)</td>
<td>No</td>
<td>$B_{wpk}$</td>
<td>$\epsilon(w)^{-c}$</td>
<td>$d - c$</td>
</tr>
</tbody>
</table>


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

\[ \lfloor \langle a, s \rangle \rceil_p = \lfloor \langle a, s \rangle + e \rceil_p \mod q, \]

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 \( \kappa \).

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 \( \kappa \)) of a given adversary \( A \) 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 \( b \) contains more than \( c \) “bad” elements is at most \( 1/2^{\kappa - c} \). Consequently, if we could find a method to transform any LWE sample with at most \( c \) “bad” elements into a valid LWR sample, then we could construct \( A' \) that could use \( A \) to successfully attack LWE in the same dimension. Specifically, \( A' \) would work as follows:

1. \( A' \) queries its LWE oracle to receive back (\( A \in \mathbb{Z}_q^{d \times w}, b \)). If \( b \) contains \( c \) or more bad elements, \( A' \) aborts.
2. ??????
3. \( A' \) sends \( A \) a sample (\( A \in \mathbb{Z}_q^{d \times w}, b \in \mathbb{Z}_p^w \)) 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, \( \lfloor b_i \rfloor_p = \lfloor \langle a_i, s \rangle \rfloor_p \). But what about the “bad” elements, where it may be the case that \( \lfloor b_i \rfloor_p \neq \lfloor \langle a_i, s \rangle \rfloor_p \). 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 \( e \) to be received along with the LWE sample (\( A, b \)). 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 \( b \) might be “bad.” With probability at least \( w^{-c} \geq 1/\text{poly}(\kappa) \), 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 \( \text{M-LWE} \) to \( \text{Ext-LWE} \); 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 \( \text{Ext-M-LWE} \) to \( \text{M-LWR} \). We stress that this section has been written to be accessible to
someone only familiar with integer lattices, by simply viewing \( \mathbb{R} \) 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 \( x \in \mathbb{Z}_q \), we define

\[
[x]_p = \lfloor (p/q) \cdot x \rfloor,
\]

where the latter \( \lfloor \cdot \rfloor \) denotes standard rounding to the nearest integer. For a ring
element \( y \in \mathbb{R}_q \), 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 \( \kappa \) 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 embed-
ding, 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

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

Letting \( e_j \in \mathbb{C}^n \) be the vector with 1 in its \( j \)th coordinate, 0 elsewhere, we
have the following orthonormal basis for \( H \). We let \( h_j = \frac{1}{\sqrt{2}}(e_j + e_{j+s_2}) \) and
\( h_{j+s_2} = \frac{1}{\sqrt{2}}(e_j - e_{j+s_2}) \). This basis also shows that \( H \) is isomorphic to \( \mathbb{R}^n \) as
an inner product space.
We also define the trace
\[ \text{Tr} \]
with \( X \). The Duality. Let \( I \) be an ideal in \( \mathcal{O}_m = \mathbb{Z}[\zeta_m] \), and refer to it as the mth cyclotomic ring. The minimal polynomial of \( \zeta_m \) over \( \mathbb{Q} \) is the mth cyclotomic polynomial \( \Phi_m(X) = \prod_{\omega \in \mathbb{Z}_m} (X - \omega) \in \mathbb{Z}[X] \), where \( \omega_m = \exp(2\pi \sqrt{-1}/m) \in \mathbb{C} \) is the principal mth complex root of unity, and the roots \( \omega_i \in \mathbb{C} \) range over all the primitive 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 \( \zeta_m \) 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 \( \text{Tr} : K \to \mathbb{Q} \) and the (field) norm \( N : K \to \mathbb{Q} \) as \( \text{Tr}(x) = \sum \sigma_j(x), N(x) = \prod \sigma_j(x) \), 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 \( L(a) = \text{Tr}(r \cdot a) \) 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 \( IJ = \{ \sum_k i_k j_k : i_k \in I, j_k \in J \} \). 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 \( I^\vee = \{ x \in K : \text{Tr}(xI) \subseteq \mathbb{Z} \} \), and we have that \( I^\vee = I^{-1} \cdot R^\vee \). If \( R \) is the \( m = 2^e p_1^{a_1} \ldots p_t^{a_t} \)-th cyclotomic ring, let \( g = \prod_{i \in \{i\}} (1 - \zeta_{p_i}) \in R \) and let \( \hat{m} = m/2 \) is even, \( \hat{m} = m \) otherwise. Then \( R^\vee = (g/\hat{m}) \). 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 coefficient.

Factorization of Ideals. Let \( q = p^r \in \mathbb{Z} \) be a prime power. In the \( m \)-th 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 \( m = \tilde{m} \cdot p^k \) where \( p \nmid \tilde{m} \). Let \( e = \varphi(p^k) \), and let \( d \) be the multiplicative order of \( p \) in \( \mathbb{Z}_{\tilde{m}}^{*} \). Note that \( d \) divides \( \varphi(\tilde{m}) = n/e \). The ideal \( qR \) then factors into the product of \( (re)^t \) \( \text{th} \) powers of \( \varphi(\tilde{m})/d = n/((de)) \) distinct prime ideals \( p_i \), i.e. \( qR = \prod p_i^{e_i} \). Each prime ideal \( p_i \) has norm \( |R/p_i| = p^d \). The factorization of \( qR \) for general \( q = p_1^{e_1} \ldots p_i^{e_i} \) then follows by recalling the unique factorization of ideals for cyclotomic rings and that \( qR = (p_1^{e_1} R) \ldots (p_i^{e_i} R) \).

**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 \( b_i \in K^d \) such that \( M = \sum_{i=1}^d I_i \cdot b_i \) for some non-zero ideals \( I_i \) 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 \( \mathcal{L} \subseteq \mathbb{R}^n \) is a discrete additive subgroup consisting of all integer linear combinations \( \sum_{i \in \mathbb{Z}_n} x_i b_i \), where \( B = (b_i) \) is called the basis.

Under the canonical embedding, a fractional ideal \( I \) with \( \mathbb{Z} \)-basis \( U = (u_1, \ldots, u_n) \) becomes a rank-\( n \) ideal lattice \( \sigma(I) \) with basis \( \{\sigma(u_1), \ldots, \sigma(u_n)\} \subset H \). 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 \( \sigma_M = (\sigma_i)_{i \in [d]} \), where \( \sigma \) is the canonical embedding. For a module lattice \( M \), the set \( \sigma_M(M) \) is a module lattice, and similarly for ideal lattices, we 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 \( \text{GapSVP} \sqrt{n} \) 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 \( \text{Mod}–\text{GapSVP} \) (i.e. \( \text{GapSVP} \) over module lattices) to within any polynomial factor. As a result, we may apply Peikert’s classical reduction from \( \text{GapSVP} \) to \( \text{BDD} \) in the module lattice setting [Pei09], using the oracle reducing \( \text{Mod}–\text{BDD} \) to \( \text{M–LWE} \) 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 \( \rho_s \) is defined as

\[
\rho_s(x) := \exp(-\pi \|x\|^2 / s^2).
\]

Normalizing this function gives the continuous Gaussian distribution \( D_s \). More generally, we can define \( DB \) as the distribution of \( Bx \) where \( x \) is sample from \( D1 \). For any \( B1, B2 \), the sum of a sample from \( DB1 \) and \( DB2 \) is distributed as \( D(B1B1^t + B2B2^t)^{1/2} \).

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

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

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

1. There is a probabilistic polynomial-time algorithm that, given \( c \in \mathbb{R}^n \) outputs a sample distributed according to \( D_{\Lambda + c,s} \).
2. The distribution of \( x \mod \Lambda \), where \( x \leftarrow D_s \), is within statistical distance \( \epsilon/2 \) of the uniform distribution over cosets of \( \Lambda \).
3. Let \( r > 0 \). Then if we choose \( x \leftarrow D_r \) and then choose \( y \leftarrow D_{\Lambda - x,s} \), we have that \( x + y \) is within statistical distance \( 8\epsilon \) 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 general 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 \( \psi \) be some distribution over \( R \). Then we define the decision version of M–LWE\(_{R,d,w,q,\psi} \) as follows.

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

When \( \psi \) is a spherical Gaussian, we sometimes write M–LWE\(_{R,d,w,q,\psi} \) to denote M–LWE\(_{R,d,w,q,D_\psi} \), and similarly for LWE and RLWE.
Under this definition, it is easy to see that \( M\text{-LWE} \) is a generalization of both RLWE and LWE. When \( d = 1 \), \( M\text{-LWE}_{R,d,w,q,\psi} \) is equivalent to RLWE\(_{R,w,q,\psi} \), while when \( R = \mathbb{Z} \), \( M\text{-LWE}_{R,d,w,q,\psi} \) is equivalent to 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 \( b \) 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 \( \text{SIVP} \) for module lattices for \( d \geq 1 \), and we also have a (potentially meaningful) classical reduction from \( \text{GapSVP} \) 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 \( \psi \) over \( \mathbb{Z} \) is said to be \( B \)-bounded if

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

For completeness, we recall the following result of Banszczzyk [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 \( \Lambda \), \( r > 0 \), and let \( x \leftarrow D_{\Lambda,r} \) we have that except with probability \( 2^{-2n} \), \( \|x\| \leq r\sqrt{n} \).

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 \( 2^k p^r \), 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 $n$, $R = O_K$ 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 M–LWR as follows: The adversary gets $(A \leftarrow R_d^{d \times w}, b \in R_p^w)$, and must distinguish between two cases. In the first case, which we refer to as the M–LWR distribution, $b = \lceil (p/q)A^t s \rceil p \in R_p^w$, where $s \leftarrow R_d^q$ uniformly at random, while in the second case, which we refer to use at the uniform case, $b = \lceil u \rceil p$, where $u \leftarrow R_q^w$ uniformly at random.

Note that if $p$ divides $q$, the distribution of $b$ is truly uniform, while if $p \nmid q$, $b$ 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 $Z \subseteq K^d$, a parameter $k$ denoting the number of hints received, and a distribution $\psi$ over $R$, the Ext–M–LWE$_{R,d,w,q,\psi,Z,k}$ problem is defined as follows: The algorithm gets to choose $z_1, \ldots, z_k \in Z$ and receives back a tuple

$$(A, b, (\text{Tr}(\langle e, z_i \rangle)))_{i \in [k]} \in R_q^{d \times w} \times R_q^w \times Q_k^q$$

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

11
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 \texttt{Ext-LWE} (and indeed has no meaning over general lattices, as \text{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 $\langle e, z \rangle$. 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 \texttt{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 $Z \subseteq \mathbb{Z}^m$.

\textbf{Definition 3.2.} For a real $\xi > 0$ and a set $Z \subseteq \mathbb{K}^w$, we say that $Z$ is of quality $(\xi, t)$ if given any $z_1, \ldots, z_t \in Z$, we can efficiently construct a unimodular matrix $U \in \mathbb{Z}^{w \times w}$ such that, if $U' \in \mathbb{Z}^{w \times (w-t)}$ is the matrix obtained from $U$ by removing its $t$ leftmost columns, then all the columns of $U'$ are orthogonal to $\text{span}(z_1, \ldots, z_t)$ and its largest singular value is at most $\xi$.

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

For this particular setting of $Z$, there exist two prior somewhat incomparable reductions for \texttt{Ext-LWE} (over general lattices), given by Alperin-Sheriff and Peikert \cite{AP12} and by Brakersi et al., respectively. \cite{BLP13}. We state them both here:

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

1. \cite{AP12}: \texttt{Ext-LWE}_{d,w,q,\alpha,k}, that reduces the advantage by a multiplicative factor of at most $q^k$.
2. \cite{BLP13}: \texttt{Ext-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.

\subsection{Extended-LWE Over Module Lattices}

Here we show that an analogous version of the \texttt{Ext-LWE} reduction found in \cite{BLP13} holds over module lattices as well. In order to prove this theorem, we follow the path of Brakerski et al. \cite{BLP13}. 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 \texttt{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 $\psi$ 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 $A \in R_q^{d \times w}, b = A's + e \in R_q^w$, where $s \leftarrow R_q^d$ uniformly at random and $e \leftarrow (0, e')$, where $e' \leftarrow \psi^{-1}$. In the second case, $b \leftarrow R_q^m$ 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 $m = \bar{m} \cdot p_i^{e_i}$, where $p \not| \bar{m}$. 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

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

**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 $p_i$ of $qR$. Recall from 2.2 that a prime ideal $p_i$ dividing $p_iR$ has norm $\|R/p_i\| = p_i^{\ell}$, 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 $d \geq 2, w, q > 1, R = \mathcal{O}_m, \psi$ 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 $a' \leftarrow R_q^d$. 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 $R \in R_q^{d \times d}$ such that $Ra' = (1, 0, \ldots, 0)^t$, and sets $U = R^{-1}$.

It also picks a uniform element $s_0 \in R_q$. It then output as the first sample $(a', s_0)$. For each of the $w − 1$ remaining samples, we first sample $(a, b)$ from the given M–LWE oracle, pick a fresh uniformly random $d \in R_q$, and output $(U(d|a), b + (s_0 \cdot d)$, where the vertical bar denotes concatenation.
Conditioned on not aborting, we will verify that the output is distributed according to \( M\text{-LWE}_{R,d,m,q,\psi} \). 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 \((a,b)\) samples are uniform and independent. In this case it is easy to see that the outputs remain uniform, since \((d|a)\) is uniform, \( U \) is invertible, and each \( b \) is uniform and independent of everything else. Secondly, if the samples are from \( M\text{-LWE} \) with respect to a secret \( s \), it is easy to verify that the outputs are distributed according to the first-is-errorless variant \( M\text{-LWE}_{R,d,w,q,\psi} \) with secret \( s' = R^t(s_0|s) \mod q \). Since \( U \) is invertible modulo \( q \) and \((s_0|s)\) is uniform, so is \( s' \), 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 \( M\text{-LWE} \). We use this corollary in the next section to reduce to \( \text{Ext-M-LWE} \) with \( t \) hints.

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

In order for our reduction to “first-is-errorless” \( M\text{-LWE} \) to be meaningful when \( d - t \) is much smaller than linear in the underlying security parameter \( \kappa \) (as it will be in the case we are interested in for our reduction to \( M\text{-LWR} \)), we will require that all of the prime ideals of \( q \) 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 \( q = 3^t \). In fact, for \( R \) of index \( 2^k \), noting that \( Z^*_{2^k} = Z_2 \oplus Z_{2^k-2} \), 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” \( M\text{-LWE} \) to \( \text{Ext-M-LWE} \). 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 \( Z \subseteq K^d \) 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 $M$–LWE to $\text{Ext} \cdot M$–LWE that reduces the advantage by at most $16\epsilon$.

Proof. The reduction proceeds as follows. Given $z_1, \ldots, z_t \in \mathbb{Z}$, we compute the unimodular $U \in \mathbb{Z}^{d \times d}$ that can be efficiently computed as in Definition 3.2, and let $U' \in \mathbb{Z}^{d \times (d-t)}$ be the matrix formed by removing the $t$ leftmost columns of $U$. We receive $m$ samples $(a \in \mathbb{R}_q^{d \times w}, b \in \mathbb{R}_q^w)$ from our “first-t-errorless” $M$–LWE oracle. We then sample $f \leftarrow D_{\alpha}(\xi I - UU')^{1/2}$. The above distribution is well-defined since $\xi^2 I - UU'$ is a positive semidefinite matrix according to our assumption on $U$.

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

$$(A' = AU', b' = \tilde{b} + c, (\text{Tr}(\langle z_i, f + c \rangle)))_{i \in [d]} \in R_q^{d \times w} \times R_q^w \times \Omega_q^t$$

We now analyze correctness of the reduction. First, we have that $U$ is invertible, so that $A' = AU'$ is uniform solely over the choice of $A$. For the remainder of the proof we condition on any fixed value of $A'$ 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 $\tilde{b} = Ub + f = A' s + Uc + f$.

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

$$\langle z_i, f + c \rangle = \langle z_i, Uc + f + c \rangle,$$

so the third component gives the inner product of the noise with each vector $z_i$, as desired.

Finally, we consider the case where the samples received by our reduction are truly uniform. Then since $b$ is truly uniform and independent, we can view it as being distributed as $b = b + e$, where $b \in \mathbb{R}_q^w$ is truly uniform and $e$’s first $t$ coordinates are 0, and the remaining coordinates are chosen independently from $D_\alpha$. Then we have that $b' = Ub + Ue + f + c$, where $Ue + f + c$ is distributed exactly as in the case above. Since $U$ is unimodular, we have that $U(b + e)$ is uniform and independent over $\mathbb{R}_q^w$. Finally, since $Ue + f + c$ 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 \( \kappa \) 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 \( \psi \) be an arbitrary coordinate-wise \( B \)-bounded distribution over \( R \) for some \( B > 0 \). Let \( p, q, w = \text{poly}(\kappa), d \in \mathbb{N} \) such that \( q \geq 4eB\text{wn}\kappa p \). Let \( Z \subseteq R^d \) be the set of vectors \( z \) such that \( \text{Tr}(\langle z, e \rangle) \) extracts the \( i \)th coefficient of the \( j \)th ring element of \( e \) for some \( i \in [n], j \in [d] \).

If \( \exists \) a probabilistic polynomial-time \( A \) succeeding with advantage \( \epsilon(\kappa) \geq \kappa^{-c} / 2 \), then there exists a probabilistic polynomial-time algorithm \( A' \) succeeding with advantage \( \epsilon(\kappa w) / 4 \geq (\kappa w)^{-c} / 4 \) in distinguishing \( \text{Ext-M-LWE}_{R,d,w,q,\psi,Z,c} \) from uniform.

4.1 The Set BAD

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

**Definition 4.2.** For \( B,p,q \geq 1 \), we define
\[
\text{BAD}_{B,q,p} := \{ x \in Z_q : \lfloor x \rfloor_p \neq \lfloor x \pm B \rfloor_p \}
\]

The reason for the terminology is that, letting one of these elements represent a coefficient with noise (from a \( B \)-bounded distribution) in an \( M\text{-LWE} \) 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 \( Z_q \) is in \( \text{BAD}_{B,q,p} \).

**Lemma 4.3.** [AKPW13, Lemma 2.7]
\[
\Pr_{x \in Z_q} [x \in \text{BAD}_{B,q,p}] \leq 2Bp/q.
\]

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

**Lemma 4.4.** Let \( B, p, q, \kappa, n, w \) satisfy the conditions in Theorem 4.1. Let \( b \leftarrow R_q^w \) uniformly at random, and let \( S = \{ b_{i,j} \in \text{BAD}_{B,q,p} \} \). Then for all \( c \geq 1 \),
\[
\Pr[|S| \geq c] \leq \kappa^{-c} / 2.
\]

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

Let \( \delta = \frac{e}{\mu} - 1 > 0 \). Then
\[
\Pr[X > c] < \left( \frac{e^{\delta \mu}}{(1+\delta)^{1+\delta}} \right) = \frac{e^{\delta \mu}}{(1+\delta)^{c}} \leq \left( \frac{e}{1+\delta} \right)^{c} = \kappa^{-c}/(2c)^{c} \leq \kappa^{-c}/2.
\]
4.2 Reduction

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

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

where the $\ell_i = (i, j_i)$ are tuples representing the $j_i$th coefficient of the $i$th element in the error vector $e \in \mathbb{R}_w^q$.

Since the trace is a “universal” $\mathbb{Q}$-linear function, there exist (efficiently computable) vectors $z_{\ell_1}, z_{\ell_2}, \ldots, z_{\ell_c}$ such that $\text{Tr}(\langle e, z_i \rangle) = e_{\ell_i}$ for all $i$. We query the $\text{Ext-M-LWE}_{R,d,w,q,\psi,Z,c}$ oracle with these vectors, receiving back $(A \in \mathbb{R}_{d}^{w}, b \in \mathbb{R}_{w}^{w}, (e_{\ell_1}, \ldots, e_{\ell_c}))$.

Let

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

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

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

Now, if we have not aborted, set $\tilde{e}_{i,j} = e_{i,j}$ if $(i, j) \in GUESSES$, $\tilde{e}_{i,j} = 0$. Let $b' := [b - \tilde{e}]_p$, where the rounding is done element-wise, and let $y \leftarrow A(A', b')$. $A'$ outputs $y$ as its guess (where 0 corresponds to $\text{Ext-M-LWE}/\text{M-LWR}$, 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 $A$ (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 $A$ is within statistical distance at most $\epsilon/2$ of the true $\text{M-LWR}$ distribution if the oracle samples $b$ from $\text{Ext-M-LWE}$, and within distance at most $\epsilon/2$ of the uniform distribution if the oracles samples $b$ uniformly and independently.

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

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

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

**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 $\text{Ext-M-LWE}_{R,d,w,q,\psi,Z,c}$ succeeding with non-negligible probability, without having to rely on $A$. 

17
We can now upper bound the probability of abort event E.1.

Lemma 4.6.

\[ \Pr[E.1 \mid \text{oracle is Ext-M-LWE}] \leq \text{negl}(\kappa) + \Pr[E.1 \mid \text{oracle is uniform}] \leq \epsilon/2 \]

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 \((1 - \epsilon/2)/\binom{n_\text{w}}{c}\).

Lemma 4.7.

\[ \Pr[\neg(E.1 \lor E.2 \lor E.3)] \geq \frac{1 - \epsilon/2}{\binom{n_\text{w}}{c}}. \]

Moreover, for any \(A^* \in \mathbb{R}^{d \times w}, b^* \in \mathbb{R}^{q \times w} \in \mathbb{Z}_q^c\), we have

\[ \Pr((A, b) = (A^*, b^*) \mid \neg E.1) = \Pr((A, b) = (A^*, b^*) \mid \neg(E.1 \lor E.2 \lor E.3)) \]

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 \(\binom{n_\text{w} - t}{c-t}\) possible distinct sets GUESSES of indices that cover all elements in BAD, while there are \(\binom{n_\text{w}}{c}\) total possible distinct sets GUESSES of indices.

Since the elements of GUESSES are chosen uniformly and independent of everything else, we have that

\[ \Pr[\neg E.2 \mid \neg E.1 \land |BAD| = t] = \frac{\binom{n_\text{w} - t}{c-t}}{\binom{n_\text{w}}{c}} \]

Next, by the definition of E.3,

\[ \Pr[\neg E.3 \mid \neg(E.1 \lor E.2) \land |BAD| = t] = 1/\binom{n_\text{w} - t}{c-t} \]

\[ \Pr[\neg E.3 \land \neg E.2 \mid \neg E.1 \land |BAD| = t] = 1/\binom{n_\text{w}}{c} \]

independently of |BAD| and thus independently of the value of A, b received from the Ext-M-LWE oracle.

\[ \Pr[\neg E.3 \land \neg E.2 \mid \neg E.1] = 1/\binom{n_\text{w}}{c}. \]

Combining the above result with Lemma 4.6 gives

\[ \Pr[\neg(E.1 \lor E.2 \lor E.3)] \geq \frac{1 - \epsilon/2}{\binom{n_\text{w}}{c}} \]

Finally, we show that conditioned on not aborting, the samples sent to \(A\) are distributed sufficiently close to the desired distributions.
Lemma 4.8. The distribution of \((A, b')\) sent to \(A\) is within statistical distance at most \(\epsilon/2\) of the true \(\mathbb{M} - \text{LWR}\) distribution if the oracle sampled \(b\) from \(\text{Ext}-\mathbb{M}-\text{LWE}\), and uniform if it sampled it uniformly at random.

Proof. Since the reduction only ever sends \((A, b')\) if none of the abort events (in particular, \(E_2\)) happen, we receive as our \(\text{Ext}-\mathbb{M}-\text{LWE}\) hints the error coefficients for every \(b_{i,j} \in \mathbb{BAD}_{B,q,p}\). Let \(y = b - \tilde{e} \in \mathbb{R}_q^w\) (with \(\tilde{e}\) as defined above).

In the case the oracle samples \(b = A's + e\) from \(\text{Ext}-\mathbb{M}-\text{LWE}\), we have that every coefficient of every ring element is either errorless or it is “good”, meaning it is not in \(\mathbb{BAD}\). As a result, we have that

\[
[y]_p = [A's]_p,
\]

so, combining with Lemmas 4.6 and 4.7, we have that it is distributed within distance \(\epsilon/2\) of \(\mathbb{M}-\text{LWR}\).

In the case the oracle samples \(b\) uniformly at random and independently, we have that \(x\) is independent of \(b\), and so \(y\) remains truly uniform, making \([y]_p\) have the identical distribution to the “uniform” distribution.

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

Combining this with Lemma 4.7 gives

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

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


