On the Hardness of Module-LWE with Binary Secret

Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, Weiqiang Wen

On the Hardness of Module-LWE with Binary Secret

Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, and Weiqiang Wen

katharina.boudgoust@irisa.fr, corentin.jeudy@irisa.fr, adeline.roux-langlois@irisa.fr, weiqiang.wen@irisa.fr

Univ Rennes, CNRS, IRISA

Abstract. We prove that the *Module Learning With Errors* (M-LWE) problem with binary secrets and rank d is at least as hard as the standard version of M-LWE with uniform secret and rank k, where the rank increases from k to $d \geq (k+1)\log_2 q + \omega(\log_2 n)$, and the Gaussian noise from α to $\beta = \alpha \cdot \Theta(n^2\sqrt{d})$, where n is the ring degree and q the modulus. Our work improves on the recent work by Boudgoust et al. in 2020 by a factor of $\sqrt{md}$ in the Gaussian noise, where m is the number of given M-LWE samples, when q fulfills some number-theoretic requirements. We use a different approach than Boudgoust et al. to achieve this hardness result by adapting the previous work from Brakerski et al. in 2013 for the *Learning With Errors* problem to the module setting. The proof applies to cyclotomic fields, but most results hold for a larger class of number fields, and may be of independent interest.

Keywords: Lattice-based cryptography · module learning with errors · binary secret

1 Introduction

Lattice-based cryptography has become more and more popular over the past two decades as lattices offer a variety of presumed hard problems as security foundations for public-key cryptographic primitives. Lattices, which are discrete subgroups of the Euclidean space, provide several computational problems that are conjectured to be hard to solve with respect to both classical and quantum computers. One central problem is the *Shortest Vector Problem* (SVP), which asks to find a shortest non-zero vector from the given lattice. SVP also appears in a decisional variant (GapSVP), and its approximate counterpart (GapSVP_γ). The latter asks to decide if the norm of such a vector is less than a threshold r or greater than γr for a factor $\gamma \geq 1$. The security of most lattice-based primitives are however based on average-case problems, such as the *Learning With Errors* (LWE) problem introduced by Regev [Reg05,Reg09]. This problem emerges in two versions: its *search* variant asks to find the secret $\mathbf{s} \in \mathbb{Z}_q^n$ given samples of the form $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e)$, where $\mathbf{a}$ is uniform over $\mathbb{Z}_q^n$ and e a small

error over $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. The *decisional* variant asks to distinguish between such samples for a uniform $\mathbf{s} \in \mathbb{Z}_q^n$, and uniformly random samples in $\mathbb{Z}_q^n \times \mathbb{T}$. We use LWE to denote the latter. The error is usually sampled from a Gaussian distribution D_α of parameter $\alpha > 0$. The appeal of the LWE problem comes from its ties with well-known lattice problems like GapSVP_γ. It enjoys both quantum [Reg05] and classical [Pei09,BLP+13] worst-case to average-case reductions from GapSVP_γ, making it a firm candidate for cryptographic constructions. The LWE problem opened the way to a wide variety of simple to advanced cryptographic primitives ranging from public-key encryption [Reg05,GPV08,MP12], fully-homomorphic encryption [BGV12,BV14,DM15], recently to non-interactive zero-knowledge proofs [PS19], and many others.

Although LWE provides provably secure cryptosystems, all these schemes lack efficiency which motivates the research around structured variants. These variants gain in efficiency by considering the ring of integers of a number field (R-LWE) [LPR10,RSW18], a ring of polynomials (P-LWE) [SSTX09] or a module over a number field (M-LWE) [BGV12,LS15]. In this work, we focus on the latter as it offers a nice security-efficiency trade-off by bridging LWE and R-LWE. Let K be a number field of degree n and R its ring of integers. We use d to denote the module rank and q for the modulus. We also define the quotient ring $R_q = R/qR$, the real tensor field $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$ and the torus $\mathbb{T}_{R^\vee} = K_\mathbb{R}/R^\vee$, where $R^\vee$ is the dual ideal of R. The secret is now chosen in $(R_q^\vee)^d$, and the error from a distribution ψ over $K_\mathbb{R}$. The *Search*-M-LWE problem asks to recover the secret $\mathbf{s} \in (R_q^\vee)^d$ from arbitrarily many samples $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s} \rangle + e \bmod R^\vee)$, for $\mathbf{a}$ uniformly random over R_q^d and e sampled from ψ. In this work, we only consider the decisional variant denoted by M-LWE, where one has to distinguish such samples for a uniformly random secret $\mathbf{s} \in (R_q^\vee)^d$, from uniformly random samples in $R_q^d \times \mathbb{T}_{R^\vee}$. It also benefits from a worst-case to average-case reduction, first shown by Langlois and Stehlé [LS15] through a quantum reduction, and recently by Boudgoust et al. [BJRW20] through a classical reduction for a module rank $d \geq 2n$, where n is the ring degree. The underlying lattice problems are though restricted to *module lattices*, which correspond to finitely generated R-modules, where R is the ring of integers of a number field.

In practice, the LWE problem is often used with a *small* secret, i.e., Gaussian (Hermite-Normal-Form-LWE) or even binary (bin-LWE). The latter corresponds to choosing the secret $\mathbf{s}$ in $\{0,1\}^n$, and it is particularly interesting as it simplifies computations and thus increases efficiency. Modulus-rank switching techniques [BLP+13,AD17,WW19] rely on using small secrets as it keeps the noise blowup to a minimum. The binary secret variant also happens to be essential for some FHE schemes as in [DM15]. First studied by Goldwasser et al. [GKPV10], it is later improved by Brakerski et al. [BLP+13] and Micciancio [Mic18] using more technical proofs. Recent work by Brakerski and Döttling [BD20] extends the hardness to more general secret distributions. The question of whether these hardness results for bin-LWE carry over to the module setting was left open. As part of the proof of the classical hardness of M-LWE, a first reduction was proposed from M-LWE to bin-M-LWE using the Rényi divergence by Boudgoust et

al. [BJRW20]. The reduction increases the module rank from k to d by roughly a $\log_2 q$ factor, which allows to preserve the complexity of an exhaustive search, while increasing the noise by a factor $n^2 d\sqrt{m}$, where m is the number of samples, n the ring degree and d the final module rank. Another very recent paper by Lin et al. [LWW20] uses the noise lossiness argument from [BD20] to prove the hardness of M-LWE for general entropic distributions.

Our contributions. In this paper, we give an alternative approach to prove the hardness of M-LWE with binary secrets over cyclotomic fields. The result is summarized in an informal way in the following. For a more formal statement, we refer to Theorem 2.

Theorem 1 (Informal). *For a cyclotomic field of degree n, the* bin-M-LWE *problem with rank d and Gaussian parameter less than β is at least as hard as* M-LWE *with rank k and Gaussian parameter α, if $d \geq (k+1)\log_2 q + \omega(\log_2 n)$ and $\beta/\alpha = \Theta(n^2\sqrt{d})$, where q is a modulus such that the cyclotomic polynomial has a specific splitting behavior in $\mathbb{Z}_q[x]$.*

Note that the increase in the noise does not depend on the number of provided bin-M-LWE samples, in contrast to [BJRW20]. In the hope of achieving better parameters than [BJRW20], which is inspired by the proof of [GKPV10], we follow the proof idea of Brakerski et al. [BLP+13] by introducing the two intermediate problems first-is-errorless M-LWE and ext-M-LWE. We first reduce M-LWE to the first-is-errorless M-LWE variant, where the first sample is not perturbed by an error. We then reduce the latter to ext-M-LWE, which can be seen as M-LWE with an extra information on the error vector $\mathbf{e}$ given by $\langle \mathbf{e}, \mathbf{z} \rangle$ for a uniformly chosen $\mathbf{z}$ in the set of binary ring elements set $\mathcal{Z} = (R_2^\vee)^d$. In the work of Alperin-Sheriff and Apon [AA16] for their reduction from M-LWE to the deterministic variant Module Learning With Rounding, the authors introduce a variant of ext-M-LWE that gives $\mathrm{Tr}(\langle \mathbf{e}, \mathbf{z} \rangle)$ to the attacker instead. This variant is not suited for our reduction due to our lossy argument in Lemma 18. The field trace does not provide enough information to reconstruct $\mathbf{N}^T\mathbf{z}$ from the hint, where $\mathbf{N}$ is our Gaussian matrix. We discuss further the differences in Section 3.2 and 3.3. We then use a lossy argument, relying on the newly derived ext-M-LWE hardness assumption and a ring version of the leftover hash lemma, to reduce ext-M-LWE to bin-M-LWE. An overview of the full reduction is provided in Figure 1.

The main challenge is the use of matrices composed of ring elements. The proof in [BLP+13, Lem. 4.7] requires the construction of unimodular matrices which is not straightforward to adapt in the module setting because of invertibility issues. The construction in Lemma 15 relies on units of the quotient ring R/qR, which are much harder to describe than the units of $\mathbb{Z}/q\mathbb{Z}$ to say the least. This is the reason why we need to control the splitting structure of the cyclotomic polynomial modulo q. Lemma 2 [LS18, Thm. 1.1] solves this issue but requires q to satisfy certain number-theoretic properties and to be sufficiently large so that all the non-zero binary ring elements are units of R_q. The second complication comes from using both the coefficient embedding and the canonical

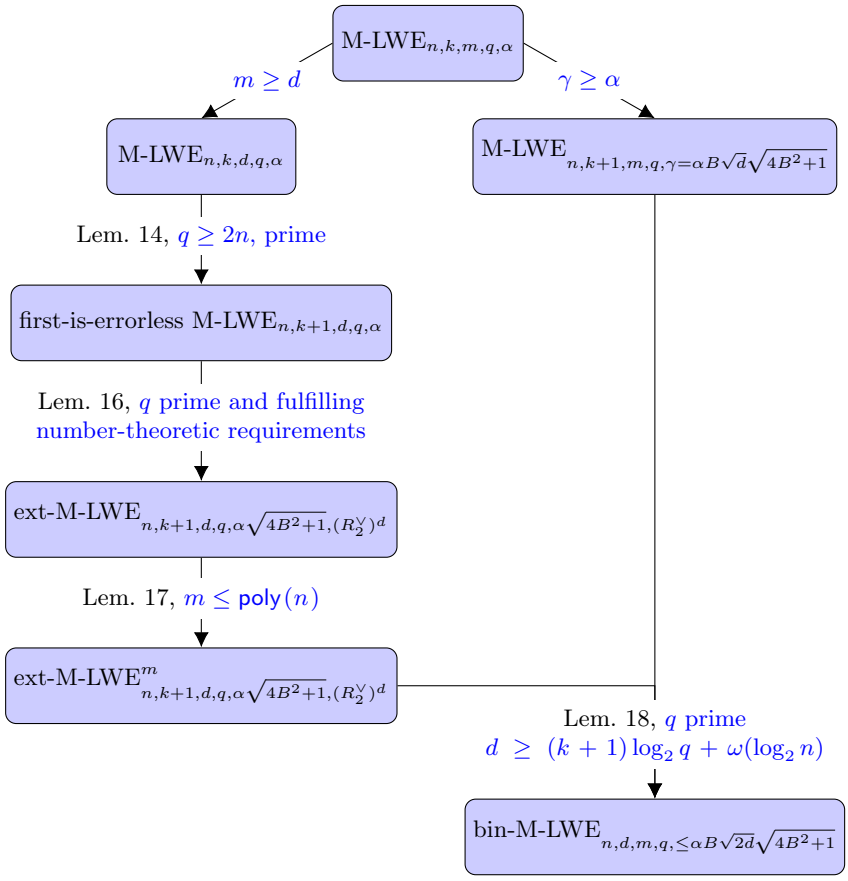

Fig. 1. Summary of the proof of Theorem 2, where $B = \max_{x\in R_2} \|\sigma(x)\|_\infty$ and σ is the canonical embedding. In cyclotomic fields, we have $B \leq n$. Note that Lemma 18 uses d samples from ext-M-LWE, where d is the module rank in bin-M-LWE. The assumptions on q concern the splitting behavior of the cyclotomic polynomial in $\mathbb{Z}_q[x]$, and are discussed in Section 3.3.

embedding. Even though some manipulations on Gaussian distributions require the use of the canonical embedding, we choose the secret to be binary in the coefficient embedding rather than the canonical embedding. As discussed in Section 3.1 for power-of-two cyclotomics, using the canonical embedding for binary secrets requires the rank d to be larger by a factor n than when using the coefficient embedding.

In the whole reduction, the ring degree n, number of samples m and modulus q are preserved, where m needs to be larger than d and q needs to be a prime satisfying certain number-theoretic properties. With the help of the modulus-switching technique of Langlois and Stehlé [LS15, Thm 4.8], we can then relax the restriction on the modulus q to be any polynomially large modu-

lus, at the expense of a loss in the Gaussian noise. The ranks must satisfy $d \geq (k+1)\log_2 q + \omega(\log_2 n)$, in the same manner as in [BJRW20]. However, our noise growth is smaller as our Gaussian parameter only increases by a factor $n\sqrt{2d}\sqrt{4n^2+1} = \Theta(n^2\sqrt{d})$ for cyclotomics. Our reduction removes the dependency in m in the noise ratio $n^2 d\sqrt{m}$ present in [BJRW20], which is more advantageous as we usually take $m = O(n \log_2 n)$ samples, and also gains an extra factor $\sqrt{d}$. As we directly show the hardness of decision bin-M-LWE one does not need the extra search-to-decision step in [BJRW20] which overall improves their classical hardness proof. Our result implies the hardness of M-LWE with a small (with respect to coefficients) secret and a moderate rank (e.g., $\omega(\log_2 n)$), which holds even with arbitrarily many samples. For a flexible choice of parameters (allowing efficiency optimizations), NIST candidates [BDK+18,DKL+18] considered M-LWE variants with a small secret and also a small rank, while restricting the number of samples to be small (e.g., linear in n) for ruling out the BKW type of attacks [KF15]. It is difficult to compare our result to the work by Lin et al. [LWW20] as their reduction does not use the coefficient embedding for the entropic secret distribution. Additionally, when bridging to LWE, the noise ratio is improved to $\sqrt{10d}$ as our construction in Lemma 15 matches the one from [BLP+13, Claim 4.6]. Our work thus matches the results from Brakerski et al. [BLP+13] when we take the ring R to be of degree 1.

The entire reduction is so far limited to cyclotomic fields due to Lemma 14 and 15. However, the other results are proven for a larger class of number fields, namely for the number fields $K = \mathbb{Q}(\zeta)$ such that their ring of integers is $R = \mathbb{Z}[\zeta]$, where ζ is an algebraic number. It ensures that R and its dual $R^\vee$ are linked by the equality $R^\vee = (f'(\zeta))^{-1}R$, where f is the minimal polynomial of ζ, and it also ensures the unique factorization of ideals. This class includes cyclotomic fields, quadratic fields $K = \mathbb{Q}(\sqrt{d})$ for square-free d with $d \neq 1 \bmod 4$, and number fields with f of square-free discriminant. These parts of the reduction can be extended to other number fields by using the quantity $B = \max_{x \in R_2} \|\sigma(x)\|_\infty$ introduced in Section 2.1, which we use throughout Section 3.3 and 3.4. The infinity norm here is simply the infinity norm over $\mathbb{C}^n$, and σ the canonical embedding. We discuss how to upper-bound B in Lemma 1, but in the case of cyclotomic fields we simply have $B \leq n$. Extending Lemma 14 and 15 to this broader class of number fields may however require additional constraints.

Open problems. In this paper, most of our results rely on the class of number fields $K = \mathbb{Q}(\zeta)$ where the ring of integers is $R = \mathbb{Z}[\zeta]$. Although this class includes all cyclotomic fields, we leave as an open problem to generalize these results to a larger class of number fields.

The leftover hash lemma used in the reduction of Lemma 18 requires the module rank d to be super-logarithmic in n. The proof of hardness of M-LWE with binary secret thus remains open for a lower module rank. In practice, a constant rank is used for increased efficiency, like the CRYSTALS-Kyber [BDK+18] candidate at the NIST standardization competition [NIS]. The interest in a lower module rank also stems from the extreme case $d = 1$ which corresponds to R-LWE. The hardness of bin-R-LWE remains an open problem.

The construction in Lemma 15 seems optimized in terms of its impact on the Gaussian parameter. However, its invertibility restricts the underlying number field, as well as the structure of the chosen modulus q. A better understanding of the unit group of R_q for general cyclotomic fields and other number fields might help relax the restrictions on the modulus q for the reduction to go through.

2 Preliminaries

Throughout the paper, q denotes a positive integer, $\mathbb{Z}$ denotes the set of integers and $\mathbb{Z}_q$ the integers modulo q. In a ring R, we write (p) for the principal ideal generated by $p \in R$, and R_p for the quotient ring $R/(p) = R/pR$. For simplicity, we denote by $[n]$ the set $\{1, \dots, n\}$ for any positive integer n. We denote the Kronecker symbol by $\delta_{i,j}$ which equals 1 if $i = j$ and 0 otherwise. The vectors are written in bold lowercase letters $\mathbf{a}$ and the matrices in bold uppercase letters $\mathbf{A}$. The transpose and Hermitian operators over vectors (resp. matrices) are respectively denoted by $\mathbf{a}^T$ (resp. $\mathbf{A}^T$) and $\mathbf{a}^\dagger$ (resp. $\mathbf{A}^\dagger$). The canonical basis of $\mathbb{C}^n$ is given by $\{\mathbf{e}_i\}_{i\in[n]}$, where $\mathbf{e}_i = [\delta_{i,k}]_{k\in[n]}$. For a vector $\mathbf{a} \in \mathbb{C}^n$, we define the matrix $\mathrm{diag}(\mathbf{a}) = [\delta_{i,j}a_i]_{i,j\in[n]}$ to be the diagonal matrix whose diagonal entries are the entries of $\mathbf{a}$. The identity matrix of size $n \times n$ is denoted by $\mathbf{I}_n$. For any $a \in \mathbb{C}^n$, we define the Euclidean norm as $\|\mathbf{a}\|_2 = \sqrt{\sum_{i\in[n]} |a_i|^2}$ and the infinity norm as $\|\mathbf{a}\|_\infty = \max_{i\in[n]} |a_i|$. We also define the *spectral norm* of any matrix $\mathbf{A} = [a_{i,j}]_{i\in[n],j\in[m]} \in \mathbb{C}^{n\times m}$ as $\|\mathbf{A}\|_2 = \max_{\mathbf{x}\in\mathbb{C}^m\setminus\{0\}} \|\mathbf{Ax}\|_2 / \|\mathbf{x}\|_2$, and the *max norm* as $\|\mathbf{A}\|_{\max} = \max_{i\in[n],j\in[m]} |a_{i,j}|$. For a complex number z, we denote by $\Re(z)$ its real component. We use the same notation for complex polynomials, where the real component is taken coefficient-wise. The statistical distance between two discrete distribution P and Q over a countable set S is defined as $\Delta(P, Q) = \frac{1}{2}\sum_{x\in S} |P(x) - Q(x)|$. It extends to continuous distributions replacing the sum by an integration. For a finite set S, we denote the uniform distribution over S by $U(S)$. The operation of sampling an element $x \in S$ according to a distribution P over S is denoted by $x \hookleftarrow P$, where the set S is implied. We also say that the random variables $X_1, \dots, X_k$ are i.i.d. from a distribution P if they are pair-wise independent and if they all are distributed according to P.

2.1 Algebraic number theory background

A complex number ζ is called an *algebraic number* if it is root of a polynomial over $\mathbb{Q}$. The monic polynomial f of minimal degree among such polynomials is called *the minimal polynomial* or *defining polynomial* of ζ, and is unique. If the minimal polynomial of ζ only has integer coefficients, then ζ is called an *algebraic integer*. A *number field* $K = \mathbb{Q}(\zeta)$ is the finite field extension of the rationals by adjoining the algebraic number ζ. Its degree is defined as the degree of the minimal polynomial of ζ. We define the tensor field $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$ which can be seen as the finite field extension of the reals by adjoining ζ.

The set of all algebraic integers in K is a ring called the *ring of integers*, and we denote it by R. We always have $\mathbb{Z}[\zeta] \subseteq R$, but only special classes of number

fields verify $\mathbb{Z}[\zeta] = R$. Among them, there are cyclotomic fields, which correspond to number fields where ζ is a primitive ν-th root of unity, for an integer ν. The ν-th cyclotomic number field has degree $n = \varphi(\nu)$, where φ is Euler's totient function. In this case, the minimal polynomial is $f = \Phi_\nu = \prod_{j\in[n]}(x - \alpha_j)$, where the α_j are the distinct primitive ν-th roots of unity. For the power-of-two cyclotomic field where $\nu = 2^{\ell+1}$, it yields $n = \varphi(\nu) = 2^\ell$, and $\Phi_\nu = x^n + 1$.

The space H. We use t_1 to denote the number of real roots of the minimal polynomial of the underlying number field, and t_2 the number of pairs of complex conjugate roots, which yields $n = t_1 + 2t_2$. The space $H \subseteq \mathbb{C}^n$ is defined by $H = \left\{\mathbf{x} \in \mathbb{R}^{t_1} \times \mathbb{C}^{2t_2} : \forall j \in [t_2], x_{t_1+t_2+j} = \overline{x_{t_1+j}}\right\}$. We can verify that H is a $\mathbb{R}$-vector space of dimension n with the columns of $\mathbf{H}$ as orthonormal basis, where

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_{t_1} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}\mathbf{I}_{t_2} & \frac{i}{\sqrt{2}}\mathbf{I}_{t_2} \\ 0 & \frac{1}{\sqrt{2}}\mathbf{I}_{t_2} & \frac{-i}{\sqrt{2}}\mathbf{I}_{t_2} \end{bmatrix}, \text{ with } \mathbf{I}_k \text{ the identity matrix of size } k.$$

Coefficient embedding. A number field $K = \mathbb{Q}(\zeta)$ of degree n can be seen as a $\mathbb{Q}$-vector space of dimension n with basis $\{1, \zeta, \ldots, \zeta^{n-1}\}$. Hence, every element $x \in K$ can be written as $x = \sum_{j=0}^{n-1} x_j \zeta^j$, with $x_j \in \mathbb{Q}$. The *coefficient embedding* is the isomorphism τ between K and $\mathbb{Q}^n$ that maps every $x \in K$ to its coefficient vector $\tau(x) = [x_0, \ldots, x_{n-1}]^T$. We also extend the coefficient embedding to $K_\mathbb{R}$, which yields an isomorphism between $K_\mathbb{R}$ and $\mathbb{R}^n$.

Canonical embedding. All the following definitions extend to $K_\mathbb{R}$ in the obvious way. A number field $K = \mathbb{Q}(\zeta)$ with defining polynomial f of degree n has exactly n field homomorphisms $\sigma_i : K \to \mathbb{C}$ that map ζ to each of the distinct roots of the defining polynomial. We denote by $\sigma_1, \ldots, \sigma_{t_1}$ the real embeddings (i.e. the embeddings that map ζ to one of the real roots of f) and $\sigma_{t_1+1}, \ldots, \sigma_{t_1+2t_2}$ the complex ones. Since f is in $\mathbb{Q}[x]$, the fundamental theorem of algebra states that the complex roots come as conjugate pairs, and therefore $\sigma_{t_1+t_2+j} = \overline{\sigma_{t_1+j}}$ for all $j \in [t_2]$. The *canonical embedding* σ is the field homomorphism from K to $\mathbb{C}^n$ defined as $\sigma(x) = \left[\sigma_1(x), \ldots, \sigma_n(x)\right]^T$, where the addition and multiplication of vectors is performed component-wise. The range of σ is a subset of H, and therefore we can map any $x \in K$ to $\mathbb{R}^n$ via the map σ_H defined by $\sigma_H(x) = \mathbf{H}^\dagger \cdot \sigma(x)$ for all $x \in K$. We also mention that the extension of σ to $K_\mathbb{R}$ is an isomorphism from $K_\mathbb{R}$ to H. Multiplication is no longer component-wise with σ_H but it can be described by a left multiplication, namely $\sigma_H(x \cdot y) = \mathbf{H}^\dagger \cdot \mathrm{diag}(\sigma(x)) \cdot \mathbf{H}\sigma_H(y)$, for any $x, y \in K$. Note that for any $x \in K$, $\mathbf{H}^\dagger \cdot \mathrm{diag}(\sigma(x)) \cdot \mathbf{H} \in \mathbb{R}^{n\times n}$, and has the $|\sigma_j(x)|$ as singular values.

We define the trace $\mathrm{Tr} : K \to \mathbb{Q}$ of K by $\mathrm{Tr}(x) = \sum_{j\in[n]} \sigma_j(x)$ for any $x \in K$. We use it to define the *dual* of R as $R^\vee = \{x \in K : \mathrm{Tr}(xR) \subseteq \mathbb{Z}\}$. For the class of number fields for which we have $R = \mathbb{Z}[\zeta]$, we have $R^\vee = \lambda^{-1}R$ where $\lambda = f'(\zeta) \in \mathbb{C}$. In particular, for power-of-two cyclotomics $\lambda = n$. We also define the norm $N : K \to \mathbb{Q}$ of K by $N(x) = \prod_{j\in[n]} \sigma_j(x)$ for any $x \in K$.

Distortion between embeddings. Both embeddings play important roles in this paper, and we recall how to go from one to the other. By applying σ to

an element $x = \sum_{i=0}^{n-1} x_i \zeta^i \in K$, we see that $\sigma(x)$ and $\tau(x)$ are linked through a linear operator which is the Vandermonde matrix of the roots of the defining polynomial f. For $j \in [n]$, we let $\alpha_j = \sigma_j(\zeta)$ be the j-th root of f. Then, we obtain that $\sigma(x) = \mathbf{V}\tau(x)$, where

$$\mathbf{V} = \begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ | & | & & | \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{bmatrix}.$$

This transformation does not necessarily carry the structure from one embedding to the other, e.g., a binary vector in the coefficient embedding need not to be binary in the canonical embedding. Changing the embedding also impacts the norm, which is captured by the inequalities $\left\|\mathbf{V}^{-1}\right\|_2^{-1} \|\tau(x)\|_2 \leq \|\sigma(x)\|_2 \leq \|\mathbf{V}\|_2 \|\tau(x)\|_2$. Hence, $\|\mathbf{V}\|_2$ and $\left\|\mathbf{V}^{-1}\right\|_2$ help approximating the distortion between both embeddings. Roşca et al. [RSW18] give additional insight on this distortion for specific number fields. Throughout this paper, we are interested in the parameter defined by $B = \max_{x \in R_2} \|\sigma(x)\|_\infty$ that is inherent to the ring. This parameter intervenes in the proof of Lemma 15 and 18, where we need an upper-bound on $\|\sigma(x)\|_\infty$, for $x \in R_2$, that is independent of x. Recall that if x is in R_2, then its coefficient vector $\tau(x)$ is in $\{0,1\}^n$. Here, we provide an upper-bound on B, that is further simplified for cyclotomic number fields.

Lemma 1. *Let K be a number field of degree n, and R its ring of integers. Let $\mathbf{V}$ be the transformation between both embeddings. Then, $B = \max_{x \in R_2} \|\sigma(x)\|_\infty \leq n \|\mathbf{V}\|_{\max}$. In particular, for cyclotomic fields, it yields $B \leq n$.*

Proof. We can express $x \in R_2$ as $x = \sum_{j=0}^{n-1} b_j \zeta^j$, with $b_j \in \{0,1\}$ for all j. Then, for $i \in [n]$, we have

$$|\sigma_i(x)| \leq \sum_{j=0}^{n-1} b_j \, |\sigma_i(\zeta)|^j = \sum_{j=0}^{n-1} b_j \, |\alpha_i|^j \leq \|\mathbf{V}\|_{\max} \sum_{j=0}^{n-1} b_j \leq n \, \|\mathbf{V}\|_{\max}.$$

Taking the maximum over all $i \in [n]$ yields $B \leq n \|\mathbf{V}\|_{\max}$. In the case of cyclotomic fields, the α_i are roots of unity and therefore, all the entries of $\mathbf{V}$ have magnitude 1. Hence $\|\mathbf{V}\|_{\max} = 1$ which yields $B \leq n$ in this case. □

Ideals and units. An *ideal* $\mathcal{I} \subseteq R$ is a non-zero additive subgroup of R that is closed under multiplication by R, i.e. for all $r, x \in R \times \mathcal{I}$, $r \cdot x \in \mathcal{I}$. An ideal $\mathcal{I}$ is *principal* if it is generated by a single element u, meaning $\mathcal{I} = uR = (u)$. A *fractional ideal* is a set $\mathcal{I}$ for which there exists an element $d \in R$ such that $d\mathcal{I}$ is an ideal of R. An ideal $\mathfrak{p} \neq R$ of R is *prime* if for all $a, b \in R$, $a \cdot b \in \mathfrak{p}$ implies that a or b is in $\mathfrak{p}$. The *product of ideals* $\mathcal{I}$ and $\mathcal{J}$ is the set of all finite sums of xy, where $x \in \mathcal{I}$ and $y \in \mathcal{J}$. We extend the field norm and define the norm of an ideal $N(\mathcal{I})$ as the index of $\mathcal{I}$ as an additive subgroup of R, which corresponds to $N(\mathcal{I}) = |R/\mathcal{I}|$. The norm is still multiplicative and

verifies $N((a)) = |N(a)|$ for any $a \in K$. For a fractional ideal $\mathcal{I}$, the norm is defined as $N(\mathcal{I}) = N(d\mathcal{I})/|N(d)|$. We also define the *dual* of an ideal $\mathcal{I}$ by $\mathcal{I}^\vee = \{x \in K : \mathrm{Tr}(x\mathcal{I}) \subseteq \mathbb{Z}\}$.
In the construction of Lemma 15, we need a condition for binary elements of $R_2 = R/(2)$ to be invertible in R_q for a specific q. To do so, we rely on the small norm condition proven in [LS18, Th. 1.1].

Lemma 2 (Th. 1.1 [LS18]). *Let K be the ν-th cyclotomic field, with $\nu = \prod_i p_i^{e_i}$ be its prime-power factorization, with $e_i \geq 1$. We denote R the ring of integers of K. Also, let $\mu = \prod_i p_i^{f_i}$ for any $f_i \in [e_i]$. Let q be a prime such that $q = 1 \bmod \mu$, and $ord_\nu(q) = \nu/\mu$, where ord_ν is the multiplicative order modulo ν. Then, any element y of $R_q = R/qR$ satisfying $0 < \|\tau(y)\|_\infty < q^{1/\varphi(\mu)}/\mathfrak{s}_1(\mu)$ is a unit in R_q, where $\mathfrak{s}_1(\mu)$ denotes the largest singular value of the Vandermonde matrix of the μ-th cyclotomic field.*

In the case where ν is a prime power, then so is μ and then [LPR13] states that $\mathfrak{s}_1(\mu) = \sqrt{\mu}$ if μ is odd, and $\mathfrak{s}_1(\mu) = \sqrt{\mu/2}$ otherwise. For more general cases, we refer to the discussions from Lyubashevsky and Seiler [LS18, Conj. 2.6]. We also refer to [LS18, Th. 2.5] that establishes the density of such primes q for specific values of ν and μ.
We also recall two results from [WW19] that we need in the proof of Lemma 14 to construct a matrix of $\mathbf{U} \in R_q^{k\times k}$ that is invertible in R_q, i.e., such that there exists a matrix $\mathbf{U}^{-1} \in R_q^{k\times k}$ that verifies $\mathbf{U}\mathbf{U}^{-1} = \mathbf{I}_k \bmod qR = \mathbf{U}^{-1}\mathbf{U}$. This requires the prime q to be unramified in the cyclotomic field, which comes down to it not dividing the discriminant Δ_K. In cyclotomics, this is equivalent to q not dividing ν. The condition from Lemma 2 subsumes this one as $q = 1 \bmod \mu$ entails that q is not a prime factor of ν. We say that the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_i \in R_q^k$ are R_q-linearly independent if for all $x_1, \ldots, x_i \in R_q$, $\sum_{j\in[i]} x_j\mathbf{a}_j = 0 \bmod qR$ implies $x_1 = \ldots = x_i = 0$.

Lemma 3 (Lem. 9 [WW19]). *Let K be the cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let q, k be positive integers such that q is a prime that verifies $q \geq n$ and $q \nmid \nu$. Then for any $i \in \{0, \ldots, k-1\}$ and R_q-linearly independent vectors $\mathbf{a}_1, \ldots, \mathbf{a}_i \in R_q^k$, the probability of sampling a vector $\mathbf{b} \hookleftarrow U(R_q^k)$ such that $\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{b}$ are R_q-linearly independent is at least $1 - \frac{n}{q}$.*

Lemma 4 (Lem. 18 [WW19]). *Let K be the cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let q, k be positive integers such that q is a prime that verifies $q \geq n$ and $q \nmid \nu$. Let $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_k] \in R_q^{k\times k}$. Then, $\mathbf{A}$ is invertible modulo qR if and only if $\mathbf{a}_1, \ldots, \mathbf{a}_k$ are R_q-linearly independent.*

2.2 Lattices

A *lattice* Λ is a discrete subgroup of $\mathbb{R}^n$. Since H is isomorphic to $\mathbb{R}^n$, we sometimes consider lattices that are discrete subgroups of H. Each lattice can be represented by a basis $\mathbf{B} = [\mathbf{b}_i]_{i\in[r]} \in \mathbb{R}^{n\times r}$, as the set of all integer linear combinations of the basis elements, i.e., $\Lambda = \sum_{i\in[r]} \mathbb{Z} \cdot \mathbf{b}_i$. The dimension of the

lattice is n and the rank is r. In this work, we only consider *full-rank* lattices, namely lattices for which $r = n$.
We define the *dual lattice* of a lattice Λ by $\Lambda^* = \{\mathbf{x} \in \mathrm{Span}(\Lambda) : \forall \mathbf{y} \in \Lambda, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}$. We denote by $\lambda_1^\infty(\Lambda)$ the *first minimum* of the lattice Λ with respect to the infinity norm, i.e., the infinity norm of a shortest non-zero vector of Λ. Any ideal $\mathcal{I}$ embeds into a lattice $\sigma(\mathcal{I})$ in H, and a lattice $\sigma_H(\mathcal{I})$ in $\mathbb{R}^n$, which we call *ideal lattices*. For an R-module $M \subseteq K^d$, $(\sigma, \ldots, \sigma)(M)$ is a lattice in H^d and $(\sigma_H, \ldots, \sigma_H)(M)$ is a lattice in $\mathbb{R}^{nd}$, both of which are called *module lattices*. The positive integer d is the module rank. To ease readability, we simply use $\mathcal{I}$ (resp. M) to denote the ideal lattice (resp. the module lattice). Note that the ideal lattice $\sigma(\mathcal{I}^\vee)$ corresponding to the dual ideal $\mathcal{I}$ is the same as the dual lattice up to complex conjugation, i.e., $\sigma(\mathcal{I}^\vee) = \overline{\sigma(\mathcal{I})^*}$. We also note that for the infinity norm, the first minimum of the module lattice $\mathcal{I}^d = \mathcal{I} \times \ldots \times \mathcal{I}$ is the same as the first minimum of the ideal lattice $\mathcal{I}$, i.e., $\lambda_1^\infty(\mathcal{I}^d) = \lambda_1^\infty(\mathcal{I})$. For a vector $\mathbf{x} \in K^d$, we define $\|\mathbf{x}\|_\infty = \max_{k \in [n], i \in [d]} |\sigma_k(x_i)|$, and $\|\mathbf{x}\|_{2,\infty} = \max_{k \in [n]} \sqrt{\sum_{i \in [d]} |\sigma_k(x_i)|^2}$.

2.3 Probabilities

Gaussian measures. For a positive definite matrix $\mathbf{\Sigma} \in \mathbb{R}^n$, a vector $\mathbf{c} \in \mathbb{R}^n$, we define the Gaussian function by $\rho_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) = \exp(-\pi(\mathbf{x}-\mathbf{c})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\mathbf{c}))$ for all $\mathbf{x} \in \mathbb{R}^n$. We extend this definition to the degenerate case, i.e., positive semi-definite, by considering the generalized Moore-Penrose inverse. For convenience, we use the same notation as the standard inverse. We then define the continuous Gaussian probability distribution by its density $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) = (\det(\mathbf{\Sigma}))^{-1/2} \rho_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x})$. By abuse of notation, we call $\mathbf{\Sigma}$ the covariance matrix, even if in theory the covariance matrix of $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}$ is $\mathbf{\Sigma}/(2\pi)$. If $\mathbf{\Sigma}$ is diagonal with diagonal vector $\mathbf{r}^2 \in (\mathbb{R}^+)^n$, we simply write $D_{\mathbf{c},\mathbf{r}}$, and if $\mathbf{c} = 0$, we omit it. When $\mathbf{\Sigma} = \alpha^2 \mathbf{I}_n$, we simplify further to $D_{\mathbf{c},\alpha}$. We then define the discrete Gaussian distribution by conditioning $\mathbf{x}$ to be in a lattice Λ, i.e. $\mathcal{D}_{\Lambda,\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) = D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) / D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\Lambda)$ for all $\mathbf{x} \in \Lambda$, and where $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\Lambda) = \sum_{\mathbf{y} \in \Lambda} D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{y})$.

The *smoothing parameter* of a lattice Λ denoted by $\eta_\varepsilon(\Lambda)$ for some $\varepsilon > 0$, introduced in [MR07], is the smallest $s > 0$ such that $\rho_{1/s}(\Lambda^* \setminus \{0\}) \leq \varepsilon$. It represents the smallest Gaussian parameter $s > 0$ such that the discrete Gaussian $\mathcal{D}_{\Lambda,\mathbf{c},s}$ behaves like a continuous Gaussian distribution. We recall the following bound on the smoothing parameter that we need throughout this paper.

Lemma 5 (Lem. 3.5 [Pei08]). *For an n-dimensional lattice Λ and $\varepsilon > 0$, we have $\eta_\varepsilon(\Lambda) \leq \sqrt{\ln(2n(1+1/\varepsilon))/\pi} / \lambda_1^\infty(\Lambda^*)$.*

Lemma 6 (Lem. 4.1 [MR07]). *Let Λ be an n-dimensional lattice, $\varepsilon > 0$, and $\alpha > \eta_\varepsilon(\Lambda)$. Then the distribution of the coset $\mathbf{e} + \Lambda$, where $\mathbf{e} \hookleftarrow D_\alpha$, is within statistical distance $\varepsilon/2$ of the uniform distribution over the cosets of Λ.*

We now extend a result on the sum of a continuous Gaussian and a discrete one to more general Gaussian distributions. In particular, the lemma works for two elliptical Gaussians, which we use in the proof of Lemma 11.

Lemma 7 (Adapted from Lem. 2.8 [LS15] & Claim 3.9 [Reg09]). *Let Λ be an n-dimensional lattice, $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{R}, \mathbf{S}$ two positive semi-definite matrices of $\mathbb{R}^{n\times n}$, and $\mathbf{T} = \mathbf{R} + \mathbf{S}$. We also define $\mathbf{U} = \left(\mathbf{R}^{-1} + \mathbf{S}^{-1}\right)^{-1}$, and we assume that $\rho_{\sqrt{\mathbf{U}^{-1}}}(\Lambda^* \setminus \{0\}) \leq \varepsilon$ for some $\varepsilon \in (0, 1/2)$. Consider the distribution Y on $\mathbb{R}^n$ obtained by adding a discrete sample from $\mathcal{D}_{\Lambda+\mathbf{a},\sqrt{\mathbf{R}}}$ and a continuous sample from $D_{\sqrt{\mathbf{S}}}$. Then we have $\Delta(Y, D_{\sqrt{\mathbf{T}}}) \leq 2\varepsilon$.*

Proof. The density function Y is given by

$$\begin{aligned}
Y(\mathbf{x}) &= \sum_{\mathbf{y}\in\Lambda+\mathbf{a}} \mathcal{D}_{\Lambda+\mathbf{a},\sqrt{\mathbf{R}}}(\mathbf{y}) D_{\sqrt{\mathbf{S}}}(\mathbf{x}-\mathbf{y}) \\
&= \frac{1}{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}(\Lambda)\sqrt{\det \mathbf{S}}} \sum_{\mathbf{y}\in\Lambda+\mathbf{a}} \rho_{\sqrt{\mathbf{R}}}(\mathbf{y})\rho_{\sqrt{\mathbf{S}}}(\mathbf{x}-\mathbf{y}) \\
&= \frac{1}{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}(\Lambda)\sqrt{\det \mathbf{S}}} \sum_{\mathbf{y}\in\Lambda+\mathbf{a}} \rho_{\sqrt{\mathbf{T}}}(\mathbf{x})\rho_{\mathbf{R}\mathbf{T}^{-1}\mathbf{x},\sqrt{\mathbf{U}}}(\mathbf{y}) \quad \text{[Pei10, Fact 2.1].} \\
&= \frac{\rho_{\sqrt{\mathbf{T}}}(\mathbf{x})}{\sqrt{\det \mathbf{T}}} \cdot \frac{\sqrt{\det \mathbf{T}}\rho_{\mathbf{R}\mathbf{T}^{-1}\mathbf{x},\sqrt{\mathbf{U}}}(\Lambda)}{\sqrt{\det \mathbf{T}}\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}(\Lambda)} \\
&= D_{\sqrt{\mathbf{T}}}(\mathbf{x}) \cdot \frac{(\sqrt{\det \mathbf{R}}\sqrt{\det \mathbf{S}}/\sqrt{\det \mathbf{T}})^{-1}\widehat{\rho_{\mathbf{x}',\sqrt{\mathbf{U}}}}(\Lambda^*)}{(\sqrt{\det \mathbf{R}})^{-1}\widehat{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}}(\Lambda^*)},
\end{aligned}$$

where $\mathbf{x}' = \mathbf{R}\mathbf{T}^{-1}\mathbf{x}$, and $\widehat{f}$ denotes the Fourier transform of f. First notice that $(\det \mathbf{R}\cdot\det \mathbf{S})/\det \mathbf{T} = 1/\det(\mathbf{R}^{-1}\mathbf{T}\mathbf{S}^{-1}) = 1/\det \mathbf{U}^{-1}$. Moreover, recalling that $\widehat{\rho_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}}(\mathbf{w}) = \sqrt{\det \mathbf{\Sigma}} e^{-2i\pi\langle \mathbf{c},\mathbf{w}\rangle}\rho_{\sqrt{\mathbf{\Sigma}^{-1}}}(\mathbf{w})$, we get

$$\left|1 - (\sqrt{\det \mathbf{U}})^{-1}\widehat{\rho_{\mathbf{x}',\sqrt{\mathbf{U}}}}(\Lambda^*)\right| \leq \rho_{\sqrt{\mathbf{U}^{-1}}}(\Lambda^* \setminus \{0\}) \leq \varepsilon.$$

For the denominator, we first notice that for two positive semi-definite matrices $\mathbf{A}$ and $\mathbf{B}$, if $\mathbf{A} - \mathbf{B}$ is positive semi-definite, then $\rho_{\sqrt{\mathbf{A}}}(\mathbf{w}) \geq \rho_{\sqrt{\mathbf{B}}}(\mathbf{w})$ for all $\mathbf{w} \in \mathbb{R}^n$. Since $\mathbf{U}^{-1} - \mathbf{R}^{-1} = \mathbf{S}^{-1}$ is positive semi-definite, it yields $\rho_{\sqrt{\mathbf{R}^{-1}}}(\Lambda^* \setminus \{0\}) \leq \rho_{\sqrt{\mathbf{U}^{-1}}}(\Lambda^* \setminus \{0\}) \leq \varepsilon$. Therefore, using the same method as above, we get

$$\left|1 - (\sqrt{\det \mathbf{R}})^{-1}\widehat{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}}(\Lambda^*)\right| \leq \rho_{\sqrt{\mathbf{R}^{-1}}}(\Lambda^* \setminus \{0\}) \leq \varepsilon.$$

which leads to

$$\frac{(\sqrt{\det \mathbf{R}}\sqrt{\det \mathbf{S}}/\sqrt{\det \mathbf{T}})^{-1}\widehat{\rho_{\mathbf{x}',\sqrt{\mathbf{U}}}}(\Lambda^*)}{(\sqrt{\det \mathbf{R}})^{-1}\widehat{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}}(\Lambda^*)} \in \left[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}\right] \subseteq [1-2\varepsilon, 1+4\varepsilon],$$

assuming that $\varepsilon < 1/2$. We thus end up with $\left|Y(\mathbf{x}) - D_{\sqrt{\mathbf{T}}}(\mathbf{x})\right| \leq 4\varepsilon D_{\sqrt{\mathbf{T}}}(\mathbf{x})$. Integration and factor $1/2$ of the statistical distance yield the lemma. □

Lemma 8 (Lem. 2.10 [BLP+13] & Thm. 3.1 [Pei10]). *Let Λ be an n-dimensional lattice, $\varepsilon \in (0, 1/2]$, and $\beta, r > 0$ such that $r \geq \eta_\varepsilon(\Lambda)$. Then the distribution of $\mathbf{x}+\mathbf{y}$, obtained by first sampling $\mathbf{x}$ from D_β, and then $\mathbf{y}$ sampled from $\mathcal{D}_{\Lambda,\mathbf{x},r}$, is within statistical distance 8ε of $\mathcal{D}_{\Lambda,\sqrt{\beta^2+r^2}}$.*

Module Gaussians. In this section we define Gaussian distributions over R-modules $M \subseteq K_{\mathbb{R}}^d$, where $K = \mathbb{Q}(\zeta)$ is a number field, R its ring of integers, and $K_{\mathbb{R}} = K \otimes_{\mathbb{Q}} \mathbb{R}$. We need to consider the real tensor field $K_{\mathbb{R}}$ as the canonical embedding is an isomorphism between $K_{\mathbb{R}}$ and H but not between R and H, nor K and H. Gaussian distributions over $K_{\mathbb{R}}$ have been introduced alongside the R-LWE problem in [LPR10], and then generalized and used in most papers dealing with structured variants of LWE. We define general Gaussian distributions over $K_{\mathbb{R}}^d$ through their embedding to $\mathbb{R}^{nd}$, namely sampling $\mathbf{y}^{(H)} \in \mathbb{R}^{nd}$ according to $D_{\sqrt{\mathbf{\Sigma}}}$ for some positive semi-definite matrix $\mathbf{\Sigma}$ in $\mathbb{R}^{nd \times nd}$ and then mapping it back to $K_{\mathbb{R}}^d$ by $\mathbf{y} = \sigma_H^{-1}(\mathbf{y}^{(H)})$. To ease readability, we denote the described distribution of $\mathbf{y} \in K_{\mathbb{R}}^d$ by $D_{\sqrt{\mathbf{\Sigma}}}$.

In the proof of Lemma 16, we also need to identify the distribution of $\mathbf{y} = \mathbf{U}\mathbf{e}$ for an arbitrary matrix $\mathbf{U}$ and a Gaussian vector $\mathbf{e} \in K_{\mathbb{R}}^d$ for which the components are independent of each other. To do so, we need the ring homomorphism $\theta : K_{\mathbb{R}}^{k \times \ell} \to \mathbb{C}^{nk \times n\ell}$ defined by

$$\theta(\mathbf{A}) = \begin{bmatrix} \mathbf{D}_{1,1} & - & \mathbf{D}_{1,\ell} \\ | & \diagdown & | \\ \mathbf{D}_{k,1} & - & \mathbf{D}_{k,\ell} \end{bmatrix}, \text{ with } \mathbf{D}_{i,j} = \mathrm{diag}(\sigma(a_{i,j})) \in \mathbb{C}^{n \times n}.$$

Lemma 9. *Let K be a number field of degree n, and d a positive integer. Let $\mathbb{S} \in \mathbb{R}^{nd \times nd}$ be a positive semi-definite matrix, and $\mathbf{U} \in K_{\mathbb{R}}^{d \times d}$. We denote $\mathbf{\Sigma} = \left(\mathbb{H}^\dagger \theta(\mathbf{U}) \mathbb{H}\right) \mathbb{S} \left(\mathbb{H}^\dagger \theta(\mathbf{U}) \mathbb{H}\right)^\dagger \in \mathbb{R}^{nd \times nd}$, where $\mathbb{H} = \mathrm{diag}(\mathbf{H}, \ldots, \mathbf{H}) \in \mathbb{C}^{nd \times nd}$, with $\mathbf{H}$ the matrix form of the basis of the space H previously defined. Then, the distribution of $\mathbf{y} = \mathbf{U}\mathbf{e}$, where $\mathbf{e} \in K_{\mathbb{R}}^d$ is distributed according to $D_{\sqrt{\mathbb{S}}}$, is exactly $D_{\sqrt{\mathbf{\Sigma}}}$.*

Proof. Let $\mathbf{e} = [e_i]_{i \in d} \in K_{\mathbb{R}}^d$ be a Gaussian vector distributed according to $D_{\sqrt{\mathbb{S}}}$. For all $i \in [d]$, we have $y_i = \sum_{j \in [d]} u_{i,j} e_j$ and thus $\sigma(y_i) = \sum_{j \in [d]} \sigma(u_{i,j}) \odot \sigma(e_j)$, where $\odot$ denotes the Hadamard product. The Hadamard product $\mathbf{a} \odot \mathbf{b}$ of two vectors $\mathbf{a}$ and $\mathbf{b}$ can also be expressed as the matrix-vector product $\mathrm{diag}(\mathbf{a}) \cdot \mathbf{b}$. It results in

$$\sigma(\mathbf{y}) = \begin{bmatrix} \sigma(y_1) \\ | \\ \sigma(y_d) \end{bmatrix} = \theta(\mathbf{U}) \sigma(\mathbf{e}),$$

where $\theta(\mathbf{U})$ is the block matrix $[\mathrm{diag}(\sigma(u_{i,j}))]_{i,j \in [d]} \in \mathbb{C}^{nd \times nd}$. As we have seen before, we can decompose σ on the basis of H and get $\sigma(y_i) = \mathbf{H}\mathbf{y}_i^{(H)}$ (respectively $\sigma(e_i) = \mathbf{H}\mathbf{e}_i^{(H)}$) for all $i \in [d]$. By using the block matrix product, we end up with

$$\sigma(\mathbf{y}) = \begin{bmatrix} \mathbf{H} & & \\ & \diagdown & \\ & & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^{(H)} \\ | \\ \mathbf{y}_d^{(H)} \end{bmatrix} = \mathbb{H}\mathbf{y}^{(H)}.$$

Thus $\mathbb{H}\mathbf{y}^{(H)} = \theta(\mathbf{U})\mathbb{H}\mathbf{e}^{(H)}$, which leads to $\mathbf{y}^{(H)} = \mathbb{H}^\dagger \theta(\mathbf{U}) \mathbb{H} \mathbf{e}^{(H)}$. Now notice that the blocks of $\mathbb{H}^\dagger \theta(\mathbf{U}) \mathbb{H}$ are the $\mathbf{H}^\dagger \mathrm{diag}(\sigma(u_{i,j}))\mathbf{H}$ which correspond to the

matrix form of the multiplication by $u_{i,j}$ in the basis of the space H and thus is in $\mathbb{R}^{n\times n}$. Hence $\mathbb{H}^\dagger\theta(\mathbf{U})\mathbb{H} \in \mathbb{R}^{nd\times nd}$.

By definition, $\mathbf{e}^{(H)}$ is distributed according to $D_{\sqrt{\mathbb{S}}}$. Thus $\mathbf{y}^{(H)}$ is also distributed along a 0-centered Gaussian over $\mathbb{R}^{nd}$, but with covariance matrix

$$\mathbf{\Sigma} = \left(\mathbb{H}^\dagger\theta(\mathbf{U})\mathbb{H}\right)\mathbb{S}\left(\mathbb{H}^\dagger\theta(\mathbf{U})\mathbb{H}\right)^\dagger.$$

□

In particular, when $\mathbb{S} = \mathrm{diag}(r_1^2,\dots,r_1^2,\dots,r_d^2,\dots,r_d^2)$ for some positive reals $r_1,\dots,r_d$, then $\sqrt{\mathbb{S}}$ commutes with $\mathbb{H}$ and the covariance simplifies to $\mathbf{\Sigma} = \mathbb{H}^\dagger\widetilde{\mathbb{U}}\widetilde{\mathbb{U}}^\dagger\mathbb{H}$, with $\widetilde{\mathbb{U}} = [\mathrm{diag}(\sigma(r_j u_{i,j}))]_{i,j\in[d]}$. We also need two other lemmata related to the inner product of $K_\mathbb{R}^d$ (which results in an element of $K_\mathbb{R}$) between a Gaussian vector and an arbitrary one. In particular, we use Lemma 11 in the proof of Lemma 18 in order to decompose a Gaussian noise into an inner product.

Lemma 10 (Lem. 2.13 [LS15]). *Let $\mathbf{r} \in (\mathbb{R}^+)^n \cap H$, $\mathbf{z} \in K^d$ fixed and $\mathbf{e} \in K_\mathbb{R}^d$ sampled from $D_{\sqrt{\mathbf{\Sigma}}}$, where $\sqrt{\mathbf{\Sigma}} = [\delta_{i,j}\mathrm{diag}(\mathbf{r})]_{i,j\in[d]} \in \mathbb{R}^{nd\times nd}$. Then $\langle \mathbf{z}, \mathbf{e}\rangle = \sum_{i\in[d]} z_i e_i$ is distributed according to $D_{\mathbf{r}'}$ with $r'_j = r_j\sqrt{\sum_{i\in[d]}|\sigma_j(z_i)|^2}$.*

Lemma 11 (Adapted from Cor. 3.10 [Reg09]). *Let $M \subset K^d$ be an R-module (yielding a module lattice), let $\mathbf{u}, \mathbf{z} \in K^d$ be fixed, and let $\beta, \gamma > 0$. Assume that $(1/\beta^2 + \|\mathbf{z}\|_{2,\infty}^2/\gamma^2)^{-1/2} \geq \eta_\varepsilon(M)$ for some $\varepsilon \in (0, 1/2)$. Then the distribution of $\langle \mathbf{z},\mathbf{v}\rangle + e$ where $\mathbf{v}$ is sampled from $\mathcal{D}_{M+\mathbf{u},\beta}$ and $e \in K_\mathbb{R}$ is sampled from D_γ, is within statistical distance at most 2ε from the elliptical Gaussian $D_\mathbf{r}$ over $K_\mathbb{R}$, where $r_j = \sqrt{\beta^2\sum_{i\in[d]}|\sigma_j(z_i)|^2 + \gamma^2}$ for $j \in [n]$.*

Proof. Consider $\mathbf{h} \in (K_\mathbb{R})^d$ distributed according to $D_{\mathbf{r}',\dots,\mathbf{r}'}$, where $\mathbf{r}'$ is given by $r'_j = \gamma/\sqrt{\sum_{i\in[d]}|\sigma_j(z_i)|^2}$ for $j \in [n]$. Then by Lemma 10, $\langle \mathbf{z},\mathbf{h}\rangle$ is distributed as D_γ and therefore $\Delta(\langle \mathbf{z},\mathbf{v}\rangle + e, D_\mathbf{r}) = \Delta(\langle \mathbf{z},\mathbf{v}+\mathbf{h}\rangle, D_\mathbf{r})$. Now, we denote $\mathbf{t}$ such that $t_j = \sqrt{\beta^2 + (r'_j)^2}$ for $j \in [n]$. Note that by assumption

$$\begin{aligned}\min_{j\in[n]} \beta r'_j/t_j &= (1/\beta^2 + \max_{j\in[n]}\sum_{i\in[d]}|\sigma_j(z_i)|^2/\gamma^2)^{-1/2}\\ &= (1/\beta^2 + \|\mathbf{z}\|_{2,\infty}^2/\gamma^2)^{-1/2} \geq \eta_\varepsilon(M).\end{aligned}$$

Lemma 7 therefore applies and yields that $\mathbf{v}+\mathbf{h}$ is distributed as $D_{\mathbf{t},\dots,\mathbf{t}}$, within statistical distance at most 2ε. By applying once more Lemma 10 and noticing that the statistical distance does not increase when applying a function (here the scalar product with $\mathbf{z}$), then we get that $\langle \mathbf{z}, \mathbf{v}+\mathbf{h}\rangle$ is distributed as $D_\mathbf{r}$ within statistical distance at most 2ε, where $r_j = t_j\sqrt{\sum_{i\in[d]}|\sigma_j(z_i)|^2} = \sqrt{\beta^2\sum_{i\in[d]}|\sigma_j(z_i)|^2 + \gamma^2}$ for $j \in [n]$. □

2.4 Ring Leftover Hash Lemma

The proof of Lemma 18 also requires a leftover hash lemma over rings, where the vector contains binary polynomials. We use the following adaption of [Mic07] proven by Boudgoust et al. [BJRW20].

Lemma 12 (Lem. 7 [BJRW20]). *Let q be prime and n, k and d be positive integers. Further, let f be the defining polynomial of degree n of the number field $K \cong \mathbb{Q}[x]/(f)$ such that its ring of integers is given by $R = \mathbb{Z}[x]/(f)$. We set $R_q = R/qR$ and $R_2 = R/2R$. Then, $\Delta((\mathbf{C}, \mathbf{Cz}), (\mathbf{C}, \mathbf{s})) \leq \frac{1}{2}\sqrt{\left(1 + \frac{q^k}{2^d}\right)^n - 1}$, where $\mathbf{C} \hookleftarrow U((R_q)^{k\times d})$, $\mathbf{z} \hookleftarrow U((R_2)^d)$ and $\mathbf{s} \hookleftarrow U((R_q)^k)$.*

2.5 Module Learning With Errors

The LWE problem over modules was first defined by Brakerski et al. [BGV12] and studied at length by Langlois and Stehlé [LS15]. We consider a number field K of degree n, R its ring of integers, and let d denote the module rank. Let ψ be a distribution on $K_\mathbb{R}$ and $\mathbf{s} \in (R_q^\vee)^d$ be a vector. We let $A_{\mathbf{s},\psi}^{(R^d)}$ denote the distribution on $(R_q)^d \times \mathbb{T}_{R^\vee}$ obtained by choosing a vector $\mathbf{a} \hookleftarrow U((R_q)^d)$, an element $e \leftarrow \psi$ and returning $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod R^\vee)$.

Definition 1. *Let q, d be positive integers with $q \geq 2$. Let Υ be a distribution on a family of distributions on $K_\mathbb{R}$. The problem* M-LWE$_{n,d,q,\Upsilon}$ *is as follows: Sample $\mathbf{s} \hookleftarrow U((R_q^\vee)^d)$ and $\psi \hookleftarrow \Upsilon$. The goal is to distinguish between arbitrarily many independent samples from $A_{\mathbf{s},\psi}^{(R^d)}$ and the same number of independent samples from $U((R_q)^d \times \mathbb{T}_{R^\vee})$. If the number of samples m is fixed, we denote it by* M-LWE$_{n,d,m,q,\Upsilon}$.

When the error distribution is a Gaussian distribution of parameter $\alpha > 0$, we write M-LWE$_{n,d,m,q,\alpha}$, and if the Gaussian is elliptical bounded by β, i.e., $D_\mathbf{r}$ for $\mathbf{r} \in (\mathbb{R}^+)^n$ such that $\|\mathbf{r}\|_\infty \leq \beta$, we write M-LWE$_{n,d,m,q,\leq\beta}$. The same goes for other variants of M-LWE. For the M-LWE problem and its variants that we introduce later, we denote by $\mathrm{Adv}[\mathcal{A}]$ the advantage of an adversary $\mathcal{A}$ in distinguishing between the two distributions of the problem.

Binary secret. Another possibility is to change the distribution of the secret. We focus on the case where the secret is chosen to be binary in the coefficient embedding. We thus define bin-M-LWE$_{n,d,m,q,\Upsilon}$ to be the M-LWE problem where the secret $\mathbf{s}$ is sampled uniformly in $(R_2^\vee)^d$. We justify this choice of embedding in Section 3.1.

3 Hardness of M-LWE with binary secret

In this section, we prove our main contribution which is a reduction from M-LWE with rank k to bin-M-LWE with rank d satisfying $d \geq (k+1)\log_2 q + \omega(\log_2 n)$,

for cyclotomic fields. The reduction preserves the modulus q, that needs to be prime satisfying number-theoretic restrictions, the ring degree n and the number of samples m, but the noise is increased by a factor of $n\sqrt{2d}\sqrt{4n^2+1}$. Our proof follows the same idea as in [BLP+13] that we adapt over modules. The noise ratio is polynomial in n, but smaller than $n^2 d\sqrt{m}$ in [BJRW20]. Not only does it no longer depend on the number of samples m, which becomes more advantageous as the typical choice for m is $m = O(n \log_2 n)$, but we also gain a factor of $\sqrt{d}$. For the reduction, m also needs to be larger than the target module rank d, and at most polynomial in n because of the hybrid argument used in Lemma 17. The reduction in Theorem 2 works for all cyclotomic fields, but most results apply for all number fields $K = \mathbb{Q}(\zeta)$ such that the ring of integers is $R = \mathbb{Z}[\zeta]$, the bottleneck being the construction in Lemma 15.

Theorem 2. *Let $\nu = \prod_i p_i^{e_i}$, K be the cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $\mu = \prod_i p_i$ and q be a prime number such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > \max(2n, \mathfrak{s}_1(\mu)^{\varphi(\mu)})$, where $\mathfrak{s}_1(\mu)$ denotes the largest singular value of the Vandermonde matrix of the μ-th cyclotomic field. Further, let k, d, m be three positive integers such that $d \geq (k+1)\log_2 q + \omega(\log_2 n)$, and $d \leq m \leq \mathsf{poly}(n)$. Let $\alpha \geq q^{-1}\sqrt{\ln(2nd(1+1/\varepsilon))/\pi}$ and $\beta \geq \alpha \cdot n\sqrt{2d}\sqrt{4n^2+1}$. Then there is a reduction from* M-LWE$_{n,k,m,q,\alpha}$ *to* bin-M-LWE$_{n,d,m,q,\leq\beta}$*, such that if $\mathcal{A}$ solves the latter with advantage* Adv$[\mathcal{A}]$*, then there exists an algorithm $\mathcal{B}$ that solves the former with advantage*

$$\mathrm{Adv}[\mathcal{B}] \geq \frac{1}{3m}\left(\mathrm{Adv}[\mathcal{A}] - \frac{1}{2}\sqrt{\left(1+\frac{q^{k+1}}{2^d}\right)^n - 1}\right) - \frac{37\varepsilon}{2}.$$

The noise ratio β/α contains three main terms. The factor n encapsulates the norm distortion between the coefficient and the canonical embedding, as well as the actual length of the binary vectors. The second term $\sqrt{2d}$ stems from the masking of $\mathbf{z}$ when introduced in the first hybrid in the proof of Lemma 18. The last factor $\sqrt{4n^2+1}$ solely represents the impact of giving information on the error in the ext-M-LWE problem.

3.1 Choice of embedding for binary secrets

As mentioned in the introduction, the variant of M-LWE using a binary secret requires the choice of an embedding in which the secret is binary. As praised in [LPR10,LPR13], the canonical embedding has nice algebraic and geometric properties that make it a good choice of embedding. However, in this section, we justify our choice of the coefficient embedding, by analyzing the set of secrets that are binary in the canonical embedding in the case of power-of-two cyclotomics. The conjugation symmetry of the canonical embedding first restricts the choice of secrets to $(\sigma^{-1}(\{0,1\}^n \cap H))^d$, where d denotes the module rank and the space H is the range of σ. In addition, the tightest worst-case to average-case reductions for M-LWE require $\mathbf{s}$ to be taken from $(R_q^\vee)^d$. However, σ^{-1} maps H

to $K_{\mathbb{R}}$ but not necessarily to R or to $R^\vee$. We thus have to further restrict the set of secrets to

$$\mathcal{Z} = (R_q^\vee \cap \sigma^{-1}(\{0, \lambda^{-1}\}^n \cap H))^d,$$

where λ is such that $R^\vee = \lambda^{-1} R$. In the case of power-of-two cyclotomics, $\lambda = n$ is real and therefore yields $\lambda\mathcal{Z} = (R_q \cap \sigma^{-1}(\{0,1\}^n \cap H))^d$.

Lagrange Basis. As opposed to R_2 which corresponds to binary vectors in the coefficient embedding, the power basis is not adapted to describe the set $\lambda\mathcal{Z}$. We thus introduce the Lagrange basis. We denote by $\alpha_j = \sigma_j(\zeta)$ the j-th root of the defining polynomial f. Recall that we assume that α_j is real for $j \in [t_1]$, and that we have $\alpha_{t_1+j} = \overline{\alpha_{t_1+t_2+j}} \in \mathbb{C}$ for $j \in [t_2]$. Applying σ_j to an element $r = \sum_{i=0}^{n-1} r_i \zeta^i \in K_{\mathbb{R}}$ comes down to evaluating the polynomial $p_r = \sum_{i=0}^{n-1} r_i x^i$ at α_j. We use this polynomial interpretation to define elements of $K_{\mathbb{R}}$ that form a basis of $\sigma^{-1}(\{0,1\}^n \cap H)$.
Lagrange interpolation defines polynomials that map a set of distinct elements to 0 and 1. Since the α_j are distinct as f is irreducible, we can apply a similar method and define

$$L_k = \prod_{j \in [n] \setminus \{k\}} \frac{x - \alpha_j}{\alpha_k - \alpha_j},$$

for $k \in [t_1]$, which is real due to the conjugation symmetry of the roots. For $k \in \{t_1 + 1, \ldots, t_1 + t_2\}$, we define

$$L_k = \prod_{j \in [n] \setminus \{k\}} \frac{x - \alpha_j}{\alpha_k - \alpha_j} + \prod_{j \in [n] \setminus \{k+t_2\}} \frac{x - \alpha_j}{\alpha_{k+t_2} - \alpha_j} = 2\mathfrak{R}\left(\prod_{j \in [n] \setminus \{k\}} \frac{x - \alpha_j}{\alpha_k - \alpha_j} \right).$$

Hence the polynomials lie in $\mathbb{R}[x]$ and we have $L_k(\alpha_j) = \delta_{k,j}$ for $(k,j) \in [t_1] \times [n]$, and $L_k(\alpha_j) = \delta_{k,j} + \delta_{k+t_2,j}$ for $(k,j) \in \{t_1 + 1, \ldots, t_1 + t_2\} \times [n]$.
Therefore, by defining the Lagrange basis $\boldsymbol{l}$ with the corresponding $l_k \cong L_k(\zeta) \in K_{\mathbb{R}}$, we have linear independence and $\sigma^{-1}(\{0,1\}^n \cap H) = \sum_{k \in [t_1+t_2]} \{0,1\} \cdot l_k$, because $\sigma(l_k) = \mathbf{e}_k$ if $k \in [t_1]$ and $\sigma(l_k) = \mathbf{e}_k + \mathbf{e}_{k+t_2}$ if $k \in \{t_1 + 1, \ldots, t_1 + t_2\}$. As far as we are aware, this is the first time that the Lagrange basis is used in the setting of structured lattice-based cryptography. We now need to determine which of these combinations lie in R_q in order to properly define the set of secrets.

Power-of-two cyclotomics. We now look at the Lagrange basis in the specific case where n is a power of two.

Lemma 13. *Let R be the cyclotomic ring of integers of degree $n = 2^\ell$. Then, for any integer $q \geq 1$, the set $R_q \cap \sigma^{-1}(\{0,1\}^n \cap H)$ contains only 0 and 1.*

Proof. Recall that in cyclotomic fields, we have $t_1 = 0$ and $t_2 = n/2$. We know that the defining polynomial is $x^n + 1$ and therefore we can re-index the roots as $\alpha_j = \exp(i(2j+1)\pi/n)$, j now ranging from 0 to $n-1$. We can therefore study the complex product. We look at the constant coefficient of L_k, i.e., $A_k = L_k(0) = 2\mathfrak{R}\left(\prod_{0 \leq j < n, j \neq k} \frac{-\alpha_j}{\alpha_k - \alpha_j}\right)$. To ease notation, we write $j \neq k$ instead

of $j \in \{0, \ldots, n-1\} \setminus \{k\}$ for the product indexes. We first look at the product for a fixed $k \in \{0, \ldots, n/2-1\}$.

$$\prod_{j \neq k}(\alpha_k - \alpha_j) = \alpha_k^{n-1} \prod_{j \neq k}(1 - \alpha_j/\alpha_k) = -\alpha_k^{-1} \prod_{j \neq k}(1 - e^{i2\pi(j-k)/n})$$
$$= -\alpha_k^{-1} \prod_{l=1}^{n-1}(1 - e^{i2\pi l/n}),$$

using the fact that $\alpha_k^n + 1 = 0$ and the circularity of the complex exponential. Yet, we also have $\prod_{l=0}^{n-1}(x - e^{i2\pi l/n}) = x^n - 1 = (x-1)\sum_{l=0}^{n-1} x^l$. By simplifying both sides by $x-1$ and then evaluating at 1, we have $\prod_{l=1}^{n-1}(1-e^{i2\pi l/n}) = \sum_{l=0}^{n-1} 1^l = n$. The product of the numerators in the definition of A_k is $(-1)^{n-1}\overline{\alpha_k}$ because we can pair all of the roots α_j with their conjugates, which gives $\alpha_j\overline{\alpha_j} = |\alpha_j|^2 = 1$, except for $\overline{\alpha_k}$. Hence, $A_k = 2\Re(-\overline{\alpha_k}/(-n/\alpha_k))$ because n is even, which yields $A_k = \frac{2}{n}$. Now we take a subset $S \subseteq \{0, \ldots, n/2-1\}$ and we study $\sum_{k \in S} L_k$. Note that the case of $S = \{0, \ldots, n/2-1\}$ corresponds to adding all the Lagrange basis elements which results in 1, and the case $S = \emptyset$ results in 0 by convention. So we now assume that $0 < |S| < n/2$. The constant coefficient of $\sum_{k \in S} L_k$ is $2|S|/n \in (0, 1)$ and is therefore not an integer. Hence, $\sum_{k \in S} L_k \notin \mathbb{Z}[x]$ which means that the element $\sum_{k \in S} l_k$ is not in R nor R_q for any $q \geq 1$.

It proves that the only binary combination of the Lagrange basis that are in R are 0 and 1, and the same conclusion is valid for R_q for any $q \geq 1$. □

Hence to preserve the complexity of a brute force attack when comparing the two embeddings, the module rank would have to be increased by a factor n in the case where we take the canonical embedding to represent binary secrets. In this case, the (dual of the) secrets are from $\{0,1\}^d$ and therefore discard most of the available ring structure as opposed to R_2^d. We remark that this issue hasn't been addressed by [LWW20]. It seems that for too narrow bounds on the entropic secret distribution, the number of available secrets is much smaller in the canonical embedding compared to the number with regard to the coefficient embedding.

3.2 First-is-errorless M-LWE

We follow the same idea as Brakerski et al. [BLP+13] by gradually giving more information to the adversary while proving that this additional information does not increase the advantage too much. We define the module version of *first-is-errorless* LWE, from [BLP+13], where the first equation is given without error. A similar definition and reduction from M-LWE are given in [AA16]. The only difference between the two reductions comes from the pre-processing step, which is simplified in our case due to the further restrictions on q of our overall reduction.

Definition 2 (First-is-errorless M-LWE). *Let K be a number field of degree n and R its ring of integers. Let q, k be positive integers. We denote by $R_q = R/qR$, $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$, and $\mathbb{T}_{R^\vee} = K_\mathbb{R}/R^\vee$ as usual.*

Let Υ be a distribution over a family of distributions over $K_{\mathbb{R}}$. The first-is-errorless *variant of the* M-LWE *problem is to distinguish between the following cases. On the one hand, the first sample is uniform over $(R_q)^k \times q^{-1}R^\vee/R^\vee$ and the rest are uniform over $(R_q)^k \times \mathbb{T}_{R^\vee}$. On the other hand, there is some unknown $\mathbf{s}$ uniformly sampled over $(R_q^\vee)^k$ and ψ sampled from Υ such that the first sample is from $A^{(R^k)}_{\mathbf{s},\{0\}}$ and the rest are distributed as $A^{(R^k)}_{\mathbf{s},\psi}$, where $\{0\}$ is the distribution that is deterministically 0.*
We denote it by first-is-errorless M-LWE$_{n,k,q,\Upsilon}$ *or, when the number of samples m is fixed,* first-is-errorless M-LWE$_{n,k,m,q,\Upsilon}$.

Lemma 14 (Adapted from Lem. 4.3 [BLP+13]). *Let K be the cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $q \geq 2n$ be a prime integer such that $q \nmid \nu$, k a positive integer, and Υ a distribution over a family of distributions over $K_{\mathbb{R}}$. There is a polynomial-time reduction from* M-LWE$_{n,k-1,q,\Upsilon}$ *to the variant* first-is-errorless M-LWE$_{n,k,q,\Upsilon}$.

Proof. The reduction first chooses $\mathbf{a}' \hookleftarrow U((R_q)^k)$ and then $\mathbf{b}_2, \ldots, \mathbf{b}_k$ i.i.d. from $U((R_q)^k)$ such that $\mathbf{a}', \mathbf{b}_2, \ldots, \mathbf{b}_k$ are R_q-linearly independent. Each time we draw a uniformly random column, the probability that the new column is R_q-linearly independent with the previous ones is at least $1 - n/q$ for $q \geq n$ by Lemma 3. Since we require $q \geq 2n$, this probability is at least $1/2$. Therefore, we only need a polynomial number of uniformly sampled columns in R_q^k to construct a matrix of $R_q^{k\times k}$ invertible modulo qR.

The preprocessing step results in a matrix $\mathbf{U} = \left[\mathbf{a}', \mathbf{b}_2, \ldots, \mathbf{b}_k\right] \in (R_q)^{k\times k}$ that is invertible modulo qR according to Lemma 4. Then, sample s_0 uniformly in $R_q^\vee$. The reduction is as follows. For the first sample, it outputs $(\mathbf{a}', q^{-1}\cdot s_0 \bmod R^\vee) \in (R_q)^k \times q^{-1}R^\vee/R^\vee$. The other samples are produced by taking $(\mathbf{a}, b) \in (R_q)^{k-1}\times\mathbb{T}_{R^\vee}$ from the M-LWE challenger, picking a fresh randomly chosen $a'' \in R_q$, and outputting $(\mathbf{U}(a''|\mathbf{a}), b + q^{-1}(s_0 \cdot a'') \bmod R^\vee) \in (R_q)^k \times \mathbb{T}_{R^\vee}$, with the vertical bar denoting concatenation. We now analyze correctness. First note that the first component is uniform over $(R_q)^k$. Indeed, $\mathbf{a}'$ is uniform over $(R_q)^k$ for the first sample, and since $\mathbf{a}$ is uniform over $(R_q)^{k-1}$, a'' is uniform over R_q, and $\mathbf{U}$ is invertible in $(R_q)^{k\times k}$, then $\mathbf{U}(a''|\mathbf{a})$ is uniform over $(R_q)^k$ as well.

If b is uniform, the first sample yields $q^{-1}s_0 \bmod R^\vee$ uniform over $q^{-1}R^\vee/R^\vee$. For the other samples, $b + q^{-1}(s_0 \cdot a'') \bmod R^\vee$ is uniform over $\mathbb{T}_{R^\vee}$ and independent of $\mathbf{U}(a''|\mathbf{a})$ but also independent from the first sample because b masks $q^{-1}(s_0 \cdot a'')$. If $b = q^{-1}\langle\mathbf{a},\mathbf{s}\rangle + e \bmod R^\vee$ for some uniform $\mathbf{s} \in (R_q^\vee)^{k-1}$ and $e \hookleftarrow \psi$ for some $\psi \hookleftarrow \Upsilon$, then $q^{-1}s_0 = q^{-1}\langle\mathbf{e}_1, (s_0|\mathbf{s})\rangle = q^{-1}\langle\mathbf{U}\mathbf{e}_1, \mathbf{U}^{-T}(s_0|\mathbf{s})\rangle = q^{-1}\langle\mathbf{a}', \mathbf{U}^{-T}(s_0|\mathbf{s})\rangle$, where $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$. For the other samples, we have

$$\begin{aligned} b + q^{-1}(s_0 \cdot a'') \bmod R^\vee &= q^{-1}\langle\mathbf{a},\mathbf{s}\rangle + q^{-1}(s_0 \cdot a'') + e \bmod R^\vee \\ &= q^{-1}\langle(a''|\mathbf{a}),(s_0|\mathbf{s})\rangle + e \bmod R^\vee \\ &= q^{-1}\langle\mathbf{U}(a''|\mathbf{a}),\mathbf{U}^{-T}(s_0|\mathbf{s})\rangle + e \bmod R^\vee. \end{aligned}$$

Note that $(s_0|\mathbf{s})$ is uniform over $(R_q^\vee)^k$ so $\mathbf{U}^{-T}(s_0|\mathbf{s})$ is also uniform over $(R_q^\vee)^k$ because $\mathbf{U}^{-T}$ is invertible in R_q. Therefore the reduction outputs samples according to first-is-errorless M-LWE with secret $\mathbf{s}' = \mathbf{U}^{-T}(s_0|\mathbf{s})$. □

3.3 Extended M-LWE

We now define the module version of the *Extended* LWE problem introduced in [BLP+13], where the adversary is allowed a hint on the errors. As opposed to [AA16], we allow for multiple secret and one single hint vector $\mathbf{z}$, as required by our final reduction of Lemma 18.

Definition 3 (Extended M-LWE). *Let K be a number field of degree n, and R its ring of integers. Let m, q, k, t be positive integers. Let $\mathcal{Z} \subseteq (R^\vee)^m$ and ψ a discrete distribution over $q^{-1}(R^\vee)^m$. The* Extended M-LWE *problem, denoted by* ext-M-LWE$^t_{n,k,m,q,\psi,\mathcal{Z}}$*, is as follows. The algorithm first samples $\mathbf{z} \in \mathcal{Z}$ and then receives a tuple $(\mathbf{A}, (\mathbf{b}_i)_{i\in[t]}, (\langle \mathbf{e}_i, \mathbf{z}\rangle)_{i\in[t]})$, over $(R_q)^{k\times m} \times \left((q^{-1}R^\vee/R^\vee)^m\right)^t \times (q^{-1}R^\vee)^t$. Its goal is to distinguish between the following cases.*
On one side, $\mathbf{A}$ is sampled uniformly over $(R_q)^{k\times m}$, and for all $i \in [t], \mathbf{e}_i \in q^{-1}(R^\vee)^m$ are independent and identically distributed from ψ, and define $\mathbf{b}_i = q^{-1}\mathbf{A}^T\mathbf{s}_i + \mathbf{e}_i \mod R^\vee$ for some uniformly chosen $\mathbf{s}_i \in (R_q^\vee)^k$.
On the other side, everything is identical except that the $\mathbf{b}_i$ are sampled uniformly over $(q^{-1}R^\vee/R^\vee)^m$, independently from $\mathbf{A}$ and the error vectors.

For simplicity in what follows, for a matrix $\mathbf{A} \in R^{m\times m}$, we denote by $\mathbf{A}^\perp \in R^{m\times(m-1)}$ the submatrix of $\mathbf{A}$ obtained by removing the leftmost column. Our reduction from first-is-errorless M-LWE to ext-M-LWE in Lemma 16 requires the construction of a matrix $\mathbf{U}_\mathbf{z} \in R^{m\times m}$, for all vectors $\mathbf{z} \in \mathcal{Z} = (R_2^\vee)^m$, satisfying several properties. This matrix allows us to transform samples from a first-is-errorless M-LWE challenger into samples that we can give to an oracle for ext-M-LWE. The largest singular value of its submatrix $\mathbf{U}_\mathbf{z}^\perp$ (when embedded with θ), controls the increase in the Gaussian parameter. We propose a construction for which we bound the largest singular value above by a quantity independent on $\mathbf{z}$, as needed in the reduction.

Lemma 15. *Let $\nu = \prod_i p_i^{e_i}$, K be the cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $\mu = \prod_i p_i$ and q be a prime number such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > \mathfrak{s}_1(\mu)^{\varphi(\mu)}$, where $\mathfrak{s}_1(\mu)$ denotes the largest singular value of the Vandermonde matrix of the μ-th cyclotomic field. Finally, let m be a positive integer, and $\mathcal{Z} = (R_2^\vee)^m$, and we recall the ring parameter $B = \max_{x\in R_2} \|\sigma(x)\|_\infty$. For all $\mathbf{z} \in \mathcal{Z}$, there is an efficiently computable matrix $\mathbf{U}_\mathbf{z} \in R^{m\times m}$ that is invertible modulo qR and that verifies the following: $\mathbf{z}$ is orthogonal to the columns of $\mathbf{U}_\mathbf{z}^\perp$, and the largest singular value of $\theta(\mathbf{U}_\mathbf{z}^\perp) \in \mathbb{C}^{mn\times(m-1)n}$ is at most $\xi = 2B$.*

Proof. Recall that for these number fields, we have $R_p^\vee = \lambda^{-1}R_p$ for any $p \in \mathbb{Z}$ with $\lambda = f'(\zeta)$. Let $\mathbf{z} \in \mathcal{Z}$ and denote $\widetilde{\mathbf{z}} = \lambda\mathbf{z} \in R_2^m$. First, we construct $\mathbf{U}_\mathbf{z}$

in the case where all the $\widetilde{z}_i$ are non-zero. To do so, we define the intermediate matrices $\mathbf{A}$, and $\mathbf{B}$ of $R^{m\times m}$, all unspecified entries being zeros:

$$\mathbf{U_z} = \begin{bmatrix} 1 & -\widetilde{z}_2 & & \\ \widetilde{z}_1 & \ddots & & \\ & & \ddots & -\widetilde{z}_m \\ & & \widetilde{z}_{m-1} & \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & \widetilde{z}_1 & & \\ & & \ddots & \\ & & & \widetilde{z}_{m-1} \end{bmatrix} + \begin{bmatrix} 0 & -\widetilde{z}_2 & & \\ & \ddots & & \\ & & \ddots & -\widetilde{z}_m \\ & & & 0 \end{bmatrix}$$

$$\mathbf{U_z^\perp} \qquad\qquad \mathbf{A}^\perp \qquad\qquad \mathbf{B}^\perp$$

The matrix $\mathbf{U_z}$ is invertible in modulo qR only if all the $\widetilde{z}_i$ (except $\widetilde{z}_m$) are in $R_q^\times$. Yet, since they are all non-zero binary polynomials (elements of R_2), we have that for all i in $[m]$, $\|\tau(\widetilde{z}_i)\|_\infty = 1$, where τ is the coefficient embedding. By Lemma 2, since q verifies the algebraic conditions taking all $f_i = 1$ and $q^{1/\varphi(\mu)}/\mathfrak{s}_1(\mu) > 1$, all the $\widetilde{z}_i$ are in $R_q^\times$.

By construction, the last $m-1$ columns of $\mathbf{U_z}$ are orthogonal to $\widetilde{\mathbf{z}}$. Let $\mathbf{U_z^\perp}$ be the submatrix of $\mathbf{U_z}$ obtained by removing the leftmost column as shown above. Since θ is a ring homomorphism, we have $\theta(\mathbf{U_z^\perp}) = \theta(\mathbf{A}^\perp) + \theta(\mathbf{B}^\perp)$. We now need to bound the spectral norm of these two matrices, and use the triangle inequality to conclude. For any vector $\mathbf{x} \in \mathbb{C}^{(m-1)n}$, we have that $\left\|\theta(\mathbf{A}^\perp)\mathbf{x}\right\|_2 = \sqrt{\sum_{i\in[m-1]} \sum_{j\in[n]} |\sigma_j(\widetilde{z}_i)|^2 \left|x_{j+n(i-1)}\right|^2} \leq B \|\mathbf{x}\|_2$, because each $\widetilde{z}_i$ is in R_2. This yields $\left\|\theta(\mathbf{A}^\perp)\right\|_2 \leq B$. A similar calculation on $\mathbf{B}^\perp$ leads to $\left\|\theta(\mathbf{B}^\perp)\right\|_2 \leq B$, thus resulting in $\left\|\theta(\mathbf{U_z^\perp})\right\|_2 \leq 2B$.

Now assume that $\widetilde{z}_{i_0}, \ldots, \widetilde{z}_m$ are zeros for some i_0 in $[m]$. If the zeros do not appear last in the vector $\widetilde{\mathbf{z}}$, we can replace $\widetilde{\mathbf{z}}$ with $\mathbf{S}\widetilde{\mathbf{z}}$, where $\mathbf{S} \in R^{m\times m}$ swaps the coordinates of $\widetilde{\mathbf{z}}$ so that the zeros appear last. Since $\mathbf{S}$ is unitary, it preserves the singular values as well as invertibility. Then, the construction remains the same except that the $\widetilde{z}_{i_0}, \ldots, \widetilde{z}_m$ on the diagonal are replaced by 1. The orthogonality is preserved, and $\left\|\theta(\mathbf{U_z^\perp})\right\|_2$ can still be bounded above by $2B$. □

Notice that when the ring is of degree 1, the constructions in the different cases match the ones from [BLP+13, Claim 4.6]. So do the singular values as $B \leq n = 1$ by Lemma 1. Also, the construction differs from the notion of quality in [AA16] due to the discrepancies between the two definitions of ext-M-LWE.

Lemma 16 (Adapted from Lem. 4.7 [BLP+13]). *Let $\nu = \prod_i p_i^{e_i}$, K be the cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $\mu = \prod_i p_i$ and q be a prime such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > \mathfrak{s}_1(\mu)^{\varphi(\mu)}$, where $\mathfrak{s}_1(\mu)$ denotes the largest singular value of the Vandermonde matrix of the μ-th cyclotomic field. Let m, k positive integers, $\mathcal{Z} = (R_2^\vee)^m$, $\varepsilon \in (0, 1/2)$ and $\alpha \geq q^{-1}\sqrt{\ln(2mn(1+1/\varepsilon))/\pi}$. Then, there is a probabilistic reduction from* first-is-errorless M-LWE$_{n,k,m,q,\alpha}$ *to* ext-M-LWE$_{n,k,m,q,\alpha\sqrt{4B^2+1},\mathcal{Z}}$ *that reduces the advantage by at most $33\varepsilon/2$, where $B = \max_{x\in R_2} \|\sigma(x)\|_\infty$.*

Note that by the transference theorems, we have $\lambda_1^\infty(R) \geq N(R)^{1/n} = 1$. So, using the fact that $(q\Lambda)^* = q^{-1}\Lambda^*$, we have

$$\lambda_1^\infty((q^{-1}(R^\vee)^m)^*) = \lambda_1^\infty(q((R^\vee)^m)^*) = q\lambda_1^\infty(((R^\vee)^m)^*) = q\lambda_1^\infty(R) \geq q,$$

which together with Lemma 5 yields $q^{-1}\sqrt{\ln(2mn(1+1/\varepsilon))/\pi} \geq \eta_\varepsilon(q^{-1}(R^\vee)^m)$.

Proof. Assume we have access to an oracle $\mathcal{O}$ for ext-M-LWE$_{n,k,m,q,\alpha\sqrt{\xi^2+1},\mathcal{Z}}$. We take m samples from the first-is-errorless challenger, resulting in

$$(\mathbf{A}, \mathbf{b}) \in (R_q)^{k\times m} \times ((q^{-1}R^\vee/R^\vee) \times \mathbb{T}_{R^\vee}^{m-1}).$$

Assume we need to provide samples to $\mathcal{O}$ for some $\mathbf{z} \in \mathcal{Z}$. By Lemma 15 we can efficiently compute a matrix $\mathbf{U_z} \in R^{m\times m}$ that is invertible modulo qR, such that its submatrix $\mathbf{U_z^\perp}$ is orthogonal to $\mathbf{z}$, and that $\theta(\mathbf{U_z^\perp})$ has largest singular value less than $\xi = 2B$. The reduction first samples $\mathbf{f} \in K_\mathbb{R}^m$ from the continuous Gaussian distribution of covariance matrix $\alpha^2(\xi^2\mathbf{I}_{mn} - \mathbb{H}^\dagger\theta(\mathbf{U_z^\perp})\theta(\mathbf{U_z^\perp})^\dagger\mathbb{H}) \in \mathbb{R}^{mn\times mn}$, where $\mathbb{H}$ is defined as in Section 2.3. Note that $\mathbb{H}$ is unitary and therefore preserves the largest singular value. The reduction then computes $\mathbf{b'} = \mathbf{U_z}\mathbf{b} + \mathbf{f}$ and samples $\mathbf{c}$ from $\mathcal{D}_{q^{-1}(R^\vee)^m - \mathbf{b'},\alpha}$, and finally gives the following to $\mathcal{O}$

$$(\mathbf{A'} = \mathbf{A}\mathbf{U_z}^T, \mathbf{b'} + \mathbf{c} \bmod R^\vee, \langle \mathbf{z}, \mathbf{f} + \mathbf{c}\rangle).$$

Note that this tuple is in $(R_q)^{k\times m} \times (q^{-1}R^\vee/R^\vee)^m \times q^{-1}R^\vee$, as required. We now prove correctness. First, consider the case where $\mathbf{A}$ is uniformly random over $R_q^{k\times m}$ and $\mathbf{b} = q^{-1}\mathbf{A}^T\mathbf{s} + \mathbf{e} \bmod R^\vee$ for some uniform $\mathbf{s} \in (R_q^\vee)^k$, and $\mathbf{e}$ sampled from $\{0\} \times D_\alpha^{m-1}$ where $\{0\}$ denotes the distribution that is deterministically 0. Since $\mathbf{U_z}$ is invertible modulo qR, $\mathbf{A'} = \mathbf{A}\mathbf{U_z}^T$ is also uniform over $(R_q)^{k\times m}$ as required. From now on we condition on an arbitrary $\mathbf{A'}$ and analyze the distribution of the remaining components. We have

$$\begin{aligned}\mathbf{b'} &= q^{-1}\mathbf{U_z}\mathbf{A}^T\mathbf{s} + \mathbf{U_z}\mathbf{e} + \mathbf{f}\\ &= q^{-1}(\mathbf{A'})^T\mathbf{s} + \mathbf{U_z}\mathbf{e} + \mathbf{f}.\end{aligned}$$

Since the first coefficient of $\mathbf{e}$ is deterministically 0 the first column is ignored in the covariance matrix, and then $\mathbf{U_z}\mathbf{e}$ is distributed as the Gaussian over $K_\mathbb{R}^m$ of covariance matrix $\alpha^2\mathbb{H}^\dagger\theta(\mathbf{U_z^\perp})\theta(\mathbf{U_z^\perp})^\dagger\mathbb{H}$ by Lemma 9. Hence the vector $\mathbf{U_z}\mathbf{e}+\mathbf{f}$ is distributed as the Gaussian over $K_\mathbb{R}^m$ of covariance matrix $\alpha^2\mathbb{H}^\dagger\theta(\mathbf{U_z^\perp})\theta(\mathbf{U_z^\perp})^\dagger\mathbb{H} + \alpha^2(\xi^2\mathbf{I}_{mn} - \mathbb{H}^\dagger\theta(\mathbf{U_z^\perp})\theta(\mathbf{U_z^\perp})^\dagger\mathbb{H})$ which is identical to $D_{\alpha\xi}^m$. Since $q^{-1}(\mathbf{A'})^T\mathbf{s} \in q^{-1}(R^\vee)^m$, the coset $q^{-1}(R^\vee)^m - \mathbf{b'}$ is the same as $q^{-1}(R^\vee)^m - (\mathbf{U_z}\mathbf{e}+\mathbf{f})$, which yields that $\mathbf{c}$ can be seen as being sampled from $\mathcal{D}_{q^{-1}(R^\vee)^m - (\mathbf{U_z}\mathbf{e}+\mathbf{f}),\alpha}$. By the remark made before the proof, we have $\alpha \geq \eta_\varepsilon(q^{-1}(R^\vee)^m)$, so by Lemma 8, the distribution of $\mathbf{U_z}\mathbf{e}+\mathbf{f}+\mathbf{c}$ is within statistical distance 8ε of $\mathcal{D}_{q^{-1}(R^\vee)^m,\alpha\sqrt{\xi^2+1}}$, which shows that the second component is correctly distributed up to 8ε. Note that $\mathbf{U_z}\mathbf{e} = \sum_{i\in[m]} e_i \cdot \mathbf{u}_i$ is in the space spanned by the columns of $\mathbf{U_z^\perp}$ because $e_1 = 0$. This yields $\langle \mathbf{z}, \mathbf{U_z}\mathbf{e}\rangle = 0$ as $\mathbf{z}$ is orthogonal to the columns of $\mathbf{U_z^\perp}$.

This proves that the third component equals $\langle \mathbf{z}, \mathbf{U_z e} + \mathbf{f} + \mathbf{c} \rangle$ and is thus correctly distributed.

Now consider the case where both $\mathbf{A}$ and $\mathbf{b}$ are uniform. First, observe that $\alpha \geq \eta_\varepsilon(q^{-1}(R^\vee)^m)$ and therefore by Lemma 6, the distribution of $(\mathbf{A}, \mathbf{b})$ is within statistical distance $\varepsilon/2$ of the distribution of $(\mathbf{A}, \mathbf{e}' + \mathbf{e})$ where $\mathbf{e}' \in (q^{-1}R^\vee/R^\vee)^m$ is uniform and $\mathbf{e}$ is distributed from $\{0\} \times D_\alpha^{m-1}$. So we can assume our input is $(\mathbf{A}, \mathbf{e}' + \mathbf{e})$. $\mathbf{A}'$ is uniform as before, and clearly independent of the other two components. Moreover, since $\mathbf{b}' = \mathbf{U_z e}' + \mathbf{U_z e} + \mathbf{f}$ and $\mathbf{U_z e}' \in q^{-1}(R^\vee)^m$, then the coset $q^{-1}(R^\vee)^m - \mathbf{b}'$ is identical to $q^{-1}(R^\vee)^m - (\mathbf{U_z e} + \mathbf{f})$. For the same reasons as above, $\mathbf{U_z e} + \mathbf{f} + \mathbf{c}$ is distributed as $\mathcal{D}_{q^{-1}(R^\vee)^m, \alpha\sqrt{\xi^2+1}}$ within statistical distance of at most 8ε, and in particular independent of $\mathbf{e}'$. So the third component is correctly distributed because once again $\langle \mathbf{z}, \mathbf{U_z e} \rangle = 0$. Finally, since $\mathbf{e}'$ is independent of the first and third components, and that $\mathbf{U_z e}'$ is uniform over $(q^{-1}R^\vee/R^\vee)^m$ as $\mathbf{U_z}$ is invertible modulo qR, it yields that the second component is uniform and independent of the other ones as required. □

Instantiation in power-of-two cyclotomics. The condition on the modulus q in Lemma 15 and 16 stems from the invertibility result from Lyubashevsky and Seiler [LS18]. This result can be simplified in the power-of-two case [LS18, Cor. 1.2] where it is conditioned on the number $\kappa > 1$ of splitting factors of $x^n + 1$ in $\mathbb{Z}_q[x]$. Choosing κ as a power of two less than $n = 2^\ell$, q now has to be a prime congruent to $2\kappa + 1$ modulo 4κ. The invertibility condition then becomes $0 < \|\tau(y)\|_\infty < q^{1/\kappa}/\sqrt{\kappa}$ for any y in R_q. The upper bound is decreasing with κ so the smaller κ, the more invertible elements. The smallest choice for κ is $\kappa = 2$, which leads to choosing a prime $q = 5 \bmod 8$. In our context, having $q^{1/2}/\sqrt{2} > 1$ is sufficient as our elements have binary coefficients. This requires $q > 2$ which is subsumed by $q = 5 \bmod 8$.

Lemma 17 (Adapted from Lem. 4.8 [BLP+13]). *Let K be a number field of degree n, R its ring of integers, and k, m, q, t be positive integers such that $t \leq poly(n)$. Let ψ be a discrete distribution over $q^{-1}(R^\vee)^m$, and $\mathcal{Z} \subseteq (R^\vee)^m$. There is a reduction from* ext-M-LWE$_{n,k,m,q,\psi,\mathcal{Z}}$ *to* ext-M-LWE$^t_{n,k,m,q,\psi,\mathcal{Z}}$ *that reduces the advantage by a factor of t.*

Proof. Let $\mathcal{O}$ be an oracle for ext-M-LWE$^t_{n,k,m,q,\psi,\mathcal{Z}}$. For each $i \in \{0, \ldots, t\}$, we denote by $\mathcal{H}_i$ the hybrid distribution defined as

$$(\mathbf{A}, (\mathbf{b}_1, \ldots, \mathbf{b}_i, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_t), (\langle \mathbf{e}_j, \mathbf{z} \rangle)_{j \in [t]}),$$

where $\mathbf{A} \hookleftarrow U(R_q^{k \times m})$, the $\mathbf{u}_j$ are independent and identically distributed (i.i.d.) from $U((q^{-1}R^\vee/R^\vee)^m)$, the $\mathbf{e}_j$ are i.i.d. from ψ, and $\mathbf{b}_j = q^{-1}\mathbf{A}^T\mathbf{s}_j + \mathbf{e}_j \bmod R^\vee$ for $\mathbf{s}_j$ i.i.d. from $U((R_q^\vee)^k)$ for every $j \in [t]$. By definition, we have $\mathrm{Adv}[\mathcal{O}] = |\Pr(\mathcal{O}(\mathcal{H}_t)) - \Pr(\mathcal{O}(\mathcal{H}_0))|$. The reduction $\mathcal{A}$ works as follows.

1. Sample $\mathbf{z}$ uniformly from $\mathcal{Z}$ and get $(\mathbf{A}, \mathbf{b}, y = \langle \mathbf{e}, \mathbf{z} \rangle)$ as input of ext-M-LWE$_{n,k,m,q,\psi,\mathcal{Z}}$.

2. Sample i^* uniformly from $[t]$.
3. Sample $\mathbf{s}_1, \ldots, \mathbf{s}_{i^*-1}$ uniformly from $(R_q^\vee)^k$, $\mathbf{e}_1, \ldots, \mathbf{e}_{i^*-1}, \mathbf{e}_{i^*+1}, \ldots, \mathbf{e}_t$ from ψ and finally $\mathbf{u}_{i^*+1}, \ldots, \mathbf{u}_t$ uniformly from $(q^{-1}R^\vee/R^\vee)^m$.
4. Compute $\mathbf{b}_j = q^{-1}\mathbf{A}^T\mathbf{s}_j + \mathbf{e}_j \bmod R^\vee$ for all $j \in [i^*-1]$, and $y_j = \langle \mathbf{e}_j, \mathbf{z} \rangle$ for all $j \in [t] \setminus \{i^*\}$.
5. Define $(\mathbf{b}'_j)_{j\in[t]}$ as $(\mathbf{b}_1, \ldots, \mathbf{b}_{i^*-1}, \mathbf{b}, \mathbf{u}_{i^*+1}, \ldots, \mathbf{u}_t)$. Then call the oracle $\mathcal{O}$ on input $\big(\mathbf{A}, (\mathbf{b}'_j)_{j\in[t]}, (y_1, \ldots, y_{i^*-1}, y, y_{i^*+1}, \ldots, y_t)\big)$, and return the same output as $\mathcal{O}$.

If $\mathbf{b}$ is uniform, then the distribution in 5. is exactly $\mathcal{H}_{i^*-1}$ whereas if $\mathbf{b}$ is M-LWE, then the distribution is $\mathcal{H}_{i^*}$. By a standard hybrid argument, the oracle can distinguish between the two for some i^* if it can distinguish between $\mathcal{H}_0$ and $\mathcal{H}_t$. So the output is correct over the randomness of i^*. Since i^* is uniformly chosen we have

$$\begin{aligned}\mathrm{Adv}[\mathcal{A}] &= |\Pr(\mathcal{A}(\mathbf{b} \text{ M-LWE})) - \Pr(\mathcal{A}(\mathbf{b} \text{ uniform}))| \\ &= \left| \sum_{i^*\in[t]} \frac{1}{t}\Pr(\mathcal{A}(\mathcal{H}_{i^*})) - \sum_{i^*\in[t]} \frac{1}{t}\Pr(\mathcal{A}(\mathcal{H}_{i^*-1})) \right| \\ &= \frac{1}{t}\mathrm{Adv}[\mathcal{O}]\end{aligned}$$

□

3.4 Reduction to bin-M-LWE

We now provide the final step of the overall reduction, by reducing to the binary secret version of M-LWE using a sequence of hybrids. The idea is to use the set $\mathcal{Z}$ of the ext-M-LWE problem as our set of secrets. The problem ext-M-LWE$^m_{n,k,d,q,\alpha,\{0\}^d}$ mentioned in the lemma statement is trivially harder than ext-M-LWE$^m_{n,k,d,q,\alpha,(R_2^\vee)^d}$, that is also why it is not specified in Figure 1.

Lemma 18 (Adapted from Lem. 4.9 [BLP+13]). *Let $K = \mathbb{Q}(\zeta)$ be a number field of degree n, such that its ring of integers is $R = \mathbb{Z}[\zeta]$, with defining polynomial f. Let q be a prime modulus. Let k, m, d be positive integers such that $d \geq k \log_2 q + \omega(\log_2 n)$ Further, let $\varepsilon \in (0, 1/2)$ and $\alpha, \gamma, \beta, \delta$ be positive reals such that $\alpha \geq q^{-1}\sqrt{2\ln(2nd(1+1/\varepsilon))/\pi}$, $\gamma = \alpha B\sqrt{d}$, $\beta = \alpha B\sqrt{2d}$, where $B = \max_{x\in R_2} \|\sigma(x)\|_\infty$, and $\delta = \frac{1}{2}\sqrt{(1+q^k/2^d)^n - 1}$. Then there is a reduction from ext-M-LWE$^m_{n,k,d,q,\alpha,(R_2^\vee)^d}$, M-LWE$_{n,k,m,q,\gamma}$ and ext-M-LWE$^m_{n,k,d,q,\alpha,\{0\}^d}$ to bin-M-LWE$_{n,d,m,q,\leq\beta}$, such that if $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$ are the algorithms obtained by applying these hybrids to an algorithm $\mathcal{A}$, then*

$$\mathrm{Adv}[\mathcal{A}] \leq \mathrm{Adv}[\mathcal{B}_1] + \mathrm{Adv}[\mathcal{B}_2] + \mathrm{Adv}[\mathcal{B}_3] + 2m\varepsilon + \delta.$$

Proof. For $x \in R^\vee$, we denote $\widetilde{x} = \lambda x \in R$ as before, where $\lambda = f'(\zeta)$. We extend this notation to vectors and matrices in the obvious way. We consider $\mathbf{z} \hookleftarrow U((R_2^\vee)^d)$ and $\mathbf{e} \in K_\mathbb{R}^m$ sampled from the continuous Gaussian $D^m_\mathbf{r}$

with parameter vector $\mathbf{r}$ with $r_j^2 = \gamma^2 + \alpha^2 \sum_i |\sigma_j(\widetilde{z}_i)|^2$. Yet, we have $\|\mathbf{r}\|_\infty = \sqrt{\gamma^2 + \alpha^2 \|\widetilde{\mathbf{z}}\|_{2,\infty}^2}$, as well as $\|\widetilde{\mathbf{z}}\|_{2,\infty}^2 \leq \sum_{i\in[d]} \|\sigma(\widetilde{z}_i)\|_\infty^2$. Recalling the parameter $B = \max_{x\in R_2} \|\sigma(x)\|_\infty$, that can be bounded above by n for cyclotomics by Lemma 1, we get $\|\mathbf{r}\|_\infty \leq \sqrt{\gamma^2 + B^2 d\alpha^2} = B\sqrt{2d}\alpha = \beta$. In addition, we sample $\mathbf{A}$ uniformly over $(R_q)^{d\times m}$ and define $\mathbf{b} = q^{-1}\mathbf{A}^T\mathbf{z} + \mathbf{e} \bmod R^\vee$.

First hybrid. We denote by H_0 the distribution of $(\mathbf{A}, \mathbf{b})$ and H_1 the distribution of $(\mathbf{A}, q^{-1}\mathbf{A}^T\mathbf{z} - \lambda\mathbf{N}^T\mathbf{z} + \widehat{\mathbf{e}} \bmod R^\vee)$, where $\mathbf{N} \hookleftarrow \mathcal{D}_{q^{-1}R^\vee,\alpha}^{d\times m}$ and $\widehat{\mathbf{e}} \hookleftarrow D_\gamma^m$. By looking at each component of the vectors we claim that $\Delta([-\mathbf{N}^T\widetilde{\mathbf{z}}+\widehat{\mathbf{e}}]_i, \mathbf{e}_i) \leq 2\varepsilon$. Indeed, $(1/\alpha^2 + \|\widetilde{\mathbf{z}}\|_{2,\infty}^2/\gamma^2)^{-1/2} \geq \alpha/\sqrt{2}$ and $\alpha/\sqrt{2} \geq \eta_\varepsilon(q^{-1}(R^\vee)^d)$ as explained for Lemma 16. If $\mathbf{n}_i \in q^{-1}(R^\vee)^d$ denotes the i-th column of $\mathbf{N}$, Lemma 11 yields the claim as $[-\mathbf{N}^T\widetilde{\mathbf{z}}+\widehat{\mathbf{e}}]_i = \langle \mathbf{n}_i, -\widetilde{\mathbf{z}}\rangle + \widehat{e}_i$, thus giving $\Delta(-\mathbf{N}^T\widetilde{\mathbf{z}}+\widehat{\mathbf{e}}, \mathbf{e}) \leq 2m\varepsilon$.

$$|\Pr(\mathcal{A}(H_0)) - \Pr(\mathcal{A}(H_1))| \leq 2m\varepsilon. \tag{1}$$

Second hybrid. We define H_2 to be the distribution of $(\widehat{\mathbf{A}}, q^{-1}\widehat{\mathbf{A}}^T\mathbf{z} - \lambda\mathbf{N}^T\mathbf{z} + \widehat{\mathbf{e}} \bmod R^\vee) = (\widehat{\mathbf{A}}, q^{-1}(\lambda\mathbf{B})^T\mathbf{C}\mathbf{z} + \widehat{\mathbf{e}} \bmod R^\vee)$ where $\mathbf{B}$ is uniformly sampled over $(R_q^\vee)^{k\times m}$, $\mathbf{C}$ uniformly sampled over $R_q^{k\times d}$ and $\widehat{\mathbf{A}} = \lambda q(q^{-1}\mathbf{C}^T\mathbf{B} + \mathbf{N} \bmod R^\vee)$. We argue that a distinguisher between H_1 and H_2 can be used to derive an adversary $\mathcal{B}_1$ for ext-M-LWE$^m_{n,k,d,q,\alpha,(R_2^\vee)^d}$ with the same advantage. To do so, $\mathcal{B}_1$ transforms the samples from the challenger of the ext-M-LWE problem to samples defined in H_1 or the ones in H_2 depending on whether or not the received samples are uniform. In the uniform case, $(\mathbf{C}, (\lambda q)^{-1}\mathbf{A}, \mathbf{N}^T\mathbf{z})$ can be efficiently transformed into a sample from H_1. Note that $(\lambda q)^{-1}\mathbf{A}$ indeed corresponds to the uniform case of ext-M-LWE, because $\mathbf{A}$ is uniform over R_q and $(\lambda q)^{-1}R_q$ can be seen as $q^{-1}R^\vee/R^\vee$. In the other case, if we apply the same transformation to the ext-M-LWE sample $(\mathbf{C}, q^{-1}\mathbf{C}^T\mathbf{B} + \mathbf{N} \bmod R^\vee, \mathbf{N}^T\mathbf{z})$, it leads to a sample from H_2. Hence, $\mathcal{B}_1$ is a distinguisher for ext-M-LWE$^m_{n,k,d,q,\alpha,(R_2^\vee)^d}$, and

$$|\Pr(\mathcal{A}(H_1)) - \Pr(\mathcal{A}(H_2))| = \mathrm{Adv}[\mathcal{B}_1]. \tag{2}$$

Third hybrid. Next we define H_3 to be the distribution of $(\widehat{\mathbf{A}}, q^{-1}\widetilde{\mathbf{B}}^T\mathbf{s} + \widehat{\mathbf{e}} \bmod R^\vee)$, where $\widetilde{\mathbf{B}} = \lambda\mathbf{B} \in R_q^{k\times m}$, and $\mathbf{s}$ is uniform over $(R_q^\vee)^k$. By the Ring Leftover Hash Lemma stated in Lemma 12, we have that $(\mathbf{C}, \mathbf{C}\widetilde{\mathbf{z}})$ is within statistical distance at most δ from $(\mathbf{C}, \widetilde{\mathbf{s}})$. By multiplying by λ^{-1} and using the fact that a function does not increase the statistical distance, we have that $\Delta((\mathbf{C}, \mathbf{C}\mathbf{z}), (\mathbf{C}, \mathbf{s})) \leq \delta$. Note that the condition $d \geq k\log_2 q + \omega(\log_2 n)$ implies $\delta \leq n^{-\omega(1)}$. This yields

$$|\Pr(\mathcal{A}(H_2)) - \Pr(\mathcal{A}(H_3))| \leq \delta. \tag{3}$$

Fourth hybrid. We then replace the second component by the uniform as we define H_4 to be the distribution of $(\widehat{\mathbf{A}}, \mathbf{u})$, with $\mathbf{u} \hookleftarrow U(\mathbb{T}_{R^\vee}^m)$. A distinguisher between H_3 and H_4 can be used to derive an adversary $\mathcal{B}_2$ for M-LWE$_{n,k,m,q,\gamma}$. For that, $\mathcal{B}_2$ applies the the efficient transformation to the samples from the M-LWE challenger, which turns $(\widetilde{\mathbf{B}}, \mathbf{u})$ into a sample from H_4 in the uniform case,

and $(\widetilde{\mathbf{B}}, q^{-1}\widetilde{\mathbf{B}}^T\mathbf{s}+\widehat{\mathbf{e}} \bmod R^\vee)$ into a sample from H_3 in the M-LWE case. Therefore, $\mathcal{B}_2$ is a distinguisher for M-LWE$_{n,k,m,q,\gamma}$ such that

$$|\Pr(\mathcal{A}(H_3)) - \Pr(\mathcal{A}(H_4))| = \mathrm{Adv}[\mathcal{B}_2]. \tag{4}$$

Last hybrid. We now change $\widehat{\mathbf{A}}$ back to uniform by defining H_5 to be the distribution of $(\mathbf{A}, \mathbf{u})$. With the same argument as for the second hybrid, we can construct an adversary $\mathcal{B}_3$ for ext-M-LWE$^m_{n,k,d,q,\alpha,\{0\}^d}$ (which corresponds to multiple-secret M-LWE without additional information on the error) based on a distinguisher between H_4 and H_5. It transforms $(\mathbf{C}, (\lambda q)^{-1}\widehat{A}, \mathbf{N}^T\mathbf{0})$ into a sample from H_4 (M-LWE case) and $(\mathbf{C}, (\lambda q)^{-1}\mathbf{A}, \mathbf{N}^T\mathbf{0})$ into a sample from H_5 (uniform case). We then get

$$|\Pr(\mathcal{A}(H_4)) - \Pr(\mathcal{A}(H_5))| = \mathrm{Adv}[\mathcal{B}_3]. \tag{5}$$

Putting Eq. 1, 2, 3, 4, 5 altogether yields the result. □

Acknowledgments

This work was supported by the European Union PROMETHEUS project (Horizon 2020 Research and Innovation Program, grant 780701). It has also received a French government support managed by the National Research Agency in the "Investing for the Future" program, under the national project RISQ P141580-2660001 / DOS0044216. Katharina Boudgoust is funded by the Direction Générale de l'Armement (Pôle de Recherche CYBER). We thank our anonymous referees of Indocrypt 2020 and CT-RSA 2021 for their thorough proof reading and constructive feedback.

References

AA16. J. Alperin-Sheriff and D. Apon. Dimension-preserving reductions from LWE to LWR. *IACR Cryptology ePrint Archive*, 2016:589, 2016.

AD17. M. R. Albrecht and A. Deo. Large modulus ring-lwe ≥ module-lwe. In *Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I*, pages 267–296, 2017.

BD20. Z. Brakerski and N. Döttling. Hardness of LWE on general entropic distributions. In *Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part II*, volume 12106 of *Lecture Notes in Computer Science*, pages 551–575. Springer, 2020.

BDK+18. J. W. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, G. Seiler, and D. Stehlé. CRYSTALS - kyber: A cca-secure module-lattice-based KEM. In *2018 IEEE European Symposium on Security and Privacy, EuroS&P 2018, London, United Kingdom, April 24-26, 2018*, pages 353–367, 2018.

BGV12. Z. Brakerski, C. Gentry, and V. Vaikuntanathan. (leveled) fully homomorphic encryption without bootstrapping. In *Innovations in Theoretical Computer Science 2012, Cambridge, MA, USA, January 8-10, 2012*, pages 309–325, 2012.

BJRW20. K. Boudgoust, C. Jeudy, A. Roux-Langlois, and W. Wen. Towards classical hardness of module-lwe: The linear rank case. *IACR Cryptol. ePrint Arch.*, 2020:1020, 2020.

BLP+13. Z. Brakerski, A. Langlois, C. Peikert, O. Regev, and D. Stehlé. Classical hardness of learning with errors. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 575–584, 2013.

BV14. Z. Brakerski and V. Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. *SIAM J. Comput.*, 43(2):831–871, 2014.

DKL+18. L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, P. Schwabe, G. Seiler, and D. Stehlé. Crystals-dilithium: A lattice-based digital signature scheme. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2018(1):238–268, 2018.

DM15. L. Ducas and D. Micciancio. FHEW: bootstrapping homomorphic encryption in less than a second. In *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I*, pages 617–640, 2015.

GKPV10. S. Goldwasser, Y. T. Kalai, C. Peikert, and V. Vaikuntanathan. Robustness of the learning with errors assumption. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 5-7, 2010. Proceedings*, pages 230–240. Tsinghua University Press, 2010.

GPV08. C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 197–206. ACM, 2008.

KF15. P. Kirchner and P.-A. Fouque. An improved BKW algorithm for LWE with applications to cryptography and lattices. In *Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I*, pages 43–62, 2015.

LPR10. V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. In *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Monaco / French Riviera, May 30 - June 3, 2010. Proceedings*, pages 1–23, 2010.

LPR13. V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. *J. ACM*, 60(6):43:1–43:35, 2013.

LS15. A. Langlois and D. Stehlé. Worst-case to average-case reductions for module lattices. *Des. Codes Cryptogr.*, 75(3):565–599, 2015.

LS18. V. Lyubashevsky and G. Seiler. Short, invertible elements in partially splitting cyclotomic rings and applications to lattice-based zero-knowledge proofs. In *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part I*, volume 10820 of *Lecture Notes in Computer Science*, pages 204–224. Springer, 2018.

LWW20. H. Lin, Y. Wang, and M. Wang. Hardness of module-lwe and ring-lwe on general entropic distributions. *IACR Cryptol. ePrint Arch.*, 2020:1238, 2020.

Mic07. D. Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Comput. Complex.*, 16(4):365–411, 2007.

Mic18. D. Micciancio. On the hardness of learning with errors with binary secrets. *Theory of Computing*, 14(1):1–17, 2018.

MP12. D. Micciancio and C. Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, April 15-19, 2012. Proceedings*, pages 700–718, 2012.

MR07. D. Micciancio and O. Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007.

NIS. NIST. Post-quantum cryptography standardization. `https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Post-Quantum-Cryptography-Standardization`.

Pei08. C. Peikert. Limits on the hardness of lattice problems in l_p norms. *Comput. Complex.*, 17(2):300–351, 2008.

Pei09. C. Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 333–342, 2009.

Pei10. C. Peikert. An efficient and parallel gaussian sampler for lattices. In *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings*, pages 80–97, 2010.

PS19. C. Peikert and S. Shiehian. Noninteractive zero knowledge for NP from (plain) learning with errors. In *Advances in Cryptology - CRYPTO 2019 - 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2019, Proceedings, Part I*, volume 11692 of *Lecture Notes in Computer Science*, pages 89–114. Springer, 2019.

Reg05. O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 84–93, 2005.

Reg09. O. Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34:1–34:40, 2009.

RSW18. M. Roşca, D. Stehlé, and A. Wallet. On the ring-lwe and polynomial-lwe problems. In *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part I*, pages 146–173, 2018.

SSTX09. D. Stehlé, R. Steinfeld, K. Tanaka, and K. Xagawa. Efficient public key encryption based on ideal lattices. In *Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings*, pages 617–635, 2009.

WW19. Y. Wang and M. Wang. Module-lwe versus ring-lwe, revisited. *IACR Cryptology ePrint Archive*, 2019:930, 2019.