On the Hardness of Module Learning With Errors with Short Distributions

Katharina Boudgoust[1], Corentin Jeudy[2,3], Adeline Roux-Langlois[2], and Weiqiang Wen[4]

katharina.boudgoust@cs.au.dk, corentin.jeudy@irisa.fr, adeline.roux-langlois@irisa.fr, weiqiang.wen@telecom-paris.fr

[1] Dept. Computer Science, Aarhus University, Aarhus, Denmark
[2] Univ Rennes, CNRS, IRISA, Rennes, France
[3] Orange Labs, Applied Crypto Group, Cesson-Sévigné, France
[4] LTCI, Telecom Paris, Institut Polytechnique de Paris, Paris, France

Abstract. The Module Learning With Errors problem (M-LWE) is a core computational assumption of lattice-based cryptography which offers an interesting trade-off between guaranteed security and concrete efficiency. The problem is parameterized by a *secret* distribution as well as an *error* distribution. There is a gap between the choices of those distributions for theoretical hardness results (standard formulation of M-LWE, i.e., uniform secret modulo q and Gaussian error) and practical schemes (small bounded secret and error). In this work, we make progress towards narrowing this gap. More precisely, we prove that M-LWE with η-bounded secret for any $2 \leq \eta \ll q$ and Gaussian error, in both its search and decision variants, is at least as hard as the standard formulation of M-LWE, provided that the module rank d is at least logarithmic in the ring degree n. We also prove that the search version of M-LWE with large uniform secret and uniform η-bounded error is at least as hard as the standard M-LWE problem, if the number of samples m is close to the module rank d and with further restrictions on η. The latter result can be extended to provide the hardness of M-LWE with uniform η-bounded secret *and* error under specific parameter conditions.

Keywords: Lattice-Based Cryptography · Module Learning With Errors · Short Distributions · Bounded Secret · Bounded Error

1 Introduction

The *Learning With Errors* (LWE) problem, introduced by Regev [Reg05], is one of the main computational assumptions for lattice-based cryptographic schemes. Given two positive integers d and q, and a secret vector $\mathbf{s} \in \mathbb{Z}_q^d$, an $\mathrm{LWE}_{d,q,\psi}$ sample is defined as $(\mathbf{a}, b = q^{-1}\langle \mathbf{a}, \mathbf{s} \rangle + e \bmod \mathbb{Z})$, where $\mathbf{a}$ is sampled from the uniform distribution over $\mathbb{Z}_q^d$, and e an error term sampled from a distribution ψ over $\mathbb{R}$.

This paper contains novel results and generalizations of existing ones already published in [BJRW20,BJRW21].

The *search* version of LWE asks to recover the secret $\mathbf{s}$ given arbitrarily many samples of the LWE distribution. Its *decision* counterpart asks to distinguish between LWE samples and the same number of samples drawn from the uniform distribution over $\mathbb{Z}_q^d \times \mathbb{T}$, where the torus is defined by $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. When the number of samples m is fixed, we use a matrix representation of the $\text{LWE}_{d,m,q,\psi}$ samples as $(\mathbf{A}, q^{-1}\mathbf{As} + \mathbf{e} \bmod \mathbb{Z})$, with $\mathbf{A}$ uniform over $\mathbb{Z}_q^{m\times d}$, and $\mathbf{e}$ sampled from ψ^m. From a theoretical standpoint, LWE is interesting for its ties with well-known lattice problems. Lattices are discrete additive subgroups of $\mathbb{R}^d$ and arise in many different areas of mathematics, such as number theory, geometry and group theory. There are several problems on lattices that are proven to be computationally hard to solve, such as the problem of finding a set of *shortest independent vectors* (SIVP). A standard relaxation of the latter, which is more suitable for building cryptography upon, consists in solving it only up to an approximation factor γ and is denoted by SIVP_γ. The caveat of this relaxation is that the hardness is only conjectured. The seminal work of Regev [Reg05,Reg09] proves a worst-case to average-case quantum reduction from SIVP_γ to LWE. It means that if there exists an efficient solver for LWE, then it can be used to construct a quantum solver for SIVP_γ in the worst case, i.e., in any Euclidean lattice. The subsequent work of Peikert [Pei09], then generalized to any polynomial modulus q by Brakerski et al. [BLP+13], dequantized the reduction to obtain fully classical worst-case to average-case reductions to LWE.

Structured Variants. Cryptographic schemes whose security proofs rely on the hardness of LWE inherently suffer from large public keys and quite intensive computations, both quadratic in the security parameter. Structured variants of LWE have been proposed in order to gain in efficiency [SSTX09,LPR13a]. In this paper, we focus on the *Module Learning With Errors* (M-LWE) problem, first defined by Brakerski et al. [BGV12] and then thoroughly studied by Langlois and Stehlé [LS15]. The formulation is similar to that of LWE where the set of integers $\mathbb{Z}$ is replaced by the ring of algebraic integers R of a number field K. This introduces a new parameter, which is the degree n of the number field. We denote by $R^\vee$ the dual ideal of R. The integer d now denotes the module rank, and q still denotes the modulus. Further, let ψ be a distribution on the tensor field $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$, and let $\mathbf{s} \in (R_q^\vee)^d$ be a secret vector, where $R_q = R/qR$. An $\text{M-LWE}_{n,d,q,\psi}$ sample is given by $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod R^\vee)$, where $\mathbf{a}$ is uniform in R_q^d, and e is sampled from ψ. The search version asks to find $\mathbf{s}$ given arbitrarily many samples, while the decision version asks to distinguish such samples from uniformly random ones over $R_q^d \times \mathbb{T}_{R^\vee}$, where the torus is $\mathbb{T}_{R^\vee} = K_\mathbb{R}/R^\vee$. We can also use a matrix formulation when the number of samples m is fixed by considering the $\text{M-LWE}_{n,d,m,q,\psi}$ distribution $(\mathbf{A}, q^{-1}\mathbf{As} + \mathbf{e} \bmod R^\vee)$ with $\mathbf{A}$ uniform in $R_q^{m\times d}$ and $\mathbf{e}$ from ψ^m. When the module rank is $d = 1$, the problem is called *Ring*-LWE (R-LWE) [LPR13a]. Just like LWE, the M-LWE problem enjoys worst-case to average-case connections from lattice problems such as SIVP_γ [LS15]. Whereas the hardness results for LWE start from general lattice problems, the set has to be restricted to *module lattices* in the case

of M-LWE, which correspond to finitely generated R-modules. Since its introduction, the M-LWE problem has attracted more and more interest as it offers a fine-grained trade-off between concrete security and efficiency, mostly by tweaking the parameters n and d. It is also extremely versatile in the sense that it allows for constructing a wide variety of cryptographic schemes. As an example, within the ongoing NIST standardization process [NIS], several finalist candidates rely on the hardness of M-LWE, e.g., the signature scheme Dilithium [DKL+18] and the key encapsulation mechanism Kyber [BDK+18]. However, these efficient schemes use different parameter settings, and in particular different distributions for the secret and error, that are not yet encompassed by theoretical proofs of hardness. In these cases, the hardness of M-LWE is argued based on the state of the art cryptanalysis and attacks on M-LWE.

Short Distributions. The standard formulation of LWE considers a large uniform secret and a Gaussian error, but in practice we tend to consider short distributions, i.e., secret or error with coefficients bounded by $\eta \ll q$. This corresponds to choosing the secret $\mathbf{s}$ in $\{0, \ldots, \eta-1\}^d$, or a discrete error distribution ψ to be over $\{0, \ldots, \eta-1\}$ instead of $\mathbb{Z}_q$. Besides gaining in efficiency, choosing a small secret plays an important role in some applications like fully homomorphic encryption [DM15] or modulus switching techniques [BLP+13,AD17a,WW19] as it keeps the noise blowup to a minimum. The LWE problem with η-bounded secret (η-LWE) has been well studied in the case of binary secret ($\eta = 2$) but the different approaches easily generalize to slightly larger secrets. A first study of 2-LWE was provided by Goldwasser et al. [GKPV10] in the context of leakage-resilient cryptography. Although their proof structure has the advantage of being easy to follow, their result suffers from a large error increase. Informally, they show a reduction from $\mathrm{LWE}_{k,q,D_\alpha}$ to 2-$\mathrm{LWE}_{d,q,D_\beta}$, where $\beta/\alpha = d^{\omega(1)}$ (super-polynomial) and $d \geq k \log_2 q + \omega(\log_2 d)$. The distribution D_r denotes a Gaussian distribution with standard deviation r (up to a factor of $\sqrt{2\pi}$). It was later improved by Brakerski et al. [BLP+13] and Micciancio [Mic18] using more technical proofs. Both of them achieve a similar dimension increase between k and d, but only increase the error by roughly $\beta/\alpha = \Omega(\sqrt{d})$. The dimension increase from k to roughly $k \log_2 q$ is reasonable as it essentially preserves the number of possible secrets. Recent work by Brakerski and Döttling [BD20a] extends the hardness results to more general secret distributions based on entropic arguments.

The hardness of LWE with η-bounded error was first studied by Micciancio and Peikert [MP13]. They proved that the LWE function $(\mathbf{s}, \mathbf{e}) \mapsto \mathbf{A}\mathbf{s} + \mathbf{e} \bmod q$ is one-way with respect to $\mathbf{e}$ uniform over $\{0, \ldots, \eta-1\}^m$, provided that the number of samples m is at most $d(1 + O(\log_2 \eta / \log_2 d))$. The one-wayness is proven under the hardness of general lattice problems over lattices of rank $O(d \log_2 \eta / \log_2 d)$. It was recently extended to non-uniform binary errors by Sun et al. [STA20], proving that the maximum number of samples must be $m = d(1 + O(p(d)/\log_2 d))$, where $p(d)$ is the probability of getting 1 from the error distribution. The proof of [MP13] corresponds to $p(d) = 1/2$.

The question of whether these hardness results carry over to structured variants, and in particular to the module case, was left open. The work on LWE with entropic secret was extended to the R-LWE case by Brakerski and Döttling [BD20b], and the module case by Lin et al. [LWW20][1]. However, no results on the hardness of M-LWE with η-bounded secret or error were known, even though they serve as hardness assumptions for most efficient M-LWE-based schemes. For example, the signature scheme Dilithium [DKL+18] in the NIST competition samples the secret and error from the uniform distribution over vectors with coefficients between -2 and 2 (security levels I and III) or between -4 and 4 (security level II).

Our Contributions. In this paper, we provide three main contributions on the hardness of M-LWE with small secret and/or error, i.e., with coefficients bounded by η. The first two contributions study the hardness of the M-LWE problem with η-bounded secret, which we denote by η-M-LWE, in both its search and decision versions respectively, for any $\eta \geq 2$. They are generalizations of the results published in our previous conference papers [BJRW20] and [BJRW21] respectively, only dealing with the special case of 2-M-LWE, which is already mentioned in one of the author's thesis [Bou21]. The third and new contribution concerns the hardness of the search version of M-LWE with η-bounded error, under more specific restrictions on η. The latter contribution can then be used to deduce the hardness of M-LWE with small secret *and* error. To the best of our knowledge, these are the first results for the hardness of M-LWE with small bounded distributions (secret or error).

Contribution 1: Computational hardness of η-M-LWE. We show a first reduction for the hardness of the search version of η-M-LWE. The formal statement can be found in Theorem 3.1. It follows the original proof structure of Goldwasser et al. [GKPV10] in the case of LWE, while achieving much better noise parameters by using the Rényi divergence instead of the statistical distance to measure the distance between two distributions. The improvement on the noise rate compared to [GKPV10] stems from the fact that the Rényi divergence only needs to be constant for the reduction to work, and not necessarily negligibly close to 1 (compared to negligibly close to 0 for the statistical distance). A similar effect arises with respect to the rank condition in comparison with Contribution 2 below. More precisely, as we use the leftover hash lemma with respect to the Rényi divergence, we can have a rank that is logarithmic in the ring degree n, instead of super-logarithmic. However, using the Rényi divergence as a measure of distribution closeness only allows us to prove the hardness of the *search* variant, denoted by η-M-SLWE. Additionally, it asks to fix the number of samples a priori.

It consists in a reduction from M-SLWE and M-LWE with rank k and Gaussian width α to η-M-SLWE with rank d and width β. The reduction preserves the ring degree n, the number of samples m and the modulus q, where q only

[1] Note that at the time of writing, the paper by Lin et al. is only accessible on ePrint and has not yet been peer-reviewed.

needs to be prime. The ranks must satisfy $d \log_2 \eta \geq k \log_2 q + \Omega(\log_2 n)$, which is due to the use of the leftover hash lemma. The Gaussian noise parameter α is also increased to β by a factor $\beta/\alpha = d\sqrt{m} \cdot n^{3/2} \log_2(n)(\eta - 1)$ in general cyclotomic fields, which can be further improved by a factor of $\sqrt{n}$ in the specific case of power-of-two cyclotomic fields.

Contribution 2: Pseudorandomness of η-M-LWE. We then provide a more involved proof of hardness for the *decision* version of η-M-LWE through a reduction from M-LWE to η-M-LWE. The thorough statement is provided in Theorem 3.2. Not only does this reduction apply to the decision versions, but it also slightly improves the noise rate of the reduction in certain parameter regimes. In particular, the noise rate no longer depends on the number of samples m, as opposed to Contribution 1. The technique follows the idea of [BLP+13] by introducing the two intermediate problems first-is-errorless M-LWE and ext-M-LWE. We first reduce the M-LWE problem 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 η-bounded ring elements $\mathcal{Z} = (R_\eta^\vee)^d$. Two other formulations of ext-M-LWE were proposed by Alperin-Sheriff and Apon [AA16], and more recently by Lyubashevsky et al. [LNS21], but neither suits our reduction due to our lossy argument in Lemma 3.5. We discuss further these differences in Section 3.2.2. Then, to reduce ext-M-LWE to η-M-LWE, we use a lossy argument, similar to that of Contribution 1 but now relying on the newly derived ext-M-LWE hardness assumption, as well as the leftover hash lemma.

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 3.2 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 2.4 [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 small norm ring elements are units of R_q.

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 the said 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 modulus, at the expense of a loss in the Gaussian noise parameter. The ranks must satisfy $d \log_2 \eta \geq (k+1) \log_2 q + \omega(\log_2 n)$, in the same manner as in Contribution 1, except that the asymptotic term is now super-logarithmic. The noise rate is now given by $n(\eta - 1)\sqrt{2d}\sqrt{4n^2(\eta-1)^2 + 1} = \Theta((\eta-1)^2 n^2 \sqrt{d})$ for cyclotomic fields. This reduction removes the dependency in m in the noise rate of Contribution 1, which can be more advantageous in certain cases as we usually take $m = \Theta(n \log_2 n)$. Additionally, when bridging to LWE, the noise ra-

tio is improved to $\sqrt{10d}$ as our construction in Lemma 3.2 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.

Contribution 3: One-wayness of M-LWE *with small error.* Our last contribution focuses on the hardness of M-SLWE when the error distribution is uniform over η-bounded elements instead of Gaussian. The complete result can be found in Theorem 4.2. It uses a different proof method from Contributions 1 and 2 by following the idea of Micciancio and Peikert [MP13] of proving the one-wayness of the M-LWE function $(\mathbf{s}, \mathbf{e}) \mapsto \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR^\vee$, with $\mathbf{e}$ uniform in $(R_\eta^\vee)^m$. To do so, we prove the one-wayness of the M-SIS function $\mathbf{e} \mapsto (\mathbf{A}')^T\mathbf{e} \bmod qR^\vee$ and use the duality between both functions to conclude. This function is inspired from the *Module Short Integer Solution* (M-SIS) problem [LS15] which asks to find a short non-zero vector $\mathbf{e} \in (R^\vee)^m$ such that $(\mathbf{A}')^T\mathbf{e} = \mathbf{0} \bmod qR^\vee$ for a public random matrix $\mathbf{A}' \in R_q^{m\times d}$. It can be generalized to an inhomogeneous version by replacing $\mathbf{0}$ by a public syndrome $\mathbf{u}$. The one-wayness of the function is ensured by two properties, namely the uninvertibility and the second preimage resistance, which we prove using statistical arguments

We obtain similar results to [MP13] in terms of the number of samples using the asymptotic approach. However, the asymptotic approach is not suited for very small values of d. To overcome this problem, we use a more fine-grained approach using tighter calculations rather than hiding constants in asymptotic notations. This leads to more complicated conditions on the parameters, especially the link between the size of the error and the number of samples. We thus evaluate this condition numerically to determine the concrete parameters that are encompassed by the result. It shows that in order to reach very small errors, e.g. binary or ternary, the module rank d has to be large enough. We can still reach a small error size η for constant module ranks, but not arbitrarily small. Additionally, to prove the hardness of M-SLWE with small error *and* secret with m samples, we need to have the hardness of M-LWE with small error and $m + d$ samples. This restriction makes it difficult to achieve small error and secret at the same time for a large enough m. We discuss this transformation in more details in Section 4.3.

The M-SLWE problem can be seen as a linear system of equations (d variables and m equations over R_q or nd variables and nm equations over $\mathbb{Z}_q$) with noise. The presence of noise or error is what makes the problem difficult to solve. The motivation is therefore to determine the threshold of noise to add to the equations above which the problem is proven hard, still under lattice assumptions. Note that the number of equations characterized by m and the distribution of the error need to be chosen carefully with respect to one another. For example, an attack by Arora and Ge [AG11] uses the m samples to build noiseless polynomial equations of degree η, where η is a bound on the error coefficients. If m is sufficiently large, root finding algorithms can perform well on the latter. In particular, if $\eta = 2$ (binary), then $m \approx d^2$ samples is enough to solve LWE in polynomial time. The attack can also be applied to M-LWE as one

equation over R_q gives n equations over $\mathbb{Z}_q$. We discuss the consequences on the parameters in Section 4.4.

Open Problems. In this paper, several results are limited to special classes of number fields, e.g. cyclotomic fields or fields $K = \mathbb{Q}(\zeta)$ for which the ring of integers is $R = \mathbb{Z}[\zeta]$. Although it covers the fields that are used in practice, it may be of independent interest to extend our results to more general fields. The first two contributions imply the hardness of M-LWE with a small secret and a moderate rank (e.g., $\Omega(\log_2 n)$ for search and $\omega(\log_2 n)$ for decision) due to the leftover hash lemma. The hardness of η-M-LWE thus remains open for lower module ranks. Practical M-LWE-based schemes use a constant rank for increased efficiency, like the CRYSTALS candidates [BDK+18,DKL+18] at the NIST standardization process. The hardness proof of η-M-LWE with η-bounded error and m samples seems to require the hardness of M-LWE with η-bounded error and $m + d$ samples. Although subexponential attacks do not apply to the case where $m = O(d)$ (even for binary errors $\eta = 2$), our proof does not encompass this range of parameters. We leave it as a major open problem to prove the hardness of M-LWE with η-bounded error for all $m = O(d)$. Finally, two of our contributions are only proven for the search version of M-LWE. One possibility (of more general interest) would be to find search-to-decision reductions for M-LWE that preserve the secret distribution or the error distribution without reducing the number of samples m too much. For the latter, a sample-preserving search-to-decision for LWE [MM11] is known, but it is yet to be extended to structured variants.

Organization. In Section 2, we introduce the notions and preliminary results that are needed in this work. Section 3 is dedicated to the proofs of Contributions 1 and 2 on the hardness of η-M-LWE, generalizing that of our earlier conference papers [BJRW20,BJRW21]. Then, in Section 4, we give the proof of Contribution 3 on the hardness of M-LWE with η-bounded error. Finally, Section 5 gives a concise view of the current landscape on the hardness of M-LWE.

2 Preliminaries

Throughout the paper, q denotes a positive integer, $\mathbb{Z}_q$ denotes the ring of integers modulo q. In a ring R, we write $\langle p \rangle$ for the principal ideal generated by $p \in R$, and R_p for the quotient ring $R/\langle p \rangle = R/pR$. For simplicity, we denote by $[n]$ the set $\{1, \ldots, n\}$ for any positive integer n. Vectors and matrices are written in bold and their transpose (resp. Hermitian) is denoted by superscript T (resp. †). We denote the Euclidean norm and infinity norm of $\mathbb{C}^n$ by $\|\cdot\|_2$ and $\|\cdot\|_\infty$ respectively. We also define the *spectral norm* of any matrix $\mathbf{A} \in \mathbb{C}^{n \times m}$ by $\|\mathbf{A}\|_2 = \max_{\mathbf{x} \in \mathbb{C}^m \setminus \{0\}} \|\mathbf{A}\mathbf{x}\|_2 / \|\mathbf{x}\|_2$, and the *max norm* as $\|\mathbf{A}\|_{\max} = \max_{i \in [n], j \in [m]} |a_{i,j}|$. The identity matrix of size n is denoted by $\mathbf{I}_n$.

2.1 Algebraic Number Theory

A number field $K = \mathbb{Q}(\zeta)$ of degree n is a finite field extension of the rational number field $\mathbb{Q}$ obtained by adjoining an algebraic number 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 of K defines a ring, called the ring of integers which we denote by R. It is always true that $\mathbb{Z}[\zeta] \subseteq R$, where this inclusion can be strict. Some of the results are restricted to the class of number fields where the equality $R = \mathbb{Z}[\zeta]$ holds. This is the case for some quadratic extensions (i.e., when $\zeta = \sqrt{d}$ with d square-free and $d \neq 1 \bmod 4$), cyclotomic fields (i.e., when ζ is a primitive root of the unity) and number fields with a defining polynomial f of square-free discriminant Δ_f.

Space $\boldsymbol{H}$. 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 $\mathbf{U}_H$ as orthonormal basis, where

$$\mathbf{U}_H = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2}\mathbf{I}_{t_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{t_2} & i\mathbf{I}_{t_2} \\ \mathbf{0} & \mathbf{I}_{t_2} & -i\mathbf{I}_{t_2} \end{bmatrix}.$$

Coefficient embedding. A number field $K = \mathbb{Q}(\zeta)$ of degree n can be seen as a vector space of dimension n over the rationals with basis $\{1, \zeta, \ldots, \zeta^{n-1}\}$, meaning that each element $x \in K$ can be written as $x = \sum_{0 \leq j \leq n-1} x_j \zeta^j$ 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 $\tau(x) = [x_0, \ldots, x_{n-1}]^T$. For simplicity, we use $\tau_k(x)$ to denote x_k. For a positive integer η, we define $R_\eta = \tau^{-1}(\{0, \ldots, \eta - 1\}^n)$, which coincides with the set of representatives of $R/\eta R$. The embedding τ can also be extended to $K_{\mathbb{R}}$, mapping it to $\mathbb{R}^n$.

Canonical embedding. Another way to embed K is to use the *canonical embedding*. K has exactly n field homomorphisms $\sigma_1, \ldots, \sigma_n$, which are characterized by the fact that they map ζ to one of the distinct roots of f. We order them so that $\sigma_1, \ldots, \sigma_{t_1}$ map to one of the real roots, and $\sigma_{t_1+1}, \ldots, \sigma_{t_1+2t_2}$ map to one of the complex roots. The *canonical embedding* is the field homomorphism from K to $\mathbb{C}^n$ defined by $\sigma(x) = [\sigma_1(x), \ldots, \sigma_n(x)]^T$, and the addition and multiplication are done component-wise. As f has rational coefficients, it holds that the complex embeddings come in conjugate pairs, and therefore the range of σ is a subset of H. We can thus map K to $\mathbb{R}^n$ with $\sigma_H = \mathbf{U}_H^\dagger \sigma$. We extend the embeddings to vectors in K^d in the natural way by concatenating the embedding vectors of each coefficient, i.e., $\tau(\mathbf{x}) = [\tau(x_1)^T, \ldots, \tau(x_d)^T]^T$ and similarly for σ and σ_H. 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 $\|\mathbf{x}\|_{2,\infty} = \max_{k\in[n]} \sqrt{\sum_{i\in[d]} |\sigma_k(x_i)|^2}$. We define the field trace of K using the canonical embedding σ as $\mathrm{Tr}(x) = \sum_{k\in[n]} \sigma_k(x)$ for all $x \in K$. We can then define the dual of R by $R^\vee = \{x \in K : \mathrm{Tr}(xR) \subseteq \mathbb{Z}\}$. We also define the field norm of K as $N(x) = \prod_{k\in[n]} \sigma_k(x)$ for all $x \in K$. By [LPR13a, Lem. 2.15], there exists λ such that for all $p \in \mathbb{Z}$, $\lambda \cdot R_p^\vee = R_p$. For $x \in R_p^\vee$, we denote by $\tilde{x} = \lambda x \in R_p$, and we extend this notation for vectors and matrices. For our purposes, the existence of this scaling factor is enough but the construction of such a factor from [LPR13a] may not be suitable for other applications. For that, we refer to the discussion and results by Roşca et al. [RSW18, Sec. 3].

Distortion between embeddings. Both embeddings play important roles in this paper, and we recall that the two embeddings are linked by the linear relation

$$\sigma(x) = \mathbf{V}\tau(x) \text{ for all } x \in K, \text{ where } \mathbf{V} = \begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ | & | & & | \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{bmatrix}$$

is the Vandermonde matrix defined by the roots $(\alpha_k)_{k\in[n]}$ of the defining polynomial f. 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 $B_\eta = \max_{x\in R_\eta} \|\sigma(x)\|_\infty$ for a positive integer η. This parameter is inherent to the ring and intervenes in the proof of Lemma 3.2 and 3.5. Here, we provide an upper-bound on B_η, that is further simplified for cyclotomic number fields. The proof is provided in Appendix A.1 for completeness.

Lemma 2.1. *Let K be a number field of degree n, R its ring of integers, and $\mathbf{V}$ the associated Vandermonde matrix. Let η be a positive integer. Then, it holds that $1 \leq B_\eta = \max_{x\in R_\eta} \|\sigma(x)\|_\infty \leq n(\eta-1)\|\mathbf{V}\|_{\max}$. In particular, for cyclotomic fields, it yields $1 \leq B_\eta \leq n(\eta-1)$.*

Multiplication matrices. The multiplication in K (or $K_\mathbb{R}$) translates into a matrix-vector multiplication once embedded with either τ, σ or σ_H. In the canonical embedding, the multiplication matrix can be easily expressed as we have that for all x and y in K, $\sigma(x \cdot y) = \sigma(x) \odot \sigma(y) = \mathrm{diag}(\sigma(x)) \cdot \sigma(y)$, where $\odot$ denotes the coefficient-wise product or Hadamard product. Therefore, the multiplication matrix is $M_\sigma(x) = \mathrm{diag}(\sigma(x))$. We can then express the multiplication

matrix with respect to σ_H as $M_{\sigma_H}(x) = \mathbf{U}_H^\dagger M_\sigma(x)\mathbf{U}_H$. In the coefficient embedding, we can still write $\tau(x \cdot y)$ as $M_\tau(x) \cdot \tau(y)$, but the expression of $M_\tau(x)$ is more involved. We differ the proof in Appendix A.1.

Lemma 2.2. *Let $K = \mathbb{Q}(\zeta)$ be a number field of degree n, and $f = x^n + \sum_{k=0}^{n-1} f_k x^k$ the minimal polynomial of ζ. Then for all x in K, it holds that*

$$M_\tau(x) = \sum_{k=0}^{n-1} \tau_k(x)\mathbf{C}^k, \text{ with } \mathbf{C} = \begin{bmatrix} 0 & \cdots & 0 & -f_0 \\ & & & -f_1 \\ & \mathbf{I}_{n-1} & & \vdots \\ & & & -f_{n-1} \end{bmatrix}$$

the companion matrix of the minimal polynomial f.

In power-of-two cyclotomic fields, we have $f = x^n + 1$ yielding that $\mathbf{C}$ is the generating nega-circulant matrix. The expression of $M_\tau(x)$ can be simplified to

$$M_\tau(x) = \begin{bmatrix} x_0 & -x_{n-1} & \cdots & -x_1 \\ x_1 & x_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -x_{n-1} \\ x_{n-1} & x_{n-2} & \cdots & x_0 \end{bmatrix} \in \mathbb{Q}^{n \times n},$$

which is itself a nega-circulant matrix, with $x_k = \tau_k(x)$. We can also translate the matrix-vector multiplication in K^d to a matrix-vector multiplication in $\mathbb{R}^{nd}$ by extending the multiplication matrix maps M_σ, M_{σ_H} and M_τ to a matrix in $K^{m \times d}$. More precisely, for a matrix $\mathbf{A} = [a_{ij}]_{(i,j)} \in K^{m \times d}$, we define the block matrix $M_\sigma(\mathbf{A}) = [M_\sigma(a_{ij})]_{(i,j)}$. We define $M_{\sigma_H}(\mathbf{A})$ and $M_\tau(\mathbf{A})$ the same way. As we need it later in this paper, we provide a way to obtain the singular values of such block matrices. This relies on a unified analysis from [Rja94] which gives conditions to obtain the eigenvalues of a matrix when described by blocks. In our setting, we end up showing that the spectral analysis of the entire block matrix $M_\tau(\mathbf{A})$ comes down to finding the singular values of the n embedded matrices $\sigma_k(\mathbf{A})$. For convenience, we write $S(\mathbf{A})$ the set of all singular values of a complex matrix $\mathbf{A}$. The proof can be found in Appendix A.1.

Lemma 2.3. *Let K be a number field of degree n, and d, m positive integers. Let $\mathbf{A}$ be a matrix in $K^{d \times m}$.*

$$S(M_\tau(\mathbf{A})) = \bigcup_{k \in [n]} S(\sigma_k(\mathbf{A})) = S(M_\sigma(\mathbf{A})) = S(M_{\sigma_H}(\mathbf{A})),$$

where $\sigma_k(\mathbf{A}) = [\sigma_k(a_{ij})]_{(i,j) \in [d] \times [m]}$. In particular, it holds that $\|M_\tau(\mathbf{A})\|_2 = \max_{k \in [n]} \|\sigma_k(\mathbf{A})\|_2$.

Ideals, units and modules. An ideal $\mathfrak{p} \neq R$ is *prime* if for all $a, b \in R$, $ab \in \mathfrak{p}$ implies that a or b is in $\mathfrak{p}$. For two ideals $\mathcal{I}$ and $\mathcal{J}$, the sum $\mathcal{I} + \mathcal{J}$ is the set

of all $x+y$, where $(x,y) \in \mathcal{I} \times \mathcal{J}$, while the product $\mathcal{I}\mathcal{J}$ is the set of all finite sums of xy, where $(x,y) \in \mathcal{I} \times \mathcal{J}$. An integer q is said to be unramified in R if the ideal $\langle q \rangle$ can be factored in a product of distinct prime ideals. 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 $N(\langle a \rangle) = |N(a)|$ for any $a \in K$.

In the construction of Lemma 3.2, we need a condition for small norm elements of R_q to be invertible for a specific q. To do so, we rely on the small norm condition proven in [LS18, Th. 1.1].

Lemma 2.4 ([LS18, Th. 1.1]). *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 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 spectral norm 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 [LPR13a] 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 then recall a result by Wang and Wang [WW19] on the linear independence of vectors in R_q^d.

Lemma 2.5 ([WW19, Lem. 9]). *Let K be a number field of degree n, and R its ring of integers. Let d, q be positive integers such that q is an unramified prime. Let i be in $\{0, \ldots, d-1\}$, and $\mathbf{a}_1, \ldots, \mathbf{a}_i \in R_q^d$ be R_q-linearly independent vectors of R_q^d. Then*

$$\begin{aligned}\mathbb{P}_{\mathbf{b} \hookleftarrow U(R_q^d)}[\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{b} \text{ are } R_q\text{-linearly independent}] \\ &= 1 - \left(1 - \prod_{k \in [r]} (1 - q^{-g_k})\right)^d \\ &\geq 1 - n/q,\end{aligned}$$

where r and the g_k are the integers from the prime ideal factorization of $\langle q \rangle$, namely $\langle q \rangle = \prod_{k \in [r]} \mathfrak{p}_k$ with $N(\mathfrak{p}_k) = q^{g_k}$.

As we use this lemma for different module ranks d, we denote by p_d this probability, i.e., $p_d = 1 - (1 - \prod_{k \in [g]}(1 - q^{-f_k}))^d$ where the modulus q and the ring R are implicit. Additionally, for $m \geq d$, we denote by $\delta(m,d)$ the probability that a uniform matrix $\mathbf{A}$ in $R_q^{m \times d}$ is singular in R_q, i.e., there is no subset S composed of d rows of $\mathbf{A}$ that are R_q-linearly independent vectors. By Lemma 2.5, this singularity probability is given by

$$\delta(m,d) = \sum_{k=0}^{d-1} \binom{m}{k} p_d^k (1-p_d)^{m-k}. \tag{1}$$

2.2 Lattices

A (full-rank) *lattice* Λ of rank n is a discrete additive subgroup of $\mathbb{R}^n$. Since H is isomorphic to $\mathbb{R}^n$, we may consider lattices that are discrete subgroups of H. Each lattice can be represented by a basis $\mathbf{B} = [\mathbf{b}_i]_{i\in[n]} \in \mathbb{R}^{n\times n}$ as the set of all integer linear combinations of the $\mathbf{b}_i$, i.e., $\Lambda = \mathbf{B}\mathbb{Z}^n$. 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., $\lambda_1^\infty(\Lambda) = \min_{\mathbf{x}\in\Lambda\setminus\{\mathbf{0}\}} \|\mathbf{x}\|_\infty$. 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. 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 if $\mathcal{I}^d$ denotes $\mathcal{I} \times \ldots \times \mathcal{I}$, then $\lambda_1^\infty(\mathcal{I}^d) = \lambda_1^\infty(\mathcal{I})$.

We recall the *Generalized Independent Vectors Problem* (GIVP) as defined in [LS15]. We can then define Id-GIVP$_\gamma^\phi$ (resp. Mod-GIVP$_\gamma^\phi$) as the restriction of GIVP$_\gamma^\phi$ to ideal lattices (resp. module lattices).

Definition 2.1. *Let N be a positive integer, $\gamma \geq 1$ a function of N, and ϕ an arbitrary function that maps a lattice of dimension N to a positive real. The* GIVP$_\gamma^\phi$ *is as follows. Given a lattice Λ of rank N, find N linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_N$ in Λ such that* $\max_{i\in[N]} \|\mathbf{v}_i\|_2 \leq \gamma \cdot \phi(\Lambda)$.

2.3 Probabilities

For a finite set S, we define $|S|$ to be its cardinality, and $U(S)$ to be the uniform probability distribution over S. The action of sampling $x \in S$ from a distribution P is denoted by $x \hookleftarrow P$. We now define two distances for probability distributions, namely the *statistical distance* Δ, and the *Rényi divergence* [Rén61,vEH14] RD. The Rényi divergence was thoroughly studied for its use in cryptography as a powerful alternative for the statistical distance measure by Bai et al. [BLR+18]. In this paper, it suffices to use the Rényi divergence of order 2 denoted by RD_2.

Definition 2.2. *Consider two discrete probability distributions P and Q over a countable set S. The* statistical distance *between P and Q is defined by $\Delta(P, Q) = \frac{1}{2}\sum_{x\in S} |P(x) - Q(x)|$. If* $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$*, we define the* Rényi divergence of order 2 *as* $\mathrm{RD}_2(P\|Q) = \sum_{x\in\mathrm{Supp}(P)} \frac{P(x)^2}{Q(x)}$*. The two definitions extend to continuous distributions by replacing the discrete sum with an integral.*

The two distances enjoy a probability preservation property, which are essential in proving our results. The Rényi divergence is also multiplicative, as proven by van Erven and Harremoës [vEH14].

Lemma 2.6. *Let P, Q be two probability distributions with* $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$ *and $E \subseteq \mathrm{Supp}(Q)$ be an arbitrary event. Then, $P(E) \leq \Delta(P,Q) + Q(E)$, and $P(E)^2 \leq \mathrm{RD}_2(P\|Q) \cdot Q(E)$. Further, let $(P_n)_{n\in\mathbb{N}}, (Q_n)_{n\in\mathbb{N}}$ be two families of independent discrete probability distributions with* $\mathrm{Supp}(P_n) \subseteq \mathrm{Supp}(Q_n)$ *for all $n \in \mathbb{N}$. It holds that*

$$\mathrm{RD}_2\left(\prod_{n\in\mathbb{N}} P_n \| \prod_{n\in\mathbb{N}} Q_n\right) = \prod_{n\in\mathbb{N}} \mathrm{RD}_2(P_n\|Q_n).$$

Leftover Hash Lemma. In this work, we use a formulation of the leftover hash lemma (LHL) that is an adaptation of the one by Micciancio [Mic07], which, instead of working with vectors over the finite field $\mathbb{Z}_q$, operates over principal ideal domains. Given a number field $K = \mathbb{Q}(\zeta)$, where the corresponding ring of integers has the form $R = \mathbb{Z}[\zeta]$, and a prime q, then R/qR is a principal ideal domain, allowing for a unique prime-ideal factorization. Further, we provide not only a bound on the statistical distance, but also on the Rényi divergence

Lemma 2.7. *Let n, k, d, q, η be positive integer with q prime. Further, let $K = \mathbb{Q}(\zeta)$ be a number field of degree n whose ring of integers is given by $R = \mathbb{Z}[\zeta]$. Then, it holds that*

$$\Delta((\mathbf{C}, \mathbf{Cz}), (\mathbf{C}, \mathbf{s})) \leq \frac{1}{2}\sqrt{\left(1 + \frac{q^k}{\eta^d}\right)^n - 1} \text{ and}$$
$$\mathrm{RD}_2((\mathbf{C}, \mathbf{Cz})\|(\mathbf{C}, \mathbf{s})) \leq \left(1 + \frac{q^k}{\eta^d}\right)^n,$$

where $\mathbf{C} \hookleftarrow U(R_q^{k\times d})$, $\mathbf{z} \hookleftarrow U(R_\eta^d)$ and $\mathbf{s} \hookleftarrow U(R_q^k)$.

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 also use $\Psi_{\leq\alpha}$ to denote the set of Gaussian distributions $D_{\mathbf{r}}$ with $\|\mathbf{r}\|_\infty \leq \alpha$.

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 $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\Lambda) = \sum_{\mathbf{y}\in\Lambda} D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{y})$.

Definition 2.3 (Sub-Gaussian Distribution). *Let n be a positive integer, and $\mathbf{x}$ a (discrete or continuous) random vector over $\mathbb{R}^n$. We say that $\mathbf{x}$ is sub-*

Gaussian with sub-Gaussian moment s, if for all unit vector $\mathbf{u} \in \mathbb{R}^n$, and all $t \in \mathbb{R}$, we have $\mathbb{E}[\exp(2\pi t\langle \mathbf{x}, \mathbf{u}\rangle)] \leq e^{\pi s^2 t^2}$.

A standard calculation shows that the discrete Gaussian distribution $\mathcal{D}_{\Lambda,s}$ is sub-Gaussian with sub-Gaussian moment s [MP12, Lem. 2.8], for any lattice Λ and $s > 0$.

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

We now give a few results related to discrete Gaussian distributions that we need in this paper. The first is due to Micciancio and Regev [MR07] and shows that above the smoothing parameter, a continuous Gaussian coset is statistically close to uniform.

Lemma 2.9 ([MR07, Lem. 4.1]). *Let Λ be lattice of rank n, $\varepsilon > 0$, 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 also need the following result on the sum of convoluted Gaussian distributions. Note that the distribution of $\mathbf{y}$ depends on $\mathbf{x}$.

Lemma 2.10 ([BLP+13, Lem. 2.10] & [Pei10, Thm. 3.1]). *Let Λ be lattice of rank n. Let $\varepsilon \in (0, 1/2]$, and $\beta, r > 0$ be 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 $\mathcal{D}_{\Lambda,\sqrt{\beta^2+r^2}}$.*

Finally, we need the Rényi divergence between two shifted discrete Gaussians.

Lemma 2.11 (Adapted from [LSS14, Lem. 4.2]). *Let Λ be lattice of rank n, $\varepsilon \in (0, 1)$, $s \geq \eta_\varepsilon(\Lambda)$, and $\mathbf{c}$ a vector of $\mathbb{R}^n$. Then,*

$$\mathrm{RD}_2(\mathcal{D}_{\Lambda,\mathbf{c},s} \| \mathcal{D}_{\Lambda,s}) \leq \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \cdot \exp\left(\frac{2\pi\|\mathbf{c}\|_2^2}{s^2}\right).$$

Gaussians over number fields. 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 [LPR13a], 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 $\mathbf{x}^{(H)} \in \mathbb{R}^{nd}$ 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 $\mathbf{x} = \sigma_H^{-1}(\mathbf{x}^{(H)})$. To ease readability, we denote the described distribution of $\mathbf{x} \in K_{\mathbb{R}}^d$ by $D_{\sqrt{\mathbf{\Sigma}}}$.

We first provide an upper bound on the spectral norm of a discrete Gaussian matrix, once embedded via $M_{\sigma_H}(\cdot)$. This combines a bound on the spectral norm of a block matrix from the spectral norm of each block, with a discrete Gaussian tail bound. Although it seems folklore, we weren't able to find a Gaussian tail bound on $\sigma(x)$ in the infinity norm for $x \hookleftarrow \mathcal{D}_{R^\vee,s}$. We therefore derive such a bound, which is based on [Pei08, Cor. 5.3] proving that $\|\sigma(x)\|_\infty \leq s \log_2 n$ with overwhelming probability. The proof can be found in Appendix A.1.

Lemma 2.12. *Let K be a number field of degree n, and R its ring of integers. Let $\mathcal{I}$ be any (fractional) ideal of R. Let m, d be positive integer, and $s > 0$. Then, for all $t \geq 0$ it holds that*

$$\mathbb{P}_{\mathbf{N} \hookleftarrow \mathcal{D}_{\mathcal{I},s}^{m \times d}} \left[\|M_{\sigma_H}(\mathbf{N})\|_2 \geq \sqrt{md} \cdot st \right] \leq 2nmd \cdot e^{-\pi t^2}.$$

Choosing $t = \log_2 n$ gives $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq s \log_2(n) \sqrt{md}$ with overwhelming probability if m, d are polynomial in n.

In the proof of Lemma 3.3, we also need 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. The proof is in Appendix A.1 for completeness.

Lemma 2.13. *Let K be a number field of degree n, and m, d positive integers. Let $\mathbf{S} \in \mathbb{R}^{nd \times nd}$ be a positive semi-definite matrix, and $\mathbf{U} \in K_{\mathbb{R}}^{m \times d}$. We denote $\mathbf{\Sigma} = M_{\sigma_H}(\mathbf{U}) \mathbf{S} M_{\sigma_H}(\mathbf{U})^T \in \mathbb{R}^{nm \times nm}$, Then, the distribution of $\mathbf{y} = \mathbf{U}\mathbf{e}$, where $\mathbf{e} \in K_{\mathbb{R}}^d$ is distributed according to $D_{\sqrt{\mathbf{S}}}$, is exactly $D_{\sqrt{\mathbf{\Sigma}}}$ over $K_{\mathbb{R}}^m$.*

We also need another lemma related to the inner product of $K_{\mathbb{R}}^d$ (which results in an element of $K_{\mathbb{R}}$) between a discrete Gaussian vector and an arbitrary one. In particular, we use Lemma 2.14 in the proof of Lemma 3.5 in order to decompose a Gaussian noise into an inner product. It generalizes [Reg09, Cor. 3.10] to the module case. A specific instance is proven in the proof of [LS15, Lem. 4.15], which is later mentioned (without proof) in [RSW18, Lem. 5.5]. We differ the proof in Appendix A.1.

Lemma 2.14 (Adapted from [Reg09, Cor. 3.10]). *Let $M \subseteq K^d$ 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 $j \in [n]$.*

2.4 Function Families

In Section 4, we prove that certain families of functions are hard to invert, or whose output are hard to distinguish from uniformly random ones. As such, we give in this section the notion of *function families* as well as the standard security properties that we desire from them. A function family $\mathcal{F}$ over a set of functions F is a probability distribution over F, where each function of F has domain X and range Y.

Definition 2.4. *Let X, Y be two sets, and F a set of functions from X to Y. Let $\mathcal{F}, \mathcal{G}$ be two function families over F. Let $\mathcal{X}$ be a probability distribution over X, and $\varepsilon \in (0,1)$.*

Indistinguishability. *$\mathcal{F}$ and $\mathcal{G}$ are ε-indistinguishable if for all PPT algorithm $\mathcal{A}$, it holds $|\mathbb{P}_{f \hookleftarrow \mathcal{F}}[\mathcal{A}(f) = 1] - \mathbb{P}_{g \hookleftarrow \mathcal{G}}[\mathcal{A}(g) = 1]| \leq \varepsilon$.*

Pseudorandomness. *$(\mathcal{F}, \mathcal{X})$ is ε-pseudorandom if for all PPT algorithm $\mathcal{A}$, it holds $\left|\mathbb{P}_{(f,x) \hookleftarrow \mathcal{F} \times \mathcal{X}}[\mathcal{A}(f, f(x)) = 1] - \mathbb{P}_{(f,y) \hookleftarrow \mathcal{F} \times U(Y)}[\mathcal{A}(f, y) = 1]\right| \leq \varepsilon$.*

Second preimage resistance. *$(\mathcal{F}, \mathcal{X})$ is ε-second preimage resistant if for all PPT algorithm $\mathcal{A}$, it holds $\mathbb{P}_{\substack{(f,x) \hookleftarrow \mathcal{F} \times \mathcal{X} \\ x' \leftarrow \mathcal{A}(f,x)}}[x \neq x' \wedge f(x) = f(x')] \leq \varepsilon$.*

Uninvertibility. *$(\mathcal{F}, \mathcal{X})$ is ε-uninvertible if for all PPT algorithm $\mathcal{A}$, it holds that $\mathbb{P}_{(f,x) \hookleftarrow \mathcal{F} \times \mathcal{X}}[\mathcal{A}(f, f(x)) = x] \leq \varepsilon$.*

One-wayness. *$(\mathcal{F}, \mathcal{X})$ is ε-one-way if for all PPT algorithm $\mathcal{A}$, it holds that $\mathbb{P}_{(f,x) \hookleftarrow \mathcal{F} \times \mathcal{X}}[f(\mathcal{A}(f, f(x))) = f(x)] \leq \varepsilon$.*

If ε is negligible in the security parameter, we omit it. We then give sufficient conditions to ensure some of these security properties.

Lemma 2.15 ([MP13, Lem. 2.2]). *Let $\mathcal{F}$ be a family of functions computable in polynomial time. Let $\mathcal{X}$ be a distribution on X. If $(\mathcal{F}, \mathcal{X})$ is ε-uninvertible and ε'-second preimage resistant, then it is also $(\varepsilon + \varepsilon')$-one-way.*

Lemma 2.16 ([MP13, Lem. 2.4]). *Let $\mathcal{F}$ be a function family with finite domain X. For $\varepsilon = \mathbb{E}_{f \hookleftarrow \mathcal{F}}[|f(X)|]/|X|$, it holds that $(\mathcal{F}, U(X))$ is ε-uninvertible, even against unbounded adversaries.*

Lemma 2.17 ([MP13, Lem. 2.5]). *Let $\mathcal{F}$ be a function family with domain X and range Y, and $\mathcal{G}$ be a family of efficiently computable functions with domain $X' \supseteq Y$. Let $\mathcal{X}$ be a distribution on X. If $(\mathcal{F}, \mathcal{X})$ is uninvertible, then so is $(\mathcal{G} \circ \mathcal{F}, \mathcal{X})$.*

We now recall the notion of *lossy function family* from [MP13]. Note that by an indistinguishability argument, if $(\mathcal{F}, \mathcal{G}, \mathcal{X})$ is a lossy function family, then so is $(\mathcal{G}, \mathcal{F}, \mathcal{X})$. In particular, by Lemma 2.15, both $(\mathcal{F}, \mathcal{X})$ and $(\mathcal{G}, \mathcal{X})$ are one-way.

Definition 2.5. *Let X, Y be two sets, and F a set of efficiently computable functions from X to Y. Let $\mathcal{F}, \mathcal{G}$ be two function families over F. Let $\mathcal{X}$ be an efficiently sampleable probability distribution over X. Then $(\mathcal{F}, \mathcal{G}, \mathcal{X})$ is a lossy function family if it holds that*

– $\mathcal{F}$ *and* $\mathcal{G}$ *are indistinguishable;*
– $(\mathcal{F}, \mathcal{X})$ *is uninvertible;*
– $(\mathcal{G}, \mathcal{X})$ *is second preimage resistant.*

2.5 Module Learning With Errors

The module variant of LWE was first defined by Brakerski et al. [BGV12] and thoroughly studied by Langlois and Stehlé [LS15]. It describes the following problem. Let K be a number field of degree n and R its ring of integers with dual $R^\vee$. Further, let d denote the rank and let ψ be a distribution on $K_\mathbb{R}$ and $\mathbf{s} \in (R_q^\vee)^d$ be a vector. We also define the torus $\mathbb{T}_{R^\vee} = K_\mathbb{R}/R^\vee$. We let $A^{\mathcal{M}}_{\mathbf{s},\psi}$ 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 $e \hookleftarrow \psi$ and returning $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod R^\vee)$.

Definition 2.6 (Module Learning With Errors). *Let q, d be positive integers with $q \geq 2$. Let Ψ be a family of distributions on $K_\mathbb{R}$. The search version* M-SLWE$_{n,d,q,\Psi}$ *is as follows: Let $\mathbf{s} \in (R_q^\vee)^d$ be secret and $\psi \in \Psi$. Given arbitrarily many samples from $A^{\mathcal{M}}_{\mathbf{s},\psi}$, the goal is to find $\mathbf{s}$. Let Υ be a distribution on a family of distributions on $K_\mathbb{R}$. Its decision version* M-LWE$_{n,d,q,\Upsilon}$ *is as follows: Choose $\mathbf{s} \hookleftarrow U((R_q^\vee)^d)$ and $\psi \hookleftarrow \Upsilon$. The goal is to distinguish between arbitrarily many independent samples from $A^{\mathcal{M}}_{\mathbf{s},\psi}$ and the same number of independent samples from $U(R_q^d \times \mathbb{T}_{R^\vee})$.*

The M-LWE problem encompasses its preceding variants LWE, corresponding to a field of degree $n = 1$, and R-LWE, corresponding to the module rank $d = 1$. We describe here the several variants and notations that we consider in this paper.

Fixed number of samples. When using the Rényi divergence as a tool to measure the distance between two probability distributions, we need to fix the number of requested samples a priori. Let m be the number of requested M-LWE samples $(\mathbf{a}_i, q^{-1}\langle \mathbf{a}_i, \mathbf{s}\rangle + e_i \bmod R^\vee)$ for $i \in [m]$, then we consider the matrix $\mathbf{A} \in R_q^{m \times d}$ whose rows are the $\mathbf{a}_i$'s and we set $\mathbf{e} = [e_1, \ldots, e_m]^T$. We obtain the representation $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod R^\vee)$. We denote it by M-LWE$_{n,d,m,q,\Upsilon}$.

Multiple secrets. Let k, m be positive integers, where m denotes the number of requested samples. In the multiple secrets version, the secret vector $\mathbf{s} \in (R_q^\vee)^d$ is replaced by a secret matrix $\mathbf{S} \in (R_q^\vee)^{d \times k}$ and the error vector $\mathbf{e} \leftarrow \psi^m$ by an error matrix $\mathbf{E} \leftarrow \psi^{m \times k}$. There is a simple polynomial-time reduction from M-LWE using a secret vector to M-LWE using a secret matrix for any k polynomially large in d via a hybrid argument, as given for instance in [Mic18, Lem. 2.9]. We denote the corresponding problem by M-LWE$^k_{n,d,m,q,\Upsilon}$.

Discrete version. As pointed out by Lyubashevsky et al. [LPR13a], sometimes it can be more convenient to work with a discrete variant, where the second component b of each sample $(\mathbf{a}, b)$ is taken from a finite set, and not from the continuous torus $\mathbb{T}_{R^\vee}$. Indeed, for the case of M-LWE, if the rounding function $\lfloor \cdot \rceil : K_\mathbb{R} \to R^\vee$ is chosen in a suitable way, see e.g. [LPR13b, Sec. 2.6],

then every sample $(\mathbf{a}, b = q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod R^\vee) \in R_q^d \times \mathbb{T}_{R^\vee}$ from $A^{\mathcal{M}}_{\mathbf{s},\psi}$ can be transformed to $(\mathbf{a}, \lfloor q \cdot b \rceil \bmod qR^\vee) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s}\rangle + \lfloor q \cdot e \rceil \bmod qR^\vee) \in R_q^d \times R_q^\vee$. We use the latter representation in Section 3.1 and 4.

Bounded secret. Another possibility is to choose a *small* secret, i.e., whose coefficients are bounded by $\eta \ll q$. Note that the bound η is with regard to the coefficient embedding τ, meaning that the secret is in $(R_\eta^\vee)^d$. We denote the corresponding problem by η-M-LWE$_{n,d,q,\chi}$. All the result can easily be extended to secrets from $(R_S^\vee)^d$, where $R_S = \tau^{-1}(S^n)$ for a set $S \subseteq \mathbb{Z}$. It would involve two quantities related to S, namely $|S|$ and $\max_{x\in S}|x|$. In particular, the results also apply to secrets that have coefficients in $S = \{-\eta, \dots, \eta\}$.

M-LWE and M-SIS function families. We now introduce M-LWE and M-SIS (Module Short Integer Solution [LS15]) with their respective function family. In most LWE-based schemes, the secret key is $(\mathbf{s}, \mathbf{e})$ and the public key is $(\mathbf{A}, \mathbf{b} = \mathbf{As} + \mathbf{e})$. Note that it is therefore important to prove one-wayness and not just uninvertibility because an adversary breaking one-wayness could compute a different secret key for the same public key, which would allow them to decrypt messages, or forge signatures. It turns out that if the parameters are chosen appropriately so that the function is second preimage resistant, the uninvertibility is then equivalent to the one-wayness. Their uninvertibility or one-wayness therefore captures the hardness of the corresponding search problem, while their pseudorandomness captures the hardness of the decision problem. We only define them with discrete inputs (i.e., discrete error for M-LWE) because they are only needed in Section 4 which studies errors in $(R_\eta^\vee)^m$.

Definition 2.7. *Let K be a number field of degree n, and R its ring of integers. Let d, q, m be positive integers, and $X \subseteq (R^\vee)^m$. The* M-SIS(n, d, m, q, X) *function family is the distribution obtained by sampling a matrix $\mathbf{A} \in R_q^{m\times d}$ uniformly at random, and outputting $f_\mathbf{A}$ defined by $f_\mathbf{A}(\mathbf{x}) = \mathbf{A}^T\mathbf{x} \bmod qR^\vee$ for all $\mathbf{x} \in X$. The* M-LWE(n, d, m, q, X) *function family is the distribution obtained by sampling $\mathbf{A} \in R_q^{m\times d}$ uniformly at random and outputting $g_\mathbf{A}$ defined by $g_\mathbf{A}(\mathbf{s}, \mathbf{e}) = \mathbf{As} + \mathbf{e} \bmod qR^\vee$ for all $(\mathbf{s}, \mathbf{e}) \in (R_q^\vee)^d \times X$.*

We now state the hardness result of M-LWE that we use in terms of the security properties of the function family.

Theorem 2.1 ([LS15, Thm. 4.7]). *Let K be the ν-th cyclotomic field of degree $n = \varphi(\nu)$. Let d, q, m be positive integers and $\alpha \in (0, 1)$ such that $m = \mathsf{poly}(\lambda)$, and q is of known factorization such that $\alpha q > 2\sqrt{d} \cdot \omega(\sqrt{\log_2 n})$. Let $\varepsilon = (nd)^{-\omega(1)}$. Assuming that Mod-*GIVP$_\gamma^{\eta_\varepsilon}$ *is (quantumly) hard, then it holds that*

$$\big(\text{M-LWE}(n, d, m, q, (R^\vee)^m), U((R_q^\vee)^d) \times \mathcal{D}^m_{R^\vee,\alpha q}\big)$$

is uninvertible, where $\gamma = \sqrt{8nd^2} \cdot \omega(\sqrt{\log_2 n})/\alpha$. If q is prime such that $q = 1 \bmod \nu$, then it is also pseudorandom.

If the module rank d is 1, we can use [LPR13a, Thm. 4.1, 5.1] instead which provides the hardness result for R-LWE. We can then choose $\alpha q > \omega(\sqrt{\log_2 n})$. This requires to change the assumption to ideal lattices, namely that Id-GIVP$_\gamma^{\eta_\varepsilon}$ is hard for $\gamma = \sqrt{n}\omega(\sqrt{\log_2 n})/\alpha$.

3 Hardness of η-M-LWE

In this section, we prove the hardness of the η-bounded secret version of M-LWE, if the module rank is (super-)logarithmic in the degree n of the underlying number field. To the best of our knowledge, this is the first result on the hardness of a structured variant of LWE with small bounded secret. We propose two independent proofs that achieve different results. The first one in Section 3.1 proves the hardness of the search version of η-M-LWE, using a more direct proof. The second one in Section 3.2 is more involved but allows for proving the hardness of the decision version of η-M-LWE as well as (slightly) improving the noise parameter.

3.1 Computational Hardness Using the Rényi Divergence

We start by proving the hardness of η-M-SLWE with a quite direct reduction. To facilitate the understanding, we illustrate the high level idea of the proof in Figure 3.1. Given an instance $(\mathbf{A}, \mathbf{Az} + \mathbf{e})$ of η-M-SLWE, our goal is to transform it into a related instance of M-SLWE defined by $(\mathbf{B}, \mathbf{Bs}+\mathbf{e}')$. Note that the secret $\mathbf{z}$ is in $(R_\eta^\vee)^d$, while the secret $\mathbf{s}$ is in $(R_q^\vee)^k$. At the core of the proof lies a lossy argument, where the public matrix $\mathbf{A}$ is replaced by a lossy matrix $\mathbf{BC}+\mathbf{N}$, which corresponds to the second part of some multiple-secrets M-LWE sample. Note that the rank of the matrix $\mathbf{B}$ is smaller than the one of $\mathbf{A}$, motivating the description *lossy*. Here, we can see that this argument does not work for R-LWE (which corresponds to M-LWE with rank 1) as it is not possible to replace the public matrix consisting of one column by a matrix of smaller rank. To argue that an adversary cannot distinguish between the two cases, we need to assume the hardness of the *decision* M-LWE problem as well. In a second step, the term $\mathbf{Nz} + \mathbf{e}$ is replaced by the new noise $\mathbf{e}'$, where the Rényi divergence between both expressions can be bounded by a constant using properties of the Rényi divergence of Gaussian distributions. Finally, the product $\mathbf{Cz}$ is replaced by the uniform secret $\mathbf{s}$, where the Rényi divergence between both elements can be bounded by a constant using Lemma 2.7. The use of the leftover hash lemma is also the reason why our reduction only works for module ranks larger than $\log_2 q + \Omega(\log_2 n)$. Informally speaking, it requires the ratio between the number of rows of $\mathbf{C}$ and its number of columns to be logarithmic in order to bound the Rényi divergence by a constant. We end up with some standard M-LWE instance, which is hard to solve due to our hardness assumption.

This consists of a reduction from M-SLWE and M-LWE with rank k to η-M-SLWE with rank $d \geq k \log_2(q)/\log_2(\eta) + \Omega(\log_2(n)/\log_2(\eta))$. It follows the

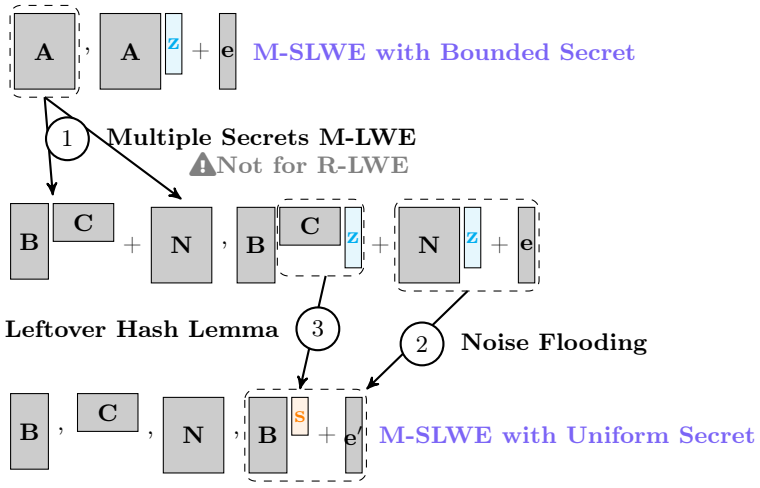

Fig. 3.1. Summary of the proof of Theorem 3.1

original proof structure of Goldwasser et al. [GKPV10], but achieves better parameters by using the Rényi divergence, while being as direct and short as the original proof. The improvement on the noise rate β/α compared to [GKPV10] comes from the fact that the Rényi divergence only needs to be constant for the reduction to work, and not necessarily negligibly close to 1 (compared to negligibly close to 0 for the statistical distance). However, using the Rényi divergence as a measure of distribution closeness requires to move to the search version of M-LWE. Overall, this reduction is restricted to number fields for which the ring of integers is $R = \mathbb{Z}[\zeta]$. Furthermore, the norm of the Vandermonde matrix $\|\mathbf{V}\|_2$ is better understood in cyclotomic fields. We study the M-LWE problem in its discrete version, as presented in Section 2.5.

Theorem 3.1. *Let $K = \mathbb{Q}(\zeta)$ be a number field of degree n such that its ring of integers is $R = \mathbb{Z}[\zeta]$. Let k, d, m, η and q be positive integers with q prime, $m, d = \mathsf{poly}(n)$ and $d \log_2 \eta \geq k \cdot \log_2 q + \Omega(\log_2 n)$. Further, let α and β be positive reals such that $\beta \geq \alpha \cdot d\sqrt{m} \cdot \|\mathbf{V}\|_2 \sqrt{n} \log_2(n)(\eta - 1)$. Let $\varepsilon = O(\frac{1}{m})$ be such that $\beta q \geq \eta_\varepsilon(R^\vee)$. There is a reduction from* $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R^\vee,\beta q}}$ *and* $\text{M-LWE}^d_{n,k,m,q,\mathcal{D}_{R^\vee,\alpha q}}$ *to* $\eta\text{-M-SLWE}_{n,d,m,q,\mathcal{D}_{R^\vee,\beta q}}$.

The degree n of K, the number of samples m and the modulus q are preserved. The reduction increases the rank of the module from k to $k \cdot \log_2 q / \log_2 \eta + \Omega(\log_2 n / \log_2 \eta)$ and the Gaussian width from αq to $\alpha q \cdot d\sqrt{m} \cdot \|\mathbf{V}\|_2 \sqrt{n} \log_2(n)(\eta - 1)$. In power-of-two cyclotomic fields, $\|\mathbf{V}\|_2 = \sqrt{n}$. In the p^k-th cyclotomic field with p an odd prime, we have $\|\mathbf{V}\|_2 = \sqrt{p^k}$. In general cyclotomic fields, we have $\|\mathbf{V}\|_2 \leq \|\mathbf{V}\|_F = (\sum_{i,j} |\alpha_i^{j-1}|^2)^{1/2} \leq n$ (as α_i is a root of unity). Also, $\text{M-LWE}_{n,k,m,q,\mathcal{D}_{R^\vee,\alpha q}}$ trivially reduces to $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R^\vee,\beta q}}$, as $\beta \geq \alpha$.

Proof. Fix any $n, k, d, m, q, \eta, \alpha, \beta$ and ε as in the statement of the theorem. Given an η-M-SLWE$_{n,d,m,q,\mathcal{D}_{R^\vee,\beta q}}$ sample $(\mathbf{A}, \mathbf{A}\cdot\mathbf{z}+\mathbf{e} \bmod qR^\vee) \in R_q^{m\times d} \times (R_q^\vee)^m$, with $\mathbf{z} \in (R_\eta^\vee)^d$ and $\mathbf{e} \hookleftarrow (\mathcal{D}_{R^\vee,\beta q})^m$, the search problem asks to find $\mathbf{s}$ and $\mathbf{e}$. In order to prove the statement, we define different hybrid distributions:

- $H_0 : (\mathbf{A}, \mathbf{Az}+\mathbf{e} \bmod qR^\vee)$, as in η-M-SLWE$_{n,d,m,q,\mathcal{D}_{R^\vee,\beta q}}$,
- $H_1 : (\mathbf{A}' = \lambda(\mathbf{BC}+\mathbf{N} \bmod qR^\vee), \mathbf{A}'\mathbf{z}+\mathbf{e} \bmod qR^\vee)$, where $\mathbf{B} \hookleftarrow U(R_q^{m\times k})$, $\mathbf{C} \hookleftarrow U((R_q^\vee)^{k\times d})$, and $\mathbf{Z} \hookleftarrow \mathcal{D}_{R^\vee,\alpha q}^{m\times d}$ and $\mathbf{z}, \mathbf{e}$ as in H_0,
- $H_2 : (\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{B}(\lambda\mathbf{Cz})+\mathbf{N}(\lambda\mathbf{z})+\mathbf{e} \bmod qR^\vee)$, where $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{z}, \mathbf{e}$ as in H_1,
- $H_3 : (\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{B}(\lambda\mathbf{Cz}) + \mathbf{e}' \bmod qR^\vee)$, where $\mathbf{e}' \hookleftarrow \mathcal{D}_{R^\vee,\beta q}^m$ and $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{z}$ as in H_2,
- $H_4 : (\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{Bs}+\mathbf{e}' \bmod qR^\vee)$, where $\mathbf{s} \hookleftarrow U((R_q^\vee)^k)$ and $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{e}'$ as in H_3.

For $i \in \{0, \ldots, 4\}$, we denote by P_i the problem of finding the secret $\mathbf{z}$ (resp. $\mathbf{s}$ in H_4), given a sample of the distribution H_i. We say that problem P_i is hard if for any probabilistic polynomial-time attacker $\mathcal{A}$ the advantage of solving P_i is negligible, thus $\mathrm{Adv}_{P_i}[\mathcal{A}(H_i) = \mathbf{z}] \leq n^{-\omega(1)}$, where n is the degree of K. The overall idea is to show that if P_4 is hard, then P_0 is hard as well.

From P_0 to P_1: By the hardness assumption of M-LWE$^d_{n,k,m,q,\mathcal{D}_{R^\vee,\alpha q}}$, the distributions H_0 and H_1 are computationally indistinguishable. By a hybrid argument, e.g., Lemma 3.4, one can reduce the single secret version to the multiple secret version while only incurring a loss factor of d in the advantage. Thus, if $\mathrm{Adv}_{\text{M-LWE}}$ is the advantage of an adversary against M-LWE$_{n,k,m,q,\mathcal{D}_{R^\vee,\alpha q}}$, it holds

$$\mathrm{Adv}_{P_0}[\mathcal{A}(H_0) = \mathbf{z}] \leq \mathrm{Adv}_{P_1}[\mathcal{A}(H_1) = \mathbf{z}] + d \cdot \mathrm{Adv}_{\text{M-LWE}},$$

where d is the number of secret vectors, i.e., the columns of the matrix $\mathbf{C}$.

From P_1 to P_2: Since more information is given in distribution H_2 than in distribution H_1, the problem P_1 is harder than P_2 and hence

$$\mathrm{Adv}_{P_1}[\mathcal{A}(H_1) = \mathbf{z}] \leq \mathrm{Adv}_{P_2}[\mathcal{A}(H_2) = \mathbf{z}].$$

From P_2 to P_3: By the probability preservation property of the Rényi divergence (Lemma 2.6), we have

$$\mathrm{Adv}_{P_2}[\mathcal{A}(H_2) = \mathbf{z}]^2 \leq \mathrm{Adv}_{P_3}[\mathcal{A}(H_3) = \mathbf{z}] \cdot \mathrm{RD}_2(H_2 \| H_3).$$

In order to compute the Rényi divergence between H_2 and H_3, we need to compute the Rényi divergence between $\mathbf{N}(\lambda\mathbf{z}) + \mathbf{e}$ and $\mathbf{e}'$. By definition of M_{σ_H}, it holds that $\|\sigma_H(\mathbf{N}\tilde{\mathbf{z}})\|_2 = \|M_{\sigma_H}(\mathbf{N})\sigma_H(\tilde{\mathbf{z}})\|_2 \leq \|M_{\sigma_H}(\mathbf{N})\|_2 \|\sigma_H(\tilde{\mathbf{z}})\|_2$. Since σ and σ_H only differ by the unitary transformation $\mathbf{U}_H$, we have that $\|\sigma_H(\tilde{\mathbf{z}})\|_2 = \|\sigma(\tilde{\mathbf{z}})\|_2 \leq \|\mathbf{V}\|_2 \|\tau(\tilde{\mathbf{z}})\|_2 \leq \|\mathbf{V}\|_2 \cdot (\eta-1)\sqrt{nd}$, as $\tilde{\mathbf{z}} \in R_\eta^d$. Finally, Lemma 2.12 for $\mathcal{I} = R^\vee$ gives the spectral bound $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha q \log_2 n\sqrt{md}$ with overwhelming probability.

This yields that each of the m coefficients of the vector $\mathbf{N}\tilde{\mathbf{z}}$ is bounded above by $\alpha q d \|\mathbf{V}\|_2 \sqrt{n} \log_2(n)(\eta-1)$ with probability $1 - 2^{-\Omega(n)}$ (when embedded with σ_H). Thus, it suffices to compute the Rényi divergence of $\mathcal{D}^m_{R^\vee,\beta q,c}$

and $\mathcal{D}^m_{R^\vee,\beta q}$, where $c \in R^\vee$ satisfies $\|\sigma_H(c)\|_2 \leq \alpha q d \|\mathbf{V}\|_2 \sqrt{n}\log_2(n)(\eta-1)$. Using that $\beta q \geq \eta_\varepsilon(R^\vee)$, the multiplicativity of the Rényi divergence (Lemma 2.6) and the Rényi divergence of shifted discrete Gaussians (Lemma 2.11), we deduce

$$\begin{aligned}\mathrm{RD}_2\left(\mathcal{D}^m_{R^\vee,\beta q,c} \| \mathcal{D}^m_{R^\vee,\beta q}\right) &= \mathrm{RD}_2\left(D_{R^\vee,\beta q,c} \| D_{R^\vee,\beta q}\right)^m \\ &\leq \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{2m} \cdot \exp\left(\frac{2\pi\|\sigma_H(c)\|_2^2}{(\beta q)^2}\right)^m.\end{aligned}$$

The way we chose β with respect to α yields with overwhelming probability that $\exp(2\pi\|\sigma_H(c)\|_2^2/(\beta q)^2)^m \leq \exp(2\pi)$. For the Rényi divergence to be bounded by a constant, we also need $\varepsilon = O(\frac{1}{m})$. Indeed, we have $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 = \left(1+\frac{4\varepsilon/1-\varepsilon}{2}\right)^2 < \exp\left(\frac{4\varepsilon}{1-\varepsilon}\right)$ as $\left(1+\frac{x}{y}\right)^y < \exp(x)$ for any $x, y > 0$. Without loss of generality, assume $\varepsilon < \frac{1}{2}$, then $\frac{1}{1-\varepsilon} < 2$ and thus, we get $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{2m} < \exp(8m\varepsilon)$ and therefore $\varepsilon = O(\frac{1}{m})$ suffices.

From P_3 to P_4: By the probability preservation property of the Rényi divergence (Lemma 2.6), we have

$$\mathrm{Adv}_{P_3}[\mathcal{A}(H_3) = \mathbf{z}]^2 \leq \mathrm{Adv}_{P_4}[\mathcal{A}(H_4) = \mathbf{s}] \cdot \mathrm{RD}_2(H_3 \| H_4).$$

The only difference between the distributions H_3 and H_4 is that the element $\lambda\mathbf{C}\mathbf{z}$ in H_3 is replaced by $\mathbf{s}$ in H_4. Our aim is to show that their Rényi divergence can be bounded by a constant. Recall that $\tilde{\mathbf{C}} = \lambda\mathbf{C} \in (R_q)^{k\times d}$. By the leftover hash lemma stated in Lemma 2.7, the Rényi divergence between the distribution $(\tilde{\mathbf{C}}, \tilde{\mathbf{C}}\tilde{\mathbf{z}})$ and the distribution $(\tilde{\mathbf{C}}, \tilde{\mathbf{s}})$ is bounded above by $(1 + q^k/\eta^d)^n$. Dividing the first and the second part of both distributions by λ preserves the Rényi divergence. As we require $d\log_2\eta \geq k\log_2 q + \Omega(\log_2 n)$, we obtain $\mathrm{RD}_2(H_3\|H_4) \leq (1 + 1/\Omega(n))^n = O(1)$ asymptotically in n.

Problem P_4: This problem is exactly the $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R^\vee,\beta q}}$ problem, as $\mathbf{C}$ and $\mathbf{N}$ are independent of $\mathbf{B}, \mathbf{s}$ and $\mathbf{e}'$. Therefore, if $\mathrm{Adv}_{\text{M-SLWE}}$ denotes the advantage of an adversary against $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R^\vee,\beta q}}$, it hold that

$$\mathrm{Adv}_{P_4}[\mathcal{A}(H_4) = \mathbf{s}] = \mathrm{Adv}_{\text{M-SLWE}}$$

Putting all equations from above together, we obtain

$$\begin{aligned}\mathrm{Adv}_{P_0}[\mathcal{A}(H_0) = \mathbf{z}] &\leq \mathrm{Adv}_{P_1}[\mathcal{A}(H_1) = \mathbf{z}] + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\ &\leq \mathrm{Adv}_{P_2}[\mathcal{A}(H_2) = \mathbf{z}] + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\ &\leq \sqrt{\mathrm{Adv}_{P_3}[\mathcal{A}(H_3) = \mathbf{z}] \cdot \mathrm{RD}_2(H_2\|H_3)} + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\ &\leq \sqrt{\sqrt{\mathrm{Adv}_{\text{M-SLWE}} \cdot \mathrm{RD}_2(H_3\|H_4)} \cdot \mathrm{RD}_2(H_2\|H_3)} \\ &\quad + d \cdot \mathrm{Adv}_{\text{M-LWE}}.\end{aligned}$$

The choice of parameters yields $\mathrm{RD}_2(H_2\|H_3), \mathrm{RD}_2(H_3\|H_4) = O(1)$, and our base assumptions give $\mathrm{Adv}_{\text{M-LWE}}, \mathrm{Adv}_{\text{M-SLWE}} \leq n^{-\omega(1)}$. It therefore proves that $\mathrm{Adv}_{P_0}[\mathcal{A}(H_0) = \mathbf{z}] \leq n^{-\omega(1)}$. □

3.2 Pseudorandomness of η-M-LWE

We now provide a more involved proof of hardness for the *decision* version of η-M-LWE. It follows the same idea as in [BLP+13] that we extend to modules. More precisely, we show a reduction from M-LWE with rank k to η-M-LWE with rank d satisfying $d\log_2\eta \geq (k+1)\log_2 q + \omega(\log_2 n)$. 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 $n(\eta-1)\sqrt{2d}\sqrt{4n^2(\eta-1)^2+1}$. In the case of general cyclotomic fields, the noise rate slightly improves on the noise rate of $d\sqrt{m}\cdot n^{3/2}\log_2(n)(\eta-1)$ from Section 3.1. We indeed improve the noise rate by a factor of roughly $\sqrt{8n}(\eta-1)/\log_2(n)\sqrt{md}$, which is advantageous whenever $m > 8n(\eta-1)^2/d\log_2^2 n$. As we wish to take η as a small constant, the condition can be met when m is sub-linear. However, in the special case of power-of-two cyclotomics, the noise rate from Section 3.1 is improved by $\sqrt{n}$. This means that this new reduction is advantageous (in terms of noise) only if $m > 8n^2(\eta-1)/d\log_2^2 n = \Theta(n^2/\log_2^3 n)$, which is now just sub-quadratic. Nonetheless, this reduction allows for proving the hardness of the decision version of η-M-LWE which is preferrable in a lot of situations. 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 3.4. The reduction in Theorem 3.2 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 3.2.

Theorem 3.2. *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 $q > \max(2n, ((\eta-1)\mathfrak{s}_1(\mu))^{\varphi(\mu)})$, where $\mathfrak{s}_1(\mu)$ denotes the largest singular value of the Vandermonde matrix of the μ-th cyclotomic field, and η a positive integer. Further, let k, d, m be positive integers such that $d\log_2\eta \geq (k+1)\log_2 q + \omega(\log_2 n)$, and $d \leq m \leq \mathsf{poly}(n)$. Let $\alpha \geq q^{-1}\sqrt{\ln(2nd(1+1/\varepsilon))/\pi}$ and $\beta \geq \alpha\cdot n(\eta-1)\sqrt{2d}\sqrt{4n^2(\eta-1)^2+1}$. Then there is a reduction from $\text{M-LWE}_{n,k,m,q,D_\alpha}$ to $\eta\text{-M-LWE}_{n,d,m,q,\Psi_{\leq\beta}}$, such that if $\mathcal{A}$ solves the latter with advantage $\mathrm{Adv}[\mathcal{A}]$, 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 $n(\eta-1)$ encapsulates the norm distortion between the coefficient and the canonical embedding, as well as the actual length of the η-bounded vectors. The second term $\sqrt{2d}$ stems from the masking of $\mathbf{z}$ when introduced in the first hybrid in the proof of Lemma 3.5. The last factor $\sqrt{4n^2(\eta-1)^2+1}$ solely represents the impact of giving information on the error in the ext-M-LWE problem. We give here an overview of the full reduction in Figure 3.2.

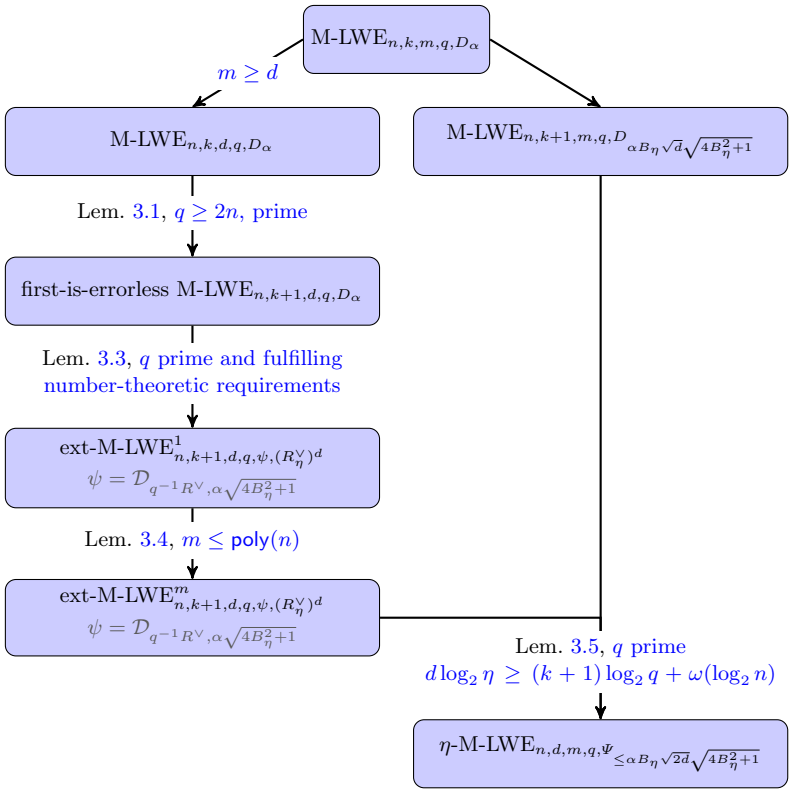

Fig. 3.2. Summary of the proof of Theorem 3.2, where $B_\eta = \max_{x \in R_\eta} \|\sigma(x)\|_\infty$ and σ is the canonical embedding. In cyclotomic fields, we have $B_\eta \leq n(\eta - 1)$. Note that Lemma 3.5 uses d samples from ext-M-LWE, where d is the module rank in η-M-LWE. The assumptions on q concern the splitting behavior of the cyclotomic polynomial in $\mathbb{Z}_q[x]$, and are discussed in Section 3.2.2.

3.2.1 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. In our case, this step is simplified and extended to general number fields, provided that the modulus q is unramified and larger than $2n$. In the ν-th cyclotomic field, this boils down to $q \nmid \nu$ and $q \geq 2n$. Further restrictions on q in our reduction encompasses these conditions.

Definition 3.1 (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, and Υ a distribution over a family of distributions over $K_\mathbb{R}$. The* first-is-errorless M-LWE$_{n,k,q,\Upsilon}$

problem is to distinguish between the following cases. On the one hand, the first sample is from $U(R_q^k \times q^{-1}R^\vee/R^\vee)$ *and the rest from* $U(R_q^k \times \mathbb{T}_{R^\vee})$*. On the other hand, there is some unknown* $\mathbf{s} \hookleftarrow U((R_q^\vee)^k)$ *and* $\psi \hookleftarrow \Upsilon$ *such that the first sample is from* $A^{\mathcal{M}}_{\mathbf{s},\{0\}}$ *and the rest are distributed as* $A^{\mathcal{M}}_{\mathbf{s},\psi}$*, where* $\{0\}$ *is the distribution that is deterministically* 0*. When the number of samples* m *is fixed, we denote it* first-is-errorless M-LWE$_{n,k,m,q,\Upsilon}$*.*

Lemma 3.1 (Adapted from [BLP+13, Lem. 4.3]). *Let* K *be a number field of degree* n*, and* R *its ring of integers. Let* $q \geq 2n$ *be an unramified prime integer,* 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. *Pre-processing:* The reduction first chooses $\mathbf{a}' \hookleftarrow U(R_q^k)$ and then independently samples $\mathbf{b}_2, \ldots, \mathbf{b}_k$ 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 2.5. 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. Note that after drawing m columns, the probability that we successfully constructed this matrix is exactly $1 - \delta(m, k)$, where $\delta(\cdot, \cdot)$ is defined in Equation 1[2].
Reduction: 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 sam-

[2] Later in the reduction, we restrict the modulus q to be a prime that splits into two prime factors of inertia degree $n/2$ in the underlying cyclotomic field. In this case, at each draw, the probability that the new column is R_q-linearly independent of the previous ones is $1 - (2q^{-n/2} - q^{-n})^k$ which is much closer to 1 than $1 - n/q$. We discuss further the $\delta(.,.)$ function in Section 4.3.

ples, 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 $\mathbf{s}' = \mathbf{U}^{-T}(s_0|\mathbf{s})$. □

3.2.2 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. A first definition of ext-M-LWE was introduced by Alperin-Sheriff and Apon [AA16] in which the hints were of the form $Tr(\langle \mathbf{z}_i, \mathbf{e} \rangle)$ for a single error vector $\mathbf{e}$ and several *hint vectors* $\mathbf{z}_i$. In our case, we allow for multiple secrets (and thus errors) and one single hint vector $\mathbf{z}$, as required by our final reduction of Lemma 3.5. Additionally, as the field trace does not provide enough information to reconstruct $\langle \mathbf{z}, \mathbf{e} \rangle$ from the hint, we instead directly give $\langle \mathbf{z}, \mathbf{e} \rangle$ as the hint. We prove that it does not make the problem easier. Another version of ext-M-LWE was recently introduced in [LNS21] in the context of lattice-based zero-knowledge proofs, where they only provide the sign $\mathrm{Sign}(\langle \mathbf{z}, \mathbf{e} \rangle)$ as an additional hint for the attacker. Again, this is not sufficient for our lossy argument in Lemma 3.5.

Definition 3.2 (Extended M-LWE). *Let K be a number field of degree n, and R its ring of integers. Let m, q, k, ℓ be positive integers. Let $\mathcal{Z} \subseteq (R^\vee)^m$ and ψ a discrete distribution over $q^{-1}R^\vee$. The* Extended M-LWE *problem, denoted by* ext-M-LWE$^{\ell}_{n,k,m,q,\psi,\mathcal{Z}}$*, is as follows. The algorithm first samples $\mathbf{z} \in \mathcal{Z}$ and then receives a tuple $(\mathbf{A}, \mathbf{B}, \mathbf{E}^T\mathbf{z})$ over $R_q^{m\times k} \times (q^{-1}R^\vee/R^\vee)^{m\times \ell} \times (q^{-1}R^\vee)^\ell$. Its goal is to distinguish between the following cases.*
On one side, $\mathbf{A}$ is sampled from $U(R_q^{m\times k})$, $\mathbf{E}$ is sampled from $\psi^{m\times\ell}$, and define $\mathbf{B} = q^{-1}\mathbf{AS} + \mathbf{E} \bmod R^\vee$ for some uniformly chosen $\mathbf{S} \in (R_q^\vee)^{k\times\ell}$. On the other side, all is identical except that $\mathbf{B}$ is sampled from $U((q^{-1}R^\vee/R^\vee)^{m\times\ell})$, independently from $\mathbf{A}$ and $\mathbf{E}$.

The parameter ℓ represents the number of given hints on independent noise vectors, and therefore the number of secret vectors (which generalizes the multiple secret version of M-LWE). The set $\mathcal{Z}$ represents the set of hints that can be given on the noise vectors. The ℓ hints are given in form of the inner product of such a fixed *hint vector* $\mathbf{z} \in \mathcal{Z}$ and the corresponding column of $\mathbf{E}$. Later, we are interested in the case where $\mathcal{Z} = (R_\eta^\vee)^m$ which is actually the set of secrets for η-M-LWE. Also, note that if $\mathcal{Z} = \{\mathbf{0}\}$, then we recover the definition of the multiple secret version of M-LWE from Section 2.5.

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 3.3 requires the construction of a matrix $\mathbf{U_z} \in R^{m\times m}$, for all vectors $\mathbf{z} \in \mathcal{Z} = (R_\eta^\vee)^m$, 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 spectral norm of its submatrix $\mathbf{U}_\mathbf{z}^\perp$ (when embedded with M_σ), controls the increase in the Gaussian parameter. We propose a construction for which we bound the spectral norm above by a quantity independent on $\mathbf{z}$, as needed in the reduction.

Lemma 3.2. *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 $\mu = \prod_i p_i$, η a positive integer and q be a prime such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > ((\eta-1)\mathfrak{s}_1(\mu))^{\varphi(\mu)}$, where $\mathfrak{s}_1(\mu)$ denotes the spectral norm of the Vandermonde matrix of the μ-th cyclotomic field. Finally, let m be a positive integer, and $\mathcal{Z} = (R_\eta^\vee)^m$. For all $\mathbf{z} \in \mathcal{Z}$, there is an efficiently computable matrix $\mathbf{U_z} \in R^{m\times m}$ that is invertible modulo qR and that verifies the following: $\mathbf{z}$ is orthogonal to the columns of $\mathbf{U}_\mathbf{z}^\perp$, and $\left\|M_\sigma(\mathbf{U}_\mathbf{z}^\perp)\right\| \leq 2B_\eta$, where $B_\eta = \max_{x\in R_\eta}\|\sigma(x)\|_\infty$.*

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 $\widetilde{\mathbf{z}} = \lambda\mathbf{z} \in R_\eta^m$. First, we construct $\mathbf{U_z}$ 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:

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

$\mathbf{U}_\mathbf{z}^\perp$ $\mathbf{A}^\perp$ $\mathbf{B}^\perp$

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 and η-bounded elements, we have that for all i in $[m]$, $0 < \|\tau(\widetilde{z}_i)\|_\infty \leq \eta - 1$, where τ is the coefficient embedding. By Lemma 2.4, since q verifies the algebraic conditions taking all $f_i = 1$ and $q^{1/\varphi(\mu)}/\mathfrak{s}_1(\mu) > \eta - 1$, 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 $\mathbf{U}_\mathbf{z}^\perp$ be the submatrix of $\mathbf{U_z}$ obtained by removing the leftmost column as shown above. Since M_σ is a ring homomorphism, we have $M_\sigma(\mathbf{U}_\mathbf{z}^\perp) = M_\sigma(\mathbf{A}^\perp) + M_\sigma(\mathbf{B}^\perp)$. 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 $\left\|M_\sigma(\mathbf{A}^\perp)\mathbf{x}\right\|_2 = \sqrt{\sum_{i\in[m-1]}\sum_{j\in[n]}|\sigma_j(\widetilde{z}_i)|^2|x_{j+n(i-1)}|^2} \leq B_\eta\|\mathbf{x}\|_2$, because each $\widetilde{z}_i$ is in R_η. This yields $\left\|M_\sigma(\mathbf{A}^\perp)\right\|_2 \leq B_\eta$. A similar calculation on $\mathbf{B}^\perp$ leads to $\left\|M_\sigma(\mathbf{B}^\perp)\right\|_2 \leq B_\eta$, thus resulting in $\left\|M_\sigma(\mathbf{U}_\mathbf{z}^\perp)\right\|_2 \leq 2B_\eta$.

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 $\left\|M_\sigma(\mathbf{U}_\mathbf{z}^\perp)\right\|_2$ can still be bounded above by $2B_\eta$. □

Notice that when the ring is of degree 1 and $\eta = 2$, the constructions in the different cases match the ones from [BLP+13, Claim 4.6]. So do the singular values as $B_\eta \leq n(\eta-1) = 1$ by Lemma 2.1. Also, the construction differs from the notion of quality in [AA16] due to the discrepancies between the two definitions of ext-M-LWE. The following lemma shows that the extended variant of M-LWE with one hint ($\ell = 1$) is at least as hard as the first-is-errorless variant of M-LWE, for carefully chosen parameters.

Lemma 3.3 (Adapted from [BLP+13, Lem. 4.7]). *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 $\mu = \prod_i p_i$, η a positive integer and q be a prime such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > ((\eta-1)\mathfrak{s}_1(\mu))^{\varphi(\mu)}$, where $\mathfrak{s}_1(\mu)$ denotes the spectral norm of the Vandermonde matrix of the μ-th cyclotomic field. Let m, k be positive integers, $\mathcal{Z} = (R_\eta^\vee)^m$, $\varepsilon \in (0, 1/2)$ and $\alpha \geq q^{-1}\sqrt{\ln(2mn(1+1/\varepsilon))/\pi}$. There is a reduction from* first-is-errorless M-LWE$_{n,k,m,q,D_\alpha}$ *to* ext-M-LWE$^1_{n,k,m,q,\psi,\mathcal{Z}}$ *that reduces the advantage by at most $33\varepsilon/2$, where $\psi = \mathcal{D}_{q^{-1}R^\vee,\alpha\sqrt{4B_\eta^2+1}}$ and $B_\eta = \max_{x\in R_\eta}\|\sigma(x)\|_\infty$.*

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

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

and thus Lemma 2.8 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$_{n,k,m,q,\alpha\sqrt{4B_\eta^2+1},\mathcal{Z}}$. 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 3.2 we can efficiently compute a matrix $\mathbf{U}_\mathbf{z} \in R^{m\times m}$ that is invertible modulo qR, such that its submatrix $\mathbf{U}_\mathbf{z}^\perp$ is orthogonal to $\mathbf{z}$, and that $\left\|M_\sigma(\mathbf{U}_\mathbf{z}^\perp)\right\|_2 \leq 2B_\eta$. The reduction first samples $\mathbf{f} \in K_\mathbb{R}^m$ from the continuous Gaussian distribution of covariance matrix $\alpha^2(4B_\eta^2\mathbf{I}_{mn} - M_{\sigma_H}(\mathbf{U}_\mathbf{z}^\perp)M_{\sigma_H}(\mathbf{U}_\mathbf{z}^\perp)^T) \in \mathbb{R}^{mn\times mn}$. The covariance matrix is well-defined because $\left\|M_{\sigma_H}(\mathbf{U}_\mathbf{z}^\perp)\right\|_2 = \left\|M_\sigma(\mathbf{U}_\mathbf{z}^\perp)\right\|_2 \leq 2B_\eta$. The reduction then computes $\mathbf{b}' = \mathbf{U}_\mathbf{z}\mathbf{b} + \mathbf{f}$ and samples $\mathbf{c}$ from $\mathcal{D}_{q^{-1}(R^\vee)^m-\mathbf{b}',\alpha}$, and finally gives the following to $\mathcal{O}$

$$(\mathbf{A}' = \mathbf{U}_\mathbf{z}\mathbf{A}, \mathbf{b}' + \mathbf{c} \bmod R^\vee, \langle \mathbf{z}, \mathbf{f}+\mathbf{c}\rangle).$$

Note that this tuple is in $(R_q)^{m\times k} \times (q^{-1}R^\vee/R^\vee)^m \times q^{-1}R^\vee$, as required. We now prove correctness. First, consider the case where $\mathbf{A}$ is uniformly random

over $R_q^{m\times k}$ and $\mathbf{b} = q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod R^\vee$ 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, $\mathbf{A}' = \mathbf{U_z}\mathbf{A}$ is also uniform over $(R_q)^{m\times k}$ as required. From now on we condition on an arbitrary $\mathbf{A}'$ and analyze the distribution of the remaining components. We have

$$\begin{aligned}\mathbf{b}' &= q^{-1}\mathbf{U_z}\mathbf{A}\mathbf{s} + \mathbf{U_z}\mathbf{e} + \mathbf{f}\\ &= q^{-1}\mathbf{A}'\mathbf{s} + \mathbf{U_z}\mathbf{e} + \mathbf{f}.\end{aligned}$$

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 continuous Gaussian over $K_\mathbb{R}^m$ of covariance matrix $\alpha^2 M_{\sigma_H}(\mathbf{U_z^\perp}) M_{\sigma_H}(\mathbf{U_z^\perp})^T$ by Lemma 2.13. Hence the vector $\mathbf{U_z}\mathbf{e} + \mathbf{f}$ is distributed as the Gaussian over $K_\mathbb{R}^m$ of covariance matrix $\alpha^2 M_{\sigma_H}(\mathbf{U_z^\perp}) M_{\sigma_H}(\mathbf{U_z^\perp})^T + \alpha^2(4B_\eta^2 \mathbf{I}_{mn} - M_{\sigma_H}(\mathbf{U_z^\perp}) M_{\sigma_H}(\mathbf{U_z^\perp})^T)$ which is identical to $D^m_{\alpha\cdot 2B_\eta}$. Since $q^{-1}\mathbf{A}'\mathbf{s} \in q^{-1}(R^\vee)^m$, 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 2.10, the distribution of $\mathbf{U_z}\mathbf{e} + \mathbf{f} + \mathbf{c}$ is within statistical distance 8ε of $\mathcal{D}_{q^{-1}(R^\vee)^m, \alpha\sqrt{4B_\eta^2+1}}$, 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 $\langle \mathbf{z}, \mathbf{U_z}\mathbf{e} + \mathbf{f} + \mathbf{c}\rangle$ 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 2.9, 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 $\mathbf{b}' = \mathbf{U_z}\mathbf{e}' + \mathbf{U_z}\mathbf{e} + \mathbf{f}$ and $\mathbf{U_z}\mathbf{e}' \in q^{-1}(R^\vee)^m$, then the coset $q^{-1}(R^\vee)^m - \mathbf{b}'$ is identical to $q^{-1}(R^\vee)^m - (\mathbf{U_z}\mathbf{e} + \mathbf{f})$. For the same reasons as above, $\mathbf{U_z}\mathbf{e} + \mathbf{f} + \mathbf{c}$ is distributed as $\mathcal{D}_{q^{-1}(R^\vee)^m, \alpha\sqrt{4B_\eta^2+1}}$ within statistical distance of at most 8ε, and in particular independent of $\mathbf{e}'$. So the third component is correctly distributed because once again $\langle \mathbf{z}, \mathbf{U_z}\mathbf{e}\rangle = 0$. Finally, since $\mathbf{e}'$ is independent of the first and third components, and that $\mathbf{U_z}\mathbf{e}'$ 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. □

Remark 3.1 (Instantiation in power-of-two cyclotomics). The condition on the modulus q in Lemma 3.2 and 3.3 stems from the invertibility result by Lyubashevsky and Seiler [LS18] stated in Lemma 2.4. 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 $q^{1/2}/\sqrt{2} > \eta - 1$ is sufficient as our elements have η-bounded coefficients. For the binary secret case $\eta = 2$, this leads to $q > 2$, which is subsumed by $q = 5 \bmod 8$.

We now use a standard hybrid argument to show that ext-M-LWE with ℓ hints is at least as hard as ext-M-LWE with one hint, at the expense of reducing the advantage by a factor of ℓ. The proof can be found in Appendix A.2 for completeness.

Lemma 3.4 (Adapted from [BLP+13, Lem. 4.8]). *Let K be a number field of degree n, R its ring of integers, and k, m, q, ℓ be positive integers such that $\ell \leq poly(n)$. Let ψ be a discrete distribution over $q^{-1}R^\vee$, and $\mathcal{Z} \subseteq (R^\vee)^m$. There is a reduction from* ext-M-LWE$^1_{n,k,m,q,\psi,\mathcal{Z}}$ *to* ext-M-LWE$^\ell_{n,k,m,q,\psi,\mathcal{Z}}$ *that reduces the advantage by a factor of ℓ.*

3.2.3 Reduction to η-M-LWE.

We now provide the final step of the overall reduction, by reducing to the M-LWE problem with η-bounded secret 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.

To facilitate understanding, we start by illustrating the high level idea of the proof of Lemma 3.5 in Figure 3.3. Given an instance of η-M-LWE by $(\mathbf{A}, \mathbf{Az}+\mathbf{e})$, our goal is to show that it is computationally indistinguishable from $(\mathbf{A}, \mathbf{b})$, where $\mathbf{b}$ is a uniformly random vector. To do so, we first decompose the error vector $\mathbf{e}$ into $-\mathbf{Nz} + \mathbf{e}'$, by using properties of Gaussian distributions. We then make use of a similar lossy argument as for the previous reduction of Section 3.1 by replacing the random matrix $\mathbf{A}$ by a lossy matrix $\mathbf{A}' = \mathbf{BC}+\mathbf{N}$. As opposed to the proof from Section 3.1, we can't simply argue with the hardness of multiple-secrets M-LWE as the second part of the sample depends on the noise matrix $\mathbf{N}$. This is the motivation for introducing the ext-M-LWE problem, where we allow for additional information with respect to the noise. We then use the same leftover hash lemma as before to replace the product $\mathbf{Cz}$ by a uniformly random vector $\mathbf{s}$. Assuming the hardness of M-LWE, the term $\mathbf{Bs}+\mathbf{e}'$ is computationally indistinguishable form a uniform vector $\mathbf{u}$. We conclude the proof by re-replacing the lossy matrix $\mathbf{A}'$ by the original uniform matrix $\mathbf{A}$.

Lemma 3.5 (Adapted from [BLP+13, Lem. 4.9]). *Let $K = \mathbb{Q}(\zeta)$ be a number field of degree n, such that its ring of integers is $R = \mathbb{Z}[\zeta]$. Let k, m, d, η and q be positive integers with q prime and $d \log_2 \eta \geq k \log_2 q + \omega(\log_2 n)$. Let $\varepsilon, \alpha, \gamma, \beta, \delta$ be reals such that $\varepsilon \in (0, 1/2)$, $\alpha \geq q^{-1}\sqrt{2\ln(2nd(1+1/\varepsilon))/\pi}$, $\gamma = \alpha B_\eta \sqrt{d}$, $\beta = \alpha B_\eta \sqrt{2d}$, where $B_\eta = \max_{x \in R_\eta} \|\sigma(x)\|_\infty$, and $\delta = \frac{1}{2}\sqrt{(1+q^k/\eta^d)^n - 1}$. There is a polynomial-time reduction from* ext-M-LWE$^m_{n,k,d,q,\psi,(R^\vee_\eta)^d}$, M-LWE$_{n,k,m,q,D_\gamma}$ *and* ext-M-LWE$^m_{n,k,d,q,\psi,\{0\}^d}$ *with $\psi = \mathcal{D}_{q^{-1}R^\vee,\alpha}$ to η-*M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$, *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*

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

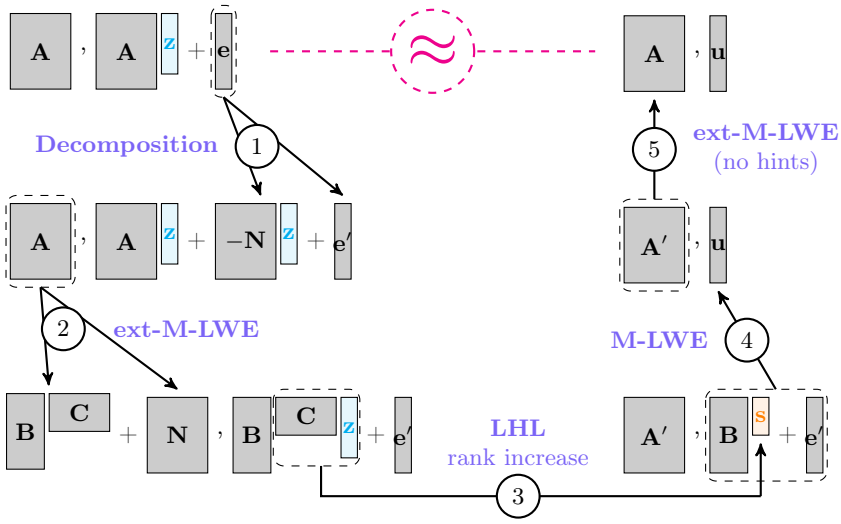

Fig. 3.3. Summary of the proof of Lemma 3.5

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$^m_{n,k,d,q,\alpha,(R^\vee_\eta)^d}$, that is also why it is not specified in Figure 3.2.

Proof. Given an η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$ sample $(\mathbf{A}, q^{-1}\mathbf{Az}+\mathbf{e} \bmod R^\vee)$, with $\mathbf{A} \hookleftarrow U(R_q^{m\times d})$, $\mathbf{z} \hookleftarrow U((R_\eta^\vee)^d)$ and $\mathbf{e} \in K_{\mathbb{R}}^m$ sampled from the continuous Gaussian $D^m_{\mathbf{r}}$ with parameter vector $\mathbf{r}$ with $r_j^2 = \gamma^2+\alpha^2\sum_i|\sigma_j(\widetilde{z}_i)|^2$. We have $\|\mathbf{r}\|_\infty = \sqrt{\gamma^2+\alpha^2\|\widetilde{\mathbf{z}}\|^2_{2,\infty}}$, as well as $\|\widetilde{\mathbf{z}}\|^2_{2,\infty} \leq \sum_{i\in[d]}\|\sigma(\widetilde{z}_i)\|^2_\infty$. Recalling the parameter $B_\eta = \max_{x\in R_\eta}\|\sigma(x)\|_\infty$, that can be bounded above by $n(\eta-1)$ for cyclotomics by Lemma 2.1, we get $\|\mathbf{r}\|_\infty \leq \sqrt{\gamma^2+B_\eta^2 d\alpha^2} = B_\eta\sqrt{2d}\alpha = \beta$. The objective is to show that $(\mathbf{A}, q^{-1}\mathbf{Az}+\mathbf{e} \bmod R^\vee)$ is computationally indistinguishable from uniform. To do so, we define different hybrid distributions as follows, and prove that each one is indistinguishable from the next.

- H_0: $(\mathbf{A}, q^{-1}\mathbf{Az}+\mathbf{e} \bmod R^\vee)$ as in η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$;
- H_1: $(\mathbf{A}, q^{-1}\mathbf{Az}-\lambda\mathbf{Nz}+\mathbf{e}' \bmod R^\vee)$, where $\mathbf{N} \hookleftarrow \mathcal{D}^{m\times d}_{q^{-1}R^\vee,\alpha}$ and $\mathbf{e}' \hookleftarrow D^m_\gamma$;
- H_2: $(\mathbf{A}', q^{-1}\mathbf{A}'\mathbf{z}-\lambda\mathbf{Nz}+\mathbf{e}' \bmod R^\vee) = (\mathbf{A}, q^{-1}(\lambda\mathbf{B})\mathbf{Cz}+\mathbf{e}' \bmod R^\vee)$, with $\mathbf{A}' = \lambda q(q^{-1}\mathbf{C}^T\mathbf{B}^T+\mathbf{N}^T \bmod R^\vee)^T$, where $\mathbf{B}$ and $\mathbf{C}$ are sampled from $U((R_q^\vee)^{m\times k})$ and $U(R_q^{k\times d})$ respectively;
- H_3: $(\mathbf{A}', q^{-1}\widetilde{\mathbf{B}}\mathbf{s}+\mathbf{e}' \bmod R^\vee)$, where $\widetilde{\mathbf{B}} = \lambda\mathbf{B} \in R_q^{m\times k}$ and $\mathbf{s} \hookleftarrow U((R_q^\vee)^k)$;
- H_4: $(\mathbf{A}', \mathbf{u})$, where $\mathbf{u} \hookleftarrow U(\mathbb{T}^m_{R^\vee})$;
- H_5: $(\mathbf{A}, \mathbf{u}) \hookleftarrow U(R_q^{m\times d} \times \mathbb{T}^m_{R^\vee})$.

From H_0 to H_1: We first claim that $\Delta([-\mathbf{N}\widetilde{\mathbf{z}}+\mathbf{e}']_i, \mathbf{e}_i) \leq 2\varepsilon$ for all $i \in [m]$. Indeed, $(1/\alpha^2+\|\widetilde{\mathbf{z}}\|^2_{2,\infty}/\gamma^2)^{-1/2} \geq \alpha/\sqrt{2}$ and $\alpha/\sqrt{2} \geq \eta_\varepsilon(q^{-1}(R^\vee)^d)$ as explained

for Lemma 3.3. If $\mathbf{n}_i \in q^{-1}(R^\vee)^d$ denotes the i-th row of $\mathbf{N}$, Lemma 2.14 yields the claim as $[-\mathbf{N}\widetilde{\mathbf{z}} + \mathbf{e}']_i = \langle \mathbf{n}_i, -\widetilde{\mathbf{z}} \rangle + e'_i$, thus giving $\Delta(-\mathbf{N}\widetilde{\mathbf{z}} + \mathbf{e}', \mathbf{e}) \leq 2m\varepsilon$.

$$|\mathbb{P}[\mathcal{A}(H_0) = 1] - \mathbb{P}[\mathcal{A}(H_1) = 1]| \leq 2m\varepsilon. \tag{2}$$

From H_1 to H_2: 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$^m_{n,k,d,q,\alpha,(R^\vee_\eta)^d}$ 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, $(\mathbf{C}, (\lambda q)^{-1}\mathbf{A}^T, \mathbf{N}\mathbf{z})$ can be efficiently transformed into a sample from H_1. Note that $(\lambda q)^{-1}\mathbf{A}^T$ 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 $q^{-1}R^\vee/R^\vee$. Additionally, the transpose operator comes from the fact that the hints are $\mathbf{N}\mathbf{z}$, which corresponds to m error vectors of size d. So the second component is indeed of size $d \times m$. In the other case, if we apply the same transformation to the ext-M-LWE sample $(\mathbf{C}^T, q^{-1}\mathbf{C}^T\mathbf{B}^T + \mathbf{N}^T \bmod R^\vee, \mathbf{N}\mathbf{z})$ where $\mathbf{B}^T$ and $\mathbf{N}^T$ are the secret and error matrix respectively, it leads to a sample from H_2. Hence, $\mathcal{B}_1$ is a distinguisher for ext-M-LWE$^m_{n,k,d,q,\alpha,(R^\vee_\eta)^d}$, and

$$|\mathbb{P}[\mathcal{A}(H_1) = 1] - \mathbb{P}[\mathcal{A}(H_2) = 1]| = \mathrm{Adv}[\mathcal{B}_1]. \tag{3}$$

From H_2 to H_3: By the Ring Leftover Hash Lemma stated in Lemma 2.7, 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 $d \log_2 \eta \geq k \log_2 q + \omega(\log_2 n)$ implies $\delta \leq n^{-\omega(1)}$. This yields

$$|\mathbb{P}[\mathcal{A}(H_2) = 1] - \mathbb{P}[\mathcal{A}(H_3) = 1]| \leq \delta. \tag{4}$$

From H_3 to H_4: 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 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 $(\widetilde{\mathbf{B}}, q^{-1}\widetilde{\mathbf{B}}\mathbf{s} + \mathbf{e}' \bmod R^\vee)$ 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

$$|\mathbb{P}[\mathcal{A}(H_3) = 1] - \mathbb{P}[\mathcal{A}(H_4) = 1]| = \mathrm{Adv}[\mathcal{B}_2]. \tag{5}$$

From H_4 to H_5: We now change $\mathbf{A}'$ back to uniform. With the same argument as before, 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 hint) based on a distinguisher between H_4 and H_5. It transforms $(\mathbf{C}^T, (\lambda q)^{-1}(\mathbf{A}')^T, \mathbf{N}\cdot\mathbf{0})$ into a sample from H_4 (M-LWE case) and $(\mathbf{C}^T, (\lambda q)^{-1}\mathbf{A}^T, \mathbf{N}\cdot\mathbf{0})$ into a sample from H_5 (uniform case). We then get

$$|\mathbb{P}[\mathcal{A}(H_4) = 1] - \mathbb{P}[\mathcal{A}(H_5) = 1]| = \mathrm{Adv}[\mathcal{B}_3]. \tag{6}$$

Putting Equations (2), (3), (4), (5), (6) altogether yields the result. □

4 Hardness of M-LWE with Small Error

In this section, we focus on the hardness of M-LWE when the error distribution is uniform over $(R_\eta^\vee)^m$ instead of Gaussian as in the standard formulation of M-LWE. All the results can easily be adapted to errors within other coefficient sets, e.g., with coefficients between $-\eta/2$ and $\eta/2$. The overall proof startegy follows the idea of Micciancio and Peikert [MP13], that we adapt to modules. It uses a different proof method as the one we used in Section 3 as it relies on proving that the M-LWE function is one-way with small uniform inputs (errors). On top of that, we provide a more fine-grained analysis to reach better parameters. The security of practical schemes is indeed driven by the ring degree n as we wish to use a small rank d for efficiency. The asymptotic approach is then not perfectly suited for achieving very small ranks d and very small error bounds η simultaneously. We therefore try to avoid asymptotic results and bounds as much as possible. Even with our approach, we cannot set d and η arbitrarily small independently of each other. We first recall the duality between M-LWE and M-SIS which allows us to switch from one to other at essentially no cost. We then prove our result in terms of M-SIS as it simplifies the analysis. We briefly discuss the practical implications of our work in Section 4.4.

4.1 Duality between M-LWE and M-SIS

It is well established that LWE and SIS are *dual* to each other, mostly from the fact that the q-periodic lattices $\Lambda_q(\mathbf{A}) = \{\mathbf{y} \in \mathbb{Z}^m\colon \mathbf{y} = \mathbf{A}\mathbf{s} \bmod q \text{ for some } \mathbf{s} \in \mathbb{Z}^d\}$ (related to LWE) and $\Lambda_q^\perp(\mathbf{A}) = \{\mathbf{x} \in \mathbb{Z}^m\colon \mathbf{A}^T\mathbf{x} = \mathbf{0} \bmod q\}$ (related to SIS) are dual up to a factor q for any matrix $\mathbf{A} \in \mathbb{Z}^{m\times d}$. This duality carries over to the module setting, which we get by extending the duality results from [MM11, Sec. 4.2] in terms of function families. For completeness, we detail the proofs in Appendix A.3. The idea when going from M-LWE to M-SIS is to cancel the secret part via a *parity check* matrix $\mathbf{B}$ that is such that $\mathbf{A}^T\mathbf{B} = \mathbf{0} \bmod qR$. The M-LWE error distribution $\mathbf{e}$ then becomes the input distribution of the M-SIS instance with matrix $\mathbf{B}' = \mathbf{B}\mathbf{U}$ where $\mathbf{U}$ simply randomizes $\mathbf{B}$. For $\mathbf{B}'$ to be well distributed, we need $\mathbf{A}$ to be non-singular which is characterized by the function $\delta(\cdot,\cdot)$ from Section 2.1. The closed-form expression in Lemma 2.5 and Equation 1 for this singularity probability from [WW19] requires q to be unramified in order to have an easier characterization of units of R_q.

Lemma 4.1 (Adapted from [MM11, Lem. 4.8]). *Let K be a number field of degree n, and R its ring of integers. Let d, q, m be positive integers such that q is an unramified prime, and $m \geq d+1$. Let $\mathcal{X}$ be a probability distribution on $(R^\vee)^m$. If $(\text{M-LWE}(n,d,m,q,(R^\vee)^m),\mathcal{X})$ is ε-uninvertible (resp. one-way, pseudorandom), then $(\text{M-SIS}(n,m-d,m,q,(R^\vee)^m),\mathcal{X})$ is ε'-uninvertible (resp. one-way, pseudorandom), with $\varepsilon' = \delta(m,m-d) + \varepsilon/(1-\delta(m,d))$.*

Lemma 4.2 (Adapted from [MM11, Lem. 4.9]). *Let K be a number field of degree n, and R its ring of integers. Let d, q, m be positive integers such*

that q is an unramified prime, and $m \geq d+1$. Let $\mathcal{X}$ be a probability distribution on $(R^\vee)^m$. If $(\text{M-SIS}(n, m-d, m, q, (R^\vee)^m), \mathcal{X})$ is ε-uninvertible (resp. one-way, pseudorandom), then $(\text{M-LWE}(n, d, m, q, (R^\vee)^m), \mathcal{X})$ is ε'-uninvertible (resp. one-way, pseudorandom), with $\varepsilon' = \delta(m,d) + \varepsilon/(1-\delta(m, m-d))$.

4.2 Hardness of M-LWE with Small Error

We now focus on proving the one-wayness of the M-LWE function family with respect to a short uniform input (i.e., error) distribution, under the assumption that the Mod-GIVP$^{\eta_\varepsilon}_{\gamma}$ problem from Definition 2.1 is hard. It therefore implies the hardness of the search version of M-LWE with small uniform error. To prove the one-wayness of the M-LWE function, we prove the result in terms of M-SIS and use Lemma 4.2 to conclude. Recall that by Lemma 2.15, it suffices to prove that M-SIS is uninvertible and second preimage resistant with respect to this specific input distribution. We actually prove the second preimage resistance of the M-SIS function, and the uninvertibility of a decomposition of the M-SIS function. We then argue that these two function families are indistinguishable based on the pseudorandomness of M-SIS (or M-LWE equivalently). The idea of the proof is summarized in Figure 4.1.

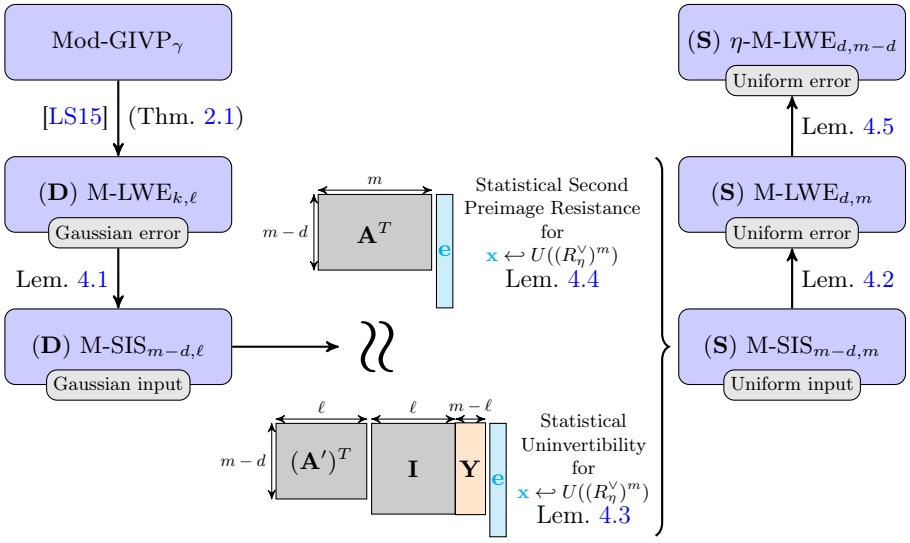

Fig. 4.1. Summary of the proof of Theorem 4.2. **S** denotes the search version, while **D** denotes the decision version. The first subscript for M-LWE denotes the rank, while the second subscript denotes the number of samples. For clarity, we removed the subscripts for the ring degree n and the modulus q as they are preserved throughout the proof. We have $\ell = m - d + k$.

4.2.1 Uninvertibility. In order to prove the uninvertibility of the function family $(\text{M-SIS}(n, m-d, m, q, (R^\vee)^m), U((R_\eta^\vee)^m))$, we decompose it into a linear (Gaussian) function family $\mathcal{L}$ and a smaller $\text{M-SIS}(n, m-d, \ell, q, (R^\vee)^\ell)$ function family with $\ell \leq m$. By Lemma 2.17, it suffices to prove the uninvertibility of $(\mathcal{L}, U((R_\eta^\vee)^m))$. We first define what we mean by linear (Gaussian) function.

Definition 4.1. *Let K be a number field of degree n, and R its ring of integers. Let ℓ, m be positive integers such that $m \geq \ell$, $s > 0$, and $X \subseteq (R^\vee)^m$. We define the function family $\mathcal{L}(\ell, m, s, X)$ obtained by sampling $\mathbf{Y}$ from $\mathcal{D}_{R,s}^{\ell \times (m-\ell)}$, and outputting $h_{\mathbf{Y}} : X \to (R^\vee)^\ell$ defined by $\forall \mathbf{x} \in X$, $h_{\mathbf{Y}}(\mathbf{x}) = [\mathbf{I}_\ell \mid \mathbf{Y}]\mathbf{x}$, where $\mid$ denotes the horizontal concatenation.*

We now use Lemma 2.16 to prove that $(\mathcal{L}(\ell, m, s, X), U(X))$ is statistically uninvertible with uniform inputs for carefully chosen parameters. In particular, the Gaussian width s has to exceed the smoothing parameter of R, and for the result to be meaningful we also need ε_3 to be negligible. This leads to involved conditions on the parameters, which we discuss in Section 4.2.3.

Lemma 4.3. *Let K be a number field of degree n, and R its ring of integers. Let ℓ, m, d be positive integers such that $m \geq \max(d, \ell)$, and $s > \eta_\varepsilon(R)$. Let η be a positive integer and $X \subseteq (R_\eta^\vee)^m$. We define the function family $\mathcal{F} = \text{M-SIS}(n, m-d, \ell, q, (R^\vee)^\ell) \circ \mathcal{L}(\ell, m, s, X)$. Then $(\mathcal{F}, U(X))$ is (statistically) ε_3-uninvertible for*

$$\varepsilon_3 = \frac{1}{|X|\sqrt{\pi n \ell}} \left((\eta - 1)\sqrt{2\pi e} \left(1 + s\sqrt{\frac{m-\ell}{\ell}} \left(C\sqrt{\ell} + C\sqrt{m-\ell} + \omega_n \right) \right) \right)^{n\ell} + 2ne^{-\pi\omega_n^2},$$

where $C > 0$ is an absolute constant, and $\omega_n = \omega(\sqrt{\log_2 n})$.

Proof. We first bound $\mathbb{E}_{h_{\mathbf{Y}} \hookleftarrow \mathcal{L}}[|h_{\mathbf{Y}}(X)|]$ and use Lemma 2.16 to conclude. Let $h_{\mathbf{Y}}$ be sampled from $\mathcal{L}(\ell, m, s, X)$. Let $\mathbf{x} = [\mathbf{x}_1^T \mid \mathbf{x}_2^T]^T \in X$, with $\mathbf{x}_1 \in (R_\eta^\vee)^\ell$, and $\mathbf{x}_2 \in (R_\eta^\vee)^{m-\ell}$. Then, $h_{\mathbf{Y}}(\mathbf{x}) = \mathbf{x}_1 + \mathbf{Y}\mathbf{x}_2$. We use the scaling factor λ as defined in Section 2.1. Thus, we have $\lambda \cdot h_{\mathbf{Y}}(\mathbf{x}) = \tilde{\mathbf{x}}_1 + \mathbf{Y}\tilde{\mathbf{x}}_2$, with $\tilde{\mathbf{x}}_1 \in R_\eta^\ell$ and $\tilde{\mathbf{x}}_2 \in R_\eta^{m-\ell}$. As $\lambda \cdot h_{\mathbf{Y}}(X)$ is isomorphic to $h_{\mathbf{Y}}(X)$, we instead bound the size of $\lambda \cdot h_{\mathbf{Y}}(X)$. As seen in Section 2.1, it holds that $\tau(\lambda h_{\mathbf{Y}}(\mathbf{x})) = \tau(\tilde{\mathbf{x}}_1) + M_\tau(\mathbf{Y})\tau(\tilde{\mathbf{x}}_2)$, and therefore

$$\|\tau(\lambda h_{\mathbf{Y}}(\mathbf{x}))\|_2 \leq \|\tau(\tilde{\mathbf{x}}_1)\|_2 + \|M_\tau(\mathbf{Y})\|_2 \cdot \|\tau(\tilde{\mathbf{x}}_2)\|_2.$$

Since $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{x}}_2$ are vectors over R_η, it holds that $\|\tau(\tilde{\mathbf{x}}_1)\|_2 \leq (\eta - 1)\sqrt{n\ell}$ and $\|\tau(\tilde{\mathbf{x}}_2)\|_2 \leq (\eta - 1)\sqrt{n(m-\ell)}$. By Lemma 2.3, we also have

$$\|M_\tau(\mathbf{Y})\|_2 = \max_{k \in [n]} \|\sigma_k(\mathbf{Y})\|_2.$$

As $\sigma_k(\mathbf{Y})$ is a centered sub-Gaussian matrix with sub-Gaussian moment s, a non-asymptotic spectral bound on sub-Gaussian random matrices due to Vershynin [Ver12] gives that for all $k \in [n]$, it holds

$$\mathbb{P}_{\mathbf{Y}}\left[\|\sigma_k(\mathbf{Y}))\|_2 > Cs(\sqrt{\ell} + \sqrt{m-\ell} + t)\right] \leq 2e^{-\pi t^2},\ t \geq 0,$$

for an absolute constants $C > 0$, ($C \approx 1/\sqrt{2\pi}$). A union bound then yields

$$\mathbb{P}_{\mathbf{Y}}\left[\|M_\tau(\mathbf{Y})\|_2 > Cs(\sqrt{\ell} + \sqrt{m-\ell} + t)\right] \leq 2n \cdot e^{-\pi t^2},\ t \geq 0,$$

For $x = \omega_n = \omega(\sqrt{\log_2 n})$, the bound becomes negligible. Hence, with probability at least $1 - 2ne^{-\pi\omega_n^2}$, we have that $\tau(\lambda h_{\mathbf{Y}}(\mathbf{x}))$ is bounded by

$$r = \sqrt{n}(\eta - 1)\left(\sqrt{\ell} + Cs\sqrt{m-\ell}(\sqrt{\ell} + \sqrt{m-\ell} + \omega_n)\right).$$

The number of integer points in the $n\ell$-dimensional ball of radius r is the dimensionless volume of the ball which is $(\sqrt{\pi}r)^{n\ell}/\Gamma(n\ell/2+1)$. Yet, it holds that $\Gamma(x+1) > \sqrt{2\pi x}(x/e)^x$. Therefore, we have that

$$\begin{aligned}
|h_{\mathbf{Y}}(X)| &\leq \frac{1}{\sqrt{\pi n\ell}}\left(\sqrt{\frac{2\pi e}{n\ell}} \cdot r\right)^{n\ell} \\
&\leq \frac{1}{\sqrt{\pi n\ell}}\left((\eta-1)\sqrt{2\pi e}\left(1 + Cs\sqrt{\frac{m-\ell}{\ell}}\left(\sqrt{\ell} + \sqrt{m-\ell} + \omega_n\right)\right)\right)^{n\ell}.
\end{aligned}$$

As the bound is independent of $\mathbf{Y}$, this provides the same bound for the expectation $\mathbb{E}_{h_{\mathbf{Y}} \hookleftarrow \mathcal{L}}[|h_{\mathbf{Y}}(X)|]$. Lemma 2.16 then yields the univertibility of $\mathcal{L}$. By Lemma 2.17, we thus obtain the uninvertibility of $\mathcal{F}$. □

4.2.2 Second Preimage Resistance of M-SIS.

We now prove the (statistical) second preimage resistance of the M-SIS function family with respect to the uniform distribution over an η-bounded domain.

Lemma 4.4. *Let K be a cyclotomic field of degree n, and R its ring of integers. Let k, q, m, η be positive integers such that q is prime. Let $X \subseteq (R_\eta^\vee)^m$. Then* (M-SIS$(n,k,m,q,X), U(X)$) *is (statistically) ε_4-second preimage resistant for*

$$\varepsilon_4 = (|X| - 1) \cdot \left(\frac{(\eta-1)\cdot n}{q}\right)^{nk}.$$

Proof. To prove it statistically, we show that for $\mathbf{A}, \mathbf{x}$ uniformly chosen, the probability that there exists $\mathbf{x}' \neq \mathbf{x}$ such that $\mathbf{A}^T\mathbf{x}' = \mathbf{A}^T\mathbf{x} \bmod qR^\vee$ is less than ε_4. Note that $\lambda X \subseteq R_\eta^m$. We then bound

$$\mathbb{P}_{\substack{\mathbf{A} \hookleftarrow U(R_q^{m\times k}) \\ \mathbf{x} \hookleftarrow U(\lambda X)}}[\exists \mathbf{x}' \in \lambda X \setminus \{\mathbf{x}\},\ \mathbf{A}^T\mathbf{x}' = \mathbf{A}^T\mathbf{x} \bmod qR].$$

Let $\mathbf{A} \in R_q^{m\times k}$ and $\mathbf{x} \in \lambda X$ be chosen uniformly at random. Fix $\mathbf{x}' \in \lambda X$ such that $\mathbf{x}' \neq \mathbf{x}$, and set $\mathbf{z} = \mathbf{x} - \mathbf{x}'$. Then $\mathbf{A}^T\mathbf{z} \bmod qR$ is uniformly distributed in $(\mathcal{I}_\mathbf{z}/qR)^k$ where $\mathcal{I}_\mathbf{z} = \langle z_1\rangle + \ldots \langle z_m\rangle + \langle q\rangle$. Hence the probability that $\mathbf{A}^T\mathbf{z} = \mathbf{0} \bmod qR$ is $|\mathcal{I}_\mathbf{z}/qR|^{-k}$. As $\mathcal{I}_\mathbf{z}$ and qR are ideals of R, we have $|\mathcal{I}_\mathbf{z}/qR| = N(qR)/N(\mathcal{I}_\mathbf{z}) = q^n/N(\mathcal{I}_\mathbf{z})$. Yet, for all $i \in [m]$, $\langle z_i\rangle \subseteq \mathcal{I}_\mathbf{z}$, so $N(\mathcal{I}_\mathbf{z})$ divides $N(\langle z_i\rangle)$. Similarly, $N(\mathcal{I}_\mathbf{z})$ divides $N(\langle q\rangle) = q^n$. Hence

$$N(\mathcal{I}_\mathbf{z}) \leq \gcd\left(q^n, N(\langle z_1\rangle), \ldots, N(\langle z_m\rangle)\right),$$

which yields the (loose) bound

$$N(\mathcal{I}_\mathbf{z}) \leq \min\left(q^n, \min_{i\in[m]:z_i\neq 0} N(\langle z_i\rangle)\right).$$

Since $\mathbf{z} \neq \mathbf{0}$, there exists $i \in [m]$ such that $z_i \neq 0$. Note that $\mathbf{z} \in \{\mathbf{a}-\mathbf{b}; (\mathbf{a},\mathbf{b}) \in (\lambda X)^2\}$ and thus has coefficients between $-(\eta-1)$ and $(\eta-1)$. It holds that

$$N(\langle z_i\rangle) = |N(z_i)| = \prod_{j\in[n]} |\sigma_j(z_i)| \leq \prod_{j\in[n]} \sum_{l=0}^{n-1} \underbrace{|\tau_l(z_i)|}_{\leq \eta-1} \cdot \underbrace{|\sigma_j(\zeta)|^l}_{=1} \leq ((\eta-1)\cdot n)^n.$$

Hence $\mathbb{P}[\mathbf{A}^T\mathbf{z} = \mathbf{0} \bmod qR] \leq ((\eta-1)\cdot n/q)^{nk}$. A union bound on $\mathbf{x}'$ concludes the proof. □

4.2.3 One-wayness of M-LWE with Small Uniform Error. Using the results from Sections 4.2.1 and 4.2.2, we can give the main theorem of this section. Under the assumption that the M-LWE function family is pseudorandom with respect to a Gaussian error distribution, it proves that the M-LWE function family is one-way with respect to a small uniform error distribution. Recall that if a function is one-way, then it is also uninvertible. Hence, this shows that the search version M-SLWE with small uniform error is at least as hard as the decision version M-LWE with Gaussian error.

Theorem 4.1. *Let K be a cyclotomic field of degree n, and R its ring of integers. Let d, q, m, k be positive integers such that q is an unramified prime, and $m > d \geq k \geq 1$. Let η be a positive integer, and $X \subseteq (R_\eta^\vee)^m$. We define $\ell = m-d+k$. Assume that (the primal function family) $(\text{M-LWE}(n,k,\ell,q,R^\ell), \mathcal{D}_{R,s}^\ell)$ is ε_1-pseudorandom for $s \geq \eta_\delta(R)$. Then $(\text{M-LWE}(n,d,m,q,X), U(X))$ is ε-one-way for*

$$\varepsilon = (d-k)\varepsilon_1 + (|X|-1)\cdot\left(\frac{(\eta-1)\cdot n}{q}\right)^{n(m-d)} + \varepsilon_3 + \mathsf{negl}(n),$$

where ε_3 is defined in the statement of Lemma 4.3.

The expression of ε actually involves the function $\delta(\cdot,\cdot)$ defined in Equation 1 when moving from M-LWE to M-SIS and back. We discuss later why these terms are negligible in the range of parameters that are needed by Theorem 4.2.

Proof. We define $\mathcal{F} = \text{M-SIS}(n, m-d, \ell, q, (R^\vee)^\ell) \circ \mathcal{L}(\ell, m, s, X)$, and $\mathcal{G} = \text{M-SIS}(n, m-d, m, q, X)$.
Indistinguishability: Using Lemma 4.1, the pseudorandomness of the M-LWE function family implies that $(\text{M-SIS}(n, m-d, \ell, q, R^\ell), \mathcal{D}_{R,s}^\ell)$ is ε_2-pseudorandom with

$$\varepsilon_2 = \delta(\ell, \ell - k) + \frac{\varepsilon_1}{1 - \delta(\ell, k)}.$$

Take $f_{\mathbf{A}} \circ h_{\mathbf{Y}}$ according to $\mathcal{F}$, and $f_{\mathbf{A}'}$ according to $\mathcal{G}$. Then $f_{\mathbf{A}} \circ h_{\mathbf{Y}}$ is the linear map $\mathbf{x} \mapsto [\mathbf{A}^T \mid \mathbf{A}^T\mathbf{Y}]\mathbf{x}$. Decomposing $\mathbf{A}'^T$ into $[(\mathbf{A}'_1)^T \mid (\mathbf{A}'_2)^T]$, with $\mathbf{A}'_1 \in R_q^{\ell \times (m-d)}, \mathbf{A}'_2 \in R_q^{(m-\ell)\times(m-d)}$, we have that $f_{\mathbf{A}'} = \mathbf{x} \mapsto [(\mathbf{A}'_1)^T \mid (\mathbf{A}'_2)^T]\mathbf{x}$. By the ε_2-pseudorandomness of M-SIS with respect to $\mathcal{D}_{R,s}^\ell$, a hybrid argument yields that $\mathcal{F}$ and $\mathcal{G}$ are $(m-\ell)\varepsilon_2$-indistinguishable.
Uninvertibility: By Lemma 4.3, it holds that $(\mathcal{F}, U(X))$ is ε_3-uninvertible, where ε_3 is defined in Lemma 4.3.
Second Preimage Resistance: By Lemma 4.4, it holds that $(\mathcal{G}, U(X))$ is ε_4-second preimage resistant for

$$\varepsilon_4 = (|X| - 1) \cdot \left(\frac{(\eta - 1) \cdot n}{q}\right)^{n(m-d)}.$$

We thus have that $(\mathcal{F}, \mathcal{G}, U(X))$ is a lossy function family, depending on $\varepsilon_2, \varepsilon_3, \varepsilon_4$. Lemma 2.15 yields that $(\mathcal{G}, U(X))$ is ε_0-one-way with $\varepsilon_0 = (m-\ell)\varepsilon_2 + \varepsilon_3 + \varepsilon_4$. Using Lemma 4.2, it gives that $(\text{M-LWE}(n, d, m, q, X), U(X))$ is ε-one-way with

$$\varepsilon = \delta(m, d) + \frac{\varepsilon_0}{1 - \delta(m, m-d)} + \mathsf{negl}(n).$$

Combining everything, we get

$$\varepsilon = \delta(m, d) + \frac{1}{1 - \delta(m, m-d)} \left((m-\ell) \left(\delta(\ell, m-d) + \frac{\varepsilon_1}{1 - \delta(\ell, k)} \right) + \varepsilon_3 + \varepsilon_4 \right) + \mathsf{negl}(n).$$

Replacing ε_3 and ε_4 by their expressions, and argueing that all the $\delta(\cdot, \cdot)$ are negligible yield the claim. □

We now combine Theorem 2.1 and 4.1 to base the one-wayness of M-LWE on module lattice problems (or ideal lattice problems if $k = 1$).

Theorem 4.2. *Let K be a cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let d, q, m, k, η be positive integers and $X = (R_\eta^\vee)^m$. Let $s = 2\sqrt{k}\omega(\sqrt{\log_2 n})$ if $k > 1$ or $s = \omega(\sqrt{\log_2 n})$ if $k = 1$. Assume that*

- *q is prime with $q = 1 \bmod \nu$, and $q > (\eta - 1) \cdot n \cdot \eta^{m/(m-d)}$*
- *$m > d \geq k \geq 1$*
- *$\varepsilon_3 \leq n^{-\omega(1)}$*

Let $\gamma = q\sqrt{8nk}$ if $k > 1$ or $\gamma = q\sqrt{n}$ if $k = 1$. Then, assuming that Mod-$\text{GIVP}_\gamma^{\eta_{\varepsilon'}}$ is hard for some $\varepsilon' > 0$, it holds that $(\text{M-LWE}(n, d, m, q, X), U(X))$ is one-way.

Let us now discuss the various conditions that are needed to apply this theorem. The lower bound on q comes from ensuring that ε_4 is negligible. Indeed, we have that $|X| = \eta^{nm}$, and therefore it suffices to have

$$\eta^m \left(\frac{(\eta - 1) \cdot n}{q} \right)^{m-d} < 1, \tag{7}$$

which can be written as $q > (\eta - 1) \cdot n \cdot \eta^{m/(m-d)}$. Hence, for $\alpha > 0$ one can choose $q > 2^{\alpha/(m-d)} \cdot (\eta - 1) \cdot n \cdot \eta^{m/(m-d)}$ which ensures $\varepsilon_4 < 2^{-\alpha n}$. The expression of ε_3 is more involved, but the idea is the same. For it to be negligible, we need

$$\frac{(\eta - 1)^\ell}{\eta^m (\pi n \ell)^{1/2n}} \left(\sqrt{2\pi e} \left(1 + Cs \sqrt{\frac{m - \ell}{\ell}} \left(\sqrt{\ell} + \sqrt{m - \ell} + \omega_n \right) \right) \right)^\ell < 1, \tag{8}$$

where $\omega_n = \omega(\sqrt{\log_2 n})$. Due to the many dependencies in m, k, d and η, it is harder to extract a concrete inequality on m given k, d and η. Instead, we evaluate the inequality with different parameters while trying to minimize η and maximize m, while ensuring $m > d \geq k \geq 1$. As we aim at proving the hardness of M-LWE with small parameters, one can evaluate Equations (7) and (8) with the goal of minimizing η, q and d, while maximizing m and making sure that $k \geq 1$ ($k \geq 2$ being preferable to rely on module lattice assumptions). It turns out that the condition is not met for all set of parameters, and η cannot be arbitrarily small for any ranks k, d. Nonetheless, we can find settings in which η is a small constant, but this might require to take d slightly larger. As expected, when $m - d$ grows for a fixed d, the error bound η must be larger as well. Table 4.1 give two example sets of parameters that verify the conditions, along with the losses $\varepsilon_3, \varepsilon_4$, one relying on ideal lattice assumptions.

n	k	d	m	η	q	ε_3	ε_4
256	1	12	13	3	$\approx 2^{31}$	$\approx 2^{-346} + 2^{-281}$	2^{-256}
256	2	11	12	10	$\approx 2^{52}$	$\approx 2^{-333} + 2^{-281}$	2^{-256}

Table 4.1. Example parameter sets reaching the conditions of Theorem 4.2. We take $\omega_n = \log_2 n$, and $s \approx \log_2 n$ if $k = 1$ and $s \approx 2\sqrt{k} \log_2 n$ if $k > 1$. Empirically, we have $C \approx 1/\sqrt{2\pi}$ as noticed for example in [MP12, Sec. 2.4]. The loss of 2^{-281} that dominates in the value of ε_3 comes from the spectral bound loss $2n \cdot e^{-\pi \log_2^2 n}$.

Remark 4.1. Note that we can provide the asymptotic behavior $\varepsilon_3 = O(s \cdot m \cdot \eta \cdot \omega_n)^{n\ell}/|X| + 2ne^{-\pi\omega_n^2}$, but this approach makes it unclear how to choose the parameters. In particular, as we can use low ranks like $d = O(1)$, we have to make sure that $k \geq 1$ and $m \geq d + 1$, which is not always possible for low values

of η. The asymptotic approach gives the more direct condition on m

$$d < m < (d-k)\left(1 + \frac{\log_2 \eta}{\log_2(C' \cdot m \cdot \omega_n^2)}\right),$$

which is much similar to the condition in [MP13]. The main difference stems from the fact that m is no longer our asymptotic parameter, which explains the presence of $\omega_n^2 = \omega(\log_2 n)$. It still remains difficult to see which parameter sets meet this condition, mostly because the constant C' can be rather large while we wish d and k to be small constants.

4.3 On Hermite Normal Form M-LWE with Small Keys

We now look at the use of our result to obtain the hardness of M-LWE where both the error and secret distribution are uniform over small elements. To do so we combine Theorem 4.2 with a Hermite Normal Form transformation for M-LWE. Langlois and Stehlé [LS15, Lem. 4.24] proposed an immediate generalization of the reduction from LWE to its Hermite Normal Form by Applebaum et al. [ACPS09] to modules. We clarify this generalization to highlight the trade-off between the loss in advantage of the reduction and the number of standard M-LWE samples that are queried by the reduction. The proof can be found in Appendix A.3 for completeness.

Lemma 4.5 (Adapted from [ACPS09,LS15]). *Let K be a number field of degree n, and R its ring of integers. Let d, q, m' be positive integers such that q is an unramified prime, and $m' \geq d \geq 1$. Let $\mathbf{s}$ be an arbitrary vector of $(R_q^\vee)^d$ and ψ a distribution over $R^\vee$. There is an efficient transformation T such that $T(A_{\mathbf{s},\psi}) = A_{\mathbf{x},\psi}$ for some $\mathbf{x}$ sampled from ψ^d, and $T(U(R_q^d \times R_q^\vee)) = U(R_q^d \times R_q^\vee)$. T can be constructed in polynomial time using m' samples from $\mathcal{D} \in \{A_{\mathbf{s},\psi}, U(R_q^d \times R_q^\vee)\}$ with probability $1 - \delta(m', d)$.*

This transformation shows a reduction from worst-case (or average-case if $\mathbf{s}$ is uniformly sampled over $(R_q^\vee)^d$ instead of arbitrary) search-M-LWE to search-HNF-M-LWE, but also from decision-M-LWE to decision-HNF-M-LWE. All the parameters are preserved except for the number of samples, as we need $m' \geq d$ extra samples to construct the transformation, i.e., construct the invertible matrix $\overline{\mathbf{A}}$ involved in the map with the corresponding $\overline{\mathbf{b}}$. To prove the hardness of HNF-M-LWE with m samples, we thus need to assume the hardness of M-LWE with $m+m'$ samples. The choice of m' allows for tweaking the success probability of the reduction, but at the expense of requiring more samples. For completeness, we now analyze this trade-off in more details.

Since the loss in advantage is decreasing when m' grows away from d, we can upper bound this loss by the one when $m' = d$. Hence

$$\delta(m', d) \leq \delta(d, d) = \sum_{k=0}^{d-1} \binom{d}{k} p_d^k (1-p_d)^{d-k} = 1 - p_d^d.$$

It may seem like this bound is not useful as $1-p_d^d$ seems to be exponentially close to 1. However, p_d is itself exponentially close to 1 in many parameter regimes. For example in the common case of cyclotomic fields with a fully splitted prime modulus, we have

$$\delta(m', d) \leq 1 - \left(1 - (1 - (1 - 1/q)^n)^d\right)^d.$$

Note that if q splits into fewer factors, then p_d becomes closer to 1 and therefore $\delta(d, d)$ becomes smaller. In particular, for an inert prime q, we have $\delta(d, d) = q^{-nd}$. The most unfavorable situation is therefore the case of fully splitted primes. Even in this particular case, it can be seen that the bound is exponentially small with respect to d for common parameters like $n = 256$, and $q = \omega(n)$. In the parameter setting of CRYSTALS for example, we have the values from Table 4.2. In the case of Kyber [BDK+18], the modulus splits into $n/2$ prime factors with inertia degree 2. This means that $p_d = 1 - (1 - (1 - q^{-2})^{n/2})^d$. Note that the parameter q required by Theorem 4.2 is however much larger, which makes the loss $\delta(d, d)$ even smaller.

	Dilithium			**Kyber**		
Security Level	2	3	5	512	768	1024
Ring degree n	256	256	256	256	256	256
Module rank d	4	5	7	2	3	4
Modulus q	8380417	8380417	8380417	3329	3329	3329
Loss $\delta(d, d)$	2^{-58}	$2^{-72.7}$	$2^{-102.2}$	$2^{-31.8}$	$2^{-47.6}$	$2^{-63.6}$

Table 4.2. Bound on the loss in advantage for concrete M-LWE parameters used in the CRYSTALS suite [DKL+18] [BDK+18] taken from the specification papers.

We then obtain the following result by combining Theorem 4.2 with Lemma 4.5.

Corollary 4.1. *Let $K, R, n, d, m, q, k, \eta, s, \gamma$ be as in Theorem 4.2. Then, assuming that Mod-*GIVP$^{\eta_{\varepsilon'}}_{\gamma}$ *is hard for some $\varepsilon' > 0$, it holds that search version η-*M-SLWE$_{n,d,q,m-d,U(R_\eta^\vee)}$ *is hard.*

4.4 A Thought on Practical Hardness

Several cryptanalytic works target the LWE problem, with sometimes increased efficiency when the parameters are small, e.g. particularly small secret, or particularly small error. They leverage either lattice reduction [LP11,LN13], combinatorial [Wag02,BKW03,KF15] or algebraic [AG11] techniques. The latter attack by Arora and Ge specifically targets LWE with small errors. It does not depend on the underlying structure, and therefore also applies to the more general case of M-LWE. The idea is to see the (search) LWE problem as solving a noisy system of equations, and transforming it into a noiseless polynomial system (where

the degree of the polynomials depend on the size of the LWE error). Then, using root finding algorithms for multivariate polynomials, one can solve the new system.

More precisely in the case of LWE with η-bounded error ($\mathbf{e} \in \mathbb{Z}_\eta^m$), the Arora and Ge attack [AG11] solves the problem in polynomial time if $m \approx \binom{d+\eta}{\eta} = \Omega(d^\eta)$, where d is the LWE dimension. For $\eta = 2$, the attack becomes subexponential for $m = \omega(d)$ and exponential for $m = O(d)$. As the attack ignores the structure, one can embed the m M-LWE equations with d unknowns over R_q into nm equations with nd unknowns over $\mathbb{Z}_q$ and apply the same attack. However, we now obtain a polynomial attack only for $nm = \Omega((nd)^\eta)$ and therefore $m = \Omega(n^{\eta-1}d^\eta)$. In practical schemes relying on M-LWE with small errors [BDK+18,DKL+18], the rank d is a small constant and n drives the security parameter. Additionally, we saw in Section 4.3 that roughly $m = m' + d$ is enough to establish the hardness of M-LWE with small secret *and* error with m' samples. For common parameters where $m' = d$ or $d+1$, we thus have $m \approx 2d \ll n^{\eta-1}d^\eta$. This is why we think that the hardness of M-LWE with both small secret and error is yet to be determined. The gap between what we proved in this section and the applicable attacks seem wide enough to improve in either direction: either by finding new attacks that require fewer samples, or by improving theoretical hardness results to allow for more samples.

5 A Quick Survey on the Hardness of M-LWE

This section aims at gathering all known results on the hardness of the M-LWE problem along with our new contributions, and comparing them whenever possible.

General Hardness. Although the M-LWE problem was originally introduced in [BGV12] for power-of-two cyclotomic fields, its hardness was first studied by Langlois and Stehlé in [LS15]. They established the hardness of the standard formulation $\text{M-LWE}_{n,d,m,q,\mathcal{D}_{R^\vee,\alpha}}$ based on the quantum hardness of $\text{Mod-GIVP}_\gamma^{\eta_\varepsilon}$ (Definition 2.1), where $\alpha = \tilde{\Omega}(d\sqrt{n}/\gamma)$. For the full statement, refer to Theorem 2.1. Although the proof for the decision version requires q to be a fully splitted prime in the cyclotomic ring, the authors gave a modulus switching reduction showing that the form of q is not restrictive if one accepts a (moderately) increased error. This reduction proves that if $\text{M-LWE}_{n,d,m,q,\Upsilon_\alpha}$ is hard, then so is $\text{M-LWE}_{n,d,m,p,\Upsilon_{\alpha'}}$ for an arbitrary modulus p. This comes at the expense of increasing the error from α to $\alpha' \geq \alpha \cdot \max(1, q/p) \cdot n^{3/4}\sqrt{d}\omega(\log_2^2 n)$. The quantum reduction from [LS15] was later used in [AA16] to derive the hardness of M-LWR (Module Learning With Rounding) which we do not cover in this discussion.

As discussed in [BGV12], the M-LWE problem offers a trade-off between security and efficiency depending on whether the parameters lean towards LWE (degree n equal to 1) or R-LWE (rank d equal to 1) respectively. In particular, establishing a hierarchy of hardness between M-LWE and the widely studied and

used R-LWE was up for debate. Albrecht and Deo [AD17a] provided a first answer by showing that R-LWE with modulus q^d is at least as hard as M-LWE with modulus q and rank d. This actually comes as a byproduct of their more general result showing a modulus-rank switching reduction from $\text{M-LWE}_{n,d,m,q,D_\alpha}$ to $\text{M-LWE}_{n,d',m,q',D_{\alpha'}}$. The moduli and ranks can be arbitrarily chosen provided that one can efficiently describe the lattice $\Lambda = {q'}^{-1}\mathbf{G}^T R^{d'} + R^d$ for a chosen $\mathbf{G} \in R^{d' \times d}$. It also requires to increase the error from α to $\alpha' \geq \sqrt{\alpha^2 + \Delta}$, where Δ depends on the size of the secret distribution and the quality of the description of Λ. As a result, the reduction becomes less interesting for very large secrets. We refer to [AD17a,AD17b] for the detailed expression of Δ. The reduction to R-LWE was later improved and generalized by Wang and Wang [WW19] to hold over all cyclotomic fields. A revision of the work by Albrecht and Deo, which can be found in [AD17b], further improved this line of work with a new analysis. Additionally, a result from Peikert and Pepin [PP19] tightly proves the hardness of M-LWE over a number field K of degree n and with rank d assuming the hardness of R-LWE over any one of a class of number field extensions K'/K with extension degree $d = [K' : K]$. Instead of showing a modulus-rank trade-off as in [AD17a], they provide a degree-rank trade-off, where the underlying ring structure is changed, while preserving the modulus q. Note that, in contrast to [AD17a], their reduction allows for an arbitrary large uniform secret.

Small Distributions Hardness. The work of this paper focuses on the hardness of M-LWE when the secret and error distributions deviate from the original formulation. The first result in this line of work was due to [LS15], which extended the reduction by Applebaum et al. [ACPS09] to modules. In particular, combined with their main proof of hardness, it can be used to obtain the hardness of $\text{M-LWE}_{n,d,m,q,\mathcal{D}_{R^\vee,\alpha}}$ with secrets drawn from $\mathcal{D}^d_{R^\vee,\alpha}$. As observed in Lemma 4.5, this is at the expense of using $m' \geq d$ M-LWE samples to construct the transformation.

Section 3 provides the first proofs of hardness for M-LWE with small bounded secret, in both the search (Section 3.1) and decision (Section 3.2) variants. As discussed, they generalize the approaches by Goldwasser et al. [GKPV10] and by Brakerski et al. [BLP+13] respectively, which are the analog results for LWE. These results can be used to derive the classical hardness of M-LWE, meaning that if one has a classical solver for M-LWE, then it can also construct a classical solver for worst-case module lattice problems. This removes the need for quantum algorithms in the reduction of [LS15], with the caveat of introducing further restrictions on the parameters. More precisely, the result of Section 3.1 can be instantiated for binary secrets ($\eta = 2$) to prove the classical hardness of M-LWE as in our previously published work [BJRW20]. The analysis in the present paper is slightly improved compared to that of the conference paper [BJRW20]. It yields a classical reduction from Mod-GapSVP_γ in module lattices of rank nk to $\text{M-LWE}_{n,d,m,p,\Psi_{\leq\alpha}}$, where $d \geq k^2 n/2 + \Omega(\log_2 n)$ and $\alpha = \widetilde{\Omega}(n^{19/4}\sqrt{m}/\gamma)$ (where α can be improved by $\sqrt{n}$ in power-of-two cyclotomic fields). We can alternatively use the result of Section 3.2 within the classical hardness proof. It

leads to a classical reduction from Mod-GapSVP$_\gamma$ in module lattices of rank nk to M-LWE$_{n,d,m,p,\Psi_{\leq\alpha}}$ for $d \geq k(k+1)n/2 + \omega(\log_2 n)$ and $\alpha = \tilde{\Omega}(n^{7/2}/\gamma)$, thus achieving a smaller error α but at the expense of a larger rank d.

Another line of work studied by Brakerski and Döttling for LWE [BD20a] and R-LWE [BD20b] was recently extended to M-LWE by Lin et al. [LWW20]. It looks at the hardness of the problem when the only requirement on the secret distribution is to contain a sufficient entropy. Although [BD20b] cannot be instantiated for η-bounded secret with η being a small constant, the result by [LWW20] on M-LWE can with certain restrictions. In particular, the entropy condition in this specific instantiation becomes a condition on the module ranks d and k that is similar to ours, i.e., $d \log_2 \eta \gtrsim k \log_2 q$. This does not come as a surprise as the proof relies more or less on the same lossy argument as ours.

Finally, prior to our work, no result was formally known about the hardness of M-LWE with unusually small uniform error. We once again stress that the rank d, number of samples m and error bound η must be cautiously chosen with respect to one another for Theorem 4.2 to apply. As mentioned, the algebraic attacks, e.g. [AG11], on this variant do not depend on the underlying structure and therefore apply for LWE as well as M-LWE. When the number of samples m covered by the proof of hardness is sufficiently larger than the rank d, the Hermite Normal Form transform from Lemma 4.5 can be used to derive the hardness of M-LWE with uniform η-bounded secret *and* error. This regime would give strong hardness guarantees for practical schemes, provided that the number of available samples is sufficient once again.

Summary. We summarize the results and the achieved parameters when clear in Table 5.1. The achievable rank d depends on the secret distribution (for [LWW20], Sections 3.1 and 3.2) or on the error size (for Section 4).

Acknowledgments

This work was supported by the European Union PROMETHEUS project (Horizon 2020 Research and Innovation Program, grant 780701). Katharina Boudgoust was funded by the Direction Générale de l'Armement (Pôle de Recherche CYBER). We thank our anonymous referees of Asiacrypt 2020, Indocrypt 2020 and CT-RSA 2021 for their throrough proof reading and constructive feedback on the original papers. We also thank Thomas Prest for making us aware of improved rank conditions when using of the leftover hash lemma in the Rényi divergence.

	[LS15] (quantum)	[BJRW20] (classical)	[LWW20]	Sec. 3.1	Sec. 3.2	Sec. 4
Field K	Cyclo	Cyclo	All	Monogenic	Cyclo	Cyclo
Rank d	All	$\Omega(n)$	(Depends)	$\Omega(\log_2 n)$	$\omega(\log_2 n)$	All
Modulus q	All (**S**) *FSP* (**D**)	Prime (**S**) *FSP* (**D**)	All (**S**) *IP* (**D**)	Prime	Prime, number-theoretic cond.	*FSP*
Secret $\mathbf{s}$	Mod q Gaussian[a]	Mod q	Entropic	η-bounded	η-bounded	Mod q η-bounded[a]
Error $\mathbf{e}$	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	η-bounded
Variant	**S**/**D**	**S**/**D**	**S**(/**D**[b])	**S**	**S**[c]/**D**	**S**

Table 5.1. Summary of the results on the hardness of the M-LWE problem. *FSP* stands for Fully Splitted Prime, *IP* for Inert Prime, and **S** denotes the search version, while **D** denotes the decision version. A fully splitted prime is a prime integer q such that qR factors into n distinct prime ideals, each of algebraic norm q. On the opposite, an inert prime is a prime integer q such that qR is a prime ideal. By abuse of language, we call monogenic the number fields $K = \mathbb{Q}(\zeta)$ for which $R = \mathbb{Z}[\zeta]$. Note that rigorously, a monogenic number field is $K = \mathbb{Q}(\zeta)$ for which $R = \mathbb{Z}[\zeta']$ for a possibly different ζ'.

[a] Obtained by the Hermite Normal Form transformation reduction from [LS15].

[b] The hardness proof for the decision version of M-LWE requires q to be an inert prime. This is a very restrictive condition as certain number fields do not contain any inert primes. For example, there exists inert primes in the ν-th cyclotomic field if and only if ν is $2, 4$ or $2^b p^k$ for $b \in \{0, 1\}$ and p an odd prime. Then, (almost) all power-of-two cyclotomic fields do not contain inert primes.

[c] The hardness of the search version is obtained from the decision version through a trivial reduction.

References

[AA16] J. Alperin-Sheriff and D. Apon. Dimension-preserving reductions from LWE to LWR. *IACR Cryptol. ePrint Arch.*, page 589, 2016.

[ACPS09] B. Applebaum, D. Cash, C. Peikert, and A. Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *CRYPTO*, volume 5677 of *Lecture Notes in Computer Science*, pages 595–618. Springer, 2009.

[AD17a] M. R. Albrecht and A. Deo. Large modulus ring-lwe $\geq$ module-lwe. In *ASIACRYPT (1)*, volume 10624 of *Lecture Notes in Computer Science*, pages 267–296. Springer, 2017.

[AD17b] M. R. Albrecht and A. Deo. Large modulus ring-lwe >= module-lwe. *IACR Cryptol. ePrint Arch.*, page 612, 2017.

[AG11] S. Arora and R. Ge. New algorithms for learning in presence of errors. In *ICALP (1)*, volume 6755 of *Lecture Notes in Computer Science*, pages 403–415. Springer, 2011.

[BD20a] Z. Brakerski and N. Döttling. Hardness of LWE on general entropic distributions. In *EUROCRYPT (2)*, volume 12106 of *Lecture Notes in Computer Science*, pages 551–575. Springer, 2020.

[BD20b] Z. Brakerski and N. Döttling. Lossiness and entropic hardness for ring-lwe. In *TCC (1)*, volume 12550 of *Lecture Notes in Computer Science*, pages 1–27. 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 *EuroS&P*, pages 353–367. IEEE, 2018.

[BGV12] Z. Brakerski, C. Gentry, and V. Vaikuntanathan. (leveled) fully homomorphic encryption without bootstrapping. In *ITCS*, pages 309–325. ACM, 2012.

[BJRW20] K. Boudgoust, C. Jeudy, A. Roux-Langlois, and W. Wen. Towards classical hardness of module-lwe: The linear rank case. In *ASIACRYPT (2)*, volume 12492 of *Lecture Notes in Computer Science*, pages 289–317. Springer, 2020.

[BJRW21] K. Boudgoust, C. Jeudy, A. Roux-Langlois, and W. Wen. On the hardness of module-lwe with binary secret. In *CT-RSA*, volume 12704 of *Lecture Notes in Computer Science*, pages 503–526. Springer, 2021.

[BKW03] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, 2003.

[BLP+13] Z. Brakerski, A. Langlois, C. Peikert, O. Regev, and D. Stehlé. Classical hardness of learning with errors. In *STOC*, pages 575–584. ACM, 2013.

[BLR+18] S. Bai, T. Lepoint, A. Roux-Langlois, A. Sakzad, D. Stehlé, and R. Steinfeld. Improved security proofs in lattice-based cryptography: Using the rényi divergence rather than the statistical distance. *J. Cryptol.*, 31(2):610–640, 2018.

[Bou21] K. Boudgoust. Theoretical hardness of algebraically structured learning with errors, 2021. https://katinkabou.github.io/Documents/Thesis_Boudgoust_Final.pdf.

[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 *EUROCRYPT (1)*, volume 9056 of *Lecture Notes in Computer Science*, pages 617–640. Springer, 2015.

[GKPV10] S. Goldwasser, Y. Tauman Kalai, C. Peikert, and V. Vaikuntanathan. Robustness of the learning with errors assumption. In *ICS*, pages 230–240. Tsinghua University Press, 2010.

[KF15] P. Kirchner and P.-A. Fouque. An improved BKW algorithm for LWE with applications to cryptography and lattices. In *CRYPTO (1)*, volume 9215 of *Lecture Notes in Computer Science*, pages 43–62. Springer, 2015.

[LN13] Mingjie Liu and Phong Q. Nguyen. Solving BDD by enumeration: An update. In *CT-RSA*, volume 7779 of *Lecture Notes in Computer Science*, pages 293–309. Springer, 2013.

[LNS21] Vadim Lyubashevsky, Ngoc Khanh Nguyen, and Gregor Seiler. Shorter lattice-based zero-knowledge proofs via one-time commitments. In *Public Key Cryptography (1)*, volume 12710 of *Lecture Notes in Computer Science*, pages 215–241. Springer, 2021.

[LP11] Richard Lindner and Chris Peikert. Better key sizes (and attacks) for lwe-based encryption. In *CT-RSA*, volume 6558 of *Lecture Notes in Computer Science*, pages 319–339. Springer, 2011.

[LPR13a] 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.

[LPR13b] V. Lyubashevsky, C. Peikert, and O. Regev. A toolkit for ring-lwe cryptography. In *EUROCRYPT*, volume 7881 of *Lecture Notes in Computer Science*, pages 35–54. Springer, 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 *EUROCRYPT (1)*, volume 10820 of *Lecture Notes in Computer Science*, pages 204–224. Springer, 2018.

[LSS14] A. Langlois, D. Stehlé, and R. Steinfeld. Gghlite: More efficient multilinear maps from ideal lattices. In *EUROCRYPT*, volume 8441 of *Lecture Notes in Computer Science*, pages 239–256. Springer, 2014.

[LWW20] H. Lin, Y. Wang, and M. Wang. Hardness of module-lwe and ring-lwe on general entropic distributions. *IACR Cryptol. ePrint Arch.*, page 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 Comput.*, 14(1):1–17, 2018.

[MM11] D. Micciancio and P. Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In *CRYPTO*, volume 6841 of *Lecture Notes in Computer Science*, pages 465–484. Springer, 2011.

[MP12] D. Micciancio and C. Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *EUROCRYPT*, volume 7237 of *Lecture Notes in Computer Science*, pages 700–718. Springer, 2012.

[MP13] D. Micciancio and C. Peikert. Hardness of SIS and LWE with small parameters. In *CRYPTO (1)*, volume 8042 of *Lecture Notes in Computer Science*, pages 21–39. Springer, 2013.

[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 *STOC*, pages 333–342. ACM, 2009.

[Pei10] C. Peikert. An efficient and parallel gaussian sampler for lattices. In *CRYPTO*, volume 6223 of *Lecture Notes in Computer Science*, pages 80–97. Springer, 2010.

[PP19] C. Peikert and Z. Pepin. Algebraically structured lwe, revisited. In *TCC (1)*, volume 11891 of *Lecture Notes in Computer Science*, pages 1–23. Springer, 2019.

[Ré61] A. Rényi. On measures of entropy and information. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I*, pages 547–561. Univ. California Press, Berkeley, Calif., 1961.

[Reg05] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93. ACM, 2005.

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

[Rja94] S. Rjasanow. Effective algorithms with circulant-block matrices. *Linear Algebra and its Applications*, 202:55–69, 1994.

[RSW18] M. Rosca, D. Stehlé, and A. Wallet. On the ring-lwe and polynomial-lwe problems. In *EUROCRYPT (1)*, volume 10820 of *Lecture Notes in Computer Science*, pages 146–173. Springer, 2018.

[SSTX09] D. Stehlé, R. Steinfeld, K. Tanaka, and K. Xagawa. Efficient public key encryption based on ideal lattices. In *ASIACRYPT*, volume 5912 of *Lecture Notes in Computer Science*, pages 617–635. Springer, 2009.

[STA20] Chao Sun, Mehdi Tibouchi, and Masayuki Abe. Revisiting the hardness of binary error LWE. In *ACISP*, volume 12248 of *Lecture Notes in Computer Science*, pages 425–444. Springer, 2020.

[vEH14] T. van Erven and P. Harremoës. Rényi divergence and kullback-leibler divergence. *IEEE Trans. Inf. Theory*, 60(7):3797–3820, 2014.

[Ver12] Roman Vershynin. *Introduction to the non-asymptotic analysis of random matrices*, page 210–268. Cambridge University Press, 2012.

[Wag02] David A. Wagner. A generalized birthday problem. In *CRYPTO*, volume 2442 of *Lecture Notes in Computer Science*, pages 288–303. Springer, 2002.

[WW19] Y. Wang and M. Wang. Module-lwe versus ring-lwe, revisited. *IACR Cryptol. ePrint Arch.*, page 930, 2019.

Appendix A Missing Proofs

A.1 Missing Proofs of Section 2

Lemma 2.1.

Proof. The lower bound is due to the fact that every non-zero element x of R has algebraic norm $N(x) \geq 1$, which implies that $\|\sigma(x)\|_\infty \geq 1$. Let x be in R_η, and $i \in [n]$. Then, it holds that

$$\begin{aligned}|\sigma_i(x)| \leq \sum_{j=0}^{n-1} \left|\tau_j(x)\sigma_i(\zeta)^j\right| &= \sum_{j=0}^{n-1} |\tau_j(x)||\alpha_i|^j \\ &\leq \|\tau(x)\|_1 \|\mathbf{V}\|_{\max} \leq n(\eta-1)\|\mathbf{V}\|_{\max}.\end{aligned}$$

Taking the maximum over all $i \in [n]$ and $x \in R_\eta$ yields $B_\eta \leq n(\eta-1)\|\mathbf{V}\|_{\max}$. 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 $B_\eta \leq n(\eta-1)$ in this case. □

Lemma 2.2.

Proof. Let $f = x^n + \sum_{k=0}^{n-1} f_k x^k$ denote the minimal polynomial of ζ, and $K = \mathbb{Q}(\zeta)$. Let $\mathbf{C}$ denote the companion matrix of f, as in the lemma statement. It is well known that the characteristic (and minimal) polynomial of the companion matrix of f is f itself. This entails that $\mathbf{C}$ has the roots of f for eigenvalues, which we denote by $\alpha_1, \ldots, \alpha_n$. Recall that the field embeddings are such that $\sigma_i(\zeta) = \alpha_i$ for all $i \in [n]$. Since the roots of f are distinct, it means that $\mathbf{C}$ is diagonalizable. More precisely, it holds that $\mathbf{C} = \mathbf{V}^{-1}\mathrm{diag}(\alpha_1, \ldots, \alpha_n)\mathbf{V} = \mathbf{V}^{-1}\mathrm{diag}(\sigma(\zeta))\mathbf{V}$. Now let x be in K. We have

$$\forall y \in K,\ \tau(xy) = \mathbf{V}^{-1}\sigma(xy) = \mathbf{V}^{-1}\mathrm{diag}(\sigma(x))\sigma(y) = \mathbf{V}^{-1}\mathrm{diag}(\sigma(x))\mathbf{V}\tau(y),$$

thus proving that $M_\tau(x) = \mathbf{V}^{-1}\mathrm{diag}(\sigma(x))\mathbf{V}$. We can then rewrite this expression in terms of the τ_k and $\mathbf{C}$ as follows.

$$\begin{aligned}\mathbf{V}^{-1}\mathrm{diag}(\sigma(x))\mathbf{V} &= \mathbf{V}^{-1}\mathrm{diag}\left(\sigma_1\left(\sum_{k=0}^{n-1}\tau_k(x)\zeta^k\right), \ldots, \sigma_n\left(\sum_{k=0}^{n-1}\tau_k(x)\zeta^k\right)\right)\mathbf{V} \\ &= \sum_{k=0}^{n-1}\tau_k(x)\mathbf{V}^{-1}\mathrm{diag}(\sigma_1(\zeta)^k, \ldots, \sigma_n(\zeta)^k)\mathbf{V} \\ &= \sum_{k=0}^{n-1}\tau_k(x)\mathbf{V}^{-1}\mathrm{diag}(\sigma(\zeta))^k\mathbf{V} \\ &= \sum_{k=0}^{n-1}\tau_k(x)\mathbf{C}^k,\end{aligned}$$

concluding the proof. □

Lemma 2.3.

Proof. For (i,j) in $[d] \times [m]$, we define the polynomial function $a_{ij}(\cdot) : t \mapsto \sum_{k=0}^{n-1} \tau_k(a_{ij}) t^k$. The way $a_{ij} \in K$ is defined, we have $a_{ij} = a_{ij}(\zeta)$. Lemma 2.2 gives $M_\tau(a_{ij}) = \sum_{k=0}^{n-1} \tau_k(a_{ij}) \mathbf{C}^k = a_{ij}(\mathbf{C})$. Finally, for $k \in [n]$, if α_k denotes $\sigma_k(\zeta)$, it holds that $a_{ij}(\alpha_k) = \sigma_k(a_{ij})$. We then define the function over complex matrices by $\mathbf{A}(t) = [a_{ij}(t)]_{(i,j)}$ for all t. By the prior observations, we get that $\mathbf{A} = \mathbf{A}(\zeta)$, $M_\tau(\mathbf{A}) = \mathbf{A}(\mathbf{C})$, and $\mathbf{A}(\alpha_k) = \sigma_k(\mathbf{A})$.

Consider $\mathbf{B}(t) = \mathbf{A}(t)^\dagger \mathbf{A}(t)$. The same reasoning holds for $\mathbf{A}(t)\mathbf{A}(t)^\dagger$. First, notice that $\mathbf{C}$ is diagonalizable with eigenvalues $\alpha_1, \ldots, \alpha_n$, as its minimal polynomial is the minimal polynomial of ζ. [Rja94] then states that $\mathbf{B}(\mathbf{C})$ is diagonalizable if and only if the n matrices $\mathbf{B}(\alpha_k)$ are diagonalizable, in which case the spectrum (set of eigenvalues) of $\mathbf{B}(\mathbf{C})$ is the union of the spectra of the $\mathbf{B}(\alpha_k)$. By construction, for every k in $[n]$, $\mathbf{B}(\alpha_k)$ is Hermitian and therefore diagonalizable. Since the eigenvalues of $\mathbf{B}(\alpha_k)$ (resp. $\mathbf{B}(\mathbf{C})$) are the square singular values of $\mathbf{A}(\alpha_k)$ (resp. $\mathbf{A}(\mathbf{C})$), we directly get that

$$S(\mathbf{A}(\mathbf{C})) = \bigcup_{k \in [n]} S(\mathbf{A}(\alpha_k)),$$

which proves the first equality.

For the third equality, recall that $M_{\sigma_H}(\mathbf{A}) = (\mathbf{I}_d \otimes \mathbf{U}_H^\dagger) M_\sigma(\mathbf{A}) (\mathbf{I}_m \otimes \mathbf{U}_H)$. Since $\mathbf{U}_H$ is unitary, we have $S(M_{\sigma_H}(\mathbf{A})) = S(M_\sigma(\mathbf{A}))$. We now prove the second equality. Recall that $M_\sigma(\mathbf{A})$ is the block matrix of size $nd \times nm$ whose block $(i,j) \in [d] \times [m]$ is $\mathrm{diag}(\sigma(a_{ij}))$. The matrix can therefore be seen as a $d \times m$ matrix with blocks of size $n \times n$. The idea is now to permute the rows and columns of $M_\sigma(\mathbf{A})$ to end up with a matrix of size $n \times n$ with blocks of size $d \times m$ only on the diagonal. For that, we define the following permutation π_k of $[nk]$ for any positive integer k. For all $i \in [nk]$, write $i - 1 = k_1^{(i)} + nk_2^{(i)}$, with $k_1^{(i)} \in \{0, \ldots, n-1\}$ and $k_2^{(i)} \in \{0, \ldots, k-1\}$. Then, define $\pi_k(i) = 1 + k_2^{(i)} + k \cdot k_1^{(i)}$. This is a well-defined permutation based on the uniqueness of the Euclidean division. We can then define the associated permutation matrix $\mathbf{P}_{\pi_k} = [\delta_{i,\pi_k(j)}]_{(i,j) \in [nk]^2} \in \mathbb{R}^{nk \times nk}$. Then, by defining $\mathbf{P}_{\pi_d}$ and $\mathbf{P}_{\pi_m}$ as described, it holds that

$$\mathbf{P}_{\pi_d} M_\sigma(\mathbf{A}) \mathbf{P}_{\pi_m}^T = \begin{bmatrix} \sigma_1(\mathbf{A}) & & \\ & \ddots & \\ & & \sigma_n(\mathbf{A}) \end{bmatrix}.$$

Since $\mathbf{P}_{\pi_d}, \mathbf{P}_{\pi_m}$ are permutation matrices, they are also unitary and therefore $S(M_\sigma(\mathbf{A})) = S(\mathbf{P}_{\pi_d} M_\sigma(\mathbf{A}) \mathbf{P}_{\pi_m}^T)$. As $\mathbf{P}_{\pi_d} M_\sigma(\mathbf{A}) \mathbf{P}_{\pi_m}^T$ is block-diagonal, it directly holds that $S(\mathbf{P}_{\pi_d} M_\sigma(\mathbf{A}) \mathbf{P}_{\pi_m}^T) = \cup_{k \in [n]} S(\sigma_k(\mathbf{A}))$, thus proving the second equality.

Finally, by taking the maximum of the sets involved in the first equality, we obtain $\|M_\tau(\mathbf{A})\|_2 = \max_{k \in [n]} \|\sigma_k(\mathbf{A})\|_2$ as claimed. □

Lemma 2.12.

Proof. First, we derive the Gaussian tail bound for a single element a. Notice that $\|M_{\sigma_H}(a)\|_2 = \|M_\sigma(a)\|_2 = \|\mathrm{diag}(\sigma(a))\|_2 = \|\sigma(a)\|_\infty$. Let $a \in \mathcal{I}$ be sampled from $\mathcal{D}_{\mathcal{I},s}$. Then $\sigma_H(a)$ is distributed according to $\mathcal{D}_{\Lambda,\gamma}$ where $\Lambda = \sigma_H(\mathcal{I})$. So $\|\sigma(a)\|_\infty = \|\mathbf{U}_H\sigma_H(a)\|_\infty \le \|\sigma_H(a)\|_\infty$. We briefly explain the last inequality. For clarity, we define $\mathbf{a} = \sigma_H(a)$. By decomposing $\mathbf{a} = [\mathbf{a}_1^T|\mathbf{a}_2^T|\tilde{\mathbf{a}}_2^T]^T$, with $\mathbf{a}_1 \in \mathbb{R}^{t_1}$ and $\mathbf{a}_2, \tilde{\mathbf{a}}_2 \in \mathbb{R}^{t_2}$, a standard calculation gives

$$\mathbf{U}_H\mathbf{a} = \frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{2}\mathbf{a}_1\\ \mathbf{a}_2 - i\tilde{\mathbf{a}}_2\\ \mathbf{a}_2 + i\tilde{\mathbf{a}}_2\end{bmatrix}.$$

Thus, $\|\mathbf{U}_H\mathbf{a}\|_\infty = \max\{\|\mathbf{a}_1\|_\infty, \|\mathbf{a}_2 + i\tilde{\mathbf{a}}_2\|_\infty/\sqrt{2}\}$. Yet $\|\mathbf{a}_1\|_\infty \le \|\mathbf{a}\|_\infty$, and for all $k \in [t_2]$, $|a_{2,k} + i\tilde{a}_{2,k}|/\sqrt{2} = \sqrt{a_{2,k}^2 + \tilde{a}_{2,k}^2}/\sqrt{2} \le \|\mathbf{a}\|_\infty$. Hence $\|\mathbf{U}_H\mathbf{a}\|_\infty \le \|\mathbf{a}\|_\infty$. By the second part of [Pei08, Cor. 5.3] for $m = 1$, $\mathbf{z} = 1$ and $\mathbf{c} = \mathbf{0}$, it holds that for all $t \ge 0$

$$\mathbb{P}_{\mathbf{a}\hookleftarrow\mathcal{D}_{\Lambda,\gamma}}[\|\mathbf{a}\|_\infty \ge st] \le 2n \cdot e^{-\pi t^2}.$$

Note that in the case where $\mathbf{c} = \mathbf{0}$, the restriction of $s \ge \eta_\varepsilon(\Lambda)$ for some $\varepsilon \le 1/(2m+1)$ is not necessary, and the calculation of the bound on the probability saves a factor of e for that reason. With the observation that $\|\sigma(a)\|_\infty \le \|\sigma_H(a)\|_\infty$ it holds

$$\mathbb{P}_{a\hookleftarrow\mathcal{D}_{\mathcal{I},s}}[\|\sigma(a)\|_\infty \le st] \ge \mathbb{P}_{a\hookleftarrow\mathcal{D}_{\mathcal{I},s}}[\|\sigma_H(a)\|_\infty \le st] \ge 1 - 2n \cdot e^{-\pi t^2}.$$

Now let $\mathbf{N}$ be sampled from $\mathcal{D}_{\mathcal{I},s}^{m\times d}$. Fix any vector $\mathbf{x} = [\mathbf{x}_1^T, \ldots, \mathbf{x}_d^T]^T \in \mathbb{C}^{nd}$, where each $\mathbf{x}_i \in \mathbb{C}^n$. It holds that $\|M_\sigma(\mathbf{N})\mathbf{x}\|_2^2 = \sum_{i\in[m]} \|\sum_{j\in[d]} M_\sigma(n_{i,j})\mathbf{x}_j\|_2^2$. Yet, for each $i \in [m]$, we have

$$\begin{aligned}\left\|\sum_{j\in[d]} M_\sigma(n_{i,j})\mathbf{x}_j\right\|_2 \le \sum_{j\in[d]} \|M_\sigma(n_{i,j})\|_2\|\mathbf{x}_j\|_2 &\le \sqrt{\sum_{j\in[d]} \|M_\sigma(n_{i,j})\|_2^2}\sqrt{\sum_{j\in[d]} \|\mathbf{x}_j\|_2^2}\\ &= \sqrt{\sum_{j\in[d]} \|M_\sigma(n_{i,j})\|_2^2}\|\mathbf{x}\|_2.\end{aligned}$$

Using the tail bound that we previously derived, a union bound on $(i,j) \in [m]\times[d]$ yields the claim. □

Lemma 2.13.

Proof. We simply use the definition of the multiplication matrix which yields that $\sigma_H(\mathbf{y}) = M_{\sigma_H}(\mathbf{U})\sigma_H(\mathbf{e})$. Then, since $\sigma_H(\mathbf{e})$ is distributed according to $D_{\sqrt{\mathbf{S}}}$, a standard fact on multi-dimensional Gaussian distributions gives that $\sigma_H(\mathbf{y})$ is Gaussian with covariance matrix $M_{\sigma_H}(\mathbf{U})\mathbf{S}M_{\sigma_H}(\mathbf{U})^T = \mathbf{\Sigma}$. □

Lemma 2.14. We need a result on the sum of independent Gaussian distributions. We therefore 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 2.14.

Lemma A.1 (Adapted from [LS15, Lem. 2.8] & [Reg09, Claim 3.9]). *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 (of Lemma A.1). The density function Y is given by

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

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. □

We also need another lemma related to the inner product of $K_{\mathbb{R}}^d$ (which results in an element of $K_{\mathbb{R}}$) between a discrete Gaussian vector and an arbitrary one. In particular, we use Lemma 2.14 in the proof of Lemma 3.5 in order to decompose a Gaussian noise into an inner product. It generalizes [Reg09, Cor. 3.10] to the module case. A specific instance is proven in the proof of [LS15, Lem. 4.15], which is later mentioned (without proof) in [RSW18, Lem. 5.5].

Lemma A.2 ([LS15, Lem. 2.13]). *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 $r'_j = r_j\sqrt{\sum_{i\in[d]}|\sigma_j(z_i)|^2}$.*

Proof (of Lemma 2.14). 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 A.2, $\langle \mathbf{z}, \mathbf{h}\rangle$ 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 A.1 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 A.2 and noticing that the statistical distance does not increase when applying a function (here the inner 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 $j \in [n]$. □

A.2 Missing Proofs of Section 3

Lemma 3.4.

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

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

where $\mathbf{A} \hookleftarrow U(R_q^{m\times k})$, 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^m, and $\mathbf{b}_j = q^{-1}\mathbf{A}\mathbf{s}_j + \mathbf{e}_j \bmod R^\vee$ for $\mathbf{s}_j$ i.i.d. from $U((R_q^\vee)^k)$ for every $j \in [\ell]$. By definition, we have $\mathrm{Adv}[\mathcal{O}] = |\mathbb{P}[\mathcal{O}(\mathcal{H}_\ell) = 1] - \mathbb{P}[\mathcal{O}(\mathcal{H}_0) = 1]|$. The reduction $\mathcal{A}$ works as follows.

1. Sample $\mathbf{z}$ from $U(\mathcal{Z})$ and get $(\mathbf{A}, \mathbf{b}, \langle \mathbf{e},\mathbf{z}\rangle)$ as input of ext-M-LWE$^1_{n,k,m,q,\psi,\mathcal{Z}}$.
2. Sample i^* from $U([\ell])$.
3. Sample $\mathbf{s}_1,\dots,\mathbf{s}_{i^*-1}$ from $U((R_q^\vee)^k)$, $\mathbf{e}_1,\dots,\mathbf{e}_{i^*-1},\mathbf{e}_{i^*+1},\dots,\mathbf{e}_\ell$ from ψ^m and finally $\mathbf{u}_{i^*+1},\dots,\mathbf{u}_\ell$ from $U((q^{-1}R^\vee/R^\vee)^m)$.

4. Compute $\mathbf{b}_j = q^{-1}\mathbf{A}\mathbf{s}_j + \mathbf{e}_j \bmod R^\vee$ for all $j \in [i^* - 1]$.
5. Define the hybrid matrix $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_{i^*-1}, \mathbf{b}, \mathbf{u}_{i^*+1}, \ldots, \mathbf{u}_\ell]$, and the error matrix $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_{i^*-1}, \mathbf{e}, \mathbf{e}_{i^*+1}, \ldots, \mathbf{e}_\ell]$. Then call the oracle $\mathcal{O}$ on input $(\mathbf{A}, \mathbf{B}, \mathbf{E}^T\mathbf{z})$, 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 $\mathcal{H}_\ell$. So the output is correct over the randomness of i^*. Since i^* is uniformly chosen we have

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

□

A.3 Missing Proofs of Section 4

Lemma 4.1.

Proof. We first describe the transformation T given in [MM11] going from M-LWE to M-SIS. Given $(\mathbf{A}, \mathbf{b}) \in R_q^{m \times d} \times (R^\vee)^m$, where $\mathbf{A}$ is uniformly sampled, T first checks if the rows of $\mathbf{A}$ generate R_q^d. If not, T returns $\perp$. By Equation (1), T aborts at this step with probability $\delta(m, d)$. We now condition on $\mathbf{A}$ being non-singular. From $\mathbf{A}$, T computes $\mathbf{B} \in R_q^{m \times (m-d)}$ whose columns generate the set of vectors $\mathbf{x} \in R_q^m$ that verify $\mathbf{A}^T\mathbf{x} = \mathbf{0} \bmod qR$. T samples $\mathbf{U} \in R_q^{(m-d)\times(m-d)}$ uniformly at random such that $\mathbf{U}$ is invertible in R_q, and define $\mathbf{B}' = \mathbf{B}\mathbf{U}$. As $\mathbf{A}$ is uniform in the set of non-singular matrices, $\mathbf{B}'$ is uniform in the set of matrices whose rows generate R_q^{m-d}. Thus, by Equation (1), $\Delta(\mathbf{B}', U(R_q^{m\times(m-d)})) \leq \delta(m, m-d)$. Finally, T computes $\mathbf{c} = \widetilde{\mathbf{B}}^T\mathbf{b} \bmod qR^\vee$, and returns $(\mathbf{B}', \mathbf{c})$.

Assume that there exists an adversary $\mathcal{A}$ that attacks the ε'-uninvertibility of M-SIS. We construct $\mathcal{B}$ that breaks the ε-univertibility of M-LWE by calling $\mathcal{A}$ on the sampled transformed by T. Consider $(\mathbf{A}, \mathbf{A}\mathbf{s}+\mathbf{e} \bmod qR^\vee)$, with $(\mathbf{s}, \mathbf{e}) \hookleftarrow U((R_q^\vee)^d) \times \mathcal{X}$. We denote E the event $\{\mathcal{B}(\mathbf{A}, \mathbf{A}\mathbf{s} + \mathbf{e} + qR^\vee) = (\mathbf{s}, \mathbf{e})\}$. Then

$$\begin{aligned}
\mathbb{P}[E] &= \mathbb{P}[\mathbf{A} \text{ non-singular}]\mathbb{P}[E|\mathbf{A} \text{ non-singular}] + \mathbb{P}[\mathbf{A} \text{ singular}] \underbrace{\mathbb{P}[E|\mathbf{A} \text{ singular}]}_{0 \text{ (abort)}} \\
&= (1 - \delta(m, d))\mathbb{P}[\mathcal{A}(\mathbf{B}', \mathbf{c}) = \mathbf{e}|\mathbf{A} \text{ non-singular}] \\
&> (1 - \delta(m, d)) \cdot (\varepsilon' - \delta(m, m-d)) \\
&= \varepsilon.
\end{aligned}$$

Indeed, by the transformation, we have

$$\begin{aligned}(\mathbf{B}')^T\mathbf{b} \bmod qR^\vee &= (\mathbf{B}')^T\mathbf{A}\mathbf{s} + (\mathbf{B}')^T\mathbf{e} \bmod qR^\vee \\ &= (\mathbf{A}^T\mathbf{B}' \bmod qR)^T\mathbf{s} + (\mathbf{B}')^T\mathbf{e} \bmod qR^\vee \\ &= (\mathbf{B}')^T\mathbf{e} \bmod qR^\vee.\end{aligned}$$

Then, $\mathcal{B}$ uses linear algebra to recover $\mathbf{s}$ from $\mathbf{b} - \mathbf{e}$. The proof for one-wayness is the same where $E = \{g_{\mathbf{A}}(\mathcal{B}(\mathbf{A}, \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR^\vee)) = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR^\vee\}$ (with $g_{\mathbf{A}}(\mathbf{s}, \mathbf{e}) = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR^\vee$). For the pseudorandomness, we define $E = \{\mathcal{B}(\mathbf{A}, \mathbf{b} \text{ uniform}) = 1\}$, $E' = \{\mathcal{B}(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR^\vee) = 1\}$, and F the event $\{\mathbf{A} \text{ non singular}\}$. It then holds that

$$\begin{aligned}&|\mathbb{P}[E] - \mathbb{P}[E']| \\ &= \mathbb{P}[\mathbf{A} \text{ non-singular}] \cdot |\mathbb{P}[E|\mathbf{A} \text{ non-singular}] - \mathbb{P}[E'|\mathbf{A} \text{ non singular}]| \\ &= (1 - \delta(m, d))\big|\mathbb{P}[\mathcal{A}(\mathbf{B}', \mathbf{c} \text{ uniform}) = 1|F] - \mathbb{P}[\mathcal{A}(\mathbf{B}', (\mathbf{B}')^T\mathbf{e} \bmod qR^\vee) = 1|F]\big| \\ &> (1 - \delta(m, d)) \cdot (\varepsilon' - \delta(m, m - d)) \\ &= \varepsilon,\end{aligned}$$

concluding the proof. □

Lemma 4.2.

Proof. The transformation T now works as follows. Given $(\mathbf{B}, \mathbf{c}) \in R_q^{m\times(m-d)} \times (R^\vee)^{m-d}$ with $\mathbf{B}$ uniformly distributed, T checks whether the rows of $\mathbf{B}$ generate R_q^{m-d}. If not, it aborts, and that with probability $\delta(m, m-d)$. Conditioning on $\mathbf{B}$ being non-singular, T computes $\mathbf{A} \in R_q^{m\times d}$ that generate $\{\mathbf{x} \in R_q^m : \mathbf{B}^T\mathbf{x} = \mathbf{0} \bmod qR\}$. The transformation then randomizes $\mathbf{A}$ by a random matrix $\mathbf{U} \in R_q^{d\times d}$ that is invertible in R_q to obtain $\mathbf{A}' = \mathbf{A}\mathbf{U}$. Similarly as in the previous proof, $\Delta(\mathbf{A}', U(R_q^{m\times d})) \leq \delta(m, d)$. Then, T finds a vector $\mathbf{b}$ such that $\mathbf{B}^T\mathbf{b} = \mathbf{c} \bmod qR^\vee$, and returns $(\mathbf{A}', \mathbf{b})$. Note that if $\mathbf{c} = \mathbf{B}^T\mathbf{e} \bmod qR^\vee$ for some $\mathbf{e} \hookleftarrow \mathcal{X}$, then $\mathbf{b} - \mathbf{e}$ is in the span of the columns of $\mathbf{A}'$ and therefore, there exists a uniform $\mathbf{s} \in (R_q^\vee)^d$ such that $\mathbf{b} - \mathbf{e} = \mathbf{A}'\mathbf{s} \bmod qR^\vee$. If $\mathbf{c}$ is uniform, we can argue that $\mathbf{b}$ is also uniform. Using the same calculations as before, we get that

$$\mathrm{Adv}[\mathcal{B}] > (1 - \delta(m, m - d)) \cdot (\varepsilon' - \delta(m, d)) = \varepsilon,$$

where $\mathrm{Adv}[\mathcal{B}]$ denotes the probability of breaking uninvertibility or one-wayness, or the absolute difference of probability in the case of pseudorandomness. □

Lemma 4.5.

Proof. Consider the distribution $\mathcal{D}$ supported over $R_q^d \times R_q^\vee$ that is either $A_{\mathbf{s},\psi}$ or $U(R_q^d \times R_q^\vee)$.
Construction: Sample independently $((\mathbf{a}_i, b_i))_{i\in[m']}$ from $\mathcal{D}$. In both cases, the first component is uniformly distributed over R_q^d. If there is no subset $S \subseteq [m']$

of size d such that the $(\mathbf{a}_i)_{i\in S}$ are R_q-linearly independent, the reduction aborts. This is equivalent to having at most $d-1$ vectors that are linearly independent within the m' available vectors. As discussed in Section 2.1, this happens with probability $\delta(m', d)$. So now, we assume that there exists a set $S \subseteq [m']$ of size d such that the $(\mathbf{a}_i)_{i\in S}$ are R_q-linearly independent. Consider the matrix $\overline{\mathbf{A}} \in R_q^{d\times d}$ whose rows are the $(\mathbf{a}_i^T)_{i\in S}$, and $\overline{\mathbf{b}} \in \mathcal{I}_q^d$ whose coefficients are the $(b_i)_{i\in S}$. By construction, $\overline{\mathbf{A}}$ is invertible in $R_q^{d\times d}$. Additionally, if $\mathcal{D} = A_{\mathbf{s},\psi}$, then $\overline{\mathbf{b}} = \overline{\mathbf{A}}\mathbf{s}+\mathbf{x} \mod qR^\vee$ for $\mathbf{x}$ sampled from ψ^d. On the other hand, if $\mathcal{D} = U(R_q^d \times R_q^\vee)$, then $\overline{\mathbf{b}}$ is uniform over $(R_q^\vee)^d$.
Reduction: The transformation T works as follows. Given $(\mathbf{a}, b)$ sampled from $\mathcal{D}$ as input:

- Compute $\mathbf{a}' = -(\overline{\mathbf{A}})^{-T} \cdot \mathbf{a} \mod qR$;
- Compute $b' = b + \langle \mathbf{a}', \overline{\mathbf{b}} \rangle \mod qR^\vee$;
- Output $(\mathbf{a}', b')$.

First, we verify that $(\mathbf{a}', b')$ indeed belongs to $R_q^d \times R_q^\vee$. Since $\overline{\mathbf{A}}$ is invertible modulo qR, then $-(\overline{\mathbf{A}})^{-T}$ is in $R_q^{d\times d}$. Therefore, $\mathbf{a}'$ is also in R_q^d. Additionally, as $\overline{\mathbf{b}} \in (R_q^\vee)^d$, $\langle \mathbf{a}', \mathbf{b} \rangle$ is in $R^\vee$. It thus holds that b' is in $R_q^\vee$.

As $-(\overline{\mathbf{A}})^{-T}$ is invertible modulo qR, and $\mathbf{a}$ is uniform in R_q^d, then a' is also uniform in R_q^d. Now, we look at the distribution of b' in both cases. First, assume that $\mathcal{D} = A_{\mathbf{s},\psi}$. Then $b = \langle \mathbf{a}, \mathbf{s} \rangle + e \mod qR^\vee$ for some $e \hookleftarrow \psi$, and $\overline{\mathbf{b}} = \overline{\mathbf{A}}\mathbf{s} + \mathbf{x} \mod qR^\vee$. It holds that

$$\begin{aligned} b' &= \langle \mathbf{a}, \mathbf{s} \rangle + e + \langle \mathbf{a}', \overline{\mathbf{A}}\mathbf{s} + \mathbf{x} \rangle \mod qR^\vee \\ &= \langle \mathbf{a} + \overline{\mathbf{A}}^T \mathbf{a}', \mathbf{s} \rangle + \langle \mathbf{a}', \mathbf{x} \rangle + e \mod qR^\vee \\ &= \langle \mathbf{a}', \mathbf{x} \rangle + e \mod qR^\vee. \end{aligned}$$

So $(\mathbf{a}', b')$ is indeed distributed according to $A_{\mathbf{x},\psi}$ for $\mathbf{x} \hookleftarrow \psi^d$ as desired. Now assume that $\mathcal{D} = U(R_q^d \times R_q^\vee)$. Then b is uniform over $R_q^\vee$ and $\overline{\mathbf{b}}$ is uniform over $(R_q^\vee)^d$. So b' is clearly uniform over $R_q^\vee$ as well, proving that $(\mathbf{a}', b')$ is uniformly distributed over $R_q^d \times R_q^\vee$ as desired. □