On the Hardness of Module Learning with Errors with Short Distributions

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

On the Hardness of Module Learning With Errors with Short Distributions

Katharina Boudgoust[1], Corentin Jeudy[2,3], Adeline Roux-Langlois[4], and Weiqiang Wen[5]

katharina.boudgoust@cs.au.dk, corentin.jeudy@irisa.fr, adeline.roux-langlois@cnrs.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] Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France
[5] LTCI, Telecom Paris, Institut Polytechnique de Paris, Paris, France

Abstract. The *Module Learning With Errors* (M-LWE) problem 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 uniform η-bounded secret for any $1 \leq \eta \ll q$ 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 search M-LWE with uniform η-bounded secret *and* error under specific parameter conditions. Overall, the results apply to all cyclotomic fields, but most of the intermediate results are proven in more general number fields.

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

This paper contains novel results and generalizations of existing ones already published in [12,13].

1 Introduction

The *Learning With Errors* (LWE) problem, introduced by Regev [46], 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 $\mathbb{R}$. 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 $\mathrm{LWE}_{d,m,q,\psi}$ samples as $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod \mathbb{Z})$, 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 [46,47] 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 [42], then generalized to any polynomial modulus q by Brakerski et al. [17], 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 [51,30]. In this paper, we focus on the *Module Learning With Errors* (M-LWE) problem, first defined by Brakerski et al. [16] and then thoroughly studied by Langlois and Stehlé [24]. 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. The integer d now denotes the module rank, and q still denotes the modulus. Further, let ψ be a distribution on the field tensor product $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$, and let $\mathbf{s} \in R_q^d$ be a secret vector, where $R_q = R/qR$. An $\mathrm{M\text{-}LWE}_{n,d,q,\psi}$ sample is given by $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s} \rangle + e \bmod R)$, 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 $R_q^d \times \mathbb{T}$, where the torus is $\mathbb{T} = K_\mathbb{R}/R$. We can also use a matrix formula-

tion when the number of samples m is fixed by considering the M-LWE$_{n,d,m,q,\psi}$ distribution $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod R)$ 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) [30]. Just like LWE, the M-LWE problem enjoys worst-case to average-case connections from lattice problems such as SIVP$_\gamma$ [24]. 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 [40], several finalist candidates relying on the hardness of M-LWE have recently been chosen for standardization, e.g., the signature scheme Dilithium [18] and the key encapsulation mechanism Kyber [10]. 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 $\mathbf{s}$ or a discrete error distribution ψ to be over $\{0, \ldots, \eta - 1\}$ (or $\{-\eta, \ldots, \eta\}$) instead of $\mathbb{Z}_q$. Besides gaining in efficiency, choosing a small secret plays an important role in some applications like fully homomorphic encryption [19] or modulus switching techniques [17,1,56] as it keeps the noise blowup to a minimum. The LWE problem with uniform bounded secret has been well studied in the case of binary secrets (i.e., secrets in $\{0, 1\}^d$), denoted by bin-LWE, but the different approaches easily generalize to slightly larger secrets. A first study of bin-LWE was provided by Goldwasser et al. [21] 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 LWE$_{k,q,D_\alpha}$ to bin-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. [17] and Micciancio [35] using more technical proofs. Both of them achieve a similar dimension increase between k and d, but only increase the error by roughly $\beta/\alpha = \Omega(\sqrt{d})$ (polynomial). 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 [14] extends the hardness results to more general secret distributions based on entropic arguments.

The hardness of LWE with error uniformly distributed below η with $\eta \ll q$ was first studied by Micciancio and Peikert [38]. They proved that the LWE

function $(\mathbf{s}, \mathbf{e}) \mapsto \mathbf{As} + \mathbf{e} \bmod q$ is one-way with respect to inputs $\mathbf{e}$ 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. [52], 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 [38] 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 secrets was extended to the R-LWE case by Brakerski and Döttling [15], and the module case by Lin et al. [26][1]. However, no results on the hardness of M-LWE with small uniform error were known, even though they serve as hardness assumptions for most efficient M-LWE-based schemes. For example, the signature scheme Dilithium [18] 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 some positive integer η. The first two contributions study the hardness of the M-LWE problem with centered η-bounded secret, which we denote by η-M-LWE, in both its search and decision versions, for any $\eta \geq 1$. They are generalizations of the results published in our previous conference papers [12] and [13] respectively, only dealing with the special case bin-M-LWE with secret coefficients in $\{0, 1\}$, which is already mentioned in one of the author's thesis [11]. As opposed to [12,13], we decide to work in the primal ring R and with a centered representation of secrets with coefficients in $\{-\eta, \ldots, \eta\}$[2] to show that the proofs still work and also to match practical uses of the M-LWE assumption like [10,18]. 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 search M-LWE with small secret *and* error. To the best of our knowledge, it is the first result for the hardness of M-LWE with small uniform error. The results apply to all cyclotomic fields, but most of the intermediate results are proven in more general number fields.

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. [21] in the case of LWE, while achieving much better noise parameters by

[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.

[2] Setting $\eta = 1$ gives ternary secrets instead of binary. We however observe that the parameters covered by the reductions for $\eta = 1$ in the centered representation match those of [12,13], and it has the upside of a larger secret space. This leads to smaller ranks d by a factor of $\log_2 3$.

using the Rényi divergence instead of the statistical distance to measure the distance between two distributions. The improvement on the noise rate compared to [21] 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 (over rings) 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, its use asks to fix the number of samples a priori.

The result 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 needs to be prime. The ranks must satisfy $d \log_2(2\eta + 1) \geq k \log_2 q + \Omega(\log_2 n)$, which is due to the use of the leftover hash lemma over rings. The Gaussian noise parameter α is also increased to β by a factor $\beta/\alpha = d\sqrt{m} \cdot n^{3/2}\eta \log_2(n)$ 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 [17] by introducing the two intermediate problems first-is-errorless M-LWE and ext-M-LWE. We start by reducing 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 vectors. Two other formulations of ext-M-LWE were proposed by Alperin-Sheriff and Apon [4], and more recently by Lyubashevsky et al. [33], 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 [17, 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 explicitly describe than the units of $\mathbb{Z}/q\mathbb{Z}$ in the sense that we do not have practical closed-form expressions. This is the reason why we need to control the splitting structure of the cyclotomic polynomial modulo q.

Lemma 2.4 [32, 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 [24, 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 $d\log_2(2\eta+1) \geq (k+1)\log_2 q + \omega(\log_2 n)$, in the same manner as in Contribution 1, except that the asymptotic term is now super-logarithmic. The noise rate is now given by $n\eta\sqrt{2d}\sqrt{4n^2\eta^2+1} = \Theta(\eta^2 n^2\sqrt{d})$ 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 $m = \Theta(n\log_2 n)$. In the special case of $\eta = 1$ and $n = 1$, we recover the same noise-ratio $\Theta(\sqrt{d})$ as in the original LWE result from Brakerski et al. [17].

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 4.1. It uses a different proof method from Contributions 1 and 2 by following the idea of Micciancio and Peikert [38] of proving the one-wayness of the M-LWE function $(\mathbf{s}, \mathbf{e}) \mapsto \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR$, with $\mathbf{e}$ uniform in S_η^m (i.e., η-bounded vectors over R of dimension m). To do so, we prove the one-wayness of the M-ISIS function $\mathbf{e} \mapsto (\mathbf{A}')^T\mathbf{e} \bmod qR$ and use the duality between both functions to conclude. This function is inspired from the *Module Short Integer Solution* (M-SIS) problem [24] which asks to find a short non-zero vector $\mathbf{e} \in R^m$ such that $(\mathbf{A}')^T\mathbf{e} = \mathbf{0} \bmod qR$ 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 [38] 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 η and the number of samples m. 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. 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. The reduction also gives a condition on the maximal number of samples m we can provide with such small uniform error. In particular, we have $m \leq d(1 + o(\log_2 \eta))$ which is similar to what is obtained in [38]. Then, 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 $m' + m''$ samples for $m' \geq d$. 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. 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 [6] 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 $\eta = 1$ (ternary), then $m \approx d^3$ 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 over rings. 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 suite [10,18] chosen for standardization by NIST.

The hardness proof of η-M-LWE with η-bounded error and m samples currently requires the hardness of M-LWE with η-bounded error and $m+d$ samples. Although there is no subexponential attack for $m = d, d+1$ (even for ternary errors $\eta = 1$) yet, our proof does not encompass this range of parameters. We leave it as a major open problem to prove the hardness of η-M-LWE with η-bounded error for $m \gtrsim d$.

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 [36] 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 [12,13]. 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 an odd prime integer, and $\mathbb{Z}_q$ denotes the ring of integers modulo q. We may occasionally identify $\mathbb{Z}_q$ with the centered set of representatives $\{-(q-1)/2, \ldots, (q-1)/2\}$. 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. $\dagger$). 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 $\mathbf{I}_n$. We use the standard Landau notations, i.e., $O(\cdot), o(\cdot), \omega(\cdot), \Omega(\cdot), \Theta(\cdot)$, and we say a function ε is negligible in n if $\varepsilon(n) = n^{-\omega(1)}$. We also say that a probability p is overwhelming if $1 - p$ is negligible in n. We use the abbreviation PPT for *probabilistic polynomial-time*.

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 field tensor product by $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R}$. 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[3]. 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. In this paper, we assume that for the number fields we consider, the ring of integers is efficiently computable and has a good basis representation. In particular, it is the case for cyclotomic fields.

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

[3] We may sometimes call this class of fields "monogenic fields", but we note that rigorously a monogenic number field is $K = \mathbb{Q}(\zeta)$ for which $R = \mathbb{Z}[\zeta']$ for a possibly different ζ'.

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 $S_\eta = \tau^{-1}(\{-\eta, \ldots, \eta\}^n)$, which can be seen as the set of polynomials with coefficients in $\{-\eta, \ldots, \eta\}$. 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 ring 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 then define the field norm of K as $N(x) = \prod_{k \in [n]} \sigma_k(x) \in \mathbb{Q}$ for all $x \in K$.

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. [50] and Blanco-Chacón [8] give

additional insight on this distortion for specific number fields. Throughout this paper, we are interested in the parameter defined by $B_\eta = \max_{x \in S_\eta} \|\sigma(x)\|_\infty$ for a positive integer η. This parameter is inherent to the ring and intervenes in the proof of Lemma 3.2, 3.5 and 4.4. Here, we provide an upper-bound on B_η, that is further simplified for cyclotomic number fields. The proof is provided in Appendix B.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 $1 \leq B_\eta = \max_{x \in S_\eta} \|\sigma(x)\|_\infty \leq n\eta \|\mathbf{V}\|_{\max}$. In particular, for cyclotomic fields, it yields $1 \leq B_\eta \leq n\eta$.*

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 defer the proof to Appendix B.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 [49] 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 B.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 $\mathfrak{p}$. In R, an ideal $\mathfrak{p}$ is prime if and only if it is maximal, implying that $R/\mathfrak{p}$ is a field. 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 $\prod_{i\in[\kappa]} \mathfrak{p}_i$. We say that q is fully splitted in R if $\kappa = n$ in the above factorization, where n is the degree of the number field. We also say that q is inert in R if $\langle q\rangle$ is prime. 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 $a \in R$.

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 [32, Th. 1.1].

Lemma 2.4 ([32, 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 by 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, for any element y of R satisfying $0 < \|\tau(y)\|_\infty < q^{1/\varphi(\mu)}/\mathfrak{s}_1(\mu)$, it holds that $y \bmod qR$ 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.*

The number theoretic conditions on q essentially say that $\langle q\rangle$ splits into $\varphi(\mu)$ distinct prime ideal factors, each of algebraic norm $q^{\varphi(\nu)/\varphi(\mu)} = q^{\nu/\mu}$. In the case where ν is a power of an odd prime, then so is μ and then [30] states that $\mathfrak{s}_1(\mu) = \sqrt{\mu}$. For more general cases, we refer to the discussions from Lyubashevsky and Seiler [32, Conj. 2.6]. We also refer the reader to [32, Th. 2.5] that discusses the existence of such primes q for specific values of ν and μ.

Remark 2.1 (Instantiation in power-of-two cyclotomics). This result is simplified in the power-of-two case [32, 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 $n = 2^\ell$, the only conditions on q are that q 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 $q = 5 \bmod 8$, meaning the ideal qR splits into two prime ideal factors of norm $q^{n/2}$. In our context in Section 3.2.2, having $q^{1/2}/\sqrt{2} > \eta$ is sufficient as our elements have η-bounded coefficients. For the ternary secret case $\eta = 1$, this leads to $q > 2$, which is subsumed by $q = 5 \bmod 8$.

We say that $\mathbf{a}_1, \ldots, \mathbf{a}_\ell \in R_q^d$ are R_q-linearly independent if and only if for all $(\lambda_i)_{i\in[\ell]} \in R_q^\ell$, $\sum_{i\in[\ell]} \lambda_i \mathbf{a}_i = 0$ implies $\lambda_i = 0$ in R_q for all $i \in [\ell]$. We now give results on the linear independence of uniform vectors of R_q^d and the singularity of uniformly random matrices. A result similar to Lemma 2.5 is present in [56] but the proof unfortunately relies on a flawed argument. We detail this matter in Appendix A and provide corrected proofs.

Lemma 2.5. *Let K be a number field, and R its ring of integers. Let d, q be positive integers such that q is an unramified prime which factors as $\langle q\rangle = \prod_{i\in[\kappa]} \mathfrak{p}_i$. Let ℓ be in $\{0, \ldots, d-1\}$, and $\mathbf{a}_1, \ldots, \mathbf{a}_\ell \in R_q^d$ be R_q-linearly independent vectors of R_q^d. Then*

$$\mathbb{P}_{\mathbf{b}\hookleftarrow U(R_q^d)}[\mathbf{a}_1, \ldots, \mathbf{a}_\ell, \mathbf{b} \text{ are } R_q\text{-linearly independent}] = \prod_{i\in[\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right).$$

As a result, for any $1 \le k \le d$, it holds that

$$\mathbb{P}_{(\mathbf{a}_i)_i \sim U(R_q^d)^k}[(\mathbf{a}_i)_{i\in[k]} \text{ are } R_q\text{-linearly independent}] = \prod_{\ell=0}^{k-1} \prod_{i\in[\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right).$$

Lemma 2.6. *Let K be a number field, and R its ring of integers. Let q be a prime integer that is unramified in R which splits as $\langle q\rangle = \prod_{i\in[\kappa]} \mathfrak{p}_i$. Let $m \ge d$ be two integers. It holds*

$$\mathbb{P}_{\mathbf{A}\sim U(R_q^{d\times m})}[\mathbf{A}\cdot R_q^m = R_q^d] \ge \prod_{\ell=0}^{d-1} \prod_{i\in[\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{m-\ell}}\right).$$

When R and q are clear from the context, for $m \ge d$, we define $\delta(m,d) = 1 - \mathbb{P}_{\mathbf{A}\sim U(R_q^{d\times m})}[\mathbf{A}\cdot R_q^m = R_q^d]$, which we use extensively throughout Section 4. We note that $\delta(m,d)$ can be upper-bounded by $\frac{d\cdot\kappa}{(\min_{i\in[\kappa]} N(\mathfrak{p}_i))^{m-d+1}}$. Hence, if q splits into only high-norm ideal factors so that $\min_{i\in[\kappa]} N(\mathfrak{p}_i) \ge n^{\omega(1/(m-d+1))}$, the probability $\delta(m,d)$ becomes negligible.

We also define $\delta'(k,d)$ to be the probability that among $k \geq d$ independent uniform columns of R_q^d, there is no subset of d of those columns that are R_q-linearly independent. Formally, we define

$$\delta'(k,d) = 1 - \mathbb{P}_{(\mathbf{a}_i)_{i\in[k]} \sim U(R_q^d)^k}[\exists S \subseteq [k], |S| = d \wedge (\mathbf{a}_i)_{i\in S} \text{ are } R_q\text{-l. i.}]$$

We note that if R_q was a field, we would have $\delta(k,d) = \delta'(k,d)$. However, in the general case, $\delta(k,d) \neq \delta'(k,d)$ as a minimal spanning set of an R_q-submodule of R_q^d is not necessarily a basis of said submodule. Additionally, note that $\delta'(d,d)$ is given by Lemma 2.5 as

$$\delta'(d,d) = 1 - \prod_{\ell=0}^{d-1} \prod_{i\in[\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right).$$

The probability $\delta'(k,d)$ is discussed in Section 4.3 as it only appears in the latter.

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 $\Lambda^* = \{\mathbf{x} \in \mathrm{Span}(\Lambda) : \forall \mathbf{y} \in \Lambda, \langle \mathbf{x}, \mathbf{y}\rangle \in \mathbb{Z}\}$. Any ideal $\mathcal{I}$ of R 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).

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 $x \hookleftarrow P$, whereas the notation $x \sim P$ means that the random variable x is distributed according to P. We now define two distances for probability distributions, namely the *statistical distance* Δ, and the *Rényi divergence* [48,53] 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. [7]. In this paper, it suffices to use the Rényi divergence of order 2 denoted by RD_2.

Definition 2.1. *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 and a data processing inequality, which are essential in proving our results.

Lemma 2.7 ([25, Lem. 4.1]). *Let P, Q be two probability distributions such that $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$ and $E \subseteq \mathrm{Supp}(Q)$ be an arbitrary event. Then, we have $P(E) \leq \Delta(P,Q) + Q(E)$, and $P(E)^2 \leq \mathrm{RD}_2(P\|Q) \cdot Q(E)$. Further, for any (possibly randomized) function f, we define P^f (resp. Q^f) the distribution obtained by sampling $x \hookleftarrow P$ (resp. Q) and outputting $f(x)$. Then, it holds that $\Delta(P^f, Q^f) \leq \Delta(P,Q)$ and $\mathrm{RD}_2(P^f\|Q^f) \leq \mathrm{RD}_2(P\|Q)$.*

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 [34], which, instead of working with vectors over the finite field $\mathbb{Z}_q$, operates over the principal ideal domain $\mathbb{Z}_q[x]$ for q prime. 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 the ideals of R/qR can be characterized via the ideals of $\mathbb{Z}_q[x]$, which is needed in the proof. Further, we provide not only a bound on the statistical distance, but also on the Rényi divergence. For completeness, we give the proof in Appendix B.1.

Lemma 2.8. *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*

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

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

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

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 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 $\mathbf{\Sigma}/(2\pi)$. We extend this definition to the degenerate case, i.e., positive semi-definite, and use the same notation for convenience[4]. 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 $\mathbf{c} = \mathbf{0}$, we omit it.

[4] In the degenerate case, the probability density function cannot be defined with respect to the Lebesgue measure as $\mathbf{\Sigma}$ is not invertible. Standard facts on non-singular Gaussian distributions can however be extended to the degenerate case by using the characteristic function which always exists and equals $\varphi_{\mathbf{x}}(\mathbf{t}) = \mathbb{E}_{\mathbf{x}}[\exp(i\mathbf{t}^T\mathbf{x})] = \exp(i\mathbf{c}^T\mathbf{t} - \pi\mathbf{t}^T\mathbf{\Sigma}\mathbf{t})$. In particular, one can easily show that the sum of two independent (potentially degenerate) Gaussians of covariance $\mathbf{\Sigma}_1, \mathbf{\Sigma}_2$ is a (potentially degenerate) Gaussian of covariance $\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2$, as needed in this paper. We also

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 coset $\Lambda + \mathbf{v}$ for $\mathbf{v} \in \mathbb{R}^n$, i.e., $\mathcal{D}_{\Lambda+\mathbf{v},\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) = D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x})/D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\Lambda + \mathbf{v})$ for all $\mathbf{x} \in \Lambda + \mathbf{v}$, and where $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\Lambda + \mathbf{v}) = \sum_{\mathbf{y} \in \Lambda + \mathbf{v}} D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{y})$.

Definition 2.2 (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 α, if for all unit vector $\mathbf{u} \in \mathbb{R}^n$, and all $t \in \mathbb{R}$, we have $\mathbb{E}[\exp(2\pi t \langle \mathbf{x}, \mathbf{u} \rangle)] \leq e^{\pi \alpha^2 t^2}$.*

A standard calculation shows that the discrete Gaussian distribution $\mathcal{D}_{\Lambda,\alpha}$ is sub-Gaussian with sub-Gaussian moment α [37, Lem. 2.8], for any lattice Λ and $\alpha > 0$. In particular, we have the following spectral bound on a discrete Gaussian matrix. The result is stated in [37, Lem. 2.9] and is derived from the works of Vershynin [54] on the non-asymptotic random matrix theory.

Lemma 2.9 ([37, Lem. 2.9]). *Let ℓ, k be positive integers, and $\alpha > 0$. Let Λ be a lattice of rank ℓ. Then, there exists a universal constant $C > 0$ such that for any $t \geq 0$, it holds $\mathbb{P}_{\mathbf{Y} \sim \mathcal{D}_{\Lambda,\alpha}^k}[\|\mathbf{Y}\|_2 > C\alpha(\sqrt{\ell} + \sqrt{k} + t)] \leq 2e^{-\pi t^2}$. Empirically, $C \approx 1/\sqrt{2\pi}$.*

The *smoothing parameter* of a lattice Λ denoted by $\eta_\varepsilon(\Lambda)$ for some $\varepsilon > 0$, introduced in [39], is the smallest $\alpha > 0$ such that $\rho_{1/\alpha}(\Lambda^*) \leq 1 + \varepsilon$. It represents the smallest Gaussian parameter $\alpha > 0$ such that the discrete Gaussian $\mathcal{D}_{\Lambda,\mathbf{c},\alpha}$ behaves like a continuous Gaussian distribution. 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 [39] and shows that above the smoothing parameter, a continuous Gaussian coset is statistically close to uniform.

Lemma 2.10 ([39, Lem. 4.1]). *Let Λ be lattice of rank n, $\varepsilon > 0$, and $\alpha \geq \eta_\varepsilon(\Lambda)$. Then the distribution of the coset $\mathbf{e} + \Lambda$, where $\mathbf{e} \sim D_\alpha$, 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 2.11 ([17, Lem. 2.10] & [43, Thm. 3.1]). *Let Λ be lattice of rank n. Let $\varepsilon \in (0, 1/2]$, and $\alpha, \beta > 0$ be such that $\alpha \geq \eta_\varepsilon(\Lambda)$. Then the distribution of $\mathbf{x} + \mathbf{y}$, obtained by first sampling $\mathbf{x}$ from D_β, and then $\mathbf{y}$ sampled from $\mathcal{D}_{\Lambda - \mathbf{x},\alpha}$, is within statistical distance 8ε of $\mathcal{D}_{\Lambda,\sqrt{\alpha^2+\beta^2}}$.*

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

---

note that one can still define a density, but with respect to a degenerate measure as $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) = (\det^+(\mathbf{\Sigma}))^{-1/2} \exp(-\pi(\mathbf{x} - \mathbf{c})^T \mathbf{\Sigma}^+ (\mathbf{x} - \mathbf{c}))$, where $\mathbf{\Sigma}^+$ is the Moore-Penrose pseudo-inverse, and $\det^+$ the pseudo-determinant.

Lemma 2.12 (Adapted from [25, Lem. 4.2]). *Let Λ be lattice of rank n, $\varepsilon \in (0,1)$, $\alpha \geq \eta_\varepsilon(\Lambda)$, and $\mathbf{c}$ a vector of $\mathbb{R}^n$. Then,*

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

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 field tensor product $K_{\mathbb{R}}$ 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 [30], 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 $D_{\sigma_H(\mathbf{c}),\sqrt{\mathbf{\Sigma}}}$, for some $\mathbf{c} \in K_{\mathbb{R}}^d$ and 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 $D_{\mathbf{c},\sqrt{\mathbf{\Sigma}}}$.

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 $x \sim \mathcal{D}_{\mathcal{I},s}$. We therefore derive such a bound, which is based on [41, Cor. 5.3] proving that $\|\sigma(x)\|_\infty \leq s \log_2 n$ with overwhelming probability. The proof can be found in Appendix B.1.

Lemma 2.13. *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 $\alpha > 0$. Then, for all $t \geq 0$ it holds that*

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

Choosing $t = \log_2 n$ gives $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha \log_2(n)\sqrt{md}$ 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 B.1 for completeness.

Lemma 2.14. *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 2.15 in the proof of Lemma 3.5 in order to decompose a

Gaussian noise into an inner product. It generalizes [47, Cor. 3.10] to the module case. A specific instance is proven in the proof of [24, Lem. 4.15], which is later mentioned (without proof) in [50, Lem. 5.5]. We defer the proof in Appendix B.1.

Lemma 2.15 (Adapted from [47, 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. In this paper, we only deal with functions that have an unambiguous and public description in some specified format. In our case, they can be represented by a public matrix $\mathbf{A}$. Hence, when we say that an adversary is given a function f as input when it is given its public representation.

Definition 2.3. *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 $|\mathbb{P}_{f \sim \mathcal{F}}[\mathcal{A}(f) = 1] - \mathbb{P}_{g \sim \mathcal{G}}[\mathcal{A}(g) = 1]| \leq \varepsilon$.*

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

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

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

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

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

Lemma 2.16 ([38, 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 2.17 ([38, Lem. 2.4]). *Let $\mathcal{F}$ be a function family with finite domain X. For $\varepsilon = \mathbb{E}_{f\sim\mathcal{F}}[|f(X)|]/|X|$, it holds that $(\mathcal{F}, U(X))$ is ε-uninvertible, even against unbounded adversaries.*

Lemma 2.18 ([38, Lem. 2.5]). *Let $\mathcal{F}$ be a function family with domain X and range Y, and $\mathcal{G}$ be an efficiently sampleable 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 [38]. 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 2.16, both $(\mathcal{F}, \mathcal{X})$ and $(\mathcal{G}, \mathcal{X})$ are one-way.

Definition 2.4. *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. [16] and thoroughly studied by Langlois and Stehlé [24]. As opposed to [24], we decide to use the primal formulation of the problem to match practical uses of the M-LWE assumption. It describes the following problem. Let K be a number field of degree n and R its ring of integers. Further, let d denote the rank and let ψ be a distribution on $K_\mathbb{R}$ and $\mathbf{s} \in R_q^d$ be a vector. We also define the torus $\mathbb{T} = K_\mathbb{R}/R$. We let $A^{\mathcal{M}}_{\mathbf{s},\psi}$ denote the distribution on $R_q^d \times \mathbb{T}$ obtained by sampling a vector $\mathbf{a} \hookleftarrow U(R_q^d)$, an element $e \hookleftarrow \psi$ and returning $(\mathbf{a}, q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod R)$.

Definition 2.5 (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 $\mathbf{s} \in R_q^d$ 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: Sample $\mathbf{s} \hookleftarrow U(R_q^d)$ 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 $U(R_q^d \times \mathbb{T})$.*

To be thorough, we should use the subscript K instead of n since there can be several number fields having the same degree n. However, to ease readability and since most of the other parameters are functions of n, we use the subscript n and keep the number field implicit. Although our results are mainly theoretical,

we note that throughout the paper n must be larger than a given security parameter, and the hardness and negligible advantages are thus expressed in terms of n rather than a proper security parameter. The advantage of an adversary $\mathcal{A}$ against M-SLWE is defined by $\mathrm{Adv}[\mathcal{A}] = \mathbb{P}[\mathcal{A}(A^{\mathcal{M}}_{\mathbf{s},\psi}) = \mathbf{s}]$ and the advantage against M-LWE by $\mathrm{Adv}[\mathcal{A}] = \left|\mathbb{P}[\mathcal{A}(A^{\mathcal{M}}_{\mathbf{s},\psi}) = 1] - \mathbb{P}[\mathcal{A}(U(Supp(A^{\mathcal{M}}_{\mathbf{s},\psi}))) = 1]\right|$. We say that the corresponding problems are ε-hard if for any PPT adversary $\mathcal{A}$, it holds that $\mathrm{Adv}[\mathcal{A}] \leq \varepsilon$. We say it is hard if ε is negligible in n.

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 $(\mathbf{a}_i, q^{-1}\langle \mathbf{a}_i, \mathbf{s}\rangle + e_i \bmod R)$ 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 $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod R)$. We denote it by $\text{M-LWE}_{n,d,m,q,\Upsilon}$.

Multiple secrets. Let k, m be positive integers, where m denotes the number of requested samples. In the multiple secrets version, the secret vector $\mathbf{s} \in R_q^d$ is replaced by a secret matrix $\mathbf{S} \in R_q^{d\times k}$ and the error vector $\mathbf{e} \sim \psi^m$ by an error matrix $\mathbf{E} \sim \psi^{m\times k}$. 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 [35, Lem. 2.9]. We denote the corresponding problem by $\text{M-LWE}^k_{n,d,m,q,\Upsilon}$.

Discrete version. As pointed out by Lyubashevsky et al. [30], 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 $\mathbb{T}$. Indeed, for the case of M-LWE, if the rounding function $\lfloor\cdot\rceil : K_{\mathbb{R}} \to R$ is chosen in a suitable way, see e.g. [31, Sec. 2.6], then every sample $(\mathbf{a}, b = q^{-1}\langle \mathbf{a}, \mathbf{s}\rangle + e \bmod R) \in R_q^d \times \mathbb{T}$ from $A^{\mathcal{M}}_{\mathbf{s},\psi}$ can be transformed to $(\mathbf{a}, \lfloor q \cdot b\rceil \bmod qR) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s}\rangle + \lfloor q \cdot e\rceil \bmod qR) \in R_q^d \times R_q$. 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 $\eta \ll (q-1)/2$. Note that the bound η is with regard to the coefficient embedding τ, meaning that the secret is in S_η^d. We denote the corresponding problem by $\eta\text{-M-LWE}_{n,d,q,\Upsilon}$.

Worst-case error. We also define the problem with worst-case error distribution when the error distribution is $D_{\mathbf{r}} \in \Psi_{\leq\alpha}$ for some possibly unknown (and possibly secret-dependent) $\mathbf{r} \in (\mathbb{R}^+)^n$. We then denote it by $\text{M-LWE}_{n,k,m,q,\Psi_{\leq\alpha}}$. We reduce to this variant in Section 3.2. Note that when the number of samples is fixed, one can use the reduction of [45, Lem. 7.2], generalizing [30, Lem. 5.16], to go to an average-case error with spherical Gaussian D_ξ. The reduction increases the noise from α to $\xi = \alpha(nm/\log_2(nm))^{1/4}$. The result is stated for R-LWE

but naturally extends to the module setting, and we therefore do not consider it new and do not include this extra step in the overall reduction.

M-LWE and M-ISIS function families. We now introduce two function families that are relevant for Section 4.

Definition 2.6. *Let K be a number field of degree n, and R its ring of integers. Let d, q, m be positive integers, and $X \subseteq R^m$. The* M-ISIS(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 $f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}^T\mathbf{x} \bmod qR$ 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 $g_{\mathbf{A}}(\mathbf{s}, \mathbf{e}) = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR$ for all $(\mathbf{s}, \mathbf{e}) \in R_q^d \times X$.*

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 S_η^m. When using the M-LWE function family in Section 4, we assume implicitly that the distribution on the first input $\mathbf{s}$ is always $U(R_q^d)$ and omit it from the notations.

In most LWE-based schemes, the secret key is $(\mathbf{s}, \mathbf{e})$ and the public key is $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e})$. 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 denote the problem corresponding to the family M-ISIS(n, d, m, q, X) by search/decision M-ISIS$_{d,m}$.

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. It improves upon our previously published works of [12,13] which dealt with secret with binary coefficients. To the best of our knowledge, [12] was the first result on the hardness of a structured variant of LWE with small uniform 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 algorithm for instance $(\mathbf{A}, \mathbf{Az}+\mathbf{e})$ of η-M-SLWE, our goal is to transform it into an algorithm for a related instance of M-SLWE defined by $(\mathbf{B}, \mathbf{Bs}+\mathbf{e}')$. Note that the secret $\mathbf{z}$ is in S_η^d, while the secret $\mathbf{s}$ is in R_q^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 2.8. 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.

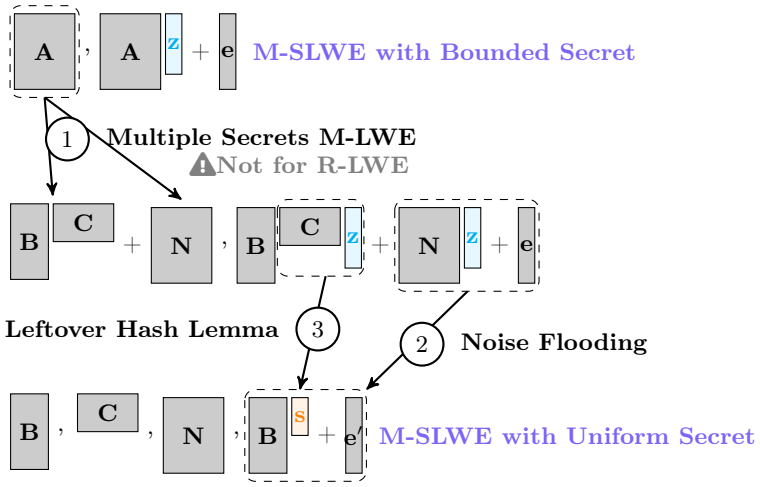

Fig. 3.1. Summary of the proof of Theorem 3.1

This consists of a reduction from M-SLWE and M-LWE with rank k to η-M-SLWE with rank $d \geq k\log_2(q)/\log_2(2\eta+1) + \Omega(\log_2(n)/\log_2(2\eta+1))$. It follows the original proof structure of Goldwasser et al. [21], 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 [21] 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 $d\log_2(2\eta+1) \geq k\cdot\log_2 q + \Omega(\log_2 n)$. Further, let α and β be positive such that $\beta \geq \alpha\cdot d\sqrt{m}\cdot\|\mathbf{V}\|_2\eta\sqrt{n}\log_2(n)$, and $\beta q \geq \eta_\varepsilon(R^m)$ for some $\varepsilon \in (0, 1/2)$. There is a PPT reduction from* $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R,\beta q}}$ *and* $\text{M-LWE}^d_{n,k,m,q,\mathcal{D}_{R,\alpha q}}$ *to* $\eta\text{-M-SLWE}_{n,d,m,q,\mathcal{D}_{R,\beta q}}$.

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 $k\log_2 q/\log_2(2\eta+1) + \Omega(\log_2 n/\log_2(2\eta+1))$ and the Gaussian width from αq to $\alpha q \cdot d\sqrt{m}\cdot\|\mathbf{V}\|_2\eta\sqrt{n}\log_2(n)$. 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, $\text{M-LWE}_{n,k,m,q,\mathcal{D}_{R,\alpha q}}$ trivially reduces to $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R,\beta q}}$, as $\gamma q = q\sqrt{\beta^2-\alpha^2}$ is above $\eta_\varepsilon(R)$ for a negligible ε, and sufficiently large so that $\mathcal{D}_{R,\gamma q}$ is efficiently sampleable.

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

- H_0 : Output $(\mathbf{A}, \mathbf{Az}+\mathbf{e} \bmod qR)$, as in $\eta\text{-M-SLWE}_{n,d,m,q,\mathcal{D}_{R,\beta q}}$;
- H_1 : Output $(\mathbf{A}' = \mathbf{BC}+\mathbf{N} \bmod qR, \mathbf{A}'\mathbf{z}+\mathbf{e} \bmod qR)$, where $\mathbf{B} \hookleftarrow U(R_q^{m\times k})$, $\mathbf{C} \hookleftarrow U(R_q^{k\times d})$, and $\mathbf{N} \hookleftarrow \mathcal{D}_{R,\alpha q}^{m\times d}$ and $\mathbf{z}, \mathbf{e}$ as in H_0;
- H_2 : Output $(\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{BCz}+\mathbf{Nz}+\mathbf{e} \bmod qR)$, where $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{z}, \mathbf{e}$ as in H_1;
- H_3 : If $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha q\sqrt{md}\log_2 n$, output $(\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{BCz}+\mathbf{Nz}+\mathbf{e} \bmod qR)$, where $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{z}, \mathbf{e}$ as in H_2;
- H_4 : If $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha q\sqrt{md}\log_2 n$, output $(\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{BCz}+\mathbf{e}' \bmod qR)$, where $\mathbf{e}' \hookleftarrow \mathcal{D}_{R,\beta q}^m$ and $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{z}$ as in H_3;
- H_5 : If $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha q\sqrt{md}\log_2 n$, output $(\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{Bs}+\mathbf{e}' \bmod qR)$, where $\mathbf{s} \hookleftarrow U(R_q^k)$ and $\mathbf{B}, \mathbf{C}, \mathbf{N}, \mathbf{e}'$ as in H_4.

For $i \in \{0, \ldots, 5\}$, we denote by P_i the problem of finding the secret $\mathbf{z}$ (resp. $\mathbf{s}$ in H_5), given a sample of the distribution H_i. Recall that problem P_i is hard if for any PPT attacker $\mathcal{A}$ the advantage of solving P_i is negligible, thus $\text{Adv}_{P_i}[\mathcal{A}] =$

$\mathbb{P}_{X\sim H_i}[\mathcal{A}(X)=\mathbf{z}] \leq n^{-\omega(1)}$, where n is the degree of K. The overall idea is to show that if P_5 is hard, then P_0 is hard as well.

From P_0 to P_1: Assuming the hardness of M-LWE$^d_{n,k,m,q,\mathcal{D}_{R,\alpha q}}$, the distributions H_0 and H_1 are computationally indistinguishable. We note that the hardness of the latter can be obtained by a hybrid argument, e.g., Lemma 3.4, from that of M-LWE$_{n,k,m,q,\mathcal{D}_{R,\alpha q}}$ with a reduction loss factor of d in the advantage. Thus, if M-LWE$_{n,k,m,q,\mathcal{D}_{R,\alpha q}}$ is $\mathrm{Adv}_{\text{M-LWE}}$-hard, it holds

$$\mathrm{Adv}_{P_0}[\mathcal{A}] \leq \mathrm{Adv}_{P_1}[\mathcal{A}] + d \cdot \mathrm{Adv}_{\text{M-LWE}},$$

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. From P_2 onwards the adversary is given more elements (namely $\mathbf{B}, \mathbf{C}, \mathbf{N}$ instead of $\mathbf{A}'$) but can simply reconstruct the M-LWE matrix from these elements. Hence, we have

$$\mathrm{Adv}_{P_1}[\mathcal{A}] \leq \mathrm{Adv}_{P_2}[\mathcal{A}].$$

From P_2 to P_3: Note that conditioned on $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha q\sqrt{md}\log_2 n$, the two distributions are identical. Yet, Lemma 2.13 for $\mathcal{I} = R$ yields the spectral bound with overwhelming probability. Hence, we have $\Delta(H_2, H_3) \leq \mathbb{P}[\|M_{\sigma_H}(\mathbf{N})\|_2 > \alpha q\sqrt{md}\log_2 n] =: p_{\text{spectral}} \leq n^{-\omega(1)}$, resulting in

$$\mathrm{Adv}_{P_2}[\mathcal{A}] \leq \mathrm{Adv}_{P_3}[\mathcal{A}] + p_{\text{spectral}}$$

From P_3 to P_4: By the probability preservation property of the Rényi divergence (Lemma 2.7), we have

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

We first explain how to bound the Rényi divergence between $\mathbf{Nz}+\mathbf{e}$ and $\mathbf{e}'$ for a fixed $(\mathbf{N}, \mathbf{z})$. First note that $\mathbf{Nz}+\mathbf{e}$ follows the distribution $\mathcal{D}_{R^m+\mathbf{Nz},\mathbf{Nz},\beta q}$. Since we have $\mathbf{Nz} \in R^m$, this distribution is exactly $\mathcal{D}_{R^m,\mathbf{Nz},\beta q}$. Then, as σ and σ_H only differ by the unitary transformation $\mathbf{U}_H$, we have that $\|\sigma_H(\mathbf{z})\|_2 = \|\sigma(\mathbf{z})\|_2 \leq \|\mathbf{V}\|_2\|\tau(\mathbf{z})\|_2 \leq \|\mathbf{V}\|_2 \cdot \eta\sqrt{nd}$, as $\mathbf{z} \in S^d_\eta$. Finally, because of our conditioning, we have $\|M_{\sigma_H}(\mathbf{N})\|_2 \leq \alpha q \log_2 n\sqrt{md}$. It then holds that $\|\sigma_H(\mathbf{Nz})\|_2 = \|M_{\sigma_H}(\mathbf{N})\sigma_H(\mathbf{z})\|_2 \leq \|M_{\sigma_H}(\mathbf{N})\|_2\|\sigma_H(\mathbf{z})\|_2 \leq \alpha q d\|\mathbf{V}\|_2\eta\sqrt{nm}\log_2(n)$. Then, using that $\beta q \geq \eta_\varepsilon(R^m)$, Lemma 2.12 yields

$$\mathrm{RD}_2\left(\mathcal{D}_{R^m,\mathbf{Nz},\beta q}\|\mathcal{D}_{R^m,\beta q}\right) \leq \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \cdot \exp\left(\frac{2\pi\|\sigma_H(\mathbf{Nz})\|_2^2}{(\beta q)^2}\right).$$

However, it holds that $\exp(2\pi\|\sigma_H(\mathbf{N}\tilde{\mathbf{z}})\|_2^2/(\beta q)^2) \leq \exp(2\pi)$ because of how we chose β with respect to α. Without loss of generality, assume $\varepsilon < \frac{1}{2}$ resulting in $\mathrm{RD}_2(\mathcal{D}_{R^m,\mathbf{Nz},\beta q}\|\mathcal{D}_{R^m,\beta q}) = O(1)$.

Next, the data processing inequality of Lemma 2.7 gives $\mathrm{RD}_2(H_3\|H_4) \leq \mathrm{RD}_2((\mathbf{N},\mathbf{z},\mathbf{e}+\mathbf{N}\mathbf{z})\|(\mathbf{N},\mathbf{z},\mathbf{e}'))$. We now bound this divergence by a constant using the previous calculation.

$$\begin{aligned}
\mathrm{RD}_2((\mathbf{N},\mathbf{z},\mathbf{e}+\mathbf{N}\mathbf{z})\|(\mathbf{N},\mathbf{z},\mathbf{e}')) &= \sum_{(\bar{\mathbf{N}},\bar{\mathbf{z}},\bar{\mathbf{e}})} \frac{\mathbb{P}[(\mathbf{N},\mathbf{z},\mathbf{e}+\mathbf{N}\mathbf{z})=(\bar{\mathbf{N}},\bar{\mathbf{z}},\bar{\mathbf{e}})]^2}{\mathbb{P}[(\mathbf{N},\mathbf{z},\mathbf{e}')=(\bar{\mathbf{N}},\bar{\mathbf{z}},\bar{\mathbf{e}})]} \\
&= \sum_{(\bar{\mathbf{N}},\bar{\mathbf{z}},\bar{\mathbf{e}})} \frac{\mathbb{P}[(\mathbf{N},\mathbf{z})=(\bar{\mathbf{N}},\bar{\mathbf{z}})]^2\mathbb{P}[\mathbf{e}+\bar{\mathbf{N}}\bar{\mathbf{z}}=\bar{\mathbf{e}}]^2}{\mathbb{P}[(\mathbf{N},\mathbf{z})=(\bar{\mathbf{N}},\bar{\mathbf{z}})]\mathbb{P}[\mathbf{e}'=\bar{\mathbf{e}}]} \\
&= \sum_{(\bar{\mathbf{N}},\bar{\mathbf{z}})} \mathbb{P}[(\mathbf{N},\mathbf{z})=(\bar{\mathbf{N}},\bar{\mathbf{z}})]\mathrm{RD}_2(\mathbf{e}+\bar{\mathbf{N}}\bar{\mathbf{z}}\|\mathbf{e}') \\
&\leq \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \cdot e^{2\pi} \sum_{(\bar{\mathbf{N}},\bar{\mathbf{z}})} \mathbb{P}[(\mathbf{N},\mathbf{z})=(\bar{\mathbf{N}},\bar{\mathbf{z}})] \\
&= \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \cdot e^{2\pi} \\
&= O(1),
\end{aligned}$$

as desired.

From P_4 to P_5: By the probability preservation property and data processing inequality of the Rényi divergence (Lemma 2.7), we have

$$\begin{aligned}
\mathrm{Adv}_{P_4}[\mathcal{A}]^2 &= \mathbb{P}_{(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})\sim H_4}[\mathcal{A}(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})=\mathbf{z}]^2 \\
&\leq \mathbb{P}_{(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})\sim H_4}[\mathcal{A}(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})=\mathbf{C}\mathbf{z}]^2 \\
&\leq \mathbb{P}_{(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})\sim H_5}[\mathcal{A}(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})=\mathbf{s}] \cdot \mathrm{RD}_2(H_4\|H_5) \\
&\leq \mathrm{Adv}_{P_5}[\mathcal{A}] \cdot \mathrm{RD}_2((\mathbf{C},\mathbf{C}\mathbf{z})\|(\mathbf{C},\mathbf{s})).
\end{aligned}$$

The first inequality follows from the fact that if $\mathcal{A}$ can find $\mathbf{z}$ from $(\mathbf{B},\mathbf{C},\mathbf{N},\mathbf{b})$, then they can also find $\mathbf{C}\mathbf{z}$, hence the inclusion of events. The second and third inequalities come from the probability preservation and data processing inequality of Lemma 2.7 respectively. By the leftover hash lemma stated in Lemma 2.8, the Rényi divergence between the distribution $(\mathbf{C},\mathbf{C}\mathbf{z})$ and the distribution $(\mathbf{C},\mathbf{s})$ is bounded above by $(1+q^k/(2\eta+1)^d)^n$. As we require $d\log_2(2\eta+1) \geq k\log_2 q + \Omega(\log_2 n)$, we obtain $\mathrm{RD}_2(H_4\|H_5) \leq (1+1/\Omega(n))^n = O(1)$ asymptotically in n.

Problem P_5: This problem is exactly the $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R,\beta q}}$ problem, as $\mathbf{C}$ and $\mathbf{N}$ are independent of $\mathbf{B},\mathbf{s}$ and $\mathbf{e}'$. So if $\text{M-SLWE}_{n,k,m,q,\mathcal{D}_{R,\beta q}}$ is $\mathrm{Adv}_{\text{M-SLWE}}$-hard, it hold that

$$\mathrm{Adv}_{P_5}[\mathcal{A}] \leq \mathrm{Adv}_{\text{M-SLWE}}$$

Putting all equations from above together, we obtain

$$\begin{aligned}
\mathrm{Adv}_{P_0}[\mathcal{A}] &\leq \mathrm{Adv}_{P_1}[\mathcal{A}] + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\
&\leq \mathrm{Adv}_{P_2}[\mathcal{A}] + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\
&\leq \mathrm{Adv}_{P_3}[\mathcal{A}] + p_{\text{spectral}} + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\
&\leq \sqrt{\mathrm{Adv}_{P_4}[\mathcal{A}] \cdot \mathrm{RD}_2(H_3 \| H_4)} + p_{\text{spectral}} + d \cdot \mathrm{Adv}_{\text{M-LWE}} \\
&\leq \sqrt{\sqrt{\mathrm{Adv}_{\text{M-SLWE}} \cdot \mathrm{RD}_2(H_4 \| H_5)} \cdot \mathrm{RD}_2(H_3 \| H_4)} \\
&\quad + p_{\text{spectral}} + d \cdot \mathrm{Adv}_{\text{M-LWE}}.
\end{aligned}$$

The choice of parameters yields $\mathrm{RD}_2(H_3\|H_4), \mathrm{RD}_2(H_4\|H_5) = O(1)$, $p_{\text{spectral}} \leq n^{-\omega(1)}$, 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 [17] 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 $d\log_2(2\eta+1) \geq (k+1)\log_2 q + \omega(\log_2 n)$. 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 $n\eta\sqrt{2d}\sqrt{4n^2\eta^2+1}$. In the case of general cyclotomic fields, the noise rate slightly improves on the noise rate of $d\sqrt{m}\cdot n^{3/2}\log_2(n)\eta$ from Section 3.1. We indeed improve the noise rate by a factor of roughly $\sqrt{8n}\eta/\log_2(n)\sqrt{md}$, which is advantageous whenever $m > 8n\eta^2/d\log_2^2 n$. 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 $m > 8n^2\eta/d\log_2^2 n = \Theta(n^2/\log_2^3 n)$, which is now just sub-quadratic. Nonetheless, this reduction allows for proving the hardness of the decision version of η-M-LWE which is preferable for the security of cryptographic applications. 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 to all number fields $K = \mathbb{Q}(\zeta)$ for which 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 ν-th cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $\mu = \prod_i p_i^{f_i}$ for some $f_i \in [e_i]$, and q be a prime number such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$, $q > (\eta\mathfrak{s}_1(\mu))^{\varphi(\mu)}$, and $q^{\nu/\mu} \geq n^{\omega(1/(k+1))}$, 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 $d\log_2(2\eta+1) \geq (k+1)\log_2 q + \omega(\log_2 n)$, and $d \leq m \leq \mathsf{poly}(n)$. Let $\alpha \geq \sqrt{n}/q \cdot \sqrt{\ln(2nm(1+\varepsilon^{-1}))/\pi}$ for*

some $\varepsilon \in (0, 1/2)$*, and* $\beta \geq \alpha \cdot n\eta\sqrt{2d}\sqrt{4n^2\eta^2+1}$*. Then there is a* PPT *reduction from* M-LWE$_{n,k,m,q,D_\alpha}$ *to* η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$*, such that if* $\mathcal{A}$ *solves the latter with advantage* Adv$[\mathcal{A}]$*, then there exists an algorithm* $\mathcal{B}$ *that solves the former with advantage*

$$\mathrm{Adv}[\mathcal{B}] \geq \frac{1}{3m}\left(\mathrm{Adv}[\mathcal{A}] - \frac{1}{2}\sqrt{\left(1+\frac{q^{k+1}}{(2\eta+1)^d}\right)^n - 1}\right) - \frac{103\varepsilon}{6} - \prod_{i\in[g]}\left(1-q^{-(k+1)\nu/\mu}\right),$$

When $\nu = 2^{\ell+1}, n = 2^\ell$*, one can take any prime* q *such that* $q = 2\kappa+1 \bmod 4\kappa$ *for some* $\kappa = 2^l$ *with* $l \in [\ell]$*, and such that* $q > (\eta\sqrt{\kappa})^\kappa$ *and* $q^{\nu/\kappa} \geq n^{\omega(1/(k+1))}$*.*

The modulus is constrained in terms of its splitting behavior. The conditions essentially mean that q splits into $\varphi(\mu)$ factors, each having algebraic norm $q^{\nu/\mu}$. This norm must be at least $n^{\omega(1/(k+1))}$ for Lemma 3.1 to go through, and q must exceed $(\eta\mathfrak{s}_1(\mu))^{\varphi(\mu)}$ so that every element of S_η is a unit in R_q. Then, the noise ratio β/α contains three main terms. The factor $n\eta$ 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 $\sqrt{4n^2\eta^2+1}$ solely represents the impact of giving information on the error in the ext-M-LWE problem.

Proof. We give an overview of the full reduction in Figure 3.2, and detail here how to combine Lemma 3.1, 3.3, 3.4 and 3.5. For clarity, we define δ_1 the loss incurred by the leftover hash lemma, namely $\delta_1 = \frac{1}{2}\sqrt{\left(1+\frac{q^{k+1}}{(2\eta+1)^d}\right)^n - 1}$, and $\delta_2 = \prod_{i\in[\varphi(\mu)]}\left(1-q^{-(k+1)\nu/\mu}\right)$. Assume there exists a PPT distinguisher $\mathcal{A}$ for η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$ which succeeds with advantage Adv$[\mathcal{A}]$. By Lemma 3.5, one can construct PPT adversaries $\mathcal{B}_1, \mathcal{B}_2$ and $\mathcal{B}_3$ for ext-M-LWE$^m_{n,k+1,d,q,\psi,S_\eta^d}$, M-LWE$_{n,k+1,m,q,D_\gamma}$ and ext-M-LWE$^m_{n,k+1,d,q,\psi,\{0\}^d}$ respectively such that

$$\mathrm{Adv}[\mathcal{B}_1] + \mathrm{Adv}[\mathcal{B}_2] + \mathrm{Adv}[\mathcal{B}_3] \geq \mathrm{Adv}[\mathcal{A}] - 2m\varepsilon - \delta_1,$$

where $\gamma = \alpha B_\eta\sqrt{d}\sqrt{4B_\eta^2+1}$ and $\psi = \mathcal{D}_{q^{-1}R,\alpha\sqrt{4B_\eta^2+1}}$. Hence, there exists $i \in \{1,2,3\}$ such that $\mathrm{Adv}[\mathcal{B}_i] \geq \frac{1}{3}(\mathrm{Adv}[\mathcal{A}] - 2m\varepsilon - \delta_1)$, otherwise the above inequality yields a contradiction.

Case $i = 1$: By Lemma 3.4, one can use $\mathcal{B}_1$ to construct a distinguisher $\mathcal{B}_1^{(1)}$ for ext-M-LWE$^1_{n,k+1,d,q,\psi,S_\eta^d}$ such that

$$\mathrm{Adv}[\mathcal{B}_1^{(1)}] = \frac{1}{m}\mathrm{Adv}[\mathcal{B}_1].$$

Then, by Lemma 3.3, the latter can be used to construct a distinguisher $\mathcal{B}_1^{(2)}$ for first-is-errorless M-LWE$_{n,k+1,d,q,D_\alpha}$ such that

$$\mathrm{Adv}[\mathcal{B}_1^{(2)}] \geq \mathrm{Adv}[\mathcal{B}_1^{(1)}] - 33\varepsilon/2.$$

Lemma 3.1 then yields a distinguisher $\mathcal{B}_1^{(3)}$ for M-LWE$_{n,k,d,q,D_\alpha}$ with advantage

$$\mathrm{Adv}[\mathcal{B}_1^{(3)}] \geq \mathrm{Adv}[\mathcal{B}_1^{(2)}] - \delta_2.$$

Since $m \geq d$, we can consider a distinguisher $\mathcal{B}_1^{(4)}$ for M-LWE$_{n,k,m,q,D_\alpha}$ which simply calls $\mathcal{B}_1^{(3)}$ on the first d samples. By combining it all, we have

$$\begin{aligned}
\mathrm{Adv}[\mathcal{B}_1^{(4)}] \geq \mathrm{Adv}[\mathcal{B}_1^{(3)}] &\geq \mathrm{Adv}[\mathcal{B}_1^{(2)}] - \delta_2 \\
&\geq \mathrm{Adv}[\mathcal{B}_1^{(1)}] - \delta_2 - 33\varepsilon/2 \\
&= \frac{1}{m}\mathrm{Adv}[\mathcal{B}_1] - \delta_2 - 33\varepsilon/2 \\
&\geq \frac{1}{3m}(\mathrm{Adv}[\mathcal{A}] - \delta_1 - 2m\varepsilon) - \delta_2 - 33\varepsilon/2 \\
&= \frac{1}{3m}(\mathrm{Adv}[\mathcal{A}] - \delta_1) - \delta_2 - 103\varepsilon/6.
\end{aligned}$$

Case $i = 2$: We construct the adversary $\mathcal{B}_2^{(1)}$ for M-LWE$_{n,k,m,q,D_\alpha}$ using $\mathcal{B}_2$ as follows. On input $(\mathbf{A}, \mathbf{b}) \in R_q^{m\times k} \times \mathbb{T}^m$, the adversary $\mathcal{B}_2^{(1)}$, samples $\mathbf{a} \hookleftarrow U(R_q^m)$, $s \hookleftarrow U(R_q)$ and $\mathbf{e}' \hookleftarrow D^m_{\sqrt{\gamma^2-\alpha^2}}$. It then calls $\mathcal{B}_2$ on input $(\mathbf{A}', \mathbf{b}') = ([\mathbf{A}|\mathbf{a}], \mathbf{b} + q^{-1}s \cdot \mathbf{a} + \mathbf{e}' \bmod R)$. If $\mathbf{b} = q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod R$, then $\mathbf{b}' = q^{-1}\mathbf{A}' \cdot [\mathbf{s}^T|s]^T + (\mathbf{e} + \mathbf{e}') \bmod R$, and $\mathbf{e} + \mathbf{e}'$ is distributed as D^m_γ. Otherwise, if $\mathbf{b}$ is uniform, then so is $\mathbf{b}'$ as $\mathbf{b}$ is independent of $q^{-1}\mathbf{a}s + \mathbf{e}'$. Hence, we have

$$\begin{aligned}
\mathrm{Adv}[\mathcal{B}_2^{(1)}] \geq \mathrm{Adv}[\mathcal{B}_2] &\geq \frac{1}{3}(\mathrm{Adv}[\mathcal{A}] - \delta_1 - 2m\varepsilon) \\
&\geq \frac{1}{3m}(\mathrm{Adv}[\mathcal{A}] - \delta_1) - \delta_2 - 103\varepsilon/6.
\end{aligned}$$

Case $i = 3$: Since ext-M-LWE$^m_{n,k+1,d,q,\psi,S_\eta^d}$ reduces to ext-M-LWE$^m_{n,k+1,d,q,\psi,\{0\}^d}$ without transforming the samples, we can assume without loss of generality that $\mathrm{Adv}[\mathcal{B}_1] \geq \mathrm{Adv}[\mathcal{B}_3]$. It then yields $\mathrm{Adv}[\mathcal{B}_1] \geq \frac{1}{3}(\mathrm{Adv}[\mathcal{A}] - 2m\varepsilon - \delta_1)$ as in Case $i = 1$.

In any case, we can construct a distinguisher $\mathcal{B}$ for M-LWE$_{n,k,m,q,D_\alpha}$ as claimed. □

3.2.1 First-is-errorless M-LWE. We follow the same idea as Brakerski et al. [17] 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 [17], where the first equation is given without error. A similar definition and reduction from M-LWE are given in [4]. The only difference between the two reductions comes from the pre-processing step, which is performed before receiving the M-LWE samples. In our case, this step is simplified and extended to general number fields, provided that the modulus q verifies certain splitting conditions. Further restrictions on q in our reduction encompasses these conditions.

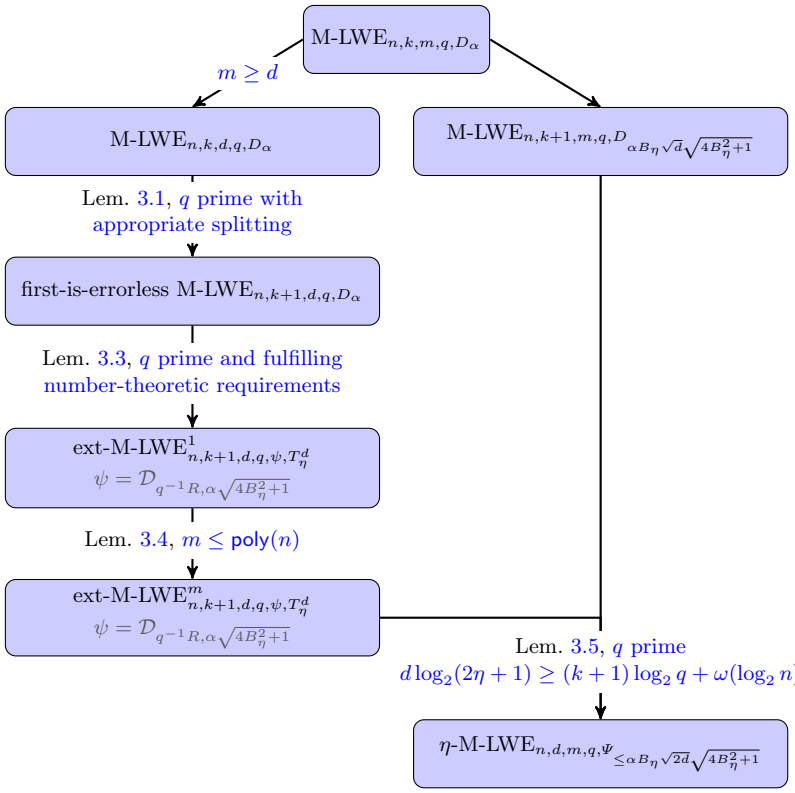

Fig. 3.2. Summary of the proof of Theorem 3.2, where $B_\eta = \max_{x \in S_\eta} \|\sigma(x)\|_\infty$ and σ is the canonical embedding. In cyclotomic fields, we have $B_\eta \leq n\eta$. Note that Lemma 3.5 uses d samples from ext-M-LWE, where d is the module rank in η-M-LWE. The assumptions on q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

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 $U(R_q^k \times q^{-1}R/R)$ and the rest from $U(R_q^k \times \mathbb{T})$. On the other hand, there is some unknown $\mathbf{s} \hookleftarrow U(R_q^k)$ 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$_{n,k,m,q,\Upsilon}$*, where only $m-1$ coefficients contain errors.*

Lemma 3.1 (Adapted from [17, Lem. 4.3]). *Let K be a number field of degree n, and R its ring of integers. Then, let k be a positive integer, and Υ a distribution over a family of distributions over $K_\mathbb{R}$. Let q be an unramified prime integer such that the smallest norm of its prime ideal factors is*

at least $n^{\omega(1/k)}$*. There is a* PPT *reduction from* M-LWE$_{n,k-1,q,\Upsilon}$ *to the variant* first-is-errorless M-LWE$_{n,k,q,\Upsilon}$*. If the number of samples* m *is fixed, it gives a* PPT *reduction from* M-LWE$_{n,k-1,m-1,q,\Upsilon}$ *to* first-is-errorless M-LWE$_{n,k,m,q,\Upsilon}$*. The reduction reduces the advantage by at most* $1-\prod_{i\in[\kappa]}(1-N(\mathfrak{p}_i)^{-k})$*, where the* $\mathfrak{p}_i$*'s are the prime ideal factors of* $\langle q\rangle$*.*

Proof. Pre-processing: The reduction first samples $\mathbf{a}' \hookleftarrow U(R_q^k)$ such that $\mathbf{a}'$ is R_q-linearly independent. As a result, $\mathbf{a}'$ is uniform among the R_q-linearly independent vectors. We first show that under the conditions of the lemma, the distribution of $\mathbf{a}'$ is statistically close to $U(R_q^k)$. Denote by S the set of vectors of R_q^k that are linearly independent, and S' its complement in R_q^k. It then holds that

$$\begin{aligned}
\Delta(U(R_q^k), U(S)) &= \frac{1}{2}\sum_{\mathbf{x}\in S}\left|\frac{1}{|R_q^k|}-\frac{1}{|S|}\right| + \frac{1}{2}\sum_{\mathbf{x}\in S'}\left|\frac{1}{|R_q^k|}-0\right| \\
&= \frac{1}{2}\left(|S|\left(\frac{1}{|S|}-\frac{1}{|R_q^k|}\right)+\frac{|S'|}{|R_q^k|}\right) \\
&= \frac{|S'|}{|R_q^k|} \\
&= 1-\mathbb{P}_{\mathbf{a}'\sim U(R_q^k)}[\mathbf{a}' \text{ is } R_q\text{-linearly independent}]
\end{aligned}$$

By Lemma 2.5 for $\ell = 0$, we have

$$\begin{aligned}
\mathbb{P}_{\mathbf{a}'\sim U(R_q^k)}[\mathbf{a}' \text{ is } R_q\text{-linearly independent}] &= \prod_{i\in[\kappa]}\left(1-\frac{1}{N(\mathfrak{p}_i)^k}\right) \\
&\geq 1-\frac{n}{\min_{i\in[\kappa]} N(\mathfrak{p}_i)^k} \\
&\geq 1-n^{-\omega(1)},
\end{aligned}$$

by assumption. Hence $\Delta(U(R_q^k), U(S)) \leq n^{-\omega(1)}$. Then, from $\mathbf{a}'$, one can efficiently complete it with $\mathbf{b}_2, \ldots, \mathbf{b}_k \in R_q^k$ such that the matrix $\mathbf{U} = [\mathbf{a}'|\mathbf{b}_2|\ldots|\mathbf{b}_k]$ is invertible in R_q. For example, this can be done by successively sampling the $\mathbf{b}_i$'s uniformly at random in R_q^k. By Lemma 2.5, the probability that the newly drawn $\mathbf{b}_{\ell+1}$ is kept is $\prod_{i\in[\kappa]} 1-N(\mathfrak{p}_i)^{-(k-\ell)} \geq 1-n^{-\omega(1)}$. It would thus require at most a polynomial number of sampled vectors.
Reduction: Then, sample s_0 uniformly in R_q. The reduction is as follows. For the first sample, it outputs $(\mathbf{a}', q^{-1}\cdot s_0 \bmod R) \in R_q^k \times q^{-1}R/R$. The other samples are produced by taking $(\mathbf{a}, b) \in R_q^{k-1}\times\mathbb{T}$ from the M-LWE challenger, picking a fresh randomly chosen $a'' \in R_q$, and outputting $(\mathbf{U}[a''|\mathbf{a}^T]^T, b+q^{-1}(s_0\cdot a'') \bmod R) \in R_q^k \times \mathbb{T}$. 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 $\mathbf{U}[a''|\mathbf{a}^T]^T$ is uniform over R_q^k as well.

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

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

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

Remark 3.1. Later in the reduction, we restrict the modulus q to be a prime that splits into few prime factors in the underlying cyclotomic field to maximize the number of invertible elements, and more precisely to be able to use Lemma A.3 without having to take superpolynomial q. In this case, one could use moduli that split into say κ factors such that $q^{-n/\kappa} \leq n^{-\omega(1)}$.

3.2.2 Extended M-LWE. We now define the module version of the *Extended* LWE problem introduced in [17], where the adversary is allowed a hint on the errors. A first definition of ext-M-LWE was introduced by Alperin-Sheriff and Apon [4] 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 [33] 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 $\mathcal{Z} \subseteq R^m$ and ψ a discrete distribution over $q^{-1}R$. The* Extended M-LWE *problem, denoted by* ext-M-LWE$^{\ell}_{n,k,m,q,\psi,\mathcal{Z}}$*, is as follows. The adversary first chooses $\mathbf{z} \in \mathcal{Z}$ and then receives a tuple $(\mathbf{A}, \mathbf{B}, \mathbf{E}^T\mathbf{z})$ over $R_q^{m\times k} \times (q^{-1}R/R)^{m\times \ell} \times (q^{-1}R)^{\ell}$. 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 $\mathbf{B} = q^{-1}\mathbf{A}\mathbf{S} + \mathbf{E} \bmod R$ for some uniformly chosen $\mathbf{S} \in R_q^{k\times \ell}$. On the other side, all is identical except that $\mathbf{B}$ is sampled from $U((q^{-1}R/R)^{m\times \ell})$, 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 hint vectors $\mathbf{z}$ that can be given. 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 $\mathcal{Z} = S_\eta^m$ 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.

Similar to the notation from [35], for a matrix $\mathbf{A} \in R^{m\times m}$, we denote by $\mathbf{A}^{]1[} \in R^{m\times(m-1)}$ 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 $\mathbf{z} \in \mathcal{Z} = S_\eta^m$, 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 $\mathbf{U}_\mathbf{z}^{]1[}$ (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 ν-th cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $\mu = \prod_i p_i^{f_i}$ for some $f_i \in [e_i]$, η a positive integer and q be a prime such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > (\eta\mathfrak{s}_1(\mu))^{\varphi(\mu)}$, 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 $\mathcal{Z} = S_\eta^m$. 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 $\mathbf{U}_\mathbf{z}^{]1[}$, and $\left\| M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[}) \right\| \leq 2B_\eta$, where $B_\eta = \max_{x\in S_\eta} \|\sigma(x)\|_\infty$.*

When $\nu = 2^{\ell+1}, n = 2^\ell$, one can take any prime q such that $q = 2\kappa+1 \bmod 4\kappa$ for some $\kappa = 2^l$ with $l \in [\ell]$, and such that $q > (\eta\sqrt{\kappa})^\kappa$.

Proof. Let $\mathbf{z} \in \mathcal{Z}$. First, we construct $\mathbf{U_z}$ in the case where all the 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:

$$\mathbf{U_z} = \begin{bmatrix} 1 & -z_2 & & \\ & z_1 & \ddots & \\ & & \ddots & -z_m \\ & & & z_{m-1} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & z_1 & & \\ & & \ddots & \\ & & & z_{m-1} \end{bmatrix} + \begin{bmatrix} 0 & -z_2 & & \\ & \ddots & \ddots & \\ & & \ddots & -z_m \\ & & & 0 \end{bmatrix}.$$

$$\mathbf{U}_\mathbf{z}^{]1[} \qquad \mathbf{A}^{]1[} \qquad \mathbf{B}^{]1[}$$

The matrix $\mathbf{U_z}$ is invertible modulo qR only if all the z_i (except 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 $[m]$, $0 < \|\tau(z_i)\|_\infty \leq \eta$, where τ is the coefficient embedding. By Lemma 2.4, since q verifies the algebraic conditions taking all $f_i = 1$ and $q^{1/\varphi(\mu)}/\mathfrak{s}_1(\mu) > \eta$,

all the z_i are in $R_q^\times$.

By construction, the last $m-1$ columns of $\mathbf{U_z}$ are orthogonal to $\mathbf{z}$. Let $\mathbf{U_z}^{]1[}$ be the submatrix of $\mathbf{U_z}$ obtained by removing the leftmost column as shown above. As observed in Lemma 2.3 for example, it holds that $\left\|M_{\sigma_H}(\mathbf{U_z}^{]1[})\right\|_2 = \left\|M_\sigma(\mathbf{U_z}^{]1[})\right\|_2$. Then, using the fact that M_σ is a ring homomorphism, we have $M_\sigma(\mathbf{U_z}^{]1[}) = M_\sigma(\mathbf{A}^{]1[}) + M_\sigma(\mathbf{B}^{]1[})$. 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 $\left\|M_\sigma(\mathbf{A}^{]1[})\mathbf{x}\right\|_2 = \sqrt{\sum_{i\in[m-1]}\sum_{j\in[n]}|\sigma_j(z_i)|^2|x_{j+n(i-1)}|^2} \leq B_\eta\|\mathbf{x}\|_2$, because each z_i is in S_η. This yields $\left\|M_\sigma(\mathbf{A}^{]1[})\right\|_2 \leq B_\eta$. A similar calculation on $\mathbf{B}^{]1[}$ leads to $\left\|M_\sigma(\mathbf{B}^{]1[})\right\|_2 \leq B_\eta$, thus resulting in $\left\|M_\sigma(\mathbf{U_z}^{]1[})\right\|_2 \leq 2B_\eta$.

Now assume that $z_{i_0}, \ldots, z_m$ are zeros for some i_0 in $[m]$. If the zeros do not appear last in the vector $\mathbf{z}$, we can replace $\mathbf{z}$ with $\mathbf{Sz}$, where $\mathbf{S} \in R^{m\times m}$ swaps the coordinates of $\mathbf{z}$ 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 $z_{i_0}, \ldots, z_m$ on the diagonal are replaced by 1. The orthogonality is preserved, and $\left\|M_\sigma(\mathbf{U_z}^{]1[})\right\|_2$ can still be bounded above by $2B_\eta$. □

Notice that when the ring is of degree 1 and $\eta = 1$, the constructions in the different cases match the ones from [17, Claim 4.6]. So do the singular values as $B_\eta \leq n\eta = 1$ by Lemma 2.1. Also, the construction differs from the notion of quality in [4] 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 [17, Lem. 4.7]). *Let $\nu = \prod_i p_i^{e_i}$, K be the ν-th cyclotomic field of degree $n = \varphi(\nu)$, and R its ring of integers. Let $\mu = \prod_i p_i^{f_i}$ for some $f_i \in [e_i]$, η a positive integer and q be a prime such that $q = 1 \bmod \mu$, $ord_\nu(q) = \nu/\mu$ and $q > (\eta\mathfrak{s}_1(\mu))^{\varphi(\mu)}$, 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, $\mathcal{Z} = S_\eta^m$, $\varepsilon \in (0, 1/2)$ and $\alpha \geq \sqrt{n}/q \cdot \sqrt{\ln(2nm(1+\varepsilon^{-1}))/\pi}$. There is a PPT reduction from the variant* 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 adversary's advantage by at most $33\varepsilon/2$, where $\psi = \mathcal{D}_{q^{-1}R,\alpha\sqrt{4B_\eta^2+1}}$ and $B_\eta = \max_{x\in S_\eta}\|\sigma(x)\|_\infty$.*

When $\nu = 2^{\ell+1}, n = 2^\ell$, one can take any prime q such that $q = 2\kappa+1 \bmod 4\kappa$ for some $\kappa = 2^l$ with $l \in [\ell]$, and such that $q > (\eta\sqrt{\kappa})^\kappa$.

Proof. First, we clarify the condition on α. Since K is a cyclotomic field, $\mathbf{B} = \mathbf{I}_m \otimes q^{-1}[\sigma_H(1)|\ldots|\sigma_H(\zeta^{n-1})]$ is a basis of the lattice $\sigma_H(q^{-1}R^m)$, and each vector has norm $\sqrt{n}/q$. As a result, the max-Euclidean norm of the Gram-Schmidt orthogonalization of $\mathbf{B}$ is at most $\sqrt{n}/q$. By [20, Lem. 3.1, Thm. 4.1], our condition on α ensures that $\alpha \geq \eta_\varepsilon(q^{-1}R^m)$ and that $\mathcal{D}_{q^{-1}R^m,\alpha}$ is efficiently sampleable.

Now, 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

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

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 $\mathbf{U_z} \in R^{m\times m}$ that is invertible modulo qR, such that its submatrix $\mathbf{U}_\mathbf{z}^{]1[}$ is orthogonal to $\mathbf{z}$, and that $\left\|M_\sigma(\mathbf{U}_\mathbf{z}^{]1[})\right\|_2 \leq 2B_\eta$. The reduction first samples $\mathbf{f} \in K_\mathbb{R}^m$ from the continuous Gaussian distribution of covariance matrix $\alpha^2(4B_\eta^2\mathbf{I}_{mn} - M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})^T) \in \mathbb{R}^{mn\times mn}$. The covariance matrix is well-defined because $\left\|M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})\right\|_2 \leq 2B_\eta$. The reduction then computes $\mathbf{b}' = \mathbf{U_z}\mathbf{b}+\mathbf{f}$ and samples $\mathbf{c}$ from $\mathcal{D}_{q^{-1}R^m-\mathbf{b}',\alpha}$ (as it is efficiently sampleable), and finally gives the following to $\mathcal{O}$

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

Note that this tuple is in $R_q^{m\times k} \times (q^{-1}R/R)^m \times q^{-1}R$, as required. We now prove correctness. First, consider the case where $\mathbf{A}$ is uniformly random over $R_q^{m\times k}$ and $\mathbf{b} = q^{-1}\mathbf{A}\mathbf{s} + \mathbf{e} \bmod R$ for some uniform $\mathbf{s} \in R_q^k$, 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, $\mathbf{e}$ is distributed according $D_{\sqrt{\mathbf{S}}}$ where $\mathbf{S} = \mathrm{diag}(\mathbf{0}_{n\times n}, \alpha^2\mathbf{I}_n, \ldots, \alpha^2\mathbf{I}_n)$. By Lemma 2.14, $\mathbf{U_z}\mathbf{e}$ is then distributed according to $D_{\sqrt{\mathbf{S}'}}$ where $\mathbf{S}' = M_{\sigma_H}(\mathbf{U_z})\mathbf{S}M_{\sigma_H}(\mathbf{U_z})$. Due to the specific form of $\mathbf{S}$, we observe that $\mathbf{S} = M_{\sigma_H}(\mathrm{diag}(0, \alpha^2, \ldots, \alpha^2))$. Using the ring homomorphism property and the form of $\mathbf{S}$, it holds that $\mathbf{S}' = M_{\sigma_H}(\alpha^2\mathbf{U}_\mathbf{z}^{]1[}(\mathbf{U}_\mathbf{z}^{]1[})^T) = \alpha^2 M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})^T$. Hence the vector $\mathbf{U_z}\mathbf{e}+\mathbf{f}$ is distributed as the Gaussian over $K_\mathbb{R}^m$ of covariance matrix $\alpha^2 M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})^T + \alpha^2(4B_\eta^2\mathbf{I}_{mn} - M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})M_{\sigma_H}(\mathbf{U}_\mathbf{z}^{]1[})^T)$ which is identical to $D^m_{\alpha\cdot 2B_\eta}$. Since $q^{-1}\mathbf{A}'\mathbf{s} \in q^{-1}R^m$, the coset $q^{-1}R^m - \mathbf{b}'$ is the same as $q^{-1}R^m - (\mathbf{U_z}\mathbf{e} + \mathbf{f})$, which yields that $\mathbf{c}$ can be seen as being sampled from $\mathcal{D}_{q^{-1}R^m-(\mathbf{U_z}\mathbf{e}+\mathbf{f}),\alpha}$. Since $\alpha \geq \eta_\varepsilon(q^{-1}R^m)$, Lemma 2.11 gives that the distribution of $\mathbf{U_z}\mathbf{e} + \mathbf{f} + \mathbf{c}$ is within statistical distance 8ε of $\mathcal{D}_{q^{-1}R^m,\alpha\sqrt{4B_\eta^2+1}}$, 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 $\mathbf{U}_\mathbf{z}^{]1[}$ 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 $\mathbf{U}_\mathbf{z}^{]1[}$, proving 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. Using that $\alpha \geq \eta_\varepsilon(q^{-1}R^m)$, Lemma 2.10 shows that the distribution of $(\mathbf{A}, \mathbf{b})$ is within statistical distance $\varepsilon/2$ of the distribution of $(\mathbf{A}, \mathbf{e}' + \mathbf{e})$ where $\mathbf{e}' \in (q^{-1}R/R)^m$ 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 $\mathbf{U_z}\mathbf{e}' \in q^{-1}R^m$, then the coset $q^{-1}R^m - \mathbf{b}'$ is identical to $q^{-1}R^m - (\mathbf{U_z}\mathbf{e} + \mathbf{f})$. For the same reasons as above, $\mathbf{U_z}\mathbf{e} + \mathbf{f} + \mathbf{c}$ is distributed as $\mathcal{D}_{q^{-1}R^m, \alpha\sqrt{4B_\eta^2+1}}$ within statistical distance of at most 8ε, and in particular independent of $\mathbf{e}'$. So the third component is correctly distributed again because $\langle \mathbf{z}, \mathbf{U_z}\mathbf{e} \rangle = 0$. Finally, since $\mathbf{e}'$ is independent of the first and third components, and that $\mathbf{U_z}\mathbf{e}'$ is uniform over $(q^{-1}R/R)^m$ as $\mathbf{U_z}$ is invertible modulo qR, it yields that the second component is uniform and independent of the other ones as required. □

The condition on the modulus q in Lemma 3.2 and 3.3 stems from the invertibility result by Lyubashevsky and Seiler [32] stated in Lemma 2.4. Recall that these conditions can be simplified in the case of power-of-two cyclotomic fields as discussed in Remark 2.1.

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 B.2 for completeness.

Lemma 3.4 (Adapted from [17, 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 $q^{-1}R$, and $\mathcal{Z} \subseteq R^m$. There is a PPT 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 $\mathbf{A}$.

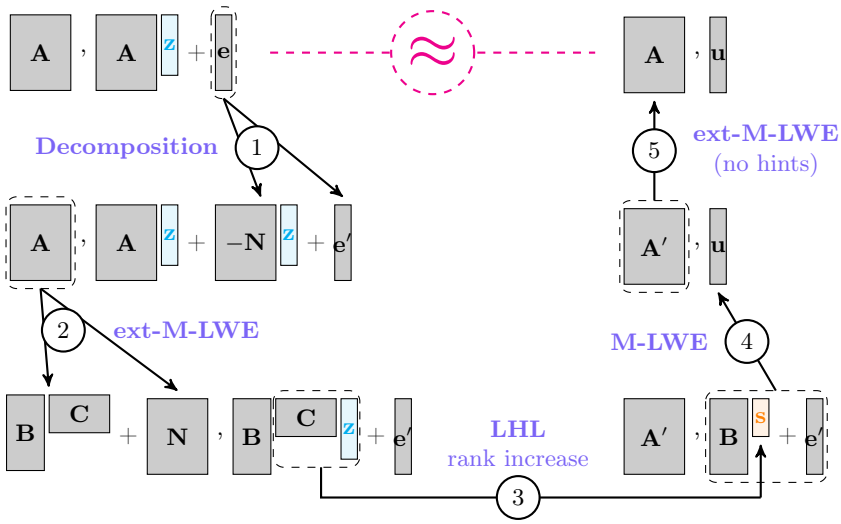

Fig. 3.3. Summary of the proof of Lemma 3.5

Lemma 3.5 (Adapted from [17, 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 $d\log_2(2\eta+1) \geq k\log_2 q + \omega(\log_2 n)$. Let $\varepsilon, \alpha, \gamma, \beta$ be such that $\varepsilon \in (0, 1/2)$, $\alpha \geq \sqrt{2}\eta_\varepsilon(q^{-1}R^d)$, $\gamma = \alpha B_\eta\sqrt{d}$, and $\beta = \alpha B_\eta\sqrt{2d}$, where $B_\eta = \max_{x\in S_\eta}\|\sigma(x)\|_\infty$. There is a PPT reduction from the problems* ext-M-LWE$^m_{n,k,d,q,\psi,S_\eta^d}$, M-LWE$_{n,k,m,q,D_\gamma}$ *and* ext-M-LWE$^m_{n,k,d,q,\psi,\{0\}^d}$ *with $\psi = \mathcal{D}_{q^{-1}R,\alpha}$ to η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$. Hence, for any distinguishing algorithm $\mathcal{A}$ for the η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$ problem, there exist algorithms $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ for* ext-M-LWE$^m_{n,k,d,q,\psi,S_\eta^d}$, M-LWE$_{n,k,m,q,D_\gamma}$ *and* ext-M-LWE$^m_{n,k,d,q,\psi,\{0\}^d}$ *respectively such that the following inequality holds.*

$$\mathrm{Adv}[\mathcal{A}] \leq \mathrm{Adv}[\mathcal{B}_1] + \mathrm{Adv}[\mathcal{B}_2] + \mathrm{Adv}[\mathcal{B}_3] + 2m\varepsilon + \frac{1}{2}\sqrt{\left(1 + \frac{q^k}{(2\eta+1)^d}\right)^n - 1}.$$

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$^m_{n,k,d,q,\alpha,S_\eta^d}$, 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 $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{z} + \mathbf{e} \bmod R)$, with $\mathbf{A} \hookleftarrow U(R_q^{m\times d})$, $\mathbf{z} \hookleftarrow U(S_\eta^d)$ and $\mathbf{e} \in K_\mathbb{R}^m$ sampled from the continuous Gaussian $D^m_\mathbf{r}$ with parameter vector $\mathbf{r}$ with $r_j^2 = \gamma^2 + \alpha^2\sum_i|\sigma_j(z_i)|^2$. We have $\|\mathbf{r}\|_\infty =$

$\sqrt{\gamma^2 + \alpha^2\|\mathbf{z}\|_{2,\infty}^2}$, as well as $\|\mathbf{z}\|_{2,\infty}^2 \leq \sum_{i\in[d]}\|\sigma(z_i)\|_\infty^2$. Recalling the parameter $B_\eta = \max_{x\in S_\eta}\|\sigma(x)\|_\infty$, that can be bounded above by $n\eta$ 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 $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{z} + \mathbf{e} \bmod R)$ 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: Output $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{z} + \mathbf{e} \bmod R)$ as in η-M-LWE$_{n,d,m,q,\Psi_{\leq\beta}}$;
- H_1: Output $(\mathbf{A}, q^{-1}\mathbf{A}\mathbf{z} - \mathbf{N}\mathbf{z} + \mathbf{e}' \bmod R)$, where $\mathbf{N} \hookleftarrow \mathcal{D}_{q^{-1}R,\alpha}^{m\times d}$ and $\mathbf{e}' \hookleftarrow D_\gamma^m$ and $\mathbf{A}, \mathbf{z}$ as in H_0;
- H_2: Output $(\mathbf{A}', q^{-1}\mathbf{A}'\mathbf{z} - \mathbf{N}\mathbf{z} + \mathbf{e}' \bmod R) = (\mathbf{A}', q^{-1}\mathbf{B}\mathbf{C}\mathbf{z} + \mathbf{e}' \bmod R)$, where $\mathbf{B} \hookleftarrow U(R_q^{m\times k})$, $\mathbf{C} \hookleftarrow U(R_q^{k\times d})$, $\mathbf{A}' = q(q^{-1}\mathbf{C}^T\mathbf{B}^T + \mathbf{N}^T \bmod R)^T$, and $\mathbf{N}, \mathbf{z}, \mathbf{e}'$ as in H_1;
- H_3: Output $(\mathbf{A}', q^{-1}\mathbf{B}\mathbf{s} + \mathbf{e}' \bmod R)$, where $\mathbf{s} \hookleftarrow U(R_q^k)$, and $\mathbf{A}', \mathbf{B}, \mathbf{e}'$ as in H_2;
- H_4: Output $(\mathbf{A}', \mathbf{u})$, where $\mathbf{u} \hookleftarrow U(\mathbb{T}^m)$, and $\mathbf{A}'$ as in H_3;
- H_5: Output $(\mathbf{A}, \mathbf{u})$, where $\mathbf{A} \hookleftarrow U(R_q^{m\times d})$ and $\mathbf{u}$ as in H_4.

From H_0 to H_1: We first claim that $\Delta([-\mathbf{N}\mathbf{z} + \mathbf{e}']_i, \mathbf{e}_i) \leq 2\varepsilon$ for all $i \in [m]$. Indeed, $(1/\alpha^2 + \|\mathbf{z}\|_{2,\infty}^2/\gamma^2)^{-1/2} \geq \alpha/\sqrt{2}$ and $\alpha/\sqrt{2} \geq \eta_\varepsilon(q^{-1}R^d)$. If $\mathbf{n}_i \in q^{-1}R^d$ denotes the i-th row of $\mathbf{N}$, Lemma 2.15 yields the claim since we have $[-\mathbf{N}\mathbf{z} + \mathbf{e}']_i = \langle \mathbf{n}_i, -\mathbf{z}\rangle + e_i'$, thus giving $\Delta(-\mathbf{N}\mathbf{z} + \mathbf{e}', \mathbf{e}) \leq 2m\varepsilon$.

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

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$^m_{n,k,d,q,\alpha,S_\eta^d}$ 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, $(\mathbf{C}^T, q^{-1}\mathbf{A}^T, \mathbf{N}\mathbf{z})$ can be efficiently transformed into a sample from H_1. Note that $q^{-1}\mathbf{A}^T$ indeed corresponds to the uniform case of ext-M-LWE, because $\mathbf{A}$ is uniform over R_q and $q^{-1}R_q$ can be seen as $q^{-1}R/R$. 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 $d \times m$ as required. In the other case, if we apply the same transformation to the ext-M-LWE sample $(\mathbf{C}^T, q^{-1}\mathbf{C}^T\mathbf{B}^T + \mathbf{N}^T \bmod R, \mathbf{N}\mathbf{z})$ where $\mathbf{B}^T$ and $\mathbf{N}^T$ are the secret and error matrix respectively, it leads to a sample from H_2. The (randomized) transformation can be described by sampling $\mathbf{e}'$ from D_γ^m and outputting $f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{x}_3) = (q\mathbf{X}_2^T, \mathbf{X}_2^T\mathbf{z} - \mathbf{x}_3 + \mathbf{e}' \bmod R)$. Hence, $\mathcal{B}_1$ is a distinguisher for ext-M-LWE$^m_{n,k,d,q,\alpha,S_\eta^d}$, and

$$|\mathbb{P}[\mathcal{A}(H_1) = 1] - \mathbb{P}[\mathcal{A}(H_2) = 1]| = \mathrm{Adv}[\mathcal{B}_1]. \tag{2}$$

From H_2 to H_3: The Ring Leftover Hash Lemma stated in Lemma 2.8 yields that $(\mathbf{C}, \mathbf{C}\mathbf{z})$ is within statistical distance at most $\delta = \frac{1}{2}\sqrt{(1 + q^k/(2\eta+1)^d)^n - 1}$

from $(\mathbf{C}, \mathbf{s})$. Note that the condition $d \log_2(2\eta + 1) \geq k \log_2 q + \omega(\log_2 n)$ implies $\delta \leq n^{-\omega(1)}$. This yields

$$|\mathbb{P}[\mathcal{A}(H_2) = 1] - \mathbb{P}[\mathcal{A}(H_3) = 1]| \leq \delta. \tag{3}$$

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 $(\mathbf{B}, \mathbf{u})$ into a sample from H_4 in the uniform case, and $(\mathbf{B}, q^{-1}\mathbf{B}\mathbf{s} + \mathbf{e}' \bmod R)$ into a sample from H_3 in the M-LWE case. The transformation is given by $g(\mathbf{X}_1, \mathbf{x}_2) = (\mathbf{X}_1\mathbf{C} + q\mathbf{N} \bmod qR, \mathbf{x}_2)$, where $\mathbf{C}, \mathbf{N}$ are sampled as in H_2. Therefore, $\mathcal{B}_2$ is a distinguisher for M-LWE$_{n,k,m,q,\gamma}$ such that

$$|\mathbb{P}[\mathcal{A}(H_3) = 1] - \mathbb{P}[\mathcal{A}(H_4) = 1]| = \mathrm{Adv}[\mathcal{B}_2]. \tag{4}$$

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 $(\mathbf{C}^T, q^{-1}(\mathbf{A}')^T, \mathbf{N} \cdot \mathbf{0})$ into a sample from H_4 (M-LWE case) and $(\mathbf{C}^T, q^{-1}\mathbf{A}^T, \mathbf{N} \cdot \mathbf{0})$ into a sample from H_5 (uniform case). The transformation samples $\mathbf{u} \hookleftarrow U(\mathbb{T}^m)$ as in H_4 and outputs $h(\mathbf{X}_1, \mathbf{X}_2, \mathbf{x}_3) = (q\mathbf{X}_2^T, \mathbf{u})$. We then get

$$|\mathbb{P}[\mathcal{A}(H_4) = 1] - \mathbb{P}[\mathcal{A}(H_5) = 1]| = \mathrm{Adv}[\mathcal{B}_3]. \tag{5}$$

Putting Equations (1), (2), (3), (4), (5) altogether yields the result. □

4 Computational 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 S_η^m instead of Gaussian as in the standard formulation of M-LWE. The overall proof strategy follows the idea of Micciancio and Peikert [38], 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 concrete 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 the M-LWE and M-ISIS function families which allows us to switch from one to other at essentially no cost. We then prove our result in terms of M-ISIS 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-ISIS

In the following, we adapt the duality results from [36, Sec. 4.2] to the module setting. To the best of our knowledge, this hasn't been formally done before. For completeness, we detail the proofs in Appendix B.3. The idea when going from M-LWE to M-ISIS 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-ISIS instance with matrix $\mathbf{B}' = \mathbf{B}\mathbf{U}$ where $\mathbf{U}$ simply randomizes $\mathbf{B}$. Note that in this paper we are considering a parameter regime such that the function family of M-ISIS is injective. In other words, solutions to M-ISIS are with a very high probability unique. This regime is sometimes referred to as *low-density* ISIS [29] or even more generally as a knapsack problem [36]. 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 upper bound derived from Lemma 2.6 for this singularity probability requires q to be unramified in order to have an easier characterization of units of R_q. Also, note that the following lemmas are only meaningful if the extra losses incurred by $\delta(\cdot,\cdot)$ are negligible, which may require to restrict the splitting of q. We elaborate on the matter in Section 4.2.3.

Lemma 4.1 (Adapted from [36, 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^m. If $(\text{M-LWE}(n, d, m, q, R^m), \mathcal{X})$ is ε-uninvertible (resp. one-way, pseudorandom), then $(\text{M-ISIS}(n, m-d, m, q, R^m), \mathcal{X})$ is ε'-uninvertible (resp. one-way, pseudorandom), with $\varepsilon' = \delta(m, m-d) + \varepsilon/(1-\delta(m,d))$ (resp. $\varepsilon' = 2\delta(m, m-d) + \varepsilon/(1-\delta(m,d))$ for pseudorandomness).*

Lemma 4.2 (Adapted from [36, 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^m. If $(\text{M-ISIS}(n, m-d, m, q, R^m), \mathcal{X})$ is ε-uninvertible (resp. one-way, pseudorandom), then $(\text{M-LWE}(n, d, m, q, R^m), \mathcal{X})$ is ε'-uninvertible (resp. one-way, pseudorandom), with $\varepsilon' = \delta(m, d) + \varepsilon/(1-\delta(m, m-d))$ (resp. $\varepsilon' = 2\delta(m, d) + \varepsilon/(1-\delta(m, m-d))$ for pseudorandomness).*

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, assuming the pseudorandomness of the M-LWE function family with Gaussian input. It therefore implies the hardness of search M-LWE with small uniform error from that of the decision version of M-LWE with Gaussian error. To prove the one-wayness of the M-LWE function, we prove the result in terms of M-ISIS and use Lemma 4.2 to conclude. Recall that by Lemma 2.16, it suffices to prove that M-ISIS is uninvertible and second preimage resistant with respect to this specific input distribution. We actually prove the second preimage resistance of the M-ISIS

function, and the uninvertibility of a decomposition of the M-ISIS function. We then argue that these two function families are indistinguishable based on the pseudorandomness of M-ISIS (or M-LWE equivalently). The idea of the proof is summarized in Figure 4.1.

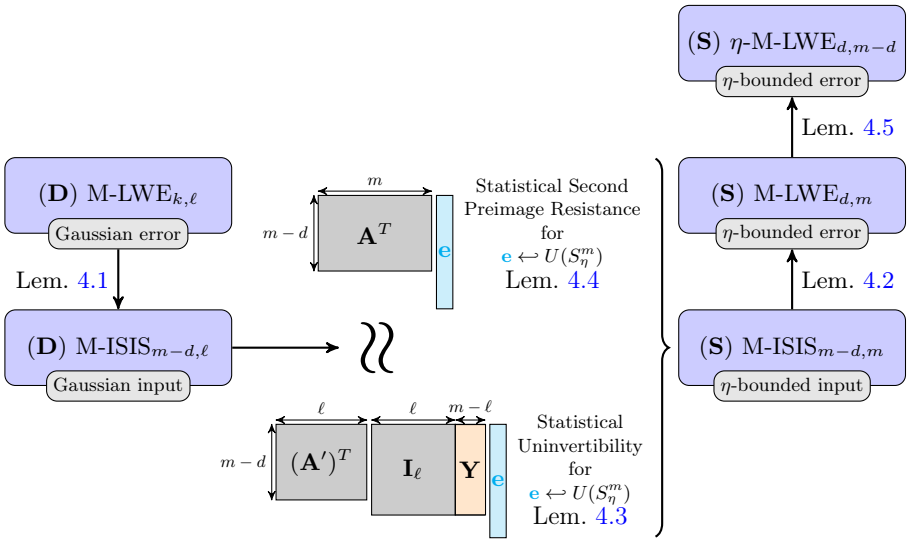

Fig. 4.1. Summary of the proof of Theorem 4.1. **S** denotes the search version, while **D** denotes the decision version. The first subscript for M-LWE denotes the rank, while the second subscript denotes the number of samples. For clarity, we removed, for both M-LWE and M-ISIS, the subscripts for the 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 $(\text{M-ISIS}(n, m-d, m, q, R^m), U(S_\eta^m))$, we decompose it into a linear (Gaussian) function family $\mathcal{L}$ and a smaller M-ISIS$(n, m-d, \ell, q, R^\ell)$ function family with $\ell \leq m$. By Lemma 2.18, it suffices to prove the uninvertibility of $(\mathcal{L}, U(S_\eta^m))$. We first define what we mean by linear (Gaussian) function.

Definition 4.1. *Let K be a number field, and R its ring of integers. Let ℓ, m be positive integers such that $m \geq \ell$, $\alpha > 0$, and $X \subseteq R^m$. We define the function family $\mathcal{L}(\ell, m, \alpha, X)$ obtained by sampling $\mathbf{Y}$ from $\mathcal{D}_{R,\alpha}^{\ell \times (m-\ell)}$, and outputting $h_{\mathbf{Y}} : X \to R^\ell$ 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 2.17 to prove that $(\mathcal{L}(\ell, m, \alpha, X), U(X))$ is statistically uninvertible with uniform inputs for carefully chosen parameters. In particu-

lar, the result is only meaningful when ε_3 is 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 $\alpha > 0$. Let η be a positive integer and $X \subseteq S_\eta^m$. We define the function family $\mathcal{F} = \text{M-ISIS}(n, m-d, \ell, q, R^\ell) \circ \mathcal{L}(\ell, m, \alpha, X)$. Then, for any $t \geq 0$, $(\mathcal{F}, U(X))$ is (statistically) ε_3-uninvertible for*

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

where $C > 0$ is an absolute constant (empirically $C \approx 1/\sqrt{2\pi}$). When $t = \omega(\sqrt{\log_2 n})$, the second term is negligible.

Proof. We first bound $\mathbb{E}_{h_\mathbf{Y} \sim \mathcal{L}}[|h_\mathbf{Y}(X)|]$ and use Lemma 2.17 to conclude. Let $h_\mathbf{Y}$ be sampled from $\mathcal{L}(\ell, m, \alpha, X)$. Let $\mathbf{x} = [\mathbf{x}_1^T \mid \mathbf{x}_2^T]^T \in X$, with $\mathbf{x}_1 \in S_\eta^\ell$, and $\mathbf{x}_2 \in S_\eta^{m-\ell}$. Then, $h_\mathbf{Y}(\mathbf{x}) = \mathbf{x}_1 + \mathbf{Y}\mathbf{x}_2$. As seen in Section 2.1, it holds that $\tau(h_\mathbf{Y}(\mathbf{x})) = \tau(\mathbf{x}_1) + M_\tau(\mathbf{Y})\tau(\mathbf{x}_2)$, and therefore

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

Since $\mathbf{x}_1$ and $\mathbf{x}_2$ are vectors over S_η, it holds that $\|\tau(\mathbf{x}_1)\|_2 \leq \eta\sqrt{n\ell}$ and that $\|\tau(\mathbf{x}_2)\|_2 \leq \eta\sqrt{n(m-\ell)}$. 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 discrete Gaussian of parameter α, Lemma 2.9 gives that for all $k \in [n]$ and all $t \geq 0$

$$\mathbb{P}_{\mathbf{Y} \sim \mathcal{D}_{R,\alpha}^{\ell \times (m-\ell)}} \left[\|\sigma_k(\mathbf{Y})\|_2 > C\alpha(\sqrt{\ell} + \sqrt{m-\ell} + t) \right] \leq 2e^{-\pi t^2},$$

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

$$\mathbb{P}_{\mathbf{Y} \sim \mathcal{D}_{R,\alpha}^{\ell \times (m-\ell)}} \left[\|M_\tau(\mathbf{Y})\|_2 > C\alpha(\sqrt{\ell} + \sqrt{m-\ell} + t) \right] \leq 2n \cdot e^{-\pi t^2}.$$

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

$$r = \sqrt{n}\eta \left(\sqrt{\ell} + C\alpha\sqrt{m-\ell}(\sqrt{\ell} + \sqrt{m-\ell} + t) \right).$$

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

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

As the bound is independent of $\mathbf{Y}$, let us temporarily denote it by B. We also define $S = \{\mathbf{Y} \in R^{\ell\times(m-\ell)} : \|M_\tau(\mathbf{Y})\| \leq C\alpha(\sqrt{\ell}+\sqrt{m-\ell}+t)\}$, and S' its complement in $R^{\ell\times(m-\ell)}$. We then have

$$\begin{aligned}\mathbb{E}[|h_{\mathbf{Y}}(X)|] &= \sum_{\mathbf{Y}'\in S}\mathbb{P}_{\mathbf{Y}}[\mathbf{Y}=\mathbf{Y}']|h_{\mathbf{Y}'}(X)| + \sum_{\mathbf{Y}'\in S'}\mathbb{P}_{\mathbf{Y}}[\mathbf{Y}=\mathbf{Y}']|h_{\mathbf{Y}'}(X)| \\ &\leq B\cdot\mathbb{P}_{\mathbf{Y}}[\mathbf{Y}\in S] + |X|\cdot\mathbb{P}_{\mathbf{Y}}[\mathbf{Y}\in S'] \\ &\leq B + |X|\cdot 2ne^{-\pi t^2},\end{aligned}$$

where the first inequality follows from the above calculations and the fact that for $\mathbf{Y}' \in S'$, we have the trivial bound $|h_{\mathbf{Y}'}(X)| \leq |X|$. Lemma 2.17 then yields the ε_3-uninvertibility of $\mathcal{L}$, with $\varepsilon_3 = B/|X| + 2ne^{-\pi t^2}$. By Lemma 2.18, we thus obtain the ε_3-uninvertibility of $\mathcal{F}$. □

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

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

$$\varepsilon_4 = (|X|-1)\cdot\left(\frac{B_{2\eta}}{q}\right)^{nk},$$

where $B_{2\eta} = \max_{x\in S_{2\eta}}\|\sigma(x)\|_\infty$.

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 $\mathbf{A}^T\mathbf{x}' = \mathbf{A}^T\mathbf{x} \bmod qR$ is less than ε_4, namely

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

Using the total probability formula and the union bound on $\mathbf{x}'$, we have the following.

$$\begin{aligned} p &= \sum_{\mathbf{x}^* \in X} \mathbb{P}_{\mathbf{x}}[\mathbf{x} = \mathbf{x}^*] \cdot \mathbb{P}_{\mathbf{A},\mathbf{x}}[\exists \mathbf{x}' \in X \setminus \{\mathbf{x}\},\ \mathbf{A}^T\mathbf{x}' = \mathbf{A}^T\mathbf{x} \bmod qR | \mathbf{x} = \mathbf{x}^*] \\ &= \sum_{\mathbf{x}^* \in X} |X|^{-1} \cdot \mathbb{P}_{\mathbf{A}}[\exists \mathbf{x}' \in X \setminus \{\mathbf{x}^*\},\ \mathbf{A}^T(\mathbf{x}' - \mathbf{x}^*) = \mathbf{0} \bmod qR] \\ &\leq |X|^{-1} \sum_{\mathbf{x}^* \in X} \sum_{\mathbf{x}' \in X \setminus \{\mathbf{x}^*\}} \mathbb{P}_{\mathbf{A}}[\mathbf{A}^T(\mathbf{x}' - \mathbf{x}^*) = \mathbf{0} \bmod qR]. \end{aligned}$$

Let $\mathbf{x}^* \in X$, $\mathbf{x}' \in X \setminus \{\mathbf{x}^*\}$, and set $\mathbf{z} = \mathbf{x}' - \mathbf{x}^*$. Then, by [34, Lem. 4.4], $\mathbf{A}^T\mathbf{z} \bmod qR$ is uniformly distributed in $(\mathcal{I}_{\mathbf{z}}/qR)^k$ over the randomness of $\mathbf{A}$, 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 $i \in [m]$, $\langle z_i \rangle \subseteq \mathcal{I}_{\mathbf{z}}$, meaning that $N(\mathcal{I}_{\mathbf{z}})$ 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 we have $\mathbf{z} \in \{\mathbf{a} - \mathbf{b}; (\mathbf{a}, \mathbf{b}) \in X^2\} \subseteq S_{2\eta}^m$. It thus holds

$$N(\langle z_i \rangle) = |N(z_i)| = \prod_{j \in [n]} |\sigma_j(z_i)| \leq B_{2\eta}^n,$$

by Lemma 2.1 where $B_{2\eta} = \max_{x \in S_{2\eta}} \|\sigma(x)\|_\infty$. Recall that in cyclotomic fields we have $B_{2\eta} \leq 2\eta n$. Hence $\mathbb{P}_{\mathbf{A}}[\mathbf{A}^T\mathbf{z} = \mathbf{0} \bmod qR] \leq (B_{2\eta}/q)^{nk}$. Going back to our original calculation, we then have $p \leq |X|^{-1}|X|(|X| - 1) \cdot (B_{2\eta}/q)^{nk} = \varepsilon_4$ which concludes the proof. □

Remark 4.1. For common choices of n, m and prime q, we heuristically observe that the ideals $\langle z_1 \rangle, \ldots, \langle z_m \rangle, \langle q \rangle$ are relatively prime with high probability, which means that $\mathcal{I}_{\mathbf{z}} = R$ in the proof above. In this case, $N(\mathcal{I}_{\mathbf{z}}) = 1$ which yields a much better bound on the probability. Since the probability sums over all the possible $\mathbf{x}'$, one would need to evaluate the proportion of $\mathbf{z}$ generated as above that verify $\mathcal{I}_{\mathbf{z}} = R$. We leave it as an open problem.

For example, consider the case of cyclotomic fields. By Lemma 2.4 (or Remark 2.1 for power-of-two conductors), if q splits into few factors and is large enough with respect to η so that $S_{2\eta} \setminus \{0\} \bmod qR \subset R_q^\times$, then we have that for all $\mathbf{z} \in S_{2\eta}^m \setminus \{\mathbf{0}\}$, $\mathcal{I}_{\mathbf{z}} = R$. Indeed, for such $\mathbf{z}$, there exists $i \in [m]$ such that $z_i \in S_{2\eta} \setminus \{0\}$ and therefore $z_i \bmod qR \in R_q^\times$. This implies that $\langle z_i \rangle + \langle q \rangle = R$ and as a result $\mathcal{I}_{\mathbf{z}} = R$. Hence, we can improve the bound on the probability to $\varepsilon_4 = (|X| - 1)/q^{nk}$ if we accept to enforce a specific splitting on q.

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 number field of degree n, and R its ring of integers. Let d, m, k be positive integers such that $m > d \geq k \geq 1$. Let q be an unramified prime such that $\min_{i\in[\kappa]} N(\mathfrak{p}_i)^{\min(m-d,k)+1} \geq n^{\omega(1)}$, where the $\mathfrak{p}_i$'s are the prime ideal factors of $\langle q\rangle$. Let η be a positive integer, and $X \subseteq S_\eta^m$. We define $\ell = m - d + k$. Assume that the function family $(\text{M-LWE}(n, k, \ell, q, R^\ell), \mathcal{D}^\ell_{R,\alpha})$ is ε_1-pseudorandom for $\alpha > 0$. Then $(\text{M-LWE}(n, d, m, q, X), U(X))$ is ε-one-way for*

$$\varepsilon = \delta(m, d) + \frac{(m-\ell)(2\delta(\ell, m-d) + \varepsilon_1/(1-\delta(\ell, k))) + \varepsilon_3 + \varepsilon_4}{1-\delta(m, m-d)},$$

where $\varepsilon_3, \varepsilon_4$ are defined in the statement of Lemma 4.3 and 4.4 respectively.

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

$$\varepsilon_2 = 2\delta(\ell, \ell-k) + \frac{\varepsilon_1}{1-\delta(\ell, k)}.$$

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-ISIS with respect to $\mathcal{D}^\ell_{R,\alpha}$, 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

$$\varepsilon_4 = (|X| - 1) \cdot \left(\frac{B_{2\eta}}{q}\right)^{n(m-d)}.$$

We thus have that $(\mathcal{F}, \mathcal{G}, U(X))$ is a lossy function family, depending on $\varepsilon_2, \varepsilon_3, \varepsilon_4$. Lemma 2.16 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

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

Combining everything, we get

$$\varepsilon = \delta(m,d) + \frac{(m-\ell)(2\delta(\ell,m-d)+\varepsilon_1/(1-\delta(\ell,k)))+\varepsilon_3+\varepsilon_4}{1-\delta(m,m-d)},$$

which yields the claim. The condition on q ensures that all the $\delta(\cdot,\cdot)$ are negligible. Indeed, as noted after Lemma 2.6, if the smallest norm N of the prime ideal factors is such that $N^{a-b+1} \geq n^{\omega(1)}$ for $a \geq b$, then $\delta(a,b) \leq n^{-\omega(1)}$. The condition on q thus yields that $\delta(m,d), \delta(m,m-d), \delta(\ell,m-d), \delta(\ell,k)$ are negligible. □

Let us now discuss the various conditions that are needed to apply this theorem in the context of cyclotomic fields. The lower bound on q comes from ensuring that ε_4 is negligible. Indeed, we have that $|X| = (2\eta+1)^{nm}$, and therefore it suffices to have

$$(2\eta+1)^m \left(\frac{2n\eta}{q}\right)^{m-d} < 1, \tag{6}$$

which can be written as $q > 2n\eta \cdot (2\eta+1)^{m/(m-d)}$. Hence, for $\lambda > 0$ one can choose $q > 2^{\lambda/(m-d)} \cdot 2n\eta \cdot (2\eta+1)^{m/(m-d)}$ which ensures $\varepsilon_4 < 2^{-\lambda n}$. Once this lower bound on q is set, one can easily find the closest prime q with an appropriate splitting as required by the theorem.

The expression of ε_3 is more involved, but the idea is the same. For it to be negligible, we need

$$\frac{\eta^\ell}{(2\eta+1)^m(\pi n\ell)^{1/2n}} \left(\sqrt{2\pi e}\left(1 + C\alpha\sqrt{\frac{m-\ell}{\ell}}\left(\sqrt{\ell}+\sqrt{m-\ell}+t\right)\right)\right)^\ell < 1, \tag{7}$$

where $t = \omega(\sqrt{\log_2 n})$. Due to the many dependencies in m, k, d and η, it is harder to extract a closed-form 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 (6) and (7) 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 modules). It turns out that the condition is not met for all sets of parameters, and η cannot be arbitrarily small for arbitrary 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 ideals ($k = 1$).

Remark 4.2. Note that we can provide the asymptotic behavior $\varepsilon_3 = O(\alpha \cdot m \cdot \eta \cdot t/\sqrt{\ell})^{n\ell}/|X| + 2ne^{-\pi t^2}$, 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

n	k	d	m	η	q	ε_3	ε_4
256	1	11	12	1	$\approx 2^{31}$	$\approx 2^{-496} + 2^{-281}$	2^{-256}
256	2	14	16	10	$\approx 2^{48}$	$\approx 2^{-360} + 2^{-281}$	2^{-256}

Table 4.1. Example parameter sets reaching the conditions of Theorem 4.1. We take $t = \log_2 n$, and $\alpha \approx \log_2 n$ if $k = 1$ and $\alpha \approx 2\sqrt{k}\log_2 n$ if $k > 1$. Empirically, we have $C \approx 1/\sqrt{2\pi}$ as noticed for example in [37, Sec. 2.4]. The loss of 2^{-281} that dominates in the value of ε_3 comes from the spectral bound loss $2n \cdot e^{-\pi t^2}$. For such parameters, one can determine the appropriate splitting of q for the $\delta(\cdot,\cdot)$ to be negligible. For the first set of parameters, if q splits into $n/2$ factors, all the $\delta(\cdot,\cdot)$ will be less than 2^{-117}, but if it splits into $n/4$ factors, they will be less than 2^{-242}. For the second set of parameters, if q is fully splitted, they can all be bounded by 2^{-136} and if q splits into $n/2$ factors, they will be less than 2^{-281}.

values of η. The asymptotic approach gives a more direct condition on m. Indeed, taking $\alpha = 2\sqrt{k}t$, and denoting by C' the asymptotic constant, we have

$$O(s \cdot m \cdot \eta \cdot t/\sqrt{\ell})^{n\ell}/|X| \leq \left(\frac{(2C'mt^2\eta\sqrt{k/\ell})^\ell}{(2\eta+1)^m}\right)^n.$$

Since $\ell > k$, we can choose the parameters to have $(2C'mt^2\eta)^\ell/(2\eta+1)^m \leq 1/2$ to have a exponentially small loss. We thus have

$$\begin{aligned}
&(2C'mt^2\eta)^\ell/(2\eta+1)^m \leq 1/2 \\
&\quad \Leftrightarrow (m-(d-k))\log_2(2C'mt^2\eta) \leq m\log_2(2\eta+1) - 1 \\
&\quad \Leftrightarrow m(\log_2(2C'mt^2\eta) - \log_2(2\eta+1)) \leq (d-k)\log_2(2C'mt^2\eta) - 1 \\
&\quad \Leftrightarrow m \leq (d-k)(1+\log_2(2\eta+1)/\log_2(2C'mt^2\eta/(2\eta+1))) \\
&\qquad\qquad - 1/\log_2(2C'mt^2\eta/(2\eta+1)).
\end{aligned}$$

This leads to a condition on m which is

$$d < m \leq (d-k)\left(1 + \frac{\log_2(2\eta+1)}{\log_2(2C' \cdot m \cdot t^2/3)}\right),$$

which is much similar to the condition in [38]. The main difference stems from the fact that m is no longer our asymptotic parameter, which explains the presence of $t^2 = \omega(\log_2 n)$. 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. Regardless, if we aim at small values of η, we see that we obtain $m \leq d(1+o(1))$. This bound seems true even with our non-asymptotic analysis.

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 4.1 with a Hermite Normal Form transformation for M-LWE. Langlois and Stehlé [24, Lem. 4.24] proposed an immediate generalization of the reduction from LWE to its Hermite Normal Form by Applebaum et al. [5] to modules. In particular, it relies on the fact that if one has access to sufficiently many M-LWE samples $(\mathbf{a}_i, b_i)$, they can find a subset of the $\mathbf{a}_i$ that form a matrix in $GL_d(R_q)$. As our proof of Section 4.2 seemingly limits the number of available samples, it is relevant for us to understand the trade-off between the quality of the reduction (in terms of loss in advantage) and the number of initial samples. More precisely, if one is limited to use $m' \geq d$ samples to construct this invertible matrix, it comes down to evaluating $\delta'(m', d)$. The proof can be found in Appendix B.3 for completeness.

Lemma 4.5 (Adapted from [5,24]). *Let K be a number field, 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^d and ψ a distribution over R. 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 $T(U(R_q^d \times R_q)) = U(R_q^d \times R_q)$. T can be constructed in polynomial time using m' samples from $\mathcal{D} \in \{A_{\mathbf{s},\psi}, U(R_q^d \times R_q)\}$ 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^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 because 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 $m = m' + m''$ samples. The choice of m' allows for tweaking the success probability of the reduction, but at the expense of requiring more samples. Let us now discuss the loss $\delta'(m', d)$. In the case of integers, $\mathbb{Z}_q$ is generally a field which yields a closed-form expression of this probability. Unfortunately, in the case of R_q, it becomes anything but trivial as explained in Appendix A. We can still obtain the following bound

$$\delta'(m', d) \leq \delta'(d, d)^{\lfloor m'/d \rfloor} = \left(1 - \prod_{\ell=0}^{d-1} \prod_{i \in [\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right)\right)^{\lfloor m'/d \rfloor},$$

which we briefly prove here.

Proof. For each $j \in [\lfloor m'/d \rfloor]$, we define $S_j = \{(j-1)d+1, \ldots, jd\}$ which is of size d. We then have

$$\begin{aligned}
\delta'(m',d) &= \mathbb{P}_{(\mathbf{a}_i)_{i\in[m']}\sim U(R_q^d)^{m'}}\left[\bigcap_{S\subseteq[m'],|S|=d} \{(\mathbf{a}_i)_{i\in S} \text{ are not } R_q\text{-l. i.}\}\right] \\
&\leq \mathbb{P}_{(\mathbf{a}_i)_{i\in[m']}\sim U(R_q^d)^{m'}}\left[\bigcap_{j\in[\lfloor m'/d \rfloor]} \{(\mathbf{a}_i)_{i\in S_j} \text{ are not } R_q\text{-l. i.}\}\right] \\
&= \prod_{j\in[\lfloor m'/d \rfloor]} \mathbb{P}_{(\mathbf{a}_i)_{i\in S_j}\sim U(R_q^d)^d}[(\mathbf{a}_i)_{i\in S_j} \text{ are not } R_q\text{-l. i.}] \\
&= \delta'(d,d)^{\lfloor m'/d \rfloor},
\end{aligned}$$

where the first inequality is just an inclusion of events, the second equality is by independence of the $\mathbf{a}_i$'s as none of the S_j overlap, and the last equality follows from the definition of $\delta'(d,d)$. □

We note that $\delta'(d,d)$ highly depends on the size and splitting of q as it is essentially dominated by $\frac{1}{\min_{i\in[\kappa]} N(\mathfrak{p}_i)}$. Hence, depending on the splitting of q, we would need to take $m' = Cd$ with C sufficiently large to make $\delta'(m',d)$ negligible. Our bound is however not tight and we expect $\delta'(m',d)$ to decrease much faster when m' grows. Unfortunately, we were not able to find a better bound on $\delta'(m',d)$ which would support this conjecture. We leave it as an interesting open problem.

As concrete examples, let us take the parameters of Table 4.1. If $q \approx 2^{31}$ splits into $n/4 = 64$ factors of inertia degree 4, for $d = 11$, we have $\delta'(d,d) \approx 2^{-118}$. As a result, $m' = 2d$ ensures $\delta'(m',d) \leq 2^{-236}$. However, if q splits into $n/2 = 128$ factors of inertia degree 2, we now have $\delta'(d,d) \approx 2^{-55}$, and we would thus need to take a larger m'.
We then obtain the following result by combining Theorem 4.1 with Lemma 4.5.

Corollary 4.1. *Let K be a number field of degree n, and R its ring of integers. Let d, m, k be positive integers such that $m > d \geq k \geq 1$. Let q be an unramified prime such that $\min_{i\in[\kappa]} N(\mathfrak{p}_i)^{\min(m-d,k)+1} \geq n^{\omega(1)}$, where the $\mathfrak{p}_i$'s are the prime ideal factors of $\langle q \rangle$. Let m' be a positive integer such that $m > m' \geq d$ and $\delta'(m',d) \leq n^{-\omega(1)}$. Then, let η be a positive integer, and $X \subseteq S_\eta^m$. We define $\ell = m - d + k$. Assuming that the decision version* M-LWE$_{n,k,q,\ell,\mathcal{D}_{R,\alpha}}$ *is ε_1-hard for $\alpha > 0$, it holds that search version* η-M-SLWE$_{n,d,q,m-m',U(T_\eta)}$ *is ε'-hard for*

$$\varepsilon' = \delta'(m',d) + \delta(m,d) + \frac{(m-\ell)(2\delta(\ell,m-d) + \varepsilon_1/(1-\delta(\ell,k))) + \varepsilon_3 + \varepsilon_4}{1-\delta(m,m-d)},$$

where $\varepsilon_3, \varepsilon_4$ are defined in the statement of Lemma 4.3 and 4.4 respectively.

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 [27,28], combinatorial [55,9,23] or algebraic [6,3] 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 ($\mathbf{e} \in \{-\eta, ..., \eta\}^m$), the Arora and Ge attack [6] solves the problem in polynomial time if $m \approx \binom{d+2\eta+1}{2\eta+1} = \Omega(d^{2\eta+1})$, where d is the LWE dimension. For $\eta = 1$, the attack becomes exponential for $m = O(d)$. It has been refined in [3] to obtain subexponential attacks whenever $m = \Omega(d \log_2 \log_2 d)$ in the uncentered binary case ($\{0, 1\}$). 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 $nm = \Omega((nd)^{2\eta+1})$ and therefore $m = \Omega(n^{2\eta} d^{2\eta+1})$. In practical schemes relying on M-LWE with small errors [10,18], the rank d is a small constant and n drives the security parameter. Additionally, we saw in Section 4.3 that roughly $m = m' + m''$ 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 $m = m' + m'' = O(d) \ll n^{2\eta} d^{2\eta+1}$. 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 can still be reduced 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 [16] for power-of-two cyclotomic fields, its hardness was first studied by Langlois and Stehlé in [24]. They established the hardness of the standard formulation M-LWE$_{n,d,m,q,\Upsilon_\alpha}$ based on the quantum hardness of Mod-GIVP$_\gamma^{\eta_\varepsilon}$ (which is a generalized version of SIVP$_\gamma$), where $\alpha = \tilde{\Omega}(d\sqrt{n}/\gamma)$. 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 M-LWE$_{n,d,m,q,\Upsilon_\alpha}$ is hard, then so is 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 $\alpha' \geq \alpha \cdot \max(1, q/p) \cdot n^{3/4}\sqrt{d}\omega(\log_2^2 n)$. Then, combined with [12, Lem. 13], one can obtain the hardness of M-LWE with error in $\mathcal{D}_{R^\vee,\alpha' q}$ for $\alpha' = \alpha\omega(\log_2 n)$ instead of Υ_α, and where $R^\vee$ is the dual of R. The quantum reduction from [24] was later used in [4] to derive the hardness of M-LWR (Module Learning With Rounding) which we do not cover in this discussion.

As discussed in [16], 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 [1] 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 M-LWE$_{n,d,m,q,D_\alpha}$ to 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 [1,2] for the detailed expression of Δ. The reduction to R-LWE was later improved and generalized by Wang and Wang [56] to hold over all cyclotomic fields. A revision of the work by Albrecht and Deo, which can be found in [2], further improved this line of work with a new analysis. Additionally, a result from Peikert and Pepin [44] 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 [1], they provide a degree-rank trade-off, where the underlying ring structure is changed, while preserving the modulus q. Note that, in contrast to [1], 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 [24], which extended the reduction by Applebaum et al. [5] to modules. In particular, combined with their main proof of hardness, it can be used to obtain the hardness of 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 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. [21] and by Brakerski et al. [17] respectively, which are the analog results for LWE. The prior version

from [12] 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 [24], with the caveat of introducing further restrictions on the parameters. More precisely, the result of our previously published work [12] proves the classical hardness of M-LWE. It yields a classical reduction from Mod-GapSVP$_\gamma$ in module lattices of rank nk to M-LWE$_{n,d,m,p,\Psi_{\leq\alpha}}$, where $d \geq k^2 n/2 + \Omega(\log_2 n)$.

Another line of work studied by Brakerski and Döttling for LWE [14] and R-LWE [15] was recently extended to M-LWE by Lin et al. [26]. It looks at the hardness of the problem when the only requirement on the secret distribution is to contain a sufficient entropy. Although [15] cannot be instantiated for η-bounded secret with η being a small constant, the result by [26] 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., $d \log_2(2\eta + 1) \gtrsim k \log_2 q$. 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 4.1 to apply. As mentioned, the algebraic attacks, e.g. [6], 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. The achievable rank d depends on the secret distribution (for [26], Sections 3.1 and 3.2) or on the error size (for Section 4).

	[24] (quantum)	[12] (classical)	[26]	Sec. 3.1	Sec. 3.2	Sec. 4
Field K of degree n	Cyclo	Cyclo	All	Monogenic	Cyclo	All
Rank d	All	$\Omega(n)$	(Depends)	$\Omega(\log_2 n)$	$\omega(\log_2 n)$	All[1]
Modulus q	All (**S**) *FSP* (**D**)	Prime (**S**) *FSP* (**D**)	All (**S**) *IP* (**D**)	Prime	Prime, number-theoretic cond.	
Secret $\mathbf{s}$	Mod q Gaussian[2]	Mod q	Entropic	η-bounded	η-bounded	Mod q η-bounded[2]
Error $\mathbf{e}$	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	η-bounded
Variant	**S**/**D**	**S**/**D**	**S**(/**D**[3])	**S**	**S**[4]/**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. 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'.

[1] Low ranks can be achieve at the expense of a larger η. However, one can still reach $\log_2 n$ ranks (instead of $\Omega(\log_2 q) + \omega(\log_2 n)$) for constant values of η.
[2] Obtained by the Hermite Normal Form transformation reduction from [24].
[3] 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 exist 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.
[4] The hardness of the search version is obtained from the decision version through a trivial reduction.

Acknowledgments

This work was supported by the European Union PROMETHEUS project (Horizon 2020 Research and Innovation Program, grant 780701), by the PEPR quantique France 2030 programme (ANR-22-PETQ-0008), and by the Danish Independent Research Council under project number 0165-00107B (C3PO). We thank our anonymous referees of Asiacrypt 2020, Indocrypt 2020, CT-RSA 2021, and the IACR Journal of Cryptology for their thorough proof reading and constructive feedback on the original papers and the present one. We thank Thomas Prest for making us aware of improved rank conditions when using of the leftover hash lemma in the Rényi divergence. We also thank Olivier Sanders for helpful discussions on linear independence probabilities.

References

[1] 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.

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

[3] Martin R. Albrecht, Carlos Cid, Jean-Charles Faugère, Robert Fitzpatrick, and Ludovic Perret. Algebraic algorithms for LWE problems. *ACM Commun. Comput. Algebra*, 49(2):62, 2015.

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

[5] 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.

[6] 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.

[7] 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.

[8] Iván Blanco-Chacón. On the RLWE/PLWE equivalence for cyclotomic number fields. *Appl. Algebra Eng. Commun. Comput.*, 33(1):53–71, 2022.

[9] 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.

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

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

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

[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.

[19] 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.

[20] C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206. ACM, 2008.

[21] 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.

[22] R. Impagliazzo and D. Zuckerman. How to recycle random bits. In *FOCS*, pages 248–253. IEEE Computer Society, 1989.

[23] 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.

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

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] V. Lyubashevsky. Lattice signatures without trapdoors. In *EUROCRYPT*, volume 7237 of *Lecture Notes in Computer Science*, pages 738–755. Springer, 2012.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

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

[35] D. Micciancio. On the hardness of learning with errors with binary secrets. *Theory Comput.*, 14(1):1–17, 2018.

[36] 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.

[37] 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.

[38] 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.

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

[40] NIST. Post-quantum cryptography standardization. https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Post-Quantum-Cryptography-Standardization.

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

[42] C. Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In *STOC*, pages 333–342. ACM, 2009.

[43] 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.

[44] 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.

[45] C. Peikert, O. Regev, and N. Stephens-Davidowitz. Pseudorandomness of ring-lwe for any ring and modulus. In *STOC*, pages 461–473. ACM, 2017.

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

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

[48] 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.

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

[50] 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.

[51] 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.

[52] 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.

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

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

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

[56] Y. Wang and M. Wang. Module-lwe versus ring-lwe, revisited. *IACR Cryptol. ePrint Arch.*, page 930, 2019. Version dated from Aug. 18th 2019.

Appendix A Singularity of Uniform Matrices

In this section, K denotes an arbitrary number field and R its ring of integers. Throughout the entire section, q is an prime integer that does not ramify in R and that splits as $qR = \prod_{i\in[\kappa]} \mathfrak{p}_i$, where $\kappa \leq n = [K : \mathbb{Q}]$. We still use R_q to define R/qR and we also define $\mathbb{F}_i = R/\mathfrak{p}_i$ for each $i \in [\kappa]$. We recall that for each $i \in [\kappa]$, $\mathbb{F}_i$ is a finite field of size $N(\mathfrak{p}_i)$, see e.g. [31, Sec. 2.5.3].

We prove several results on the probability that a uniformly random matrix $\mathbf{A} \in R_q^{d\times d}$ is invertible in R_q. For a ring A and an integer d, we denote by $GL_d(A)$ the set of matrices of $A^{d\times d}$ that are invertible in A. We note that such results were provided in [56]. However, the proofs were based on a flawed argument which was that a vector of R_q^d which is linearly independent (with itself) must contain a coefficient in $R_q^\times$. This is not the case as a vector of R_q^d consisting only of zero divisors can still be linearly independent. We give more details in Section A.2. We also provide new corrected proofs for their results which essentially rely on analyzing each residue modulo $\mathfrak{p}_i$.

A.1 Preliminaries

The ring R_q is a finite commutative ring. As such, an element of R_q is either a unit or a zero divisor. We denote the set of units by $R_q^\times$, and the set of zero divisors by $Z(R_q) = \{r \in R_q : \exists s \in R_q \setminus \{0\}, rs = 0 \bmod qR\}$. By the Chinese Remainder Theorem [30, Lem. 2.12], there exists an isomorphism θ between R_q and $\oplus_{i\in[\kappa]}\mathbb{F}_i$ such that for all $r \in R$, $\theta(r \bmod qR) = (r \bmod \mathfrak{p}_1, \ldots, r \bmod \mathfrak{p}_\kappa)$. Note that the direct sum $\oplus_{i\in[\kappa]}\mathbb{F}_i$ is here canonically isomorphic to the direct product, which also corresponds to the Cartesian product $\times_{i\in[\kappa]}\mathbb{F}_i$ with coordinate-wise operations. We thus identify the elements in the range of θ as vectors. Also, $R_q^\times$ is isomorphic to $\oplus_{i\in[\kappa]}\mathbb{F}_i^\times = \oplus_{i\in[\kappa]}\mathbb{F}_i \setminus \{0\}$.

In what follows, we consider vectors and matrices over R_q. Since R_q is not a field, we cannot use regular linear algebra results. Instead we use results from module theory to obtain similar results over R_q. Although many of the following may be folklore for the reader who is familiar with module theory, we provide proofs for completeness. First, we give a characterization of invertible matrices in R_q using the determinant.

Lemma A.1. *Let* $\mathbf{A} \in R_q^{d\times d}$. *Then:* $\mathbf{A} \in GL_d(R_q) \Leftrightarrow \det \mathbf{A} \in R_q^\times$.

Proof. Assume $\mathbf{A} \in GL_d(R_q)$. Then, there exists $\mathbf{B} \in R_q^{d\times d}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_d$. Then $(\det \mathbf{A})(\det \mathbf{B}) = \det \mathbf{AB} = \det \mathbf{I}_d = 1$. Hence, $\det \mathbf{A} \in R_q^\times$. Reciprocally, assume that $\det \mathbf{A} \in R_q^\times$. Since R_q is a commutative ring, we have

$$Com(\mathbf{A})^T \cdot \mathbf{A} = \mathbf{A} \cdot Com(\mathbf{A})^T = (\det \mathbf{A})\mathbf{I}_d,$$

where $Com(\mathbf{A})$ is the comatrix of $\mathbf{A}$. Since $\det \mathbf{A} \in R_q^\times$, it holds that

$$(\det \mathbf{A})^{-1}Com(\mathbf{A})^T \cdot \mathbf{A} = \mathbf{A} \cdot (\det \mathbf{A})^{-1}Com(\mathbf{A})^T = \mathbf{I}_d,$$

thus proving that $\mathbf{A} \in GL_d(R_q)$. □

We recall that k vectors $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ of R_q^d are R_q-linearly independent if and only if for all $(\lambda_1, \ldots, \lambda_k) \in R_q^k$, if $\sum_{j\in[k]} \lambda_j \mathbf{a}_j = \mathbf{0} \bmod qR$, then $\lambda_j = 0 \bmod qR$ for all $j \in [k]$. We first show that R_q-linear independence can be analyzed from the $\mathbb{F}_i$-linear independence of the residues.

Lemma A.2. *Let $\mathbf{a}_1, \ldots, \mathbf{a}_\ell$ be vectors of R_q^d. Then $(\mathbf{a}_1, \ldots, \mathbf{a}_\ell)$ are R_q-linearly independent if and only if for all $i \in [\kappa]$, $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i)$ are $\mathbb{F}_i$-linearly independent.*

Proof. We extend the CRT isomorphism θ to vectors coefficient-wise. For $j \in [\ell]$ and $i \in [\kappa]$, we denote $\mathbf{a}_j \bmod \mathfrak{p}_i$ by $\mathbf{a}_j^{(i)}$ for clarity.

First, assume that for all $i \in [\kappa]$, $(\mathbf{a}_1^{(i)}, \ldots, \mathbf{a}_\ell^{(i)})$ are $\mathbb{F}_i$-linearly independent. Let $(\lambda_1, \ldots, \lambda_\ell)$ be in R_q^ℓ such that $\sum_{j\in[\ell]} \lambda_j \mathbf{a}_j = \mathbf{0} \bmod qR$. By applying θ, we have that for all $i \in [\kappa]$

$$\sum_{j\in[\ell]} (\lambda_j \bmod \mathfrak{p}_i)\mathbf{a}_j^{(i)} = \mathbf{0} \bmod \mathfrak{p}_i.$$

By assumption, it gives that for all $j \in [\ell]$ and all $i \in [\kappa]$, $\lambda_j \bmod \mathfrak{p}_i = 0$. Then, for all $j \in [\ell]$, it holds

$$\lambda_j = \theta^{-1}(\lambda_j \bmod \mathfrak{p}_1, \ldots, \lambda_j \bmod \mathfrak{p}_\kappa) = \theta^{-1}(0, \ldots, 0) = 0,$$

where the equalities are in R_q. Hence, $(\mathbf{a}_1, \ldots, \mathbf{a}_\ell)$ are R_q-linearly independent.

Reciprocally, assume that $(\mathbf{a}_1, \ldots, \mathbf{a}_\ell)$ are R_q-linearly independent. Let $i \in [\kappa]$ and let $(\mu_1, \ldots, \mu_\ell)$ be in $\mathbb{F}_i^\ell$ such that $\sum_{j\in[\ell]} \mu_j \mathbf{a}_j^{(i)} = \mathbf{0} \bmod \mathfrak{p}_i$. For each $j \in [\ell]$, define $\lambda_j = \theta^{-1}(0, \ldots, 0, \mu_j, 0, \ldots, 0)$, where μ_j is at position i. Then,

$$\theta\left(\sum_{j\in[\ell]} \lambda_j \mathbf{a}_j\right) = \left(\mathbf{0}, \ldots, \mathbf{0}, \sum_{j\in[\ell]} \mu_j \mathbf{a}_j^{(i)}, \mathbf{0}, \ldots, \mathbf{0}\right) = (\mathbf{0}, \ldots, \mathbf{0}).$$

Hence, $\sum_{j\in[\ell]} \lambda_j \mathbf{a}_j = \mathbf{0} \bmod qR$ and by assumption, it holds that $\lambda_j = 0 \bmod qR$ for all $j \in [\ell]$. As a result, it gives that $\mu_j = 0 \bmod \mathfrak{p}_i$ for all $j \in [\ell]$, thus proving that $(\mathbf{a}_1^{(i)}, \ldots, \mathbf{a}_\ell^{(i)})$ are $\mathbb{F}_i$-linearly independent. Being valid for all $i \in [\kappa]$, it yields the claim. □

We can now prove that a square matrix is invertible if and only if its columns are R_q-linearly independent, as claimed in [56, Lem. 18]. The following can be used to show that if $\ell > d$, then ℓ vectors of R_q^d cannot be R_q-linearly independent.

Lemma A.3. *Let $\mathbf{A} = [\mathbf{a}_1 | \ldots | \mathbf{a}_d]$ be in $R_q^{d\times d}$. It then holds that $\mathbf{A} \in GL_d(R_q)$ if and only if $(\mathbf{a}_1, \ldots, \mathbf{a}_d)$ are R_q-linearly independent.*

Proof. First, by contraposition, assume that $(\mathbf{a}_1, \ldots, \mathbf{a}_d)$ are not R_q-linearly independent. Hence, there exists $(\lambda_1, \ldots, \lambda_d) \in R_q^d \setminus \{\mathbf{0}\}$ such that $\sum_{j\in[d]} \lambda_j \mathbf{a}_j =$

$\mathbf{0} \bmod qR$. There exists some $j_0 \in [d]$ such that $\lambda_{j_0} \neq 0 \bmod qR$. We then have

$$\begin{aligned}
\lambda_{j_0} \cdot \det \mathbf{A} &= \det\left([\mathbf{a}_1|\dots|\mathbf{a}_{j_0-1}|\lambda_{j_0}\mathbf{a}_{j_0}|\mathbf{a}_{j_0+1}|\dots|\mathbf{a}_d]\right) \\
&= \det\left(\left[\mathbf{a}_1|\dots|\mathbf{a}_{j_0-1}|\lambda_{j_0}\mathbf{a}_{j_0} + \sum_{j\in[d]\setminus\{j_0\}} \lambda_j\mathbf{a}_j|\mathbf{a}_{j_0+1}|\dots|\mathbf{a}_d\right]\right) \\
&= \det\left([\mathbf{a}_1|\dots|\mathbf{a}_{j_0-1}|\mathbf{0}|\mathbf{a}_{j_0+1}|\dots|\mathbf{a}_d]\right) \\
&= 0 \bmod qR.
\end{aligned}$$

This proves that $\det \mathbf{A} \in Z(R_q)$ and thus $\det \mathbf{A} \notin R_q^\times$. By Lemma A.1, it holds that $\mathbf{A} \notin GL_d(R_q)$.

Now assume that $(\mathbf{a}_1, \dots, \mathbf{a}_d)$ are R_q-linearly independent. Then, Lemma A.2 yields that for all $i \in [\kappa]$, $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \dots, \mathbf{a}_d \bmod \mathfrak{p}_i)$ are $\mathbb{F}_i$-linearly independent. Let i be in $[\kappa]$. Since $\mathbb{F}_i$ is a field and that $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \dots, \mathbf{a}_d \bmod \mathfrak{p}_i)$ are $\mathbb{F}_i$-linearly independent, then the matrix $\mathbf{A} \bmod \mathfrak{p}_i$ is in $GL_d(\mathbb{F}_i)$. Hence $\det \mathbf{A} \neq 0 \bmod \mathfrak{p}_i$, which proves that $(\det \mathbf{A}) \bmod \mathfrak{p}_i \in \mathbb{F}_i^\times$. Being valid for all $i \in [\kappa]$, the Chinese Remainder Theorem yields $\det \mathbf{A} \in R_q^\times$. Hence, by Lemma A.1, $\mathbf{A} \in GL_d(R_q)$. □

A.2 Linear Independence in Uniform Matrices — Lemma 2.5

As a warm-up for proving Lemma 2.5, we first analyze the probability that a random vector of R_q^d is R_q-linearly independent. The formula given in footnote 9 of [56] provides only a lower bound. This formula relies on the flawed observation that a vector is R_q-linearly independent if and only if it contains a coefficient in $R_q^\times$. The argument provided for this claim is as follows. If $\mathbf{a} = [a_1|\dots|a_d]^T \in Z(R_q)^d$, then there exist $y_1, \dots, y_d \in (R_q \setminus \{0\})^d$ such that $a_i \cdot y_i = 0 \bmod qR$ for all $i \in [d]$. Then, by defining $\lambda = \prod_{i\in[d]} y_i$, we get $\lambda\mathbf{a} = \mathbf{0} \bmod qR$. However, the authors claim at this point that $\lambda \neq 0 \bmod qR$, which has no reason to be the case. We provide the following lemma to show the contrary, as well as a concrete counterexample that satisfies the conditions of the lemma.

Lemma A.4. *Let K be a number field, and R its ring of integers. Let q be a prime integer that is not inert*[5] *in R. Let r, s be elements of R such that $r \bmod qR \in Z(R_q)$, $s \bmod qR \in Z(R_q)$ and $\langle r\rangle + \langle s\rangle = R$. Then, the vector $[r \bmod qR, s \bmod qR]^T \in Z(R_q)^2$ is R_q-linearly independent.*

Proof. By assumption, r and s are coprime and therefore there exists u, v in R such that $u\cdot r + v\cdot s = 1$. Hence $(u \bmod qR)(r \bmod qR) + (v \bmod qR)(s \bmod qR) = 1 \bmod qR$. Define $\mathbf{x} = [r \bmod qR, s \bmod qR]^T$. By assumption $\mathbf{x} \in Z(R_q)^2$. Let $\lambda \in R_q$ be such that $\lambda\mathbf{x} = 0 \bmod qR$. It implies that $\lambda(r \bmod qR) =$

[5] If q is inert, then R_q is a field, and therefore $Z(R_q) = \{0\}$, meaning that there does not exist r, s satisfying the conditions.

$0 \bmod qR$ and $\lambda(s \bmod qR) = 0 \bmod qR$. As a result, it holds

$$\begin{aligned}\lambda &= \lambda \cdot ((u \bmod qR)(r \bmod qR) + (v \bmod qR)(s \bmod qR)) \bmod qR \\ &= (u \bmod qR) \cdot \lambda(r \bmod qR) + (v \bmod qR) \cdot \lambda(s \bmod qR) \\ &= 0 \bmod qR\end{aligned}$$

It thus proves that the only common annihilator of $r \bmod qR$ and $s \bmod qR$ is 0, which in other terms means $\mathbf{x}$ is R_q-linearly independent. □

The conditions of the previous lemma can easily be met. As a concrete example, in the cyclotomic field of conductor $\nu = 256$, for $q = 257$, one can check that the elements $r = \zeta + 3$ and $s = \zeta + 6$ verify the conditions of the above lemma. Such counterexamples can easily be found by enumerating the first few zero divisors, or by sampling a few zero divisors at random.

We then recall the following lemma which we need to prove Lemma 2.5. We provide a proof for completeness.

Lemma A.5. *Let r be a random variable that is uniformly distributed over R_q. Then, the random variables $(r \bmod \mathfrak{p}_i)_{i\in[\kappa]}$ are independent and uniformly distributed over their respective support $\mathbb{F}_i$.*

Proof. Denote by $\mathbf{r} = \theta(r)$, which is the random vector composed of the $r \bmod \mathfrak{p}_i$. Since θ is an isomorphism, it holds that $\mathbf{r} \sim U(\oplus_{i\in[\kappa]}\mathbb{F}_i)$. Let $\mathbf{s}$ be the random vector over $\oplus_{i\in[\kappa]}\mathbb{F}_i$ such that the coordinates are independent and uniform over each $\mathbb{F}_i$ respectively. Let $\mathbf{r}' \in \oplus_{i\in[\kappa]}\mathbb{F}_i$. It holds

$$\mathbb{P}_{\mathbf{r}}[\mathbf{r} = \mathbf{r}'] = q^{-n} = \prod_{i\in[\kappa]} N(\mathfrak{p}_i)^{-1} = \prod_{i\in[\kappa]} \mathbb{P}_{s_i}[s_i = r'_i] = \mathbb{P}_{\mathbf{s}}[\mathbf{s} = \mathbf{r}'].$$

This proves that $\mathbf{r}$ and $\mathbf{s}$ are identical random vectors, which yields that the coordinates of $\mathbf{r}$ are independent and uniform over each $\mathbb{F}_i$. □

We now focus on proving Lemma 2.5, where the first part was claimed in [56, Lem. 9]. First, we note that we cannot naively use the analysis over residues to obtain a proof of [56, Lem. 19]. The latter argues that if $(\mathbf{a}_1, \ldots, \mathbf{a}_\ell) \in (R_q^d)^\ell$ are R_q-l. i. with $\ell \le d$, then one can always extract a submatrix in $GL_\ell(R_q)$ by selecting some subset of the rows. However, this is not generally true for such matrices over R_q. This is due to the fact that a minimal spanning set of an R_q-submodule of R_q^d is not necessarily a basis of said submodule. When analyzing the problem in the residues, we will obtain submatrices in $GL_\ell(\mathbb{F}_i)$. But it is not guaranteed that these submatrices correspond to the same subsets of rows, in which case we cannot re-combine the residues. We show that we do not need this fact to prove Lemma 2.5. For convenience, we now abbreviate "R_q-linearly independent" by "R_q-l. i.".

Proof (of Lemma 2.5). By Lemma A.2 and A.5, we can analyze the residues and first determine $\mathbb{P}_{\mathbf{b}^{(i)}\sim U(\mathbb{F}_i^d)}[(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i, \mathbf{b}^{(i)})$ are $\mathbb{F}_i$-l. i.$]$ for each

$i \in [\kappa]$ individually. Let i be in $[\kappa]$. Since $(\mathbf{a}_1, \ldots, \mathbf{a}_\ell)$ are R_q-linearly independent, then again the residues $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i)$ are $\mathbb{F}_i$-linearly independent by Lemma A.2. It yields that two linear combinations of $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i)$ that are equal must have equal coefficients. Hence, there are $|\mathbb{F}_i|^\ell$ distinct linear combinations of the $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i)$.

As a result, if we denote by S_i the set of $\mathbf{b}^{(i)} \in \mathbb{F}_i^d$ that are not in the $\mathbb{F}_i$-span of $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i)$, we have $|S_i| = |\mathbb{F}_i|^d - |\mathbb{F}_i|^\ell$. Note that when $\ell = 0$, there are $|\mathbb{F}_i|^0 = 1$ vectors in the span of $\mathbf{0}$. In this case, we obtain $|S_i| = |\mathbb{F}_i|^d - 1$ which is coherent with the previous formula. Since we work over a field $\mathbb{F}_i$, and thus a vector space, if $\mathbf{b}^{(i)}$ is in S_i, then $(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i, \mathbf{b}^{(i)})$ are $\mathbb{F}_i$-linearly independent (which is not necessarily true in a module). It then yields

$$\begin{aligned}\mathbb{P}_{\mathbf{b}^{(i)} \sim U(\mathbb{F}_i^d)}[(\mathbf{a}_1 \bmod \mathfrak{p}_i, \ldots, \mathbf{a}_\ell \bmod \mathfrak{p}_i, \mathbf{b}^{(i)}) \text{ are } \mathbb{F}_i\text{-l. i.}] &= \mathbb{P}_{\mathbf{b}^{(i)} \sim U(\mathbb{F}_i^d)}[\mathbf{b}^{(i)} \in S_i] \\ &= \frac{|S_i|}{|\mathbb{F}_i|^d} \\ &= 1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}.\end{aligned}$$

By combining the residues, we obtain

$$\mathbb{P}_{\mathbf{b} \sim U(R_q^d)}[(\mathbf{a}_1, \ldots, \mathbf{a}_\ell, \mathbf{b}) \text{ are } R_q\text{-l. i.}] = \prod_{i \in [\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right). \tag{8}$$

We note that we can still prove the relaxed lower bound claimed in [56]. Since all the prime ideal factors $\mathfrak{p}_i$ are above q, we have $N(\mathfrak{p}_i) = q^{f_i} \geq q$ for some $f_i \geq 1$. It yields

$$\prod_{i \in [\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right) \geq \left(1 - \frac{1}{q^{d-\ell}}\right)^\kappa \geq 1 - \frac{\kappa}{q^{d-\ell}} \geq 1 - \frac{\kappa}{q},$$

where the second inequality holds by the Bernoulli inequality, and the last follows from the fact that $d - \ell \geq 1$. Since κ is at most the degree n of the number field K, we can lower bound the probability by $1 - n/q$.

Finally, let us prove the last formula. Fix $\ell \in \{1, \ldots, k-1\}$. We study the following probability

$$\mathbb{P}_{(\mathbf{a}_j)_{j \in [\ell+1]} \sim U(R_q^d)^{\ell+1}}[(\mathbf{a}_j)_{j \in [\ell+1]} \text{ are } R_q\text{-l. i.}].$$

It can be decomposed by the total probability formula as

$$\sum_{(\bar{\mathbf{a}}_j)_{j \in [\ell]}} \mathbb{P}_{(\mathbf{a}_j)_{j \in [\ell]}}[(\mathbf{a}_j)_{j \in [\ell]} = (\bar{\mathbf{a}}_j)_{j \in [\ell]}] \cdot \mathbb{P}_{\mathbf{a}_{\ell+1}}[((\bar{\mathbf{a}}_j)_{j \in [\ell]}, \mathbf{a}_{\ell+1}) \text{ are } R_q\text{-l. i.}].$$

Note that if the summands $(\bar{\mathbf{a}}_j)_j$ are not R_q-linearly independent, it directly gives that $((\bar{\mathbf{a}}_j)_{j \in [\ell]}, \mathbf{a}_{\ell+1})$ cannot be R_q-linearly independent and that the second

probability is therefore 0. Hence, we consider the sum over the $(\bar{\mathbf{a}}_j)_j$ that are R_q-linearly independent. By the first formula, the second probability over $\mathbf{a}_{\ell+1}$ is given by $\prod_{i\in[\kappa]}(1-1/N(\mathfrak{p}_i)^{d-\ell})$. As it does not depend on the summand, we have that the probability is exactly

$$\mathbb{P}_{(\mathbf{a}_j)_{j\in[\ell]}\sim U(R_q^d)^\ell}[(\mathbf{a}_j)_{j\in[\ell]} \text{ are } R_q\text{-l. i.}] \cdot \prod_{i\in[\kappa]}\left(1-\frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right).$$

Using Equation (8) with $\ell = 0$ for initialization, induction on ℓ then yields

$$\mathbb{P}_{(\mathbf{a}_i)_{i\in[k]}\sim U(R_q^d)^k}\left[(\mathbf{a}_i)_{i\in[k]} \text{ are } R_q\text{-l. i.}\right] = \prod_{\ell=0}^{k-1}\prod_{i\in[\kappa]}\left(1-\frac{1}{N(\mathfrak{p}_i)^{d-\ell}}\right),$$

as desired. □

We note that the number of columns is $k \leq d$. Additionally, in R_q we still have the fact that a linearly independent family of vectors of R_q^d cannot contain more that d vectors. As a result, when $k > d$, we have to analyze the following probability

$$\mathbb{P}_{(\mathbf{a}_i)_{i\in[k]}\sim U(R_q^d)^k}[\exists S \subseteq [k], |S| = d \wedge (\mathbf{a}_i)_{i\in S} \text{ are } R_q\text{-l. i.}]$$

As explained, even if there exists subsets $S_i \subseteq [k]$ with $|S_i| = d$ and $(\mathbf{a}_j \bmod \mathfrak{p}_i)_{j\in S_i}$ are $\mathbb{F}_i$-l. i., there is no guarantee that all the S_i are equal.

A.3 Singularity of Uniform Matrices — Lemma 2.6

In this section, we can show that the equality of the S_i is not necessary to guarantee that the columns form a spanning set of R_q^d.

Proof (of Lemma 2.6). We prove the lower bound by inclusion of events. Let $\mathbf{A} = [\mathbf{a}_1|\dots|\mathbf{a}_m] \in R_q^{d\times m}$ be such that

$$\forall i \in [\kappa], \exists S_i \subseteq [m], |S_i| = d \wedge (\mathbf{a}_j \bmod \mathfrak{p}_i)_{j\in S_i} \text{ are } \mathbf{F}_i\text{-l. i.}. \tag{9}$$

We show that this guarantees that $\mathbf{A}\cdot R_q^m = R_q^d$. Consider the CRT basis $\lambda_1,\dots,\lambda_\kappa$ defined by $\lambda_i = \theta^{-1}(\mathbf{e}_i)$ where $(\mathbf{e}_1,\dots,\mathbf{e}_\kappa)$ is the canonical basis of $\oplus_{i\in[\kappa]}\mathbb{F}_i$. We index each set S_i as $S_i = \{j_1^{(i)},\dots,j_d^{(i)}\}$. We then construct the vectors $\mathbf{b}_\ell = \sum_{i\in[\kappa]}\lambda_i \mathbf{a}_{j_\ell^{(i)}}$ for all $\ell \in [d]$. We now prove that $(\mathbf{b}_\ell)_{\ell\in[d]}$ are R_q-linearly independent. Let $(\mu_\ell)_{\ell\in[d]} \in R_q^d$ be such that $\sum_{\ell\in[d]}\mu_\ell\mathbf{b}_\ell = \mathbf{0} \bmod qR$. Let $i^* \in [\kappa]$. It holds that

$$\mathbf{0} = \sum_{\ell\in[d]}(\mu_\ell \bmod \mathfrak{p}_{i^*})(\mathbf{b}_\ell \bmod \mathfrak{p}_{i^*}) = \sum_{\ell\in[d]}(\mu_\ell \bmod \mathfrak{p}_{i^*})(\mathbf{a}_{j_\ell^{(i^*)}} \bmod \mathfrak{p}_{i^*}),$$

where the equality is over $\mathbf{F}_{i^*}$. The last equality follows from the construction of the $\mathbf{b}_\ell$ as $\mathbf{b}_\ell \bmod \mathfrak{p}_{i^*} = \sum_{i\in[\kappa]} \delta_{i,i^*}(\mathbf{a}_{j_\ell^{(i)}} \bmod \mathfrak{p}_{i^*}) = \mathbf{a}_{j_\ell^{(i^*)}} \bmod \mathfrak{p}_{i^*}$. By definition of S_{i^*}, the $(\mathbf{a}_{j_\ell^{(i^*)}} \bmod \mathfrak{p}_{i^*})_{\ell\in[d]}$ are $\mathbf{F}_{i^*}$-linearly independent and therefore $\mu_\ell \bmod \mathfrak{p}_{i^*} = 0$. Being true for all i^*, it proves by Lemma A.2 that $(\mathbf{b}_\ell)_{\ell\in[d]}$ are R_q-linearly independent.

We can then write $\mathbf{B} = [\mathbf{b}_1|\dots|\mathbf{b}_d] = \mathbf{AC} \in R_q^{d\times d}$ where the matrix $\mathbf{C} \in R_q^{m\times d}$ is composed of the λ_i at the correct positions. More precisely, the ℓ-th column of $\mathbf{C}$ is $\sum_{i\in[\kappa]} \lambda_i \mathbf{e}'_{j_\ell^{(i)}} \in R_q^m$, where the $\mathbf{e}'_1, \dots, \mathbf{e}'_m$ represent the canonical basis of R_q^m. By Lemma A.3, since the columns of $\mathbf{B}$ are R_q-linearly independent, then $\mathbf{B} \in GL_d(R_q)$. Hence, for all $\mathbf{y} \in R_q^d$, there exists $\mathbf{x}' \in R_q^d$ such that $\mathbf{Bx}' = \mathbf{y} \bmod qR$. By defining $\mathbf{x} = \mathbf{Cx}' \in R_q^m$, we have $\mathbf{Ax} = \mathbf{y} \bmod qR$. This proves that $\mathbf{A} \cdot R_q^m = R_q^d$.

As a result, for any matrix $\mathbf{A}$ verifying (9), it holds that $\mathbf{A} \cdot R_q^m = R_q^d$. By inclusion of events, we have

$$\begin{aligned}&\mathbb{P}_{\mathbf{A}\sim U(R_q^{d\times m})}[\mathbf{A} \cdot R_q^m = R_q^d] \\ &\quad \geq \mathbb{P}_{\mathbf{A}\sim U(R_q^{d\times m})}[\forall i \in [\kappa], \exists S_i \subseteq [m], |S_i| = d \wedge (\mathbf{a}_j \bmod \mathfrak{p}_i)_{j\in S_i} \text{ are } \mathbf{F}_i\text{-l. i.}].\end{aligned}$$

We now evaluate the right-hand side. By Lemma A.5, the latter probability equals

$$\prod_{i\in\kappa} \mathbb{P}_{\mathbf{A}^{(i)}\sim U(\mathbf{F}_i^{d\times m})}[\exists S_i \subseteq [m], |S_i| = d \wedge (\mathbf{a}_j^{(i)} \bmod \mathfrak{p}_i)_{j\in S_i} \text{ are } \mathbf{F}_i\text{-l. i.}]$$

Let $i \in [\kappa]$. Since $\mathbf{F}_i$ is a field, we have the following

$$\begin{aligned}&\mathbb{P}_{\mathbf{A}^{(i)}\sim U(\mathbf{F}_i^{d\times m})}[\exists S_i \subseteq [m], |S_i| = d \wedge (\mathbf{a}_j^{(i)} \bmod \mathfrak{p}_i)_{j\in S_i} \text{ are } \mathbf{F}_i\text{-l. i.}] \\ &\qquad = \mathbb{P}_{\mathbf{A}^{(i)}\sim U(\mathbf{F}_i^{d\times m})}[\text{Column-Rank}(\mathbf{A}^{(i)}) = d] \\ &\qquad = \mathbb{P}_{\mathbf{A}^{(i)}\sim U(\mathbf{F}_i^{d\times m})}[\text{Row-Rank}(\mathbf{A}^{(i)}) = d] \\ &\qquad = \mathbb{P}_{(\mathbf{a}'_i)_{i\in[d]}\sim U(\mathbf{F}_i^m)^d}[(\mathbf{a}'_i)_{i\in[d]} \text{ are } \mathbf{F}_i\text{-l. i.}]. \\ &\qquad = \prod_{\ell=0}^{d-1} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{m-\ell}}\right),\end{aligned}$$

where the last inequality is the special case of Lemma 2.5 over the residue field $\mathbf{F}_i$. Combining it all yields

$$\mathbb{P}_{\mathbf{A}\sim U(R_q^{d\times m})}[\mathbf{A} \cdot R_q^m = R_q^d] \geq \prod_{\ell=0}^{d-1} \prod_{i\in[\kappa]} \left(1 - \frac{1}{N(\mathfrak{p}_i)^{m-\ell}}\right).$$

□

Appendix B Missing Proofs

B.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 S_η, and $i \in [n]$. Then, it holds that

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

Taking the maximum over all $i \in [n]$ and $x \in S_\eta$ yields $B_\eta \leq n\eta\|\mathbf{V}\|_{\max}$. 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 $B_\eta \leq n\eta$ 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. [49] 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 2.8. We begin with stating some lemmas that we need for the proof. The first two bound the Rényi divergence and statistical distance, respectively, if the second distribution is the uniform distribution over the support of the first.

Lemma B.1. *Let P be a probability distribution and Q be the uniform distribution over its support* $\mathrm{Supp}(P)$*. It holds*

$$\mathrm{RD}_2(P\|Q) = |\mathrm{Supp}(P)| \cdot \mathbb{P}[P = P'],$$

where $P \sim P'$ are independent and identically distributed.

Proof. By the definition of the Rényi divergence, it yields

$$\begin{aligned}\mathrm{RD}_2(P\|Q) = \sum_{x\in\mathrm{Supp}(P)} \frac{P(x)^2}{Q(x)} &= |\mathrm{Supp}(P)| \cdot \sum_{x\in\mathrm{Supp}(P)} P(x)^2 \\ &= |\mathrm{Supp}(P)| \cdot \mathbb{P}[P = P'].\end{aligned}$$

□

The following result has been attributed to Rackoff by Impagliazzo and Zuckerman [22].

Lemma B.2 ([22, Claim 2][34, Lem. 4.3]). *Let P be a probability distribution and Q be the uniform distribution over its support* $\mathrm{Supp}(P)$*. It holds*

$$\Delta(P, Q) \le \frac{1}{2}\sqrt{|\mathrm{Supp}(P)| \cdot \mathbb{P}[P = P'] - 1},$$

where $P \sim P'$ are independent and identically distributed.

We also adapt [34, Lem. 4.4] from vectors to matrices over a finite ring.

Lemma B.3. *Let A be a finite ring and k, d be positive integers. Further, take an arbitrary vector $\mathbf{z} = (z_j)_{j\in[d]} \in A^d$. If $\mathbf{C} \sim U(A^{k\times d})$, then $\mathbf{Cz}$ is uniformly distributed over the module $\langle z_1, \ldots, z_d\rangle^k$. In particular, the probability that $\mathbf{Cz} = \mathbf{0}$ is exactly $\frac{1}{|\langle z_1,\ldots,z_d\rangle|^k}$.*

Proof. Let $\mathbf{z} \in A^d$. For $\mathbf{b} \in A^k$ we define $T_{\mathbf{b}} = \{\mathbf{C} \in A^{k\times d} \colon \mathbf{Cz} = \mathbf{b}\}$. Notice that the probability that $\mathbf{Cz} = \mathbf{b}$ over the uniform random choice of $\mathbf{C}$ is exactly $\frac{|T_{\mathbf{b}}|}{|A|^{k\cdot d}}$. If $\mathbf{b} \notin \langle z_1, \ldots, z_d\rangle^k$, then $T_{\mathbf{b}} = \emptyset$ and hence $\mathbb{P}_{\mathbf{C}\sim U(A^{k\times d})}[\mathbf{Cz} = \mathbf{b}] = 0$. We now show that all $\mathbf{b} \in \langle z_1, \ldots, z_d\rangle^k$ have the same probability. Let $\mathbf{b}$ be an arbitrary element of $\langle z_1, \ldots, z_d\rangle^k$, i.e., it can be represented as $\mathbf{Cz} = \mathbf{b}$ for some fixed $\mathbf{C} \in A^{k\times d}$. It follows that $\mathbf{C}' \in T_{\mathbf{b}}$ if and only if $\mathbf{C}' - \mathbf{C} \in T_{\mathbf{0}}$. Further, the mapping $\mathbf{C}' \mapsto \mathbf{C}' - \mathbf{C}$ is a bijection between $T_{\mathbf{b}}$ and $T_{\mathbf{0}}$, which implies that $|T_{\mathbf{b}}| = |T_{\mathbf{0}}|$. This shows that all $\mathbf{b} \in \langle z_1, \ldots, z_d\rangle^k$ have the same probability, completing the proof. □

Proof (of Lemma 2.8). Let P be the distribution that samples $\mathbf{C} \hookleftarrow U(R_q^{k\times d})$ and $\mathbf{z} \hookleftarrow U(S_\eta^d)$ and outputs $(\mathbf{C}, \mathbf{Cz}) \in R_q^{k\times d} \times R_q^k$. Let Q be the uniform distribution over the support of P, i.e., it samples $\mathbf{C} \hookleftarrow U(R_q^{k\times d})$ and $\mathbf{s} \hookleftarrow U(R_q^k)$, and outputs $(\mathbf{C}, \mathbf{s}) \in R_q^{k\times d} \times R_q^k$. Note that $|\mathrm{Supp}(P)| = q^{nk(d+1)}$.

In the following we bound the collision probability of P and then we simply apply Lemma B.1 and B.2 (with the finite ring R_q) to conclude the proof.

For $\mathbf{C}, \mathbf{C}' \sim U(R_q^{k\times d})$ and $\mathbf{z}, \mathbf{z}' \sim U(S_\eta^d)$ it yields

$$\begin{aligned}\mathbb{P}[\mathbf{C} = \mathbf{C}' \wedge \mathbf{Cz} = \mathbf{C}'\mathbf{z}'] &= \mathbb{P}[\mathbf{C} = \mathbf{C}'] \cdot \mathbb{P}[\mathbf{Cz} = \mathbf{C}'\mathbf{z}' | \mathbf{C} = \mathbf{C}'] \\ &= \frac{1}{|R_q|^{k\cdot d}} \cdot \mathbb{P}[\mathbf{C}(\mathbf{z} - \mathbf{z}') = \mathbf{0}].\end{aligned}$$

By Lemma B.3 over the random choice of $\mathbf{C}$ and the size of the finite ring R_q, we can further transform this equation to

$$\begin{aligned}\frac{1}{q^{n\cdot k\cdot d}} \cdot \mathbb{P}[\mathbf{C}(\mathbf{z} - \mathbf{z}') = \mathbf{0}] &= \frac{1}{q^{nkd}} \cdot \sum_{I\in\mathcal{I}} \frac{\mathbb{P}\left[\langle z_1 - z_1', \ldots, z_d - z_d'\rangle^k = I^k\right]}{|I|^k} \\ &\leq \frac{1}{q^{nkd}} \cdot \sum_{I\in\mathcal{I}} \frac{\mathbb{P}\left[\langle z_1 - z_1', \ldots, z_d - z_d'\rangle^k \subseteq I^k\right]}{|I|^k} \\ &= \frac{1}{q^{(nk)\cdot(d+1)}} \cdot \sum_{I\in\mathcal{I}} \frac{q^{nk}}{|I|^k} \cdot \prod_{j\in[d]} \mathbb{P}\left[(z_j - z_j') \in I\right],\end{aligned}$$

where $\mathcal{I}$ denotes the set of all ideals in R_q and we conditioned on the ideal $\langle z_1 - z_1', \ldots, z_k - z_k'\rangle$.

We now specify $\mathcal{I}$. For $K = \mathbb{Q}(\zeta)$, let f be the minimal polynomial of ζ and let $f = \prod_{i\in[\kappa]} f_i$ be its factorization in irreducible polynomials in $\mathbb{Z}_q[x]$. As $\mathbb{Z}_q$ is a field, $\mathbb{Z}_q[x]$ is a principal ideal domain. The ideal correspondence theorem in commutative algebra states that every ideal in R_q corresponds to an ideal in $\mathbb{Z}_q[x]$ containing $\langle f\rangle$. As each ideal in $\mathbb{Z}_q[x]$ itself is principal, thus of the form $\langle g\rangle$ for a polynomial $g \in \mathbb{Z}[x]$, this is equivalent to g dividing f. Hence, we know that the ideals of R_q are given by $\mathcal{I} = \{\langle f_G\rangle \colon G \subseteq \{1, \ldots, \kappa\}\}$, where we define $f_G = \prod_{i\in G} f_i$. By convention, we say that the empty set $\emptyset$ defines the constant polynomial $f_\emptyset = 1$. For any f_G, it holds that

$$\begin{aligned}\mathbb{P}\left[(z_j - z_j') \in \langle f_G\rangle\right] &= \mathbb{P}\left[z_j = z_j' \bmod f_G\right] \\ &\leq \max_{\tilde{z}} \mathbb{P}\left[z_j \bmod f_G = \tilde{z}\right] \leq \frac{1}{(2\eta+1)^{\deg(f_G)}},\end{aligned}$$

where the maximum is taken over all $\tilde{z} \in R$ with $\deg(\tilde{z}) < \deg(f_G)$. As explained in [34], the last inequality follows from the fact that for any fixed value of the $n - \deg(f_G)$ highest degree coefficients of z, the map $z \mapsto z \bmod f_G$ is a bijection

between sets of size $(2\eta+1)^{\deg(f_G)}$. We then get

$$\frac{q^{nk}}{|\langle f_G\rangle|^k}\prod_{j\in[d]}\mathbb{P}\left[(z_j-z'_j)\in\langle f_G\rangle\right]\leq\frac{q^{nk}}{\left(q^{n-\deg(f_G)}\right)^k}\left(\frac{1}{(2\eta+1)^{\deg(f_G)}}\right)^d$$
$$=\left(\frac{q^k}{(2\eta+1)^d}\right)^{\deg(f_G)}.$$

Adding up over all ideals we can deduce

$$\sum_{\langle f_G\rangle\in\mathcal{I}}\frac{q^{nk}}{|\langle f_G\rangle|^k}\cdot\prod_{j\in[d]}\mathbb{P}\left[(z_j-z'_j)\in\langle f_G\rangle\right]\leq\sum_{G\subseteq\{1,\ldots,\kappa\}}\left(\frac{q^k}{(2\eta+1)^d}\right)^{\deg(f_G)}$$
$$=\prod_{i\in[\kappa]}\left(1+\left(\frac{q^k}{(2\eta+1)^d}\right)^{\deg(f_i)}\right)$$
$$\leq\prod_{i\in[\kappa]}\left(1+\frac{q^k}{(2\eta+1)^d}\right)^{\deg(f_i)}$$
$$=\left(1+\frac{q^k}{(2\eta+1)^d}\right)^n.$$

Putting everything together, it holds

$$\mathbb{P}[P=P']\leq\left(1+\frac{q^k}{(2\eta+1)^d}\right)^n q^{-nk(d+1)},$$

where $P\sim P'$ are independent and identically distributed. Using Lemma B.1 and B.2 together with $|\mathrm{Supp}(P)|=q^{nk(d+1)}$ completes the proof. □

Lemma 2.13.

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 $\mathcal{D}_{\mathcal{I},\alpha}$. Then $\sigma_H(a)$ is distributed according to $\mathcal{D}_{\Lambda,\alpha}$ where $\Lambda=\sigma_H(\mathcal{I})$. So $\|\sigma(a)\|_\infty=\|\mathbf{U}_H\sigma_H(a)\|_\infty\leq\|\sigma_H(a)\|_\infty$. 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 $\|\mathbf{a}_1\|_\infty\leq\|\mathbf{a}\|_\infty$, and for all $k\in[t_2]$, $|a_{2,k}+i\tilde{a}_{2,k}|/\sqrt{2}=\sqrt{a_{2,k}^2+\tilde{a}_{2,k}^2}/\sqrt{2}\leq\|\mathbf{a}\|_\infty$. Hence $\|\mathbf{U}_H\mathbf{a}\|_\infty\leq$

$\|\mathbf{a}\|_\infty$. By the second part of [41, Cor. 5.3] for $m = 1$, $\mathbf{z} = 1$ and $\mathbf{c} = \mathbf{0}$, it holds that for all $t \geq 0$

$$\mathbb{P}_{\mathbf{a}\sim\mathcal{D}_{\Lambda,\alpha}}[\|\mathbf{a}\|_\infty \geq \alpha t] \leq 2n \cdot e^{-\pi t^2}.$$

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

$$\mathbb{P}_{a\sim\mathcal{D}_{\mathcal{I},\alpha}}[\|\sigma(a)\|_\infty \leq \alpha t] \geq \mathbb{P}_{a\sim\mathcal{D}_{\mathcal{I},\alpha}}[\|\sigma_H(a)\|_\infty \leq \alpha t] \geq 1 - 2n \cdot e^{-\pi t^2}.$$

Now let $\mathbf{N}$ be sampled from $\mathcal{D}_{\mathcal{I},\alpha}^{m\times d}$. 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

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

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

Lemma 2.14.

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 $M_{\sigma_H}(\mathbf{U})\mathbf{S}M_{\sigma_H}(\mathbf{U})^T = \mathbf{\Sigma}$. We note that it still applies in the degenerate case. In particular, the result still holds when $\mathbf{S}, \mathbf{U}$ are not full-rank, and also when $m > d$ which automatically results in $\mathbf{\Sigma}$ being singular. □

Lemma 2.15. 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 2.15.

Lemma B.4 (Adapted from [24, Lem. 2.8] & [47, Claim 3.9]). *Let Λ be an n-dimensional lattice, $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{R}, \mathbf{S}$ two positive definite matrices of $\mathbb{R}^{n\times n}$, and $\mathbf{T} = \mathbf{R}+\mathbf{S}$. We define $\mathbf{U} = \left(\mathbf{R}^{-1}+\mathbf{S}^{-1}\right)^{-1}$, and assume that $\rho_{\sqrt{\mathbf{U}^{-1}}}(\Lambda^*) \leq 1+\varepsilon$ 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 B.4). The density function Y is given by

$$\begin{aligned}
Y(\mathbf{x}) &= \sum_{\mathbf{y}\in\Lambda+\mathbf{a}} \mathcal{D}_{\Lambda+\mathbf{a},\sqrt{\mathbf{R}}}(\mathbf{y}) D_{\sqrt{\mathbf{S}}}(\mathbf{x}-\mathbf{y}) \\
&= \frac{1}{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}(\Lambda)\sqrt{\det \mathbf{S}}} \sum_{\mathbf{y}\in\Lambda+\mathbf{a}} \rho_{\sqrt{\mathbf{R}}}(\mathbf{y})\rho_{\sqrt{\mathbf{S}}}(\mathbf{x}-\mathbf{y}) \\
&= \frac{1}{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}(\Lambda)\sqrt{\det \mathbf{S}}} \sum_{\mathbf{y}\in\Lambda+\mathbf{a}} \rho_{\sqrt{\mathbf{T}}}(\mathbf{x})\rho_{\mathbf{R}\mathbf{T}^{-1}\mathbf{x},\sqrt{\mathbf{U}}}(\mathbf{y}) \quad [43, \text{Fact } 2.1]. \\
&= \frac{\rho_{\sqrt{\mathbf{T}}}(\mathbf{x})}{\sqrt{\det \mathbf{T}}} \cdot \frac{\sqrt{\det \mathbf{T}}\rho_{\mathbf{R}\mathbf{T}^{-1}\mathbf{x},\sqrt{\mathbf{U}}}(\Lambda)}{\sqrt{\det \mathbf{S}}\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}(\Lambda)} \\
&= D_{\sqrt{\mathbf{T}}}(\mathbf{x}) \cdot \frac{(\sqrt{\det \mathbf{R}}\sqrt{\det \mathbf{S}}/\sqrt{\det \mathbf{T}})^{-1}\widehat{\rho_{\mathbf{x}',\sqrt{\mathbf{U}}}}(\Lambda^*)}{(\sqrt{\det \mathbf{R}})^{-1}\widehat{\rho_{-\mathbf{a},\sqrt{\mathbf{R}}}}(\Lambda^*)},
\end{aligned}$$

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

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

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 $\rho_{\sqrt{\mathbf{A}}}(\mathbf{w}) \ge \rho_{\sqrt{\mathbf{B}}}(\mathbf{w})$ for all $\mathbf{w}\in\mathbb{R}^n$. Since $\mathbf{U}^{-1}-\mathbf{R}^{-1} = \mathbf{S}^{-1}$ is positive semi-definite, it yields $\rho_{\sqrt{\mathbf{R}^{-1}}}(\Lambda^* \setminus \{0\}) \le \rho_{\sqrt{\mathbf{U}^{-1}}}(\Lambda^* \setminus \{0\}) \le \varepsilon$. Therefore, using the same method as above, we get

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

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 $\left|Y(\mathbf{x}) - D_{\sqrt{\mathbf{T}}}(\mathbf{x})\right| \le 4\varepsilon D_{\sqrt{\mathbf{T}}}(\mathbf{x})$. 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 2.15 in the proof of Lemma 3.5 in order to decompose a Gaussian noise into an inner product. It generalizes [47, Cor. 3.10] to the module case. A specific instance is proven in the proof of [24, Lem. 4.15], which is later mentioned (without proof) in [50, Lem. 5.5].

Lemma B.5 ([24, 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 2.15). 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 B.5, $\langle \mathbf{z}, \mathbf{h}\rangle$ 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

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

Lemma B.4 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 B.5 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 $j \in [n]$. □

B.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 $\mathbf{A} \sim U(R_q^{m\times k})$, the $\mathbf{u}_j$ are independent and identically distributed (i.i.d.) from $U((q^{-1}R/R)^m)$, the $\mathbf{e}_j$ are i.i.d. from ψ^m, and $\mathbf{b}_j = q^{-1}\mathbf{A}\mathbf{s}_j + \mathbf{e}_j \bmod R$ for $\mathbf{s}_j$ i.i.d. from $U(R_q^k)$ 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 $\mathbf{z} \hookleftarrow U(\mathcal{Z})$ 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 $i^* \hookleftarrow U([\ell])$.
3. Sample $\mathbf{s}_1,\dots,\mathbf{s}_{i^*-1} \hookleftarrow U(R_q^k)$, $\mathbf{e}_1,\dots,\mathbf{e}_{i^*-1},\mathbf{e}_{i^*+1},\dots,\mathbf{e}_\ell \hookleftarrow \psi^m$ and finally $\mathbf{u}_{i^*+1},\dots,\mathbf{u}_\ell \hookleftarrow U((q^{-1}R/R)^m)$.
4. Compute $\mathbf{b}_j = q^{-1}\mathbf{A}\mathbf{s}_j + \mathbf{e}_j \bmod R$ for all $j \in [i^*-1]$.
5. Define the hybrid matrix $\mathbf{B} = [\mathbf{b}_1,\dots,\mathbf{b}_{i^*-1},\mathbf{b},\mathbf{u}_{i^*+1},\dots,\mathbf{u}_\ell]$, and the error matrix $\mathbf{E} = [\mathbf{e}_1,\dots,\mathbf{e}_{i^*-1},\mathbf{e},\mathbf{e}_{i^*+1},\dots,\mathbf{e}_\ell]$. 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

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

□

B.3 Missing Proofs of Section 4

Lemma 4.1.

Proof. We start by describing the transformation T of [36] to move from M-LWE to M-ISIS. Given $(\mathbf{A}, \mathbf{b}) \in R_q^{m \times d} \times R^m$, 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 the quantity defined in Section 2.1, T aborts at this step with probability $\delta(m, d)$ (which can be upper bound from Lemma 2.6). 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}. Again, by definition of $\delta(\cdot, \cdot)$, we get $\Delta(\mathbf{B}', U(R_q^{m\times(m-d)})) \leq \delta(m, m-d)$. Finally, T computes $\mathbf{c} = \mathbf{B}'^T\mathbf{b} \bmod qR$, and returns $(\mathbf{B}', \mathbf{c})$.

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

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

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

$$
\begin{aligned}
&|\mathbb{P}[E] - \mathbb{P}[E']| \\
&= \mathbb{P}[\mathbf{A} \text{ non-singular}] \cdot |\mathbb{P}[E|\mathbf{A} \text{ non-singular}] - \mathbb{P}[E'|\mathbf{A} \text{ non singular}]| \\
&= (1-\delta(m,d))\big|\mathbb{P}[\mathcal{A}(\mathbf{B}', \mathbf{c} \text{ uniform}) = 1|F] - \mathbb{P}[\mathcal{A}(\mathbf{B}', (\mathbf{B}')^T\mathbf{e} \bmod qR) = 1|F]\big| \\
&> (1-\delta(m,d)) \cdot (\varepsilon' - 2\delta(m, m-d)) \\
&= \varepsilon,
\end{aligned}
$$

concluding the proof. □

Lemma 4.2.

Proof. The transformation T now works as follows. Given $(\mathbf{B}, \mathbf{c}) \in R_q^{m\times(m-d)} \times R^{m-d}$ 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 $\mathbf{A}' = \mathbf{AU}$. 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 $\mathbf{B}^T\mathbf{b} = \mathbf{c} \bmod qR$, and returns $(\mathbf{A}', \mathbf{b})$. Note that if $\mathbf{c} = \mathbf{B}^T\mathbf{e} \bmod qR$ 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 $\mathbf{s} \in R_q^d$ such that $\mathbf{b}-\mathbf{e} = \mathbf{A}'\mathbf{s} \bmod qR$. 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 $R_q^d \times R_q$ that is either $A_{\mathbf{s},\psi}$ or $U(R_q^d \times R_q)$.
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. By the quantity defined 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 $\overline{\mathbf{b}} \in R_q^d$ 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 $\overline{\mathbf{b}} = \overline{\mathbf{A}}\mathbf{s}+\mathbf{x} \bmod qR^\vee$ for $\mathbf{x}$

sampled from ψ^d. On the other hand, if $\mathcal{D} = U(R_q^d \times R_q)$, then $\overline{\mathbf{b}}$ is uniform over R_q^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 $b' = b + \langle \mathbf{a}', \overline{\mathbf{b}} \rangle \mod qR$;
- Output $(\mathbf{a}', b')$.

First, we verify that $(\mathbf{a}', b')$ indeed belongs to $R_q^d \times R_q$. 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 $\overline{\mathbf{b}} \in R_q^d$, $\langle \mathbf{a}', \overline{\mathbf{b}} \rangle$ is in R. It thus holds that b' is in R_q.

As $-(\overline{\mathbf{A}})^{-T}$ is invertible modulo qR, and $\mathbf{a}$ is uniform in R_q^d, then $\mathbf{a}'$ 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 $b = \langle \mathbf{a}, \mathbf{s} \rangle + e \mod qR$ for some $e \hookleftarrow \psi$, and $\overline{\mathbf{b}} = \overline{\mathbf{A}}\mathbf{s} + \mathbf{x} \mod qR$. It holds that

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

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