# Hardness of M-LWE with General Distributions and Applications to Leaky Variants

Katharina Boudgoust[1], Corentin Jeudy[2], Erkan Tairi[3⋆], and Weiqiang Wen[4]

katharina.boudgoust@lirmm.fr, corentin.jeudy@orange.com, erkan.tairi@berkeley.edu, weiqiang.wen@telecom-paris.fr

[1] CNRS, Univ Montpellier, LIRMM
[2] Orange Labs, Applied Crypto Group
[3] University of California, Berkeley
[4] Télécom Paris, Institut Polytechnique de Paris

**Abstract.** The Module Learning With Errors (M-LWE) problem has become a fundamental hardness assumption for lattice-based cryptography. It offers an attractive trade-off between strong robustness guarantees, sometimes directly based on worst-case lattice problems, and efficiency of the subsequent cryptographic primitives. Different flavors of M-LWE have then been introduced towards improving performance. Such variants look at different secret-error distributions and might allow for additional hints on the secret-error vector. Existing hardness results however only cover restricted classes of said distributions, or are tailored to specific leakage models. This lack of generality hinders the design of efficient and versatile cryptographic schemes, as each new distribution or leakage model requires a separate and nontrivial hardness evaluation.

In this work, we address this limitation by establishing the hardness of search M-LWE under *general distributions*. As a first step, we show that M-LWE remains hard when the *error vector* follows an arbitrary bounded distribution with sufficient entropy, with some restriction on the number of samples. Building on this, we then reduce to the Hermite Normal Form (HNF) where the *secret-error vector* follows said arbitrary distribution. Overall, our result shows that the actual shape of the distribution does not matter, as long as it keeps sufficient entropy.

To demonstrate the versatility of our framework, we further analyze a range of leakage scenarios. By examining the residual entropy given the leakage, we show that our results of M-LWE with general distributions encompass various types of leakage. More precisely, we cover exact and approximate linear hints which are widely used in recent cryptographic designs, as well as quadratic, and even non-algebraic forms, some of which were not yet covered by any theoretical hardness guarantees. The generality of our results aims at facilitating future cryptographic designs and security analyses.

**Keywords:** Module Learning With Errors · General Distributions · Entropy · Leakage

## 1 Introduction

After being introduced in the pioneering work by Regev [Reg05], the Learning With Errors (LWE) problem has induced a number of cryptographic applications through its versatility and robustness. Its connection to worst-case lattice problems is a source of confidence in the security of the cryptosystems that rely on it. While said systems were lacking efficiency, the introduction of structured variants [SSTX09, LPR10, BGV12, LS15, PP19] like Module Learning With Errors (M-LWE) [LS15] allowed for new optimizations and trade-off avenues, while preserving some trustworthy guarantees of robustness. Since then, many primitives have become practical thanks to these theoretical advances. The most undeniable examples are the recently standardized key encapsulation and signature schemes ML-KEM [BDK+18] and ML-DSA [DKL+18], whose security is

⋆ Work done while the author was at ENS Paris.

© IACR 2026. This article is the full version of the version to be published by Springer-Verlag in the proceedings of PKC 2026.

directly based on the hardness of M-LWE. Similarly, a plethora of other primitives benefited from the versatility of M-LWE like advanced encryption schemes including fully homomorphic encryption [MAM+24], zero-knowledge arguments [KLSS23], threshold schemes [DKM+24], etc. The caveat, however, is that the genericity of M-LWE opens up to very different parameter selections, all of which must be carefully assessed to ensure the best balance between efficiency and security.

Concretely, for a ring $R$, the M-LWE problem consists in recovering a secret vector $\mathbf{s}$ in an $R$-module given some noisy linear system $(\mathbf{A}, \mathbf{t} = \mathbf{A}\mathbf{s}+\mathbf{e} \bmod q)$ with uniformly random $\mathbf{A}$, where $\mathbf{e}$ is a somewhat short error, or distinguishing such a pair $(\mathbf{A}, \mathbf{t})$ from random. Several parameters can then be tweaked and yield very different regimes both in terms of efficiency and exposure to efficient attacks: the degree $n$ of $R$, the rank $d$ of the module (i.e., the dimension of $\mathbf{s}$), the modulus $q$, the number of samples $m$ (i.e., the dimension of $\mathbf{e}$), but also the distributions of $\mathbf{s}$ and $\mathbf{e}$. Far from being mere theoretical constraints, hasty parameter selection can lead to security issues rendering the M-LWE problem easy to solve (e.g., large modulus, noise-to-sample ratio too small, small lattice dimension $nd$, etc.). In particular, the distributions of $\mathbf{s}$ and $\mathbf{e}$ vary from one construction to the next so that they are tailored to make said constructions as efficient as possible. This multiplicity of distributions contrasts with the very few theoretical results that support them with proven mathematical evidence. Although not too concerning for practical constructions (as the hardness is estimated using the well-known lattice estimator [APS15]), it still questions the theoretical foundation of said systems and somewhat weakens our confidence in the robustness of M-LWE in such new parameter settings.

To make up for this lack of theoretical arguments, many papers have proven results for specific choices of distributions, trying to cover the most widely used ones. Although they do not cover the range of parameters used in practical schemes, they still act as a confidence booster that nothing is fundamentally wrong with these distributions and, if attacks arise, one could always increase the parameters to fall in the scope of the reduction. The first such results focused on the secret distribution [GKPV10, BLP+13, Mic18]. They proved that it could be taken to be uniform binary (or with a small bound) while being as hard as the standard formulation of LWE (i.e., with $\mathbf{s}$ uniform over $\mathbb{Z}_q^d$ and $\mathbf{e}$ from a discrete Gaussian distribution). They were later generalized to M-LWE in [BJRW20, BJRW21, BJRW23], however, only the uniform distribution with small coefficients were covered. A study initiated by Brakerski and Döttling [BD20a] aimed at abstracting the actual shape of the secret distribution while still being able to prove the hardness of LWE based on well-known assumptions. More precisely, the hardness was proven making the least amount of assumptions on the secret distribution beyond the entropy it carried. They were able to show that, as long as the secret distribution had sufficient min-entropy, LWE in that setting is no easier than the standard formulation. This result was later extended in [BD20b] to the ring setting proving that R-LWE [LPR10] with arbitrary (albeit with sufficient entropy) distributions is at least as hard as (a generalized version of) NTRU [HPS98]. Both these approaches were later extended to the case of M-LWE in [LWZW20] (reduction from standard M-LWE) and [BJRW22] (reduction from Module-NTRU). Although these provide evidence that the actual shape of the secret distribution is not so important in the hardness of (R/M-)LWE, they all rely on Gaussian error distributions, and thus do not cover variants with different error distributions.

Changing the error distribution has shown to be a little more delicate, as it is the cornerstone of LWE. Indeed, even if no error leads to a trivial problem, having an error distribution that is too short or too sparse is also insecure as shown for example through algebraic attacks [AG11, STA20]. More precisely, there is an intricate constraint between the error distribution (and its sparsity) and the number of samples $m$. This constraint actually comes up in the first reduction provided by [MP13], where they showed LWE with uniform small errors stays as hard as LWE with Gaussian errors, if $m = d + O(d/\log d)$ where $d$ is the secret dimension. The specific case of binary errors was then further investigated in [STA20], where they did not require a uniform distribution. They obtained different requirements on $m$ depending on the sparsity of the distribution. The former was later extended to the module setting in [BJRW23] showing similar constraints for uniform bounded errors. But again, only variants of (M-)LWE with uniform bounded errors or sparse binary errors were provided with this evidence of robustness, which falls short of generally covering the distributions used across the cryptographic literature. Taking the opposite view, wandering off the distributions covered by reductions or estimated using [APS15] is often frowned upon, thus limiting the range of available distributions to parameterize LWE with.

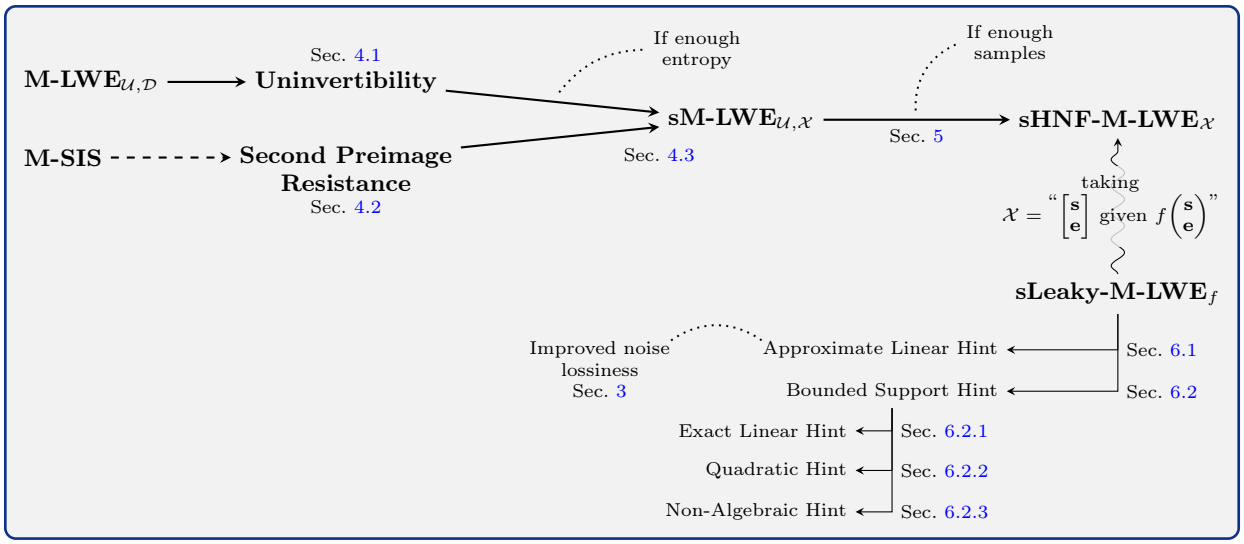


**Fig. 1.1.** High-level overview of our results and proof structure. The prefix "**s**" refers to the search variant. The subscript $\mathcal{U}$ designates a uniform distribution of secrets modulo $q$, while $\mathcal{D}$, $\mathcal{X}$ are error distributions representing the starting distribution (for which hardness of M-LWE is assumed) and target distribution (for which we want to prove hardness of M-LWE), respectively. In the case of HNF-M-LWE, $\mathcal{X}$ is the joint distribution of $(\mathbf{s}, \mathbf{e})$. The dashed arrow signifies it is optional and can be removed by relying on the statistical version of Lemma 4.3. The squiggly arrow means Leaky-M-LWE$_f$ is interpreted as HNF-M-LWE$_{\mathcal{X}}$ for $\mathcal{X}$ being the conditional distribution of $(\mathbf{s}, \mathbf{e})$ given the leakage $f(\mathbf{s}, \mathbf{e})$. The angled arrows specify which families of leakage functions $f$ are tackled in this paper and where.

Meanwhile, several works have introduced so-called *hint* or *extended* variants of LWE (and structured counterparts) so as to give more flexibility and improve or unlock certain applications. This has been done not only in the context of reductions [BLP+13, BJRW21, BJRW23, BK25], but also in many different cryptographic constructions [AP12, ALS16, AA16, LNS21, MKMS22, LN22, DKL+23, KLSS23, HPS23, MS23, WLL24, ENP24, PS24]. These variants not only give $(\mathbf{A}, \mathbf{t})$ to the adversary, but also give out *hints* or *leakages* $f(\mathbf{s}, \mathbf{e})$ for a leakage function $f$ that vary depending on the application. Although these variants generally lead to more efficient primitives, one has to ensure that the hint or leakage does not provide the adversary sufficient information to solve LWE. Indeed, the leakage changes the distribution of $(\mathbf{s}, \mathbf{e})$ from the adversary's perspective because this side information reduces the original secret-error space. In some cases, reductions from known flavours of LWE are proven, but always adopt a tailored approach to cover their specific variant. For that, one of the popular solutions is to characterize the distribution of $(\mathbf{s}, \mathbf{e})$ conditioned on some specific leaked value $f(\mathbf{s}, \mathbf{e}) = c$ (see e.g., [KLSS23, Lem. 7] or [ENP24, Lem. 1]). This is a non-trivial task in general, which is why most results, to our knowledge, impose Gaussian distributions in order to rely on their nice convolution properties. On the other hand, certain variants are not supported by any theoretical insights, which makes it unclear whether the hints hinder security. Overall, because the distribution of $(\mathbf{s}, \mathbf{e})$ changes due to the leakage, there is no general hardness result covering these cases.

### 1.1 Our Contributions

In this paper, we provide a general study of the hardness of M-LWE with arbitrary error and secret distributions. Such a global study is especially motivated by the recent introduction of various leaky M-LWE variants, and was, prior to our work, was lacking. In particular, we provide the first hardness results for non-linear leakage. We provide a detailed description of our contributions below, with a high-level pictorial description of our results given in Figure 1.1.

We start by focusing on general error distributions in Section 4. We generalize the approach of [MP13, STA20,BJRW23] from bounded uniform to general bounded error distributions, placing only minimal amount of requirements on it. Analogous to the series of works initiated by [BD20a] for the secret distribution, we show that one can prove the hardness of M-LWE with bounded error distribution $\mathcal{X}$ from standard M-LWE, solely based on the entropy of $\mathcal{X}$. The restriction of $\mathcal{X}$ being bounded comes naturally in the M-LWE problem as one usually uses small errors to obtain more efficient schemes. In addition to this bound, we obtain a limitation on the number of samples. Abstracting the distribution unfortunately makes this limitation on the number of samples harder to pin down than with the uniform over hypercube case covered in [BJRW23]. Nevertheless, the combination of both these small restrictions helps ruling out trivial attacks for extreme distributions (for example when all the entropy would lie in a single coefficient). We discuss further these restrictions in Section 4. Overall, the main constraint on the error distribution comes from its entropy. It then yields interesting insights on the hardness of M-LWE: the concrete shape of the error distribution does not matter too much, provided it has sufficient entropy. Our proof also suggests that as long as the entropy stays the same, the hardness of M-LWE for two different error distributions does not drastically change, even if the two distributions are not cryptographically close (in terms of statistical distance, Rényi divergence or other metrics). It then opens for attractive perspectives for cryptographic designs where one could opt for a more suitable error distribution and arguing the underlying hardness based on the entropy alone. Our main result is summarized with the following informal theorem, and the full statement is provided in Theorem 4.1. Following [MP13,BJRW23], the high-level proof strategy is to show uninvertibility (Section 4.1) and second-preimage-resistance (Section 4.2), which imply the hardness of the search variant of M-LWE with general error distribution $\mathcal{X}$. Second-preimage resistance can either be obtained statistically, or computationally assuming the hardness of M-SIS. Uninvertibility can be guaranteed assuming M-LWE with an error distribution $\mathcal{D}$ for which hardness has already been established.

**Theorem 1.1 (Informal).** *Let $m > d > k \geq 1$, modulus $q$ and ring $R$ be of degree $n$, and assume decision* M-LWE *with secret distribution $U(R_q^k)$ and error distribution $\mathcal{D}$ over $R^{m-d+k}$ and* M-SIS *in dimension $m-d$ are hard. Then, for a distribution $\mathcal{X}$ over $R^m$, the search* M-LWE *with secret distribution $U(R_q^d)$ and error distribution $\mathcal{X}$ is hard provided that*

$$H_\infty(\mathcal{X}) \gtrsim \lambda + \log_2\left(\mathrm{Vol}(\mathcal{B}_{n(m-d+k)}(r_{\mathcal{D},\mathcal{X}}))\right),$$

*where the $n(m-d+k)$-hyperball radius $r_{\mathcal{D},\mathcal{X}}$ depends on a norm bound on $\mathcal{X}$ and spectral norm bound on $\mathcal{D}$. Here, $\lambda$ denotes the security parameter and $\gtrsim$ refers to an asymptotic behavior.*

By setting $\mathcal{D}$ to the discrete Gaussian and $\mathcal{X}$ to the bounded uniform distributions, we recover the results of [BJRW23]. We then extend this result to the Hermite Normal Form (HNF) of M-LWE in Section 5. The HNF is often the preferred regime for M-LWE as it allows for replacing a large uniform secret by a secret of small norm, hence improving the efficiency of the subsequent cryptographic primitive. This regime is usually described by the fact that the secret distribution is the same as the error distribution, i.e., $\mathbf{t} = \mathbf{A}\mathbf{s}+\mathbf{e}$ for $\mathbf{s}$ drawn from $\mathscr{D}^d$ and $\mathbf{e}$ from $\mathscr{D}^m$. However, this suggests that all the entries of $\mathbf{s}$ and $\mathbf{e}$ are independent and identically distributed according to $\mathscr{D}$. A more natural and general formulation is to define $\mathbf{t} = [\mathbf{A} \mid \mathbf{I}_m]\mathbf{x}$ for $\mathbf{x}$ drawn from a distribution $\mathcal{X}$ over $R^{m+d}$. Here, $\mathbf{I}_m$ denotes the $m$-dimensional identity matrix over $R$. In that case, the vector $\mathbf{x}$ can be seen as the concatenation $[\mathbf{s} \mid \mathbf{e}]$, but no independence condition or specific shape is required for $\mathcal{X}$. Reducing M-LWE with $\mathbf{s}$ uniform in $R_q^d$ and $\mathbf{e}$ sampled from $\mathcal{X}$ to this formulation with $\mathbf{x}$ drawn from $\mathcal{X}$ is fairly common [ACPS09, LS15, BJRW23]. A small subtlety however arises in the reduction when $\mathcal{X}$ is not invariant under permutation, e.g., not of the form $\mathcal{X} = \mathscr{D}^{m+d}$. Indeed, the reduction interprets $\mathbf{e}$ in the former problem as $\mathbf{x}$ in the latter. But for that, it requires to identify an invertible submatrix of $\mathbf{A}$ and to permute $\mathbf{e}$ so that the new "secret part" (the first $d$ entries of $\mathbf{x}$) matches with the rows of this submatrix. As a result, the distribution of the resulting $\mathbf{x}$ is $\mathbf{P}\mathcal{X}$ for a random permutation matrix $\mathbf{P}$. Albeit equivalent to the standard HNF formulation for permutation-invariant distributions, it shows that the HNF transform needs to account for this permutation at the risk of yielding an insecure problem. We thus introduce the *permuted* HNF-M-LWE (pHNF-M-LWE) problem with distribution $\mathcal{X}$ as the problem of finding $\mathbf{x}$ given $\mathbf{A}$ and $\mathbf{t} = [\mathbf{A} \mid \mathbf{I}_m]\mathbf{P}\mathbf{x}$ for $\mathbf{x}$ drawn from $\mathcal{X}$, and random permutation matrix $\mathbf{P}$. Our proof then shows that

pHNF-M-LWE with rank $d$, $m$ samples and secret-error distribution $\mathcal{X}$ is at least as hard as M-LWE with rank $d$, $m+d$ samples, and error distribution $\mathcal{X}$. We also provide an alternative (looser) reduction to avoid this permutation intricacy and reduce to the regular formulation of HNF-M-LWE.

At this stage, we proved in Sections 4+5 that HNF-M-LWE with secret-error distribution $\mathcal{X}$ and M-LWE with error distribution $\mathcal{X}$ are at least as hard as the standard formulation of M-LWE if $\mathcal{X}$ has sufficient entropy. Even though not considered surprising results by themselves, they required substantial technical rigorousness to be proven for M-LWE. Besides giving more freedom in the choice of distributions to instantiate M-LWE with, our findings also provide a systematic study of *leaky* and *hint* variants of M-LWE by simply relying on entropy considerations. Concretely, if one is given $f(\mathbf{x})$ in addition to $[\mathbf{A} \mid \mathbf{I}_m]\mathbf{x}$, we can consider the conditional distribution of $\mathbf{x}$ given $f(\mathbf{x})$. Having abstracted the distribution in our proof then allows to encompass these oddly-shaped distributions as long as they are bounded (which follows from the bound on $\mathbf{x}$) and that they have sufficient entropy. Proving hardness then comes down to estimating the conditional entropy of $\mathbf{x}$ given $f(\mathbf{x})$, i.e., how much information $f(\mathbf{x})$ provides on $\mathbf{x}$. For that we use the average conditional min-entropy, denoted $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x}))$, which is well suited for security arguments, as detailed in Section 6.

Computing this quantity is non-trivial in general, but one can still derive meaningful lower bounds that are sufficient in some cases. We start in Section 6.1 with studying the setting covered in [AP12,KLSS23,ENP24] of approximate linear hints $f(\mathbf{x}) = \mathbf{M}\mathbf{x} + \mathbf{f}$, with $\mathbf{M}$ arbitrary (possibly adversarial) and $\mathbf{f}$ Gaussian. Our result provides the hardness of this variant even when $\mathbf{x}$ does not follow a Gaussian distribution. As we explained, previous works required $\mathbf{x}$ to be Gaussian in order to characterize the conditional distribution of $\mathbf{x}$ given $f(\mathbf{x})$ using convolution techniques. In our case, we can lower bound $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x} + \mathbf{f})$ by $H_\infty(\mathbf{x}) - O(\sqrt{nk}\|\mathbf{M}\mathbf{x}\|_2/\sigma)$, where $n$ is the ring degree, $k$ is the number of hints and $\sigma$ is the Gaussian parameter of $\mathbf{f}$, and without requiring $\mathbf{x}$ to be Gaussian. The latter point was identified as a limitation in the Raccoon signature scheme [dPKPR24], where the hardness of *Hint*-M-LWE [KLSS23] required Gaussian distributions which were incompatible with the sums of uniforms of [dPKPR24]. As a side contribution of potentially independent interest, detailed in Section 3, we generalize the notion of noise lossiness from [BD20a] and results thereof to what we call *linear noise lossiness*. We indeed show that their "*flooding at the source*" technique is not always optimal. At a high level, they decompose the Gaussian $\mathbf{f}$ into $\mathbf{M}\mathbf{f}_1 + \mathbf{f}_2$, argue that $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x} + \mathbf{f}) \geq \widetilde{H_\infty}(\mathbf{x} \mid \mathbf{x} + \mathbf{f}_1)$, and then choose the Gaussian parameter of $\mathbf{f}_1$ sufficiently large to flood $\mathbf{x}$. This Gaussian decomposition and inequality are a bit loose, and we show that one can directly use $\mathbf{f}$ to hide the leaked information $\mathbf{M}\mathbf{x}$. Getting rid of these first two steps removes unnecessary restrictions, giving more flexibility for the trade-off between the size of the mask $\mathbf{f}$ and the remaining guaranteed min-entropy of $\mathbf{x}$.

We then show in Section 6.2 that our approach also covers more general leakage variants, some of which were not proven hard prior to our work. We mentioned that the key methodology is to lower bound $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x}))$. This can be done for arbitrary leakage functions, as long as the range of $f(\mathbf{x})$ is bounded. Indeed, we have $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x})) \geq H_\infty(\mathbf{x}) - \log_2|f(\mathrm{Supp}(\mathcal{X}))|$. It means that for sufficiently small leakage space, one can derive a meaningful hardness result solely based on the entropy of $\mathcal{X}$. We note that if the entropy $H_\infty(\mathbf{x})$ of $\mathbf{x}$ is large enough to begin with, one can cover leakage that has an exponential leakage space, i.e., $\log_2|f(\mathrm{Supp}(\mathcal{X}))| = \Omega(\lambda)$, where $\lambda$ is the security parameter. Conversely, such large leakage spaces are precluded in proofs based on guessing arguments which are used when no better alternative is known (e.g., [MS23]). Additionally, this bound on $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x}))$ is agnostic to the specific shape of $f$ which then allows us to cover many different types of leakage. We describe in the paper how to cover existing variants with exact linear hints [ALS16,AA16,BJRW21,LN22,WLL24], i.e., $f$ linear (possibly adversarial), quadratic hints [MS23], but also non-algebraic hints [LNS21]. In particular, until now, each of these either required a tailored reduction or was not covered by any theoretical hardness evidence (e.g., Known-Covariance variant in [MS23]). We explain in details in Appendix A how to instantiate it to provide the first theoretical hardness evidence for the (module version of the) Known-Covariance variant proposed in [MS23]. To summarize, our results encompass most leakage variants, to our knowledge, by providing a unified methodology for studying their hardness.

### 1.2 Limitations and Open Problems

Similar to many other theoretical reductions, our results do not cover parameter ranges considered in practice. This is partly due to the generality of our results, which may lead to looser arguments than the ones focusing on specific settings. We thus interpret our findings as confidence booster in the studied variants of M-LWE. In particular, we think the impact of our work is most important when looking at variants of M-LWE where no prior reduction was known, such as non-linear yet algebraic leakage, e.g., [MS23], or non-algebraic leakage, e.g., from side-channel measurements. Tightening our results would reduce the gap between theory and practice, which we leave as an interesting research direction.

Our results hold for general number fields, including the case of degree $n = 1$, thus covering the unstructured case of LWE. Nonetheless, our analysis necessitates the evaluation of norms on the multiplication matrices of ring elements, which is better understood in (power-of-two) cyclotomic fields. On the opposite extreme, we highlight that our results, as stated in Theorem 1.1, require $d > 1$ and hence do not cover the ring setting. This was already left open for the general small secret distribution case. Indeed, previous papers rely on a low-rank factorization which is naturally not rank-preserving. We note that even [BD20b], which focuses on R-LWE with entropic secret, inherently cannot reach small secrets.

Perhaps the most relevant limitation of our work is that our main result (Theorem 4.1) only guarantees the hardness of the *search* variant of M-LWE with general error (and secret) distributions. This restriction is also present in prior works that study M-LWE with small error distributions [MP13, STA20, BJRW23]. Moreover, we note that the existing search-to-decision reductions for M-LWE are not error preserving for ring degree $n > 1$ [MM11, LS15]. We believe it is an important and challenging open problem to extend our results to the *decision* variant.

## 2 Preliminaries

We use $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$ to denote the set of natural numbers, the ring of integers, the field of reals and the field of complex numbers respectively. For two integers $a \leq b$, we define $[a, b] = \{a, \dots, b\}$ and $[b] = [1, b]$. For a positive integer $q$ and a ring $R$, we define the quotient ring $R_q = R/qR$. For a ring $R$, we denote by $R^\times$ its unit group, and we abuse notation by defining $\mathbb{N}^\times = \mathbb{N} \setminus \{0\}$ to be the set of positive integers. Vectors and matrices over a ring $R$ are written in bold lowercase and bold uppercase respectively. We use $\|\mathbf{x}\|_p$ to denote the usual $\ell_p$ norm over $\mathbb{C}^n$. Recall that for $p > r \geq 1$, we have $\|\cdot\|_p \leq \|\cdot\|_r \leq n^{1/r-1/p}\|\cdot\|_p$. We use $\mathcal{B}_{p,n}(r)$ to denote the closed $n$-dimensional $\ell_p$-norm ball of radius $r$, i.e., $\mathcal{B}_{p,n}(r) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_p \leq r\}$. For $\alpha, \beta \geq 1$, we define the $(\alpha, \beta)$-induced matrix norm of a matrix $\mathbf{A} \in \mathbb{C}^{n\times m}$ by $\|\mathbf{A}\|_{\alpha,\beta} = \sup_{\mathbf{x}\neq\mathbf{0}} \|\mathbf{A}\mathbf{x}\|_\beta / \|\mathbf{x}\|_\alpha$. This norm has simpler expressions for some known pairs $(\alpha, \beta)$. For example, we have

$$\|\mathbf{A}\|_{1,1} = \max_{j\in[m]} \sum_{i\in[n]} |a_{i,j}| \qquad \|\mathbf{A}\|_{2,\infty} = \max_{i\in[n]} \sqrt{\sum_{j\in[m]} |a_{i,j}|^2}$$

$$\|\mathbf{A}\|_{\infty,\infty} = \max_{i\in[n]} \sum_{j\in[m]} |a_{i,j}| \qquad \|\mathbf{A}\|_{1,2} = \max_{j\in[m]} \sqrt{\sum_{i\in[n]} |a_{i,j}|^2}$$

$$\|\mathbf{A}\|_{2,2} = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})},$$

where $\lambda_{\max}(\mathbf{A}^T\mathbf{A})$ denotes the largest eigenvalue. When $\alpha = \beta$, we abbreviate the matrix norm by $\|\cdot\|_{\alpha,\beta} =: \|\cdot\|_\alpha$. In particular, $\|\cdot\|_{2,2}$ is generally noted $\|\cdot\|_2$ in the literature and called the spectral norm.

### 2.1 Algebraic Number Theory

A number field $K = \mathbb{Q}(\zeta)$ is a field extension of $\mathbb{Q}$ of finite degree $n$ obtained by adjoining an algebraic number $\zeta$, i.e., $\zeta$ is a root of a rational polynomial. The number $\zeta$ is called an algebraic integer if it is root of an *integer* polynomial. The field $K$ is isomorphic to $\mathbb{Q}[x]/\langle f\rangle$ where $f$ is the minimal polynomial of $\zeta$.

The set of algebraic integers within $K$ defines a ring $R$ called the ring of integers of $K$. We also define the ring $K_{\mathbb{R}} = K \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{R}[x]/\langle f \rangle$. For an ideal $\mathcal{I}$ of $R$, we define its norm $N(\mathcal{I})$ as the index of $\mathcal{I}$ in $R$, i.e, $N(\mathcal{I}) = |R/\mathcal{I}|$. We call a prime integer $q$ unramified in $R$ if all the distinct prime ideal factors of the ideal $qR$ have ramification index 1, that is $qR = \prod_{i\in[\kappa]} \mathfrak{p}_i$ where $\mathfrak{p}_1, \ldots, \mathfrak{p}_\kappa$ are distinct prime ideals of $R$.

The number field $K$ can be seen as an $n$-dimensional $\mathbb{Q}$-vector space with basis $\{\zeta^j\}_{j\in[0,n-1]}$, meaning that $a \in K$ can be expressed as $a = \sum_{0\le j<n} a_j\zeta^j$ with $a_j \in \mathbb{Q}$. We can therefore define the coefficient embedding $\tau$ such that $\tau(a) = [a_0|\ldots|a_{n-1}]^T$. We sometimes use $\tau_j(a) = a_j$ for simplicity. It then holds that the norms of $\mathbb{R}^n$ induce norms on $K$ (and $K_{\mathbb{R}}, R$) by $\|a\| := \|\tau(a)\|$. Multiplication in $K$ also maps to a matrix-vector multiplication of the coefficient embeddings as $\tau(ab) = M_\tau(a)\tau(b)$ where $M_\tau$ is a ring isomorphism defined by $M_\tau : a \in K \mapsto [\tau(a) \mid \tau(a\zeta) \mid \ldots \mid \tau(a\zeta^{n-1})] \in \mathbb{Q}^{n\times n}$

A popular class of number fields is that of cyclotomic fields. The $m$-th cyclotomic field is $K_m = \mathbb{Q}(\zeta_m)$ where $\zeta_m$ is a primitive $m$-th root of unity, e.g., $\zeta_m = \exp(i2\pi/m)$. The minimal polynomial, called the $m$-th cyclotomic polynomial, is $\Phi_m = \prod_{j\in\mathbb{Z}_m^\times}(x - \zeta_m^j)$. It has degree $n = \varphi(m)$. We define the Vandermonde matrix of $K_m$ as $\mathbf{V}_m = [\zeta_m^{jk}]_{j\in\mathbb{Z}_m^\times, k\in[0,n-1]}$. For example, when $m = 2^{\ell+1}$, $\Phi_m = x^n + 1$ with $n = 2^\ell$, and $\mathbf{V}_m = \sqrt{n}\mathbf{U}$ with $\mathbf{U}$ unitary. In that case, we call $K_m$ a power-of-two cyclotomic field. In the latter $M_\tau(a)$ is the nega-circulant matrix with first column $\tau(a)$. In cyclotomic fields, we also define the conjugate of an element $a$ by $a^* = \sigma_{-1}(a)$ where $\sigma_{-1}$ is the field automorphism of $K$ defined by $\zeta \mapsto \zeta^{-1}$. Conjugation corresponds to transposing $M_\tau$, i.e., $M_\tau(a)^T = M_\tau(a^*)$. As a result, in a power-of-two cyclotomic field of degree $n$, we have $\tau(a^*) = [a_0 \mid -a_{n-1} \mid \ldots \mid -a_1]^T$.

We extend these notions to vectors and matrices over $K$. Namely, for $\mathbf{a} \in K^d$, $\tau(\mathbf{a})$ is the concatenation of the coefficient embeddings of its entries. Similarly, $M_\tau(\mathbf{A})$ is the block matrix composed of the $M_\tau([\mathbf{A}]_{i,j})$. Vector and matrix norms are then extended to $K^d$ and $K^{d\times k}$ through these embedding $\tau(\mathbf{a})$ and $M_\tau(\mathbf{A})$ respectively. We also define the conjugate of a vector or matrix by taking the conjugate of each entry and transposing the result, i.e., if $\mathbf{x} = [x_1 \mid \ldots \mid x_d]^T \in K^d$, then $\mathbf{x}^*$ is the row vector $[x_1^* \mid \ldots \mid x_d^*]$.

As we generally work with elements of $R_q$, we sometimes need to ensure these elements are invertible in the quotient ring $R_q$. Necessary conditions are provided for example in [LS18] in cyclotomic rings, where it suffices that $q$ is sufficiently large compared to the $\ell_2$ or $\ell_\infty$ norm of $y \in R_q$. We generalize their invertibility result in cyclotomic rings to be expressed in terms of the $\ell_p$ norm for an arbitrary $p$. As $R_q$ is isomorphic to $\mathbb{Z}_q[x]/\langle\Phi_m(x)\rangle$ through $\zeta \mapsto x$, we analyze invertibility in the latter.

**Lemma 2.1 ([LS18, Thm. 1.1] adapted).** *Let $m = \prod_i p_i^{e_i}$ for $p_i$ distinct primes and $e_i \in \mathbb{N}^\times$. Let $z = \prod_i p_i^{f_i}$ for arbitrary $f_i \in [1, e_i]$. Let $q$ be a prime such that $q = 1 \bmod z$ and $ord_m(q) = m/z$. Let $y \in \mathbb{Z}_q[x]/\langle\Phi_m(x)\rangle$ such that*

$$\exists p \in \mathbb{N}^\times, 0 < \|\tau(y)\|_p < \frac{\varphi(z)^{\min(1/2,1/p)}}{\|\mathbf{V}_z\|_2} q^{1/\varphi(z)}.$$

*Then $y \in (\mathbb{Z}_q[x]/\langle\Phi_m(x)\rangle)^\times$.*

*Proof.* By [LS18, Thm. 2.3], the way $m, z, q$ are set ensures the cyclotomic polynomial factors as $\Phi_m(x) \bmod q = \prod_{i=1}^{\varphi(z)}(x^{m/z} - r_i)$ where the $x^{m/z} - r_i$ are distinct and irreducible in $\mathbb{Z}_q[x]$. We then have $\Phi_z(x) \bmod q = \prod_{i=1}^{\varphi(z)}(x - r_i)$. Now let $y$ be as in the lemma statement. For $i \in [0, \varphi(m)/\varphi(z) - 1]$, define

$$t_i = \sum_{j=0}^{\varphi(z)-1} \tau_{j\varphi(m)/\varphi(z)+i}(y) \cdot x^j,$$

Note that $\varphi(m)/\varphi(z) = m/z$, which we later use for conciseness. By [LS18, Lem. 3.2], if one of the $t_i$ is invertible in the fully-splitting ring $\mathbb{Z}_q[x]/\langle\Phi_z(x)\rangle$, then $y$ is invertible in $\mathbb{Z}_q[x]/\langle\Phi_m(x)\rangle$. Because $y \neq 0$, there exists $k$ such that $t_k \neq 0$. Additionally, we have $\|\tau(t_k)\|_p^p = \sum_{j=0}^{\varphi(z)-1}\left|\tau_{jm/z+k}(y)\right|^p \le \sum_{j=0}^{\varphi(m)-1}|\tau_j(y)|^p = \|\tau(y)\|_p^p$. Hence, we get

$$0 < \|\tau(t_k)\|_p < \frac{\varphi(z)^{\min(1/2,1/p)}}{\|\mathbf{V}_z\|_2} q^{1/\varphi(z)},$$

Assume towards contradiction $t_k$ is not invertible in $\mathbb{Z}_q[x]/\langle \Phi_z(x)\rangle$. By the Chinese Remainder Theorem, it yields there exists $i \in [1, \varphi(z)]$ such that $t_k(r_i) = t_k \bmod x - r_i = 0$. Define the ideal $\mathcal{I} = \{s \in \mathbb{Z}[x]/\langle \Phi_z(x)\rangle : s(r_i) = s \bmod x - r_i = 0 \bmod q\}$. $\mathcal{I}$ is indeed clearly an additive group. Also, because $x - r_i$ divides $\Phi_z(x) \bmod q$, $\mathcal{I}$ is also an ideal of $\mathbb{Z}[x]/\langle \Phi_z(x)\rangle$. By the CRT representation, it also follows that $N(\mathcal{I}) = |(\mathbb{Z}[x]/\langle \Phi_z(x)\rangle)/\mathcal{I}| = q$.

Looking at the ideal lattice $\sigma(\mathcal{I})$ in the canonical embedding, we have by [PR07, Lem. 6.2] that $\lambda_1^{p'}(\sigma(\mathcal{I})) \geq \varphi(z)^{1/p'} N(\mathcal{I})^{1/\varphi(z)} = \varphi(z)^{1/p'} q^{1/\varphi(z)}$ for any $p' \in \mathbb{N}^\times$, where $\lambda_1^{p'}(\mathcal{L}) = \min_{\mathbf{x}\in\mathcal{L}\setminus\{\mathbf{0}\}} \|\mathbf{x}\|_{p'}$ is the first minimum of the lattice $\mathcal{L}$ in $\ell_{p'}$ norm. We however need a similar bound for the ideal lattice $\tau(\mathcal{I})$ in the coefficient embedding. We know that $\sigma = \mathbf{V}_z \tau$ where $\mathbf{V}_z$ is the Vandermonde matrix of the $z$-th cyclotomic field $K_z$.

Let $w \in \mathcal{I} \setminus \{0\}$ such that $\mathbf{w} = \tau(w)$ has norm $\|\mathbf{w}\|_p = \lambda_1^p(\tau(\mathcal{I}))$. If $p > 2$, then

$$\lambda_1^p(\tau(\mathcal{I})) = \|\mathbf{w}\|_p \geq \varphi(z)^{\frac{1}{p}-\frac{1}{2}} \|\mathbf{w}\|_2 \geq \frac{\varphi(z)^{\frac{1}{p}-\frac{1}{2}}}{\|\mathbf{V}_z\|_2} \|\sigma(w)\|_2 \geq \frac{\varphi(z)^{\frac{1}{p}-\frac{1}{2}}}{\|\mathbf{V}_z\|_2} \varphi(z)^{1/2} q^{1/\varphi(z)},$$

which gives $\lambda_1^p(\tau(\mathcal{I})) \geq \frac{\varphi(z)^{1/p}}{\|\mathbf{V}_z\|_2} q^{1/\varphi(z)}$. If $p \leq 2$, we simply have $\lambda_1^p(\tau(\mathcal{I})) \geq \lambda_1^2(\tau(\mathcal{I})) \geq \frac{\varphi(z)^{1/2}}{\|\mathbf{V}_z\|_2} q^{1/\varphi(z)}$. Combining both cases then simply gives $\lambda_1^p(\tau(\mathcal{I})) \geq \frac{\varphi(z)^{\min(1/2,1/p)}}{\|\mathbf{V}_z\|_2} q^{1/\varphi(z)}$.

We then have $t_k \in \mathcal{I}$ and $0 < \|\tau(t_k)\|_p < \lambda_1^p(\tau(\mathcal{I}))$, which is a contradiction. It then means $t_k$ is invertible in $\mathbb{Z}_q[x]/\langle \Phi_z(x)\rangle$ and in turn that $y$ is invertible in $\mathbb{Z}_q[x]/\langle \Phi_m(x)\rangle$ as desired. □

We can also specialize the result in the case of power-of-two cyclotomics akin [LS18, Cor. 1.2] as follows.

**Lemma 2.2 ([LS18, Cor. 1.2] adapted).** *Let $n \geq \kappa > 1$ be powers of 2 and $q = 2\kappa + 1 \bmod 4\kappa$ a prime. Let $y \in \mathbb{Z}_q[x]/\langle x^n + 1\rangle$ such that $0 < \|\tau(y)\|_p < \kappa^{\min(0,1/p-1/2)} q^{1/\kappa}$ for some $p \in \mathbb{N}^\times$, then $y \in (\mathbb{Z}_q[x]/\langle x^n + 1\rangle)^\times$.*

### 2.2 Gaussians

For $s > 0$, we define the centered Gaussian function of parameter $s$ by $\rho_s : \mathbf{x} \in \mathbb{R}^n \mapsto \exp(-\pi\|\mathbf{x}\|_2^2/s^2)$. We can then define the discrete Gaussian over $\mathbb{Z}^n$ through its probability mass function $\mathcal{D}_{\mathbb{Z}^n,s} = \frac{\rho_s}{\rho_s(\mathbb{Z}^n)} \mathbf{1}_{\mathbb{Z}^n}$, where $\mathbf{1}_{\mathbb{Z}^n}$ is the indicator function of $\mathbb{Z}^n$ and $\rho_s(\mathbb{Z}^n) = \sum_{\mathbf{y}\in\mathbb{Z}^n} \rho_s(\mathbf{y})$. We extend the discrete Gaussian distribution to (vectors of) ring elements as $\mathcal{D}_{R^d,s} = \tau^{-1}(\mathcal{D}_{\mathbb{Z}^{nd},s})$, i.e., the vector of $\mathbb{Z}^{nd}$ of concatenated coefficient embeddings is sampled from $\mathcal{D}_{\mathbb{Z}^{nd},s}$, where $n$ denotes the degree of $R$. We define the smoothing parameter of $\mathbb{Z}^n$, introduced in [MR04] for arbitrary lattices, as $\eta_\varepsilon(\mathbb{Z}^n) = \min\{s > 0 : \rho_{1/s}(\mathbb{Z}^n) = 1+\varepsilon\}$ for a given $\varepsilon > 0$. It was recently shown in [EWY23] that $\eta_\varepsilon(\mathbb{Z}^n) \approx \sqrt{\ln(2n/\varepsilon)/\pi}$. We also recall the standard tail bound in $\ell_2$ norm from [Ban93, Lem. 1.5]. We only specify it to $\mathbb{Z}^n$ but note it holds for all $n$-dimensional lattices.

**Lemma 2.3 ([Ban93, Lem. 1.5]).** *Let $n \in \mathbb{N}^\times$ and $s > 0$. For all $t \geq 1/\sqrt{2\pi}$, it holds $\mathbb{P}_{\mathbf{x}\sim\mathcal{D}_{\mathbb{Z}^n,s}}[\|\mathbf{x}\|_2 > ts\sqrt{n}] < (t\sqrt{2\pi e}e^{-\pi t^2})^n$.*

### 2.3 Entropy

For a distribution $\mathcal{X}$ on a support $X$, we define its min-entropy by $H_\infty(\mathcal{X}) = -\log_2 \max_{x'\in X} \mathbb{P}_{x\sim\mathcal{X}}[x = x']$. Throughout the paper, we interchangeably consider the min-entropy of $\mathcal{X}$ and the min-entropy of a random variable $x$ distributed according to $\mathcal{X}$. We also consider the average predictability through the average conditional min-entropy of distribution $\mathcal{X}$ conditioned on a distribution $\mathcal{Y}$ supported over $Y$ by

$$\widetilde{H_\infty}(\mathcal{X} \mid \mathcal{Y}) = -\log_2 \left( \mathbb{E}_{y\sim\mathcal{Y}} \left[ \max_{x'\in X} \mathbb{P}_{x\sim\mathcal{X}}[x = x' \mid y] \right] \right)$$

$$-\log_2\left(\sum_{y'\in Y}\mathbb{P}_{y\sim\mathcal{Y}}[y=y']\max_{x'\in X}\mathbb{P}_{\substack{x\sim\mathcal{X}\\ y\sim\mathcal{Y}}}[x=x'\mid y=y']\right).$$

We note that other definitions of conditional entropy are possible, but the average one is better suited for security arguments. We elaborate on this subject in Section 6. By abuse of language, we may use conditional entropy or residual entropy to designate the average conditional min-entropy. We recall the following results of [DORS08] on the conditional entropy, which we use in Section 6.2.

**Lemma 2.4 ([DORS08, Lem. 2.2]).** *Let $\mathcal{X},\mathcal{Y},\mathcal{Z}$ be random variables. Then*

1. *For any $\delta>0$, the conditional entropy $H_\infty(\mathcal{X}\mid\mathcal{Y}=y)$ is at least $\widetilde{H_\infty}(\mathcal{X}\mid\mathcal{Y})-\log_2(1/\delta)$ with probability $1-\delta$ over the choice of $y$.*
2. *If $\mathcal{Y}$ takes at most $N$ values, then $\widetilde{H_\infty}(\mathcal{X}\mid\mathcal{Y})\geq H_\infty(\mathcal{X})-\log_2 N$ and $\widetilde{H_\infty}(\mathcal{X}\mid\mathcal{Y},\mathcal{Z})\geq\widetilde{H_\infty}(\mathcal{X}\mid\mathcal{Z})-\log_2 N$.*

### 2.4 Assumptions

We now introduce the *Module Learning With Errors* problem as formalized in [LS15]. We however leave free the choice of secret and error distributions which is particularly relevant for our work.

**Definition 2.1 (M-LWE).** *Let $R$ be the ring of integers of a number field of degree $n$, and $d,m,q$ be in $\mathbb{N}^\times$. Let $\mathscr{D}_s$ be a secret distribution on $R^d$, and $\mathscr{D}_e$ an error distribution on $R^m$. The search version of the Module Learning With Errors problem, denoted $\text{M-LWE}_{n,d,m,q,\mathscr{D}_s,\mathscr{D}_e}$, asks to recover $\mathbf{s}$ given $\mathbf{A}\hookleftarrow U(R_q^{m\times d})$, and $\mathbf{b}=\mathbf{A}\mathbf{s}+\mathbf{e}\bmod qR$ for $\mathbf{s}\hookleftarrow\mathscr{D}_s$ and $\mathbf{e}\hookleftarrow\mathscr{D}_e$. The advantage of an adversary $\mathcal{A}$ is then $Adv[\mathcal{A}]=\mathbb{P}[\mathcal{A}(\mathbf{A},\mathbf{b})=\mathbf{s}]$. The decision version asks to distinguish such pairs $(\mathcal{P}_0)$ $(\mathbf{A},\mathbf{b}=\mathbf{A}\mathbf{s}+\mathbf{e}\bmod qR)$ from $(\mathcal{P}_1)$ $(\mathbf{A},\mathbf{u})$ where $\mathbf{u}\hookleftarrow U(R_q^m)$. The adversary's advantage in this case is $Adv[\mathcal{A}]=\left|\mathbb{P}_{(\mathbf{A},\mathbf{b})\sim\mathcal{P}_0}[\mathcal{A}(\mathbf{A},\mathbf{b})=1]-\mathbb{P}_{(\mathbf{A},\mathbf{u})\sim\mathcal{P}_1}[\mathcal{A}(\mathbf{A},\mathbf{u})=1]\right|$.*

Our proof, in its computational variant, leverages the *Module Short Integer Solution* problem, formalized in [LS15], which we extend to arbitrary $\ell_p$ norms.

**Definition 2.2 (M-SIS).** *Let $R$ be the ring of integers of a number field of degree $n$, and $d,m,q$ be in $\mathbb{N}^\times$. Let $p\in[1,\infty]$ and $\beta>0$. The Module Short Integer Solution problem, denoted $\text{M-SIS}^p_{n,d,m,q,\beta}$, asks to find $\mathbf{x}\in R^m$ such that $\mathbf{A}^T\mathbf{x}=\mathbf{0}\bmod qR$ and $0<\|\mathbf{x}\|_p\leq\beta$ given $\mathbf{A}\hookleftarrow U(R_q^{m\times d})$. The advantage of an adversary is $Adv[\mathcal{A}]=\mathbb{P}[\mathbf{A}^T\mathbf{x}=\mathbf{0}\wedge 0<\|\mathbf{x}\|_p\leq\beta:\mathbf{x}\leftarrow\mathcal{A}(\mathbf{A})]$.*

We say that a problem $P$ (among search/decision M-LWE and M-SIS) is $\varepsilon$-hard if for all probabilistic polynomial time adversary, $\text{Adv}[\mathcal{A}]\leq\varepsilon$. We sometimes refer to $\varepsilon$ as the hardness bound of problem $P$.

### 2.5 Function Families

Our proof of hardness in Section 4 focuses on the security properties of M-LWE when seen as function family. More generally, a function family $\mathcal{F}$ over a set of functions which all have domain $X$ and range $Y$ is simply a probability distribution over $F$. The M-LWE problem defined in Section 2.4 can be interpreted as a function family whose distribution relies on the distribution of $\mathbf{A}$. Moreover, we define another function family, called M-Knap, which is closely related to the M-LWE function family. As all functions can be unambiguously represented by a matrix $\mathbf{A}$, we say that an adversary is given the function $f$ as input when it is given its public representation, i.e., $\mathbf{A}$.

**Definition 2.3.** *Let $n,d,m,q\in\mathbb{N}^\times$, and $R$ be the ring of integers of a number field of degree $n$, and $X\subseteq R^m$. The $\text{M-Knap}(n,d,m,q,X)$ function family is the distribution obtained by sampling a matrix $\mathbf{A}\hookleftarrow U(R_q^{m\times d})$, 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 $\text{M-LWE}(n,d,m,q,X)$ function family is the distribution obtained by sampling $\mathbf{A}\hookleftarrow U(R_q^{m\times d})$ 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 present the security properties we need for arbitrary function families, as well as some useful results which are relevant to our proof. The hardness of the M-LWE problem from Section 2.4 can then be expressed as security properties of M-LWE$(n, d, m, q, X)$. Concretely, the one-wayness captures the hardness of the corresponding search problem, while the pseudorandomness captures the hardness of the decision problem.

**Definition 2.4.** *Let $X, Y$ be two sets, and $F$ be 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 $\varepsilon$-indistinguishable if for all PPT algorithm $\mathcal{A}$, it holds $|\mathbb{P}_{f\sim\mathcal{F}}[\mathcal{A}(f) = 1] - \mathbb{P}_{g\sim\mathcal{G}}[\mathcal{A}(g) = 1]| \leq \varepsilon$.*

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

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

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

*Each of the probabilities is implicitly reliant on the randomness $\mathcal{A}$ itself.*

We now give useful sufficient conditions to ensure some of these security properties. In particular, we extend the results of [MP13, Lem. 2.4] and [STA20, Lem. 6] to general distributions, i.e., not limited uniform distributions or unbalanced Bernoulli distributions.

**Lemma 2.5 ([STA20, Lem. 6] adapted).** *Let $\mathcal{F}$ be a function family with finite domain $X$. Let $\mathcal{X}$ be a distribution on $X$. For $\varepsilon = 2^{-H_\infty(\mathcal{X})}\mathbb{E}_{f\sim\mathcal{F}}[|f(X)|]$, it holds that $(\mathcal{F}, \mathcal{X})$ is $\varepsilon$-uninvertible, even against unbounded adversaries.*

*Proof.* We follow the proof method from [STA20, Lem. 6]. It holds that

$$\begin{aligned}
&\mathbb{P}_{(f,x)\sim\mathcal{F}\times\mathcal{X}}[\mathcal{A}(f, f(x)) = x] \\
&\qquad = \sum_{g\in\mathrm{Supp}(\mathcal{F})} \mathbb{P}_f[f = g]\mathbb{P}_{x\sim\mathcal{X}}[\mathcal{A}(g, g(x)) = x] \\
&\qquad = \sum_{g\in\mathrm{Supp}(\mathcal{F})} \mathbb{P}_f[f = g] \sum_{x'\in\mathrm{Supp}(\mathcal{X})} \mathbb{P}_x[x = x']\mathbb{P}_x[\mathcal{A}(g, g(x')) = x'].
\end{aligned}$$

Yet, for a fixed function $f$ and a fixed $y = f(x)$, the best attack for an adversary is to choose the element with the highest conditional probability. So in the latter sum, we have

$$\begin{aligned}
&\mathbb{P}[\mathcal{A}(g, g(x')) = x'] \\
&= \mathbb{P}[x \text{ is the preimage with the highest conditional probability in } g^{-1}(g(x))].
\end{aligned}$$

We then have

$$\sum_{x'\in\mathrm{Supp}(\mathcal{X})} \mathbb{P}_x[x = x']\mathbb{P}_x[\mathcal{A}(g, g(x')) = x']$$

$$
\begin{aligned}
&= \sum_{y \in g(X)} \frac{\max_{x' \in g^{-1}(y)} \mathbb{P}_x[x = x']}{\sum_{x'' \in g^{-1}(y)} \mathbb{P}_x[x = x'']} \cdot \sum_{x'' \in g^{-1}(y)} \mathbb{P}_x[x = x''] \\
&= \sum_{y \in g(X)} \max_{x' \in g^{-1}(y)} \mathbb{P}_x[x = x'] \\
&\leq \sum_{y \in g(X)} \max_{x' \in \mathrm{Supp}(\mathcal{X})} \mathbb{P}_x[x = x'] \\
&= 2^{-H_\infty(\mathcal{X})} |g(X)|.
\end{aligned}
$$

As a result, it yields

$$
\begin{aligned}
\mathbb{P}_{(f,x) \sim \mathcal{F} \times \mathcal{X}}[\mathcal{A}(f, f(x)) = x] &\leq \sum_{g \in \mathrm{Supp}(\mathcal{F})} \mathbb{P}_f[f = g] \cdot 2^{-H_\infty(\mathcal{X})} |g(X)| \\
&= 2^{-H_\infty(\mathcal{X})} \mathbb{E}_{f \sim \mathcal{F}}[|f(X)|] = \varepsilon,
\end{aligned}
$$

as claimed. □

**Lemma 2.6 ([MP13, Lem. 2.2]).** *Let $\mathcal{F}$ be a family of functions computable in polynomial time. Let $\mathcal{X}$ be a distribution on $X$. If $(\mathcal{F}, \mathcal{X})$ is $\varepsilon$-uninvertible and $\varepsilon'$-second preimage resistant, then it is also $(\varepsilon + \varepsilon')$-one-way.*

**Lemma 2.7 ([MP13, 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 $\varepsilon$-uninvertible, then so is $(\mathcal{G} \circ \mathcal{F}, \mathcal{X})$.*

Lastly, we recall the duality result between the M-LWE and M-Knap function families [MM11], generalized to the module setting in [BJRW23]. We only recall the results we need for our reduction and refer to [BJRW23, Lem. 4.1 & 4.2] for a more complete statement.

**Lemma 2.8.** *Let $n, d, m, k, \ell, q \in \mathbb{N}^\times$, and $R$ be the ring of integers of a number field of degree $n$, such that $q$ is an unramified prime in $R$, and $m \geq d + 1$. For any $a, b \in \mathbb{N}^\times$ with $a \geq b$, we define $\delta(a, b) = 1 - \Pr_{\mathbf{A} \sim U(R_q^{b \times a})}[\mathbf{A} R_q^a = R_q^b]$. Let $\mathcal{X}_1, \mathcal{X}_2$ be probability distributions over $X_1 \subseteq R^\ell$ and $X_2 \subseteq R^m$ respectively. If $(\text{M-LWE}(n, k, \ell, q, R^\ell), \mathcal{X}_1)$ is $\varepsilon_1$-pseudorandom, it then holds that $(\text{M-Knap}(n, \ell - k, \ell, q, R^\ell), \mathcal{X}_1)$ is $\varepsilon_1'$-pseudorandom where we have $\varepsilon_1' = 2\delta(\ell, \ell - k) + \varepsilon_1/(1 - \delta(\ell, k))$. Additionally, if $(\text{M-Knap}(n, m - d, m, q, R^m), \mathcal{X}_2)$ is $\varepsilon_2$-one-way, then $(\text{M-LWE}(n, d, m, q, R^m), \mathcal{X}_2)$ is $\varepsilon_2'$-one-way with $\varepsilon_2' = \delta(m, d) + \varepsilon_2/(1 - \delta(m, m - d))$.*

## 3 Linear Noise Lossiness

In this section, we consider the conditional min-entropy of a random vector $\mathbf{x}$ given a linear evaluation of $\mathbf{x}$, i.e., $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x} + \mathbf{f})$. More concretely, the matrix $\mathbf{M}$ can be adversarially chosen, while the mask vector $\mathbf{f}$ follows a prescribed distribution. This can be seen as a generalization of the *noise lossiness* notion of [BD20a] which was formulated for $\mathbf{M}$ being the identity matrix $\mathbf{I}$.

The overall goal of [BD20a] is the same as ours: to analyze the conditional min-entropy $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x} + \mathbf{f})$ for some given matrix $\mathbf{M}$. To do so, they assume that $\mathbf{f}$ follows a Gaussian distribution and, using Gaussian properties, they decompose the mask $\mathbf{f}$ into $\mathbf{M}\mathbf{f}_1 + \mathbf{f}_2$. This decomposition is possible only if the Gaussian parameter of $\mathbf{f}$ is larger than $\|\mathbf{M}\|_2$ times the Gaussian parameter of $\mathbf{f}_1$. Note that in [BD20a], all considered Gaussians $\mathbf{f}, \mathbf{f}_1$ and $\mathbf{f}_2$ are continuous, but the analysis can be generalized to discrete Gaussians using discrete convolution theorems, e.g., [DGPY20, Lem. 3] or [GMPW20, Thm. 3.1]. Then, they show $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x} + \mathbf{f}) \geq \widetilde{H_\infty}(\mathbf{x} + \mathbf{f}_1)$, which means it is enough to consider the conditional entropy of $\mathbf{x}$ given $\mathbf{x} + \mathbf{f}_1$. This approach is termed *flooding at the source* and motivated by the idea that it is less costly to hide $\mathbf{x}$ with $\mathbf{f}_1$ than hiding $\mathbf{M}\mathbf{x}$ with $\mathbf{f}$.

We observe that this approach is in general not optimal, and we can in fact generalize their result to directly bound $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x}+\mathbf{f})$ without relying on a Gaussian decomposition and (loose) inequalities like $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x}+\mathbf{f}) \geq \widetilde{H_\infty}(\mathbf{x}+\mathbf{f}_1)$. Hence, in our approach, the parameter for $\mathbf{f}$ does not need to be larger than $\|\mathbf{M}\|_2$, which allows for different trade-offs between the size of the mask $\mathbf{f}$ and the amount of entropy left in $\mathbf{x}$. Put differently, we challenge the idea that flooding at the source is most effective.

Even though our proof closely follows the original proof of [BD20a], we think this observation merits some spotlight and might be of independent interest.

**Lemma 3.1.** *Let $k, n \in \mathbb{N}^\times$ and $\mathbf{M} \in \mathbb{Z}^{n\times k}$ be a matrix. Let $\mathbf{x}$ be a random variable over $X \subseteq \mathbb{Z}^k$, and let $\mathbf{f}$ be a random variable following a discrete distribution $\chi$ over $\mathbb{Z}^n$. We denote by $\mathbf{M}X$ the support of $\mathbf{M}\mathbf{x}$, and by $Y$ the support of $\mathbf{M}\mathbf{x}+\mathbf{f}$. It then holds that*

$$\widetilde{H_\infty}(\mathbf{x}|\mathbf{M}\mathbf{x}+\mathbf{f}) \geq H_\infty(\mathbf{x}) - \log_2\left(\sum_{\mathbf{y}\in Y} \max_{\widehat{\mathbf{x}}\in\mathbf{M}X} \mathbb{P}_{\mathbf{f}\sim\chi}[\mathbf{f}=\mathbf{y}-\widehat{\mathbf{x}}]\right).$$

*The equality holds when $\mathbf{x}$ follows a uniform distribution. The result also holds true for continuous distributions by replacing the discrete sum with integrals and using density functions.*

*Proof.* We follow the blueprint of [BD20a, Lem. 5.1] with a few tweaks to generalize it. Writing the definition of the conditional entropy yields

$$\begin{aligned}
\widetilde{H_\infty}(\mathbf{x}|\mathbf{M}\mathbf{x}+\mathbf{f}) &= -\log_2\left(\mathbb{E}_{\mathbf{y}}[\max_{\mathbf{x}^\star\in X} \mathbb{P}_{\mathbf{x},\mathbf{f}}[\mathbf{x}=\mathbf{x}^\star|\mathbf{M}\mathbf{x}+\mathbf{f}=\mathbf{y}]]\right)\\
&= -\log_2\left(\sum_{\mathbf{y}\in Y} \mathbb{P}_{\mathbf{x},\mathbf{f}}[\mathbf{M}\mathbf{x}+\mathbf{f}=\mathbf{y}] \cdot \max_{\mathbf{x}^\star\in X} \mathbb{P}_{\mathbf{x},\mathbf{f}}[\mathbf{x}=\mathbf{x}^\star|\mathbf{M}\mathbf{x}+\mathbf{f}=\mathbf{y}]\right)\\
&= -\log_2\left(\sum_{\mathbf{y}\in Y} \max_{\mathbf{x}^\star\in X} \mathbb{P}_{\mathbf{x},\mathbf{f}}[\mathbf{x}=\mathbf{x}^\star \wedge \mathbf{M}\mathbf{x}+\mathbf{f}=\mathbf{y}]\right)\\
&= -\log_2\left(\sum_{\mathbf{y}\in Y} \max_{\mathbf{x}^\star\in X} \mathbb{P}_{\mathbf{x},\mathbf{f}}[\mathbf{M}\mathbf{x}+\mathbf{f}=\mathbf{y} \mid \mathbf{x}=\mathbf{x}^\star] \cdot \mathbb{P}_{\mathbf{x}}[\mathbf{x}=\mathbf{x}^\star]\right)\\
&\geq H_\infty(\mathbf{x}) - \log_2\left(\sum_{\mathbf{y}\in Y} \max_{\mathbf{x}^\star\in X} \mathbb{P}_{\mathbf{f}}[\mathbf{M}\mathbf{x}^\star+\mathbf{f}=\mathbf{y}]\right)\\
&= H_\infty(\mathbf{x}) - \log_2\left(\sum_{\mathbf{y}\in Y} \max_{\widehat{\mathbf{x}}\in\mathbf{M}X} \mathbb{P}_{\mathbf{f}}[\mathbf{f}=\mathbf{y}-\widehat{\mathbf{x}}]\right).
\end{aligned}$$

The first equalities follow by definition of the conditional entropy and conditional probabilities. The inequality stems from the fact that $\mathbb{P}_{\mathbf{x}}[\mathbf{x}=\mathbf{x}^\star] \leq 2^{-H_\infty(\mathbf{x})}$ by definition of the min-entropy. The equality then holds when $\mathbf{x}$ follows when $\mathbf{x}$ follows a uniform distribution. Finally, the last equality uses the fact that for all $\mathbf{y} \in Y$, $\max_{\mathbf{x}^\star\in X} \mathbb{P}_{\mathbf{f}}[\mathbf{f}=\mathbf{y}-\mathbf{M}\mathbf{x}^\star] = \max_{\widehat{\mathbf{x}}\in\mathbf{M}X} \mathbb{P}_{\mathbf{f}}[\mathbf{f}=\mathbf{y}-\widehat{\mathbf{x}}]$. □

We then bound the sum using the following lemma, which adapts [BD20a, Lem. 5.4] to a discrete Gaussian mask $\mathbf{f}$. Note that moving to the discrete Gaussian case requires the Gaussian parameter to be above the smoothing parameter of the integer lattice. Combining this bound with Lemma 3.1 then leads to Lemma 3.3 below. We recover [BD20a, Lem. 5.4] by setting $\mathbf{M} = \mathbf{I}$.

**Lemma 3.2 ([BD20a, Lem. 5.4] adapted).** *Let $n \in \mathbb{N}^\times$, and $\sigma \geq \eta_\varepsilon(\mathbb{Z}^n)$ for some $\varepsilon > 0$. Let $\widehat{X} \subseteq \mathbb{Z}^n$ be a bounded set, and define $B = \max_{\widehat{\mathbf{x}}\in\widehat{X}} \|\widehat{\mathbf{x}}\|_2$. Let $\mathbf{f} \sim \mathcal{D}_{\mathbb{Z}^n,\sigma}$. Then, it holds that*

$$\sum_{\mathbf{y}\in\mathbb{Z}^n} \max_{\widehat{\mathbf{x}}\in\widehat{X}} \mathbb{P}_{\mathbf{f}\sim\mathcal{D}_{\mathbb{Z}^n,\sigma}}[\mathbf{f}=\mathbf{y}-\widehat{\mathbf{x}}] \leq (1+\varepsilon)e^{\sqrt{2\pi n}B/\sigma}.$$

*Proof.* We here follow the same proof as for [BD20a, Lem. 5.4] which we adapt to discrete Gaussians over $\mathbb{Z}^n$ instead of $q$-periodic continuous Gaussians. Fix some $\sigma' > \sigma$. We have the following inequalities.

$$\begin{aligned}\sum_{\mathbf{y}\in\mathbb{Z}^n} \max_{\widehat{\mathbf{x}}\in\widehat{X}} \mathbb{P}_{\mathbf{f}\sim\mathcal{D}_{\mathbb{Z}^n,\sigma}}[\mathbf{f} = \mathbf{y} - \widehat{\mathbf{x}}] &= \frac{1}{\rho_\sigma(\mathbb{Z}^n)} \sum_{\mathbf{y}\in\mathbb{Z}^n} \max_{\widehat{\mathbf{x}}\in\widehat{X}} \rho_\sigma(\mathbf{y} - \mathbf{e}^*) \\ &\leq \frac{1}{\rho_\sigma(\mathbb{Z}^n)} \sum_{\mathbf{y}\in\mathbb{Z}^n} \max_{\widehat{\mathbf{x}}\in\widehat{X}} e^{\pi \frac{\|\widehat{\mathbf{x}}\|_2^2}{\sigma'^2-\sigma^2}} \rho_{\sigma'}(\mathbf{y}) \\ &\leq e^{\pi \frac{B^2}{\sigma'^2-\sigma^2}} \frac{\rho_{\sigma'}(\mathbb{Z}^n)}{\rho_\sigma(\mathbb{Z}^n)} \\ &= e^{\pi \frac{B^2}{\sigma'^2-\sigma^2}} \left(\frac{\sigma'}{\sigma}\right)^n \frac{\rho_{1/\sigma'}(\mathbb{Z}^n)}{\rho_{1/\sigma}(\mathbb{Z}^n)}.\end{aligned}$$

The first equality holds due to the fact that $\mathbf{y} - \widehat{\mathbf{x}} \in \mathbb{Z}^n$. The first inequality follows by a routine calculation (see the proof of [BD20a, Lem. 2.5]) and the second inequality uses the fact that $\widehat{\mathbf{x}}$ is always bounded by $B$. Finally, the last equality stems from the Poisson summation formula. Then, because $\sigma' > \sigma \geq \eta_\varepsilon(\mathbb{Z}^n)$, it holds that the Gaussian mass ratio is bounded above by $1 + \varepsilon$. Setting $\sigma' = \sigma\sqrt{1+\eta}$ for a free variable $\eta > 0$ gives

$$\sum_{\mathbf{y}\in\mathbb{Z}^n} \max_{\widehat{\mathbf{x}}\in\widehat{X}} \mathbb{P}_{\mathbf{f}\sim\mathcal{D}_{\mathbb{Z}^n,\sigma}}[\mathbf{f} = \mathbf{y} - \widehat{\mathbf{x}}] \leq (1+\varepsilon) e^{\pi \frac{B^2}{\eta\sigma^2}} (1+\eta)^{n/2} \leq (1+\varepsilon) e^{\frac{\pi B^2}{\eta\sigma^2} + \frac{n\eta}{2}}.$$

Minimizing the expression over $\eta$ by choosing $\eta = \sqrt{2\pi/n}B/\sigma$ then gives the bound of $(1+\varepsilon)e^{\sqrt{2\pi n}B/\sigma}$ as claimed. □

**Lemma 3.3 (Linear Noise Lossiness).** *Let $k, n \in \mathbb{N}^\times$ and $\mathbf{M} \in \mathbb{Z}^{n\times k}$ be a matrix. Let $\sigma \geq \eta_\varepsilon(\mathbb{Z}^n)$ for some $\varepsilon > 0$. Let $\mathbf{x}$ be a random variable over $\mathbb{Z}^k$ such that $\|\mathbf{Mx}\|_2 \leq B$ is always verified. Let $\mathbf{f}$ follow $\mathcal{D}_{\mathbb{Z}^n,\sigma}$. It holds that*

$$\widetilde{H_\infty}(\mathbf{x}|\mathbf{Mx} + \mathbf{f}) \geq H_\infty(\mathbf{x}) - \sqrt{2\pi n}\frac{B}{\sigma}\log_2 e - \log_2(1+\varepsilon).$$

## 4 M-LWE With General Error Distributions

In this section, we prove the hardness of search M-LWE for *general error distributions* under certain constraints we clarify later on. The high-level proof strategy, albeit generalized to our broader setting, follows the blueprint of [MP13, BJRW23]. It goes as follows: proving hardness of the search M-LWE problem with general bounded error distribution comes down to proving the one-wayness of the corresponding M-LWE function family (Definition 2.3). Instead of doing so directly, we first prove the one-wayness of the related M-Knap function family (Definition 2.3) and then invoke the duality of M-Knap and M-LWE (Lemma 2.8). By Lemma 2.6, one-wayness of M-Knap is implied by its uninvertibility (Section 4.1) and its second preimage resistance (Section 4.2).

Essentially, our results only require the error distribution to have sufficient min-entropy and to output elements of bounded norm. Previous works on the hardness of M-LWE for *general secret distributions*, e.g., [BD20b, BJRW22], did not inherently require any norm bound, which then might be interpreted as an unnecessary requirement in our case. However, because we focus on *error distributions*, this bound is intuitively crucial to ensure hardness. For example, if one considers the pathological case where all the entropy lies in too few coefficients, many samples would be noiseless which may be sufficient to solve this particular M-LWE instance. However, note that in that case, the coefficients containing all the entropy would be extremely big. The latter would then require very large parameters so that our result can apply and ensure the hardness of such an edge M-LWE instance. In particular, it would require an extremely large lattice dimension (or secret dimension) which would essentially compensate for the noiseless equations. Although our result still covers these pathological cases, it does so with very large parameters. For more reasonable cases, one can interpret this norm bound as a safeguard to spread the entropy more evenly and,

in a sense, smooth out these extreme edge cases. Typical applications of M-LWE naturally come attached to such norm constraints, e.g., in order to ensure correctness of an encryption scheme. As such, this is a reasonable assumption to make on the error distribution in addition to its min-entropy.

### 4.1 Uninvertibility

We start by proving the uninvertibility of the M-Knap function family. Actually, we first show it in Lemma 4.1 for a slightly different function family. Then, in a second step, the uninvertiblity of the M-Knap family function is proven in Lemma 4.2 by assuming the hardness of decision M-LWE (with a "standard" error distribution). The slightly different function family is obtained by composing a smaller-dimensional M-Knap with a family $\mathcal{G}$ of linear functions over $R$. Previous works [MP13, BJRW23] defined this linear function family with a Gaussian distribution in order to connect the result to worst-case to average-case reduction for M-LWE. The distribution $\mathcal{D}$ used to define $\mathcal{G}$ essentially corresponds to the error distribution for which one assumes M-LWE is hard. To give more modularity to our result, we do not fix $\mathcal{D}$ to be a Gaussian so that one can choose the starting M-LWE assumption they deem reasonable. For example, one could reasonably assume M-LWE with bounded uniform noise is hard, which has been well studied from a theoretical and cryptanalytic perspective and used in countless cryptographic applications. The final result would then show that M-LWE with general noise distributions is at least as hard as M-LWE with a bounded uniform noise. We thus start by defining the function family $\mathcal{G}$.

**Definition 4.1.** *Let $n, \ell, m$ be in $\mathbb{N}^\times$ such that $m \geq \ell$. Let $R$ be the ring of integers of a number field of degree $n$. Let $\mathcal{D}$ be a distribution on $R^\ell$, and let $X \subseteq R^m$. We define the function family $\mathcal{G}(n, \ell, m, \mathcal{D}, X)$ obtained by sampling the columns of a matrix $\mathbf{Y} \in R^{\ell \times (m-\ell)}$ from $\mathcal{D}$ and outputting $h_{\mathbf{Y}} : X \to R^\ell$ defined by $\forall \mathbf{x} \in X$, $h_{\mathbf{Y}}(\mathbf{x}) = [\mathbf{I}_\ell | \mathbf{Y}]\mathbf{x}$.*

We now show that under some norm requirements on $\mathcal{D}$ and the input distribution $\mathcal{X}$ of $\mathcal{G}$, we can show that $(\mathcal{G}(n, \ell, m, \mathcal{D}, X), \mathcal{X})$ is statistically uninvertible, provided $\mathcal{X}$ has sufficient entropy. By composition, Lemma 2.7 gives uninvertibility of M-Knap $\circ\, \mathcal{G}$ which is relevant to later show one-wayness of M-Knap.

**Lemma 4.1.** *Let $n, \ell, m, d$ be in $\mathbb{N}^\times$ such that $m \geq \max(d, \ell)$. Let $R$ be the ring of integers of a number field of degree $n$. Let $\mathcal{D}$ be a distribution on $R^\ell$, and, for any $k \in \mathbb{N}^\times$, let $\mathcal{D}^k$ be the distribution over $R^{\ell \times k}$ where all the columns are sampled from $\mathcal{D}$. We also assume that for any[5] $\alpha, \beta \in [1, \infty]$, there exists $B_{\alpha,\beta} \geq 0$ and $p_{\alpha,\beta} \in [0, 1]$ such that*

$$\mathbb{P}_{\mathbf{Y} \sim \mathcal{D}^{m-\ell}}[\|M_\tau([\mathbf{I}_\ell | \mathbf{Y}])\|_{\alpha,\beta} > B_{\alpha,\beta}] \leq p_{\alpha,\beta}.$$

*Then, let $X \subseteq R^m$ be a bounded set, and denote by $B_\alpha = \max_{\mathbf{x} \in X} \|\mathbf{x}\|_\alpha$. Let $\mathcal{X}$ be a distribution such that $\mathrm{Supp}(\mathcal{X}) = X$. We define the function family $\mathcal{F}' = \text{M-Knap}(n, m - d, \ell, q, R^\ell) \circ \mathcal{G}(n, \ell, m, \mathcal{D}, X)$. Then, $(\mathcal{F}', \mathcal{X})$ is (statistically) $\varepsilon'_{inv}$-uninvertible for*

$$\varepsilon'_{inv} = 2^{-H_\infty(\mathcal{X})} \left( \left| \mathcal{B}_{\beta, n\ell}(B_{\alpha,\beta} B_\alpha) \cap \mathbb{Z}^{n\ell} \right| + |X| \cdot p_{\alpha,\beta} \right).$$

*Proof.* We first bound $\mathbb{E}_{h_{\mathbf{Y}} \sim \mathcal{G}}[|h_{\mathbf{Y}}(X)|]$ and use Lemma 2.5 to conclude. Let $h_{\mathbf{Y}}$ be sampled from $\mathcal{G}(n, \ell, m, \mathcal{D}, X)$. Let $\mathbf{x} \in X$. Then, $h_{\mathbf{Y}}(\mathbf{x}) = [\mathbf{I}_\ell | \mathbf{Y}]\mathbf{x}$. We then have $\|h_{\mathbf{Y}}(\mathbf{x})\|_\beta \leq \|[\mathbf{I}_\ell | \mathbf{Y}]\|_{\alpha,\beta} \|\mathbf{x}\|_\alpha$. Because $X$ is bounded and all norms are equivalent on a finitely dimensional vector space, it holds that the bound $B_\alpha$ exists and is finite for any $\alpha$. By assumption, it then holds that $\|\mathbf{x}\|_\alpha \leq B_\alpha$, and

$$\mathbb{P}_{\mathbf{Y} \sim \mathcal{D}^{m-\ell}} \left[ \|M_\tau([\mathbf{I}_\ell | \mathbf{Y}])\|_{\alpha,\beta} > B_{\alpha,\beta} \right] \leq p_{\alpha,\beta}.$$

Hence, with probability at least $1 - p_{\alpha,\beta}$, we have that the $\ell_\beta$ norm of $\tau(h_{\mathbf{Y}}(\mathbf{x}))$ is bounded by $B_{\alpha,\beta} B_\alpha$. The number of integer points in the $n\ell$-dimensional $\ell_\beta$-ball then gives an upper bound on $|h_{\mathbf{Y}}(X)|$. It means we have

$$|h_{\mathbf{Y}}(X)| \leq \left| \mathcal{B}_{\beta, n\ell}(B_{\alpha,\beta} B_\alpha) \cap \mathbb{Z}^{n\ell} \right| =: K.$$

[5] Recall that all norms are equivalent in finite-dimensional vector spaces. Hence, if there exists $B_{\alpha_0,\beta_0}, p_{\alpha_0,\beta_0}$ for one pair $(\alpha_0, \beta_0)$, it gives the existence for all pairs $(\alpha, \beta)$.

We note that if the radius $B_{\alpha,\beta}B_\alpha$ is large enough compared to $n\ell$, the value of $K$ is well approximated by the volume of the $\ell_\beta$-ball, which is

$$\mathrm{Vol}(\mathcal{B}_{\beta,n\ell}(B_{\alpha,\beta}B_\alpha)) = \frac{\left(2B_{\alpha,\beta}B_\alpha \cdot \Gamma\left(\frac{1}{\beta}+1\right)\right)^{n\ell}}{\Gamma\left(\frac{n\ell}{\beta}+1\right)}.$$

We note that we could have trivially chosen $K = |X|$ but this is not tight because $h_{\mathbf{Y}}$ is not injective. We also define $S = \{\mathbf{Y} \in R^{\ell\times(m-\ell)} : \|M_\tau([\mathbf{I}_\ell|\mathbf{Y}])\|_{\alpha,\beta} \le B_{\alpha,\beta}\}$, and $S'$ its complement in $R^{\ell\times(m-\ell)}$. We then have

$$\begin{aligned}\mathbb{E}\left[|h_{\mathbf{Y}}(X)|\right] &= \sum_{\mathbf{Y}'\in S} \mathbb{P}_{\mathbf{Y}}[\mathbf{Y}=\mathbf{Y}']|h_{\mathbf{Y}'}(X)| + \sum_{\mathbf{Y}'\in S'} \mathbb{P}_{\mathbf{Y}}[\mathbf{Y}=\mathbf{Y}']|h_{\mathbf{Y}'}(X)| \\ &\le K \cdot \mathbb{P}_{\mathbf{Y}}[\mathbf{Y}\in S] + |X| \cdot \mathbb{P}_{\mathbf{Y}}[\mathbf{Y}\in S'] \\ &\le K + |X| \cdot p_{\alpha,\beta},\end{aligned}$$

where the first inequality follows from the above calculations and the fact that for $\mathbf{Y}' \in S'$, we have the trivial bound $|h_{\mathbf{Y}'}(X)| \le |X|$. Lemma 2.5 then yields the $\varepsilon'_{\mathrm{inv}}$-uninvertibility of $\mathcal{G}$, with $\varepsilon'_{\mathrm{inv}} = 2^{-H_\infty(\mathcal{X})}(K + |X| \cdot p_{\alpha,\beta})$. By Lemma 2.7, we thus obtain the $\varepsilon'_{\mathrm{inv}}$-uninvertibility of $\mathcal{F}'$.

□

*Remark 4.1.* The quantity $2^{-H_\infty(\mathcal{X})}|\mathrm{Supp}(\mathcal{X})|$ is always larger than 1, as we have $|\mathrm{Supp}(\mathcal{X})| = 2^{H_\infty(U(X))}$. Hence $2^{-H_\infty(\mathcal{X})}|\mathrm{Supp}(\mathcal{X})| = 2^{H_\infty(U(X))-H_\infty(\mathcal{X})}$. We then use the fact that maximal entropy is achieved for the uniform distribution for a fixed support. As a result, one needs not only to set the bound $B_{\alpha,\beta}$ so that $p_{\alpha,\beta}$ is negligible (e.g., $\le 2^{-\lambda}$), but also to compensate the entropy difference between $\mathcal{X}$ and $U(X)$.

Our result of Lemma 4.1 is naturally parameterized by $(\alpha,\beta)$ so as to remain as generic as possible, and cover more distributions $\mathcal{X}$. Certain distributions may be more relevant to analyze in the $\alpha = \infty$ metric than $\alpha = 2$ for example. One could even optimize over $(\alpha,\beta)$ provided that they have an efficient way of computing or approximating $B_{\alpha,\beta}, B_\alpha, p_{\alpha,\beta}$ and the number of integer points in the $\ell_\beta$ ball of a given radius. Indeed, a stronger statement would be

$$\varepsilon'_{\mathrm{inv}} \le 2^{-H_\infty(\mathcal{X})} \cdot \inf_{\alpha,\beta}\left(\left|\mathcal{B}_{\beta,n\ell}(B_{\alpha,\beta}B_\alpha) \cap \mathbb{Z}^{n\ell}\right| + |X| \cdot p_{\alpha,\beta}\right).$$

As estimating all these quantities is non trivial for arbitrary values of $(\alpha,\beta)$, we only give in Appendix C a few example parameterizations for $\alpha,\beta$ (and $\mathcal{D}$), as well as full examples in $(1,2)$-norm in Appendix A and B. We typically restrict our examples to $\beta \in \{1,2,\infty\}$ for which we know closed-form or close approximations of the number of integer points in the $\ell_\beta$ ball[6]. We then restrict to the values of $\alpha$ for which we have an expression of $\|\cdot\|_{\alpha,\beta}$. But we again insist that our result can be parameterized differently if needed.

So far, we have proven the uninvertibility of a composition between M-Knap with rank $m-d$ and dimension $\ell$ and $\mathcal{G}$. However, to eventually prove hardness of M-LWE with rank $d$ and $m$ samples, we would need the uninvertibility of M-Knap with rank $m-d$ and dimension $m > \ell$. We now show that we can obtain the latter by combining Lemma 4.1 with an indistinguishability argument based on M-LWE with error distribution $\mathcal{D}$.

**Lemma 4.2.** *Let $n,k,q,d,m$ be in $\mathbb{N}^\times$ with $m > d \ge k \ge 1$ and let $\ell = m-d+k$. Let $\mathcal{D}$ be a distribution over $R^\ell$, and $X \subseteq R^m$. Assume* M-LWE$_{n,k,\ell,q,U(R_q^d),\mathcal{D}}$ *is $\varepsilon_{\text{M-LWE}}$-hard. Define the function families $\mathcal{F}' =$* M-Knap$(n,m-d,\ell,q,R^\ell)\circ\mathcal{G}(n,\ell,m,\mathcal{D},X)$ *(as in Lemma 4.1), and $\mathcal{F} =$* M-Knap$(n,m-d,m,q,X)$*. If $\mathcal{F}'$ is $\varepsilon'_{inv}$-uninvertible, then $\mathcal{F}$ is $\varepsilon_{inv}$-uninvertible, where*

$$\varepsilon_{inv} = \varepsilon'_{inv} + (m-\ell)\left(2\delta(\ell,ell-k) + \varepsilon_{\text{M-LWE}}/(1-\delta(\ell,k))\right).$$

[6] We have $\left|\mathcal{B}_{1,N}(r)\cap\mathbb{Z}^N\right| = \sum_{i=0}^{\min(N,\lfloor r\rfloor)} 2^i\binom{N}{i}\binom{\lfloor r\rfloor}{i}$, $\left|\mathcal{B}_{\infty,N}(r)\cap\mathbb{Z}^N\right| = (2\lfloor r\rfloor+1)^N$, and $\left|\mathcal{B}_{2,N}(r)\cap\mathbb{Z}^N\right| \le \mathrm{Vol}(\mathcal{B}_{2,N}(r+\sqrt{N}/2))$. When $r \gg N$, we also have $\left|\mathcal{B}_{p,N}(r)\cap\mathbb{Z}^N\right| \approx \mathrm{Vol}(\mathcal{B}_{p,N}(r))$.

*Proof.* First, assuming $\text{M-LWE}_{n,k,\ell,q,U(R_q^d),\mathcal{D}}$ is $\varepsilon_{\text{M-LWE}}$-hard entails that the M-LWE function family is $\varepsilon_{\text{M-LWE}}$-pseudorandom. Based on the duality result Lemma 2.8, we get that $(\text{M-Knap}(n, m-d, \ell, q, R^\ell), \mathcal{D})$ is $\varepsilon_{\text{rand}}$-pseudorandom with $\varepsilon_{\text{rand}} = 2\delta(\ell, m-d) + \varepsilon_{\text{M-LWE}}/(1-\delta(\ell,k))$ because $m-d = \ell - k$.

Then, it holds that $\mathcal{F}$ and $\mathcal{F}'$ are indistinguishable based on $\varepsilon_{\text{rand}}$. Indeed, assume an adversary $\mathcal{A}$ breaks the $(m-\ell)\varepsilon_{\text{rand}}$-indistinguishability between $\mathcal{F}$ and $\mathcal{F}'$. Take $f_{\mathbf{A}'} \circ h_{\mathbf{Y}}$ according to $\mathcal{F}'$, and $f_{\mathbf{A}}$ according to $\mathcal{F}$. Then $f_{\mathbf{A}'} \circ h_{\mathbf{Y}}$ is the linear map $\mathbf{x} \mapsto [\mathbf{A}'^T | \mathbf{A}'^T\mathbf{Y}]\mathbf{x}$. Decomposing $\mathbf{A}^T$ into $[\mathbf{A}_1^T | \mathbf{A}_2^T]$, 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 $f_{\mathbf{A}} = \mathbf{x} \mapsto [\mathbf{A}_1^T | \mathbf{A}_2^T]\mathbf{x}$. By assumption, it means that $\mathcal{A}$ can distinguish $(\mathbf{A}'^T, \mathbf{A}'^T\mathbf{Y} \bmod qR)$ from uniform. By a hybrid argument, it means $\mathcal{A}$ can distinguish $(\mathbf{A}'^T, \mathbf{A}'^T\mathbf{y} \bmod qR)$ for $\mathbf{y} \sim \mathcal{D}$ from uniform with advantage at least $\varepsilon_{\text{rand}}$, thus breaking the $\varepsilon_{\text{rand}}$-pseudorandomness of $(\text{M-Knap}(n, m-d, \ell, q, R^\ell), \mathcal{D})$. By assumption, we then have a contradiction, which means that $\mathcal{F}$ and $\mathcal{F}'$ are $(m-\ell)\varepsilon_{\text{rand}}$-indistinguishable.

By indistinguishability, the properties of $\mathcal{F}$ and $\mathcal{F}'$ transfer to one another with an additive loss of $(m-\ell)\varepsilon_{\text{rand}}$. As such, applying Lemma 4.1 yields that $(\mathcal{F}, \mathcal{X})$ is $(\varepsilon'_{\text{inv}} + (m-\ell)\varepsilon_{\text{rand}})$-uninvertible as claimed. □

### 4.2 Second Preimage Resistance

We now prove the second preimage resistance of the M-Knap function family with respect to an arbitrary distribution. We provide the proof in both statistical and computational hardness regimes to allow for more flexibility depending on the final application.

**Lemma 4.3.** *Let $n, h, q, m$ be in $\mathbb{N}^\times$ such that $q$ is prime. Let $R$ be the ring of integers of a number field of degree $n$, and $X \subseteq R^m$. We also define $\mathcal{N} = \max_{\mathbf{x},\mathbf{x}'\in X, \mathbf{x}\neq\mathbf{x}'} N(\langle x_1 - x'_1, \ldots, x_m - x'_m, q\rangle)$. Let $\mathcal{X}$ be a distribution such that $\text{Supp}(\mathcal{X}) = X$. Then $(\text{M-Knap}(n, h, m, q, X), \mathcal{X})$ is $\varepsilon_{spr}$-second preimage resistant with respect to unbounded adversaries for*

$$\varepsilon_{spr} = \frac{\mathcal{N}^h}{q^{nh}}(|X| - 1).$$

*If $X$ is a bounded set and $\mathcal{X}$ is efficiently sampleable, $(\text{M-Knap}(n, h, m, q, X), \mathcal{X})$ is also $\varepsilon_{\text{M-SIS}}$-second preimage resistant with respect to PPT adversaries provided $\text{M-SIS}^2_{n,h,m,q,\beta_2}$ is $\varepsilon_{\text{M-SIS}}$-hard for $\beta_2 = \max\{\|\mathbf{x} - \mathbf{x}'\|_2; (\mathbf{x}, \mathbf{x}') \in X^2\}$. The latter statement generalizes to $\text{M-SIS}^p$ by changing the definition of $\beta_2$ accordingly.*

*Proof.* Observing that for any (unbounded) adversaries $\mathcal{A}$, we have

$$\text{Adv}_{\text{spr}}[\mathcal{A}] \leq \mathbb{P}_{\substack{\mathbf{A}\hookleftarrow U(R_q^{m\times h}) \\ \mathbf{x}\hookleftarrow\mathcal{X}}}[\exists \mathbf{x}'' \in X \setminus \{\mathbf{x}\},\ \mathbf{A}^T\mathbf{x}'' = \mathbf{A}^T\mathbf{x} \bmod qR] =: P,$$

by simple inclusion of event and where $\text{Adv}_{\text{spr}}[\mathcal{A}]$ is the advantage of $\mathcal{A}$ in breaking the second-preimage resistance. As $P$ no longer depends on the adversary's behavior or power, we can say that $(\text{M-Knap}(n, h, m, q, X), \mathcal{X})$ is $P$-second preimage resistant with respect to unbounded adversaries. We now need to estimate or bound $P$. We then show that for $\mathbf{A}$ uniformly chosen and $\mathbf{x}$ drawn from $\mathcal{X}$, 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 $\varepsilon_{\text{spr}}$. Using the total probability formula and the union bound on $\mathbf{x}''$, we have the following.

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

Let $\mathbf{x}' \in X$, $\mathbf{x}'' \in X \setminus \{\mathbf{x}'\}$. Then, by [Mic04, Lem. 4.4], $\mathbf{A}^T(\mathbf{x}'' - \mathbf{x}') \bmod qR$ is uniformly distributed in $(\mathcal{I}_{\mathbf{x}''-\mathbf{x}'}/qR)^h$ over the randomness of $\mathbf{A}$, where $\mathcal{I}_{\mathbf{x}''-\mathbf{x}'} = \langle x''_1 - x'_1, \ldots, x''_m - x'_m, q\rangle$. Hence the probability that

$\mathbf{A}^T(\mathbf{x}''-\mathbf{x}') = \mathbf{0} \bmod qR$ is exactly $|\mathcal{I}_{\mathbf{x}''-\mathbf{x}'}/qR|^{-h}$. As $\mathcal{I}_{\mathbf{x}''-\mathbf{x}'}$ and $qR$ are ideals of $R$, we have $|\mathcal{I}_{\mathbf{x}''-\mathbf{x}'}/qR| = N(qR)/N(\mathcal{I}_{\mathbf{x}''-\mathbf{x}'}) = q^n/N(\mathcal{I}_{\mathbf{x}''-\mathbf{x}'})$. Hence, we have

$$P \leq \sum_{\mathbf{x}'\in X} \mathbb{P}_{\mathbf{x}}[\mathbf{x}=\mathbf{x}'] \sum_{\mathbf{x}''\in X\setminus\{\mathbf{x}'\}} \frac{N(\mathcal{I}_{\mathbf{x}''-\mathbf{x}'})^h}{q^{nh}} \leq \sum_{\mathbf{x}'\in X} \mathbb{P}_{\mathbf{x}}[\mathbf{x}=\mathbf{x}'](|X|-1)\frac{\mathcal{N}^h}{q^{nh}}$$
$$= (|X|-1)\frac{\mathcal{N}^h}{q^{nh}} = \varepsilon_{\mathrm{spr}},$$

where the last inequality comes from the definition of $\mathcal{N}$, then giving the first statement.

For the second statement, we simply note that second-preimage resistance is implied by collision-resistance for an efficiently sampleable distribution $\mathcal{X}$. Yet the collision-resistance of M-Knap$(n,h,m,q,X)$ is implied by M-SIS$^p_{n,h,m,q,\beta_p}$ in $\ell_p$ norm for any $p \geq 1$, and where $\beta_p = \max\{\|\mathbf{x}-\mathbf{x}'\|_p; (\mathbf{x},\mathbf{x}') \in X^2\}$. This is obtained directly from [Ajt96], but we sketch it here for completeness. Assume an adversary $\mathcal{A}$ breaks the collision-resistance of M-Knap$(n,h,m,q,X)$ with advantage $\varepsilon$. The reduction is given an instance $\mathbf{A} \hookleftarrow U(R_q^{m\times h})$ of M-SIS$^p_{n,h,m,q,\beta_p}$. It sends $\mathbf{A}$ to $\mathcal{A}$ as the description of the function, who then returns $(\mathbf{x},\mathbf{x}') \in X^2$ such that $\mathbf{A}^T\mathbf{x} = \mathbf{A}^T\mathbf{x}' \bmod qR$ and $\mathbf{x} \neq \mathbf{x}'$. Then $\mathbf{A}^T(\mathbf{x}-\mathbf{x}') = \mathbf{0} \bmod qR$ and $0 < \|\mathbf{x}-\mathbf{x}'\|_p \leq \beta_p$ by definition of $\beta_p$, thus solving the M-SIS instance. The reduction thus has advantage $\varepsilon$ in solving M-SIS$^p_{n,h,m,q,\beta_p}$. Considering $p=2$ gives the theorem statement, but we insist it is more general. In particular, if $\varepsilon_{\text{M-SIS},p}$ denotes the hardness bound of M-SIS$^p_{n,h,m,q,\beta_p}$, then $\varepsilon_{\mathrm{spr}} \leq \min_p \varepsilon_{\text{M-SIS},p}$. □

We observe that the statistical bound is most likely not tight. First, the union bound on $\mathbf{x}''$ may be too loose, yielding a daunting factor $|X|-1$. Second, in many cases, we can ensure $\mathcal{I}_{\mathbf{z}} = R$ (which has norm 1) for all $\mathbf{z}$ by placing minor restrictions on $\mathcal{X}$ (or rather $X$) and $q$. A sufficient condition is that at least one of the $x_i - x'_i$ is a unit in $R_q$. Such a condition can be obtained in cyclotomic rings from Lemma 2.1. For example, if $q$ splits in $\kappa$ factor in a power-of-two cyclotomic ring, having

$$\max_{\mathbf{x},\mathbf{x}'\in X^2} \min_{\substack{i\in[m]\\ \text{st. } x_i\neq x'_i}} \|x_i - x'_i\|_p < \kappa^{\min(0,1/p-1/2)} q^{1/\kappa},$$

is enough to ensure $\mathcal{N} = 1$. Indeed, let $\mathbf{x} \neq \mathbf{x}'$ be in $X$, then $0 < \min_i \|x_i - x'_i\|_p < \kappa^{\min(0,1/p-1/2)} q^{1/\kappa}$. Lemma 2.2 then ensures that at least one difference $x_i - x'_i$ is invertible in $R_q$. Hence $\langle x_i - x'_i, q\rangle = R$ and as such $\mathcal{I}_{\mathbf{x}-\mathbf{x}'} = R$ which has norm 1. In such situations, we can thus have $\varepsilon_{\mathrm{spr}} = (|X|-1)/q^{nh}$. This would require limiting the splitting of $q$, but it was shown in [CHK+21] it is not so damaging in practical cryptographic applications anyway. We note that the bound $\varepsilon_{\mathrm{spr}}$ does not directly depend on the distribution $\mathcal{X}$ nor $H_\infty(\mathcal{X})$. This is because our proof method (union bound and maximize the norm of $\mathcal{I}_{\mathbf{z}}$) is closer in essence to the collision resistance (which does not depend on the input distribution) than it is to the second-preimage resistance. A finer analysis may lead to a quantity that depends on the entropy though.

The computational bound requires to assume the hardness of M-SIS, which, at first glance, may seem like a circular argument when trying to prove the second-preimage resistance of the M-Knap function family. The main difference lies in the fact that M-SIS, which relates to the collision resistance, is agnostic to the input distribution $\mathcal{X}$, provided the latter is bounded. Although we do not want to assume specific properties of $\mathcal{X}$ to keep our result general, the hypothesis of it being bounded is reasonable in most use cases and applications.

### 4.3 One-Wayness of the M-LWE Function

Combining Lemma 2.6 to Lemmas 4.2 and 4.3 gives the one-wayness of the M-Knap function family with input distribution $\mathcal{X}$. Combined with the duality results stated in Lemma 2.8, one obtains a reduction from (decision) M-LWE with noise distribution $\mathcal{D}$ to (search) M-LWE with noise distribution $\mathcal{X}$.

**Theorem 4.1.** *Let $\lambda$ be a security parameter, and $n,d,m,k,q$ be in $\mathbb{N}^\times$ such that $m > d \geq k \geq 1$ and let $\ell = m-d+k$. Let $R$ be the ring of integers of a number field of degree $n$. Let $\mathcal{D}$ be a distribution over $R^\ell$*

*such that for* $\alpha, \beta \in [1, \infty]$ *there exists* $B_{\alpha,\beta} \geq 0$ *and* $p_{\alpha,\beta} \in [0,1]$ *such that* $\mathbb{P}_{\mathbf{Y} \sim \mathcal{D}^{m-\ell}}[\|M_\tau([\mathbf{I}_\ell | \mathbf{Y}])\|_{\alpha,\beta} > B_{\alpha,\beta}] \leq p_{\alpha,\beta}$. *Then, let* $X \subseteq R^m$ *a bounded set, and* $\mathcal{X}$ *an efficiently sampleable distribution with support* $X$. *Denote by* $B_\alpha = \max_{\mathbf{x} \in X} \|\mathbf{x}\|_\alpha$, *and* $\beta_p = \max\{\|\mathbf{x} - \mathbf{x}'\|_p ; (\mathbf{x}, \mathbf{x}') \in X^2\}$ *for some* $p \in \mathbb{N}^\times$. *We further assume that* $R$ *is unramified above the prime* $q$, *and such that* $\min_{i \in [\kappa]} N(\mathfrak{p}_i)^{\min(m-d,k)+1} \geq \lambda^{\omega(1)}$, *where the* $\mathfrak{p}_i$*'s are the prime ideal factors of* $\langle q \rangle$.
*Then, if decision* $\text{M-LWE}_{n,k,\ell,q,U(R_q^k),\mathcal{D}}$ *is* $\varepsilon_{\text{M-LWE}}$*-hard and* $\text{M-SIS}^p_{n,m-d,m,q,\beta_p}$ *is* $\varepsilon_{\text{M-SIS}}$*-hard, it holds that search* $\text{M-LWE}_{n,d,m,q,U(R_q^d),\mathcal{X}}$ *is* $\varepsilon$*-hard for*

$$\begin{aligned}\varepsilon &= \delta(m,d) + \frac{(m-\ell)(2\delta(\ell, m-d) + \varepsilon_{\text{M-LWE}}/(1-\delta(\ell,k))) + \varepsilon'_{inv} + \varepsilon_{\text{M-SIS}}}{1 - \delta(m, m-d)} \\ &\approx (d-k)\varepsilon_{\text{M-LWE}} + \varepsilon_{\text{M-SIS}} + 2^{-H_\infty(\mathcal{X})} \left( \left| \mathcal{B}_{\beta,n\ell}(B_{\alpha,\beta} B_\alpha) \cap \mathbb{Z}^{n\ell} \right| + |X| \cdot p_{\alpha,\beta} \right),\end{aligned}$$

*where* $\varepsilon'_{inv}$ *as in Lemma 4.1,* $\mathcal{B}_{\beta,n\ell}(r)$ *is the closed* $\ell_\beta$ *ball of radius* $r$, $\delta(a,b) = 1 - \mathbb{P}_{\mathbf{A} \sim U(R_q^{b \times a})}[\mathbf{A} R_q^a = R_q^b]$, *and* $\approx$ *smoothes out terms negligible in* $\lambda$.

The formula for $\varepsilon$ includes the term in $\varepsilon'_{\text{inv}}$ from Lemma 4.1, which may be difficult to grasp. Concretely, if $\varepsilon_{\text{M-LWE}}$ and $\varepsilon_{\text{M-SIS}}$ are negligible, having $\varepsilon$ negligible requires $\mathcal{X}$ to have sufficient entropy to overcome the factor in parenthesis. Assume for simplicity that the bound $B_{\alpha,\beta}$ is enforced[7] in $\mathcal{D}$ so that $p_{\alpha,\beta} = 0$. As mentioned in Section 4.1, if the radius $r$ is sufficiently large compared to the dimension $N$, then $\left|\mathcal{B}_{\beta,N}(r) \cap \mathbb{Z}^N\right| \approx \text{Vol}(\mathcal{B}_{\beta,N}(r)) = (2r\Gamma(\beta^{-1}+1))^N / \Gamma(N\beta^{-1}+1)$. With this approximation, the entropy requirement to make $\varepsilon$ negligible becomes

$$H_\infty(\mathcal{X}) \gtrsim \lambda + n\ell \log_2\left(2B_{\alpha,\beta} B_\alpha \Gamma(\beta^{-1}+1)\right) - \log_2 \Gamma(n\ell\beta^{-1}+1).$$

*Proof.* We define $\mathcal{F} = \text{M-Knap}(n, m-d, m, q, X)$. Lemma 4.2 in combination with Lemma 4.1 directly yields the $\varepsilon_{\text{inv}}$-uninvertibility of $(\mathcal{F}, \mathcal{X})$, where

$$\varepsilon_{\text{inv}} = \left( 2^{-H_\infty(\mathcal{X})} \left( \left| \mathcal{B}_{\beta,n\ell}(B_{\alpha,\beta} B_\alpha) \cap \mathbb{Z}^{n\ell} \right| + |X| p_{\alpha,\beta} \right) + (m-\ell) \left( 2\delta(\ell, m-d) + \frac{\varepsilon_{\text{M-LWE}}}{1-\delta(\ell,k)} \right) \right).$$

In the computational statement of Lemma 4.3, it holds that $(\mathcal{F}, \mathcal{X})$ is $\varepsilon_{\text{M-SIS}}$-second preimage resistant. Lemma 2.6 then yields that $(\mathcal{F}, \mathcal{X})$ is $\varepsilon_{\text{ow}}$-one-way with $\varepsilon_{\text{ow}} = \varepsilon_{\text{inv}} + \varepsilon_{\text{M-SIS}}$. Lemma 2.8 then proves $(\text{M-LWE}(n,d,m,q,X), \mathcal{X})$ is $\varepsilon$-one-way with $\varepsilon = \delta(m,d) + \varepsilon_{\text{ow}}/(1-\delta(m,m-d))$. As one-wayness corresponds to the search variant, we get that search $\text{M-LWE}_{n,d,m,q,U(R_q^d),\mathcal{X}}$ is $\varepsilon$-hard.

Combining the different expressions of $\varepsilon_{\text{ow}}$ and $\varepsilon_{\text{inv}}$ gives the claimed expression of $\varepsilon$ from the statement. The final approximation comes from the fact that the $\delta(\cdot,\cdot)$ terms are negligible in $\lambda$ based on the requirements on $q$. Indeed, as noted in [BJRW23, Lem. 2.6], if the smallest norm $N$ of the prime ideal factors is such that $N^{a-b+1} \geq \lambda^{\omega(1)}$ for $a \geq b$, then $\delta(a,b) \leq \lambda^{-\omega(1)}$. The condition on the $\mathfrak{p}_i$ thus yields that $\delta(m,d), \delta(m, m-d), \delta(\ell, m-d), \delta(\ell,k)$ are negligible. □

Observe that we opted for the computational second preimage resistance using the M-SIS assumption as it is well understood and likely to give better parameters. One can easily use the statistical second preimage resistance from Lemma 4.3 and obtain a similar statement by replacing $\varepsilon_{\text{M-SIS}}$ with $\varepsilon_{\text{spr}} = \mathcal{N}^{m-d}(|X| - 1)/q^{n(m-d)}$. We also insist that although we present our result over modules, it also covers the unstructured case by selecting $n = 1$ so that $R = \mathbb{Z}$. This in particular removes certain painstaking subtleties of modules, especially on the duality result of Lemma 2.8 where the $\delta(a,b)$ have a closed-form expression when $q$ is prime, and on the statistical version of Lemma 4.3 which becomes considerably simpler and tighter.

Prior works [MP13,STA20,BJRW23] deduce a somewhat explicit constraint on the number of samples $m$. Typically, they require $m = d(1+o(1))$ for very small error (e.g., binary). Reaching $m = 2d$ is possible but takes a toll on other parameters. Unfortunately, one likely cannot reach typical choices of $m$ in cryptographic

[7] This comes down to truncating $\mathcal{D}$ to verify the norm bound. Note that this may be enforced at key generation in cryptographic applications, e.g., encryption schemes.

constructions where one needs at least $O(d \log_2 q)$ samples. Not being able to prove the hardness in such regimes is however not a surprise for unusually small errors as one can possibly obtain subexponential attacks in those regimes. Our result unfortunately suffers from similar constraints, albeit less explicitly. As we do not specify the actual error distribution, the constraint on $m$ does not show up in closed form but rather as the entropy requirement mentioned above. Plugging a specific choice of $\mathcal{X}$ and computing its entropy would then transform the entropy condition into a constraint on $m$ which would also depend on other parameters of the distribution $\mathcal{X}$.

In order to show that our result is applicable to concretely interesting distributions despite its genericity, we provide example parameter choices to prove the hardness of M-LWE with a sparse ternary error in Appendix B.

## 5 Hermite Normal Form Transform

At this stage, we have proven that, if $\mathcal{X}$ has sufficient entropy, we can change the error distribution from $\mathcal{D}$ to $\mathcal{X}$ (albeit with different dimensions). But this is still for a secret distribution that is uniform over $R_q$. We thus use a Hermite Normal Form transformation reduction in order to use $\mathcal{X}$ for *both* the secret and error, i.e., $[\mathbf{s}|\mathbf{e}] \sim \mathcal{X}$. This transform is well-known [ACPS09] and quite standard, both from a theoretical standpoint and a practical cryptanalytic aspect. However, it is mostly applied to product distributions, that is $\mathcal{X} = \mathcal{Y}^m$ where all the entries are independently sampled from the same distribution $\mathcal{Y}$ over $R$. As we target a more general class of distributions, we need to account for this change in the reduction which naturally gives rise to the following variant of HNF-M-LWE which we call *permuted*-HNF-M-LWE. Concretely, it is similar to HNF-M-LWE, i.e., find $\mathbf{x} = [\mathbf{s}|\mathbf{e}] \sim \mathcal{X}'$ given $(\mathbf{A}, [\mathbf{A}|\mathbf{I}]\mathbf{x})$, but for $\mathcal{X}'$ being a random permutation of $\mathcal{X}$.

**Definition 5.1 (Permuted HNF).** *Let $n, d, m, q$ be in $\mathbb{N}^\times$. Let $R$ be the ring of integers of a number field of degree $n$, and $\mathcal{X}$ be a distribution supported on a set $X \subseteq R^{m+d}$. The search version of the permuted HNF Module Learning With Errors problem, denoted* pHNF-M-LWE$_{n,d,m,q,\mathcal{X}}$*, is as follows: Given $\mathbf{A} \hookleftarrow U(R_q^{m \times d})$, and $\mathbf{b} = [\mathbf{A} \mid \mathbf{I}_m] \cdot \mathbf{P} \cdot \mathbf{x} \bmod qR$ for some $\mathbf{x} \hookleftarrow \mathcal{X}$ and $\mathbf{P} \in \{0,1\}^{(m+d)\times(m+d)}$ a random permutation matrix, find $\mathbf{x}$. The decision variant asks to distinguish such pairs $(\mathbf{A}, \mathbf{b} = [\mathbf{A} \mid \mathbf{I}_m] \cdot \mathbf{P} \cdot \mathbf{x} \bmod qR)$ from $(\mathbf{A}, \mathbf{u})$ where $\mathbf{u} \hookleftarrow U(R_q^m)$.*

*Remark 5.1.* If $\mathcal{X} = \mathcal{Y}^{m+d}$ is a product distribution, identically and independently sampling every entry, $\mathbf{Px}$ follows exactly the same distribution as $\mathbf{x}$ and we hence recover the standard formulation of M-LWE in Hermite Normal Form [LS15].

We now show that pHNF-M-LWE with distribution $\mathcal{X}$ is no easier than M-LWE with noise drawn from $\mathcal{X}$, at the expense of reducing the number of samples. The associated reduction loss features the quantity $\delta'(m, d)$ defined as $\mathbb{P}_{(\mathbf{a}_i)_{i\in[m]} \sim U(R_q^d)^m}[\nexists S \subseteq [m], |S| = d \wedge (\mathbf{a}_i)_{i\in S} \text{ are } R_q\text{-linearly independent}]$, and introduced in [BJRW23]. Because this is a non-standard HNF formulation, we also show that the regular HNF-M-LWE version with distribution $\mathcal{X}$ is also no easier than M-LWE with noise drawn from $\mathcal{X}$, with the same sample constraints, but with a slightly looser reduction (featuring $\delta(d, d)$ defined in Theorem 4.1 instead of $\delta'(m, d)$). As the loss is still acceptable in that case, the latter reduction is still preferable in cases where the distribution $\mathcal{X}$ is not permutation-invariant so as to have a more standard HNF form. We use the notation HNF-M-LWE$_{n,d,m,q,\mathcal{X}}$ for the standard HNF variant, i.e., recovering $\mathbf{x}$ from $[\mathbf{A} \mid \mathbf{I}]\mathbf{x}$ or distinguishing the latter from uniform.

**Theorem 5.1.** *Let $n, d, m, q$ be in $\mathbb{N}^\times$, such that $m > d \geq 1$. Let $R$ be the ring of integers of a number field of degree $n$, and assume that $q$ is prime and unramified in $R$. Let $\mathscr{D}_s$ be a distribution over $R^d$ and $\mathcal{X}$ a distribution over a set $X \subseteq R^m$. There is a* PPT *reduction from* M-LWE$_{n,d,m,q,\mathscr{D}_s,\mathcal{X}}$ *to* pHNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$*. More concretely, if $\varepsilon_{\text{M-LWE}}$ and $\varepsilon_{\text{pHNF-M-LWE}}$ denote the corresponding hardness bounds, it holds that $\varepsilon_{\text{pHNF-M-LWE}} \leq (1-\delta'(m,d))^{-1}\varepsilon_{\text{M-LWE}}$. The reduction holds for both the search and decision variants. Similarly, there is a* PPT *reduction from* M-LWE$_{n,d,m,q,\mathscr{D}_s,\mathcal{X}}$ *to the regular* HNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$ *such that $\varepsilon_{\text{HNF-M-LWE}} \leq (1-\delta(d,d))^{-1}\varepsilon_{\text{M-LWE}}$.*

*Proof.* We construct a reduction for the search variant. Let $(\mathbf{A}, \mathbf{b}) \in R_q^{m\times d} \times R_q^m$ be an instance of M-LWE$_{n,d,m,q,\mathscr{D}_s,\mathcal{X}}$. The reduction first checks if there is a subset $S \subseteq [m]$ of size $d$ such that the rows of $\mathbf{A}$ indexed by $S$ are $R_q$-linearly independent. If no such subset $S$ exists, then the reduction aborts. We denote this abort probability with

$$\begin{aligned}\delta'(m,d) &= \mathbb{P}_{(\mathbf{a}_i)_{i\in[m]}\sim U(R_q^d)^m}\left[\nexists S \subseteq [m], |S| = d \wedge (\mathbf{a}_i^T)_{i\in S} \text{ are } R_q\text{-l.i.}\right] \\ &= 1 - \mathbb{P}_{(\mathbf{a}_i)_{i\in[m]}\sim U(R_q^d)^m}\left[\exists S \subseteq [m], |S| = d \wedge (\mathbf{a}_i^T)_{i\in S} \text{ are } R_q\text{-l.i.}\right].\end{aligned}$$

Assume that there exists such a subset $S \subseteq [m]$ of size $d$ such that the rows $(\mathbf{a}_i^T)_{i\in S}$ are $R_q$-linearly independent. Let $T = [m] \setminus S$. We let $\sigma_S$ be the unique permutation of $[m]$ which maps $S$ to $[d]$ and $T$ to $[d+1, m]$ while preserving the non-decreasing order, i.e., for all $i, j \in S$, $i > j \Rightarrow \sigma_S(i) > \sigma_S(j)$ and for all $i, j \in T$, $i > j \Rightarrow \sigma_S(i) > \sigma_S(j)$. We let $\mathbf{P}'$ be the permutation matrix corresponding to $\sigma_S$. Then, the reduction samples $\mathbf{P}_S$ a random permutation of $[d]$ and $\mathbf{P}_T$ a random permutation of $[d+1, m]$, and defines $\mathbf{P} = \mathrm{diag}(\mathbf{P}_S, \mathbf{P}_T)\mathbf{P}'$ which is a permutation matrix by construction. We then compute

$$\begin{bmatrix}\mathbf{A}_S \\ \mathbf{A}_T\end{bmatrix} = \mathbf{P}\mathbf{A}, \quad \text{and} \quad \begin{bmatrix}\mathbf{b}_S \\ \mathbf{b}_T\end{bmatrix} = \mathbf{P}\mathbf{b},$$

where $\mathbf{A}_S \in R_q^{d\times d}$, and $\mathbf{b}_S \in R_q^d$. Because of how we constructed $\mathbf{P}$, the rows of $\mathbf{A}_S$ are the $(\mathbf{a}_i^T)_{i\in S}$ (in permuted order by $\mathbf{P}_S$) which thus means they are $R_q$-linearly independent. Consequently, $\mathbf{A}_S$ is invertible (this is true even if $R_q$ is not a field, as shown for example in [BJRW23, Lem. A.3]). We note that if $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod qR$, then it holds that $\mathbf{b}_S = \mathbf{A}_S\mathbf{s} + \mathbf{e}_S$, for $\mathbf{s} \hookleftarrow \mathscr{D}_s$ and $\mathbf{e}_S \in R^d$ is the top subvector of $\mathbf{P}\mathbf{e}$. On the other hand, if $\mathbf{b}$ is uniform in $R_q^m$, then $\mathbf{b}_S$ is uniform in $R_q^d$.

The reduction defines $\mathbf{A}' = -\mathbf{A}_T\mathbf{A}_S^{-1} \bmod qR$ and $\mathbf{b}' = \mathbf{b}_T + \mathbf{A}'\mathbf{b}_S \bmod qR$. The reduction sends the pair $(\mathbf{A}', \mathbf{b}')$ to the pHNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$ oracle. In the search case, it receives back some $\mathbf{e}^\star \sim \mathcal{X}$. It then computes $[\mathbf{e}_S^{\star T} \mid \mathbf{e}_T^{\star T}]^T = \mathbf{P}\mathbf{e}^\star$ and then $\mathbf{s}^\star = \mathbf{A}_S^{-1}(\mathbf{b}_S - \mathbf{e}_S^\star) \bmod qR$. It finally returns $\mathbf{s}^\star$ as the M-LWE solution. In the decision case, the reduction returns the same output as that of the oracle.

Let us now analyze the correctness of the reduction. Since $\mathbf{A}_S \in GL_d(R_q)$, it holds that $\mathbf{A}'$ is uniform in $R_q^{m-d\times d}$. Moreover, since $\mathbf{A}$ is uniform in $R_q^{m\times d}$, if the reduction does not abort (i.e., if there exists such a subset $S$) then the rows selected in $S$ will be random, and hence $\mathbf{P}$ is a random permutation. Furthermore, in the search case we have that

$$\begin{aligned}\mathbf{b}' &= \mathbf{b}_T + \mathbf{A}'\mathbf{b}_S \bmod qR \\ &= \mathbf{A}_T\mathbf{s} + \mathbf{e}_T + \mathbf{A}'(\mathbf{A}_S\mathbf{s} + \mathbf{e}_S) \bmod qR \\ &= \mathbf{A}_T\mathbf{s} - \mathbf{A}_T\mathbf{A}_S^{-1}\mathbf{A}_S\mathbf{s} + \mathbf{A}'\mathbf{e}_S + \mathbf{e}_T \bmod qR \\ &= \mathbf{A}'\mathbf{e}_S + \mathbf{e}_T \bmod qR \\ &= [\mathbf{A}' \mid \mathbf{I}_{m-d}] \cdot [\mathbf{e}_S \mid \mathbf{e}_T] \bmod qR \\ &= [\mathbf{A}' \mid \mathbf{I}_{m-d}] \cdot \mathbf{P} \cdot \mathbf{e} \bmod qR,\end{aligned}$$

which is correctly distributed as $\mathbf{e}$ is drawn from $\mathcal{X}$. Hence, if the oracle successfully returns $\mathbf{e}^\star = \mathbf{e}$, it holds that $\mathbf{s}^\star = \mathbf{A}_S^{-1}(\mathbf{A}_S\mathbf{s} + \mathbf{e}_S - \mathbf{e}_S) \bmod qR = \mathbf{s}$, which is indeed the M-LWE secret. In the decision case, if $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$ then we have seen that $\mathbf{b}'$ is well-distributed. And if $\mathbf{b}$ is uniform, then $\mathbf{b}_T$ is uniform and independent of $\mathbf{A}'\mathbf{b}_S$ which proves that $\mathbf{b}'$ is also uniform in $R_q^{m-d}$.

Therefore, given an oracle $\mathcal{O}$ for pHNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$, we can construct a PPT adversary $\mathcal{A}$, such that

$$\begin{aligned}\mathbb{P}\left[\mathcal{A}(\mathbf{A}, \mathbf{b}) = \mathbf{s}\right] &= \mathbb{P}\left[E\right]\mathbb{P}\left[\mathcal{A}(\mathbf{A}, \mathbf{b}) = \mathbf{s} \mid E\right] + \mathbb{P}\left[\neg E\right]\mathbb{P}\left[\mathcal{A}(\mathbf{A}, \mathbf{b}) = (\mathbf{s}, \mathbf{e}) \mid \neg E\right] \\ &= (1 - \delta'(m,d))\mathbb{P}\left[\mathcal{O}(\mathbf{A}', \mathbf{b}') = \mathbf{e}\right],\end{aligned}$$

where the event $E = \{\exists S \subseteq [m], |S| = d \wedge (\mathbf{a}_i^T)_{i\in S} \text{ are } R_q\text{-linearly independent}\}$. This implies that $\varepsilon_{\text{M-LWE}} \geq (1 - \delta'(m,d))\varepsilon_{\text{pHNF-M-LWE}}$, and hence, we obtain

$$\varepsilon_{\text{pHNF-M-LWE}} \leq \frac{1}{1 - \delta'(m,d)}\varepsilon_{\text{M-LWE}},$$

as claimed. Similarly, for the decision variant, we have

$$\begin{aligned}
&\mathrm{Adv}_{\text{M-LWE}}[\mathcal{A}] \\
&\quad = |\mathbb{P}[\mathcal{A}(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}) = 1] - \mathbb{P}[\mathcal{A}(\mathbf{A}, \mathbf{b} \text{ unif}) = 1]| \\
&\quad = |\mathbb{P}[E]\mathbb{P}[\mathcal{A}(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}) = 1 \mid E] - \mathbb{P}[E]\mathbb{P}[\mathcal{A}(\mathbf{A}, \mathbf{b} \text{ unif}) = 1 \mid E]| \\
&\quad = \mathbb{P}[E]|\mathbb{P}[\mathcal{O}(\mathbf{A}', \mathbf{b}' = [\mathbf{A}' \mid \mathbf{I}_{m-d}]\mathbf{P}\mathbf{e}) = 1] - \mathbb{P}[\mathcal{O}(\mathbf{A}', \mathbf{b}' \text{ unif}) = 1]| \\
&\quad = (1 - \delta'(m, d))\mathrm{Adv}_{\text{pHNF-M-LWE}}[\mathcal{O}],
\end{aligned}$$

which leads to

$$\varepsilon_{\text{pHNF-M-LWE}} \leq \frac{1}{1 - \delta'(m, d)} \varepsilon_{\text{M-LWE}},$$

as desired.

The reduction to HNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$ proceeds exactly the same way, except for the initial check. The reduction instead simply checks whether the first $d$ rows of $\mathbf{A}$ are $R_q$-linearly independent, and aborts otherwise. If this is the case, then $\mathbf{A} = [\mathbf{A}_1^T \mid \mathbf{A}_2^T]^T$ with $\mathbf{A}_1$ invertible. Constructing $\mathbf{A}' = -\mathbf{A}_2\mathbf{A}_1^{-1} \bmod qR$ and $\mathbf{b}' = \mathbf{b}_2 + \mathbf{A}'\mathbf{b}_1 \bmod qR$ then defines a valid instance of HNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$. Indeed, using the same arguments as above, $\mathbf{A}'$ is uniform, and if $\mathbf{b}$ is uniform, so is $\mathbf{b}'$. Then, if $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$, we have $\mathbf{b}' = (\mathbf{A}_2\mathbf{s} + \mathbf{e}_2) - \mathbf{A}_2\mathbf{A}_1^{-1}(\mathbf{A}_1\mathbf{s} + \mathbf{e}_1) = [\mathbf{A}' \mid \mathbf{I}_{m-d}]\mathbf{e}$ where $\mathbf{e}$ is the original M-LWE error following $\mathcal{X}$. When computing the advantages, we simply change the formulas above by considering the event $E = \{(\mathbf{a}_i^T)_{i\in[d]} \text{ are } R_q\text{-linearly independent}\}$. The probability of $E$ is then $1 - \delta'(d, d)$, where $\delta'(d, d) = \delta(d, d) = \prod_{i\in[0,d-1]} \prod_{j\in[\kappa]}(1 - N(\mathfrak{p}_j)^{d-i})$. It then yields the claimed result. □

In [BJRW23], the authors show the following bound $\delta'(m, d) \leq \delta(d, d)^{\lfloor m/d \rfloor} = (1 - \prod_{0\leq i<d} \prod_{j\in[\kappa]}(1 - N(\mathfrak{p}_j)^{-(d-i)}))^{\lfloor m/d \rfloor}$, where the $\mathfrak{p}_j$ are the $\kappa$ prime ideal factors of $\langle q \rangle$. This quantity may not be negligible, but since it only appears multiplicatively, this term does not incur a noticeable security loss for typical parameters. The same holds for the reduction to HNF-M-LWE. Finally, combining Theorem 5.1 with our result of Theorem 4.1, we obtain that HNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$ (or pHNF-M-LWE$_{n,d,m-d,q,\mathcal{X}}$) is at least as hard as M-LWE$_{n,k,m-d+k,q,U(R_q^k),\mathcal{D}}$, conditioned on all the parameters and distributions constraints of Theorem 4.1 being met. As the HNF reduction consumes $d$ samples, meaningful applications of this reduction would require starting with $m \geq 2d$, meaning one would need to instantiate Theorem 4.1 to reach sufficiently many samples. We mention in Section 4.3 that this regime is attainable, and a concrete application can be found in Appendix A.

## 6 Hardness of M-LWE with Leakage

Our results from Sections 4 and 5 prove that under some careful parameter conditions, the search M-LWE problem with $\mathbf{x} = [\mathbf{s}|\mathbf{e}] \sim \mathcal{X}$ remains at least as hard as the standard formulation of M-LWE provided $\mathcal{X}$ is a bounded distribution with sufficient min-entropy. These constraints show that for a given bound and a given min-entropy, the specific shape of the distribution of $\mathbf{x}$ is somewhat irrelevant in the hardness of M-LWE. Additionally, as explored in this section, this allows us to encompass many variants of M-LWE with leakage. If one considers the problem of finding $\mathbf{x} \sim \mathcal{X}$ given $[\mathbf{A} \mid \mathbf{I}]\mathbf{x}$ and some leakage $f(\mathbf{x})$, we can see it as finding $\mathbf{x} \sim \mathcal{X}_f$ given $[\mathbf{A} \mid \mathbf{I}]\mathbf{x}$ where $\mathcal{X}_f$ would be the conditional distribution $\mathcal{X}$ conditioned to some specific value for $f(\mathbf{x})$.

More concretely, for every possible leaked value $c \in f(\mathrm{Supp}(\mathcal{X})) =: Y$, one can define the conditional distribution $\mathcal{X}_{f,c}$ as the distribution of $\mathbf{x}$ conditioned on $f(\mathbf{x}) = c$. We could then apply our result to error distribution $\mathcal{X}_{f,c}$ which has entropy $H_\infty(\mathcal{X}_{f,c}) = H_\infty(\mathbf{x} \mid f(\mathbf{x}) = c) = -\log_2 \max_{\mathbf{x}'} \mathbb{P}[\mathbf{x} = \mathbf{x}' \mid f(\mathbf{x}) = c]$. However, this would only give hardness guarantee for a specific value of $c$. The natural direction would be to take the worst possible value for $c$, i.e., the one giving the smallest entropy. In that context, it would mean computing the worst-case conditional min-entropy defined by

$$H_\infty(\mathbf{x} \mid f(\mathbf{x})) = \min_{c\in Y} H_\infty(\mathcal{X}_{f,c}) = -\log_2 \max_{\mathbf{x}',c} \mathbb{P}[\mathbf{x} = \mathbf{x}' \mid f(\mathbf{x}) = c].$$

However, assessing the security based on this worst-case conditional min-entropy is too strict. Indeed, if an adversary $\mathcal{A}$ had significant advantage only for leaked values $c$ which happen with negligible probability, it means the overall advantage of $\mathcal{A}$ in breaking the cryptographic scheme would be negligible anyway. It means we can smooth out these corner cases in security arguments, and only deal with values $c$ that happen with non-negligible probability. We note that for these values, it is possible to sample $\mathbf{x}$ a posteriori, i.e., sampling $\mathbf{x}$ such that $f(\mathbf{x}) = c$ for a fixed $c$, as long as $\mathcal{X}$ is efficiently sampleable and $f$ is efficiently computable. It comes from the requirement that for such $c$, we may need $\mathcal{X}_{f,c}$ to be efficiently sampleable. For that, the challenger could use a rejection sampling strategy and sample polynomially many $\mathbf{x}$ from $\mathcal{X}$ until it finds one that verifies $f(\mathbf{x}) = c$. As $c$ occurs with non-negligible probability, it would require at most a polynomial number of samples $\mathbf{x}$. We note that the leakage functions we consider in this section satisfy this efficiency requirement.

Nevertheless, because the leaked value $f(\mathbf{x})$ is not under adversarial control (as $\mathbf{x}$ and $f(\mathbf{x})$ are chosen/-computed by the challenger), it means that averaging over the value of $c$ suffices for security arguments as explained also in [DORS03, App. B]. We thus consider the average conditional min-entropy $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x}))$ which quantifies more faithfully the hardness of M-LWE with leakage using the result of previous sections. It is based on the fact that given the leaked value $f(\mathbf{x}) = c$, the predictability of $\mathbf{x}$ by the adversary is at most $\max_{\mathbf{x}'} \mathbb{P}[\mathbf{x} = \mathbf{x}' \mid f(\mathbf{x}) = c]$. On average, the adversary's chance of successfully predicting $\mathbf{x}$ is then $\mathbb{E}_c[\max_{\mathbf{x}'} \mathbb{P}[\mathbf{x} = \mathbf{x}' \mid f(\mathbf{x}) = c]]$. The average conditional min-entropy then corresponds to the bit-security of this average predictability, i.e.,

$$\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x})) = -\log_2 \mathbb{E}_c[\max_{\mathbf{x}'} \mathbb{P}[\mathbf{x} = \mathbf{x}' \mid f(\mathbf{x}) = c]].$$

*Remark 6.1.* Both the worst-case and average conditional min-entropy can be delicate to estimate or tightly bound in the general case. Nevertheless, we show that certain generic bounds (Lemma 6.1) can still be used to derive meaningful conclusions to apply our hardness result. We also showed that widely used distributions like discrete Gaussians can lead to tighter bounds (Lemma 3.3) in the context of noisy linear leakage. Obtaining better evaluations of this quantity for various distributions and leakage functions then readily improves the hardness guarantees using our result.

In what follows, we apply our result to this average conditional min-entropy in order to obtain an average-case hardness which is sufficient for cryptographic applications. If one desires to deal with specific leaked values $c$, then we note that it is possible to directly apply our result if one is able to compute $H_\infty(\mathcal{X}_{f,c})$ as described above. One can also determine from the conditional min-entropy that $H_\infty(\mathcal{X}_{f,c})$ is sufficiently large compared to $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x}))$ to apply the result, with the desired probability over the choice of $c$, i.e., for values $c$ that are likely to happen as we described above. For that, one can use [DORS08, Lem. 2.2], which is recalled in Item 1 Lemma 2.4, to relate the two quantities.

Overall, if we are further able to show the average conditional min-entropy $\widetilde{H_\infty}((\mathbf{s}, \mathbf{e}) \mid f(\mathbf{s}, \mathbf{e}))$ is above the threshold for the hardness of our reduction, it would in turn provide hardness of M-LWE with leakage function $f$. We now formalize this with the following M-LWE with leakage assumption.

**Definition 6.1 (Leaky-M-LWE).** *Let $n, d, m, q$ be in $\mathbb{N}^\times$. Let $R$ be the ring of integers of a number field of degree $n$, and $\mathcal{X}$ be a distribution supported on a set $X \subseteq R^{m+d}$. Further, let $f \colon X \to Y$ be a function with range $Y$ called the leakage space. The search version of the Leaky Module Learning With Errors problem, denoted as* Leaky-M-LWE$_{n,d,m,q,\mathcal{X},f}$*, is as follows: Given $\mathbf{A} \hookleftarrow U(R_q^{m \times d})$, $\mathbf{b} = [\mathbf{A} \mid \mathbf{I}_m] \cdot \mathbf{x} \bmod qR$ and $f(\mathbf{x})$ for some $\mathbf{x} \hookleftarrow \mathcal{X}$, find $\mathbf{x}$.*

In this general formulation, we leave unspecified how the function $f$ is defined. One can for instance assume that it is honestly chosen from some distribution, or that the adversary has the right to choose $f$ within some given constraints. Here, the choice of the adversary might even come *after* having seen the matrix $\mathbf{A}$. Observe however that many choices of $f$ yield a totally trivial problem. For example, if $f$ is a linear invertible function, then the leakage allows for recovering $\mathbf{x}$ efficiently. For an arbitrary bijective function $f$, the problem becomes easy as long as $f^{-1}$ is efficiently computable. As in this case the conditional min-entropy is zero, our reduction does not cover the easy instances.

In the rest of this section we summarize different leakage variants used in the literature and show how our work encompasses them with the Leaky-M-LWE assumption. This demonstrates the full potential of our results. In particular, our work provides the first hardness results for M-LWE with non-linear leakage. As these variants were introduced primarily to improve cryptographic constructions, our framework targets the same applications by providing further confidence in the underlying leakage variants.

### 6.1 Tailored Approach for Approximate Linear Leakage

We start by looking at a special case of Leaky-M-LWE, where the leakage is a noisy linear equation, often called Hint-M-LWE in the literature. In particular, the leakage function $f$ is defined by $f(\mathbf{x}) = \mathbf{M}\mathbf{x} + \mathbf{f}$ for some adversarially chosen matrix $\mathbf{M} \in R^{k\times m+d}$ and honestly sampled Gaussian noise $\mathbf{f}$. Note that the adversary knows $\mathbf{M}$, but not $\mathbf{f}$[8]. Several works have used these variants to construct more efficient primitives such as functional encryption [MKMS22], updatable encryption [HPS23], zero-knowledge proofs [KLSS23], threshold signatures [ENP24], or laconic cryptography [DKL+23]. However, all existing reductions from standard M-LWE proposed in these works only function for $\mathbf{x}$ coming from a discrete Gaussian distribution. Ours is the first to work for general noise distributions $\mathcal{X}$. Indeed, they essentially study the conditional distribution of $\mathbf{x}$ given $\mathbf{M}\mathbf{x}+\mathbf{f}$. When both $\mathbf{x}$ and $\mathbf{f}$ are Gaussian, convolution theorems akin to [GMPW20, Thm. 3.1] can be used to characterize the leakage distribution and in turn the conditional distribution. However, if $\mathbf{x}$ does not follow the prescribed Gaussian distribution, this approach based on convolution fails as the leakage distribution becomes harder to analyze. In our case, instead of studying the conditional distribution, we only look at the conditional entropy, which is all we need to apply our hardness result. This allows for generalizing to different distributions for $\mathbf{x}$ while guaranteeing hardness, which was not known before.

For that, we rely on our generalization of [BD20a] provided in Section 3 and Lemma 3.3. It shows that as long as $\mathbf{M}\mathbf{x}$ is bounded by some value $B$ in $\ell_2$ norm, one can mask it with a discrete Gaussian of parameter $\sigma$ resulting in

$$\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{M}\mathbf{x}+\mathbf{f}) \geq H_\infty(\mathbf{x}) - \sqrt{2\pi nk}\frac{B}{\sigma}\log_2 e - \log_2(1+\varepsilon),$$

provided that $\sigma \geq \eta_\varepsilon(R^k) = \eta_\varepsilon(\mathbb{Z}^{nk})$ ($\approx \sqrt{\ln(2nk/\varepsilon)/\pi}$ by [EWY23]). As we only place a mild requirement on $\|\mathbf{M}\mathbf{x}\|_2$, our approach is then better suited to capture non-Gaussian distributions for $\mathbf{x}$. In particular, if $\sigma \geq B\sqrt{nk}$, then the entropy loss is only $\sqrt{2\pi}\log_2 e + \log_2(1+\varepsilon)$.

This tailored approach is thus the combination of two key ingredients: (1) our generalization of the noise lossiness for better trade-offs, and (2) the fact that we can prove the hardness solely based on evaluating this conditional min-entropy. This opens for different choices of distributions for $\mathbf{x}$ in M-LWE variants with approximate linear leakage. This is for example the case of [dPEK+23], which constructs a side-channel resistant signature scheme, although we warn that our reduction does not cover the practical parameter regime they consider.

### 6.2 Leakage With Bounded Support

Although arbitrary leakage functions may be difficult to analyze in general, our approach of abstracting the secret-error distribution comes down to evaluating the residual entropy given the leakage as explained above. We can thus use Lemma 2.4 to obtain a lower bound on said entropy, provided that the range of $f$ is finite and sufficiently small compared (in log) to the initial entropy of $\mathcal{X}$.

**Lemma 6.1.** *Assuming $f$ takes at most $N$ different values, we have $\widetilde{H_\infty}(\mathbf{x} \mid f(\mathbf{x})) \geq H_\infty(\mathbf{x}) - \log_2 N$. It also holds that*

$$\mathbb{P}_{y\sim f(\mathcal{X})}\left[H_\infty(\mathbf{x} \mid f(\mathbf{x}) = y) \geq H_\infty(\mathbf{x}) - \log_2 N - \log_2(1/\delta)\right] \geq 1-\delta.$$

[8] Previous Extended-LWE assumptions, e.g., [AP12], considered $\mathbf{M}$ to be (a single row vector) chosen by the challenger from a discrete Gaussian distribution.

*Proof.* The first statement directly holds from Item 2 of Lemma 2.4. Combined with Item 1, it yields the second claim. □

*Remark 6.2.* It provides a systematic way of evaluating the hardness of the Leaky-M-LWE problem simply by computing or bounding the size of the leakage space $|Y|$. In the context of arbitrary but bounded-size leakage, an alternative strategy could be to guess the leaked value. More concretely, given a regular M-LWE instance $(\mathbf{A}, \mathbf{b})$, the reduction guesses $c$ such that $f(\mathbf{x}) = c$ and calls the Leaky-M-LWE oracle on $(\mathbf{A}, \mathbf{b}, c)$. This guessing approach is commonly used when no tailored technique is known and entails a reduction security loss that depends on $|Y|$ as well. This is why it is limited to the situations where $|Y|$ is polynomial. Our approach allows us to go beyond polynomial-sized leakage spaces, as long as $H_\infty(\mathbf{x})$ is sufficiently larger than $\log_2|Y|$. More concretely, if we take the example of a leakage with $2^\lambda$ possible values with $\lambda$ the security parameter (e.g., $\lambda = 128$), we only need $\lambda$ extra bits of entropy to begin with. This can be easily achieved in typical scenarios in lattice-based cryptography (e.g., for $\mathbf{x}$ of dimension 256, going from $\mathbf{x}$ binary to $\mathbf{x}$ ternary gives $256\log_2(3/2) \geq 128$ extra bits of entropy).

*Remark 6.3.* Notice that this generic approach would give worse results for approximate linear leakage than the tailored one described in Section 6.1. Indeed, assume we consider the Gaussian of width $\sigma$ for $\mathbf{f}$ to be truncated in $\ell_\infty$ norm at $t\sigma$ for some $t > 0$. We would have $\|\mathbf{Mx} + \mathbf{f}\|_\infty \leq \|\mathbf{Mx}\|_\infty + t\sigma \leq B_\infty + t\sigma$. Hence, the entropy bound would be $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{Mx}+\mathbf{f}) \geq H_\infty(\mathbf{x}) - nk\log_2(2(B_\infty + t\sigma)+1)$. As a result, the larger the mask (i.e., larger $\sigma$), the more entropy for $\mathbf{x}$ we would need to guarantee hardness, which is counter-intuitive. This demonstrates the looseness of the generic approach in certain use cases. For Hint-M-LWE as defined in the works mentioned in Section 6.1, the tailored approach based on Lemma 3.3 then follows the intuition that the larger the mask, the less entropy we need in the first place as the hint would give less and less information on $\mathbf{x}$.

The general approach is most interesting for leakage functions that do not fit the approximate linear hint framework, and for which no prior hardness result was known. We detail below a few concrete examples including exact linear hints, as well as quadratic leakage which, to our knowledge, was not yet tackled by any theoretical hardness result.

**6.2.1 Exact Linear Leakage.** First, we tackle several variants from the so-called Extended-M-LWE family that use exact linear hints without modular reduction, i.e., when $f$ is a linear function over $\mathbb{Z}$ (or $R$). In that context, the elements constituting the leakage can be assumed short (meaning they do not entail wrap-around modulo $q$ anyway). Examples of such variants have been used to construct functional encryption [ALS16], zero-knowledge proofs [LN22], side-channel resistant signatures [dPEK+23], or used in reductions [AA16, BJRW21]. Exact linear hints are typically expressed via $f(\mathbf{x}) = \mathbf{M}\tau(\mathbf{x})$ for a short adversarially chosen $\mathbf{M} \in \mathbb{Z}^{k\times n(m+d)}$. Note that this encompasses leakage over the ring $f(\mathbf{x}) = \mathbf{M}'\mathbf{x}$ by defining $\mathbf{M} = M_\tau(\mathbf{M}')$. We can then bound the cardinal of the support of the leakage space in a similar fashion as in Lemma 4.1 using induced $(\alpha, \beta)$-norms.

**Lemma 6.2 (Generic Exact Linear Leakage).** *The cardinal of the range of $f$ is bounded by $\min_{\alpha,\beta}\left|\mathcal{B}_{\beta,k}(B_{\alpha,\beta}B_\alpha) \cap \mathbb{Z}^k\right|$, where $B_{\alpha,\beta} = \|\mathbf{M}\|_{\alpha,\beta}$ and $B_\alpha = \max_{\mathbf{x}\in X}\|\mathbf{x}\|_\alpha$.*

If the bound exceeds $|X|$, then the residual entropy bound would be $H_\infty(\mathcal{X}) - \log_2|X| \leq 0$ which is vacuous. Reaching this bound of $|X|$ is possible for example if $\ker(\mathbf{M}) = \{\mathbf{0}\}$, i.e., $f$ is injective. It means that $f$ can be inverted and that Leaky-M-LWE$_{n,d,m,q,\mathcal{X},f}$ is easy in this case. It then limits the number of hints $k$ that can be safely given to the adversary. We now explain how Lemma 6.2 helps covering the hardness of the aforementioned works. We distinguish between the adversarially and non-adversarially chosen case.

*Adversarial Exact Linear Hints.* In both [AA16, BJRW21], the linear function is only on the error part (we thus do not need the HNF transform of Section 5) and most importantly adversarial, meaning $\mathbf{M}$ is chosen by the adversary in a certain set. Indeed, [BJRW21] considers $f(\mathbf{e}) = \mathbf{z}^T\mathbf{e}$ over the ring for a short adversarially-chosen $\mathbf{z}$. In that case, we can express it as above with $\mathbf{M} = [\mathbf{0} \mid M_\tau(\mathbf{z}^T)] \in \mathbb{Z}^{n\times n(m+d)}$ (the

**0** part corresponding to the secret part of $\mathbf{x}$). Let us now bound $\|\mathbf{M}\|_{2,\infty}$ as an example. It corresponds to the maximal $\ell_2$ norm of the rows. Yet, the $i$-th row of $M_\tau(\mathbf{z}^T)$ is the $i$-th column of $M_\tau(\mathbf{z}^T)^T = M_\tau(\mathbf{z}^{*T})$ which is exactly $\tau(\mathbf{z}^* x^i)$. We then have $B_{2,\infty} = \max_{0\le i<n} \left\|\mathbf{z}^* x^i\right\|_2$, which simplifies to $\|\mathbf{z}\|_2$ in power-of-two cyclotomic fields. It means the size of the leakage space is bounded by $(2\lfloor \max_{\mathbf{e}} \|\mathbf{e}\|_2 \cdot \max_i \left\|\mathbf{z}^* x^i\right\|_2 \rfloor + 1)^n$, which decreases the min-entropy by at most the logarithm of that. We note that as opposed to [BJRW21], our case covers general number fields, and does not restrict the error distribution to a continuous Gaussian. Also, it does not require the explicit construction of a short unimodular matrix as in [BJRW21, Lem. 15].

In the general formulation of [AA16], they define $f(\mathbf{e}) = [\mathrm{Tr}(\langle \mathbf{z}_i \,,\, \mathbf{e}\rangle)]_{i\in[k]}$, which is linear by linearity of the field trace and takes values in $\mathbb{Q}^k$. We can write it as $f(\mathbf{e}) = \mathbf{Z}^T \widetilde{\mathbf{V}}^T \widetilde{\mathbf{V}} \tau(\mathbf{e})$, where $\widetilde{\mathbf{V}} = \mathbf{I}_m \otimes \mathbf{V}$, for the Vandermonde matrix $\mathbf{V}$ of the field, and $\mathbf{Z} = [\tau(\mathbf{z}_1) \mid \ldots \mid \tau(\mathbf{z}_k)]$. In the power-of-two cyclotomic case, a standard calculation shows that $[\mathbf{V}^T\mathbf{V}]_{i,j} = 0$ if $i + j \notin \{0, n\}$ and $(-1)^{(i+j)/n} \cdot n$ otherwise. Hence, defining $\mathbf{M} = \mathbf{Z}^T(\mathbf{I}_m \otimes \mathbf{V}^T\mathbf{V})$ yields a matrix in $\mathbb{Z}^{k\times nm}$ which can be used in Lemma 6.2 with an appropriate bound. For example, in the power-of-two cyclotomic case, one gets $\|\mathbf{M}\|_{2,\infty} = n \cdot \max_{i\in[k]} \|\mathbf{z}_i\|_2$.

The set of vectors $\{\mathbf{z}_i\}_{i\in[k]}$ they use in their paper however leads to a much simpler situation. Although they formulate the assumption with the trace, they only use it for adversarial vectors $\mathbf{z}_i$ where each one essentially extracts a single coefficient of $\mathbf{e}$. More concretely, they use $\mathbf{M} = [\mathbf{0} \mid \mathbf{Z}] \in \{0,1\}^{k\times n(m+d)}$ where $\mathbf{Z} \in \{0,1\}^{k\times nm}$ is of rank $k$ and has a single 1 in each row (i.e., up to permutation of the columns, $\mathbf{Z} = [\mathbf{I}_k \mid \mathbf{0}]$). In that case, we can use the $(\infty,\infty)$-norm and get $B_{\infty,\infty} = 1$. Hence, the leakage space is directly bounded by $(2\max_{\mathbf{e}} \|\mathbf{e}\|_\infty + 1)^k$. Alternatively, as it comes down to leaking $k$ out of $nm$ coefficients, it may be analyzed differently for specific error distributions. For example, if all the coefficients $\tau_i(\mathbf{e})$ are independently sampled from distributions $\psi_i$, leaking $\{\tau_i(\mathbf{e}) \colon i \in S\}$ then entails that the average and worst-case conditional entropy coincide and equal $\widetilde{H}_\infty(\mathbf{e} \mid f(\mathbf{e})) = H_\infty(\mathbf{e} \mid f(\mathbf{e})) = \sum_{i\in[nm]\setminus S} H_\infty(\psi_i)$. In the case of a discrete Gaussian error for example, this would simply be $H_\infty(\mathcal{D}_{\mathbb{Z}^{nm-k},s}) \ge (nm-k)\log_2 s$.

*Non-adversarial Exact Linear Hints.* Cryptographic constructions also use the exact linear hints but with a non-adversarial function [ALS16, LN22, dPEK+23]. In that case, the matrix $\mathbf{M}$ is controlled by the challenger and follows a prescribed distribution, which can allow for better fine-tuning. In [dPEK+23], the hints are described as $\mathbf{M}\tau(\mathbf{x})$ which is directly covered by Lemma 6.2. The situation in [ALS16] is exactly the same except they consider the unstructured problem LWE (i.e., when $n = 1$).

In [LN22], the hint has a slightly different shape than in [AA16, BJRW21], but most importantly does not contain adversarially chosen elements[9]. In their case, the function is $f(\mathbf{x}) = \langle \tau(c\mathbf{x}) \,,\, \tau(\mathbf{y})\rangle$, where $\mathbf{y}$ is the mask of $c\mathbf{s}$ in the zero-knowledge proof, and $c$ is the public challenge generated as a hash output. This can then be written as $\mathbf{M}\tau(\mathbf{x})$ with $\mathbf{M} = \tau(\mathbf{y})^T(\mathbf{I}_{m+d} \otimes M_\tau(\mathbf{c})) = \tau(c^*\mathbf{y})^T$. Using the $(2,\infty)$-norm (which comes down to a Cauchy-Schwarz inequality), we have $B_{2,\infty} = \|c^*\mathbf{y}\|_2 \le \eta \cdot t\sigma_2\sqrt{n(d+m)}$, where $\eta$ is an integer defining the challenge space, and $\sigma_2$ is the discrete Gaussian parameter of $\mathbf{y}$ tailcutted at $t$. The residual entropy is then at least $H_\infty(\mathbf{x}) - \log_2(2\eta t\sigma_2\sqrt{n(d+m)} \cdot \max_{\mathbf{x}\in X} \|\mathbf{x}\|_2 + 1)$. In most cases, this leakage decreases the entropy by at most $\log_2 q$ bits.

**6.2.2 Quadratic Leakage.** Our result also provides a way to prove the hardness of M-LWE beyond linear leakage. Variants of LWE with non-linear hints have been recently introduced, like quadratic leakage in [MS23, PS24] to propose threshold encryption and threshold FHE. In that context, the leakage function is $f(\mathbf{x}) = \mathbf{x}^T\mathbf{F}\mathbf{x} + \mathbf{f}^T\mathbf{x} + f$, possibly with automorphism, or the evalutation thereof. In the threshold PKE construction of [MS23], the authors introduce the Known-Norm LWE assumption where the adversary learns $f(\mathbf{x}) = \|\mathbf{x}\|_2^2$. This can be generalized to the module setting by giving out $f(\mathbf{x}) = \|\tau(\mathbf{x})\|_2^2$. However, in the structured case, the authors instead generalize the hint to $f(\mathbf{x}) = \langle \mathbf{x} \,,\, \mathbf{x}\rangle_{K_\mathbb{R}}$ which corresponds to the autocorrelation (which they call covariance) of $\mathbf{x}$ defined by $\langle \mathbf{x} \,,\, \mathbf{x}\rangle_{K_\mathbb{R}} = \sum_{i\in[m+d]} x_i^* x_i = \mathbf{x}^*\mathbf{x}$. When $K$ is a cyclotomic field, the conjugate is the automorphism of $K$ defined by $\zeta \mapsto \zeta^{-1}$, which also stabilizes the ring

[9] This was generalized to adversarially chosen $c$ and $\mathbf{y}$ in [WLL24]. Our result can easily handle this case in a similar way as long as $c$ and $\mathbf{y}$ are bounded. In the worst case, they can be considered modulo $q$, reducing the entropy by at most $\log_2 q$.

of integers $R$. It means for $\mathbf{x} \in R^{m+d}$, $f(\mathbf{x}) \in R$. The resulting problem is referred to as Known-Covariance M-LWE (which they only use in the ring setting, i.e., $d = 1$). In power-of-two cyclotomics, we can clearly see it is no harder than the Known-Norm case as the constant coefficient of $\mathbf{x}^*\mathbf{x}$ is $\|\tau(\mathbf{x})\|_2^2$.

The case of Known-Norm M-LWE can be easily handled by guessing $\|\tau(\mathbf{x})\|_2^2$ which is a single (polynomially-bounded) integer, but this method fails for the Known-Covariance M-LWE problem as it would require guessing an entire ring element. Our approach can take over in this setting by bounding the size of the range of $f$. Using similar bounds as for the exact linear hints, we have $\max_{\mathbf{x}} \|\mathbf{x}^*\mathbf{x}\|_\infty \leq \max_{\mathbf{x}} \|M_\tau(\mathbf{x}^*)\|_{2,\infty} \|\mathbf{x}\|_2 = \max_{\mathbf{x}} \max_i \left\|\mathbf{x}x^i\right\|_2 \|\mathbf{x}\|_2$. Focusing on the power-of-two cyclotomic case, we obtain a bound of $\max_{\mathbf{x}} \|\mathbf{x}\|_2^2$. This would give a leakage space of at most $(2 \max_{\mathbf{x}} \|\mathbf{x}\|_2^2 + 1)^n$ elements. We however notice that $f(\mathbf{x})$ is self-adjoint, i.e., $f(\mathbf{x})^* = f(\mathbf{x})$. Hence, $\tau_{n/2}(f(\mathbf{x})) = 0$ and for all $i \in [1, n/2 - 1], \tau_{n-i}(f(\mathbf{x})) = -\tau_i(f(\mathbf{x}))$. Hence, $f(\mathbf{x})$ takes at most $(2B + 1)^{n/2}$ values, where $B = \max_{\mathbf{x}} \|\mathbf{x}\|_2^2$.

As a result, the average conditional min-entropy is at least $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{x}^*\mathbf{x}) \geq H_\infty(\mathbf{x}) - \frac{n}{2} \log_2(2B+1)$. This could possibly be refined as it appears that although the constant coefficient of $\mathbf{x}^*\mathbf{x}$ is $\|\mathbf{x}\|_2^2$, the other ones tend to be much smaller. Also, we note that this conditional distribution is invariant under permutation as $(\mathbf{Px})^*(\mathbf{Px}) = \mathbf{x}^*\mathbf{P}^T\mathbf{Px} = \mathbf{x}^*\mathbf{x}$. In that case the pHNF-M-LWE and regular HNF-M-LWE coincide.

We give more details on how to instantiate our result to prove the hardness of Known-Covariance M-LWE in Appendix A. In particular, we give a parameter set ensuring close to 128 bits of security to show the applicability of our result. Albeit with impractical parameters, it shows that there is nothing fundamentally weak with leaking $\mathbf{x}^*\mathbf{x}$.

**6.2.3 Non-Algebraic Leakage.** Our result can also be used to handle non-algebraic leakages contrarily to the ones we covered in the previous sections. A simple example is that of the Extended-M-LWE problem formulated in [LNS21] where $f(\mathbf{x}) = \text{sign}(\langle \tau(c\mathbf{x}) , \tau(\mathbf{y}) \rangle)$, where $\mathbf{y}$ is a mask for $c\mathbf{x}$ in the zero-knowledge proof, and $c$ is the public challenge obtained as a hash output. This can be seen as a weaker version of the Extended-M-LWE problem from [LN22] mentioned in Section 6.2.1, which now just leaks the sign of $\langle \tau(c\mathbf{x}) , \tau(\mathbf{y}) \rangle$ instead of the full inner product. As a result, the leakage space is simply $\{-1, 1\}$ yielding a residual entropy of at least $H_\infty(\mathbf{x}) - 1$. The authors provide a reduction from LWE in the unstructured case which strongly relies on the fact that both $c$ and $\mathbf{y}$ are controlled by the challenger. Our result extends it (albeit most likely loosely) to the structured case of M-LWE and to adversarial $\mathbf{y}$ and $c$.

More generally, non-algebraic leakage often comes up in side-channel analysis, for example through timing measurements, power traces, or electromagnetic readings. These are generally non-linear in the secret data, and even most likely non-algebraic. Sophisticated machinery is then needed to transform these non-algebraic leakage into information that is more relevant to solving M-LWE (such as randomness bit, hamming weight, etc.). We only mention that our result is general enough to cover these situations, again provided one can lower-bound the average conditional min-entropy. Because leakage incurred by insecure implementations is not easy to model, we leave the application of our result to such use-cases as an interesting research direction. In particular, modeling the leakage exploited in side-channel attacks would provide a deeper understanding of the actual M-LWE assumption underlying such vulnerable systems.

## Acknowledgments

We thank the anonymous reviewers for their helpful feedback. This work was supported by the French government through the National Research Agency, under the projects ANR-21-ASTR-0016 *AMIRAL*, ANR-22-PECY-003 *SecureCompute*, ANR-24-CE48-4293 *HELO* and ANR-25-CE39-4214-01 *RELATE*, by the Austrian Science Fund (FWF) Project J4879-N, by the Bakar Funds and Peder Sather Funds.

## References

AA16. Jacob Alperin-Sheriff and Daniel Apon. Dimension-preserving reductions from LWE to LWR. Cryptology ePrint Archive, Report 2016/589, 2016.

ACPS09. Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In Shai Halevi, editor, *CRYPTO 2009*, volume 5677 of *LNCS*, pages 595–618. Springer, Berlin, Heidelberg, August 2009.

AG11. Sanjeev Arora and Rong Ge. New algorithms for learning in presence of errors. In Luca Aceto, Monika Henzinger, and Jiri Sgall, editors, *ICALP 2011, Part I*, volume 6755 of *LNCS*, pages 403–415. Springer, Berlin, Heidelberg, July 2011.

Ajt96. Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *28th ACM STOC*, pages 99–108. ACM Press, May 1996.

ALS16. Shweta Agrawal, Benoît Libert, and Damien Stehlé. Fully secure functional encryption for inner products, from standard assumptions. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part III*, volume 9816 of *LNCS*, pages 333–362. Springer, Berlin, Heidelberg, August 2016.

AP12. Jacob Alperin-Sheriff and Chris Peikert. Circular and KDM security for identity-based encryption. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 334–352. Springer, Berlin, Heidelberg, May 2012.

APS15. Martin R. Albrecht, Rachel Player, and Sam Scott. On the concrete hardness of learning with errors. Cryptology ePrint Archive, Report 2015/046, 2015.

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

BD20a. Zvika Brakerski and Nico Döttling. Hardness of LWE on general entropic distributions. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part II*, volume 12106 of *LNCS*, pages 551–575. Springer, Cham, May 2020.

BD20b. Zvika Brakerski and Nico Döttling. Lossiness and entropic hardness for ring-LWE. In Rafael Pass and Krzysztof Pietrzak, editors, *TCC 2020, Part I*, volume 12550 of *LNCS*, pages 1–27. Springer, Cham, November 2020.

BDK+18. Joppe W. Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. CRYSTALS - Kyber: A CCA-secure module-lattice-based KEM. In *2018 IEEE European Symposium on Security and Privacy*, pages 353–367. IEEE Computer Society Press, April 2018.

BGV12. Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In Shafi Goldwasser, editor, *ITCS 2012*, pages 309–325. ACM, January 2012.

BJRW20. Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, and Weiqiang Wen. Towards classical hardness of module-LWE: The linear rank case. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part II*, volume 12492 of *LNCS*, pages 289–317. Springer, Cham, December 2020.

BJRW21. Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, and Weiqiang Wen. On the hardness of module-LWE with binary secret. In Kenneth G. Paterson, editor, *CT-RSA 2021*, volume 12704 of *LNCS*, pages 503–526. Springer, Cham, May 2021.

BJRW22. Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, and Weiqiang Wen. Entropic hardness of module-LWE from module-NTRU. In Takanori Isobe and Santanu Sarkar, editors, *INDOCRYPT 2022*, volume 13774 of *LNCS*, pages 78–99. Springer, Cham, December 2022.

BJRW23. Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, and Weiqiang Wen. On the hardness of module learning with errors with short distributions. *Journal of Cryptology*, 36(1):1, January 2023.

BK25. Katharina Boudgoust and Hannah Keller. Module learning with errors with truncated matrices. In Ruben Niederhagen and Markku-Juhani O. Saarinen, editors, *Post-Quantum Cryptography - 16th International Workshop, PQCrypto 2025, Part I*, pages 255–277. Springer, Cham, April 2025.

BLP+13. Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th ACM STOC*, pages 575–584. ACM Press, June 2013.

CHK+21. Chi-Ming Marvin Chung, Vincent Hwang, Matthias J. Kannwischer, Gregor Seiler, Cheng-Jhih Shih, and Bo-Yin Yang. NTT multiplication for NTT-unfriendly rings. *IACR TCHES*, 2021(2):159–188, 2021.

DGPY20. Léo Ducas, Steven Galbraith, Thomas Prest, and Yang Yu. Integral matrix gram root and lattice gaussian sampling without floats. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part II*, volume 12106 of *LNCS*, pages 608–637. Springer, Cham, May 2020.

DKL+18. Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, Peter Schwabe, Gregor Seiler, and Damien Stehlé. CRYSTALS-Dilithium: A lattice-based digital signature scheme. *IACR TCHES*, 2018(1):238–268, 2018.

DKL+23. Nico Döttling, Dimitris Kolonelos, Russell W. F. Lai, Chuanwei Lin, Giulio Malavolta, and Ahmadreza Rahimi. Efficient laconic cryptography from learning with errors. In Carmit Hazay and Martijn Stam,

editors, *EUROCRYPT 2023, Part III*, volume 14006 of *LNCS*, pages 417–446. Springer, Cham, April 2023.

DKM+24. Rafaël Del Pino, Shuichi Katsumata, Mary Maller, Fabrice Mouhartem, Thomas Prest, and Markku-Juhani O. Saarinen. Threshold raccoon: Practical threshold signatures from standard lattice assumptions. In Marc Joye and Gregor Leander, editors, *EUROCRYPT 2024, Part II*, volume 14652 of *LNCS*, pages 219–248. Springer, Cham, May 2024.

DORS03. Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. Cryptology ePrint Archive, Report 2003/235, 2003.

DORS08. Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM journal on computing*, 38(1):97–139, 2008.

dPEK+23. R. del Pino, T. Espitau, S. Katsumata, M. Maller, F. Mouhartem, T. Prest, M. Rossi, and M.-J. Saarinen. *Raccoon: A Side-Channel Secure Signature Scheme (Specifications)*, 2023. Available at https://raccoonfamily.org/wp-content/uploads/2023/07/raccoon.pdf.

dPKPR24. Rafaël del Pino, Shuichi Katsumata, Thomas Prest, and Mélissa Rossi. Raccoon: A masking-friendly signature proven in the probing model. In Leonid Reyzin and Douglas Stebila, editors, *CRYPTO 2024, Part I*, volume 14920 of *LNCS*, pages 409–444. Springer, Cham, August 2024.

ENP24. Thomas Espitau, Guilhem Niot, and Thomas Prest. Flood and submerse: Distributed key generation and robust threshold signature from lattices. In Leonid Reyzin and Douglas Stebila, editors, *CRYPTO 2024, Part VII*, volume 14926 of *LNCS*, pages 425–458. Springer, Cham, August 2024.

EWY23. Thomas Espitau, Alexandre Wallet, and Yang Yu. On gaussian sampling, smoothing parameter and application to signatures. In Jian Guo and Ron Steinfeld, editors, *ASIACRYPT 2023, Part VII*, volume 14444 of *LNCS*, pages 65–97. Springer, Singapore, December 2023.

GKPV10. Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the learning with errors assumption. In Andrew Chi-Chih Yao, editor, *ICS 2010*, pages 230–240. Tsinghua University Press, January 2010.

GMPW20. Nicholas Genise, Daniele Micciancio, Chris Peikert, and Michael Walter. Improved discrete gaussian and subgaussian analysis for lattice cryptography. In Aggelos Kiayias, Markulf Kohlweiss, Petros Wallden, and Vassilis Zikas, editors, *PKC 2020, Part I*, volume 12110 of *LNCS*, pages 623–651. Springer, Cham, May 2020.

HPS98. Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NTRU: A ring-based public key cryptosystem. In *Third Algorithmic Number Theory Symposium (ANTS)*, volume 1423 of *LNCS*, pages 267–288. Springer, June 1998.

HPS23. Calvin Abou Haidar, Alain Passelègue, and Damien Stehlé. Efficient updatable public-key encryption from lattices. In Jian Guo and Ron Steinfeld, editors, *ASIACRYPT 2023, Part V*, volume 14442 of *LNCS*, pages 342–373. Springer, Singapore, December 2023.

KLSS23. Duhyeong Kim, Dongwon Lee, Jinyeong Seo, and Yongsoo Song. Toward practical lattice-based proof of knowledge from hint-MLWE. In Helena Handschuh and Anna Lysyanskaya, editors, *CRYPTO 2023, Part V*, volume 14085 of *LNCS*, pages 549–580. Springer, Cham, August 2023.

LN22. Vadim Lyubashevsky and Ngoc Khanh Nguyen. BLOOM: Bimodal lattice one-out-of-many proofs and applications. In Shweta Agrawal and Dongdai Lin, editors, *ASIACRYPT 2022, Part IV*, volume 13794 of *LNCS*, pages 95–125. Springer, Cham, December 2022.

LNS21. Vadim Lyubashevsky, Ngoc Khanh Nguyen, and Gregor Seiler. Shorter lattice-based zero-knowledge proofs via one-time commitments. In Juan Garay, editor, *PKC 2021, Part I*, volume 12710 of *LNCS*, pages 215–241. Springer, Cham, May 2021.

LPR10. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Henri Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 1–23. Springer, Berlin, Heidelberg, May / June 2010.

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

LS18. Vadim Lyubashevsky and Gregor Seiler. Short, invertible elements in partially splitting cyclotomic rings and applications to lattice-based zero-knowledge proofs. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part I*, volume 10820 of *LNCS*, pages 204–224. Springer, Cham, April / May 2018.

LWZW20. Hao Lin, Mingqiang Wang, Jincheng Zhuang, and Yang Wang. Hardness of module-LWE and ring-LWE on general entropic distributions. Cryptology ePrint Archive, Report 2020/1238, 2020.

MAM+24. Anisha Mukherjee, Aikata, Ahmet Can Mert, Yongwoo Lee, Sunmin Kwon, Maxim Deryabin, and Sujoy Sinha Roy. ModHE: Modular homomorphic encryption using module lattices potentials and limitations. *IACR TCHES*, 2024(1):527–562, 2024.

Mic04. Daniele Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions from worst-case complexity assumptions. Cryptology ePrint Archive, Report 2004/286, 2004.

Mic18. Daniele Micciancio. On the hardness of learning with errors with binary secrets. Cryptology ePrint Archive, Report 2018/988, 2018.

MKMS22. Jose Maria Bermudo Mera, Angshuman Karmakar, Tilen Marc, and Azam Soleimanian. Efficient lattice-based inner-product functional encryption. In Goichiro Hanaoka, Junji Shikata, and Yohei Watanabe, editors, *PKC 2022, Part II*, volume 13178 of *LNCS*, pages 163–193. Springer, Cham, March 2022.

MM11. Daniele Micciancio and Petros Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In Phillip Rogaway, editor, *CRYPTO 2011*, volume 6841 of *LNCS*, pages 465–484. Springer, Berlin, Heidelberg, August 2011.

MP13. Daniele Micciancio and Chris Peikert. Hardness of SIS and LWE with small parameters. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 21–39. Springer, Berlin, Heidelberg, August 2013.

MR04. Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th FOCS*, pages 372–381. IEEE Computer Society Press, October 2004.

MS23. Daniele Micciancio and Adam Suhl. Simulation-secure threshold PKE from LWE with polynomial modulus. Cryptology ePrint Archive, Report 2023/1728, 2023.

PP19. Chris Peikert and Zachary Pepin. Algebraically structured LWE, revisited. In Dennis Hofheinz and Alon Rosen, editors, *TCC 2019, Part I*, volume 11891 of *LNCS*, pages 1–23. Springer, Cham, December 2019.

PR07. Chris Peikert and Alon Rosen. Lattices that admit logarithmic worst-case to average-case connection factors. In David S. Johnson and Uriel Feige, editors, *39th ACM STOC*, pages 478–487. ACM Press, June 2007.

PS24. Alain Passelègue and Damien Stehlé. Low communication threshold fully homomorphic encryption. In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024, Part I*, volume 15484 of *LNCS*, pages 297–329. Springer, Singapore, December 2024.

Reg05. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 84–93. ACM Press, May 2005.

SSTX09. Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In Mitsuru Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 617–635. Springer, Berlin, Heidelberg, December 2009.

STA20. Chao Sun, Mehdi Tibouchi, and Masayuki Abe. Revisiting the hardness of binary error LWE. In Joseph K. Liu and Hui Cui, editors, *ACISP 20*, volume 12248 of *LNCS*, pages 425–444. Springer, Cham, November / December 2020.

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

WLL24. Zhedong Wang, Qiqi Lai, and Feng-Hao Liu. Ring/module learning with errors under linear leakage - hardness and applications. In Qiang Tang and Vanessa Teague, editors, *PKC 2024, Part II*, volume 14602 of *LNCS*, pages 275–304. Springer, Cham, April 2024.

## A Application: Hardness of Known-Covariance M-LWE

We explain in details how to instantiate our result to obtain the hardness of the Known-Covariance M-LWE problem introduced in [MS23]. To show it is possible to have hard parameter regimes, we propose one parameter set for which our reduction reaches a security target $\lambda = 128$. Other parameterization could likely lead to better parameters. We also note that the parameters we propose are much larger than what is used in practical applications. In particular, as our result does not cover the ring setting $d = 1$, we analyze the hardness of the module version only.

As explained in Section 6.2.2, the Known-Covariance M-LWE problem asks to recover $\mathbf{x} \in R^{2d}$ given $\mathbf{A} \hookleftarrow U(R_q^{d\times d})$, $\mathbf{b} = [\mathbf{A} \mid \mathbf{I}_d]\mathbf{x} \bmod qR$ and $\mathbf{x}^*\mathbf{x}$. We also showed the average conditional min-entopy $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{x}^*\mathbf{x})$ was lower-bounded by $H_\infty(\mathbf{x}) - \frac{n}{2}\log_2(2\max_{\mathbf{x}\in X}\|\mathbf{x}\|_2^2 + 1)^{n/2}$, where $n$ is the ring degree, and $X$ the support of $\mathbf{x}$. We now need to plug these in our results of Theorem 4.1 and 5.1.

We make similar choices as their integer construction based on Known-Norm LWE, namely with a discrete Gaussian distribution for $\mathbf{x}$ and with $\mathbf{A}$ being a square matrix (secret and error having the same dimension). We let $R$ be the power-of-two cyclotomic ring of degree $n$, and define $m = 2d$. We also consider $\mathbf{x}$ being drawn from $\mathcal{D}_{R^m,s}$.

**Min-Entropy.** For simplicity, we assume the discrete Gaussian is truncated in $\ell_2$ norm. We choose the tailcut bound so that it is verified with overwhelming probability. As a result, both versions are equivalent up to a negligible reduction loss. More concretely, we define $X = \tau^{-1}(\mathcal{B}_{2,nm}(t_2 s\sqrt{nm}) \cap \mathbb{Z}^{nm})$ for $t_2 \geq 1/\sqrt{2\pi}$ that is such that $(t_2\sqrt{2\pi e}e^{-\pi t_2^2})^{nm} = 2^{-\lambda}$. The probability mass function of this truncated Gaussian is then $\frac{\rho_s}{\rho_s(X)} \cdot \mathbf{1}_X$, where $\mathbf{1}_X$ is the indicator function of $X$. The union bound combined with Lemma 2.3 then gives

$$\frac{\rho_s(X)}{\rho_s(\mathbb{Z}^{nm})} = \mathbb{P}_{\mathbf{x}\sim\mathcal{D}_{R^m,s}}[\|\mathbf{x}\|_2 \leq t_2 s\sqrt{nm}] \geq 1 - 2^{-\lambda}.$$

We can then compute $H_\infty(\mathbf{x})$ and the lower bound on $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{x}^*\mathbf{x})$. First, we have that $H_\infty(\mathbf{x}) = -\log_2 \max_{\mathbf{x}\in X} \rho_s(\mathbf{x})/\rho_s(X) = \log_2 \rho_s(X)$. By the inequality above, we then obtain $H_\infty(\mathbf{x}) \geq \log_2(1 - 2^{-\lambda}) + \log_2 \rho_s(\mathbb{Z}^{nm})$. By the Poisson summation formula, we directly get $\rho_s(\mathbb{Z}^{nm}) = s^{nm}\rho_{1/s}(\mathbb{Z}^{nm}) \geq s^{nm}$. Hence, $H_\infty(\mathbf{x}) \geq \log_2(1 - 2^{-\lambda}) + nm\log_2 s$. Then, by definition of $X$, we get $\max_{\mathbf{x}\in X}\|\mathbf{x}\|_2^2 = \lfloor t_2^2 s^2 nm \rfloor$. Overall, it yields

$$\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{x}^*\mathbf{x}) \geq nm\log_2 s - \frac{n}{2}\log_2\left(2\lfloor t_2^2 s^2 nm\rfloor + 1\right) + \log_2(1 - 2^{-\lambda}).$$

**Uninvertibility.** We consider $\mathcal{D} = \mathcal{D}_{R^\ell,\sigma}$ truncated in $\ell_2$ norm at $t\sigma\sqrt{n\ell}$. In a similar fashion as above, we take a tailcut $t$ so that $(t\sqrt{2\pi e}e^{-\pi t^2})^{n\ell} = 2^{-\lambda}$. We then study the $(1,2)$-norm of $M_\tau([\mathbf{I}_\ell \mid \mathbf{Y}])$ where the $(m-\ell)$ columns of $\mathbf{Y}$ are drawn from $\mathcal{D}$. It holds that

$$\|M_\tau([\mathbf{I}_\ell \mid \mathbf{Y}])\|_{1,2} = \max\left(1, \|M_\tau(\mathbf{Y})\|_{1,2}\right) \leq t\sigma\sqrt{n\ell}.$$

Indeed, in the power-of-two cyclotomic ring, we have that all the columns of $M_\tau(a)$ for $a \in R$ have the same $\ell_2$ norm which equals $\|a\|_2$. We then have $\|M_\tau(\mathbf{Y})\|_{1,2} = \max_{i\in[m-\ell]}\|\mathbf{Y}\mathbf{e}_i\|_2 \leq t\sigma\sqrt{n\ell}$, where $\mathbf{e}_i$ has a 1 at entry $i$ and 0 everywhere else.

It allows us to define $B_{1,2} = t\sigma\sqrt{n\ell}$ and $p_{1,2} = 0$. Also, by definition of $X$, we can also define $B_1 = t_2 snm$ as $\max_{\mathbf{x}}\|\mathbf{x}\|_1 \leq \max_{\mathbf{x}}\sqrt{nm}\|\mathbf{x}\|_2 = t_2 snm$. Noting that $\left|\mathcal{B}_{2,N}(r) \cap \mathbb{Z}^N\right| \leq \mathrm{Vol}(\mathcal{B}_{2,N}(r + \sqrt{N}/2))$, Lemma 4.1 yields the $\varepsilon'_{\mathrm{inv}}$-uninvertibility needed for Theorem 4.1, with

$$\varepsilon'_{\mathrm{inv}} = 2^{-\widetilde{H_\infty}(\mathbf{x}|\mathbf{x}^*\mathbf{x})}\frac{\left(\sqrt{\pi}(t_2 ts\sigma nm\sqrt{n\ell} + \sqrt{n\ell}/2)\right)^{n\ell}}{\Gamma(n\ell/2+1)}$$

$$\leq \frac{1}{1-2^{-\lambda}} \cdot \frac{\left(2\lfloor t_2^2 s^2 nm \rfloor + 1\right)^{n/2}}{s^{nm}} \cdot \frac{\left(\sqrt{\pi}(t_2 ts\sigma nm\sqrt{n\ell} + \sqrt{n\ell}/2)\right)^{n\ell}}{\Gamma(n\ell/2+1)},$$

where the inequality stems from that on $\widetilde{H_\infty}(\mathbf{x} \mid \mathbf{x}^*\mathbf{x})$.

**Hardness of Known-Covariance M-LWE.** We can now apply Theorem 4.1. Defining $\beta_2 = 2t_2 s\sqrt{nm}$, and assuming $q$ satisfy the correct number-theoretic requirements, we obtain the $\varepsilon$-hardness of $\text{M-LWE}_{n,d,m,q,U(R_q^d),\mathcal{X}}$, with $\varepsilon \approx (d-k)\varepsilon_{\text{M-LWE}} + \varepsilon_{\text{M-SIS}} + \varepsilon'_{\text{inv}}$, where $\varepsilon_{\text{M-SIS}}$ is the hardness bound of $\text{M-SIS}^2_{n,d,m,q,\beta_2}$ and $\varepsilon_{\text{M-LWE}}$ that of $\text{M-LWE}_{n,k,\ell,q,U(R_q^k),\mathcal{D}}$. Theorem 5.1 then ensures that the problem $\text{pHNF-M-LWE}_{n,d,d,q,\mathcal{X}}$ is $\varepsilon/(1-\delta'(m,d))$-hard. Notice that the conditional distribution $\mathcal{X}$ is invariant under permutation due to the fact that we have $(\mathbf{Px})^*(\mathbf{Px}) = \mathbf{x}^*\mathbf{P}^T\mathbf{Px} = \mathbf{x}^*\mathbf{x}$ for any permutation matrix $\mathbf{P}$. It then holds $\text{pHNF-M-LWE}_{n,d,d,q,\mathcal{X}}$ corresponds to the standard HNF formulation $\text{HNF-M-LWE}_{n,d,d,q,\mathcal{X}}$. The latter is then exactly the Known-Covariance M-LWE problem. Plugging the expression of $\varepsilon'_{\text{inv}}$ and using the fact that $m = 2d$ (and thus $\ell = d+k$) entails that Known-Covariance $\text{M-LWE}_{n,d,d,q,\mathcal{D}_{R^{2d},s}}$ is $\varepsilon'$-hard with

$$\begin{aligned}
\varepsilon' &= \frac{\delta(2d,d)}{1-\delta'(2d,d)} + \frac{(d-k)\left(2\delta(d+k,d) + \frac{\varepsilon_{\text{M-LWE}}}{1-\delta(d+k,k)}\right) + \varepsilon'_{\text{inv}} + \varepsilon_{\text{M-SIS}}}{(1-\delta'(2d,d))(1-\delta(2d,d))} \\
&\approx (d-k)\varepsilon_{\text{M-LWE}} + \varepsilon_{\text{M-SIS}} \\
&\quad + \left( \frac{\left(\sqrt{\pi}(2t_2 ts\sigma n^{3/2} d\sqrt{d+k} + \sqrt{n(d+k)}/2)\right)^{d+k} \sqrt{2\lfloor 2t_2^2 s^2 nd \rfloor + 1}}{s^{2d} \cdot \Gamma(n(d+k)/2+1)^{1/n}} \right)^n .
\end{aligned}$$

**Parameter Selection.** We now look at the last term (corresponding to $\varepsilon'_{\text{inv}}$) which we write as $\epsilon^n$. Assuming all parameters but $d$ are fixed, we asymptotically have that the $\epsilon$ is approximately $(2\sqrt{2\pi e}tt_2\sigma nd)^{d+k}\sqrt{d}/s^{d-k}$. Thence, if $s$ is too small compared to $d$, this term will be larger than 1 and lead to a vacuous bound. We then decide to choose $s = \Omega(2\sqrt{2\pi e}tt_2\sigma nd)$ and then search for the smallest $d$ that ensures $\epsilon \leq 2^{-\lambda/n}$. We propose an example parameter set with $s = 4\sqrt{2\pi e}tt_2\sigma nd$, but we note that a larger $s$ would lead to a smaller rank $d$. We indeed clearly see that as $s$ gets larger, the minimal rank $d$ gets smaller. There is then a trade-off between $s$ (and thus the modulus $q$) and the rank $d$. Once $d$ (and $s$) are found, we can set the modulus $q$ so that $\text{M-SIS}_{n,d,2d,q,2t_2 s\sqrt{nm}}$ and $\text{M-LWE}_{n,k,d+k,q,U(R_q^k),\mathcal{D}}$ are hard. The modulus $q$ is chosen as a prime that splits in $n$ prime ideals which is sufficient for the $\delta(\cdot,\cdot)$ to be negligible (these values become smaller when $q$ has fewer and fewer prime ideals factors of $\langle q \rangle$). We note that a larger $s$ will necessitate a larger $q$, which will decrease the security of the starting M-LWE assumption. One may need to increase $k$ to compensate. As $k$ impacts $\varepsilon'$ more marginally, it should still be possible to find parameters in the same order of magnitude. We give in Table A.1 example parameters, and detail in Table A.2 the values of $\delta(\cdot,\cdot), \delta'(\cdot,\cdot), \varepsilon'_{\text{inv}}, \varepsilon_{\text{M-LWE}}, \varepsilon_{\text{M-SIS}}, \varepsilon'$. We provide an estimation script[10] written in Sage which finds the smallest $d$ and estimates the final hardness.

| $n$ | $k$ | $q$ | $\sigma$ | $t$ | $d$ | $s$ | $t_2$ |
|---|---|---|---|---|---|---|---|
| 256 | 5 | 532843009 | 6 | 0.41463 | 222 | 958593.521 | 0.41014 |

**Table A.1.** Example parameter set for the hardness of Known-Covariance $\text{M-LWE}_{n,d,d,q,\mathcal{D}_{R^{2d},s}}$.

We aimed for a reasonably small asymptotic constant when setting $s$, which then leads to a high rank of $d = 222$. Albeit far from practical, it shows that there is nothing fundamentally weak in the Known-Covariance M-LWE assumption as we are able to find parameters for which it is hard. We also insist that these

[10] https://github.com/mlwegeneraldistributions/mlwe-general-distributions

| $\delta(2d,d)$ | $\delta(d+k,d)$ | $\delta(d+k,k)$ | $\delta'(2d,d)$ | $\varepsilon_{\text{M-LWE}}$ | $\varepsilon_{\text{M-SIS}}$ | $\varepsilon'_{\text{inv}}$ | $\varepsilon'$ |
|---|---|---|---|---|---|---|---|
| $2^{-6448}$ | $2^{-158}$ | $2^{-6454}$ | $2^{-26}$ | $2^{-131.0}$ | $2^{-14131.8}$ | $2^{-216.7}$ | $2^{-123.3}$ |

**Table A.2.** Upper bounds on the losses and hardness estimates for the hardness of Known-Covariance $\text{M-LWE}_{n,d,d,q,\mathcal{D}_{R^{2d},s}}$

parameters are large because of our proof methodology. Indeed, the hardness for entropic noise distribution comes with limitations on the number of samples $m$. As shown in [BJRW23] for uniform noise (and coherently with unstructured versions [MP13, STA20]), the number of samples is limited to $m = d(1 + o(1))$. When adopting a fine-grained analysis of the parameter constraints instead of asymptotic, one can still reach $m = 2d$ but at the expense of other parameters (like the error bound) which grow exponentially with $m - d$. The number of samples of $m = 2d$ required for the HNF transform is then drastically impacting the overall parameters.

## B Application: Hardness of M-LWE with Sparse Error

In order to demonstrate the natural requirements on the number of samples entailed by our proof, we focus in this section on the hardness of M-LWE with a sparse error distribution. These regimes are indeed prone to efficient (even polynomial-time) attacks when the number of samples $m$ given is too large and the density of the error distribution is too low. We can still instantiate our result to cover very low density regimes, but the rank (or equivalently lattice dimension) needs to be extremely large to ensure hardness. We note that the sparse error regime was studied from a hardness perspective in [STA20] where they looked at an error distribution where all the coefficients were taken from a Bernoulli distribution of parameter $p$. The uniform case of [MP13] corresponds to $p = 1/2$, but they showed it could be adapted to lower values of $p$, at the expense of limitting more and more the number of samples. In our case, we take a slightly different approach by looking at ternary errors with a fixed Hamming weight $w$. We thus consider $S_1 = \tau^{-1}(\{-1, 0, 1\}^n)$, and define $X = \{\mathbf{x} \in S_1^m : \|\mathbf{x}\|_1 = w\}$. As we wish to establish a correlation between the density and the required lattice dimension, we equivalently choose the density $\rho$ and set $w = \lfloor \rho nm \rfloor$. We then consider $\mathcal{X} = U(X)$, meaning $H_\infty(\mathcal{X}) = |X|^{-1} = 2^{-w}\binom{nm}{w}^{-1}$.

We again consider $\mathcal{D} = \mathcal{D}_{R^\ell,\sigma}$ truncated in $\ell_2$ norm at $t\sigma\sqrt{n\ell}$ as in Appendix A. The $(1, 2)$-norm of $M_\tau([\mathbf{I}_\ell \mid \mathbf{Y}])$ is then bounded by $B_{1,2} = t\sigma\sqrt{n\ell}$, and we have $p_{1,2} = 0$. Also, by definition of $X$, we directly have $B_1 = w$. Lemma 4.1 then yields the $\varepsilon'_{\text{inv}}$-uninvertibility with

$$\begin{aligned}\varepsilon'_{\text{inv}} &= \frac{1}{2^w\binom{nm}{w}} \cdot \frac{\left(\sqrt{\pi}(wt\sigma\sqrt{n\ell} + \sqrt{n\ell}/2)\right)^{n\ell}}{\Gamma(n\ell/2+1)} \\ &= \frac{1}{2^{\lfloor \rho nm \rfloor}\binom{nm}{\lfloor \rho nm \rfloor}} \cdot \frac{\left(\sqrt{\pi n\ell}(t\sigma\lfloor \rho nm \rfloor + 1/2)\right)^{n\ell}}{\Gamma(n\ell/2+1)},\end{aligned}$$

The second-preimage resistance also follows from $\text{M-SIS}^1_{n,m-d,m,q,2w}$. Since M-SIS in $\ell_1$ norm is unusual, we can instead rely on $\text{M-SIS}^2_{n,m-d,m,q,2\sqrt{w}}$ as we directly get $\max_{\mathbf{x},\mathbf{x}'\in X}\|\mathbf{x} - \mathbf{x}'\|_2 = 2\max_{\mathbf{x}\in X}\|\mathbf{x}\|_2 = 2\sqrt{w}$. Theorem 4.1 then gives the $\varepsilon$-hardness of $\text{M-LWE}_{n,d,m,q,U(R_q^d),\mathcal{X}}$ with $\varepsilon \approx (d-k)\varepsilon_{\text{M-LWE}} + \varepsilon_{\text{M-SIS}} + \varepsilon'_{\text{inv}}$.

We can now fix $n, k, \ell, \sigma, q$ and, for variable densities $\rho$, we search for the minimal ranks $d$ achieving $\lambda = 128$ bits of security. Fixing $k$ and $\ell$ naturally makes $m = d + \ell - k$ uniquely determined by $d$. We give a few examples of the minimal rank $d$ for different densities in Table B.1, fixing $n = 256$, $k = 2$, $\ell = 4$, $\sigma = 5$ and $q = 2113$. These examples are computed with our estimation script[11] written in Sage.

[11] https://github.com/mlwegeneraldistributions/mlwe-general-distributions

| $\rho$ | 0.5 | 0.1 | 0.05 | 0.01 | 0.001 | 0.0001 | 0.00001 |
|---|---|---|---|---|---|---|---|
| $d$ | 41 | 104 | 173 | 624 | 4416 | 33866 | 273045 |
| $\varepsilon$ | $2^{-138.3}$ | $2^{-136.9}$ | $2^{-136.2}$ | $2^{-131.2}$ | $2^{-131.1}$ | $2^{-128.3}$ | $2^{-125.4}$ |

**Table B.1.** Example parameters for the minimal module rank $d$ for different densities $\rho$. The reported value $\varepsilon$ is the obtained hardness bound for M-LWE$_{n,d,m,q,U(R_q^d),\mathcal{X}}$. The deterioration in security comes from the hybrid argument term $(d-k)\varepsilon_{\text{M-LWE}}$ which grows with $d$.

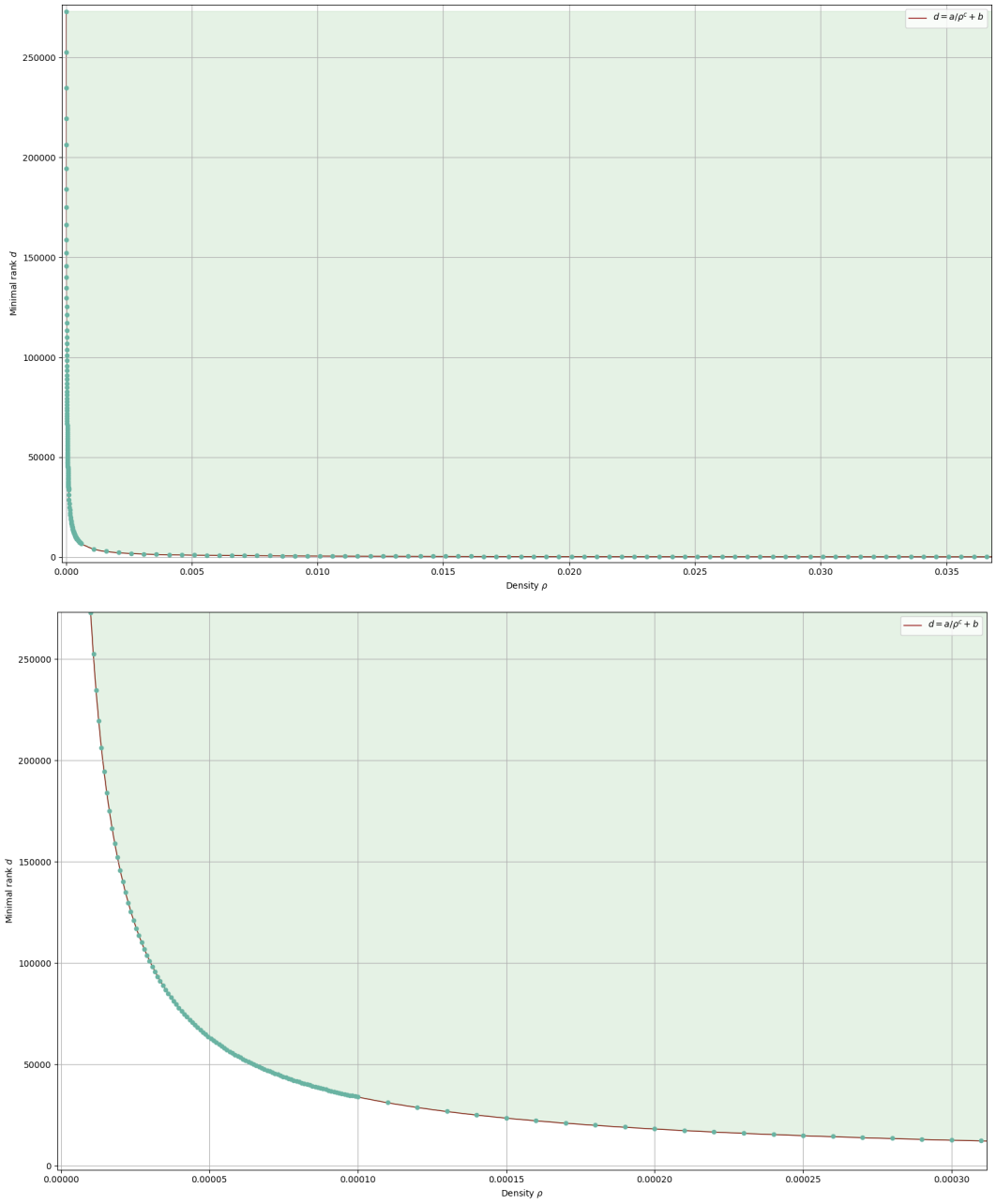


**Fig. B.1.** Interpolation of the minimal rank $d$ as a function of the density $\rho$. The interpolation curve is $d = 7.932694 \cdot \rho^{-0.907169} + 43.677687$. The colored area corresponds to the zone where M-LWE$_{n,d,m,q,U(R_q^d),\mathcal{X}}$ is hard.

Although the parameters are far from practical for very low densities, we observe a somewhat inverse relation between $\rho$ and $d$, i.e., $d \approx a/\rho^b + c$ especially for low values of $\rho$, as depicted in Figure B.1. Interpolation for densities ranging between $\rho = 10^{-5}$ and $\rho = 1/2$ gives $(a, b, c) \approx (7.933, 0.907, 43.678)$. The interpolation is realized with our Sage script mentioned above. It then seems that for a given dimension $d$, the M-LWE problem with $m = d + o(1)$ samples and sparse ternary errors remains hard for densities $\rho > (a/(d-c))^{1/b}$. We however leave further experiments for future work.

# C Examples of $(\alpha, \beta)$-norms for Uninvertibility

We here give a few other examples of choices of $(\alpha, \beta)$ pairs to instantiate Lemma 4.1 (and thus Theorem 4.1) with.

## C.1 Example in $(2, 2)$-norm.

Our result encompasses the result of [BJRW23] in the $(2, 2)$-norm, i.e., the spectral norm. We indeed observe that $\|M_\tau([\mathbf{I}|\mathbf{Y}])\|_2 = (1 + \|\mathbf{Y}\|_2^2)^{1/2}$ and then bound $\|\mathbf{Y}\|_2$ using results from random matrix theory. Their analysis was carried for a discrete Gaussian distribution $\mathcal{D}$ in the Minkowski embedding (so as to plug their reduction to worst-case to average-case reductions), but the approach could be extended to any sub-Gaussian distribution using [Ver12, Thm. 4.4.5]. The bound depends on the sub-Gaussian norm $\|\mathcal{D}\|_{\psi_2}$. For example, choosing the coefficients of $\mathbf{Y}$ from a centered binomial distribution gives a sub-Gaussian norm of $1/\sqrt{\ln 3}$ for each coefficient. The ring setting however invalidates the entries independence condition necessary for [Ver12, Thm. 4.4.5]. One can instead bound the norm of $\mathbf{Y}$ in each field embedding $\sigma_k$ (which then depends on the distortion between the coefficient and canonical embeddings) and apply the union bound. The latter gives a proven bound on $\|M_\tau(\mathbf{Y})\|_2$ using [BJRW23, Lem. 2.3], but is unfortunately very loose. Concrete schemes instead rely on heuristically derived bounds.

Again taking the example of centered binomial coefficients gives a (heuristic) bound on $\|M_\tau([\mathbf{I}_\ell \mid \mathbf{Y}]\|_2$ of $B_{2,2} = \sqrt{1 + (\sqrt{n\ell} + \sqrt{n(m-\ell)} + t)^2/\ln(3)}$ for a chosen tailcut $t$ defining $p_{2,2} = 2e^{-t^2}$. Then, if the radius $r = B_2 \cdot B_{2,2}$ is large enough, then the number of integer points within the $\ell_2$ ball of radius $r$ is well approximated by its volume $(\sqrt{\pi}r)^{n\ell}/\Gamma(n\ell/2+1)$. Otherwise, one can use the more conservative bound of $(\sqrt{\pi}(r + \sqrt{n\ell}/2))^{n\ell}/\Gamma(n\ell/2+1)$.

## C.2 Example in $(1, 1)$-norm.

As mentioned, other distributions $\mathcal{X}$ and $\mathcal{D}$ may be better analyzed in different norms than the $(2, 2)$-norm. Another natural regime for M-LWE to improve efficiency is that of sparse errors. In this situation, the $\ell_1$ norm seems the more natural metric. As described in Appendix B, one can choose the $(1, 2)$-norm with $\mathcal{D} = \mathcal{D}_{R^\ell,\sigma}$. Another choice could be the $(1, 1)$-norm for which closed-form formulae can be derived. Pursuing this choice, one could opt for a distribution $\mathcal{D}$ that is reasonable in terms of hardness, but better suited for the $(1, 1)$-norm.

Let us consider $\mathcal{D}$ to be the centered binomial distribution with parameter 1, i.e., each coefficient of every vector entry is 0 with probability $1/2$ and $\pm 1$ each with probability $1/4$. We first have $\|M_\tau([\mathbf{I}_\ell \mid \mathbf{Y}])\|_{1,1} = \max(1, \|M_\tau(\mathbf{Y})\|_{1,1})$. In power-of-two cyclotomics, it also holds that $\|M_\tau(\mathbf{Y})\|_{1,1} = \max_{j\in[m-\ell]}\|\tau(\mathbf{y}_j)\|_1$. As $\mathbf{y}_j$ is drawn from $\mathcal{D}$, we can easily see that $\|\tau(\mathbf{y}_j)\|_1$ follows the (uncentered) binomial distribution of parameter $(n\ell, 1/2)$. As a result,

$$\forall K \in [0, n\ell], \mathbb{P}_{\mathbf{y}\sim\mathcal{D}}[\|\mathbf{y}\|_1 \le K] = \sum_{i=0}^{K} \binom{n\ell}{i} 2^{-n\ell}.$$

Then, each column of $\mathbf{Y}$ being sampled independently from the others, we have

$$\forall K \in [1, n\ell], \mathbb{P}_{\mathbf{Y}\sim\mathcal{D}^{m-\ell}}[\|M_\tau([\mathbf{I}_\ell \mid \mathbf{Y}])\|_{1,1} > K] = 1 - \frac{1}{2^{n\ell(m-\ell)}} \left(\sum_{i=0}^{K} \binom{n\ell}{i}\right)^{m-\ell}.$$

We can then define $B_{1,1} = K$ and $p_{1,1} = 1 - 2^{-n\ell(m-\ell)}(\sum_{i=0}^{K}\binom{n\ell}{i})^{m-\ell}$ for a tailcut parameter $K$. As described in Appendix B, choosing $\mathcal{X} = U(\{\mathbf{x} \in S_1^m : \|\mathbf{x}\|_1 = w\})$ directly yields $B_1 = w$. We then obtain

$$\varepsilon'_{\text{inv}} = \frac{1}{2^w\binom{nm}{w}} \sum_{i=0}^{\min(Kw,n\ell)} \binom{n\ell}{i}\binom{Kw}{i}2^i + \left(1 - \frac{1}{2^{n\ell(m-\ell)}}\left(\sum_{i=0}^{K}\binom{n\ell}{i}\right)^{m-\ell}\right).$$

For every given $(n, \ell, m, w)$, one can then optimize over $K$ to minimize this expression. We observe a similar behavior between the density $\rho = w/nm$ and the minimal rank $d$ that ensures $\varepsilon'_{\text{inv}} \leq 2^{-\lambda}$, namely a somewhat inverse relation. However, because the starting assumption is slightly weaker (because of using centered binomial instead of discrete Gaussian), the example presented in Appendix B gives tighter parameters.

### C.3 Example in $(2, \infty)$-norm.

Keeping $\mathcal{D}$ to be a discrete Gaussian distribution, we can study the $(2, \infty)$-norm as well by looking at the rows of $[\mathbf{I}_\ell|\mathbf{Y}]$. Using the Gaussian tail bound on the rows then yields $B_{2,\infty} = \sqrt{1 + t^2s^2n(m-\ell)}$, and $p_{2,\infty} = \ell(t\sqrt{2\pi e}e^{-\pi t^2})^{n(m-\ell)}$. If $X$ is contained in the $\ell_2$-ball of radius $B_2$, then we can set

$$\varepsilon'_{\text{inv}} = 2^{-H_\infty(\mathcal{X})}\left((2\left\lfloor B_2\sqrt{1 + t^2s^2n(m-\ell)}\right\rfloor + 1)^{n\ell} + |X|\ell(t\sqrt{2\pi e}e^{-\pi t^2})^{n(m-\ell)}\right).$$