Lossiness and Entropic Hardness for Ring-LWE

Zvika Brakerski[1]* and Nico Döttling[2]**

[1] Weizmann Institute of Science
[2] CISPA Helmholtz Center for Information Security

Abstract. The hardness of the Ring Learning with Errors problem (RLWE) is a central building block for efficiency-oriented lattice-based cryptography. Many applications use an "entropic" variant of the problem where the so-called "secret" is not distributed uniformly as prescribed but instead comes from some distribution with sufficient min-entropy. However, the hardness of the entropic variant has not been substantiated thus far.

For standard LWE (not over rings) entropic results are known, using a "lossiness approach" but it was not known how to adapt this approach to the ring setting. In this work we present the first such results, where entropic security is established either under RLWE or under the Decisional Small Polynomial Ratio (DSPR) assumption which is a mild variant of the NTRU assumption.

In the context of general entropic distributions, our results in the ring setting essentially match the known lower bounds (Bolboceanu et al., Asiacrypt 2019; Brakerski and Döttling, Eurocrypt 2020).

1 Introduction

Lyubashevsky, Peikert and Regev [16, 17] introduced the Ring Learning with Errors (RLWE) problem as a structured variant of the celebrated LWE problem [24]. RLWE (and similar variants such as ideal/polynomial LWE [28]) are by now an indispensable tool for constructing efficient lattice-based cryptographic primitives, such as public-key encryption, key agreement and signatures. It is appealing to use RLWE-based cryptographic primitives since they are usually more succinct and efficient than their non-ring counterparts. Translating a cryptographic construction from LWE to RLWE is often straightforward, and indeed many LWE based constructions have RLWE counterparts that enjoy a higher level of efficiency (at the cost of only enjoying hardness respective to a special class of lattices instead of all lattices as in LWE).

The focus of this work is *entropic hardness*, which is an important property of LWE-based cryptography [2, 4, 6, 8, 18] that so far resisted translation to the

* Supported by the Binational Science Foundation (Grant No. 2016726), and by the European Union Horizon 2020 Research and Innovation Program via ERC Project REACT (Grant 756482) and via Project PROMETHEUS (Grant 780701).

** This work is partially funded by the Helmholtz Association within the project "Trustworthy Federated Data Analytics" (TFDA) (funding number ZT-I-OO1 4)

RLWE regime. Entropic hardness is the property of the LWE problem (and hopefully also RLWE) to remain hard even when the so called "LWE secret" is not sampled from the prescribed distribution, but instead is sampled from some distribution with sufficient min-entropy. This is relevant in the context of key-leakage (see e.g. [12] for a survey), and in a number of other applications which use RLWE with a key that is not sampled according to the prescribed distribution. These include implementations of fully homomorphic encryption such as [5, 9, 10, 25] and even some of the candidates in the NIST post-quantum cryptography contest [19].

The question of entropic security for RLWE is therefore highly motivated. Nevertheless, very little was known about its security prior to this work. The only work we are aware of in this context is by Bolboceanu et al. [3], which introduced a non-standard assumption that they call HLBDD. They prove the hardness of entropic RLWE for a class of distributions that they call k-wise independent, based on the hardness of HLBDD and standard RLWE. This solution has a number of drawbacks in not addressing general entropic distributions, being applicable only in certain rings (it requires that the ring has CRT representation) and making a new assumption.

One would have hoped that it would be possible to use similar methods to those used in the context of LWE also for RLWE. After all, the structure of the problems is very similar. However, the same barrier seemed to have stopped all prior attempts. In a nutshell, it is the failure to find a proper analog *lossiness argument* in the ring setting. This term refers to a family of proof techniques that underlie all known entropic hardness results [2, 4, 6, 8, 18]. We explain this barrier in more detail below.

We recall that in standard LWE, an instance is composed of a random matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times N}$ with $N \gg n$, and a vector $\mathbf{y} = \mathbf{s}\mathbf{A} + \mathbf{e}$, where $\mathbf{s}$ is the "LWE secret" and $\mathbf{e}$ is a noise vector (usually sampled from a Gaussian). The goal is to find the vector $\mathbf{s}$, or in the *decisional* version of the problem to distinguish $(\mathbf{A}, \mathbf{y})$ from uniform. The RLWE problem is a structured variant of the above, usually defined using elements from the ring of integers of an algebraic number field (and its dual). For the purpose of this work, it will be instructive to consider an equivalent (and in fact more general) formulation of RLWE that does not refer to algebraic number theory at all and takes great resemblance to the above LWE description. Let us rewrite the above LWE instance as follows, consider the case where $N = n \cdot m$. We can break the matrix $\mathbf{A}$ into square blocks s.t. $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_m]$ and consider the LWE instance as a sequence of blocks of the form $\{(\mathbf{A}_i, \mathbf{y}_i = \mathbf{s}\mathbf{A}_i + \mathbf{e}_i)\}_{i=1}^m$. RLWE instances can be presented in the same way, except the matrices $\mathbf{A}_i$ are no longer uniform, but instead are drawn from a distribution over *structured* matrices.[3] Throughout this work we will attempt to state our results and techniques in terms of this *Structured LWE* formulation as much as possible, without specifying the exact structure of the

[3] Essentially this structure represents the multiplication of an element a from the (dual) of a ring of integers by an element from the ring of integers.

matrices $\mathbf{A}_i$, and the instantiations to the special case of number fields will follow as straightforward corollaries.

A lossiness argument for LWE hinges on the observation that the entropic LWE distribution is computationally indistinguishable from one where the matrices $\mathbf{A}_i$ are not uniform, but instead are distributed as $\mathbf{A}_i = \mathbf{B} \cdot \mathbf{C}_i + \mathbf{F}_i$, where $\mathbf{B} \in \mathbb{Z}_q^{n \times k}$ (note that the same $\mathbf{B}$ is used for all i) for $k \ll n$, $\mathbf{C}_i \in \mathbb{Z}_q^{k \times n}$, and $\mathbf{F}_i$ is small noise. Indistinguishability is established by decisional LWE. This step makes the matrices $\mathbf{A}_i$ "close to low-rank". Furthermore, now $\mathbf{y}_i = \mathbf{s}\mathbf{A}_i = \mathbf{s}\mathbf{B}\mathbf{C}_i + \mathbf{s}\mathbf{F}_i + \mathbf{e}_i$. From this point the methods diverge somewhat, let us stick to the approach of [2, 4] that we follow in this paper. In these works, it is shown that *even information theoretically* $\mathbf{s}$ cannot be recovered, essentially because the adversary only has access to $\mathbf{s}\mathbf{B}$, which has dimension k and therefore does not contain much information, and to the terms $\mathbf{s}\mathbf{F}_i + \mathbf{e}_i$, where it is shown that the entropy in $\mathbf{e}_i$ masks the information about $\mathbf{s}$.

Trying to apply this argument in the structured LWE setting runs into a problem. The matrices $\mathbf{A}_i$ are no longer uniform but instead have some (efficiently verifiable) structure. Therefore, we need to find a distribution that is both indistinguishable from the structured $\mathbf{A}_i$ distribution, and has lossiness properties as above. In the context of the structure that is imposed by RLWE, this seems hopeless since the structure imposed by the ring does not allow the $\mathbf{A}_i$ matrices to be close to low-rank for general rings.[4] In this work we overcome this barrier.

1.1 Our Results

We present a new approach to achieve lossiness that generalizes the "closeness to low-rank" approach, but that can be applied for general RLWE (and possibly other structured LWE variants). Concretely, we observe that it suffices to replace $\mathbf{A}_i$ with a matrix whose span contains short vectors. That is, we will set $\mathbf{A}_i = \mathbf{H} \cdot \mathbf{Z}_i$, where $\mathbf{H}$ is an invertible matrix that is sampled once and used for all i, and the $\mathbf{Z}_i$ come from a distribution over low-norm matrices. The exact norm that we use depends on the underlying ring, but for the purposes of this overview, it suffices to think of $\mathbf{Z}_i$ as a matrix where all entries are shorter than some bound $\ll q$. We observe that the matrices $\mathbf{H} \cdot \mathbf{Z}_i$ are neither low rank nor close to low-rank, however they become close to low rank under a (common) *basis-change* corresponding to the matrix $\mathbf{H}$.[5] The level of lossiness will be dictated by the properties of $\mathbf{Z}_i$: the lower the norm of $\mathbf{Z}_i$, the more lossiness is obtained. We note that we can assume that $\mathbf{H}$ itself has a short inverse, that we denote by $\mathbf{Z}_0$ (we explain below that this does not actually impose an additional restriction).

[4] The work of [3] can be viewed as targeting a special case where this is possible, the case where the ring decomposes into a "CRT representation". This requires making a non-standard assumption like their HLBDD assumption which only applies to that special setting.

[5] In fact, under this basis change the matrices are even close to the 0 matrix, which has the lowest possible rank.

We show that this notion is both sufficient for proving entropic security, and that there exist such lossy distributions that are indistinguishable from uniform under standard assumptions.

The DSPR and NTRU Assumptions. We notice that the assumption as described above closely resembles the Decisional Small Polynomial Ratio (DSPR) [15] and NTRU assumptions [11]. Both assumptions are defined over polynomial rings and have very similar syntax. Both essentially assert that over some polynomial ring, there is a distribution over ring elements s.t. when sampling $f, g_1, \ldots, g_m$ from this distribution, it holds that $g_1/f, \ldots, g_m/f$ are jointly indistinguishable from a set of uniformly random ring elements. The NTRU cryptosystem uses a specific and very short distributions for f, g_i (over polynomials with $\{-1, 0, +1\}$ coefficients) and DSPR considers a Gaussian distribution (say with some Gaussian parameter γ) which will be easier to use.[6] The assumption becomes weaker as γ increases. As observed by Stehlé and Steinfeld [27], when the distributions become wide enough ($\gamma \gtrsim \sqrt{q}$), this assumption is actually implied by RLWE. For other parameter regimes, however, DSPR appears to provide a lower level of security compared to RLWE, at least with respect to state of the art attacks [13]. Translating the above into the structured LWE terminology, we can define $\mathbf{Z}_0$ as the matrix that corresponds to the operator of multiplying by f, and $\mathbf{Z}_i$ as the matrix that corresponds to multiplying by g_i. Intuitively the parameter γ can be thought of as a measure for the smallness of the elements in the $\mathbf{Z}_i$ matrices. We note that since the polynomial rings are commutative, the matrices $\mathbf{H}, \mathbf{Z}_0, \mathbf{Z}_i$ all commute with each other in the actual instantiation. However, we will not require this property.

Lastly, we point out that while RLWE enjoys a worst-case to average-case hardness reduction [16,23], such reduction is not known for NTRU/DSPR with small γ. Hence there is a tradeoff between the quality of the result obtained and the hardness of the assumption that we need to make.

Noise Lossiness and Entropic Security Under DSPR. We follow the approach of [4] and consider the notion of *noise lossiness* of a distribution of secrets $\mathcal{S}$, which is defined to be the conditional smooth min-entropy of a sample from $\mathcal{S}$ conditioned on learning its perturbation by Gaussian noise. Formally:

$$\nu_\sigma(\mathcal{S}) = \tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}) \ , \tag{1}$$

where $\mathbf{e}$ is Gaussian with parameter σ. We also recall that [4] show a general relation between noise lossiness and entropy

$$\nu_\sigma(\mathcal{S}) \gtrsim \tilde{H}_\infty(\mathcal{S}) - n\log(q/\sigma) \ . \tag{2}$$

We show that similarly to LWE, the hardness of entropic RLWE on a given secret distribution $\mathcal{S}$ is also related to its noise lossiness. We present our result in the context of RLWE in power-of-2 cyclotomic number fields, but the result is modular and applies to RLWE on any ring that has reasonable regularity condition. See also discussion below.

[6] We will consider Gaussians over the *canonical embedding* of ring elements into Euclidean space.

Theorem 1.1 (Informal). *Assume DSPR with parameter γ. Let $\mathcal{S}$ be a distribution s.t. for some σ' it holds that $\nu_{\sigma'}(\mathcal{S}) \gtrsim n\log(\gamma\mathsf{poly}(n)) + \omega(\log\lambda)$. Then Entropic RLWE in a power-of-2 cyclotomic with secret distribution $\mathcal{S}$ and Gaussian noise parameter $\sigma \approx \sigma' \cdot \mathsf{poly}(n) \cdot \sqrt{m}$ is hard.*

Plugging in Eq. (2), we get that for general entropic distributions we require average min-entropy of roughly $\tilde{H}_\infty(\mathcal{S}) \gtrsim n\log(q/\sigma) + O(n\log(nm\gamma)) + \omega(\log\lambda)$ in order to achieve entropic hardness. We note that better bounds on noise lossiness are known for "short" distributions, where the entropy requirement can go almost all the way down to $n\log\gamma$, which allows to show entropic hardness for many low-norm distributions but unfortunately is still insufficient for the widely used setting where the secret is chosen as a ring element with binary coefficients. Indeed, even in our results, we need to make stronger DSPR assumptions as we wish to deal with secrets of lower entropy. We believe that there is an inherent difficulty for proving hardness of such distributions without making extreme hardness assumptions.

Our results are stated in a rather general form. We present a notion of "structured LWE" problem, which captures standard LWE, RLWE and potentially other problems, and present lossiness results based on a "matrix DSPR" assumption, assuming that the matrix distributions in the DSPR instance satisfy some (mild) non-degeneracy conditions. Proving that the non-degeneracy conditions indeed hold is the only place where the specifics of the number field are required. The aforementioned [27] fortunately implies these required conditions for power-of-2 cyclotomic number fields. We believe that the proof can be generalized to other number fields (especially cyclotomics) but this would require essentially repeating the [27] proofs in more generality which we feel is tangent to the purpose of this work.

Since our paper is written in a modular manner, it suffices to simply prove the non-degeneracy conditions in Section 6 in order to obtain entropic hardness results for other variants of structured LWE, be it RLWE in other number fields (or a different embedding) or other forms of the problem completely.

To conclude, let us discuss the applicability of our techniques to the so called "module LWE" problem [1, 5, 14, 22]. Module-LWE interpolates between LWE and ring-LWE and is appealing in the practical context as it may offer superior security benefits over RLWE with minimal additional computational cost. Viewed as a structured LWE problem, in module-LWE the matrix $\mathbf{A}$ is simply a block matrix, where each block is an independent RLWE matrix. Our methods apply to such matrices as well, under a matrix DSPR assumption. We can instantiate matrix DSPR under RLWE-like assumptions, but we do not know of variants of this assumption that rely on module-LWE-like structures. A complete module LWE analog of our result would require introducing such an analog.

1.2 Our Techniques

As explained above, in order to prove security for entropic structured LWE, we rely on the assumption that we can replace the uniform $\mathbf{A}_i$ with $\mathbf{A}_i = \mathbf{H} \cdot \mathbf{Z}_i$,

where $\mathbf{Z}_i$ are short, and there exists $\mathbf{Z}_0$ (also short) s.t. $\mathbf{HZ}_0 = \mathbf{I} \pmod q$. We note that a survey by Peikert [21] uses a similar method when sketching a proof that the hardness of NTRU implies that of RLWE. Namely, replacing the a_i elements in RLWE samples with NTRU values, and arguing that the RLWE secret should become information-theoretically irrecoverable. One can view our method as putting together a rigorous variant of Peikert's arguments, and showing that it is possible to obtain lossiness for various entropic distributions.

We start by examining the distribution of the $\mathbf{y}_i$ values after substituting $\mathbf{A}_i = \mathbf{H} \cdot \mathbf{Z}_i$. We have

$$\mathbf{y}_i = \mathbf{sA}_i + \mathbf{e}_i = \mathbf{sH} \cdot \mathbf{Z}_i + \mathbf{e}_i \ .$$

We now take the approach of "flooding at the source" [4]. The idea is to "bring the noise closer to the secret" and show that all structured LWE blocks in fact depend on a noisy version of the secret, which allows to apply noise lossiness. Specifically, the technique that is used is Gaussian decomposition. Using Gaussian decomposition it is possible to show that if the $\mathbf{e}_i$ Gaussians are wide enough relative to the norm of the $\mathbf{Z}_i$ matrices, it is possible to find $\mathbf{e}$ s.t. for all i, $\mathbf{e}_i = \mathbf{eZ}_i + \mathbf{e}'_i$, where $\mathbf{e}$ and all $\mathbf{e}'_i$ are independent. This essentially follows from the covariance-additivity of Gaussian vectors, which can be carried over to discrete Gaussians as well.

Plugging this decomposed Gaussian into the equation for $\mathbf{y}_i$, we get

$$\mathbf{sH} \cdot \mathbf{Z}_i + \mathbf{eZ}_i + \mathbf{e}'_i = (\mathbf{sH} + \mathbf{e})\mathbf{Z}_i + \mathbf{e}'_i \ .$$

This implies that all information about $\mathbf{s}$ is captured in the term $\mathbf{sH}+\mathbf{e} \pmod q$.

We now note that already at this point we can derive a non-trivial entropic result. Let us denote $\mathbf{s}' = \mathbf{sH}$, and notice that since $\mathbf{H}$ is invertible, the entropy of $\mathbf{s}'$ is the same as that of $\mathbf{s}$ and recovering $\mathbf{s}$ is information-theoretically equivalent to recovering $\mathbf{s}'$. Now, essentially by definition, the probability of recovering $\mathbf{s}'$ is exactly captured by its noise lossiness. Specifically, if the noise lossiness is super-logarithmic then $\mathbf{s}'$ is not recoverable. Since we can relate noise lossiness to entropy (recall Eq. (2)) we have

$$\nu_\sigma(\mathbf{s}') \gtrsim \tilde{H}_\infty(\mathbf{s}') - n\log(q/\sigma) = \tilde{H}_\infty(\mathbf{s}) - n\log(q/\sigma) \ ,$$

where σ is the Gaussian parameter of $\mathbf{e}$. Therefore, so long as it holds that $\tilde{H}_\infty(\mathbf{s}) \gtrsim n\log(q/\sigma) + \omega(\lambda)$, then we have entropic security for RLWE with secret coming from the distribution of $\mathbf{s}$. This is indeed a non-trivial bound which may be useful in certain settings (e.g. when we only know the entropy of the distribution of $\mathbf{s}$ but do not know any other properties), but in many cases we would like to take into account additional properties of the distribution that reduce the large gap of $n\log(q/\sigma)$ between noise lossiness and entropy. However, in the current analysis we can say very little about the distribution of $\mathbf{s}'$ given the distribution of $\mathbf{s}$ (other than the entropy being preserved). We therefore proceed to show a connection that directly relates to the noise lossiness of $\mathbf{s}$ itself.

Recall that we deduced that all information about $\mathbf{s}$ is captured in the term $\mathbf{sH}+\mathbf{e} \pmod q$. Since $\mathbf{H}$ is invertible, we can multiply the equation by its inverse

$\mathbf{Z}_0$ (on the right) to obtain $\mathbf{s}+\mathbf{e}\mathbf{Z}_0 \pmod q$. We conclude that even information theoretically, an attacker can only recover $\mathbf{s}+\mathbf{e}\mathbf{Z}_0 \pmod q$, where $\mathbf{e}$ is Gaussian and $\mathbf{Z}_0$ is a low-norm matrix which is known to the attacker.

We wish to show that $\mathbf{s}+\mathbf{e}\mathbf{Z}_0 \pmod q$ does not leak much information about $\mathbf{s}$. We can see that some information can in fact be leaked. For example, if $\mathbf{s}$ is short, then the reduction modulo q does not have any effect, and the adversary can learn $\mathbf{s}+\mathbf{e}\mathbf{Z}_0$ (as a value over the integers), this in particular allows to learn the coset of $\mathbf{s}$ relative to the lattice spanned by the rows of $\mathbf{Z}_0$ (henceforth we refer to it as the "$\mathbf{Z}_0$ lattice"). This is essentially the reason why our techniques don't carry over to the setting of very low norm $\mathbf{s}$ – this would require sampling $\mathbf{Z}_0$ from a very narrow distribution that would imply very strong and unrealistic parameters for our DSPR assumption.

Instead, we show that essentially all the entropy that can be gained by the adversary, beyond the "usual" noise lossiness, is indeed proportional to learning a coset of the $\mathbf{Z}_0$ lattice. The number of such cosets is $\approx \gamma^n$, and thus the loss in entropy of $n \log \gamma$ in Theorem 1.1.

To see this, we consider the distribution: $(\mathbf{Z}_0, \mathbf{s}+\tilde{\mathbf{e}}, c)$, where $\tilde{\mathbf{e}}$ is a spherical discrete Gaussian over the integers, and c indicates a coset of $\tilde{\mathbf{e}}$ with respect to the $\mathbf{Z}_0$ lattice. We show that there is a (randomized) process that takes this distribution as input, and outputs $(\mathbf{Z}_0, \mathbf{s}+\mathbf{e}\mathbf{Z}_0)$. This means that the adversary cannot learn about $\mathbf{s}$ from $(\mathbf{Z}_0, \mathbf{s}+\mathbf{e}\mathbf{Z}_0)$ more than it can from $(\mathbf{Z}_0, \mathbf{s}+\tilde{\mathbf{e}}, c)$. The latter, just by definition, translates to the noise lossiness of $\mathbf{s}$ (with respect to the Gaussian parameter of $\tilde{\mathbf{e}}$), minus the "leakage" that is imposed by providing the adversary the value c. Since this value is a coset indicator, this leakage is bounded.

To generate $(\mathbf{Z}_0, \mathbf{s}+\mathbf{e}\mathbf{Z}_0)$ from $(\mathbf{Z}_0, \mathbf{s}+\tilde{\mathbf{e}}, c)$, we use the Gaussian convolution theorem of Peikert [20]. This theorem shows that it is possible to sample the term $\mathbf{e}\mathbf{Z}_0$, which is just a (non spherical) discrete Gaussian over the $\mathbf{Z}_0$ lattice, in two steps: first sampling from a Gaussian over the integer lattice, and then "rounding" the sample into the $\mathbf{Z}_0$ lattice. The rounding step only requires to know the coset of the first step (in order to cancel it out). Setting the parameters appropriately, the theorem can be used and the result follows.

In order to be able to apply Gaussian decomposition and also the Gaussian convolution theorem, we rely on probabilistic properties of the $\mathbf{Z}$ matrices, in particular their minimal and maximal singular values. The properties required in order for our method to go through turn out not to hold with high probability, but rather only with some fixed inverse-polynomial probability. We thus introduce a notion of "sometimes lossiness" and show that it suffices for proving entropic hardness. In Section 5 we show how to obtain entropic hardness based on probabilistic properties of the $\mathbf{Z}$ matrices. Then in Section 6 we show that these properties hold for RLWE over power-of-two cyclotomics, using properties proved in [27].

1.3 Paper Organization

We try to keep the discussion abstract and use the notion of "structured LWE" as much as we can. Eventually we state our result in terms of properties that need to hold for the structured LWE problem at hand, and show that the RLWE/DSPR instantiation indeed possesses these properties. Standard preliminaries in information theory, lattices and algebraic number theory are provided in Section 2. Section 3 introduces the entropic structured LWE (entSLWE) problem and shows that mild form of (entropic) hardness for entSLWE with relatively few samples implies full-fledged (entropic) hardness. Section 4 presents a notion of lossiness that we call "sometimes lossiness" and shows how it is used to prove (entropic) hardness, then a sometimes lossy distribution is constructed in Section 5 based on an abstract problem we call Decisional Small Ratio (DSR) problem. Finally Section 6 shows how to instantiate all required building blocks in the RLWE setting.

2 Notation and Definitions

We will denote the security parameter by λ. We say a function $\nu(\lambda)$ is negligible if $\nu(\lambda) \in \lambda^{-\omega(1)}$. We will generally denote row vectors by $\mathbf{x}$ and column vectors by $\mathbf{x}^\top$. We will denote the L_2 norm of a vector $\mathbf{x}$ by $\|\mathbf{x}\| = \sqrt{\sum_i x_i^2}$ and the L_∞ norm by $\|\mathbf{x}\|_\infty = \max_i |x_i|$.

Let X, Y be two discrete random variables defined on a common support $\mathcal{X}$. We define the statistical distance between X and Y as

$$\Delta(X,Y) = \frac{1}{2}\sum_{x\in\mathcal{X}} |\Pr[X=x] - \Pr[Y=x]|.$$

Consider a real valued matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$, assume for convenience that $m \geq n$. The singular values of $\mathbf{A}$ are the square roots of the eigenvalues of the positive semidefinite (PSD) matrix $\mathbf{A}\mathbf{A}^\top$. We will denote the largest singular value of $\mathbf{A}$ by $\sigma_{max}(\mathbf{A})$. The *spectral norm* of $\mathbf{A}$ is $\sigma_{max}(\mathbf{A})$. It holds that

$$\sigma_{max}(\mathbf{A}) = \max_{\mathbf{x}\in\mathbb{R}^m\setminus\{\mathbf{0}\}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} .$$

2.1 Min-Entropy

Let $\mathbf{x}$ be a discrete random variable supported on a set X and $\mathbf{z}$ be a possibly (continuous) random variable supported on a (measurable) set Z. The conditional min-entropy $\tilde{H}_\infty(\mathbf{x}|\mathbf{z})$ of $\mathbf{x}$ given $\mathbf{z}$ is defined by

$$\tilde{H}_\infty(\mathbf{x}|\mathbf{z}) = -\log\left(\mathsf{E}_{\mathbf{z}'}\left[\max_{\mathbf{x}'\in X}\Pr[\mathbf{x}=\mathbf{x}'|\mathbf{z}=\mathbf{z}']\right]\right).$$

In the case that $\mathbf{z}$ is continuous, this becomes

$$\tilde{H}_\infty(\mathbf{x}|\mathbf{z}) = -\log\left(\int_{\mathbf{z}'} p_{\mathbf{z}}(\mathbf{z}')\max_{\mathbf{x}'\in X}\Pr[\mathbf{x}=\mathbf{x}'|\mathbf{z}=\mathbf{z}']\right),$$

where $p_{\mathbf{z}}(\cdot)$ is the probability density of $\mathbf{z}$.

For an $\epsilon > 0$ we define the ϵ-smooth min-entropy $\tilde{H}^{\epsilon}_{\infty}(\mathbf{x}|\mathbf{z})$ as the maximum over all $\tilde{H}^{\epsilon}_{\infty}(\mathbf{x}'|\mathbf{z}')$ for which $(\mathbf{x}', \mathbf{z}')$ is ϵ-close to $(\mathbf{x}, \mathbf{z})$ in statistical distance.

2.2 Leftover Hashing

We recall a version of the generalized leftover hash lemma [7,24].

Lemma 2.1. *Let $\mathbb{G}$ be a finite Abelian group, and $\mathcal{Y}$ be a finite set. Let $\ell \geq \log(|\mathbb{G}|) + \log(|\mathcal{Y}|) + \omega(\log(\lambda))$ be an integer. Let $g_1, \ldots, g_\ell \leftarrow_{\$} \mathbb{G}$ be chosen uniformly at random. Further let $\mathbf{x} \leftarrow_{\$} \{0,1\}^{\ell}$ be chosen uniformly at random. Let Y be a random variable supported on $\mathcal{Y}$ which is possibly correlated with $\mathbf{x}$ but independent of the g_i. Then it holds that $(g_1, \ldots, g_\ell, \sum_i x_i g_i, Y)$ is statistically close to $(g_1, \ldots, g_\ell, u, Y)$, where $u \leftarrow_{\$} \mathbb{G}$ is chosen uniformly at random.*

2.3 Lattices and Gaussians

Lattices. We recall the standard facts about lattices. A lattice $\Lambda \subseteq \mathbb{R}^m$ is the set of all integer-linear combinations of a set of linearly independent basis-vectors, i.e. for every lattice Λ there exists a full-rank matrix $\mathbf{B} \in \mathbb{R}^{k \times m}$ such that $\Lambda = \Lambda(\mathbf{B}) = \{\mathbf{z} \cdot \mathbf{B} \mid \mathbf{z} \in \mathbb{Z}^k\}$. We call k the rank of Λ and $\mathbf{B}$ a basis of Λ, and we say that Λ is full-rank if $k = m$. For a lattice $\Lambda \subseteq \mathbb{R}^m$, the dual lattice Λ^* is defined by $\Lambda^* = \{\mathbf{x} \in \mathsf{Span}(\Lambda) \mid \forall \mathbf{z} \in \Lambda : \langle \mathbf{z}, \mathbf{x} \rangle \in \mathbb{Z}\}$.

We say that a lattice is q-ary if $(q\mathbb{Z})^m \subseteq \Lambda \subseteq \mathbb{Z}^m$. In particular, for every q-ary lattice Λ there exists a matrix $\mathbf{A} \in \mathbb{Z}_q^{k \times m}$ such that $\Lambda = \Lambda_q(\mathbf{A}) = \{\mathbf{y} \in \mathbb{Z}^m \mid \exists \mathbf{x} \in \mathbb{Z}_q^k : \mathbf{y} = \mathbf{x} \cdot \mathbf{A} \bmod q\}$. We also define the lattice $\Lambda_q^{\perp}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{Z}^m \mid \mathbf{A} \cdot \mathbf{y} = 0 \bmod q\}$.

Gaussians. The Gaussian function $\rho_\sigma : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\rho_\sigma(\mathbf{x}) = e^{-\pi \cdot \frac{\|\mathbf{x}\|^2}{\sigma^2}}.$$

For a a non-singular matrix $\mathbf{B}$ we define $\rho_{\mathbf{B}}(\mathbf{x}) = \rho(\mathbf{x}\mathbf{B}^{-1})$.

The continuous gaussian distribution $D_{\mathbf{B}}$ on $\mathbb{R}^n$ has the probability density function $\rho_{\mathbf{B}}(\mathbf{x})/\rho_{\mathbf{B}}(\mathbb{R}^n)$. We call $\mathbf{\Sigma} = \mathbf{B}^{\top}\mathbf{B}$ the covariance matrix of the gaussian $D_{\mathbf{B}}$. For a lattice Λ, the discrete gaussian distribution $D_{\Lambda,\mathbf{B}}$ supported on Λ has the probability mass function $\rho_{\mathbf{B}}(\mathbf{x})/\rho_{\mathbf{B}}(\Lambda)$.

For a lattice Λ and a positive real $\epsilon > 0$, the *smoothing parameter* $\eta_\epsilon(\Lambda)$ is defined to be the smallest real number s for which $\rho_{1/s}(\Lambda^* \backslash \{0\}) \leq \epsilon$. For a matrix $\mathbf{B}$ we write $\mathbf{B} \geq \eta_\epsilon(\Lambda)$ if $\eta_\epsilon(\Lambda \mathbf{B}^{-1}) \leq 1$.

The following claim follows routinely from the definition of the smoothing parameter.

Claim. Let $\Lambda \subseteq \mathbb{R}^n$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ be a matrix with largest singular value $\sigma_{max}(\mathbf{V})$. It holds that $\eta_\epsilon(\Lambda \cdot \mathbf{V}) \leq \sigma_{max}(\mathbf{V}) \cdot \eta_\epsilon(\Lambda)$.

The following proposition allows us to decompose spherical gaussians with respect to a matrix $\mathbf{F}$.

Proposition 2.2 ([4], Proposition 3.2). *Let $\mathbf{F} \in \mathbb{R}^{n\times m}$ be an arbitrary matrix with spectral norm σ_F. Let $\sigma, \sigma_1 > 0$ be s.t. $\sigma > \sigma_1 \cdot \sigma_F$. Let $\mathbf{e}_1 \sim D^n_{\sigma_1}$ and let $\mathbf{e}_2 \sim D_{\sqrt{\mathbf{\Sigma}}}$ for $\mathbf{\Sigma} = \sigma^2\mathbf{I} - \sigma_1^2\mathbf{F}^\top\mathbf{F}$. Then the random variable $\mathbf{e} = \mathbf{e}_1\mathbf{F} + \mathbf{e}_2$ is distributed according to D^m_σ.*

2.4 Noise Lossiness

The noise lossiness of a distribution $\mathcal{S}$ measures how much information is lost about a sample of $\mathcal{S}$ when adding gaussian noise. Another way to think about noise lossiness is as a measure of how bad $\mathcal{S}$ performs as a Euclidean error-correcting code. The following definition of noise lossiness slightly deviates from the definition given in [4] by considering potentially non-spherical gaussians.

Definition 2.3 (Noise Lossiness). *Fix a matrix $\mathbf{B} \in \mathbb{R}^{n\times n}$. Let $\mathcal{S} \subseteq \mathbb{Z}^n_q$ be a distribution of secrets and let $\sigma > 0$ be a gaussian parameter. We define the noise-lossiness $\nu_{\sigma\mathbf{B}}(\mathcal{S})$ by*

$$\nu_{\sigma\mathbf{B}}(\mathcal{S}) = \tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e})$$

where $\mathbf{s} \leftarrow_\$ \mathcal{S}$ and $\mathbf{e} \leftarrow_\$ D_{\sigma\mathbf{B}}$.

In [4] the following bounds for the noise lossiness of distributions were provided.

Lemma 2.4 (Noise-Lossiness for General Entropic Distributions). *Let $0 < \sigma \leq q\sqrt{\pi/\ln(4n)}$ be a gaussian parameter and let $\mathcal{S}$ be any distribution on $\mathbb{Z}^n_q$. Then it holds that*

$$\nu_\sigma(\mathcal{S}) \geq \tilde{H}_\infty(\mathcal{S}) - n \cdot \log(q/\sigma) - 1$$

Lemma 2.5 (Noise-Lossiness for Short Distributions). *Let $\sigma > 0$ be a gaussian parameter and let $\mathcal{S}$ be a r-bounded distribution on $\mathbb{Z}^n_q$. Then it holds that*

$$\nu_\sigma(\mathcal{S}) \geq \tilde{H}_\infty(\mathcal{S}) - \sqrt{2\pi n}\log(e) \cdot \frac{r}{\sigma}.$$

2.5 Algebraic Number Fields

We will briefly reiterate some basics about algebraic number fields and the Learning with Errors Problem over Rings. See e.g., [16,17] for more details.

An algebraic number field K is a finite extension of the rationals $\mathbb{Q}$, every number field can be constructed via $\mathbb{Q}(\xi) = \mathbb{Q}[X]/(f(X))$ where $f \in \mathbb{Q}[X]$ is a monic irreducible polynomial and ξ is a root of f. The degree n of K is defined to be the degree of f and K can be seen as an n-dimensional $\mathbb{Q}$-vectorspace.

The number fields most relevant to us are power-of-two cyclotomics. For this instantiation the polynomial f is of the form $f = X^n + 1$ where n is a power of two.

A number field K of degree n has n embeddings, that is injective ring homomorphisms into the complex numbers $\mathbb{C}$, usually denoted by $\sigma_i : \mathsf{K} \to \mathbb{C}$. Each σ_i is defined by sending ξ to one of the roots of f in $\mathbb{C}$.

The embeddings σ_i come in conjugate pairs, there are s_1 real embeddings and $2s_2$ complex conjugate embeddings with $n = s_1 + 2s_2$. We can define the space $H \subseteq \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$ by

$$H = \{(x_1, \ldots, x_n) \in \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2} \mid \forall j \in [s_2] : x_{s_1+s_2+j} = \overline{x_{s_1+j}}\}.$$

It can be shown that the space H is isomorphic to $\mathbb{R}^n$ as an inner product space. Let $\Theta : H \to \mathbb{R}^n$ be this isomorphism. Moreover, the space H is isomorphic as a ring to the field tensor product $\mathsf{K}_\mathbb{R} = \mathsf{K} \otimes_\mathbb{Q} \mathbb{R}$. Let $\bar{\Theta} : \mathsf{K}_\mathbb{R} \to \mathbb{R}^n$ be the metric isomorphism which takes $\mathsf{K}_\mathbb{R}$ to $\mathbb{R}^n$, i.e. $\bar{\Theta}$ is just the concatenation of σ and Θ.

The *canonical embedding* $\sigma : \mathsf{K} \to H$ is given by $\sigma(x) = (\sigma_1(x), \ldots, \sigma_n(x))$. It can be shown that σ is a ring-homomorphism, where both addition and multiplication on $\mathbb{C}^n$ are defined component-wise. The canonical embedding induces a geometry on K, that is we can define a eucilidean norm on K via the euclidean norm on $\mathbb{C}^n$, concretely for $x \in \mathsf{K}$ we define $\|x\| = \|\sigma(x)\|$. Note that $\|\sigma(x)\| = \|\Theta(\sigma(x))\|$.

While $\|\cdot\|$ immediately satisfies the triangle inequality, in the canonical embedding also the following multiplicative inequality holds: For all $x, y \in \mathsf{K}_\mathbb{R}$ it holds that $\|x \cdot y\| \leq \|x\|_\infty \cdot \|y\|$. Here, $\|\cdot\|_\infty$ is the L_∞ norm defined by $\|x\|_\infty = \max_i |\sigma_i(x)|$. We will also use the inequality $\|x \cdot y\|_\infty \leq \|x\|_\infty \cdot \|y\|_\infty$.

We can define a gaussian distribution $D_{\mathsf{K}_\mathbb{R}, \sqrt{\mathbf{\Sigma}}}$ via the gaussian distribution $D_{\sqrt{\mathbf{\Sigma}}}$ on $\mathbb{R}^n$, i.e. we set $D_{\mathsf{K}_\mathbb{R}, \sqrt{\mathbf{\Sigma}}} = \bar{\Theta}^{-1}(D_{\sqrt{\mathbf{\Sigma}}})$.

A element $x \in \mathsf{K}$ is called *algebraic integer*, if the minimal polynomial of x has integer coefficients. For a number field K we denote by $\mathsf{R} \subseteq \mathsf{K}$ the set of all algebraic integers in K, which can be shown to be a sub-ring of K. For the special case that K is a cyclotomic, it holds that $\mathsf{R} = \mathbb{Z}[\xi]$.

Since R is a finitely generated $\mathbb{Z}$-module, it holds that $\Lambda = \bar{\Theta}(\mathsf{R}) \subseteq \mathbb{R}^n$ is a lattice. We let $\mathbf{L}$ denote some basis for this lattice and we denote $\mathbf{B} = \mathbf{L}^{-1}$. In this notation, multiplication by the matrix $\mathbf{B}$ maps a $\mathbf{x} \in \lambda$ to an integer vector, i.e. $\mathbf{xB} \in \mathbb{Z}^n$, which is exactly the coefficient vector of the ring element with respect to the basis $\mathbf{L}$. We define the smoothing parameter $\eta_\epsilon(\mathsf{R})$ of R to be $\eta_\epsilon(\Lambda)$.

Gaussian distributions over K, or more precisely over $\mathsf{K}_\mathbb{R}$ are defined as follows. Given a Gaussian distribution $D_{\sqrt{\mathbf{\Sigma}}}$ over $\mathbb{R}^n$, we map it to $\mathsf{K}_\mathbb{R}$ via $\bar{\Theta}^{-1}$. The resulting distribution is the Gaussian with parameter $\sqrt{\mathbf{\Sigma}}$ over $\mathsf{K}_\mathbb{R}$.

2.6 Ring-LWE

Let q be a modulus and R be a ring of integers of a number field K. We will briefly define the (non-dual) decisional Ring Learning with Errors (Ring-LWE) problem in Hermite form for an error-distribution χ supported on R is defined as follows. We discuss other versions of the Ring LWE problem in Section 3. We use a definition provided by Peikert [21, Section 4.4.1] which is slightly different

from the one in [16] but easier to work with. See discussion in [21, Section 4.4.1] for details.

Definition 2.6 (Decisional Ring-LWE (Hermite Form)). *Let* $\mathbf{s} \leftarrow_\$ \chi$. *Given* m *samples* $(\mathbf{a}_i, \mathbf{b}_i) \in \mathsf{R}_q \times \mathsf{R}_q$, *the task is to decide whether the* $\mathbf{b}_i$ *are of the form* $\mathbf{b}_i = \mathbf{a}_i\mathbf{s} + \mathbf{e}_i$ *for errors* $\mathbf{e}_i \leftarrow_\$ \chi$ *or if the* $\mathbf{b}_i$ *are chosen uniformly at random from* R_q.

Lyubashevsky, Peikert and Regev [16] provided a worst-to-average case reduction for the Ring LWE problem relative to worst-case problems in ideal lattices. In particular, they show that if the error distribution χ is an appropriate gaussian, then the Ring LWE search problem is as hard as the approximate shortest vector problem in worst case ideal lattices. Furthermore, [16] provide a search-to-decision reduction which bases the hardness of decisional Ring LWE on the search variant.

3 (Entropic) Structured LWE

In this section, we define a version of LWE which we call *structured LWE*. Structured LWE generalizes both standard and ring-LWE.

We will only consider the search version of structured LWE in this work.

Definition 3.1 (Entropic Structured Learning with Errors). *Let* q *be a modulus and* n, k *be integers. Let* $\mathcal{M}$ *be a distribution of matrices on* $\mathbb{Z}_q^{n\times n}$ *and* Υ *be a distribution of error-distributions on* $\mathbb{R}^n$. *Furthermore, let* $\mathcal{S}$ *be a distribution on* $\mathbb{Z}_q^n$. *The goal of the* $\mathsf{entSLWE}(q, k, \mathcal{M}, \Upsilon, \mathcal{S})$ *problem is to find a secret* $\mathbf{s} \leftarrow_\$ \mathcal{S}$ *given* k *samples* $((\mathbf{A}_1, \mathbf{y}_1), \ldots, (\mathbf{A}_k, \mathbf{y}_k))$, *where* $\chi \leftarrow_\$ \Upsilon$ *is an error distribution and for all* $i \in [k]$ *we have* $\mathbf{A}_i \leftarrow_\$ \mathcal{M}$, $\mathbf{e}_i \leftarrow_\$ \chi$ *and* $\mathbf{y}_i \leftarrow \mathbf{s}\mathbf{A}_i + \mathbf{e}_i$.

If $\mathbf{A} = (\mathbf{A}_1, \ldots, \mathbf{A}_k)$, $\mathbf{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$ and $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ where $\mathbf{y}_i = \mathbf{s}\mathbf{A}_i + \mathbf{e}_i$, we will use the shorthand $\mathbf{y} = \mathbf{s}\mathbf{A} + \mathbf{e}$. This in fact corresponds to the standard matrix multiplication and vector addition if we identify $\mathbf{A}$ with to be the horizontal concatenation of all $\mathbf{A}_i$ and $\mathbf{e}$ the horizontal concatenation of all $\mathbf{e}_i$. If an unbounded number of samples are given (via an oracle), then we will omit the parameter k. We note that Regev's LWE is obtained when $\mathcal{M}, \mathcal{S}$ are uniform and Υ is Gaussian. The Ring-LWE instantiation is discussed in Section 3.2 below.

We will consider two different hardness notions for $\mathsf{entSLWE}$. In the standard notion, we require that no PPT adversary find the secret $\mathbf{s}$ with non-negligible probability.

Definition 3.2 (Standard Hardness). *Let* q, n, k, $\mathcal{M}$, Υ *and* $\mathcal{S}$ *be as above. We say that the* $\mathsf{entSLWE}(q, k, \mathcal{M}, \Upsilon, \mathcal{S})$ *problem is (standard-) hard, if it holds for every PPT adversary* $\mathcal{A}$ *that*

$$\Pr[\mathcal{A}(\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e}) = \mathbf{s}] < \mathsf{negl}(\lambda),$$

where $\chi \leftarrow_\$ \Upsilon$, $\mathbf{A} \leftarrow_\$ \mathcal{M}^k$, $\mathbf{s} \leftarrow_\$ \mathcal{S}$ *and* $\mathbf{e} \leftarrow_\$ \chi^k$.

We call the second notion mild hardness. In essence, the success probability of an adversary which breaks mild hardness only depends on the choice of $\mathbf{s}$ and $\mathbf{e}$, but not on the choice of $\mathbf{A}$.

Definition 3.3 (Mild Hardness). *Let q, n, k, $\mathcal{M}$, Υ and $\mathcal{S}$ be as above. We say that the problem $\mathsf{entSLWE}(q, k, \mathcal{M}, \Upsilon, \mathcal{S})$ is mildly hard, if for every PPT adversary $\mathcal{A}$ and every negligible function ν it holds that*

$$\Pr_{\mathbf{s},\mathbf{e},\chi}[\Pr_{\mathbf{A}}[\mathcal{A}(\mathbf{A}, \mathbf{sA} + \mathbf{e}) = \mathbf{s}] > 1 - \nu] < \mathsf{negl}(\lambda).$$

In this work we will focus on the notion of mild hardness. While this seems like a restriction at first glance, it follows by a routine amplification argument that, given an unbounded number of samples, mild hardness implies standard hardness.

Lemma 3.4. *Let q, n, $\mathcal{M}$ and $\mathcal{S}$ be as above and let Υ be a distribution of error-distributions. If $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon, \mathcal{S})$ is mildly hard, then it is also standard hard.*

Proof. Assume towards contradiction there was a PPT search adversary $\mathcal{A}$ with non-negligible success probability ϵ' against standard hardness of the problem $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon, \mathcal{S})$. For notational convenience, the adversary $\mathcal{A}$ obtains its samples via an oracle $\mathcal{O}_{\mathbf{s},\chi}$, which has $\mathbf{s}$ and χ hardwired. When queried, $\mathcal{O}_{\mathbf{s},\chi}$ chooses $\mathbf{A} \leftarrow_\$ \mathcal{M}$ and $\mathbf{e} \leftarrow_\$ \chi$ and outputs a sample $(\mathbf{A}, \mathbf{sA} + \mathbf{e})$. Let $\epsilon = 1/\mathsf{poly}(\lambda)$ be such that $\epsilon(\lambda) = \epsilon'(\lambda)$ infinitely often. We will construct an adversary $\mathcal{B}$ against the mild hardness of $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon, \mathcal{S})$ as follows.

Algorithm $\mathcal{B}^{\mathcal{O}_{\mathbf{s},\chi}}$

- For $i = 1, \ldots, 2\lambda/\epsilon$:
  - Compute $\mathbf{s}_i \leftarrow \mathcal{A}^{\mathcal{O}_{\mathbf{s},\chi}}(1^\lambda)$.
  - Query λ additional samples and test whether $\mathbf{s}_i$ is a valid solution, if so output $\mathbf{s} \leftarrow \mathbf{s}_i$
- If none of the $\mathbf{s}_i$ passed the check, output $\perp$.

Assume that $\mathbf{y}_i = \mathbf{sA}_i + \mathbf{e}_i$ for all $i \in [k]$. We will now analyze the success probability of $\mathcal{B}$. Say that a pair $(\mathbf{s}, \chi)$ is *good*, if it holds that

$$\Pr[\mathcal{A}^{\mathcal{O}_{\mathbf{s},\chi}}(1^\lambda) = \mathbf{s}] \geq \epsilon/2,$$

where the probability is taken over the remaining random choices of $\mathcal{O}$ and the random coins of $\mathcal{A}$. By a Markov inequality, it holds that

$$\Pr_{\mathbf{s},\chi}[(\mathbf{s}, \chi) \text{ good}] = \Pr_{\mathbf{s},\chi}[\Pr[\mathcal{A}^{\mathcal{O}_{\mathbf{s},\chi}(\cdot)}(1^\lambda) = \mathbf{s}] \geq \epsilon/2] \geq \epsilon/2.$$

Now, fix a good $(\mathbf{s}, \chi)$. We will bound the probability that all iterations of $\mathcal{B}$ fail to compute $\mathbf{s}$. Once we have fixed $\mathbf{s}$ and χ, all iterations use independent

random coins, and thus their outcomes are independent. Consequently, it holds that

$$\begin{aligned}\Pr[\forall i \in [2\lambda/\epsilon] : \mathcal{A}^{\mathcal{O}_{\mathbf{s},\chi}(\cdot)}(1^\lambda) \neq \mathbf{s}] &= \prod_{i=1}^{2\lambda/\epsilon} \Pr[\mathcal{A}^{\mathcal{O}_{\mathbf{s},\chi}(\cdot)}(1^\lambda) \neq \mathbf{s}] \\ &\leq (1-\epsilon/2)^{2\lambda/\epsilon} \\ &\leq \exp(-\epsilon/2 \cdot 2\lambda/\epsilon) \\ &= \exp(-\lambda),\end{aligned}$$

which is negligible. We can conclude that

$$\Pr_{\mathbf{s}}[\Pr[\mathcal{B}((\mathbf{A}_i, \mathbf{s}\mathbf{A}_i + \mathbf{e}_i)_{i\in[k]}) = \mathbf{s}] > 1 - \exp(-\lambda)] \geq \epsilon/2,$$

which means that $\mathcal{B}$ breaks the mild hardness of $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon, \mathcal{S})$.

3.1 Rerandomization

Lemma 3.4 holds given an unbounded number of samples. We will now consider statistical rerandomization procedures which allow to generate an unbounded number of samples $(\mathbf{A}_i, \mathbf{s}\mathbf{A}_i + \mathbf{e}_i)$ from a fixed number of samples. A typical artifact of statistical re-randomization is that if one starts with a bounded number of samples for a fixed error distribution χ, then the rerandomized samples will have an error that comes from a distribution of error distributions. We provide a simple rerandomization procedure which takes random subset sums over the input samples. While the norm of errors in the output distribution will be bounded, these errors will not follow a *nice* distribution.

Lemma 3.5. *Let $k \geq \log(|\mathbb{G}|) + n\log(q) + \omega(\log(\lambda))$, let Φ be an error distribution on $\mathbb{Z}^n$. The distribution of error-distributions $\Upsilon_{\Phi,bin}$ is defined as follows: A distribution $\chi \leftarrow_\$ \Upsilon_{\Phi,bin}$ is determined by k elements $\mathbf{e}_1, \ldots, \mathbf{e}_k \in \mathbb{Z}^n$ chosen from Φ. To sample from the distribution χ, choose a $\mathbf{x} \leftarrow_\$ \{0,1\}^k$ uniformly at random and output $\sum_i x_i \mathbf{e}_i$.*

If $\mathsf{entSLWE}(q, k, \mathcal{M}, \Phi, \mathcal{S})$ is mildly hard, then $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon_{\Phi,bin}, \mathcal{S})$ is also mildly hard.

Note that if the distribution Φ is B-bounded, then $\Upsilon_{\Phi,\mathrm{bin}}$ is kB-bounded.

Proof. The reduction proceeds via statistical rerandomization. Let $\mathcal{A}$ be an adversary against the mild hardness of $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon_{\Phi,\mathrm{bin}}, \mathcal{S})$. We will construct an adversary $\mathcal{B}$ against the mild hardness of $\mathsf{entSLWE}(q, k, \mathcal{M}, \Phi, \mathcal{S})$. More concretely, assume there is a negligible function ν and a non-negligible function ϵ such that

$$\Pr_{\mathbf{s},\mathbf{e},\chi}[\Pr_{\mathbf{A}}[\mathcal{A}(\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e}) = \mathbf{s}] > 1 - \nu] > \epsilon.$$

The adversary $\mathcal{B}$ proceeds as follows.

Algorithm $\mathcal{B}$

- Input: k samples $(\mathbf{A}_1, \mathbf{y}_1), \ldots, (\mathbf{A}_k, \mathbf{y}_k)$.
- Setup an oracle $\mathcal{O}$, which when queried chooses a uniformly random $\mathbf{x} \in \{0,1\}^k$ and outputs $(\sum_i \mathbf{x}_i \mathbf{A}_i, \sum_i \mathbf{x}_i \mathbf{y}_i)$.
- Compute and output $\mathbf{s} \leftarrow \mathcal{A}^{\mathcal{O}(\cdot)}(1^\lambda)$

We will now show that $\mathcal{B}$ faithfully simulates the oracle $\mathcal{O}$ of the problem $\mathsf{entSLWE}(q, \mathcal{M}, \Upsilon_{\Phi,\mathrm{bin}}, \mathcal{S})$. Assume that $\mathbf{y}_i = \mathbf{s}\mathbf{A}_i + \mathbf{e}_i$. Then the rerandomized sample

$$(\sum_i \mathbf{x}_i \mathbf{A}_i, \sum_i \mathbf{x}_i \mathbf{y}_i = \mathbf{s}(\sum_i \mathbf{x}_i \mathbf{A}_i) + \sum_i \mathbf{x}_i \mathbf{e}_i)$$

has an error term $\mathbf{e}^* = \sum_i \mathbf{x}_i \mathbf{e}_i$ which follows a distribution χ of $\Upsilon_{\Phi,\mathrm{bin}}$, where χ is defined by $\mathbf{e}_1, \ldots, \mathbf{e}_k \in \mathbb{Z}^n$. Note that $\mathbf{e}^*$ is supported on $\mathbb{Z}_q^n$. Thus, by the leftover hash lemma (Lemma 2.1), the distribution of $\sum_i \mathbf{x}_i \mathbf{A}_i$ is statistically close to uniform in $\mathbb{G}$ given the $\mathbf{e}^*$ and we conclude that the distribution of the samples generated by $\mathcal{O}$ is statistically close to the correct distribution, which concludes the proof.

3.2 Ring-LWE as Structured LWE

Recall the conventions and properties from algebraic number theory as described in Section 2.5, and the definition of RLWE from Section 2.6 (note that we use the simpler definition that does not use the so-called dual-ring). In particular recall that the ring of integers R of the number field K is a finitely generated $\mathbb{Z}$-module. Since the number field K is mapped into $\mathbb{R}^n$ via the mapping $\bar{\Theta}$, this mapping allows to cast R as a lattice Λ. We denote the basis of this lattice by $\mathbf{L}$ and its inverse by $\mathbf{B} = \mathbf{L}^{-1}$. The mapping $\mathbf{B} \circ \bar{\Theta}$ therefore maps from K to $\mathbb{R}^n$ such that the image of R is $\mathbb{Z}^n$.

Let $a \in \mathsf{R}$. Since multiplication with a is a linear function, there exists a matrix $\mathbf{A}_a \in \mathbb{Z}^{n \times n}$, such that for all $s \in \mathsf{R}$, if $\mathbf{s} \in \mathbb{Z}^n$ is the vector representation of s according to the aforementioned mapping, then $\mathbf{A}_a \mathbf{s}$ is the vector representation of $a \cdot s \in \mathsf{R}$ according to the above mapping. A Gaussian distribution with parameter $\sqrt{\mathbf{\Sigma}}$ over the field is mapped by $\mathbf{B} \circ \bar{\Theta}$ to a Gaussian over $\mathbb{R}^n$ with parameter $\sigma\mathbf{B}$.

Therefore, a Ring-LWE equation of the form $as + e$, with $a, s \in \mathsf{R}_q = \mathsf{R}/q\mathsf{R}$ is translated by the mapping $\mathbf{B} \circ \bar{\Theta}$ (which is efficiently computable and efficiently invertible given $\mathbf{B}$) into the linear equation $\mathbf{A}_a \mathbf{s} + \mathbf{e} \pmod q$, where $\mathbf{A}_a \in \mathbb{Z}_q^{n \times n}$, $\mathbf{s} \in \mathbb{Z}_q^n$ and e is sampled from the distribution $\chi = D_{\sigma\mathbf{B}}$.

Therefore given a Ring-LWE instance, we can convert it into a structured LWE instance with the aforementioned parameters, so that solving the structured-LWE instance will also imply a solution to the original Ring-LWE instance. The "quality" of the translation relies on the properties of the matrix $\mathbf{B}$, i.e. on how good of a basis for R we can obtain. We discuss the properties of $\mathbf{B}$ in the case of power-of-two cyclotomic number fields in Section 6.

4 Sometimes Lossiness and Hardness of Entropic Structured LWE

We will first define a new lossiness notion which we call *Sometimes Lossiness*. This notion will serve as our main tool to establish hardness of entropic generalized LWE problems. Recall the definitions of smooth min-entropy (see Section 2.1).

Definition 4.1. *Let q, n, k be integers. Let $\mathcal{X}$ be a distribution on $(\mathbb{Z}_q^{n\times n})^k$, $\mathcal{S}$ be a distribution on $\mathbb{Z}_q^n$ and χ be an error-distribution on $\mathbb{Z}_q^n$. We say that $\mathcal{X}$ is a sometimes lossy pseudorandom distribution for $\mathcal{S}$ and χ if there exists negligible function ϵ, a $\kappa = \omega(\log(\lambda))$ and a $\delta \geq 1/\mathsf{poly}(\lambda)$ such that the following properties hold.*

- ***Pseudorandomness:*** *$\mathcal{X}$ is computationally indistinguishable from $\mathcal{M}^k$.*
- ***Sometimes Lossiness:*** *It holds that*

$$\Pr_{\mathbf{A}\leftarrow_\$\mathcal{X}}[\tilde{H}_\infty^\epsilon(\mathbf{s}|\mathbf{A}, \mathbf{sA}+\mathbf{e}) \geq \kappa] \geq \delta,$$

where $\mathbf{s} \leftarrow_\$ \mathcal{S}$ and $\mathbf{e} \leftarrow_\$ \chi^k$.

4.1 From Sometimes Lossiness to the Hardness of Entropic Structured LWE

We will now show that a sometimes lossy pseudorandom distribution $\mathcal{X}$ for a distribution of secrets $\mathcal{S}$ and an error distribution χ implies that hardness of $\mathsf{entSLWE}(q, k, \mathcal{M}, \chi, \mathcal{S})$.

Theorem 4.2. *Let $\mathcal{S}$ be a distribution of secrets and let χ be an error distribution. Assume there exists a sometimes lossy pseudorandom distribution $\mathcal{X}$ on $(\mathbb{Z}_q^{n\times n})^k$. Then $\mathsf{entSLWE}(q, k, \mathcal{M}, \chi, \mathcal{S})$ is mildly hard.*

Proof. Let $\delta = 1/\mathsf{poly}(\lambda)$ be as in Definition 4.1. Set $\ell = \lambda/\delta = \mathsf{poly}(\lambda)$. By a standard hybrid argument, it holds that

$$(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)}) \approx_c (\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{\ell}),$$

where $\mathbf{A}^{(i)} \leftarrow_\$ \mathcal{X}$ and $\mathbf{U}^{(i)} \leftarrow_\$ \mathcal{M}^k$ for all $i = 1, \ldots, \ell$. Our argument will make use of the fact that by our choice of ℓ, *some* of the $\mathbf{A}^{(i)}$ must be lossy, except with some negligible probability.

Assume towards contradiction that $\mathsf{entSLWE}(q, k, \mathcal{M}, \chi, \mathcal{S})$ is not mildly hard, i.e. there exists a PPT adversary $\mathcal{A}$ against $\mathsf{entSLWE}(q, k, \mathcal{M}, \chi, \mathcal{S})$ such that

$$\Pr_{\mathbf{s},\mathbf{e}}[\Pr_{\mathbf{A}}[\mathcal{A}(\mathbf{A}, \mathbf{sA}+\mathbf{e}) = \mathbf{s}] > 1-\nu] > \epsilon,$$

where $\mathbf{s} \leftarrow_\$ \mathcal{S}$, $\mathbf{A} \leftarrow_\$ \mathcal{X}$, $\mathbf{e} \leftarrow_\$ \chi$, $\nu = \nu(\lambda)$ is negligible and $\epsilon \geq 1/\mathsf{poly}(\lambda)$.

We will use $\mathcal{A}$ to construct a distinguisher $\mathcal{D}$ which distinguishes the random variables $(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)})$ and $(\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(\ell)})$ with non-negligible advantage. Let $N = \lambda/\epsilon = \mathsf{poly}(\lambda)$. The distinguisher $\mathcal{D}$ is given as follows.

$\mathcal{D}(\mathbf{A}_1, \ldots, \mathbf{A}_\ell)$:
For $i = 1, \ldots, \ell$:
- For $j = 1, \ldots, N$:
  - Choose $\mathbf{s}_{i,j} \leftarrow_\$ \mathcal{S}$ and $\mathbf{e}_{i,j} \leftarrow_\$ \chi^k$
  - Compute $\mathbf{s}'_{i,j} \leftarrow \mathcal{A}(\mathbf{A}^{(i)}, \mathbf{s}_{i,j}\mathbf{A}^{(i)} + \mathbf{e}_{i,j})$
- If for all $j \in [N]$ it holds that $\mathbf{s}'_{i,j} \neq \mathbf{s}_{i,j}$, abort and output 1.

Output 0.

We will now analyze the distinguishing advantage of $\mathcal{D}$.

1. First assume that $\mathcal{A}$'s input is $(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)})$, where each $\mathbf{A}^{(i)}$ is chosen from $\mathcal{X}$. Since the $\mathbf{A}^{(i)}$ are all independent and $\mathcal{X}$ is sometimes lossy for $\mathcal{S}$ and χ, recalling that $\ell = \lambda/\epsilon$ it holds that

$$\begin{aligned}\Pr_{\mathbf{A}^{(1)},\ldots,\mathbf{A}^{(\ell)}}[\forall i \in [\ell] : \tilde{H}_\infty(\mathbf{s}|\mathbf{s}\mathbf{A}^{(i)} + \mathbf{e}) < \kappa] &= \prod_{i=1}^{\ell} \Pr_{\mathbf{A}^{(i)}}[\tilde{H}_\infty(\mathbf{s}|\mathbf{s}\mathbf{A}^{(i)} + \mathbf{e}) < \kappa] \\ &\leq (1-\epsilon)^\ell \leq e^{-\epsilon\ell} = e^{-\lambda},\end{aligned}$$

   which is negligible. Consequently, there exists an index $i \in [\ell]$ such that $\tilde{H}_\infty(\mathbf{s}|\mathbf{s}\mathbf{A}^{(i)} + \mathbf{e}) \geq k$, except with negligible probability over the choice of the $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)}$. Thus, fix $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)}$ for which there exists an index $i^* \in [\ell]$ with $\tilde{H}_\infty(\mathbf{s}|\mathbf{s}\mathbf{A}^{(i^*)} + \mathbf{e}) \geq k$. Now, since $\mathbf{s}_{i^*,1}, \ldots, \mathbf{s}_{i^*,N} \leftarrow_\$ \mathcal{S}$, it holds by a union-bound that

$$\begin{aligned}\Pr[\exists j \in [N] : \mathcal{A}(\mathbf{A}^{(i^*)}, \mathbf{s}_{i^*,j}\mathbf{A}^{(i^*)} + \mathbf{e}_{i^*,j}) = \mathbf{s}_{i^*,j}] &\\ \leq N \cdot \Pr[\mathcal{A}(\mathbf{A}^{(i^*)}, \mathbf{s}\mathbf{A}^{(i^*)} + \mathbf{e}) = \mathbf{s}] &\\ \leq N \cdot 2^{-\tilde{H}_\infty(\mathbf{s}|\mathbf{A}^{(i^*)}, \mathbf{s}\mathbf{A}^{(i^*)} + \mathbf{e})} &\\ \leq N \cdot 2^{-\kappa}, &\end{aligned}$$

   where $\mathbf{s} \leftarrow_\$ \mathcal{S}$ and $\mathbf{e} \leftarrow_\$ \chi$. The term $N \cdot 2^{-\kappa}$ is negligible as $N = \mathsf{poly}(\lambda)$ and $\kappa = \omega(\log(\lambda))$. Consequently, it follows that in the computation of $\mathcal{D}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)})$ in the i^*-th iteration of the outer loop it will hold that $\mathbf{s}'_{i^*,j} \neq \mathbf{s}_{i^*,j}$ for all $j \in [N]$, except with negligible probability over the choice of $\mathbf{s}_{i^*,1}, \ldots, \mathbf{s}_{i^*,N}$ and $\mathbf{e}_{i^*,1}, \ldots, \mathbf{e}_{i^*,N}$. This will cause $\mathcal{D}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)})$ to output 1.
   All together, we conclude that in case $(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)})$ is chosen from $\mathcal{X}^\ell$, it holds that $\mathcal{D}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)}) = 1$, except with negligible probability over the choice of $(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\ell)})$ and the random coins of $\mathcal{D}$.
2. Now assume that $\mathcal{A}$'s input is $(\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(\ell)})$, where each $\mathbf{U}_i$ is chosen from $\mathcal{M}^k$. We will show that with high probability over the choice of the $(\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(\ell)})$ and the random coins of $\mathcal{D}$, for every iteration i there will be an index j such that $\mathbf{s}'_{i,j} = \mathbf{s}_{i,j}$, which will cause $\mathcal{D}(\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(\ell)})$ to output 0.
   Now fix an $i^* \in [\ell]$. Define the event $\mathsf{BAD}(\mathbf{s}, \mathbf{e})$ by

$$\mathsf{BAD}(\mathbf{s}, \mathbf{e}) :\Leftrightarrow \Pr_{\mathbf{U}}[\mathcal{A}(\mathbf{U}, \mathbf{s}\mathbf{U} + \mathbf{e}) = \mathbf{s}] \leq 1 - \nu,$$

where $\mathbf{U} \leftarrow_{\$} \mathcal{M}$. Recall that since we assume that $\mathcal{A}$ breaks mild hardness it holds that $\Pr_{\mathbf{s},\mathbf{e}}[\mathsf{BAD}(\mathbf{s},\mathbf{e})] \leq 1-\epsilon$. We will now bound the probability that all $(\mathbf{s}_{i^*,1},\mathbf{e}_{i^*,1}),\ldots,(\mathbf{s}_{i^*,N},i^*,N)$ are bad. Since all the pairs $(\mathbf{s}_{i^*,1},\mathbf{e}_{i^*,1}),\ldots,(\mathbf{s}_{i^*,N},i^*,N)$ are independent, it holds that

$$\begin{aligned}\Pr[\forall j \in [N] : \mathsf{BAD}(\mathbf{s}_{i^*,j},\mathbf{e}_{i^*,j})] &= \prod_{j\in[N]} \Pr[\mathsf{BAD}(\mathbf{s}_{i^*,j},\mathbf{e}_{i^*,j})] \\ &\leq (1-\epsilon)^N \leq \exp(-\epsilon\cdot N) = \exp(-\lambda),\end{aligned}$$

where we have used that $N = \lambda/\epsilon$. Consequently, it holds with overwhelming probability $1-\exp(-\lambda)$ that at least one $\mathbf{s}_{i^*,j}$ is not bad. Thus, fix $(\mathbf{s}_{i^*,1},\mathbf{e}_{i^*,1}),\ldots,(\mathbf{s}_{i^*,N},i^*,N)$ such that there is an index j^* such that the pair $(\mathbf{s}_{i^*,j^*},\mathbf{e}_{i^*,j^*})$ is not bad, i.e. $\Pr_{\mathbf{U}}[\mathcal{A}(\mathbf{U},\mathbf{s}_{i^*,j^*}\mathbf{U}+\mathbf{e}_{i^*,j^*}) = \mathbf{s}_{i^*,j^*}] > 1-\nu$. It follows that

$$\begin{aligned}&\Pr_{\mathbf{U}^{(i^*)}}[\exists j \in [N] : \mathcal{A}(\mathbf{U}^{(i^*)},\mathbf{s}_{i^*,j}\mathbf{U}^{(i^*)}+\mathbf{e}_{i^*,j}) = \mathbf{s}_{i^*,j}] \\ &\quad\geq \Pr_{\mathbf{U}^{(i^*)}}[\mathcal{A}(\mathbf{U}^{(i^*)},\mathbf{s}_{i^*,j^*}\mathbf{U}^{(i^*)}+\mathbf{e}_{i^*,j^*}) = \mathbf{s}_{i^*,j^*}] \\ &\quad\geq 1-\nu,\end{aligned}$$

which is overwhelming. We can conclude that, it happens with at most negligible probability over the choice of the $\mathbf{s}_{i^*,1},\ldots,\mathbf{s}_{i^*,N}$, $\mathbf{e}_{i^*,1},\ldots,\mathbf{e}_{i^*,N}$ and $\mathbf{U}^{(i^*)}$ that the i^*-th iteration of the outer loop *does not* results in an abort with output 1.
A union-bound over all $i^* \in [\ell]$ yields that with at most negligible probability over the choice of the $\mathbf{U}^{(1)},\ldots,\mathbf{U}^{(\ell)}$ and the random coins of $\mathcal{D}$ that in the computation of $\mathcal{D}(\mathbf{U}^{(1)},\ldots,\mathbf{U}^{(\ell)})$ any of the ℓ iterations of the outer loop results in an abort with output 1. By construction of $\mathcal{D}$, this means that $\mathcal{D}(\mathbf{U}^{(1)},\ldots,\mathbf{U}^{(\ell)}) = 0$ with overwhelming probability.

Putting everything together, we conclude that

$$\begin{aligned}&\Pr[\mathcal{D}(\mathbf{A}^{(1)},\ldots,\mathbf{A}^{(\ell)}) = 1] - \Pr[\mathcal{D}(\mathbf{U}^{(1)},\ldots,\mathbf{U}^{(\ell)}) = 1] \\ &\quad= \Pr[\mathcal{D}(\mathbf{A}^{(1)},\ldots,\mathbf{A}^{(\ell)}) = 1] + \Pr[\mathcal{D}(\mathbf{U}^{(1)},\ldots,\mathbf{U}^{(\ell)}) = 0] - 1 \\ &\quad= 1 - \mathsf{negl}(\lambda),\end{aligned}$$

Thus, $\mathcal{D}$ distinguishes $\mathcal{X}$ and $\mathcal{M}^k$ with advantage close to 1, which contradicts the assumption that $\mathcal{X}$ and $\mathcal{M}$ are computationally indistinguishable. This concludes the proof.

5 Construction of Sometimes Lossy Distributions

In this section we will construct sometimes lossy distributions from a somewhat general problem we call Decisional Small Ratio (DSR) problem. In Section 6 we will show that DSR can be instantiated with by the Decisional Small Polynomial

Ratio (DSPR) assumption (which is related to the NTRU problem) or the standard RLWE assumption, leading to sometimes lossy distributions with different parameters.

Definition 5.1 (Decisional Small Ratio (DSR) Assumption). *Let q be a modulus and k, n be integers and let $\mathcal{M}$ be a distribution of matrices on $\mathbb{Z}_q^{n\times n}$. Let Ψ be a distribution on $(\mathbb{Z}_q^{n\times n})^\times \times \mathbb{Z}_q^{n\times nk}$. The DSR assumption for q, n, k, $\mathcal{M}$ and Ψ postulates that*

$$\mathbf{H}\cdot\mathbf{Z} \approx_c \mathbf{U},$$

where $(\mathbf{Z}_0, \mathbf{Z}) \leftarrow_\$ \Psi$, $\mathbf{H}$ is the $\mathbb{Z}_q$-inverse of $\mathbf{Z}_0 \mod q$ and $\mathbf{U} \leftarrow_\$ \mathcal{M}^k$.

The DSR assumption generalizes the Decisional Small Polynomial Ration (DSPR) assumption [15], which itself is a generalization of the decisional NTRU assumption. We will show that under certain conditions the DSR assumption implies a sometimes lossy mode for LWE.

In our analysis, we will make use of the following smoothing lemma and convolution theorem.

Lemma 5.2 ([24, Claim 3.9]). *Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice and let $\sigma \geq \sqrt{2}\eta_\epsilon(\Lambda)$. Let $\mathbf{e} \sim D_{\Lambda,\sigma}$ be a discrete gaussian and $\mathbf{e}' \sim D_{\mathbb{R}^n,\sigma}$ be a continuous gaussian. Then $\mathbf{e}+\mathbf{e}'$ is 4ϵ close to $D_{\mathbb{R}^n,\sqrt{2}\sigma}$.*

Theorem 5.3 ([20, Thm 3.1]). *Let $\mathbf{\Sigma}_1, \mathbf{\Sigma}_2 > 0$ be two positive definite matrices such that $\mathbf{\Sigma} = \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2 > \mathbf{0}$ and $\mathbf{\Sigma}_1^{-1} + \mathbf{\Sigma}_2^{-1} > \mathbf{0}$. Let Λ_1, Λ_2 be two lattices such that $\sqrt{\mathbf{\Sigma}_1} \geq \eta_\epsilon(\Lambda_1)$ and $\sqrt{\Lambda_2} \geq \eta_\epsilon(\Lambda_2)$ for some $\epsilon > 0$. Let $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^n$ be arbitrary. Consider the following sampling procedure for $\mathbf{x} \in \Lambda_2 + \mathbf{c}_2$.*

- *Choose $\mathbf{x}_1 \leftarrow_\$ D_{\Lambda_1+\mathbf{c}_1,\sqrt{\mathbf{\Sigma}_1}}$.*
- *Choose $\mathbf{x} \leftarrow_\$ \mathbf{x}_1 + D_{\Lambda_2+\mathbf{c}_2-\mathbf{x}_1,\sqrt{\mathbf{\Sigma}_2}}$.*

Then it holds that the marginal distribution of $\mathbf{x}$ is within statistical distance 8ϵ to $D_{\Lambda_2+\mathbf{c}_2}$.

Lemmas 5.4,5.5 and 5.6 will be used to prove Theorem 5.7, the main technical result of this section.

Convention: In the following lemmas, always assume the following: q is a modulus, n is an integer and $\mathbf{B} \in \mathbb{R}^{n\times n}$. Moreover let $\Lambda = \Lambda(\mathbf{B}^{-1})$ and set $s = \eta_\epsilon(\Lambda)$.

Lemma 5.4 (Blockwise Gaussian Decomposition). *Let $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_k) \in \mathbb{R}^{n\times nk}$, where for all i $\mathbf{F}_i \in \mathbb{R}^{n\times n}$ and set $\mathbf{F}' = (\mathbf{B}\mathbf{F}_1\mathbf{B}^{-1}, \ldots, \mathbf{B}\mathbf{F}_k\mathbf{B}^{-1})$. Assume that the largest singular value of $\mathbf{F}'$ is $\sigma_{\mathbf{F}'}$. Let $\sigma, \sigma_1 > 0$ be such that $\sigma \geq \sigma_{\mathbf{F}'} \cdot \sigma_1$. There exists a distribution Ψ on $\mathbb{R}^{nk}$, such that if $\mathbf{e}' \sim D_{\sigma_1\cdot\mathbf{B}}$ and $\mathbf{e}'' \sim \Psi$ are independent, then $\mathbf{e} = \mathbf{e}'\mathbf{F} + \mathbf{e}''$ is distributed according to $D^k_{\sigma\mathbf{B}}$.*

Proof. Let $\mathbf{\Sigma} = \sigma^2\mathbf{I} - \sigma_1^2{\mathbf{F}'}^\top\mathbf{F}'$. Let $\mathbf{f}' \sim D_{\sigma_1\mathbf{I}} = D^n_{\sigma_1}$ and $\mathbf{f}'' \sim D_{\sqrt{\mathbf{\Sigma}}}$. By Proposition 2.2 it holds that $\mathbf{f} = \mathbf{f}'\mathbf{F}' + \mathbf{f}''$ is distributed according to $D^{nk}_\sigma = D^k_{\sigma\mathbf{I}}$.

Write $\mathbf{f} = (\mathbf{f}_1, \ldots, \mathbf{f}_k)$ and $\mathbf{f}'' = (\mathbf{f}''_1, \ldots, \mathbf{f}''_k)$. Then it holds for all i

$$\mathbf{f}_i = \mathbf{f}' \cdot \mathbf{F}'_i + \mathbf{f}'' = \mathbf{f}'\mathbf{B}\mathbf{F}_i\mathbf{B}^{-1} + \mathbf{f}''_i.$$

Multiplying both sides with $\mathbf{B}$ yields

$$\mathbf{f}_i\mathbf{B} = \mathbf{f}'\mathbf{B}\mathbf{F} + \mathbf{f}''_i\mathbf{B}.$$

Now notice that $\mathbf{f}'\mathbf{B}$ is distributed according to $D_{\sigma_1\mathbf{B}}$ and for all $i \in [k]$ it holds that $\mathbf{f}_i\mathbf{B}$ is distributed according to $D_{\sigma\mathbf{B}}$. Note that $\mathbf{e}'$ and $\mathbf{f}'\mathbf{B}$ are identically distributed, and also $\mathbf{e}_i$ and $\mathbf{f}_i\mathbf{B}$ are identically distributed for all $i \in [k]$. Setting Ψ to be the distribution of the $\mathbf{f}''\mathbf{B}$ the result follows.

Lemma 5.5 (Continuous to Discrete). *Let $\mathbf{Z}_0 \in \mathbb{Z}^{n\times n}$. Let τ_2 be the largest singular value of $\mathbf{Z}'_0 = \mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}$. Assume that $\sigma > \sqrt{2}\tau_2\eta_\epsilon(\mathbf{B}^{-1})$. Let $\mathbf{f} \sim D_{\sqrt{2}\sigma\mathbf{B}}$ and $\mathbf{e} \sim D_{\Lambda(\mathbf{Z}_0),\sigma\cdot\mathbf{B}}$. Let $\mathcal{S}$ be a random variable supported on $\mathbb{Z}_q^n$. Then it holds that*

$$\tilde{H}^{4\epsilon}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{f}\mathbf{Z}_0^{-1}) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}\mathbf{Z}_0^{-1}).$$

Proof. Let $\tilde{\mathbf{e}}' \sim D_{\sigma\mathbf{I}}$ be a spherical continuous gaussian and let $\tilde{\mathbf{e}}$ be distributed according to $D_{\Lambda(\mathbf{B}^{-1}\mathbf{Z}'_0),\sigma\mathbf{I}}$. By Claim 2.3 we have that $\tau_1 \cdot \eta_\epsilon(\mathbf{B}^{-1}) \geq \sigma_{max}(\mathbf{Z}'_0) \cdot \eta_\epsilon(\Lambda(\mathbf{B}^{-1})) \geq \eta_\epsilon(\Lambda(\mathbf{B}^{-1} \cdot \mathbf{Z}'_0))$. Now let $\tilde{\mathbf{f}} \sim D_{\sqrt{2}\sigma\mathbf{I}}$. Then it holds by Lemma 5.2 that $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{e}} + \tilde{\mathbf{e}}'$ are 4ϵ close.

Now note that by the definition of $\mathbf{e}$ we have that $\mathbf{e}$ and $\tilde{\mathbf{e}} \cdot \mathbf{B}$ are identically distributed, also $\mathbf{f}$ and $\tilde{\mathbf{f}} \cdot \mathbf{B}$ are identically distributed. Setting $\mathbf{e}' = \tilde{\mathbf{e}}'\mathbf{B}$, we obtain that $\mathbf{f}\mathbf{Z}_0^{-1}$ and $\mathbf{e}\mathbf{Z}_0^{-1} + \mathbf{e}'\mathbf{Z}_0^{-1}$ are 4ϵ-close. We can conclude that

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}\mathbf{Z}_0^{-1}) = \tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}\mathbf{Z}_0^{-1}, \mathbf{e}') \leq \tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}\mathbf{Z}_0^{-1}+\mathbf{e}'\mathbf{Z}_0^{-1}) \leq \tilde{H}^{4\epsilon}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{f}\mathbf{Z}_0^{-1}).$$

Lemma 5.6 (Discrete to Continuous). *Let $\mathbf{f} \leftarrow_\$ D_{\mathbb{Z}^n,\sqrt{2}\sigma\cdot\mathbf{B}}$ and $\mathbf{e} \leftarrow_\$ D_{\sigma\mathbf{B}}$, then it holds that*

$$\tilde{H}^{8\epsilon}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{f}) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}).$$

Proof. Let $\mathbf{e}'$ be distributed according to $D_{\mathbb{Z}^n-\mathbf{e},\sigma\mathbf{B}}$. Then it holds by Theorem 5.3 that the statistical distance between $\mathbf{e} + \mathbf{e}'$ and $\mathbf{f}$ is smaller than 8ϵ.

Theorem 5.7. *Let $\mathbf{Z}_0 \in \mathbb{Z}^{n\times n}$ and for $i \in [k]$ $\mathbf{Z}_i \in \mathbb{Z}^{n\times n}$ be matrices and let $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_k) \in \mathbb{Z}_q^{n\times nk}$ be the matrix obtained by concatenating the $\mathbf{Z}_i$. Further let $\mathbf{Z}_0^{-1} \in \mathbb{Q}^{n\times n}$ be the rational inverse of $\mathbf{Z}_0$ and $\mathbf{H} \in \mathbb{Z}_q^{n\times n}$ be the $\mathbb{Z}_q$-inverse of $\mathbf{Z}_0 \mod q$.*

Define the matrix $\mathbf{Z}'_0 = \mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}$ and $\mathbf{Z}' = (\mathbf{B}\mathbf{Z}_1\mathbf{B}^{-1}, \ldots, \mathbf{B}\mathbf{Z}_k\mathbf{B}^{-1})$. Let τ_1 be the largest singular value of ${\mathbf{Z}'_0}^{-1}\mathbf{Z}'$ and τ_2 be the largest singular value of $\mathbf{Z}'_0$. For a $\sigma > \tau_2\eta_\epsilon(\Lambda(\mathbf{B}^{-1}))$ let $\sigma_0 \geq 2^{3/2}\sigma \cdot \tau_1$. Then it holds that

$$\tilde{H}^{20\epsilon}_\infty(\mathbf{s}|\mathbf{s}\mathbf{H}\mathbf{Z} + \mathbf{e}_0) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}) - n\log(\tau_2),$$

where $\mathbf{e}_0 \leftarrow_\$ D^k_{\sigma_0\mathbf{B}}$ and $\mathbf{e} \leftarrow_\$ D_{\sigma\mathbf{B}}$.

Proof. Fix a distribution of secrets $\mathcal{S}$ and let $\mathbf{s} \leftarrow_{\$} \mathcal{S}$. Let $\sigma_1 = \sigma_0/\tau_1 \geq 2^{3/2}\sigma$

Since the largest singular value of ${\mathbf{Z}'_0}^{-1}\mathbf{Z}'$ is τ_1, by Lemma 5.4 there exists a distribution Ψ over $\mathbb{R}^{nk}$ such that we can equivalently sample $\mathbf{e}_0$ by $\mathbf{e}_0 = \mathbf{e}_1\mathbf{Z}_0^{-1}\mathbf{Z} + \mathbf{e}'_1$, where $\mathbf{e}_1 \sim D_{\sigma_1\mathbf{B}}$ and $\mathbf{e}'_1 \sim \Psi$. Consequently, we can write

$$\mathbf{y} = \mathbf{sHZ} + \mathbf{e}_0 = \mathbf{sHZ} + \mathbf{e}_1\mathbf{Z}_0^{-1}\mathbf{Z} + \mathbf{e}'_1 = (\mathbf{sH} + \mathbf{e}_1\mathbf{Z}_0^{-1})\mathbf{Z} + \mathbf{e}'_1.$$

Thus, since $\mathbf{y}$ can be computed from $\mathbf{sH} + \mathbf{e}_1\mathbf{Z}_0^{-1}$ and $\mathbf{e}'_1$ it follows that

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{sHZ} + \mathbf{e}_0) = \tilde{H}_\infty(\mathbf{s}|\mathbf{sH} + \mathbf{e}_1\mathbf{Z}_0^{-1}, \mathbf{e}'_1) = \tilde{H}_\infty(\mathbf{s}|\mathbf{sH} + \mathbf{e}_1\mathbf{Z}_0^{-1}),$$

where the second equality follows as $\mathbf{e}'_1$ is independent from $\mathbf{s}$ and $\mathbf{e}_1$.

Now let $\sigma_2 = \sigma_1/\sqrt{2} \geq 2\sigma$ and let $\mathbf{e}_2 \sim D_{\Lambda(\mathbf{Z}_0),\sigma_2\mathbf{B}}$ be a discrete gaussian. By Lemma 5.5 it holds that

$$\tilde{H}^{4\epsilon}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_1\mathbf{Z}_0^{-1}) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_2\mathbf{Z}_0^{-1}).$$

Now, since $\mathbf{H}$ is the $\mathbb{Z}_q$-inverse of $\mathbf{Z}_0 \mod q$, multiplying $\mathbf{sH} + \mathbf{e}_2\mathbf{Z}_0^{-1}$ by $\mathbf{Z}_0$ yields

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{sH} + \mathbf{e}_2\mathbf{Z}_0^{-1}) = \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_2).$$

Now let $\sigma_3 = \sigma_2/\sqrt{2} \geq \sqrt{2}\sigma$, $\mathbf{e}_3 \sim D_{\mathbb{Z}^n,\sigma_3\mathbf{B}}$ and $\mathbf{e}'_3 \sim D_{\Lambda(\mathbf{Z}_0)-\mathbf{e}_3,\sigma_3\mathbf{B}}$. Setting $\Lambda_2 = \mathbb{Z}^n$ and $\Lambda_1 = \Lambda(\mathbf{Z}_0)$ in Theorem 5.3 and noting that $\sigma_3 > \sigma > \eta_\epsilon(\Lambda(\mathbf{B}^{-1}))$ we obtain that the statistical distance between $\mathbf{e}_2$ and $\mathbf{e}_3 + \mathbf{e}'_3$ is at most 8ϵ.

It follows that

$$\tilde{H}^{8\epsilon}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_2) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3 + \mathbf{e}'_3) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3, \mathbf{e}'_3).$$

Since $\mathbf{e}'_3$ is distributed according to $D_{\Lambda(\mathbf{Z}_0)-\mathbf{e}_3,\sigma_3}$, it only depends on $\mathbf{e}_3 \mod \Lambda(\mathbf{Z}_0)$. Thus

$$\begin{aligned}\tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3, \mathbf{e}'_3) &\geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3) - H_0(\mathbf{e}'_3)\\ &\geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3) - \log(\det(\mathbf{Z}_0))\\ &\geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3) - n \cdot \log(\tau_2),\end{aligned}$$

as $|\mathbb{Z}^n/\Lambda(\mathbf{Z}_0)| = \det(\mathbf{Z}_0) = \det(\mathbf{Z}'_0) \leq n \cdot \log(\tau)$ (as the largest singular value of $\mathbf{Z}'_0$ is τ_2).

Finally as $\sigma_3/\sqrt{2} = \sigma > \eta_\epsilon(\mathbf{B}^{-1})$, by Lemma 5.5 we can bound

$$\tilde{H}^{8\epsilon}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_3) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}),$$

where $\mathbf{e} \leftarrow_{\$} D_{\sigma\mathbf{B}}$. Putting everything together, we obtain that

$$\tilde{H}^{20\epsilon}_\infty(\mathbf{s}|\mathbf{sHZ} + \mathbf{e}_0) \geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}) - n \cdot \log(\tau_2).$$

We can now summarize the results of this section in the following theorem.

Theorem 5.8. *Let $\tau_1, \tau_2 > 0$. Let Ψ be a distribution on $(\mathbb{Z}^{n\times n})^\times \times \mathbb{Z}^{n\times nk}$ and assume the Decisional Small Ratio assumption holds for Ψ. Assume further that if $(\mathbf{Z}_0, \mathbf{Z}) \leftarrow_{\$} \Psi$ then*

- $\sigma_{max}(\mathbf{B}\mathbf{Z}_0^{-1}\mathbf{Z}\mathbf{B}^{-1}) \leq \tau_1$ *where* $\mathbf{Z}_0^{-1}$ *is the rational inverse of* $\mathbf{Z}_0$.
- $\sigma_{max}(\mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}) \leq \tau_2$

with probability at least δ *over the choice of* $(\mathbf{Z}_0, \mathbf{Z})$. *Define the distribution* $\mathcal{X}$ *on* $\mathbb{Z}_q^{n\times m}$ *by* $\mathbf{HZ}$, *where* $(\mathbf{Z}_0, \mathbf{Z}) \leftarrow_\$ \Psi$ *and* $\mathbf{H} \in \mathbb{Z}^{n\times n}$ *is the* $\mathbb{Z}_q$*-inverse of* $\mathbf{Z}_0$. *Let* $\sigma > \tau_2\eta_\epsilon(\Lambda(\mathbf{B}^{-1}))$ *and* $\sigma_0 > 2^{3/2}\tau_1\sigma$. *Now let* $\chi = D_{\sigma_0\mathbf{B}}$. *Further assume that* $\nu_{\sigma\mathbf{B}}(\mathcal{S}) \geq n\log(\tau_2) + \omega(\log(\lambda))$.

Then $\mathcal{X}$ *is a sometimes lossy pseudorandom distribution for* $\mathcal{S}$ *and error distribution* χ.

By combining Theorems 5.8 and 4.2 we obtain the following corollary.

Corollary 5.9. *Assume that the conditions of Theorem 5.8 are satisfied. Then* $\mathsf{entSLWE}(q, k, \mathcal{M}, D_{\sigma_0\mathbf{B}}, \mathcal{S})$ *is mildly hard.*

6 Instantiation for RLWE over Power-of-Two Cyclotomics

In this Section, we will instantiate the results of Section 5 for Ring LWE over power-of-two cyclotomics. That is, we will construct a sometimes lossy pseudorandom distribution in this setting.

Throughout this section let $\mathbf{B} \in \mathbb{R}^{n\times n}$ be a basis-change matrix as described in Section 3.2.

First recall the Decisional Small Polynomial Ratio (DSPR) problem, as defined by Lopez-Alt et al. [15]. The DSPR problem is in fact a generalization of the NTRU problem.

Definition 6.1 (Decisional Small Polynomial Ratio problem (DSPR)). *Let* R *be a ring of integers of a number field* K *and let* q *be a modulus. Let* $\gamma > 0$. *Let* $\mathbf{g} \leftarrow_\$ D_{\mathsf{R},\gamma}$ *and* $\mathbf{f} \leftarrow_\$ D_{\mathsf{R},\gamma}$ *conditioned on* $\mathbf{f} \mod q \in \mathsf{R}_q^\times$. *Let* $\mathbf{h}$ *be the* R_q*-inverse of* $\mathbf{f}$. *The DSPR problem for distribution* $D_{\mathsf{R},\gamma}$ *asks to distinguish* $\mathbf{hg} \in \mathsf{R}_q$ *from a uniformly random* $\mathbf{a} \leftarrow_\$ \mathsf{R}_q$.

We will make use of the following Lemmas and Theorems of Stehlé and Steinfeld [27].

Theorem 6.2 shows that if the a gaussian χ is sufficiently wide, then ring elements $\mathbf{hg}$ are actually statistically close to a uniform $\mathbf{a} \leftarrow_\$ \mathsf{R}_q$.

Theorem 6.2 ([27, Theorem 3.2 restated]). *Let* $n \geq 8$ *be a power of 2 such that* $\Phi = X^n + 1$ *splits into* n *linear factors modulo a prime* $q \geq 5$. *Let* $0 < \alpha < 1/3$ *and assume that* $\gamma \geq n \cdot \sqrt{\ln(8nq)} \cdot q^{1/2+\alpha}$ *and that* $\mathbf{f}, \mathbf{g} \leftarrow_\$ D_{\mathsf{R}_q^\times,\gamma}$. *Let* $\mathbf{h}$ *be the* R_q*-inverse of* $\mathbf{f}$. *Then it holds that* $\mathbf{hg}$ *is within statistical distance* $2^{10n} \cdot q^{-\alpha n}$ *of the uniform distribution on* $\mathsf{R}_q^\times$.

Lemma 6.3 ([27, Lemma 3.5 restated]). *Let* $n \geq 8$ *be a power of 2 such that* $\Phi = X^n + 1$ *splits into* n *linear factors modulo* $q \geq 5$. *Let* $\gamma \geq \sqrt{n \cdot \ln(2n(1+n^2))/\pi} \cdot q^{1/n}$. *Then it holds that*

$$\Pr_{\mathbf{f}\leftarrow_\$ D_{\mathsf{R},\gamma}}[\mathbf{f} \notin \mathsf{R}_q^\times] \leq n(1/q + 2/n^2).$$

Lemma 6.4 ([27, Lemma 2.8 restated]). *Let* R *be a ring of integers. Then it holds for any* $\gamma \geq \eta_\epsilon(\mathsf{R})$ *that*

$$\Pr_{\mathbf{f} \leftarrow_\$ D_{\mathsf{R},\gamma}}[\|\mathbf{f}\| \geq \gamma \log(n)\sqrt{n}] \leq \mathsf{negl}(\lambda)$$

Lemma 6.5 ([27, Lemma 4.1 restated]). *Let* $n \geq 8$ *be a power of 2,* $\Phi = X^n + 1$ *and* $\mathsf{R} = \mathbb{Z}[X]/(\Phi)$*. For any* $\gamma \geq 8n\eta_\epsilon(\mathsf{R})$ *it holds that*

$$\Pr_{\mathbf{f} \leftarrow_\$ D_{\mathsf{R},\gamma}}\left[\|\mathbf{f}^{-1}\| \geq \frac{24\sqrt{n}}{\gamma}\right] \leq 1/2$$

We will now establish the hardness of an instance of the DSR problem, assuming RLWE and either the DSPR problem or Theorem 6.2. Let χ be a B-bounded error distribution on R and let $\gamma > 0$ be a gaussian parameter. Define the distribution Ψ as follows:

- Choose $\mathbf{f}, \mathbf{g} \leftarrow_\$ D_{\mathsf{R},\gamma}$ such that $\mathbf{f} \mod q \in \mathsf{R}_q^\times$.
- Choose $\mathbf{e}_1, \ldots, \mathbf{e}_k \leftarrow_\$ \chi$ and $\mathbf{e}'_1, \ldots, \mathbf{e}'_k \leftarrow_\$ \chi$
- For all $i \in [k]$ set $\mathbf{z}_i = \mathbf{g} \cdot \mathbf{e}_i + \mathbf{f} \cdot \mathbf{e}'_i$.
- Let $\mathbf{Z}_0$ be the multiplication matrix of $\mathbf{f}$ and for all $i \in [k]$ let $\mathbf{Z}_i$ be the multiplication matrix of $\mathbf{z}_i$
- Set $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_k)$
- Output $(\mathbf{Z}_0, \mathbf{Z})$

We will now show that the distribution Ψ is a sometimes lossy pseudorandom distribution. Recall that by Theorem 5.8 it is sufficient to bound the maximal singular values of $\mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}$, $\mathbf{B}\mathbf{Z}_0^{-1}\mathbf{Z}\mathbf{B}^{-1}$ and establish that the DSR assumption for Ψ holds. We will start by showing that if the Ring LWE assumption for error distribution χ and the DSPR assumptions hold, then the DSR assumption holds for Ψ.

Lemma 6.6. *Assuming both DSPR for distribution* $D_{\mathsf{R},\gamma}$ *and RLWE for error-distribution* χ*, it follows that DSR for distribution* Ψ *is hard. Moreover, if* $\chi = D_{\mathsf{R}_q^\times,\gamma}$ *and the conditions of Theorem 6.2 are met, then the DSPR assumption is not necessary.*

Proof. Let $\mathbf{h}$ be the R_q-inverse of $\mathbf{f}$. Observe that $\mathbf{y}_i = \mathbf{h}\mathbf{z}_i = \mathbf{h}\mathbf{g} \cdot \mathbf{e}_i + \mathbf{e}'_i$. Under the DSPR assumption we can replace $\mathbf{h}\mathbf{g}$ by a uniformly random $\mathbf{a} \in \mathsf{R}_q$. It then follows by a simple hybrid argument that for all i $\mathbf{y}_i = \mathbf{a}\mathbf{e}_i + \mathbf{e}'_i$ is indistinguishable from a uniformly random $\mathbf{u}_i$ under Hermite RLWE for error distribution χ'.

Likewise, if the conditions of Theorem 6.2 are met, $\mathbf{h}\mathbf{g}$ is statistically close to a uniformly random $\mathbf{a} \in \mathsf{R}_q^\times$. It follows again via a hybrid argument that for all i $\mathbf{y}_i = \mathbf{a}\mathbf{e}_i + \mathbf{e}'_i$ is indistinguishable from a uniformly random $\mathbf{u}_i$ under Hermite RLWE for error distribution χ'. Note that RLWE also holds if we condition on $\mathbf{a} \in \mathsf{R}_q^\times$, as this event happens with significant probability.

The following technical lemma lets us bound the maximal singular value of a matrix $\mathbf{Z}'$ by bounding the singular values of *blocks* of $\mathbf{Z}'$.

Lemma 6.7. *Let $\mathbf{Z}' = (\mathbf{Z}'_1|\dots|\mathbf{Z}'_m) \in \mathbb{R}^{n\times n\cdot k}$ be a block matrix where each $\mathbf{Z}'_i \in \mathbb{R}^{n\times n}$. Assume that it holds for all i that $\sigma_{max}(\mathbf{Z}'_i) \leq \gamma$. Then it holds that $\sigma_{max}(\mathbf{Z}') \leq \sqrt{k}\cdot\gamma$.*

Proof. Fix any vector $\mathbf{x} = (\mathbf{x}_1,\dots,\mathbf{x}_k) \in \mathbb{R}^{nk}$, where the $\mathbf{x}_i \in \mathbb{R}^n$. Then it holds that

$$\|\mathbf{Z}'\mathbf{x}\| = \|\sum_{i=1}^{k}\mathbf{Z}'_i\mathbf{x}_i\| \leq \sum_{i=1}^{k}\|\mathbf{Z}'_i\mathbf{x}_i\| \leq \sum_{i=1}^{k}\gamma\|\mathbf{x}_i\| \leq \gamma\sqrt{k}\cdot\sqrt{\sum_{i=1}^{k}\|\mathbf{x}_i\|^2} = \gamma\sqrt{k}\cdot\|\mathbf{x}\|,$$

where the last inequality follows from the relationship between the L_1 and L_2 norms. It follows that $\sigma_{max}(\mathbf{Z}') \leq \sqrt{k}\cdot\gamma$

Lemma 6.8 bounds the maximal singular values of $\mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}$ and $\mathbf{B}\mathbf{Z}_0^{-1}\mathbf{Z}\mathbf{B}^{-1}$

Lemma 6.8. *Let $\gamma > \max\{\sqrt{n\cdot\ln(2n(1+n^2))/\pi}\cdot q^{1/n}, 8n\eta_\epsilon(\mathsf{R})\}$ and assume that χ is B-bounded. Let $(\mathbf{Z}_0,\mathbf{Z}) \leftarrow_\$ \Psi$. It holds that*

- *$\mathbf{Z}_0$ is invertible in $\mathbb{Z}_q^{n\times n}$*
- *$\sigma_{max}(\mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}) \leq O(\gamma\log(n)\sqrt{n})$*
- *$\sigma_{max}(\mathbf{B}\mathbf{Z}_0^{-1}\mathbf{Z}\mathbf{B}^{-1}) \leq O(n\log(n)\sqrt{k}B)$*

except with probability $1/2+o(1)$ over the choice of $(\mathbf{Z}_0,\mathbf{Z})$.

Proof. We will bound the maximal singular value of $\mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}$ by $\|\sigma(\mathbf{f})\|_\infty$. Likewise, we will bound the maximal singular values of the $\mathbf{B}\mathbf{Z}_0^{-1}\mathbf{Z}_i\mathbf{B}^{-1}$ via $\|\sigma(\mathbf{f}^{-1}\mathbf{z}_i)\|_\infty$. The bound on the maximal singular value of $\mathbf{B}\mathbf{Z}\mathbf{B}^{-1}$ will follow by Lemma 6.7.

Note that

- It holds by Lemma 6.3 that $\mathbf{f}$ in invertible in $\mathsf{R}_q^\times$, except with probability $n/q + 2/n = O(1/n)$.
- It holds by Lemma 6.4 and a union bound that $\|\sigma(\mathbf{f})\| \leq \gamma\log(n)\sqrt{n}$ and $\|\sigma(\mathbf{z}_i)\| \leq \gamma\log(n)\sqrt{n}$ for all $i\in[k]$, except with negligible probability.
- By Lemma 6.5 we have that $\|\sigma(\mathbf{f}^{-1})\| \leq 24\sqrt{n}/\gamma$, except with probability $\leq 1/2$.

Consequently, all 3 items hold, except with probability $1/2+o(1)$. Moreover, since the $\mathbf{e}_i$ and $\mathbf{e}'_i$ are distributed according to χ and χ is B-bounded, it holds that for all $i\in[k]$ that $\|\sigma(\mathbf{e}_i)\| \leq B$ and $\|\sigma(\mathbf{e}'_i)\| \leq B$.

Thus, we have that

$$\sigma_{max}(\mathbf{B}\mathbf{Z}_0\mathbf{B}^{-1}) \leq \|\sigma(\mathbf{f})\|_\infty \leq \|\sigma(\mathbf{f})\| \leq \gamma\log(n)\sqrt{n} = O(\gamma\log(n)\sqrt{n}).$$

Moreover, it holds for all i that

$$\begin{aligned}\sigma_{max}(\mathbf{B}\mathbf{Z}_0^{-1}\mathbf{Z}_i\mathbf{B}^{-1}) &\leq \|\sigma(\mathbf{f}^{-1}\mathbf{z}_i)\|_\infty \\ &= \|\sigma(\mathbf{f}^{-1}\mathbf{g}\mathbf{e}_i + \mathbf{e}'_i)\|_\infty \\ &\leq \|\sigma(\mathbf{f}^{-1})\|_\infty \cdot \|\sigma(\mathbf{g})\|_\infty \cdot \|\sigma(\mathbf{e}_i)\|_\infty + \|\sigma(\mathbf{e}'_i)\|_\infty \\ &\leq \|\sigma(\mathbf{f}^{-1})\| \cdot \|\sigma(\mathbf{g})\| \cdot \|\sigma(\mathbf{e}_i)\| + \|\sigma(\mathbf{e}'_i)\| \\ &\leq 24n\log(n) \cdot B = O(n\log(n)B)\end{aligned}$$

By Lemma 6.7 we conclude that $\sigma_{max}(\mathbf{B}\mathbf{Z}\mathbf{B}^{-1}) \leq O(n\log(n)\sqrt{k}B)$.

We can now summarize the results of this section in our main theorem by combining Lemma 6.8 with Corollary 5.9.

Theorem 6.9. *Assume that DSPR with parameter γ and Ring LWE with a B-bounded noise distribution χ holds. Let $\mathcal{S}$ be a distribution s.t. for some σ it holds that $\nu_\sigma(\mathcal{S}) \geq n\log(\gamma \cdot \log(n)\sqrt{n}) + \omega(\log\lambda)$. Then Entropic Ring LWE for power-of-two cyclotomics with k samples, secret distribution $\mathcal{S}$ and Gaussian noise parameter $\sigma_0 \geq O(\sigma n\log(n)B\sqrt{k})$ is mildly hard.*

By Theorem 6.2 we know that we can drop the DSPR assumption provided that $\gamma \geq \mathsf{poly}(n)q^{1/2+\alpha}$ for an arbitrarily small constant α. This translates to the stronger requirement that $\nu_\sigma(\mathcal{S}) \geq (1/2+\alpha)n\log(q) + O(n\log(n))$. Thus, the distribution $\mathcal{S}$ must have at more than $(1/2+\alpha)n\log(q) + O(n\log(n))$ min-entropy to begin with. However, note that if $\mathcal{S}$ is an r-bounded distribution, where $r \geq \mathsf{poly}(n)q^{1/2+\alpha}$, then Lemma 2.5 tells us if σ is a $\mathsf{poly}(n)$ factor larger than r, we have essentially $\nu_\sigma(\mathcal{S}) \approx \tilde{H}_\infty(\mathcal{S})$ and the requirements can be met.

References

1. Albrecht, M.R., Deo, A.: Large modulus ring-LWE ≥ module-LWE. In: Takagi, T., Peyrin, T. (eds.) Advances in Cryptology – ASIACRYPT 2017, Part I. Lecture Notes in Computer Science, vol. 10624, pp. 267–296. Springer, Heidelberg, Germany, Hong Kong, China (Dec 3–7, 2017)
2. Alwen, J., Krenn, S., Pietrzak, K., Wichs, D.: Learning with rounding, revisited - new reduction, properties and applications. In: Canetti, R., Garay, J.A. (eds.) Advances in Cryptology – CRYPTO 2013, Part I. Lecture Notes in Computer Science, vol. 8042, pp. 57–74. Springer, Heidelberg, Germany, Santa Barbara, CA, USA (Aug 18–22, 2013)
3. Bolboceanu, M., Brakerski, Z., Perlman, R., Sharma, D.: Order-lwe and the hardness of ring-lwe with entropic secrets. Asiacrypt (2019), https://eprint.iacr.org/2018/494
4. Brakerski, Z., Döttling, N.: Hardness of LWE on general entropic distributions. In: Canteaut, A., Ishai, Y. (eds.) Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part II. Lecture Notes in Computer Science, vol. 12106, pp. 551–575. Springer (2020), https://doi.org/10.1007/978-3-030-45724-2_19

5. Brakerski, Z., Gentry, C., Vaikuntanathan, V.: (Leveled) fully homomorphic encryption without bootstrapping. In: Goldwasser, S. (ed.) ITCS 2012: 3rd Innovations in Theoretical Computer Science. pp. 309–325. Association for Computing Machinery, Cambridge, MA, USA (Jan 8–10, 2012)
6. Brakerski, Z., Langlois, A., Peikert, C., Regev, O., Stehlé, D.: Classical hardness of learning with errors. In: Boneh, D., Roughgarden, T., Feigenbaum, J. (eds.) 45th Annual ACM Symposium on Theory of Computing. pp. 575–584. ACM Press, Palo Alto, CA, USA (Jun 1–4, 2013)
7. Dodis, Y., Ostrovsky, R., Reyzin, L., Smith, A.D.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. SIAM J. Comput. 38(1), 97–139 (2008)
8. Goldwasser, S., Kalai, Y.T., Peikert, C., Vaikuntanathan, V.: Robustness of the learning with errors assumption. In: Yao, A.C.C. (ed.) ICS 2010: 1st Innovations in Computer Science. pp. 230–240. Tsinghua University Press, Tsinghua University, Beijing, China (Jan 5–7, 2010)
9. Halevi, S., Shoup, V.: Algorithms in helib. In: Garay, J.A., Gennaro, R. (eds.) Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I. Lecture Notes in Computer Science, vol. 8616, pp. 554–571. Springer (2014), http://dx.doi.org/10.1007/978-3-662-44371-2_31
10. Halevi, S., Shoup, V.: Bootstrapping for helib. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I. Lecture Notes in Computer Science, vol. 9056, pp. 641–670. Springer (2015), http://dx.doi.org/10.1007/978-3-662-46800-5_25
11. Hoffstein, J., Pipher, J., Silverman, J.H.: NTRU: A ring-based public key cryptosystem. In: ANTS. Lecture Notes in Computer Science, vol. 1423, pp. 267–288. Springer (1998)
12. Kalai, Y.T., Reyzin, L.: A survey of leakage-resilient cryptography. IACR Cryptol. ePrint Arch. 2019, 302 (2019)
13. Kirchner, P., Fouque, P.A.: Revisiting lattice attacks on overstretched ntru parameters. In: Coron, J.S., Nielsen, J.B. (eds.) Advances in Cryptology – EUROCRYPT 2017. pp. 3–26. Springer International Publishing, Cham (2017)
14. Langlois, A., Stehlé, D.: Worst-case to average-case reductions for module lattices. Des. Codes Cryptogr. 75(3), 565–599 (2015), https://doi.org/10.1007/s10623-014-9938-4
15. López-Alt, A., Tromer, E., Vaikuntanathan, V.: On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption. In: Karloff, H.J., Pitassi, T. (eds.) 44th Annual ACM Symposium on Theory of Computing. pp. 1219–1234. ACM Press, New York, NY, USA (May 19–22, 2012)
16. Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. In: Gilbert, H. (ed.) Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, French Riviera, May 30 - June 3, 2010. Proceedings. Lecture Notes in Computer Science, vol. 6110, pp. 1–23. Springer (2010), https://doi.org/10.1007/978-3-642-13190-5_1
17. Lyubashevsky, V., Peikert, C., Regev, O.: A toolkit for ring-lwe cryptography. In: Johansson, T., Nguyen, P.Q. (eds.) Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of

Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings. Lecture Notes in Computer Science, vol. 7881, pp. 35–54. Springer (2013), https://doi.org/10.1007/978-3-642-38348-9_3
18. Micciancio, D.: On the hardness of learning with errors with binary secrets. Theory of Computing 14(1), 1–17 (2018)
19. NIST: Post-quantum cryptography standardization, https://csrc.nist.gov/Projects/Post-Quantum-Cryptography
20. Peikert, C.: An efficient and parallel Gaussian sampler for lattices. In: Rabin, T. (ed.) Advances in Cryptology – CRYPTO 2010. Lecture Notes in Computer Science, vol. 6223, pp. 80–97. Springer, Heidelberg, Germany, Santa Barbara, CA, USA (Aug 15–19, 2010)
21. Peikert, C.: A decade of lattice cryptography. Foundations and Trends in Theoretical Computer Science 10(4), 283–424 (2016), https://doi.org/10.1561/0400000074, specific references are to the ePrint version https://eprint.iacr.org/2015/939
22. Peikert, C., Pepin, Z.: Algebraically structured LWE, revisited. In: TCC 2019: 17th Theory of Cryptography Conference, Part I. pp. 1–23. Lecture Notes in Computer Science, Springer, Heidelberg, Germany (Mar 2019)
23. Peikert, C., Regev, O., Stephens-Davidowitz, N.: Pseudorandomness of ring-lwe for any ring and modulus. In: Hatami, H., McKenzie, P., King, V. (eds.) Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017. pp. 461–473. ACM (2017), https://doi.org/10.1145/3055399.3055489
24. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. In: Gabow, H.N., Fagin, R. (eds.) 37th Annual ACM Symposium on Theory of Computing. pp. 84–93. ACM Press, Baltimore, MA, USA (May 22–24, 2005)
25. Microsoft SEAL. http://sealcrypto.org (Oct 2018), microsoft Research, Redmond, WA.
26. Stehlé, D., Steinfeld, R.: Making NTRU as secure as worst-case problems over ideal lattices. In: Paterson, K.G. (ed.) Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings. Lecture Notes in Computer Science, vol. 6632, pp. 27–47. Springer (2011), https://doi.org/10.1007/978-3-642-20465-4_4
27. Stehlé, D., Steinfeld, R.: Making ntruencrypt and ntrusign as secure as standard worst-case problems over ideal lattices. IACR Cryptol. ePrint Arch. 2013, 4 (2013), http://eprint.iacr.org/2013/004, preliminary version in [26]
28. Stehlé, D., Steinfeld, R., Tanaka, K., Xagawa, K.: Efficient public key encryption based on ideal lattices. In: Matsui, M. (ed.) Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings. Lecture Notes in Computer Science, vol. 5912, pp. 617–635. Springer (2009), https://doi.org/10.1007/978-3-642-10366-7_36