Hardness of LWE on General Entropic Distributions

Zvika Brakerski [*] Nico Döttling[†]

Abstract

The hardness of the Learning with Errors (LWE) problem is by now a cornerstone of the cryptographic landscape. In many of its applications the so called "LWE secret" is not sampled uniformly, but comes from a distribution with some min-entropy. This variant, known as "Entropic LWE", has been studied in a number of works, starting with Goldwasser et al. (ICS 2010). However, so far it was only known how to prove the hardness of Entropic LWE for secret distributions supported inside a ball of small radius.

In this work we resolve the hardness of Entropic LWE with arbitrary long secrets, in the following sense. We show an entropy bound that guarantees the security of arbitrary Entropic LWE. This bound is higher than what is required in the ball-bounded setting, but we show that this is essentially tight. Tightness is shown unconditionally for highly-composite moduli, and using black-box impossibility for arbitrary moduli.

Technically, we show that the entropic hardness of LWE relies on a simple to describe lossiness property of the distribution of secrets itself. This is simply the probability of recovering a random sample from this distribution s, given $s + e$, where e is Gaussian noise (i.e. the quality of the distribution of secrets as an error correcting code for Gaussian noise). We hope that this characterization will make it easier to derive entropic LWE results more easily in the future. We also use our techniques to show new results for the ball-bounded setting, essentially showing that under a strong enough assumption even polylogarithmic entropy suffices.

1 Introduction

Lattice-based cryptography has emerged in the last few decades as one of the most important developments in cryptography. Lattice-based cryptographic schemes have been shown to achieve functionalities that are unknown under any other cryptographic structure (such as fully homomorphic encryption [Gen09, BV11], attribute-based encryption for circuits [GVW13] and many others). At the same time, it is possible in many cases to show strong security properties such as worst-case to average-case hardness results [Ajt96, AD97, MR04, Reg05] that relate the hardness of breaking the cryptographic scheme to that of solving approximate short-vector problems in worst-case lattices, a problem that resists algorithmic progress even when when use of quantum computers is considered.

Much of the progress in advancing lattice-based cryptography can be attributed to the hardness of the Learning with Errors (LWE) problem, introduced by Regev [Reg05]. This problem can be stated in a very clean linear-algebraic syntax, which allows to utilize it for applications very easily,

[*]Weizmann Institute of Science, Israel, zvika.brakerski@weizmann.ac.il. 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).

[†]CISPA Helmholtz Center for Information Security, doettling@cispa.saarland.

and at the same time was shown to enjoy worst-case hardness as explained above. An instance of the LWE problem has the following form. It is parameterized by a dimension n and modulus $q \gg n$. Consider the following distribution. Sample a (public) random matrix $\mathbf{A} \in \mathsf{Z}_q^{n \times m}$, for arbitrary $m = \mathsf{poly}(n)$, and a (secret) random vector $\mathbf{s} \in \mathsf{Z}_q^n$, and output $(\mathbf{A}, \mathbf{y})$, where $\mathbf{y} = \mathbf{s}\mathbf{A} + \mathbf{e}$ (mod q), and $\mathbf{e}$ is a noise vector selected from some distribution (often a Gaussian with parameter $\sigma \ll q$). The goal of the LWE solver is to find $\mathbf{s}$ given $(\mathbf{A}, \mathbf{y})$, where m can be as large as the adversary desires. In the most straightforward use of this assumption for cryptography (suggested in Regev's original paper), $(\mathbf{A}, \mathbf{y})$ are used as public key for an encryption scheme, and $\mathbf{s}$ is the secret key. Similar roles are assumed in other cryptographic constructions.

Goldwasser et al. [GKPV10] initiated a study on the hardness of LWE when $\mathbf{s}$ is not chosen uniformly at random. This study was motivated by the desire to achieve an *entropic* notion of security that will allow to guarantee that the problem remains hard even if some information about $\mathbf{s}$ is leaked. They showed that if $\mathbf{s}$ is sampled from a binary distribution (i.e. supported over $\{0,1\}^n$), then LWE remains hard so long as $\mathbf{s}$ has sufficient entropy. In fact, sampling $\mathbf{s}$ from a (possibly sparse) binary distribution is attractive in other contexts such as constructing efficient post-quantum cryptographic objects [NIS], minimizing noise blowup in homomorphic encryption [BGV12], classical worst-case to average-case reduction [BLP+13] and proving hardness for the so-called Learning with Rounding (LWR) problem [BPR12, BGM+16]. Progress on understanding entropic LWE in the binary setting was made in subsequent works [BLP+13, Mic18].

However, the question of hardness of LWE on imperfect secret distributions carries significance beyond the binary setting. If we consider the key-leakage problem, then changing the honest key distribution to be binary just for the sake of improving robustness against key-leakage carries a heavy cost in the performance and security features in case no leakage occurs. An entropic hardness result for the general uniform setting is thus a natural question. Furthermore, for a problem as important as LWE, the mere scientific understanding of the robustness of the problem to small changes in the prescribed distributions and parameters stands as a self-supporting goal.

Alas, it appears that current approaches provide no insight for the general setting. Existing results can be extended beyond the binary setting so long as the norm of the vectors $\mathbf{s}$ is bounded, i.e. so long as the secret distribution is contained within some small enough ball, as was made explicit by Alwen et al. [AKPW13]. However this appeared to be an artifact of the proof technique and it was speculated by some that a general entropic LWE result should exist. Exploring the hardness of general entropic LWE is the goal of this work.

1.1 Our Results

We relate the hardness of Entropic LWE for arbitrary distributions to a basic property of the distribution, specifically to how *bad* the distribution performs as an error correcting code against Gaussian noise. Specifically, let $\mathcal{S}$ be some distribution over secrets in Z_q^n. Recall the notion of conditional smooth min-entropy $\tilde{H}_\infty$ and define the *noise lossiness* of $\mathcal{S}$ as

$$\nu_\sigma(\mathcal{S}) = \tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}) = -\log\left(\Pr_{\mathbf{s},\mathbf{e}}[\mathcal{A}^*(\mathbf{s}+\mathbf{e}) = \mathbf{s}]\right) , \tag{1}$$

where $\mathbf{s}$ is sampled from $\mathcal{S}$ and $\mathbf{e}$ is (continuous, say) Gaussian noise with parameter σ, and $\mathcal{A}^*$ is the optimal maximal likelihood decoder for $\mathbf{s}$, namely $\mathcal{A}^*(\mathbf{y}) = \arg\max_{\mathbf{s}} \Pr_{\mathbf{s},\mathbf{e}}[\mathbf{s}|\mathbf{y} = \mathbf{s}+\mathbf{e}]$. This notion is a min-entropy analogue to the notion of *equivocation* for Shannon-entropy, and can be seen as a *guaranteed information loss* of a gaussian channel (rather than average information loss).

We advocate for noise lossiness as a new and natural measure for a distribution and show that it allows to get a good handle on the entropic LWE question. We do this by showing that distributions with sufficiently high noise lossiness lead to hard instances of Entropic LWE (under assumptions, see details below). We then show that high min-entropy implies (some limited level of) noise lossiness, which allows us to derive hardness results for general Entropic LWE. We furthermore show that results for distributions supported inside a ball can also be derived using our technique and show that noise lossiness of such distributions is larger than that of general distributions.[1] Finally, we show that our bounds for the general entropic setting are essentially tight. See below for details.

Noise Lossiness Implies Entropic LWE Hardness (Section 4). We show that high noise lossiness implies entropic hardness. Our result relies on the hardness of the decision version of LWE (with "standard" secret distribution). Whereas the variant we discussed so far is the search variant, which asserts that finding $\mathbf{s}$ given $(\mathbf{A}, \mathbf{y})$ should be hard, the decision variant dLWE asserts that it is computationally hard to even distinguish $(\mathbf{A}, \mathbf{y})$ from $(\mathbf{A}, \mathbf{u})$ where $\mathbf{u} \in \mathsf{Z}_q^m$ is uniform. The hardness of decision LWE immediately implies hardness for search LWE, and the converse is also true but not for every noise distribution and via a reduction that incurs some cost. This is also the case in the entropic setting. By default when we refer to (Entropic) LWE in this work, we refer to the search version. We will mention explicitly when referring to the decision version.

Our results in this setting are as follows.

Theorem 1.1 (Main Theorem, Informal)**.** *Assume that decision LWE with dimension k, modulus q and Gaussian noise parameter γ is hard. Let $\mathcal{S}$ be a distribution over Z_q^n with $\nu_{\sigma_1}(\mathcal{S}) \geq k \log(q) + \omega(\log \lambda)$ for some parameter σ_1, then Entropic LWE with secret distribution $\mathcal{S}$ and Gaussian noise parameter $\sigma \approx \sigma_1 \gamma \sqrt{m}$ is hard.*

Our actual theorem is even more expressive on two aspects. First, while the above result applies for search Entropic LWE for all values of q, but in some cases, e.g. when q is prime, it also applies to decision Entropic LWE. Second, in the case where $\mathcal{S}$ is supported inside a ball, the term $k \log(q)$ can be relaxed to roughly $k \log(\gamma r)$ where r is the radius of the ball (this only applies to the search version).

We note that we incur a loss in noise that depends on $\sqrt{m}$, i.e. depends on the number of LWE samples. This is inherent in our proof technique, but using known statistical or computational rerandomization results, this dependence can be replaced by dependence on n, γ.

As explained above, most of our results imply hardness for search Entropic LWE and do not directly imply hardness for the decision version (albeit search-to-decision reductions can be applied, as we explained below). We note that this is an artifact of the applicability of our proof technique even in cases where the decision problem is not hard at all. We view this as a potentially useful property which may find future applications. To illustrate, consider the setting where the distributions of $\mathbf{s}$ and $\mathbf{e}$, as well as the modulus q, are all even. (Indeed, usually we consider the coordinates of $\mathbf{e}$ to be continuous Gaussians or a discrete Gaussians over Z, but one may be interested in a setting where they are, say, discrete Gaussian over $2\mathsf{Z}$.) In this setting, decision LWE is trivially easy, but search LWE remains hard. Our techniques (as detailed in the technical overview below) naturally extend to this setting and can be used to prove entropic hardness in this case as well.

[1] In fact, noise lossiness provides a simple intuitive explanation on why ball-bounded distributions with given min-entropy yield harder Entropic LWE instances than general ones. This is due to the fact that packing the same number of elements in a small ball necessarily makes it harder to go back to the point of origin once noise is added.

In the standard regime of parameters, where $\mathbf{e}$ is a continuous Gaussian, we can derive the hardness of the decision problem using known search-to-decision reductions. The most generic version, as in e.g. [Reg05], runs in time $q \cdot \mathsf{poly}(n)$ but in many cases the dependence on q can be eliminated [Pei09, MM11]. In particular we note that in the ball-bounded setting, search-to-decision does not incur dependence on q.

Noise-Lossiness and Entropy (Section 5). We analyze the relation between noise-lossiness and min-entropy of a distribution both in the general setting and in the ball-bounded setting. We derive the following bounds.

Lemma 1.2 (Noise-Lossiness of General Distributions)**.** *Let $\mathcal{S}$ be a general distribution over Z_q^n, then $\nu_\sigma(\mathcal{S}) \geq \tilde{H}_\infty(\mathbf{s}) - n\log(q/\sigma) - 1$.*

Lemma 1.3 (Noise-Lossiness of Small Distributions.)**.** *Let $\mathcal{S}$ be a distribution over Z_q^n which is supported only inside a ball of radius r, then $\nu_\sigma(\mathcal{S}) \geq \tilde{H}_\infty(\mathbf{s}) - \sqrt{2\pi}\log(e) \cdot \sqrt{n}r/\sigma$.*

Putting these results together with our main theorem, we get general Entropic LWE hardness whenever $\tilde{H}_\infty(\mathbf{s}) \;\&\; k\log(q) + n\log(q\gamma\sqrt{m}/\sigma)$. In the r-ball-bounded setting we require entropy $\tilde{H}_\infty(\mathbf{s}) \;\&\; k\log(\gamma r) + \sqrt{2\pi}\log(e)\gamma\sqrt{nm}r/\sigma$.[2] Note that if we make the very strong (yet not implausible) assumption that LWE is sub-exponentially secure, then we can use complexity leveraging and choose k to be polylogarithmic, we can choose σ to be large enough that the second term vanishes, and we get entropic hardness even with $\tilde{H}_\infty(\mathbf{s})$ which is polylogarithmic in the security parameter, in particular independent of $\log(q)$.

Tightness (Sections 6 and 7). We provide two tightness results. The first one is essentially a restatement of a bound that was shown in the Ring-LWE setting by Bolboceanu et al. [BBPS19]. It is unconditional, but requires q to have a factor of a proper size.

Theorem 1.4 (Counterexample for Entropic LWE, Informal [BBPS19])**.** *Let n, q, σ be LWE parameters. Then if there exists p s.t. $p|q$ and $p \approx \sigma\sqrt{n}$, then there exists a distribution $\mathcal{S}$ with min-entropy roughly $n\log(q/\sigma)$, such that Entropic LWE is insecure with respect to $\mathcal{S}$.*

However, the above requires that q has a factor of appropriate size. One could wonder whether one can do better for a prime q. While we do not have an explicit counterexample here, we can show that proving such a statement (i.e. security for Entropic LWE with entropy below roughly $n\log(q/\sigma)$) cannot be done by a black-box reduction to a standard "game based" assumption. In particular if the reduction can only access the adversary and to the distribution of secrets as black-box, then the entropy bound $n\log(q/\sigma)$ applies even for prime q.

Theorem 1.5 (Barrier for Entropic LWE, Informal)**.** *Let n, q, σ be LWE parameters. Then there is no black-box reduction from Entropic LWE with entropy $\ll n\log(q/\sigma)$ to any game-based cryptographic assumption.*

[2] In the ball-bounded setting, our main improvement over [AKPW13, Appendix B] is that our entropy bound is independent of q. This is due to our use of Hermite normal form. Beyond this important difference, our flooding method and that of [AKPW13] are asymptotically similar *in the ball-bounded setting*. Our method of flooding at the source, however, is a general method that performs at least as well as the state of the art in the ball-bounded setting, and also implies tight results in the unbounded setting.

1.2 Technical Overview

We provide a technical overview of our main contributions.

The Lossiness Approach to Entropic LWE. The starting point of our proof is the lossiness approach. This approach (in some small variants) was used in all existing hardness results for Entropic LWE [GKPV10]. However, prior works were only able to use it for norm-bounded secrets. We show a minor yet crucial modification that allows to relate the hardness of Entropic LWE to the noise-lossiness of the noise distribution.

Fix parameters n, q, σ and recall that the adversary is given $(\mathbf{A}, \mathbf{y})$, where $\mathbf{A}$ is uniform, $\mathbf{y} = \mathbf{sA} + \mathbf{e} \pmod q$, $\mathbf{s}$ sampled from $\mathcal{S}$ and $\mathbf{e}$ is a (continuous) Gaussian with parameter σ. The lossiness approach replaces the uniform matrix $\mathbf{A}$ with an "LWE matrix" of the form: $\mathbf{BC} + \mathbf{F}$, where $\mathbf{B} \in \mathsf{Z}_q^{n \times k}$, $\mathbf{C} \in \mathsf{Z}_q^{k \times m}$ are uniform, and $k \ll n, m$, and where $\mathbf{F}$ is a matrix whose every element is a (discrete) Gaussian with parameter γ. The *decisional* LWE assumption with dimension k, modulus q and noise parameter γ asserts that $\mathbf{BC}+\mathbf{F}$ is computationally indistinguishable from a uniform matrix, and therefore the adversary should also be able to recover $\mathbf{s}$ when $(\mathbf{A}, \mathbf{y})$ is generated using $\mathbf{A} = \mathbf{BC} + \mathbf{F}$. At this point, the vector $\mathbf{y}$ is distributed as

$$\mathbf{y} = \mathbf{sA} + \mathbf{e} = \mathbf{sBC} + \mathbf{sF} + \mathbf{e} \ .$$

The strategies on how to continue from here diverge. The [GKPV10] approach is to say that when $\mathbf{s}$ is confined inside a ball, and when $\mathbf{e}$ is a wide enough Gaussian, then the value $\mathbf{sF} + \mathbf{e}$ is "essentially independent" of $\mathbf{s}$. This is sometimes referred to as "noise flooding" since the noise $\mathbf{e}$ "floods" the value $\mathbf{sF}$ and minimizes its effect. This allows to apply the leftover hash lemma to argue that $\mathbf{sB}$ is statistically close to a uniform $\mathbf{s}'$ and obtain a new "standard" LWE instance. The [BLP+13, Mic18] approaches can be viewed as variants of this method, where the argument on $\mathbf{sF} + \mathbf{e}$ is refined in non-trivial ways to achieve better parameters.

This type of argument cannot work for the general setting (i.e. when $\mathbf{s}$ is not short) since in this case $\mathbf{sF} + \mathbf{e}$ can reveal noticeable information about $\mathbf{s}$. For example, if $\mathbf{s}$ is a multiple of some large enough factor then the noise $\mathbf{e}$ can just be rounded away (indeed this will be the starting point for our tightness result, as we explain further below).

Our approach therefore is to resort to a weaker claim. We do not try to change $\mathbf{y}$ into a form of standard LWE, but instead all we show is that $\mathbf{y}$ *loses information* about $\mathbf{s}$. Namely, we will show that even information-theoretically it is not possible to recover $\mathbf{s}$ from $(\mathbf{A}, \mathbf{y})$. This approach was taken, for example, by Alwen et al. [AKPW13], but they were unable to show lossiness for the general setting. The reason, essentially, is that they also use a refined version of noise flooding, one that did not require that $\mathbf{e}$ completely floods $\mathbf{sF}$, only slightly perturb it. We can call it "gentle flooding" for the purpose of this work. A similar argument was used in [DM13] to establish hardness of LWE with uniform errors from a short interval.

We note that in all flooding methods, it is beneficial if $\mathbf{F}$ contains small values as much as possible. Therefore in order to show hardness for $\mathbf{s}$ with as low entropy as possible, the parameter γ is to be taken as small as possible, while still supporting the hardness of distinguishing $\mathbf{BC} + \mathbf{F}$ from uniform.

Our Approach: Gentle Flooding at the Source. Our approach can be viewed in hindsight as a very simple modification of the lossiness / flooding approach, that results in a very clean

statement, and the characterization of the noise lossiness as the "right" parameter for the hardness of Entropic LWE.

We take another look at the term $\mathbf{sF}+\mathbf{e}$ and recall that our goal is to use $\mathbf{e}$ to lose information about $\mathbf{s}$. Clearly, if $\mathbf{e}$ was of the form $\mathbf{e_1F}$, then things would be more approachable since then we would simply have $(\mathbf{s}+\mathbf{e_1})\mathbf{F}$, and we will simply need to argue about the lossiness of $\mathbf{s}$ under additive Gaussian noise (which is exactly our notion of noise lossiness for the distribution $\mathcal{S}$). Our observation is that even though $\mathbf{e}$ does not have this form, the properties of the Gaussian distribution allow to present $\mathbf{e}$ as $\mathbf{e}=\mathbf{e_1F}+\mathbf{e_2}$, where $\mathbf{e_1},\mathbf{e_2}$ are independent random variables (but the *distribution* of $\mathbf{e_2}$ depends on $\mathbf{F}$). This is easiest to analyze when $\mathbf{e}$ is a continuous Gaussian, which is the approach we take in this work.[3]

It can be shown essentially by definition that the sum of two independent Gaussian vectors with covariance matrices $\mathbf{\Sigma_1},\mathbf{\Sigma_2}$ is a Gaussian with covariance matrix $\mathbf{\Sigma_1}+\mathbf{\Sigma_2}$. It follows that if we choose $\mathbf{e_1}$ to be a spherical Gaussian with parameter σ_1 then $\mathbf{e_1F}$ will have covariance matrix $\sigma_1\mathbf{F}^T\mathbf{F}$. Therefore if we choose $\mathbf{e_2}$ to be an aspherical Gaussian with covariance $\sigma\mathbf{I}-\sigma_1\mathbf{F}^T\mathbf{F}$, we get that $\mathbf{e}=\mathbf{e_1F}+\mathbf{e_2}$ is indeed a spherical σ Gaussian. There is an important emphasis here, the matrix $\sigma\mathbf{I}-\sigma_1\mathbf{F}^T\mathbf{F}$ must be a valid covariance matrix, i.e. positive semidefinite. To guarantee this, we must set the ratio σ/σ_1 to be at least the largest singular value of the matrix $\mathbf{F}$. Standard results on singular values of Gaussian matrices imply that the largest singular value is roughly $\sqrt{m}\gamma$, which governs the ratio between σ_1 and σ. We stress again that $\mathbf{e_1}$ and $\mathbf{e_2}$ are independent random variables.

Once we established the decomposition of the Gaussian, we can write $\mathbf{y}$ as

$$\mathbf{y}=\mathbf{sA}+\mathbf{e}=\mathbf{sBC}+(\mathbf{s}+\mathbf{e_1})\mathbf{F}+\mathbf{e_2}\ .$$

Now, our noise lossiness term $\nu_{\sigma_1}(\mathcal{S})=\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e_1})$ naturally emerges. Note that $\mathbf{y}$ cannot provide more information about $\mathbf{s}$ than the two variables $(\mathbf{sB},\mathbf{s}+\mathbf{e_1})$. Since the former contains only $k\log q$ bits, it follows that if the noise lossiness is sufficiently larger than $k\log q$, then naturally $\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e_1},\mathbf{sB})$ is non-trivial (we need $\omega(\log\lambda)$ where λ is the security parameter), which implies that finding $\mathbf{s}$ is information theoretically hard. Thus the hardness of Entropic (search) LWE is established.

If in addition $\mathbf{B}$ can serve as an extractor (this is the case when the modulus q is prime, or when the $\mathcal{S}$ is binary), then we can make a stronger claim, that $\mathbf{sB}$ is statistically close to uniform, and then apply (standard) LWE again in order to obtain hardness for Entropic dLWE directly.

Finally, we notice that for norm-bounded distributions we can improve the parameters further by using LWE in *Hermite Normal Form* (HNF) which has been shown to be equivalent to standard LWE in [ACPS09]. HNF LWE allows to argue that $\mathbf{BC}+\mathbf{F}$ is indistinguishable from uniform even when the elements of $\mathbf{B}$ are also sampled from a Gaussian with parameter γ (same as $\mathbf{F}$). Using HNF, we can further bound the entropy loss caused by the term $\mathbf{sB}$ and achieve a bound that is independent of q, and only depends on γ,r,σ. We can only apply this technique for Entropic search LWE.

For the complete analysis and formal statement of the result, see Section 4.

Computing The Noise Lossiness. We briefly explain the intuition behind the noise lossiness computation. The exact details require calculation and are detailed in Section 5.

[3] It can be shown and is by now standard that the hardness of LWE is essentially equivalent whether $\mathbf{e}$ is a continuous Gaussian, discrete Gaussian, or "rounded" Gaussian [Pei10].

For the sake of this overview, let us consider only "flat" distributions, i.e. ones that are uniform over a set of K strings (and thus have min-entropy $\log K$). We will provide an upper bound on the probability $\Pr_{\mathbf{s},\mathbf{e}}[\mathcal{A}^*(\mathbf{s}+\mathbf{e})=\mathbf{s}]$ from Eq. (1), which will immediately translate to a bound on the noise-lossiness.

For general distributions, we note that we can write

$$\Pr_{\mathbf{s},\mathbf{e}}[\mathcal{A}^*(\mathbf{s}+\mathbf{e})=\mathbf{s}] = \int_{\mathbf{y}} \Pr_{\mathbf{s},\mathbf{e}}[\mathbf{s}+\mathbf{e}=\mathbf{y} \wedge \mathcal{A}^*(\mathbf{y})=\mathbf{s}]dy \ ,$$

where the integral is over the entire q-cube (we use integral since we use a continuous distribution for $\mathbf{e}$, but a calculation with discrete Gaussian will be very similar). Note that the expression $\Pr_{\mathbf{s},\mathbf{e}}[\mathbf{s}+\mathbf{e}=\mathbf{y} \wedge \mathcal{A}^*(\mathbf{y})=\mathbf{s}]$ can be written as $\Pr_{\mathbf{s},\mathbf{e}}[\mathcal{A}^*(\mathbf{y})+\mathbf{e}=\mathbf{y} \wedge \mathcal{A}^*(\mathbf{y})=\mathbf{s}]$, which can then be decomposed since the event $\mathcal{A}^*(\mathbf{y})+\mathbf{e}=\mathbf{y}$ depends only on $\mathbf{e}$ and the event $\mathcal{A}^*(\mathbf{y})=\mathbf{s}$ depends only on $\mathbf{s}$ (recall that $\mathbf{y}$ is fixed at this point). We can therefore write

$$\Pr_{\mathbf{s},\mathbf{e}}[\mathcal{A}^*(\mathbf{s}+\mathbf{e})=\mathbf{s}] = \int_{\mathbf{y}} \Pr_{\mathbf{e}}[\mathbf{e}=\mathbf{y}-\mathcal{A}^*(\mathbf{y})] \cdot \Pr_{\mathbf{s}}[\mathcal{A}^*(\mathbf{y})=\mathbf{s}]dy \ .$$

Now, for all $\mathbf{y}$ it holds that $\Pr_{\mathbf{s}}[\mathcal{A}^*(\mathbf{y})=\mathbf{s}] \leq 1/K$, simply since $\mathcal{A}^*(\mathbf{y})$ is a fixed value. It also holds that $\Pr_{\mathbf{e}}[\mathbf{e}=\mathbf{y}-\mathcal{A}^*(\mathbf{y})]$ is bounded by the maximum value of the Gaussian mass function, which is $1/\sigma^n$. We get that

$$\Pr_{\mathbf{s},\mathbf{e}}[\mathcal{A}^*(\mathbf{s}+\mathbf{e})=\mathbf{s}] \leq \frac{1}{K\sigma^n}\int_{\mathbf{y}} dy = \frac{q^n}{K\sigma^n} \ ,$$

and Lemma 1.2 follows.

For the setting of Lemma 1.3, recall that $\mathcal{S}$ is supported only over r-norm-bounded vectors. Note that the analysis above is correct up to and including the conclusion that $\Pr_{\mathbf{s}}[\mathcal{A}^*(\mathbf{y})=\mathbf{s}] \leq 1/K$. Furthermore, $\mathcal{A}^*(\mathbf{y})$ must return a value in the support of $\mathcal{S}$, that is small. We therefore remain with the challenge of bounding $\int_{\mathbf{y}} \Pr_{\mathbf{e}}[\mathbf{e}=\mathbf{y}-\mathcal{A}^*(\mathbf{y})]dy$, when we are guaranteed that $\|\mathcal{A}^*(\mathbf{y})\| \leq r$. We can deduce that this can only induce a minor perturbation to the $\mathbf{e}$ Gaussian. Using Gaussian tail bounds the result follows.

Tightness. The result of [BBPS19] (Theorem 1.4 above) is quite straightforward in our setting (they showed a ring variant which is somewhat more involved). The idea to choose $\mathcal{S}$ to be uniform over the set of all vectors that are multiples of p (or in the [BBPS19] terminology, uniform over an ideal dividing the ideal q). This distribution has min-entropy $n\log(q/p) \approx n\log(q/\sigma)$ (since $p \approx \sigma$), and it clearly leads to an insecure LWE instance since the instance can be taken modulo p in order to recover the noise, and then once the noise is removed the secret can easily be recovered.

The above argument seems to "unfairly" rely on the structure of the modulus q, and one could hope that for prime q, which has no factors, a better result can be achieved. We extend a methodology due to Wichs [Wic13] to show that if such a result exists then it will require non-black-box use of the adversary and/or the sampler for the distribution $\mathcal{S}$. Consider a black-box reduction that given access to an entropic LWE adversary $\mathcal{A}$ and a sampler for $\mathcal{S}$ (we overload the notation and denote the sampler by $\mathcal{S}$ as well), manages to solve some hard problem, e.g. solve a standard LWE instance.

We show that it is possible to efficiently (jointly) simulate $\mathcal{A},\mathcal{S}$, such that in the eyes of a reduction they are indistinguishable from a real high-entropy distribution $\mathcal{S}$ and an adversary $\mathcal{A}$

that solves Entropic LWE on it, thus leading to an efficient unconditional algorithm for said hard problem. The basic idea relies on the natural intuition that it is hard to generate a "valid" LWE instance without knowing the value of $\mathbf{s}$ that is being used. While this intuition is false in many situations, we show that in the entropic setting with black-box reductions it can be made formal.

Specifically, consider $\mathcal{S}$ that is just a uniform distribution over a set of K randomly chosen strings (note that this distribution does not have an efficient sampler, but a black-box reduction is required to work in such a setting as well, and we will show how to simulate $\mathcal{S}$ efficiently). The adversary $\mathcal{A}$, upon receiving an instance $(\mathbf{A}, \mathbf{y})$ first checks that $\mathbf{A}$ is full rank (otherwise return $\perp$), and if so it brute-forces $\mathbf{s}$ out of $\mathbf{y}$ by trying all possible $\mathbf{s}^*$ in the support of $\mathcal{S}$, and if there is one for which $\mathbf{y} - \mathbf{s}^*\mathbf{A} \pmod q$ is short (i.e. of the length that we expect from noise with Gaussian parameter σ), then return a random such $\mathbf{s}^*$ as answer (otherwise return $\perp$). This is a valid adversary for Entropic LWE and therefore it should allow the reduction to solve the hard problem.

Now, let us show how to simulate $\mathcal{S}, \mathcal{A}$ efficiently. The idea is to rely on the intuition that the reduction cannot generate valid LWE instances with values of $\mathcal{S}$ that it does not know, and since the distribution is sparse, the reduction cannot generate strings in the support of $\mathcal{S}$ in any way except calling the $\mathcal{S}$ sampler. Furthermore, since the reduction can only make polynomially many queries to the sampler, there are only polynomially many options for $\mathbf{s}$ for which it can generate valid LWE instances, and our efficient implementation of $\mathcal{A}$ can just check these polynomially many options. (Note that throughout this intuitive outline we keep referring to *valid* Entropic LWE instances, the above argument actually fails without a proper notion of validity as will be explained below.)

Concretely, we will simulate the adversary using a *stateful* procedure, i.e. one that keeps state. However, in the eyes of the reduction this will simulate the original stateless adversary and therefore will suffice for our argument. We will simulate $\mathcal{S}$ using "lazy sampling". Whenever the reduction makes a call to $\mathcal{S}$, we will just sample a new random string $\mathbf{s}$, and save the new sample to its internal state. When a query $(\mathbf{A}, \mathbf{y})$ to $\mathcal{A}$ is made, then we first check that $\mathbf{A}$ is indeed full rank (otherwise return $\perp$), and if it is the case, go over all vectors $\mathbf{s}^*$ that we generated so far (and are stored in the state), and check whether $\mathbf{y} - \mathbf{s}^*\mathbf{A} \pmod q$ is short (in the same sense as above, i.e. of the length that we expect from noise with Gaussian parameter σ). If it is the case then a random such $\mathbf{s}^*$ is returned as the Entropic LWE answer. If the scan did not reveal any adequate candidate, then return $\perp$.

We want to argue that the above simulates the stateless process. The first step is to show that if there is no $\mathbf{s}^*$ in the state and thus our simulated adversary returns $\perp$, then the inefficient adversary would also have returned $\perp$ with all but negligible probability. Secondly, noticing that when our simulated adversary does return a value $\mathbf{s}^*$, this $\mathbf{s}^*$ is a value that the reduction already received as a response to a $\mathcal{S}$ query, and only one such $\mathbf{s}^*$ exists. In fact, both of these concerns boil down to properly defining a notion of validity of the Entropic LWE instance that will prevent both of these concerns.

To this end, we notice that the original inefficient adversary return a non-$\perp$ value only on instances where $\mathbf{A}$ is full rank, and there exists a short $\mathbf{e}^*$ and value $\mathbf{s}^*$ in the support of $\mathcal{S}$ such that $\mathbf{y} = \mathbf{s}^*\mathbf{A} + \mathbf{e}^*$. We will prove that it is not possible to find an instance which is valid for $\mathbf{s}$ in the support of $\mathcal{S}$ which has not been seen by the reduction. This will address both concerns and can be proven since the unseen elements of $\mathcal{S}$ are just randomly sampled strings, so we can think of the vectors as sampled after the matrix $\mathbf{A}$ is determined. The probability of a random vector $\mathbf{s}$ to be s.t. $\mathbf{y} - \mathbf{s}\mathbf{A}$ is σ-short, where $\mathbf{y}$ is arbitrary and $\mathbf{A}$ is full rank, is roughly $(\sigma/q)^n$.

This translates to the cardinality K of $\mathcal{S}$ being as large as (roughly) $n \log(q/\sigma)$ and still allowing to apply the union bound. The result thus follows.

Maybe somewhat interestingly, while our security proofs for entropic LWE are technically similar to *converse coding theorems* [Sha48, W+59], our barrier result resembles the random coding arguments used to prove the *coding theorem* [Sha48, Sha49].

1.3 Acknowledgement

We would like to thank the reviewers of Eurocrypt 2020 and Katharina Boudgoust for their helpful comments.

2 Preliminaries

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 $\|\mathbf{x}\|_\infty = \max_i |x_i|$.

We denote by $\mathsf{T}_q = \mathsf{R}/q\mathsf{Z}$ be the real torus of scale q. We can embed $\mathsf{Z}_q = \mathsf{Z}/q\mathsf{Z}$ into T_q in the natural way. T_q is an abelian group and therefore a Z-algebra. Thus multiplication of vectors from T_q^n with Z-matrices is well-defined. T_q is however not a Z_q-algebra. We will represent T_q elements by their central residue class representation in $[-q/2, q/2)$.

For a continuous random variable $\mathbf{x}$, we will denote the probability-density function of $\mathbf{x}$ by $p_{\mathbf{x}}(\cdot)$. We will denote the probability density of $\mathbf{x}$ conditioned on an event E by $p_{\mathbf{x}|E}(\cdot)$. 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 $\Delta(X,Y) = \sum_{x \in \mathcal{X}} |\Pr[X = x] - \Pr[Y = x]|$. Likewise, if X and Y are two continuous random variables defined on a measurable set $\mathcal{X}$, we define the statistical distance between X and Y as $\Delta(X,Y) = \int_{x \in \mathcal{X}} |p_X(x) - p_Y(x)|$.

Random Matrices Let p be a prime modulus. Let $\mathbf{A} \leftarrow_\$ \mathsf{Z}_p^{n \times m}$ be chosen uniformly at random. Then the probability that $\mathbf{A}$ is not invertible (i.e. does not have an invertible column-submatrix)

$$\Pr[\mathbf{A} \text{ not invertible}] = 1 - \prod_{i=0}^{n-1} (1 - p^{i-m}) \leq p^{n-m}.$$

For an arbitrary modulus q, a matrix $\mathbf{A}$ is invertible if and only if it is invertible modulo all prime factors p_i of q. As we can bound the number of prime factors of q by $\log(q)$, we get for an $\mathbf{A} \leftarrow_\$ \mathsf{Z}_p^{n \times m}$ that

$$\Pr[\mathbf{A} \text{ not invertible}] \leq \log(q) \cdot 2^{n-m}.$$

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 $\mathbf{z}$.

2.2 Leftover Hashing

We recall the generalized leftover hash lemma [DORS08, Reg05].

Lemma 2.1. *Let q be a modulus and let n, k be integers. Let $\mathbf{s}$ be a random variable de ned on Z_q^n and let $\mathbf{B} \leftarrow_\$ \mathsf{Z}_q^{n\times k}$ be chosen uniformly random. Furthermore let Y be a random-variable (possibly) correlated with $\mathbf{s}$. Then, given that either q is prime or $\mathbf{s}$ is supported on $\{-1, 0, 1\}^n$ it holds that*

$$\Delta((\mathbf{B}, \mathbf{sB}, Y), (\mathbf{B}, \mathbf{u}, Y)) \leq \sqrt{q^k \cdot 2^{-H_\infty(\mathbf{s}|Y)}}.$$

The proof of the LHL relies on the fact the the function $H_{\mathbf{B}} : \mathcal{X} \to \mathsf{Z}_q^k$ given by $\mathbf{x} \mapsto \mathbf{sB}$ is a universal hash function. This can either be achieved by choosing q to be prime or restricting the domain $\mathcal{X} \subseteq \mathsf{Z}_q^n$ appropriately, e.g. to $\{-1, 0, 1\}^n$.

2.3 Gaussians

Continuous Gaussians A matrix $\mathbf{\Sigma} \in \mathsf{R}^{n\times n}$ is called positive definite, if it holds for every $\mathbf{x} \in \mathsf{R}^n \backslash \{\mathbf{0}\}$ that $\mathbf{x\Sigma x}^\top > 0$. For every positive definite matrix $\mathbf{\Sigma}$ there exists a unique positive definite matrix $\sqrt{\mathbf{\Sigma}}$ such that $(\sqrt{\mathbf{\Sigma}})^2 = \mathbf{\Sigma}$.

For a parameter $\sigma > 0$ define the n-dimensional gaussian function $\rho_\sigma : \mathsf{R}^n \to (0, 1]$ by

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

For a positive definite matrix $\mathbf{\Sigma} \in \mathsf{R}^{n\times n}$, define the function $\rho_{\sqrt{\mathbf{\Sigma}}} : \mathsf{R}^n \to (0, 1]$ by

$$\rho_{\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) := e^{-\pi\mathbf{x}\mathbf{\Sigma}^{-1}\mathbf{x}^\top}.$$

For a scalar $\sigma > 0$, we will define

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

The total measure of $\rho_{\sqrt{\mathbf{\Sigma}}}$ over R^n is

$$\rho_{\sqrt{\mathbf{\Sigma}}}(\mathsf{R}^n) = \int_{\mathsf{R}^n} \rho_{\sqrt{\mathbf{\Sigma}}}(\mathbf{x})d\mathbf{x} = \sqrt{\det(\mathbf{\Sigma})}.$$

In the scalar case this becomes

$$\rho_\sigma(\mathsf{R}^n) = \int_{\mathsf{R}^n} \rho_\sigma(\mathbf{x})d\mathbf{x} = \sigma^n.$$

Normalizing $\rho_{\sqrt{\mathbf{\Sigma}}}$ by $\rho_{\sqrt{\mathbf{\Sigma}}}(\mathsf{R}^n)$ yields the probability density for the continuous gaussian distribution $D_{\sqrt{\mathbf{\Sigma}}}$ over R^n.

For a discrete set $S \subseteq \mathsf{R}^n$ we define $\rho_{\sqrt{\mathbf{\Sigma}}}(S)$ by

$$\rho_{\sqrt{\mathbf{\Sigma}}}(S) := \sum_{\mathbf{s} \in S} \rho_{\sqrt{\mathbf{\Sigma}}}(\mathbf{s}).$$

In particular, for a integer q we have

$$\rho_{\sqrt{\mathbf{\Sigma}}}(q\mathsf{Z}^n) = \sum_{\mathbf{z} \in q\mathsf{Z}^n} \rho_{\sqrt{\mathbf{\Sigma}}}(\mathbf{z}).$$

For a gaussian $x \sim D_\sigma$ we get the tail-bound

$$\Pr[|x| \geq t] \leq 2 \cdot e^{-\frac{t^2}{2\sigma^2}}.$$

As a simple consequence we get $\Pr[|x| \geq (\log(\lambda)) \cdot \sigma] \leq \mathsf{negl}(\lambda)$.

Discrete Gaussians We say a random variable x defined on Z follows the discrete gaussian distribution $D_{\mathsf{Z},\sigma}$ for a parameter $\sigma > 0$, if the probability mass function of x is given by

$$\Pr[x = x'] = \frac{\rho_\sigma(x')}{\rho_\sigma(\mathsf{Z})}$$

for every $x' \in \mathsf{Z}$.

Modular Gaussians For a modulus q, we also define the q-periodic gaussian function $\tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}$: by

$$\tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) := \sum_{\mathbf{z} \in q\mathsf{Z}^n} \rho_{q,\sqrt{\mathbf{\Sigma}}}(\mathbf{x} - \mathbf{z}).$$

We define $\tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}(\mathsf{T}_q^n)$ by

$$\tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}(\mathsf{T}_q^n) := \tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}([-q/2, q/2)^n) = \int_{[-q/2,q/2)^n} \tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}(\mathbf{x}) d\mathbf{x} = \rho_{\sqrt{\mathbf{\Sigma}}}(\mathsf{R}^n).$$

Consequently, normalizing $\tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}$ by $\tilde{\rho}_{q,\sqrt{\mathbf{\Sigma}}}(\mathsf{T}_q^n)$ yields a probability density on T_q^n. We call the corresponding distribution $D_{\sqrt{\mathbf{\Sigma}}} \mod q$ a modular gaussian. A $\mathbf{x} \sim D_{\sqrt{\mathbf{\Sigma}}} \mod q$ can be sampled by sampling and $\mathbf{x}' \leftarrow_\$ D_{\sqrt{\mathbf{\Sigma}}}$ and computing $\mathbf{x} \leftarrow \mathbf{x}' \mod q$.

In order to prove our strong converse coding theorems, we need various upper bounds for the periodic gaussian function. We will use the following variant of the smoothing lemma of Micciancio and Regev [MR04][4].

Lemma 2.2 (Smoothing Lemma [MR04])**.** *Let $\epsilon > 0$. Given that $\frac{1}{\sigma} \geq \sqrt{\frac{\ln(2n(1+1/\epsilon))}{\pi}} \cdot \frac{1}{q}$, then it holds that*

$$\rho_\sigma(q\mathsf{Z}^n \backslash \{\mathbf{0}\}) \leq \epsilon.$$

[4]We use the smoothing lemma with the parameter $s = 1/\sigma$ and the lattice $\Lambda = \frac{1}{q}\mathsf{Z}^n$. Note that for this lattice it holds that $\lambda_n = 1/q$.

Lemma 2.3. *The periodic gaussian function $\tilde{\rho}_{q,\sigma}$ assumes its maximum at $q \cdot \mathsf{Z}^n$. In particular, it holds for all $\mathbf{x} \in \mathsf{R}^n$ that $\tilde{\rho}_{q,\sigma}(\mathbf{x}) \leq \tilde{\rho}_{q,\sigma}(\mathbf{0})$.*

Proof. We can write $\tilde{\rho}_{q,\sigma}$ as

$$\tilde{\rho}_{q,\sigma}(\mathbf{x}) = f_{\mathbf{x}}(q\mathsf{Z}^n),$$

where $f_{\mathbf{x}}(\mathbf{z}) = \rho_\sigma(\mathbf{z} - \mathbf{x})$. The Poisson summation formula allows us to express $f_{\mathbf{x}}(q\mathsf{Z}^n)$ by

$$f_{\mathbf{x}}(q\mathsf{Z}^n) = \det(\frac{1}{q}\mathsf{Z}^n) \cdot \hat{f}_{\mathbf{x}}(\frac{1}{q}\mathsf{Z}^n) = \frac{1}{q^n}\hat{f}_{\mathbf{x}}(\frac{1}{q}\mathsf{Z}^n)$$

Since $\hat{f}_{\mathbf{x}}(\boldsymbol{\omega}) = e^{-2\pi i \cdot \langle \mathbf{x}, \boldsymbol{\omega} \rangle}\sigma^n \rho_{1/\sigma}(\boldsymbol{\omega})$, we can write $f_{\mathbf{x}}(q\mathsf{Z}^n)$ as

$$f_{\mathbf{x}}(q\mathsf{Z}^n) = \left(\frac{\sigma}{q}\right)^n \cdot \sum_{\omega \in \frac{1}{q}\mathsf{Z}^n} e^{-2\pi i \cdot \langle \mathbf{x}, \boldsymbol{\omega} \rangle} \rho_{1/\sigma}(\boldsymbol{\omega}),$$

we can bound

$$\begin{aligned}
\tilde{\rho}_{q,\sigma}(\mathbf{x}) &= |\tilde{\rho}_{q,\sigma}(\mathbf{x})| \\
&= |f_{\mathbf{x}}(q\mathsf{Z}^n)| \\
&= \left| \left(\frac{\sigma}{q}\right)^n \cdot \sum_{\omega \in \frac{1}{q}\mathsf{Z}^n} e^{-2\pi i \cdot \langle \mathbf{x}, \boldsymbol{\omega} \rangle} \rho_{1/\sigma}(\boldsymbol{\omega}) \right| \\
&\leq \left(\frac{\sigma}{q}\right)^n \cdot \sum_{\omega \in \frac{1}{q}\mathsf{Z}^n} |e^{-2\pi i \cdot \langle \mathbf{x}, \boldsymbol{\omega} \rangle}| \cdot |\rho_{1/\sigma}(\boldsymbol{\omega})| \\
&= \left(\frac{\sigma}{q}\right)^n \cdot \sum_{\omega \in \frac{1}{q}\mathsf{Z}^n} \rho_{1/\sigma}(\boldsymbol{\omega}) \\
&= f_{\mathbf{0}}(q\mathsf{Z}^n) \\
&= \tilde{\rho}_{q,\sigma}(\mathbf{0})
\end{aligned}$$

The first equality holds as $\tilde{\rho}_{q,\sigma}(\mathbf{x}) > 0$ and the inequality holds by an application of the triangle inequality. This concludes the proof. □

Lemma 2.4. *If $\frac{q}{\sigma} \geq \sqrt{\frac{\ln(4n)}{\pi}}$, then it holds for all $\mathbf{x} \in \mathsf{R}^n$ that*

$$\tilde{\rho}_{q,\sigma}(\mathbf{x}) \leq 2.$$

Proof. Choosing $\epsilon = 1$ in Lemma 2.2 yields

$$\rho_\sigma(q\mathsf{Z}^n \backslash \{\mathbf{0}\}) = 1$$

as $\frac{1}{\sigma} \geq \sqrt{\frac{\ln(2n(1+1/\epsilon))}{\pi}} \cdot \frac{1}{q} = \sqrt{\frac{\ln(4n)}{\pi}} \cdot \frac{1}{q}$. Consequently, we obtain

$$\begin{aligned}
\hat{\rho}_{q,\sigma}(\mathbf{0}) &= \sum_{\mathbf{z} \in q\mathsf{Z}^n} \hat{\rho}_\sigma(\mathbf{z}) \\
&= \rho_{\sigma_1}(q\mathsf{Z}^n) \\
&= 1 + \rho_{\sigma_1}(q\mathsf{Z}^n \backslash \{\mathbf{0}\}) \\
&\leq 2.
\end{aligned}$$

The claim follows from the fact above. □

We will use the following estimate for shifted gaussians.

Lemma 2.5. *Let $\sigma_2 > \sigma_1 > 0$. Then it holds for all $\mathbf{x} \in \mathsf{R}^n$ and $\mathbf{t} \in \mathsf{R}^n$ that*

$$\rho_{\sigma_1}(\mathbf{x} - \mathbf{t}) \leq e^{\pi \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2}} \cdot \rho_{\sigma_2}(\mathbf{x}).$$

Moreover, the same holds for the q-periodic gaussian function $\hat{\rho}_{q\mathsf{Z}^n,\sigma_1}$, i.e.

$$\hat{\rho}_{q\mathsf{Z}^n,\sigma_1}(\mathbf{x} - \mathbf{t}) \leq e^{\pi \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2}} \cdot \hat{\rho}_{q\mathsf{Z}^n,\sigma_2}(\mathbf{x}).$$

Proof. By a routine calculation,

$$\rho_{\sigma_1}(\mathbf{x} - \mathbf{t}) \leq e^{\pi \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2}} \cdot \rho_{\sigma_2}(\mathbf{x})$$

is equivalent to

$$\frac{\|\mathbf{x} - \mathbf{t}\|^2}{\sigma_1^2} - \frac{\|\mathbf{x}\|^2}{\sigma_2^2} + \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2} \geq 0.$$

Now it holds that

$$\begin{aligned}
\frac{\|\mathbf{x} - \mathbf{t}\|^2}{\sigma_1^2} - \frac{\|\mathbf{x}\|^2}{\sigma_2^2} + \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2} &= \frac{\|\mathbf{x}\|^2}{\sigma_1^2} - \frac{2}{\sigma_1^2}\langle \mathbf{x}, \mathbf{t} \rangle + \frac{\|\mathbf{t}\|^2}{\sigma_1^2} - \frac{\|\mathbf{x}\|^2}{\sigma_2^2} + \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2} \\
&= \frac{\sigma_2^2 - \sigma_1^2}{\sigma_1^2 \sigma_2^2}\|\mathbf{x}\|^2 - \frac{2}{\sigma_1^2}\langle \mathbf{x}, \mathbf{t} \rangle + \frac{\sigma_2^2}{\sigma_1^2(\sigma_2^2 - \sigma_1^2)}\|\mathbf{t}\|^2 \\
&= \frac{1}{\sigma_1^2}\left\| \sqrt{1 - (\sigma_1/\sigma_2)^2}\mathbf{x} - \frac{1}{\sqrt{1 - (\sigma_1/\sigma_2)^2}}\mathbf{t} \right\|^2 \\
&\geq 0.
\end{aligned}$$

To prove the second statement, note that

$$\begin{aligned}
\hat{\rho}_{q\mathsf{Z}^n,\sigma_1}(\mathbf{x} - \mathbf{t}) &= \sum_{\mathbf{z} \in q\mathsf{Z}^n} \rho_{\sigma_1}(\mathbf{x} - \mathbf{t} - \mathbf{z}) \\
&\leq \sum_{\mathbf{z} \in q\mathsf{Z}^n} e^{\pi \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2}} \cdot \rho_{\sigma_2}(\mathbf{x} - \mathbf{z}) \\
&= e^{\pi \frac{\|\mathbf{t}\|^2}{\sigma_2^2 - \sigma_1^2}} \cdot \hat{\rho}_{q\mathsf{Z}^n,\sigma_2}(\mathbf{x}),
\end{aligned}$$

which concludes the proof. □

2.4 Learning with Errors

The learning with errors (LWE) problem was defined by Regev [Reg05]. The search problem $\mathsf{LWE}(n, m, q, \chi)$, for $n, m, q \in \mathbb{N}$ and for a distribution χ supported over the torus $\mathbb{T}_q$ is to find $\mathbf{s}$ given $(\mathbf{A}, \mathbf{sA}+\mathbf{e})$, where $\mathbf{A} \leftarrow_\$ \mathbb{Z}_q^{n\times m}$ is chosen uniformly random and $\mathbf{e} \leftarrow_\$ \chi^m$ is chosen according to χ^m. The decisional version $\mathsf{dLWE}(n, m, q, \chi)$ asks to distinguish between the distributions $(\mathbf{A}, \mathbf{sA} + \mathbf{e})$ and $(\mathbf{A}, \mathbf{u} + \mathbf{e})$, where $\mathbf{A}$, $\mathbf{s}$ and $\mathbf{e}$ are as in the search version and $\mathbf{u} \leftarrow_\$ \mathbb{Z}_q^m$ is chosen uniformly random. We also consider the hardness of solving dLWE for *any* $m = \mathsf{poly}(n \log q)$. This problem is denoted $\mathsf{dLWE}(n, q, \chi)$. The matrix version of this problem asks to distinguish $(\mathbf{A}, \mathbf{S} \cdot \mathbf{A} + \mathbf{E})$ from $(\mathbf{A}, \mathbf{U})$, where $\mathbf{S} \leftarrow_\$ \mathbb{Z}_q^{k\times n}$, $\mathbf{E} \leftarrow_\$ \chi^{k\times m}$ and $\mathbf{U} \leftarrow \mathbb{Z}_q^{k\times m}$. The hardness of the matrix version for any $k = \mathsf{poly}(n)$ can be established from $\mathsf{dLWE}_{n,m,q,\chi}$ via a routine hybrid-argument. Moreover, Applebaum et al. [ACPS09] showed that if the error-distribution χ is supported on $\mathbb{Z}_q$, then the matrix $\mathbf{S}$ can also be chosen from $\chi^{k\times m}$ without affecting the hardness of the problem.

As shown in [Reg05], the $\mathsf{LWE}(n, q, \chi)$ problem with χ being a continuous Gaussian distribution with parameter $\sigma = \alpha q \geq 2\sqrt{n}$ is at least as hard as approximating the shortest independent vector problem (SIVP) to within a factor of $\gamma = \tilde{O}(n/\alpha)$ in *worst case* dimension n lattices. This is proven using a quantum reduction. Classical reductions (to a slightly different problem) exist as well [Pei09, BLP+13] but with somewhat worse parameters. The best known (classical or quantum) algorithms for these problems run in time $2^{\mathcal{O}(n/\log\gamma)}$, and in particular they are conjectured to be intractable for $\gamma = \mathsf{poly}(n)$.

Regev also provided a search-to-decision reduction which bases the hardness of the decisional problem $\mathsf{dLWE}(n, q, \chi)$ on the search version $\mathsf{LWE}(n, q, \chi)$ whenever q is prime of polynomial size. This reduction has been generalized to more general classes of moduli [Pei09, BLP+13]. Moreover, there exists a *sample preserving* reduction which [MM11] which bases the hardness of $\mathsf{dLWE}(n, m, q, \chi)$ on $\mathsf{LWE}(n, m, q, \chi)$ for certain moduli q without affecting the number of samples m.

Finally, Peikert [Pei10] provided a randomized rounding algorithm which allows to base the hardness of $\mathsf{LWE}(n, m, q, D_{\mathbb{Z},\sigma'})$ (i.e. LWE with a discrete gaussian error $D_{\mathbb{Z},\sigma'}$) on $\mathsf{LWE}(n, m, q, D_\sigma)$ (continuous gaussian error), where σ' is only slightly larger than σ.

2.5 Entropic LWE

We will now consider LWE with entropic secrets, entropic LWE for short. In this variant, we allow the distribution of secrets $\mathcal{S}$ to be chosen from a family of distributions $\bar{\mathcal{S}} = \{\mathcal{S}_i\}_i$. This captures the idea the distribution of secrets can be worst-case from a certain family.

Definition 2.6 (Entropic LWE)**.** *Let $q = q(\lambda)$ be a modulus and $n, m = \mathsf{poly}(\lambda)$. Let χ be an error-distribution on $\mathbb{T}_q$. Let $\bar{\mathcal{S}} = \mathcal{S}(\lambda, q, n, m)$ be a family of distributions on $\mathbb{Z}_q^n$. We say that the search problem* $\mathsf{ent\text{-}LWE}(q, n, m, \bar{\mathcal{S}}, \chi)$ *is hard, if it holds for every PPT adversary $\mathcal{A}$ and every $\mathcal{S} \in \bar{\mathcal{S}}$ that*

$$\Pr[\mathcal{A}(1^\lambda, \mathbf{A}, \mathbf{s} \cdot \mathbf{A} + \mathbf{e}) = \mathbf{s}] \leq \mathsf{negl}(\lambda),$$

where $\mathbf{A} \leftarrow_\$ \mathbb{Z}_q^{m\times n}$, $\mathbf{s} \leftarrow_\$ \mathcal{S}$ and $\mathbf{e} \leftarrow_\$ \chi^m$. Likewise, we say that the decisional problem $\mathsf{ent\text{-}dLWE}(q, n, m, \bar{\mathcal{S}}, \chi)$ *is hard, if it holds for every PPT distinguisher $\mathcal{D}$ and every $\mathcal{S} \in \bar{\mathcal{S}}$ that*

$$|\Pr[\mathcal{D}(1^\lambda, \mathbf{A}, \mathbf{sA} + \mathbf{e}) = 1] - \Pr[\mathcal{D}(1^\lambda, \mathbf{u} + \mathbf{e}) = 1]| \leq \mathsf{negl}(\lambda),$$

where $\mathbf{A} \leftarrow_\$ \mathbb{Z}_q^{m\times n}$, $\mathbf{s} \leftarrow_\$ \mathcal{S}$, $\mathbf{e} \leftarrow_\$ \chi^m$ and $\mathbf{u} \leftarrow_\$ \mathbb{Z}_q^m$.

3 Probability-Theoretic Tools

3.1 Singular Values of Discrete Gaussian Matrices

Consider a real valued matrix $\mathbf{A} \in \mathsf{R}^{n \times m}$, 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 $\mathbf{A}\mathbf{A}^\top$. They are denoted $\sigma_1(\mathbf{A}) \geq \cdots \geq \sigma_n(\mathbf{A}) \geq 0$. The *spectral norm* of $\mathbf{A}$ is $\sigma_1(\mathbf{A})$, and we will also denote it by σ_A. It holds that

$$\sigma_{\mathbf{A}} = \sigma_1(\mathbf{A}) = \max_{\mathbf{x} \in \mathsf{R}^m \setminus \{\mathbf{0}\}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|x\|} \ .$$

We will be interested in the of discrete Gaussian matrices.

Proposition 3.1 ([MP12, Lemma 2.8, 2.9])**.** *Let $\mathbf{F} \sim D_{\mathsf{Z},\gamma}^{n \times m}$, assume for convenience that $m \geq n$. Then with all but 2^{-m} probability it holds that $\sigma_{\mathbf{F}} \leq \gamma \cdot C \cdot \sqrt{m}$, where C is a global constant.*

3.2 Decomposition Theorem for Continuous Gaussians

The following proposition is an immediate corollary of the properties of (continuous) Gaussian vectors. We provide a proof for the sake of completeness.

Proposition 3.2. *Let $\mathbf{F} \in \mathsf{Z}^{n \times m}$ be an arbitrary matrix with spectral norm σ_F. Let $\sigma, \sigma_1 > 0$ be s.t. $\sigma > \sigma_1 \cdot \sigma_F$. Let $\mathbf{e}_1 \sim D_{\sigma_1}^n$ 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_σ^m.*

Proof. First note that Σ is positive definite: It holds for any $\mathbf{x} \in \mathsf{R}^m \setminus \{\mathbf{0}\}$ that

$$\mathbf{x}\mathbf{\Sigma}\mathbf{x}^\top = \sigma^2\|\mathbf{x}\|^2 - \sigma_1^2\|\mathbf{x}\mathbf{F}\|^2 \geq \sigma^2\|\mathbf{x}\| - \sigma^2\sigma_{\mathbf{F}}^2\|\mathbf{x}\|^2 \geq (\sigma^2 - \sigma_1^2\sigma_{\mathbf{F}}^2) \cdot \|\mathbf{x}\|^2 > 0,$$

as $\sigma > \sigma_1 \cdot \sigma_{\mathbf{F}}$. Since $\mathbf{e}_1, \mathbf{e}_2$ are independent Gaussian vectors, they are also jointly Gaussian, and therefore $\mathbf{e}$ is also a Gaussian vector. Since $\mathbf{e}_1, \mathbf{e}_2$ have expectation 0, then so does $\mathbf{e}$. The covariance matrix of $\mathbf{e}$ is given by a direct calculation, recalling that $\mathbf{e}_1, \mathbf{e}_2$ are independent:

$$\begin{aligned}
\mathsf{E}[\mathbf{e}^\top\mathbf{e}] &= \mathsf{E}[\mathbf{F}^\top\mathbf{e}^\top\mathbf{e}\mathbf{F}] + \mathsf{E}[\mathbf{e}_2^\top\mathbf{e}_2] \\
&= \mathbf{F}^\top\sigma_1^2\mathbf{I}\mathbf{F} + \mathbf{\Sigma} \\
&= \sigma_1^2\mathbf{F}^\top\mathbf{F} + \sigma^2\mathbf{I} - \sigma_1^2\mathbf{F}^\top\mathbf{F} \\
&= \sigma^2\mathbf{I} \ ,
\end{aligned}$$

and the statement follows. □ □

4 Hardness of Entropic LWE with Gaussian Noise

In this Section we will establish our main result, the hardness of entropic search LWE with continuous gaussian noise. Using standard techniques, we can conclude that entropic search LWE with discrete gaussian noise is also hard. Finally for suitable moduli a search-to-decision reduction can be used to establish the hardness of entropic decisional LWE.

Theorem 4.1. *Let C be the global constant from Proposition 3.1. Let $q = q(\lambda)$ be a modulus and $n, m = \mathsf{poly}(\lambda)$ where $m \geq n$, and let $r, \gamma, \sigma_1 > 0$. Let $\mathbf{s}$ be a random variable on Z_q^n distributed according to some distribution $\mathcal{S}$. Let $\mathbf{e}_1 \sim D_{\sigma_1} \mod q$ be an error term. Assume that $\mathbf{s}$ is r-bounded, where we assume that $r = q$ if no bound for $\mathbf{s}$ is known. Further assume that*

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_1) \geq k \cdot \log(\min\{2C \cdot \gamma \cdot \sqrt{n}r, q\}) + \omega(\log(\lambda))$$

Let $\sigma > C \cdot \sqrt{m} \cdot \gamma \cdot \sigma_1$. Then the search problem $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, D_\sigma)$ is hard, provided that $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$ is hard.

Furthermore, if $\tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_1) \geq k \cdot \log(q) + \omega(\log(\lambda))$ and we have that either q is prime or $\mathbf{s} \in \{0,1\}^n$, then the decisional problem $\mathsf{ent\text{-}dLWE}(q, n, m, \mathcal{S}, D_\sigma)$ is hard, provided that $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$ and $\mathsf{dLWE}(q, k, m, D_\sigma)$ are hard.

Proof. Let $\mathcal{A}$ be a search adversary against $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, D_\sigma)$. Throughout this proof, as in the theorem statement, C is the global constant from Proposition 3.1. Let $\mathcal{W}$ be a distribution on $\mathsf{Z}_q^{n \times k}$, which depending on the setting will either be the uniform distribution on $\mathsf{Z}_q^{n \times k}$ or a discrete gaussian distribution $D_{\mathsf{Z},\gamma}^{n \times k}$. Consider the following hybrid LWE-distributions.

- $\mathcal{H}_0$:
  - $\mathbf{s} \leftarrow_\$ \mathcal{S}$
  - $\mathbf{A} \leftarrow_\$ \mathsf{Z}_q^{n \times m}$
  - $\mathbf{e} \leftarrow_\$ D_\sigma^m$
  - Output $(\mathbf{A}, \mathbf{sA} + \mathbf{e})$
- $\mathcal{H}_1$:
  - $\mathbf{s} \leftarrow_\$ \mathcal{S}$
  - $\mathbf{B} \leftarrow_\$ \mathcal{W}$, $\mathbf{C} \leftarrow_\$ \mathsf{Z}_q^{k \times m}$, $\mathbf{F} \leftarrow_\$ D_{\mathsf{Z},\gamma}^{n \times m}$
  - $\mathbf{A} \leftarrow \mathbf{B} \cdot \mathbf{C} + \mathbf{F}$
  - $\mathbf{e} \leftarrow_\$ D_\sigma^m$
  - Output $(\mathbf{A}, \mathbf{sA} + \mathbf{e})$
- $\mathcal{H}_2$:
  - $\mathbf{s} \leftarrow_\$ \mathcal{S}$
  - $\mathbf{B} \leftarrow_\$ \mathcal{W}$, $\mathbf{C} \leftarrow_\$ \mathsf{Z}_q^{k \times m}$, $\mathbf{F} \leftarrow_\$ D_{\mathsf{Z},\gamma}^{n \times m}$
  - If $\|\mathbf{B}\| > C \cdot \gamma \cdot \sqrt{n}$ or $\|\mathbf{F}\| > C \cdot \gamma \cdot \sqrt{m}$ output $\perp$
  - $\mathbf{A} \leftarrow \mathbf{B} \cdot \mathbf{C} + \mathbf{F}$
  - $\mathbf{e} \leftarrow_\$ D_\sigma^m$
  - Output $(\mathbf{A}, \mathbf{sA} + \mathbf{e})$
- $\mathcal{H}_3$:
  - $\mathbf{s} \leftarrow_\$ \mathcal{S}$

- $\mathbf{B} \leftarrow_\$ \mathcal{W}$, $\mathbf{C} \leftarrow_\$ \mathsf{Z}_q^{k\times m}$, $\mathbf{F} \leftarrow_\$ D_{\mathsf{Z},\gamma}^{n\times m}$
- If $\|\mathbf{B}\| > C\cdot\gamma\cdot\sqrt{n}$ or $\|\mathbf{F}\| > C\cdot\gamma\cdot\sqrt{m}$ output $\perp$
- $\mathbf{A} \leftarrow \mathbf{B}\cdot\mathbf{C} + \mathbf{F}$
- Set $\mathbf{\Sigma} \leftarrow \sigma^2\mathbf{I} - \sigma_1^2\mathbf{F}^T\mathbf{F}$
- $\mathbf{e}_1 \leftarrow_\$ D_\sigma^n$, $\mathbf{e}_2 \leftarrow_\$ D_{\sqrt{\mathbf{\Sigma}}}^m$
- $\mathbf{e} \leftarrow \mathbf{e}_1\mathbf{F} + \mathbf{e}_2$
- Output $(\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e})$

First note that $\mathcal{H}_0$ is identical to the $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, D_\sigma)$-experiment.

Next, it follows directly by the hardness of $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$ that $\mathcal{H}_0$ and $\mathcal{H}_1$ are computationally indistinguishable. Specifically, if $\mathcal{W}$ is the uniform distribution on $\mathsf{Z}_q^{n\times k}$ it follows by the standard form of $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$, whereas if $\mathcal{W}$ is $D_{\mathsf{Z},\gamma}^{n\times k}$ we use the Hermite form of $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$.

We claim that $\mathcal{H}_1$ and $\mathcal{H}_2$ are statistically close. To see this, note that conditioned on $\|\mathbf{B}\| \leq C\cdot\gamma\cdot\sqrt{n}$ and $\|\mathbf{F}\| \leq C\cdot\gamma\cdot\sqrt{m}$ the two experiments are identically distributed. Thus we can bound the statistical distance between $\mathcal{H}_1$ and $\mathcal{H}_2$ by $\Pr[\|\mathbf{B}\| > C\cdot\gamma\cdot\sqrt{n} \text{ or } \|\mathbf{F}\| > C\cdot\gamma\cdot\sqrt{m}]$. As $\mathbf{B} \sim D_{\mathsf{Z},\gamma}^{n\times k}$, it holds by Proposition 3.1 that $\|\mathbf{B}\| \leq C\cdot\gamma\cdot\sqrt{n}$ except with probability 2^{-n}. Likewise, as $\mathbf{F} \sim D_{\mathsf{Z},\gamma}^{n\times m}$, it also holds by Proposition 3.1 that $\|\mathbf{F}\| \leq C\cdot\gamma\cdot\sqrt{m}$ except with probability 2^{-m}. Thus, a union bound yields

$$\Pr[\|\mathbf{B}\| > C\cdot\gamma\cdot\sqrt{n} \text{ or } \|\mathbf{F}\| > C\cdot\gamma\cdot\sqrt{m}] \leq 2^{-n} + 2^{-m} \leq 2\cdot 2^{-n},$$

i.e. the statistical distance between $\mathcal{H}_1$ and $\mathcal{H}_2$ is at most $2\cdot 2^{-n}$.

Finally, we claim that $\mathcal{H}_2$ and $\mathcal{H}_3$ are identically distributed. As $\|\mathbf{F}\| \leq C\cdot\gamma\cdot\sqrt{m}$ and $\sigma > C\cdot\sqrt{m}\cdot\gamma\cdot\sigma_1$, Proposition 3.2 yields that the distribution of $\mathbf{e}_1\mathbf{F} + \mathbf{e}_2$ is exactly D_σ^m and the claim follows.

We will now show that it holds for any search adversary $\mathcal{A}$ that if $\Pr[\mathcal{A}(\mathbf{s}\mathbf{A} + \mathbf{e}) = \mathbf{s}] < \mathsf{negl}(\lambda)$ where $(\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e}) \leftarrow_\$ \mathcal{H}_3$. Consequently, by the above we can then argue that the same holds if $(\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e}) \leftarrow_\$ \mathcal{H}_0$, which means that search problem $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, D_\sigma)$ is hard, concluding the proof for the first statement of the theorem. To do so, we will bound the conditional min-entropy of $\mathbf{s}$ given $\mathbf{A}$ and $\mathbf{s}\mathbf{A} + \mathbf{e}_1\mathbf{F} + \mathbf{e}_2$. We first observe that we can compute $\mathbf{y} = \mathbf{s}\mathbf{A} + \mathbf{e}_1\mathbf{F} + \mathbf{e}_2$ given $\mathbf{s}\mathbf{B} \in \mathsf{Z}_q^k$, $\mathbf{s} + \mathbf{e}_1 \in \mathsf{T}_q^n$ and $\mathbf{e}_2 \in \mathsf{R}^m$, as well as the (fixed) matrices $\mathbf{C} \in \mathsf{Z}_q^{k\times m}$ and $\mathbf{F} \in \mathsf{Z}^{n\times m}$. It holds that

$$\begin{aligned}\mathbf{s}\mathbf{A} + \mathbf{e}_1\mathbf{F} + \mathbf{e}_2 &= \mathbf{s}\mathbf{B}\mathbf{C} + \mathbf{s}\mathbf{F} + \mathbf{e}_1\mathbf{F} + \mathbf{e}_2 \\ &= (\mathbf{s}\mathbf{B})\mathbf{C} + (\mathbf{s} + \mathbf{e}_1)\mathbf{F} + \mathbf{e}_2.\end{aligned}$$

Assume that we can describe $\mathbf{s}\mathbf{B}$ using ℓ bits. We can then bound

$$\begin{aligned}\tilde{H}_\infty(\mathbf{s}|\mathbf{A}, \mathbf{s}\mathbf{A} + \mathbf{e}) &\geq \tilde{H}_\infty(\mathbf{s}|\mathbf{B}, \mathbf{C}, \mathbf{F}, \mathbf{s}\mathbf{B}, \mathbf{s} + \mathbf{e}_1, \mathbf{e}_2) \\ &= \tilde{H}_\infty(\mathbf{s}|\mathbf{s}\mathbf{B}, \mathbf{s} + \mathbf{e}_1) \\ &\geq \tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_1) - \ell\end{aligned}$$

where the equality follows from the fact that $\mathbf{e}_2$ is independent of everything else and the second inequality follows from the entropy chain-rule. We will now distinguish two cases, depending on whether $\mathbf{s}$ is short or not.

1. In the first case, assume that $\mathcal{W}$ is the discrete gaussian distribution $D^{n\times k}_{\mathsf{Z},\gamma}$, thus it holds that $\|\mathbf{B}\| \leq C\cdot\gamma\cdot\sqrt{n}$ except with negligible probability. Moreover, assume that we have a bound $\|\mathbf{s}\| \leq r$. Then it holds that $\|\mathbf{sB}\| \leq C\cdot\gamma\cdot\sqrt{n}r$. We can bound the number of $\mathbf{z}\in\mathsf{Z}^k$ with $\|\mathbf{z}\| \leq C\cdot\gamma\cdot\sqrt{n}r$ by $(2C\cdot\gamma\cdot\sqrt{n}r)^k$. Consequently, we can describe $\mathbf{sB}$ using $\ell = k\cdot\log(2C\cdot\gamma\cdot\sqrt{n}r)$ bits.

2. If we have no bound on $\|\mathbf{s}\|$, we can generically describe $\mathbf{sB}$ using $\ell = k\cdot\log(q)$ bits as $\mathbf{sB}\in\mathsf{Z}^k_q$.

As by assumption we have that

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e_1}) \geq k\cdot\log(\min\{2C\cdot\gamma\cdot\sqrt{n}r, q\}) + \omega(\log(\lambda)),$$

it follows that

$$\begin{aligned}\Pr[\mathcal{A}(\mathbf{sA}+\mathbf{e})=\mathbf{s}] &\leq 2^{-\tilde{H}_\infty(\mathbf{s}|\mathbf{A},\mathbf{sA}+\mathbf{e})}\\ &\leq 2^{-\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e_1})+\ell}\\ &\leq 2^{-\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e_1})+k\cdot\log(\min\{2C\cdot\gamma\cdot\sqrt{n}r,q\})}\\ &\leq 2^{-\omega(\log(\lambda))},\end{aligned}$$

which is negligible.This concludes the proof of the first part of the statement.

To prove the second part of the statement, consider the following additional hybrids, where we will set $\mathcal{W}$ to be the uniform distribution on $\mathsf{Z}^{n\times k}_q$.

- $\mathcal{H}_4$:
  - $\mathbf{s}\leftarrow_\$ \mathcal{S}$
  - $\mathbf{s}^*\leftarrow_\$ \mathsf{Z}^k_q$
  - $\mathbf{B}\leftarrow_\$ \mathsf{Z}^{n\times k}_q$, $\mathbf{C}\leftarrow_\$ \mathsf{Z}^{k\times m}_q$, $\mathbf{F}\leftarrow_\$ D^{n\times m}_{\mathsf{Z},\gamma}$
  - If $\|\mathbf{B}\| > C\cdot\gamma\cdot\sqrt{n}$ or $\|\mathbf{F}\| > C\cdot\gamma\cdot\sqrt{m}$ output $\perp$
  - $\mathbf{A}\leftarrow\mathbf{B}\cdot\mathbf{C}+\mathbf{F}$
  - Set $\mathbf{\Sigma}\leftarrow\sigma^2\mathbf{I}-\sigma_1^2\mathbf{F}^T\mathbf{F}$
  - $\mathbf{e_1}\leftarrow_\$ D^n_\sigma$, $\mathbf{e_2}\leftarrow_\$ D^m_{\sqrt{\mathbf{\Sigma}}}$
  - $\mathbf{e}\leftarrow\mathbf{e_1F}+\mathbf{e_2}$
  - Output $(\mathbf{A}, \mathbf{s}^*\mathbf{C}+(\mathbf{s}+\mathbf{e_1})\mathbf{F}+\mathbf{e_2})$
- $\mathcal{H}_5$:
  - $\mathbf{s}\leftarrow_\$ \mathcal{S}$
  - $\mathbf{s}^*\leftarrow_\$ \mathsf{Z}^k_q$
  - $\mathbf{B}\leftarrow_\$ \mathsf{Z}^{n\times k}_q$, $\mathbf{C}\leftarrow_\$ \mathsf{Z}^{k\times m}_q$, $\mathbf{F}\leftarrow_\$ D^{n\times m}_{\mathsf{Z},\gamma}$
  - If $\|\mathbf{B}\| > C\cdot\gamma\cdot\sqrt{n}$ or $\|\mathbf{F}\| > C\cdot\gamma\cdot\sqrt{m}$ output $\perp$
  - $\mathbf{A}\leftarrow\mathbf{B}\cdot\mathbf{C}+\mathbf{F}$
  - $\mathbf{e}\leftarrow_\$ D^m_\sigma$

– Output $(\mathbf{A}, \mathbf{s}^*\mathbf{C} + \mathbf{sF} + \mathbf{e})$

We will now show that $\mathcal{H}_3$ and $\mathcal{H}_4$ are statistically close via the leftover hash lemma. Note that the only difference between $\mathcal{H}_3$ and $\mathcal{H}_4$ is that in $\mathcal{H}_4$ we have replaced $\mathbf{sB}$ by a uniformly random $\mathbf{s}^*$. Moreover, the only other term depending on $\mathbf{s}$ is $\mathbf{s} + \mathbf{e}_1$. Consequently, we can bound the statistical distance between $\mathcal{H}_3$ and $\mathcal{H}_4$ by

$$\begin{aligned}\Delta(\mathcal{H}_3, \mathcal{H}_4) &\leq \Delta((\mathbf{B}, \mathbf{sB}, \mathbf{s} + \mathbf{e}_1), (\mathbf{B}, \mathbf{s}^*, \mathbf{s} + \mathbf{e}_1)) \\ &\leq \sqrt{q^k \cdot 2^{-\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}_1)}} \\ &\leq \sqrt{2^{k \cdot \log(q)} \cdot 2^{-k \cdot \log(q) + -\omega(\log(\lambda))}} \\ &= 2^{-\omega(\log(\lambda))},\end{aligned}$$

which is negligible. The second inequality follows by the generalized leftover hash lemma, whereas the third inequality follows from the assumption $\tilde{H}_\infty(\mathbf{s}|\mathbf{s} + \mathbf{e}_1) \geq k \cdot \log(q) + \omega(\log(\lambda))$. Note that we can apply the leftover hash lemma whenever q is prime or $\mathbf{s}$ is binary.

Next, we claim that $\mathcal{H}_4$ and $\mathcal{H}_5$ are identically distributed. To see this, note that all we did was reversing the decomposition of $\mathbf{e} = \mathbf{e}_1\mathbf{F} + \mathbf{e}_2$. Thus, $\mathcal{H}_4$ and $\mathcal{H}_5$ are identically distributed by the same argument as $\mathcal{H}_2$ and $\mathcal{H}_3$ are identically distributed.

Now assume there was a PPT distinguisher $\mathcal{D}$ which distinguishes $\mathsf{ent\text{-}dLWE}$ with non-negligible advantage. By the above argument, such a distinguisher must also have non-negligible advantage in distinguishing $(\mathbf{A}, \mathbf{y})$ from $(\mathbf{A}, \mathbf{u} + \mathbf{e})$, where $(\mathbf{A}, \mathbf{y}) \leftarrow_\$ \mathcal{H}_5$, $\mathbf{u} \leftarrow_\$ \mathsf{Z}_q^m$ and $\mathbf{e} \leftarrow_\$ D_\sigma$. From such as distinguisher we can construct a distinguisher $\mathcal{D}'$ against $\mathsf{dLWE}(q, k, m, D_\sigma)$ as follows. $\mathcal{D}'$ gets as input a matrix $\mathbf{C} \in \mathsf{Z}_q^{k \times m}$ and a vector $\mathbf{z} \in \mathsf{Z}_q^m$. $\mathcal{D}'$ proceeds as follows:

- $\mathbf{s} \leftarrow_\$ \mathcal{S}$
- $\mathbf{B} \leftarrow_\$ D_{\mathsf{Z},\gamma}^{n \times k}$, $\mathbf{F} \leftarrow_\$ D_{\mathsf{Z},\gamma}^{n \times m}$
- If $\|\mathbf{B}\| > C \cdot \gamma \cdot \sqrt{n}$ or $\|\mathbf{F}\| > C \cdot \gamma \cdot \sqrt{m}$ output $\perp$
- $\mathbf{A} \leftarrow \mathbf{B} \cdot \mathbf{C} + \mathbf{F}$
- $\mathbf{y} = \mathbf{z} + \mathbf{sF}$
- Output $\mathcal{D}(\mathbf{A},)$

We claim that $\mathcal{D}'$ has the same advantage as $\mathcal{D}$. First consider the case that the input of $\mathcal{D}'$ is of the form $(\mathbf{C}, \mathbf{z} = \mathbf{s}^*\mathbf{C} + \mathbf{e})$, where $\mathbf{D} \leftarrow_\$ \mathsf{Z}_q^{k \times m}$, $\mathbf{s}^* \leftarrow_\$ \mathsf{Z}_q^k$ and $\mathbf{e} \leftarrow_\$ D_\sigma$. Then it holds that

$$\mathbf{y} = \mathbf{z} + \mathbf{sF} = \mathbf{s}^*\mathbf{C} + \mathbf{sF} + \mathbf{e},$$

i.e. $(\mathbf{A}, \mathbf{y})$ is distributed according to $\mathcal{H}_5$.

On the other hand, if the input of $\mathcal{D}'$ is distributed according to $(\mathbf{C}, \mathbf{z} = \mathbf{u} + \mathbf{e})$, then it holds that $\mathbf{y} = \mathbf{z} + \mathbf{sF} \equiv \mathbf{u}' + \mathbf{e}$ for a uniformly random $\mathbf{u}' \leftarrow_\$ \mathsf{Z}_q^m$. Consequently $(\mathbf{A}, \mathbf{y})$ has the same distribution as $(\mathbf{A}, \mathbf{u} + \mathbf{e})$, where $(\mathbf{A}, \mathbf{y}) \leftarrow_\$ \mathcal{H}_5$, $\mathbf{u} \leftarrow_\$ \mathsf{Z}_q^m$ and $\mathbf{e} \leftarrow_\$ D_\sigma$. We conclude that $\mathcal{D}'$ has the same advantage as $\mathcal{D}$, which contradicts the hardness of $\mathsf{dLWE}(q, k, m, D_\sigma)$. This concludes the proof. □

5 Noise-Lossiness for Modular Gaussians

In this Section, we will compute the noise lossiness for general high-minentropy distributions. We further show that considerable improvements can be achieved when considering short distributions. Our Lemmas in this Section can be seen as *strong converse coding theorems* for gaussian channels. I.e. if a distribution codes above a certain information rate, then information must be lost and noise lossiness quantifies how much information is lost. The following lemma will allow us to bound $\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e})$ by suitably bounding $\max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y}-\mathbf{s}^*)$.

Lemma 5.1. *Let $q \in \mathsf{N}$ be a modulus and $x\ n, m \in \mathsf{N}$ with $m > n$. Let $\mathbf{s}$ be a random variable on Z_q^k with min-entropy $\tilde{H}_\infty(\mathbf{s})$. Let χ be a noise distribution over R^n and let $\mathbf{e} \sim \chi$. Then it holds that*

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}) \geq \tilde{H}_\infty(\mathbf{s}) - \log\left(\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y}-\mathbf{s}^*)d\mathbf{y}\right)$$

in the case that χ is continuous and

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}) \geq \tilde{H}_\infty(\mathbf{s}) - \log\left(\sum_{\mathbf{y}} \max_{\mathbf{s}^*} \Pr_{\mathbf{e}}[\mathbf{e}=\mathbf{y}-\mathbf{s}^*]\right)$$

in the case that χ is discrete. Moreover, if $\mathbf{s}$ is a at distribution then equality holds.

Proof. The lemma follows from the following derivation in the continuous case. The discrete case follows analogously.

$$\begin{aligned}
\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}) &= -\log\left(\mathop{\mathsf{E}}_{\mathbf{y}}[\max_{\mathbf{s}^*\in S} \Pr_{\mathbf{s},\mathbf{e}}[\mathbf{s}=\mathbf{s}^*|\mathbf{s}+\mathbf{e}=\mathbf{y}]]\right) \\
&= -\log\left(\int_{\mathbf{y}} p_{\mathbf{s}+\mathbf{e}}(\mathbf{y}) \cdot \max_{\mathbf{s}^*} \Pr_{\mathbf{s},\mathbf{e}}[\mathbf{s}=\mathbf{s}^*|\mathbf{s}+\mathbf{e}=\mathbf{y}]d\mathbf{y}\right) \\
&= -\log\left(\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{s},\mathbf{s}+\mathbf{e}}(\mathbf{s}^*,\mathbf{y})d\mathbf{y}\right) \\
&= -\log\left(\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{s}+\mathbf{e}|\mathbf{s}=\mathbf{s}^*}(\mathbf{y}) \cdot \underbrace{\Pr[\mathbf{s}=\mathbf{s}^*]}_{\leq 2^{-H_\infty(\mathbf{s})}} d\mathbf{y}\right) \\
&\geq \tilde{H}_\infty(\mathbf{s}) - \log\left(\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y}-\mathbf{s}^*)d\mathbf{y}\right) .
\end{aligned}$$

To see that equality holds for flat distributions, note that in this case we have $\Pr[\mathbf{s}=\mathbf{s}^*] = 2^{-H_\infty(\mathbf{s})}$. □

5.1 General High Entropy Secrets

We first turn to the case of general high-entropy secrets and prove the following lemma.

Lemma 5.2. *Let n be an integer, let q be a modulus and σ_1 be a parameter for a gaussian. Assume that*

$$\frac{q}{\sigma_1} \geq \sqrt{\frac{\ln(4n)}{\pi}}.$$

Let $\mathbf{s}$ *be a random variable on* Z_q^n *and* $\mathbf{e}_1 \sim D_{\sigma_1} \mod q$. *Then it holds that*

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}_1) \geq \tilde{H}_\infty(\mathbf{s}) - n \cdot \log(q/\sigma_1) - 1$$

We remark that the requirement $\frac{q}{\sigma_1} \geq \sqrt{\frac{\ln(4n)}{\pi}}$ is made for technical reasons, but we restrict ourselves to keep the proof simple. We also remark that this condition is essentially trivially fulfilled by interesting parameter choices.

We can instantiate Theorem 4.1 with Lemma 5.2 obtaining the following corollary.

Corollary 5.3. *Let* C *be a global constant. Let* $q = q(\lambda)$ *be a modulus and let* $n, m, k = \mathsf{poly}(\lambda)$. *Let* $\gamma, \sigma_1 > 0$. *Assume that* $\mathcal{S}$ *is a distribution on* Z_q^n *with* $\tilde{H}_\infty(\mathbf{s}) > k \cdot \log(q) + n \cdot \log(q/\sigma_1) + \omega(\log(\lambda))$. *Now let* $\sigma > C \cdot \sqrt{m} \cdot \gamma\sigma_1$. *Then* $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, D_\sigma)$ *is hard, provided that* $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$ *is hard.*

of Lemma 5.2. It holds that

$$\begin{aligned}
\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y}-\mathbf{s}^*)d\mathbf{y} &= \frac{1}{\rho_{\sigma_1}(\mathsf{R}^n)} \int_{\mathbf{y}} \max_{\mathbf{s}^*} \hat{\rho}_{q\mathsf{Z}^n,\sigma_1}(\mathbf{y}-\mathbf{s}^*)d\mathbf{y} \\
&\leq \frac{1}{\rho_{\sigma_1}(\mathsf{R}^n)} \cdot \int_{\mathbf{y}} 2d\mathbf{y} \\
&= 2 \cdot \frac{q^n}{\rho_{\sigma_1}(\mathsf{R}^n)} \\
&= 2 \cdot \frac{q^n}{\sigma_1^n},
\end{aligned}$$

where the $\hat{\rho}_{q\mathsf{Z}^n,\sigma_1}(\mathbf{y}-\mathbf{s}^*) \leq 2$ follows by Lemma 2.4 as $\frac{q}{\sigma_1} \geq \sqrt{\frac{\ln(4n)}{\pi}}$. We can conclude by Lemma 5.1 that

$$\begin{aligned}
\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}) &\geq \tilde{H}_\infty(\mathbf{s}) - \log\left(\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y}-\mathbf{s}^*)d\mathbf{y}\right) \\
&\geq \tilde{H}_\infty(\mathbf{s}) - n \cdot \log(q/\sigma_1) - 1.
\end{aligned}$$

□

5.2 Short Secrets

We will now turn to the case where the secret has bounded norm.

Lemma 5.4. *Let* n *be an integer, let* q *be a modulus and* σ_1 *be a parameter for a gaussian. Assume that* $\mathbf{s}$ *is a random-variable on* Z_q^n *such that* $\|\mathbf{s}\| \leq r$ *for a parameter* $r = r(\lambda)$. *Let* $\mathbf{e}_1 \sim D_{\sigma_1} \mod q$ *Then it holds that*

$$\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}_1) \geq \tilde{H}_\infty(\mathbf{s}) - \sqrt{2\pi n} \cdot \frac{r}{\sigma_1} \log(e).$$

In particular, if $\sigma_1 > \sqrt{n} \cdot r$, then $\tilde{H}_\infty(\mathbf{s}|\mathbf{s}+\mathbf{e}_1) \geq \tilde{H}_\infty(\mathbf{s}) - \pi \log(e)$. We can instantiate Theorem 4.1 with Lemma 5.4 obtaining the following corollary.

Corollary 5.5. *Let C be a global constant. Let $q = q(\lambda)$ be a modulus and let $n, m, k = \mathsf{poly}(\lambda)$. Let $\gamma = \gamma(\lambda) > 0$ and $\sigma_1 = \sigma_1(\lambda) > 0$. Assume that $\mathcal{S}$ is a r-bounded distribution with $\tilde{H}_\infty(\mathbf{s}) > k \cdot \log(2C \cdot \gamma \cdot \sigma_1) + \sqrt{2\pi n} \cdot \frac{r}{\sigma_1} \log(e) + \omega(\log(\lambda))$. Now let $\sigma > C \cdot \sqrt{m} \sigma_1 \cdot \gamma$. Then $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, D_\sigma)$ is hard, provided that $\mathsf{dLWE}(q, k, D_{\mathsf{Z},\gamma})$ is hard.*

of Lemma 5.4. Fix some $\sigma_2 > \sigma_1$. Since it holds that $\|\mathbf{s}\| \leq r$, it holds that

$$\begin{aligned}
\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y} - \mathbf{s}^*) d\mathbf{y} &= \frac{1}{\rho_{\sigma_1}(\mathsf{R}^n)} \int_{\mathbf{y}} \max_{\mathbf{s}^*} \hat{\rho}_{q\mathsf{Z}^n, \sigma_1}(\mathbf{y} - \mathbf{s}^*) d\mathbf{y} \\
&\leq \frac{1}{\rho_{\sigma_1}(\mathsf{R}^n)} \int_{\mathbf{y}} \max_{\mathbf{s}^*} e^{\pi \frac{\|\mathbf{s}^*\|^2}{\sigma_2^2 - \sigma_1^2}} \cdot \hat{\rho}_{q\mathsf{Z}^n, \sigma_2}(\mathbf{y}) d\mathbf{y} \\
&\leq \frac{1}{\rho_{\sigma_1}(\mathsf{R}^n)} \cdot e^{\pi \frac{r^2}{\sigma_2^2 - \sigma_1^2}} \cdot \int_{\mathbf{y}} \hat{\rho}_{q\mathsf{Z}^n, \sigma_2}(\mathbf{y}) d\mathbf{y} \\
&= e^{\pi \frac{r^2}{\sigma_2^2 - \sigma_1^2}} \cdot \frac{\rho_{\sigma_2}(\mathsf{R}^n)}{\rho_{\sigma_1}(\mathsf{R}^n)} \\
&= e^{\pi \frac{r^2}{\sigma_2^2 - \sigma_1^2}} \cdot \left(\frac{\sigma_2}{\sigma_1}\right)^n
\end{aligned}$$

Now, setting $\sigma_2 = \sigma_1 \cdot \sqrt{1 + \eta}$ we get that

$$\begin{aligned}
\int_{\mathbf{y}} \max_{\mathbf{s}^*} p_{\mathbf{e}}(\mathbf{y} - \mathbf{s}^*) d\mathbf{y} &\leq e^{\pi \frac{r^2}{\sigma_2^2 - \sigma_1^2}} \cdot \left(\frac{\sigma_2}{\sigma_1}\right)^n \\
&= e^{\pi \frac{r^2}{\eta \sigma_1^2}} \cdot (1 + \eta)^{n/2} \\
&\leq e^{\pi \frac{r^2}{\eta \sigma_1^2} + \frac{n\eta}{2}}
\end{aligned}$$

By Lemma 5.1, we can conclude that

$$\tilde{H}_\infty(\mathbf{s} | \mathbf{s} + \mathbf{e}_1) \geq \tilde{H}_\infty(\mathbf{s}) - \left(\pi \frac{r^2}{\eta \sigma_1^2} + \frac{n\eta}{2}\right) \log(e).$$

Recall that η is still a free parameter. This expression is minimized by choosing $\eta = \sqrt{\frac{2\pi}{n}} \frac{r}{\sigma_1}$, which yields

$$\tilde{H}_\infty(\mathbf{s} | \mathbf{s} + \mathbf{e}_1) \geq \tilde{H}_\infty(\mathbf{s}) - \sqrt{2\pi n} \cdot \frac{r}{\sigma_1} \log(e).$$

□

6 Tightness of the Result

In this Section, we will show that for general moduli and general min-entropy distributions our result is tight up to polynomial factors.

For a modulus q and a noise parameter σ, we will provide an example of a distribution $\mathbf{s}$ with min-entropy $\approx n \cdot \log(q/\sigma)$, such that $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{X}, \chi)$ is easy. For this counter-example, the choice of the modulus q is critical.

Lemma 6.1. *Let $q = q(\lambda)$ be a modulus such that q has a divisor p of size $|p| > 2B+1$, let $n, m =$ poly(λ) and let χ be a B-bounded error-distribution. Define the distribution $\mathcal{S}$ to be the uniform distribution on $p \cdot \mathsf{Z}_q^n$. Then there exists an efficient algorithm $\mathcal{A}$ that solves* ent-LWE$(q, n, m, \mathcal{S}, \chi)$

Corollary 6.2. *There exist moduli q and distributions $\mathcal{S}$ with min-entropy $\geq n \cdot (\log(q/\sigma) - \log(\log(\lambda)))$ such that* ent-LWE$(q, n, m, \mathcal{S}, D_\sigma)$ *is easy.*

The corollary follows from Lemma 6.1 by choosing p such that $p = 2\log(\lambda)\cdot\sigma+1$ and noting that a gaussian of parameter σ is $\log(\lambda)\cdot\sigma$ bounded, except with negligible probability. Moreover, for this choice of p the distribution $\mathcal{S}$ in Lemma 6.1 has min-entropy $n\cdot\log(q/p) \geq n\cdot\log(q/\sigma) - 2\log(\log(\lambda))$

Proof of Lemma 6.1. Assume that reduction modulor p computes a *central* residue class representation in $[-p/2, p/2]$. The algorithm $\mathcal{A}$ proceeds as follows.

$\mathcal{A}(\mathbf{A}, \mathbf{y})$:

- Compute $\mathbf{e} \leftarrow \mathbf{y} \mod p$
- Solve the equation system $\mathbf{s} \cdot \mathbf{A} = \mathbf{y} - \mathbf{e}$ for $\mathbf{s}$, e.g. via Gaussian elimination.
- Output $\mathbf{s}$

To see that the algorithm $\mathcal{A}$ is correct, note that

$$\mathbf{y} \mod p = (\mathbf{s} \cdot \mathbf{A} + \mathbf{e}) \mod p = (p \cdot \mathbf{r} \cdot \mathbf{A} + \mathbf{e}) \mod p = \mathbf{e}$$

as $p \geq 2B$ and $\|\mathbf{e}\| \leq B$.

□

7 Barriers for Entropic LWE

In the last Section we provided an attack on entropic LWE when the min-entropy of the secret is below $n\cdot\log(q/\sigma)$ for a worst-case choice of the modulus q. On might still hope that for more benign choices of the modulus q this problem might be hard in this entropy regime. In this section we will provide a barrier for the hardness of entropic LWE in this regime for any modulus. In particular, we will show that for entropies below $n \cdot \log(q/\sigma)$, the hardness of entropic LWE does not follow from any standard assumption in a black-box way. This leaves open the possibility that in this regime the hardness of entropic LWE may be established from more exotic *knowledge assumptions*. To establish our result, we will use a framework developed by Wichs [Wic13].

7.1 Simulatable Attacks

We first recall the notion of cryptographic games as a way to characterize cryptographic standard assumptions due to Haitner and Holenstein [HH09]. This chracterization captures essentially all falsifiable assumptions [Nao03] used in cryptography, such as LWE.

Definition 7.1 (Cryptographic Games [HH09])**.** *A cryptographic game $\mathcal{C} = (\Gamma, c)$ is defined by a (possibly inefficient) randomized machine Γ, called the challenger, and a constant $c \in [0, 1)$. On*

input a security parameter 1^λ, the challenger interacts with an attack $\mathcal{A}(1^\lambda)$ and outputs a bit b. Denote this by $\Gamma(1^\lambda) \quad \mathcal{A}(1^\lambda)$. The advantage of an attacker $\mathcal{A}$ against $\mathcal{C}$ is de ned by

$$\mathsf{Adv}_{\mathcal{C}}^{\mathcal{A}}(1^\lambda) = \Pr[(\Gamma(1^\lambda) \quad \mathcal{A}(1^\lambda)) = 1] - c.$$

We say that a cryptographic game $\mathcal{C}$ is secure if for all PPT attackers $\mathcal{A}$ the advantage $\mathsf{Adv}_{\mathcal{C}}^{A}(\lambda)$ is negligible.

Definition 7.2 (Black-Box Reduction)**.** *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be cryptographic games. A black-box reduction deriving the security of $\mathcal{C}_2$ from the security of $\mathcal{C}_1$ is an oracle PPT-machine $\mathcal{B}^{(\cdot)}$ for which there are constants c, λ_0 such that for all $\lambda \geq \lambda_0$ and all (possibly ine cient, non-uniform) attackers $\mathcal{A}_\lambda$ with advantage $\mathsf{Adv}_{\mathcal{C}_1}^{\mathcal{A}_\lambda}(\lambda) \geq 1/2$, we have $\mathsf{Adv}_{\mathcal{C}_2}^{\mathcal{B}^{\mathcal{A}_\lambda}}(\lambda) \geq \lambda^{-c}$.*

We remark that the choice of the constant $1/2$ for the advantage of $\mathcal{A}_\lambda$ is arbitrary and can be replaced by a non-negligible function (depending $\mathcal{A}_\lambda$). We now recall the notion of *simulatable attacks* [Wic13].

Definition 7.3 (Simulatable Attacks [Wic13])**.** *An ϵ-simulatable attack on an assumption $\mathcal{C}$ is a tuple $(\mathcal{A}, \mathsf{Sim})$ such that $\mathcal{A}$ is a stateless, non-uniform possibly ine cient attacker against $\mathcal{C}$, and Sim is a stateful PPT simulator. We require the following two properties to hold.*

- *The (ine cient) attacker $\mathcal{A}$ successfully breaks $\mathcal{C}$ with advantage $1 - \mathsf{negl}(\lambda)$.*
- *For every (possibly ine cient) oracle machine $\mathcal{M}^{(\cdot)}$ making at most q queries to its oracle it holds that*

  $$|\Pr[\mathcal{M}^{\mathcal{A}(1^\lambda,1)}(1^\lambda) = 1] - \Pr[\mathcal{M}^{\mathsf{Sim}(1^\lambda)} = 1]| \leq \mathsf{poly}(q) \cdot \epsilon.$$

  where the probabilities are taken over all the random choices involved.

We use the shorthand simulatable attack for ϵ-simulatable attack with some negligible ϵ.

We remark that for reasons of conceptual simplicity Wichs [Wic13] required the advantage of the simulatable adversary $\mathcal{A}$ to be 1. But it can easily be verified that Theorem 7.4 below also works with our slightly relaxed notion which allows the unbounded adversary to have advantage $1 - \mathsf{negl}(\lambda)$. The following theorem by Wichs [Wic13] shows that the existence of a simulatable attack for some assumption $\mathcal{C}_1$ implies that there cannot by a reduction $\mathcal{B}$ which reduces the hardness of $\mathcal{C}_1$ to any standard assumption $\mathcal{C}_2$, where $\mathcal{C}_1$ and $\mathcal{C}_2$ are cryptographic games in the sense of Definition 7.1.

Theorem 7.4 ([Wic13] Theorem 4.2)**.** *If there exists a simulatable attack against some assumption $\mathcal{C}_1$ and there is a black-box reduction $\mathcal{B}$ reducing the security of $\mathcal{C}_1$ to some assumption $\mathcal{C}_2$, then $\mathcal{C}_2$ is not secure.*

The idea for the proof of this theorem is simple: If an attack $\mathcal{A}$ against $\mathcal{C}_1$ is simulatable, then the behavior of $\mathcal{B}^{\mathsf{Sim}}$ will be indistinguishable from $\mathcal{B}^{\mathcal{A}}$. But since $\mathcal{A}$ breaks $\mathcal{C}_1$, it holds that $\mathcal{B}^{\mathcal{A}}$ breaks $\mathcal{C}_2$. Therefore, the efficient algorithm $\mathcal{B}^{\mathsf{Sim}}$ must also break $\mathcal{C}_2$, implying that $\mathcal{C}_2$ is insecure.

7.2 A Simulatable Attack for Entropic LWE

We will now provide a simulatable attack against entropic (search-)LWE. The attack consists of a pair of a min-entropy distribution $\mathcal{S}$ and an attacker $\mathcal{A}$. Since we want to prove a result for general min-entropy distributions, we assume that both the adversary and the min-entropy distribution $\mathcal{S}$ are adversarially chosen. Thus, we can consider the distribution $\mathcal{S}$ as running a coordinated attack with the attacker $\mathcal{A}$. More importantly, any black-box reduction $\mathcal{B}$ reducing the entropic LWE to a standard assumption will only have black-box access to the distribution $\mathcal{S}$. We remark that, to the best of our knowledge, currently all reductions in the realm of leakage resilient cryptography only make black-box use of the distribution. Making effective non-black box use of an adversarially chosen sampling circuit seems out of reach for current techniques. Assume in the following that $m \geq 2n$ and let χ be a B-bounded error distribution. Furthermore let k be a positive integer. Consider the following attacker, consisting of the adversary $\mathcal{A}$ and the distribution $\mathcal{S}$.

- The distribution $\mathcal{S}$ is a flat distribution on a set S of size 2^k, where the set S is chosen uniformly random.
- $\mathcal{A}_S(\mathbf{A}, \mathbf{y})$: Given a pair $(\mathbf{A}, \mathbf{y})$, the attacker $\mathcal{A}$ proceeds as follows:
  - Check if the matrix $\mathbf{A}$ has an invertible column-submatrix, if not abort and output $\perp$ (this check can be performed efficiently using linear algebra).
  - Compute a set $I \subseteq [m]$ of size n such that the column-submatrix $\mathbf{A}_I$ is invertible (where $\mathbf{A}_I$ is obtained by dropping all columns of $\mathbf{A}$ that do not have indices in I).
  - Set $\mathbf{A}' = \mathbf{A}_I$ and $\mathbf{y}' = \mathbf{y}_I$ (i.e. $\mathbf{y}'$ is $\mathbf{y}$ projected to the coordinates in I)
  - Initialize a set $S' = \emptyset$
  - For every $\mathbf{s} \in S$, check if $\|\mathbf{y} - \mathbf{sA}\|_\infty \leq B$, if so include $\mathbf{s}$ in the set S'.
  - Choose an $\mathbf{s} \leftarrow_\$ S'$ uniformly random and output $\mathbf{s}$

First observe that whenever the matrix $\mathbf{A}$ has an invertible submatrix, then $\mathcal{A}$ does have advantage 1. The probability that $\mathbf{A}$ does not have an invertible submatrix is at most $\log(q) \cdot 2^{n-m} = \log(q) \cdot 2^{-n}$, which is negligible (see Section 2). Consequently, $\mathcal{A}$ breaks $\mathsf{ent\text{-}LWE}(q, n, m, \mathcal{S}, \chi)$ with probability $1 - \mathsf{negl}(\lambda)$.

We will now provide our simulator for the adversary $\mathcal{A}$ and the distribution $\mathcal{S}$. The simulator jointly simulates the distribution $\mathcal{S}$ and the attacker $\mathcal{A}$, i.e. from the interface of an oracle machine $\mathcal{B}$ it holds that $\mathsf{Sim}(1^\lambda, \cdot, \cdot)$ simulates $(\mathcal{S}(\cdot), \mathcal{A}(\cdot))$. The advantage of the simulator stems from having a joint view of the samples provided so far and the inputs of the adversary $\mathcal{A}$. The main idea of our simulator is that is samples the set S lazily and keeps track of all the samples S^* it gave out so far. When provided with an instance $(\mathbf{A}, \mathbf{y})$, it will perform the same check as $\mathcal{A}$ but restricted to the set X' and therefore run in time $O(q)$. Recall that the simulator is stateful.

- Simulator $\mathsf{Sim}(1^\lambda, \cdot, \cdot)$:
  - Initialize a set $S^* = \emptyset$.
  - Whenever a sample is queried from $\mathcal{S}$, choose $\mathbf{s} \leftarrow_\$ Z_q^n$ uniformly random, include $\mathbf{s}$ in the set S^* and output $\mathbf{s}$.
  - Whenever an instance is provided to $\mathcal{A}$, do the following:

* Initialize a set $S' = \emptyset$
* Check for every $\mathbf{s} \in S^*$, check if $\|\mathbf{y} - \mathbf{sA}\|_\infty \leq B$, if so include $\mathbf{s}$ in the set S'.
* Choose an $\mathbf{s} \leftarrow_\$ S'$ uniformly random and output $\mathbf{s}$.

We will now show that the simulator Sim simulates the attack $(\mathcal{A}, \mathcal{X})$ with negligible error. We need the following lemma.

Lemma 7.5. *Let $\mathbf{z} \leftarrow_\$ \mathsf{Z}_q^n$ be distributed uniformly random. Then it holds that*

$$\Pr[\|\mathbf{z}\|_\infty \leq B] \leq ((2B+1)/q)^n.$$

Proof. Since all the components z_i of $\mathbf{z}$ are distributed uniformly and independently, it holds that

$$\Pr[\|\mathbf{z}\|_\infty \leq B] = \prod_{i=1}^{n} \Pr[|z_i| \leq B] \leq ((2B+1)/q)^n.$$

□

Theorem 7.6. *Let $\chi = \chi(\lambda)$ be a B-bounded error-distribution. Further, let $k < n \cdot \log(q/(2B+1)) - \omega(\log(\lambda))$ be an integer. Let $\bar{\mathcal{S}}$ be the family of all distributions on Z_q^n with min-entropy at most k. Then, if there is a reduction $\mathcal{B}$ from $\mathsf{ent\text{-}LWE}(q, n, m, \bar{\mathcal{S}}, \chi)$ to any cryptographic game $\mathcal{C}$, then $\mathcal{C}$ is not secure.*

Proof. We will show that $(\mathcal{S}, \mathcal{A})$ is a simulatable attack for $\mathsf{ent\text{-}LWE}(q, n, m, \bar{\mathcal{S}}, \chi)$, where $\mathcal{S} \in \bar{\mathcal{S}}$. The claim then follows immediately by Theorem 7.4. We will prove the statement by a hybrid argument. Fix a (possibly inefficient) machine $\mathcal{B}$, assume that $\mathcal{B}$ queries $\mathcal{S}$ at most ℓ_1 times and $\mathcal{A}$ at most ℓ_2 times. Let $\mathcal{H}_0(1^\lambda) = \mathcal{B}^{\mathcal{S},\mathcal{A}}(1^\lambda)$. For $i = 1, \ldots, \ell_2$ we define the following hybrids $\mathcal{H}_i$.

$\mathcal{H}_i$: $\mathcal{H}_i$ behaves identical to $\mathcal{B}^{\mathsf{Sim}}(1^\lambda, \cdot, \cdot)$ until $\mathcal{B}$ makes its $i+1$-st call to the attacker oracle. At this point, take the set S^* computed by Sim and complement it to a set S by choosing the missing elements uniformly random and fixing them. Moreover, starting from this call answer every call to the attacker oracle by $\mathcal{A}_S(\cdot)$.

Note first that $\mathcal{H}_{\ell_2}$ is identically distributed to $\mathcal{B}^{\mathsf{Sim}}(1^\lambda)$. Also observe that from the view of $\mathcal{B}$ the experiments $\mathcal{H}_i$ and $\mathcal{H}_{i+1}$ are identically distributed until the $i+1$-st query to the attacker oracle.

Now fix the any state of $\mathcal{B}$ and Sim when $\mathcal{B}$ sends its $i+1$-st query $(\mathbf{A}, \mathbf{y})$ and assume that $\mathcal{B}$ has observed outputs from $\tilde{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_{\ell_1'}\}$ from $\mathcal{S}$ with $\ell_1' \leq \ell_1$. At this point $\mathcal{H}_i$ samples elements $\mathbf{s}'_{\ell_1'+1}, \ldots, \mathbf{s}'_K$ uniformly at random. We claim that if it holds for all $j \in \{\ell+1, \ldots, K\}$ that

$$\|\mathbf{s}'_j\mathbf{A} - \mathbf{y}\| > B,$$

then the outputs of $\mathcal{A}_S$ and Sim are identical. To see this, note that if this condition holds, then both $\mathcal{A}_S$ and Sim will either output a uniformly random element from $\tilde{S} = \{\mathbf{s} \in X^* \mid \|\mathbf{sA} - \mathbf{y}\| \leq B\}$ or $\perp$ if the set $\tilde{S}$ is empty.

Thus, to bound the statistical distance between $\mathcal{H}_i$ and $\mathcal{H}_{i+1}$ it is sufficient to bound the probability of the above event. Note that we can sample $\mathbf{s}'_{\ell_1'+1}, \ldots, \mathbf{s}'_{2^k}$ by choosing the $\mathbf{s}'_j$ from

$\mathsf{Z}_q^n \backslash \tilde{S}$ uniformly random without replacement. This implies that for each index $j \in \{\ell_1' + 1, \ldots, 2^k\}$, the marginal distribution of $\mathbf{s}_j'$ is uniform over $\mathsf{Z}_q^n \backslash \tilde{S}$.

But we can sample $\mathbf{s}$ uniformly from $\mathsf{Z}_q^n \backslash \tilde{S}$ by sampling $\mathbf{s}$ uniformly from Z_q^n and rejecting if $\mathbf{s} \in \tilde{S}$. Thus, a $\mathbf{s}'$ sampled uniformly from $\mathsf{Z}_q^n \backslash \tilde{S}$ has statistical distance at most $|\tilde{S}|/q^n \leq \ell_1/q^n$ from a uniformly random $\mathbf{s} \leftarrow_{\$} \mathsf{Z}_q^n$. Consequently by Lemma 7.5, it holds for each index j that

$$\Pr_S[\|\mathbf{s}_j'\mathbf{A} - \mathbf{y}\| \leq B] \leq ((2B+1)/q)^n + \ell_1/q^n.$$

A union-bound over all indices $j \in \{\ell_1', \ldots, 2^k\}$ yields that

$$\Pr_S[\exists j \in \{\ell_1', \ldots, 2^k\} : \|\mathbf{s}_j'\mathbf{A} - \mathbf{y}\| \leq B] \leq (2^k - \ell_1) \cdot (((2B+1)/q)^n + \ell_1/q^n))$$

We can conclude that the statistical distance between $\mathcal{H}_0$ and $\mathcal{H}_{\ell_2}$ is at most

$$\begin{aligned}
\delta &= \ell_2 \cdot (2^k - \ell_1) \cdot (((2B+1)/q)^n + \ell_1/q^n)) \\
&\leq 2\ell_2 \cdot 2^k \cdot ((2B+1)/q)^n \\
&\leq 2\ell_2 \cdot 2^{-\omega(\log(\lambda))} \\
&= \mathsf{negl}(\lambda).
\end{aligned}$$

Here the first inequality follows from $\ell_1 \leq (2B+1)^n$ (as ℓ_1 is polynomial) and the second inequality follows from our assumption that $k < n \cdot \log(q/(2B+1)) - \omega(\log(\lambda))$. This concludes the proof. □

References

[ACPS09] Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In Shai Halevi, editor, *Advances in Cryptology { CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 595–618, Santa Barbara, CA, USA, August 16–20, 2009. Springer, Heidelberg, Germany.

[AD97] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *29th Annual ACM Symposium on Theory of Computing*, pages 284–293, El Paso, TX, USA, May 4–6, 1997. ACM Press.

[Ajt96] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *28th Annual ACM Symposium on Theory of Computing*, pages 99–108, Philadephia, PA, USA, May 22–24, 1996. ACM Press.

[AKPW13] Joël Alwen, Stephan Krenn, Krzysztof Pietrzak, and Daniel Wichs. Learning with rounding, revisited - new reduction, properties and applications. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology { CRYPTO 2013, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 57–74, Santa Barbara, CA, USA, August 18–22, 2013. Springer, Heidelberg, Germany.

[BBPS19] Madalina Bolboceanu, Zvika Brakerski, Renen Perlman, and Devika Sharma. Order-lwe and the hardness of ring-lwe with entropic secrets. Asiacrypt, 2019. https://eprint.iacr.org/2018/494.

[BGM+16] Andrej Bogdanov, Siyao Guo, Daniel Masny, Silas Richelson, and Alon Rosen. On the hardness of learning with rounding over small modulus. In Eyal Kushilevitz and Tal Malkin, editors, *TCC 2016-A: 13th Theory of Cryptography Conference, Part I*, volume 9562 of *Lecture Notes in Computer Science*, pages 209–224, Tel Aviv, Israel, January 10–13, 2016. Springer, Heidelberg, Germany.

[BGV12] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In Shafi Goldwasser, editor, *ITCS 2012: 3rd Innovations in Theoretical Computer Science*, pages 309–325, Cambridge, MA, USA, January 8–10, 2012. Association for Computing Machinery.

[BLP+13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th Annual ACM Symposium on Theory of Computing*, pages 575–584, Palo Alto, CA, USA, June 1–4, 2013. ACM Press.

[BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology { EUROCRYPT 2012*, volume 7237 of *Lecture Notes in Computer Science*, pages 719–737, Cambridge, UK, April 15–19, 2012. Springer, Heidelberg, Germany.

[BV11] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. In Rafail Ostrovsky, editor, *52nd Annual Symposium on Foundations of Computer Science*, pages 97–106, Palm Springs, CA, USA, October 22–25, 2011. IEEE Computer Society Press.

[DM13] Nico Döttling and Jörn Müller-Quade. Lossy codes and a new variant of the learning-with-errors problem. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology { EUROCRYPT 2013*, volume 7881 of *Lecture Notes in Computer Science*, pages 18–34, Athens, Greece, May 26–30, 2013. Springer, Heidelberg, Germany.

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

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In Michael Mitzenmacher, editor, *41st Annual ACM Symposium on Theory of Computing*, pages 169–178, Bethesda, MD, USA, May 31 – June 2, 2009. ACM Press.

[GKPV10] Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the learning with errors assumption. In Andrew Chi-Chih Yao, editor, *ICS 2010: 1st Innovations in Computer Science*, pages 230–240, Tsinghua University, Beijing, China, January 5–7, 2010. Tsinghua University Press.

[GVW13] Sergey Gorbunov, Vinod Vaikuntanathan, and Hoeteck Wee. Attribute-based encryption for circuits. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th Annual ACM Symposium on Theory of Computing*, pages 545–554, Palo Alto, CA, USA, June 1–4, 2013. ACM Press.

[HH09] Iftach Haitner and Thomas Holenstein. On the (im)possibility of key dependent encryption. In Omer Reingold, editor, *TCC 2009: 6th Theory of Cryptography Conference*, volume 5444 of *Lecture Notes in Computer Science*, pages 202–219. Springer, Heidelberg, Germany, March 15–17, 2009.

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

[MM11] Daniele Micciancio and Petros Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In Phillip Rogaway, editor, *Advances in Cryptology { CRYPTO 2011*, volume 6841 of *Lecture Notes in Computer Science*, pages 465–484, Santa Barbara, CA, USA, August 14–18, 2011. Springer, Heidelberg, Germany.

[MP12] Daniele Micciancio and Chris Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology { EUROCRYPT 2012*, volume 7237 of *Lecture Notes in Computer Science*, pages 700–718, Cambridge, UK, April 15–19, 2012. Springer, Heidelberg, Germany.

[MR04] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th Annual Symposium on Foundations of Computer Science*, pages 372–381, Rome, Italy, October 17–19, 2004. IEEE Computer Society Press.

[Nao03] Moni Naor. On cryptographic assumptions and challenges (invited talk). In Dan Boneh, editor, *Advances in Cryptology { CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 96–109, Santa Barbara, CA, USA, August 17–21, 2003. Springer, Heidelberg, Germany.

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

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

[Pei10] Chris Peikert. An efficient and parallel Gaussian sampler for lattices. In Tal Rabin, editor, *Advances in Cryptology { CRYPTO 2010*, volume 6223 of *Lecture Notes in Computer Science*, pages 80–97, Santa Barbara, CA, USA, August 15–19, 2010. Springer, Heidelberg, Germany.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th Annual ACM Symposium on Theory of Computing*, pages 84–93, Baltimore, MA, USA, May 22–24, 2005. ACM Press.

[Sha48] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[Sha49] Claude Elwood Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.

[W+59] J Wolfowitz et al. Strong converse of the coding theorem for semicontinuous channels. *Illinois Journal of Mathematics*, 3(4):477–489, 1959.

[Wic13] Daniel Wichs. Barriers in cryptography with weak, correlated and leaky sources. In Robert D. Kleinberg, editor, *ITCS 2013: 4th Innovations in Theoretical Computer Science*, pages 111–126, Berkeley, CA, USA, January 9–12, 2013. Association for Computing Machinery.