On the Hardness of Learning with Rounding over Small Modulus *

Andrej Bogdanov[1], Siyao Guo[1], Daniel Masny[2], Silas Richelson[3], and Alon Rosen[4]

[1] Dept. of Computer Science and Engineering, Chinese Univ. of Hong Kong
{andrejb,syguo}@cse.cuhk.edu.hk
[2] Horst-Görtz Inst. for IT Security and Faculty of Mathematics,
Ruhr-Universität Bochum
daniel.masny@ruhr-uni-bochum.de
[3] Dept. of Electrical Engineering and Computer Science, MIT
silas.richelson@gmail.com
[4] Efi Arazi School of Computer Science, IDC Herzliya
alon.rosen@idc.ac.il

Abstract. We show the following reductions from the learning with errors problem (LWE) to the learning with rounding problem (LWR): (1) Learning the secret and (2) distinguishing samples from random strings is at least as hard for LWR as it is for LWE for efficient algorithms if the number of samples is no larger than $O(q/Bp)$, where q is the LWR modulus, p is the rounding modulus, and the noise is sampled from any distribution supported over the set $\{-B, \ldots, B\}$.
Our second result generalizes a theorem of Alwen, Krenn, Pietrzak, and Wichs (CRYPTO 2013) and provides an alternate proof of it. Unlike Alwen et al., we do not impose any number theoretic restrictions on the modulus q. The first result also extends to variants of LWR and LWE over polynomial rings. The above reductions are sample preserving and run in time $\mathrm{poly}(n, q, m)$.
As additional results we show that (3) distinguishing any number of LWR samples from random strings is at least as hard as LWE whose noise distribution is uniform over the integers in the range $[-q/2p, \ldots, q/2p)$ provided q is a multiple of p and (4) the "noise flooding" technique for converting faulty LWE noise to a discrete Gaussian distribution can be applied whenever $q = \Omega(B\sqrt{m})$.

* Part of this work done while authors were visiting IDC Herzliya, supported by the European Research Council under the European Union's Seventh Framework Programme (FP 2007-2013), ERC Grant Agreement n. 307952. The first and second author were supported in part by RGC GRF grants CUHK410112 and CUHK410113. The third author was supported by DFG Research Training Group GRK 1817/1. The fifth author was supported by ISF grant no.1255/12 and by the ERC under the EU's Seventh Framework Programme (FP/2007-2013) ERC Grant Agreement n. 307952. Work in part done while the author was visiting the Simons Institute for the Theory of Computing, supported by the Simons Foundation and by the DIMACS/Simons Collaboration in Cryptography through NSF grant #CNS-1523467.

1 Introduction

1.1 Learning with Rounding

The learning with rounding (LWR) problem, introduced by Banerjee, Peikert, and Rosen [BPR12], concerns the cryptographic properties of the function $f_{\mathbf{s}} \colon \mathbb{Z}_q^n \to \mathbb{Z}_p$ given by

$$f_{\mathbf{s}}(\mathbf{x}) = \lfloor \langle \mathbf{x}, \mathbf{s} \rangle \rceil_p = \lfloor (p/q) \cdot \langle \mathbf{x}, \mathbf{s} \rangle \rceil$$

where $\mathbf{s} \in \mathbb{Z}_q^n$ is a secret key, $\langle \mathbf{x}, \mathbf{s} \rangle$ is the inner product of $\mathbf{x}$ and $\mathbf{s}$ mod q, and $\lfloor \cdot \rceil$ denotes the closest integer. In this work we are interested in the algorithmic hardness of the tasks of learning the secret $\mathbf{s}$ and of distinguishing $f_{\mathbf{s}}$ from a random function given uniform and independent samples of the form $(\mathbf{x}, f_{\mathbf{s}}(\mathbf{x}))$.

Learning with rounding was proposed as a deterministic variant of the learning with errors (LWE) problem [Reg05]. In this problem $f_{\mathbf{s}}$ is replaced by the randomized function $g_{\mathbf{s}} \colon \mathbb{Z}_q^n \to \mathbb{Z}_q$ given by $g_{\mathbf{s}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{s} \rangle + e$, where e is sampled from some error distribution over $\mathbb{Z}_q$ independently for every input $\mathbf{x} \in \mathbb{Z}_q^n$.

In spite of the superficial similarities between the two problems, the cryptographic hardness of LWE is much better understood. Extending works of Regev [Reg05], Peikert [Pei09], and others, Brakerski et al. [BLP+13] gave a polynomial-time reduction from finding an approximate shortest vector in an arbitrary lattice to the task of distinguishing $g_{\mathbf{s}}$ from a random function given access to uniform and independent samples $(\mathbf{x}, g_{\mathbf{s}}(\mathbf{x}))$ when e is drawn from the discrete Gaussian distribution of sufficiently large standard deviation. Their reduction is versatile in two important aspects. First, it is meaningful for any modulus q that exceeds the standard deviation of the noise. Second, it does not assume a bound on the number of samples given to the distinguisher.

In contrast, the hardness of the learning with rounding problem has only been established for restricted settings of the parameters. In their work Banerjee, Peikert, and Rosen show that if $f_{\mathbf{s}}$ can be efficiently distinguished from a random function given m random samples with advantage δ, then so can $g_{\mathbf{s}}$ with advantage $\delta - O(mBp/q)$, where the noise e is supported on the range of integers $\{-B, \ldots, B\}$ modulo q. From here one can conclude the hardness of distinguishing $f_{\mathbf{s}}$ from a random function given m random samples assuming the hardness of learning with errors, but only when the modulus q is of an exponential order of magnitude in the security parameter.

Alwen et al. [AKPW13] give a reduction from LWE to the same problem assuming that $q_{\max}$ is at least as large as $2nmBp$ and $q_{\max}^2$ does not divide q, where $q_{\max}$ is the largest prime divisor of q. This reduction can be meaningful even for values of q that are polynomially related to the security parameter. For example, when q is a prime number then the improvement over the reduction of Banerjee, Peikert, and Rosen is substantial.

However, the result of Alwen et al. does not apply to all (sufficiently large) values of the modulus q. For example it does not cover values of q that are powers of two. In this case the rounding function is particularly natural as it outputs the first $\log p$ significant bits of q in binary representation. Moreover, rounding with a small prime q necessarily introduces noticeable bias, consequently requiring some form of deterministic extraction. Finally, the work of Alwen et al. does not include a treatment of the significantly more efficient ring variant of LWR.

1.2 Our results

We establish the cryptographic hardness of the function $f_{\mathbf{s}}$ in the following three settings:

One-wayness: In Theorem 1 in Section 2 we show that any algorithm that recovers the secret $\mathbf{s}$ from m independent random samples of the form $(\mathbf{x}, f_{\mathbf{s}}(\mathbf{x}))$ with probability at least ε also recovers the secret $\mathbf{s}$ from m independent random samples of the form $(\mathbf{x}, \lfloor g_{\mathbf{s}}(\mathbf{x}) \rceil_p)$ with probability at least $\varepsilon^2/(1+2Bp/q)^m$. Therefore, if the function $G(\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{s}) = (\mathbf{x}_1, \ldots, \mathbf{x}_m, g_{\mathbf{s}}(\mathbf{x}_1), \ldots, g_{\mathbf{s}}(\mathbf{x}_m))$ is one-way under some B-bounded distribution (*i.e.* if the search version of LWE is hard) then we conclude that

$$F(\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{s}) = (\mathbf{x}_1, \ldots, \mathbf{x}_m, f_{\mathbf{s}}(\mathbf{x}_1), \ldots, f_{\mathbf{s}}(\mathbf{x}_m))$$

is also one-way, as long as $q \geq 2mBp$.

In Theorem 2 in Section 2.2 we show that the ring variants of the LWE and LWR problems (defined in that section) are related in an analogous manner.

Pseudorandomness: In Theorem 3 in Section 3 we show that if there exists an efficient distinguisher that tells apart m independent random samples $(\mathbf{x}, g_{\mathbf{s}}(\mathbf{x}))$ from m independent random samples of the form $(\mathbf{x}, \lfloor u \rceil_p)$, then LWE secrets can be learned efficiently assuming $q \geq 2mBp$.
In particular, when p divides q, the above function F is a pseudorandom generator assuming the hardness of learning with errors.
Theorem 3 improves upon several aspects of the work of Alwen et al.: First, we do not impose any number-theoretic restrictions on q; second, they require the stronger condition $q \geq 2nmBp$; third, unlike theirs, our reduction is sample preserving; and fourth, we believe our proof is considerably simpler. On the other hand, the complexity of their reduction has a better dependence on the modulus q and the distinguishing probability.

Hardness of learning from samples with uniform noise: In Theorem 5 in Section 4 we give an efficient reduction that takes as input independent random samples of the form $(\mathbf{x}, g_{\mathbf{s}}(\mathbf{x}))$ and produces independent random samples of the form $(\mathbf{x}, f_{\mathbf{s}}(\mathbf{x}))$ provided that p divides q and the noise e of $g_{\mathbf{s}}$ is uniformly distributed over the integers in the range $[-q/2p, \ldots, q/2p)$. Therefore if $f_{\mathbf{s}}$ can be distinguished efficiently from a random function for any number of independent random samples, so can $g_{\mathbf{s}}$. The learning with errors problem under this noise distribution is not known to be as hard as the learning with errors problem with discrete Gaussian noise when the number of samples is unbounded in terms of q and n. The existence of a reduction to the case of discrete Gaussian noise is an interesting open problem.

Noise flooding: In addition, our technique allows for an improved analysis of noise flooding. The noise flooding technique is ubiquitous in the LWE cryptographic literature. Roughly speaking, it is used to rerandomize a faulty sample $\big(\mathbf{x}, \langle \mathbf{x}, \mathbf{s} \rangle + e_{\mathsf{bad}}\big)$ into one of the form $\big(\mathbf{x}, \langle \mathbf{x}, \mathbf{s} \rangle + e_{\mathsf{good}}\big)$ where e_{good} is distributed according to the error distribution implicit in $g_{\mathbf{s}}(\cdot)$, while e_{bad} is not. Most of the time, the desired error distribution is a discrete Gaussian over $\mathbb{Z}_q$ whereas e_{bad} is some arbitrary B-bounded element in $\mathbb{Z}_q$. The most common method is to draw a fresh Gaussian error e and set $e_{\mathsf{good}} = e_{\mathsf{bad}} + e$ which results in the distribution of e_{good} being within statistical distance B/σ of the desired Gaussian. However, this requires choosing parameters in order to ensure that $B/\sigma \geq B/q$ is small. In particular, it requires setting q to be larger than any polynomial in the security parameter. Even worse, often the bound B is polynomially related to the standard deviation σ' of another discrete Gaussian used in the construction. This means that q/σ' also grows faster than any polynomial in the security parameter, which is not ideal as the quantity q/σ' corresponds to the strength of assumption one is making on the hardness of the underlying lattice problem. In Section 5 we use techniques from Section 2 to give a simple proof that noise flooding can be used whenever $q = \Omega\big(B\sqrt{m}\big)$. In particular, it can be used even when q is polynomial in the security parameter.

Conventions We write $x \leftarrow X$ for a uniform sample from the set X, $R(\mathbf{x})$ for the function $(R(x_1), \ldots, R(x_n))$, and $\mathbb{Z}_q^{n*}$ for the set of vectors in $\mathbb{Z}_q^n$ which are not zero-divisors. Namely, $\mathbb{Z}_q^{n*} = \{\mathbf{x} \in \mathbb{Z}_q^n \colon \gcd(x_1, \ldots, x_n, q) = 1\}$. All algorithms are assumed to be randomized.

2 One-wayness of LWR

In this section we prove the following theorem. We say a distribution over $\mathbb{Z}_q$ is *B-bounded* if it is supported over the interval of integers $\{-B, \ldots, B\}$, where $B \leq (q-1)/2$. We say a B-bounded distribution e is *balanced* if $\Pr[e \leq 0] \geq 1/2$ and $\Pr[e \geq 0] \geq 1/2$.

Theorem 1. *Let p, q, n, m, and B be integers such that $q > 2pB$. For every algorithm* Learn*,*

$$\Pr_{\mathbf{A},\mathbf{s},\mathbf{e}}[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{A}\mathbf{s} + \mathbf{e} \rceil_p) = \mathbf{s}] \geq \frac{\Pr_{\mathbf{A},\mathbf{s}}[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{A}\mathbf{s} \rceil_p) = \mathbf{s}]^2}{(1 + 2pB/q)^m},$$

where $\mathbf{A} \leftarrow \mathbb{Z}_q^{m \times n}$, the noise $\mathbf{e}$ is independent over all m coordinates, B-bounded and balanced in each coordinate, and $\mathbf{s}$ is chosen from any distribution supported on $\mathbb{Z}_q^{n*}$.

The assumptions made on the secret and error distribution in Theorem 1 are extremely mild. The condition $\mathbf{s} \in \mathbb{Z}_q^{n*}$ is satisfied for at least a $1 - O(1/2^n)$ fraction of secrets $s \leftarrow \mathbb{Z}_q^n$. While a B-bounded error

distribution may not be balanced, it can always be converted to a $2B$-bounded and balanced error distribution by a suitable constant shift. The discrete Gaussian distribution of standard deviation σ is $e^{-\Omega(t^2)}$-statistically close to being $t\sigma$-bounded and balanced for every $t \geq 1$.

Theorem 2 in Section 2.2 concerns the ring variants of the LWR and LWE problems and will be proved in an analogous manner.

We now outline the proof of Theorem 1. Let $\mathsf{X}_{\mathbf{s}}$ denote the distribution of a single LWR sample $\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p$ where $\mathbf{a} \leftarrow \mathbb{Z}_q^n$ and $\mathsf{Y}_{\mathbf{s}}$ denote the distribution of a single rounded LWE sample $\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rceil_p$. To prove Theorem 1 we will fix $\mathbf{s}$ and look at the ratio of probabilities of any possible instance under the product distributions $\mathsf{X}_{\mathbf{s}}^m$ and $\mathsf{Y}_{\mathbf{s}}^m$, respectively. If this ratio was always bounded by a sufficiently small quantity K,[5] then it would follow that the success probability of any search algorithm for LWR does not deteriorate by more than a factor of $1/K$ when it is run on rounded LWE instances instead.

While it happens that there are exceptional instances for which the ratio of probabilities under $\mathsf{X}_{\mathbf{s}}^m$ and $\mathsf{Y}_{\mathbf{s}}^m$ can be large, our proof of Theorem 1 will show that such instances cannot occur too often under the rounded LWE distribution and therefore does not significantly affect the success probability of the inversion algorithm. This can be showed by a standard probabilistic analysis, but we opt instead to work with a measure of distributions that is particularly well suited for bounding ratios of probabilities: the Rényi divergence.

The role of Rényi divergence in our analysis accounts for our quantitative improvement over the result of Banerjee, Peikert, and Rosen, who used the measure of statistical distance in its place. Rényi divergence has been used in a related context: Bai, Langlois, Lepoint, Stehlé and Steinfeld [BLL+15] use it to obtain tighter bounds for several lattice-based primitives.

2.1 Proof of Theorem 1

Given two distributions X and Y over Ω, the power of their Rényi divergence[6] is $\mathrm{RD}_2(\mathsf{X} \| \mathsf{Y}) = \mathrm{E}_{a \leftarrow \mathsf{X}}[\Pr[\mathsf{X} = a] / \Pr[\mathsf{Y} = a]]$.

Lemma 1. *Let $\mathsf{X}_{\mathbf{s}}$ be the distribution of a single LWR sample and let $\mathsf{Y}_{\mathbf{s}}$ be that of a single rounded LWE sample. Assume $B < q/2p$. For every $\mathbf{s} \in \mathbb{Z}_q^{n*}$ and every noise distribution that is B-bounded and balanced, $\mathrm{RD}_2\big(\mathsf{X}_{\mathbf{s}} \| \mathsf{Y}_{\mathbf{s}}\big) \leq 1 + 2Bp/q$.*

Proof. By the definition of Rényi divergence,

$$\mathrm{RD}_2\big(\mathsf{X}_{\mathbf{s}} \| \mathsf{Y}_{\mathbf{s}}\big) = \mathrm{E}_{\mathbf{a} \leftarrow \mathbb{Z}_q^n} \frac{\Pr\big[\mathsf{X}_{\mathbf{s}} = (\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p)\big]}{\Pr\big[\mathsf{Y}_{\mathbf{s}} = (\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p)\big]} = \mathrm{E}_{\mathbf{a} \leftarrow \mathbb{Z}_q^n} \frac{1}{\Pr_e\big[\lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rceil_p = \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p\big]}.$$

Let $\mathsf{BAD}_{\mathbf{s}}$ be the set $\big\{\mathbf{a} \in \mathbb{Z}_q^n : \big| \langle \mathbf{a}, \mathbf{s} \rangle - \frac{q}{p} \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p \big| < B\big\}$. These are the $\mathbf{a}$ for which $\langle \mathbf{a}, \mathbf{s} \rangle$ is dangerously close to the rounding boundary. When $\mathbf{a} \notin \mathsf{BAD}_{\mathbf{s}}$, $\Pr_e\big[\lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rceil_p = \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p\big] = 1$. Since $\gcd(s_1, \ldots, s_n, q) = 1$, the inner product $\langle \mathbf{a}, \mathbf{s} \rangle$ is uniformly distributed over $\mathbb{Z}_q$, so $\Pr[\mathbf{a} \in \mathsf{BAD}_{\mathbf{s}}] \leq (2B-1)p/q$. When $\mathbf{a} \in \mathsf{BAD}_{\mathbf{s}}$, the event $\lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rceil_p = \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p$ still holds at least in one of the two cases $e \leq$ or $e \geq 0$. By our assumptions on the noise distribution, $\Pr_e\big[\lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rceil_p = \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p\big] \geq 1/2$. Conditioning over the event $\mathbf{a} \in \mathsf{BAD}_{\mathbf{s}}$, we conclude that

$$\mathrm{RD}_2(\mathsf{X}_{\mathbf{s}} \| \mathsf{Y}_{\mathbf{s}}) \leq 1 \cdot \Pr[\mathbf{a} \notin \mathsf{BAD}_{\mathbf{s}}] + 2 \cdot \Pr[\mathbf{a} \in \mathsf{BAD}_{\mathbf{s}}] \leq 1 + \frac{2Bp}{q}.$$

□

To complete the proof of Theorem 1 we need two elementary properties of Rényi divergence.

Claim. For any two distributions X and Y, (1) $\mathrm{RD}_2(\mathsf{X}^m \| \mathsf{Y}^m) = \mathrm{RD}_2(\mathsf{X} \| \mathsf{Y})^m$ and (2) for any event E, $\Pr[\mathsf{Y} \in E] \geq \Pr[\mathsf{X} \in E]^2 / \mathrm{RD}_2(\mathsf{X} \| \mathsf{Y})$.

[5] Levin [Lev86] calls this condition K-domination.

[6] Rényi divergences [vEH14] are a class of measures parametrized by a real number $\alpha > 1$. The definition we give specializes α to 2, which is sufficient for our analysis.

Proof. Property (1) follows immediately from independence of the m samples. Property (2) is the Cauchy-Schwarz inequality applied to the functions

$$f(a) = \frac{\Pr[\mathsf{X} = a]}{\sqrt{\Pr[\mathsf{Y} = a]}}; \text{ and } g(a) = \sqrt{\Pr[\mathsf{Y} = a]}.$$

□

Proof (Proof of Theorem 1). Fix $\mathbf{s}$ such that $\gcd(\mathbf{s}, q) = 1$ and the randomness of Learn. By Lemma 1 and part (1) of Claim 2.1, $\mathrm{RD}_2(\mathsf{X}_{\mathbf{s}}^m \| \mathsf{Y}_{\mathbf{s}}^m) \le (1 + 2Bp/q)^m$. Letting E be the event $\{(\mathbf{A}, \mathbf{y}) \colon \mathsf{Learn}(\mathbf{A}, \mathbf{y}) = \mathbf{s}\}$, by part (2) of Claim 2.1,

$$\Pr_{\mathbf{A},\mathbf{e}}[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{As} + \mathbf{e} \rceil_p) = \mathbf{s}] \ge \frac{\Pr_{\mathbf{A}}[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{As} \rceil_p) = \mathbf{s}]^2}{(1 + 2pB/q)^m}.$$

To obtain the theorem, we average over $\mathbf{s}$ and and the randomness of Learn and apply the Cauchy-Schwarz inequality. □

2.2 Hardness over Rings

For many applications it is more attractive to use a ring version of LWR (RLWR). Banerjee, Peikert, and Rosen [BPR12] introduced it together with LWR. It brings the advantage of reducing the entropy of $\mathbf{A}$ for same sized $\lfloor \mathbf{As} + \mathbf{e} \rceil_p$. In the following theorem, we give a variant of Theorem 1 for the RLWR based on the hardness of ring LWE. This theorem is not needed for the remaining sections of the paper.

Theorem 2. *Let p, q, n, k, B be integers such that $q > 2pB$. Let R_q be the ring $\mathbb{Z}_q[x]/g(x)$ where g is a polynomial of degree n over $\mathbb{Z}_q$ and f be an arbitrary function over R_q. For every algorithm Learn,*

$$\Pr_{\mathbf{a},s,\mathbf{e}}[\mathsf{Learn}(\mathbf{a}, \lfloor \mathbf{a}s + \mathbf{e} \rceil_p) = f(s)] \ge \frac{\Pr_{\mathbf{a},s}[\mathsf{Learn}(\mathbf{a}, \lfloor \mathbf{a}s \rceil_p) = f(s)]^2}{(1 + 2pB/q)^{nk}},$$

where $\mathbf{a} \leftarrow R_q^k$, the noise $\mathbf{e}$ is independent over all k coordinates, B-bounded and balanced in each coordinate, and s is chosen from any distribution supported on the set of all units in R_q.

An element in $R_q = \mathbb{Z}_q[x]/g(x)$ can be represented as a polynomial (in x) of degree less than n with coefficients in $\mathbb{Z}_q$. Here, for $a \in R_q$, $\lfloor a \rceil_p$ is an element in $\mathbb{Z}_p[x]/g(x)$ obtained by applying the function $\lfloor \cdot \rceil_p$ to each of coefficient of a separately. A distribution over ring R_q is B-bounded and balanced if every coefficient is drawn independently from a B-bounded and balanced distribution over $\mathbb{Z}_q$.

The bound in Theorem 2 matches the bound in Theorem 1 since k can be chosen such that nk is on the order of m. Theorem 2 follows from Claim 2.1 and the following variant of Lemma 1.

Lemma 2. *Assume $B < q/2p$. For every unit $s \in R_q$ and noise distribution χ that is B-bounded and balanced over R_q, $\mathrm{RD}_2\big(\mathsf{X}_s \| \mathsf{Y}_s\big) \le \big(1 + 2pB/q\big)^n$ where X_s is the random variable $\big(a, \lfloor a \cdot s \rceil_p\big)$ and Y_s is the random variable $\big(a, \lfloor a \cdot s \rceil_p + e\big)$ with $a \leftarrow R_q$ and $e \leftarrow \chi$.*

Since the proof is very similar to the proof of Lemma 1, we defer it to Appendix A.

3 Pseudorandomness of LWR

In this section we prove the following Theorem. We will implicitly assume that algorithms have access to the prime factorization of the modulus q throughout this section.

Theorem 3. *For every $\varepsilon > 0$, n, m, $q > 2pB$, and algorithm Dist such that*

$$\big|\Pr_{\mathbf{A},\mathbf{s}}\big[\mathsf{Dist}\big(\mathbf{A}, \lfloor \mathbf{As} \rceil_p\big) = 1\big] - \Pr_{\mathbf{A},\mathbf{u}}\big[\mathsf{Dist}\big(\mathbf{A}, \lfloor \mathbf{u} \rceil_p\big) = 1\big]\big| \ge \varepsilon, \tag{1}$$

where $\mathbf{A} \leftarrow \mathbb{Z}_q^{m\times n}$, $\mathbf{s} \leftarrow \{0,1\}^n$ *and* $\mathbf{u} \leftarrow \mathbb{Z}_q^m$ *there exists an algorithm* Learn *that runs in time polynomial in* n, m, *the number of divisors of* q, *and the running time of* Dist *such that*

$$\Pr_{\mathbf{A},\mathbf{s}}\big[\mathsf{Learn}\big(\mathbf{A}, \mathbf{A}\mathbf{s}+\mathbf{e}\big) = \mathbf{s}\big] \geq \Big(\frac{\varepsilon}{4qm} - \frac{2^n}{p^m}\Big)^2 \cdot \frac{1}{(1+2Bp/q)^m} \tag{2}$$

for any noise distribution $\mathbf{e}$ *that is* B*-bounded and balanced in each coordinate.*

One unusual aspect of this theorem is that the secret is a uniformly distributed *binary* string in $\mathbb{Z}_q^n$. This assumption can be made essentially without loss of generality: Brakerski et al. [BLP+13] show that under discrete Gaussian noise, learning a binary secret in $\{0,1\}^n$ from LWE samples is as hard as learning a secret uniformly sampled from $\mathbb{Z}_q^{\Omega(n/\log q)}$. The assumption (1) can also be stated with $\mathbf{s}$ sampled uniformly from $\mathbb{Z}_q^n$: In Section 3.4 we show that distinguishing LWR samples from random ones is no easier for uniformly distributed secrets than it is for any other distribution on secrets, including the uniform distribution over binary secrets. (When q is prime, the proof of Theorem 3 can be carried out for $\mathbf{s}$ uniformly distributed over $\mathbb{Z}_q^n$ so these additional steps are not needed.)

To prove Theorem 3 we follow a sequence of standard steps originating from Yao [Yao82], Goldreich and Levin [GL89]: In Lemma 3 we convert the distinguisher Dist into a predictor that given a sequence of LWR samples and a label $\mathbf{a}$ guesses the inner product $\langle \mathbf{a}, \mathbf{s}\rangle$ in $\mathbb{Z}_q$ with significant advantage. In Lemma 4 we show how to use this predictor to efficiently learn the entries of the vector $\mathbf{s}$ modulo q' for some divisor $q' > 1$ of q. If the entries of the secret $\mathbf{s}$ are bits, $\mathbf{s}$ is then fully recovered given LWR samples. By Theorem 1 the learner's advantage does not deteriorate significantly when the LWR samples are replaced by LWE samples.

Our proof resembles the work of Micciancio and Mol [MM11] who give, to the best of our knowledge, the only sample preserving search-to-decision reduction for LWE (including its variants). Unlike our theorem, theirs imposes certain number-theoretic restrictions on q. Also, while Micciancio and Mol work with a problem that is "dual" to LWE, we work directly with LWR samples.

3.1 Predicting the Inner Product

Lemma 3. *For all* ε *(possibly negative),* n, m, q, *every polynomial-time function* R *over* $\mathbb{Z}_q$, *and every algorithm* Dist *such that*

$$\Pr_{\mathbf{A},\mathbf{s}}\big[\mathsf{Dist}\big(\mathbf{A}, R(\mathbf{A}\mathbf{s})\big) = 1\big] - \Pr_{\mathbf{A},\mathbf{u}}\big[\mathsf{Dist}\big(\mathbf{A}, R(\mathbf{u})\big) = 1\big] = \varepsilon,$$

there exists an algorithm Pred *whose running time is polynomial in its input size and the running time of* Dist *such that*

$$\Pr_{\mathbf{A},\mathbf{s},\mathbf{a}}\big[\mathsf{Pred}\big(\mathbf{A}, R(\mathbf{A}\mathbf{s}), \mathbf{a}\big) = \langle \mathbf{a}, \mathbf{s}\rangle\big] = \frac{1}{q} + \frac{\varepsilon}{mq}.$$

where the probabilities are taken over $\mathbf{A} \leftarrow \mathbb{Z}_q^{m\times n}$, $\mathbf{u} \leftarrow \mathbb{Z}_q^m$, *the random coins of the algorithms, and secret* $\mathbf{s}$ *sampled from an arbitrary distribution.*

Here, $R(\mathbf{y})$ is the vector obtained by applying R to every coordinate of the vector $\mathbf{y}$.

Proof. Consider the following algorithm Pred. On input $(\mathbf{A}, \mathbf{b})$=$((\mathbf{a}_1, b_1), \ldots, (\mathbf{a}_m, b_m))$ $(\mathbf{a}_j \in \mathbb{Z}_q^n,\ b_j \in \mathbb{Z}_q)$ and $\mathbf{a} \in \mathbb{Z}_q^n$:

1. Sample a random index $i \leftarrow \{1, \ldots, m\}$ and a random $c \leftarrow \mathbb{Z}_q$.
2. Obtain $\mathbf{A}', \mathbf{b}'$ from $\mathbf{A}, \mathbf{b}$ by replacing $\mathbf{a}_i$ with $\mathbf{a}$, b_i with $R(c)$, and every b_j for $j > i$ with an independent element of the form $R(u_j), u_j \leftarrow \mathbb{Z}_q$.
3. If $\mathsf{Dist}(\mathbf{A}', \mathbf{b}') = 1$, output c. Otherwise, output a uniformly random element in $\mathbb{Z}_q$.

Let $\mathbf{h}_i = \big(R(\langle \mathbf{a}_1, \mathbf{s}\rangle), \ldots, R(\langle \mathbf{a}_i, \mathbf{s}\rangle), R(u_{i+1}), \ldots, R(u_m)\big) \in \mathbb{Z}_p^m$, for i ranging from 0 to m. Then $\mathbf{h}_m = R(\mathbf{A}\mathbf{s})$ and $\mathbf{h}_0 = R(\mathbf{u})$ so by the assumption on Dist it follows that

$$\mathrm{E}_i\Big[\Pr_{\mathbf{A},\mathbf{s},\mathbf{u}}\big[\mathsf{Dist}\big(\mathbf{A}, \mathbf{h}_i\big) = 1\big] - \Pr_{\mathbf{A},\mathbf{s},\mathbf{u}}\big[\mathsf{Dist}\big(\mathbf{A}, \mathbf{h}_{i-1}\big) = 1\big]\Big] = \frac{\varepsilon}{m}.$$

Conditioned on the choice of i,

$$\begin{aligned}
&\Pr\big[\mathsf{Pred}(\mathbf{A},\mathbf{b},\mathbf{a}) = \langle \mathbf{a},\mathbf{s}\rangle\big] \\
&= \Pr\big[\mathsf{Dist}(\mathbf{A}',\mathbf{b}') = 1 \text{ and } c = \langle \mathbf{a},\mathbf{s}\rangle\big] + \frac{1}{q}\cdot\Pr\big[\mathsf{Dist}(\mathbf{A}',\mathbf{b}') \neq 1\big] \\
&= \frac{1}{q}\cdot\Pr\big[\mathsf{Dist}(\mathbf{A}',\mathbf{b}') = 1 \mid c = \langle \mathbf{a},\mathbf{s}\rangle\big] + \frac{1}{q}\cdot\Pr\big[\mathsf{Dist}(\mathbf{A}',\mathbf{b}') \neq 1\big] \\
&= \frac{1}{q} + \frac{1}{q}\cdot\big(\Pr\big[\mathsf{Dist}(\mathbf{A}',\mathbf{b}') = 1 \big| c = \langle \mathbf{a},\mathbf{s}\rangle\big] - \Pr\big[\mathsf{Dist}(\mathbf{A}',\mathbf{b}') = 1\big]\big)
\end{aligned}$$

when $\mathbf{b} = R(\mathbf{As})$, the distribution $(\mathbf{A}',\mathbf{b}')$ is the same as $(\mathbf{A},\mathbf{h}_{i-1})$ while $(\mathbf{A}',\mathbf{b}')$ conditioned on $c = \langle \mathbf{a},\mathbf{s}\rangle$ is the same as $(\mathbf{A},\mathbf{h}_i)$. Averaging over i yields the desired advantage of Pred. □

3.2 Learning the Secret

Lemma 4. *There exists an oracle algorithm* List *such that for every algorithm* Pred *satisfying* $|\Pr[\mathsf{Pred}(\mathbf{a}) = \langle \mathbf{a},\mathbf{s}\rangle] - 1/q| \geq \varepsilon$, $\mathsf{List}^{\mathsf{Pred}}(\varepsilon)$ *outputs a list of entries* $(q',\mathbf{s}')$ *containing at least one such that* $q' > 1$, q' *divides* q, *and* $\mathbf{s}' = \mathbf{s} \bmod q'$ *in time polynomial in* n, $1/\varepsilon$, *and the number of divisors of* q *with probability at least* $\varepsilon/4$. *The probabilities are taken over* $\mathbf{a} \leftarrow \mathbb{Z}_q^n$, *any distribution on* $\mathbf{s}$, *and the randomness of the algorithms.*

When q is a prime number, the conclusion of the theorem implies that the list must contain the secret $\mathbf{s}$. When q is a composite, the assumption does not in general guarantee full recovery of $\mathbf{s}$. For example, the predictor $\mathsf{Pred}(a) = \langle \mathbf{a},\mathbf{s}\rangle \bmod q'$ has advantage $\varepsilon = (q'-1)/q$ but does not distinguish between pairs of secrets that are congruent modulo q'. In this case List cannot hope to learn any information on $\mathbf{s}$ beyond the value $\mathbf{s}$ modulo q'.

The proof of Lemma 4 makes use of the following result of Akavia, Goldwasser, and Safra [AGS03] on learning heavy Fourier coefficients, extending work of Kushilevitz, Mansour, and others. Recall that the Fourier coefficients of a function $h\colon \mathbb{Z}_q^n \to \mathbb{C}$ are the complex numbers $\hat{h}(\mathbf{a}) = \mathrm{E}_{\mathbf{x}\leftarrow\mathbb{Z}_q^n}[h(\mathbf{x})\omega^{-\langle \mathbf{a},\mathbf{x}\rangle}]$, where $\omega = e^{2\pi i/q}$ is a primitive $q-$th root of unity. Our functions of interest all map into the unit complex circle $\mathbb{T} = \{c \in \mathbb{C}\colon |c| = 1\}$, so we specialize the result to this setting.

Theorem 4 (Akavia et al. [AGS03]). *There is an algorithm* AGS *that given query access to a function* $h\colon \mathbb{Z}_q^n \to \mathbb{T}$ *outputs a list of size at most* $2/\varepsilon^2$ *which contains all* $\mathbf{a} \in \mathbb{Z}_q^n$ *such that* $|\hat{h}(\mathbf{a})| \geq \varepsilon$ *in time polynomial in* n, $\log q$, *and* $1/\varepsilon$ *with probability at least* $1/2$.

We will also need the following property of the Fourier transform of random variables. For completeness the proof is given below.

Claim. For every random variable Z over $\mathbb{Z}_q$ there exists a nonzero r in $\mathbb{Z}_q$ such that $|\mathrm{E}[\omega^{rZ}]| \geq |\Pr[Z = 0] - 1/q|$.

Proof (Proof of Lemma 4). We first replace Pred by the following algorithm: Sample a uniformly random unit (invertible element) u from $\mathbb{Z}_q^*$ and output $u^{-1}\mathsf{Pred}(u\mathbf{a})$. This transformation does not affect the advantage of Pred but ensures that for fixed $\mathbf{s}$ and randomness of Pred, the value $\mathrm{E}_{\mathbf{a}}[\omega^{r(\mathsf{Pred}(\mathbf{a})-\langle \mathbf{a},\mathbf{s}\rangle)}]$ is the same for all r with the same $\gcd(r,q)$.

Algorithm List works as follows: For every divisor $r < q$ of q run AGS with oracle access to the function $h_r(\mathbf{a}) = \omega^{r\cdot\mathsf{Pred}(\mathbf{a})}$ and output $(q' = q/r, \mathbf{s}'/r \bmod q')$ for every $\mathbf{s}'$ in the list produced by AGS.

We now assume Pred satisfies the assumption of the lemma and analyze List. By Claim 3.2 there exists a nonzero $r \in \mathbb{Z}_q$ such that $|\mathrm{E}[\omega^{r(\mathsf{Pred}(\mathbf{a})-\langle \mathbf{a},\mathbf{s}\rangle)}]| \geq \varepsilon$. By Markov's inequality and the convexity of the absolute value, with probability at least $\varepsilon/2$ over the choice of $\mathbf{s}$ and the randomness of Pred $|\mathrm{E}_{\mathbf{a}}[\omega^{r(\mathsf{Pred}(\mathbf{a})-\langle \mathbf{a},\mathbf{s}\rangle)}]|$ is at least $\varepsilon/2$. We fix $\mathbf{s}$ and the randomness of Pred and assume this is the case. By our discussion on Pred, the expectation of interest is the same for all r with the same $\gcd(r,q)$, so we may and will assume without loss of generality that r is a divisor of q.

Since $\mathrm{E}_{\mathbf{a}}[\omega^{r(\mathsf{Pred}(\mathbf{a})-\langle \mathbf{a},\mathbf{s}\rangle)}] = \hat{h}_r(r\mathbf{s})$, by Theorem 4, the r-th run of AGS outputs $r\mathbf{s}$ with probability at least $1/2$. Since $(r\mathbf{s})/r \bmod q' = \mathbf{s} \bmod q'$ it follows that the entry $(q', \mathbf{s} \bmod q')$ must appear in the output of List with probability at least $(1/2)(\varepsilon/2) = \varepsilon/4$. Regarding time complexity, List makes a call to AGS for every divisor of q except q, so its running time is polynomial in n and the number of divisors of q. □

Proof (Proof of Claim 3.2). Let $\varepsilon = \Pr[Z = 0] - 1/q$ and $h(a) = q(\Pr[Z = a] - \Pr[U = a])$, where $U \leftarrow \mathbb{Z}_q$ is a uniform random variable. By Parseval's identity from Fourier analysis,

$$\sum_{r \in \mathbb{Z}_q} |\hat{h}(r)|^2 = \mathrm{E}_{a \leftarrow \mathbb{Z}_q}[h(a)^2] \geq \frac{1}{q} h(0)^2 = q\varepsilon^2.$$

On the left hand side, after normalizing we obtain that $\hat{h}(r) = \mathrm{E}[\omega^{-rZ}] - \mathrm{E}[\omega^{-rU}]$. Therefore $\hat{h}(0) = 0$, so $|\hat{h}(r)|^2 = |\mathrm{E}[\omega^{-rZ}]|^2$ must be at least as large as $q\varepsilon^2/(q-1)$ for at least one nonzero value of r, giving a slightly stronger conclusion than desired. □

3.3 Proof of Theorem 3

On input $(\mathbf{A}, \mathbf{b})$, algorithm Learn runs $\mathsf{List}^{\mathsf{Pred}(\mathbf{A}, \lfloor \mathbf{b} \rceil_p, \cdot)}(\varepsilon/2qm)$ and outputs any $\mathbf{s} \in \{0,1\}^n$ appearing in the list such that $\lfloor \mathbf{As} \rceil_p = \lfloor \mathbf{b} \rceil_p$ (or the message `fail` if no such $\mathbf{s}$ exists). By Theorem 1,

$$\Pr[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{As} + \mathbf{e} \rceil_p) = \mathbf{s}] \geq \frac{\Pr[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{As} \rceil_p) = \mathbf{s}]^2}{(1 + 2Bp/q)^m}.$$

For $\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{As} \rceil_p)$ to output $\mathbf{s}$ it is sufficient that $\mathbf{s}$ appears in the output of $\mathsf{List}^{\mathsf{Pred}(\mathbf{A}, \lfloor \mathbf{As} \rceil_p, \cdot)}(\varepsilon/2qm)$ and that no other $\mathbf{s}' \in \{0,1\}^n$ satisfies $\lfloor \mathbf{As}' \rceil_p = \lfloor \mathbf{As} \rceil_p$. By Lemmas 3 and 4, the list contains $\mathbf{s} \bmod q'$ for some q' with probability at least $\varepsilon/4qm$. Since $\mathbf{s}$ is binary, $\mathbf{s} \bmod q' = \mathbf{s}$. By a union bound, the probability that some $\lfloor \mathbf{As}' \rceil_p = \lfloor \mathbf{As} \rceil_p$ for some $\mathbf{s}' \neq \mathbf{s}$ is at most $2^n p^{-m}$ and so

$$\Pr[\mathsf{Learn}(\mathbf{A}, \lfloor \mathbf{As} + \mathbf{e} \rceil_p) = \mathbf{s}] \geq \frac{(\varepsilon/4qm - 2^n p^{-m})^2}{(1 + 2Bp/q)^m}.$$

3.4 Rerandomizing the Secret

Lemma 5. *Let S be any distribution supported on $\mathbb{Z}_q^{n*}$. For every function R on $\mathbb{Z}_q$, there is a polynomial-time transformation that (1) maps the distribution $(\mathbf{A}, R(\mathbf{As}))_{\mathbf{A} \leftarrow \mathbb{Z}_q^{m \times n}, \mathbf{s} \leftarrow S}$ to $(\mathbf{A}, R(\mathbf{As}))_{\mathbf{A} \leftarrow \mathbb{Z}_q^{m \times n}, \mathbf{s} \leftarrow \mathbb{Z}_q^{n*}}$ and (2) maps the distributon $(\mathbf{A}, R(\mathbf{u}))_{\mathbf{A} \leftarrow \mathbb{Z}_q^{m \times n}, \mathbf{u} \leftarrow \mathbb{Z}_q^m}$ to itself.*

In particular, it follows that the distinguishing advantage (1) can be preserved when the secret is chosen uniformly from $\mathbb{Z}_q^{n*}$ instead of uniformly from $\{0,1\}^n - \{0^n\}$. The sets $\mathbb{Z}_q^{n*}$ and $\{0,1\}^n - \{0^n\}$ can be replaced by $\mathbb{Z}_q^n$ and $\{0,1\}^n$, respectively, if we allow for failure with probability $O(2^{-n})$.

To prove Lemma 5 we need a basic fact from algebra. We omit the easy proof.

Claim. Multiplication by an $n \times n$ invertible matrix over $\mathbb{Z}_q$ is a transitive action on $\mathbb{Z}_q^{n*}$.

Proof (Proof of Lemma 5). Choose a uniformly random invertible matrix $\mathbf{P} \in \mathbb{Z}_q^{n \times n}$ and apply the map $f(\mathbf{a}, b) = (\mathbf{Pa}, b)$ to every row. Clearly this map satisfies the second condition. For the first condition, we write $f(\mathbf{a}, R(\langle \mathbf{a}, \mathbf{s} \rangle)) = (\mathbf{Pa}, R(\langle \mathbf{a}, \mathbf{s} \rangle))$, which is identically distributed as $(\mathbf{a}, R(\langle \mathbf{a}, \mathbf{P}^{-t}\mathbf{s} \rangle))$. By Claim 3.4, for every $\mathbf{s}$ in the support of S the orbit of $\mathbf{P}^{-t}\mathbf{s}$ is $\mathbb{Z}_q^{n*}$, so by symmetry $\mathbf{P}^{-t}\mathbf{s}$ is uniformly random in $\mathbb{Z}_q^{n*}$. Therefore the first condition also holds. □

4 Reduction from LWE with Uniform Errors to LWR

When the number of LWR samples is not a priori bounded, we show that the pseudorandomness (resp. one-wayness) of LWR follows from the pseudorandomness (resp. one-wayness) of LWE with a uniform noise distribution over the range of integers $[-\frac{q}{2p}, \ldots, \frac{q}{2p})$. We use a rejection sampling based approach to reject LWE samples which are likely to be rounded to the wrong value in $\mathbb{Z}_p$. This comes at the cost of throwing away samples, and indeed the sample complexity of our reduction grows with q.

Theorem 5. *Let p and q be integers such that p divides q. Then there is a reduction R with query access to independent samples such that for every $\mathbf{s} \in \mathbb{Z}_q^{n*}$:*

- *Given query access to samples $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)$, $\mathbf{a} \leftarrow \mathbb{Z}_q^n, e \leftarrow [-\frac{q}{2p}, \ldots, \frac{q}{2p}) \subset \mathbb{Z}_q$, R outputs samples from the distribution $(\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p)$, $\mathbf{a} \leftarrow \mathbb{Z}_q^n$,*
- *Given query access to uniform samples $(\mathbf{a}, u)$, $\mathbf{a} \leftarrow \mathbb{Z}_q^n, u \leftarrow \mathbb{Z}_q$, R outputs a uniform sample $(\mathbf{a}, v)$, $\mathbf{a} \leftarrow \mathbb{Z}_q^n$, $v \leftarrow \mathbb{Z}_p$.*

In both cases, the expected running time and sample complexity of the reduction is $O(q/p)$.

Proof. We view the set $(q/p)\mathbb{Z}_p$ as a subset of $\mathbb{Z}_q$. The reduction R queries its oracle until it obtains the first sample $(\mathbf{a}, b) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ such that b is in the set $(q/p)\mathbb{Z}_p$ and outputs $(\mathbf{a}, (p/q)b) \in \mathbb{Z}_q^n \times \mathbb{Z}_p$. In both cases of interest b is uniformly distributed in $\mathbb{Z}_q$, so the expected number of query calls until success is q/p.

When the queried samples are uniformly distributed in $\mathbb{Z}_q^n \times \mathbb{Z}_q$, the output is also uniformly distributed in $\mathbb{Z}_q^n \times \mathbb{Z}_p$. For queried samples of the form $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)$, we calculate the probability mass function of the output distribution. For every possible output $(\mathbf{a}', b')$, we have

$$\begin{aligned}
\Pr\left[\mathsf{R} \text{ outputs } (\mathbf{a}', b')\right] &= \Pr\left[\mathbf{a} = \mathbf{a}' \text{ and } \langle \mathbf{a}, \mathbf{s} \rangle + e = b' \;\middle|\; \langle \mathbf{a}, \mathbf{s} \rangle + e \in (q/p)\mathbb{Z}_p\right] \\
&= \Pr_{\mathbf{a}}[\mathbf{a} = \mathbf{a}'] \cdot \frac{\Pr_e\left[\langle \mathbf{a}, \mathbf{s} \rangle + e = (q/p)b' \;\middle|\; \mathbf{a} = \mathbf{a}'\right]}{\Pr_e\left[\langle \mathbf{a}, \mathbf{s} \rangle + e \in (q/p)\mathbb{Z}_p \;\middle|\; \mathbf{a} = \mathbf{a}'\right]} \\
&= q^{-n} \cdot \begin{cases} \frac{p/q}{p/q}, \text{ if } (q/p)b' - \langle \mathbf{a}', \mathbf{s} \rangle \in \left[-\frac{q}{2p}, \ldots, \frac{q}{2p}\right) \\ 0, \quad \text{otherwise.} \end{cases} \\
&= \begin{cases} q^{-n}, \text{ if } b' = \lfloor \langle \mathbf{a}', \mathbf{s} \rangle \rceil_p \\ 0, \quad \text{otherwise.} \end{cases}
\end{aligned}$$

This is the probability mass function of the distribution $(\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rceil_p)$, as desired. □

5 Noise Flooding

In this section, let χ_σ denote the discrete Gaussian distribution on $\mathbb{Z}_q$ with standard deviation σ: $\chi_\sigma(x)$ is proportional to $\exp\left(-\pi(x/\sigma)^2\right)$. Often in applications of LWE, one is given a sample $(\mathbf{a}, b)$ with $b = \langle \mathbf{a}, \mathbf{s} \rangle + e$ for $e \leftarrow \chi_\sigma$ and by performing various arithmetic operations obtains a new pair $(\mathbf{a}', b')$ with $b' = \langle \mathbf{a}', \mathbf{s}' \rangle + e'$. Sometimes, the noise quantity e' obtained is not distributed according to a Gaussian, but is only subject to an overall bound on its absolute value. If the proof of security needs $(\mathbf{a}', b')$ to be an LWE instance, then sometimes the "noise flooding" technique is used where a fresh Gaussian $x \leftarrow \chi_{\sigma'}$ is drawn and b' is set to $\langle \mathbf{a}', \mathbf{s}' \rangle + e' + x$. As long as $e' + \chi_{\sigma'} \approx_{\mathsf{s}} \chi_{\sigma'}$ the resulting $(\mathbf{a}', b')$ is statistically close to a fresh LWE instance. This technique in some form or another appears in many places, for example [AIK11,GKPV10,DGK+10,OPW11]. Unfortunately, $e' + \chi_{\sigma'} \approx_{\mathsf{s}} \chi_{\sigma'}$ requires q to be large and so the applications also carry this requirement. In this section we bound the continuous analogue of Rényi divergence between $e' + \chi_{\sigma'}$ and $\chi_{\sigma'}$ and show that the noise flooding technique can be used even when q is polynomial in the security parameter, as long as the number of samples is also bounded.

We remark that our main result in this section, Corollary 1, follows from general results in prior work which bound the Rényi divergence between Gaussians. For example, Lemma 4.2 of [LSS14] implies Corollary 1 below. However, we are unaware of a theorem in the literature with a simple statement which subsumes Corollary 1. We include a proof for completeness.

Claim. Let Ψ_α be the continuous Gaussian on $\mathbb{R}$ with standard deviation α: $\Psi_\alpha(x) = \alpha^{-1} e^{-\pi(x/\alpha)^2}$. Then for any $\beta \in \mathbb{R}$,

$$\mathrm{RD}_2(\beta + \Psi_\alpha \| \Psi_\alpha) = e^{2\pi(\beta/\alpha)^2}.$$

Proof. We have

$$\begin{aligned}\mathrm{RD}_2(\beta + \Psi_\alpha \| \Psi_\alpha) &= \alpha^{-1} \int_{-\infty}^{\infty} e^{-\left(\pi/\alpha^2\right)\left[2(x-\beta)^2 - x^2\right)} dx \\ &= \alpha^{-1} \cdot e^{2\pi\left(\beta/\alpha\right)^2} \int_{-\infty}^{\infty} e^{-\left(\pi/\alpha^2\right)\left[(x-2\beta)^2\right]} dx \\ &= e^{2\pi\left(\beta/\alpha\right)^2}.\end{aligned}$$

We have used the substitution $u = x - 2\beta$ and the identity $\int_{\mathbb{R}} e^{-\pi c u^2} du = c^{-1/2}$ for all $c > 0$. □

Corollary 1. *Fix $m, q, k \in \mathbb{Z}$, a bound B, and a standard deviation σ such that $B < \sigma < q$. Moreover, let $e \in \mathbb{Z}_q$ be such that $|e| \le B$. If $\sigma = \Omega\big(B\sqrt{m/\log k}\big)$, then*

$$\mathrm{RD}_2\big((e + \chi_\sigma)^m \| \chi_\sigma^m\big) = \mathrm{poly}(k)$$

where X^m denotes m independent samples from X.

Proof. Rényi divergence cannot grow by applying a function to both distributions. Since the discrete Gaussians $e + \chi_\sigma$ and χ_σ are obtained from the continuous Gaussians $\beta + \Psi_\alpha$ and Ψ_α by scaling and rounding, where $\beta = |e|/q$ and $\alpha = \sigma/q$, we see that

$$\mathrm{RD}_2\big(e + \chi_\sigma \| \chi_\sigma\big) \le \mathrm{RD}_2\big(\beta + \Psi_\alpha \| \Psi_\alpha\big) = \exp\big(2\pi(\beta/\alpha)^2\big) \le \exp\big(2\pi(B/\sigma)^2\big).$$

Therefore, $\mathrm{RD}_2\big((e + \chi_\sigma)^m \| \chi_\sigma^m\big) \le \exp\big(2\pi m(B/\sigma)^2\big)$, and the result follows. □

Acknowledgement. We would like to thank Damien Stehlé for sharing an early version of [BLL+15] and useful suggestions, Daniele Micciancio Nick Genise for insightful comments and for pointing out a flaw in an earlier version of the manuscript. We also thank the anonymous TCC 2016A reviewers for useful suggestions.

References

[AGS03] Adi Akavia, Shafi Goldwasser, and Shmuel Safra. Proving hard-core predicates using list decoding. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 146–157. IEEE Computer Society, 2003.

[AIK11] Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. How to garble arithmetic circuits. In *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 120–129, 2011.

[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 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 57–74. Springer, 2013.

[BLL+15] Shi Bai, Adeline Langlois, Tancrède Lepoint, Damien Stehlé, and Ron Steinfeld. Improved security proofs in lattice-based cryptography: using the rényi divergence rather than the statistical distance. *IACR Cryptology ePrint Archive*, 2015:483, 2015.

[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, *STOC*, pages 575–584. ACM, 2013.

[BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, April 15-19, 2012. Proceedings*, volume 7237 of *Lecture Notes in Computer Science*, pages 719–737. Springer, 2012.

[DGK+10] Yevgeniy Dodis, Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Public-key encryption schemes with auxiliary inputs. In *Theory of Cryptography, 7th Theory of Cryptography Conference, TCC 2010, Zurich, Switzerland, February 9-11, 2010. Proceedings*, pages 361–381, 2010.

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

[GL89] Oded Goldreich and Leonid A. Levin. A hard-core predicate for all one-way functions. In David S. Johnson, editor, *Proceedings of the 21st Annual ACM Symposium on Theory of Computing, May 14-17, 1989, Seattle, Washigton, USA*, pages 25–32. ACM, 1989.

[Lev86] Leonid A Levin. Average case complete problems. *SIAM J. Comput.*, 15(1):285–286, February 1986.

[LSS14] Adeline Langlois, Damien Stehlé, and Ron Steinfeld. Gghlite: More efficient multilinear maps from ideal lattices. In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, volume 8441 of *Lecture Notes in Computer Science*, pages 239–256. Springer, 2014.

[MM11] Daniele Micciancio and Petros Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In Rogaway [Rog11], pages 465–484.

[OPW11] Adam O'Neill, Chris Peikert, and Brent Waters. Bi-deniable public-key encryption. In Rogaway [Rog11], pages 525–542.

[Pei09] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In Michael Mitzenmacher, editor, *STOC*, pages 333–342. ACM, 2009.

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

[Rog11] Phillip Rogaway, editor. *Advances in Cryptology - CRYPTO 2011 - 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings*, volume 6841 of *Lecture Notes in Computer Science*. Springer, 2011.

[vEH14] Tim van Erven and Peter Harremoës. Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014.

[Yao82] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 80–91, 1982.

A Proof of Lemma 2

Proof. By the definition of Rényi divergence,

$$\begin{aligned}\mathrm{RD}_2\big(\mathsf{X}_s\|\mathsf{Y}_s\big) &= \mathrm{E}_{a\leftarrow R_q}\frac{\Pr\big(\mathsf{X}_s=(a,\lfloor a\cdot s\rceil_p)\big)}{\Pr\big(\mathsf{Y}_s=(a,\lfloor a\cdot s\rceil_p)\big)}\\ &= \mathrm{E}_{a\leftarrow R_q}\frac{1}{\Pr_{e\leftarrow\chi}\big(\lfloor a\cdot s+e\rceil_p=\lfloor a\cdot s\rceil_p\big)}.\end{aligned}$$

We define the set $\mathsf{border}_{p,q}(B)=\big\{x\in\mathbb{Z}_q:\big|x-\frac{q}{p}\lfloor x\rceil_p\big|<B\big\}$. For a ring element $a\in R_q$, we use a_i denote the ith coefficient in the power basis. For $t=0,\ldots,n$ and for any $t\in\{0,\ldots,n\}$, we define the set $\mathsf{BAD}_{s,t}=\big\{a\in R_q:|\{i\in[n],(a\cdot s)_i\in\mathsf{border}_{p,q}(B)\}|=t\big\}\big\}$. These are the a for which $a\cdot s$ has exactly t coefficients which are dangerously close to the rounding boundary. Fix arbitrary t and $a\in\mathsf{BAD}_{s,t}$. For any $i\in[n]$ such that $(a\cdot s)_i\notin\mathsf{border}_{p,q}(B)$, $\Pr_{e_i}[\lfloor(a\cdot s)_i+e_i\rceil_p=\lfloor(a\cdot s)_i\rceil_p]=1$. For any $i\in[n]$ such that $(a\cdot s)_i\in\mathsf{border}_{p,q}(B)$, the event $\lfloor(a\cdot s)_i+e_i\rceil_p=\lfloor(a\cdot s)_i\rceil_p$ still holds in one of the two cases $e_i\in[-B,\ldots,0]$ and $e_i\in[0,\ldots,B]$. By the assumption on the noise distribution $\Pr_{e_i}[\lfloor(a\cdot s)_i+e_i\rceil_p=\lfloor(a\cdot s)_i\rceil_p]\geq 1/2$. Because e is independent over all coefficients and a has exactly t coefficients in $\mathsf{border}_{p,q}(B)$, $\Pr_{e\leftarrow\chi}\big(\lfloor a\cdot s+e\rceil_p=\lfloor a\cdot s\rceil_p\big)\geq\frac{1}{2^t}$. Because s is a unit in R_q so that $a\cdot s$ is uniform over R_q and $\Pr[a\in\mathsf{BAD}_{s,t}]\leq\binom{n}{t}\left(1-\frac{|\mathrm{border}_{p,q}(B)|}{q}\right)^{n-t}\left(\frac{|\mathrm{border}_{p,q}(B)|}{q}\right)^{t}$. Conditioning over the event $a\in\mathsf{BAD}_{s,t}$, we conclude

$$\mathrm{RD}_2\big(\mathsf{X}_s\|\mathsf{Y}_s\big)\leq\sum_{t=0}^{n}2^t\cdot\Pr[a\in\mathsf{BAD}_{s,t}]=\left(1+\frac{|\mathrm{border}_{p,q}(B)|}{q}\right)^{n}.$$

The desired conclusion follows from $|\mathsf{border}_{p,q}(B)|\leq 2pB$. □