Improved Security Proofs in Lattice-Based Cryptography: Using the Rényi Divergence Rather Than the Statistical Distance

Shi Bai[1(✉)], Adeline Langlois[2,3], Tancrède Lepoint[4], Damien Stehlé[1], and Ron Steinfeld[5]

[1] Laboratoire LIP (U. Lyon, CNRS, ENSL, INRIA, UCBL), ENS de Lyon, Lyon, France
{shi.bai,damien.stehle}@ens-lyon.fr
http://perso.ens-lyon.fr/shi.bai, http://perso.ens-lyon.fr/damien.stehle
[2] EPFL, Lausanne, Switzerland
adeline.langlois@epfl.ch
http://lasec.epfl.ch/alangloi/
[3] CNRS/IRISA, Rennes, France
[4] CryptoExperts, Paris, France
tancrede.lepoint@cryptoexperts.com
https://www.cryptoexperts.com/tlepoint/
[5] Faculty of Information Technology, Monash University, Clayton, Australia
ron.steinfeld@monash.edu
http://users.monash.edu.au/rste/

Abstract. The Rényi divergence is a measure of closeness of two probability distributions. We show that it can often be used as an alternative to the statistical distance in security proofs for lattice-based cryptography. Using the Rényi divergence is particularly suited for security proofs of primitives in which the attacker is required to solve a search problem (e.g., forging a signature). We show that it may also be used in the case of distinguishing problems (e.g., semantic security of encryption schemes), when they enjoy a public sampleability property. The techniques lead to security proofs for schemes with smaller parameters, and sometimes to simpler security proofs than the existing ones.

1 Introduction

Let D_1 and D_2 be two non-vanishing probability distributions over a common measurable support X. Let $a \in (1, +\infty)$. The *Rényi divergence* [Rén61,EH12] (RD for short) $R_a(D_1\|D_2)$ of order a between D_1 and D_2 is defined as the $((a-1)$th root of the) expected value of $(D_1(x)/D_2(x))^{a-1}$ over the randomness of x sampled from D_1. For notational convenience, our definition of the RD is the exponential of the classical definition [EH12]. The RD is an alternative to the statistical distance (SD for short) $\Delta(D_1, D_2) = \frac{1}{2}\sum_{x\in X}|D_1(x) - D_2(x)|$ as measure of distribution closeness, where we replace the difference in SD, by the ratio in RD. RD enjoys several properties that are analogous of those enjoyed

T. Iwata and J.H. Cheon (Eds.): ASIACRYPT 2015, Part I, LNCS 9452, pp. 3–24, 2015.
DOI: 10.1007/978-3-662-48797-6_1

by SD, where addition in the property of SD is replaced by multiplication in the analogous property of RD (see Subsect. 2.3).

SD is ubiquitous in cryptographic security proofs. One of its most useful properties is the so-called *probability preservation property*: For any measurable event $E \subseteq X$, we have $D_2(E) \geq D_1(E) - \Delta(D_1, D_2)$. RD enjoys the analogous property $D_2(E) \geq D_1(E)^{\frac{a}{a-1}} / R_a(D_1 \| D_2)$. If the event E occurs with significant probability under D_1, and if the SD (resp. RD) is small, then the event E also occurs with significant probability under D_2. These properties are particularly handy when the success of an attacker against a given scheme can be described as an event whose probability should be negligible, e.g., the attacker outputs a new valid message-signature pair for a signature scheme. If in the attacker succeeds with good probability in the real scheme based on distribution D_1, then it also succeeds with good probability in the simulated scheme (of the security proof) based on distribution D_2.

To make the SD probability preservation property useful, it must be ensured that the SD $\Delta(D_1, D_2)$ is smaller than any $D_1(E)$ that the security proof must handle. Typically, the quantity $D_1(E)$ is assumed to be greater than some success probability lower bound ε, which is of the order of $1/\mathrm{poly}(\lambda)$ where λ refers to the security parameter, or even $2^{-o(\lambda)}$ if the proof handles attackers whose success probabilities can be sub-exponentially small (which we believe better reflects practical objectives). As a result, the SD $\Delta(D_1, D_2)$ must be $< \varepsilon$ for the SD probability preservation property to be relevant. Similarly, the RD probability preservation property is non-vacuous when the RD $R_a(D_1 \| D_2)$ is $\leq \mathrm{poly}(1/\varepsilon)$. In many cases, the latter seems less demanding than the former: in all our applications of RD, the RD between D_1 and D_2 is small while their SD is too large for the SD probability preservation to be applicable. In fact, as we will see in Subsect. 2.3, the RD becomes sufficiently small to be useful before the SD when $\sup_x D_1(x)/D_2(x)$ tends to 1. This explains the superiority of the RD in several of our applications.

Although RD seems more amenable than SD for search problems, it seems less so for distinguishing problems. A typical cryptographic example is semantic security of an encryption scheme. Semantic security requires an adversary $\mathcal{A}$ to distinguish between the encryption distributions of two plaintext messages of its choosing: the distinguishing advantage $\mathrm{Adv}_{\mathcal{A}}(D_1, D_2)$, defined as the difference of probabilities that $\mathcal{A}$ outputs 1 using D_1 or D_2, should be large. In security proofs, algorithm $\mathcal{A}$ is often called on distributions D_1' and D_2' that are close to D_1 and D_2 (respectively). If the SDs between D_1 and D_1' and D_2 and D_2' are both bounded from above by ε, then, by the SD probability preservation property (used twice), we have $\mathrm{Adv}_{\mathcal{A}}(D_1', D_2') \geq \mathrm{Adv}_{\mathcal{A}}(D_1, D_2) - 2\varepsilon$. As a result, SD can be used to distinguishing problems in a similar fashion as for search problems. The multiplicativity of the RD probability preservation property seems to prevent RD from being applicable to distinguishing problems.

We replace the statistical distance by the Rényi divergence in several security proofs for lattice-based cryptographic primitives. *Lattice-based cryptography* is a relatively recent cryptographic paradigm in which cryptographic primitives are shown at least as secure as it is hard to solve standard problems over lattices (see the survey [MR09]). Security proofs in lattice-based cryptography involve different types of distributions, often over infinite sets, such as continuous Gaussian distributions and Gaussian distributions with lattice supports. The RD seems particularly well suited to quantify the closeness of Gaussian distributions. Consider for example two continuous distributions over the reals, both with standard deviation 1, but one with center 0 and the other one with center c. Their SD is linear in c, so that c must remain extremely small for the SD probability preservation property to be useful. On the other hand, their RD of order $a = 2$ is bounded as $\exp(O(c^2))$ so that the RD preservation property remains useful event for slightly growing c.

RD was first used in lattice-based cryptography by [LPR13], in the decision to search reduction for the Ring Learning With Errors problem (which serves as a security foundation for many asymptotically fast primitives). It was then exploited in [LSS14] to decrease the parameters of the Garg et al. (approximation to) cryptographic multilinear maps [GGH13]. In the present work, we present a more extensive study of the power of RD in lattice-based cryptography, by showing several independent applications of RD. In some cases, it leads to security proofs allowing to take smaller parameters in the cryptographic schemes, hence leading to efficiency improvements. In other cases, this leads to alternative security proofs that are conceptually simpler.

Our applications of RD also include distinguishing problems. To circumvent the aforementioned a priori limitation of the RD probability preservation property for distinguishing problems, we propose an alternative approach that handles a class of distinguishing problems, enjoying a special property that we call *public sampleability*. This public sampleability allows to estimate success probabilities via Hoeffding's bound.

The applications we show in lattice-based cryptography are as follows:

- Smaller signatures for the Hash-and-Sign GPV signature scheme [GPV08].
- Smaller storage requirement for the Fiat-Shamir BLISS signature scheme [DDLL13, PDG14, Duc14].
- Alternative proof that the Learning With Errors (LWE) problem with noise chosen uniformly in an interval is no easier than the Learning With Errors problem with Gaussian noise [DMQ13]. Our reduction does not require the latter problem to be hard, and it is hence marginally more general as it also applies to distributions with smaller noises. Further, our reduction preserves the LWE dimension n, and is hence tighter than the one from [DMQ13] (the latter degrades the LWE dimension by a constant factor).[1]
- Smaller parameters in the dual-Regev encryption scheme from [GPV08].

[1] Note that LWE with uniform noise in a small interval is also investigated in [MP13], with a focus on the number of LWE samples. The reduction from [MP13] does not preserve the LWE reduction either.

We think RD is likely to have further applications in lattice-based cryptography, for search and for distinguishing problems.

Related Works. The framework for using RD in distinguishing problems was used in [LPSS14], in the context of the k-LWE problem (a variant of LWE in which the attacker is given extra information). In [PDG14], Pöpplemann, Ducas and Güneysu used the Kullback-Leibler divergence (which is the RD of order $a = 1$) to lower the storage requirement of [DDLL13]. Asymptotically, using the Kullback-Leibler divergence rather than SD only leads to a constant factor improvement. Our approach allows bigger savings in the case where the number of signature queries is limited, as explained in Sect. 3.

Very recently, Bogdanov *et al.* [BGM+15] adapted parts of our RD-based hardness proof for LWE with noise uniform in a small interval, to the Learning With Rounding problem. In particular, they obtained a substantial improvement over the hardness results of [BPR12, AKPW13].

Road-Map. In Sect. 2, we provide necessary background on lattice-based cryptography, and on the Rényi divergence. In Sect. 3, we use RD to improve some lattice-based signature scheme parameters. Section 4 contains the description of the framework in which we can use RD for distinguishing problems, which we apply to the dual-Regev encryption scheme. In Sect. 5, we describe an alternative hardness proof for LWE with noise uniformly chosen in an interval.

Notations. If x is a real number, we let $\lfloor x \rceil$ denote a closest integer to x. The notation ln refers to the natural logarithm and the notation log refers to the base 2 logarithm. We define $\mathbb{T} = ([0,1],+)$. For an integer q, we let $\mathbb{Z}_q$ denote the ring of integers modulo q. We let $\mathbb{T}_q$ denote the group $\mathbb{T}_q = \{i/q \bmod 1 : i \in \mathbb{Z}\} \subseteq \mathbb{T}$. Vectors are denoted in bold. If $\boldsymbol{b}$ is a vector in $\mathbb{R}^d$, we let $\|\boldsymbol{b}\|$ denote its Euclidean norm. By default, all our vectors are column vectors.

If D is a probability distribution, we let $\mathrm{Supp}(D) = \{x : D(x) \neq 0\}$ denote its support. For a set X of finite weight, we let $U(X)$ denote the uniform distribution on X. The statistical distance between two distributions D_1 and D_2 over a countable support X is $\Delta(D_1, D_2) = \frac{1}{2}\sum_{x\in X} |D_1(x) - D_2(x)|$. This definition is extended in the natural way to continuous distributions. If $f : X \to \mathbb{R}$ takes non-negative values, then for all countable $Y \subseteq X$, we define $f(Y) = \sum_{y\in Y} f(y) \in [0, +\infty]$. For any vector $\boldsymbol{c} \in \mathbb{R}^n$ and any real $s > 0$, the (spherical) Gaussian function with standard deviation parameter s and center $\boldsymbol{c}$ is defined as follows: $\forall \boldsymbol{x} \in \mathbb{R}^n, \rho_{s,\boldsymbol{c}}(\boldsymbol{x}) = \exp(-\pi\|\boldsymbol{x} - \boldsymbol{c}\|^2/s^2)$. The Gaussian distribution is $D_{s,\boldsymbol{c}} = \rho_{s,\boldsymbol{c}}/s^n$. When $\boldsymbol{c} = \boldsymbol{0}$, we may omit the subscript $\boldsymbol{c}$.

We use the usual Landau notations. A function $f(\lambda)$ is said negligible if it is $\lambda^{-\omega(1)}$. A probability $p(\lambda)$ is said overwhelming if it is $1 - \lambda^{-\omega(1)}$.

The distinguishing advantage of an algorithm $\mathcal{A}$ between two distributions D_0 and D_1 is defined as $\mathrm{Adv}_{\mathcal{A}}(D_0, D_1) = |\Pr_{x\hookleftarrow D_0}[\mathcal{A}(x) = 1] - \Pr_{x\hookleftarrow D_1}[\mathcal{A}(x) = 1]|$, where the probabilities are taken over the randomness of the input x and the internal randomness of $\mathcal{A}$. Algorithm $\mathcal{A}$ is said to be an (ε, T)-distinguisher if it runs in time $\leq T$ and if $\mathrm{Adv}_{\mathcal{A}}(D_0, D_1) \geq \varepsilon$.

2 Preliminaries

We assume the reader is familiar with standard cryptographic notions, as well as with lattices and lattice-based cryptography. We refer to [Reg09a, MR09] for introductions on the latter topic.

2.1 Lattices

A (full-rank) n-dimensional *Euclidean lattice* $\Lambda \subseteq \mathbb{R}^n$ is the set of all integer linear combinations $\sum_{i=1}^{n} x_i \boldsymbol{b}_i$ of some $\mathbb{R}$-basis $(\boldsymbol{b}_i)_{1 \leq i \leq n}$ of $\mathbb{R}^n$. In this setup, the tuple $(\boldsymbol{b}_i)_i$ is said to form a $\mathbb{Z}$-basis of Λ. For a lattice Λ and any $i \leq n$, the ith successive minimum $\lambda_i(\Lambda)$ is the smallest radius r such that Λ contains i linearly independent vectors of norm at most r. The dual Λ^* of a lattice Λ is defined as $\Lambda^* = \{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{y}^t \Lambda \subseteq \mathbb{Z}^n\}$.

The (spherical) *discrete Gaussian distribution over a lattice* $\Lambda \subseteq \mathbb{R}^n$, with standard deviation parameter $s > 0$ and center $\boldsymbol{c}$ is defined as:

$$\forall \boldsymbol{x} \in \Lambda, D_{\Lambda,s,\boldsymbol{c}} = \frac{\rho_{s,\boldsymbol{c}}(\boldsymbol{x})}{\rho_{s,\boldsymbol{c}}(\Lambda)}.$$

When the center is $\mathbf{0}$, we omit the subscript $\boldsymbol{c}$.

The *smoothing parameter* [MR07] of an n-dimensional lattice Λ with respect to $\varepsilon > 0$, denoted by $\eta_\varepsilon(\Lambda)$, is the smallest $s > 0$ such that $\rho_{1/s}(\Lambda^* \setminus \{0\}) \leq \varepsilon$. We use the following properties.

Lemma 2.1 ([MR07, Lemma 3.3]). *Let Λ be an n-dimensional lattice and $\varepsilon > 0$. Then*

$$\eta_\varepsilon(\Lambda) \leq \sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}} \cdot \lambda_n(\Lambda).$$

Lemma 2.2 (Adapted from [GPV08, Lemma 5.3]). *Let $m, n \geq 1$ and q a prime integer, with $m \geq 2n \ln q$. For $A \in \mathbb{Z}_q^{n \times m}$ we define $A^\perp$ as the lattice $\{\boldsymbol{x} \in \mathbb{Z}^m : A\boldsymbol{x} = \mathbf{0} \mod q\}$. Then,*

$$\forall \varepsilon < 1/2 : \Pr_{A \hookleftarrow U(\mathbb{Z}_q^{n \times m})} \left[\eta_\varepsilon(A^\perp) \geq 4\sqrt{\frac{\ln(4m/\varepsilon)}{\pi}} \right] \leq q^{-n}.$$

Lemma 2.3 (Adapted from [GPV08, Corollary 2.8]). *Let Λ, Λ' be n-dimensional lattices with $\Lambda' \subseteq \Lambda$ and $\varepsilon \in (0, 1/2)$. Then for any $\boldsymbol{c} \in \mathbb{R}^n$ and $s \geq \eta_\varepsilon(\Lambda')$ and any $x \in \Lambda/\Lambda'$ we have*

$$(D_{\Lambda,s,c} \bmod \Lambda')(x) \in \left[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}\right] \cdot \frac{\det(\Lambda)}{\det(\Lambda')}.$$

2.2 The SIS and LWE Problems

The Small Integer Solution (SIS) problem was introduced by Ajtai in [Ajt96]. It serves as a security foundation for numerous cryptographic primitives, including, among many others, hash functions [Ajt96] and signatures [GPV08, DDLL13].

Definition 2.4. *Let $m \geq n \geq 1$ and $q \geq 2$ be integers, and β a positive real. The* $\mathrm{SIS}_{n,m,q,\beta}$ *problem is as follows: given $A \hookleftarrow U(\mathbb{Z}_q^{n \times m})$, the goal is to find $\boldsymbol{x} \in \mathbb{Z}^m$ such that $A\boldsymbol{x} = \boldsymbol{0} \mod q$ and $0 < \|\boldsymbol{x}\| \leq \beta$.*

The SIS problem was proven by Ajtai [Ajt96] to be at least as hard as some standard worst-case problems over Euclidean lattices, under specific parameter constraints. We refer to [GPV08] for an improved (and simplified) reduction.

The Learning With Errors (LWE) problem was introduced in 2005 by Regev [Reg05, Reg09b]. LWE is also extensively used as a security foundation, for encryption schemes [Reg09b, GPV08], fully homomorphic encryption schemes [BV11], and pseudo-random functions [BPR12, AKPW13], among many others. Its definition involves the following distribution. Let χ be a distribution over $\mathbb{T}$, $q \geq 2$, $n \geq 1$ and $\boldsymbol{s} \in \mathbb{Z}_q^n$. A sample from $A_{\boldsymbol{s},\chi}$ is of the form $(\boldsymbol{a}, b) \in \mathbb{Z}_q^n \times \mathbb{T}$, with $\boldsymbol{a} \hookleftarrow U(\mathbb{Z}_q^n)$, $b = \frac{1}{q}\langle \boldsymbol{a}, \boldsymbol{s} \rangle + e$ and $e \hookleftarrow \chi$.

Definition 2.5. *Let χ be a distribution over $\mathbb{T}$, $q \geq 2$, and $m \geq n \geq 1$. The search variant* $\mathrm{sLWE}_{n,q,\chi,m}$ *of the* LWE *problem is as follows: given m samples from $A_{\boldsymbol{s},\chi}$ for some $\boldsymbol{s} \in \mathbb{Z}_q^n$, the goal is to find $\boldsymbol{s}$. The decision variant* $\mathrm{LWE}_{n,q,\chi,m}$ *consists in distinguishing between the distributions $(A_{\boldsymbol{s},\chi})^m$ and $U(\mathbb{Z}_q^n \times \mathbb{T})^m$, where $\boldsymbol{s} \hookleftarrow U(\mathbb{Z}_q^n)$.*

In some cases, it is convenient to use an error distribution χ whose support is $\mathbb{T}_q$. In these cases, the definition of LWE is adapted in that $U(\mathbb{Z}_q^n \times \mathbb{T})$ is replaced by $U(\mathbb{Z}_q^n \times \mathbb{T}_q)$. Note also that for a fixed number of samples m, we can represent the LWE samples using matrices. The $\boldsymbol{a_i}$'s form the rows of a matrix A uniform in $\mathbb{Z}_q^{m \times n}$, and the scalar product is represented by the product between A and $\boldsymbol{s}$.

Regev [Reg09b] gave a quantum reduction from standard worst-case problems over Euclidean lattices to sLWE and LWE, under specific parameter constraints. Classical (but weaker) reductions have later been obtained (see [Pei09, BLP+13]). We will use the following sample-preserving search to decision reduction for LWE.

Theorem 2.6 (Adapted from [MM11, Proposition 4.10]). *If $q \leq \mathrm{poly}(m, n)$ is prime and the error distribution χ has support in $\mathbb{T}_q$, then there exists a reduction from* $\mathrm{sLWE}_{n,q,\chi,m}$ *to* $\mathrm{LWE}_{n,q,\chi,m}$ *that is polynomial in n and m.*

2.3 The Rényi Divergence

For any two discrete probability distributions P and Q such that $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$ and $a \in (1, +\infty)$, we define the Rényi divergence of order a by

$$R_a(P\|Q) = \left(\sum_{x \in \mathrm{Supp}(P)} \frac{P(x)^a}{Q(x)^{a-1}} \right)^{\frac{1}{a-1}}.$$

We omit the a subscript when $a = 2$. We define the Rényi divergences of orders 1 and $+\infty$ by

$$R_1(P\|Q) = \exp\left(\sum_{x \in \mathrm{Supp}(P)} P(x) \log \frac{P(x)}{Q(x)} \right) \quad \text{and} \quad R_\infty(P\|Q) = \max_{x \in \mathrm{Supp}(P)} \frac{P(x)}{Q(x)}.$$

The definitions are extended in the natural way to continuous distributions. The divergence R_1 is the (exponential of) the Kullback-Leibler divergence.

For any fixed P, Q, the function $a \mapsto R_a(P\|Q) \in (0, +\infty]$ is non-decreasing, continuous over $(1, +\infty)$, tends to $R_\infty(P\|Q)$ when a grows to infinity, and if $R_a(P\|Q)$ is finite for some a, then $R_a(P\|Q)$ tends to $R_1(P\|Q)$ when a tends to 1 (we refer to [EH12] for proofs). A direct consequence is that if $P(x)/Q(x) \leq c$ for all $x \in \mathrm{Supp}(P)$ and for some constant c, then $R_a(P\|Q) \leq R_\infty(P\|Q) \leq c$. In the same setup, we have $\Delta(P, Q) \leq c/2$.

The following properties can be considered the multiplicative analogues of those of the SD. We refer to [EH12, LSS14] for proofs.

Lemma 2.7. *Let $a \in [1, +\infty]$. Let P and Q denote distributions with $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$. Then the following properties hold:*

- ***Log. Positivity:*** $R_a(P\|Q) \geq R_a(P\|P) = 1$.
- ***Data Processing Inequality:*** $R_a(P^f\|Q^f) \leq R_a(P\|Q)$ *for any function f, where P^f (resp. Q^f) denotes the distribution of $f(y)$ induced by sampling $y \hookleftarrow P$ (resp. $y \hookleftarrow Q$).*
- ***Multiplicativity:*** *Assume P and Q are two distributions of a pair of random variables (Y_1, Y_2). For $i \in \{1, 2\}$, let P_i (resp. Q_i) denote the marginal distribution of Y_i under P (resp. Q), and let $P_{2|1}(\cdot|y_1)$ (resp. $Q_{2|1}(\cdot|y_1)$) denote the conditional distribution of Y_2 given that $Y_1 = y_1$. Then we have:*
  - $R_a(P\|Q) = R_a(P_1\|Q_1) \cdot R_a(P_2\|Q_2)$ *if Y_1 and Y_2 are independent.*
  - $R_a(P\|Q) \leq R_\infty(P_1\|Q_1) \cdot \max_{y_1 \in X} R_a(P_{2|1}(\cdot|y_1)\|Q_{2|1}(\cdot|y_1))$.
- ***Probability Preservation:*** *Let $A \subseteq \mathrm{Supp}(Q)$ be an arbitrary event. If $a \in (1, +\infty)$, then $Q(A) \geq P(A)^{\frac{a}{a-1}}/R_a(P\|Q)$. Further, we have*

$$Q(A) \geq P(A)/R_\infty(P\|Q).$$

Let P_1, P_2, P_3 be three distributions with $\mathrm{Supp}(P_1) \subseteq \mathrm{Supp}(P_2) \subseteq \mathrm{Supp}(P_3)$. Then we have:

- ***Weak Triangle Inequality:***

$$R_a(P_1\|P_3) \leq \begin{cases} R_a(P_1\|P_2) \cdot R_\infty(P_2\|P_3), \\ R_\infty(P_1\|P_2)^{\frac{a}{a-1}} \cdot R_a(P_2\|P_3) \end{cases} \quad \text{if} \quad a \in (1, +\infty).$$

Getting back to the setup in which $P(x)/Q(x) \leq c$ for all $x \in \mathrm{Supp}(P)$ and for some constant c, the RD probability preservation property above is relevant even for large c, whereas the analogous SD probability preservation property starts making sense only when $c < 2$.

Pinsker's inequality is the analogue of the probability preservation property for $a = 1$: for an arbitrary event $A \subseteq \mathrm{Supp}(Q)$, we have $Q(A) \geq P(A) - \sqrt{\ln R_1(P\|Q)/2}$ (see [PDG14, Lemma 1] for a proof). Analogously to the statistical distance, this probability preservation property is useful for unlikely events A only if $\ln R_1(P\|Q)$ is very small. We refer to Subsect. 3.1 for additional comments on this property.

2.4 RD Bounds

We will use the following result, adapted from [LSS14].

Lemma 2.8. *For any n-dimensional lattice $\Lambda \subseteq \mathbb{R}^n$ and $s > 0$, let P be the distribution $D_{\Lambda,s,\boldsymbol{c}}$ and Q be the distribution $D_{\Lambda,s,\boldsymbol{c}'}$ for some fixed $\boldsymbol{c}, \boldsymbol{c}' \in \mathbb{R}^n$. If $\boldsymbol{c}, \boldsymbol{c}' \in \Lambda$, let $\varepsilon = 0$. Otherwise, fix $\varepsilon \in (0,1)$ and assume that $s \geq \eta_\varepsilon(\Lambda)$. Then, for any $a \in (1, +\infty)$:*

$$R_a(P\|Q) \in \left[\left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{\frac{2}{a-1}}, \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{\frac{2}{a-1}}\right] \cdot \exp\left(a\pi\frac{\|\boldsymbol{c}-\boldsymbol{c}'\|^2}{s^2}\right).$$

It may be checked that also $R_1(P\|Q)$ is of the order of $\exp(\|\boldsymbol{c}'-\boldsymbol{c}\|^2/s^2)$, $R_\infty(P\|Q) = +\infty$ and $\Delta(P,Q)$ is of the order of $\|\boldsymbol{c}'-\boldsymbol{c}\|^2/s^2$. In that setup, the RD of order $a = \infty$ is useless, and the probability preservation properties of the SD and RD of order $a = 1$ lead to interesting bounds for events occurring only when $\|\boldsymbol{c}'-\boldsymbol{c}\|/s = o(\varepsilon)$. Oppositely, for any $a \in (1, +\infty)$, the probability preservation property for the RD of order $a \in (1, +\infty)$ may be used with $\|\boldsymbol{c}'-\boldsymbol{c}\|/s = O(\sqrt{\log(1/\varepsilon)})$ while still leading to probabilistic lower bounds of the order of $\varepsilon^{O(1)}$.

As we have already seen, if two distributions are close in a uniform sense, then their RD is small. We observe the following immediate consequence of Lemma 2.3, that allows replacing the SD with the RD in the context of smoothing arguments, in order to save on the required parameter s. In applications of Lemma 2.3, it is customary to use $s \geq \eta_\varepsilon(\Lambda')$ with $\varepsilon \leq 2^{-\lambda}$, in order to make the distribution $D_{\Lambda/\Lambda',s,\boldsymbol{c}} = D_{\Lambda,s,\boldsymbol{c}} \mod \Lambda'$ within SD $2^{-\lambda}$ of the uniform distribution $U(\Lambda/\Lambda')$. This translates via Lemma 2.1 to use $s = \Omega(\sqrt{\lambda + \log n} \cdot \lambda_n(\Lambda'))$. Whereas if using an RD bound $R_\infty(D_{\Lambda/\Lambda',s,\boldsymbol{c}}\|U_{\Lambda/\Lambda'}) = O(1)$ suffices for the application, one can take $\varepsilon = O(1)$ in the corollary below, which translates to just $s = \Omega(\sqrt{\log n} \cdot \lambda_n(\Lambda'))$, saving a factor $\Theta(\sqrt{\lambda})$.

Lemma 2.9. *Let Λ, Λ' be n-dimensional lattices with $\Lambda' \subseteq \Lambda$ and $\varepsilon \in (0, 1/2)$. Let $D_{\Lambda/\Lambda',s,c}$ denote the distribution on Λ/Λ' induced by sampling from $D_{\Lambda,s,c}$*

and reducing modulo Λ', and let $U_{\Lambda/\Lambda'}$ denote the uniform distribution on Λ/Λ'. Then for any $c \in \mathbb{R}^n$ and $s \geq \eta_\varepsilon(\Lambda')$ and any $x \in \Lambda/\Lambda'$ we have

$$R_\infty(D_{\Lambda/\Lambda',s,c} \| U_{\Lambda/\Lambda'}) \leq \frac{1+\varepsilon}{1-\varepsilon}.$$

3 Application to Lattice-Based Signature Schemes

In this section, we use the RD to improve the security proofs of the GPV and BLISS signature schemes [GPV08, DDLL13], allowing to take smaller parameters for any fixed security level.

3.1 Sampling Discrete Gaussians and the BLISS Signature Scheme

We show that the use of RD in place of SD leads to significant savings in the required precision of integers sampled according to a discrete Gaussian distribution in the security analysis of lattice-based signature schemes. These savings consequently lower the precomputed table storage for sampling discrete Gaussians with the method described in [DDLL13, PDG14]. In Table 1, we provide a numerical comparison of RD and SD based on an instantiation of BLISS-I.

Discrete Gaussian Sampling. In the BLISS signature scheme [DDLL13] (and similarly in earlier variants [Lyu12]), each signature requires the signing algorithm to sample $O(n)$ independent integers from the 1-dimensional discrete Gaussian distribution $D_{\mathbb{Z},s}$, where $s = O(m)$ is the standard deviation parameter (here the variable m denotes a parameter related to the underlying lattice dimension, and is typically in the order of several hundreds)[2].

In [DDLL13], a particularly efficient sampling algorithm for $D_{\mathbb{Z},s}$ is presented. To produce a sample from $D_{\mathbb{Z},s}$, this algorithm samples about $\ell = \lceil \log(0.22 s^2(1+2\tau s)) \rceil$ Bernoulli random variables of the form $B_{\exp(-\pi 2^i/s^2)}$, where $i = 0, \ldots, \ell - 1$ and $\tau = O(\sqrt{\lambda})$ is the tail-cut factor for the Gaussian. To sample those Bernoulli random variables, the authors of [DDLL13] use a precomputed table of the probabilities $c_i = \exp(-\pi 2^i/s^2)$, for $i = 1, \ldots, \ell$. Since these probabilities are real numbers, they must be truncated to some bit precision p in the precomputed table, so that truncated values $\tilde{c}_i = c_i + \varepsilon_i$ are stored, where $|\varepsilon_i| \leq 2^{-p} c_i$ are the truncation errors.

In previous works, the precision was determined by an analysis either based on the statistical distance (SD) [DDLL13] or the Kullback-Leibler divergence (KLD) [PDG14]. In this section, we review and complete these methods, and we propose an RD-based analysis that leads to bigger savings, asymptotically and in practice (see Table 1). More precisely, we give sufficient lower bounds on the precision p to ensure security on the scheme implemented with truncated values against adversaries succeeding with probability $\geq \varepsilon$ and making $\leq q_s$ signing

[2] Note that [Lyu12, DDLL13] consider the unnormalized Gaussian function $\rho'_{\sigma,c}(\boldsymbol{x}) = \exp(-\|\boldsymbol{x} - \boldsymbol{c}\|/(2\sigma^2))$ instead of $\rho_{s,c}$. We have $\rho_{s,c} = \rho'_{\sigma,c}$ when $\sigma = s/\sqrt{2\pi}$.

queries. For any adversary, the distributions Φ' and Φ denote the signatures in the view of the adversary in the untruncated (resp. truncated) cases.

SD-based analysis [DDLL13]. Any forging adversary $\mathcal{A}$ with success probability $\geq \varepsilon$ on the scheme implemented with truncated Gaussian has a success probability $\varepsilon' \geq \varepsilon - \Delta(\Phi, \Phi')$ against the scheme implemented with perfect Gaussian sampling. We select parameters to handle adversaries with success probabilities $\geq \varepsilon/2$ against the untruncated scheme; we can set the required precision p so that $\Delta(\Phi, \Phi') \leq \varepsilon/2$. Each signature requires $\ell \cdot m$ samples from the Bernoulli random variables $(B_{\tilde{c}_i})_i$. To ensure security against q_s signing queries, each of the truncated Bernoulli random variables $B_{\tilde{c}_i}$ should be within SD $\Delta(\Phi, \Phi')/(\ell \cdot m \cdot q_s)$ of the desired B_{c_i} (by the union bound). Using $\Delta(B_{\tilde{c}_i}, B_{c_i}) = |\varepsilon_i| \leq 2^{-p} c_i \leq 2^{-p-1}$ leads to a precision requirement

$$p \geq \log(\ell \cdot m \cdot q_s / \Delta(\Phi, \Phi')) \geq \log(\ell \cdot m \cdot q_s / \varepsilon). \tag{1}$$

The overall precomputed table is hence of bit-size $L_{\mathrm{SD}} = p \cdot \ell \geq \log(\ell \cdot m \cdot q_s/\varepsilon) \cdot \ell$.

Note that in [DDLL13], the authors omitted the term $\ell \cdot m \cdot q_s$ in their analysis: they only ensured that $\Delta(B_{\tilde{c}_i}, B_{c_i}) \leq \varepsilon$, leading to the requirement that $p \geq \log(1/\varepsilon)$.

One may also set the precision p_i depending on i for $0 \leq i \leq \ell - 1$. It is sufficient to set

$$2^{-p_i} c_i = 2^{-p_i} \exp(-\pi 2^i / s^2) \leq (\varepsilon/2)/(\ell \cdot m \cdot q_s).$$

Hence the precision p_i is

$$p_i \geq \log\left(\frac{\ell \cdot m \cdot q_s}{\varepsilon} \cdot \exp(-\pi 2^i / s^2)\right) + 1. \tag{2}$$

The bit-size of the overall precomputed table can be computed as a sum of the above p_i's. Using the symmetry of the Bernoulli variable, we can further drop the bit-size of the precomputed table.

KLD-based analysis [PDG14]. In [PDG14], Pöppelman, Ducas and Güneysu replace the SD-based analysis by a KLD-based analysis (i.e., using the RD of order $a = 1$) to reduce the precision p needed in the precomputed table. They show that any forging adversary $\mathcal{A}$ with success probability ε on the scheme implemented with truncated Gaussian has a success probability $\varepsilon' \geq \varepsilon - \sqrt{\ln R_1(\Phi \| \Phi')/2}$ on the scheme implemented with perfect Gaussian (see remark at the end of Subsect. 2.3). By the multiplicative property of the RD over the $\ell \cdot m \cdot q_s$ independent samples needed for signing q_s times, we get that $R_1(\Phi \| \Phi') \leq (\max_{i=1,\ldots,\ell} R_1(B_{\tilde{c}_i} \| B_{c_i}))^{\ell \cdot m \cdot q_s}$. Now, we have:

$$\begin{aligned}
\ln R_1(B_{\tilde{c}_i} \| B_{c_i}) &= (1 - c_i - \varepsilon_i) \ln \frac{1 - c_i - \varepsilon_i}{1 - c_i} + (c_i + \varepsilon_i) \ln \frac{c_i + \varepsilon_i}{c_i} \\
&\leq -(1 - c_i - \varepsilon_i) \frac{\varepsilon_i}{1 - c_i} + (c_i + \varepsilon_i) \frac{\varepsilon_i}{c_i} \;=\; \frac{\varepsilon_i^2}{(1 - c_i) c_i}.
\end{aligned}$$

Using $|\varepsilon_i| \le 2^{-p} c_i$ and $1 - c_i \ge 1/2$, we obtain $\ln R_1(B_{\tilde{c}_i} \| B_{c_i}) = 2^{-2p+1} \frac{c_i}{1-c_i} \le 2^{-2p}$. Therefore, we obtain $\varepsilon' \ge \varepsilon - \sqrt{\ell \cdot m \cdot q_s \cdot 2^{-2p}}$. We can select parameters such that $\sqrt{\ell \cdot m \cdot q_s \cdot 2^{-2p+1}} \le \varepsilon/2$. This leads to a precision requirement

$$p \ge \frac{1}{2} \log \left(\frac{\ell \cdot m \cdot q_s}{\varepsilon^2} \right) + \frac{1}{2} . \tag{3}$$

The overall precomputed table is hence of bit-size $L_{\mathrm{KLD}} \ge (\log(\ell \cdot m \cdot q_s/\varepsilon^2)/2 + 1/2) \cdot \ell$. This KLD-based analysis may save some storage if ε is not too small.

Note that in [PDG14], the authors selected $\varepsilon = 1/2$ and $\ell \cdot m \cdot q_s = 2^\lambda$ where λ is the desired bit-security, and hence obtained $p \ge \lambda/2 + 1$.

One may also set the precision p_i depending on i. It is sufficient to set

$$p_i \ge \frac{1}{2} \log \left(\frac{\ell \cdot m \cdot q_s}{\varepsilon^2} \cdot \frac{c_i}{1 - c_i} \right) + 1. \tag{4}$$

Using symmetry, we may assume $c_i \le 1/2$.

R_∞-based analysis. The probability preservation property of the Rényi divergence from Lemma 2.7 is multiplicative for $a > 1$ (rather than additive for $a = 1$). Here we use the order $a = \infty$. This property gives that any forging adversary $\mathcal{A}$ having success probability ε on the scheme implemented with truncated Gaussian sampling has a success probability $\varepsilon' \ge \varepsilon / R_\infty(\Phi \| \Phi')$ on the scheme implemented with perfect Gaussian. If $R = R_\infty(\Phi \| \Phi') \le O(1)$, then $\varepsilon' = \Omega(\varepsilon)$. By the multiplicative property of the RD (over the $\ell \cdot m \cdot q_s$ samples needed for signing q_s times), we have $R_\infty(\Phi \| \Phi') \le R_\infty(B_{\tilde{c}_i} \| B_{c_i})^{\ell \cdot m \cdot q_s}$. By our assumption that $c_i \le 1/2$, we have $R_\infty(B_{\tilde{c}_i} \| B_{c_i}) = 1 + |\varepsilon_i|/c_i \le 1 + 2^{-p}$. Therefore, we get $\varepsilon' \ge \varepsilon/(1 + 2^{-p})^{\ell \cdot m \cdot q_s}$. We select parameters to get adversaries with success probabilities $\ge \varepsilon/2$ against the untruncated scheme and set the precision so that $(1 + 2^{-p})^{\ell \cdot m \cdot q_s} \le 2$. This yields an approximated precision requirement

$$p \ge \log(\ell \cdot m \cdot q_s). \tag{5}$$

Note above estimate may not be accurate unless $\ell \cdot m \cdot q_s$ is much smaller than 2^p. Hence we may also require that $p \ge \log(\ell \cdot m \cdot q_s) + C$ for some constant C. This condition essentially eliminates the term ε from the precision needed by the SD-based and KLD-based analyses.[3] Overall, we get a precomputed table of bit-size $L_{\mathrm{RD}} = \log(\ell \cdot m \cdot q_s) \cdot \ell$.

R_a-based analysis. We may also consider R_a-based analysis for general $a > 1$. It should noted that the reductions here are not tight: for R_a-based analysis with $a > 1$, the probability preservation shows $\varepsilon' > \varepsilon^{a/(a-1)}/R_a(\Phi \| \Phi')$. The Rényi

[3] Note that the resulting precision is not independent of ε. The parameters $m = m(\varepsilon)$ and $\ell = \ell(\varepsilon)$ are chosen in [DDLL13] so that any forging adversary has success probability at most ε on the scheme implemented with perfect Gaussian sampling.

divergence can be computed by

$$(R_a(\Phi\|\Phi'))^{a-1} = \frac{(1-c_i-\varepsilon_i)^a}{(1-c_i)^{a-1}} + \frac{(c_i+\varepsilon_i)^a}{c_i^{a-1}}$$
$$= (1-c_i-\varepsilon_i)\left(1-\frac{\varepsilon_i}{1-c_i}\right)^{a-1} + (c_i+\varepsilon_i)\left(1+\frac{\varepsilon_i}{c_i}\right)^{a-1}.$$

If a is much smaller than 2^p, we get

$$(R_a(\Phi\|\Phi'))^{a-1} \approx (1-c_i-\varepsilon_i)\left(1-\frac{(a-1)\varepsilon_i}{1-c_i}\right) + (c_i+\varepsilon_i)\left(1+\frac{(a-1)\varepsilon_i}{c_i}\right)$$
$$= 1 + \frac{\varepsilon_i^2(a-1)}{c_i(1-c_i)} \le 1 + 2^{-2p}(a-1)\frac{c_i}{1-c_i} \le 1 + 2^{-2p}(a-1).$$

For instance if we take $a = 2$, we have $R_2(\Phi\|\Phi') \le 1 + 2^{-2p}$ and hence $\varepsilon' \ge \varepsilon^2/R_2(\Phi\|\Phi')$. To get a success probability lower bound $\varepsilon^2/2$, it is sufficient to set

$$p \ge \frac{1}{2}\log(\ell \cdot m \cdot q_s). \tag{6}$$

On the other hand, if a is much larger than 2^p, then we have

$$(R_a(\Phi\|\Phi'))^{a-1} = (1-c_i-\varepsilon_i)\left(1-\frac{\varepsilon_i}{1-c_i}\right)^{a-1} + (c_i+\varepsilon_i)\left(1+\frac{\varepsilon_i}{c_i}\right)^{a-1}$$
$$\approx (c_i+\varepsilon_i)\exp\left(\frac{(a-1)\varepsilon_i}{c_i}\right).$$

Hence the Rényi divergence

$$R_a(\Phi\|\Phi') \approx (c_i+\varepsilon_i)^{1/(a-1)}\exp\left(\frac{\varepsilon_i}{c_i}\right) \approx 1 + \frac{\varepsilon_i}{c_i}.$$

As $a \to \infty$, $R_a(\Phi\|\Phi') \to 1 + 2^{-p}$.

Thus if the tightness of the reduction is not a concern, using R_a with small a reduces the precision requirement. Furthermore, we can amplify the success probability of the forger on the truncated Gaussian from ε' to some $\varepsilon'' > \varepsilon'$.

Numerical Examples. In Table 1, we consider a numerical example which gives the lower bound on the precision for the scheme BLISS-I (with $\varepsilon = 2^{-128}$) when allowing up to $q_s = 2^{64}$ sign queries to the adversary. For the BLISS-I parameters, we use $m = 1024$, $\ell = 29$, $s = \lceil\sqrt{2\pi}\cdot 254\cdot\sqrt{1/(2\ln 2)}\rceil = 541$ and $\tau = 13.4/\sqrt{2\pi} \approx 5.4$). The reductions in the table are tight, except for R_2 (as ε' in the reduction does not depend directly on ε but on ε^2), and we are a little bit loose for the R_∞ case.

When instantiating BLISS-I with the parameters of Table 1, the table bit-size can be reduced from about 6000 bits to about 1200 bits by using R_2 in place of SD. If the tightness of the reduction is concerned, we may use R_∞ instead, which leads to a table of about 2300 bits.

Table 1. Comparison of the precision to handle adversaries with success probability $\geq \varepsilon$ making $\leq q_s$ sign queries to BLISS-I. Our Rényi-based parameters are on the last two lines.

	Lower bound on the precision p	Example p	Table bit-size
SD (Eq. (1))	$p \geq \log(\ell \cdot m \cdot q_s/\varepsilon)$	$p \geq 207$	6003
SD (Eq. (2))	$p_i \geq \log(\ell \cdot m \cdot q_s \cdot e^{-\pi 2^i/s^2}/\varepsilon) + 1$	–	4598
KLD (Eq. (3))	$p \geq \log(\ell \cdot m \cdot q_s/\varepsilon^2)/2 + 1/2$	$p \geq 168$	4872
KLD (Eq. (4))	$p_i \geq \log(\ell \cdot m \cdot q_s/\varepsilon^2 \cdot c_i/(1-c_i))/2 + 1$	–	3893
R_∞ (Eq. (5))	$p \geq \log(\ell \cdot m \cdot q_s)$	$p \geq 79$	2291
R_2 (Eq. (6))	$p \geq \log(\ell \cdot m \cdot q_s)/2$	$p \geq 40$	1160

3.2 GPV Signature Scheme

The RD can also be used to reduce the parameters obtained via the SD-based analysis of the GPV signature scheme in [GPV08].

In summary, the signature and the security proof from [GPV08] work as follows. The signature public key is a matrix $A \in \mathbb{Z}_q^{n \times m}$ with n linear in the security parameter λ, $q = \text{poly}(n)$, and $m = O(n \log q)$. The private signing key is a short basis matrix T for the lattice $A^\perp = \{\boldsymbol{x} \in \mathbb{Z}^m : A \cdot \boldsymbol{x} = \boldsymbol{0} \mod q\}$, whose last successive minimum satisfies $\lambda_m(A^\perp) \leq O(1)$ when $m = \Omega(n \log q)$ (see [GPV08]). A signature $(\boldsymbol{\sigma}, s)$ on a message M is a short vector $\boldsymbol{\sigma} \in \mathbb{Z}^m$ and a random salt $s \in \{0,1\}^\lambda$, such that $A \cdot \boldsymbol{\sigma} = H(M, s) \mod q$, where H is a random oracle hashing into $\mathbb{Z}_q^n$. The short vector $\boldsymbol{\sigma}$ is sampled by computing an arbitrary vector $\boldsymbol{t}$ satisfying $A \cdot \boldsymbol{t} = H(M, s) \mod q$ and using T along with a Gaussian sampling algorithm (see [GPV08, BLP+13]) to produce a sample from $\boldsymbol{t} + D_{A^\perp, r, -\boldsymbol{t}}$.

The main idea in the security proof from the SIS problem [GPV08] is based on simulating signatures without T, by sampling $\boldsymbol{\sigma}$ from $D_{\mathbb{Z}^m, r}$ and then programming the random oracle H at (M, s) according to $H(M, s) = A \cdot \boldsymbol{\sigma} \mod q$. As shown in [GPV08, Lemma 5.2], the conditional distribution of $\boldsymbol{\sigma}$ given $A \cdot \boldsymbol{\sigma} \mod q$ is exactly the same in the simulation and in the real scheme. Therefore, the SD between the simulated signatures and the real signatures is bounded by the SD between the marginal distribution D_1 of $A \cdot \boldsymbol{\sigma} \mod q$ for $\boldsymbol{\sigma} \hookleftarrow D_{\mathbb{Z}^m, r}$ and $U(\mathbb{Z}_q^m)$. This SD for one signature is bounded by ε if $r \geq \eta_\varepsilon(A^\perp)$. This leads, over the q_s sign queries of the attacker, in the SD-based analysis of [GPV08], to take $\varepsilon = O(2^{-\lambda} q_s^{-1})$ and thus $r = \Omega(\sqrt{\lambda + \log q_s})$ (using Lemma 2.2), in order to handle attackers with success probability $2^{-o(\lambda)}$.

Now, by Lemma 2.9, we have that the RD $R_\infty(D_1 \| U)$ is bounded by $1 + c \cdot \varepsilon$ for one signature, for some constant c. By the multiplicativity property of Lemma 2.7, over q_s queries, it is bounded by $(1 + c\varepsilon)^{q_s}$. By taking $\varepsilon = O(q_s^{-1})$, we obtain overall an RD bounded as $O(1)$ between the view of the attacker in the real attack and simulation, leading to a security proof with respect to SIS but with a smaller $r = \Omega(\sqrt{\log \lambda + \log(nq_s)})$. When the number of sign queries q_s allowed

to the adversary is much smaller than 2^λ, this leads to significant parameter savings, because SIS's β is reduced and hence n, m, q may be set smaller for the same security parameter λ.

4 Rényi Divergence and Distinguishing Problems

In this section, we prove Theorem 4.1 which allows to use the RD for distinguishing problems, and we show how to apply it to the dual-Regev encryption scheme.

4.1 Problems with Public Sampleability

A general setting one comes across in analyzing the security of cryptographic schemes has the following form. Let P denote a decision problem that asks to distinguish whether a given x was sampled from distribution X_0 or X_1, defined as follows:

$$X_0 = \{x : r \hookleftarrow \Phi, x \hookleftarrow D_0(r)\}, \quad X_1 = \{x : r \hookleftarrow \Phi, x \hookleftarrow D_1(r)\}.$$

Here r is some parameter that is sampled from the same distribution Φ in both X_0 and X_1. The parameter r then determines the conditional distributions $D_0(r)$ and $D_1(r)$ from which x is sampled in X_0 and X_1, respectively, given r. Now, let P' denote another decision problem that is defined similarly to P, except that in P' the parameter r is sampled from a different distribution Φ' (rather than Φ). Given r, the conditional distributions $D_0(r)$ and $D_1(r)$ are the same in P' as in P. Let X'_0 (resp. X'_1) denote the resulting marginal distributions of x in problem P'. Now, in the applications we have in mind, the distributions Φ' and Φ are "close" in some sense, and we wish to show that this implies an efficient reduction between problems P' and P, in the usual sense that every distinguisher with efficient run-time T and non-negligible advantage ε against P implies a distinguisher for P' with efficient run-time T' and non-negligible advantage ε'. In the classical situation, if the SD $\Delta(\Phi, \Phi')$ between Φ' and Φ is negligible, then the reduction is immediate. Indeed, for $b \in \{0, 1\}$, if p_b (resp. p'_b) denotes the probability that a distinguisher algorithm $\mathcal{A}$ outputs 1 on input distribution X_b (resp. X'_b), then we have, from the SD probability preservation property, that $|p'_b - p_b| \leq \Delta(\Phi, \Phi')$. As a result, the advantage $\varepsilon' = |p'_1 - p'_0|$ of $\mathcal{A}$ against P' is bounded from below by $\varepsilon - 2\Delta(\Phi, \Phi')$ which is non-negligible (here $\varepsilon = |p_1 - p_0|$ is the assumed non-negligible advantage of $\mathcal{A}$ against P).

Unfortunately, for general decision problems P, P' of the above form, it seems difficult to obtain an RD-based analogue of the above SD-based argument, in the weaker setting when the SD $\Delta(\Phi, \Phi')$ is non-negligible, but the RD $R = R(\Phi \| \Phi')$ is small. Indeed, the probability preservation property of the RD in Lemma 2.7 does not seem immediately useful in the case of general decision problems P, P'. With the above notations, it can be used to conclude that $p'_b \geq p_b^2/R$ but this does not allow us to usefully relate the advantages $|p'_1 - p'_0|$ and $|p_1 - p_0|$.

Nevertheless, we now make explicit a special class of "publicly sampleable" problems P, P' for which such a reduction can be made. In such problems, it is possible to efficiently sample from both distributions $D_0(r)$ (resp. $D_1(r)$) given the single sample x from the unknown $D_b(r)$. This technique is implicit in the application of RD in the reductions of [LPR13], and we abstract it and make it explicit in the following.

Theorem 4.1. *Let Φ, Φ' denote two distributions with* $\mathrm{Supp}(\Phi) \subseteq \mathrm{Supp}(\Phi')$*, and $D_0(r)$ and $D_1(r)$ denote two distributions determined by some parameter $r \in \mathrm{Supp}(\Phi')$. Let P, P' be two decision problems defined as follows:*

- *Problem P: Distinguish whether input x is sampled from distribution X_0 or X_1, where*

$$X_0 = \{x : r \hookleftarrow \Phi, x \hookleftarrow D_0(r)\}, \quad X_1 = \{x : r \hookleftarrow \Phi, x \hookleftarrow D_1(r)\}.$$

- *Problem P': Distinguish whether input x is sampled from distribution X'_0 or X'_1, where*

$$X'_0 = \{x : r \hookleftarrow \Phi', x \hookleftarrow D_0(r)\}, \quad X'_1 = \{x : r \hookleftarrow \Phi', x \hookleftarrow D_1(r)\}.$$

Assume that $D_0(\cdot)$ and $D_1(\cdot)$ satisfy the following public sampleability *property: there exists a sampling algorithm S with run-time T_S such that for all (r, b), given any sample x from $D_b(r)$:*

- *$\mathsf{S}(0, x)$ outputs a fresh sample distributed as $D_0(r)$ over the randomness of S,*
- *$\mathsf{S}(1, x)$ outputs a fresh sample distributed as $D_1(r)$ over the randomness of S.*

Then, given a T-time distinguisher $\mathcal{A}$ for problem P with advantage ε, we can construct a distinguisher $\mathcal{A}'$ for problem P' with run-time and distinguishing advantage respectively bounded from above and below by (for any $a \in (1, +\infty]$):

$$O\left(\frac{1}{\varepsilon^2}\log\left(\frac{R_a(\Phi\|\Phi')}{\varepsilon^{a/(a-1)}}\right)\cdot(T_S + T)\right) \text{ and } \frac{\varepsilon}{4 \cdot R_a(\Phi\|\Phi')}\cdot\left(\frac{\varepsilon}{2}\right)^{\frac{a}{a-1}}.$$

Proof. Distinguisher $\mathcal{A}'$ is given an input x sampled from $D_b(r)$ for some r sampled from Φ' and some unknown $b \in \{0, 1\}$. For an ε' to be determined later, it runs distinguisher $\mathcal{A}$ on $N = O(\varepsilon^{-2}\log(1/\varepsilon'))$ independent inputs sampled from $D_0(r)$ and $D_1(r)$ calling algorithm S on $(0, x)$ and $(1, x)$ to obtain estimates $\hat{p}_0$ and $\hat{p}_1$ for the acceptance probabilities $p_0(r)$ and $p_1(r)$ of $\mathcal{A}$ given as inputs samples from $D_0(r)$ and $D_1(r)$ (with the r fixed to the value used to sample the input x of $\mathcal{A}'$). By the choice of N and the Hoeffding bound, the estimation errors $|\hat{p}_0 - p_0|$ and $|\hat{p}_1 - p_1|$ are $< \varepsilon/8$ except with probability $< \varepsilon'$ over the randomness of S. Then, if $\hat{p}_1 - \hat{p}_0 > \varepsilon/4$, distinguisher $\mathcal{A}'$ runs $\mathcal{A}$ on input x and returns whatever $\mathcal{A}$ returns, else distinguisher $\mathcal{A}'$ returns a uniformly random bit. This completes the description of distinguisher $\mathcal{A}'$.

Let $\mathcal{S}_1$ denote the set of r's such that $p_1(r) - p_0(r) \geq \varepsilon/2$, $\mathcal{S}_2$ denote the set of r's that are not in $\mathcal{S}_1$ and such that $p_1(r) - p_0(r) \geq 0$, and $\mathcal{S}_3$ denote all the remaining r's. Then:

- If $r \in \mathcal{S}_1$, then except with probability $< \varepsilon'$ over the randomness of S, we will have $\hat{p}_1 - \hat{p}_0 > \varepsilon/4$ and thus $\mathcal{A}'$ will output $\mathcal{A}(x)$. Thus, in the case $b = 1$, we have $\Pr[\mathcal{A}'(x) = 1 | r \in \mathcal{S}_1] \geq p_1(r) - \varepsilon'$ and in the case $b = 0$, we have $\Pr[\mathcal{A}'(x) = 1 | r \in \mathcal{S}_1] \leq p_0(r) + \varepsilon'$.
- Assume that $r \in \mathcal{S}_2$. Let $u(r)$ be the probability over the randomness of S that $\hat{p}_1 - \hat{p}_0 > \varepsilon/4$. Then $\mathcal{A}'$ will output $\mathcal{A}(x)$ with probability $u(r)$ and a uniform bit with probability $1 - u(r)$. Thus, in the case $b = 1$, we have $\Pr[\mathcal{A}'(x) = 1 | r \in \mathcal{S}_2] = u(r) \cdot p_1(r) + (1 - u(r))/2$, and in the case $b = 0$, we have $\Pr[\mathcal{A}'(x) = 1 | r \in \mathcal{S}_2] = u(r) \cdot p_0(r) + (1 - u(r))/2$.
- If $r \in \mathcal{S}_3$, except with probability $< \varepsilon'$ over the randomness of S, we have $\hat{p}_1 - \hat{p}_0 < \varepsilon/4$ and $\mathcal{A}'$ will output a uniform bit. Thus, in the case $b = 1$, we have $\Pr[\mathcal{A}'(x) = 1 | r \in \mathcal{S}_3] \geq 1/2 - \varepsilon'$, and in the case $b = 0$, we have $\Pr[\mathcal{A}'(x) = 1 | r \in \mathcal{S}_3] \leq 1/2 + \varepsilon'$.

Overall, the advantage of $\mathcal{A}'$ is bounded from below by:

$$\sum_{r \in \mathcal{S}_1} \Phi'(r) \left(p_1(r) - p_0(r) - 2\varepsilon'\right) + \sum_{r \in \mathcal{S}_2} \Phi'(r) u(r) \left(p_1(r) - p_0(r)\right) - \sum_{r \in \mathcal{S}_3} \Phi'(r) 2\varepsilon'$$
$$\geq \Phi'(\mathcal{S}_1) \cdot \frac{\varepsilon}{2} - 2\varepsilon'.$$

Without loss of generality, we may assume that the advantage of $\mathcal{A}$ is positive. By an averaging argument, the set $\mathcal{S}_1$ has probability $\Phi(\mathcal{S}_1) \geq \varepsilon/2$ under distribution Φ. Hence, by the RD probability preservation property (see Lemma 2.7), we have $\Phi'(\mathcal{S}_1) \geq (\varepsilon/2)^{\frac{a}{a-1}} / R_a(\Phi \| \Phi')$. The proof may be completed by setting $\varepsilon' = (\varepsilon/8) \cdot (\varepsilon/2)^{\frac{a}{a-1}} / R_a(\Phi \| \Phi')$. □

4.2 Application to Dual-Regev Encryption

Let m, n, q, χ be as in Definition 2.5 and Φ denote a distribution over $\mathbb{Z}_q^{m \times n}$. We define the LWE variant $\mathrm{LWE}_{n,q,\chi,m}(\Phi)$ as follows: Sample $A \hookleftarrow \Phi$, $\boldsymbol{s} \hookleftarrow U(\mathbb{Z}_q^n)$, $\boldsymbol{e} \hookleftarrow \chi^m$ and $\boldsymbol{u} \hookleftarrow U(\mathbb{T}^m)$; The goal is to distinguish between the distributions $(A, \frac{1}{q} A\boldsymbol{s} + \boldsymbol{e})$ and $(A, \boldsymbol{u})$ over $\mathbb{Z}_q^{m \times n} \times \mathbb{T}^m$. Note that standard LWE is obtained by taking $\Phi' = U(\mathbb{Z}_q^{m \times n})$.

As an application to Theorem 4.1, we show that LWE with non-uniform and possibly statistically correlated $\boldsymbol{a}_i$'s of the samples $(\boldsymbol{a}_i, b_i)$'s (with b_i either independently sampled from $U(\mathbb{T})$ or close to $\langle \boldsymbol{a}_i, \boldsymbol{s} \rangle$ for a secret vector $\boldsymbol{s}$) remains at least as hard as standard LWE, as long as the RD $R(\Phi \| U)$ remains small, where Φ is the joint distribution of the given $\boldsymbol{a}_i$'s and U denotes the uniform distribution.

To show this result, we first prove in Corollary 4.2 that there is a reduction from $\mathrm{LWE}_{n,q,\chi,m}(\Phi')$ to $\mathrm{LWE}_{n,q,\chi,m}(\Phi)$ using Theorem 4.1 if $R_a(\Phi \| \Phi')$ is small enough. We then describe in Corollary 4.3 how to use this first reduction to obtain smaller parameters for the dual-Regev encryption. This allows us to save an $\Omega(\sqrt{\lambda / \log \lambda})$ factor in the Gaussian deviation parameter r used for secret key generation in the dual-Regev encryption scheme [GPV08], where λ refers to the security parameter.

Corollary 4.2. *Let Φ and Φ' be two distributions over $\mathbb{Z}_q^{m\times n}$ with* $\mathrm{Supp}(\Phi) \subseteq \mathrm{Supp}(\Phi')$. *If there exists a distinguisher $\mathcal{A}$ against* $\mathrm{LWE}_{n,q,\chi,m}(\Phi)$ *with run-time T and advantage $\varepsilon = o(1)$, then there exists a distinguisher $\mathcal{A}'$ against* $\mathrm{LWE}_{n,q,\chi,m}(\Phi')$ *with run-time* $T' = O(\varepsilon^{-2}\log\frac{R_a(\Phi\|\Phi')}{\varepsilon^{a/(a-1)}}\cdot(T+\mathrm{poly}(m,\log q)))$ *and advantage* $\Omega\left(\frac{\varepsilon^{1+a/(a-1)}}{R_a(\Phi\|\Phi')}\right)$, *for any* $a\in(1,+\infty]$.

Proof. Apply Theorem 4.1 with $r = A \in \mathbb{Z}_q^m$, $x = (A, \boldsymbol{b}) \in \mathbb{Z}_q^{m\times n}\times\mathbb{T}^m$, $D_0(r) = (A, A\cdot\boldsymbol{s}+\boldsymbol{e})$ with $\boldsymbol{s}\hookleftarrow U(\mathbb{Z}_q^n)$ and $\boldsymbol{e}\hookleftarrow\chi^m$, and $D_1(r) = (A,\boldsymbol{u})$ with $\boldsymbol{u}\hookleftarrow U(\mathbb{Z}_q^m)$. The sampling algorithm S is such that $\mathsf{S}(0,x)$ outputs $(A, A\cdot\boldsymbol{s}'+\boldsymbol{e}')$ for $\boldsymbol{s}'\hookleftarrow U(\mathbb{Z}_q^n)$ and $\boldsymbol{e}'\hookleftarrow\chi^m$, while $\mathsf{S}(1,x)$ outputs $(A,\boldsymbol{u}')$ with $\boldsymbol{u}'\hookleftarrow U(\mathbb{Z}_q^m)$. □

We recall that the dual-Regev encryption scheme has a general public parameter $A\in\mathbb{Z}_q^{m\times n}$, a secret key of the form $sk = \boldsymbol{x}$ with $\boldsymbol{x}\hookleftarrow D_{\mathbb{Z}^m,r}$ and a public key of the form $\boldsymbol{u} = A^t\boldsymbol{x} \mod q$. A ciphertext for a message $M\in\{0,1\}$ is obtained as follows: Sample $\boldsymbol{s}\hookleftarrow U(\mathbb{Z}_q^n)$, $\boldsymbol{e}_1\hookleftarrow\chi^m$ and $e_2\hookleftarrow\chi$; return ciphertext $(c_1,c_2) = (\frac{1}{q}A\boldsymbol{s}+\boldsymbol{e}_1, \frac{1}{q}\langle\boldsymbol{u},\boldsymbol{s}\rangle + e_2 + \frac{M}{2}) \in \mathbb{T}^m\times\mathbb{T}$.

Corollary 4.3. *Suppose that q is prime, $m\geq 2n\log q$ and $r\geq 4\sqrt{\log(12m)/\pi}$. If there exists an adversary against the IND-CPA security of the dual-Regev encryption scheme with run-time T and advantage ε, then there exists a distinguishing algorithm for* $\mathrm{LWE}_{n,q,\chi,m+1}$ *with run-time* $O((\varepsilon')^{-2}\log(\varepsilon')^{-1}\cdot(T+\mathrm{poly}(m)))$ *and advantage* $\Omega((\varepsilon')^2)$, *where* $\varepsilon' = \varepsilon - 2q^{-n}$.

Proof. The IND-CPA security of the dual-Regev encryption scheme as described above is at least as hard as $\mathrm{LWE}_{n,q,\chi,m+1}(\Phi)$ where Φ is obtained by sampling $A\hookleftarrow U(\mathbb{Z}_q^{m\times n})$, $\boldsymbol{u}\hookleftarrow A^t\cdot D_{\mathbb{Z}^m,r} \mod q$ and returning the $(m+1)\times n$ matrix obtained by appending $\boldsymbol{u}^t$ at the bottom of A. We apply Corollary 4.2 with $\Phi' = U(\mathbb{Z}_q^{(m+1)\times n})$.

Since q is prime, if A is full rank, then the multiplication by A^t induces an isomorphism between the quotient group $\mathbb{Z}^m/A^\perp$ and $\mathbb{Z}_q^n$, where $A^\perp = \{\boldsymbol{x}\in\mathbb{Z}^m : A^t\cdot\boldsymbol{x} = \boldsymbol{0} \mod q\}$. By Lemma 2.2, we have $\eta_{1/3}(A^\perp)\leq 4\sqrt{\log(12m)/\pi}\leq r$, except for a fraction $\leq q^{-n}$ of the A's. Let BAD denote the union of such bad A's and the A's that are not full rank. We have $\Pr[BAD]\leq 2q^{-n}$.

By the multiplicativity property of Lemma 2.7, we have:

$$R_\infty(\Phi\|\Phi')\leq \max_{A\notin BAD} R_\infty(D_{\mathbb{Z}^m,r} \bmod A^\perp \| U_{\mathbb{Z}^m/A^\perp}).$$

Thanks to Lemma 2.9, we know that the latter is ≤ 2. The result now follows from Corollary 4.2. □

In all applications we are aware of, the parameters satisfy $m\leq\mathrm{poly}(\lambda)$ and $q^{-n}\leq 2^{-\lambda}$, where λ refers to the security parameter. The $r=\Omega(\sqrt{\log\lambda})$ bound of our Corollary 4.3, that results from using $\delta = 1/3$ in the condition $r\geq\eta_\delta(A^\perp)$ in the RD-based smoothing argument of the proof above, improves on the corresponding bound $r=\Omega(\sqrt{\lambda})$ that results from the requirement to use

$\delta = O(2^{-\lambda})$ in the condition $r \geq \eta_\delta(A^\perp)$ in the SD-based smoothing argument of the proof of [GPV08, Theorem 7.1], in order to handle adversaries with advantage $\varepsilon = 2^{-o(\lambda)}$ in both cases. Thus our RD-based analysis saves a factor $\Omega(\sqrt{\lambda/\log\lambda})$ in the choice of r, and consequently of a^{-1} and q. (The authors of [GPV08] specify a choice of $r = \omega(\sqrt{\log\lambda})$ for their scheme because they use in their analysis the classical "no polynomial attacks" security requirement, corresponding to assuming attacks with advantage $\varepsilon = \lambda^{-O(1)}$, rather than the stronger $\varepsilon = \omega(2^{-\lambda})$ but more realistic setting we take.)

5 Application to LWE with Uniform Noise

The LWE problem with noise uniform in a small interval was introduced in [DMQ13]. In that article, the authors exhibit a reduction from LWE with Gaussian noise, which relies on a new tool called *lossy codes*. The main proof ingredients are the construction of lossy codes for LWE (which are lossy for the uniform distribution in a small interval), and the fact that lossy codes are pseudorandom.

We note that the reduction from [DMQ13] needs the number of LWE samples to be bounded by $\mathrm{poly}(n)$ and that it degrades the LWE dimension by a constant factor. The parameter β (when the interval of the noise is $[-\beta, \beta]$) should be at least $mn^\sigma\alpha$ where α is the LWE Gaussian noise parameter and $\sigma \in (0, 1)$ is an arbitrarily small constant.

We now provide an alternative reduction from the $\mathrm{LWE}_{n,q,D_\alpha,m}$ distinguishing problem to the $\mathrm{LWE}_{n,q,U([-\beta,\beta]),m}$ distinguishing problem, and analyze it using RD. Our reduction preserves the LWE dimension n, and is hence tighter than the one from [DMQ13]. We also require that $\beta = \Omega(m\alpha)$.

Theorem 5.1. *Let $m \geq n \geq 1$ and with $q \leq \mathrm{poly}(m, n)$ prime. Let $\alpha, \beta > 0$ be real numbers with $\beta = \Omega(m\alpha)$. Then there is a polynomial-time reduction from $\mathrm{LWE}_{n,q,D_\alpha,m}$ to $\mathrm{LWE}_{n,q,\phi,m}$, with $\phi = \frac{1}{q}\lfloor qU([-\beta,\beta])\rceil$.*

Proof. In the proof, we let U_β denote the distribution $U([-\beta, \beta])$, to ease notations. Our reduction relies on four steps:

- A reduction from $\mathrm{LWE}_{n,q,D_\alpha,m}$ to $\mathrm{LWE}_{n,q,\psi,m}$ with $\psi = D_\alpha + U_\beta$,
- A reduction from $\mathrm{LWE}_{n,q,\psi,m}$ to $\mathrm{sLWE}_{n,q,\psi,m}$,
- A reduction from $\mathrm{sLWE}_{n,q,\psi,m}$ to $\mathrm{sLWE}_{n,q,U_\beta,m}$,
- A reduction from $\mathrm{sLWE}_{n,q,U_\beta,m}$ to $\mathrm{LWE}_{n,q,U_\beta,m}$.

First step. The reduction is given m elements $(\boldsymbol{a}_i, b_i) \in \mathbb{Z}_q^n \times \mathbb{T}$, all drawn from $A_{\boldsymbol{s},D_\alpha}$ (for some $\boldsymbol{s}$), or all drawn from $U(\mathbb{Z}_q^n \times \mathbb{T})$. The reduction consists in adding independent samples from U_β to each b_i. The reduction maps the uniform distribution to itself, and $A_{\boldsymbol{s},D_\alpha}$ to $A_{\boldsymbol{s},\psi}$.

Second step. Reducing the distinguishing variant of LWE to its search variant is direct.

Third step. The reduction from $\mathrm{sLWE}_{n,q,\psi,m}$ to $\mathrm{sLWE}_{n,q,U_\beta,m}$ is vacuous: by using the RD (and in particular the probability preservation property of Lemma 2.7), we show that an oracle solving $\mathrm{sLWE}_{n,q,U_\beta,m}$ also solves $\mathrm{sLWE}_{n,q,\psi,m}$.

Lemma 5.2. *Let α, β be real numbers with $\alpha \in (0, 1/e)$ and $\beta \geq \alpha$. Let $\psi = D_\alpha + U_\beta$. Then*

$$R_2(U_\beta \| \psi) = \frac{\alpha}{\beta} \int_0^\beta \frac{1}{\int_{-\beta}^{\beta} e^{\frac{-\pi(x-y)^2}{\alpha^2}} dy} dx \;\leq\; 1 + 16\frac{\alpha}{\beta}\sqrt{\ln(1/\alpha)/\pi}.$$

Proof. The density function of ψ is the convolution of the density functions of D_α and U_β:

$$f_\psi(x) = \frac{1}{2\alpha\beta} \int_{-\beta}^{\beta} e^{\frac{-\pi(x-y)^2}{\alpha^2}} dy.$$

Using Rényi of order 2, we have:

$$R_2(U_\beta \| \psi) = \int_{-\beta}^{\beta} \frac{\frac{1}{(2\beta)^2}}{\frac{1}{2\alpha\beta}\int_{-\beta}^{\beta} e^{\frac{-\pi(x-y)^2}{\alpha^2}} dy} dx = \frac{\alpha}{\beta} \int_0^\beta \frac{1}{\int_{-\beta}^{\beta} e^{\frac{-\pi(x-y)^2}{\alpha^2}} dy} dx.$$

The denominator in the integrand is a function for $x \in [0, \beta]$.

$$\phi(x) = \alpha - \int_{\beta+x}^{\infty} \exp(\frac{-\pi y^2}{\alpha^2})\, \mathrm{d}y - \int_{\beta-x}^{\infty} \exp(\frac{-\pi y^2}{\alpha^2})\, \mathrm{d}y.$$

For standard Gaussian, we use the following tail bound [CDS03]:

$$\frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-x^2/2} dx \leq \frac{1}{2} e^{-z^2/2}.$$

Then we have

$$\phi(x) \geq \alpha \left(1 - \frac{1}{2}\exp\left(\frac{-\pi(\beta+x)^2}{\alpha^2}\right) - \frac{1}{2}\exp\left(\frac{-\pi(\beta-x)^2}{\alpha^2}\right)\right).$$

Taking the reciprocal of above, we use the first-order Taylor expansion. Note here

$$t(x) = \frac{1}{2}\exp\left(\frac{-\pi(\beta+x)^2}{\alpha^2}\right) + \frac{1}{2}\exp\left(\frac{-\pi(\beta-x)^2}{\alpha^2}\right).$$

We want to bound the function $t(x)$ by a constant $c \in (0, 1)$. Here $t(x)$ is not monotonic. We take the maximum of the first-half and the maximum of the second-half of $t(x)$. An upper bound ($\beta \geq \alpha$) is:

$$t(x) \leq \frac{1}{2} e^{-\pi\beta^2/\alpha^2} + \frac{1}{2} =: \sigma_{\alpha,\beta} + \frac{1}{2} < 1.$$

We then use the fact that $\frac{1}{1-t(x)} = 1 + \frac{1}{1-t(x)}t(x) \le 1 + \frac{1}{1-2\sigma_{\alpha,\beta}}t(x)$ to bound the Rényi divergence of order 2.

$$\begin{aligned}
R_2(U_\beta\|\psi) &= \frac{\alpha}{\beta}\int_0^\beta \frac{1}{\phi(x)}\mathrm{d}x \\
&\le \frac{1}{\beta}\int_0^\beta \frac{1}{1-\frac{1}{2}\exp\left(\frac{-\pi(\beta+x)^2}{\alpha^2}\right)-\frac{1}{2}\exp\left(\frac{-\pi(\beta-x)^2}{\alpha^2}\right)}\mathrm{d}x \\
&\le \frac{1}{\beta}\int_0^\beta \left(1+\frac{1}{1-2\sigma_{\alpha,\beta}}\exp\left(\frac{-\pi(\beta+x)^2}{\alpha^2}\right)\right. \\
&\quad \left.+\frac{1}{1-2\sigma_{\alpha,\beta}}\exp\left(\frac{-\pi(\beta-x)^2}{\alpha^2}\right)\right)\mathrm{d}x \\
&= 1+\frac{1}{(1-2\sigma_{\alpha,\beta})\beta}\int_0^{2\beta}\exp\left(\frac{-\pi x^2}{\alpha^2}\right)\mathrm{d}x \\
&= 1+\frac{1}{2(1-2\sigma_{\alpha,\beta})\beta}\int_{-2\beta}^{2\beta}\exp\left(\frac{-\pi x^2}{\alpha^2}\right)\mathrm{d}x \\
&= 1+\frac{\alpha}{(1-2\sigma_{\alpha,\beta})\beta}(1-2D_\alpha(2\beta)) \le 1+\frac{1}{1-2\sigma_{\alpha,\beta}}\frac{\alpha}{\beta}.
\end{aligned}$$

Hence we have the bound,

$$R_2(U_\beta\|\psi) \le 1+\frac{1}{1-e^{-\pi\beta^2/\alpha^2}}\frac{\alpha}{\beta}.$$ □

We use Lemma 5.2 with m samples and $\beta = \Omega(m\alpha)$ to ensure that the mth power of the RD is ≤ 2. The RD multiplicativity and probability preservation properties (see Lemma 2.7) imply that $\varepsilon' \ge \varepsilon^2/R_2^m(U_\beta\|\phi)$; hence if an oracle solves $\mathrm{sLWE}_{n,q,U_\beta,m}$ with probability ε, then it also solves $\mathrm{sLWE}_{n,q,\psi,m}$ with probability $\ge \varepsilon^2/2$.

Fourth step. We reduce $\mathrm{sLWE}_{n,q,U_\beta,m}$ with continuous noise U_β to $\mathrm{sLWE}_{n,q,\phi,m}$ with discrete noise $\phi = \frac{1}{q}\lfloor qU_\beta \rceil$ with support contained in $\mathbb{T}_q$, by rounding to the nearest multiple of $\frac{1}{q}$ any provided b_i (for $i \le m$). We reduce $\mathrm{sLWE}_{n,q,\phi,m}$ to $\mathrm{LWE}_{n,q,\phi,m}$ by invoking Theorem 2.6. □

6 Open Problems

Our results show the utility of the Rényi divergence in several areas of lattice-based cryptography. However, they also suggest some natural open problems, whose resolution could open up further applications. In particular, can we extend the applicability of RD to more general distinguishing problems than those satisfying our 'public sampleability' requirement? This may extend our results further. For instance, can we use RD-based arguments to prove the hardness of LWE with uniform noise without using the search to decision reduction of [MM11]? This may allow the proof to apply also to Ring-LWE with uniform noise.

Acknowledgments. We thank Léo Ducas, Vadim Lyubashevsky and Fabrice Mouhartem for useful discussions. This work has been supported in part by ERC Starting Grant ERC-2013-StG-335086-LATTAC, an Australian Research Fellowship (ARF) from the Australian Research Council (ARC), and ARC Discovery Grants DP0987734, DP110100628 and DP150100285. This work has been supported in part by the European Union's H2020 Programme under grant agreement number ICT-644209.

References

[Ajt96] Ajtai, M.: Generating hard instances of lattice problems (extended abstract). In: Proceedings of of STOC, pp. 99–108. ACM (1996)

[AKPW13] Alwen, J., Krenn, S., Pietrzak, K., Wichs, D.: Learning with rounding, revisited. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 57–74. Springer, Heidelberg (2013)

[BGM+15] Bogdanov, A., Guo, S., Masny, D., Richelson, S., Rosen, A.: On the hardness of learning with rounding over small modulus. Cryptology ePrint Archive, Report 2015/769 (2015). http://eprint.iacr.org/

[BLP+13] Brakerski, Z., Langlois, A., Peikert, C., Regev, O., Stehlé, D.: Classical hardness of learning with errors. In: Procedings of STOC, pp. 575–584. ACM (2013)

[BPR12] Banerjee, A., Peikert, C., Rosen, A.: Pseudorandom functions and lattices. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 719–737. Springer, Heidelberg (2012)

[BV11] Brakerski, Z., Vaikuntanathan, V.: Efficient fully homomorphic encryption from (standard) LWE, In: Proceedings of FOCS, pp. 97–106. IEEE Computer Society Press (2011)

[CDS03] Chiani, M., Dardari, D., Simon, M.K.: New exponential bounds and approximations for the computation of error probability in fading channels. IEEE Trans. Wireless. Comm. **2**(4), 840–845 (2003)

[DDLL13] Ducas, L., Durmus, A., Lepoint, T., Lyubashevsky, V.: Lattice signatures and bimodal gaussians. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 40–56. Springer, Heidelberg (2013)

[DMQ13] Döttling, N., Müller-Quade, J.: Lossy codes and a new variant of the learning-with-errors problem. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 18–34. Springer, Heidelberg (2013)

[Duc14] Ducas, L.: Accelerating Bliss: the geometry of ternary polynomials. Cryptology ePrint Archive, Report 2014/874 (2014). http://eprint.iacr.org/

[EH12] van Erven, T., Harremoës, P.: Rényi divergence and Kullback-Leibler divergence. CoRR, abs/1206.2459 (2012)

[GGH13] Garg, S., Gentry, C., Halevi, S.: Candidate multilinear maps from ideal lattices. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 1–17. Springer, Heidelberg (2013)

[GPV08] Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: Proceedings of STOC, pp. 197–206. ACM (2008)

[LPR13] Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. J. ACM **60**(6), 43 (2013)

[LPSS14] Ling, S., Phan, D.H., Stehlé, D., Steinfeld, R.: Hardness of k-LWE and applications in traitor tracing. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014, Part I. LNCS, vol. 8616, pp. 315–334. Springer, Heidelberg (2014)

[LSS14] Langlois, A., Stehlé, D., Steinfeld, R.: GGHLite: more efficient multilinear maps from ideal lattices. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 239–256. Springer, Heidelberg (2014)

[Lyu12] Lyubashevsky, V.: Lattice signatures without trapdoors. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 738–755. Springer, Heidelberg (2012)

[MM11] Micciancio, D., Mol, P.: Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 465–484. Springer, Heidelberg (2011)

[MP13] Micciancio, D., Peikert, C.: Hardness of SIS and LWE with small parameters. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 21–39. Springer, Heidelberg (2013)

[MR07] Micciancio, D., Regev, O.: Worst-case to average-case reductions based on Gaussian measures. SIAM J. Comput. **37**(1), 267–302 (2007)

[MR09] Micciancio, D., Regev, O.: Lattice-based cryptography. In: Bernstein, D.J., Buchmann, J., Dahmen, E. (Eds), Post-Quantum Cryptography, pp. 147–191. Springer, Heidelberg (2009)

[PDG14] Pöppelmann, T., Ducas, L., Güneysu, T.: Enhanced lattice-based signatures on reconfigurable hardware. In: Batina, L., Robshaw, M. (eds.) CHES 2014. LNCS, vol. 8731, pp. 353–370. Springer, Heidelberg (2014)

[Pei09] Peikert, C.: Public-key cryptosystems from the worst-case shortest vector problem. In: Proceedings of STOC, pp. 333–342. ACM (2009)

[Reg05] Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. In: Proceedings of STOC, pp. 84–93 (2005)

[Reg09a] Regev, O.: Lecture notes of lattices in computer science, taught at the Computer Science Tel Aviv University, (2009). http://www.cims.nyu.edu/regev

[Reg09b] Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. J. ACM **56**(6), 1–40 (2009)

[Rén61] Rényi, A.: On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 547–561 (1961)