Improved Security Evaluation Techniques for Imperfect Randomness from Arbitrary Distributions

Takahiro Matsuda[1], Kenta Takahashi[2],
Takao Murakami[1], and Goichiro Hanaoka[1]

[1] National Institute of Advanced Industrial Science and Technology (AIST),
Tokyo, Japan
{t-matsuda,takao-murakami,hanaoka-goichiro}@aist.go.jp
[2] Hitachi, Ltd., Yokohama, Japan
kenta.takahashi.bw@hitachi.com

Abstract. Dodis and Yu (TCC 2013) studied how the security of cryptographic primitives that are secure in the "ideal" model in which the distribution of a randomness is the uniform distribution, is degraded when the ideal distribution of a randomness is switched to a "real-world" (possibly biased) distribution that has some lowerbound on its min-entropy or collision-entropy. However, in many constructions, their security is guaranteed only when a randomness is sampled from some non-uniform distribution (such as Gaussian in lattice-based cryptography), in which case we cannot directly apply the results by Dodis and Yu.
In this paper, we generalize the results by Dodis and Yu using the *Rényi divergence*, and show how the security of a cryptographic primitive whose security is guaranteed when the ideal distribution of a randomness is a general (possibly non-uniform) distribution Q, is degraded when the distribution is switched to another (real-world) distribution R. More specifically, we derive two general inequalities regarding the Rényi divergence of R from Q and an adversary's advantage against the security of a cryptographic primitive. As applications of our results, we show (1) an improved reduction for switching the distributions of distinguishing problems with public samplability, which is simpler and much tighter than the reduction by Bai et al. (ASIACRYPT 2015), and (2) how the differential privacy of a mechanism is degraded when its randomness comes from not an ideal distribution Q but a real-world distribution R. Finally, we show methods for approximate-sampling from an arbitrary distribution Q with some guaranteed upperbound on the Rényi divergence (of the distribution R of our sampling methods from Q).

Keywords: Rényi divergence, security evaluation, security reduction.

1 Introduction

1.1 Background and Motivation

Most cryptographic primitives such as encryption and signature schemes, are defined using a probabilistic algorithm that internally generates and uses ran-

domness, and their security is typically defined and analyzed assuming that the randomness used by the algorithm is sampled from some pre-determined "ideal" distribution. Let us call it an ideal model. For example, in the case of encryption and signature schemes, their key generation algorithm is typically defined as a probabilistic algorithm that takes a randomness chosen from the uniform distribution as input, and we evaluate their security by estimating the probability of any possible adversary (with some resource constraint, e.g. running time, memory size, the number of oracle queries) violating the security of the considered schemes is sufficiently small. However, randomness available in the real world may not necessarily come from the ideal distribution with which the security of cryptographic primitives is analyzed. It is often the case that randomness used for generating some secret parameter (such as a secret key) could be biased and/or estimating its exact distribution could be difficult, for example, a situation of using randomness generated based on some physical phenomena (radiation, thermal noise, etc.) [5], a situation in which its partial information is possibly leaked, or a situation of using randomness generated from biometric information [8], to name a few. Even if a cryptographic primitive is guaranteed to be secure in the ideal model via a formal security proof, the security of the primitive is no longer guaranteed in the real world when such a "real-world" randomness is used.

Regarding such "ideal" vs. "real-world" randomness problem, Dodis and Yu [11] studied how the security of a cryptographic primitive in the ideal model where the distribution of its randomness is the uniform distribution $\mathcal{U}$, is degraded when the distribution is switched to another ("real-world") possibly biased distribution $\mathcal{R}$. In particular, they showed that for all cryptographic primitives categorized as *unpredictability applications* (e.g. one-way functions, message authentication codes, and signature schemes) and for some (but not all) cryptographic primitives categorized as *indistinguishability applications* satisfying the so-called "square-friendly" property [4, 9, 11] (e.g. pseudorandom functions and IND-CPA secure encryption schemes), their security is not totally lost even if the distribution of a randomness is switched to a real-world distribution $\mathcal{R}$ that satisfies some entropy criteria. More specifically, Dodis and Yu showed two inequalities that show how an adversary's advantage against the security of a cryptographic primitive could increase when the min-entropy or collision-entropy of the real-world distribution $\mathcal{R}$ is decreased, compared to the ideal model in which its distribution is the uniform distribution $\mathcal{U}$ and has the maximum entropy.

However, an ideal distribution, which we denote by $\mathcal{Q}$ throughout this paper, of a randomness used by cryptographic primitives is in general not necessarily the uniform distribution. For example, there are constructions in lattice-based cryptography in which a secret key is sampled from the (discrete) Gaussian distribution (e.g. [16, 2]), and randomness (a noise vector) used in the encryption procedure is chosen according to a biased distribution so that 0 appears more often than other values (e.g. [15]). When implementing these constructions in practice, again the real-world distribution $\mathcal{R}$ of a randomness may not necessarily follow the ideal distribution $\mathcal{Q}$. However, for these constructions, we cannot

directly apply the results by Dodis and Yu [11], since their results are restricted to the case in which the ideal distribution $\mathcal{Q}$ of a randomness is the uniform distribution.

The main motivation of our work is to generalize and extend the results by Dodis and Yu [11], so that we can apply the analogues of their results to a wider class of distributions as the ideal distribution $\mathcal{Q}$.

1.2 Our Results

As mentioned above, we generalize and extend the results by Dodis and Yu [11] so that the analogues of their results can be applied to a wider class of distributions as the ideal distribution $\mathcal{Q}$ of a randomness. The main tool we use in this paper is the *Rényi divergence* [21, 23], which is a measure of divergence between distributions, and has recently been found useful in security evaluations of cryptographic primitives [3, 22, 6, 1, 19].

Our results are summarized as follows:

- In Section 3, we show two general lemmas that serve as the main tools throughout the paper, which are inequalities on two expectations each taken over arbitrary distribution $\mathcal{Q}$ and over $\mathcal{R}$, respectively (where intuitively, $\mathcal{Q}$ is an "ideal" distribution and $\mathcal{R}$ is a "real-world" distribution), and involve the Rényi divergence of $\mathcal{R}$ from $\mathcal{Q}$. These lemmas are generalizations of the lemmas shown by Dodis and Yu [11], who showed similar inequalities involving the min-entropy and collision-entropy of the "real-world" distribution $\mathcal{R}$, and theirs can only handle the case where the "ideal" distribution $\mathcal{Q}$ is the uniform distribution.
- Based on our general lemmas, in Section 4, we show general techniques for evaluating security of a cryptographic primitive (or, we use the term "application" following the style of [11] from here on) in case the distribution of a parameter (such as a secret key and/or a randomness) is switched from an ideal distribution $\mathcal{Q}$ to an arbitrary "real-world" distribution $\mathcal{R}$, using the Rényi divergence. As in [11], we show two types of results, one regarding unpredictability applications and the other regarding "square-friendly" indistinguishability applications. These results are generalizations of the corresponding results by Dodis and Yu [11], where their results only capture the case in which the ideal distribution $\mathcal{Q}$ is the uniform distribution.
- In Section 5, we show two applications of the above general results: one application from our general security evaluation technique for square-friendly indistinguishability applications from Section 4, and the other application from one of our lemmas in Section 3.
  - Our first application is for switching the distribution of a problem instance in distinguishing problems that satisfy the property called *public samplability*, formalized by Bai et al. [3]. Using the Rényi divergence, they showed a reduction from the hardness of a problem in this class to the hardness of the same problem but in which the distribution of a parameter behind a problem instance is switched from an original

distribution $\mathcal{Q}$ to another distribution $\mathcal{R}$. We show that distinguishing problems with public samplability are square-friendly in the sense of [4, 9, 11], thereby we can apply our above result on the general security evaluation technique of square-friendly applications under switching distributions to obtain a quantitatively improved reduction. Although our results are not applicable to the case in which the order α of the Rényi divergence is less than 2, our result gives a simpler and much tighter reduction than that of [3] for all $\alpha \geq 2$. Concretely, if we compare the ratio of the running time and the advantage of the reduction algorithm (which is sometimes called the "work factor", and a smaller value means a tighter reduction) for the same order of the Rényi divergence, the work factor of our reduction (solving the problem under distribution $\mathcal{Q}$) is always at least $\mathcal{O}(\epsilon^{-2})$ times smaller than that of the reduction shown in [3], where ϵ denotes the advantage of an underlying adversary (solving the problem under distribution $\mathcal{R}$). For the details, see Section 5.1.

- As the second application, we show that how differential privacy [14, 12, 13] of a mechanism is degraded when the randomness used by the mechanism comes from not an ideal distribution $\mathcal{Q}$ but a real-world distribution $\mathcal{R}$, using the Rényi divergence of order ∞. It is typical that non-uniform distributions that are uncommon in the constructions of cryptographic primitives (e.g. the Laplace distribution, the matrix Bingham distribution [7]), are used in the literature of differential privacy. Thus, although simple, we believe that this result is useful. For the details, see Section 5.2.

– Finally, motivated by the difficulty of sampling randomness from non-uniform distributions in computer implementations in practice, and in the light of the usefulness and versatility of the Rényi divergence in cryptography, in Section 6, we show two methods for approximate sampling from an arbitrary distribution $\mathcal{Q}$ by using a uniformly chosen random string via the inversion sampling (a.k.a. inverse transform sampling), with the guarantee that the Rényi divergence of the distribution of our sampling method (which we denote by $\mathcal{R}$) from the target ideal distribution $\mathcal{Q}$, is upperbounded. We show two results: one for the Rényi divergence of order 2 and the other for the Rényi divergence of order ∞.

We remark that previously, Yao and Li [24] showed some generalization of Dodis and Yu's lemmas [11] using *Rényi entropy* (which incorporates min-entropy and collision-entropy as special cases), and corresponding techniques for evaluating security of unpredictability and square-friendly indistinguishability applications, in a similar way we do in this paper. Interestingly, [24] uses the Hölder inequality as a main tool, which we also use for showing one of our technical lemmas in Section 3. Like [11], however, the results of [24] are only applicable to the case in which the ideal distribution is the uniform distribution (and the real-world distribution has some high Rényi entropy), and our results in this paper are more general than their main theorems [24, Theorems 3.2 and 3.3] in the sense that the latter can be derived from ours. On the other hand,

Yao and Li also studied the application of their results to a setting where the real-world distribution only has some high *computational* version of Rényi entropy, which is a setting we do not explore in this work. It would be interesting to investigate whether results with computational variants of Rényi divergence analogous to [24] can be established.

1.3 Paper Organization

The rest of the paper is organized as follows. In Section 2, we review basic notation and the definitions used in the paper. In Section 3, we show two general lemmas that are used throughout the subsequent sections. In Section 4, we show two general techniques for evaluating the security of cryptographic primitives, one for unpredictability applications and the other for "square-friendly" indistinguishability applications. In Section 5, we show two applications of the results from the previous sections, one for an improved reduction for a class of distinguishing problems called distinguishing problems with public samplability, and the other for differential privacy. In Section 6, we propose two inversion sampling methods for arbitrary discrete distributions, with some guaranteed upperbounds on the Rényi divergence.

2 Preliminaries

In this section, we review the basic notation and the definitions for the Rényi divergence and entropy, and some useful lemmas.

2.1 Basic Notation

$\mathbb{N}$, $\mathbb{Z}_{\geq 0}$, $\mathbb{R}$, and $\mathbb{R}_{\geq 0}$ denote the set of all natural numbers, all non-negative integers, all real numbers, and all non-negative real numbers, respectively. For $n \in \mathbb{N}$, we define $[n] := \{1, \ldots, n\}$. If S is a finite set, then "$|S|$" denotes its size, and "$x \leftarrow_{\mathtt{R}} S$" denotes that x is chosen uniformly at random from S. If $\mathcal{X}$ is a distribution (over some set), then "$x \leftarrow_{\mathtt{R}} \mathcal{X}$" denotes that x is chosen according to the distribution $\mathcal{X}$, and "$[\mathcal{X}]$" denotes the support of $\mathcal{X}$, i.e. $[\mathcal{X}] := \{x | \Pr[\mathcal{X} = x] > 0\}$. In this paper, we only treat discrete distributions.

If A is a probabilistic algorithm, then "$\mathsf{A}(x)$" denotes the distribution of A's output when it takes x as input and uses an internal randomness chosen according to some prescribed distribution, and if we need to specify a particular randomness r used by A, we denote it by "$\mathsf{A}(x; r)$" (in which case the computation of A is deterministic that takes x and r as input).

2.2 Hölder Inequality

Here, we recall the Hölder inequality.

Lemma 1 (Hölder Inequality). *Let $n \in \mathbb{N}$, and let $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ be sequences of real numbers. Let $\alpha, \beta \in (1, \infty)$ be real numbers such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Then, it holds that*

$$\sum_{i \in [n]} |a_i \cdot b_i| \leq \Big(\sum_{i \in [n]} |a_i|^\alpha \Big)^{\frac{1}{\alpha}} \cdot \Big(\sum_{i \in [n]} |b_i|^\beta \Big)^{\frac{1}{\beta}}.$$

Note that the case of $\alpha = \beta = 2$ implies the Cauchy-Schwarz inequality.

2.3 Rényi Divergence

Here, we recall the definition of Rényi divergence in the form typically used in cryptography.[3]

Definition 1. *Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions such that $[\mathcal{R}] \subseteq [\mathcal{Q}]$, and let $\alpha > 1$ be a real number. The* Rényi divergence of order α *(or α-Rényi divergence, for short) of the distribution $\mathcal{R}$ from the distribution $\mathcal{Q}$, denoted by $\mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q})$, is defined by*

$$\mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q}) := \Big(\sum_{z \in [\mathcal{Q}]} \frac{\Pr[\mathcal{R} = z]^\alpha}{\Pr[\mathcal{Q} = z]^{\alpha - 1}} \Big)^{\frac{1}{\alpha - 1}},$$

and the ∞-Rényi divergence of $\mathcal{R}$ from $\mathcal{Q}$ is given by

$$\mathtt{RD}_\infty(\mathcal{R} \| \mathcal{Q}) := \max_{z \in [\mathcal{Q}]} \frac{\Pr[\mathcal{R} = z]}{\Pr[\mathcal{Q} = z]}.$$

It is known that the Rényi divergence is non-decreasing in its order, and not less than 1 when $\alpha > 1$ (see [23]). Thus, for any distributions $\mathcal{Q}$ and $\mathcal{R}$ and $1 < \alpha < \alpha'$, we have $1 \leq \mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q}) \leq \mathtt{RD}_{\alpha'}(\mathcal{R} \| \mathcal{Q})$.

It is also known that the Rényi divergence enjoys several (multiplicative) analogues of the properties satisfied by the statistical distance (see [3, Lemma 2.9]). Here, we recall the so-called probability preservation property of the Rényi divergence.

Lemma 2 (Probability Preservation). *Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the same set X such that $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. Then, for all $E \subseteq X$ and $\alpha \in (1, \infty)$, it holds that*

$$\Pr[\mathcal{R} \in E] \leq \min\Big\{ \Big(\mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q}) \cdot \Pr[\mathcal{Q} \in E] \Big)^{\frac{\alpha - 1}{\alpha}}, \ \mathtt{RD}_\infty(\mathcal{R} \| \mathcal{Q}) \cdot \Pr[\mathcal{Q} \in E] \Big\}.$$

2.4 Entropy

Here, we recall the definitions of entropy.

[3] In a non-cryptographic context, it is typical to define the α-Rényi divergence as the logarithm of the quantity $\mathtt{RD}_\alpha$ defined here [23].

Definition 2. *Let $\mathcal{X}$ (resp. $\mathcal{Y}$) be a distribution defined over a set X (resp. Y).*

- *The* min-entropy *of $\mathcal{X}$, denoted by $\mathbf{H}_\infty(\mathcal{X})$, is defined by*
$$\mathbf{H}_\infty(\mathcal{X}) := -\log_2\Big(\max_{x\in X}\Pr[\mathcal{X}=x]\Big).$$
- *The* average collision-entropy *of $\mathcal{X}$ given $\mathcal{Y}$, denoted by $\mathbf{H}_2(\mathcal{X}|\mathcal{Y})$, is defined by*
$$\mathbf{H}_2(\mathcal{X}|\mathcal{Y}) := -\log_2\Big(\mathop{\mathbf{E}}_{y\leftarrow_{\mathrm{R}}\mathcal{Y}}\Big[\sum_{x\in X}\Pr[\mathcal{X}=x|\mathcal{Y}=y]^2\Big]\Big).$$

3 General Lemmas for Switching Distributions

In this section, we show two lemmas that are used as the main tools for showing our results in the subsequent sections. Our lemmas are generalizations of the lemmas shown by Dodis and Yu [11]. Thus, for reference we first recall their lemmas in Section 3.1. We then show our lemmas in Section 3.2.

3.1 Lemmas by Dodis and Yu

Dodis and Yu [11] showed the following lemmas. Actually, they only state the lemmas for functions taking bitstrings as input, but the lemmas straightforwardly generalize for functions with any domain. Thus, we state such versions.

Lemma 3 (Lemma 1 in [11]). *Let X be a finite set, $\mathcal{R}$ be a distribution over X, and $\mathcal{U}$ be the uniform distribution over X. Then, for any (deterministic) non-negative function $f : X \to \mathbb{R}_{\geq 0}$, we have*
$$\mathbf{E}[f(\mathcal{R})] \leq |X| \cdot 2^{-\mathbf{H}_\infty(\mathcal{R})} \cdot \mathbf{E}[f(\mathcal{U})].$$

Lemma 4 (Adapted from Lemmas 5 and 7 in [11][4]). *Let X and Y be finite sets, and $(\mathcal{R},\mathcal{S})$ be a joint distribution over $X \times Y$. Let $\mathcal{U}$ be the uniform distribution over X. Then, for any (deterministic) real-valued function $f : X \times Y \to \mathbb{R}$, we have*
$$\Big|\mathbf{E}[f(\mathcal{R},\mathcal{S})]\Big| \leq \sqrt{|X| \cdot 2^{-\mathbf{H}_2(\mathcal{R}|\mathcal{S})} \cdot \mathbf{E}[f(\mathcal{U},\mathcal{S})^2]} \qquad \textit{and}$$
$$\Big|\mathbf{E}[f(\mathcal{R},\mathcal{S})] - \mathbf{E}[f(\mathcal{U},\mathcal{S})]\Big| \leq \sqrt{(|X| \cdot 2^{-\mathbf{H}_2(\mathcal{R}|\mathcal{S})} - 1) \cdot \mathbf{E}[f(\mathcal{U},\mathcal{S})^2]}.$$

3.2 Our Lemmas

Our first lemma is as follows.

Lemma 5. *Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the same set X such that $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. Then, for any (deterministic) real-valued function $f : X \to \mathbb{R}$ and any $\alpha \in (1,\infty)$, we have*
$$\Big|\mathbf{E}[f(\mathcal{R})]\Big| \leq \min\left\{\left(\mathtt{RD}_\alpha(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}\Big[|f(\mathcal{Q})|^{\frac{\alpha}{\alpha-1}}\Big]\right)^{\frac{\alpha-1}{\alpha}},\ \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}\Big[|f(\mathcal{Q})|\Big]\right\}. \tag{1}$$

[4] Lemma 7 in [11] is attributed to Barak et al. [4].

Relation to Dodis and Yu's Lemma. Before providing the proof, let us remark that the above lemma is a generalization of Lemma 3 and the first inequality in Lemma 4. To see this, note that for any distribution $\mathcal{R}$ over some set X and the uniform distribution $\mathcal{U}$ over X, we have $\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{U}) = \max_{x\in X} \frac{\Pr[\mathcal{R}=x]}{\Pr[\mathcal{U}=x]} = |X| \cdot 2^{-\mathbf{H}_\infty(\mathcal{R})}$, and thus Lemma 3 can be obtained by setting $\mathcal{Q} = \mathcal{U}$ in our lemma for non-negative functions. Also, let $\mathcal{S}$ be an arbitrary distribution over some set Y that forms a joint distribution $(\mathcal{R}, \mathcal{S})$ over $X \times Y$. Then, we have

$$\begin{aligned}
\mathtt{RD}_2\Big((\mathcal{R},\mathcal{S})\|(\mathcal{U},\mathcal{S})\Big) &= \sum_{(x,y)\in X\times Y} \frac{\Pr[\mathcal{R}=x\wedge\mathcal{S}=y]^2}{\Pr[\mathcal{U}=x\wedge\mathcal{S}=y]} \\
&= \sum_{(x,y)\in X\times Y} \frac{(\Pr[\mathcal{R}=x|\mathcal{S}=y]\cdot\Pr[\mathcal{S}=y])^2}{\frac{1}{|X|}\cdot\Pr[\mathcal{S}=y]} \\
&= |X| \cdot \sum_{y\in Y} \Pr[\mathcal{S}=y] \cdot \Big(\sum_{x\in X} \Pr[\mathcal{R}=x|\mathcal{S}=y]^2\Big) \\
&= |X| \cdot 2^{-\mathbf{H}_2(\mathcal{R}|\mathcal{S})},
\end{aligned}$$

and thus, the first inequality in Lemma 4 can be obtained by setting $\mathcal{R}$ in our lemma to be $(\mathcal{R}, \mathcal{S})$ explained here, setting $\mathcal{Q} = (\mathcal{U}, \mathcal{S})$, and then invoking our lemma for general real-valued functions and $\alpha = 2$.

Proof of Lemma 5. For each $z \in [\mathcal{Q}]$, let $r_z := \Pr[\mathcal{R} = z]$ and $q_z := \Pr[\mathcal{Q} = z]$. The bound regarding the ∞-Rényi divergence can be shown as follows:

$$\begin{aligned}
\Big|\mathbf{E}[f(\mathcal{R})]\Big| \le \sum_{z\in[\mathcal{Q}]} r_z \cdot |f(z)| &\overset{(*)}{\le} \sum_{z\in[\mathcal{Q}]} \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \cdot q_z \cdot |f(z)| \\
&= \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}\Big[|f(\mathcal{Q})|\Big],
\end{aligned}$$

where the inequality (*) uses the probability preservation property (Lemma 2), which implies $r_z \le \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \cdot q_z$.

The bound for a general $\alpha \in (1, \infty)$ does not simply follow from the probability preservation property, but can be shown using the Hölder inequality (Lemma 1). Specifically, we have

$$\begin{aligned}
\Big|\mathbf{E}[f(\mathcal{R})]\Big| &\le \sum_{z\in[\mathcal{Q}]} \Big(r_z \cdot q_z^{-\frac{\alpha-1}{\alpha}}\Big) \cdot q_z^{\frac{\alpha-1}{\alpha}} \cdot |f(z)| \\
&\overset{(*)}{\le} \Big(\sum_{z\in[\mathcal{Q}]} \Big(r_z \cdot q_z^{-\frac{\alpha-1}{\alpha}}\Big)^\alpha\Big)^{\frac{1}{\alpha}} \cdot \Big(\sum_{z\in[\mathcal{Q}]} \Big(q_z^{\frac{\alpha-1}{\alpha}} \cdot |f(z)|\Big)^{\frac{\alpha}{\alpha-1}}\Big)^{\frac{\alpha-1}{\alpha}} \\
&= \Big(\sum_{z\in[\mathcal{Q}]} r_z^\alpha \cdot q_z^{-(\alpha-1)}\Big)^{\frac{1}{\alpha}} \cdot \Big(\sum_{z\in[\mathcal{Q}]} q_z \cdot |f(z)|^{\frac{\alpha}{\alpha-1}}\Big)^{\frac{\alpha-1}{\alpha}} \\
&= \Big(\mathtt{RD}_\alpha(\mathcal{R}\|\mathcal{Q})^{\alpha-1}\Big)^{\frac{1}{\alpha}} \cdot \Big(\mathbf{E}\Big[|f(\mathcal{Q})|^{\frac{\alpha}{\alpha-1}}\Big]\Big)^{\frac{\alpha-1}{\alpha}},
\end{aligned}$$

where the inequality (*) is due to the Hölder inequality. Note that the rightmost is equivalent to the first bound in Eq. (1). □ (**Lemma 5**)

Note that if the range of a function f is $[0,1]$, then $f(z)^a \leq f(z)$ holds for every $a \geq 1$, and thus the inequalities in Lemma 5 can be slightly simplified. For our purpose, it is useful to formally state it as a corollary, which can be seen as a generalization of the probability preservation property (Lemma 2).

Corollary 1 (Special Case of Lemma 5). *Let $\mathcal{Q}$ and $\mathcal{R}$ be the same as in Lemma 5. Then, for any (deterministic) function $f : X \to [0,1]$ and any $\alpha \in (1,\infty)$, we have*

$$\mathbf{E}[f(\mathcal{R})] \leq \min\Big\{ \Big(\mathtt{RD}_\alpha(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}[f(\mathcal{Q})]\Big)^{\frac{\alpha-1}{\alpha}}, \ \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}[f(\mathcal{Q})] \Big\}.$$

Our second lemma is as follows. Similarly to our first lemma, the lemma here is a generalization of the second inequality in Lemma 4.

Lemma 6. *Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the same set X such that $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. Then, for any (deterministic) real-valued function $f : X \to \mathbb{R}$, we have*

$$\Big|\mathbf{E}[f(\mathcal{R})] - \mathbf{E}[f(\mathcal{Q})]\Big| \leq \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \mathbf{E}[f(\mathcal{Q})^2]}. \tag{2}$$

Proof of Lemma 6. Let $c \in \mathbb{R}_{\geq 0}$. Using the same notation as in the proof of Lemma 5, we have

$$\begin{aligned}
\Big|\mathbf{E}[f(\mathcal{R})] - c \cdot \mathbf{E}[f(\mathcal{Q})]\Big| &= \Big| \sum_{z \in [\mathcal{Q}]} \Big(\frac{r_z}{\sqrt{q_z}} - c \cdot \sqrt{q_z}\Big) \cdot \sqrt{q_z} \cdot f(z) \Big| \\
&\leq \sqrt{\sum_{z \in [\mathcal{Q}]} \Big(\frac{r_z}{\sqrt{q_z}} - c \cdot \sqrt{q_z}\Big)^2} \cdot \sqrt{\sum_{z \in [\mathcal{Q}]} q_z \cdot f(z)^2} \\
&= \sqrt{\sum_{z \in [\mathcal{Q}]} \frac{r_z^2}{q_z} - 2c \cdot \sum_{z \in [Q]} r_z + c^2 \cdot \sum_{z \in [\mathcal{Q}]} q_z} \cdot \sqrt{\mathbf{E}[f(\mathcal{Q})^2]} \\
&= \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 2c + c^2) \cdot \mathbf{E}[f(\mathcal{Q})^2]},
\end{aligned}$$

where the inequality is due to the Cauchy-Schwarz inequality (Lemma 1 with $\alpha = \beta = 2$), and the last equality uses $\sum_{z \in [\mathcal{Q}]} r_z = \sum_{z \in [\mathcal{Q}]} q_z = 1$. Then, Eq. (2) is obtained by taking $c = 1$.[5] □ (**Lemma 6**)

4 General Security Evaluation Techniques via Rényi Divergence

In this section, we show general techniques for evaluating security in case the distribution of a parameter (e.g. a secret key, randomness etc.) in a security

[5] We can also obtain the proof of the case $\alpha = 2$ of our first lemma by setting $c \in \{0, 2\}$ in this proof. Setting other values for c does not seem to give us any merit.

game is switched from an "ideal" distribution $\mathcal{Q}$ to an arbitrary "real-world" distribution $\mathcal{R}$, using the Rényi divergence.

Specifically, in Section 4.1, we first recall the definition of an abstract security game in the style of [11] that abstractly captures most security games used in cryptography, in particular unpredictability applications and indistinguishability applications. There, we also recall the notion of "square-security" [4, 11]. It plays an important role when showing results for "square-friendly" applications, which is a class of applications including all unpredictability applications and many indistinguishability applications.

Then, in Sections 4.2 and 4.3, we show general results on how the security of applications in the "ideal" model in which a parameter is drawn from an ideal distribution $\mathcal{Q}$, is "degraded" in the "real-world" model in which a parameter is drawn from an arbitrary distribution $\mathcal{R}$. Our result for unpredictability applications is given in Section 4.2, and our result for square-friendly indistinguishability applications is given in Section 4.3.

4.1 Definitions

Abstract Security Game. We define a general type of cryptographic applications in the same manner as [11]. The security of a cryptographic application Π is defined via an interactive *security game* between a probabilistic adversary A and a probabilistic challenger $\mathsf{C}(r)$, where C is fixed by the definition of Π, and $r \in X$ is a "parameter" [6] in the security game that is drawn from some distribution, which we wish to switch to another distribution using the Rényi divergence. The game can have an arbitrary structure, and after the interaction with the adversary A, the challenger $\mathsf{C}(r)$ outputs a bit. If $\mathsf{C}(r)$ outputs 1 (resp. 0), A is said to win (resp. lose) the game. As usual, we consider two types of cryptographic applications: *unpredictability* applications and *indistinguishability* applications. The former type captures applications in which it is hard for an adversary to compute some value (e.g. a preimage of a one-way function, forging a signature on a fresh message), and the latter type captures applications in which it is hard for an adversary to guess the challenge bit chosen by the challenger (e.g. security of a pseudorandom function, IND-CPA security of an encryption scheme).

Given a particular parameter $r \in X$, let $\mathtt{Win}_{\mathsf{A}}(r)$ be the probability that A wins in the security game played with the challenger $\mathsf{C}(r)$, where the probability is over the choice of the randomness consumed by A and $\mathsf{C}(r)$. Then, we define the *advantage* $\mathtt{Adv}_{\mathsf{A}}(r)$ *of* A *on* r (against particular C fixed by an application Π) as follows:

$$\mathtt{Adv}_{\mathsf{A}}(r) := \begin{cases} \mathtt{Win}_{\mathsf{A}}(r) & \text{(for unpredictability applications)} \\ 2 \cdot \mathtt{Win}_{\mathsf{A}}(r) - 1 & \text{(for indistinguishability applications)} \end{cases}.$$

[6] In [11], r was called a "secret key". Since r can be any value sampled in the security game, we call it just a "parameter".

The actual advantage of an adversary in a security game is defined by taking (the absolute value of) the expectation over the choice of the parameter r in the game. In this paper, as in [11], we will treat ordinary security and "square-security", the latter of which takes the expectation of the squared value of the advantage and plays an important role for the results on indistinguishability applications that are "square-friendly". The (square-)security of Π in case the parameter r is chosen according to a distribution $\mathcal{X}$, is called (square-)security in the $\mathcal{X}$-model.

Definition 3 (Security and Square-Security (Adapted from [11])). *Let $\mathcal{X}$ be a distribution over the parameter space X. We say that an application Π is*

- (T, ϵ)-secure in the $\mathcal{X}$-model, *if for all adversaries* A *with resource*[7] *T, it holds that* $|\,\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})]| \leq \epsilon$.
  $|\,\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})]|$ *is called the* advantage of A in the $\mathcal{X}$-model.
- (T, σ)-square-secure in the $\mathcal{X}$-model, *if for all adversaries* A *with resource T, it holds that* $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})^2] \leq \sigma$.
  $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})^2]$ *is called the* square-advantage of A in the $\mathcal{X}$-model.

4.2 General Result for Unpredictability Applications

Our result for unpredictability applications is stated as follows.

Theorem 1. *Let Π be an unpredictability application. Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the parameter space X satisfying $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. Then, for any adversary* A *against the security of Π, it holds that*[8]

$$\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] \leq F_{\mathcal{Q}\to\mathcal{R}}\Big(\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big),$$

where the function $F_{\mathcal{Q}\to\mathcal{R}}(\cdot)$ is defined by

$$F_{\mathcal{Q}\to\mathcal{R}}(\epsilon) := \min\left\{\begin{array}{c} \min_{\alpha\in(1,\infty)}\Big(\mathtt{RD}_\alpha(\mathcal{R}\|\mathcal{Q})\cdot\epsilon\Big)^{\frac{\alpha-1}{\alpha}}, \\ \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot\epsilon, \\ \epsilon + \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q})-1)\cdot\epsilon} \end{array}\right\}.$$

In particular, if Π is (T, ϵ)-secure in the $\mathcal{Q}$-model, then Π is also $(T, F_{\mathcal{Q}\to\mathcal{R}}(\epsilon))$-secure in the $\mathcal{R}$-model.

This theorem shows how the security of an unpredictability application under a real-world distribution $\mathcal{R}$ is guaranteed in terms of its security under an ideal distribution $\mathcal{Q}$, via the Rényi divergence. In particular, this theorem gives us

[7] Resource of an adversary abstractly models all of an adversary's efficiency measures, e.g. the running time, the circuit size, the number of oracle queries, etc.

[8] Note that for unpredictability applications, the absolute value of an adversary A's advantage can be removed.

an implication to the standard asymptotic-style security: If an unpredictability application is secure in the $\mathcal{Q}$-model in the asymptotic sense (i.e. any efficient adversary's advantage in the $\mathcal{Q}$-model is bounded by a negligible function of a security parameter), it remains secure in the $\mathcal{R}$-model as long as the Rényi divergence of some order $\alpha \in (1, \infty]$ is bounded by a polynomial of the security parameter.

Proof of Theorem 1. Let A be any adversary against the security of Π. Since Π is an unpredictability application, the range of $\mathtt{Adv}_{\mathsf{A}}(\cdot)$ is $[0, 1]$. Thus, A's advantage in the $\mathcal{R}$-model (resp. $\mathcal{Q}$-model) is $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]$ (resp. $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]$). Then, $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] \leq \min_{\alpha \in (1,\infty)}(\mathtt{RD}_{\alpha}(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})])^{\frac{\alpha-1}{\alpha}}$ and $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] \leq \mathtt{RD}_{\infty}(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]$, are obtained by applying Corollary 1 to the advantage function $\mathtt{Adv}_{\mathsf{A}}(\cdot)$.

To complete the proof, it remains to show $\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] \leq \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})] + \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]}$. By the triangle inequality, we have

$$\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] \leq \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})] + \Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] - \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big|.$$

Regarding the second term in the right hand side, due to Lemma 6, we have

$$\begin{aligned}\Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] - \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big| &\leq \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})^2]} \\ &\leq \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]},\end{aligned}$$

where we use $\mathtt{Adv}_{\mathsf{A}}(\cdot)^2 \leq \mathtt{Adv}_{\mathsf{A}}(\cdot)$, which is in turn because its range is $[0, 1]$. Combining the two inequalities yields the desired inequality. □ (**Theorem 1**)

4.3 General Result for Square-Friendly Indistinguishability Applications

Here, we show our general result for "square-friendly" indistinguishability applications, which is done via the notion of square-security.

We first show how the security of any application (including both unpredictability and indistinguishability applications) under a real-world distribution $\mathcal{R}$ is guaranteed from its square-security (and ordinary security) under an ideal distribution $\mathcal{Q}$, via the 2-Rényi divergence $\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q})$.

Lemma 7. *Let Π be an (unpredictability/indistinguishability) application. Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the parameter space X satisfying $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. Then, for any adversary A against the security of Π, it holds that*

$$\Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]\Big| \leq G_{\mathcal{Q}\to\mathcal{R}}\Big(\ \Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big|,\ \ \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})^2]\ \Big),$$

where the function $G_{\mathcal{Q}\to\mathcal{R}}(\cdot,\cdot)$ is defined by

$$G_{\mathcal{Q}\to\mathcal{R}}(\epsilon, \sigma) := \min\Big\{\ \sqrt{\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) \cdot \sigma},\ \epsilon + \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \sigma}\ \Big\}. \quad (3)$$

In particular, if Π is (simultaneously) (T, ϵ)-secure and (T, σ)-square-secure in the $\mathcal{Q}$-model, then Π is also $(T, G_{\mathcal{Q}\to\mathcal{R}}(\epsilon, \sigma))$-secure in the $\mathcal{R}$-model.[9]

Proof of Lemma 7. Let A be any adversary against the security of Π. Then, applying the first bound in Eq. (1) in Lemma 5 with $\alpha = 2$ to the advantage function $\mathtt{Adv}_{\mathsf{A}}(\cdot)$, we immediately obtain $|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]| \le \sqrt{\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})^2]}$.

To complete the proof, it remains to show $|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]| \le |\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]| + \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})^2]}$. By the triangle inequality, we have

$$\Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]\Big| \le \Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big| + \Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] - \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big|.$$

Regarding the second term in the right hand side, due to Lemma 6, we have

$$\Big|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] - \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]\Big| \le \sqrt{(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) - 1) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})^2]}.$$

Combining the two inequalities yields the desired inequality. □ (**Lemma 7**)

Next, we would like to establish the implication of the security of an indistinguishability application to its square-security, but unfortunately it is known that for some indistinguishability applications, their square-security is not necessarily implied by the ordinary security. Fortunately, however, the works of Barak et al. [4] and Dodis and Yu [11] showed that for some indistinguishability applications in which the so-called "double-run trick" is applicable, ordinary (non-square) security *does* imply its corresponding square-security. Dodis and Yu formalized a sufficient condition for such indistinguishability applications as what they call *simulatability*. It is this property that makes indistinguishability applications square-friendly. We recall the definition here.

Definition 4 (Simulatability [11]). *Consider an indistinguishability application Π in the security game of which possibly there is a "failure predicate"*[10] *F (that is efficiently checkable by both an adversary A and the challenger $\mathsf{C}(r)$) such that $\mathsf{C}(r)$ regards A as winning the game if A succeeds in guessing the challenge bit* and *does not violate F, while if A violates F, the challenger flips a random coin on behalf of A and uses it to decide if A wins the game or not*[11]*. We say that Π is* (T', T, γ)-simulatable*, if for any parameter r and any adversary A whose resource is T and that never violates the failure predicate F, there exists an adversary B (against the security of Π) with resource T' (for some $T' \ge T$) such that:*

[9] Note that the first bound does not involve the (non-square) advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})]|$, and hence is true regardless of the (non-square) security of Π in the $\mathcal{Q}$-model.

[10] A failure predicate models the restrictions in a security game that typically prevent an adversary from winning the game trivially, e.g., submitting the challenge ciphertext as a decryption query in the IND-CCA security game of an encryption scheme.

[11] This is to offset an adversary's advantage in case it violates the failure predicate F. How an adversary's advantage is affected in case it violates the failure predicate F is not explicit in the definition of [11], and thus we adopt (seemingly) the most natural choice which is also convenient for our purpose.

1. *The execution between* B *and "real"* $\mathsf{C}(r)$ *defines two independent executions between a copy* A_i *of* A *and a "simulated" challenger* $\mathsf{C}_i(r)$*, for* $i = 1, 2$*. In particular, except reusing the same* r*,* A_1*,* $\mathsf{C}_1(r)$*,* A_2*, and* $\mathsf{C}_2(r)$ *use fresh and independent randomness, including independent challenge bits* b_1 *and* b_2*.*
2. *The challenge bit* b *used by the "real"* $\mathsf{C}(r)$ *is equal to the challenge bit* b_2 *used by the "simulated"* C_2*.*
3. *Before making its guess* b' *of the challenge bit* b*,* B *learns the values* b_1*,* b'_1*, and* b'_2*, where each* b'_i *denotes* A_i*'s guess for* b_i*.*
4. *The probability of* B *violating the failure predicate* F *is at most* γ*.*

Though it might look somewhat complicated, as noted in [11], simulatability is satisfied by many natural indistinguishability applications, such as IND-CPA and IND-CCA security of encryption schemes, (weak) pseudorandom functions. Looking ahead, in Section 5.1, we will see another example of indistinguishability applications with simulatability, which is called "distinguishing problems with public samplability" formalized by Bai et al. [3].

We now show that for indistinguishability applications that satisfy simulatability as defined above, their square-security is indeed implied by the ordinary security. This is a generalized version of [11, Lemma 4].

Lemma 8. *Let* $\mathcal{X}$ *be a distribution over the parameter space* X*. If an indistinguishability application* Π *is* (T', T, γ)*-simulatable (with* $T' \geq T$*), then for any adversary* A *with resource* T *against the security of* Π*, there exists another adversary* B *with resource* T' *against the security of* Π*, such that*

$$\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})^2] \leq \left|\mathbf{E}[\mathtt{Adv}_{\mathsf{B}}(\mathcal{X})]\right| + \gamma.$$

In particular, if Π *is* (T', ϵ)*-secure in the* $\mathcal{X}$*-model and* (T', T, γ)*-simulatable, then* Π *is also* $(T, \epsilon + \gamma)$*-square-secure in the* $\mathcal{X}$*-model.*

Proof of Lemma 8. This theorem can be shown via the "double-run trick" [4, 11]. Let Π be an indistinguishability application that is (T, T', γ)-simulatable. Let A be any adversary with resource T against the security of Π, and let B be the adversary (corresponding to A and the challenger $\mathsf{C}(r)$, where r is chosen according to $\mathcal{X}$) with resource T' against the security of Π, which is guaranteed to exist by the (T', T, γ)-simulatability of Π. We augment B as an adversary against the security of Π so that when B successfully finishes the two executions of A (without violating the failure predicate), if $b'_1 = b_1$ then B sets $b' := b'_2$, otherwise sets $b' := 1 - b'_2$, and outputs b' as its guess for the challenge bit, which we denote by b_2. Let $\mathtt{F}$ be the event that B violates the failure predicate. Then, due to the (T', T, γ)-simulatability, both of B's simulations of the challenger $\mathsf{C}(r)$ for A are perfect as long as $\mathtt{F}$ does not happen, and whether $\mathtt{F}$ happens is independent of the choice of r and whether $b'_1 = b_1$ or $b'_2 = b_2$ occurs. Hence, B's

advantage on a fixed parameter r can be calculated as follows:

$$\begin{aligned}
&\mathtt{Adv}_{\mathsf{B}}(r) = 2 \cdot \mathtt{Win}_{\mathsf{B}}(r) - 1 = 2 \cdot \Pr[b' = b_2] - 1 \\
&= 2 \cdot \Big(\Pr[b'_1 = b_1 \wedge b'_2 = b_2 \wedge \overline{\mathtt{F}}] + \Pr[b'_1 \neq b_1 \wedge b'_2 \neq b_2 \wedge \overline{\mathtt{F}}] + \frac{1}{2} \Pr[\mathtt{F}] \Big) - 1 \\
&= (1 - \Pr[\mathtt{F}]) \cdot \Big(2 \cdot \Pr[b'_1 = b_1 \wedge b'_2 = b_2] + 2 \cdot \Pr[b'_1 \neq b_1 \wedge b'_2 \neq b_2] - 1 \Big).
\end{aligned}$$

Here, $\Pr[b'_1 = b_1 \wedge b'_2 = b_2]$ (resp. $\Pr[b'_1 \neq b_1 \wedge b'_2 \neq b_2]$) corresponds to the probability (which does not include the choice of r) that A wins (resp. loses) the game played with $\mathsf{C}(r)$ twice, and thus is equal to $\mathtt{Win}_{\mathsf{A}}(r)^2$ (resp. $(1-\mathtt{Win}_{\mathsf{A}}(r))^2$). Hence, we have

$$\begin{aligned}
\mathtt{Adv}_{\mathsf{B}}(r) &= (1 - \Pr[\mathtt{F}]) \cdot \Big(2 \cdot \mathtt{Win}_{\mathsf{A}}(r)^2 + 2 \cdot (1 - \mathtt{Win}_{\mathsf{A}}(r))^2 - 1 \Big) \\
&= (1 - \Pr[\mathtt{F}]) \cdot \Big(2 \cdot \mathtt{Win}_{\mathsf{A}}(r) - 1 \Big)^2 \\
&= (1 - \Pr[\mathtt{F}]) \cdot \mathtt{Adv}_{\mathsf{A}}(r)^2.
\end{aligned}$$

From this equality, $\Pr[\mathtt{F}] \leq \gamma$, and the fact that the square-advantage is at most 1, we obtain

$$\Big| \mathbf{E}[\mathtt{Adv}_{\mathsf{B}}(\mathcal{X})] \Big| = (1 - \Pr[\mathtt{F}]) \cdot \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})^2] \geq \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{X})^2] - \gamma,$$

which is equivalent to the inequality stated in the theorem. □ (**Lemma 8**)

Combining Lemma 8 with Lemma 7, we obtain our general result for square-friendly indistinguishability applications. Specifically, the following theorem shows how the security of an indistinguishability application satisfying simulatability under a real-world distribution $\mathcal{R}$ is guaranteed in terms of its security under an ideal distribution $\mathcal{Q}$, via the 2-Rényi divergence $\mathtt{RD}_2(\mathcal{R} \| \mathcal{Q})$.

Theorem 2. *Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the parameter space X satisfying $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. If an indistinguishability application Π is (T', T, γ)-simulatable (with $T' \geq T$), then for any adversary A with resource T against the security of Π, there exists an adversary B with resource T' against the security of Π, such that*

$$\Big| \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})] \Big| \leq G_{\mathcal{Q} \to \mathcal{R}} \Big(\ \Big| \mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{Q})] \Big|, \ \Big| \mathbf{E}[\mathtt{Adv}_{\mathsf{B}}(\mathcal{Q})] \Big| + \gamma \ \Big),$$

where the function $G_{\mathcal{Q} \to \mathcal{R}}(\cdot, \cdot)$ is defined as in Eq. (3) of Lemma 7. In particular, if Π is (T', ϵ)-secure in the $\mathcal{Q}$-model and (T', T, γ)-simulatable, then Π is also $(T, G_{\mathcal{Q} \to \mathcal{R}}(\epsilon, \epsilon + \gamma))$-secure in the $\mathcal{R}$-model.

It would be an interesting question whether our general result for indistinguishability applications can be extended to those without simulatability.

5 Applications

In this section, we show some applications of our results from Sections 3 and 4.

Specifically, in Section 5.1, we show an improved reduction for a class of distinguishing problems, called *distinguishing problems with public samplability* formalized by Bai et al. [3], using our results for indistinguishability applications with simulatability given in Section 4. Next, in Section 5.2, we show how one of the general lemmas shown in Section 3 is useful for assessing the differential privacy [14, 12, 13] of a privacy mechanism in which randomness (a.k.a. "noise") comes from a "real-world" distribution in terms of its differential privacy with an ideal randomness distribution.

5.1 Tighter Reduction for Distinguishing Problems with Public Samplability

In [3], Bai et al. formalized a class of distinguishing problems called *distinguishing problems with public samplability*. Informally, a distinguishing problem is said to have public samplability if given a problem instance x which is generated according to one of distributions $\mathcal{D}_0(r)$ or $\mathcal{D}_1(r)$ where r denotes a parameter chosen from some distribution common to both $\mathcal{D}_0$ and $\mathcal{D}_1$, we can efficiently sample a "fresh" sample from both $\mathcal{D}_0(r)$ and $\mathcal{D}_1(r)$, regardless of whether the original x comes from $\mathcal{D}_0(r)$ or $\mathcal{D}_1(r)$. One example of such a problem is the learning with errors (LWE) problem, which is a problem to decide, given a matrix/vector pair $(A, \mathbf{b})$, whether the vector $\mathbf{b}$ is of the form $\mathbf{b} = A \cdot \mathbf{s} + \mathbf{e}$ where $\mathbf{s}$ is a secret vector and $\mathbf{e}$ is a small "noise" vector, or $\mathbf{b}$ is chosen uniformly at random. It has public samplability because given a problem instance $x = (A, \mathbf{b})$, one can sample fresh LWE problem instances having the same A. (In this example, r is the matrix A.)

Bai et al. showed a reduction that reduces the hardness of a distinguishing problem with public samplability to the hardness of the same problem in which the distribution of a parameter r is changed to another distribution, using the Rényi divergence.

Our result in this subsection is a tighter reduction than the one by Bai et al.. To formally show our result and give a comparison, we first recall the formal definition of distinguishing problems with public samplability, and then recall the result by Bai et al..

Definition 5 (Distinguishing Problem with Public Samplability [3]). *A* distinguishing problem *is a type of indistinguishability application and consists of a tuple $D = (X, \mathcal{D}_0, \mathcal{D}_1)$ where X is the parameter space, and $\mathcal{D}_0$ and $\mathcal{D}_1$ are (possibly probabilistic) functions with domain X. In the security game of D, the challenger $\mathsf{C}(r)$ (which receives a parameter $r \in X$ as input) first picks the challenge bit $b \in \{0, 1\}$ uniformly at random, samples $x \leftarrow_{\mathtt{R}} \mathcal{D}_b(r)$, and gives x to an adversary A. When A returns its guess b' for b, $\mathsf{C}(r)$ decides that A wins (resp. loses) the game if $b' = b$ (resp. $b' \neq b$) and outputs 1 (resp. 0).*

We say that a distinguishing problem $D = (X, \mathcal{D}_0, \mathcal{D}_1)$ is publicly samplable*, if there exists a probabilistic algorithm S (called the sampling algorithm) satisfying the following properties:*

- S *takes a bit b and a sample x (output by $\mathcal{D}_0$ or $\mathcal{D}_1$) as input, and outputs some value x'.*[12]
- *For any $(r, b) \in X \times \{0, 1\}$ and any values x output by $\mathcal{D}_b(r)$,*
  - *The output of $\mathsf{S}(0, x)$ is distributed identically to a fresh sample chosen according to the distribution $\mathcal{D}_0(r)$.*
  - *The output of $\mathsf{S}(1, x)$ is distributed identically to a fresh sample chosen according to the distribution $\mathcal{D}_1(r)$.*

Theorem 3 (Theorem 4.2 of [3]). *Let $D = (X, \mathcal{D}_0, \mathcal{D}_1)$ be a distinguishing problem with public samplability, and let S be the corresponding sampling algorithm whose running time is T_{S}. Let $\mathcal{Q}$ and $\mathcal{R}$ be distributions over the parameter space X such that $[\mathcal{R}] \subseteq [\mathcal{Q}] \subseteq X$. Then, for any adversary A against the security of D with running time T_{A} and advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]| = \epsilon_{\mathsf{A}}$ in the $\mathcal{R}$-model, and for any $\alpha \in (1, \infty]$, there exists another adversary B against the security of D with running time T_{B} and advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{B}}(\mathcal{Q})]| = \epsilon_{\mathsf{B}}$ in the $\mathcal{Q}$-model, such that:*

$$T_{\mathsf{B}} \leq \frac{64}{\epsilon_{\mathsf{A}}^2} \log_2 \left(\frac{8 \cdot \mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q})}{\epsilon_{\mathsf{A}}^{2 + \frac{1}{\alpha - 1}}} \right) \cdot (T_{\mathsf{A}} + T_{\mathsf{S}}), \quad \textit{and}$$
$$\epsilon_{\mathsf{B}} \geq \frac{1}{2^{3 + \frac{1}{\alpha - 1}} \cdot \mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q})} \cdot \epsilon_{\mathsf{A}}^{2 + \frac{1}{\alpha - 1}}.$$

Now, we show our result.

Theorem 4. *Let $D = (X, \mathcal{D}_0, \mathcal{D}_1)$, $\mathcal{R}$, and $\mathcal{Q}$ be the same as in Theorem 3. Then, for any adversary A against the security of D with running time T_{A} and advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]| = \epsilon_{\mathsf{A}}$ in the $\mathcal{R}$-model, there exists another adversary B against the security of D with running time T_{B} and advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{B}}(\mathcal{Q})]| = \epsilon_{\mathsf{B}}$ in the $\mathcal{Q}$-model, such that:*

$$T_{\mathsf{B}} = 2T_{\mathsf{A}} + T_{\mathsf{S}} + \tau,$$
$$\epsilon_{\mathsf{B}} \geq \frac{1}{\mathtt{RD}_2(\mathcal{R} \| \mathcal{Q})} \cdot \epsilon_{\mathsf{A}}^2 \quad \left(\geq \frac{1}{\mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q})} \cdot \epsilon_{\mathsf{A}}^2 \ \textit{ for any } \alpha \in [2, \infty] \right), \tag{4}$$

where τ represents some (small) constant independent of T_{A} and ϵ_{A}.

Note that the Rényi divergence is non-decreasing regarding the order α (see [23]), and thus $\mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q}) \leq \mathtt{RD}_{\alpha'}(\mathcal{R} \| \mathcal{Q})$ holds for all $\alpha < \alpha'$. Thus, Eq. (4) implies $\epsilon_{\mathsf{B}} \geq \frac{1}{\mathtt{RD}_\alpha(\mathcal{R} \| \mathcal{Q})} \cdot \epsilon_{\mathsf{A}}^2$ for every $\alpha \in [2, \infty]$ as well. Hence, although our reduction is not applicable for $\alpha \in (1, 2)$, otherwise ours strictly improves and is much simpler

[12] We stress that S is not given as input the parameter r used to generate a sample x, but may instead infer whatever it needs to know from x for generating x'.

and tighter than the reduction of Bai et al. [3] for every $\alpha \in [2, \infty]$, both in terms of the running time and the distinguishing advantage of the reduction algorithm B. More concretely, the ratio $T_{\mathsf{B}} \cdot \epsilon_{\mathsf{B}}^{-1}$ (sometimes called the *work factor*, and a smaller value means a tighter reduction) of our reduction is $\approx 2 \cdot \mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) \cdot T_{\mathsf{A}} \cdot \epsilon_{\mathsf{A}}^{-2}$, while that of Bai et al. is as large as $\widetilde{O}(\mathtt{RD}_\alpha(\mathcal{R}\|\mathcal{Q}) \cdot T_{\mathsf{A}} \cdot \epsilon_{\mathsf{A}}^{-(4+\frac{1}{\alpha-1})})$.[13] It is $\widetilde{O}(\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) \cdot T_{\mathsf{A}} \cdot \epsilon_{\mathsf{A}}^{-5})$ for $\alpha = 2$ and $\widetilde{O}(\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \cdot T_{\mathsf{A}} \cdot \epsilon_{\mathsf{A}}^{-4})$ for $\alpha = \infty$.

Proof of Theorem 4. We will show that a distinguishing problem with public samplability satisfies $(2T+T_{\mathsf{S}}+\tau, T, 0)$-simulatability in the sense of Definition 4, and then invoke Theorem 2 to conclude the proof.

To see that any distinguishing problem with public samplability $D = (X, \mathcal{D}_0, \mathcal{D}_1)$ satisfies $(2T + T_{\mathsf{S}} + \tau, T, 0)$-simulatability, consider an adversary A against the security of D with running time T, and consider the corresponding adversary B' against the security of D for showing $(2T + T_{\mathsf{S}} + \tau, T, 0)$-simulatability, which interacts with the challenger $\mathsf{C}(r)$ as follows:

B' is initially given a sample x chosen according to $\mathcal{D}_b(r)$, where b is the challenge bit chosen by $\mathsf{C}(r)$ in the security game of B' (and r is sampled according to $\mathcal{Q}$, possibly unknown to B'). Then, B' picks the challenge bit $b_1 \in \{0, 1\}$ in the "first run" for A uniformly at random, and generates $x_1 \leftarrow_{\mathtt{R}} \mathsf{S}(b_1, x)$. B' then executes A twice, first with input x_1 and second with input x, where for each execution B' uses a fresh randomness for A. Let b'_1 (resp. b'_2) be the output of A in the first (resp. second) run of A.

By design, the running time of this B' is $2T + T_{\mathsf{S}} + \tau$ for some small τ independent of A. Furthermore, due to the property of S, B' simulates the challenger $\mathsf{C}(r)$ perfectly for A in both of the executions, so that the challenge bit for A in the first (resp. second) execution is b_1 (resp. b). Also, there is no notion of failure predicate in a distinguishing problem. Consequently, B' satisfies all the properties of $(2T + T_{\mathsf{S}} + \tau, T, 0)$-simulatability.

Then, by Theorem 2, for any adversary A against the security of D with running time T_{A} and advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{A}}(\mathcal{R})]| = \epsilon_{\mathsf{A}}$ in the $\mathcal{R}$-model, there exists another adversary B against the security of D with running time $T_{\mathsf{B}} = 2T_{\mathsf{A}} + T_{\mathsf{S}} + \tau$ and advantage $|\mathbf{E}[\mathtt{Adv}_{\mathsf{B}}(\mathcal{Q})]| = \epsilon_{\mathsf{B}}$ in the $\mathcal{Q}$-model satisfying $\epsilon_{\mathsf{A}} \le \sqrt{\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) \cdot \epsilon_{\mathsf{B}}}$. This inequality is equivalent to Eq. (4). □ (**Theorem 4**)

5.2 Switching Distributions in Differential Privacy

Intuitively, a differentially private mechanism (for some statistical task) takes a data set D as input, and uses its internal randomness r (typically called "noise" in the context of differential privacy) to produce a "sanitized" version of a true

[13] If we adopt the approach of Micciancio and Walter [18] that regards ("running time") · ("advantage")$^{-2}$ (which corresponds to the steps needed to solve a distinguishing problem with a constant advantage) of the best adversary as the "bit security" of a problem, the difference between our reduction and that of Bai et al. will be even larger.

answer computed from D so that it is hard to tell whether any individual's data was included in D. It is often the case that the distribution of a randomness r used in a differentially private mechanism is not the uniform distribution, e.g. the Laplace distribution [12].

Here, we would like to consider the problem of how differential privacy in the setting where a randomness is drawn from an ideal distribution $\mathcal{Q}$ is degraded if we use a randomness drawn from another distribution $\mathcal{R}$. Using one of our lemmas in Section 3, we show a simple technique to assess differential privacy under such switching of distributions of a randomness via the ∞-Rényi divergence.

We note that the connection between differential privacy and the ∞-Rényi divergence is almost immediate from their definitions, and has already been mentioned in existing works (say, [17]). However, we are not aware of any work that formally states a statement in the form that we show below. We also note that our result only covers the case where randomness distributions are discrete, while many works on differential privacy use continuous distributions.

Below, we recall the definition of a differentially private mechanism and then give our result. (We adopt the so-called *approximate differential privacy* [13].)

Let $n \geq 1$ and let $\mathcal{D}$ be the data space. We say that two data sets $D, D' \in \mathcal{D}^n$ are *neighboring* if D and D' have exactly one distinct entry. As in [10], we parameterize differential privacy with not only the privacy budget (ϵ and δ) but also the distribution of randomness used by a mechanism.

Definition 6. *Let $n \geq 1$ and $\mathcal{D}$ be as above. Let $\mathsf{M} : \mathcal{D}^n \to R$ be a probabilistic algorithm whose randomness space is some finite set X. Let $\epsilon, \delta \geq 0$ be real numbers, and let $\mathcal{X}$ be a distribution over X. We say that M satisfies $(\mathcal{X}, \epsilon, \delta)$-*differential privacy *if for every neighboring data sets $D, D' \in \mathcal{D}^n$ and for every $T \subseteq R$, we have*

$$\Pr_{r \leftarrow_{\mathrm{R}} \mathcal{X}}[\mathsf{M}(D; r) \in T] \leq e^{\epsilon} \cdot \Pr_{r \leftarrow_{\mathrm{R}} \mathcal{X}}[\mathsf{M}(D'; r) \in T] + \delta.$$

Theorem 5. *Let $n \geq 1$ and $\mathcal{D}$ be as above, and let M be a probabilistic algorithm whose randomness space is some set X. Let $\mathcal{R}$ and $\mathcal{Q}$ be distributions such that $[\mathcal{R}] = [\mathcal{Q}] \subseteq X$. Let $\epsilon, \delta \geq 0$. Then, if M satisfies $(\mathcal{Q}, \epsilon, \delta)$-differential privacy, then M also satisfies $(\mathcal{R}, \epsilon', \delta')$-differential privacy, where*

$$\epsilon' = \epsilon + \ln \mathtt{RD}_{\infty}(\mathcal{R} \| \mathcal{Q}) + \ln \mathtt{RD}_{\infty}(\mathcal{Q} \| \mathcal{R}) \qquad \text{and} \qquad \delta' = \mathtt{RD}_{\infty}(\mathcal{R} \| \mathcal{Q}) \cdot \delta.$$

Proof of Theorem 5. Fix arbitrarily neighboring data sets $D, D' \in \mathcal{D}^n$ and $T \subseteq R$. For each $r \in X$, define $f_D(r) := \Pr[\mathsf{M}(D; r) \in T]$ and $f_{D'}(r) := \Pr[\mathsf{M}(D'; r) \in T]$. Note that the range of these functions is $[0, 1]$, and we have $\mathbf{E}[f_D(\mathcal{R})] = \Pr_{r \leftarrow_{\mathrm{R}} \mathcal{R}}[\mathsf{M}(D; r) \in T]$, $\mathbf{E}[f_D(\mathcal{Q})] = \Pr_{r \leftarrow_{\mathrm{R}} \mathcal{Q}}[\mathsf{M}(D; r) \in T]$, and we have similar

equations for $f_{D'}$. Now, for D, D', T, we have

$$\begin{aligned}
&\Pr_{r\leftarrow_{\mathrm{R}}\mathcal{R}}[\mathsf{M}(D;r)\in T] = \mathbf{E}[f_D(\mathcal{R})]\\
&\overset{(*)}{\leq} \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot\mathbf{E}[f_D(\mathcal{Q})]\\
&\overset{(\dagger)}{\leq} \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot\Big(e^\epsilon\cdot\mathbf{E}[f_{D'}(\mathcal{Q})]+\delta\Big)\\
&\overset{(\ddagger)}{\leq} \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot e^\epsilon\cdot\mathtt{RD}_\infty(\mathcal{Q}\|\mathcal{R})\cdot\mathbf{E}[f_{D'}(\mathcal{R})]+\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot\delta\\
&= \mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot e^\epsilon\cdot\mathtt{RD}_\infty(\mathcal{Q}\|\mathcal{R})\cdot\Pr_{r\leftarrow_{\mathrm{R}}\mathcal{R}}[\mathsf{M}(D';r)\in T]+\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})\cdot\delta,
\end{aligned}$$

where the inequalities (*) and (‡) use Corollary 1 and the inequality (†) uses the $(\mathcal{Q},\epsilon,\delta)$-differential privacy of M that implies $\mathbf{E}[f_D(\mathcal{Q})] = \Pr_{r\leftarrow_{\mathrm{R}}\mathcal{Q}}[\mathsf{M}(D;r)\in T] \leq e^\epsilon\cdot\Pr_{r\leftarrow_{\mathrm{R}}\mathcal{Q}}[\mathsf{M}(D';r)\in T]+\delta = e^\epsilon\cdot\mathbf{E}[f_{D'}(\mathcal{Q})]+\delta$. Since the choice of D, D', and T is arbitrary, we can conclude that M satisfies $(\mathcal{R},\epsilon',\delta')$-differential privacy with the claimed ϵ' and δ'. □ (**Theorem 5**)

6 Approximate Sampling with Guaranteed Rényi Divergence Bound using Uniform Randomness

There are a number of (not necessarily cryptographic) applications in which we wish to sample random elements from distributions that are not the uniform distribution, e.g. the discrete Gaussian distribution in lattice-based cryptography (e.g. [20]), the Laplace distribution [14, 12] (and other complicated distributions such as the matrix Bingham distribution [7]) in the literature of differential privacy, to name a few. However, it is not always easy (and sometimes impossible) for computers to sample a randomness that exactly follows a target distribution. Thus, a lot of efforts have been made for approximately sampling a randomness from the target distribution using a randomness drawn from the uniform distribution (over bitstrings), so that the sampling method is implementable by computers. One of the basic approaches used for such approximate sampling of a randomness is the *inversion sampling* method (a.k.a. inverse transform sampling), which is the focus in this section.

We propose two computer-friendly inversion sampling methods for an arbitrary discrete distribution $\mathcal{Q}$ using a randomness drawn from the uniform distribution over bitstrings.

- The first method, given in Section 6.1, has the guarantee that the actual distribution $\mathcal{R}$ of a randomness sampled by our method has a guarantee that the 2-Rényi divergence $\mathcal{R}$ from $\mathcal{Q}$ is upperbounded by some number that depends on the size of the support of the distribution and the bit-length of the randomness. More concretely, when using an n-bit string for each sampling from a distribution $\mathcal{Q}$ the size of whose support is m, then the distribution $\mathcal{R}$ of our first sampling method guarantees $\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) \leq 1 + m/2^n$.

– The second method, given in Section 6.2, has a similar property to the first method, but it has a guaranteed ∞-Rényi divergence bound. Concretely, under the same setting as above, our second sampling method guarantees $\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) \leq (1+\sqrt{m/2^n})^2$.

Throughout this section, for simplicity, we work with distributions over $[m]$ for some $m \in \mathbb{N}$ (but our proposals straightforwardly generalize to distributions with an arbitrary finite support). Note that in this case, a distribution $\mathcal{D}$ can be identified with an m-dimensional vector $(p_1, \ldots, p_m) \in [0,1]^m$ such that $\sum_{i\in[m]} p_i = 1$, where $p_i := \Pr[\mathcal{D} = i]$ for each $i \in [m]$.

Our Approach. Recall that the inversion sampling method is based on the inverse of the cumulative distribution function (CDF) of a distribution $\mathcal{D} = (p_1, \ldots, p_m)$. More specifically, let $c_0 = 0$ and $c_i := \Pr[1 \leq \mathcal{D} \leq i] = \sum_{j\in[i]} p_i$ for all $i \in [m]$, then given a uniformly random value x in the interval $[0,1)$, the sampling method outputs k such that $c_{k-1} \leq x < c_k$. Hence, the problem is reduced to showing how to construct a table of the CDF of a distribution.

Given a target distribution $\mathcal{Q} = (q_1, \ldots, q_m)$, our approach is to consider an approximated version $\mathcal{R} = (r_1, \ldots, r_m)$ of $\mathcal{Q}$ such that (1) each r_i can be described by an n-bit string, and (2) the α-Rényi divergence ($\alpha \in \{2, \infty\}$) of $\mathcal{R}$ from $\mathcal{Q}$ has an upperbound dependent on m and n. Note that (1) means that each r_i is of the form $R_i/2^n$ for some $R_i \in \mathbb{Z}_{\geq 0}$ with $R_i \leq N$ and it holds that $\sum_{i\in[m]} R_i = 2^n$, which in turn implies that any value of the CDF of $\mathcal{R}$ can be expressed by an n-bit string, and thus $\mathcal{R}$ can be exactly sampled by using a uniformly random n-bit string. Hence, to achieve the goal, it is sufficient to show how to construct such $\mathcal{R}$ given $\mathcal{Q}$.

The high-level structure for both of our proposed methods is common, and quite simple and intuitive. For convenience, instead of working with a distribution, we work with its scaled-up version, i.e. a vector $(Q_1, \ldots, Q_m) = (2^n \cdot q_1, \ldots, 2^n \cdot q_m) \in ([0, 2^n])^m$.

1. From the original vector $(Q_1, \ldots, Q_m)$, we construct its "tail-cut" version $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$. That is, if some value Q_i is too small, $\widetilde{Q}_i$ is set as 0, while the suppressed values are distributed (added) to the non-zero positions in $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$ so that $\sum_{i\in[m]} \widetilde{Q}_i = 2^n$ holds.
2. We construct an *integer* vector $(R_1, \ldots, R_m) \in (\mathbb{Z}_{\geq 0})^m$ from $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$, so that each R_i is "close" to $\widetilde{Q}_i$ and $\sum_{i\in[m]} R_i = 2^n$ holds. The resulting integer vector $(R_1, \ldots, R_m)$ is a scaled-up version of our desired distribution $\mathcal{R} = (r_1, \ldots, r_m)$. Note that the distribution $\mathcal{R}$ obtained in this way satisfies the property that each r_i can be represented by an n-bit string.

Although simple in a high-level structure, the details are quite different between our first and second proposed methods due to the difference between the 2-Rényi divergence and the ∞-Rényi divergence. For each method, we have to carefully choose the definition of $(\widetilde{Q}_i)_{i\in[m]}$ (in particular, the threshold for the tail cutting), and how to approximate $(\widetilde{Q}_i)_{i\in[m]}$ by the integer vector $(R_i)_{i\in[m]}$,

so that we have the desired upperbound of the α-Rényi divergence ($\alpha \in \{2, \infty\}$). For the details, see the actual proofs.

Supporting Lemma. In the proofs of both of our sampling methods, we will use the following supporting lemma, whose proof is given in Appendix A.

Lemma 9. *Let $k \in \mathbb{N}$, and let $A = (a_1, \ldots, a_k) \in (\mathbb{R}_{\geq 0})^k$ be a vector satisfying $\sum_{i \in [k]} a_i \in \mathbb{N}$. Then, there exists a constructive procedure for constructing a vector $B = (b_1, \ldots, b_k) \in (\mathbb{Z}_{\geq 0})^k$ satisfying the following two properties:*

1. *$\sum_{i \in [k]} b_i = \sum_{i \in [k]} a_i$.*
2. *For each $i \in [k]$, let $d_i := b_i - a_i$. Then, $|d_i| \leq 1$ holds for all $i \in [k]$, and $|d_i - d_j| \leq 1$ holds for all $i, j \in [k]$.*

6.1 Approximate Sampling with a 2-Rényi-Divergence Bound

The following theorem captures our first sampling method.

Theorem 6. *Let $n, m \in \mathbb{N}$. Let $\mathcal{Q} = (q_1, \ldots, q_m)$ be a distribution whose support is $[m]$. Then, there is a constructive procedure for constructing a distribution $\mathcal{R} = (r_1, \ldots, r_m)$ over $[m]$ satisfying the following two properties:*

- **(Samplable Using Uniform Random Bits):** *Each r_i can be described by using at most n-bits. Namely, for all $i \in [m]$, r_i is of the form $r_i = \frac{R_i}{2^n}$, where $R_i \in \mathbb{Z}_{\geq 0}$ and $R_i \leq 2^n$.*
- **(Upperbound of 2-Rényi Divergence):** *The 2-Rényi divergence of $\mathcal{R}$ from $\mathcal{Q}$ is upperbounded as follows:*

$$\mathtt{RD}_2(\mathcal{R} \| \mathcal{Q}) \leq 1 + \frac{m}{2^n}. \tag{5}$$

Proof of Theorem 6. If $m = 1$, then $q_1 = 1$, and thus by defining $r_1 := 1$, the theorem trivially holds. Hence, from here on we assume $m \geq 2$, i.e. the support of $\mathcal{Q}$ contains at least two elements. Then, first of all, note that for proving the theorem, it is sufficient to consider the case that $\mathcal{Q} = (q_1, \ldots, q_m)$ is "ordered" in the following way:

$$0 < q_1 \leq \cdots \leq q_m < 1. \tag{6}$$

Specifically, for a general "non-ordered" distribution $\mathcal{Q}' = (q'_1, \ldots, q'_m)$ with support $[m]$, let $\pi : [m] \to [m]$ be a permutation such that $\mathcal{Q} = (q_1, \ldots, q_m) = (q'_{\pi^{-1}(1)}, \ldots, q'_{\pi^{-1}(m)})$ satisfying Eq. (6). Then, we construct a distribution $\mathcal{R} = (r_1, \ldots, r_m)$ satisfying the two properties guaranteed by the theorem with respect to the ordered distribution $\mathcal{Q}$, and finally obtain the desired distribution $\mathcal{R}' = (r'_1, \ldots, r'_m)$ by defining $r'_i = r_{\pi(i)}$ for every $i \in [m]$. Then, $\mathcal{R}'$ obtained in this way satisfies the two properties with respect to the original distribution $\mathcal{Q}'$: π preserves the first property of $\mathcal{R}$, and we have $\mathtt{RD}_2(\mathcal{R}' \| \mathcal{Q}') = \sum_{i \in [m]} \frac{r'^2_i}{q'_i} = \sum_{i \in [m]} \frac{r^2_{\pi(i)}}{q_{\pi(i)}} = \sum_{i \in [m]} \frac{r_i^2}{q_i} = \mathtt{RD}_2(\mathcal{R} \| \mathcal{Q})$. Hence, we can focus on the case that $\mathcal{Q}$ is ordered in the sense of Eq. (6).

Let $N := 2^n$, and let $Q_i := N \cdot q_i$ for all $i \in [m]$. Then, $\sum_{i\in[m]} Q_i = N$ holds due to the fact that $\mathcal{Q}$ is a probability distribution, and Eq. (6) implies

$$0 < Q_1 \leq \cdots \leq Q_m < N. \tag{7}$$

Next, we note that showing how to construct a distribution $\mathcal{R} = (r_1, \ldots, r_m)$ satisfying the desired properties with respect to $\mathcal{Q}$, is reduced to showing how to construct an integer vector $(R_1, \ldots, R_m) \in (\mathbb{Z}_{\geq 0})^m$ satisfying $\sum_{i\in[m]} R_i = N$ and

$$Z := \sum_{i\in[m]} \frac{(R_i - Q_i)^2}{Q_i} \leq m. \tag{8}$$

To see this, define the probability distribution $\mathcal{R} = (r_1, \ldots, r_m)$ by $r_i = R_i/N$ for every $i \in [m]$ (which guarantees $\sum_{i\in[m]} r_i = 1$). Then, we have

$$\begin{aligned}
\mathtt{RD}_2(\mathcal{R}\|\mathcal{Q}) = \sum_{i\in[m]} \frac{r_i^2}{q_i} = \sum_{i\in[m]} \frac{(R_i/N)^2}{Q_i/N} &= \frac{1}{N} \sum_{i\in[m]} \frac{((R_i - Q_i) + Q_i)^2}{Q_i} \\
&= \frac{1}{N}\Big(\sum_{i\in[m]} \frac{(R_i - Q_i)^2}{Q_i} + 2 \cdot \sum_{i\in[m]} R_i - \sum_{i\in[m]} Q_i \Big) \\
&\overset{(*)}{=} \frac{1}{N}(Z + 2N - N) = 1 + \frac{Z}{N},
\end{aligned}$$

where the equality (*) is due to $\sum_{i\in[m]} R_i = \sum_{i\in[m]} Q_i = N$. Since $N = 2^n$, the right hand side of the above equality is exactly that of Eq. (5) if $Z \leq m$.

Hence, our task is to show how to construct a vector $(R_1, \ldots, R_m) \in (\mathbb{Z}_{\geq 0})^m$ satisfying $\sum_{i\in[m]} R_i = N$ and Eq. (8). To this end, we introduce the following values m^* and S:

$$\begin{aligned}
m^* &:= \max\Big\{ \ell \in \{0\} \cup [m] \;\Big|\; Q_\ell \leq \frac{1}{2} \;\wedge\; \sum_{i\in[\ell]} Q_i \leq \frac{m - \ell - 1}{2} \Big\}, \\
S &:= \sum_{i\in[m^*]} Q_i,
\end{aligned}$$

where for convenience we define $Q_0 := 0$. Note that the definition of m^* implies $m^* \leq m - 1$. Indeed, $m = m^*$ cannot hold because this and the condition $\sum_{i\in[m^*]} Q_i \leq \frac{m-m^*-1}{2}$ imply $\sum_{i\in[m]} Q_i < 0$, which contradicts $\sum_{i\in[m]} Q_i = N$. Furthermore, due to the definitions of m^* and S, the following inequalities hold, which will be used later in the proof:

Lemma 10.

$$Q_{m^*+1} > \min\Big\{ \frac{1}{2}, \frac{m - m^*}{2(m^* + 1)} \Big\} \quad \textit{and} \quad S \leq \min\Big\{ \frac{m^*}{2}, \frac{m - m^* - 1}{2} \Big\}.$$

Proof of Lemma 10. The definitions of m^* and S directly imply (a) $S = \sum_{i\in[m^*]} Q_i \leq \frac{m-m^*-1}{2}$, (b) $Q_{m^*} \leq \frac{1}{2}$, and (c) either $Q_{m^*+1} > \frac{1}{2}$ or $\sum_{i\in[m^*+1]} Q_i$

$> \frac{m-m^*}{2}$. Here, the condition (b) and Eq. (7) imply $S \leq m^* \cdot Q_{m^*} \leq \frac{m^*}{2}$. Combining this with the condition (a), we immediately obtain $S \leq \min\{\frac{m^*}{2}, \frac{m-m^*-1}{2}\}$.

It remains to show $Q_{m^*+1} > \min\{\frac{1}{2}, \frac{m-m^*}{2(m^*+1)}\}$. Assume towards a contradiction that $Q_{m^*+1} \leq \frac{1}{2}$ and $Q_{m^*+1} \leq \frac{m-m^*}{2(m^*+1)}$ simultaneously hold. Then, on the one hand, $Q_{m^*+1} \leq \frac{1}{2}$ and the condition (c) imply $\sum_{i\in[m^*+1]} Q_i > \frac{m-m^*}{2}$. On the other hand, $Q_{m^*+1} \leq \frac{m-m^*}{2(m^*+1)}$ and Eq. (7) imply $\sum_{i\in[m^*+1]} Q_i \leq (m^*+1) \cdot Q_{m^*+1} \leq \frac{m-m^*}{2}$, and thus we have reached a contradiction. Hence, we can conclude that $Q_{m^*+1} > \min\{\frac{1}{2}, \frac{m-m^*}{2(m^*+1)}\}$ holds as well. □ (**Lemma 10**)

Now, as an intermediate step for constructing the desired vector $(R_1, \ldots, R_m)$, we consider the following modified vector $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$, which is the "tail-cut" version of $(Q_1, \ldots, Q_m)$, such that for every $i \in [m]$:

$$\widetilde{Q}_i := \begin{cases} 0 & \text{if } 1 \leq i \leq m^* \\ Q_i + \frac{S}{m-m^*} & \text{if } m^*+1 \leq i \leq m \end{cases}.$$

(We note that the above definition covers the case of $m^* = 0$, which implies $S = 0$ and thus $\widetilde{Q}_i = Q_i$ for all $i \in [m]$.) Note that $0 \leq \widetilde{Q}_i \leq N$ for all $i \in [m]$, and they preserve the sum N of the original vector $(Q_1, \ldots, Q_m)$:

$$\begin{aligned} \sum_{i\in[m]} \widetilde{Q}_i = \sum_{i=m^*+1}^{m} \widetilde{Q}_i = \sum_{i=m^*+1}^{m} \left(Q_i + \frac{S}{m-m^*}\right) &= \sum_{i=m^*+1}^{m} Q_i + S \\ &= \sum_{i=m^*+1}^{m} Q_i + \sum_{i=1}^{m^*} Q_i = N. \end{aligned}$$

Our target vector $(R_1, \ldots, R_m)$ is constructed by approximating the above defined modified vector $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$ by integers. Specifically, we define $R_1 = \cdots = R_{m^*} = 0$. The remaining values R_i for $i \geq m^*+1$, are constructed by using the supporting lemma (Lemma 9). Specifically, by setting $k := m - m^*$ and $a_i := \widetilde{Q}_{m^*+i}$ for every $i \in [m-m^*]$, we have a vector $A = (a_1, \ldots, a_k)$ satisfying $\sum_{i\in[k]} a_i = \sum_{i=m^*+1}^{m} \widetilde{Q}_i = N$. Then, we apply Lemma 9 to this vector A and obtain a vector $B = (b_1, \ldots, b_k) \in (\mathbb{Z}_{\geq 0})^k$, from which we define $R_{m^*+i} := b_i$ for every $i \in [k] = [m-m^*]$. Note that the first property guaranteed by Lemma 9 implies $\sum_{i\in[k]} b_i = \sum_{i=m^*+1}^{m} R_i = N$, and the second property of the lemma guarantees that we have $|d_i - d_j| \leq 1$ for all $i, j \in [m-m^*]$, where $d_i := b_i - a_i = R_{m^*+i} - \widetilde{Q}_{m^*+i}$ for each $i \in [m-m^*]$.

So far, we have defined the vector $(R_1, \ldots, R_m) \in (\mathbb{Z}_{\geq 0})^m$ that satisfies $\sum_{i\in[m]} R_i = \sum_{i=m^*+1}^{m} R_i = N$. Hence, it remains to show that Eq. (8), i.e.,

$Z \leq m$, is satisfied. Calculating Z gives us the following inequality:

$$Z = \sum_{i=1}^{m} \frac{(R_i - Q_i)^2}{Q_i} = \sum_{i=1}^{m^*} \frac{(0 - Q_i)^2}{Q_i} + \sum_{i=m^*+1}^{m} \frac{(R_i - Q_i)^2}{Q_i}$$
$$\leq S + \frac{1}{Q_{m^*+1}} \cdot \sum_{i=m^*+1}^{m} (R_i - Q_i)^2 \tag{9}$$

where the inequality uses $S = \sum_{i=1}^{m^*} Q_i$ and $Q_{m^*+1} \leq Q_i$ for every $i \geq m^* + 1$.

In order to show that Eq. (9) is further upperbounded by m, we wish to upperbound the sum $\sum_{i=m^*+1}^{m}(R_i - Q_i)^2$. For doing this, we do some preparation. Notice that for every $i \in [m - m^*]$, we have

$$R_{m^*+i} - Q_{m^*+i} = R_{m^*+i} - \widetilde{Q}_{m^*+i} + \widetilde{Q}_{m^*+i} - Q_{m^*+i} = d_i + \frac{S}{m - m^*}.$$

Due to the second property satisfied by a vector obtained via Lemma 9, for all $i, j \in [m - m^*]$, we have

$$\Big| \; (R_{m^*+i} - Q_{m^*+i}) - (R_{m^*+j} - Q_{m^*+j}) \; \Big| = |d_i - d_j| \leq 1. \tag{10}$$

Moreover, we have

$$\sum_{i=m^*+1}^{m} (R_i - Q_i) = \sum_{i=1}^{m} (R_i - Q_i) - \sum_{i=1}^{m^*} (R_i - Q_i) \overset{(*)}{=} N - N + \sum_{i=1}^{m^*} Q_i \overset{(\dagger)}{=} S, \tag{11}$$

where the equality (*) uses $\sum_{i=1}^{m} R_i = \sum_{i=1}^{m} Q_i = N$ and the property that $R_i = 0$ for all $i \in [m^*]$, and the equality (†) is due to the definition of S.

We now use the following supporting lemma to upperbound $\sum_{i=m^*+1}^{m}(R_i - Q_i)^2$, whose proof is given in Appendix B.

Lemma 11. *Let $k \in \mathbb{N}$. Let $A = (a_1, \dots a_k) \in \mathbb{R}^k$ be any vector satisfying $|a_i - a_j| \leq 1$ for all $i, j \in [k]$. Let $\alpha := \sum_{i \in [k]} a_i$. Then,*

$$\sum_{i \in [k]} a_i^2 \leq \frac{\alpha^2}{k} + \frac{k}{4}.$$

Let $k := m - m^*$ and $a'_i := (R_{m^*+i} - Q_{m^*+i})$ for every $i \in [k]$. By Eq. (11), we have $\alpha := \sum_{i \in [k]} a'_i = \sum_{i=m^*+1}^{m}(R_i - Q_i) = S \leq \frac{m-m^*-1}{2} \leq \frac{m-m^*}{2}$. Also, Eq. (10) guarantees that $|a'_i - a'_j| \leq 1$ holds for all $i, j \in [k]$. Then, by applying Lemma 11 to the vector $(a'_1, \dots, a'_k)$, we obtain

$$\sum_{i=m^*+1}^{m} (R_i - Q_i)^2 = \sum_{i \in [k]} a_i'^2 \leq \frac{\alpha^2}{k} + \frac{k}{4} = \frac{S^2}{m - m^*} + \frac{m - m^*}{4} \leq \frac{m - m^*}{2}.$$

Using this inequality and Lemma 10 in Eq. (9), we have

$$\begin{aligned}
Z &\leq S + \frac{1}{Q_{m^*+1}} \cdot \sum_{i=m^*+1}^{m} (R_i - Q_i)^2 \\
&\leq \min\left\{ \frac{m^*}{2},\ \frac{m-m^*-1}{2} \right\} + \max\left\{ 2, \frac{2(m^*+1)}{m-m^*} \right\} \cdot \frac{m-m^*}{2} \\
&= \min\left\{ \frac{m^*}{2},\ \frac{m-m^*-1}{2} \right\} + \max\left\{ m-m^*,\ m^*+1 \right\} \\
&= \begin{cases} m - \frac{m^*}{2} & \text{if } m^* \leq \frac{m-1}{2} \\ \frac{m+m^*+1}{2} & \text{if } m^* > \frac{m-1}{2} \end{cases}.
\end{aligned}$$

Recall that we have $0 \leq m^* \leq m-1$. Hence, regardless of the value m^*, we have $Z \leq m$, as required. This completes the proof of the theorem. □ (**Theorem 6**)

6.2 Approximate Sampling with a ∞-Rényi-Divergence Bound

The following theorem captures our second sampling method.

Theorem 7. *Let $n, m \in \mathbb{N}$. Let $\mathcal{Q} = (q_1, \ldots, q_m)$ be a distribution whose support is $[m]$. Then, there is a constructive procedure for constructing a distribution $\mathcal{R} = (r_1, \ldots, r_m)$ over $[m]$ satisfying the following two properties:*

- **(Samplable Using Uniform Random Bits):** *Each r_i can be described by using at most n-bits. Namely, for all $i \in [m]$, r_i is of the form $r_i = \frac{R_i}{2^n}$, where $R_i \in \mathbb{Z}_{\geq 0}$ and $R_i \leq 2^n$.*
- **(Upperbound of ∞-Rényi Divergence):** *The ∞-Rényi divergence of $\mathcal{R}$ from $\mathcal{Q}$ is upperbounded as follows:*

$$\mathtt{RD}_\infty(\mathcal{R} \| \mathcal{Q}) \leq \left(1 + \sqrt{\frac{m}{2^n}}\right)^2. \tag{12}$$

Proof of Theorem 7. If $m = 1$, then $q_1 = 1$, and thus by defining $r_1 := 1$, the theorem trivially holds. Hence, from here on we assume $m \geq 2$. Then, with exactly the same reason as in the proof of Theorem 6, it is sufficient to consider the case that $\mathcal{Q} = (q_1, \ldots, q_m)$ satisfies the "ordered" condition $0 < q_1 \leq \cdots \leq q_m < 1$.

Let $N = 2^n$, and let $Q_i = N \cdot q_i$ for all $i \in [m]$. Then, $\sum_{i \in [m]} Q_i = N$ holds due to the fact that $\mathcal{Q}$ is a probability distribution, and the "ordered" condition implies

$$0 < Q_1 \leq \cdots \leq Q_m < N. \tag{13}$$

Next, we note that due to the definition of the ∞-Rényi divergence, showing how to construct a distribution $\mathcal{R} = (r_1, \ldots, r_m)$ satisfying Eq. (12), is equivalent to showing how to construct an integer vector $(R_1, \ldots, R_m) \in (\mathbb{Z}_{\geq 0})^m$ satisfying $\sum_{i \in [m]} R_i = N$ and

$$\max_{i \in [m]} \frac{R_i}{Q_i} \leq \left(1 + \sqrt{\frac{m}{N}}\right)^2, \tag{14}$$

since it holds that $\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q}) = \max_{i\in[m]} \frac{r_i}{q_i} = \max_{i\in[m]} \frac{R_i/N}{Q_i/N} = \max_{i\in[m]} \frac{R_i}{Q_i}$.

To this end, we introduce the following values S^*, m^*, and S:

$$S^* := \frac{\sqrt{m}}{\sqrt{N}+\sqrt{m}} \cdot N,$$
$$m^* := \max\Big\{\ \ell \in \{0\}\cup[m] \ \Big|\ \sum_{i\in[\ell]} Q_i \le S^* < \sum_{i\in[\ell+1]} Q_i\ \Big\},$$
$$S := \sum_{i\in[m^*]} Q_i,$$

where for convenience we define $Q_0 := 0$. Here, the definition of S^* may look somewhat sudden and bizarre. This is the value of x that minimizes the function $f(x) := \frac{N}{N-x} + \frac{m}{x}$ in the interval $x \in (0, N)$, so that $f(S^*) = (1+\sqrt{\frac{m}{N}})^2$ holds (which can be checked by considering the zero of the first-order derivative of f). Note that this is the desired upperbound of $\max_{i\in[m]} \frac{R_i}{Q_i}$ to be shown.

Our proof from here on is heading to showing how to bound $\mathtt{RD}_\infty(\mathcal{R}\|\mathcal{Q})$ by using the above minimum. We note that the definition of m^* implies that m^* is strictly smaller than m, because $\sum_{i\in[m^*]} Q_i \le S^* < N = \sum_{i\in[m]} Q_i$. Furthermore, the definitions of m^* and S imply the following inequality, which will be used later in the proof.

Lemma 12.

$$Q_{m^*+1} \ge \frac{S^*}{m}. \tag{15}$$

Proof of Lemma 12. If $m^* = 0$, then we have $Q_1 > S^*$, and thus $Q_{m^*+1} \ge \frac{S^*}{m}$ is trivially satisfied. Hence, from here on we consider the case $m^* \ge 1$.

By the definitions of m^* and S, we have $S = \sum_{i\in[m^*]} Q_i \le S^* < \sum_{i\in[m^*+1]} Q_i = S + Q_{m^*+1}$. This implies $Q_{m^*+1} > S^* - S$. Furthermore, we also have $Q_{m^*+1} \ge \frac{S}{m^*}$ because otherwise (i.e., $Q_{m^*+1} < \frac{S}{m^*}$) we have $0 < Q_1 \le \cdots \le Q_{m^*} < \frac{S}{m^*}$, which implies $\sum_{i\in[m^*]} Q_i = Q_1 + \cdots + Q_{m^*} < m^* \cdot \frac{S}{m^*} = S$, contradicting the definitions of m^* and S.

So far, we have seen $Q_{m^*+1} > S^* - S$ and $Q_{m^*+1} \ge \frac{S}{m^*}$, equivalently,

$$Q_{m^*+1} \ge \max\Big\{\ S^* - S,\ \frac{S}{m^*}\ \Big\}. \tag{16}$$

We now show that Eq. (15) holds regardless of the values m^* and S. This is shown by considering the following two cases covering all possibilities:

Case $S \le \frac{m-1}{m} \cdot S^*$: Note that

$$S \le \frac{m-1}{m} \cdot S^* \iff S \le \Big(1 - \frac{1}{m}\Big) \cdot S^* \iff S^* - S \ge \frac{S^*}{m}.$$

Hence, by Eq. (16), we have $Q_{m^*+1} \ge S^* - S \ge \frac{S^*}{m}$.

Case $S > \frac{m-1}{m} \cdot S^*$: By dividing both sides of the condition of this case by $m^* \geq 1$, we obtain

$$\frac{S}{m^*} > \frac{m-1}{m^*} \cdot \frac{S^*}{m} \overset{(*)}{\geq} \frac{S^*}{m},$$

where the inequality (*) uses $\frac{m-1}{m^*} \geq 1$, which holds because of the condition $m^* < m$. Hence, by Eq. (16), we have $Q_{m^*+1} \geq \frac{S}{m^*} > \frac{S^*}{m}$.

As seen above, $Q_{m^*+1} \geq \frac{S^*}{m}$ holds in any case. □ (**Lemma 12**)

Now, as an intermediate step for constructing the desired vector $(R_1, \ldots, R_m)$, we consider the following modified vector $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$, which is the "tail-cut" version of $(Q_1, \ldots, Q_m)$, such that for every $i \in [m]$:

$$\widetilde{Q}_i := \begin{cases} 0 & \text{if } 1 \leq i \leq m^* \\ Q_i \cdot \frac{N}{N-S} & \text{if } m^*+1 \leq i \leq m \end{cases}.$$

(We note that the above definition covers the case of $m^* = 0$, which implies $S = 0$ and thus $\widetilde{Q}_i = Q_i$ for all $i \in [m]$.) Note that $0 \leq \widetilde{Q}_i \leq N$ for all $i \in [m]$, and they preserve the sum N of the original vector $(Q_1, \ldots, Q_m)$:

$$\begin{aligned} \sum_{i \in [m]} \widetilde{Q}_i = \sum_{i=m^*+1}^{m} \widetilde{Q}_i = \frac{N}{N-S} \cdot \sum_{i=m^*+1}^{m} Q_i &= \frac{N}{N-S} \cdot \Big(\sum_{i \in [m]} Q_i - \sum_{i \in [m^*]} Q_i \Big) \\ &= \frac{N}{N-S} \cdot (N-S) = N. \end{aligned}$$

Our target vector $(R_1, \ldots, R_m)$ is constructed by approximating the above defined modified vector $(\widetilde{Q}_1, \ldots, \widetilde{Q}_m)$ by integers, in the same manner as what we do in the proof of Theorem 6. Specifically, we define $R_1 = \cdots = R_{m^*} = 0$. The remaining values R_i for $i \geq m^*+1$, are constructed by using the supporting lemma (Lemma 9). Specifically, by setting $k := m - m^*$ and $a_i := \widetilde{Q}_{m^*+i}$ for every $i \in [m - m^*]$, we have a vector $A = (a_1, \ldots, a_k)$ satisfying $\sum_{i \in [k]} a_i = \sum_{i=m^*+1}^{m} \widetilde{Q}_i = N$. Then, we apply Lemma 9 to this vector A and obtain a vector $B = (b_1, \ldots, b_k) \in (\mathbb{Z}_{\geq 0})^k$, from which we define $R_{m^*+i} := b_i$ for every $i \in [k] = [m - m^*]$. Note that the first property guaranteed by Lemma 9 implies $\sum_{i \in [k]} b_i = \sum_{i=m^*+1}^{m} R_i = N$, and the second property of the lemma guarantees that we have $|d_i| \leq 1$ for all $i \in [m - m^*]$, where $d_i := b_i - a_i = R_{m^*+i} - \widetilde{Q}_{m^*+i}$ for each $i \in [m - m^*]$.

So far, we have defined the vector $(R_1, \ldots, R_m) \in (\mathbb{Z}_{\geq 0})^m$ that satisfies $\sum_{i \in [m]} R_i = \sum_{i=m^*+1}^{m} R_i = N$. Hence, it remains to show $\max_{i \in [m]} \frac{R_i}{Q_i} \leq (1 + \sqrt{\frac{m}{N}})^2$. To this end, we use the following lemma as an intermediate step.

Lemma 13. *For every $i \in [m]$, we have*

$$\frac{R_i}{Q_i} \leq \frac{N}{N-S} + \frac{1}{Q_{m^*+1}}. \tag{17}$$

Proof of Lemma 13. For $i \in [m^*]$, we have $\frac{R_i}{Q_i} = 0$ due to $R_i = 0$, and thus Eq. (17) is trivially satisfied.

For showing the remaining case, fix any $i \in \{m^* + 1, \ldots, m\}$. Recall that $|R_i - \widetilde{Q}_i| \leq 1$ holds due to the second property of the vector obtained from Lemma 9, and thus we have $R_i \leq \widetilde{Q}_i + 1$. Dividing both sides of this inequality by $Q_i > 0$, we have

$$\frac{R_i}{Q_i} \leq \frac{\widetilde{Q}_i}{Q_i} + \frac{1}{Q_i} \overset{(*)}{=} \frac{N}{N-S} + \frac{1}{Q_i} \overset{(\dagger)}{\leq} \frac{N}{N-S} + \frac{1}{Q_{m^*+1}},$$

where the equality (*) uses the definition of $\widetilde{Q}_i$ for $i \in \{m^* + 1, \ldots, m\}$, and the inequality (†) uses $Q_{m^*+1} \leq Q_i$ for all $i \in \{m^* + 1, \ldots, m\}$, which is due to the "ordered" condition (Eq. (13)). The above shows that Eq. (17) is satisfied for $i \in \{m^* + 1, \ldots, m\}$ as well. □ (**Lemma 13**)

Now, combining Lemmas 12 and 13, we obtain

$$\max_{i \in [m]} \frac{R_i}{Q_i} \leq \frac{N}{N-S} + \frac{m}{S^*} \overset{(*)}{\leq} \frac{N}{N-S^*} + \frac{m}{S^*} \overset{(\dagger)}{=} \left(1 + \sqrt{\frac{m}{N}}\right)^2,$$

where the inequality (*) is due to $S \leq S^*$, and the equality (†) is just a direct calculation. (As mentioned earlier, S^* is the value minimizing the function $f(x) = \frac{N}{N-x} + \frac{m}{x}$ in the domain $0 < x < N$ such that we have $f(S^*) = (1 + \sqrt{\frac{m}{N}})^2$.) This completes the proof of the theorem. □ (**Theorem 7**)

Acknowledgement. The authors would like to thank the anonymous reviewers of PKC 2019 for their helpful comments.

References

1. E. Alkim, L. Ducas, T. Pöppelmann, and P. Schwabe. Post-quantum key exchange - A new hope. In *Proc. of USENIX Security 2016*, pages 327–343. USENIX Association, 2016.
2. B. Applebaum, D. Cash, C. Peikert, and A. Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *Proc. of CRYPTO 2009*, volume 5677 of *LNCS*, pages 595–618. Springer, 2009.
3. S. Bai, A. Langlois, T. Lepoint, D. Stehlé, and R. Steinfeld. Improved security proofs in lattice-based cryptography: Using the Rényi divergence rather than the statistical distance. In *Proc. of ASIACRYPT 2015 Part 1*, volume 9452 of *LNCS*, pages 3–24. Springer, 2015.
4. B. Barak, Y. Dodis, H. Krawczyk, O. Pereira, K. Pietrzak, F.-X. Standaert, and Y. Yu. Leftover hash lemma, revisited. In *Proc. of CRYPTO 2011*, volume 6841 of *LNCS*, pages 1–20. Springer, 2011.
5. B. Barak, R. Shaltiel, and E. Tromer. True random number generators secure in a changing environment. In *Proc. of CHES 2003*, volume 2779 of *LNCS*, pages 166–180. Springer, 2003.

6. A. Bogdanov, S. Guo, D. Masny, S. Richelson, and A. Rosen. On the hardness of learning with rounding over small modulus. In *Proc. of TCC 2016-A Part 1*, volume 9562 of *LNCS*, pages 209–224. Springer, 2016.
7. K. Chaudhuri, A.D. Sarwate, and K. Sinha. Near-optimal differentially private principal components. In *Proc. of NIPS 2012*, pages 998–1006, 2012.
8. Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM J. Comput.*, 38(1):97–139, 2008.
9. Y. Dodis, T. Ristenpart, and S.P. Vadhan. Randomness condensers for efficiently samplable, seed-dependent sources. In *Proc. of TCC 2012*, volume 7194 of *LNCS*, pages 618–635. Springer, 2012.
10. Y. Dodis and Y. Yao. Privacy with imperfect randomness. In *Proc. of CRYPTO 2015 Part 2*, volume 9216 of *LNCS*, pages 463–482. Springer, 2015.
11. Y. Dodis and Y. Yu. Overcoming weak expectations. In *Proc. of TCC 2013*, volume 7785 of *LNCS*, pages 1–22. Springer, 2013.
12. C. Dwork. Differential privacy. In *ICALP 2006 Part 2*, volume 4052 of *LNCS*, pages 1–12. Springer, 2006.
13. C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *Proc. of EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 486–503. Springer, 2006.
14. C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Proc. of TCC 2006*, volume 3876 of *LNCS*, pages 265–284. Springer, 2006.
15. C. Gentry and S. Halevi. Implementing Gentry's fully-homomorphic encryption scheme. In *Proc. of EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 129–148. Springer, 2011.
16. C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proc. of STOC 2008*, pages 197–206. ACM, 2008.
17. D. Micciancio and M. Walter. Gaussian sampling over the integers: Efficient, generic, constant-time. In *Proc. of CRYPTO 2017 Part 2*, volume 10402 of *LNCS*, pages 455–485. Springer, 2017.
18. D. Micciancio and M. Walter. On the bit security of cryptographic primitives. In *Proc. of EUROCRYPT 2018 Part 1*, volume 10820 of *LNCS*, pages 3–28. Springer, 2018.
19. T. Prest. Sharper bounds in lattice-based cryptography using the Rényi divergence. In *Proc. of ASIACRYPT 2017 Part 1*, volume 10624 of *LNCS*, pages 347–374. Springer, 2017.
20. O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proc. of STOC 2005*, pages 84–93. ACM, 2005.
21. A. Rényi. On measures of entropy and information. In *Proc. of Fourth Berkeley Symp. on Math. Statist. and Prob.*, volume 1, pages 547–561. Univ. of Calif. Press, 1961.
22. K. Takashima and A. Takayasu. Tighter security for efficient lattice cryptography via the Rényi divergence of optimized orders. In *Proc. of ProvSec 2015*, volume 9451 of *LNCS*, pages 412–431. Springer, 2015.
23. E. van Erven and P. Harremoës. Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. Information Theory*, 60(7):3797–3820, 2014.
24. Y. Yao and Z. Li. Overcoming weak expectations via the Rényi entropy and the expanded computational entropy. In *Proc. of ICITS 2013*, volume 8317 of *LNCS*, pages 162–178. Springer, 2013.

A Proof of Lemma 9

Let $k \in \mathbb{N}$ and $A = (a_1, \ldots, a_k) \in (\mathbb{R}_{\geq 0})^k$ such that $\sum_{i \in [k]} a_i \in \mathbb{N}$. For each $i \in [k]$, let $\delta_i := \lceil a_i \rceil - a_i$. Also, define

$$\Delta := \sum_{i \in [k]} \delta_i = \sum_{i \in [k]} \lceil a_i \rceil - \sum_{i \in [k]} a_i.$$

Note that since $\sum_{i \in [k]} a_i \in \mathbb{N}$ and $\delta_i \in [0, 1)$ for every $i \in [k]$, the definition of Δ implies $\Delta \in \mathbb{Z}_{\geq 0}$ and $\Delta < k$. Furthermore, let $\mathcal{S}_{\mathtt{up}}$ and $\mathcal{S}_{\mathtt{low}}$ be subsets of $[k]$ satisfying the following four conditions:

- (1) $\mathcal{S}_{\mathtt{up}} \cup \mathcal{S}_{\mathtt{low}} = [k]$
- (2) $\mathcal{S}_{\mathtt{up}} \cap \mathcal{S}_{\mathtt{low}} = \emptyset$
- (3) $|\mathcal{S}_{\mathtt{up}}| = \Delta$
- (4) $\max\{\delta_i | i \in \mathcal{S}_{\mathtt{low}}\} \leq \min\{\delta_i | i \in \mathcal{S}_{\mathtt{up}}\}$

Using them, define the vector $B = (b_1, \ldots, b_k)$ such that for every $i \in [k]$,

$$b_i := \begin{cases} \lceil a_i \rceil & \text{if } i \in \mathcal{S}_{\mathtt{low}} \\ \lceil a_i \rceil - 1 & \text{if } i \in \mathcal{S}_{\mathtt{up}} \end{cases}.$$

By definition, every b_i is an integer. Since $a_i \in \mathbb{R}_{\geq 0}$ for every $i \in [k]$, we have $b_i \in \mathbb{Z}_{\geq 0}$ for every $i \in \mathcal{S}_{\mathtt{low}}$. Note also that by the definitions of Δ and $\mathcal{S}_{\mathtt{up}}$, we have $|\{i \in [k] | \delta_i > 0\}| \geq \Delta = |\mathcal{S}_{\mathtt{up}}|$, and thus $a_i > 0$ holds for every $i \in \mathcal{S}_{\mathtt{up}}$, which implies $b_i = \lceil a_i \rceil - 1 \geq 0$ for every $i \in \mathcal{S}_{\mathtt{up}}$. Hence, we have $B = (b_1, \ldots, b_k) \in (\mathbb{Z}_{\geq 0})^k$.

In the following we confirm that the vector B defined above satisfies both of the properties. Regarding the first property, we have

$$\begin{aligned} \sum_{i \in [k]} b_i &= \sum_{i \in \mathcal{S}_{\mathtt{low}}} \lceil a_i \rceil + \sum_{i \in \mathcal{S}_{\mathtt{up}}} (\lceil a_i \rceil - 1) \\ &\overset{(*)}{=} \sum_{i \in [k]} \lceil a_i \rceil - \Delta = \sum_{i \in [k]} (a_i + \delta_i) - \sum_{i \in [k]} \delta_i = \sum_{i \in [k]} a_i, \end{aligned}$$

where the equality (*) uses $|\mathcal{S}_{\mathtt{up}}| = \Delta$. Hence, B satisfies the first property.

It remains to show that B satisfies the second property. For each $i \in [k]$, let

$$d_i := b_i - a_i = \begin{cases} \delta_i & \text{if } i \in \mathcal{S}_{\mathtt{low}} \\ \delta_i - 1 & \text{if } i \in \mathcal{S}_{\mathtt{up}} \end{cases}.$$

Recall that $\delta_i \in [0, 1)$ holds for every $i \in [k]$. Thus, we have $|d_i| \leq 1$ for all $i \in [k]$. Furthermore, for every $(i, j) \in [k]^2$, we have

$$|d_i - d_j| = \begin{cases} |\delta_i - \delta_j| & \text{if } (i, j) \in (\mathcal{S}_{\mathtt{low}})^2 \text{ or } (i, j) \in (\mathcal{S}_{\mathtt{up}})^2 \\ |1 - (\delta_j - \delta_i)| & \text{if } (i, j) \in \mathcal{S}_{\mathtt{low}} \times \mathcal{S}_{\mathtt{up}} \\ |\delta_i - \delta_j - 1| & \text{if } (i, j) \in \mathcal{S}_{\mathtt{up}} \times \mathcal{S}_{\mathtt{low}} \end{cases}.$$

From the above, it is immediate that $|d_i - d_j| \leq 1$ holds for the cases $(i,j) \in (\mathcal{S}_{\mathtt{low}})^2$ and $(i,j) \in (\mathcal{S}_{\mathtt{up}})^2$. Also, for the case $(i,j) \in \mathcal{S}_{\mathtt{low}} \times \mathcal{S}_{\mathtt{up}}$, we have $\delta_i \leq \delta_j$ due to the condition (4) of $\mathcal{S}_{\mathtt{low}}$ and $\mathcal{S}_{\mathtt{up}}$, and thus we have $|1 - (\delta_j - \delta_i)| \leq 1$. Similarly, for the case $(i,j) \in \mathcal{S}_{\mathtt{up}} \times \mathcal{S}_{\mathtt{low}}$, we have $\delta_i \geq \delta_j$, and thus we have $|\delta_i - \delta_j - 1| \leq 1$. Hence, we have $|d_i - d_j| \leq 1$ for any pair $(i,j) \in [k]^2$. This shows that the vector B satisfies the second property as well. □ (**Lemma 9**)

B Proof of Lemma 11

Fix arbitrarily a number $k \in \mathbb{N}$ and a vector $(a_1, \ldots, a_n) \in \mathbb{R}^k$ satisfying $|a_i - a_j| \leq 1$ for all $i, j \in [k]$, and let $\alpha := \sum_{i \in [k]} a_i$. We will show that $\sum_{i \in [k]} a_i^2 \leq \frac{\alpha^2}{k} + \frac{k}{4}$ holds, which proves the lemma.

Let $a_{\min} := \min\{a_i\}_{i \in [k]}$, and $\delta_i := a_i - a_{\min}$ for each $i \in [k]$. Note that due to the given condition of the vector $(a_1, \ldots, a_k)$, $\delta_i \in [0,1]$ holds for all $i \in [k]$. We also have

$$\begin{aligned} & \alpha = \sum_{i \in [k]} a_i = \sum_{i \in [k]} \Big(a_{\min} + \delta_i \Big) = k a_{\min} + \sum_{i \in [k]} \delta_i \\ \iff \quad & \sum_{i \in [k]} \delta_i = \alpha - k a_{\min}. \end{aligned} \tag{18}$$

Furthermore, for each $i \in [k]$, we have

$$\begin{aligned} a_i^2 = (a_{\min} + \delta_i)^2 &= a_{\min}^2 + 2 a_{\min} \delta_i + \delta_i^2 \\ &\leq a_{\min}^2 + (2 a_{\min} + 1) \cdot \delta_i, \end{aligned} \tag{19}$$

where the inequality uses $\delta_i^2 \leq \delta_i$, which is due to $\delta_i \in [0,1]$.

Now, consider the sum of squares $\sum_{i \in [k]} a_i^2$. We have

$$\begin{aligned} \sum_{i \in [k]} a_i^2 &\overset{(*)}{\leq} \sum_{i \in [k]} \Big(a_{\min}^2 + (2 a_{\min} + 1) \cdot \delta_i \Big) \\ &= k a_{\min}^2 + (2 a_{\min} + 1) \cdot \sum_{i \in [k]} \delta_i \\ &\overset{(\dagger)}{=} k a_{\min}^2 + (2 a_{\min} + 1) \cdot (\alpha - k a_{\min}) \\ &= -k \Big(a_{\min} - \Big(\frac{\alpha}{k} - \frac{1}{2} \Big) \Big)^2 + \frac{\alpha^2}{k} + \frac{k}{4} \\ &\leq \frac{\alpha^2}{k} + \frac{k}{4}, \end{aligned}$$

where the inequality (*) uses Eq. (19), and the equality (†) uses Eq. (18). □ (**Lemma 11**)