# Probability Preservation Property with Relative Error and Its Applications

Yuanyuan Gao[1,2,3] and Kunpeng Wang[1,2,3]
[1]*State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 100093, China*
[2]*Data Assurance and Communication Security Research Center, Chinese Academy of Sciences, Beijing, 100093, China*
[3]*School of Cyber Security, University of Chinese Academy of Sciences, Beijing, 100049, China*

Keywords: Probability Preservation, Rényi Divergence, Preimage Sampleable Functions, Gaussian Sampling Algorithm, Smoothing Parameter.

Abstract: Probability preservation property plays an important part in security proofs of lattice-based cryptography, which bounds the closeness of two probability distributions. Recent works revolve around different measures. We reform probability preservation properties with relative error which simplify analysis of the security reductions of preimage sampleable functions (PSFs) via different measures and demonstrate Rényi divergence with order $\infty$ ($RD_\infty$) can coordinate performance with security well. We apply $RD_\infty$-based reduction to PSFs over lattices, which reduces the smoothing parameter of Gaussian sampling algorithm by a factor $\mathbf{O}(\sqrt{\lambda})$ without security loss. We further extend the optimized parameter to the secret extraction of identity-based encryption (IBE) over the general lattices by Gentry et al. in STOC 2008 and NTRU lattices proposed by Ducas et al. in Asiacrypt 2014. As a consequence, the size of secret key can be shortened by a factor $\mathbf{O}(\sqrt{\lambda})$ accordingly.

## 1 INTRODUCTION

**Background.** An essential tool named probability preservation property in cryptography is the use of divergence measures to prove the security of lattice-based cryptographic schemes. To measure the closeness, the classical statistical distance (SD) is naturally adopted in probability preservation property.

As an introductory example, the SD-based probability preservation property works as follows: Let $P$ ($Q$) be the probability distributions in the real (ideal) scheme. If the adversary succeeds with non-negligible probability under the distribution $P$, and if the SD between $P$ and $Q$ is negligible, then the adversary succeeds with non-negligible probability under the distribution $Q$. Similar to the case of SD, Kullback-Leibler Divergence (KL) (Pöppelmann et al., 2014; Ducas et al., 2014) and Rényi divergence (RD) (Langlois et al., 2014; Bai et al., 2015; Takashima and Takayasu, 2015; Prest, 2017) enjoy analogical property and have been demonstrated to enable security reductions with better parameters than SD for search problems. A new metric called max-log distance was introduced in (Micciancio and Walter, 2017). The new measure is closely related to the standard notion of relative error and the Rényi divergence of order $\infty$.

**Our Contribution.** The contribution of this paper is twofold. In theory, we propose RD-based probability preservation properties with relative error in the reduction proof of one-way preimage sampleable functions. In practice, we improve the Gaussian sampling algorithm (Gentry et al., 2008) and the secret key extraction of the IBE scheme in (Gentry et al., 2008; Ducas et al., 2014) with the Rényi divergence.

Firstly, we explicitly express the RD-based probability preservation property with relative error by two theorems. Our theoretical result is not limited to the Gaussian distribution and can be extended to other distributions by means of Taylor series expansion. Furthermore, it is convenient for analysis with our uniform representations. We give an immediate conclusion that $RD_\infty$ outperforms other measures which enjoys both the good performance and the tight reduction.

Then with the above technique, we improve the smoothing parameter of Gaussian sampling algorithm (Gentry et al., 2008) and the secret key size of the IBE



Gao, Y. and Wang, K.
Probability Preservation Property with Relative Error and Its Applications.
DOI: 10.5220/0006717204610468
In *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018)*, pages 461-468
ISBN: 978-989-758-282-0

scheme in (Gentry et al., 2008; Ducas et al., 2014). In the security proof of the Gaussian sampling algorithm (Gentry et al., 2008), the parameter named tolerance in the Gaussian sampling algorithm can be relaxed, which results in the smaller smoothing parameter by $\mathbf{O}(\sqrt{\lambda})$. Thus shorter vectors are outputted without security loss. Accordingly, the user secret key size in the IBE scheme (Gentry et al., 2008; Ducas et al., 2014) is shortened by a factor $\mathbf{O}(\sqrt{\lambda})$.

**Paper Organization.** This paper is organized as follows. In Sect 2, we introduce some basic definitions and properties about discrete Gaussian distributions on lattices and measures of probability distributions. In Sect 3, We provide the security reductions of one-way PSFs with relative error in order to give straightforward analysis and comprehensive comparisons. In Sect 4, we improve the smoothing parameter of Gaussian sampling algorithm and the user secret keys of the IBE schemes over general lattices and NTRU lattices. Finally, Sect 5 is the conclusion.

# 2 PRELIMINARIES

## 2.1 Notation

Some notations throughout the paper are listed below.

The security parameter is $\lambda$. A negligible function, denoted by $negl(\lambda)$, is a function $f(\lambda)$ such that $f(\lambda) = \lambda^{-\omega(1)}$. The probability $1 - negl(\lambda)$ indicates that it is overwhelming. $f(\lambda) = \omega(\sqrt{\log\lambda})$ denotes that $f(\lambda)$ grows asymptotically faster than $\sqrt{\log\lambda}$. $f(\lambda) = \widetilde{O}(g(\lambda))$ denotes $f(\lambda) = O(g(\lambda)\cdot\log^c\lambda)$. $m = poly(n)$ shows that $m$ is polynomial times of $n$.

We denote vector by lower-case bold letters (e.g., $\mathbf{v}$) and denote matrices by upper-case bold letters (e.g., $\mathbf{B}$). The $i$th column vector of a matrix $\mathbf{B}$ is denoted $\mathbf{b}_i$. For a vector $\mathbf{v} \in \mathbb{R}^m$, $\|\mathbf{v}\| = \sqrt{<\mathbf{v},\mathbf{v}>}$ denotes its Euclidean norm. The norm of a matrix $\mathbf{B} \in \mathbb{R}^{m\times n}$ is the maximal norm of its columns: $\|\mathbf{B}\| = \max_{i=1}^{n}\|\mathbf{b}_i\|$.

## 2.2 Lattices and Gaussian

Let $\mathbf{B} = \{\boldsymbol{b}_1, ..., \mathbf{b_n}\} \subseteq \mathbb{R}^m$ consist of $n$ linearly independent vectors. The lattice $\Lambda$ generated by $B$ is defined as $\Lambda = \mathcal{L}(\mathbf{B}) = \{\sum_{i=1}^{n} z_i\mathbf{b_i} : \mathbf{z} \in \mathbb{Z}^n\}$, where $\mathbf{B}$ is a basis of the lattice, the rank of the lattice is $n$, and the dimension of the lattice is $m$. When $m = n$, the lattice is full rank. In Gaussian sampling algorithm, we work on non-full rank lattice.

For a lattice $\Lambda$ and any $i \le n$, the $i$th successive minimum $\lambda_i(\Lambda)$ is the smallest radius $r$ such that the lattice points inside a ball of radius $r$ span a space of dimension $i$.

For a full-rank lattice $\Lambda$, its dual lattice $\Lambda^*$ is defined as $\Lambda^* = \{\mathbf{x} \in \mathbb{R}^m : \forall \mathbf{v} \in \Lambda, \langle\mathbf{x},\mathbf{v}\rangle \in \mathbb{Z}\}$.

We usually use q-ary integer lattice in the cryptography primitive constructions below.

$$\Lambda^{\perp}(\mathbf{A}) = \{\mathbf{z} \in \mathbb{Z}^m : \mathbf{Az} = 0 \mod q\}$$

$$\Lambda(\mathbf{A}^t) = \{\mathbf{z} \in \mathbb{Z}^{\mathbf{m}} : \exists \mathbf{s} \in \mathbb{Z}^{\mathbf{n}} \quad \mathbf{s.t.} \quad \mathbf{z} = \mathbf{A^t s} \mod \mathbf{q}\}.$$

$\Lambda^{\perp}(\mathbf{A})$ and $\Lambda(\mathbf{A}^t)$ up to a $q$ scaling factor are dual lattices: $q\cdot\Lambda^{\perp}(\mathbf{A})^* = \Lambda(\mathbf{A}^t)$ and $q\cdot\Lambda(\mathbf{A}^t)^* = \Lambda^{\perp}(\mathbf{A})$.

For any $\mathbf{u} \in \mathbb{Z}^n$ , $\mathbf{x}$ as an integral solution to $\mathbf{Ax} = \mathbf{u} \mod \mathbf{q}$, the coset is defined $\Lambda^{\perp}_{\mathbf{u}}(\mathbf{A}) = \{\mathbf{z} \in \mathbb{Z}^m : \mathbf{Az} = \mathbf{u} \mod q\} = \Lambda^{\perp}(\mathbf{A}) + \mathbf{x}$.

The preimage sampled in the Gaussian algorithm can be viewed as coset given the basis of $\Lambda^{\perp}(\mathbf{A})$ and the outputs conditional distribution is identical to preimage from Gaussian distribution.

For any vector $\mathbf{c} \in \mathbb{R}^m$ and any real $s > 0$, the (spherical) Gaussian function with standard deviation parameter $s$ and center $\mathbf{c}$ is defined as: $\forall \mathbf{x} \in \mathbb{R}^n, \rho_{s,\mathbf{c}}(\mathbf{x}) = \exp(-\pi\|\mathbf{x}-\mathbf{c}\|/s^2)$.

The (spherical) discrete Gaussian distribution over a lattice $\Lambda \subseteq \mathbb{R}^m$, with standard deviation parameter $s > 0$ and center $\mathbf{c}$ is defined as: $\forall \mathbf{x} \in \Lambda, D_{\Lambda,s,\mathbf{c}} = \frac{\rho_{s,\mathbf{c}}(\mathbf{x})}{\rho_{s,\mathbf{c}}(\Lambda)}$, where Gaussian mass $\rho_{s,\mathbf{c}}(\Lambda) = \sum_{\mathbf{x}\in\Lambda}\rho_{s,\mathbf{c}}(\mathbf{x})$.

$D_{\Lambda,s,\mathbf{c}}$ is merely a normalization factor and it is simply proportional to $\rho_{s,\mathbf{c}}(\mathbf{x})$.

Micciancio and Regev (Micciancio and Regev, 2004) proposed a lattice quantity called the smoothing parameter. We will recall some fundamental properties of the smoothing parameter.

**Lemma 1** ((Gentry et al., 2008), Lemma 3.1)**.** *For any m-dimensional lattice $\Lambda$ with basis $\mathbf{B}$ and $\varepsilon > 0$, we have $\eta_\varepsilon \le \sqrt{\frac{\ln(2m(1+1/\varepsilon))}{\pi}}\cdot\|\widetilde{\mathbf{B}}\|$, where $\|\widetilde{\mathbf{B}}\| = \max_{i=1}^{m}\|\widetilde{\mathbf{b_i}}\|$ denotes the maximal length of the Gram-Schmidt orthogonalized vectors $\widetilde{\mathbf{b_i}}$ of the ordered basis $B = \{\mathbf{b_i}\}$. In particular, for any $\omega(\sqrt{\log m})$ function, there is a negligible $\varepsilon(m)$ for which $\eta_\varepsilon \le \|\widetilde{\mathbf{B}}\|\cdot\omega(\sqrt{\log m})$.*

Many works of cryptography over lattices rely on the discrete Gaussian probability distributions. When $s > \eta_\varepsilon(\Lambda)$, it has the following properties.

**Lemma 2** ((Gentry et al., 2008), Lemma 5.2)**.** *Let $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$, $\varepsilon \in (0, 1/2)$, $s \ge \eta_\varepsilon(\Lambda^{\perp}(\mathbf{A}))$, then for $\mathbf{e} \sim D_{\mathbb{Z}^m}$, the distribution of the syndrome $\mathbf{u} = \mathbf{Ae} \mod q$ is within statistical distance $2\varepsilon$ of uniform over $\mathbb{Z}_q^n$.*

*Furthermore, let $\mathbf{u} \in \mathbb{Z}_q^n$ be fixed and $\mathbf{t} \in \mathbb{Z}^m$ be an arbitrary solution to $\mathbf{At} = \mathbf{u} \mod q$, then given $\mathbf{Ae} =$*

$\mathbf{u} \mod q$, *the conditional distribution of* $\mathbf{e} \sim D_{\mathbb{Z}^m}$ *is* $\mathbf{t} + D_{\Lambda^\perp,s,-\mathbf{t}}$.

The tail inequality of Gaussian distribution in the next lemma shows a sample from $D_{\mathbb{Z}^m,s}$ is at most $s\sqrt{m}$ away from the center with overwhelming probability.

**Lemma 3** ((Micciancio and Regev, 2004), Lemma 4.4)**.** *For any n-dimensional lattice* $\Lambda$, $\mathbf{c} \in \mathbb{R}^m$, $\varepsilon \in (0,1)$ *and* $s > \eta_\varepsilon(\Lambda)$, $\mathbf{c} \in span(\Lambda)$, $\Pr_{x \leftarrow D_{\Lambda,s,\mathbf{c}}}[\|\mathbf{x}-\mathbf{c}\| \geq s\sqrt{m}] \leq \frac{1-\varepsilon}{1+\varepsilon} \cdot 2^{-m}$.

NTRU lattice is a lattice with special algebra structure. For a power-of two integer $N$, a positive integer $q$, and $f, g \in \mathcal{R}$, $\mathcal{R} \triangleq \mathbb{Z}[x]/(x^N+1)$, let $\mathbf{h} = \mathbf{g} * \mathbf{f}^{-1} \mod q$, NTRU lattice is defined as $\Lambda_{\mathbf{h},q} = \{(\mathbf{u},\mathbf{v}) \in \mathcal{R}^2 | \mathbf{u} + \mathbf{v} * \mathbf{h} = 0 \mod q\}$.

The last lemma provides the theoretical lower bound of the maximal length of Gram-Schmidt orthogonalized vectors over NTRU lattices.

**Lemma 4** ((Ducas et al., 2014), Lemma 2)**.** *Let* $\mathbf{B_{f,g}}$ *be a basis of NTRU lattice,* $\mathbf{b_1}, \cdots, \mathbf{b_{2N}}$ *be the row vectors of* $B_{f,g}$*, then* $\|\widetilde{\mathbf{B}}_{\mathbf{f},\mathbf{g}}\| = \max\{\|\widetilde{\mathbf{b}}_1\|, \|\widetilde{\mathbf{b}}_{N+1}\|\}$.

## 2.3 Measures of Distributions Closeness

Referring to (Rényi, 1961; van Erven and Harremoës, 2014), the definitions are given as follows. For convenience, our definition of the RD is the exponential of the usual definition used in information theory(van Erven and Harremoës, 2014).

The support of the distribution $P$ is defined as $Supp(P) = \{x : P(x) \neq 0\}$. $\mathbf{u} \sim U(\mathbb{Z}^n)$ denotes the distribution of $\mathbf{u} \in \mathbb{Z}^n$ is uniform.

Let $P$ and $Q$ be two distributions over the countable support $Supp(P) \subseteq Supp(Q)$.

The Statistical distance between $P$ and $Q$ is defined as

$$\Delta(P\|Q) = \frac{1}{2} \sum_{x \in Supp(P)} |P(x) - Q(x)|.$$

The Kullback-Leibler Divergence of $Q$ from $P$ is defined as:

$$D_{KL}(P\|Q) = \sum_{x \in Supp(P)} P(x) \ln \frac{P(x)}{Q(x)}.$$

For $\alpha \in (1, +\infty)$, the Rényi divergence of order $\alpha$ is defined as:

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

For $\alpha = 1$ and $\alpha = \infty$, Rényi divergences are respectively defined as:

$$R_1(P\|Q) = \exp\left(\sum_{x \in Supp(P)} P(x) \cdot \ln \frac{P(x)}{Q(x)}\right)$$

and

$$R_\infty(P\|Q) = \max_{x \in Supp(P)} \frac{P(x)}{Q(x)}.$$

SD denotes the difference of two probability distributions, RD denotes the ratio of two probability distributions. RD and KL have similar properties with SD as follows:

**Lemma 5** ((van Erven and Harremoës, 2014; Langlois et al., 2014; Bai et al., 2015))**.** *Let P and Q denote probability distributions over the countable set* $Supp(P) \subseteq Supp(Q)$, $\alpha \in [1, +\infty]$. *For any function f:*

- **Additivity.**

$$\Delta(\prod_i P_i \| \prod_i Q_i) = \sum_i \Delta(P_i\|Q_i).$$
$$KL(\prod_i P_i \| \prod_i Q_i) = \sum_i KL(P_i\|Q_i).$$

- **Multiplicativity.**

$$R_\alpha(\prod_i P_i \| \prod_i Q_i) = \prod_i R_\alpha(P_i\|Q_i).$$

- **Data Processing Inequality.**

$$\Delta(P^f\|Q^f) \leq \Delta(P\|Q)$$
$$KL(P^f\|Q^f) \leq KL(P\|Q)$$
$$R_\alpha(P^f\|Q^f) \leq R_\alpha(P\|Q)$$

*where* $P^f$ $(resp.Q^f)$ *denotes the distribution of* $f(x)$ *induced by sampling* $x \leftarrow P (resp.x \leftarrow Q)$.

**Lemma 6** (**Probability Preservation** (van Erven and Harremoës, 2014; Langlois et al., 2014; Bai et al., 2015))**.** *Let P (Q) be the probability distributions over* $Supp(P) \subseteq Supp(Q)$, $A \subseteq Supp(Q)$ *be an arbitrary event.*

*SD-based probability preservation property:*

$$Q(A) \geq P(A) - 2\Delta(P\|Q)$$

*KL-based probability preservation property:*

$$|Q(A) - P(A)| \leq \frac{1}{\sqrt{2}} \sqrt{D_{KL}(P\|Q)}$$

*RD-based probability preservation:*

$$Q(A) \geq \frac{(P(A))^{\frac{\alpha}{\alpha-1}}}{R_\alpha(P\|Q)}$$
$$Q(A) \geq \frac{P(A)^2}{R_2(P\|Q)}$$
$$Q(A) \geq P(A) - \sqrt{\ln R_1(P\|Q)/2}$$
$$Q(A) \geq \frac{P(A)}{R_\infty(P\|Q)}.$$

Since $R_1(P\|Q)$ is the exponential of the Kullback-Leibler divergence. For a random event set $A \subseteq Supp(P) \subseteq Supp(Q)$, then we have $Q(A) \geq P(A) - \sqrt{\ln R_1(P\|Q)/2}$.

# 3 SECURITY REDUCTION

Inspired by the Taylor series expansion technique in (Pöppelmann et al., 2014) and the idea of relative error in (Micciancio and Walter, 2017), we provide the bound of two probability distributions closeness with relative error. Then, we obtain probability preservation properties with relative error. It is easy to analyse and compare the bound of the success probability with our uniform representations.

## 3.1 Bound of Distributions Closeness

We describe the bound of closeness of two probability distributions with relative error by Taylor series, where we give a tight factor by analyzing the tail-cut of the Taylor expansion explicitly.

**Theorem 1** (Bound of Distributions Closeness)**.** *Let $P$ and $Q$ be two probability distributions over the same countable $S$. Assume that for any $x \in S$, there exists $|P(x)-Q(x)| \leq \delta(x)Q(x)$, where $\delta(x) \in (0, \frac{1}{\alpha})$ and $\alpha > 2$. Then we have:*

$$R_\alpha(P\|Q) \leq 1+\alpha\max_{x\in S}\delta(x)^2.$$

$$R_2(P\|Q) \leq 1+\max_{x\in S}\delta(x)^2.$$

$$R_1(P\|Q) \leq 1+\max_{x\in S}\delta(x)^2.$$

$$R_\infty(P\|Q) \leq 1+\max_{x\in S}\delta(x).$$

We also present the SD and KL bounds of distributions closeness with relative error.

SD bound:

$$\Delta(P\|Q) \leq \frac{1}{2}\max_{x\in S}\delta(x).$$

KL bound:

$$D_{KL}(P\|Q) \leq \max_{x\in S}\delta(x)^2.$$

Main proofs of theorem are given in Appendix A. In fact, as the special relation that $R_1(P\|Q)$ is the exponential of KL, the derivation of the bound of distributions closeness via $R_1(P\|Q)$ relies on the computation of KL. The proof of KL was proposed in (Pöppelmann et al., 2014). In essence, it is identical to the work of (Prest, 2017). We extend to other closeness measures and different parameter settings. The essential derivation process amounts to preprocessing distributions closeness bound.

## 3.2 Bound of Success Probability

Let $P$ ($Q$) be the probability distributions over the same countable set $S$ in the real (ideal) scheme. For a probabilistic polynomial-time adversary $\mathcal{A}$ against the scheme, if $\mathcal{A}$ succeeds in the attack, $Suc_{\mathcal{A}}(f(x)) = 1$, otherwise $Suc_{\mathcal{A}}(f(x)) = 0$, where f is any efficiently computable function, then the advantage of adversary $\mathcal{A}$ succeeding in the attack is defined as $\mathrm{Adv}_{\mathcal{A}}(P,Q) = |\Pr_{x\leftarrow P}[Suc_{\mathcal{A}}(f(x))=1] - \Pr_{x\leftarrow Q}[Suc_{\mathcal{A}}(f(x))=1]|$.

Assume that $P$ is the probability distribution over $S$ in the real scheme, $Q$ is the probability distribution over $S$ in the ideal scheme. If $\mathcal{A}$ succeeds with significant probability in the real scheme and $Adv_{\mathcal{A}}(P,Q)$ is negligible, then $\mathcal{A}$ succeeds with significant probability in the ideal scheme. In other words, if the adversary succeeds with negligible probability in the ideal scheme and $Adv_{\mathcal{A}}(P,Q)$ is negligible, then it also succeeds with negligible probability in the real scheme. The property is named the probability preservation property. According to Theorem 1 and Lemma 6, we acquire Theorem 2 by the substitution method.

**Theorem 2** (Bound of Success Probability)**.** *Let $P$ ($Q$) be the probability distribution over $S$ in the real (ideal) scheme. Assume that for any $x \in S$, there exists $|P(x)-Q(x)| \leq \delta(x)Q(x)$, where $\delta(x) \in (0, \frac{1}{\alpha})$ and $\alpha > 2$. $\mathcal{A}$ is a probabilistic polynomial-time adversary making at most $q$ queries an oracle. If $\mathcal{A}$ succeeds in the attack, $Suc_{\mathcal{A}}(f(x)) = 1$, otherwise $Suc_{\mathcal{A}}(f(x)) = 0$, where $f$ is a one-way trapdoor function. Let $\varepsilon_P = \Pr_{x\leftarrow P}[Suc_{\mathcal{A}}(f(x))=1]$, $\varepsilon_Q = \Pr_{x\leftarrow Q}[Suc_{\mathcal{A}}(f(x))=1]$, then*

*RD-based probability preservation property:*

$$RD_\alpha : \varepsilon_Q \geq \frac{(\varepsilon_P)^{\frac{\alpha}{\alpha-1}}}{(1+\alpha\max_{x\in S}\delta(x)^2)^q}.$$

$$RD_2 : \varepsilon_Q \geq \frac{\varepsilon_P^2}{(1+\max_{x\in S}\delta(x)^2)^q}.$$

$$RD_1 : \varepsilon_Q \geq \varepsilon_P - \sqrt{\frac{q}{2}}\max_{x\in S}\delta(x).$$

$$RD_\infty : \varepsilon_Q \geq \frac{\varepsilon_P}{(1+\max_{x\in S}\delta(x))^q}.$$

SD-based probability preservation property:

$$\varepsilon_Q \geq \varepsilon_P - \frac{q}{2}\max_{x\in S}\delta(x).$$

KL-based probability preservation property:

$$|\varepsilon_Q - \varepsilon_P| \leq \sqrt{\frac{q}{2}}\max_{x\in S}\delta(x).$$

From all the formulas above in the Theorem 2, We can give straightforward analysis and comprehensive comparisons. The times of query $q$, the order of $\alpha$ and the max relative error $\max_{x\in S}\delta(x)$ act on the tightness of the reduction. The goal in (Takashima and Takayasu, 2015; Prest, 2017) is to achieve $R_\alpha(P\|Q)^q = \Omega(1)$ ( e.g. $(1+\alpha\max_{x\in S}\delta(x)^2)^q \leq e^{1/4} \leq \sqrt{2}$ ). Our method is different from theirs. We deal with $q$ in the next section, which is closely related to iteration times of sampling algorithm.

In the analysis of the SD-based probability preservation property with relative error, the success probabilities in the ideal and real schemes demand for the same magnitude $\mathbf{O}(2^{-\lambda})$, the advantage is limited to the magnitude $\mathbf{O}(2^{-\lambda})$, so the $\max_{x\in S}\delta(x)$ is limited to the magnitude $\mathbf{O}(2^{-\lambda})$.

In the analysis of the KL-based probability preservation property with relative error, if the success probability of the adversary is the magnitude $\mathbf{O}(2^{-\lambda})$ in both the ideal and real schemes, the $\max_{x\in S}\delta(x)$ is limited to the magnitude $\mathbf{O}(2^{-\lambda})$. Comparing to SD, $\max_{x\in S}\delta(x)$ in the KL-based reduction can be larger by $\sqrt{\frac{q}{2}}$ than $\max_{x\in S}\delta(x)$ in the SD-based reduction remaining the same security, which is one of the main contributions of works in (Pöppelmann et al., 2014).

In the analysis of the RD-based probability preservation property with relative error, the security becomes looser when the order is $\alpha$, particularly to $order = 2$. The security remains almost the same when the order of RD is $\infty$.

We give a comprehensive comparison of distribution and security parameters in Table 1, where $\max_{x\in S}\delta(x)$ range is $\mathbf{O}(2^{-\lambda})$ in SD and KL rows and $\max_{x\in S}\delta(x)$ range is $\mathbf{O}(2^{-\lambda})$ and $\mathbf{O}(\frac{1}{\lambda})$ in RD rows. We can observe the performance and security parameters by the different bounds of $\max_{x\in S}\delta(x)$. From the column named security, it shows that when $order = \infty$, the security is almost the same when $\max_{x\in S}\delta(x)$ relaxes to $\frac{1}{\lambda}$.

We come to a conclusion that $RD_\infty$ gains both the optimized performance and same security magnitude, which relax the relative error of two probability distribution from $2^{-\lambda}$ to $\frac{1}{\lambda}$. In summary, probability preservation property with relative error is easy to analyse and compare.

# 4 OPTIMIZATION

In this section, we optimize the smoothing parameter of Gaussian sampling algorithm by $RD_\infty$-based probability preservation property. We improve the user key size of IBEs over general lattices and NTRU lattices by optimizing the smoothing parameter.

## 4.1 Optimization of Gaussian Sampling Algorithm

A collection of one-way preimage sampleable functions is defined as follows (Gentry et al., 2008):

- $TrapGen(1^\lambda)$: $(a,t) \leftarrow TrapGen(1^\lambda)$, where $a$ is a function description: $f_a : D_\lambda \rightarrow R_\lambda$, $D_\lambda$ is domain of $f_a$, $R_\lambda$ is range of $f_a$, $t$ is trapdoor information for $f_a$.
- $SampleDom(1^\lambda)$: $x \leftarrow SampleDom(1^\lambda)$, where x is from some distribution over $D_\lambda$, for which the distribution of $f_a(x)$ is uniform over $R_\lambda$.
- $SamplePre(1^\lambda)$: $x \leftarrow SamplePre(1^\lambda)$, for $\forall y \in R_\lambda$, $SamplePre(t,y)$ samples from the distribution of $x \leftarrow SampleDom(1^\lambda)$, given $f_a(x) = y$.
- one-wayness without trapdoor: For any probabilistic poly-time algorithm $\mathcal{A}$, $\mathbf{Pr}[\mathcal{A}(1^\lambda, a, y) \in f_a^{-1}(y)] \leq negl(\lambda)$.

These properties of probability distribution defined above only hold in the ideal trapdoor functions construction. The properties in the real trapdoor functions construction will be relaxed.

PSFs over lattices based on the average-case hardness of **ISIS** are $f_{\mathbf{A}} = \mathbf{Ae} = \mathbf{u} \mod q$, where the domain is $\{\mathbf{e} \in \mathbb{Z}^m : \|\mathbf{e}\| < s\sqrt{m}\}$ and the range is $\mathbb{Z}^n$. They have two essential properties as follows: $\mathbf{u} \in \mathbb{Z}^{n\times m}$ is statistically close to uniform and the distribution of $\mathbf{e}$ is statistically close to $D_{\mathbb{Z}^m,s}$. According to Lemma 2, Given a good basis $\mathbf{T} \subset \Lambda^\perp(A)$, preimage $\mathbf{e}$ can be sampled by the trapdoor inversion algorithm and $\mathbf{e} \sim D_{\Lambda^\perp,s}$ statistically.

We require $\mathbf{e}$ is distributed according to the exact Gaussian distribution $D_{\mathbb{Z}^m,s}$ in the ideal scheme and define its distribution as $Q$. However, we get $\mathbf{e}$ from the trapdoor inversion algorithm in the real scheme and define its distribution as $P$. First, choose via linear algebra $\mathbf{t} \in \mathbb{Z}^m$ such that $\mathbf{At} = \mathbf{u} \mod q$. Then sample $\mathbf{v} \sim D_{\Lambda^\perp,s,-t}$ statistically using Gaussian sampling algorithm and output $\mathbf{e} = \mathbf{t} + \mathbf{v}$. Referring to proof of algorithm, we have the output probability distribution of Gaussian sampling algorithm $P \in [(\frac{1-\varepsilon}{1+\varepsilon})^m, (\frac{1+\varepsilon}{1-\varepsilon})^m] \cdot D_{\mathbb{Z}^m,s}$, where $\mathbf{e}$ is the output of the trapdoor inversion algorithm, as detailed in

Table 1: Comparisons of distribution and security parameters.

| number | measures | max relative error $\max\delta(x)$ | security $\varepsilon_P$ |
|---|---|---|---|
| 1 | SD | $2^{-\lambda}$ | $2\cdot 2^{-\lambda}$ |
| 2 | KL | $2^{-\lambda}$ | $(1+\frac{\sqrt{2}}{2})\cdot 2^{-\lambda}$ |
| 3 | $RD_\infty$ | $2^{-\lambda}$ | $(1+2^{-\lambda})\cdot 2^{-\lambda}$ |
| 4 | $RD_\infty$ | $\frac{1}{\lambda}$ | $(1+\frac{1}{\lambda})\cdot 2^{-\lambda}$ |
| 5 | $RD_2$ | $2^{-\lambda}$ | $\sqrt{2^{-\lambda}(1+2^{-2\lambda})}\approx 2^{-\frac{\lambda}{2}}$ |
| 6 | $RD_2$ | $\frac{1}{\lambda}$ | $\sqrt{2^{-\lambda}(1+\frac{1}{\lambda^2})}\approx 2^{-\frac{\lambda}{2}}$ |
| 7 | $RD_\alpha$ | $2^{-\lambda}$ | $(2^{-\lambda}(1+\alpha 2^{-2\lambda}))^{\frac{\alpha-1}{\alpha}}\approx 2^{-\frac{(\alpha-1)\lambda}{\alpha}}$ |
| 8 | $RD_\alpha$ | $\frac{1}{\lambda}$ | $(2^{-\lambda}(1+\frac{1}{\alpha\lambda^2}))^{\frac{\alpha-1}{\alpha}}\approx 2^{-\frac{(\alpha-1)\lambda}{\alpha}}$ |

(Gentry et al., 2008). The relative error to the desired distribution is therefore bounded by $(\frac{1+\varepsilon}{1-\varepsilon})^m-1$, $\delta(x=\mathbf{e})=\frac{P(x)-Q(x)}{Q(x)}=((\frac{1+\varepsilon}{1-\varepsilon})^m-1)\approx 1+\frac{2m\varepsilon}{1-\varepsilon}-1=\frac{2m\varepsilon}{1-\varepsilon}\leq\max_{x\in\Lambda_{\mathbf{u}}^{\perp}}\delta(x)$.

According to Lemma 1, We have $\eta_\varepsilon\leq\sqrt{\frac{\ln(2m(1+1/\varepsilon))}{\pi}}\cdot\|\widetilde{\mathbf{B}}\|$. It is easy to get $\eta_\varepsilon$ from $\varepsilon$. When $\max_{x\in\Lambda_{\mathbf{u}}^{\perp}}\delta(x)$ relaxes to $\frac{1}{\lambda}$ (even approximating to $1/2$), it is easy to find out that the tolerance $\varepsilon$ is shortened to $\frac{1}{2m\lambda}$ (when $\max_{x\in\Lambda_{\mathbf{u}}^{\perp}}\delta(x)$ reaches the upper bound 1/2, $\varepsilon$ reaches $\frac{1}{4m}$) and the security parameter of $RD_\infty$ remains $(1+\frac{1}{\lambda})\cdot 2^{-\lambda}$ on the 4th row of Table 1. So the smoothing parameter decreases from $\sqrt{\frac{\ln(2m(1+2m\cdot 2^{\lambda}))}{\pi}}\cdot\|\widetilde{\mathbf{B}}\|$ to $\sqrt{\frac{\ln(2m(1+2m\lambda))}{\pi}}\cdot\|\widetilde{\mathbf{B}}\|$ (when $\max_{x\in\Lambda_{\mathbf{u}}^{\perp}}\delta(x)$ approximates the upper bound 1/2, the smoothing parameter approximates $\sqrt{\frac{\ln(2m(1+4m))}{\pi}}\cdot\|\widetilde{\mathbf{B}}\|$). We come to a conclusion the smoothing parameter declines by a factor $\mathbf{O}(\sqrt{\lambda})$ via $RD_\infty$ and the security remains the same.

## 4.2 Optimization of IBE

We give two illustrative examples of instantiations over general lattices and NTRU lattices. We start with the IBE system in(Gentry et al., 2008), then we extend to the IBE system in(Ducas et al., 2014).

IBE consists of four algorithm: IBESetup, IBEExtract, IBEEnc, IBEDec.

- $IBESetup(1^n)$: $(A,T)\leftarrow IBESetup(1^n)$, where $T$ is trapdoor of function $f_A$. The Master public key is $A$, the secret key is $T$.
- $IBEExtract(A,T,id)$: $e\leftarrow IBEExtract(A,T,id)$ such that $Ae=u=H(id)$, where $e$ is user secret key for the user $id$ and $H$ is an oracle.
- $IBEEnc(A,id,b)$: $(p,c)\leftarrow DualEnc(u,b)$.
- $IBEDec(e,(p,c))$: $b\leftarrow DualDec(e,(p,c))$.

In (Gentry et al., 2008) preimage sampleable (trapdoor) function is $f_{\mathbf{A}}(\mathbf{e})=\mathbf{Ae} \mod q$, where $\mathbf{A}\in\mathbb{Z}_q^{m\times n}$ is a random matrix and $\mathbf{e}\in\mathbb{Z}_q^m$ is short vector. In the algorithm $IBEExtract(\mathbf{A},\mathbf{T},\mathbf{id})$, the decryption key $\mathbf{e}\leftarrow f^{-1}(u)$ is obtained by using the preimage sampler with trapdoor $\mathbf{T}$. $\mathbf{e}$ is the output of Gaussian sampling algorithm with short basis $\mathbf{T}$. According to the tail inequality Lemma 3, we have $\mathbf{e}:\|\mathbf{e}\|\leq s\cdot\sqrt{m}$, so the size of user secret key $\mathbf{e}$ can be reduced by a factor $\mathbf{O}(\sqrt{\lambda})$ as the smoothing parameter is optimized.

Because the discrete Gaussian sampling described in (Gentry et al., 2008) involves sequential iterations and the Gram-Schmidt orthogonalized operations, it is rather inefficient. So making adjustments to some of tunable parameters of the main operations about trapdoor generation and inversion algorithms may provide better combinations of efficiency and concrete security.

In (Ducas et al., 2014) Gaussian sampling algorithm of NTRU lattices is applied to sampling $(\mathbf{s_1},\mathbf{s_2})$ such that $\mathbf{s_1}+\mathbf{s_2}*\mathbf{h}=\mathbf{t}$ by a short basis $\mathbf{B}$ of $\Lambda_{\mathbf{h},q}$. The output of Gaussian sampling algorithm has the length $s\cdot\sqrt{2N}$, where $s>\eta_\varepsilon$. The improvements involve the tolerance $\varepsilon$ and the Gram-Schmidt norm of the trapdoor. According to Lemma 1, the optimization of the tolerance is orthogonal to the optimization the Gram-Schmidt norm of the basis. $B$ is a short basis of $\Lambda_{\mathbf{h},q}$. Referring to Lemma 4 the Gram-Schmidt norm computation of the trapdoor basis is simplified and public key size is reduced owing to the structure of ring. With RD $\varepsilon$ can be reduced to $\frac{1}{4N\lambda}$. In summary, two optimization methods can be combined well. The user secret key size is lowered by a factor $\mathbf{O}(\sqrt{\lambda})$ in the terms of our optimization.

# 5 CONCLUSION

In this work, we reform preservation property of one-way preimage sampleable functions with relative error. We give straightforward analysis and comprehensive comparisons of probability and show that $RD_\infty$ has improved performance without security loss. Furthermore, we optimize the smoothing parameter of Gaussian sampling algorithm by $RD_\infty$-based probability preservation property. Finally, we improve the user key size of IBEs over general lattices and NTRU lattices by optimized the smoothing parameter. We can improve them with shorter parameters and better efficiency without compromising the security.

# ACKNOWLEDGEMENT

The authors would like to thank anonymous reviewers for their helpful comments and suggestions. The work of this paper was supported by the National Natural Science Foundation of China (Grants Y610112103).

# REFERENCES

Bai, S., Langlois, A., Lepoint, T., Stehlé, D., and Steinfeld, R. (2015). Improved security proofs in lattice-based cryptography: Using the rényi divergence rather than the statistical distance. In *Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part I*, pages 3–24.

Ducas, L., Lyubashevsky, V., and Prest, T. (2014). Efficient identity-based encryption over NTRU lattices. In *Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II*, pages 22–41.

Gentry, C., Peikert, C., and Vaikuntanathan, V. (2008). Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 197–206.

Langlois, A., Stehlé, D., and Steinfeld, R. (2014). Gghlite: More efficient multilinear maps from ideal lattices. In *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, pages 239–256.

Micciancio, D. and Regev, O. (2004). Worst-case to average-case reductions based on gaussian measures. In *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*, pages 372–381.

Micciancio, D. and Walter, M. (2017). Gaussian sampling over the integers: Efficient, generic, constant-time. In *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part II*, pages 455–485.

Pöppelmann, T., Ducas, L., and Güneysu, T. (2014). Enhanced lattice-based signatures on reconfigurable hardware. In *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, pages 353–370.

Prest, T. (2017). Sharper bounds in lattice-based cryptography using the rényi divergence. *IACR Cryptology ePrint Archive*, 2017:480.

Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 547–561, Berkeley, Calif. University of California Press.

Takashima, K. and Takayasu, A. (2015). Tighter security for efficient lattice cryptography via the rényi divergence of optimized orders. In *Provable Security - 9th International Conference, ProvSec 2015, Kanazawa, Japan, November 24-26, 2015, Proceedings*, pages 412–431.

van Erven, T. and Harremoës, P. (2014). Rényi divergence and kullback-leibler divergence. *IEEE Trans. Information Theory*, 60(7):3797–3820.

# APPENDIX *A*

We now provide **Proof of Theorem** 1.

*Proof.* 1) Firstly, we prove the RD bound with the order $\alpha$.

We set $f(P(x)) = \frac{P(x)^\alpha}{Q(x)^{\alpha-1}}$, then give the partial derivative $\partial^n f/\partial P(x)^n$ at $P(x) = Q(x)$,

$$f(P(x))|_{P(x)=Q(x)} = \frac{P(x)^\alpha}{Q(x)^{\alpha-1}}|_{P(x)=Q(x)} = Q(x),$$

$$f^{(1)}(P(x))|_{P(x)=Q(x)} = \alpha\frac{P(x)^{\alpha-1}}{Q(x)^{\alpha-1}}|_{P(x)=Q(x)} = \alpha,$$

$$f^{(2)}(P(x))|_{P(x)=Q(x)} = \alpha(\alpha-1)\frac{P(x)^{\alpha-2}}{Q(x)^{\alpha-1}}|_{P(x)=Q(x)}$$

$$= \alpha(\alpha-1)\frac{1}{Q(x)}$$

according to Taylor series expansion, it holds that

$$\begin{aligned} f(P(x)) &= f(P(x))|_{P(x)=Q(x)} + \\ &\quad f^{(1)}(P(x))|_{P(x)=Q(x)}(P(x)-Q(x)) \\ &\quad + \frac{f^{(2)}(P(x))|_{P(x)=Q(x)}}{2!} \\ &\quad (P(x)-Q(x))^2 + R(P(x)) \\ &= Q(x) + \alpha(P(x)-Q(x)) + \\ &\quad \frac{\alpha(\alpha-1)(Q(x)-P(x))^2}{2!Q(x)} + R(P(x)) \\ &= Q(x) + \alpha(P(x)-Q(x)) + \\ &\quad \frac{\alpha(\alpha-1)}{2}Q(x)(\frac{P(x)-Q(x)}{Q(x)})^2 \\ &\quad + R(P(x)), \end{aligned}$$

where $R(P(x))$ is the tail-cut of the Taylor expansion.

$$\begin{aligned} R(P(x) &= \sum_{i=3}^{\infty} C(\alpha,i)Q(x)(\frac{P(x)-Q(x)}{Q(x)})^i \\ &= \sum_{i=3}^{\infty} \frac{C(\alpha,i)}{\alpha(\alpha-1)}(\frac{P(x)-Q(x)}{Q(x)})^{i-2}Q(x) \\ &\quad \alpha(\alpha-1)(\frac{P(x)-Q(x)}{Q(x)})^2. \end{aligned}$$

As $\frac{C(\alpha,i)}{\alpha(\alpha-1)} < \frac{\alpha^{i-2}}{2}$ and $(\frac{P(x)-Q(x)}{Q(x)})^{i-2} \leq \frac{1}{\alpha^{i-2}}$, $\max_{x\in X}(R(P(x))) \leq \frac{1}{2}\sum_{x\in S} Q(x)(\frac{P(x)-Q(x)}{Q(x)})^2$. As $P(x)$ and $Q(x)$ sum over $x \in S$, we have

$$\begin{aligned} \sum_{x\in S} f(P(x)) &\leq \sum_{x\in S} Q(x) + \alpha\sum_{x\in S}(P(x)-Q(x)) + \\ &\quad \alpha(\alpha-1)\sum_{x\in S} Q(x)(\frac{(P(x)-Q(x)}{Q(x)})^2. \end{aligned}$$

Since $S$ is the support of both $P$ and $Q$, it holds that $\sum_{x\in S} P(x) = \sum_{x\in S} Q(x) = 1$. So $\sum_{x\in S} f(P(x)) \leq (1+\alpha(\alpha-1)\sum_{x\in S}\delta(x)^2 Q(x)$.

$$\begin{aligned} R_\alpha(P\|Q) &= (\sum_{x\in S}\frac{P(x)^\alpha}{Q(x)^{\alpha-1}})^{\frac{1}{\alpha-1}} \\ &\leq (1+\alpha(\alpha-1)\sum_{x\in S}\delta(x)^2 Q(x))^{\frac{1}{\alpha-1}} \\ &= 1+\frac{1}{\alpha-1}(\alpha(\alpha-1)\sum_{x\in S}\delta(x)^2 Q(x)) \\ &\quad + R', \end{aligned}$$

where the tail-cut of the Taylor expansion $R' < 0$, so we have

$$R_\alpha(P\|Q) \leq 1+\alpha(\max_{x\in S}\delta(x))^2.$$

By the same deduction, we get $R_2(P\|Q) = \sum_{x\in S}\frac{P(x)^2}{Q(x)} = 1+\sum_{x\in S}\delta(x)^2 Q(x) \leq 1+(\max_{x\in S}\delta(x))^2$ when the order is 2.

2) Then, we prove the RD bound with the order 1 and $\infty$. $R_1(P\|Q)$ is the exponential of the Kullback-Leibler divergence. $R_1(P\|Q) = \exp D_{KL}(P\|Q) \approx 1+D_{KL}(P\|Q) = 1+\sum_{x\in S} Q(x)(\delta(x))^2 = 1+(\max_{x\in S}\delta(x))^2$. $R_\infty(P\|Q) = \max_{x\in S}\frac{P(x)}{Q(x)} \leq \max_{x\in S}\frac{(1+\delta(x))Q(x)}{Q(x)} \leq 1+\max_{x\in S}\delta(x)$.

□

Finally we give the SD bound of distributions closeness with relative error.

*Proof.* $\Delta(P\|Q) = \frac{1}{2}\sum_{x\in S}|P(x)-Q(x)| \leq \frac{1}{2}\sum_{x\in S}\delta(x)\cdot Q(x) \leq \frac{1}{2}\sum_{x\in S}\max\delta(x)\cdot Q(x) = \frac{1}{2}\max_{x\in S}\delta(x)$.

Since $S$ is the support of $Q$, it holds that $\sum_{x\in S} Q(x) = 1$, therefore the last equation holds.

□