Simulation of Random Variables under Rényi Divergence Measures of All Orders

Lei Yu*, Vincent Y. F. Tan†

*†Department of Electrical and Computer Engineering, †Department of Mathematics,
National University of Singapore, Singapore
Email: *leiyu@nus.edu.sg, †vtan@nus.edu.sg

Abstract—The random variable simulation problem consists in using a $k$-dimensional i.i.d. random vector $X^k$ with distribution $P_X^k$ to simulate an $n$-dimensional i.i.d. random vector $Y^n$ so that its distribution is approximately $Q^n_Y$. In contrast to previous works, in this paper we use the Rényi divergences of all orders to measure the level of approximation. We also characterize the asymptotics of normalized Rényi divergences as well as the Rényi conversion rates. The latter are defined as the supremum of $\frac{a}{g}$ that the Rényi divergences vanish asymptotically. In addition, when the Rényi parameter is in $(0,1)$, the Rényi conversion rates equal the ratio of the Shannon entropies $H(P_X)/H(Q_Y)$, which is consistent with traditional results, in which the Rényi parameter is equal to 1. When the Rényi parameter is in $[1,\infty)$, the Rényi conversion rates are, in general, smaller than the TV conversion rates, in which the Rényi divergences vanish asymptotically.

In addition to the standard Rényi divergence, we consider two new divergences — the max-Rényi divergence $D^\text{max}_\alpha(P,Q)$ and the sum-Rényi divergence $D^+_{\alpha}(P,Q)$, and use them (with order infinity) as approximation measures. These two measures with order infinity are very strong since for any $\epsilon > 0$, $D^\text{max}_\alpha(P,Q) \leq \epsilon$ or $D^+_{\alpha}(P,Q) \leq \epsilon$ implies $e^{-\epsilon} \leq \frac{H(P_X)}{H(Q_Y)} \leq e^{\epsilon}$, $\forall x$.

I. INTRODUCTION

How can we use a $k$-dimensional i.i.d. random vector $X^k$ with distribution $P_X^k$ to simulate an $n$-dimensional i.i.d. random vector $Y^n$ so that its distribution is approximately $Q^n_Y$? This is so-called random variable simulation problem or distribution approximation problem [1]. In [1] and [2], the total variation (TV) distance and the Bhattacharyya coefficient (the Rényi divergence of order $\frac{1}{2}$) were respectively used to measure the level of approximation. In these papers, the asymptotic conversion rate $R$ was studied, which is defined as the supremum of $\frac{a}{g}$ such that the employed measure vanishes asymptotically as the dimensions $n$ and $k$ tend to infinity.

For both the TV distance and the Bhattacharyya coefficient, the asymptotic (first-order) conversion rates are the same, and both equal to the ratio of the Shannon entropies $H(P_X)/H(Q_Y)$. Furthermore, in [2], Kumagai and Hayashi also investigated the asymptotic second order conversion rate. Note that by Pinsker’s inequality [3], the Bhattacharyya coefficient (the Rényi divergence of order $\frac{1}{2}$) is stronger than the TV distance, i.e., if the Bhattacharyya coefficient tends to 1 (or the Rényi divergence of order $\frac{1}{2}$ tends to 0), then the TV distance tends to 0. In this paper, we strengthen the TV distance and the Bhattacharyya coefficient by considering Rényi divergences of orders in $[0,\infty]$.

As two important special cases of the distribution approximation problem, the resolvability and intrinsic randomness problems have been extensively studied in literature, e.g., [1], [4]–[8].

A. Main Contributions

For the distribution approximation problem, we use the Rényi divergences $D_\alpha(P_Y^n\|Q_Y^n)$ and $D_\alpha(Q_Y^n\|P_Y^n)$ of orders in $[0,\infty]$ as the distance measure between the simulated and target output distributions $P_Y^n$ and $Q_Y^n$. We characterize the asymptotics of Rényi divergences, as well as the Rényi conversion rates, which are defined as the supremum of $\frac{a}{g}$ to guarantee that the Rényi divergences vanish asymptotically. Interestingly, when the Rényi parameter is in $(0,1]$ for the measure $D_\alpha(P_Y^n\|Q_Y^n)$ and in $(0,1)$ for the measure $D_\alpha(Q_Y^n\|P_Y^n)$, the Rényi conversion rates are just equal to the ratio of the Shannon entropies $\frac{H(P_X)}{H(Q_Y)}$. This is consistent with the existing results in [2]. In contrast if the Rényi parameter is in $[1,\infty]$ for the measure $D_\alpha(P_Y^n\|Q_Y^n)$ and in $[1,\infty]$ for the measure $D_\alpha(Q_Y^n\|P_Y^n)$, the Rényi conversion rates are, in general, larger than $\frac{H(P_X)}{H(Q_Y)}$. Furthermore, we also consider two new divergences — the max-Rényi divergence and the sum-Rényi divergence, and use them as the measure of level of approximation. The obtained expressions for the Rényi conversion rates involve Rényi entropies of various orders, of approximation. The obtained expressions for the Rényi conversion rates involve Rényi entropies of various orders, even negative orders. To the best of our knowledge, this is the first time to give an operational interpretation of the Rényi entropies of negative orders. Furthermore, our results can be applied to information-theoretic security, e.g., secret key generation. Owing to space limitation, we refer readers to the extended version [9] for more details.

B. Notation and Information Distance Measures

The set of probability measures on $\mathcal{X}$ is denoted as $\mathcal{P}(\mathcal{X})$. We use $T^n_{\alpha}(x) := \frac{1}{n} \sum_{i=1}^{n} 1\{x_i = x\}$ to denote the type (empirical distribution) of a sequence $x^n$, $T_X$ to denote a type of sequences in $\mathcal{X}^n$. For a type $T_X$, the type class (set of sequences having the same type $T_X$) is denoted by $\mathcal{T}_X$. The set of types of sequences in $\mathcal{X}^n$ is denoted as $\mathcal{P}^{(n)}(\mathcal{X}) := \{T^n_x : x^n \in \mathcal{X}^n\}$. For a function $f : \mathcal{X} \to \mathcal{Y}$, and any subset $A \subset \mathcal{X}$, define $f(A) := \{f(x) : x \in A\}$, and $f^{-1}(y) := \{x \in \mathcal{X} : f(x) = y\}$. Finally, we write $f(n) \leq g(n)$ if $\lim\sup_{n \to \infty} \frac{1}{n} \log \frac{f(n)}{g(n)} \leq 0$. In addition, $f(n) \geq g(n)$ means $f(n) \leq g(n)$ and $g(n) \leq f(n)$. We use $\delta_n, \delta'_n, \delta_n^\ast$ to denote generic sequences tending to zero as $n \to \infty$. For $a \in \mathbb{R}$, $[a]^+ := \max\{a,0\}$ denotes positive clipping. We define $\frac{a}{n} := +\infty$, if $a > 0$; $-\infty$, if $a < 0$; 0, if $a = 0$. 


For a distribution \( P_X \in \mathcal{P}(\mathcal{X}) \), the Shannon entropy and the Rényi entropy of order \( \alpha \in [0, +\infty[ \) are respectively defined as

\[
H(P_X) := -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x), \quad \text{and} \\
H_\alpha(P_X) := \frac{1}{1 - \alpha} \log \sum_{x \in \mathcal{X}} P_X(x)\,^\alpha,
\]

where throughout, \( \log \) is to the natural base \( e \). It is known that \( \lim_{\alpha \to 1} H_\alpha(P_X) = H(P_X) \) so a special case of the Rényi entropy is the usual Shannon entropy. Fix distributions \( P_X, Q_X \in \mathcal{P}(\mathcal{X}) \). Then the relative entropy and the Rényi divergence of order \( \alpha \geq 0 \) are respectively defined as

\[
D(P_X\|Q_X) := \sum_{x \in \mathcal{X}} P_X(x) \log \frac{P_X(x)}{Q_X(x)}, \quad \text{and} \\
D_\alpha(P_X\|Q_X) := \frac{1}{1 - \alpha} \log \sum_{x \in \mathcal{X}} P_X(x)\,^\alpha Q_X(x)^{1-\alpha}.
\]

It is known that \( \lim_{\alpha \to 1} D_\alpha(P_X\|Q_X) = D(P_X\|Q_X) \) so a special case of the Rényi divergence is the usual relative entropy.

We define the max-Rényi divergence as \( D_{\max}(P, Q) := \max \{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \), and the sum-Rényi divergence as \( D_{\sum}(P, Q) := D_\alpha(P\|Q) + D_\alpha(Q\|P) \). The sum-Rényi divergence reduces to Jeffrey's divergence \( D(P) + D(Q) \) in the sense that for any sequences of distribution pairs \( \{(P(n), Q(n))\}_{n=1}^{\infty} \), \( D_{\max}(P(n), Q(n)) \to 0 \) implies that \( D_{\sum}(P(n), Q(n)) \to 0 \) and vice versa. In this paper, we only consider the max-Rényi divergence. For \( \alpha = \infty \), \( D_{\max}(P, Q) = \sup_{A \subset \mathcal{X}} |\log P(A) - \log Q(A)| \). This expression is similar to the definition of TV distance, hence we call \( D_{\max} \) as the logarithmic variation distance.\(^1\)

**Lemma 1.** The following properties hold.

1. \( D_{\max} \) is a metric. Similarly, \( D_{\sum} \) is also a metric.
2. \( D_{\max}(P, Q) \leq \epsilon \iff e^{-\epsilon} \leq \frac{P(x)}{Q(x)} \leq e^\epsilon \) for \( x \in \mathcal{X} \).
3. For any \( f \), \( D_{\max}(f(P), f(Q)) \leq \epsilon \iff e^{-\epsilon} \leq \frac{f(P(x))}{f(Q(x))} \leq e^\epsilon \).
4. \( D_{\max}(P_X P_{Y|X}, Q_X P_{Y|X}) = D_{\max}(P_X, Q_X) \).

**C. Problem Formulation**

We consider the distribution approximation problem, which can be described as follows. We are given a target “output” distribution \( Q_Y \) that we would like to simulate. At the same time, we are given \( k \)-length sequence of a memoryless source \( X^k \sim P_X \). We would like to design a function \( f : \mathcal{X}^k \to \mathcal{Y}^n \) such that the distance, according to some divergence measure, of the simulated distribution \( P_{Y^n} \) with \( Y^n := f(X^k) \) and \( n \) independent copies of the target distribution \( Q_Y^n \) is minimized.

\(^1\)In literature, the Rényi entropy is defined for orders \( \alpha \in [0, +\infty[ \), but here we define it for orders \( \alpha \in [-\infty, +\infty[ \). This is due to the fact that our results involve Rényi entropies of all real orders, even negative orders.\(^2\)

\(^2\)In [11], \( D_{\max}(P, Q) \leq \epsilon \) is called \((\epsilon, 0)\)-closeness.

Here we let \( n = \lceil kR \rceil \), where \( R \) is a fixed positive number known as the rate. We assume the alphabets \( \mathcal{X} \) and \( \mathcal{Y} \) are finite. There are now two fundamental questions associated to this simulation task: (i) As \( k \to \infty \), what is the asymptotic level of approximation as a function of \( (R, P_X, Q_Y) \)? (ii) As \( k \to \infty \), what is the maximum rate \( R \) such that the discrepancy between the distribution \( P_{Y^n} \) and \( Q_Y^n \) tends to zero? In contrast to previous works on this problem [11, 2], here we employ the Rényi divergences \( D_\alpha(P_{Y^n}\|Q_Y^n) \), \( D_\alpha(Q_Y^n\|P_{Y^n}) \), or \( D_{\max}(P_{Y^n}, Q_Y^n) \) of all orders \( \alpha \in [0, \infty[ \) to measure the discrepancy between \( P_{Y^n} \) and \( Q_Y^n \).

**II. RÉNYI DISTRIBUTION APPROXIMATION**

**A. Asymptotics of Rényi Divergences**

We first characterize the asymptotics of Rényi divergences \( D_\alpha(P_{Y^n}\|Q_Y^n) \) and \( D_\alpha(Q_Y^n\|P_{Y^n}) \).

**Theorem 1** (Asymptotics of \( \frac{1}{n} D_\alpha(P_{Y^n}\|Q_Y^n) \)). For any \( \alpha \in [0, \infty[ \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \inf_{f \in \mathcal{F}} D_\alpha(P_{Y^n}\|Q_Y^n) = \sup_{\epsilon \in [0, 1]} \left\{ t H_{\frac{\alpha}{\alpha + 1}}(Q_Y) - \frac{t}{R} H_{\frac{\alpha}{\alpha + 1}}(P_X) \right\},
\]

**Theorem 2** (Asymptotics of \( \frac{1}{n} D_\alpha(Q_Y^n\|P_{Y^n}) \)). For any \( \alpha \in [0, \infty[ \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \inf_{f \in \mathcal{F}} D_\alpha(Q_Y^n\|P_{Y^n}) = \left\{ \begin{array}{ll}
\sup_{\epsilon \in (0, \infty)} \left\{ \frac{t}{R} H_{\frac{\alpha}{\alpha + 1}}(Q_Y) \right\}, & \alpha \in [1, \infty] \text{ and } R \neq \frac{H_0(P_X)}{H_0(Q_Y)}; \\
\frac{\alpha}{\alpha - 1} \sup_{\epsilon \in [0, 1]} \left\{ t H_{\frac{\alpha}{\alpha + 1}}(Q_Y) \right\}, & \alpha \in (0, 1); \\
0, & \alpha = 0.
\end{array} \right.
\]

**Remark 1.** For \( \alpha \in [1, \infty[ \) and \( R = \frac{H_0(P_X)}{H_0(Q_Y)} \), the asymptotic behavior of \( \frac{1}{n} \inf_{f \in \mathcal{F}} D_\alpha(Q_Y^n\|P_{Y^n}) \) depends on how fast \( n \) converges to \( R \). In this paper, we set \( n = \lceil kR \rceil \), i.e., the fastest case. For this case, \( \frac{1}{n} \inf_{f \in \mathcal{F}} D_\alpha(Q_Y^n\|P_{Y^n}) = \infty \), if \( kR \notin \mathbb{N} \); and \( \frac{1}{n} \inf_{f \in \mathcal{F}} D_\alpha(Q_Y^n\|P_{Y^n}) = \frac{1}{n} D_\alpha(Q_Y^n\|P_{Y^n}) \), if \( kR \in \mathbb{N} \), where \( \{P_X\} \) and \( \{Q_Y\} \) respectively denote the resulting sequences after sorting \( P_X \) and \( Q_Y \) in descending order.

The proof of Theorem 1 is provided in Section III. Theorem 2 is proven by a similar method, hence the proof is omitted here and provided in [9].

**B. Rényi Conversion Rates**

As shown in the theorems above, when the code rate is large, the normalized Rényi divergences \( \frac{1}{n} D_\alpha(P_{Y^n}\|Q_Y^n) \) and \( \frac{1}{n} D_\alpha(Q_Y^n\|P_{Y^n}) \) converge to a positive number; however when the code rate is small enough, the normalized Rényi divergences converge to zero. The threshold rate, named Rényi conversion rate, is very important, since it represents the
maximum possible rate under the condition that the distribution induced by the code well approximates the target distribution. We characterize the Rényi conversion rates for \( \frac{1}{n} D_n(P^n_Y \| Q^n_Y) \) and \( \frac{1}{n} D_n(Q^n_Y \| P^n_Y) \) in the following two corollaries.

**Corollary 1** (Rényi Conversion Rate). For any \( \alpha \in [0, \infty] \),
\[
\sup \{ R : \frac{1}{n} D_n(P^n_Y \| Q^n_Y) \to 0 \} = \begin{cases} 
\inf_{\varepsilon \in (0,1)} \frac{H(P_X)}{\sup_{Q \sim (\alpha-1)^{X^n}(\alpha^{n-1})T^n}(\alpha^{n-1})T^n}, & \alpha \in (0,1), \\
\alpha = 0, & \alpha = 1, \infty 
\end{cases}
\]  
(5)

**Remark 2.** The result for \( \alpha = 1 \) (i.e., the relative entropy case) was first shown by Han [1]. Corollary 1 is the generalization of [1] to the (normalized) Rényi divergence of all orders \( \alpha \in [0, \infty) \). Besides, the result, including first order and second order rates, for the unnormalized Rényi divergence \( D_n(P^n_Y \| Q^n_Y) \) with \( \alpha = \frac{1}{2} \) was given by Kumagai and Hayashi [2]. For the result for the unnormalized Rényi divergence with \( \alpha \in (0, \frac{1}{2}) \) can be obtained by combining two observations: 1) the achievable for \( D_\frac{1}{2}(P^n_Y \| Q^n_Y) \) implies the achievability for \( \alpha \in (0, \frac{1}{2}) \); 2) by Pinsker’s inequality [3], the result under the TV distance measure [1] implies the converse for \( \alpha \in (0, \frac{1}{2}) \). The achievability for orders \( \alpha \in (0,1) \cup (1, \infty) \) and the converse for orders \( \alpha \in [0,1) \cup (1, \infty] \) are new.

**Remark 3.** \( D_n(P_{Y|X=x^n} \| P_{Y|X=x'^n}) \leq \epsilon \) for all neighboring database instances \( x, x' \) is called \( \epsilon \)-Rényi differential privacy of order \( \alpha \) [12], and the special case with \( \alpha = \infty \) is well known as \( \epsilon \)-differential privacy [13], where \( X \) represents public data and \( Y \) represents private data. In the two corollaries above, this measure is applied to the random variable simulation problem, and we show the “necessary and sufficient condition” for \( \lim_{n \to \infty} \frac{1}{n} D_n \leq \epsilon \) for any \( \epsilon > 0 \).

**Corollary 2** (Rényi Conversion Rate). For any \( \alpha \in [0, \infty] \),
\[
\sup \{ R : \frac{1}{n} D_n(Q^n_Y \| P^n_Y) \to 0 \} = \begin{cases} 
\min \{ \inf_{\varepsilon \in (0,\infty)} \frac{H(P_X)}{\sup_{Q \sim (\alpha-1)^{X^n}(\alpha^{n-1})T^n}(\alpha^{n-1})T^n}, & \alpha \in [1, \infty], \\
\frac{H(P_X)}{H(Q_Y)}, & \alpha \in (0,1), \\
\infty, & \alpha = 0 
\end{cases}
\]  
(6)

**Remark 4.** Our results for all orders \( \alpha \in [0, \infty] \) are new.

One can find simple interpretations for the special cases \( \alpha = 0 \) in Corollaries 1 and 2, by noting that \( D_0(P \| Q) = -\log Q(\text{support}(P)) \). Furthermore, we also consider the Rényi divergence \( D_\alpha^{\text{max}}(P_{Y^n}, Q_{Y^n}) \). The proof of the following theorem is provided in Section IV.

**Theorem 3** (Rényi Conversion Rate). For \( \alpha = \infty \), we have
\[
\sup \{ R : \frac{1}{n} D_\infty^{\text{max}}(P_{Y^n}, Q_{Y^n}) \to 0 \} = \sup \{ R : D_\infty^{\text{max}}(P_{Y^n}, Q_{Y^n}) \to 0 \} = \min_{\beta \in [-\infty, \infty]} \frac{H_\beta(P_X)}{H_\beta(Q_Y)}.
\]  
(7)

\[ \frac{1}{n} D_{1+s}(P_{Y^n} \| Q_{Y^n}) \leq \frac{1}{s} \max_{T_X} \left\{ -(1+s) D(T_X \| P_X) + s D(T_Y \| Q_Y) + s H(T_Y) - H(T_X) \right\} \bigg|_{T_Y = g(T_X)} + \delta_n. \]  
(8)
For each $T_X$, choose $g(T_X)$ as the $T_Y$ that minimizes the expression in (8). Then we can get
\[
\limsup_{n \to \infty} \frac{1}{n} D_{1+s}(P_{Y^n} || Q_{Y^n}) = \limsup_{n \to \infty} \max_{T_X} \min_{T_Y} \left\{ -\frac{1+s}{s} D(T_X || P_X) + D(T_Y || Q_Y) + |H(T_Y) - H(T_X)|^+ \right\}
\]
\begin{equation}
\geq \min_{\tilde{P}_Y \in \mathcal{P}(\tilde{Y})} \max_{P_{X^n} \in \mathcal{P}(X^n)} \left\{ -\frac{1+s}{s} D(\tilde{P}_X || P_X) + D(\tilde{P}_Y || Q_Y) + |H(\tilde{P}_Y) - H(\tilde{P}_X)|^+ \right\}
\end{equation}
\begin{equation}
= \sup_{t \in [0,1]} \left\{ t H_{\tilde{P}_Y} - t H_{\tilde{P}_X} + \frac{1+s}{s} D(\tilde{P}_X || P_X) + D(\tilde{P}_Y || Q_Y) \right\} + \delta_n + \delta'_n,
\end{equation}
where (17) follows from the derivation (10)-(11).

Similarly we can prove the converse part for other values of $\alpha$.

IV. PROOF OF THEOREM 3

We first prove the following bounds for the normalized and unnormalized Rényi conversion rates for general simulation problem (the seed and target distributions are not limited to product distributions). For general distributions $P_X$ and $Q_Y$, we use $P_{X^n}$ to approximate $Q_{Y^n}$. Define $F_{P_{X^n}}(j) = P_{X^n} (x^n : -\frac{1}{\pi} \log P_X(x^n) < j)$ and $F_{P_{X^n}}^{-1}(j) := \sup \{ j : F_{P_{X^n}}(j) \leq \theta \}$. For $Q_{Y^n}$, we define $F_{Q_{Y^n}}$ and $F_{Q_{Y^n}}^{-1}$ similarly. Then we have the following bounds.

Lemma 3. We have
\[
\sup \left\{ R : \frac{1}{n} \log \sup_{j \geq 0} \frac{F_{P_{X^n}}(\binom{n}{j} \theta - j)}{F_{Q_{Y^n}}(j)} \leq 1 \right\}
\sup \left\{ R : \frac{1}{n} \log \sup_{j \geq 0} \frac{1 - F_{Q_{Y^n}}(j)}{1 - F_{P_{X^n}}(j)} \leq 1 \right\}
\geq \sup \left\{ R : \frac{1}{n} D_{\infty}(P_{X^n} || Q_{Y^n}) \rightarrow 0 \right\}
\geq \sup \left\{ R : D_{\infty}(P_{X^n} || Q_{Y^n}) \rightarrow 0 \right\}
\geq \sup \left\{ R : \liminf_{n \to \infty} \inf_{\theta \in [0,1]} \left\{ \frac{k}{n} F_{P_{X^n}}^{-1}(\theta) - F_{Q_{Y^n}}^{-1}(\theta) \right\} \right\}.
\]

Now we turn back to proving Theorem 3. Consider product distributions $P_{X^n}$ and $Q_{Y^n}$. Then
\[
F_{P_{X^n}}(j) = F_{P_X}(x^n : \sum_{x \in \mathcal{X}} T_X(x) \log P_X(x) < j)
\geq e^{-k \min \frac{1}{n} \log \sup_{j \geq 0} \frac{F_{P_{X^n}}(\binom{n}{j} \theta - j)}{F_{Q_{Y^n}}(j)}} \leq 0
\]
and
\[
1 - F_{P_{X^n}}(j) \geq e^{-k \min \frac{1}{n} \log \sup_{j \geq 0} \frac{1 - F_{Q_{Y^n}}(j)}{1 - F_{P_{X^n}}(j)}} \leq 0
\]
respectively imply for any $j > J$ with $J_{\min} := \max \left\{ \frac{1}{\pi} H_{\tilde{P}_X}(P_X), H_{\tilde{P}_Y}(Q_Y) \right\}$.

Therefore,
\[
\inf \left\{ \tilde{P}_X : \sum \tilde{P}_X(x) \log P_X(x) < J \right\} \geq \frac{1}{R} D(\tilde{P}_X || P_X)
\geq \inf \left\{ \tilde{P}_Y : \sum \tilde{P}_Y(y) \log Q_Y(y) < J \right\} D(\tilde{P}_Y || Q_Y)
\]

By this lemma, we have
\[
\sum_{y^n \in T_{Y^n}} \left| f^{-1}(y^n) \cap \mathcal{T}_{T^n} \right|^{1+s} \geq |T_{T^n}| \left( \frac{1+s}{s} |T_{T^n}| \right)^{1+s} \left\{ 1 \left[ |T_{T^n}| \geq |T_{T^n}| \right] + |T_{T^n}| \left\{ 1 \left[ |T_{T^n}| < |T_{T^n}| \right] \right\} \right.
\]
\begin{equation}
\geq e^{(1+s)H(T_X) - s n H(T_X)} \left\{ 1 \left[ H(T_X) \geq H(T_Y) \right] + e^{n H(T_X)} \right\}.
\end{equation}

Therefore,
\[
\frac{1}{n} D_{1+s}(P_{Y^n} || Q_{Y^n}) \geq \max_{T_Y} \min_{T_X} \left\{ -\frac{1+s}{s} D(T_X || P_X) + D(T_Y || Q_Y) + |H(T_Y) - H(T_X)|^+ \right\}
\geq \sup_{t \in [0,1]} \left\{ t H_{\tilde{P}_Y} - t H_{\tilde{P}_X} + \frac{1+s}{s} D(\tilde{P}_X || P_X) + D(\tilde{P}_Y || Q_Y) \right\} + \delta_n + \delta'_n,
\]
where (17) follows from the derivation (10)-(11).
and for any $j < J_{\text{max}}$ with $J_{\text{max}} := \min \{ \frac{1}{R} H_{-\infty}(P_X), H_{-\infty}(Q_Y) \},$

$$\min_{\tilde{P}_X := \sum_y \tilde{P}_X(y) \log P_X(x) \geq Y} \frac{1}{R} D(\tilde{P}_X || P_X) \leq \min_{\tilde{P}_Y := \sum_y \tilde{P}_Y(y) \log Q_Y(y) \geq Y} D(\tilde{P}_Y || Q_Y). \quad (21)$$

Observe that all the constraint functions in the optimization problems in (20) and (21) are linear. Hence Slater’s constraint qualification is satisfied. On the other hand, all these optimization problems are convex optimization problems. Hence strong duality holds [17], which implies that (20) and (21) are equivalent to any $j > J_{\text{min}}$,

$$\max_{t \in [0, \infty]} \left\{ \frac{t}{R} H_{1+t}(P_X) - t \right\} \geq \max_{t \in [0, \infty]} \left\{ t H_{1+t}(Q_Y) - t \right\},$$

and for any $j < J_{\text{max}},$

$$\max_{t \in [0, \infty]} \left\{ -\frac{t}{R} H_{1-t}(P_X) + t \right\} \leq \max_{t \in [0, \infty]} \left\{ -t H_{1-t}(Q_Y) + t \right\}.$$

Observe that $t H_{1+t}(P_X)$ is concave in $t,$ hence it can be shown that these two inequalities are equivalent to $\frac{1}{R} H_{3}(P_X) \geq H_{3}(Q_Y), \forall \beta \in [-\infty, \infty].$ Therefore, we obtain the converse part (upper bound).

Now we prove the achievable part (lower bound). Assume $R < \min_{j \in [-\infty, \infty]} H_{3}(P_X).$

Observe that $\max_{t \in [0, \infty]} \left\{ \frac{1}{R} H_{1+t}(P_X) - t \right\}$ and $\max_{t \in [0, \infty]} \left\{ t H_{1+t}(Q_Y) - t \right\}$ are strictly convex and strictly decreasing in $j$ for $j < \frac{1}{R} H(P_X)$ and $j < H(Q_Y),$ respectively. $\max_{t \in [0, \infty]} \left\{ -\frac{1}{R} H_{1-t}(P_X) + t \right\}$ and $\max_{t \in [0, \infty]} \left\{ -(t H_{1-t}(Q_Y) + t) \right\}$ are strictly convex and strictly increasing in $j$ for $j > \frac{1}{R} H(P_X)$ and $j > H(Q_Y),$ respectively. Based on these observations, we can conclude that there exists some $\epsilon > 0$ such that for any $j \geq 0,$ either $\max_{t \in [0, \infty]} \left\{ \frac{1}{R} H_{1+t}(P_X) - t \right\} \geq \max_{t \in [0, \infty]} \left\{ t H_{1+t}(Q_Y) - t \right\} + \epsilon$ or $\max_{t \in [0, \infty]} \left\{ \frac{1}{R} H_{1-t}(P_X) + t \right\} \leq \max_{t \in [0, \infty]} \left\{ -(t H_{1-t}(Q_Y) + t) \right\} - \epsilon$ holds. That is, for any $j_{\text{min}} < J < J_{\text{max}},$

$$\min_{\tilde{P}_X := \sum_x \tilde{P}_X(x) \log P_X(x) \geq Y} \frac{1}{R} D(\tilde{P}_X || P_X) \geq \min_{\tilde{P}_Y := \sum_y \tilde{P}_Y(y) \log Q_Y(y) \geq Y} D(\tilde{P}_Y || Q_Y) + \epsilon \quad (22)$$

or

$$\min_{\tilde{P}_X := \sum_x \tilde{P}_X(x) \log P_X(x) \geq Y} \frac{1}{R} D(\tilde{P}_X || P_X) \leq \min_{\tilde{P}_Y := \sum_y \tilde{P}_Y(y) \log Q_Y(y) \geq Y} D(\tilde{P}_Y || Q_Y) - \epsilon. \quad (23)$$

By (18) and (19), we know (22) and (23) respectively imply

$$\lim_{n \to \infty} \frac{1}{n} \log \sup_{j \geq 0} \frac{F_{P_{R_k}}(R_j)}{F_{Q_{R_k}}(j - \epsilon)} \leq -\epsilon,$$

$$\lim_{n \to \infty} \frac{1}{n} \log \sup_{j \geq 0} \frac{1 - F_{Q_{R_k}}(j - \epsilon)}{1 - F_{P_{R_k}}(R_j)} \leq -\epsilon,$$

which in turn respectively imply

$$\lim_{n \to \infty} \inf_{\theta \in [0, 1)} \left\{ \frac{1}{R} F_{P_{R_k}}^{\theta - \epsilon} - F_{Q_{R_k}}^{1 - \theta} - \epsilon \right\} \geq \epsilon,$$

$$\lim_{n \to \infty} \inf_{\theta \in [0, 1)} \left\{ \frac{1}{R} F_{P_{R_k}}^{1 - \theta - \epsilon} - F_{Q_{R_k}}^{\theta} - \epsilon \right\} \geq \epsilon.$$

Since $F_{P_{R_k}}^{\theta}(\theta)$ is nondecreasing in $\theta,$ we have both the two inequalities above imply

$$\lim_{k \to \infty} \inf_{\theta \in [0, 1)} \left\{ \frac{1}{R} F_{P_{R_k}}^{\theta - \epsilon} - F_{Q_{R_k}}^{1 - \theta} - \epsilon \right\} \geq \epsilon. \quad (24)$$

Therefore, (24) always holds. Observe that $F_{P_{R_k}}^{\theta}(\theta) \in [H_{-\infty}(P_X), H_{-\infty}(P_X)]$ is bounded for any $\theta \in [0, 1),$ hence (24) also holds if $R$ is replaced with $\frac{R}{2}.$ Combining this with Lemma 3 completes the proof for the lower bound.

Acknowledgments

The authors are supported in part by a Singapore National Research Foundation (NRF) National Cybersecurity R&D Grant (R-263-000-C07-281 and NRF2015NCRNCR003-006), and in part by a National Natural Science Foundation of China (NSFC) under Grant (61631017).

REFERENCES

[1] T. S. Han. Information-spectrum methods in information theory. Springer, 2003.
[2] W. Kamagai and M. Hayashi. Second-order asymptotics of conversions of distributions and entangled states based on rayleigh-normal probability distributions. IEEE Trans. Inf. Theory, 63(3):1829–1857, 2017.
[3] T. Van Erven and P. Harremoës. Rényi divergence and Kulback-Leibler divergence. IEEE Trans. Inf. Theory, 60(7):3797–3820, 2014.
[4] T. Han and S. Verdú. Approximation theory of output statistics. IEEE Trans. Inf. Theory, 39(3):752–772, 1993.
[5] M. Hayashi. General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel. IEEE Trans. Inf. Theory, 52(4):1562–1575, 2006.
[6] M. Hayashi. Exponential decreasing rate of leaked information in universal random privacy amplification. IEEE Trans. Inf. Theory, 57(6):3989–4001, 2011.
[7] L. Yu and V. Y. F. Tan. Rényi resolvability and its applications to the wiretap channel. arXiv preprint 1707.00810, 2017.
[8] S. Vembu and S. Verdú. Generating random bits from an arbitrary source: Fundamental limits. IEEE Trans. Inf. Theory, 41(5):1322–1332, 1995.
[9] L. Yu and V. Y. F. Tan. Simulation of random variables under Rényi divergence measures of all orders. arXiv preprint arXiv:1805.12451, 2018.
[10] H. Jeffreys. An invariant form for the prior probability in estimation problems. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, pages 453–461, 1946.
[11] P. Cuff and L. Yu. Differential privacy as a mutual information constraint. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pages 43–54. ACM, 2016.
[12] I. Mironov. Rényi differential privacy. In Computer Security Foundations Symposium (CSF), 2017 IEEE 30th, pages 263–275. IEEE, 2017.
[13] C. Dwork. Differential privacy: A survey of results. In International Conference on Theory and Applications of Models of Computation, pages 1–19. Springer, 2008.
[14] F. di Pietro Calmon and N. Fawaz. Privacy against statistical inference. In Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on, pages 1401–1408. IEEE, 2012.
[15] I. Csiszár and J. Körner. Information Theory: Coding Theorems for Discrete Memoryless Systems. Cambridge University Press, 2011.
[16] M. Sion. On general minimax theorems. Pacific J. Math, 8(1):171–176, 1958.
[17] S. Boyd and L. Vandenberghe. Convex optimization. Cambridge university press, 2004.