Variable-Length Intrinsic Randomness Allowing Positive Value of the Average Variational Distance

Jun Yoshizawa
Waseda University
Email: junbadchel0313@suou.waseda.jp

Shota Saito
Waseda University
Email: shota@aoni.waseda.jp

Toshiyasu Matsushima
Waseda University
Email: toshimat@waseda.jp

Abstract—This paper considers the problem of variable-length intrinsic randomness. We propose the average variational distance as the performance criterion from the viewpoint of a dual relationship with the problem formulation of variable-length resolvability. Previous study has derived the general formula of the ϵ-variable-length resolvability. We derive the general formula of the ϵ-variable-length intrinsic randomness. Namely, we characterize the supremum of the mean length under the constraint that the value of the average variational distance is smaller than or equal to a constant ϵ. Our result clarifies a dual relationship between the general formula of ϵ-variable-length resolvability and that of ϵ-variable-length intrinsic randomness. We also derive a lower bound of the quantity characterizing our general formula.

I. INTRODUCTION

The problem of random number generation is one of the important research topics in *Shannon theory*. This problem is divided into

- (i) the problem of resolvability (e.g. [1], [3], [7], [10], [11]),
- (ii) the problem of intrinsic randomness (e.g. [1], [2], [8], [9]).

For these problems, the *variational distance* is a major criterion used to measure the difference between the probability distribution generated by a mapping from a *coin distribution* [1] and a *target distribution* [1]. Further, those problems are divided into

- (A) the case of fixed-length (e.g. [1], [3], [7], [8]),
- (B) the case of variable-length (e.g. [1], [2], [9], [10], [11]).

Investigating a duality between resolvability and intrinsic randomness is one of the important research topics. For the problems of fixed-length resolvability ((i) & (A)) and fixed-length intrinsic randomness ((ii) & (A)), a duality of those general formulae has been studied. One way to capture the dual relationship of the general formulae is to see them from the viewpoint of the *smooth Rényi entropy* [6]. For the problem of fixed-length resolvability, Uyematsu [7] has characterized the general formula by using the *smooth Rényi entropy of order zero* [6]. On the other hand, for the problem of fixed-length intrinsic randomness, Uyematsu and Kunimatsu [8] have characterized the general formula by using the *smooth Rényi entropy of order infinity* [6].

For the problem of variable-length resolvability ((i) & (B)), Yagi and Han [10], [11] have characterized the infimum of the mean length allowing positive value of the variational distance.

However, the dual problem formulation to this problem has not been discussed yet.

This paper considers the problem of variable-length intrinsic randomness ((ii) & (B)) and discusses the duality with the work by Yagi and Han [10], [11]. From the viewpoint of a dual relationship with the problem formulation of variable-length resolvability, we propose the average variational distance. This is the expectation of the variational distance between the probability distribution generated by a mapping and the uniform distribution for each length, where the expectation is taken with respect to the length. As the main result, we characterize the supremum of the mean length allowing positive value of the average variational distance.

We can see a duality between the general formula by Yagi and Han [10], [11] and our general formula from the viewpoint of the smooth Rényi entropy. The general formula of Yagi and Han [10], [11] is related to the *smooth Rényi entropy of order* $\alpha \in (0,1)$ [6] (cf. [4], [5], [11]). On the other hand, our general formula is related to the sub-probability distribution which achieves the infimum of the *smooth Rényi entropy of order* $\alpha \in (1,\infty)$ [6] (cf. [4]).

It is worth noticing that our problem formulation is different from the original formulation introduced by Vembu and Verdú [9]. Vembu and Verdú [9] and Han [1], [2] have derived the general formula of the supremum of the mean length under the constraint that the value of the variational distance is equal to zero. Their variational distance measures the supremum of the difference between a *conditional probability distribution* generated by a mapping given each length and a uniform distribution. On the other hand, we consider one probability distribution on all lengths, because we consider the average variational distance. Therefore, our problem formulation is different from their problem formulation.

The organization of this paper is as follows. In Sec. II, we introduce the result of variable-length intrinsic randomness by Han [1], [2]. In Sec. III, we state the result of ϵ-variable-length resolvability by Yagi and Han [10], [11]. In Sec. IV, we describe ϵ-variable-length intrinsic randomness. Specifically, Sec. IV-A introduces the problem formulation of our study. In Sec. IV-B, we describe the general formula of the ϵ-variable-length intrinsic randomness. Further, we state a lower bound of the quantity characterizing the ϵ-variable-length intrinsic randomness. In Sec. V, we prove our results. In Sec. VI, we discuss the dual relationship between the result by Yagi

and Han [10], [11] and our result. In Sec. VII, we consider the second-order general formula. Finally, in Sec. VIII, we summarize this paper.

II. VARIABLE-LENGTH INTRINSIC RANDOMNESS: REVIEW

Let $\mathcal{X}$ be a finite or countably infinite alphabet and $\mathcal{X}^n$ be the n-th Cartesian product of $\mathcal{X}$. Let X^n be a random variable taking a value in $\mathcal{X}^n$ and x^n be a realization of X^n. Let $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ be a *general source* [1]. We denote by $\mathcal{P}(\mathcal{X}^n)$ a set of probability distribution P_{X^n} on $\mathcal{X}^n$. We do not impose any assumptions such as stationarity or ergodicity. Let $\mathcal{U} = \{0, 1, \ldots, K-1\}$ be a finite alphabet of size K, where K is an integer greater than or equal to 2. For any nonnegative integer m, $U^{(m)}$ denotes a random variable distributed uniformly on $\mathcal{U}^m$ and $u^{(m)}$ denotes a realization of $U^{(m)}$, where m is called the *length* of $U^{(m)}$. Let $\mathcal{U}^*$ be the set of all finite strings taken from $\mathcal{U}$, including the null string Λ whose length is zero, i.e., $\mathcal{U}^* = \{\Lambda, 0, 1, 00, \ldots\}$. A mapping φ_n is defined as $\varphi_n : \mathcal{X}^n \to \mathcal{U}^*$. Let $l(\varphi_n(x^n))$ be the length of $\varphi_n(x^n)$. Given m and φ_n, the set $\mathcal{D}_m$ is defined as

$$\mathcal{D}_m = \{x^n \in \mathcal{X}^n \mid l(\varphi_n(x^n)) = m\}.$$

Given φ_n, the set $\mathcal{J}(\varphi_n)$ is defined as

$$\mathcal{J}(\varphi_n) = \{m \in \mathbb{Z}_{\geq 0} \mid \mathbb{P}[l(\varphi_n(X^n)) = m] > 0\},$$

where $\mathbb{Z}_{\geq 0}$ is the set of nonnegative integers. We denote by[1] $\iota_{P_X}(x) := \log \frac{1}{P_X(x)}$. The *variational distance* between two probability distributions P_X and Q_X is defined as $d(P_X, Q_X) := \frac{1}{2} \sum_{x \in \mathcal{X}} |P_X(x) - Q_X(x)|$. A probability distribution $P_{X_m^n}$ is defined as

$$P_{X_m^n}(x^n) = \frac{P_{X^n}(x^n)}{\mathbb{P}[X^n \in \mathcal{D}_m]} \quad (x^n \in \mathcal{D}_m).$$

Previous studies such as [1], [2], and [9] investigated the *problem of variable-length intrinsic randomness* defined as follows.

Definition 1 ([1], [2]): A rate R is said to be *i-achievable* if there exists a mapping $\varphi_n : \mathcal{X}^n \to \mathcal{U}^*$ satisfying

$$\limsup_{n \to \infty} \sup_{m \in \mathcal{J}(\varphi_n)} d(P_{\varphi_n(X_m^n)}, P_{U^{(m)}}) = 0, \qquad (1)$$

$$\liminf_{n \to \infty} \frac{1}{n} \mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))] \geq R,$$

where $\mathbb{E}_{P_{X^n}}[\cdot]$ denotes the expectation with respect to the distribution P_{X^n}.

The *variable-length intrinsic randomness* [1], [2] is defined as follows.

Definition 2 ([1], [2]):

$$S_{\mathrm{i}}(\mathbf{X}) := \sup\{R \mid R \text{ is i-achievable}\}.$$

The following result was given by Han [1], [2].

Theorem 1 ([1], [2]): For any general source $\mathbf{X}$,

$$S_{\mathrm{i}}(\mathbf{X}) = \liminf_{n \to \infty} \frac{1}{n} H(P_{X^n}),$$

where $H(P_{X^n})$ is the *entropy*.

[1] In this paper, logarithms are of base K.

III. ϵ-VARIABLE-LENGTH RESOLVABILITY: REVIEW

Let L_n be a random variable taking a length m. Let $U^{(L_n)}$ be the *variable-length uniform random number* [10], [11], where the probability distribution is defined as

$$P_{U^{(L_n)}}(u^{(m)}, m) = \mathbb{P}[U^{(L_n)} = u^{(m)}, L_n = m] = \frac{\mathbb{P}[L_n = m]}{K^m},$$

for all $u^{(m)} \in \mathcal{U}^m$. Therefore, $U^{(m)}$ is uniformly distributed over $\mathcal{U}^m$ given $L_n = m$. A mapping ϕ_n is defined as $\phi_n : \mathcal{U}^* \to \mathcal{X}^n$.

Previous studies such as [10] and [11] investigated the *problem of ϵ-variable-length resolvability* defined as follows.

Definition 3 ([10], [11]): Given $\epsilon \in [0, 1)$, a rate R is said to be *r(ϵ)-achievable* if there exists a variable-length uniform random number $U^{(L_n)}$ and a mapping $\phi_n : \mathcal{U}^* \to \mathcal{X}^n$ satisfying

$$\limsup_{n \to \infty} d(P_{\phi_n(U^{(L_n)})}, P_{X^n}) \leq \epsilon,$$

$$\limsup_{n \to \infty} \frac{1}{n} \mathbb{E}_{P_{L_n}}[L_n] \leq R.$$

The *ϵ-variable-length resolvability* [10], [11] is defined as follows.

Definition 4 ([10], [11]):

$$S_{\mathrm{r}}(\epsilon|\mathbf{X}) := \inf\{R \mid R \text{ is r}(\epsilon)\text{-achievable}\}.$$

The following quantity was defined by Koga and Yamamoto [5].

Definition 5 ([5]): Given $\epsilon \in [0, 1)$, $G_{[\epsilon]}(\mathbf{X})$ is defined as

$$G_{[\epsilon]}(\mathbf{X}) = \lim_{\tau \downarrow 0} \limsup_{n \to \infty} \frac{1}{n} G_{[\epsilon+\tau]}(X^n),$$

where $G_{[\epsilon+\tau]}(X^n)$ is defined as

$$G_{[\epsilon+\tau]}(X^n) = \inf_{A_n : \mathbb{P}[X^n \in A_n] \geq 1-\epsilon-\tau} \sum_{x^n \in A_n} P_{X^n}(x^n)$$
$$\cdot \log \frac{\mathbb{P}[X^n \in A_n]}{P_{X^n}(x^n)}.$$

The following result was given by Yagi and Han [10], [11] (cf. [5]).

Theorem 2 ([10], [11]): For any general source $\mathbf{X}$,

$$S_{\mathrm{r}}(\epsilon|\mathbf{X}) = G_{[\epsilon]}(\mathbf{X}) \quad (\epsilon \in [0, 1)).$$

Remark 1: The study [5] derived the general formula of *weak variable-length source coding* allowing ϵ-error probability. The general formula is also characterized by $G_{[\epsilon]}(\mathbf{X})$.

IV. ϵ-VARIABLE-LENGTH INTRINSIC RANDOMNESS

A. Problem Formulation

In this study, let $\mathcal{X}$ be a finite alphabet. In the problem of variable-length intrinsic randomness, the probability distribution of L_n is defined as $P_{L_n}(m) = \mathbb{P}[X^n \in \mathcal{D}_m]$. Therefore, the probability distribution of $U^{(L_n)}$ is defined as

$$P_{U^{(L_n)}}(u^{(m)}, m) = \mathbb{P}[U^{(L_n)} = u^{(m)}, L_n = m] = \frac{\mathbb{P}[X^n \in \mathcal{D}_m]}{K^m},$$

for all $u^{(m)} \in \mathcal{U}^m$. The performance criteria are the *average variational distance* and the mean length. The average variational distance between $P_{\varphi_n(X^n)}$ and $P_{U^{(L_n)}}$ is defined as

$$
\begin{aligned}
&\bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}}) \\
&= \sum_{m \in \mathcal{J}(\varphi_n)} P_{L_n}(m) \frac{1}{2} \sum_{u \in \mathcal{U}^m} |P_{\varphi_n(X_m^n)}(u) - P_{U^{(m)}}(u)| \\
&= \frac{1}{2} \sum_{m \in \mathcal{J}(\varphi_n)} \mathbb{P}[X^n \in \mathcal{D}_m] \sum_{u \in \mathcal{U}^m} \left| \frac{P_{\varphi_n(X^n)}(u)}{\mathbb{P}[X^n \in \mathcal{D}_m]} - \frac{1}{K^m} \right| \\
&= \frac{1}{2} \sum_{m \in \mathcal{J}(\varphi_n)} \sum_{u \in \mathcal{U}^m} \left| P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in \mathcal{D}_m]}{K^m} \right|.
\end{aligned}
$$

We define the *problem of ϵ-variable-length intrinsic randomness*.

Definition 6: Given $\epsilon \in [0, 1)$, a rate R is said to be $i(\epsilon)$-*achievable* if there exists a mapping $\varphi_n : \mathcal{X}^n \to \mathcal{U}^*$ satisfying

$$\limsup_{n \to \infty} \bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}}) \le \epsilon, \tag{2}$$

$$\liminf_{n \to \infty} \frac{1}{n} \mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))] \ge R. \tag{3}$$

Remark 2: The variational distance (1) in Definition 1 measures the supremum of the difference between a conditional probability distribution given each length generated by a mapping and a uniform distribution. On the other hand, the average variational distance (2) in Definition 6 measures the difference between a probability distribution generated by a mapping and a probability distribution of variable-length uniform random number $U^{(L_n)}$. Therefore, unlike Definition 1, we consider one probability distribution on all lengths in Definition 6.

The *ϵ-variable-length intrinsic randomness* is defined as follows.

Definition 7:

$$S_i(\epsilon|\mathbf{X}) := \sup\{R \mid R \text{ is } i(\epsilon)\text{-achievable}\}.$$

Our concern is to investigate $S_i(\epsilon|\mathbf{X})$ for a general source $\mathbf{X}$.

B. Main Results

The following set plays an important role in producing our main results.

Definition 8: Given $\delta \in [0, 1)$, $\mathcal{Q}_\delta(\mathcal{X}^n)$ is defined as the set of *sub-probability distribution* Q_{X^n} satisfying the following conditions:

$$
\begin{aligned}
&Q_{X^n}(x^n) > 0, \ (\forall x^n \in \{x^n \in \mathcal{X}^n \mid P_{X^n}(x^n) > 0\}), \\
&Q_{X^n}(x^n) \le P_{X^n}(x^n), \ (\forall x^n \in \mathcal{X}^n), \\
&\sum_{x^n \in \mathcal{X}^n} Q_{X^n}(x^n) = 1 - \delta.
\end{aligned}
$$

Next, using $\mathcal{Q}_\epsilon(\mathcal{X}^n)$, we introduce a new quantity.

Definition 9: Given $\epsilon \in [0, 1)$, $G^{[\epsilon]}(\mathbf{X})$ is defined as

$$G^{[\epsilon]}(\mathbf{X}) = \lim_{\tau \downarrow 0} \liminf_{n \to \infty} \frac{1}{n} G^{[\epsilon+\tau]}(X^n),$$

where $G^{[\epsilon+\tau]}(X^n)$ is defined as

$$G^{[\epsilon+\tau]}(X^n) = \sup_{Q_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)} \mathbb{E}_{P_{X^n}}[\iota_{Q_{X^n}}(X^n)].$$

The following theorem is the main result in this paper.

Theorem 3: For any general source $\mathbf{X}$,

$$S_i(\epsilon|\mathbf{X}) = G^{[\epsilon]}(\mathbf{X}) \quad (\epsilon \in [0, 1)).$$

Proof: The proofs of the direct part and the converse part are in Section V-A and Section V-B, respectively. ∎

In Sec. VI, we discuss the dual relationship between the general formula in Theorem 2 by Yagi and Han [10], [11] and our general formula in Theorem 3.

Remark 3: Instead of (2), we consider the next condition:

$$\bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}}) \le \epsilon \quad (\forall n \ge n_0)$$

for some $n_0 \in \mathbb{N}$. We define $\tilde{S}_i(\epsilon|\mathbf{X})$ as ϵ-variable-length intrinsic randomness corresponding to this condition. Then, we have

$$\tilde{S}_i(\epsilon|\mathbf{X}) = \liminf_{n \to \infty} \frac{1}{n} G^{[\epsilon]}(X^n).$$

The following theorem characterizes the lower bound of $G^{[\epsilon]}(\mathbf{X})$.

Theorem 4: For any general source $\mathbf{X}$,

$$\underline{H}_\epsilon(\mathbf{X}) \le G^{[\epsilon]}(\mathbf{X}) \quad (\epsilon \in [0, 1)),$$

where the quantity of the lower bound is defined as

$$\underline{H}_\epsilon(\mathbf{X}) = \sup\left\{ R \mid \limsup_{n \to \infty} \mathbb{P}\left[\frac{1}{n} \iota_{P_{X^n}}(X^n) \le R \right] \le \epsilon \right\}.$$

Proof: See Section V-C. ∎

Remark 4: The quantity $\underline{H}_\epsilon(\mathbf{X})$ characterizes ϵ-*fixed-length intrinsic randomness* [1].

V. PROOFS OF MAIN RESULTS

A. Proof of the Direct Part of Theorem 3

For any $\tau > 0$, $n \in \mathbb{N}$, and $\gamma > 0$, there exist a $\tilde{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$ satisfying

$$
\begin{aligned}
\mathbb{E}_{P_{X^n}}[\iota_{\tilde{Q}_{X^n}}(X^n)] &> \sup_{Q_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)} \mathbb{E}_{P_{X^n}}[\iota_{Q_{X^n}}(X^n)] - \gamma \\
&= G^{[\epsilon+\tau]}(X^n) - \gamma. \tag{4}
\end{aligned}
$$

[Definitions of notation]

- R_j is defined as $R_j = 3\gamma j$, $(j = 0, 1, 2, \dots)$.
- I_j is defined as $I_j = [R_j, R_{j+1})$, $(j = 0, 1, 2, \dots)$.
- $S_n^{(j)} \subset \mathcal{X}^n$ is defined as

$$S_n^{(j)} = \left\{ x^n \in \mathcal{X}^n \mid \frac{1}{n} \iota_{\tilde{Q}_{X^n}}(x^n) \in I_j \right\}, \ (j = 0, 1, 2, \dots). \tag{5}$$

- J is defined as $J = \{0, 1, 2, \dots\}$.
- We partition J into J_1 and J_2 defined as

$$J_1 = \{j \ge 1 \mid \mathbb{P}[X^n \in S_n^{(j)}] \ge K^{-n\gamma R_j}\}, \tag{6}$$

$$J_2 = \{0\} \cup \{j \ge 1 \mid \mathbb{P}[X^n \in S_n^{(j)}] < K^{-n\gamma R_j}\}. \tag{7}$$

For $x^n \in S_n^{(j)}$, it holds that

$$\tilde{Q}_{X^n}(x^n) \le K^{-nR_j} = K^{-n\gamma R_j} K^{-n(1-\gamma)R_j}.$$

Then, for $j \in J_1$ from (6), it follows that

$$\tilde{Q}_{X_n}(x^n) \le \mathbb{P}[X^n \in S_n^{(j)}]K^{-n(1-\gamma)R_j}. \qquad (8)$$

[Construction of the mapping]

We use the following Lemma 1. The proof of this lemma is similar to that of Lemma 2.2 in [2].

Lemma 1: Let $R > 0$, $a > 0$ be any constants, $\gamma > 0$ be an arbitrarily small constant, $A_n \subset \mathcal{X}^n$ be an arbitrarily set, and $c \ge \mathbb{P}[X^n \in A_n]$ be an arbitrarily constant. Suppose that the probability distribution $P_{X^n} \in \mathcal{P}(\mathcal{X}^n)$ satisfies the condition

$$P_{X^n}(x^n) \le cK^{-n(a+\gamma)R} \quad (\forall x^n \in A_n). \qquad (9)$$

Then, there exists a mapping $\varphi_n : A_n \to \mathcal{U}^{\lfloor naR \rfloor}$ such that

$$\frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor naR \rfloor}}\left|P_{\varphi_n(X^n)}(u) - \frac{c}{K^{\lfloor naR \rfloor}}\right|$$
$$\le cK^{-n\gamma R} + \frac{1}{2}(c - \mathbb{P}[X^n \in A_n]). \qquad (10)$$

Proof: See Appendix A. ∎

Remark 5: It is easy to check from the way of the proof of Lemma 1 that Lemma 1 holds even if we replace the probability distribution P_{X^n} and $\mathbb{P}[X^n \in A_n]$ defined by P_{X^n} by sub-probability distributions.

We use Lemma 1 with $R = R_j$, $a = 1 - 2\gamma$, $A_n = S_n^{(j)}$, $c = \mathbb{P}[X^n \in S_n^{(j)}](\ge \tilde{\mathbb{Q}}[X^n \in S_n^{(j)}])$. From (8), there exists a mapping $\varphi_n^{(j)} : S_n^{(j)} \to \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}$ such that

$$\frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}\left|\tilde{Q}_{\varphi_n^{(j)}(X^n)}(u) - \frac{\mathbb{P}[X^n \in S_n^{(j)}]}{K^{\lfloor n(1-2\gamma)R_j \rfloor}}\right|$$
$$\le \mathbb{P}[X^n \in S_n^{(j)}]K^{-n\gamma R_j} + \frac{1}{2}(\mathbb{P}[X^n \in S_n^{(j)}] - \tilde{\mathbb{Q}}[X^n \in S_n^{(j)}]), \qquad (11)$$

where $\tilde{\mathbb{Q}}[X^n \in S_n^{(j)}]) = \sum_{x^n \in S_n^{(j)}} \tilde{Q}_{X^n}(x^n)$.

Next, we construct the mapping $\varphi_n : \mathcal{X}^n \to \mathcal{U}^*$ by

$$\varphi_n(x^n) := \begin{cases} \varphi_n^{(j)}(x^n), & (x^n \in S_n^{(j)}, j \in J_1), \\ \Lambda, & (\text{otherwise}). \end{cases} \qquad (12)$$

Therefore, it holds that

$$\mathcal{J}(\varphi_n) = \{0\} \cup \{\lfloor n(1-2\gamma)R_j \rfloor \mid j \in J_1\}. \qquad (13)$$

[Evaluation of the average variational distance]

From $\tilde{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$, there exists a $\{\epsilon^{(j)}\}_{j=1}^{|J|}$, $\{\tau^{(j)}\}_{j=1}^{|J|}$ such that

$$\epsilon^{(1)} + \epsilon^{(2)} + \cdots + \epsilon^{(|J|)} = \epsilon, \qquad (14)$$
$$\tau^{(1)} + \tau^{(2)} + \cdots + \tau^{(|J|)} = \tau, \qquad (15)$$
$$\mathbb{P}[X^n \in S_n^{(j)}] - \tilde{\mathbb{Q}}[X^n \in S_n^{(j)}] = \epsilon^{(j)} + \tau^{(j)}. \qquad (16)$$

For $j \in J_1$, we have

$$\frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}\left|\tilde{Q}_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in S_n^{(j)}]}{K^{\lfloor n(1-2\gamma)R_j \rfloor}}\right|$$

$$\overset{(a)}{\le} \mathbb{P}[X^n \in S_n^{(j)}]K^{-n\gamma R_j} + \frac{1}{2}(\mathbb{P}[X^n \in S_n^{(j)}] - \tilde{\mathbb{Q}}[X^n \in S_n^{(j)}])$$

$$\overset{(b)}{=} \mathbb{P}[X^n \in S_n^{(j)}]K^{-n\gamma R_j} + \frac{\epsilon^{(j)} + \tau^{(j)}}{2}, \qquad (17)$$

where (a) follows from (11) and (12), (b) follows from (16). Then, it holds that

$$\frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in S_n^{(j)}]}{K^{\lfloor n(1-2\gamma)R_j \rfloor}}\right|$$

$$\le \frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}|P_{\varphi_n(X^n)}(u) - \tilde{Q}_{\varphi_n(X^n)}(u)|$$

$$+ \frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}\left|\tilde{Q}_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in S_n^{(j)}]}{K^{\lfloor n(1-2\gamma)R_j \rfloor}}\right|$$

$$\overset{(c)}{\le} \frac{1}{2}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}|P_{\varphi_n(X^n)}(u) - \tilde{Q}_{\varphi_n(X^n)}(u)|$$

$$+ \mathbb{P}[X^n \in S_n^{(j)}]K^{-n\gamma R_j} + \frac{\epsilon^{(j)} + \tau^{(j)}}{2}, \qquad (18)$$

where (c) follows from (17). On the other hand, for $j \in J_2$, it follows that

$$\frac{1}{2}\sum_{u \in \mathcal{U}^0}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in \mathcal{D}_0]}{K^0}\right|$$

$$= \frac{1}{2}|P_{\varphi_n(X^n)}(\Lambda) - \mathbb{P}[X^n \in \mathcal{D}_0]|$$

$$= \frac{1}{2}|\mathbb{P}[X^n \in \mathcal{D}_0] - \mathbb{P}[X^n \in \mathcal{D}_0]| = 0. \qquad (19)$$

Next, for sufficient large number $n \in \mathbb{N}$, we prove that the length $\lfloor n(1-2\gamma)R_j \rfloor$ differs for each $j \in J_1$. For $n \ge \frac{1}{(1-2\gamma)3\gamma}$, we have

$$n(1-2\gamma)3\gamma(j+1) - 1 \ge n(1-2\gamma)3\gamma j. \qquad (20)$$

From the definition of the *floor function*, it holds that

$$\lfloor n(1-2\gamma)R_{j+1} \rfloor > n(1-2\gamma)3\gamma(j+1) - 1, \qquad (21)$$
$$\lfloor n(1-2\gamma)R_j \rfloor \le n(1-2\gamma)3\gamma j. \qquad (22)$$

By substituting (21) and (22) for (20), we have

$$\lfloor n(1-2\gamma)R_{j+1} \rfloor > \lfloor n(1-2\gamma)R_j \rfloor.$$

Therefore, we obtain the following fact: (♠) *For $n \ge \frac{1}{(1-2\gamma)3\gamma}$, the length $\lfloor n(1-2\gamma)R_j \rfloor$ differs for each $j \in J_1$.*

For $n \ge \frac{1}{(1-2\gamma)3\gamma}$, the combination of (12), (13), (18), (19), and (♠) yields

$$\bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}})$$

$$= \frac{1}{2}\sum_{m \in \mathcal{J}(\varphi_n)}\sum_{u \in \mathcal{U}^m}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in \mathcal{D}_m]}{K^m}\right|$$

$$= \frac{1}{2}\sum_{m \in \mathcal{J}(\varphi_n)}\sum_{u \in \mathcal{U}^m}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[l(\varphi_n(X^n)) = m]}{K^m}\right|$$

$$= \frac{1}{2}\sum_{j \in J_1}\sum_{u \in \mathcal{U}^{\lfloor n(1-2\gamma)R_j \rfloor}}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n \in S_n^{(j)}]}{K^{\lfloor n(1-2\gamma)R_j \rfloor}}\right|$$

$$\leq \frac{1}{2} \sum_{j\in J_1} \sum_{u\in\mathcal{U}^{\lfloor n(1-2\gamma)R_j\rfloor}} |P_{\varphi_n(X^n)}(u) - \tilde{Q}_{\varphi_n(X^n)}(u)|$$

$$+ \sum_{j\in J_1} \mathbb{P}[X^n \in S_n^{(j)}]K^{-n\gamma R_j} + \sum_{j\in J_1} \frac{\epsilon^{(j)} + \tau^{(j)}}{2}. \quad (23)$$

First, we evaluate the first term on the right-hand side of (23).

$$\frac{1}{2} \sum_{j\in J_1} \sum_{u\in\mathcal{U}^{\lfloor n(1-2\gamma)R_j\rfloor}} |P_{\varphi_n(X^n)}(u) - \tilde{Q}_{\varphi_n(X^n)}(u)|$$

$$= \frac{1}{2} \sum_{j\in J_1} \sum_{u\in\mathcal{U}^{\lfloor n(1-2\gamma)R_j\rfloor}} \left| \sum_{x^n:\varphi_n(x^n)=u} (P_{X^n}(x^n) - \tilde{Q}_{X^n}(x^n)) \right|$$

$$\leq \frac{1}{2} \sum_{j\in J_1} \sum_{u\in\mathcal{U}^{\lfloor n(1-2\gamma)R_j\rfloor}} \sum_{x^n:\varphi_n(x^n)=u} |P_{X^n}(x^n) - \tilde{Q}_{X^n}(x^n)|$$

$$\overset{(d)}{\leq} \frac{1}{2} \sum_{x^n\in\mathcal{X}^n} |P_{X^n}(x^n) - \tilde{Q}_{X^n}(x^n)| \overset{(e)}{=} \frac{\epsilon+\tau}{2}, \quad (24)$$

where (d) follows from (12) and (♠), (e) follows from $\tilde{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$. Next, we evaluate the second term on the right-hand side of (23).

$$\sum_{j\in J_1} \mathbb{P}[X^n \in S_n^{(j)}]K^{-n\gamma R_j} \leq \sum_{j\in J\setminus\{0\}} K^{-3n\gamma^2 j}$$

$$= \frac{K^{-3n\gamma^2}}{1 - K^{-3n\gamma^2}}. \quad (25)$$

Finally, we evaluate the third term on the right-hand side of (23).

$$\sum_{j\in J_1} \frac{\epsilon^{(j)} + \tau^{(j)}}{2} \leq \sum_{j\in J} \frac{\epsilon^{(j)} + \tau^{(j)}}{2} \overset{(f)}{=} \frac{\epsilon+\tau}{2}, \quad (26)$$

where (f) follows from (14) and (15). By substituting (24), (25), and (26) for (23), we have, for $n \geq \frac{1}{(1-2\gamma)3\gamma}$,

$$\bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}}) \leq \epsilon + \tau + \frac{K^{-3n\gamma^2}}{1 - K^{-3n\gamma^2}}.$$

By letting $\tau \downarrow 0$, we have

$$\limsup_{n\to\infty} \bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}}) \leq \epsilon. \quad (27)$$

[Evaluation of the mean length]

For $n \geq \frac{1}{(1-2\gamma)3\gamma}$, it follows that

$$\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))]$$

$$= \sum_{m\in\mathcal{J}(\varphi_n)} m\mathbb{P}[l(\varphi_n(X^n)) = m]$$

$$\overset{(g)}{=} \sum_{j\in J_1} \lfloor n(1-2\gamma)R_j \rfloor \mathbb{P}[X^n \in S_n^{(j)}]$$

$$\geq \sum_{j\in J_1} (n(1-2\gamma)R_j - 1)\mathbb{P}[X^n \in S_n^{(j)}]$$

$$\geq \sum_{j\in J_1} n(1-2\gamma)R_j\mathbb{P}[X^n \in S_n^{(j)}] - \sum_{j\in J} \mathbb{P}[X^n \in S_n^{(j)}]$$

$$= \sum_{j\in J_1} n(1-2\gamma)(R_{j+1} - 3\gamma)\mathbb{P}[X^n \in S_n^{(j)}] - 1$$

$$\geq n(1-2\gamma) \sum_{j\in J_1} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}] - 3n\gamma(1-2\gamma) - 1$$

$$= n(1-2\gamma) \sum_{j\in J} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}]$$

$$- n(1-2\gamma) \sum_{j\in J_2} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}] - 3n\gamma(1-2\gamma) - 1$$

$$= n(1-2\gamma) \sum_{j\in J} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}]$$

$$- n(1-2\gamma) \sum_{j\in J_2\setminus\{0\}} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}]$$

$$- n(1-2\gamma)R_1\mathbb{P}[X^n \in S_n^{(0)}] - 3n\gamma(1-2\gamma) - 1$$

$$\geq n(1-2\gamma) \sum_{j\in J} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}]$$

$$- n(1-2\gamma) \sum_{j\in J_2\setminus\{0\}} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}] - 6n\gamma(1-2\gamma) - 1, \quad (28)$$

where (g) follows from (12). We evaluate the second term on the right-hand side of (28).

$$\sum_{j\in J_2\setminus\{0\}} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}] \overset{(h)}{<} 3\gamma \sum_{j=1}^{\infty} (j+1)K^{-3n\gamma^2 j}$$

$$= \frac{3\gamma K^{-3n\gamma^2}}{1 - K^{-3n\gamma^2}} + \frac{3\gamma K^{-3n\gamma^2}}{(1 - K^{-3n\gamma^2})^2} \leq \frac{6\gamma K^{-3n\gamma^2}}{(1 - K^{-3n\gamma^2})^2}, \quad (29)$$

where (h) follows from (7). By substituting (29) for (28), it holds that

$$\frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))] \geq (1-2\gamma) \sum_{j\in J} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}]$$

$$- \frac{6\gamma(1-2\gamma)K^{-3n\gamma^2}}{(1 - K^{-3n\gamma^2})^2} - 6\gamma(1-2\gamma) - \frac{1}{n}. \quad (30)$$

We evaluate the first term on the right-hand side of (30).

$$(1-2\gamma) \sum_{j\in J} R_{j+1}\mathbb{P}[X^n \in S_n^{(j)}]$$

$$\overset{(i)}{>} (1-2\gamma) \sum_{j\in J} \sum_{x^n\in S_n^{(j)}} P_{X^n}(x^n)\frac{1}{n}\iota_{\tilde{Q}_{X^n}}(x^n)$$

$$= (1-2\gamma)\frac{1}{n}\mathbb{E}_{P_{X^n}}[\iota_{\tilde{Q}_{X^n}}(X^n)], \quad (31)$$

where (i) follows from (5). The combination of (31) and (30) yields

$$\frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))]$$

$$\geq (1-2\gamma)\frac{1}{n}\mathbb{E}_{P_{X^n}}[\iota_{\tilde{Q}_{X^n}}(X^n)] - \frac{6\gamma(1-2\gamma)K^{-3n\gamma^2}}{(1 - K^{-3n\gamma^2})^2}$$

$$- 6\gamma(1-2\gamma) - \frac{1}{n}$$

$$\overset{(j)}{\geq} (1-2\gamma)\frac{1}{n}G^{[\epsilon+\tau]}(X^n) - \frac{(1-2\gamma)\gamma}{n}$$

$$- \frac{6\gamma(1-2\gamma)K^{-3n\gamma^2}}{(1 - K^{-3n\gamma^2})^2} - 6\gamma(1-2\gamma) - \frac{1}{n},$$

where (j) follows from (4). Therefore, it follows that

$$\liminf_{n\to\infty}\frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))]$$
$$\geq (1-2\gamma)\liminf_{n\to\infty}\frac{1}{n}G^{[\epsilon+\tau]}(X^n) - 6\gamma(1-2\gamma).$$

Since $\gamma > 0$ and $\tau > 0$ are arbitrary, this indicates that

$$\liminf_{n\to\infty}\frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))] \geq G^{[\epsilon]}(\mathbf{X}). \tag{32}$$

From (27) and (32), R satisfying $R < G^{[\epsilon]}(\mathbf{X})$ is i(ϵ)-achievable. Hence, we have $S_i(\epsilon|\mathbf{X}) \geq G^{[\epsilon]}(\mathbf{X})$.

B. Proof of the Converse Part of Theorem 3

Suppose that R is i(ϵ)-achievable, i.e., suppose that there exists a mapping $\varphi_n : \mathcal{X}^n \to \mathcal{U}^*$ satisfying (2) and (3). From (2), for $\tau > 0$, there exists an $n_0 \in \mathbb{N}$ such that

$$\frac{1}{2}\sum_{m\in\mathcal{J}(\varphi_n)}\sum_{u\in\mathcal{U}^m}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right| \leq \epsilon + \tau$$

for all $n \geq n_0$. There exists a sequence $\{\epsilon^{(i)}\}_{i=1}^{|\mathcal{J}(\varphi_n)|}$, $\{\tau^{(i)}\}_{i=1}^{|\mathcal{J}(\varphi_n)|}$ such that

$$\epsilon^{(1)} + \epsilon^{(2)} + \cdots + \epsilon^{(|\mathcal{J}(\varphi_n)|)} \leq \epsilon, \tag{33}$$
$$\tau^{(1)} + \tau^{(2)} + \cdots + \tau^{(|\mathcal{J}(\varphi_n)|)} \leq \tau, \tag{34}$$

for $m \in \mathcal{J}(\varphi_n)$,

$$\frac{1}{2}\sum_{u\in\mathcal{U}^m}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right| = \epsilon^{(m)} + \tau^{(m)}.$$

Next, the set $A_n^{(m)} \subset \mathcal{X}^n$ for $m \in \mathcal{J}(\varphi_n)$ is defined as follows:

$$A_n^{(m)} = \left\{x^n\in\mathcal{X}^n \mid P_{X^n}(x^n) \geq \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}, \varphi_n(x^n)\in\mathcal{U}^m\right\}.$$

Moreover, we define the set $V_m \supset \varphi_n(A_n^{(m)})$ by

$$V_m = \left\{u\in\mathcal{U}^m \mid P_{\varphi_n(X^n)}(u) \geq \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right\}.$$

For $n \geq n_0$ and $m \in \mathcal{J}(\varphi_n)$, it holds that

$$\epsilon^{(m)} + \tau^{(m)} = \frac{1}{2}\sum_{u\in\mathcal{U}^m}\left|P_{\varphi_n(X^n)}(u) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right|$$
$$= \frac{1}{2}\sum_{u\in\mathcal{U}^m}\left|\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right|$$
$$= \frac{1}{2}\sum_{u\in V_m}\left(\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right)$$
$$+ \frac{1}{2}\sum_{u\in\mathcal{U}^m\setminus V_m}\left(\frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m} - \sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n)\right)$$
$$= \frac{1}{2}\sum_{u\in V_m}\left(\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right)$$
$$+ \frac{1}{2}\left(\mathbb{P}[X^n\in\mathcal{D}_m] - \sum_{u\in V_m}\frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right)$$

$$- \frac{1}{2}\left(\mathbb{P}[X^n\in\mathcal{D}_m] - \sum_{u\in V_m}\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n)\right)$$
$$= \sum_{u\in V_m}\left(\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right)$$
$$\overset{(k)}{\geq} \sum_{u\in\varphi_n(A_n^{(m)})}\left(\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right)$$
$$= \sum_{u\in\varphi_n(A_n^{(m)})}\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{|\varphi_n(A_n^{(m)})|\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}$$
$$\geq \sum_{u\in\varphi_n(A_n^{(m)})}\sum_{x^n:\varphi_n(x^n)=u}P_{X^n}(x^n) - \frac{|A_n^{(m)}|\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}$$
$$\geq \sum_{x^n\in A_n^{(m)}}P_{X^n}(x^n) - \frac{|A_n^{(m)}|\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}$$
$$= \sum_{x^n\in A_n^{(m)}}\left(P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right),$$

where (k) follows from $V_m \supset \varphi_n(A_n^{(m)})$. Hence, from (33) and (34), it follows that

$$\sum_{m\in\mathcal{J}(\varphi_n)}\sum_{x^n\in A_n^{(m)}}\left(P_{X^n}(x^n) - \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}\right) \leq \epsilon + \tau. \tag{35}$$

From (35), the definition of $A_n^{(m)}$, and the definition of $\mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$, there exists a $\hat{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$ such that

$$\hat{Q}_{X^n}(x^n) \leq \frac{\mathbb{P}[X^n\in\mathcal{D}_m]}{K^m}, \tag{36}$$

for all $m \in \mathcal{J}(\varphi_n)$, and $x^n \in \mathcal{X}^n$ satisfying $\varphi_n(x^n) \in \mathcal{U}^m$. Hence, for $n \geq n_0$, it holds that

$$\frac{1}{n}G^{[\epsilon+\tau]}(X^n)$$
$$\overset{(l)}{\geq} \frac{1}{n}\mathbb{E}_{P_{X^n}}[\iota_{\hat{Q}_{X^n}}(X^n)]$$
$$= \frac{1}{n}\sum_{m\in\mathcal{J}(\varphi_n)}\sum_{x^n\in\mathcal{D}_m}P_{X^n}(x^n)\iota_{\hat{Q}_{X^n}}(x^n)$$
$$\overset{(m)}{\geq} \frac{1}{n}\sum_{m\in\mathcal{J}(\varphi_n)}\sum_{x^n\in\mathcal{D}_m}P_{X^n}(x^n)\log\frac{K^m}{\mathbb{P}[X^n\in\mathcal{D}_m]}$$
$$\geq \frac{1}{n}\sum_{m\in\mathcal{J}(\varphi_n)}\sum_{x^n\in\mathcal{D}_m}mP_{X^n}(x^n)$$
$$= \frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))],$$

where (l) follows from $\hat{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$, (m) follows from (36). Therefore, it follows that

$$\liminf_{n\to\infty}\frac{1}{n}G^{[\epsilon+\tau]}(X^n) \geq \liminf_{n\to\infty}\frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))].$$

Since $\tau > 0$ is arbitrary, this indicates that

$$G^{[\epsilon]}(\mathbf{X}) \geq \liminf_{n\to\infty}\frac{1}{n}\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))]. \tag{37}$$

By (3) and (37), $G^{[\epsilon]}(\mathbf{X}) \geq R$. Hence, we have $S_{\mathrm{i}}(\epsilon|\mathbf{X}) \leq G^{[\epsilon]}(\mathbf{X})$.

C. Proof of Theorem 4

For any $\gamma > 0$, we define $R_0 = \underline{H}_\epsilon(\mathbf{X}) - \gamma$. From the definition of $\underline{H}_\epsilon(\mathbf{X})$, it holds that

$$\limsup_{n \to \infty} \mathbb{P}\left[\frac{1}{n} \iota_{P_{X^n}}(X^n) \leq R_0\right] \leq \epsilon. \tag{38}$$

For $\tau > 0$, there exists an $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}\left[\frac{1}{n} \iota_{P_{X^n}}(X^n) \leq R_0\right] \leq \limsup_{n \to \infty} \mathbb{P}\left[\frac{1}{n} \iota_{P_{X^n}}(X^n) \leq R_0\right] + \tau, \tag{39}$$

for all $n \geq n_0$. From (38) and (39), for all $n \geq n_0$, it follows that

$$\mathbb{P}\left[\frac{1}{n} \iota_{P_{X^n}}(X^n) \leq R_0\right] \leq \epsilon + \tau. \tag{40}$$

The set $A_n \subset \mathcal{X}^n$ is defined by

$$A_n = \left\{ x^n \in \mathcal{X}^n \mid \frac{1}{n} \iota_{P_{X^n}}(X^n) \leq R_0 \right\}.$$

For $x^n \in A_n$, it holds that

$$P_{X^n}(x^n) \geq K^{-nR_0}. \tag{41}$$

Next, we define the sub-probability $\tilde{Q}_{X^n}$ by

$$\tilde{Q}_{X^n}(x^n) = \begin{cases} \frac{K^{-nR_0}}{|A_n|}, & (x^n \in A_n), \\ P_{X^n}(x^n), & (x^n \notin A_n). \end{cases} \tag{42}$$

From (41) and (42), for $x^n \in \mathcal{X}^n$, it follows that

$$\tilde{Q}_{X^n}(x^n) \leq P_{X^n}(x^n), \tag{43}$$

which implies that

$$\frac{1}{n} \iota_{\tilde{Q}_{X^n}}(x^n) = \frac{1}{n} \log \frac{|A_n|}{K^{-nR_0}} = R_0 + \frac{1}{n} \log |A_n|, \tag{44}$$

for $x^n \in A_n$. On the other hand, for $x^n \notin A_n$, we have

$$\frac{1}{n} \iota_{\tilde{Q}_{X^n}}(x^n) = \frac{1}{n} \iota_{P_{X^n}}(x^n) > R_0. \tag{45}$$

From (44) and (45), for $x^n \in \mathcal{X}^n$, it follows that

$$\frac{1}{n} \iota_{\tilde{Q}_{X^n}}(x^n) \geq R_0. \tag{46}$$

Further, for $n \geq n_0$, it holds that

$$\sum_{x^n \in \mathcal{X}^n} \tilde{Q}_{X^n}(x^n) = \sum_{x^n \in A_n} \tilde{Q}_{X^n}(x^n) + \sum_{x^n \notin A_n} \tilde{Q}_{X^n}(x^n)$$
$$= \sum_{x^n \in A_n} \frac{K^{-nR_0}}{|A_n|} + \sum_{x^n \notin A_n} P_{X^n}(x^n)$$
$$= K^{-nR_0} + \mathbb{P}\left[\frac{1}{n} \iota_{P_{X^n}}(X^n) > R_0\right]$$
$$\overset{(n)}{\geq} K^{-nR_0} + 1 - \epsilon - \tau \geq 1 - \epsilon - \tau, \tag{47}$$

where (n) follows from (40). From the definition of $\mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$, (43), and (47), for $n \geq n_0$, there exists a $\bar{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$ such that

$$\iota_{\bar{Q}_{X^n}}(x^n) \geq \iota_{\tilde{Q}_{X^n}}(x^n) \geq \iota_{P_{X^n}}(x^n), \tag{48}$$

for all $x^n \in \mathcal{X}^n$. From (46) and (48), for any $n \geq n_0$ and $x^n \in \mathcal{X}^n$, it follows that

$$\frac{1}{n} \iota_{\bar{Q}_{X^n}}(x^n) \geq R_0. \tag{49}$$

Hence, we have

$$\underline{H}_\epsilon(\mathbf{X}) - \gamma = R_0 \overset{(o)}{\leq} \frac{1}{n} \min_{x^n \in \mathcal{X}^n} \iota_{\bar{Q}_{X^n}}(x^n)$$
$$= \frac{1}{n} \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) \min_{x^n \in \mathcal{X}^n} \iota_{\bar{Q}_{X^n}}(x^n)$$
$$\leq \frac{1}{n} \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) \iota_{\bar{Q}_{X^n}}(x^n)$$
$$\overset{(p)}{\leq} \frac{1}{n} \sup_{Q_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)} \mathbb{E}_{P_{X^n}}[\iota_{Q_{X^n}}(X^n)]$$
$$= \frac{1}{n} G^{[\epsilon+\tau]}(X^n),$$

where (o) follows from (49), (p) follows from $\bar{Q}_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$. Since this formula holds for $n \geq n_0$ and arbitrary $\gamma > 0$ and $\tau > 0$, we have $\underline{H}_\epsilon(\mathbf{X}) \leq G^{[\epsilon]}(\mathbf{X})$.

VI. Discussion

In this section, we discuss a duality between the general formula in Theorem 2 and our general formula in Theorem 3 from the viewpoint of the smooth Rényi entropy.

The study [4] clarified the sub-probability distribution $\mathbf{q}^*$, which achieves the infimum of the *smooth Rényi entropy of order* $\alpha \in (0, 1)$ [6]. In view of the condition of the infimum in Theorem 2, this sub-probability distribution $\mathbf{q}^*$ is related to the sub-probability distribution $\frac{P_{X^n}(x^n)}{\mathbb{P}[X^n \in A_n]}$ of $G_{[\epsilon]}(\mathbf{X})$.

On the other hand, the sub-probability distribution $Q_{X^n} \in \mathcal{Q}_{\epsilon+\tau}(\mathcal{X}^n)$ of $G^{[\epsilon]}(\mathbf{X})$ in Theorem 3 is related to $\mathbf{q}^\dagger$ [4], where $\mathbf{q}^\dagger$ is the sub-probability distribution achieving the infimum of the *smooth Rényi entropy of order* $\alpha \in (1, \infty)$ [6].

Therefore, we observe a duality between the general formula of ϵ-variable-length resolvability discussed in [10] and [11] and that of ϵ-variable-length intrinsic randomness discussed in this paper.

VII. Second-Order Variable-Length Intrinsic Randomness

We define the *problem of* (ϵ, R)-*variable-length intrinsic randomness*.

Definition 10: Given $\epsilon \in [0, 1)$ and $R \geq 0$, a second-order rate L is said to be $i(\epsilon, R)$-*achievable* if there exists a mapping $\varphi_n : \mathcal{X}^n \to \mathcal{U}^*$ satisfying

$$\limsup_{n \to \infty} \bar{d}(P_{\varphi_n(X^n)}, P_{U^{(L_n)}}) \leq \epsilon,$$

$$\liminf_{n \to \infty} \frac{1}{\sqrt{n}}(\mathbb{E}_{P_{X^n}}[l(\varphi_n(X^n))] - nR) \geq L.$$

The (ϵ, R)-*variable-length intrinsic randomness* is defined as follows.

Definition 11:

$$T_{\mathrm{i}}(\epsilon, R|\mathbf{X}) := \sup\{L \mid L \text{ is i}(\epsilon, R)\text{-achievable}\}.$$

We establish the second-order general formula.

Theorem 5: For any general source $\mathbf{X}$,

$$T_{\mathrm{i}}(\epsilon, R|\mathbf{X}) = \lim_{\tau \downarrow 0} \liminf_{n \to \infty} \frac{1}{\sqrt{n}}(G^{[\epsilon+\tau]}(X^n) - nR)$$
$$(\epsilon \in [0, 1), R \geq 0).$$

Proof: The theorem can be proven analogously to Theorem 3 with due modifications. ∎

VIII. CONCLUSION

We have investigated the problem of ϵ-variable-length intrinsic randomness. The contribution of this paper is to derive the general formula when we allow positive value of the average variational distance and the lower bound of the value characterizing ϵ-variable-length intrinsic randomness. Further, by comparing the previous result by Yagi and Han [10], [11] and our result, we have clarified the dual relationship between the ϵ-variable-length resolvability and the ϵ-variable-length intrinsic randomness.

APPENDIX A
PROOF OF LEMMA 1

We will use some notations for this proof.

- $U_n := U^{(\lfloor naR \rfloor)}$
- $M_n := K^{\lfloor naR \rfloor}$
- $u_i := u_i^{(\lfloor naR \rfloor)} \quad (i = 1, 2, \ldots, M_n)$

We construct sets $A(i) \subset A_n$ $(i = 1, 2, \ldots, M_n)$. First, for $u_1 \in \mathcal{U}^{\lfloor naR \rfloor}$, construct a subset $A(1) \subset A_n$ so as to satisfy the following conditions

$$\sum_{x^n \in A(1)} P_{X^n}(x^n) \leq c \cdot P_{U_n}(u_1) = \frac{c}{M_n}$$

and, for any $\hat{x}^n \in A_n \setminus A(1)$,

$$c \cdot P_{U_n}(u_1) < \sum_{x^n \in A(1)} P_{X^n}(x^n) + P_{X^n}(\hat{x}^n).$$

Next, for $u_2 \in \mathcal{U}^{\lfloor naR \rfloor}$, construct a subset $A(2) \subset A_n \setminus A(1)$ so as to satisfy the following conditions

$$\sum_{x^n \in A(2)} P_{X^n}(x^n) \leq c \cdot P_{U_n}(u_2) = \frac{c}{M_n}$$

and, for any $\hat{x}^n \in A_n \setminus A(1) \cup A(2)$,

$$c \cdot P_{U_n}(u_2) < \sum_{x^n \in A(2)} P_{X^n}(x^n) + P_{X^n}(\hat{x}^n).$$

In an analogous manner, also for $u_3 \in \mathcal{U}^{\lfloor naR \rfloor}$, construct a subset $A(3) \subset A_n \setminus A(1) \cup A(2)$, and so on. Then, i_0 is defined as the number of final step of this precedure. For i_0, we consider two cases.

1) case of $i_0 = M_n - 1$:

The set $A(i) \subset A_n \setminus \bigcup_{j=1}^{i-1} A(j)$ $(i = 1, 2, \ldots, M_n - 1)$ satisfies the following conditions:

$$\sum_{x^n \in A(i)} P_{X^n}(x^n) \leq c \cdot P_{U_n}(u_i) = \frac{c}{M_n} \tag{50}$$

and, for any $\hat{x}^n \in A_n \setminus \bigcup_{j=1}^{i} A(j)$,

$$c \cdot P_{U_n}(u_i) < \sum_{x^n \in A(i)} P_{X^n}(x^n) + P_{X^n}(\hat{x}^n). \tag{51}$$

On the other hand, the set $A(M_n)$ is defined as follows:

$$A(M_n) := A_n \setminus \bigcup_{i=1}^{M_n - 1} A(i).$$

From (9) and (51), for any $i = 1, 2, \ldots, M_n - 1$ and $\hat{x}^n \in A_n \setminus \bigcup_{j=1}^{i} A(j)$,

$$c \cdot P_{U_n}(u_i) < \sum_{x^n \in A(i)} P_{X^n}(x^n) + cK^{-n(a+\gamma)R}$$
$$= \mathbb{P}[X^n \in A(i)] + cK^{-n(a+\gamma)R}.$$

Hence, it holds that

$$\mathbb{P}[X^n \in A(i)] > \frac{c}{M_n} - cK^{-n(a+\gamma)R}. \tag{52}$$

We define the mapping $\varphi_n : A_n \to \mathcal{U}^{\lfloor naR \rfloor}$ by

$$\varphi_n(x^n) = u_i \quad (x^n \in A(i) \ (i = 1, 2, \ldots, M_n)). \tag{53}$$

Then, it follows that

$$\frac{1}{2} \sum_{u \in \mathcal{U}^{\lfloor naR \rfloor}} \left| P_{\varphi_n(X^n)}(u) - \frac{c}{K^{\lfloor naR \rfloor}} \right|$$
$$\leq \frac{1}{2} \sum_{i=1}^{M_n} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)|$$
$$= \frac{1}{2} \sum_{i=1}^{M_n - 1} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)|$$
$$+ \frac{1}{2}|P_{\varphi_n(X^n)}(u_{M_n}) - c \cdot P_{U_n}(u_{M_n})|. \tag{54}$$

We evaluate the second term on the right-hand side of (54).

$$|P_{\varphi_n(X^n)}(u_{M_n}) - c \cdot P_{U_n}(u_{M_n})|$$
$$= \left| \mathbb{P}[X^n \in A_n] - \sum_{i=1}^{M_n - 1} P_{\varphi_n(X^n)}(u_i) - c\left(1 - \sum_{i=1}^{M_n - 1} P_{U_n}(u_i)\right) \right|$$
$$\leq (c - \mathbb{P}[X^n \in A_n]) + \left| \sum_{i=1}^{M_n - 1} \left(P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)\right) \right|$$
$$\leq (c - \mathbb{P}[X^n \in A_n]) + \sum_{i=1}^{M_n - 1} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)|. \tag{55}$$

The combination of (54) and (55) yields

$$\frac{1}{2} \sum_{u \in \mathcal{U}^{\lfloor naR \rfloor}} \left| P_{\varphi_n(X^n)}(u) - \frac{c}{K^{\lfloor naR \rfloor}} \right|$$

$$\leq \sum_{i=1}^{M_n-1} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)| + \frac{1}{2}(c - \mathbb{P}[X^n \in A_n]). \tag{56}$$

Further, we evaluate the first term on the right-hand side of (56).

$$\sum_{i=1}^{M_n-1} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)|$$

$$\overset{(q)}{=} \sum_{i=1}^{M_n-1} |\mathbb{P}[X^n \in A(i)] - c \cdot P_{U_n}(u_i)|$$

$$\overset{(r)}{=} \sum_{i=1}^{M_n-1} \left(\frac{c}{M_n} - \mathbb{P}[X^n \in A(i)] \right) \overset{(s)}{<} \sum_{i=1}^{M_n-1} cK^{-n(a+\gamma)R}$$

$$\leq M_n cK^{-n(a+\gamma)R} = K^{\lfloor naR \rfloor} cK^{-n(a+\gamma)R} \leq cK^{-n\gamma R}, \tag{57}$$

where (q) follows from (53), (r) follows from (50), (s) follows from (52). By substituting (57) for (56), we have

$$\frac{1}{2} \sum_{u \in \mathcal{U}^{\lfloor naR \rfloor}} \left| P_{\varphi_n(X^n)}(u) - \frac{c}{K^{\lfloor naR \rfloor}} \right|$$

$$\leq cK^{-n\gamma R} + \frac{1}{2}(c - \mathbb{P}[X^n \in A_n]). \tag{58}$$

2) case of $i_0 < M_n - 1$:

The set $A(i) \subset A_n \setminus \bigcup_{j=1}^{i-1} A(j)$ $(i = 1, 2, \ldots, i_0)$ satisfies the following conditions:

$$\sum_{x^n \in A(i)} P_{X^n}(x^n) \leq c \cdot P_{U_n}(u_i) = \frac{c}{M_n}$$

and, for any $\hat{x}^n \in A_n \setminus \bigcup_{j=1}^{i} A(j)$,

$$c \cdot P_{U_n}(u_i) < \sum_{x^n \in A(i)} P_{X^n}(x^n) + P_{X^n}(\hat{x}^n).$$

On the other hand, the set $A(i_0 + 1)$ is defined as follows:

$$A(i_0 + 1) := A_n \setminus \bigcup_{i=1}^{i_0} A(i).$$

Moreover, define the set $A(i)$ $(i = i_0 + 2, \ldots, M_n)$ by

$$A(i) := \phi. \tag{59}$$

Then, for any $i = 1, 2, \ldots, M_n$, it holds that

$$\mathbb{P}[X^n \in A(i)] \leq c \cdot P_{U_n}(u_i) = \frac{c}{M_n}. \tag{60}$$

We define the mapping $\varphi_n : A_n \to \mathcal{U}^{\lfloor naR \rfloor}$ by

$$\varphi_n(x^n) = u_i \quad (x^n \in A(i) \ (i = 1, 2, \ldots, M_n)). \tag{61}$$

Then, it follows that

$$\frac{1}{2} \sum_{u \in \mathcal{U}^{\lfloor naR \rfloor}} \left| P_{\varphi_n(X^n)}(u) - \frac{c}{K^{\lfloor naR \rfloor}} \right|$$

$$\leq \frac{1}{2} \sum_{i=1}^{M_n-1} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)|$$

$$+ \frac{1}{2} |P_{\varphi_n(X^n)}(u_{M_n}) - c \cdot P_{U_n}(u_{M_n})|. \tag{62}$$

We evaluate the first term on the right-hand side of (62).

$$\frac{1}{2} \sum_{i=1}^{M_n-1} |P_{\varphi_n(X^n)}(u_i) - c \cdot P_{U_n}(u_i)|$$

$$\overset{(t)}{=} \frac{1}{2} \sum_{i=1}^{M_n-1} |\mathbb{P}[X^n \in A(i)] - c \cdot P_{U_n}(u_i)|$$

$$\overset{(u)}{=} \frac{1}{2} \sum_{i=1}^{M_n-1} (c \cdot P_{U_n}(u_i) - \mathbb{P}[X^n \in A(i)])$$

$$\overset{(v)}{=} \frac{c(M_n - 1)}{2M_n} - \frac{1}{2}\mathbb{P}[X^n \in A_n], \tag{63}$$

where (t) follows from (61), (u) follows from (60), (v) follows from the construction of $A(i)$. Next, we evaluate the second term on the right-hand side of (62). From the analogous calculation of the first term on the right-hand side of (62), it holds that

$$\frac{1}{2}|P_{\varphi_n(X^n)}(u_{M_n}) - c \cdot P_{U_n}(u_{M_n})|$$

$$= \frac{1}{2}|\mathbb{P}[X^n \in A(M_n)] - c \cdot P_{U_n}(u_{M_n})|$$

$$= \frac{1}{2}(c \cdot P_{U_n}(u_{M_n}) - \mathbb{P}[X^n \in A(M_n)]) \overset{(w)}{=} \frac{c}{2M_n}, \tag{64}$$

where (w) follows from $i_0 + 1 < M_n$ and (59). By substituting (63) and (64) for (62), we have

$$\frac{1}{2} \sum_{u \in \mathcal{U}^{\lfloor naR \rfloor}} \left| P_{\varphi_n(X^n)}(u) - \frac{c}{K^{\lfloor naR \rfloor}} \right|$$

$$\leq \frac{c(M_n - 1)}{2M_n} - \frac{1}{2}\mathbb{P}[X^n \in A_n] + \frac{c}{2M_n} = \frac{1}{2}(c - \mathbb{P}[X^n \in A_n]). \tag{65}$$

By (58) and (65), we can prove (10).

ACKNOWLEDGMENT

REFERENCES

[1] T. S. Han, *Information-Spectrum Methods in Information Theory*, Springer, 2003.

[2] T. S. Han, "Theorems on the variable-length intrinsic randomness," *IEEE Trans. Inf. Theory*, vol.46, no.6, pp.2108-2116, Sept. 2000.

[3] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol.39, no.3, pp.752-771, May. 1993.

[4] H. Koga, "Characterization of the smooth Rényi entropy using majorization," *ITW*, Sevelle, Spain, pp.1-5, Sept. 2013.

[5] H. Koga and H. Yamamoto, "Asymptotic properties on codeword length of an optimal FV code for general sources," *IEEE Trans. Inf. Theory*, vol.51, no.4 pp.1546-1555, Apr. 2005.

[6] R. Renner and S. Wolf, "Smooth Rényi entropy and applications," *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, USA, page 232, June-July, 2004.

[7] T. Uyematsu, "Relating source coding and resolvabillity: a direct approach," *Proc. IEEE Int. Symp. Inf. Theory*, Austin, Texas, pp.1350-1354, June 2010.

[8] T. Uyematsu and S. Kunimatsu, "A new unified method for intrinsic randomness problems of general sources," *ITW*, Sevelle, Spain, pp.9-13, Sept. 2013.

[9] S. Vembu and S. Verdú, "Generating random bits from an arbitrary source: Fundamental limits," *IEEE Trans. Inf. Theory*, vol.41, no.5, pp.1322-1332, Sept. 1995.

[10] H. Yagi and T. S. Han, "Variable-length resolvability for general sources," *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, pp.1748-1752, June 2017.

[11] H. Yagi and T. S. Han, "Variable-length resolvability for general sources and channels," [Online]. Available: https://arxiv.org/abs/1701.08712