entropy

MDPI

Article

Exponential Strong Converse for One Helper Source Coding Problem †

Yasutada Oohama

Department of Communication Engineering and Informatics, University of Electro-Communications, Tokyo 182-8585, Japan; oohama@uec.ac.jp; Tel.: +81-42-443-5358

† This paper is an extended version of our paper published in 2015 IEEE International Symposium on Information Theory (ISIT) .

Received: 12 March 2019; Accepted: 31 May 2019; Published: 5 June 2019

Abstract: We consider the one helper source coding problem posed and investigated by Ahlswede, Körner and Wyner. Two correlated sources are separately encoded and are sent to a destination where the decoder wishes to decode one of the two sources with an arbitrary small error probability of decoding. In this system, the error probability of decoding goes to one as the source block length n goes to infinity. This implies that we have a strong converse theorem for the one helper source coding problem. In this paper, we provide the much stronger version of this strong converse theorem for the one helper source coding problem. We prove that the error probability of decoding tends to one exponentially and derive an explicit lower bound of this exponent function.

Keywords: one helper source coding problem; strong converse theorem; exponent of correct probability of decoding

1. Introduction

For single or multi terminal source encoding systems, the converse coding theorems state that, at any data compression rates below the fundamental theoretical limit of the system, the error probability of decoding *can not go to zero* when the block length n of the codes tends to infinity.

In this paper, we study the one helper source coding problem posed and investigated by Ahlswede, Körner [1] and Wyner [2]. We call the above source coding system (the AKW system). The AKW system is shown in Figure 1.

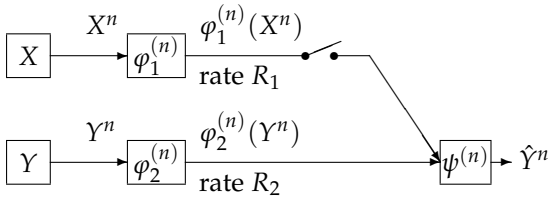

Figure 1. Source encoding with or without side information at the decoder.

In this figure, the AKW system corresponds to the case where the switch is closed. In Figure 1, the sequence (X^n, Y^n) represents independent copies of a pair of dependent random variables (X, Y) which take values in the finite sets $\mathcal{X}, \mathcal{Y}$, respectively. We assume that (X, Y) has a probability distribution denoted by p_{XY}. For each $i = 1, 2$, the encoder $\varphi_i^{(n)}$ outputs a binary sequence which appears at a rate R_i bits per input symbol. The decoder function $\psi^{(n)}$ observes $\varphi_1^{(n)}(X^n)$ and $\varphi_2^{(n)}(Y^n)$ to output a sequence

$\hat{Y}^n := \psi^{(n)}(\varphi_1^{(n)}(X^n), \varphi_2^{(n)}(Y^n))$, which is an estimation of Y^n. When the switch is open, it is well known that the minimum transmission rate R_2 such that the error probability $P_e^{(n)} := \Pr\{Y^n \neq \hat{Y}^n\}$ of decoding tends to zero as n tends to infinity is given by $H(Y)$. Csiszár and Longo [3] proved that, if $R_2 < H(Y)$, then the correct probability $P_c^{(n)} := \Pr\{Y^n = \hat{Y}^n\}$ of decoding decay exponentially and derived the optimal exponent function. When the switch is open and $R_1 > H(X)$, Slepian and Wolf [4] proved that $H(Y|X)$ is the minimum transmission rate R_2 such that the error probability $\Pr\{Y^n \neq \hat{Y}^n\}$ of decoding tends to zero as n tends to infinity. Oohama and Han [5] proved that, if $R_2 < H(Y|X)$, then the correct probability $P_c^{(n)} := \Pr\{Y^n = \hat{Y}^n\}$ of decoding decay exponentially and derived the optimal exponent function.

In this paper, we consider the strong converse theorem in the case where the switch is closed and $0 < R_1 < H(X)$. Let $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$ be the rate region of the AKW system. This region consists of the rate pair (R_1, R_2) such that the error provability of decoding goes to zero as n tends to infinity. The rate region was determined by Ahlswede, Körner [1] and Wyner [2]. On the converse coding theorem, Ahlswede et al. [6] proved that, if (R_1, R_2) is outside the rate region, then, $\mathrm{P}_c^{(n)}$ must tends to zero as n tends to infinity. Gu and Effors [7] examined a speed of convergence for $\mathrm{P}_c^{(n)}$ to tend to zero as $n \to \infty$ by carefully checking the proof of Ahlswede et al. [6]. However, they could not obtain a result on an explicit form of the exponent function with respect to the code length n.

Our main results on the strong converse theorem for the AKW system are as follows. For the AKW system, we prove that, if (R_1, R_2) is outside the rate region $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$, $\mathrm{P}_c^{(n)}$ must go to zero exponentially and derive an explicit lower bound of this exponent. This result corresponds to Theorem 3. As a corollary from this theorem, we obtain the strong converse result, which is stated in Corollary 2. This result states that we have an outer bound with $O(1/\sqrt{n})$ gap from the rate region $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$.

To derive our result, we use a new method called the recursive method. This method, which is a new method introduced by the author, includes a certain recursive algorithm for a single letterization of exponent functions. In a standard argument of proving converse coding theorems, single letterization methods based on the chain rule of the entropy functions are used. In general, the functions representing multi letter characterizations of exponent functions do not have the chain rule property. In such cases, the recursive method is quite useful for deriving single letterized bounds. The recursive method is a general powerful tool to prove strong converse theorems for several coding problems in information theory. In fact, the recursive method plays important roles in deriving exponential strong converse exponent for communication systems treated in [8–12].

On the strong converse theorem for the one helper source coding problem, we have two recent other works [13,14]. The above two works proved the strong converse theorem using different methods from our method. In [13], Watanabe found a relationship between the AKW system and the Gray–Wyner network. Using this relationship and the second order rate region for the Gray–Wyner network obtained by him [15], Watanabe established the strong converse theorem for the AKW system. In [14], Liu et al. introduced a new method to derive sharp strong converse bounds via a reverse hypercontractivity. Using this method, they obtained an outer bound of the rate region for the AKW system with $O(1/\sqrt{n})$ gap from the rate region. Furthermore, in [14], an extension of the AKW system to the case of Gaussian source and quadratic distortion is investigated, obtaining an outer bound with $O(1/\sqrt{n})$ gap from the rate distortion region for the extended source coding system. In his resent paper [16], Liu showed a lower bound (converse) on the dispersion of AWK as the variance of the linear combination of information densities.

The strong converse theorems seem to be regarded just as a mathematical problem and have been investigated mainly from theoretical interest. Recently, Watanabe and Oohama [17] have found an interesting security problem, which has a close connection with the strong converse theorem for the AKW system. Furthermore, Oohama and Santoso [18] and Santoso and Oohama [19] clarify that the exponential strong converse theorem obtained by this paper plays an essential role in deriving a strong

sufficient secure condition for the privacy amplification in their new theoritical model of side channel attacks to the Shannon chipher systems. From the above two cases, we expect that exponential strong converse theorems for multiterminal source networks will serve as a strong tool to several information theoretical security problems.

2. Problem Formulation

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets and $\{(X_t, Y_t)\}_{t=1}^{\infty}$ be a stationary discrete memoryless source. For each $t = 1, 2, \cdots$, the random pair (X_t, Y_t) takes values in $\mathcal{X} \times \mathcal{Y}$, and has a probability distribution

$$p_{XY} = \{p_{XY}(x, y)\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}}.$$

We write n independent copies of $\{X_t\}_{t=1}^{\infty}$ and $\{Y_t\}_{t=1}^{\infty}$, respectively as

$$X^n = X_1, X_2, \cdots, X_n \text{ and } Y^n = Y_1, Y_2, \cdots, Y_n.$$

We consider a communication system depicted in Figure 2. This communication system corresponds to the case where the switch is closed in Figure 1. Data sequences X^n and Y^n are separately encoded to $\varphi_1^{(n)}(X^n)$ and $\varphi_2^{(n)}(Y^n)$ and those are sent to the information processing center. At the center, the decoder function $\psi^{(n)}$ observes $(\varphi_1^{(n)}(X^n), \varphi_2^{(n)}(Y^n))$ to output the estimation $\hat{Y}^n$ of Y^n. The encoder functions $\varphi_1^{(n)}$ and $\varphi_2^{(n)}$ are defined by

$$\left.\begin{aligned} \varphi_1^{(n)} &: \mathcal{X}^n \to \mathcal{M}_1 = \{1, 2, \cdots, M_1\} \\ \varphi_2^{(n)} &: \mathcal{Y}^n \to \mathcal{M}_2 = \{1, 2, \cdots, M_2\} \end{aligned}\right\}, \tag{1}$$

where for each $i = 1, 2$, $\|\varphi_i^{(n)}\|$ $(= M_i)$ stands for the range of cardinality of $\varphi_i^{(n)}$. The decoder function $\psi^{(n)}$ is defined by

$$\psi^{(n)} : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{Y}^n. \tag{2}$$

The error probability of decoding is

$$\mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) = \Pr\left\{\hat{Y}^n \neq Y^n\right\}, \tag{3}$$

where $\hat{Y}^n = \psi^{(n)}(\varphi_1^{(n)}(X^n), \varphi_2^{(n)}(Y^n))$. A rate pair (R_1, R_2) is ε-*achievable* if, for any $\delta > 0$, there exists a positive integer $n_0 = n_0(\varepsilon, \delta)$ and a sequence of triples $\{(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})\}_{n\geq n_0}$ such that, for $n \geq n_0$,

$$\frac{1}{n}\log\|\varphi_i^{(n)}\| \leq R_i + \delta \text{ for } i = 1, 2, \mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \leq \varepsilon.$$

For $\varepsilon \in (0, 1)$, the rate region $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ is defined by

$$\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}) := \{(R_1, R_2) : (R_1, R_2) \text{ is } \varepsilon\text{-achievable for } p_{XY}\}.$$

Furthermore, define

$$\mathcal{R}_{\mathrm{AKW}}(p_{XY}) := \bigcap_{\varepsilon\in(0,1)} \mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}).$$

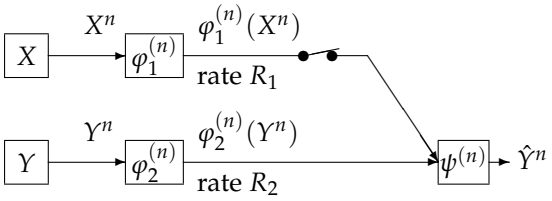

Figure 2. One helper source coding system [21].

We can show that the two rate regions $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|\ p_{XY})$, $\varepsilon \in (0,1)$ and $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$ satisfy the following property.

Property 1.

(a) The regions $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$, $\varepsilon \in (0,1)$, and $\mathcal{R}_{\mathrm{AKW}}(\ p_{XY})$ are closed convex sets of $\mathbb{R}_+^2$, where

$$\mathbb{R}_+^2 := \{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0\}.$$

(b) $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ has another form using (n,ε)-rate region $\mathcal{R}_{\mathrm{AKW}}(n,\varepsilon|p_{XY})$, the definition of which is as follows. We set

$$\begin{aligned}\mathcal{R}_{\mathrm{AKW}}(n,\varepsilon|p_{XY}) = \{(R_1,R_2) :\ &\textit{There exists } (\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \textit{ such that}\\ &\frac{1}{n}\log||\varphi_i^{(n)}|| \leq R_i, i = 1,2, \mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \leq \varepsilon\}.\end{aligned}$$

Using $\mathcal{R}_{\mathrm{AKW}}(n,\varepsilon|p_{XY})$, $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ can be expressed as

$$\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}) = \mathrm{cl}\left(\bigcup_{m\geq 1}\bigcap_{n\geq m}\mathcal{R}_{\mathrm{AKW}}(n,\varepsilon|p_{XY})\right).$$

Proof of this property is given in Appendix A. It is well known that $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$ was determined by Ahlswede, Körner and Wyner. To describe their result, we introduce an auxiliary random variable U taking values in a finite set $\mathcal{U}$. We assume that the joint distribution of (U,X,Y) is

$$p_{UXY}(u,x,y) = p_U(u)p_{X|U}(x|u)p_{Y|X}(y|x).$$

The above condition is equivalent to $U \leftrightarrow X \leftrightarrow Y$. Define the set of probability distribution $p = p_{UXY}$ by

$$\mathcal{P}(p_{XY}) := \{p_{UXY} : |\mathcal{U}| \leq |\mathcal{X}| + 1, U \leftrightarrow X \leftrightarrow Y\}.$$

Set

$$\begin{aligned}\mathcal{R}(p) &:= \{(R_1,R_2) : R_1, R_2 \geq 0\ R_1 \geq I_p(X;U), R_2 \geq H_p(Y|U)\},\\ \mathcal{R}(p_{XY}) &:= \bigcup_{p\in\mathcal{P}(p_{XY})}\mathcal{R}(p).\end{aligned}$$

We can show that the region $\mathcal{R}(p_{XY})$ satisfies the following property.

Property 2.

(a) The region $\mathcal{R}(p_{XY})$ is a closed convex subset of $\mathbb{R}_+^2$.

(b) For any p_{XY}, we have

$$\min_{(R_1,R_2)\in\mathcal{R}(p_{XY})} (R_1+R_2) = H_p(Y). \tag{4}$$

The minimum is attained by $(R_1,R_2)=(0,H_p(Y))$. This result implies that

$$\mathcal{R}(p_{XY}) \subseteq \{(R_1,R_2) : R_1+R_2 \geq H_p(Y)\} \cap \mathbb{R}_+^2.$$

Furthermore, the point $(0,H_p(Y))$ always belongs to $\mathcal{R}(p_{XY})$.

Property 2 part a is a well known property. Proof of Property 2 part b is easy. Proofs of Property 2 parts a and b are omitted. A typical shape of the rate region $\mathcal{R}(p_{XY})$ is shown in Figure 3.

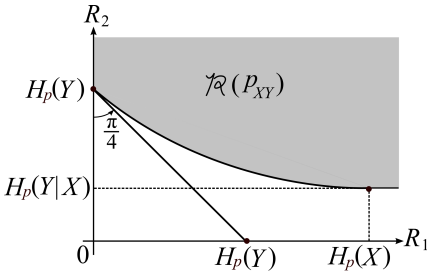

Figure 3. A typical shape of $\mathcal{R}(p_{XY})$.

The rate region $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$ was determined by Ahlswede and Körner [1] and Wyner [2]. Their results are the following.

Theorem 1 (Ahlswede, Körner [1] and Wyner [2])**.**

$$\mathcal{R}_{\mathrm{AKW}}(p_{XY}) = \mathcal{R}(p_{XY}).$$

On the converse coding theorem, Ahlswede et al. [6] obtained the following.

Theorem 2 (Ahlswede et al. [6])**.** *For each fixed $\varepsilon\in(0,1)$, we have*

$$\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}) = \mathcal{R}(p_{XY}).$$

Gu and Effors [7] examined a speed of convergence for $\mathrm{P}_{\mathrm{e}}^{(n)}$ to tend to 1 as $n\to\infty$ by carefully checking the proof of Ahlswede et al. [6]. However, they could not obtain a result on an explicit form of the exponent function with respect to the code length n.

Our aim is to find an explicit form of the exponent function for the error probability of decoding to tend to one as $n\to\infty$ when $(R_1,R_2)\notin\mathcal{R}_{\mathrm{AKW}}(p_{XY})$. To examine this quantity, we define the following quantity. Set

$$\begin{aligned}
&\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}) := 1-\mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}),\\
&G^{(n)}(R_1,R_2|p_{XY}) := \min_{\substack{(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}):\\ (1/n)\log\|\varphi_i^{(n)}\|\leq R_i, i=1,2}} \left(-\frac{1}{n}\right)\log \mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}).\\
&G(R_1,R_2|p_{XY}) := \lim_{n\to\infty} G^{(n)}(R_1,R_2|p_{XY}),\\
&\mathcal{G}(p_{XY}) := \{(R_1,R_2,G) : G\geq G(R_1,R_2|p_{XY})\}.
\end{aligned}$$

By time sharing, we have that

$$G^{(n+m)}\left(\frac{nR_1+mR_1'}{n+m},\frac{nR_2+mR_2'}{n+m}\middle|p_{XY}\right)\le\frac{nG^{(n)}(R_1,R_2|p_{XY})+mG^{(m)}(R_1',R_2'|p_{XY})}{n+m}. \tag{5}$$

Choosing $R=R'$ in the inequality (5), we obtain the following subadditivity property on $\{G^{(n)}(R_1,R_2|p_{XY})\}_{n\ge1}$:

$$G^{(n+m)}(R_1,R_2|p_{XY})\le\frac{nG^{(n)}(R_1,R_2|p_{XY})+mG^{(m)}(R_1,R_2|p_{XY})}{n+m},$$

from which this, and Fekete's subadditive lemma, we have that $G^{(n)}(R_1,R_2|p_{XY})$ exists and satisfies the following:

$$\lim_{n\to\infty}G^{(n)}(R_1,R_2|p_{XY})=\inf_{n\ge1}G^{(n)}(R_1,R_2|p_{XY}).$$

The exponent function $G(R_1,R_2|p_{XY})$ is a convex function of (R_1,R_2). In fact, from the inequality (5), we have that for any $\alpha\in[0,1]$

$$G(\alpha R_1+\bar{\alpha}R_1',\alpha R_2+\bar{\alpha}R_2'|p_{XY})\le\alpha G(R_1,R_2|p_{XY})+\bar{\alpha}G(R_1',R_2'|p_{XY}).$$

The region $\mathcal{G}(p_{XY})$ is also a closed convex set. Our main aim is to find an explicit characterization of $\mathcal{G}(p_{XY})$. In this paper, we derive an explicit outer bound of $\mathcal{G}(p_{XY})$ whose section by the plane $G=0$ coincides with $\mathcal{R}_{\mathrm{AKW}}(p_{XY})$.

3. Main Results

In this section, we state our main result. We first explain that the region $\mathcal{R}(p_{XY})$ can be expressed with a family of supporting hyperplanes. To describe this result, we define a set of probability distributions on $\mathcal{U}\times\mathcal{X}\times\mathcal{Y}$ by

$$\mathcal{P}_{\mathrm{sh}}(p_{XY}):=\{p=p_{UXY}:|\mathcal{U}|\le|\mathcal{X}|,U\leftrightarrow X\leftrightarrow Y\}.$$

For $\mu\ge0$, define

$$R^{(\mu)}(p_{XY}):=\min_{p\in\mathcal{P}_{\mathrm{sh}}(p_{XY})}\left\{\mu I_p(X;U)+\bar{\mu}H_p(Y|U)\right\}.$$

Furthermore, define

$$\mathcal{R}_{\mathrm{sh}}(p_{XY}):=\bigcap_{\mu\in[0,1]}\{(R_1,R_2):\mu R_1+\bar{\mu}R_2\ge R^{(\mu)}(p_{XY})\}.$$

Then, we have the following property.

Property 3.

(a) The bound $|\mathcal{U}|\le|\mathcal{X}|$ is sufficient to describe $R^{(\mu)}(p_{XY})$.
(b) For every $\mu\in[0,1]$, we have

$$\min_{(R_1,R_2)\in\mathcal{R}(p_{XY})}\{\mu R_1+\bar{\mu}R_2\}=R^{(\mu)}(p_{XY}). \tag{6}$$

(c) For any p_{XY}, we have

$$\mathcal{R}_{\mathrm{sh}}(p_{XY})=\mathcal{R}(p_{XY}). \tag{7}$$

Property 3 part a is stated as Lemma A1 in Appendix B. Proof of this lemma is given in this appendix. Proofs of Property 3 parts b and c are given in Appendix C. Set

$$\mathcal{Q}(p_{Y|X}) := \{q = q_{UXY} : |\mathcal{U}| \leq |\mathcal{X}|, U \leftrightarrow X \leftrightarrow Y, p_{Y|X} = q_{Y|X}\}.$$

For $(\mu, \alpha) \in [0,1]^2$, and for $q = q_{UXY} \in \mathcal{Q}(p_{Y|X})$, define

$$\begin{aligned}
&\omega_{q|p_X}^{(\mu,\alpha)}(x,y|u) := \bar{\alpha} \log \frac{q_X(x)}{p_X(x)} + \alpha \left[\mu \log \frac{q_{X|U}(x|u)}{p_X(x)} + \bar{\mu} \log \frac{1}{q_{Y|U}(y|u)}\right],\\
&f_{q|p_X}^{(\mu,\alpha)}(x,y|u) := \exp\left\{-\omega_{q|p_X}^{(\mu,\alpha)}(x,y|u)\right\},\\
&\Omega^{(\mu,\alpha)}(q|p_X) := -\log \mathrm{E}_q\left[\exp\left\{-\omega_{q|p_X}^{(\mu,\alpha)}(X,Y|U)\right\}\right], \Omega^{(\mu,\alpha)}(p_{XY}) := \min_{q \in \mathcal{Q}(p_{Y|X})} \Omega^{(\mu,\alpha)}(q|p_X),\\
&F^{(\mu,\alpha)}(\mu R_1 + \bar{\mu} R_2 | p_{XY}) := \frac{\Omega^{(\mu,\alpha)}(p_{XY}) - \alpha(\mu R_1 + \bar{\mu} R_2)}{2 + \alpha\bar{\mu}},\\
&F(R_1, R_2 | p_{XY}) := \sup_{(\mu,\alpha) \in [0,1]^2} F^{(\mu,\alpha)}(\mu R_1 + \bar{\mu} R_2 | p_{XY}).
\end{aligned}$$

We next define a function serving as a lower bound of $F(R_1, R_2|p_{XY})$. For $\lambda \geq 0$ and for $p_{UXY} \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$, define

$$\begin{aligned}
&\tilde{\omega}_p^{(\mu)}(x,y|u) := \mu \log \frac{p_{X|U}(x|u)}{p_X(x)} + \bar{\mu} \log \frac{1}{p_{Y|U}(y|u)},\\
&\tilde{\Omega}^{(\mu,\lambda)}(p) := -\log \mathrm{E}_p\left[\exp\left\{-\lambda \tilde{\omega}_p^{(\mu)}(X,Y|U)\right\}\right], \tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) := \min_{p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})} \tilde{\Omega}^{(\mu,\lambda)}(p).
\end{aligned}$$

Furthermore, set

$$\begin{aligned}
&\underline{F}^{(\mu,\lambda)}(\mu R_1 + \bar{\mu} R_2 | p_{XY}) := \frac{\tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) - \lambda(\mu R_1 + \bar{\mu} R_2)}{2 + \lambda(5 - \mu)},\\
&\underline{F}(R_1, R_2 | p_{XY}) := \sup_{\lambda \geq 0, \mu \in [0,1]} \underline{F}^{(\mu,\lambda)}(\mu R_1 + \bar{\mu} R_2 | p_{XY}).
\end{aligned}$$

We can show that the above functions satisfy the following property.

Property 4.

(a) *The cardinality bound $|\mathcal{U}| \leq |\mathcal{X}|$ in $\mathcal{Q}(p_{Y|X})$ is sufficient to describe the quantity $\Omega^{(\mu,\alpha)}(p_{XY})$. Furthermore, the cardinality bound $|\mathcal{U}| \leq |\mathcal{X}|$ in $\mathcal{P}_{\mathrm{sh}}(p_{XY})$ is sufficient to describe the quantity $\tilde{\Omega}^{(\mu,\lambda)}(p_{XY})$.*

(b) *For any $R_1, R_2 \geq 0$, we have*

$$F(R_1, R_2 | p_{XY}) \geq \underline{F}(R_1, R_2 | p_{XY}).$$

(c) *For any $p = p_{UXY} \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$ and any $(\mu, \lambda) \in [0,1]^2$, we have*

$$0 \leq \tilde{\Omega}^{(\mu,\lambda)}(p) \leq \mu \log |\mathcal{X}| + \bar{\mu} \log |\mathcal{Y}|. \tag{8}$$

(d) *Fix any $p = p_{UXY} \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$ and $\mu \in [0,1]$. For $\lambda \in [0,1]$, we define a probability distribution $p^{(\lambda)} = p_{UXY}^{(\lambda)}$ by*

$$p^{(\lambda)}(u,x,y) := \frac{p(u,x,y)\exp\left\{-\lambda\tilde{\omega}_p^{(\mu)}(x,y|u)\right\}}{\mathrm{E}_p\left[\exp\left\{-\lambda\tilde{\omega}_p^{(\mu)}(X,Y|U)\right\}\right]}.$$

Then, for $\lambda \in [0,1/2]$, $\tilde{\Omega}^{(\mu,\lambda)}(p)$ *is twice differentiable. Furthermore, for* $\lambda \in [0,1/2]$, *we have*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{\Omega}^{(\mu,\lambda)}(p) = \mathrm{E}_{p^{(\lambda)}}\left[\tilde{\omega}_p^{(\mu)}(X,Y|U)\right], \frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\tilde{\Omega}^{(\mu,\lambda)}(p) = -\mathrm{Var}_{p^{(\lambda)}}\left[\tilde{\omega}_p^{(\mu)}(X,Y|U)\right].$$

The second equality implies that $\tilde{\Omega}^{(\mu,\lambda)}(p|p_{XY})$ *is a concave function of* $\lambda \in [0,1/2]$.

(e) For every $(\mu,\lambda) \in [0,1] \times [0,1/2]$, *define*

$$\rho^{(\mu,\lambda)}(p_{XY}) := \max_{\substack{(\nu,p)\in[0,\lambda]\times\mathcal{P}_{\mathrm{sh}}(p_{XY}):\\ \tilde{\Omega}^{(\mu,\lambda)}(p)=\tilde{\Omega}^{(\mu,\lambda)}(p_{XY})}} \mathrm{Var}_{p^{(\nu)}}\left[\tilde{\omega}_p^{(\mu)}(X,Y|U)\right],$$

and set

$$\rho = \rho(p_{XY}) := \max_{(\mu,\lambda)\in[0,1]\times[0,1/2]} \rho^{(\mu,\lambda)}(p_{XY}).$$

Then, we have $\rho(p_{XY}) < \infty$. *Furthermore, for any* $(\mu,\lambda) \in [0,1] \times [0,1/2]$, *we have*

$$\tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) \geq \lambda R^{(\mu)}(p_{XY}) - \frac{\lambda^2}{2}\rho(p_{XY}). \tag{9}$$

(f) For every $\tau \in (0,(1/2)\rho(p_{XY}))$, *the condition* $(R_1+\tau, R_2+\tau) \notin \mathcal{R}(p_{XY})$ *implies*

$$\underline{F}(R_1,R_2|p_{XY}) > \frac{\rho(p_{XY})}{4}\cdot g^2\left(\frac{\tau}{\rho(p_{XY})}\right) > 0,$$

where g *is the inverse function of* $\vartheta(a) := a + (5/4)a^2, a \geq 0$.

Property 3 part a is stated as Lemma A2 in Appendix B. Proof of this lemma is given in this appendix. Proof of Property 4 part b is given in Appendix D. Proofs of Property 4 parts c, d, e, and f are given in Appendix E.

Our main result is the following.

Theorem 3. *For any* $R_1, R_2 \geq 0$, *any* p_{XY}, *and for any* $(\varphi_1^{(n)}, \varphi_1^{(n)}, \psi^{(n)})$ *satisfying* $(1/n)\log||\varphi_i^{(n)}|| \leq R_i, i = 1,2$, *we have*

$$\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \leq 5\exp\left\{-nF(R_1,R_2|p_{XY})\right\}. \tag{10}$$

It can be seen from Property 4 parts b and f that $F(R_1,R_2|p_{XY})$ is strictly positive if (R_1,R_2) is outside the rate region $\mathcal{R}(p_{XY})$. Hence, by Theorem 3, we have that, if (R_1,R_2) is outside the rate region, then the error probability of decoding goes to one exponentially and its exponent is not below $F(R_1,R_2|p_{XY})$. It immediately follows from Theorem 3 that we have the following corollary.

Corollary 1.

$$G(R_1,R_2|p_{XY}) \geq F(R_1,R_2|p_{XY}),$$
$$\mathcal{G}(p_{XY}) \subseteq \overline{\mathcal{G}}(p_{XY}) = \left\{(R_1,R_2,G) : G \geq F(R_1,R_2|p_{XY})\right\}.$$

Proof of Theorem 3 will be given in the next section. The exponent function at rates outside the rate region was derived by Oohama and Han [5] for the separate source coding problem for correlated sources [4]. The techniques used by them is a method of types [20], which is not useful to prove Theorem 3. Some novel techniques based on the information spectrum method introduced by Han [22] are necessary to prove this theorem.

From Theorem 3 and Property 4 part e, we can obtain an explicit outer bound of $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ with an asymptotically vanishing deviation from $\mathcal{R}_{\mathrm{AKW}}(p_{XY}) = \mathcal{R}(p_{XY})$. The strong converse theorem established by Ahlswede et al. [6] immediately follows from this corollary. To describe this outer bound, for $\kappa > 0$, we set

$$\mathcal{R}(p_{XY}) - \kappa(1,1) := \{(R_1 - \kappa, R_2 - \kappa) : (R_1, R_2) \in \mathcal{R}(p_{XY})\},$$

which serves as an outer bound of $\mathcal{R}(p_{XY})$. For each fixed $\varepsilon \in (0,1)$, we define $\kappa_n = \kappa_n(\varepsilon, \rho(p_{XY}))$ by

$$\begin{aligned} \kappa_n &:= \rho(p_{XY})\vartheta\left(\sqrt{\frac{4}{n\rho(p_{XY})}\log\left(\frac{5}{1-\varepsilon}\right)}\right) \\ &\overset{(a)}{=} 2\sqrt{\frac{\rho(p_{XY})}{n}\log\left(\frac{5}{1-\varepsilon}\right)} + \frac{5}{n}\log\left(\frac{5}{1-\varepsilon}\right). \end{aligned} \tag{11}$$

Step (a) follows from $\vartheta(a) = a + (5/4)a^2$. Since $\kappa_n \to 0$ as $n \to \infty$, we have the smallest positive integer $n_0 = n_0(\varepsilon, \rho(p_{XY}))$ such that $\kappa_n \le (1/2)\rho(p_{XY})$ for $n \ge n_0$. From Theorem 3 and Property 4 part e, we have the following corollary.

Corollary 2. *For each fixed $\varepsilon \in (0,1)$, we choose the above positive integer $n_0 = n_0(\varepsilon, \rho(p_{XY}))$. Then, for any $n \ge n_0$, we have*

$$\mathcal{R}_{\mathrm{AKW}}(n, \varepsilon|p_{XY}) \subseteq \mathcal{R}(p_{XY}) - \kappa_n(1,1).$$

The above result together with

$$\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}) = \mathrm{cl}\left(\bigcup_{m\ge 1}\bigcap_{n\ge m}\mathcal{R}_{\mathrm{AKW}}(n, \varepsilon|p_{XY})\right)$$

yields that, for each fixed $\varepsilon \in (0,1)$, we have

$$\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}) = \mathcal{R}_{\mathrm{AKW}}(p_{XY}) = \mathcal{R}(p_{XY}).$$

This recovers the strong converse theorem proved by Ahlswede et al. [6].

Proof of this corollary will be given in the next section.

4. Proof of the Main Result

Let (X^n, Y^n) be a pair of random variables from the information source. We set $S = \varphi_1^{(n)}(X^n)$. Joint distribution $p_{SX^nY^n}$ of (S, X^n, Y^n) is given by

$$p_{SX^nY^n}(s, x^n, y^n) = p_{S|X^n}(s|x^n)\prod_{t=1}^{n} p_{X_tY_t}(x_t, y_t).$$

It is obvious that $S \leftrightarrow X^n \leftrightarrow Y^n$. Then, we have the following lemma, which is well known as a single shot infomation spectrum bound.

Lemma 1. *For any $\eta > 0$ and for any $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ satisfying $(1/n)\log||\varphi_i^{(n)}|| \le R_i, i = 1, 2$, we have*

$$\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \le p_{SX^nY^n}\Bigg\{ 0 \ge \frac{1}{n}\log\frac{\hat{q}_{SX^nY^n}(S, X^n, Y^n)}{p_{SX^nY^n}(S, X^n, Y^n)} - \eta, \tag{12}$$

$$0 \ge \frac{1}{n}\log\frac{Q_{X^n}(X^n)}{p_{X^n}(X^n)} - \eta, \tag{13}$$

$$R_1 \ge \frac{1}{n}\log\frac{\tilde{Q}_{X^n|S}(X^n|S)}{p_{X^n}(X^n)} - \eta, \tag{14}$$

$$R_2 \ge \frac{1}{n}\log\frac{1}{p_{Y^n|S}(Y^n|S)} - \eta\Bigg\} + 4\mathrm{e}^{-n\eta}. \tag{15}$$

The probability distributions appearing in the three inequalities (12), (13), and (14) in the right members of (15) have a property that we can select them as arbitrary. In (12), we can choose any probability distribution $\hat{q}_{SX^nY^n}$ on $\mathcal{S}\times\mathcal{X}^n\times\mathcal{Y}^n$. In (13), we can choose any distribution Q_{X^n} on $\mathcal{X}^n$. In (14), we can choose any stochastic matrix $\tilde{Q}_{X^n|U^n}$: $\mathcal{X}^n \to \mathcal{U}^n$.

This lemma can be proved by a standard argument in the information spectrum method [22]. The detail of the proof is given in Appendix F. Next, we single letterize the four information spectrum quantities inside the first term in the right members of (15) in Lemma 1 to obtain the following lemma.

Lemma 2. *For any $\eta > 0$ and for any $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ satisfying $(1/n)\log||\varphi_i^{(n)}|| \le R_i, i = 1, 2$, we have*

$$\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \le p_{SX^nY^n}\Bigg\{0 \ge \frac{1}{n}\sum_{t=1}^{n}\log\frac{Q_{X_t}(X_t)}{p_{X_t}(X_t)} - \eta, \tag{16}$$

$$R_1 \ge \frac{1}{n}\sum_{t=1}^{n}\log\frac{\tilde{Q}_{X_t|SX^{t-1}}(X_t|S, X^{t-1})}{p_{X_t}(X_t)} - \eta, \tag{17}$$

$$R_2 \ge \frac{1}{n}\sum_{t=1}^{n}\log\frac{1}{p_{Y_t|SX^{t-1}Y^{t-1}}(Y_t|S, X^{t-1}, Y^{t-1})} - 2\eta\Bigg\} + 4\mathrm{e}^{-n\eta},$$

where for each $t = 1, 2, \cdots, n$, the probability distribution Q_{X_t} on $\mathcal{X}$ appearing in (16) and the stochastic matrix $\tilde{Q}_{X_t|SX^{t-1}} : \mathcal{M}_1 \times \mathcal{X}^{t-1} \to \mathcal{X}$ appearing in (17) have a property that we can choose their values arbitrary.

Proof. In (12) in Lemma 1, we choose $\hat{q}_{SX^nY^n}$ having the form

$$\hat{q}_{SX^nY^n}(S, X^n, Y^n) = p_S(S)\prod_{t=1}^{n}\left\{p_{X_t|SX^{t-1}Y^t}(X_t|S, X^{t-1}, Y^t)p_{Y_t|SY^{t-1}}(Y_t|S, Y^{t-1})\right\}.$$

In (13) in Lemma 1, we choose Q_{X^n} having the form

$$Q_{X^n}(X^n) = \prod_{t=1}^{n}Q_{X_t}(X_t).$$

We further note that

$$\frac{\tilde{Q}_{X^n|S}(X^n|S)}{p_{X^n}(X^n)} = \prod_{t=1}^{n}\frac{\tilde{Q}_{X_t|SX^{t-1}}(X_t|S, X^{t-1})}{p_{X_t}(X_t)}, p_{Y^n|S}(Y^n|S) = \prod_{t=1}^{n}p_{Y_t|SY^{t-1}}(Y_t|S, Y^{t-1}).$$

Then, the bound (15) in Lemma 1 becomes

$$
\begin{aligned}
\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \le p_{SX^nY^n}\Bigg\{ & 0 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{p_{Y_t|SY^{t-1}}(Y_t|S, Y^{t-1})}{p_{Y_t|SX^{t-1}Y^{t-1}}(Y_t|S, X^{t-1}, Y^{t-1})} - \eta, \\
& 0 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{Q_{X_t}(X_t)}{p_{X_t}(X_t)} - \eta, \\
& R_1 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{\tilde{Q}_{X_t|SX^{t-1}}(X_t|S, X^{t-1})}{p_{X_t}(X_t)} - \eta, \\
& R_2 \ge \frac{1}{n}\sum_{t=1}^{n} \frac{1}{p_{Y_t|SY^{t-1}}(Y_t|S, Y^{t-1})} - \eta \Bigg\} + 4\mathrm{e}^{-n\eta} \\
\le p_{SX^nY^n}\Bigg\{ & 0 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{Q_{X_t}(X_t)}{p_{X_t}(X_t)} - \eta, \\
& R_1 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{\tilde{Q}_{X_t|SX^{t-1}}(X_t|S, X^{t-1})}{p_{X_t}(X_t)} - \eta, \\
& R_2 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{1}{p_{Y_t|SX^{t-1}Y^{t-1}}(Y_t|S, X^{t-1}, Y^{t-1})} - 2\eta \Bigg\} + 4\mathrm{e}^{-n\eta},
\end{aligned}
$$

completing the proof. □

As in the standard converse coding argument, we identify auxiliary random variables, based on the bound in Lemma 2. The following lemma is necessary for such identification.

Lemma 3. *Suppose that, for each $t = 1, 2, \cdots, n$, the joint distribution $p_{SX^tY^t}$ of the random vector SX^tY^t is a marginal distribution of $p_{SX^nY^n}$. Then, we have the following Markov chain:*

$$SX^{t-1} \leftrightarrow X_t \leftrightarrow Y_t \tag{18}$$

or equivalently that $I(Y_t; SX^{t-1}|X_t) = 0$. Furthermore, we have the following Markov chain:

$$Y^{t-1} \leftrightarrow SX^{t-1} \leftrightarrow (X_t, Y_t) \tag{19}$$

or equivalently that $I(X_tY_t; Y^{t-1}|SX^{t-1}) = 0$. The above two Markov chains are equivalent to the following one long Markov chain:

$$Y^{t-1} \leftrightarrow SX^{t-1} \leftrightarrow X_t \leftrightarrow Y_t. \tag{20}$$

Proof of this lemma is given in Appendix G. For $t = 1, 2, \cdots, n$, set $\mathcal{U}_t := \mathcal{M}_1 \times \mathcal{X}^{t-1}$. Define a random variable $U_t \in \mathcal{U}_t$ by $U_t := (S, X^{t-1})$. From Lemmas 2 and 3, we identify auxiliary random variables to obtain the following lemma.

Lemma 4. *For any $\eta > 0$ and for any $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ satisfying $(1/n)\log||\varphi_i^{(n)}|| \le R_i, i = 1, 2$, we have*

$$\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \le p_{SX^nY^n}\left\{0 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{Q_{X_t}(X_t)}{p_{X_t}(X_t)} - \eta, \right. \tag{21}$$

$$R_1 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{\tilde{Q}_{X_t|U_t}(X_t|U_t)}{p_{X_t}(X_t)} - \eta, \tag{22}$$

$$\left. R_2 \ge \frac{1}{n}\sum_{t=1}^{n} \log \frac{1}{p_{Y_t|U_t}(Y_t|U_t)} - 2\eta \right\} + 4\mathrm{e}^{-n\eta}, \tag{23}$$

where, for each $t = 1, 2, \cdots, n$, the probability distribution Q_{X_t} on $\mathcal{X}$ appearing in (21) and the stochastic matrix $\tilde{Q}_{X_t|U_t} : \mathcal{U}_t \to \mathcal{X}$ appearing in (22) have a property that we can choose their values arbitrary.

Now, the challenge is that, although the quantities inside the first term in the right members of (23) in Lemma 4 have n sum of information spectrum quantities, the measure $p_{SX^nY^n}$ does not have an i.i.d. structure in general. To resolve this, we first use the large deviation theory to upper bound the first quantity in the right members of (23). For each $t = 1, 2, \cdots, n$, set $\underline{Q}_t := (Q_{X_t}, \tilde{Q}_{X_t|U_t})$. Let $\underline{\mathcal{Q}}_t$ be a set of all $\underline{Q}_t$. We define a quantity which serves as an exponential upper bound of $\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$. Let $\mathcal{P}^{(n)}(p_{XY})$ be a set of all probability distributions $p_{SX^nY^n}$ on $\mathcal{M}_1 \times \mathcal{X}^n \times \mathcal{Y}^n$ having a form:

$$p_{SX^nY^n}(s, x^n, y^n) = p_{S|X^n}(s|x^n) \prod_{t=1}^{n} p_{XY}(x_t, y_t)$$
$$\text{for } (s, x^n, y^n) \in \mathcal{M}_1 \times \mathcal{X}^n \times \mathcal{Y}^n.$$

For simplicity of notation, we use the notation $p^{(n)}$ for $p_{SX^nY^n} \in \mathcal{P}^{(n)}(p_{XY})$. For each $t = 1, 2, \cdots, n$, $p_{U_tX_tY_t} = p_{SX^tY_t}$ is a marginal distribution of $p^{(n)}$. For $t = 1, 2, \cdots, n$, we simply write $p_t = p_{U_tX_tY_t}$. For $\mu \in [0, 1]$, $\alpha \in [0, 1)$, $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$, and $\underline{Q}^n \in \underline{\mathcal{Q}}^n$, we define

$$\Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n | p_{XY}) := -\log \mathrm{E}_{p^{(n)}} \left[\prod_{t=1}^{n} \frac{p_{X_t}^{\bar{\alpha}}(X_t)}{Q_{X_t}^{\bar{\alpha}}(X_t)} \frac{p_{X_t}^{\mu\alpha}(X_t) p_{Y_t|U_t}^{\mu\alpha}(Y_t|U_t)}{\tilde{Q}_{X_t|U_t}^{\mu\alpha}(X_t|U_t)} \right],$$

where for each $t = 1, 2, \cdots, n$, the probability distribution Q_{X_t} and the conditional probability distribution $\tilde{Q}_{X_t|U_t}$ appearing in the definition of $\Omega^{(\mu,\theta)}(p^{(n)}, \underline{Q}^n)$ can be chosen as arbitrary.

The following is well known as the Cramèr's bound in the large deviation principle.

Lemma 5. *For any real valued random variable Z and any $\alpha \geq 0$, we have*

$$\Pr\{Z \geq a\} \leq \exp\left[-\left(\alpha a - \log \mathrm{E}[\exp(\alpha Z)] \right) \right].$$

By Lemmas 4 and 5, we have the following proposition.

Proposition 1. *For any $(\mu, \alpha) \in [0, 1]^2$ any $\underline{Q}^n \in \underline{\mathcal{Q}}^n$, and any $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log ||\varphi_i^{(n)}|| \leq R_i, i = 1, 2$, there exists $p^{(n)} \in \mathcal{P}^{(n)}(W_1, W_2)$ such that*

$$\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \leq 5 \exp\left\{ -n \left[2 + \alpha\bar{\mu} \right]^{-1} \left[\frac{1}{n} \Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n | p_{XY}) - \alpha(\mu R_1 + \bar{\mu} R_2) \right] \right\}.$$

Proof. By Lemma 4, for $(\mu, \alpha) \in [0, 1]^2$, we have the following chain of inequalities:

$$
\begin{aligned}
&\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}) \le p_{SX^nY^n}\left\{0 \ge \left[\frac{1}{n}\sum_{t=1}^{n}\log\frac{Q_{X_t}^{\bar{\alpha}}(X_t)}{p_{X_t}^{\bar{\alpha}}(X_t)} - \bar{\alpha}\eta\right],\right.\\
&\quad \alpha\mu R_1 \ge \frac{1}{n}\sum_{t=1}^{n}\log\frac{\tilde{Q}_{X_t|U_t}^{\alpha\mu}(X_t|U_t)}{p_{X_t}^{\alpha\mu}(X_t)} - \alpha\mu\eta,\\
&\quad \left.\alpha\bar{\mu}R_2 \ge \frac{1}{n}\sum_{t=1}^{n}\log\frac{1}{p_{Y_t|U_t}^{\alpha\bar{\mu}}(Y_t|U_t)} - 2\alpha\bar{\mu}\eta\right\} + 4\mathrm{e}^{-n\eta}\\
&\le p_{SX^nY^n}\left\{\alpha(\mu R_1+\bar{\mu}R_2)+(1+\alpha\bar{\mu})\eta \ge -\frac{1}{n}\sum_{t=1}^{n}\log\left[\frac{p_{X_t}^{\bar{\alpha}}(X_t)}{Q_{X_t}^{\bar{\alpha}}(X_t)}\frac{p_{X_t}^{\mu\alpha}(X_t)p_{Y_t|U_t}^{\bar{\mu}\alpha}(Y_t|U_t)}{\tilde{Q}_{X_t|U_t}^{\mu\alpha}(X_t|U_t)}\right]\right\} + 4\mathrm{e}^{-n\eta}\\
&= p_{SX^nY^n}\left\{\frac{1}{n}\sum_{t=1}^{n}\log\left[\frac{p_{X_t}^{\bar{\alpha}}(X_t)}{Q_{X_t}^{\bar{\alpha}}(X_t)}\frac{p_{X_t}^{\mu\alpha}(X_t)p_{Y_t|U_t}^{\alpha}(Y_t|U_t)}{\tilde{Q}_{X_t|U_t}^{\mu\alpha}(X_t|U_t)}\right] \ge -\left[\alpha(\mu R_1+\bar{\mu}R_2)+(1+\alpha\bar{\mu})\eta\right]\right\} + 4\mathrm{e}^{-n\eta}\\
&\overset{(a)}{\le} \exp\left[n\left\{\alpha(\mu R_1+\bar{\mu}R_2)+(1+\alpha\bar{\mu})\eta - \frac{1}{n}\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY})\right\}\right] + 4\mathrm{e}^{-n\eta}. \qquad (24)
\end{aligned}
$$

Step (a) follows from Lemma 5. When $\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY}) \le n\alpha(\mu R_1+\bar{\mu}R_2)$, the bound we wish to prove is obvious. In the following argument, we assume that $\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY}) > n\alpha(\mu R_1+\bar{\mu}R_2)$. We choose η so that

$$
-\eta = \alpha(\mu R_1+\bar{\mu}R_2)+(1+\alpha\bar{\mu})\eta - \frac{1}{n}\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY}). \qquad (25)
$$

Solving (25) with respect to η, we have

$$
\eta = \frac{(1/n)\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY}) - \alpha(\mu R_1+\bar{\mu}R_2)}{2+\alpha\bar{\mu}}.
$$

For this choice of η and (24), we have

$$
\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}) \le 5\mathrm{e}^{-n\eta} = 5\exp\left\{-n\left[2+\alpha\bar{\mu}\right]^{-1}\left[\frac{1}{n}\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY}) - \alpha(\mu R_1+\bar{\mu}R_2)\right]\right\},
$$

completing the proof. □

Set

$$
\underline{\Omega}^{(\mu,\alpha)}(p_{XY}) := \inf_{n\ge1}\min_{p^{(n)}\in\mathcal{P}^{(n)}}\max_{\underline{Q}^n\in\underline{\mathcal{Q}}^n}\frac{1}{n}\Omega^{(\mu,\alpha)}(p^{(n)},\underline{Q}^n|p_{XY}).
$$

By Proposition 1, we have the following corollary.

Corollary 3. *For any* $(\mu,\alpha)\in[0,1]^2$ *and any* $(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)})$ *satisfying* $(1/n)\log||\varphi_i^{(n)}|| \le R_i, i=1,2$, *we have*

$$
\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)},\varphi_2^{(n)},\psi^{(n)}) \le 5\exp\left\{-n\left[\frac{\underline{\Omega}^{(\mu,\alpha)}(p_{XY}) - \alpha(\mu R_1+\bar{\mu}R_2)}{2+\alpha\bar{\mu}}\right]\right\}.
$$

We shall call $\underline{\Omega}^{(\mu,\alpha)}(p_{XY})$ the communication potential. The above corollary implies that the analysis of $\underline{\Omega}^{(\mu,\alpha)}(\,p_{XY})$ leads to an establishment of a strong converse theorem for the one helper source coding problem. Note here that $\underline{\Omega}^{(\mu,\alpha)}(p_{XY})$ is still a multi letter quantity. However, we successfully

single letterize this quantity. This result which will be stated later in Proposition 2 is a mathematical core of our main result.

In the following argument, we drive an explicit lower bound of $\underline{\Omega}^{(\mu,\alpha)}(p_{XY})$. For each $t = 1, 2, \cdots, n$, set $u_t = (s, x^{t-1}) \in \mathcal{U}_t$ and

$$\mathcal{F}_t := (p_{X_t}, p_{X_tY_t|U_t}, \underline{Q}_t), \quad \mathcal{F}^t := \{\mathcal{F}_i\}_{i=1}^t.$$

For $t = 1, 2, \cdots, n$, define a function of $(u_t, x_t, y_t) \in \mathcal{U}_t \times \mathcal{X} \times \mathcal{Y}$ by

$$f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t|u_t) := \frac{p_{X_t}^{\bar{\alpha}}(x_t)}{Q_{X_t}^{\bar{\alpha}}(x_t)} \frac{p_{X_t}^{\mu\alpha}(x_t) p_{Y_t|U_t}^{\alpha}(y_t|u_t)}{\tilde{Q}_{X_t|U_t}^{\mu\alpha}(x_t|u_t)}.$$

By definition, we have

$$\exp\left\{-\Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n|p_{XY})\right\} = \sum_{s,x^n,y^n} p_{SX^nY^n}(s, x^n, y^n) \prod_{t=1}^n f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t|u_t).$$

For each $t = 1, 2, \cdots, n$, we define the probability distribution

$$p_{SX^tY^t;\mathcal{F}^t}^{(\mu,\alpha)} := \left\{p_{SX^tY^t;\mathcal{F}^t}^{(\mu,\alpha)}(s, x^t, y^t)\right\}_{(s,x^t,y^t)\in\mathcal{M}_1\times\mathcal{X}^t\times\mathcal{Y}^t}$$

by

$$p_{SX^tY^t;\mathcal{F}^t}^{(\mu,\alpha)}(s, x^t, y^t) := C_t^{-1} p_{SX^tY^t}(s, x^t, y^t) \prod_{i=1}^t f_{\mathcal{F}_i}^{(\mu,\alpha)}(x_i, y_i|u_i),$$

where

$$C_t := \sum_{s,x^t,y^t} p_{SX^tY^t}(s, x^t, y^t) \prod_{i=1}^t f_{\mathcal{F}_i}^{(\mu,\alpha)}(x_i, y_i)$$

are constants for normalization. For $t = 1, 2, \cdots, n$, define

$$\Phi_t^{(\mu,\alpha)} := C_t C_{t-1}^{-1}, \tag{26}$$

where we define $C_0 = 1$. Then, we have the following lemma.

Lemma 6. *For each $t = 1, 2, \cdots, n$, and for any $(s, x^t, y^t) \in \mathcal{M}_1 \times \mathcal{X}^t \times \mathcal{Y}^t$, we have*

$$\begin{aligned} &p_{SX^tY^t;\mathcal{F}^t}^{(\mu,\alpha)}(s, x^t, y^t) \\ &= (\Phi_t^{(\mu,\alpha)})^{-1} p_{SX^{t-1}Y^{t-1};\mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}, y^{t-1}) p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t|u_t). \end{aligned} \tag{27}$$

Furthermore, we have

$$\Phi_t^{(\mu,\alpha)} = \sum_{s,x^t,y^t} p_{SX^{t-1}Y^{t-1};\mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}, y^{t-1}) p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t|u_t). \tag{28}$$

Proof of this lemma is given in Appendix H. Define

$$p_{U_t;\mathcal{F}^{t-1}}^{(\mu,\alpha)}(u_t) = p_{SX^{t-1};\mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}) := \sum_{y^{t-1}} p_{SX^{t-1}Y^{t-1};\mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}, y^{t-1}).$$

Then, we have the following lemma, which is a key result to derive a single letterized lower bound of $\underline{\Omega}^{(\mu,\alpha)}(p_{XY})$.

Lemma 7. *For any $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$ and any $\underline{Q}^n \in \underline{\mathcal{Q}}^n$, we have*

$$\Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n | p_{XY}) = (-1) \sum_{t=1}^{n} \log \Phi_t^{(\mu,\alpha)}, \tag{29}$$

$$\Phi_t^{(\mu,\alpha)} = \sum_{u_t, x_t, y_t} p_{U_t;\mathcal{F}^{t-1}}^{(\mu,\alpha)}(u_t) p_{X_t|U_t}(x_t|u_t) p_{Y_t|X_t}(y_t|x_t) f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t|u_t). \tag{30}$$

Proof. We first prove (29). From (26), we have

$$\log \Phi_t^{(\mu,\alpha)} = -\log C_t + \log C_{t-1}. \tag{31}$$

Furthermore, by definition, we have

$$\Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n | p_{XY}) = -\log C_n, C_0 = 1. \tag{32}$$

From (31) and (32), (29) is obvious. We next prove (30). We first observe that for $(s, x^t, y^t) \in \mathcal{S} \times \mathcal{X}^t \times \mathcal{Y}^t$ and for $t = 1, 2, \cdots, n$,

$$\begin{aligned} &p_{X_t Y_t | S X^{t-1} Y^{t-1}}(x_t, y_t | s, x^{t-1}, y^{t-1}) = p_{X_t | S X^{t-1} Y^{t-1}}(x_t | s, x^{t-1}, y^{t-1}) p_{Y_t | S X^t Y^{t-1}}(y_t | s, x^t, y^{t-1}) \\ &\overset{(a)}{=} p_{X_t | S X^{t-1}}(x_t | s, x^{t-1}) p_{Y_t | X_t}(y_t | x_t). \end{aligned}$$

Step (a) follows from Lemma 3. Then, by Lemma 6, we have

$$\begin{aligned} \Phi_t^{(\mu,\alpha)} &= \sum_{s, x^t, y^t} p_{S X^{t-1} Y^{t-1}; \mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}, y^{t-1}) p_{X_t Y_t | S X^{t-1} Y^{t-1}}(x_t, y_t | s, x^{t-1}, y^{t-1}) f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t | u_t) \\ &= \sum_{s, x^t, y_t} \left\{ \sum_{y^{t-1}} p_{S X^{t-1} Y^{t-1}; \mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}, y^{t-1}) \right\} p_{X_t | S X^{t-1}}(x_t | s, x^{t-1}) p_{Y_t | X_t}(y_t | x_t) f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t | u_t) \\ &= \sum_{s, x^t, y_t} p_{S X^{t-1}; \mathcal{F}^{t-1}}^{(\mu,\alpha)}(s, x^{t-1}) p_{X_t | S X^{t-1}}(x_t | s, x^{t-1}) p_{Y_t | X_t}(y_t | x_t) f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t | u_t), \end{aligned}$$

completing the proof. □

The following proposition is a mathematical core to prove our main result.

Proposition 2. *For any $\mu \in [0, 1]$ and any $\alpha \geq 0$, we have*

$$\underline{\Omega}^{(\mu,\alpha)}(p_{XY}) \geq \Omega^{(\mu,\alpha)}(p_{XY}).$$

Proof: Set

$$\begin{aligned} \mathcal{Q}_n(p_{Y|X}) &:= \{q = q_{UXY} : |\mathcal{U}| \leq |\mathcal{M}_1||\mathcal{X}^{n-1}||\mathcal{Y}^{n-1}|, q_{Y|X} = p_{Y|X}, U \leftrightarrow X \leftrightarrow Y\}, \\ \hat{\Omega}_n^{(\mu,\alpha)}(p_{XY}) &:= \min_{q \in \mathcal{Q}_n(p_{Y|X})} \Omega^{(\mu,\alpha)}(q | p_{XY}). \end{aligned}$$

For each $t = 1, 2, \cdots, n$, we define $q_t = q_{U_t X_t Y_t Z_t}$ by

$$q_{U_t}(u_t) = p_{U_t;\mathcal{F}^{t-1}}^{(\mu,\alpha)}(u_t), q_{X_t Y_t | U_t}(x_t, y_t | u_t) = p_{X_t | U_t}(x_t | u_t) p_{Y|X}(y_t | x_t). \tag{33}$$

Equation (33) implies that $q_t = q_{U_t X_t Y_t} \in \mathcal{Q}_n(p_{Y|X})$. Furthermore, for each $t = 1, 2, \cdots, n$, we choose $\underline{Q}_t = (Q_{X_t}, \tilde{Q}_{X_t|U_t})$ appearing in

$$f_{\mathcal{F}_t}^{(\mu,\alpha)}(x_t, y_t|u_t) = \frac{p_{X_t}^{\bar{\alpha}}(x_t)}{Q_{X_t}^{\bar{\alpha}}(x_t)} \frac{p_{X_t}^{\mu\alpha}(x_t) p_{Y_t|U_t}^{\alpha}(y_t|u_t)}{\tilde{Q}_{X_t|U_t}^{\mu\alpha}(x_t|u_t)}$$

such that $\underline{Q}_t = (Q_{X_t}, \tilde{Q}_{X_t|U_t}) = (q_{X_t}, q_{X_t|U_t})$. For this choice of $\underline{Q}_t$, we have the following chain of inequalities:

$$\begin{aligned}
\Phi_t^{(\mu,\alpha)} &\overset{(a)}{=} \mathrm{E}_{q_t}\left[f_{\mathcal{F}_t}^{(\mu,\theta)}(X_t, Y_t|U_t)\right] \overset{(b)}{=} \mathrm{E}_{q_t}\left[\frac{p_{X_t}^{\bar{\alpha}}(X_t)}{q_{X_t}^{\bar{\alpha}}(X_t)} \frac{p_{X_t}^{\mu\alpha}(X_t) p_{Y_t|U_t}^{\alpha}(Y_t|U_t)}{q_{X_t|U_t}^{\mu\alpha}(X_t|U_t)}\right] = \mathrm{E}_{q_t}\left[f_{q_t|p_{X_t}}^{(\mu,\alpha)}(X_t, Y_t|U_t)\right] \\
&= \exp\left\{-\Omega^{(\mu,\alpha)}(q_t|p_{X_t})\right\} \overset{(c)}{=} \exp\left\{-\Omega^{(\mu,\alpha)}(q_t|p_X)\right\} \\
&\overset{(d)}{\leq} \exp\left\{-\hat{\Omega}_n^{(\mu,\alpha)}(p_{XY})\right\} \overset{(e)}{=} \exp\left\{-\Omega^{(\mu,\alpha)}(p_{XY})\right\}. \qquad (34)
\end{aligned}$$

Step (a) follows from Lemma 7 and (33). Step (b) follows from the choice $(Q_{X_t}, \tilde{Q}_{X_t|U_t}) = (q_{X_t}, q_{X_t|U_t})$ of $(Q_{X_t}, \tilde{Q}_{X_t|U_t})$ for $t = 1, 2, \cdots, n$. Step (c) follows from $p_{X_t} = p_X$ for $t = 1, 2, \cdots, n$. Step (d) follows from $q_t \in \mathcal{Q}_n(p_{Y|X})$ and the definition of $\hat{\Omega}_n^{(\mu,\alpha)}(p_{XY})$. Step (e) follows from Property 4 part a. Hence, we have the following:

$$\max_{\underline{Q}^n \in \underline{\mathcal{Q}}^n} \frac{1}{n}\Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n|p_{XY}) \geq \frac{1}{n}\Omega^{(\mu,\alpha)}(p^{(n)}, \underline{Q}^n|p_{XY}) \overset{(a)}{=} -\frac{1}{n}\sum_{t=1}^{n} \log \Phi_t^{(\mu,\alpha)} \overset{(b)}{\geq} \Omega^{(\mu,\alpha)}(p_{XY}). \qquad (35)$$

Step (a) follows from Lemma 7. Step (b) follows from (34). Since (35) holds fo any $n \geq 1$ and any $p_{SX^nY^n}$ satisfying $S \leftrightarrow X^n \leftrightarrow Y^n$, we have that, for any $(\mu, \alpha) \in [0,1]^2$,

$$\underline{\Omega}^{(\mu,\alpha)}(p_{XY}) \geq \Omega^{(\mu,\alpha)}(p_{XY}).$$

Thus, Proposition 2 is proved. □

Proof of Theorem 3. For any $(\mu, \alpha) \in [0,1]^2$, for any $R_1, R_2 \geq 0$ and for any $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ satisfying $(1/n)\log||\varphi_i^{(n)}|| \leq R_i, i = 1, 2$, we have the following:

$$\begin{aligned}
&\frac{1}{n}\log\left\{\frac{5}{\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})}\right\} \overset{(a)}{\geq} \frac{\underline{\Omega}^{(\mu,\alpha)}(p_{XY}) - \alpha(\mu R_1 + \bar{\mu} R_2)}{2 + \alpha\bar{\mu}} \overset{(b)}{\geq} \frac{\Omega^{(\mu,\alpha)}(p_{XY}) - \alpha(\mu R_1 + \bar{\mu} R_2)}{2 + \alpha\bar{\mu}} \\
&= F^{(\mu,\alpha)}(\mu R_1 + \bar{\mu} R_2|p_{XY}).
\end{aligned}$$

Step (a) follows from Corollary 3. Step (b) follows from Proposition 2. Since the above bound holds for any $\mu \in [0,1]$ and any $\alpha \geq 0$, we have

$$\frac{1}{n}\log\left\{\frac{5}{\mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})}\right\} \geq F(R_1, R_2|p_{XY}).$$

Thus, (10) in Theorem 3 is proved. □

Proof of Corollary 2: Since g is an inverse function of ϑ, the definition (11) of κ_n is equivalent to

$$g\left(\frac{\kappa_n}{\rho(p_{XY})}\right) = \sqrt{\frac{4}{n\rho(p_{XY})}\log\left(\frac{5}{1-\varepsilon}\right)}. \qquad (36)$$

By the definition of $n_0 = n_0(\varepsilon, \rho(p_{XY}))$, we have that $\kappa_n \le (1/2)\rho(p_{XY})$ for $n \ge n_0$. We assume that, for $n \ge n_0$, $(R_1, R_2) \in \mathcal{R}_{\mathrm{AKW}}(n, \varepsilon | p_{XY})$. Then, there exists a sequence $\{(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})\}_{n \ge n_0}$ such that, for $n \ge n_0$, we have

$$\frac{1}{n} \log ||\varphi_i^{(n)}|| \le R_i, i = 1, 2, 1 - \varepsilon \le \mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}). \tag{37}$$

Then, by Theorem 3, we have

$$1 - \varepsilon \le \mathrm{P_c}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \le 5 \exp\{-nF(R_1, R_2 | p_{XY})\} \tag{38}$$

for any $n \ge n_0(\varepsilon, \rho(p_{XY}))$. From (38), we have that for $n \ge n_0(\varepsilon, \rho(\ p_{XY}))$,

$$F(R_1, R_2 | p_{XY}) \le \frac{1}{n} \log \left(\frac{5}{1 - \varepsilon} \right) \overset{\text{(a)}}{=} \frac{\rho(p_{XY})}{4} \cdot g^2 \left(\frac{\kappa_n}{\rho(p_{XY})} \right). \tag{39}$$

Step (a) follows from (36). Hence, by Property 4 part e, we have that, under $\kappa_n \le (1/2)\rho(p_{XY})$, the inequality (39) implies

$$(R_1, R_2) \in \mathcal{R}(p_{XY}) + \kappa_n(1, 1). \tag{40}$$

Since (40) holds for any $n \ge n_0$ and $(R_1, R_2) \in \mathcal{R}_{\mathrm{AKW}}(n, \varepsilon | p_{XY})$, we have

$$\mathcal{R}_{\mathrm{AKW}}(n, \varepsilon | p_{XY}) \subseteq \mathcal{R}(p_{XY}) + \kappa_n(1, 1) \text{ for } n \ge n_0,$$

completing the proof. □

5. One Helper Problem Studied by Wyner

We consider a communication system depicted in Figure 4. Data sequences X^n, Y^n, and Z^n, respectively are separately encoded to $\varphi_1^{(n)}(X^n)$, $\varphi_2^{(n)}(Y^n)$, and $\varphi_3^{(n)}(Z^n)$. The encoded data $\varphi_1^{(n)}(X^n)$ and $\varphi_2^{(n)}(Y^n)$ are sent to the information processing center 1. The encoded data $\varphi_1^{(n)}(X^n)$ and $\varphi_3^{(n)}(Z^n)$ are sent to the information processing center 2. At center 1, the decoder function $\psi^{(n)}$ observes $(\varphi_1^{(n)}(X^n), \varphi_2^{(n)}(Y^n))$ to output the estimation $\hat{Y}^n$ of Y^n. At center 2, the decoder function $\phi^{(n)}$ observes $(\varphi_1^{(n)}(X^n), \varphi_3^{(n)}(Z^n))$ to output the estimation $\hat{Z}^n$ of Z^n. The error probability of decoding is

$$\mathrm{P_e}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)}) = \Pr\{\hat{Y}^n \ne Y^n \text{ or } \hat{Z}^n \ne Z^n\},$$

where $\hat{Y}^n = \psi^{(n)}(\ \varphi_1^{(n)}(X^n), \varphi_2^{(n)}(Y^n))$ and $\hat{Z}^n = \psi^{(n)}(\ \varphi_1^{(n)}(X^n), \varphi_3^{(n)}(Z^n))$.

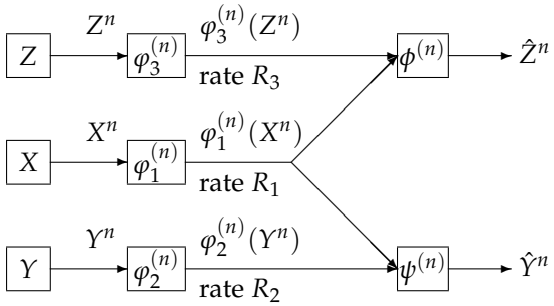

Figure 4. One helper source coding system investigated by Wyner.

A rate triple (R_1, R_2, R_3) is *ε-achievable* if, for any $\delta > 0$, there exist a positive integer $n_0 = n_0(\varepsilon, \delta)$ and a sequence of three encoders and two decoder functions $\{(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)})\}_{n \geq n_0}$ such that, for $n \geq n_0(\varepsilon, \delta)$,

$$\frac{1}{n} \log ||\varphi_i^{(n)}|| \leq R_i + \delta \text{ for } i = 1, 2, 3, \mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)}) \leq \varepsilon.$$

The rate region $\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ})$ is defined by

$$\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}) := \{(R_1, R_2, R_3) : (R_1, R_2, R_3) \text{ is } \varepsilon\text{-achievable for } p_{XYZ}\}.$$

Furthermore, define

$$\mathcal{R}_{\mathrm{W}}(p_{XYZ}) := \bigcap_{\varepsilon \in (0,1)} \mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}).$$

We can show that the two rate regions $\mathcal{R}_{\mathrm{W}}(\varepsilon|\ p_{XYZ})$, $\varepsilon \in (0,1)$ and $\mathcal{R}_{\mathrm{W}}(p_{XYZ})$ satisfy the following property.

Property 5.

(a) The regions $\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ})$, $\varepsilon \in (0,1)$, and $\mathcal{R}_{\mathrm{W}}(\ p_{XYZ})$ are closed convex sets of $\mathbb{R}_+^3$.
(b) We set

$$\begin{aligned}\mathcal{R}_{\mathrm{W}}(n, \varepsilon|p_{XYZ}) = \{&(R_1, R_2, R_3) : \text{ There exists } (\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}) \text{ such that} \\ &\frac{1}{n} \log ||\varphi_i^{(n)}|| \leq R_i, i = 1, 2, 3, \mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}) \leq \varepsilon\},\end{aligned}$$

which is called the (n, ε)-rate region. Using $\mathcal{R}_{\mathrm{W}}(n, \varepsilon|p_{XYZ})$, $\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ})$ can be expressed as

$$\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}) = \mathrm{cl}\left(\bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{\mathrm{W}}(n, \varepsilon|p_{XYZ})\right).$$

It is well known that $\mathcal{R}_{\mathrm{W}}(p_{XYZ})$ was determined by Wyner. To describe his result, we introduce an auxiliary random variable U taking values in a finite set $\mathcal{U}$. We assume that the joint distribution of (U, X, Y, Z) is

$$p_{UXY}(u, x, y, z) = p_U(u) p_{X|U}(x|u) p_{YZ|X}(y, z|x).$$

The above condition is equivalent to $U \leftrightarrow X \leftrightarrow YZ$. Define the set of probability distribution on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ by

$$\mathcal{P}(p_{XYZ}) := \{p = p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 2, U \leftrightarrow X \leftrightarrow YZ\}.$$

Set

$$\begin{aligned}\mathcal{R}(p) := \ &\{(R_1, R_2, R_3) : R_1, R_2, R_3 \geq 0, \\ &R_1 \geq I_p(X; U), R_2 \geq H_p(Y|U), R_3 \geq H_p(Z|U)\}, \\ \mathcal{R}(p_{XYZ}) := \ &\bigcup_{p \in \mathcal{P}(p_{XYZ})} \mathcal{R}(p).\end{aligned}$$

We can show that the region $\mathcal{R}(p_{XYZ})$ satisfies the following property.

Property 6.

(a) The region $\mathcal{R}(p_{XYZ})$ is a closed convex subset of $\mathbb{R}_+^3$.

(b) For any p_{XYZ}, and any $\gamma \in [0,1]$, we have

$$\min_{(R_1,R_2,R_3)\in\mathcal{R}(p_{XY})} (R_1 + \bar{\gamma}R_2 + \gamma R_3) = \bar{\gamma}H_p(Y) + \gamma H_p(Z). \tag{41}$$

The minimun is attained by $(R_1, R_2, R_3) = (0, H_p(Y), H_p(Z))$. This result implies that

$$\mathcal{R}(p_{XYZ}) \subseteq \left[\bigcap_{\gamma\in[0,1]} \{(R_1,R_2,R_3) : R_1 + \bar{\gamma}R_2 + \gamma R_3 \geq \bar{\gamma}H_p(Y) + \gamma H_p(Z)\}\right] \cap \mathbb{R}^3_+.$$

Furthermore, the point $(0, H_p(Y), H_p(Z))$ always belongs to $\mathcal{R}(p_{XYZ})$.

The rate region $\mathcal{R}_{\mathrm{W}}(p_{XYZ})$ was determined by Wyner [2]. His result is the following.

Theorem 4 (Wyner [2])**.**

$$\mathcal{R}_{\mathrm{W}}(p_{XYZ}) = \mathcal{R}(p_{XYZ}).$$

On the strong converse theorem, Csiszár and Körner [20] obtained the following.

Theorem 5 (Csiszár and Körner [20])**.** *For each fixed $\varepsilon \in (0,1)$, we have*

$$\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}) = \mathcal{R}(p_{XYZ}).$$

To examine a rate of convergence for the error probability of decoding to tend to one as $n \to \infty$ for $(R_1, R_2, R_3) \notin \mathcal{R}_{\mathrm{W}}(p_{XYZ})$, we define the following quantity. Set

$$\begin{aligned}
&\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)}) := 1 - \mathrm{P}_{\mathrm{e}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)}),\\
&G^{(n)}(R_1,R_2,R_3|p_{XYZ}) := \min_{\substack{(\varphi_1^{(n)},\varphi_2^{(n)},\varphi_3^{(n)},\\ \psi^{(n)},\phi^{(n)}):\\ (1/n)\log\|\varphi_i^{(n)}\| \\ \leq R_i, i=1,2,3}} \left(-\frac{1}{n}\right) \log \mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)}),\\
&G(R_1,R_2,R_3|p_{XYZ}) := \lim_{n\to\infty} G^{(n)}(R_1,R_2,R_3|p_{XYZ}),\\
&\mathcal{G}(p_{XYZ}) := \{(R_1,R_2,R_3,G) : G \geq G(R_1,R_2,R_3|p_{XYZ})\}.
\end{aligned}$$

By time sharing, we have that

$$\begin{aligned}
&G^{(n+m)}\left(\frac{nR_1 + mR_1'}{n+m}, \frac{nR_2 + mR_2'}{n+m}, \frac{nR_2 + mR_2'}{n+m}\middle| p_{XYZ}\right)\\
&\leq \frac{nG^{(n)}(R_1,R_2,R_3|p_{XYZ}) + mG^{(m)}(R_1',R_2',R_3'|p_{XYZ})}{n+m}.
\end{aligned} \tag{42}$$

Choosing $R = R'$ in (42), we obtain the following subadditivity property on $\{G^{(n)}(R_1,R_2,R_3|p_{XYZ})\}_{n\geq 1}$:

$$G^{(n+m)}(R_1,R_2,R_3|p_{XYZ}) \leq \frac{nG^{(n)}(R_1,R_2,R_3|p_{XYZ}) + mG^{(m)}(R_1,R_2,R_3|p_{XYZ})}{n+m},$$

from which we have that $G(R_1,R_2,R_3|p_{XYZ})$ exists and satisfies the following:

$$G(R_1,R_2,R_3|p_{XYZ}) = \inf_{n\geq 1} G^{(n)}(R_1,R_2,R_3|p_{XYZ}).$$

The exponent function $G(R_1, R_2, R_3|p_{XYZ})$ is a convex function of (R_1, R_2, R_3). In fact, by time sharing, we have that

$$G^{(n+m)}\left(\frac{nR_1+mR_1'}{n+m}, \frac{nR_2+mR_2'}{n+m}, \frac{nR_2+mR_2'}{n+m}\middle| p_{XYZ}\right) \leq \frac{nG^{(n)}(R_1, R_2, R_3|p_{XYZ}) + mG^{(m)}(R_1', R_2', R_3'|p_{XYZ})}{n+m},$$

from which we have that for any $\alpha \in [0,1]$

$$G(\alpha R_1 + \bar{\alpha}R_1', \alpha R_2 + \bar{\alpha}R_2', \alpha R_3 + \bar{\alpha}R_3'|p_{XYZ}) \leq \alpha G(R_1, R_2, R_3|p_{XYZ}) + \bar{\alpha}G(R_1', R_2', R_3'|p_{XYZ}).$$

The region $\mathcal{G}(p_{XYZ})$ is also a closed convex set. Our main aim is to find an explicit characterization of $\mathcal{G}(p_{XYZ})$. In this paper, we derive an explicit outer bound of $\mathcal{G}\ (p_{XYZ})$ whose section by the plane $G = 0$ coincides with $\mathcal{R}_{\mathrm{W}}(p_{XYZ})$. We first explain that the region $\mathcal{R}(p_{XYZ})$ has another expression using the supporting hyperplane. We define two sets of probability distributions on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ by

$$\mathcal{P}_{\mathrm{sh}}(p_{XYZ}) := \{p = p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|, U \leftrightarrow X \leftrightarrow YZ\},$$
$$\mathcal{Q}(p_{YZ|X}) := \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|, p_{YZ|X} = q_{YZ|X}, U \leftrightarrow X \leftrightarrow YZ\}.$$

For $(\mu, \gamma) \in [0,1]^2$, set

$$R^{(\mu,\gamma)}(p_{XYZ}) := \max_{p \in \mathcal{P}_{\mathrm{sh}}(p_{XYZ})} \left\{\mu I_p(X;U) + \bar{\mu}(\bar{\gamma}H_p(Y|U) + \gamma H_p(Z|U))\right\}.$$

Furthermore, define

$$\mathcal{R}_{\mathrm{sh}}(p_{XYZ}) = \bigcap_{(\mu,\gamma)\in[0,1]^2} \{(R_1, R_2, R_3) : \mu R_1 + \bar{\mu}(\bar{\gamma}R_2 + \gamma R_3) \geq R^{(\mu,\gamma)}(p_{XYZ})\}.$$

Then, we have the following property.

Property 7.

(a) The bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $R^{(\mu)}(\ p_{XYZ})$.
(b) For every $(\mu, \gamma) \in [0,1]^2$, we have

$$\min_{(R_1,R_2,R_3)\in\mathcal{R}(p_{XYZ})} \{\mu R_1 + \bar{\mu}(\bar{\gamma}R_2 + \gamma R_3)\} = R^{(\mu,\gamma)}(p_{XYZ}).$$

(c) For any p_{XYZ}, we have

$$\mathcal{R}_{\mathrm{sh}}(p_{XYZ}) = \mathcal{R}(p_{XYZ}). \tag{43}$$

For $(\mu,\gamma,\alpha)\in[0,1]^3$, and for $q=q_{UXYZ}\in\mathcal{Q}(p_{YZ|X})$, define

$$\omega_{q|p_X}^{(\mu,\gamma,\alpha)}(x,y,z|u):=\bar{\alpha}\log\frac{q_X(x)}{p_X(x)}+\alpha\left[\mu\log\frac{q_{X|U}(x|u)}{p_X(x)}+\bar{\mu}\left(\bar{\gamma}\log\frac{1}{q_{Y|U}(y|u)}+\gamma\log\frac{1}{q_{Z|U}(z|u)}\right)\right],$$
$$f_{q|p_X}^{(\mu,\gamma,\alpha)}(x,y,z|u):=\exp\left\{-\omega_{q|p_X}^{(\mu,\gamma,\alpha)}(x,y,z|u)\right\},$$
$$\Omega^{(\mu,\gamma,\alpha)}(q|p_X):=-\log\mathrm{E}_q\left[f_{q|p_X}^{(\mu,\gamma,\alpha)}(X,Y,Z|U)\right],\Omega^{(\mu,\gamma,\alpha)}(p_{XYZ}):=\min_{q\in\mathcal{Q}(p_{YZ|X})}\Omega^{(\mu,\gamma,\alpha)}(q|p_X),$$
$$F^{(\mu,\gamma,\alpha)}(\mu R_1+\bar{\gamma}R_2+\gamma R_3):=\frac{\Omega^{(\mu,\gamma,\alpha)}(p_{XYZ})-\alpha[\mu R_1+\bar{\mu}(\bar{\gamma}R_2+\gamma R_3)]}{2+\alpha\bar{\mu}},$$
$$F(R_1,R_2,R_3|p_{XYZ}):=\sup_{(\mu,\gamma,\alpha)\in[0,1]^3,}F^{(\mu,\gamma,\alpha)}(\mu R_1+\bar{\mu}(\bar{\gamma}R_2+\gamma R_3)|p_{XYZ}).$$

We next define a function serving as a lower bound of $F(R_1,R_2,R_3|p_{XYZ})$. For each $p=p_{UXYZ}\in\mathcal{P}_{\mathrm{sh}}(p_{XYZ})$, define

$$\tilde{\omega}_p^{(\mu,\gamma)}(x,y,z|u):=\mu\log\frac{p_{X|U}(x|u)}{p_X(x)}+\bar{\mu}\left(\bar{\gamma}\log\frac{1}{p_{Y|U}(y|u)}+\gamma\log\frac{1}{p_{Z|U}(z|u)}\right),$$
$$\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p):=-\log\mathrm{E}_p\left[\exp\left\{-\lambda\omega_p^{(\mu,\gamma)}(X,Y,Z|U)\right\}\right],\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p_{XYZ}):=\min_{p\in\mathcal{P}_{\mathrm{sh}}(p_{XYZ})}\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p).$$

Furthermore, set

$$\underline{F}^{(\mu,\gamma,\lambda)}(\mu R_1+\bar{\gamma}R_2+\gamma R_3|p_{XYZ}):=\frac{\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p_{XYZ})-\lambda[\mu R_1+\bar{\mu}(\bar{\gamma}R_2+\gamma R_3)]}{2+\lambda(5-\mu)},$$
$$\underline{F}(R_1,R_2,R_3|p_{XYZ}):=\sup_{\substack{(\mu,\gamma)\in[0,1]^2,\\ \lambda\geq 0}}\underline{F}^{(\mu,\gamma,\lambda)}(\mu R_1+\bar{\mu}\bar{\gamma}R_2+\gamma R_3|p_{XYZ}).$$

We can show that the above functions and sets satisfy the following property.

Property 8.

(a) The cardinality bound $|\mathcal{U}|\leq|\mathcal{X}|$ in $\mathcal{Q}(p_{Y|X})$ is sufficient to describe the quantity $\Omega^{(\mu,\alpha)}(p_{XY})$. Furthermore, the cardinality bound $|\mathcal{U}|\leq|\mathcal{X}|$ in $\mathcal{Q}(p_{YZ|X})$ is sufficient to describe the quantity $\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p_{XYZ})$.

(b) For any $R_1,R_2,R_3\geq 0$, we have

$$F(R_1,R_2,R_3|p_{XYZ})\geq\underline{F}(R_1,R_2,R_3|p_{XYZ}).$$

(c) For any $p=p_{UXY}\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$ and any $(\mu,\gamma,\lambda)\in[0,1]^3$, we have

$$0\leq\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p)\leq\mu\log|\mathcal{X}|+\bar{\mu}\log(|\mathcal{Y}|^{\bar{\gamma}}|\mathcal{Z}|^{\gamma}).\tag{44}$$

(d) Fix any $p=p_{UXYZ}\in\mathcal{P}_{\mathrm{sh}}(p_{XYZ})$ and $(\mu,\gamma)\in[0,1]^2$. We define a probability distribution $p^{(\lambda)}=p_{UXYZ}^{(\lambda)}$ by

$$p^{(\lambda)}(u,x,y,z):=\frac{p(u,x,y,z)\exp\left\{-\lambda\omega_p^{(\mu,\gamma)}(x,y,z|u)\right\}}{\mathrm{E}_p\left[\exp\left\{-\lambda\omega_p^{(\mu,\gamma)}(X,Y,Z|U)\right\}\right]}.$$

Then, for $\lambda \in [0,1/2]$, $\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p)$ *is twice differentiable. Furthermore, for* $\lambda \in [0,1/2]$, *we have*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p) = \mathrm{E}_{p^{(\lambda)}}\left[\omega_p^{(\mu,\gamma)}(X,Y,Z|U)\right], \frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p) = -\mathrm{Var}_{p^{(\lambda)}}\left[\omega_p^{(\mu,\gamma)}(X,Y,Z|U)\right].$$

The second equality implies that $\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p)$ *is a concave function of* $\lambda \in [0,1/2]$.

(e) For $(\mu,\gamma,\lambda) \in [0,1]^2 \times [0,1/2]$, *define*

$$\rho^{(\mu,\gamma,\lambda)}(p_{XYZ}) := \max_{\substack{(\nu,p)\in[0,\lambda]\times\mathcal{P}_{\mathrm{sh}}(p_{XYZ}):\\ \tilde{\Omega}^{(\mu,\gamma,\lambda)}(p)=\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p_{XYZ})}} \mathrm{Var}_{p^{(\nu)}}\left[\tilde{\omega}_p^{(\mu,\gamma)}(X,Y,Z|U)\right],$$

and set

$$\rho = \rho(p_{XYZ}) := \max_{(\mu,\gamma,\lambda)\in[0,1]^2\times[0,1/2]} \rho^{(\mu,\gamma,\lambda)}(p_{XYZ}).$$

Then, we have $\rho(p_{XYZ}) < \infty$. *Furthermore, for any* $(\mu,\gamma,\lambda) \in [0,1]^2 \times [0,1/2]$, *we have*

$$\tilde{\Omega}^{(\mu,\gamma,\lambda)}(p_{XYZ}) \geq \lambda R^{(\mu,\gamma)}(p_{XYZ}) - \frac{\lambda^2}{2}\rho(p_{XYZ}).$$

(f) For every $\tau \in (0,(1/2)\rho(p_{XYZ}))$, *the condition* $(R_1+\tau, R_2+\tau, R_3+\tau) \notin \mathcal{R}(p_{XYZ})$ *implies*

$$F(R_1,R_2,R_3|p_{XYZ}) > \frac{\rho(p_{XYZ})}{4}\cdot g^2\left(\frac{\tau}{\rho(p_{XYZ})}\right) > 0.$$

Since proofs of the results stated in Property 8 are quite parallel with those of the results stated in Property 4, we omit them. Our main result is the following.

Theorem 6. *For any* $R_1, R_2, R_3 \geq 0$, *any* p_{XYZ}, *and for any* $(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)})$ *satisfying* $(1/n)\log||\varphi_i^{(n)}|| \leq R_i, i = 1,2,3$, *we have*

$$\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \varphi_3^{(n)}, \psi^{(n)}, \phi^{(n)}) \leq 7\exp\left\{-nF(R_1,R_2,R_3|p_{XYZ})\right\}.$$

It follows from Theorem 6 and Property 8 part d) that, if (R_1,R_2,R_3) is outside the capacity region, then the error probability of decoding goes to one exponentially and its exponent is not below $F(R_1,R_2,R_3|p_{XYZ})$. It immediately follows from Theorem 3 that we have the following corollary.

Corollary 4.

$$G(R_1,R_2,R_3|p_{XYZ}) \geq F(R_1,R_2,R_3|p_{XYZ}),$$
$$\mathcal{G}(p_{XYZ}) \subseteq \overline{\mathcal{G}}(p_{XYZ}) = \left\{(R_1,R_2,R_3,G) : G \geq F(R_1,R_2,R_3|p_{XYZ})\right\}.$$

Proof of Theorem 6 is quite parallel with that of Theorem 3. We omit the detail of the proof. From Theorem 6 and Property 8 part e, we can obtain an explicit outer bound of $\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ})$ with an asymptotically vanishing deviation from $\mathcal{R}_{\mathrm{W}}(p_{XYZ}) = \mathcal{R}(p_{XYZ})$. The strong converse theorem established by Csiszár and Körner [20] immediately follows from this corollary. To describe this outer bound, for $\kappa > 0$, we set

$$\mathcal{R}(p_{XYZ}) - \kappa(1,1,1) := \{(R_1-\kappa, R_2-\kappa, R_3-\kappa) : (R_1,R_2,R_3) \in \mathcal{R}(p_{XYZ})\},$$

which serves as an outer bound of $\mathcal{R}(p_{XYZ})$. For each fixed $\varepsilon \in (0,1)$, we define $\tilde{\kappa}_n = \tilde{\kappa}_n(\varepsilon, \rho(p_{XYZ}))$ by

$$\tilde{\kappa}_n := \rho(p_{XY})\vartheta\left(\sqrt{\frac{4}{n\rho(p_{XY})}\log\left(\frac{7}{1-\varepsilon}\right)}\right) \tag{45}$$
$$\overset{(a)}{=} 2\sqrt{\frac{\rho(p_{XY})}{n}\log\left(\frac{7}{1-\varepsilon}\right)} + \frac{5}{n}\log\left(\frac{7}{1-\varepsilon}\right).$$

Step (a) follows from $\vartheta(a) = a + (5/4)a^2$. Since $\tilde{\kappa}_n \to 0$ as $n \to \infty$, we have the smallest positive integer $n_1 = n_1(\varepsilon, \rho(p_{XYZ}))$ such that $\tilde{\kappa}_n \le (1/2)\rho(p_{XYZ})$ for $n \ge n_1$. From Theorem 6 and Property 8 part e, we have the following corollary.

Corollary 5. *For each fixed $\varepsilon \in (0,1)$, we choose the above positive integer $n_1 = n_1(\varepsilon, \rho(p_{XYZ}))$. Then, for any $n \ge n_1$, we have*

$$\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}) \subseteq \mathcal{R}(p_{XYZ}) - \tilde{\kappa}_n(0,1,1).$$

The above result together with

$$\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}) = \mathrm{cl}\left(\bigcup_{m\ge 1}\bigcap_{n\ge m}\mathcal{R}_{\mathrm{W}}(n,\varepsilon|p_{XYZ})\right)$$

yields that for each fixed $\varepsilon \in (0,1)$, we have

$$\mathcal{R}_{\mathrm{W}}(\varepsilon|p_{XYZ}) = \mathcal{R}_{\mathrm{W}}(p_{XYZ}) = \mathcal{R}(p_{XYZ}).$$

This recovers the strong converse theorem proved by Csiszár and Körner [20].

Proof of this corollary is quite parallel with that of Corollary 2. We omit the detail.

Funding: This research received no external funding

Acknowledgments: The author is very grateful to Shun Watanabe and Shigeaki Kuzuoka for their helpful comments.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Properties of the Rate Regions

In this appendix, we prove Property 1. Property 1 part a can easily be proved by the definitions of the rate distortion regions. We omit the proofs of this part. In the following argument, we prove the part b.

Proof of Property 1. Proof of Property 1 part b: We set

$$\underline{\mathcal{R}}_{\mathrm{AKW}}(m,\varepsilon|p_{XY}) = \bigcap_{n\ge m}\mathcal{R}_{\mathrm{AKW}}(n,\varepsilon|p_{XY}).$$

By the definitions of $\underline{\mathcal{R}}_{\mathrm{AKW}}(m,\varepsilon|p_{XY})$ and $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$, we have that $\underline{\mathcal{R}}_{\mathrm{AKW}}(m,\varepsilon|p_{XY}) \subseteq \mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ for $m \ge 1$. Hence, we have that

$$\bigcup_{m\ge 1}\underline{\mathcal{R}}_{\mathrm{AKW}}(m,\varepsilon|p_{XY}) \subseteq \mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}). \tag{A1}$$

We next assume that $(R_1, R_2) \in \mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$. Set

$$\mathcal{R}^{(\delta)}_{\mathrm{AKW}}(\varepsilon|p_{XY}) := \{(R_1+\delta, R_2+\delta) : (R_1, R_2) \in \mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})\}.$$

Then, by the definitions of $\mathcal{R}_{\mathrm{AKW}}(n, \varepsilon\,|p_{XY})$ and $\mathcal{R}_{\mathrm{AKW}}(\,\varepsilon|p_{XY})$, we have that, for any $\delta > 0$, there exists $n_0(\varepsilon, \delta)$ such that for any $n \geq n_0(\varepsilon, \delta)$, $(R_1+\delta, R_2+\delta) \in \mathcal{R}_{\mathrm{AKW}}(n, \varepsilon|p_{XY})$, which implies that

$$\begin{aligned}\mathcal{R}^{(\delta)}_{\mathrm{AKW}}(\varepsilon|p_{XY}) &\subseteq \bigcap_{n \geq n_0(\varepsilon,\delta)} \mathcal{R}_{\mathrm{AKW}}(n, \varepsilon|p_{XY}) = \underline{\mathcal{R}}_{\mathrm{AKW}}(n_0(\varepsilon,\delta), \varepsilon|p_{XY}) \\ &\subseteq \mathrm{cl}\left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\mathrm{AKW}}(m, \varepsilon|p_{XY})\right). \end{aligned} \tag{A2}$$

Here, we assume that there exists a pair (R_1, R_2) belonging to $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ such that

$$(R_1, R_2) \notin \mathrm{cl}\left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\mathrm{AKW}}(m, \varepsilon|p_{XY})\right). \tag{A3}$$

Since the set on the right-hand side of (A3) is a closed set, we have

$$(R_1+\delta, R_2+\delta) \notin \mathrm{cl}\left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\mathrm{AKW}}(m, \varepsilon|p_{XY})\right) \tag{A4}$$

for some small $\delta > 0$. On the other hand, we have $(R_1+\delta, R_2+\delta) \in \mathcal{R}^{(\delta)}_{\mathrm{AKW}}(\varepsilon|p_{XY})$, which contradicts (A2). Thus, we have

$$\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\mathrm{AKW}}(m, \varepsilon|p_{XY}) \subseteq \mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY}) \subseteq \mathrm{cl}\left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\mathrm{AKW}}(m, \varepsilon|p_{XY})\right). \tag{A5}$$

Note here that $\mathcal{R}_{\mathrm{AKW}}(\varepsilon|p_{XY})$ is a closed set. Then, from (A5), we conclude that

$$\mathcal{R}_{\mathrm{AKW}}(\varepsilon|W) = \mathrm{cl}\left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{\mathrm{AKW}}(m, \varepsilon|p_{XY})\right) = \mathrm{cl}\left(\bigcup_{m \geq 1} \bigcap_{n \geq m} \mathcal{R}_{\mathrm{AKW}}(n, \varepsilon|p_{XY})\right),$$

completing the proof. □

Appendix B. Cardinality Bound on Auxiliary Random Variables

We first prove the following lemma.

Lemma A1.

$$\begin{aligned}\underline{R}^{(\mu)}(p_{XY}) &:= \min_{p \in \mathcal{P}(p_{XY})} \{\mu I_p(X;U) + \bar{\mu} H_p(Y|U)\} \\ &= R^{(\mu)}(p_{XY}) := \min_{p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})} \{\mu I_p(X;U) + \bar{\mu} H_p(Y|U)\}.\end{aligned}$$

Proof. We bound the cardinality $|\mathcal{U}|$ of U to show that the bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $\underline{R}^{(\mu)}(p_{XY})$. Observe that

$$p_X(x) = \sum_{u \in \mathcal{U}} p_U(u) p_{X|U}(x|u), \tag{A6}$$

$$\mu I_p(X;U) + \bar{\mu} H_p(Y|U) = \sum_{u \in \mathcal{U}} p_U(u) \pi(p_{X|U}(\cdot|u)), \tag{A7}$$

where

$$\pi(p_{X|U}(\cdot|u)) := \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p_{X|U}(x|u)p_{Y|X}(y|x)\log\left\{\frac{p^{\mu}_{X|U}(x|u)}{p^{\mu}_{X}(x)}\left[\sum_{\tilde{x}\in\mathcal{X}} p_{Y|X}(y|\tilde{x})p_{X|U}(\tilde{x}|u)\right]^{-\bar{\mu}}\right\}.$$

For each $u \in \mathcal{U}$, $\pi(p_{X|U}(\cdot|u))$ is a continuous function of $p_{X|U}(\cdot|u)$. Then, by the support lemma, $|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$ is sufficient to express $|\mathcal{X}| - 1$ values of (A6) and one value of (A7). □

Next, we prove the following lemma.

Lemma A2. *The cardinality bound $|\mathcal{U}| \leq |\mathcal{X}|$ in $\mathcal{Q}(p_{Y|X})$ is sufficient to describe the quantity $\Omega^{(\mu,\alpha)}(p_{XY})$. The cardinality bound $|\mathcal{U}| \leq |\mathcal{X}|$ in $\mathcal{P}_{\mathrm{sh}}(p_{XY})$ is sufficient to describe the quantity $\tilde{\Omega}^{(\mu,\lambda)}(p_{XY})$.*

Proof. We first bound the cardinality $|\mathcal{U}|$ of U in $\mathcal{Q}(p_{Y|X})$ to show that the bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $\Omega^{(\mu,\alpha)}(p_{XY})$. Observe that

$$q_X(x) = \sum_{u\in\mathcal{U}} q_U(u)q_{X|U}(x|u), \tag{A8}$$

$$\exp\left\{-\Omega^{(\mu,\alpha)}(q|p_X)\right\} = \sum_{u\in\mathcal{U}} q_U(u)\Pi^{(\mu,\alpha)}(q_X, q_{XY|U}(\cdot,\cdot|u)), \tag{A9}$$

where

$$\Pi^{(\mu,\alpha)}(q_X, q_{XY|U}(\cdot,\cdot|u)) := \sum_{\substack{(x,y)\\ \in\mathcal{X}\times\mathcal{Y}}} q_{XY|U}(x,y|u)\exp\left\{-\omega^{(\mu,\alpha)}_{q|p_X}(x,y|u)\right\}.$$

The value of q_X included in $\Pi^{(\mu,\alpha)}(q_X, q_{XY|U}(\cdot,\cdot|u))$ must be preserved under the reduction of $\mathcal{U}$. For each $u \in \mathcal{U}$, $\Pi^{(\mu,\alpha)}(q_X, q_{XY|U}(\cdot,\cdot|u))$ is a continuous function of $q_{XY|U}(\,\cdot,\cdot|u)$. Then, by the support lemma, $|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$ is sufficient to express $|\mathcal{X}| - 1$ values of (A8) and one value of (A9). We next bound the cardinality $|\mathcal{U}|$ of U in $\mathcal{P}_{\mathrm{sh}}(p_{XY})$ to show that the bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $\tilde{\Omega}^{(\mu,\lambda)}(p_{XY})$. Observe that

$$p_X(x) = \sum_{u\in\mathcal{U}} p_U(u)p_{X|U}(x|u), \tag{A10}$$

$$\exp\left\{-\tilde{\Omega}^{(\mu,\lambda)}(p)\right\} = \sum_{u\in\mathcal{U}} p_U(u)\tilde{\Pi}^{(\mu,\lambda)}(p_X, p_{XY|U}(\cdot,\cdot|u)), \tag{A11}$$

where

$$\tilde{\Pi}^{(\mu,\lambda)}(p_X, p_{XY|U}(\cdot,\cdot|u)) := \sum_{\substack{(x,y)\\ \in\mathcal{X}\times\mathcal{Y}}} p_{XY|U}(x,y|u)\exp\left\{-\lambda\tilde{\omega}^{(\mu)}_{p}(x,y|u)\right\}.$$

The value of p_X included in $\tilde{\Pi}^{(\mu,\lambda)}(p_X, p_{XY|U}(\cdot,\cdot|u))$ must be preserved under the reduction of $\mathcal{U}$. For each $u \in \mathcal{U}$, $\tilde{\Pi}^{(\mu,\lambda)}(p_X, p_{XY|U}(\cdot,\cdot|u))$ is a continuous function of $p_{XY|U}(\cdot,\cdot|u)$. Then, by the support lemma, $|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$ is sufficient to express $|\mathcal{X}| - 1$ values of (A10) and one value of (A11). □

Appendix C. Supporting Hyperplain Expressions of $\mathcal{R}(p_{XY})$

In this appendix we prove Property 3 parts (b), (c). We first prove the part (b).

Proof of Property 3 part b: For any $\mu \geq 0$, we have the following chain of inequalities:

$$\begin{aligned}&\min_{(R_1,R_2)\in\mathcal{R}(p_{XY})}\{\mu R_1+\bar{\mu}R_2\}\\&=\min_{p\in\mathcal{P}(p_{XY})}\{\mu I_p(X;U)+\bar{\mu}H_p(Y|U)\}\overset{(a)}{=}\min_{p\in\mathcal{P}_{\mathrm{sh}}(p_{XY})}\{\mu I_p(X;U)+\bar{\mu}H_p(Y|U)\}=R^{(\mu)}(p_{XY}).\end{aligned}$$

Step (a) follows from Lemma A1 stating that the cardinality bound $|\mathcal{U}| \leq |\mathcal{X}|+1$ in $\mathcal{P}(p_{XY})$ can be reduced to that $|\mathcal{U}| \leq |\mathcal{X}|$ in $\mathcal{P}_{\mathrm{sh}}(p_{XY})$. □

We next prove part c. We first prepare a lemma useful to prove this property. From the convex property of the region $\mathcal{R}(p_{XY})$, we have the following lemma.

Lemma A3. *Suppose that $(\hat{R}_1,\hat{R}_2)$ does not belong to $\mathcal{R}(p_{XY})$. Then, there exist $\epsilon>0$ and $\mu_0 \geq 0$ such that for any $(R_1,R_2)\in\mathcal{R}(p_{XY})$ we have*

$$\mu_0(R_1-\hat{R}_1)+\overline{\mu_0}(R_2-\hat{R}_2)-\epsilon \geq 0.$$

Proof of this lemma is omitted here. Lemma A3 is equivalent to the fact that if the region $\mathcal{R}(p_{XY})$ is a convex set; then, for any point $(\hat{R}_1,\hat{R}_2)$ outside the region $\mathcal{R}(p_{XY})$, there exists a line which separates the point $(\hat{R}_1,\hat{R}_2)$ from the region $\mathcal{R}(p_{XY})$.

Proof of Property 3 part c: We first prove $\underline{\mathcal{R}}_{\mathrm{sh}}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$. We assume that $(\hat{R}_1,\hat{R}_2)\notin\mathcal{R}(p_{XY})$. Then, by Lemma A3, there exist $\epsilon>0$ and $\mu_0\geq 0$ such that for any $(R_1,R_2)\in\mathcal{R}(p_{XY})$, we have

$$\mu_0\hat{R}_1+\overline{\mu_0}\hat{R}_2 \leq \mu_0 R_1+\overline{\mu_0}R_2-\epsilon.$$

Then, we have

$$\begin{aligned}&\mu_0\hat{R}_1+\overline{\mu_0}\hat{R}_2 \leq \min_{(R_1,R_2)\in\mathcal{R}(p_{XY})}\{\mu_0R_1+\overline{\mu_0}R_2\}-\epsilon\overset{(a)}{=}\min_{p\in\mathcal{P}(p_{XY})}\{\mu_0 I_p(U;X)+\overline{\mu_0}H_p(Y|U)\}-\epsilon\\&\leq\min_{p\in\mathcal{P}_{\mathrm{sh}}(p_{XY})}\{\mu_0 I_p(U;X)+\overline{\mu_0}H_p(Y|U)\}-\epsilon=R^{(\mu_0)}(p_{XY})-\epsilon.\end{aligned}\tag{A12}$$

Step (a) follows from the definition of $\mathcal{R}(p_{XY})$. The inequality (A12) implies that $(\hat{R}_1,\hat{R}_2)\notin\mathcal{R}_{\mathrm{sh}}(p_{XY})$. Thus $\mathcal{R}_{\mathrm{sh}}(p_{XY})\subseteq\mathcal{R}(p_{XY})$ is concluded. □

Appendix D. Proof of Property 4 Part b

In this appendix, we prove Property 4 part b. Fix $q=q_{UXY}\in\mathcal{Q}(p_{Y|X})$ and $p=p_{UXY}=(p_{U|X},p_{XY})\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$ arbitrary. For $\beta\geq 0$, $p\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$, and $q_{Y|U}$ induced by q, define

$$\begin{aligned}\hat{\omega}^{(\mu)}_{p,q_{Y|U}}(x,y|u)&:=\mu\log\frac{p_{X|U}(x|u)}{p_X(x)}+\bar{\mu}\log\frac{1}{q_{Y|U}(y|u)},\\\hat{\Omega}^{(\mu,\beta)}(p,q_{Y|U})&:=-\log\mathrm{E}_p\left[\exp\left\{-\beta\hat{\omega}^{(\mu)}_{p,q_{Y|U}}(X,Y|U)\right\}\right].\end{aligned}$$

Then, we have the following two lemmas.

Lemma A4. *For any $\mu\in[0,1]$, $\alpha\in[0,1)$, and any $q=q_{UXY}\in\mathcal{Q}(p_{Y|X})$, there exists $p=p_{UXY}\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$ such that*

$$\Omega^{(\mu,\alpha)}(q|p_X)\geq\bar{\alpha}\hat{\Omega}^{(\mu,\frac{\alpha}{\bar{\alpha}})}(p,q_{Y|U}).\tag{A13}$$

Lemma A5. *For any μ,α satisfying $\mu\in[0,1]$, $\alpha\in[0,1/2)$, any $p=p_{UXY}\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$, and any stochastic matrix $q_{Y|U}$ induced by $q_{UXY}\in\mathcal{Q}(p_{Y|X})$, we have*

$$\hat{\Omega}^{(\mu,\frac{\alpha}{\bar{\alpha}})}(p,q_{Y|U}) \geq \frac{1-2\alpha}{\bar{\alpha}}\tilde{\Omega}^{(\mu,\frac{\alpha}{1-2\alpha})}(p). \tag{A14}$$

From Lemmas A4 and A5, we have the following corollary.

Corollary A1. *For any μ,α satisfying $\mu\in[0,1]$, $\alpha\in[0,1/2)$, and any $q=q_{UXY}\in\mathcal{Q}(p_{Y|X})$, there exists $p=p_{UXY}\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$ such that*

$$\Omega^{(\mu,\alpha)}(q|p_X) \geq (1-2\alpha)\tilde{\Omega}^{(\mu,\frac{\alpha}{1-2\alpha})}(p). \tag{A15}$$

From (A15), we have that for any $\mu\in[0,1]$, $\alpha\in[0,1/2)$, we have

$$\Omega^{(\mu,\alpha)}(p_{XY}) \geq (1-2\alpha)\tilde{\Omega}^{(\mu,\frac{\alpha}{1-2\alpha})}(p_{XY}). \tag{A16}$$

Proof of Lemma A4: We fix $(\mu,\alpha)\in[0,1]^2$ arbitrary. For each $q=q_{UXY}\in\mathcal{Q}(p_{Y|X})$, we choose $p=p_{UXY}\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$ so that $p_{U|X}=q_{U|X}$. Then, we have the following:

$$\begin{aligned}
&\exp\left\{-\Omega^{(\mu,\alpha)}(q|p_X)\right\} = \mathrm{E}_q\left[\frac{p_X^{\bar{\alpha}}(X)}{q_X^{\bar{\alpha}}(X)}\left\{\frac{p_X^{\mu\alpha}(X)q_{Y|U}^{\bar{\mu}\alpha}(Y|U)}{q_{X|U}^{\mu\alpha}(X|U)}\right\}\right]\\
&= \mathrm{E}_q\left[\left\{\frac{p_{UX}(U,X)}{q_{UX}(U,X)}\right\}^{\bar{\alpha}}\left\{\frac{p_X^{\mu\frac{\alpha}{\bar{\alpha}}}(X)q_{Y|U}^{\bar{\mu}\frac{\alpha}{\bar{\alpha}}}(Y|U)}{p_{X|U}^{\mu\frac{\alpha}{\bar{\alpha}}}(X|U)}\right\}^{\bar{\alpha}}\left\{\frac{p_{X|U}^{\mu}(X|U)}{q_{X|U}^{\mu}(X|U)}\right\}^{\alpha}\right]\\
&\overset{(a)}{\leq}\left(\mathrm{E}_q\left[\frac{p_{UX}(U,X)}{q_{UX}(U,X)}\frac{p_X^{\mu\frac{\alpha}{\bar{\alpha}}}(X)q_{Y|U}^{\bar{\mu}\frac{\alpha}{\bar{\alpha}}}(Y|U)}{p_{X|U}^{\mu\frac{\alpha}{\bar{\alpha}}}(X|U)}\right]\right)^{\bar{\alpha}}\left(\mathrm{E}_q\left[\frac{p_{X|U}^{\mu}(X|U)}{q_{X|U}^{\mu}(X|U)}\right]\right)^{\alpha}\\
&= \exp\left\{-\bar{\alpha}\hat{\Omega}^{(\mu,\frac{\alpha}{\bar{\alpha}})}(p,q_{Y|U})\right\}A^{\alpha},
\end{aligned} \tag{A17}$$

where we set

$$A := \mathrm{E}_q\left[\frac{p_{X|U}^{\mu}(X|U)}{q_{X|U}^{\mu}(X|U)}\right].$$

Step (a) follows from Hölder's inequality. From (A17), we can see that it suffices to show $A\leq 1$ to complete the proof. When $\mu=1$, we have $A=1$. When $\mu\in[0,1)$, we apply Hölder's inequality to A to obtain

$$A = \mathrm{E}_q\left[\frac{p_{X|U}^{\mu}(X|U)}{q_{X|U}^{\mu}(X|U)}\right] \leq \left(\mathrm{E}_q\left[\frac{p_{X|U}(X|U)}{q_{X|U}(X|U)}\right]\right)^{\mu} = 1.$$

Hence, we have (A13) in Lemma A4. □

Proof of Lemma A5: We fix $\mu\in[0,1]$, $\alpha\in[0,1/2)$, arbitrary. For any $p=p_{UXY}\in\mathcal{P}_{\mathrm{sh}}(p_{XY})$, and any $q=q_{UXY}\in\mathcal{Q}(p_{Y|X})$, we have the following chain of inequalities:

$$\exp\left\{-\hat{\Omega}^{(\mu,\frac{\alpha}{\bar{\alpha}})}(p,q_{Y|U})\right\} = \mathrm{E}_p\left[\left\{\frac{p_X^{\mu\frac{\alpha}{1-2\alpha}}(X)p_{Y|U}^{\bar{\mu}\frac{\alpha}{1-2\alpha}}(Y|U)}{p_{X|U}^{\mu\frac{\alpha}{1-2\alpha}}(X|U)}\right\}^{\frac{1-2\alpha}{\bar{\alpha}}}\left\{\frac{q_{Y|U}^{\bar{\mu}}(Y|U)}{p_{Y|U}^{\bar{\mu}}(Y|U)}\right\}^{\frac{\alpha}{\bar{\alpha}}}\right]$$

$$\overset{(a)}{\leq} \exp\left\{-\frac{1-2\alpha}{\bar{\alpha}}\tilde{\Omega}^{(\mu,\frac{\alpha}{1-2\alpha})}(p)\right\}\left(\mathrm{E}_p\left[\frac{q_{Y|U}^{\bar{\mu}}(Y|U)}{p_{Y|U}^{\bar{\mu}}(Y|U)}\right]\right)^{\frac{\alpha}{\bar{\alpha}}} = \exp\left\{-\frac{1-2\alpha}{\bar{\alpha}}\tilde{\Omega}^{(\mu,\frac{\alpha}{1-2\alpha})}(p)\right\}B^{\frac{\alpha}{\bar{\alpha}}}, \quad \text{(A18)}$$

where we set

$$B := \mathrm{E}_q\left[\frac{q_{Y|U}^{\bar{\mu}}(Y|U)}{p_{Y|U}^{\bar{\mu}}(Y|U)}\right].$$

Step (a) follows from Hölder's inequality. From (A18), we can see that it suffices to show $B \leq 1$ to complete the proof. In a manner quite smilar to the proof of $A \leq 1$ in the proof of (A13) in Lemma A4, we can show that $B \leq 1$. Thus, we have (A14) in Lemma A5. □

Proof of Property 4 part b: We evaluate lower bounds of $F(R_1,R_2|p_{XY})$ to obtain the following chain of inequalities:

$$F(R_1,R_2|p_{XY}) \overset{(a)}{\geq} \sup_{\substack{\mu\in[0,1],\\ \alpha\in[0,1/2)}} \frac{(1-2\alpha)\tilde{\Omega}^{(\mu,\frac{\alpha}{1-2\alpha})}(p_{XY}) - \alpha(\mu R_1+\bar{\mu}R_2)}{2+\alpha\bar{\mu}}$$

$$= \sup_{\substack{\mu\in[0,1],\\ \alpha\in[0,1/2),\\ \lambda=\frac{\alpha}{1-2\alpha}}} \frac{(1-2\alpha)\tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) - \alpha(\mu R_1+\bar{\mu}R_2)}{2+\alpha\bar{\mu}}$$

$$\overset{(b)}{=} \sup_{\substack{\mu\in[0,1],\\ \alpha=\frac{\lambda}{1+2\lambda},\lambda\geq 0}} \frac{(1-2\alpha)\tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) - \alpha(\mu R_1+\bar{\mu}R_2)}{2+\alpha\bar{\mu}}$$

$$\overset{(c)}{=} \sup_{\mu\in[0,1],\lambda\geq 0} \frac{\tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) - \lambda(\mu R_1+\bar{\mu}R_2)}{2+\lambda(5-\mu)} = \sup_{\mu\in[0,1],\lambda\geq 0} \underline{F}^{(\mu,\alpha)}(\mu R_1+\bar{\mu}R_2|p_{XY}). \quad \text{(A19)}$$

Step (a) follows from the definition of $F(R_1,R_2|p_{XY})$ and (A16) in Corollary A1. Steps (b) and (c) follow from that

$$\alpha\in[0,1/2), \lambda = \frac{\alpha}{1-2\alpha} \Leftrightarrow \lambda\geq 0, \alpha = \frac{\lambda}{1+2\lambda}.$$

From (A19), we have

$$F(R_1,R_2|p_{XY}) \geq \sup_{\mu\in[0,1],\lambda\geq 0} \underline{F}^{(\mu,\lambda)}(\mu R_1+\bar{\mu}R_2|p_{XY}) = \underline{F}(R_1,R_2|p_{XY}),$$

completing the proof. □

Appendix E. Proof of Property 4 Parts c, d, e, and f

In this appendix, we prove Property 4 parts c, d, e, and f. We first prove part c and then prove parts d and e. We finally prove part f.

Proof of Property 4 part c: We first prove the second inequality in (8) in part c. We first observe that

$$\exp[-\tilde{\Omega}^{(\mu,\lambda)}(p)] = \mathrm{E}_p\left[\frac{p_X^{\mu\lambda}(X)p_{Y|U}^{\bar{\mu}\lambda}(Y|U)}{p_{X|U}^{\mu\lambda}(X|U)}\right]. \tag{A20}$$

Let $\bar{p}_X$ be the uniform distribution on $\mathcal{X}$ and let $\bar{p}_Y$ be the uniform distribution on $\mathcal{Y}$. On lower bound of $\exp[-\tilde{\Omega}^{(\mu,\lambda)}(p)]$ for $p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$ and $(\mu,\lambda) \in [0,1]^2$, we have the following chain of inequalities:

$$\begin{aligned}
\exp[-\tilde{\Omega}^{(\mu,\lambda)}(p)] &= \frac{1}{|\mathcal{X}|^{\mu\lambda}|\mathcal{Y}|^{\bar{\mu}\lambda}}\mathrm{E}_p\left[p_{X|U}^{-\mu\lambda}(X|U)\left\{\frac{p_X(X)}{\bar{p}_X(X)}\right\}^{\mu\lambda}\left\{\frac{p_{Y|U}(Y|U)}{\bar{p}_Y(Y)}\right\}^{\bar{\mu}\lambda}\right] \\
&\overset{(a)}{\geq} \frac{1}{|\mathcal{X}|^{\mu}|\mathcal{Y}|^{\bar{\mu}}}\mathrm{E}_p\left[\left\{\frac{\bar{p}_X(X)}{p_X(X)}\right\}^{-\mu\lambda}\left\{\frac{\bar{p}_Y(Y)}{p_{Y|U}(Y|U)}\right\}^{-\bar{\mu}\lambda}\right] \\
&\overset{(b)}{\geq} \frac{1}{|\mathcal{X}|^{\mu}|\mathcal{Y}|^{\bar{\mu}}}\left(\mathrm{E}_p\left[\frac{\bar{p}_X(X)}{p_X(X)}\right]\right)^{-\mu\lambda}\left(\mathrm{E}_p\left[\frac{\bar{p}_Y(Y)}{p_{Y|U}(Y|U)}\right]\right)^{-\bar{\mu}\lambda} = \frac{1}{|\mathcal{X}|^{\mu}|\mathcal{Y}|^{\bar{\mu}}}.
\end{aligned} \tag{A21}$$

Step (a) follows from that $\lambda \in [0,1]$ and $p_{X|U}(x|u) \leq 1$ for any $(u,x) \in \mathcal{U} \times \mathcal{X}$. Step (b) follows from the reverse Hölder's inequality. The bound (A21) implies the second inequality in (8). We next show that $\tilde{\Omega}^{(\mu,\lambda)}(p) \geq 0$ for $\lambda \in [0,1]$. On upper bounds of $\exp[-\tilde{\Omega}^{(\mu,\lambda)}(p)]$ for $p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$ and $\lambda \in [0,1]$, we have the following chain of inequalities:

$$\exp[-\tilde{\Omega}^{(\mu,\lambda)}(p)] \overset{(a)}{\leq} \mathrm{E}_p\left[\left\{\frac{p_X(X)}{p_{X|U}(X|U)}\right\}^{\mu\lambda}\right] \overset{(b)}{\leq} \left\{\mathrm{E}_p\left[\frac{p_X(X)}{p_{X|U}(X|U)}\right]\right\}^{\mu\lambda} = 1. \tag{A22}$$

Step (a) follows from (A20) and $p_{Y|U}(y|u) \leq 1$ for any $(u,y) \in \mathcal{U} \times \mathcal{Y}$. Step (b) follows from $\mu\lambda \in [0,1]$ and Hölder's inequality. □

Proof of Property 4 parts d and e: We first prove that, for each $p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$ and $\mu \in [0,1]$, $\tilde{\Omega}^{(\mu,\lambda)}(p)$ is twice differentiable for $\lambda \in [0,1/2]$. For simplicity of notations, set

$$\begin{aligned}
&\underline{a} := (u,x,y), \underline{A} := (U,X,Y), \underline{\mathcal{A}} := \mathcal{U} \times \mathcal{X} \times \mathcal{Y}, \\
&\tilde{\omega}_p^{(\mu)}(x,y|u) := \varsigma(\underline{a}), \tilde{\Omega}^{(\mu,\lambda)}(p) := \xi(\lambda).
\end{aligned}$$

Then, we have

$$\tilde{\Omega}^{(\mu,\lambda)}(p) = \xi(\lambda) = -\log\left[\sum_{\underline{a}\in\underline{\mathcal{A}}} p_{\underline{A}}(\underline{a})\mathrm{e}^{-\lambda\varsigma(\underline{a})}\right]. \tag{A23}$$

The quantity $p^{(\lambda)}(\underline{a}) = p_{\underline{A}}^{(\lambda)}(\underline{a}), \underline{a} \in \mathcal{A}$ has the following form:

$$p^{(\lambda)}(\underline{a}) = \mathrm{e}^{\xi(\lambda)}p(\underline{a})\mathrm{e}^{-\lambda\varsigma(\underline{a})}. \tag{A24}$$

By simple computations, we have

$$\xi'(\lambda) = \mathrm{e}^{\xi(\lambda)} \left[\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma(\underline{a}) \mathrm{e}^{-\lambda \varsigma(\underline{a})} \right] = \sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma(\underline{a}),$$

$$\xi''(\lambda) = -\mathrm{e}^{2\xi(\lambda)} \left[\sum_{\underline{a},\underline{b} \in \underline{\mathcal{A}}} p(\underline{a}) p_{\underline{A}}(\underline{b}) \frac{\{\varsigma(\underline{a}) - \varsigma(\underline{b})\}^2}{2} \mathrm{e}^{-\lambda\{\varsigma(\underline{a}) + \varsigma(\underline{b})\}} \right]$$

$$= -\sum_{\underline{a},\underline{b} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) p^{(\lambda)}(\underline{b}) \frac{\{\varsigma(\underline{a}) - \varsigma(\underline{b})\}^2}{2} = -\sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma^2(\underline{a}) + \left[\sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma(\underline{a}) \right]^2 \leq 0. \tag{A25}$$

On upper bound of $-\xi''(\lambda) \geq 0$ for $\lambda \in [0, 1/2]$, we have the following chain of inequalities:

$$-\xi''(\lambda) \overset{\text{(a)}}{\leq} \sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma^2(\underline{a}) \overset{\text{(b)}}{=} \sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^2(\underline{a}) \mathrm{e}^{-\lambda \varsigma(\underline{a}) + \xi(\lambda)} = \mathrm{e}^{\xi(\lambda)} \sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \sqrt{\mathrm{e}^{-2\lambda \varsigma(\underline{a})}} \sqrt{\varsigma^4(\underline{a})}$$

$$\overset{\text{(c)}}{\leq} \sqrt{\mathrm{e}^{2\xi(\lambda) - \xi(2\lambda)}} \sqrt{\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^4(\underline{a})} \overset{\text{(d)}}{\leq} \sqrt{\mathrm{e}^{2\xi(\lambda)}} \sqrt{\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^4(\underline{a})}. \tag{A26}$$

Step (a) follows from (A25). Step (b) follows from (A24). Step (c) follows from Cauchy–Schwarz inequality and (A23). Step (d) follows from that $\xi(2\lambda) \geq 0$ for $2\lambda \in [0, 1]$. Note that $\xi(\lambda)$ exists for $\lambda \in [0, 1/2]$. Furthermore, we have the following:

$$\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^4(\underline{a}) < \infty.$$

Hence, by (A26), $\xi''(\lambda)$ exists for $\lambda \in [0, 1/2]$. We next prove part e. We derive the lower bound (9) of $\tilde{\Omega}^{(\mu,\lambda)}(p_{XY})$. Fix any $(\mu, \lambda) \in [0, 1] \times [0, 1/2]$ and any $p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$. By the Taylor expansion of $\xi(\lambda) = \tilde{\Omega}^{(\mu,\lambda)}(p)$ with respect to λ around $\lambda = 0$, we have that for any $p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$ and for some $\nu \in [0, \lambda]$

$$\tilde{\Omega}^{(\mu,\lambda)}(p) = \xi(0) + \xi'(0)\lambda + \frac{1}{2}\xi''(\nu)\lambda^2 = \lambda \mathrm{E}_p \left[\tilde{\omega}_p^{(\mu)}(X, Y|U) \right] - \frac{\lambda^2}{2} \mathrm{Var}_{p^{(\nu)}} \left[\tilde{\omega}_p^{(\mu)}(X, Y|U) \right]$$

$$\overset{\text{(a)}}{\geq} \lambda R^{(\mu)}(p_{XY}) - \frac{\lambda^2}{2} \mathrm{Var}_{p^{(\nu)}} \left[\tilde{\omega}_p^{(\mu)}(X, Y, Z|U) \right]. \tag{A27}$$

Step (a) follows from $p \in \mathcal{P}_{\mathrm{sh}}(p_{XY})$,

$$\mathrm{E}_p \left[\tilde{\omega}_p^{(\mu)}(X, Y|U) \right] = \mu I_p(X; U) + \bar{\mu} H_p(Y|U),$$

and the definition of $R^{(\mu)}(p_{XY})$. Let $(\nu_{\mathrm{opt}}, p_{\mathrm{opt}}) \in [0, \lambda] \times \mathcal{P}_{\mathrm{sh}}(p_{XY})$ be a pair which attains $\rho^{(\mu,\lambda)}(p_{XY})$. By this definition, we have that

$$\tilde{\Omega}^{(\mu,\lambda)}(p_{\mathrm{opt}}) = \tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) \tag{A28}$$

and that, for any $\nu \in [0, \lambda]$,

$$\mathrm{Var}_{p_{\mathrm{opt}}^{(\nu)}} \left[\omega_{p_{\mathrm{opt}}}^{(\mu)}(X, Y|U) \right] \leq \mathrm{Var}_{p_{\mathrm{opt}}^{(\nu_{\mathrm{opt}})}} \left[\omega_{p_{\mathrm{opt}}}^{(\mu)}(X, Y|U) \right] = \rho^{(\mu,\lambda)}(p_{XY}). \tag{A29}$$

On lower bounds of $\Omega^{(\mu,\lambda)}(p_{XY})$, we have the following chain of inequalities:

$$\tilde{\Omega}^{(\mu,\lambda)}(p_{XY}) \overset{\text{(a)}}{=} \tilde{\Omega}^{(\mu,\lambda)}(p_{\text{opt}}) \overset{\text{(b)}}{\geq} \lambda R^{(\mu)}(p_{XY}) - \frac{\lambda^2}{2}\text{Var}_{p^{(\nu)}_{\text{opt}}}\left[\tilde{\omega}^{(\mu)}_{p_{\text{opt}}}(X,Y|U)\right]$$
$$\overset{\text{(c)}}{\geq} \lambda R^{(\mu)}(p_{XY}) - \frac{\lambda^2}{2}\rho^{(\mu,\lambda)}(p_{XY}) \overset{\text{(d)}}{\geq} \lambda R^{(\mu)}(p_{XY}) - \frac{\lambda^2}{2}\rho(p_{XY}).$$

Step (a) follows from (A28). Step (b) follows from (A27). Step (c) follows from (A29). Step (d) follows from the definition of $\rho(p_{XY})$. □

To prove part f, we use the following lemma.

Lemma A6. *When $\tau \in (0,(1/2)\rho]$, the maximum of*

$$\frac{1}{2+5\lambda}\left\{-\frac{\rho}{2}\lambda^2+\tau\lambda\right\}$$

for $\lambda \in (0,1/2]$ is attained by the positive λ_0 satisfying

$$\vartheta(\lambda_0) := \lambda_0 + \frac{5}{4}\lambda_0^2 = \frac{\tau}{\rho}. \tag{A30}$$

Let $g(a)$ be the inverse function of $\vartheta(a)$ for $a \geq 0$. Then, the condition of (A30) is equivalent to $\lambda_0 = g(\frac{\tau}{\rho})$. The maximum is given by

$$\frac{1}{2+5\lambda_0}\left\{-\frac{\rho}{2}\lambda_0^2+\tau\lambda_0\right\} = \frac{\rho}{4}\lambda_0^2 = \frac{\rho}{4}g^2\left(\frac{\tau}{\rho}\right).$$

By an elementary computation, we can prove this lemma. We omit the detail.

Proof of Property 4 part f: By the hyperplane expression $\mathcal{R}_{\text{sh}}(p_{XY})$ of $\mathcal{R}(p_{XY})$ stated Property 3 part b, we have that, when $(R_1+\tau, R_2+\tau) \notin \mathcal{R}(p_{XY})$, we have

$$R^{(\mu_0)}(p_{XY}) - (\mu_0 R_1 + \overline{\mu_0} R_2) > \tau \tag{A31}$$

for some $\mu_0 \in [0,1]$. Then, for each positive τ, we have the following chain of inequalities:

$$\underline{F}(R_1,R_2|p_{XY}) \geq \sup_{\lambda\in(0,1/2]} \underline{F}^{(\mu_0,\lambda)}(\mu_0 R_1 + \overline{\mu_0} R_2|p_{XY}) = \sup_{\lambda\in(0,1/2]} \frac{\tilde{\Omega}^{(\mu_0,\lambda)}(p_{XY}) - \lambda(\mu_0 R_1 + \overline{\mu_0} R_2)}{2+\lambda(5-\mu_0)}$$
$$\overset{\text{(a)}}{\geq} \sup_{\lambda\in(0,1/2]} \frac{1}{2+5\lambda}\left\{-\frac{\rho}{2}\lambda^2 + \lambda R^{(\mu_0)}(p_{XY}) - \lambda(\mu_0 R_1 + \overline{\mu_0} R_2)\right\}$$
$$\overset{\text{(b)}}{>} \sup_{\lambda\in(0,1/2]} \frac{1}{2+5\lambda}\left\{-\frac{\rho}{2}\lambda^2+\tau\lambda\right\} \overset{\text{(c)}}{=} \frac{\rho}{4}g^2\left(\frac{\tau}{\rho}\right).$$

Step (a) follows from Property 4 part d. Step (b) follows from (A31). Step (c) follows from Lemma A6. □

Appendix F. Proof of Lemma 1

To prove Lemma 1, we prepare a lemma. Set

$$\mathcal{A}_n := \left\{(s,x^n,y^n) : \frac{1}{n}\log\frac{p_{SX^nY^n}(s,x^n,y^n)}{\hat{q}_{SX^nY^n}(s,x^n,y^n)} \geq -\eta\right\}.$$

Furthermore, set

$$\tilde{\mathcal{B}}_n := \left\{ x^n : \frac{1}{n} \log \frac{p_{X^n}(x^n)}{Q_{X^n}(x^n)} \geq -\eta \right\}, \mathcal{B}_n := \tilde{\mathcal{B}}_n \times \mathcal{M}_1 \times \mathcal{Y}^n, \mathcal{B}_n^{\mathrm{c}} := \tilde{\mathcal{B}}_n^{\mathrm{c}} \times \mathcal{M}_1 \times \mathcal{Y}^n,$$
$$\tilde{\mathcal{C}}_n := \{(s, x^n) : s = \varphi_1^{(n)}(x^n), \tilde{Q}_{X^n|S}(x^n|s) \leq M_1 \mathrm{e}^{n\eta} p_{X^n}(x^n)\}, \mathcal{C}_n := \tilde{\mathcal{C}}_n \times \mathcal{Y}^n, \mathcal{C}_n^{\mathrm{c}} := \tilde{\mathcal{C}}_n^{\mathrm{c}} \times \mathcal{Y}^n,$$
$$\mathcal{D}_n := \{(s, x^n, y^n) : s = \varphi_1^{(n)}(x^n), p_{Y^n|S}(y^n|s) \geq (1/M_2)\mathrm{e}^{-n\eta}\},$$
$$\mathcal{E}_n := \{(s, x^n, y^n) : s = \varphi_1^{(n)}(x^n), \psi^{(n)}(\varphi_1^{(n)}(x^n), \varphi_2^{(n)}(y^n)) = y^n\}.$$

Then, we have the following lemma.

Lemma A7.

$$p_{SX^nY^n}(\mathcal{A}_n^{\mathrm{c}}) \leq \mathrm{e}^{-n\eta}, p_{SX^nY^n}(\mathcal{B}_n^{\mathrm{c}}) \leq \mathrm{e}^{-n\eta}, p_{SX^nY^n}(\mathcal{C}_n^{\mathrm{c}}) \leq \mathrm{e}^{-n\eta}, p_{SX^nY^n}(\mathcal{D}_n^{\mathrm{c}} \cap \mathcal{E}_n) \leq \mathrm{e}^{-n\eta}.$$

Proof. We first prove the first inequality.

$$p_{SX^nY^n}(\mathcal{A}_n^{\mathrm{c}}) = \sum_{(s,x^n,y^n) \in \mathcal{A}_n^{\mathrm{c}}} p_{SX^nY^n}(s, x^n, y^n)$$
$$\overset{\text{(a)}}{\leq} \sum_{(s,x^n,y^n) \in \mathcal{A}_n^{\mathrm{c}}} \mathrm{e}^{-n\eta} \hat{q}_{SX^nY^n}(s, x^n, y^n) = \mathrm{e}^{-n\eta} \hat{q}_{SX^nY^n}(\mathcal{A}_n^{\mathrm{c}}) \leq \mathrm{e}^{-n\eta}.$$

Step (a) follows from the definition of $\mathcal{A}_n$. In the second inequality, we have

$$p_{SX^nY^n}(\mathcal{B}_n^{\mathrm{c}}) = p_{X^n}(\tilde{\mathcal{B}}_n^{\mathrm{c}}) = \sum_{x^n \in \tilde{\mathcal{B}}_n^{\mathrm{c}}} p_{X_n}(x^n) \overset{\text{(a)}}{\leq} \sum_{x^n \in \tilde{\mathcal{B}}_n^{\mathrm{c}}} \mathrm{e}^{-n\eta} Q_{X^n}(x^n) = \mathrm{e}^{-n\eta} Q_{X^n}(\tilde{\mathcal{B}}_n^{\mathrm{c}}) \leq \mathrm{e}^{-n\eta}.$$

Step (a) follows from the definition of $\mathcal{B}_n$. We next prove the third inequality:

$$p_{SX^nY^n}(\mathcal{C}_n^{\mathrm{c}}) = p_{SX^n}(\tilde{\mathcal{C}}_n^{\mathrm{c}}) = \sum_{s \in \mathcal{M}_1} \sum_{\substack{x^n : \varphi_1^{(n)}(x^n) = s \\ p_{X^n}(x^n) \leq (1/M_1)\mathrm{e}^{-n\eta} \\ \times \tilde{Q}_{X^n|S}(x^n|s)}} p_{X^n}(x^n)$$
$$\leq \frac{1}{M_1} \mathrm{e}^{-n\eta} \sum_{s \in \mathcal{M}_1} \sum_{\substack{x^n : \varphi_1^{(n)}(x^n) = s \\ p_{X^n}(x^n) \leq (1/M_1)\mathrm{e}^{-n\eta} \\ \times \tilde{Q}_{X^n|S}(x^n|s)}} \tilde{Q}_{X^n|S}(x^n|s) \leq \frac{1}{M_1} \mathrm{e}^{-n\eta} |\mathcal{M}_1| = \mathrm{e}^{-n\eta}.$$

Finally, we prove the fourth inequality. We first observe that

$$p_S(s) = \sum_{x^n : \varphi_1^{(n)}(x^n) = s} p_{X^n}(x^n), \; p_{X^n|S}(x^n|s) = \frac{p_{X^n}(x^n)}{p_S(s)}.$$

We have the following chain of inequalities:

$$p_{SX^nY^n}\left(\mathcal{D}_n^{\mathrm{c}} \cap \mathcal{E}_n\right) = \sum_{s\in\mathcal{M}_1} p_S(s) \sum_{x^n:\varphi_1^{(n)}(x^n)=s} p_{X^n|S}(x^n|s) \sum_{\substack{y^n:\psi^{(n)}(s,\varphi_2^{(n)}(y^n))=y^n \\ p_{Y^n|S}(y^n|s)\le(1/M_2)\mathrm{e}^{-n\eta}}} p_{Y^n|X^n}(y^n|x^n)$$

$$= \sum_{s\in\mathcal{M}_1} p_S(s) \sum_{\substack{y^n:\psi^{(n)}(s,\varphi_2^{(n)}(y^n))=y^n \\ p_{Y^n|S}(y^n|s)\le(1/M_2)\mathrm{e}^{-n\eta}}} p_{Y^n|S}(y^n|s)$$

$$\le \sum_{s\in\mathcal{M}_1} p_S(s)\frac{1}{M_2}\mathrm{e}^{-n\eta}\left|\left\{y^n : \psi^{(n)}(s,\varphi_2^{(n)}(y^n)) = y^n\right\}\right| \overset{\text{(a)}}{\le} \sum_{s\in\mathcal{M}_1} p_S(s)\frac{1}{M_2}\mathrm{e}^{-n\eta}M_2 = \mathrm{e}^{-n\eta}.$$

Step (a) follows from that the number of y^n correctly decoded does not exceed M_2. □

Proof of Lemma 1: By definition, we have

$$p_{SX^nY^n}\left(\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\right) = p_{SX^nY^n}\left\{\begin{array}{r} \frac{1}{n}\log\frac{p_{SX^nY^n}(S,X^n,Y^n)}{\hat{q}_{SX^nY^n}(S,X^n,Y^n)} \ge -\eta, \\ 0 \ge \frac{1}{n}\log\frac{Q_{X^n}(X^n)}{p_{X^n}(X^n)} - \eta, \\ \frac{1}{n}\log M_1 \ge \frac{1}{n}\log\frac{\tilde{Q}_{X^n|S}(X^n|S)}{p_{X^n}(X^n)} - \eta, \\ \frac{1}{n}\log M_2 \ge \frac{1}{n}\log\frac{1}{p_{Y^n|S}(Y^n|S)} - \eta \end{array}\right\}.$$

Then, for any $(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)})$ satisfying $(1/n)\log||\varphi_i^{(n)}|| \le R_i, i = 1, 2$, we have

$$p_{SX^nY^n}\left(\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\right) \le p_{SX^nY^n}\left\{\begin{array}{r} \frac{1}{n}\log\frac{p_{SX^nY^n}(S,X^n,Y^n)}{\hat{q}_{SX^nY^n}(S,X^n,Y^n)} \ge -\eta, \\ 0 \ge \frac{1}{n}\log\frac{Q_{X^n}(X^n)}{p_{X^n}(X^n)} - \eta, \\ R_1 \ge \frac{1}{n}\log\frac{\tilde{Q}_{X^n|S}(X^n|S)}{p_{X^n}(X^n)} - \eta, \\ R_2 \ge \frac{1}{n}\log\frac{1}{p_{Y^n|S}(Y^n|S)} - \eta \end{array}\right\}.$$

Hence, it suffices to show

$$\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) \le p_{SX^nY^n}\left(\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\right) + 4\mathrm{e}^{-n\eta}$$

to prove Lemma 1. By definition, we have $\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) = p_{SX^nY^n}\left(\mathcal{E}_n\right)$. Then, we have the following.

$$\begin{aligned}
&\mathrm{P}_{\mathrm{c}}^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}) = p_{SX^nY^n}\left(\mathcal{E}_n\right) \\
&= p_{SX^nY^n}\left(\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n\right) + p_{SX^nY^n}\left(\left[\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\right]^{\mathrm{c}} \cap \mathcal{E}_n\right) \\
&\le p_{SX^nY^n}\left(\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\right) + p_{SX^nY^n}\left(\mathcal{A}_n^{\mathrm{c}}\right) + p_{SX^nY^n}\left(\mathcal{B}_n^{\mathrm{c}}\right) + p_{SX^nY^n}\left(\mathcal{C}_n^{\mathrm{c}}\right) + p_{SX^nY^n}\left(\mathcal{D}_n^{\mathrm{c}} \cap \mathcal{E}_n\right) \\
&\overset{\text{(a)}}{\le} p_{SX^nY^n}\left(\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\right) + 4\mathrm{e}^{-n\eta}.
\end{aligned}$$

Step (a) follows from Lemma A7. □

Appendix G. Proof of Lemma 3

In this appendix, we prove Lemma 3.

Proof of Lemma 3: We first prove the Markov chain $SX^{t-1} \leftrightarrow X_t \leftrightarrow Y_t$ in (18) in Lemma 3. We have the following chain of inequalities:

$$\begin{aligned} &I(Y_t; SX^{t-1}|X_t) = H(Y_t|X_t) - H(Y_t|SX^{t-1}X_t) \le H(Y_t|X_t) - H(Y_t|SX^n) \\ &\overset{\text{(a)}}{=} H(Y_t|X_t) - H(Y_t|X^n) \overset{\text{(b)}}{=} H(Y_t|X_t) - H(Y_t|X_t) = 0. \end{aligned}$$

Step (a) follows from that $S = \varphi_1^{(n)}(X^n)$ is a function of X^n. Step (b) follows from the memoryless property of the information source $\{(X_t, Y_t)\}_{t=1}^{\infty}$. Next, we prove the Markov chain $Y^{t-1} \leftrightarrow SX^{t-1} \leftrightarrow (X_t, Y_t)$ in (19) in Lemma 3. We have the following chain of inequalities:

$$\begin{aligned} &I(X_tY_t; Y^{t-1}|SX^{t-1}) = H(Y^{t-1}|SX^{t-1}) - H(Y^{t-1}|SX^{t-1}X_tY_t) \le H(Y^{t-1}|X^{t-1}) - H(Y^{t-1}|X^nSY_t) \\ &\overset{\text{(a)}}{=} H(Y^{t-1}|X^{t-1}) - H(Y^{t-1}|X^nY_t) \overset{\text{(b)}}{=} H(Y^{t-1}|X^{t-1}) - H(Y^{t-1}|X^{t-1}Y_t) = 0. \end{aligned}$$

Step (a) follows from that $S = \varphi_1^{(n)}(X^n)$ is a function of X^n. Step (b) follows from the memoryless property of the information source $\{(X_t, Y_t)\}_{t=1}^{\infty}$. □

Appendix H. Proof of Lemma 6

In this appendix, we prove Lemma 6.

Proof of Lemma 6. By the definition of $p^{(\mu,\alpha)}_{SX^tY^t;\mathcal{F}^t}(s, x^t, y^t)$, for $t = 1, 2, \cdots, n$, we have

$$p^{(\mu,\alpha)}_{SX^tY^t;\mathcal{F}^t}(s, x^t, y^t) = C_t^{-1} p_{SX^tY^t}(s, x^t, y^t) \prod_{i=1}^{t} f^{(\mu,\alpha)}_{\mathcal{F}_i}(x_i, y_i|u_i). \tag{A32}$$

Then, we have the following chain of equalities:

$$\begin{aligned} &p^{(\mu,\alpha)}_{SX^tY^t;\mathcal{F}^t}(s, x^t, y^t) \overset{\text{(a)}}{=} C_t^{-1} p_{SX^tY^t}(s, x^t, y^t) \prod_{i=1}^{t} f^{(\mu,\alpha)}_{\mathcal{F}_i}(x_i, y_i|u_i) \\ &= C_t^{-1} p_{SX^{t-1}Y^{t-1}}(s, x^{t-1}, y^{t-1}) \prod_{i=1}^{t-1} f^{(\mu,\alpha)}_{\mathcal{F}_i}(x_i, y_i|u_i) \\ &\quad \times p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f^{(\mu,\alpha)}_{\mathcal{F}_t}(x_t, y_t|u_t) \\ &\overset{\text{(b)}}{=} C_t^{-1} C_{t-1} p^{(\mu,\alpha)}_{SX^{t-1}Y^{t-1}}(s, x^{t-1}, y^{t-1}) p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f^{(\mu,\alpha)}_{\mathcal{F}_t}(x_t, y_t|u_t) \\ &= (\Phi_t^{(\mu,\alpha)})^{-1} p^{(\mu,\alpha)}_{SX^{t-1}Y^{t-1};\mathcal{F}^{t-1}}(s, x^{t-1}, y^{t-1}) p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f^{(\mu,\alpha)}_{\mathcal{F}_t}(x_t, y_t|u_t). \end{aligned} \tag{A33}$$

Steps (a) and (b) follow from (A32). From (A33), we have

$$\Phi_t^{(\mu,\alpha)} p^{(\mu,\alpha)}_{SX^tY^t;\mathcal{F}^t}(s, x^t, y^t) \tag{A34}$$

$$= p^{(\mu,\alpha)}_{SX^{t-1}Y^{t-1};\mathcal{F}^{t-1}}(s, x^{t-1}, y^{t-1}) p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f^{(\mu,\alpha)}_{\mathcal{F}_t}(x_t, y_t|u_t). \tag{A35}$$

Taking summations of (A34) and (A35) with respect to s, x^t, y^t, we obtain

$$\Phi_t^{(\mu,\alpha)} = \sum_{s,x^t,y^t} p^{(\mu,\alpha)}_{SX^{t-1}Y^{t-1};\mathcal{F}^{t-1}}(s, x^{t-1}, y^{t-1}) p_{X_tY_t|SX^{t-1}Y^{t-1}}(x_t, y_t|s, x^{t-1}, y^{t-1}) f^{(\mu,\alpha)}_{\mathcal{F}_t}(x_t, y_t|u_t),$$

completing the proof. □

References

1. Ahlswede, R.F.; Körner, J. Source coding with side information and a converse for degraded broadcast channels. *IEEE Trans. Inf. Theory* **1975**, *21*, 629–637.
2. Wyner, A.D. On source coding with side information at the decoder. *IEEE Trans. Inf. Theory* **1975**, *21*, 294–300.
3. Csiszár, I.; Longo, G. On the exponent function for source coding and for testing simple statistical hypotheses. *Studia Sci. Math. Hungar* **1971**, *6*, 181–191
4. Slepian, D.; Wolf, J.K. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory* **1973**, *19*, 471–480.
5. Oohama, Y.; Han, T.S. Universal coding for the Slepian-wolf data compression system and the strong converse theorem. *IEEE Trans. Inf. Theory* **1994**, *40*, 1908–1919.
6. Ahlswede, R.; Gács, P.; Körner, J. Bounds on conditional probabilities with applications in multi-user communication. *Probab. Theory Relat. Fields* **1976**, *34*, 157–177.
7. Gu, W.; Effors, M. A strong converse for a collection of network source coding problems. In Proceedings of the IEEE International Symposium on Information Theory, Soul, Korea, 28 June–3 July 2009; pp. 2356–2320.
8. Oohama, Y. Strong converse exponent for degraded broadcast channels at rates outside the capacity region. In Proceedings of the 2015 IEEE International Symposium on Information Theory, Hong Kong, China, 14–19 June 2015; pp. 939–943.
9. Oohama, Y. Strong converse theorems for degraded broadcast channels with feedback. In Proceedings of the 2015 IEEE International Symposium on Information Theory. Hong Kong, China, 14–19 June 2015; pp. 2510– 2514.
10. Oohama, Y. Exponent function for asymmetric broadcast channels at rates outside the capacity region. In Proceedings of the 2016 IEEE International Symposium on Information Theory and its Applications. Monterey, CA, USA, 30 October–2 Novomber 2016; pp. 568–572.
11. Oohama, Y. New Strong Converse for Asymmetric Broadcast Channels. Available online: https://arxiv.org/pdf/1604.02901.pdf. (accessed on 31 May 2019)
12. Oohama, Y. Exponential strong converse for source coding with side information at the decoder. *Entropy* **2018**, *20*, 352.
13. Watanabe, S. A converse bound on Wyner-Ahlswede-Körner network via Gray–Wyner network. In Proceedings of the 2017 IEEE Information Theory Workshop (ITW), Kaohsiung, Taiwan, 6–10 November 2017; pp. 81–85.
14. Liu, J.; van Handel, R.; Verdu, S. Beyond the blowing-up lemma: Sharp converses via reverse hypercontractivity. In Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 943–947
15. Watanabe, S. Second-order region for Gray–Wyner network. *IEEE Trans. Inform. Theory* **2017**, *63*, 1006–1018.
16. Liu, J. Dispersion bound for the Wyner-Ahlswede-Körner network via reverse hypercontractivity on types. In Proceedings of the 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA, 17–22 June 2018; pp. 1854–1858.
17. Watanabe, S.; Oohama, Y. Privacy amplification theorem for bounded storage eavesdropper. In Proceedings of 2012 IEEE Information Theory Workshop (ITW), Lausanne, Switzerland, 3–7 September 2012; pp. 177–181.
18. Oohama, Y.; Santoso, B. Information Theoretic Security for Side-Channel Attacks to the Shannon Cipher System. Available online: https://arxiv.org/pdf/1801.02563v5.pdf. (accessed on 31 May 2019)
19. Santoso, B.; Oohama, Y. Information Theoretic Security for Shannon Cipher System under Side-Channel Attacks. *Entropy* **2019**, *21*, 469.
20. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Cambridge University Press: London, UK, 1981.
21. Oohama, Y. Exponent Function for One Helper Source Coding Problem at Rates outside the Rate Region. *Arxiv* **2015**, arXiv:1504.05891.
22. Han, T.S. *Information-Spectrum Methods in Information Theory*; Springer Nature Switzerland AG: Basel, Switzerland, 2002.