

Article

Exponential Strong Converse for Source Coding with Side Information at the Decoder [†]

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 2016 International Symposium on Information Theory and Its Applications, Monterey, CA, USA, 6–9 November 2016; pp. 171–175.

Received: 31 January 2018; Accepted: 20 April 2018; Published: 8 May 2018



Abstract: We consider the rate distortion problem with side information at the decoder posed and investigated by Wyner and Ziv. Using side information and encoded original data, the decoder must reconstruct the original data with an arbitrary prescribed distortion level. The rate distortion region indicating the trade-off between a data compression rate R and a prescribed distortion level Δ was determined by Wyner and Ziv. In this paper, we study the error probability of decoding for pairs of (R, Δ) outside the rate distortion region. We evaluate the probability of decoding such that the estimation of source outputs by the decoder has a distortion not exceeding a prescribed distortion level Δ . We prove that, when (R, Δ) is outside the rate distortion region, this probability goes to zero exponentially and derive an explicit lower bound of this exponent function. On the Wyner–Ziv source coding problem the strong converse coding theorem has not been established yet. We prove this as a simple corollary of our result.

Keywords: source coding with side information at the decoder; the rate distortion region; exponent function outside the rate distortion region; strong converse theorem

1. Introduction

For single or multi terminal source coding 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 *cannot go to zero* when the block length n of the codes tends to infinity. On the other hand, the strong converse theorems state that, at any transmission rates exceeding the fundamental theoretical limit, the error probability of decoding *must go to one* when n tends to infinity. The former converse theorems are sometimes called the weak converse theorems to distinguish them with the strong converse theorems.

In this paper, we study the strong converse theorem for the rate distortion problem with side information at the decoder posed and investigated by Wyner and Ziv [1]. We call the above source coding system the Wyner and Ziv source coding system (the WZ system). The WZ system is shown in Figure 1. In this figure, the WZ system corresponds to the case where the switch is close. 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} and \mathcal{Y} , respectively. We assume that (X, Y) has a probability distribution denoted by p_{XY} . The encoder $\varphi^{(n)}$ outputs a binary sequence which appears at a rate R bits per input symbol. The decoder function $\psi^{(n)}$ observes $\varphi^{(n)}(X^n)$ and Y^n to output a sequence Z^n . The t -th component Z_t of Z^n for $t = 1, 2, \dots, n$ take values in the finite reproduction alphabet \mathcal{Z} . Let $d : \mathcal{X} \times \mathcal{Z} \rightarrow [0, \infty)$ be an arbitrary distortion measure on $\mathcal{X} \times \mathcal{Z}$. The distortion between $x^n \in \mathcal{X}^n$ and $z^n \in \mathcal{Z}^n$ is defined by

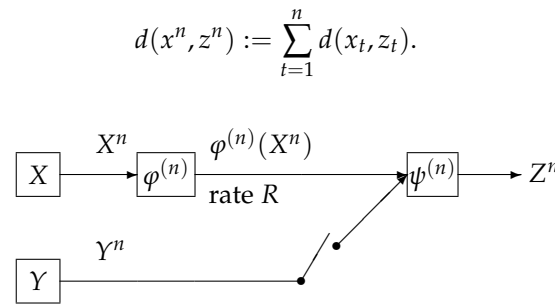


Figure 1. Source encoding with a fidelity criterion with or without side information at the decoder.

In general, we have two criteria on $d(X^n, Z^n)$. One is the excess-distortion probability of decoding defined by

$$P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) := \Pr \left\{ \frac{1}{n} d(X^n, Z^n) \geq \Delta \right\}. \quad (1)$$

The other is the average distortion defined by

$$\begin{aligned} \Delta_n &:= E \left[\frac{1}{n} d(X^n, Z^n) \right] := \sum_{(x^n, z^n) \in \mathcal{X}^n \times \mathcal{Z}^n} \left[\frac{1}{n} \sum_{k=1}^n d(x_k, z_k) \right] \Pr \{ X^n = x^n, Z^n = z^n \} \\ &= \frac{1}{n} \sum_{k=1}^n \sum_{(x_k, z_k) \in \mathcal{X} \times \mathcal{Z}} d(x_k, z_k) \Pr \{ X_k = x_k, Z_k = z_k \}. \end{aligned}$$

A pair (R, Δ) is ε -achievable for p_{XY} if there exist a sequence of pairs $\{(\varphi^{(n)}, \psi^{(n)})\}_{n \geq 1}$ such that for any $\delta > 0$ and any n with $n \geq n_0 = n_0(\varepsilon, \delta)$

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R + \delta, \quad P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq \varepsilon,$$

where $\|\varphi^{(n)}\|$ stands for the range of cardinality of $\varphi^{(n)}$. The rate distortion region $\mathcal{R}_{WZ}(\varepsilon|p_{XY})$ is defined by

$$\mathcal{R}_{WZ}(\varepsilon|p_{XY}) = \{ (R, \Delta) : (R, \Delta) \text{ is } \varepsilon\text{-achievable for } p_{XY} \}.$$

Furthermore, set

$$\mathcal{R}_{WZ}(p_{XY}) := \bigcap_{\varepsilon > 0} \mathcal{R}_{WZ}(\varepsilon|p_{XY}).$$

On the other hand, we can define a rate distortion region based on the average distortion criterion, a formal definition of which is the following. A pair (R, Δ) is *achievable* for p_{XY} if there exist a sequence of pairs $\{(\varphi^{(n)}, \psi^{(n)})\}_{n \geq 1}$ such that for any $\delta > 0$ and any n with $n \geq n_0 = n_0(\delta)$,

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R + \delta, \quad \Delta^{(n)} \leq \Delta + \delta.$$

The rate distortion region $\tilde{\mathcal{R}}_{WZ}(p_{XY})$ is defined by

$$\tilde{\mathcal{R}}_{WZ}(p_{XY}) := \{ (R, \Delta) : (R, \Delta) \text{ is achievable for } p_{XY} \}.$$

If the switch is open, then the side information is not available to the decoder. In this case the communication system corresponds to the source coding for the discrete memoryless source (DMS) specified with p_X . We define the rate distortion region $\tilde{\mathcal{R}}_{DMS}(p_X)$ in a similar manner to the definition

of $\tilde{\mathcal{R}}_{WZ}(p_{XY})$. We further define the region $\mathcal{R}_{DMS}(\varepsilon|p_X)$, $\varepsilon \in (0, 1)$ and $\mathcal{R}_{DMS}(p_X)$, respectively in a similar manner to the definition of $\mathcal{R}_{WZ}(\varepsilon|p_{XY})$ and $\mathcal{R}_{WZ}(p_{XY})$.

Previous works on the characterizations of $\tilde{\mathcal{R}}_{DMS}(p_X)$, $\mathcal{R}_{DMS}(\varepsilon|p_X)$, $\varepsilon \in (0, 1)$, and $\mathcal{R}_{DMS}(p_X)$ are shown in Table 1. Shannon [2] determined $\tilde{\mathcal{R}}_{DMS}(p_X)$. Subsequently, Wolfowitz [3] proved that $\tilde{\mathcal{R}}_{DMS}(p_X) = \mathcal{R}_{DMS}(p_X)$. Furthermore, he proved the strong converse theorem. That is, if $(R, \Delta) \notin \tilde{\mathcal{R}}_{DMS}(p_X)$, then for any sequence $\{(\varphi^{(n)}, \psi^{(n)})\}_{n=1}^{\infty}$ of encoder and decoder functions satisfying the condition

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\varphi^{(n)}\| \leq R, \quad (2)$$

we have

$$\lim_{n \rightarrow \infty} P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = \lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n_k} d(X^n, Z^n) \geq \Delta \right\} = 1. \quad (3)$$

The above strong converse theorem implies that, for any $\varepsilon \in (0, 1)$,

$$\tilde{\mathcal{R}}_{DMS}(p_X) = \mathcal{R}_{DMS}(p_X) = \mathcal{R}_{DMS}(\varepsilon|p_X).$$

Csiszár and Körner proved that in Equation (3), the probability $P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta)$ converges to one exponentially and determined the optimal exponent as a function of (R, Δ) .

The previous works on the coding theorems for the WZ system are summarized in Table 1. The rate distortion region $\tilde{\mathcal{R}}_{WZ}(p_{XY})$ was determined by Wyner and Ziv [1]. Csiszár and Körner [4] proved that $\tilde{\mathcal{R}}_{WZ}(p_{XY}) = \mathcal{R}_{WZ}(p_{XY})$. On the other hand, we have had no result on the strong converse theorem for the WZ system.

Main results of this paper are summarized in Table 1. For the WZ system, we prove that if (R, Δ) is out side the rate distortion region $\mathcal{R}_{WZ}(p_{XY})$, then we have that for any sequence $\{(\varphi^{(n)}, \psi^{(n)})\}_{n=1}^{\infty}$ of encoder and decoder functions satisfying the condition in Equation (2), the quantity $P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta)$ goes to zero exponentially and derive an explicit lower bound of this exponent function. This result corresponds to Theorem 3 in Table 1. As a corollary from this theorem, we obtain the strong converse result, which is stated in Corollary 2 in Table 1. This results states that we have an outer bound with $O(1/\sqrt{n})$ gap from the rate distortion region $\mathcal{R}_{WZ}(p_{XY})$.

Table 1. Previous results on the converse coding theorems for DMS, WZ. Main results in the present paper on WZ are also included.

	Characterization of the Rate Distortion Region	Strong Converse	Strong Converse Exponent
DMS	Shannon [2] (1959) (Explicit form of $\tilde{\mathcal{R}}_{DMS}(p_X)$) Wolfowitz [3] (1966) ($\tilde{\mathcal{R}}_{DMS}(p_X) = \mathcal{R}_{DMS}(p_X)$)	Wolfowitz [3] (1966) ($\mathcal{R}_{DMS}(\varepsilon p_X) = \mathcal{R}_{DMS}(p_X)$ for any $\varepsilon \in (0, 1)$)	Csiszár and Körner [4] (1981) (The optimal exponent)
WZ	Wyner and Ziv [1] (1976) (Explicit form of $\tilde{\mathcal{R}}_{WZ}(p_{XY})$) Csiszár and Körner [4] (1981) ($\tilde{\mathcal{R}}_{WZ}(p_{XY}) = \mathcal{R}_{WZ}(p_{XY})$)	Corollary 2 (Outer bound with $O(1/\sqrt{n})$ gap from the rate distortion region, $\mathcal{R}_{WZ}(\varepsilon p_{XY}) = \mathcal{R}_{WZ}(p_{XY})$ for any $\varepsilon \in (0, 1)$)	Theorem 3 (Lower bound F of the opt. exp. G)

To derive our result, we use a new method called the recursive method. This 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 [5–8].

2. Source Coding with Side Information at the Decoder

In the following argument, the operations $E_p[\cdot]$ and $\text{Var}_p[\cdot]$, respectively, stand for the expectation and the variance with respect to a probability distribution p . When the value of p is obvious from the context, we omit the suffix p in those operations to simply write $E[\cdot]$ and $\text{Var}[\cdot]$. 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 $t = 1, 2, \dots$, 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

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

We consider a communication system depicted in Figure 2. Data sequences X^n is separately encoded to $\varphi^{(n)}(X^n)$ and is sent to the information processing center. At the center the decoder function $\psi^{(n)}$ observes $\varphi^{(n)}(X^n)$ and Y^n to output the estimation Z^n of X^n . The encoder function $\varphi^{(n)}$ is defined by

$$\varphi^{(n)} : \mathcal{X}^n \rightarrow \mathcal{M}_n = \{1, 2, \dots, M_n\}. \quad (4)$$

Let \mathcal{Z} be a reproduction alphabet. The decoder function $\psi^{(n)}$ is defined by

$$\psi^{(n)} : \mathcal{M}_n \times \mathcal{Y}^n \rightarrow \mathcal{Z}^n. \quad (5)$$

Let $d : \mathcal{X} \times \mathcal{Z} \rightarrow [0, \infty)$ be an arbitrary distortion measure on $\mathcal{X} \times \mathcal{Z}$. The distortion between $x^n \in \mathcal{X}^n$ and $z^n \in \mathcal{Z}^n$ is defined by

$$d(x^n, z^n) := \sum_{t=1}^n d(x_t, z_t).$$

The excess-distortion probability of decoding is

$$P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = \Pr \left\{ \frac{1}{n} d(X^n, Z^n) \geq \Delta \right\}, \quad (6)$$

where $Z^n = \psi^{(n)}(\varphi^{(n)}(X^n), Y^n)$. The average distortion $\Delta^{(n)}$ between X^n and Z^n is defined by

$$\Delta^{(n)} := \frac{1}{n} E[d(X^n, Z^n)] := \frac{1}{n} \sum_{t=1}^n E[d(X_t, Z_t)].$$

In the previous section, we gave the formal definitions of $\mathcal{R}_{WZ}(\varepsilon | p_{XY})$, $\varepsilon \in (0, 1)$, $\mathcal{R}_{WZ}(p_{XY})$, and $\tilde{\mathcal{R}}_{WZ}(p_{XY})$. We can show that the above three rate distortion regions satisfy the following property.

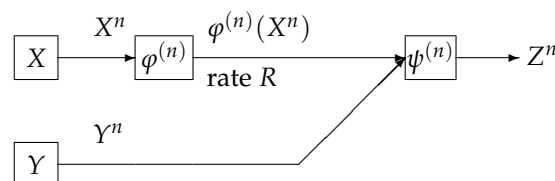


Figure 2. Wyner–Ziv source coding system.

Property 1.

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

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

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

$$\mathcal{R}_{WZ}(n, \varepsilon|p_{XY}) = \{(R, \Delta) : \text{There exists } (\varphi^{(n)}, \psi^{(n)}) \text{ such that} \\ \frac{1}{n} \log \|\varphi^{(n)}\| \leq R, \quad P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq \varepsilon\},$$

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

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

where $\text{cl}(\cdot)$ stands for the closure operation.

Proof of this property is given in Appendix A.

It is well known that $\tilde{\mathcal{R}}_{WZ}(p_{XY})$ was determined by Wyner and Ziv [1]. To describe their result we introduce auxiliary random variables U and Z , respectively, taking values in finite sets \mathcal{U} and \mathcal{Z} . We assume that the joint distribution of (U, X, Y, Z) is

$$p_{UXYZ}(u, x, y, z) = p_U(u)p_{X|U}(x|u)p_{Y|X}(y|x)p_{Z|UY}(z|u, y).$$

The above condition is equivalent to

$$U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z.$$

Define the set of probability distribution $p = p_{UXYZ}$ by

$$\mathcal{P}(p_{XY}) := \{p = p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 1, U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}, \\ \mathcal{P}^*(p_{XY}) := \{p = p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 1, U \leftrightarrow X \leftrightarrow Y, Z = \phi(U, Y) \text{ for some } \phi : \mathcal{U} \times \mathcal{Y} \rightarrow \mathcal{Z}\}.$$

By definitions, it is obvious that $\mathcal{P}^*(p_{XY}) \subseteq \mathcal{P}(p_{XY})$. Set

$$\mathcal{R}(p) := \{(R, \Delta) : R, \Delta \geq 0, R \geq I_p(X; U|Y), \Delta \geq E_p d(X, Z)\}, \\ \mathcal{R}(p_{XY}) := \bigcup_{p \in \mathcal{P}(p_{XY})} \mathcal{R}(p), \mathcal{R}^*(p_{XY}) := \bigcup_{p \in \mathcal{P}^*(p_{XY})} \mathcal{R}(p).$$

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

Property 2.

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

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

$$\mathcal{R}(p_{XY}) = \mathcal{R}^*(p_{XY}).$$

Proof of Property 2 is given in Appendix C. In Property 2 Part (b), $\mathcal{R}(p_{XY})$ is regarded as another expression of $\mathcal{R}^*(p_{XY})$. This expression is useful for deriving our main result. The rate region $\mathcal{R}_{WZ}(p_{XY})$ was determined by Wyner and Ziv [1]. Their result is the following:

Theorem 1 (Wyner and Ziv [1]).

$$\tilde{\mathcal{R}}_{WZ}(p_{XY}) = \mathcal{R}^*(p_{XY}) = \mathcal{R}(p_{XY}).$$

On $\mathcal{R}_{WZ}(p_{XY})$, Csiszár and Körner [4] obtained the following result.

Theorem 2 (Csiszár and Körner [4]).

$$\mathcal{R}_{WZ}(p_{XY}) = \tilde{\mathcal{R}}_{WZ}(p_{XY}) = \mathcal{R}^*(p_{XY}) = \mathcal{R}(p_{XY}).$$

We are interested in an asymptotic behavior of the error probability of decoding to tend to one as $n \rightarrow \infty$ for $(R, \Delta) \notin \mathcal{R}_{WZ}(p_{XY})$. To examine the rate of convergence, we define the following quantity. Set

$$\begin{aligned} P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &:= 1 - P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta), \\ G^{(n)}(R, \Delta | p_{XY}) &:= \min_{\substack{(\varphi^{(n)}, \psi^{(n)}): \\ (1/n) \log \|\varphi^{(n)}\| \leq R}} \left(-\frac{1}{n} \right) \log P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta). \end{aligned}$$

By time sharing, we have that

$$G^{(n+m)} \left(\frac{nR + mR'}{n+m}, \frac{n\Delta + m\Delta'}{n+m} \middle| p_{XY} \right) \leq \frac{nG^{(n)}(R, \Delta | p_{XY}) + mG^{(m)}(R', \Delta' | p_{XY})}{n+m}. \quad (7)$$

Choosing $R = R'$ and $\Delta = \Delta'$ in Equation (7), we obtain the following subadditivity property on $\{G^{(n)}(R, \Delta | p_{XY})\}_{n \geq 1}$:

$$G^{(n+m)}(R, \Delta | p_{XY}) \leq \frac{nG^{(n)}(R, \Delta | p_{XY}) + mG^{(m)}(R, \Delta | p_{XY})}{n+m},$$

which together with Fekete's lemma yields that $G^{(n)}(R, \Delta | p_{XY})$ exists and satisfies the following:

$$\lim_{n \rightarrow \infty} G^{(n)}(R, \Delta | p_{XY}) = \inf_{n \geq 1} G^{(n)}(R, \Delta | p_{XY}).$$

Set

$$\begin{aligned} G(R, \Delta | p_{XY}) &:= \lim_{n \rightarrow \infty} G^{(n)}(R, \Delta | p_{XY}), \\ \mathcal{G}(p_{XY}) &:= \{(R, \Delta, G) : G \geq G(R, \Delta | p_{XY})\}. \end{aligned}$$

The exponent function $G(R, \Delta | p_{XY})$ is a convex function of (R, Δ) . In fact, from Equation (7), we have that for any $\alpha \in [0, 1]$

$$G(\alpha R + \bar{\alpha} R', \alpha \Delta + \bar{\alpha} \Delta' | p_{XY}) \leq \alpha G(R, \Delta | p_{XY}) + \bar{\alpha} G(R', \Delta' | p_{XY}),$$

where $\bar{\alpha} = 1 - \alpha$. 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 $\mathcal{R}_{WZ}(p_{XY})$.

3. Main Results

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

$$\begin{aligned}\mathcal{P}_{\text{sh}}(p_{XY}) &:= \{p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|, U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}. \\ \mathcal{Q} &:= \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|\}.\end{aligned}$$

Let $\bar{\mu} = 1 - \mu$. We set

$$\begin{aligned}R^{(\mu)}(p_{XY}) &:= \min_{p \in \mathcal{P}_{\text{sh}}(p_{XY})} \{\bar{\mu} I_p(X; U|Y) + \mu E_p d(X; Z)\}, \\ \mathcal{R}_{\text{sh}}(p_{XY}) &:= \bigcap_{\mu \in [0,1]} \{(R, \Delta) : \bar{\mu} R + \mu \Delta \geq R^{(\mu)}(p_{XY})\}.\end{aligned}$$

Then, we have the following property:

Property 3. For any p_{XY} , we have

$$\mathcal{R}_{\text{sh}}(p_{XY}) = \mathcal{R}(p_{XY}). \quad (8)$$

Proof of Property 3 is given in Appendix D. For $\mu \in [0, 1]$ and $\lambda, \alpha \geq 0$, define

$$\begin{aligned}\omega_{q||p}^{(\mu, \lambda)}(x, y, z|u) &:= \log \left[\frac{q_X(x) q_{Y|XU}(y|x, u) q_{Z|UYX}(z|u, y, x)}{p_X(x) p_{Y|X}(y|x) q_{Z|UY}(z|u, y)} \right] + \lambda \left[\bar{\mu} \log \frac{q_{X|YU}(x|y, u)}{p_{X|Y}(x|y)} + \mu d(x, z) \right], \\ \Omega^{(\mu, \lambda, \alpha)}(q|p_{XY}) &:= -\log E_q \left[\exp \left\{ -\alpha \omega_{q||p}^{(\mu, \lambda)}(X, Y, Z|U) \right\} \right], \\ \Omega^{(\mu, \lambda, \alpha)}(p_{XY}) &:= \min_{q \in \mathcal{Q}} \Omega^{(\mu, \lambda, \alpha)}(q|p_{XY}), \\ F^{(\mu, \lambda, \alpha)}(\bar{\mu} R + \mu \Delta | p_{XY}) &:= \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY}) - \lambda \alpha (\bar{\mu} R + \mu \Delta)}{1 + (4 + \lambda \bar{\mu}) \alpha}.\end{aligned}$$

Furthermore, set

$$\begin{aligned}F(R, \Delta | p_{XY}) &:= \sup_{\mu \in [0,1], \lambda, \alpha \geq 0} F^{(\mu, \lambda, \alpha)}(\bar{\mu} R + \mu \Delta | p_{XY}), \\ \bar{\mathcal{G}}(p_{XY}) &:= \{(R, \Delta, G) : G \geq F(R, \Delta | p_{XY})\}.\end{aligned}$$

We next define a functions serving as a lower bound of $F(R, \Delta | p_{XY})$. For each $p = p_{UXYZ} \in \mathcal{P}_{\text{sh}}(p_{XY})$, define

$$\begin{aligned}\tilde{\omega}_p^{(\mu)}(x, y, z|u) &:= \left[\bar{\mu} \log \frac{p_{X|YU}(x|y, u)}{p_{X|Y}(x|y)} + \mu d(x, z) \right], \\ \tilde{\Omega}^{(\mu, \lambda)}(p) &:= -\log E_p \left[\exp \left\{ -\lambda \omega_p^{(\mu)}(X, Y, Z|U) \right\} \right].\end{aligned}$$

Furthermore, set

$$\begin{aligned}\tilde{\Omega}^{(\mu, \lambda)}(p_{XY}) &:= \min_{p \in \mathcal{P}_{\text{sh}}(p_{XY})} \tilde{\Omega}^{(\mu, \lambda)}(p), \\ \tilde{F}^{(\mu, \lambda)}(\bar{\mu} R + \mu \Delta | p_{XY}) &:= \frac{\tilde{\Omega}^{(\mu, \lambda)}(p_{XY}) - \lambda (\bar{\mu} R + \mu \Delta)}{5 + \lambda(1 + \bar{\mu})},\end{aligned}$$

$$\tilde{F}(R, \Delta | p_{XY}) := \sup_{\lambda \geq 0, \mu \in [0,1]} \tilde{F}^{(\mu, \lambda)}(\bar{\mu}R + \mu\Delta | p_{XY}).$$

We can show that the above functions satisfies the following properties:

Property 4.

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

(b) For any $R, \Delta \geq 0$, we have

$$F(R, \Delta | p_{XY}) \geq \tilde{F}(R, \Delta | p_{XY}).$$

(c) Fix any $p = p_{UXY} \in \mathcal{P}_{\text{sh}}(p_{XY})$ and $\mu \in [0, 1]$. For $\lambda \in [0, 1]$, $\tilde{\Omega}^{(\mu, \lambda)}(p)$ exists and is nonnegative. For $p = p_{UXYZ} \in \mathcal{P}_{\text{sh}}(p_{XY})$, define a probability distribution $p^{(\lambda)} = p_{UXYZ}^{(\lambda)}$ by

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

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

$$\begin{aligned} \frac{d}{d\lambda} \tilde{\Omega}^{(\mu, \lambda)}(p) &= \mathbb{E}_{p^{(\lambda)}} \left[\omega_p^{(\mu)}(X, Y, Z | U) \right], \\ \frac{d^2}{d\lambda^2} \tilde{\Omega}^{(\mu, \lambda)}(p) &= -\text{Var}_{p^{(\lambda)}} \left[\tilde{\omega}_p^{(\mu)}(X, Y, Z | U) \right]. \end{aligned}$$

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

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

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

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 $(\mu, \lambda) \in [0, 1] \times [0, 1/2]$,

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

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

$$\tilde{F}(R, \Delta | p_{XY}) > \frac{\rho(p_{XY})}{10} \cdot g^2 \left(\frac{\tau}{\rho(p_{XY})} \right) > 0,$$

where g is the inverse function of $\vartheta(a) := a + (1/5)a^2$, $a \geq 0$.

Proof of Property 4 Part (a) is given in Appendix B. Proof of Property 4 Part (b) is given in Appendix E. Proofs of Property 4 Parts (c), (d), and (e) are given in Appendix F.

Our main result is the following:

Theorem 3. For any $R, \Delta \geq 0$, any p_{XY} , and for any $(\varphi^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log \|\varphi^{(n)}\| \leq R$, we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq 5 \exp \{-nF(R, \Delta|p_{XY})\}. \quad (10)$$

It follows from Theorem 3 and Property 4 Part (d) that if (R, Δ) is outside the rate distortion region, then the error probability of decoding goes to one exponentially and its exponent is not below $F(R, \Delta|p_{XY})$.

It immediately follows from Theorem 3 that we have the following corollary.

Corollary 1. For any $R, \Delta \geq 0$ and any p_{XY} , we have

$$G(R, \Delta|p_{XY}) \geq F(R, \Delta|p_{XY}). \quad (11)$$

Furthermore, for any p_{XY} , we have

$$\mathcal{G}(p_{XY}) \subseteq \overline{\mathcal{G}}(p_{XY}) := \{(R, \Delta, G) : G \geq F(R, \Delta|p_{XY})\}. \quad (12)$$

Proof of Theorem 3 will be given in the next section. The exponent function in the case of $\Delta = 0$ can be obtained as a corollary of the result of Oohama and Han [9] for the separate source coding problem of correlated sources [10]. The techniques used by them is a method of types [4], which is not useful for proving Theorem 3. In fact, when we use this method, it is very hard to extract a condition related to the Markov chain condition $U \leftrightarrow X \leftrightarrow Y$, which the auxiliary random variable $U \in \mathcal{U}$ must satisfy when (R, Δ) is on the boundary of the set $\mathcal{R}(p_{XY})$. Some novel techniques based on the information spectrum method introduced by Han [11] are necessary to prove this theorem.

From Theorem 3 and Property 4 Part (e), we can obtain an explicit outer bound of $\mathcal{R}_{WZ}(\varepsilon|p_{XY})$ with an asymptotically vanishing deviation from $\mathcal{R}_{WZ}(p_{XY}) = \mathcal{R}(p_{XY})$. The strong converse theorem immediately follows from this corollary. To describe this outer bound, for $\kappa > 0$, we set

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

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

$$\begin{aligned} \kappa_n &:= \rho(p_{XY}) \vartheta \left(\sqrt{\frac{10}{n\rho(p_{XY})} \log \left(\frac{5}{1-\varepsilon} \right)} \right) \\ &\stackrel{(a)}{=} \sqrt{\frac{10\rho(p_{XY})}{n} \log \left(\frac{5}{1-\varepsilon} \right)} + \frac{2}{n} \log \left(\frac{5}{1-\varepsilon} \right). \end{aligned} \quad (13)$$

Step (a) follows from $\vartheta(a) = a + (1/5)a^2$. Since $\kappa_n \rightarrow 0$ as $n \rightarrow \infty$, we have the smallest positive integer $n_0 = n_0(\varepsilon, \rho(p_{XY}))$ such that $\kappa_n \leq (1/2)\rho(p_{XY})$ for $n \geq n_0$. 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 $n \geq n_0$, we have

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

The above result together with

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

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

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

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

The direct part of coding theorem, i.e., the inclusion of $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}_{WZ}(\varepsilon|p_{XY})$ was established by Csiszár and Körner [4]. They proved a weak converse theorem to obtain the inclusion $\mathcal{R}_{WZ}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$. Until now, we have had no result on the strong converse theorem. The above corollary stating the strong converse theorem for the Wyner–Ziv source coding problem implies that a long standing open problem since Csiszár and Körner [4] has been resolved.

4. Proof of the Main Results

In this section, we prove Theorem 3 and Corollary 2. We first present a lemma which upper bounds the correct probability of decoding by the information spectrum quantities. We set

$$S_n := \varphi^{(n)}(X^n), Z^n := \psi_n(\varphi^{(n)}(X^n), Y^n).$$

It is obvious that

$$S_n \leftrightarrow X^n \leftrightarrow Y^n, X^n \leftrightarrow (S_n, Y^n) \leftrightarrow Z^n.$$

Then, we have the following:

Lemma 1. For any $\eta > 0$ and for any $(\varphi^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log \|\varphi^{(n)}\| \leq R$, we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq p_{S_n X^n Y^n Z^n} \left\{ \begin{aligned} &\eta \geq \frac{1}{n} \log \frac{Q_{X^n}^{(i)}(X^n)}{p_{X^n}(X^n)}, \end{aligned} \right. \quad (14)$$

$$\eta \geq \frac{1}{n} \log \frac{Q_{Y^n|S_n X^n}^{(ii)}(Y^n|S_n, X^n)}{p_{Y^n|X^n}(Y^n|X^n)}, \quad (15)$$

$$\eta \geq \frac{1}{n} \log \frac{Q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n, Y^n, Z^n)}{p_{X^n|S_n Y^n}(X^n|S_n, Y^n)}, \quad (16)$$

$$R + \eta \geq \frac{1}{n} \log \frac{Q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n, Y^n)}{p_{X^n|Y^n}(X^n|Y^n)}, \quad (17)$$

$$\Delta \geq \frac{1}{n} \log \exp \{d(X^n, Z^n)\} + 4e^{-n\eta}. \quad (18)$$

The probability distribution and stochastic matrices appearing in the right members of Equation (18) have a property that we can select them arbitrary. In Equation (14), we can choose any probability distribution $Q_{X^n}^{(i)}$ on \mathcal{X}^n . In Equation (15), we can choose any stochastic matrix $Q_{Y^n|S_n X^n}^{(ii)} : \mathcal{M}_n \times \mathcal{X}^n \rightarrow \mathcal{Y}^n$. In Equation (16), we can choose any stochastic matrix $Q_{X^n|S_n Y^n Z^n}^{(iii)} : \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{Z}^n \rightarrow \mathcal{X}^n$. In Equation (17), we can choose any stochastic matrix $Q_{X^n|S_n Y^n}^{(iv)} : \mathcal{M}_n \times \mathcal{Y}^n \rightarrow \mathcal{X}^n$.

Proof of this lemma is given in Appendix G.

Lemma 2. Suppose that, for each $t = 1, 2, \dots, n$, the joint distribution $p_{S_n X^t Y^n}$ of the random vector $S_n X^t Y^n$ is a marginal distribution of $p_{S_n X^n Y^n}$. Then, for $t = 1, 2, \dots, n$, we have the following Markov chain:

$$X_t \leftrightarrow S_n X^{t-1} Y_t^n \leftrightarrow Y^{t-1} \quad (19)$$

or equivalently that $I(X_t; Y^{t-1} | S_n X^{t-1} Y_t^n) = 0$.

Proof of this lemma is given in Appendix H. For $t = 1, 2, \dots, n$, set $u_t := (s, x^{t-1}, y_{t+1}^n)$. Let $U_t := (S_n, X^{t-1}, Y_{t+1}^n)$ be a random vector taking values in $\mathcal{M}_n \times \mathcal{X}^{t-1} \times \mathcal{Y}_{t+1}^n$. From Lemmas 1 and 2, we have the following:

Lemma 3. For any $\eta > 0$ and for any $(\varphi^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log \|\varphi^{(n)}\| \leq R$, we have the following:

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &\leq p_{S_n X^n Y^n Z^n} \left\{ \right. \\ \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{X_t}^{(i)}(X_t)}{p_{X_t}(X_t)}, \\ \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{Y_t|U_t X_t}^{(ii)}(Y_t|U_t, X_t)}{p_{Y_t|X_t}(Y_t|X_t)}, \\ \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{X_t|U_t Y_t Z_t}^{(iii)}(X_t|U_t, Y_t, Z_t)}{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}, \\ R + \eta &\geq \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{X_t|U_t Y_t}^{(iv)}(X_t|U_t, Y_t)}{p_{X_t|Y_t}(X_t|Y_t)}, \\ \Delta &\geq \frac{1}{n} \sum_{t=1}^n \log e^{d(X_t, Z_t)} \left. \right\} + 4e^{-n\eta}, \end{aligned} \quad (20)$$

where for each $t = 1, 2, \dots, n$, the following probability distribution and stochastic matrices:

$$Q_{X_t}^{(i)}, Q_{Y_t|U_t X_t}^{(ii)}, Q_{X_t|U_t Y_t Z_t}^{(iii)}, \text{ and } Q_{X_t|U_t Y_t}^{(iv)}$$

appearing in the first term in the right members of Equation (21) have a property that we can choose their values arbitrary.

Proof. On the probability distributions appearing in the right members of Equation (18), we take the following choices. In Equation (14), we choose $Q_{X^n}^{(i)}$ so that

$$Q_{X^n}^{(i)}(X^n) = \prod_{t=1}^n Q_{X_t}^{(i)}(X_t). \quad (21)$$

In Equation (15), we choose $Q_{Y^n|S_n X^n}^{(ii)}$ so that

$$Q_{Y^n|S_n X^n}^{(ii)}(Y^n|S_n, X^n) = \prod_{t=1}^n Q_{Y_t|S_n X^t Y_{t+1}^n}^{(ii)}(Y_t|S_n, X^t, Y_{t+1}^n) = \prod_{t=1}^n Q_{Y_t|X_t U_t}^{(ii)}(Y_t|U_t, X_t). \quad (22)$$

In Equation (16), we choose $Q_{X^n|S_n Y^n Z^n}^{(iii)}$ so that

$$\begin{aligned} Q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n, Y^n, Z^n) &= \prod_{t=1}^n Q_{X_t|S_n X^{t-1} Y_t^n Z_t}^{(iii)}(X_t|S_n X^{t-1}, Y_t^n, Z_t) \\ &= \prod_{t=1}^n Q_{X_t|U_t Y_t Z_t}^{(iii)}(X_t|U_t, Y_t, Z_t). \end{aligned} \quad (23)$$

In Equation (16), we note that

$$\begin{aligned} p_{X^n|S_n Y^n}(X^n|S_n, Y^n) &= \prod_{t=1}^n p_{X_t|S_n X^{t-1} Y^n}(X_t|S_n, X^{t-1}, Y^n) \stackrel{(a)}{=} \prod_{t=1}^n p_{X_t|S_n X^{t-1} Y_t^n}(X_t|S_n, X^{t-1}, Y_t^n) \\ &= \prod_{t=1}^n p_{X_t|U_t Y_t}(X_t|U_t, Y_t). \end{aligned} \quad (24)$$

Step (a) follows from Lemma 2. In Equation (17), we choose $Q_{X^n|S_n Y^n}^{(iv)}$ so that

$$Q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n, Y^n) = \prod_{t=1}^n Q_{X_t|S_n X^{t-1} Y_t^n}^{(iv)}(X_t|S_n, X^{t-1}, Y_t^n) = \prod_{t=1}^n Q_{X_t|U_t Y_t}^{(iv)}(X_t|U_t, Y_t). \quad (25)$$

From Lemma 1 and Equations (21)–(25), we have the bound of Equation (21) in Lemma 3. \square

To evaluate an upper bound of Equation (21) in Lemma 3, we use the following lemma, which is well known as the Cramér's bound in the large deviation principle.

Lemma 4. For any real valued random variable A and any $\theta \geq 0$, we have

$$\Pr\{A \leq a\} \leq \exp[(\theta a + \log E[\exp(-\theta A)])].$$

Here, we define a quantity which serves as an exponential upper bound of $P_c^{(n)}(\varphi^{(n)}, \psi^{(n)})$. For each $t = 1, 2, \dots, n$, let \underline{Q}_t be a set of all

$$\underline{Q}_t = (Q_{X_t}^{(i)}, Q_{Y_t|U_t X_t}^{(ii)}, Q_{X_t|U_t Y_t Z_t}^{(iii)}, Q_{X_t|U_t Y_t}^{(iv)}).$$

Set

$$\underline{Q}^n := \prod_{t=1}^n \underline{Q}_t, \underline{Q}^n := \{\underline{Q}_t\}_{t=1}^n \in \underline{Q}^n.$$

Let $\mathcal{P}^{(n)}(p_{XY})$ be a set of all probability distributions $p_{S_n X^n Y^n Z^n}$ on $\mathcal{M}_n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ having the form:

$$p_{S_n X^n Y^n Z^n}(s, x^n, y^n, z^n) = p_{S_n|X^n}(s|x^n) \left\{ \prod_{t=1}^n p_{X_t Y_t}(x_t, y_t) \right\} p_{Z^n|Y^n S_n}(z^n|y^n, s).$$

For simplicity of notation, we use the notation $p^{(n)}$ for $p_{S_n X^n Y^n Z^n} \in \mathcal{P}^{(n)}(p_{XY})$. We assume that $p_{U_t X_t Y_t Z_t} = p_{S_n X^n Y_t^n Z_t}$ is a marginal distribution of $p^{(n)}$. For $t = 1, 2, \dots, n$, we simply write $p_t = p_{U_t X_t Y_t Z_t}$. For $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$ and $\underline{Q}^n \in \underline{Q}^n$, we define

$$\begin{aligned} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n|p_{XY}) &:= -\log E_{p^{(n)}} \left[\left(\prod_{t=1}^n \left\{ \frac{p_{X_t}(X_t)}{Q_{X_t}^{(i)}(X_t)} \frac{p_{Y_t|X_t}(Y_t|X_t)}{Q_{Y_t|X_t U_t}^{(ii)}(Y_t|X_t, U_t)} \right\} \right)^\theta \right. \\ &\quad \times \left. \left(\prod_{t=1}^n \left\{ \frac{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{Q_{X_t|U_t Y_t Z_t}^{(iii)}(X_t|U_t, Y_t, Z_t)} \right\}^\theta \right) \left(\prod_{t=1}^n \left\{ \left(\frac{p_{X_t|Y_t}(X_t|Y_t)}{Q_{X_t|Y_t U_t}^{(iv)}(X_t|U_t, Y_t)} \right)^{\bar{\mu}} e^{-\mu d(X_t, Z_t)} \right\}^{\lambda \theta} \right) \right], \end{aligned}$$

where, for each $t = 1, 2, \dots, n$, the following probability distribution and stochastic matrices:

$$Q_{X_t}^{(i)}, Q_{X_t|U_t Y_t}^{(ii)}, Q_{X_t|U_t Y_t Z_t}^{(iii)}, Q_{Y_t|X_t U_t}^{(iv)}$$

appearing in the definition of $\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY})$ are chosen so that they are induced by the joint distribution $q_t = q_{U_t X_t Y_t Z_t} \in \mathcal{Q}_t$.

By Lemmas 3 and 4, we have the following proposition:

Proposition 1. For any $\mu \in [0, 1]$, $\lambda, \theta \geq 0$, any $q^n \in \mathcal{Q}^n$, and any $(\varphi^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log \|\varphi^{(n)}\| \leq R$, we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq 5 \exp \left\{ -n [1 + (3 + \lambda \bar{\mu})\theta]^{-1} \left[\frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) - \lambda \theta (\bar{\mu} R + \mu \Delta) \right] \right\}.$$

Proof. When $\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \leq n \lambda \theta (\bar{\mu} R + \mu \Delta)$, the bound we wish to prove is obvious. In the following argument, we assume that $\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) > n \lambda \theta (\bar{\mu} R + \mu \Delta)$. We define five random variables $A_i, i = 1, 2, \dots, 5$ by

$$\begin{aligned} A_1 &= \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{X_t}^{(i)}(X_t)}{p_{X_t}(X_t)}, A_2 = \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{Y_t|X_t U_t}^{(ii)}(Y_t | X_t, U_t)}{p_{Y_t|X_t}(Y_t | X_t)}, \\ A_3 &= \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{X_t|U_t Y_t Z_t}^{(iii)}(X_t | U_t, Y_t, Z_t)}{p_{X_t|U_t Y_t}(X_t | U_t, Y_t)}, \\ A_4 &= \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{X_t|U_t Y_t}^{(iv)}(X_t | U_t, Y_t)}{p_{X_t|Y_t}(X_t | Y_t)}, A_5 = \frac{1}{n} \sum_{t=1}^n \log e^{d(X_t, Z_t)}. \end{aligned}$$

By Lemma 3, for any $(\varphi^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log \|\varphi^{(n)}\| \leq R$, we have

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &\leq p_{S_n X^n Y^n Z^n} \{A_i \leq \eta \text{ for } i = 1, 2, 3, A_4 \leq R + \eta, A_5 \leq \Delta\} + 4e^{-n\eta} \\ &\leq p_{S_n X^n Y^n Z^n} \{A_1 + A_2 + A_3 + \lambda(\bar{\mu} A_4 + \mu A_5) \leq \lambda(\bar{\mu} R + \mu \Delta) + (3 + \lambda \bar{\mu})\eta\} + 4e^{-n\eta} \\ &= p_{S_n X^n Y^n Z^n} \{A \leq a\} + 4e^{-n\eta}, \end{aligned} \quad (26)$$

where we set

$$\begin{aligned} A &:= A_1 + A_2 + A_3 + \lambda(\bar{\mu} A_4 + \mu A_5), \\ a &:= \lambda(\bar{\mu} R + \mu \Delta) + (3 + \lambda \bar{\mu})\eta. \end{aligned}$$

Applying Lemma 4 to the first term in the right member of Equation (26), we have

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &\leq \exp \left[\left(\theta a + \log E_{p^{(n)}}[\exp(-\theta A)] \right) \right] + 4e^{-n\eta} \\ &= \exp \left[n \left\{ \lambda \theta (\bar{\mu} R + \mu \Delta) + \theta (3 + \lambda \bar{\mu})\eta - \frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \right\} \right] + 4e^{-n\eta}. \end{aligned} \quad (27)$$

We choose η so that

$$-\eta = \lambda \theta (\bar{\mu} R + \mu \Delta) + \theta (3 + \lambda \bar{\mu})\eta - \frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}). \quad (28)$$

Solving Equation (28) with respect to η , we have

$$\eta = \frac{\frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) - \lambda \theta (\bar{\mu} R + \mu \Delta)}{1 + (3 + \lambda \bar{\mu})\theta}.$$

For this choice of η and Equation (27), we have

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) &\leq 5e^{-n\eta} \\ &= 5 \exp \left\{ -n [1 + (3 + \lambda \bar{\mu})\theta]^{-1} \left[\frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) - \lambda \theta (\bar{\mu} R + \mu \Delta) \right] \right\}, \end{aligned}$$

completing the proof. \square

Set

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

By Proposition 1, we have the following corollary.

Corollary 3. For any $\mu \in [0, 1]$, $\lambda \geq 0$, for any $\theta \geq 0$, and for any $(\varphi^{(n)}, \psi^{(n)})$ satisfying $(1/n) \log \|\varphi^{(n)}\| \leq R$, we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq 5 \exp \left\{ -n \left[\frac{\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY}) - \lambda \theta (\bar{\mu} R + \mu \Delta)}{1 + (3 + \lambda \bar{\mu})\theta} \right] \right\}.$$

We shall call $\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY})$ the communication potential. The above corollary implies that the analysis of $\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY})$ leads to an establishment of a strong converse theorem for Wyner–Ziv source coding problem. In the following argument, we drive an explicit lower bound of $\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY})$. We use a new technique we call *the recursive method*. The recursive method is a powerful tool to drive a single letterized exponent function for rates below the rate distortion function. This method is also applicable to prove the exponential strong converse theorems for other network information theory problems [5–7]. Set

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

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

$$\begin{aligned} &f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t) \\ &:= \left\{ \frac{p_{X_t}(x_t)}{Q_{X_t}^{(i)}(x_t)} \frac{p_{Y_t|X_t}(y_t | x_t)}{Q_{Y_t|X_t}^{(ii)}(y_t | x_t, u_t)} \frac{p_{X_t|U_t Y_t}(x_t | u_t, y_t)}{Q_{X_t|U_t Y_t}^{(iii)}(x_t | u_t, y_t, z_t)} \left(\frac{p_{X_t|Y_t}(x_t | y_t)}{Q_{X_t|Y_t}^{(iv)}(x_t | u_t, y_t)} \right)^{\lambda \bar{\mu}} e^{-\lambda \mu d(x_t, z_t)} \right\}^{\theta}. \end{aligned}$$

By definition, we have

$$\begin{aligned} &\exp \left\{ -\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \right\} \\ &= \sum_{s, y^n} p_{S_n Y^n}(s, y^n) \sum_{x^n, z^n} p_{X^n Z^n | S_n Y^n}(x^n, z^n | s, y^n) \prod_{t=1}^n f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (29)$$

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

$$p_{X^t Z^t | S_n Y^n; \mathcal{F}^t}^{(\mu, \lambda, \theta)} := \left\{ p_{X^t Z^t | S_n Y^n; \mathcal{F}^t}^{(\mu, \lambda, \theta)}(x^t, z^t | s, y^n) \right\}_{(x^t, z^t, s, y^n) \in \mathcal{X}^t \times \mathcal{Z}^t \times \mathcal{M}_n \times \mathcal{Y}^n}$$

by

$$p_{X^t Z^t | S_n Y^n; \mathcal{F}^t}^{(\mu, \lambda, \theta)}(x^t, z^t | s, y^n) := C_t^{-1}(s, y^n) p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \prod_{i=1}^t f_{\mathcal{F}_i}^{(\mu, \lambda, \theta)}(x_i, y_i, z_i | u_i)$$

where

$$C_t(s, y^n) := \sum_{x^t, z^t} p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \prod_{i=1}^t f_{\mathcal{F}_i}^{(\mu, \lambda, \theta)}(x_i, y_i, z_i | u_i) \quad (30)$$

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

$$\Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n) := C_t(s, y^n) C_{t-1}^{-1}(s, y^n), \quad (31)$$

where we define $C_0(s, y^n) = 1$ for $(s, y^n) \in \mathcal{M}_n \times \mathcal{Y}^n$. Then, we have the following lemma:

Lemma 5. For each $t = 1, 2, \dots, n$, and for any $(s, y^n, x^t, z^t) \in \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{X}^t \times \mathcal{Z}^t$, we have

$$p_{X^t Z^t | S_n Y^n, \mathcal{F}^t}(x^t, z^t | s, y^n) = (\Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n))^{-1} p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}(x^{t-1}, z^{t-1} | s, y^n) \\ \times p_{X_t Z_t | S_n X^{t-1} Y^n}(x_t, z_t | s, x^{t-1}, z^{t-1}, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \quad (32)$$

$$\Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n) = \sum_{x^t, z^t} p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}(x^{t-1}, z^{t-1} | s, y^n) \\ \times p_{X_t Z_t | S_n X^{t-1} Y^n}(x_t, z_t | s, x^{t-1}, z^{t-1}, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \quad (33)$$

Furthermore, we have

$$\exp \left\{ -\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \right\} = \sum_{s, y^n} p_{S_n Y^n}(s, y^n) \prod_{t=1}^n \Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n). \quad (34)$$

The equality in Equation (34) in Lemma 5 is obvious from Equations (29)–(31). Proofs of Equations (32) and (33) in this lemma are given in Appendix I. Next, we define a probability distribution of the random pair (S_n, Y^n) taking values in $\mathcal{M}_n \times \mathcal{Y}^n$ by

$$p_{S_n Y^n, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n) = \tilde{C}_t^{-1} p_{S_n Y^n}(s, y^n) \prod_{i=1}^t \Phi_{i, \mathcal{F}^i}^{(\mu, \lambda, \theta)}(s, y^n), \quad (35)$$

where \tilde{C}_t is a constant for normalization given by

$$\tilde{C}_t = \sum_{s, y^n} p_{S_n Y^n}(s, y^n) \prod_{i=1}^t \Phi_{i, \mathcal{F}^i}^{(\mu, \lambda, \theta)}(s, y^n).$$

For $t = 1, 2, \dots, n$, define

$$\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} := \tilde{C}_t \tilde{C}_{t-1}^{-1}, \quad (36)$$

where we define $\tilde{C}_0 = 1$. Set

$$p_{S_n X^t Y^n Z_t | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(s, x^t, y^n, z_t) = p_{U_t X_t Y_t Z_t | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(u_t, x_t, y_t, z_t) \\ := \sum_{y^{t-1}, z^{t-1}} p_{S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(s, y^n) p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n). \quad (37)$$

Then, we have the following:

Lemma 6.

$$\exp \left\{ -\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \right\} = \prod_{t=1}^n \Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}, \quad (38)$$

$$\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} = \sum_{u_t, x_t, y_t, z_t} p_{U_t X_t Y_t Z_t | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(u_t, x_t, y_t, z_t) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \quad (39)$$

Proof. By the equality Equation (34) in Lemma 5, we have

$$\exp \left\{ -\Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \right\} = \tilde{C}_n = \prod_{t=1}^n \tilde{C}_t \tilde{C}_{t-1}^{-1} \stackrel{(a)}{=} \prod_{t=1}^n \Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}. \quad (40)$$

Step (a) follows from the definition in Equation (36) of $\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \nu, \theta)}$. We next prove Equation (39) in Lemma 6. Multiplying $\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} = \tilde{C}_t / \tilde{C}_{t-1}$ to both sides of Equation (35), we have

$$\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} p_{S_n Y^n | \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n) \quad (41)$$

$$= \tilde{C}_{t-1}^{-1} p_{S_n Y^n}(s, y^n) \prod_{i=1}^t \Phi_{i, \mathcal{F}^i}^{(\mu, \lambda, \theta)}(s, y^n) \\ = p_{S_n Y^n | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(s, y^n) \Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n). \quad (42)$$

Taking summations of Equations (41) and (42) with respect to (s, y^n) , we have

$$\begin{aligned} \Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} &= \sum_{s, y^n} p_{S_n Y^n | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(s, y^n) \Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n) \\ &\stackrel{(a)}{=} \sum_{s, y^n} p_{S_n Y^n | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(s, y^n) \sum_{x^t, z^t} p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t) \\ &= \sum_{s, x^t, y_t^t, z_t^t} \sum_{y^{t-1}, z^{t-1}} p_{S_n Y^n | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(s, y^n) p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \\ &\stackrel{(b)}{=} \sum_{u_t, x_t, y_t, z_t} p_{U_t X_t Y_t Z_t | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(u_t, x_t, y_t, z_t) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \end{aligned}$$

Step (a) follows from Equation (33) in Lemma 5. Step (b) follows from the definition in Equation (37) of $p_{U_t X_t Y_t Z_t | \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}$. \square

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

Proposition 2. For $\theta \in [0, 1)$, we choose the parameter α such that

$$\alpha = \frac{\theta}{1 - \theta} \Leftrightarrow \theta = \frac{\alpha}{1 + \alpha}. \quad (43)$$

Then, for any $\lambda \geq 0, \mu \in [0, 1]$ and for any $\theta \in [0, 1)$, we have

$$\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY}) \geq \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY})}{1 + \alpha}. \quad (44)$$

Proof. Set

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

Then, by Lemma 6, we have

$$\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} = \sum_{u_t, x_t, y_t, z_t} p_{U_t X_t Y_t Z_t; \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(u_t, x_t, y_t, z_t) f_{\mathcal{F}^t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t).$$

For each $t = 1, 2, \dots, n$, we recursively choose $q_t = q_{U_t X_t Y_t Z_t}$ so that $q_{U_t X_t Y_t Z_t} = p_{U_t X_t Y_t Z_t; \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}$ and choose $Q_{X_t}^{(i)}$, $Q_{Y_t|X_t U_t}^{(ii)}$, $Q_{X_t|U_t Y_t Z_t}^{(iii)}$, and $Q_{X_t|Y_t U_t}^{(iv)}$ appearing in

$$\begin{aligned}f_{\mathcal{F}^t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t) \\ := \left\{ \frac{p_{X_t}(x_t)}{Q_{X_t}^{(i)}(x_t)} \frac{p_{Y_t|X_t}(y_t|x_t)}{Q_{Y_t|X_t U_t}^{(ii)}(y_t|x_t, u_t)} \frac{p_{X_t|U_t Y_t}(x_t|u_t, y_t)}{Q_{X_t|U_t Y_t Z_t}^{(iii)}(x_t|u_t, y_t, z_t)} \left(\frac{p_{X_t|Y_t}(x_t|y_t)}{Q_{X_t|Y_t U_t}^{(iv)}(x_t|u_t, y_t)} \right)^{\lambda \bar{\mu}} e^{-\lambda \mu d(x_t, z_t)} \right\}^{\theta}\end{aligned}$$

such that they are the distributions induced by $q_{U_t X_t Y_t Z_t}$. Then, for each $t = 1, 2, \dots, n$, we have the following chain of inequalities:

$$\begin{aligned}\Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} &= \mathbb{E}_{q_t} \left[\left\{ \frac{p_{X_t}(X_t)}{q_{X_t}(X_t)} \frac{p_{Y_t|X_t}(Y_t|X_t)}{q_{Y_t|X_t U_t}(Y_t|X_t, U_t)} \frac{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{q_{X_t|U_t Y_t Z_t}(X_t|U_t, Y_t, Z_t)} \frac{p_{X_t|Y_t}^{\lambda \bar{\mu}}(X_t|Y_t) e^{-\lambda \mu d(X_t, Z_t)}}{q_{X_t|Y_t U_t}^{\lambda \bar{\mu}}(X_t|U_t, Y_t)} \right\}^{\theta} \right] \\ &= \mathbb{E}_{q_t} \left[\left\{ \frac{p_{X_t}(X_t)}{q_{X_t}(X_t)} \frac{p_{Y_t|X_t}(Y_t|X_t)}{q_{Y_t|X_t U_t}(Y_t|X_t, U_t)} \frac{q_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{q_{X_t|U_t Y_t Z_t}(X_t|U_t, Y_t, Z_t)} \frac{p_{X_t|Y_t}^{\lambda \bar{\mu}}(X_t|Y_t) e^{-\lambda \mu d(X_t, Z_t)}}{q_{X_t|Y_t U_t}^{\lambda \bar{\mu}}(X_t|U_t, Y_t)} \right\}^{\theta} \right. \\ &\quad \times \left. \left\{ \frac{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{q_{X_t|U_t Y_t}(X_t|U_t, Y_t)} \right\}^{\theta} \right] \tag{45} \\ &\stackrel{(a)}{\leq} \left(\mathbb{E}_{q_t} \left[\left\{ \frac{p_{X_t}(X_t)}{q_{X_t}(X_t)} \frac{p_{Y_t|X_t}(Y_t|X_t)}{q_{Y_t|X_t U_t}(Y_t|X_t, U_t)} \frac{q_{Z_t|U_t Y_t}(Z_t|U_t, Y_t)}{q_{Z_t|U_t X_t Y_t}(Z_t|U_t, X_t, Y_t)} \frac{p_{X_t|Y_t}^{\lambda \bar{\mu}}(X_t|Y_t) e^{-\lambda \mu d(X_t, Z_t)}}{q_{X_t|Y_t U_t}^{\lambda \bar{\mu}}(X_t|U_t, Y_t)} \right\}^{\frac{\theta}{1-\theta}} \right] \right)^{1-\theta} \\ &\quad \times \left(\mathbb{E}_{q_t} \left\{ \frac{p_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{q_{X_t|U_t Y_t}(X_t|U_t, Y_t)} \right\}^{\theta} \right) = \exp \left\{ -(1-\theta) \Omega^{(\mu, \lambda, \frac{\theta}{1-\theta})}(q_t|p_{XY}) \right\} \\ &\stackrel{(b)}{=} \exp \left\{ -\frac{\Omega^{(\mu, \lambda, \alpha)}(q_t|p_{XY})}{1+\alpha} \right\} \stackrel{(c)}{\leq} \exp \left\{ -\frac{\hat{\Omega}_n^{(\mu, \lambda, \alpha)}(p_{XY})}{1+\alpha} \right\} \stackrel{(d)}{=} \exp \left\{ -\frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY})}{1+\alpha} \right\}.\end{aligned}$$

Step (a) follows from Hölder's inequality and the following identity:

$$\frac{q_{X_t|U_t Y_t}(X_t|U_t, Y_t)}{q_{X_t|U_t Y_t Z_t}(X_t|U_t, Y_t, Z_t)} = \frac{q_{Z_t|U_t Y_t}(Z_t|U_t, Y_t)}{q_{Z_t|U_t X_t Y_t}(Z_t|U_t, X_t, Y_t)} \quad \text{for } t = 1, 2, \dots, n.$$

Step (b) follows from Equation (43). Step (c) follows from the definition of $\hat{\Omega}_n^{(\mu, \lambda, \alpha)}(p_{XY})$. Step (d) follows from that by Property 4 Part (a), the bound $|\mathcal{U}| \leq |\mathcal{X}|$, is sufficient to describe $\hat{\Omega}_n^{(\mu, \lambda, \alpha)}(p_{XY})$. Hence, we have the following:

$$\begin{aligned}
& \max_{q^n \in \mathcal{Q}^n} \frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \\
& \geq \frac{1}{n} \Omega^{(\mu, \lambda, \theta)}(p^{(n)}, \underline{Q}^n | p_{XY}) \stackrel{(a)}{=} -\frac{1}{n} \sum_{t=1}^n \log \Lambda_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)} \stackrel{(b)}{\geq} \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY})}{1 + \alpha}.
\end{aligned} \quad (46)$$

Step (a) follows from Equation (38) in Lemma 6. Step (b) follows from Equation (45). Since Equation (46) holds for any $n \geq 1$ and any $p^{(n)} \in \mathcal{P}^{(n)}(p_{XY})$, we have

$$\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY}) \geq \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY})}{1 + \alpha}.$$

Thus, we have Equation (44) in Proposition 2. \square

Proof of Theorem 3: For $\theta \in [0, 1)$, set

$$\alpha = \frac{\theta}{1 - \theta} \Leftrightarrow \theta = \frac{\alpha}{1 + \alpha}. \quad (47)$$

Then, we have the following:

$$\begin{aligned}
& \frac{1}{n} \log \left\{ \frac{5}{P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta)} \right\} \stackrel{(a)}{\geq} \frac{\underline{\Omega}^{(\mu, \lambda, \theta)}(p_{XY}) - \lambda \theta (\bar{\mu} R + \mu \Delta)}{1 + \theta(3 + \lambda \bar{\mu})} \\
& \stackrel{(b)}{\geq} \frac{\frac{1}{1 + \alpha} \Omega^{(\mu, \lambda, \alpha)}(p_{XY}) - \frac{\lambda \alpha}{1 + \alpha} (\bar{\mu} R + \mu \Delta)}{1 + \frac{\alpha}{1 + \alpha} (3 + \lambda \bar{\mu})} = \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY}) - \lambda \alpha (\bar{\mu} R + \mu \Delta)}{1 + \alpha + \alpha(3 + \lambda \bar{\mu})} \\
& = F^{(\mu, \lambda, \alpha)}(\bar{\mu} R + \mu \Delta | p_{XY}).
\end{aligned}$$

Step (a) follows from Corollary 3. Step (b) follows from Proposition 2 and Equation (47). Since the above bound holds for any positive $\lambda \geq 0$, $\mu \in [0, 1]$, and $\alpha \geq 0$, we have

$$\frac{1}{n} \log \left\{ \frac{5}{P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta)} \right\} \geq F(R, \Delta | p_{XY}).$$

Thus, Equation (10) in Theorem 3 is proved. \square

Proof of Corollary 2: Since g is an inverse function of ϑ , the definition in Equation (13) of κ_n is equivalent to

$$g\left(\frac{\kappa_n}{\rho(p_{XY})}\right) = \sqrt{\frac{10}{n\rho(p_{XY})} \log\left(\frac{5}{1 - \varepsilon}\right)}. \quad (48)$$

By the definition of $n_0 = n_0(\varepsilon, \rho(p_{XY}))$, we have that $\kappa_n \leq (1/2)\rho(p_{XY})$ for $n \geq n_0$. We assume that for $n \geq n_0$, $(R, \Delta) \in \mathcal{R}_{WZ}(n, \varepsilon | p_{XY})$. Then, there exists a sequence $\{(\varphi^{(n)}, \psi^{(n)})\}_{n \geq n_0}$ such that for $n \geq n_0$, we have

$$\frac{1}{n} \log \|\varphi^{(n)}\| \leq R, \quad P_e^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq \varepsilon.$$

Then, by Theorem 3, we have

$$1 - \varepsilon \leq P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) \leq 5 \exp\{-nF(R, \Delta | p_{XY})\} \quad (49)$$

for any $n \geq n_0$. We claim that for $n \geq n_0$, we have $(R + \kappa_n, \Delta + \kappa_n) \in \mathcal{R}(p_{XY})$. To prove this claim, we suppose that $(R + \kappa_{n^*}, \Delta + \kappa_{n^*})$ does not belong to $\mathcal{R}(p_{XY})$ for some $n^* \geq n_0$. Then, we have the following chain of inequalities:

$$\begin{aligned} 5 \exp [-n^* F(R, \Delta | p_{XY})] &\stackrel{(a)}{<} 5 \exp \left[-\frac{n^* \rho(p_{XY})}{10} \cdot g^2 \left(\frac{\kappa_{n^*}}{\rho(p_{XY})} \right) \right] \\ &\stackrel{(b)}{=} 5 \exp \left[-\frac{n^* \rho(p_{XY})}{10} \frac{10}{n^* \rho(p_{XY})} \log \left(\frac{5}{1-\varepsilon} \right) \right] = 5 \exp \left[\log \left(\frac{1-\varepsilon}{5} \right) \right] = 1 - \varepsilon. \end{aligned} \quad (50)$$

Step (a) follows from $\kappa_{n^*} \leq (1/2)\rho(p_{XY})$ and Property 4 Part (e). Step (b) follows from Equation (48). The bound of Equation (50) contradicts Equation (49). Hence, we have $(R + \kappa_n, \Delta + \kappa_n) \in \mathcal{R}(p_{XY})$ or equivalent to

$$(R, \Delta) \in \mathcal{R}(p_{XY}) - \kappa_n(1, 1)$$

for $n \geq n_0$, which implies that for $n \geq n_0$,

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

completing the proof. \square

5. Conclusions

For the WZ system, we have derived an explicit lower bound of the optimal exponent function on the correct probability of decoding for $(R, \Delta) \notin \mathcal{R}_{WZ}(p_{XY})$. We have described this result in Theorem 3. The determination problem of the optimal exponent remains to be resolved. This problem is our future work.

In this paper, we have treated the case where \mathcal{X} and \mathcal{Y} are finite sets. Extension of Theorem 3 to the case where \mathcal{X} and \mathcal{Y} are arbitrary sets is also our future work. Wyner [12] investigated the characterization of the rate distortion region in the case where \mathcal{X} and \mathcal{Y} are general sets and $\{(X_t, Y_t)\}_{t=1}^\infty$ is a correlated stationary memoryless source. This work may provide a good tool to investigate the second future work.

Acknowledgments: I am very grateful to Shun Watanabe for his helpful comments.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Properties of the Rate Distortion 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 Part (b).

Proof of Property 1 Part (b): We set

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

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

$$\bigcup_{m \geq 1} \underline{\mathcal{R}}_{WZ}(m, \varepsilon | p_{XY}) \subseteq \mathcal{R}_{WZ}(\varepsilon | p_{XY}). \quad (A1)$$

We next assume that $(R, \Delta) \in \mathcal{R}_{WZ}(\varepsilon | p_{XY})$. Set

$$\mathcal{R}_{WZ}^{(\delta)}(\varepsilon | p_{XY}) := \{(R + \delta, \Delta) : (R, \Delta) \in \mathcal{R}_{WZ}(\varepsilon | p_{XY})\}.$$

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

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

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

$$(R, \Delta) \notin \text{cl} \left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{WZ}(m, \varepsilon|p_{XY}) \right). \quad (\text{A3})$$

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

$$(R + \delta, \Delta) \notin \text{cl} \left(\bigcup_{m \geq 1} \underline{\mathcal{R}}_{WZ}(m, \varepsilon|p_{XY}) \right) \quad (\text{A4})$$

for some small $\delta > 0$. Note that $(R + \delta, \Delta) \in \mathcal{R}_{WZ}^{(\delta)}(\varepsilon|p_{XY})$. Then, Equation (A4) contradicts Equation (A2). Thus, we have

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

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

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

completing the proof. \square

Appendix B. Cardinality Bound on Auxiliary Random Variables

Set

$$\begin{aligned}R^{(\mu)}(p_{XY}) &:= \min_{q \in \mathcal{P}_{\text{sh}}(p_{XY})} \{ \bar{\mu} I_q(X; U|Y) + \mu E_q d(X, Z) \}, \\ \underline{R}^{(\mu)}(p_{XY}) &:= \min_{q \in \mathcal{P}(p_{XY})} \{ \bar{\mu} I_q(X; U|Y) + \mu E_q d(X, Z) \},\end{aligned}$$

Since $\mathcal{P}_{\text{sh}}(p_{XY}) \subseteq \mathcal{P}(p_{XY})$, it is obvious that

$$R^{(\mu)}(p_{XY}) \geq \underline{R}^{(\mu)}(p_{XY}). \quad (\text{A6})$$

We first prove the following lemma.

Lemma A1.

$$R^{(\mu)}(p_{XY}) = \underline{R}^{(\mu)}(p_{XY}).$$

To prove Lemma A1, we set

$$\begin{aligned}\mathcal{P}_1(p_{XY}) &:= \{ q_1 = q_{UXY} : |\mathcal{U}| \leq |\mathcal{X}|, q_{XY} = p_{XY}, U \leftrightarrow X \leftrightarrow Y \}, \\ \mathcal{Q}_1(p_{XY}) &:= \{ q_1 = q_{UXY} : |\mathcal{U}| \leq |\mathcal{X}| + 1, q_{XY} = p_{XY}, U \leftrightarrow X \leftrightarrow Y \}, \\ \mathcal{Q}_2(q_{UXY}) &:= \{ q_2 = q_{Z|UXY} : q_{UXYZ} = (q_{UXY}, q_2), X \leftrightarrow (U, Y) \leftrightarrow Z \}.\end{aligned}$$

By definition, it is obvious that

$$\mathcal{P}_{\text{sh}}(p_{XY}) = \{q = (q_1, q_2) : q_1 \in \mathcal{P}_1(p_{XY}), q_2 \in \mathcal{Q}_2(q_1)\}, \quad (\text{A7})$$

$$\mathcal{P}(p_{XY}) = \{q = (q_1, q_2) : q_1 \in \mathcal{Q}_1(p_{XY}), q_2 \in \mathcal{Q}_2(q_1)\}. \quad (\text{A8})$$

Proof of Lemma A1: Since we have Equation (A6), it suffices to show $R^{(\mu)}(p_{XY}) \leq \underline{R}^{(\mu)}(p_{XY})$ to prove Lemma A1. We bound the cardinality $|\mathcal{U}|$ of U to show that the bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $\underline{R}^{(\mu)}(p_{XY})$. We first observe that, by Equation (A8), we have

$$\begin{aligned} \underline{R}^{(\mu)}(p_{XY}) &= \min_{q_1 \in \mathcal{Q}_1(p_{XY})} \min_{q_2 \in \mathcal{Q}_2(q_1)} \left\{ \bar{\mu} I_{q_1}(X; U|Y) + \mu E_{(q_1, q_2)} d(X, Z) \right\} \\ &= \min_{q_1 \in \mathcal{Q}_1(p_{XY})} \left\{ \bar{\mu} I_{q_1}(X; U|Y) + \mu \min_{q_2 \in \mathcal{Q}_2(q_1)} E_{(q_1, q_2)} d(X, Z) \right\} \\ &= \min_{q_1 \in \mathcal{Q}_1(p_{XY})} \left\{ \bar{\mu} I_{q_1}(X; U|Y) + \mu E_{(q_1, q_2^*(q_1))} d(X, Z) \right\}, \end{aligned}$$

where

$$q_2^* = q_2^*(q_1) = q_{Z|UY}^* = \{q_{Z|UY}^*(z|u, y)\}_{(u, y, z) \in \mathcal{U} \times \mathcal{Y} \times \mathcal{Z}}$$

is a conditional probability distribution that attains the following optimization problem:

$$\min_{q_2 \in \mathcal{Q}_2(q_1)} E_{(q_1, q_2)} d(X, Z).$$

Observe that

$$p_X(x) = \sum_{u \in \mathcal{U}} q_U(u) q_{X|U}(x|u), \quad (\text{A9})$$

$$\bar{\mu} I_{q_1}(X; U|Y) + \mu E_{(q_1, q_2^*)}(X, Z) = \sum_{u \in \mathcal{U}} p_U(u) \pi(q_{X|U}(\cdot|u)), \quad (\text{A10})$$

where

$$\begin{aligned} \pi(q_{X|U}(\cdot|u)) &:= \sum_{(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}} q_{X|U}(x|u) p_{Y|X}(y|x) q_{Z|UY}^*(z|u, y) \\ &\quad \times \log \left\{ \frac{q_{X|U}^{\bar{\mu}}(x|u)}{p_X^{\bar{\mu}}(x)} \frac{p_Y^{\bar{\mu}}(y) e^{-\mu d(x, z)}}{\left[\sum_{\tilde{x} \in \mathcal{X}} p_{Y|X}(y|\tilde{x}) q_{X|U}(\tilde{x}|u) \right]^{\bar{\mu}}} \right\}. \end{aligned}$$

For each $u \in \mathcal{U}$, $\pi(q_{X|U}(\cdot|u))$ is a continuous function of $q_{X|U}(\cdot|u)$. Then, by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$$

is sufficient to express $|\mathcal{X}| - 1$ values of Equation (A9) and one value of Equation (A10). \square

Next, we give a proof of Property 4 Part (a).

Proof of Property 4 Part (a): We first bound the cardinality $|\mathcal{U}|$ of U in \mathcal{Q} to show that the bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $\Omega^{(\mu, \lambda, \alpha)}(p_{XY})$. Observe that

$$q_X(x) = \sum_{u \in \mathcal{U}} q_U(u) q_{X|U}(x|u), \quad (\text{A11})$$

$$\exp \left\{ -\Omega^{(\mu, \lambda, \alpha)}(q|p_{XY}) \right\} = \sum_{u \in \mathcal{U}} q_U(u) \Pi^{(\mu, \lambda, \alpha)}(q_X, q_{XYZ|U}(\cdot, \cdot, \cdot|u)), \quad (\text{A12})$$

where

$$\Pi^{(\mu, \lambda, \alpha)}(q_X, q_{XYZ|U}(\cdot, \cdot, \cdot|u)) := \sum_{\substack{(x,y,z) \\ \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}} q_{XYZ|U}(x, y, z|u) \exp \left\{ -\alpha \omega_{q||p}^{(\mu, \lambda)}(x, y, z|u) \right\}.$$

The value of q_X included in $\Pi^{(\mu, \lambda, \alpha)}(q_X, q_{XYZ|U}(\cdot, \cdot, \cdot|u))$ must be preserved under the reduction of \mathcal{U} . For each $u \in \mathcal{U}$, $\Pi^{(\mu, \lambda)}(q_X, q_{XYZ|U}(\cdot, \cdot, \cdot|u))$ is a continuous function of $q_{XYZ|U}(\cdot, \cdot, \cdot|u)$. Then, by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$$

is sufficient to express $|\mathcal{X}| - 1$ values of Equation (A11) and one value of Equation (A12). We next bound the cardinality $|\mathcal{U}|$ of U in $\mathcal{P}_{\text{sh}}(p_{XY})$ to show that the bound $|\mathcal{U}| \leq |\mathcal{X}|$ is sufficient to describe $\tilde{\Omega}^{(\mu, \lambda)}(p_{XY})$. Observe that

$$p_X(x) = \sum_{u \in \mathcal{U}} p_U(u) p_{X|U}(x|u), \quad (\text{A13})$$

$$\exp \left\{ -\tilde{\Omega}^{(\mu, \lambda)}(p) \right\} = \sum_{u \in \mathcal{U}} p_U(u) \tilde{\Pi}^{(\mu, \lambda)}(p_X, p_{XYZ|U}(\cdot, \cdot, \cdot|u)), \quad (\text{A14})$$

where

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

The value of p_X included in $\tilde{\Pi}^{(\mu, \lambda)}(p_X, p_{XYZ|U}(\cdot, \cdot, \cdot|u))$ must be preserved under the reduction of \mathcal{U} . For each $u \in \mathcal{U}$, $\tilde{\Pi}^{(\mu, \lambda)}(p_X, p_{XYZ|U}(\cdot, \cdot, \cdot|u))$ is a continuous function of $p_{XYZ|U}(\cdot, \cdot, \cdot|u)$. Then, by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$$

is sufficient to express $|\mathcal{X}| - 1$ values of Equation (A13) and one value of Equation (A14). \square

Appendix C. Proof of Property 2

In this Appendix, we prove Property 2. Property 2 Part (a) is a well known property. Proof of this property is omitted here. We only prove Property 2 Part (b).

Proof of Property 2 Part (b): Since $\mathcal{P}^*(p_{XY}) \subseteq \mathcal{P}(p_{XY})$, it is obvious that $\mathcal{R}^*(p_{XY}) \subseteq \mathcal{R}(p_{XY})$. Hence it suffices to prove that $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}^*(p_{XY})$. We assume that $(R, \Delta) \in \mathcal{R}(p_{XY})$. Then, there exists $p \in \mathcal{P}(p_{XY})$ such that

$$R \geq I_p(U; X|Y) \text{ and } \Delta \geq E_p d(X, Z). \quad (\text{A15})$$

On the second inequality in Equation (A15), we have the following:

$$\begin{aligned} \Delta \geq E_p d(X, Z) &= \sum_{(u,y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u, y) \left[\sum_{z \in \mathcal{Z}} p_{Z|UY}(z|u, y) \left(\sum_{x \in \mathcal{X}} d(x, z) p_{X|UY}(x|u, y) \right) \right] \\ &\geq \sum_{(u,y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u, y) \left[\min_{z \in \mathcal{Z}} \left(\sum_{x \in \mathcal{X}} d(x, z) p_{X|UY}(x|u, y) \right) \right] \\ &= \sum_{(u,y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u, y) \left(\sum_{x \in \mathcal{X}} d(x, z^*) p_{X|UY}(x|u, y) \right), \end{aligned} \quad (\text{A16})$$

where $z^* = z^*(u, y)$ is one of the minimizers of the function

$$\sum_{x \in \mathcal{X}} d(x, z) p_{X|UY}(x|u, y).$$

Define $\phi : \mathcal{U} \times \mathcal{Y} \rightarrow \mathcal{Z}$ by $\phi(u, y) = z^*$. We further define $q = q_{UXYZ}$ by $q_{UXY} = p_{UXY}$, $q_Z = q_{\phi(U, Y)}$. It is obvious that

$$q \in \mathcal{P}^*(p_{XY}) \text{ and } R \geq I_p(X; U|Y) = I_q(X; U|Y). \quad (\text{A17})$$

Furthermore, from Equation (A16), we have

$$\Delta \geq \mathbb{E}_p d(X, Z) \geq \mathbb{E}_q d(X, \phi(U, Y)). \quad (\text{A18})$$

From Equations (A17) and (A18), we have $(R, \Delta) \in \mathcal{R}^*(p_{XY})$. Thus $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}^*(p_{XY})$ is proved. \square

Appendix D. Proof of Property 3

In this Appendix, we prove Property 3. From Property 2 Part (a), we have the following lemma.

Lemma A2. Suppose that $(\hat{R}, \hat{\Delta})$ does not belong to $\mathcal{R}(p_{XY})$. Then, there exist $\epsilon > 0$ and $\mu^* \in [0, 1]$ such that for any $(R, \Delta) \in \mathcal{R}(p_{XY})$ we have

$$\bar{\mu}^*(R - \hat{R}) + \mu^*(\Delta - \hat{\Delta}) - \epsilon \geq 0.$$

Proof of this lemma is omitted here. Lemma A2 is equivalent to the fact that if the region $\mathcal{R}(p_{XY})$ is a convex set, then for any point $(\hat{R}, \hat{\Delta})$ outside the region $\mathcal{R}(p_{XY})$, there exists a line which separates the point $(\hat{R}, \hat{\Delta})$ from the region $\mathcal{R}(p_{XY})$. Lemma A2 will be used to prove Equation (8) in Property 3.

Proof of Equation (8) in Property 3: We first recall the following definitions of $\mathcal{P}(p_{XY})$ and $\mathcal{P}_{\text{sh}}(p_{XY})$:

$$\begin{aligned} \mathcal{P}(p_{XY}) &:= \{p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}| + 1, U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}, \\ \mathcal{P}_{\text{sh}}(p_{XY}) &:= \{p_{UXYZ} : |\mathcal{U}| \leq |\mathcal{X}|, U \leftrightarrow X \leftrightarrow Y, X \leftrightarrow (U, Y) \leftrightarrow Z\}. \end{aligned}$$

We prove $\mathcal{R}_{\text{sh}}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$. We assume that $(\hat{R}, \hat{\Delta}) \notin \mathcal{R}(p_{XY})$. Then, by Lemma A2, there exist $\epsilon > 0$ and $\mu^* \in [0, 1]$ such that for any $(R, \Delta) \in \mathcal{R}(p_{XY})$, we have

$$\bar{\mu}^* \hat{R} + \mu^* \hat{\Delta} \leq \bar{\mu}^* R + \mu^* \Delta - \epsilon.$$

Hence, we have

$$\begin{aligned} \bar{\mu}^* \hat{R} + \mu^* \hat{\Delta} &\leq \min_{(R, \Delta) \in \mathcal{R}(p_{XY})} \{\bar{\mu}^* R + \mu^* \Delta\} - \epsilon \\ &\stackrel{(a)}{=} \min_{p \in \mathcal{P}(p_{XY})} \{\bar{\mu}^* I_p(U; X|Y) + \mu^* \mathbb{E}_p d(X, Z)\} - \epsilon \\ &\leq \min_{p \in \mathcal{P}_{\text{sh}}(p_{XY})} \{\bar{\mu}^* I_p(U; X|Y) + \mu^* \mathbb{E}_p d(X, Z)\} - \epsilon = R^{(\mu^*)}(p_{XY}) - \epsilon. \end{aligned} \quad (\text{A19})$$

Step (a) follows from the definition of $\mathcal{R}(p_{XY})$. The inequality in Equation (A19) implies that $(\hat{R}, \hat{\Delta}) \notin \mathcal{R}_{\text{sh}}(p_{XY})$. Thus, $\mathcal{R}_{\text{sh}}(p_{XY}) \subseteq \mathcal{R}(p_{XY})$ is concluded. We next prove $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}_{\text{sh}}(p_{XY})$. We assume that $(R, \Delta) \in \mathcal{R}(p_{XY})$. Then, there exists $q \in \mathcal{P}(p_{XY})$ such that

$$R \geq I_q(X; U|Y), \Delta \geq \mathbb{E}_q d(X, Z). \quad (\text{A20})$$

Then, for each $\mu \in [0, 1]$ and for $(R, \Delta) \in \mathcal{R}(p_{XY})$, we have the following chain of inequalities:

$$\begin{aligned} \bar{\mu}R + \mu\Delta &\stackrel{(a)}{\geq} \bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z) \\ &\geq \min_{q \in \mathcal{P}(p_{XY})} [\bar{\mu}I_q(X; U|Y) + \mu E_q d(X, Z)] = \underline{R}^{(\mu)}(p_{XY}) \stackrel{(b)}{=} R^{(\mu)}(p_{XY}). \end{aligned}$$

Step (a) follows from Equation (A20). Step (b) follows from Lemma A1. Hence, we have $\mathcal{R}(p_{XY}) \subseteq \mathcal{R}_{\text{sh}}(p_{XY})$. \square

Appendix E. Proof of Property 4 Part (b)

In this Appendix, we prove Property 4 Part (b). We have the following lemma.

Lemma A3. For any $\mu \in [0, 1]$, $\lambda \geq 0$, $\alpha \in [0, \frac{1}{\lambda+1}]$ and any $q = q_{UXY} \in \mathcal{Q}(p_{Y|X})$, there exists $p = p_{UXYZ} \in \mathcal{P}_{\text{sh}}(p_{XY})$ such that

$$\Omega^{(\mu, \gamma, \alpha)}(q|p_{XY}) \geq \alpha \tilde{\Omega}^{(\mu, \lambda)}(p). \quad (\text{A21})$$

This implies that for any $\mu \in [0, 1]$, $\lambda \geq 0$, $\alpha \in [0, \frac{1}{\lambda+1}]$, we have

$$\Omega^{(\mu, \gamma, \alpha)}(p_{XY}) \geq \alpha \tilde{\Omega}^{(\mu, \lambda)}(p_{XY}). \quad (\text{A22})$$

Proof. Since Equation (A22) is obvious from Equation (A21), we only prove Equation (A21). We consider the case where (μ, λ, α) satisfies $(\mu, \lambda) \in [0, 1]$, $\lambda \geq 0$, and $\alpha \in [0, \frac{1}{1+\lambda}]$. In this case, we have

$$\lambda \bar{\mu} \frac{\alpha}{\bar{\alpha}} \leq \frac{\lambda \alpha}{\bar{\alpha}} \leq \frac{\frac{\lambda}{1+\lambda}}{1 - \frac{1}{1+\lambda}} = 1. \quad (\text{A23})$$

For each $q = q_{UXYZ} \in \mathcal{Q}$, we choose $p = p_{UXYZ} \in \mathcal{P}_{\text{sh}}(p_{XY})$ so that $p_{U|X} = q_{U|X}$ and $p_{Z|UY} = q_{Z|UY}$. Then, for any (μ, λ, α) satisfying $\mu \in [0, 1]$, $\lambda \geq 0$, and $\alpha \in [0, \frac{1}{1+\lambda}]$, we have the following chain of inequalities:

$$\begin{aligned} &\exp \left\{ -\Omega^{(\mu, \lambda, \alpha)}(q|p_{XY}) \right\} \\ &= E_q \left[\left\{ \frac{p_X(X)}{q_X(X)} \frac{p_{Y|X}(Y|X)}{q_{Y|XU}(Y|X, U)} \frac{q_{Z|UY}(Z|U, Y)}{q_{Z|UYX}(Z|U, Y, X)} \frac{p_{X|Y}^{\lambda \bar{\mu}}(X|Y) e^{-\lambda \mu d(X, Z)}}{q_{X|UY}^{\lambda \bar{\mu}}(X|U, Y)} \right\}^\alpha \right] \\ &= E_q \left[\left\{ \frac{p_{UXYZ}(X, Y, Z, U)}{q_{UXYZ}(X, Y, Z, U)} \frac{p_{X|Y}^{\lambda \bar{\mu}}(X|Y) e^{-\lambda \mu d(X, Z)}}{p_{X|UY}^{\lambda \bar{\mu}}(X|U, Y)} \right\}^\alpha \left\{ \frac{p_{X|UY}(X|U, Y)}{q_{X|UY}(X|U, Y)} \right\}^{\lambda \bar{\mu} \alpha} \right] \\ &\stackrel{(a)}{\leq} \left(E_q \left[\frac{p_{UXYZ}(X, Y, Z, U)}{q_{UXYZ}(X, Y, Z, U)} \frac{p_{X|Y}^{\lambda \bar{\mu}}(X|Y) e^{-\lambda \mu d(X, Z)}}{p_{X|UY}^{\lambda \bar{\mu}}(X|U, Y)} \right]^\alpha \left(E_q \left[\left\{ \frac{p_{X|UY}(X|U, Y)}{q_{X|UY}(X|U, Y)} \right\}^{\lambda \bar{\mu} \frac{\alpha}{\bar{\alpha}}} \right] \right)^{\bar{\alpha}} \right) \\ &= \exp \left\{ -\alpha \tilde{\Omega}^{(\mu, \lambda)}(p) \right\} \left(E_q \left[\left\{ \frac{p_{X|UY}(X|U, Y)}{q_{X|UY}(X|U, Y)} \right\}^{\lambda \bar{\mu} \frac{\alpha}{\bar{\alpha}}} \right] \right)^{\bar{\alpha}} \stackrel{(b)}{\leq} \exp \left\{ -\alpha \tilde{\Omega}^{(\mu, \lambda)}(p) \right\}. \quad (\text{A24}) \end{aligned}$$

Step (a) follows from Hölder's inequality. Step (b) follows from Equation (A23) and Hölder's inequality. \square

Proof of Property 4 Part (b): We have the following chain of inequalities:

$$\begin{aligned}
 F(R, \Delta | p_{XY}) &= \sup_{\mu \in [0,1], \lambda \geq 0} \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY}) - \lambda \alpha (\bar{\mu} R + \mu \Delta)}{1 + (4 + \bar{\mu} \lambda) \alpha} \\
 &\geq \sup_{\mu \in [0,1], \lambda \geq 0} \sup_{\alpha \in [0, \frac{1}{1+\bar{\mu}}]} \frac{\Omega^{(\mu, \lambda, \alpha)}(p_{XY}) - \lambda \alpha (\bar{\mu} R + \mu \Delta)}{1 + (4 + \bar{\mu} \lambda) \alpha} \\
 &\stackrel{(a)}{\geq} \sup_{\mu \in [0,1], \lambda \geq 0} \sup_{\alpha \in [0, \frac{1}{1+\bar{\mu}}]} \frac{\alpha [\tilde{\Omega}^{(\mu, \lambda)}(p_{XY}) - \lambda (\bar{\mu} R + \mu \Delta)]}{1 + (4 + \bar{\mu} \lambda) \alpha} \\
 &\stackrel{(b)}{=} \sup_{\mu \in [0,1], \lambda \geq 0} \frac{\tilde{\Omega}^{(\mu, \lambda)}(p_{XY}) - \lambda (\bar{\mu} R + \mu \Delta)}{5 + \lambda (1 + \bar{\mu})} = \tilde{F}(R, \Delta | p_{XY}).
 \end{aligned}$$

Step (a) follows from Lemma A3. Step (b) follows from

$$\sup_{\alpha \in [0, \frac{1}{1+\bar{\mu}}]} \frac{\alpha}{1 + (4 + \bar{\mu} \lambda) \alpha} = \frac{1}{5 + \lambda (1 + \bar{\mu})},$$

completing the proof. \square

Appendix F. Proof of Property 4 Parts (c), (d), and (e)

In this Appendix, we prove Property 4 Parts (c), (d), and (e). We first prove Parts (c) and (d).

Proof of Property 4 Parts (c) and (d): We first prove Part (c). We first show that for each $p \in \mathcal{P}_{\text{sh}}(p_{XY})$ and $\mu \in [0, 1]$, $\tilde{\Omega}^{(\mu, \lambda)}(p)$ exists for $\lambda \in [0, 1]$. On a lower bound of $\exp[-\tilde{\Omega}^{(\mu, \lambda)}(p)]$, for $\lambda \in [0, 1]$, we have the following:

$$\exp[-\tilde{\Omega}^{(\mu, \lambda)}(p)] = \sum_{\substack{(u, x, y, z) \\ \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}} p_{UXYZ}(u, x, y, z) \left[\frac{p_{X|Y}(x|y)}{p_{X|UY}(x|u, y)} \right]^{\bar{\mu} \lambda} e^{-\mu \lambda d(x, z)} \quad (\text{A25})$$

$$\stackrel{(a)}{\geq} \sum_{\substack{(u, x, y, z) \\ \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Y}}} p_{UXYZ}(u, x, y, z) p_{X|Y}(x|y) e^{-d(x, z)}. \quad (\text{A26})$$

Step (a) follows from that, for $\mu, \lambda \in [0, 1]$,

$$\left[\frac{p_{X|Y}(x|y)}{p_{X|UY}(x|u, y)} \right]^{\bar{\mu} \lambda} e^{-\mu \lambda d(x, z)} \geq p_{X|Y}(x|y) e^{-d(x, z)}.$$

It is obvious that the lower bound of Equation (A26) of $\exp[-\tilde{\Omega}^{(\mu, \lambda)}(p)]$ takes some positive value. This implies that $\tilde{\Omega}^{(\mu, \lambda)}(p)$ exists for $\lambda \in [0, 1]$. 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 $\lambda \in [0, 1]$, we have the following chain of inequalities:

$$\begin{aligned}
\exp[-\tilde{\Omega}^{(\mu,\lambda)}(p)] &\stackrel{(a)}{\leq} \sum_{\substack{(u,x,y,z) \\ \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}} p_{UXYZ}(u,x,y,z) \left[\frac{p_{X|Y}(x|y)}{p_{X|UY}(x|u,y)} \right]^{\bar{\mu}\lambda} \\
&= \sum_{\substack{(u,x,y) \\ \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y}}} p_{UY}(u,y) p_{X|UY}^{1-\bar{\mu}\lambda}(x|u,y) p_{X|Y}^{\bar{\mu}\lambda}(x|y) \\
&\stackrel{(b)}{\leq} \sum_{(u,y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u,y) \left(\sum_{x \in \mathcal{X}} p_{X|UY}(x|u,y) \right)^{1-\bar{\mu}\lambda} \left(\sum_{x \in \mathcal{X}} p_{X|Y}(x|y) \right)^{\bar{\mu}\lambda} \\
&= \sum_{(u,y) \in \mathcal{U} \times \mathcal{Y}} p_{UY}(u,y) = 1.
\end{aligned} \tag{A27}$$

Step (a) follows from Equation (A25) and $e^{-\mu\lambda d(x,z)} \leq 1$. Step (b) follows from $\bar{\mu}\lambda \in [0, 1]$ and Hölder's inequality. We next prove that that, for each $p \in \mathcal{P}_{\text{sh}}(p_{XY})$ and $\mu \in [0, 1]$, $\tilde{\Omega}^{(\mu,\lambda)}(p)$ is twice differentiable for $\lambda \in [0, 1/2]$. For simplicity of notations, set

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

Then, we have

$$\tilde{\Omega}^{(\mu,\lambda)}(p) = \zeta(\lambda) = -\log \left[\sum_{\underline{a} \in \underline{\mathcal{A}}} p_{\underline{A}}(\underline{a}) e^{-\lambda \varsigma(\underline{a})} \right]. \tag{A28}$$

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

$$p^{(\lambda)}(\underline{a}) = e^{\zeta(\lambda)} p(\underline{a}) e^{-\lambda \varsigma(\underline{a})}. \tag{A29}$$

By simple computations, we have

$$\begin{aligned}
\zeta'(\lambda) &= e^{\zeta(\lambda)} \left[\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma(\underline{a}) e^{-\lambda \varsigma(\underline{a})} \right] = \sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma(\underline{a}), \\
\zeta''(\lambda) &= -e^{2\zeta(\lambda)} \left[\sum_{\underline{a}, \underline{b} \in \underline{\mathcal{A}}} p(\underline{a}) p(\underline{b}) \frac{\{\varsigma(\underline{a}) - \varsigma(\underline{b})\}^2}{2} e^{-\lambda \{\varsigma(\underline{a}) + \varsigma(\underline{b})\}} \right] \\
&= - \sum_{\underline{a}, \underline{b} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) p^{(\lambda)}(\underline{b}) \frac{\{\varsigma(\underline{a}) - \varsigma(\underline{b})\}^2}{2} = - \sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma^2(\underline{a}) + \left[\sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma(\underline{a}) \right]^2 \leq 0.
\end{aligned} \tag{A30}$$

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

$$\begin{aligned}
-\zeta''(\lambda) &\stackrel{(a)}{\leq} \sum_{\underline{a} \in \underline{\mathcal{A}}} p^{(\lambda)}(\underline{a}) \varsigma^2(\underline{a}) \stackrel{(b)}{=} \sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^2(\underline{a}) e^{-\lambda \varsigma(\underline{a}) + \zeta(\lambda)} \\
&= e^{\zeta(\lambda)} \sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \sqrt{e^{-2\lambda \varsigma(\underline{a})}} \sqrt{\varsigma^4(\underline{a})} \stackrel{(c)}{\leq} \sqrt{e^{2\zeta(\lambda) - \zeta(2\lambda)}} \sqrt{\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^4(\underline{a})} \\
&\stackrel{(d)}{\leq} \sqrt{e^{2\zeta(\lambda)}} \sqrt{\sum_{\underline{a} \in \underline{\mathcal{A}}} p(\underline{a}) \varsigma^4(\underline{a})}.
\end{aligned} \tag{A31}$$

Step (a) follows from Equation (A30). Step (b) follows from Equation (A29). Step (c) follows from Cauchy-Schwarz inequality and Equation (A28). Step (d) follows from that $\zeta(2\lambda) \geq 0$ for $2\lambda \in [0, 1]$. Note that $\zeta(\lambda)$ exists for $\lambda \in [0, 1/2]$. Furthermore, we have the following:

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

Hence, by Equation (A31), $\zeta''(\lambda)$ exists for $\lambda \in [0, 1/2]$. We next prove Part (d). We derive the lower bound Equation (9) of $\tilde{\Omega}^{(\mu, \lambda)}(p_{XY})$. Fix any $(\mu, \lambda) \in [0, 1] \times [0, 1/2]$ and any $p \in \mathcal{P}_{\text{sh}}(p_{XY})$. By the Taylor expansion of $\zeta(\lambda) = \tilde{\Omega}^{(\mu, \lambda)}(p)$ with respect to λ around $\lambda = 0$, we have that, for any $p \in \mathcal{P}_{\text{sh}}(p_{XY})$ and for some $\nu \in [0, \lambda]$,

$$\begin{aligned} \tilde{\Omega}^{(\mu, \lambda)}(p) &= \zeta(0) + \zeta'(0)\lambda + \frac{1}{2}\zeta''(\nu)\lambda^2 \\ &= \lambda \mathbb{E}_p \left[\tilde{\omega}_p^{(\mu)}(X, Y, Z|U) \right] - \frac{\lambda^2}{2} \text{Var}_{p^{(\nu)}} \left[\tilde{\omega}_p^{(\mu)}(X, Y, Z|U) \right] \\ &\stackrel{(a)}{\geq} \lambda R^{(\mu)}(p_{XY}) - \frac{\lambda^2}{2} \text{Var}_{p^{(\nu)}} \left[\tilde{\omega}_p^{(\mu)}(X, Y, Z|U) \right]. \end{aligned} \quad (\text{A32})$$

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

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

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

By this definition, we have that

$$\tilde{\Omega}^{(\mu, \lambda)}(p_{\text{opt}}) = \tilde{\Omega}^{(\mu, \lambda)}(p_{XY}) \quad (\text{A33})$$

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

$$\text{Var}_{p_{\text{opt}}^{(\nu)}} \left[\tilde{\omega}_{p_{\text{opt}}}^{(\mu)}(X, Y, Z|U) \right] \leq \text{Var}_{p_{\text{opt}}^{(\nu_{\text{opt}})}} \left[\tilde{\omega}_{p_{\text{opt}}}^{(\mu)}(X, Y, Z|U) \right] = \rho^{(\mu, \lambda)}(p_{XY}). \quad (\text{A34})$$

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

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

Step (a) follows from Equation (A33). Step (b) follows from Equation (A32). Step (c) follows from Equation (A34). Step (d) follows from the definition of $\rho(p_{XY})$. \square

We next prove Part (e). For the proof we use the following lemma.

Lemma A4. When $\tau \in (0, (1/2)\rho]$, the maximum of

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

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

$$\vartheta(\lambda_0) := \frac{1}{5}\lambda_0^2 + \lambda_0 = \frac{\tau}{\rho}. \quad (\text{A35})$$

Let $g(a)$ be the inverse function of $\vartheta(a)$ for $a \geq 0$. Then, the condition of Equation (A35) is equivalent to $\lambda_0 = g(\frac{\tau}{\rho})$. The maximum is given by

$$\frac{1}{5+2\lambda_0} \left\{ -\frac{1}{2}\rho\lambda_0^2 + \tau\lambda_0 \right\} = \frac{\rho}{10}\lambda_0^2 = \frac{\rho}{10}g^2 \left(\frac{\tau}{\rho} \right).$$

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

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

$$\bar{\mu}^*(R + \tau) + \mu^*(\Delta + \tau) < R^{(\mu^*)}(p_{XY})$$

or equivalent to

$$R^{(\mu^*)}(p_{XY}) - (\bar{\mu}^*R + \mu^*\Delta) > \tau \quad (\text{A36})$$

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

$$\begin{aligned} \tilde{F}(R, \Delta | p_{XY}) &\geq \sup_{\lambda \in (0, 1/2]} \tilde{F}^{(\mu^*, \lambda)}(\bar{\mu}^*R + \mu^*\Delta | p_{XY}) \\ &= \sup_{\lambda \in (0, 1/2]} \frac{\tilde{\Omega}^{(\mu^*, \lambda)}(p_{XY}) - \lambda(\bar{\mu}^*R + \mu^*\Delta)}{5 + \lambda(1 + \bar{\mu}^*)} \\ &\stackrel{(a)}{\geq} \sup_{\lambda \in (0, 1/2]} \frac{1}{5 + 2\lambda} \left\{ -\frac{1}{2}\rho\lambda^2 + \lambda R^{(\mu^*)}(p_{XY}) - \lambda(\bar{\mu}^*R + \mu^*\Delta) \right\} \\ &\stackrel{(b)}{>} \sup_{\lambda \in (0, 1/2]} \frac{1}{5 + 2\lambda} \left\{ -\frac{1}{2}\rho\lambda^2 + \tau\lambda \right\} \stackrel{(c)}{=} \frac{\rho}{10}g^2 \left(\frac{\tau}{\rho} \right). \end{aligned}$$

Step (a) follows from Property 4 Part (d). Step (b) follows from Equation (A36). Step (c) follows from Lemma A4. \square

Appendix G. Proof of Lemma 1

To prove Lemma 1, we prepare a lemma. Set

$$\begin{aligned} \mathcal{A}_n &:= \left\{ x^n : \frac{1}{n} \log \frac{p_{X^n}(x^n)}{Q_{X^n}^{(i)}(x^n)} \geq -\eta \right\}, \\ \mathcal{A}_n &:= \tilde{\mathcal{A}}_n \times \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{Z}^n, \mathcal{A}_n^c := \tilde{\mathcal{A}}_n^c \times \mathcal{M}_n \times \mathcal{Y}^n \times \mathcal{Z}^n, \\ \tilde{\mathcal{B}}_n &:= \left\{ (s, x^n, y^n) : \frac{1}{n} \log \frac{p_{Y^n|X^n}(y^n|x^n)}{Q_{Y^n|X^n S_n}^{(ii)}(y^n|x^n, s)} \geq -\eta \right\}, \\ \mathcal{B}_n &:= \tilde{\mathcal{B}}_n \times \mathcal{Z}^n, \mathcal{B}_n^c := \tilde{\mathcal{B}}_n^c \times \mathcal{Z}^n, \\ \mathcal{C}_n &:= \left\{ (s, x^n, y^n, z^n) : \frac{1}{n} \log \frac{p_{X^n|S_n Y^n}(x^n|s, y^n)}{Q_{X^n|S_n Y^n Z^n}^{(iii)}(x^n|s, y^n, z^n)} \geq -\eta \right\}, \\ \tilde{\mathcal{D}}_n &:= \{(s, x^n, y^n) : s = \varphi^{(n)}(x^n), Q_{X^n|S_n Y^n}^{(iv)}(x^n|s, y^n) \leq M_n e^{n\eta} p_{X^n|Y^n}(x^n|y^n)\}, \\ \mathcal{D}_n &:= \tilde{\mathcal{D}}_n \times \mathcal{Z}^n, \mathcal{D}_n^c := \tilde{\mathcal{D}}_n^c \times \mathcal{Z}^n. \end{aligned}$$

Then, we have the following lemma:

Lemma A5.

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{A}_n^c) &\leq e^{-n\eta}, p_{S_n X^n Y^n Z^n}(\mathcal{B}_n^c) \leq e^{-n\eta}, \\ p_{S_n X^n Y^n Z^n}(\mathcal{C}_n^c) &\leq e^{-n\eta}, p_{S_n X^n Y^n Z^n}(\mathcal{D}_n^c) \leq e^{-n\eta}. \end{aligned}$$

Proof. We first prove the first inequality. We have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{A}_n^c) &= p_{X^n}(\tilde{\mathcal{A}}_n^c) = \sum_{x^n \in \tilde{\mathcal{A}}_n^c} p_{X^n}(x^n) \\ &\stackrel{(a)}{\leq} \sum_{x^n \in \tilde{\mathcal{A}}_n^c} e^{-n\eta} Q_{X^n}^{(i)}(x^n) \leq e^{-n\eta} \sum_{x^n} Q_{X^n}^{(i)}(x^n) = e^{-n\eta}. \end{aligned}$$

Step (a) follows from the definition of \mathcal{A}_n . We next prove the second inequality. We have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{B}_n^c) &= p_{S_n X^n Y^n}(\tilde{\mathcal{B}}_n^c) \stackrel{(a)}{=} \sum_{(s, x^n, y^n) \in \tilde{\mathcal{B}}_n^c} p_{S_n X^n}(s, x^n) p_{Y^n|X^n}(y^n|x^n) \\ &\stackrel{(b)}{\leq} \sum_{(s, x^n, y^n) \in \tilde{\mathcal{B}}_n^c} e^{-n\eta} p_{S_n X^n}(s, x^n) Q_{Y^n|S_n X^n}^{(ii)}(y^n|s, x^n) \\ &\leq e^{-n\eta} \sum_{s, x^n, y^n} p_{S_n X^n}(s, x^n) Q_{Y^n|S_n X^n}^{(ii)}(y^n|s, x^n) = e^{-n\eta}. \end{aligned}$$

Step (a) follows from the Markov chain $S_n \leftrightarrow X^n \leftrightarrow Y^n$. Step (b) follows from the definition of \mathcal{B}_n . On the third inequality, we have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{C}_n^c) &\stackrel{(a)}{=} \sum_{(s, x^n, y^n, z^n) \in \mathcal{C}_n^c} p_{X^n|S_n Y^n}(x^n|s, y^n) p_{S_n Y^n Z^n}(s, y^n, z^n) \\ &\stackrel{(b)}{\leq} \sum_{(s, x^n, y^n, z^n) \in \mathcal{C}_n^c} e^{-n\eta} Q_{X^n|S_n Y^n Z^n}^{(iii)}(x^n|s, y^n, z^n) p_{S_n Y^n Z^n}(s, y^n, z^n) \\ &\leq e^{-n\eta} \sum_{s, x^n, y^n, z^n} Q_{X^n|S_n Y^n Z^n}^{(iii)}(x^n|s, y^n, z^n) p_{S_n Y^n Z^n}(s, y^n, z^n) = e^{-n\eta}. \end{aligned}$$

Step (a) follows from the Markov chain $X^n \leftrightarrow S_n Y^n \leftrightarrow Z^n$. Step (b) follows from the definition of \mathcal{C}_n . We finally prove the fourth inequality. We have the following chain of inequalities:

$$\begin{aligned} p_{S_n X^n Y^n Z^n}(\mathcal{D}_n^c) &= p_{S_n X^n Y^n}(\tilde{\mathcal{D}}_n^c) = \sum_{s \in \mathcal{M}_n} \sum_{\substack{(x^n, y^n): \varphi^{(n)}(x^n) = s, \\ p_{X^n|Y^n}(x^n|y^n) \\ \leq (1/M_n)e^{-n\eta} \\ \times Q_{X^n|S_n, Y^n}^{(iv)}(x^n|s, y^n)}} p_{X^n|Y^n}(x^n|y^n) p_{Y^n}(y^n) \\ &\leq \frac{e^{-n\eta}}{M_n} \sum_{s \in \mathcal{M}_n} \sum_{\substack{(x^n, y^n): \varphi^{(n)}(x^n) = s, \\ p_{X^n|Y^n}(x^n|y^n) \\ \leq (1/M_n)e^{-n\eta} \\ \times Q_{X^n|S_n, Y^n}^{(iv)}(x^n|s, y^n)}} Q_{X^n|S_n, Y^n}^{(iv)}(x^n|s, y^n) p_{Y^n}(y^n) \\ &\leq \frac{e^{-n\eta}}{M_n} \sum_{s \in \mathcal{M}_n} \sum_{x^n, y^n} Q_{X^n|S_n, Y^n}^{(iv)}(x^n|s, y^n) p_{Y^n}(y^n) = e^{-n\eta}. \end{aligned}$$

□

Proof of Lemma 1: We set

$$\mathcal{E}_n := \left\{ (s, x^n, y^n, z^n) : \frac{1}{n} d(X^n, Z^n) \leq \Delta \right\}.$$

Set $R^{(n)} := (1/n) \log M_n$. By definition, we have

$$\begin{aligned}
& p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) \\
&= p_{S_n X^n Y^n Z^n} \left\{ \eta \geq \frac{1}{n} \log \frac{Q_{X^n}^{(i)}(X^n)}{p_{X^n}(X^n)}, \right. \\
&\quad \eta \geq \frac{1}{n} \log \frac{Q_{Y^n|X^n S}^{(ii)}(Y^n|X^n S)}{p_{Y^n|X^n}(Y^n|X^n)}, \\
&\quad \eta \geq \frac{1}{n} \log \frac{Q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n Y^n Z^n)}{p_{X^n|S_n Y^n}(X^n|S_n Y^n)}, \\
&\quad R^{(n)} + \eta \geq \frac{1}{n} \log \frac{Q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n Y^n)}{p_{X^n|Y^n}(X^n|Y^n)}, \\
&\quad \left. \Delta \geq \frac{1}{n} \log \exp \{d(X^n, Z^n)\} \right\}. \tag{A37}
\end{aligned}$$

Then, for any $(\varphi^{(n)}, \psi^{(n)})$ satisfying

$$R^{(n)} = \frac{1}{n} \log M_n \leq R,$$

we have

$$\begin{aligned}
& p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) \\
&\leq p_{S_n X^n Y^n Z^n} \left\{ \eta \geq \frac{1}{n} \log \frac{Q_{X^n}^{(i)}(X^n)}{p_{X^n}(X^n)}, \right. \\
&\quad \eta \geq \frac{1}{n} \log \frac{Q_{Y^n|X^n S}^{(ii)}(Y^n|X^n S)}{p_{Y^n|X^n}(Y^n|X^n)}, \\
&\quad \eta \geq \frac{1}{n} \log \frac{Q_{X^n|S_n Y^n Z^n}^{(iii)}(X^n|S_n Y^n Z^n)}{p_{X^n|S_n Y^n}(X^n|S_n Y^n)}, \\
&\quad R + \eta \geq \frac{1}{n} \log \frac{Q_{X^n|S_n Y^n}^{(iv)}(X^n|S_n Y^n)}{p_{X^n|Y^n}(X^n|Y^n)}, \\
&\quad \left. \Delta \geq \frac{1}{n} \log \exp \{d(X^n, Z^n)\} \right\}. \tag{A38}
\end{aligned}$$

Hence, it suffices to show

$$P_c^{(n)}(\varphi_1^{(n)}, \varphi_2^{(n)}, \psi^{(n)}; \Delta) \leq p_{S_n X^n Y^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) + 4e^{-n\eta}$$

to prove Lemma 1. By definition, we have

$$P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = p_{S_n X^n Y^n Z^n} (\mathcal{E}_n).$$

Then, we have the following:

$$\begin{aligned}
& P_c^{(n)}(\varphi^{(n)}, \psi^{(n)}; \Delta) = p_{S_n X^n Y^n Z^n} (\mathcal{E}_n) \\
&= p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) + p_{S_n X^n Y^n Z^n} ([\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n]^c \cap \mathcal{E}_n) \\
&\leq p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) + p_{S_n X^n Y^n Z^n} (\mathcal{A}_n^c) + p_{S_n X^n Y^n Z^n} (\mathcal{B}_n^c) \\
&\quad + p_{S_n X^n Y^n Z^n} (\mathcal{C}_n^c) + p_{S_n X^n Y^n Z^n} (\mathcal{D}_n^c) \\
&\stackrel{(a)}{\leq} p_{S_n X^n Y^n Z^n} (\mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n \cap \mathcal{E}_n) + 4e^{-n\eta}.
\end{aligned}$$

Step (a) follows from Lemma A5. \square

Appendix H. Proof of Lemma 2

In this Appendix, we prove Lemma 2.

Proof of Lemma 2: We have the following chain of inequalities:

$$\begin{aligned} I(X_t; Y^{t-1} | S_n X_{t-1} Y_t^n) &= H(Y^{t-1} | S_n X_{t-1} Y_t^n) - H(Y^{t-1} | S_n X_t Y_t^n) \\ &\leq H(Y^{t-1} | X^{t-1}) - H(Y^{t-1} | S_n X^n Y_t^n) \stackrel{(a)}{=} H(Y^{t-1} | X^{t-1}) - H(Y^{t-1} | X^n Y_t^n) \\ &\stackrel{(b)}{=} H(Y^{t-1} | X^{t-1}) - H(Y^{t-1} | X^{t-1}) = 0. \end{aligned}$$

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

Appendix I. Proof of Lemma 5

In this Appendix, we prove Equations (32) and (33) in Lemma 5.

Proofs of Equations (32) and (33) in Lemma 5: By the definition of $p_{X^t Z^t | S_n Y^n, q^t}^{(\mu, \lambda, \theta)}(x^t, z^t | s, y^n)$, for $t = 1, 2, \dots, n$, we have

$$p_{X^t Z^t | S_n Y^n, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(x^t, z^t | s, y^n) = C_t^{-1}(s, y^n) p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \prod_{i=1}^t f_{\mathcal{F}_i}^{(\mu, \lambda, \theta)}(x_i, y_i, z_i | u_i). \quad (\text{A39})$$

Then, we have the following chain of equalities:

$$\begin{aligned} p_{X^t Z^t | S_n Y^n, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(x^t, z^t | s, y^n) &\stackrel{(a)}{=} C_t^{-1}(s, y^n) p_{X^t Z^t | S_n Y^n}(x^t, z^t | s, y^n) \prod_{i=1}^t f_{\mathcal{F}_i}^{(\mu, \lambda, \theta)}(x_i, y_i, z_i | u_i) \\ &= C_t^{-1}(s, y^n) p_{X^{t-1} Z^{t-1} | S_n Y^n}(x^{t-1}, z^{t-1} | s, y^n) \prod_{i=1}^{t-1} f_{\mathcal{F}_i}^{(\mu, \lambda, \theta)}(x_i, y_i, z_i | u_i) \\ &\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t) \\ &\stackrel{(b)}{=} \frac{C_{t-1}(s, y^n)}{C_t(s, y^n)} p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t) \\ &= (\Phi_t^{(\mu, \lambda, \theta)}(s, y^n))^{-1} p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (\text{A40})$$

Steps (a) and (b) follow from Equation (A39). From Equation (A40), we have

$$\Phi_{t, q^t}^{(\mu, \lambda, \theta)}(s, y^n) p_{X^t Z^t | S_n Y^n}^{(\mu, \lambda, \theta)}(x^t, z^t | s, y^n) \quad (\text{A41})$$

$$\begin{aligned} &= p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (\text{A42})$$

Taking summations of Equation (A41) and (A42) with respect to x^t, z^t , we obtain

$$\begin{aligned} \Phi_{t, \mathcal{F}^t}^{(\mu, \lambda, \theta)}(s, y^n) &= \sum_{x^t, z^t} p_{X^{t-1} Z^{t-1} | S_n Y^n, \mathcal{F}^{t-1}}^{(\mu, \lambda, \theta)}(x^{t-1}, z^{t-1} | s, y^n) \\ &\quad \times p_{X_t | Z_t | X^{t-1} Z^{t-1} S_n Y^n}(x_t, z_t | x^{t-1}, z^{t-1}, s, y^n) f_{\mathcal{F}_t}^{(\mu, \lambda, \theta)}(x_t, y_t, z_t | u_t), \end{aligned}$$

completing the proof. \square

References

1. Wyner, A.D.; Ziv, J. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inf. Theory* **1976**, *22*, 1–10. [[CrossRef](#)]
2. Shannon, C.E. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.* **1959**, *7*, 142–163.
3. Wolfowitz, J. Approximation with a fidelity criterion. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1966; Volume 1, pp. 566–573.
4. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Academic: London, UK, 1981.
5. Oohama, Y. Exponent function for one helper source coding problem at rates outside the rate region. In *Proceedings of the 2015 IEEE International Symposium on Information Theory*, Hong Kong, China, 14–19 June 2015; pp. 1575–1579.
6. 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.
7. 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.
8. 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 November 2016; pp. 568–572.
9. 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. [[CrossRef](#)]
10. Slepian, D.; Wolf, J.K. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory* **1973**, *19*, 471–480. [[CrossRef](#)]
11. Han, T.S. *Information-Spectrum Methods in Information Theory*; Springer: Berlin, Germany, 2002; The Japanese edition was published by Baifukan-Publisher, Tokyo, 1998.
12. Wyner, A.D. The rate-distortion function for source coding with side information at the decoder-II: General sources. *Inf. Control* **1978**, *38*, 60–80. [[CrossRef](#)]



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).