Weak Variable-Length Slepian-Wolf Coding with Linked Encoders for Mixed Sources

符号器に相関がある場合の混合情報源に対する弱可変長Slepian-Wolf符号化

Akisato KIMURA

Advisor : Professor Tomohiko UYEMATSU

Dept. of Electrical and Electronic Eng.
Tokyo Institute of Technology

February 2000

Contents

Chapter 1

Introduction

Coding problems for correlated information sources were first investigated by Slepian and Wolf [1], where sequences from two correlated sources are separately encoded, sent to a single decoder and decoded with a arbitrarily small probability of error. Their result is now regarded as a starting point of multiterminal information theory, and many variations of their source coding problems have been investigated. After the original proof of coding theorem by Slepian and Wolf, Cover [2] showed a simple proof based on the random coding argument called *bin coding*. Recently, Miyake and Kanaya [3] extended the theorem to the class of non-ergodic or non-stationary sources called *general source* by using the method developed by Han and Verdú [4, 5].

Slepian and Wolf have considered the case where neither of the encoders can observe the encoded sequence from the other encoder. On the other hand, Kaspi and Berger [6], Ericson and Körner [7] have studied the case, where one of two encoders can observe not only the sequence from its own source but also the encoded sequence generated from the other encoder. Later, Oohama [8] has investigated more general case, where there are some mutual linkages between two encoders of the coding system proposed by Slepian and Wolf. However, in the above mentioned coding systems, only fixed-length source coding is investigated, and variable-length source coding still remains open in spite of their significance. In coding problem for single general source, it is known that weak variable-length code, i.e. variable-length code having vanishing error, may achieve lower coding rate than fixed-length code [9]. Hence, we can expect a similar result for correlated sources.

In this paper, we investigate weak variable-length coding problems for correlated two sources which is a special case of the system considered by

Oohama. We clarify the achievable rate region for mixed sources characterized by two ergodic sources, and show that this region is strictly wider than that of fixed-length code in the system proposed by Slepian and Wolf. The organization of this paper is as follows: In Section II, we describe some coding systems for correlated sources, and describe the formulation of the problem. In Section III, we show the main result without proof, and give the proofs in Section IV.

Chapter 2

Coding Systems for Correlated Sources

2.1 Basic Definitions

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite or infinitely countable sets and $\mathcal{B}$ be a binary set. Without loss of generality, we assume that $\mathcal{X} = \mathcal{Y} = \{1, 2, \cdots\}$ and $\mathcal{B} = \{0, 1\}$. Let $\mathcal{B}^*$ be a set of all sequences of finite length. Let $(\boldsymbol{X}, \boldsymbol{Y}) = \{(X_j, Y_j)\}_{j=1}^{\infty}$ be a stationary-ergodic process of random variables (X_j, Y_j) $(j = 1, 2, \cdots)$ which takes values in $\mathcal{X} \times \mathcal{Y}$. Then, both $\boldsymbol{X} = \{X_j\}_{j=1}^{\infty}$ and $\boldsymbol{Y} = \{Y_j\}_{j=1}^{\infty}$ are stationary-ergodic processes. We shall call $\boldsymbol{X}$ or $\boldsymbol{Y}$ *ergodic source*, and $(\boldsymbol{X}, \boldsymbol{Y})$ *correlated ergodic source*. We define a *correlated mixed source* $(\boldsymbol{X}, \boldsymbol{Y})$ by the following distribution:

$$P_n(\boldsymbol{x}, \boldsymbol{y}) = \alpha P_n^{(1)}(\boldsymbol{x}, \boldsymbol{y}) + (1 - \alpha) P_n^{(2)}(\boldsymbol{x}, \boldsymbol{y}) \qquad \forall (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n,$$

where $0 < \alpha < 1$ and $P_n^{(i)}$ $(i = 1, 2)$ are the distributions of the jointly ergodic process $\left(X_{(i)}^n, Y_{(i)}^n\right) = \left\{\left(X_{(i)j}, Y_{(i)j}\right)\right\}_{j=1}^{n}$. Further, we introduce the notation $(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)}) = \left\{\left(X_{(i)j}, Y_{(i)j}\right)\right\}_{j=1}^{\infty}$ $(i = 1, 2)$. For an ergodic source $\boldsymbol{X}$, we define a *entropy rate* by

$$\begin{aligned} H(\boldsymbol{X}) &\triangleq \lim_{n \to \infty} \frac{1}{n} H(X^n) \\ &= \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n), \end{aligned}$$

where $H(X_1, X_2, \cdots, X_n)$ denotes the standard information-theoretic quantity as defined in [10]. Similarly, for a correlated ergodic source $(\boldsymbol{X}, \boldsymbol{Y})$ we denote a *joint entropy rate* and a *conditional entropy rate* by

$$\begin{aligned} H(\boldsymbol{X},\boldsymbol{Y}) &\triangleq \lim_{n\to\infty}\frac{1}{n}H(X^n,Y^n) \\ &= \lim_{n\to\infty}\frac{1}{n}H(X_1,X_2,\cdots,X_n,Y_1,Y_2,\cdots,Y_n), \\ H(\boldsymbol{X}|\boldsymbol{Y}) &\triangleq \lim_{n\to\infty}\frac{1}{n}H(X^n|Y^n), \\ H(\boldsymbol{Y}|\boldsymbol{X}) &\triangleq \lim_{n\to\infty}\frac{1}{n}H(Y^n|X^n), \end{aligned}$$

respectively. In what follows, all logarithms and exponentials are to the base two.

2.2 Slepian-Wolf Coding System

Slepian and Wolf [1] have studied the coding problem for correlated two sources, where sequences from correlated sources are separately encoded, sent to a single decoder and decoded (Figure 2.1). We call this code a *Slepian-Wolf code (SW code)*. First, we introduce fixed-length SW code.

Definition 1: A sequence $\{(\varphi_n^{(1)},\varphi_n^{(2)},\varphi_n^{-1})\}_{n=1}^{\infty}$ of triples is called a *fixed-length SW code*, if the encoders $\varphi_n^{(1)}:\mathcal{X}^n\to\mathcal{M}_n^{(1)}$, $\varphi_n^{(2)}:\mathcal{Y}^n\to\mathcal{M}_n^{(2)}$ and the decoder $\varphi_n^{-1}:\mathcal{M}_n^{(1)}\times\mathcal{M}_n^{(2)}\to\mathcal{X}^n\times\mathcal{Y}^n$ satisfy

$$\lim_{n\to\infty}\Pr\{\varphi_n^{-1}(\varphi_n^{(1)}(X^n),\varphi_n^{(2)}(Y^n))\neq(X^n,Y^n)\}=0, \tag{2.1}$$

where $\mathcal{M}_n^{(1)}=\{1,2,\cdots,M_n^{(1)}\}$ and $\mathcal{M}_n^{(2)}=\{1,2,\cdots,M_n^{(2)}\}$. □

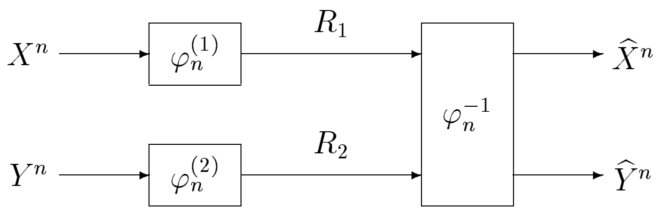

Figure 2.1: Slepian-Wolf coding system

Definition 2: A rate pair (R_1, R_2) is called an *achievable fixed-length SW rate pair*, if there exists a fixed-length SW code which satisfies

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log M_n^{(1)} &\le R_1,\\ \limsup_{n\to\infty}\frac{1}{n}\log M_n^{(2)} &\le R_2.\end{aligned}$$

□

Definition 3 (Achievable fixed-length SW rate region):

$$\begin{aligned}\mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y}) = \{&(R_1,R_2) :\\ &(R_1,R_2) \text{ is the achievable fixed-length SW rate pair}\},\end{aligned}$$

we call $\mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y})$ as *fixed-length SW rate region.* □

Miyake and Kanaya [3] investigated the fixed-length SW rate region, and showed the following result.

Theorem 1 [3]:

$$\begin{aligned}\mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y}) = \{(R_1,R_2) \;:\; & R_1 \ge \overline{H}^{HV}(\boldsymbol{X}|\boldsymbol{Y}),\\ & R_2 \ge \overline{H}^{HV}(\boldsymbol{Y}|\boldsymbol{X}),\\ & R_1 + R_2 \ge \overline{H}^{HV}(\boldsymbol{X},\boldsymbol{Y})\},\end{aligned}$$

where $\overline{H}^{HV}(\boldsymbol{X},\boldsymbol{Y})$ is the *joint sup-entropy rate* [5] defined by

$$\overline{H}^{HV}(\boldsymbol{X},\boldsymbol{Y}) \triangleq \inf\left\{\alpha : \lim_{n\to\infty}\Pr\left\{\frac{1}{n}\log\frac{1}{P_n(X^n,Y^n)} > \alpha\right\} = 0\right\},$$

$\overline{H}^{HV}(\boldsymbol{X}|\boldsymbol{Y})$ and $\overline{H}^{HV}(\boldsymbol{Y}|\boldsymbol{X})$ are the *conditional sup-entropy rate* [5] defined by

$$\begin{aligned}\overline{H}^{HV}(\boldsymbol{X}|\boldsymbol{Y}) &\triangleq \inf\left\{\alpha : \lim_{n\to\infty}\Pr\left\{\frac{1}{n}\log\frac{1}{P_n(X^n|Y^n)} > \alpha\right\} = 0\right\},\\ \overline{H}^{HV}(\boldsymbol{Y}|\boldsymbol{X}) &\triangleq \inf\left\{\alpha : \lim_{n\to\infty}\Pr\left\{\frac{1}{n}\log\frac{1}{P_n(Y^n|X^n)} > \alpha\right\} = 0\right\},\end{aligned}$$

respectively. □

We obtain the following corollary immediately from the nature of sup-entropy rate.

Corollary 1 [5]: If $(\boldsymbol{X},\boldsymbol{Y})$ is a correlated ergodic source, then

$$\begin{aligned}\mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y}) = \{(R_1,R_2) \ : \ & R_1 \geq H(\boldsymbol{X}|\boldsymbol{Y}), \\ & R_2 \geq H(\boldsymbol{Y}|\boldsymbol{X}), \\ & R_1+R_2 \geq H(\boldsymbol{X},\boldsymbol{Y})\}.\end{aligned}$$

Further, if $(\boldsymbol{X},\boldsymbol{Y})$ is a correlated mixed source, then

$$\begin{aligned}\mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y}) = \big\{(R_1,R_2) : & \\ & R_1 \geq \max\left(H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}), H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)})\right), \\ & R_2 \geq \max\left(H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}), H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)})\right), \\ & R_1+R_2 \geq \max\left(H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)}), H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)})\right)\big\}.\end{aligned}$$

□

2.3 Slepian-Wolf Coding System with Linked Encoders

Oohama [8] studied the coding problem for correlated sources, where two separate encoders of the SW code are mutually linked (Figure 2.2). We call this code a *Slepian-Wolf code with linked encoders (SWL code)*. Especially, when both switches are closed, we call it a *SWL-I code*. Also, when both switches are open, we call it a *SWL-II code*. First, we define weak variable-length SWL-I code.

Definition 4: A sequence $\{(\varphi_n^{(11)},\varphi_n^{(12)},\varphi_n^{(21)},\varphi_n^{(22)},\varphi_n^{-1})\}_{n=1}^{\infty}$ of sets is called a *weak variable-length SWL-I code*, if the encoders

$$\begin{aligned}\varphi_n^{(11)} &: \mathcal{X}^n \to \mathcal{B}^*, \\ \varphi_n^{(12)} &: \mathcal{X}^n \times \varphi_n^{(21)}(\mathcal{Y}^n) \to \mathcal{B}^*, \\ \varphi_n^{(21)} &: \mathcal{Y}^n \to \mathcal{B}^*, \\ \varphi_n^{(22)} &: \mathcal{Y}^n \times \varphi_n^{(11)}(\mathcal{X}^n) \to \mathcal{B}^*\end{aligned}$$

and the decoder

$$\varphi_n^{-1} : \mathcal{B}^* \times \mathcal{B}^* \times \mathcal{B}^* \times \mathcal{B}^* \to \mathcal{X}^n \times \mathcal{Y}^n$$

satisfy the following conditions:

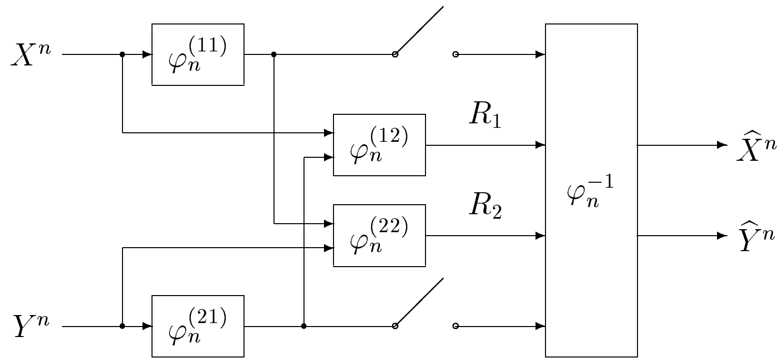

Figure 2.2: Slepian-Wolf coding system with linked encoders

1. The images of $\varphi_n^{(11)}$, $\varphi_n^{(12)}$, $\varphi_n^{(21)}$ and $\varphi_n^{(22)}$ are all prefix sets.
2. $$\lim_{n\to\infty} \Pr\{\varphi_n^{-1}(\varphi_n^{(11)}(X^n), \varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)), \varphi_n^{(21)}(Y^n), \varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n))) \neq (X^n, Y^n)\} = 0. \quad (2.2)$$

□

Definition 5: A rate pair (R_1, R_2) is called an *achievable weak variable-length SWL-I rate pair*, if there exists a weak variable-length SWL-I code which satisfies

$$\left.\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(11)}(X^n))] &= 0, \\ \limsup_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(21)}(Y^n))] &= 0, \\ \limsup_{n\to\infty} \frac{1}{n} E\left[l(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)))\right] &\leq R_1, \\ \limsup_{n\to\infty} \frac{1}{n} E\left[l(\varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n)))\right] &\leq R_2, \end{aligned}\right\} \quad (2.3)$$

where $E[\cdot]$ denotes the expected value and $l : \mathcal{B}^* \to \{0, 1, \cdots\}$ denotes the length function. □

Remark: In [8], Oohama considered the fixed-length coding under the different constraints of rates for SWL-I code. Especially, (2.3) is replaced by

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}E\left[l(\varphi_n^{(11)}(X^n))+l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)))\right] &\leq R_1,\\ \limsup_{n\to\infty}\frac{1}{n}E\left[l(\varphi_n^{(21)}(Y^n))+l(\varphi_n^{(22)}(Y^n,\varphi_n^{(11)}(X^n)))\right] &\leq R_2.\end{aligned}$$

Definition 6: (Achievable weak variable-length SWL-I rate region)

$$\begin{aligned}&\mathcal{R}^*_{SWL-I}(\boldsymbol{X},\boldsymbol{Y})=\{(R_1,R_2):\\ &\qquad (R_1,R_2)\text{ is the achievable weak variable-length SWL-I rate pair}\},\end{aligned}$$

we call $\mathcal{R}^*_{SWL-I}(\boldsymbol{X},\boldsymbol{Y})$ as *weak variable-length SWL-I rate region.* □

Similarly, we define weak variable-length SWL-II code.

Definition 7: A sequence $\{(\varphi_n^{(11)},\varphi_n^{(12)},\varphi_n^{(21)},\varphi_n^{(22)},\varphi_n^{-1})\}_{n=1}^{\infty}$ of sets is called a *weak variable-length SWL-II code*, if the encoders

$$\begin{aligned}\varphi_n^{(11)} &: \mathcal{X}^n\to\mathcal{B}^*,\\ \varphi_n^{(12)} &: \mathcal{X}^n\times\varphi_n^{(21)}(\mathcal{Y}^n)\to\mathcal{B}^*,\\ \varphi_n^{(21)} &: \mathcal{Y}^n\to\mathcal{B}^*,\\ \varphi_n^{(22)} &: \mathcal{Y}^n\times\varphi_n^{(11)}(\mathcal{X}^n)\to\mathcal{B}^*,\end{aligned}$$

and the decoder

$$\varphi_n^{-1}:\mathcal{B}^*\times\mathcal{B}^*\to\mathcal{X}^n\times\mathcal{Y}^n$$

satisfy the following conditions:

1. The images of $\varphi_n^{(11)}$, $\varphi_n^{(12)}$, $\varphi_n^{(21)}$ and $\varphi_n^{(22)}$ are all prefix sets.
2. $$\begin{aligned}\lim_{n\to\infty}\Pr\{\varphi_n^{-1}(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)),&\\ \varphi_n^{(22)}(Y^n,\varphi_n^{(11)}(X^n)))\neq(X^n,Y^n)\}&=0.\end{aligned}$$

□

Definition 8: A rate pair (R_1,R_2) is called an *achievable weak variable-length SWL-II rate pair*, if there exists a weak variable-length SWL-II code which satisfies (2.3). □

Definition 9: (Achievable weak variable-length SWL-II rate region)

$$\mathcal{R}^*_{SWL-II}(\boldsymbol{X},\boldsymbol{Y}) = \{(R_1,R_2): \\ (R_1,R_2) \text{ is the achievable weak variable-length SWL-II rate pair}\}$$

we call $\mathcal{R}^*_{SWL-II}(\boldsymbol{X},\boldsymbol{Y})$ as *weak variable-length SWL-II rate region.* □

Chapter 3

Main Results

In this chapter, we shall investigate the weak variable-length SWL rate region for correlated mixed sources. Before we describe our main result, we impose a following assumption for the correlated mixed source.

Assumption: A correlated mixed source $(\boldsymbol{X}, \boldsymbol{Y})$ satisfies

$$(1) \quad H(X_{(1)1}, Y_{(1)1}) < \infty, \quad H(X_{(2)1}, Y_{(2)1}) < \infty$$

and at least one of the following conditions $(2) - (4)$:

$$\begin{aligned}
&(2) \quad H(\boldsymbol{X}_{(1)}) \neq H(\boldsymbol{X}_{(2)}) \\
&(3) \quad H(\boldsymbol{Y}_{(1)}) \neq H(\boldsymbol{Y}_{(2)}) \\
&(4) \quad H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) \neq H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})
\end{aligned}$$

$\square$

(1) in Assumption indicates the finiteness of the entropy. $(2) - (4)$ indicate that two ergodic sources $(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ and $(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$ can be discriminated by the entropy rate.

The next theorem is our main result.

Theorem 2: If $(\boldsymbol{X}, \boldsymbol{Y})$ is a correlated mixed source satisfying Assumption, then

$$\begin{aligned}
\mathcal{R}^*_{SWL-I}(\boldsymbol{X}, \boldsymbol{Y}) = \Big\{ (R_1, R_2) : \\
& R_1 \geq \alpha H(\boldsymbol{X}_{(1)} | \boldsymbol{Y}_{(1)}) + (1-\alpha) H(\boldsymbol{X}_{(2)} | \boldsymbol{Y}_{(2)}), \\
& R_2 \geq \alpha H(\boldsymbol{Y}_{(1)} | \boldsymbol{X}_{(1)}) + (1-\alpha) H(\boldsymbol{Y}_{(2)} | \boldsymbol{X}_{(2)}), \\
& R_1 + R_2 \geq \alpha H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) + (1-\alpha) H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)}) \Big\}.
\end{aligned}$$

$\square$

According to Theorem 2 and Corollary 1, $\mathcal{R}^*_{SWL-I}(\boldsymbol{X},\boldsymbol{Y}) \supset \mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y})$ for any correlated mixed source $(\boldsymbol{X},\boldsymbol{Y})$ satisfying Assumption, i.e. weak variable-length SWL-I code can achieve strictly lower coding rate than fixed-length SW code.

As a special case of Theorem 2, we immediately obtain the weak variable-length SWL-I rate region for a correlated ergodic source.

Corollary 2: If $(\boldsymbol{X},\boldsymbol{Y})$ is a correlated ergodic source, then

$$\begin{aligned}\mathcal{R}^*_{SWL-I}(\boldsymbol{X},\boldsymbol{Y}) = \{(R_1,R_2) \quad : \quad & R_1 \geq H(\boldsymbol{X}|\boldsymbol{Y}),\\ & R_2 \geq H(\boldsymbol{Y}|\boldsymbol{X}),\\ & R_1+R_2 \geq H(\boldsymbol{X},\boldsymbol{Y})\}.\end{aligned}$$

$\square$

Comparing Corollary 1 and 2, we have $\mathcal{R}^*_{SWL-I}(\boldsymbol{X},\boldsymbol{Y}) = \mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y})$ for any correlated ergodic source $(\boldsymbol{X},\boldsymbol{Y})$, i.e. weak variable-length SWL code can achieve no less coding rate than fixed-length SW code.

In a similar manner as Theorem 2, we obtain the weak variable-length SWL-II rate region for a correlated mixed source.

Corollary 3: If $(\boldsymbol{X},\boldsymbol{Y})$ is a correlated mixed source satisfying Assumption, then

$$\begin{aligned}\mathcal{R}^*_{SWL-II}(\boldsymbol{X},\boldsymbol{Y}) = \Big\{(R_1,R_2) :\\ & R_1 \geq \alpha H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}),\\ & R_2 \geq \alpha H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) + (1-\alpha)H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}),\\ & R_1+R_2 \geq \alpha H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)})\Big\}.\end{aligned}$$

$\square$

Comparing Corollary 1 and 3, we have $\mathcal{R}^*_{SWL-II}(\boldsymbol{X},\boldsymbol{Y}) = \mathcal{R}^*_{SWL-I}(\boldsymbol{X},\boldsymbol{Y})$ for any correlated mixed source $(\boldsymbol{X},\boldsymbol{Y})$ satisfying Assumption, i.e. weak variable-length SWL-II code can achieve the same coding rate as weak variable-length SWL-I code.

Chapter 4

Proof of Theorems

4.1 Proof of Theorem 2

(a) Converse part

For simplicity, we first define the following notations:

$$\begin{aligned}
H^*(\boldsymbol{X}) &\stackrel{\triangle}{=} \alpha H(\boldsymbol{X}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}),\\
H^*(\boldsymbol{Y}) &\stackrel{\triangle}{=} \alpha H(\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{Y}_{(2)}),\\
H^*(\boldsymbol{X},\boldsymbol{Y}) &\stackrel{\triangle}{=} \alpha H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)}),\\
H^*(\boldsymbol{X}|\boldsymbol{Y}) &\stackrel{\triangle}{=} \alpha H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}),\\
H^*(\boldsymbol{Y}|\boldsymbol{X}) &\stackrel{\triangle}{=} \alpha H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) + (1-\alpha)H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}).
\end{aligned}$$

It should be noted that all these values are finite from Assumption.

According to [5, Theorem 1.13] there exists a weak variable-length code $\{(\widehat{\varphi}_n, \widehat{\varphi}_n^{-1})\}_{n=1}^{\infty}$ such that the encoder $\widehat{\varphi}_n : \mathcal{Y}^n \to \mathcal{B}^*$ and the decoder $\widehat{\varphi}_n^{-1} : \mathcal{B}^* \to \mathcal{Y}^n$ satisfy

$$\limsup_{n\to\infty} \frac{1}{n} E[l(\widehat{\varphi}_n(Y^n))] \le H^*(\boldsymbol{Y}) < \infty, \tag{4.1}$$

$$\lim_{n\to\infty} \Pr\{\widehat{\varphi}_n^{-1}(\widehat{\varphi}_n(Y^n)) \ne Y^n\} = 0. \tag{4.2}$$

Then, for a given weak variable-length SWL-I code $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$, we construst a variable-length code for a correlated mixed source

$(\boldsymbol{X}, \boldsymbol{Y})$. Define a sequence of codes $\{(\psi_n, \psi_n^{-1})\}_{n=1}^{\infty}$ ($\psi_n : \mathcal{X}^n \times \mathcal{Y}^n \to \mathcal{B}^*$, $\psi_n^{-1} : \mathcal{B}^* \to \mathcal{X}^n \times \mathcal{Y}^n$) as follows (see also Figure 2):

$$\begin{aligned} \psi_n(\boldsymbol{x}, \boldsymbol{y}) &\stackrel{\triangle}{=} \varphi_n^{(11)}(\boldsymbol{x}) * \varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) \\ &\qquad * \varphi_n^{(21)}(\boldsymbol{y}) * \widehat{\varphi}_n(\boldsymbol{y}), \\ \psi_n^{-1}(s_1 * s_2 * s_3 * s_4) &\stackrel{\triangle}{=} \varphi_n^{-1}(s_1, s_2, s_3, \varphi_n^{(22)}(\widehat{\varphi}_n^{-1}(s_4), s_1)), \end{aligned} \tag{4.3}$$

for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, where $*$ represents a concatination and

$$\begin{aligned} s_1 &\stackrel{\triangle}{=} \varphi_n^{(11)}(\boldsymbol{x}), \\ s_2 &\stackrel{\triangle}{=} \varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})), \\ s_3 &\stackrel{\triangle}{=} \varphi_n^{(21)}(\boldsymbol{y}), \\ s_4 &\stackrel{\triangle}{=} \widehat{\varphi}_n(\boldsymbol{y}). \end{aligned}$$

Note that both $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ and $\{(\widehat{\varphi}_n, \widehat{\varphi}_n^{-1})\}_{n=1}^{\infty}$ are weak variable-length codes. Since the images of $\varphi_n^{(11)}$, $\varphi_n^{(12)}$, $\varphi_n^{(21)}$ and $\widehat{\varphi}_n$ are all prefix sets, the image of ψ_n is also a prefix set. Further, from (2.2) and (4.2), the error probability of this code can be bounded as

$$\begin{aligned} &\Pr\{\psi_n^{-1}(\psi_n(X^n, Y^n)) \neq (X^n, Y^n)\} \\ &= \Pr\{\widehat{\varphi}_n^{-1}(\widehat{\varphi}_n(Y^n)) \neq Y^n \text{ or } \varphi_n^{-1}(\varphi_n^{(11)}(X^n), \varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)), \\ &\qquad \varphi_n^{(21)}(Y^n), \varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n))) \neq (X^n, Y^n)\} \\ &\leq \Pr\{\widehat{\varphi}_n^{-1}(\widehat{\varphi}_n(Y^n)) \neq Y^n\} + \Pr\{\varphi_n^{-1}(\varphi_n^{(11)}(X^n), \varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)), \\ &\qquad \varphi_n^{(21)}(Y^n), \varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n))) \neq (X^n, Y^n)\} \\ &\to 0 \quad (n \to \infty). \end{aligned}$$

This implies that $\{(\psi_n, \psi_n^{-1})\}_{n=1}^{\infty}$ is a weak variable-length code. Hence, according to [5, Theorem 1.13] or [9, Theorem 4.1] it must satisfy

$$\limsup_{n\to\infty} \frac{1}{n} E[l(\psi_n(X^n, Y^n))] \geq H^*(\boldsymbol{X}, \boldsymbol{Y}). \tag{4.4}$$

Hence, we obtain

$$
\begin{aligned}
&H^*(\boldsymbol{X},\boldsymbol{Y}) \\
&\le \limsup_{n\to\infty}\frac{1}{n}E[l(\psi_n(X^n,Y^n))] \\
&\overset{\text{(a)}}{=} \limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(11)}(X^n))+l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n))) \\
&\qquad\qquad\qquad +l(\varphi_n^{(21)}(Y^n))+l(\widehat{\varphi}_n(Y^n))] \\
&\le \limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(11)}(X^n))]+\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)))] \\
&\qquad +\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(21)}(Y^n))]+\limsup_{n\to\infty}\frac{1}{n}E[l(\widehat{\varphi}_n(Y^n))] \\
&\overset{\text{(b)}}{=} \limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)))]+\limsup_{n\to\infty}\frac{1}{n}E[l(\widehat{\varphi}_n(Y^n))] \\
&\overset{\text{(c)}}{\le} \limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)))]+H^*(\boldsymbol{Y})
\end{aligned}
$$

where the equality (a) from (4.3), the equality (b) from (2.3) and the equality (c) from (4.1). This implies

$$
\begin{aligned}
\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)))] &\ge H^*(\boldsymbol{X},\boldsymbol{Y})-H^*(\boldsymbol{Y}) \\
&= H^*(\boldsymbol{X}|\boldsymbol{Y}).
\end{aligned}
$$

Therefore, if a weak variable-length SWL code $\{(\varphi_n^{(11)},\varphi_n^{(12)},\varphi_n^{(21)},\varphi_n^{(22)},\varphi_n^{-1})\}_{n=1}^{\infty}$ satisfies

$$
\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n,\varphi_n^{(21)}(Y^n)))]=R_1,
$$

then $R_1\ge H^*(\boldsymbol{X}|\boldsymbol{Y})$. In a similar manner, we can prove that if

$$
\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(22)}(Y^n,\varphi_n^{(11)}(X^n)))]=R_2,
$$

then $R_2\ge H^*(\boldsymbol{Y}|\boldsymbol{X})$. Further, $R_1+R_2\ge H^*(\boldsymbol{X},\boldsymbol{Y})$ is obvious from (4.4).

(b) Achievability part

Before the proof of this part, we introduce two lemmas:

Lemma 1 (Asymptotic Equipartition Property (AEP)) [10]: Assume that a correlated ergodic source $(\boldsymbol{X},\boldsymbol{Y})$ satisfies $H(X_1)<\infty,\ H(Y_1)<\infty$ and $H(X_1,Y_1)<\infty$. Then, for arbitrary $\varepsilon>0$ and $\delta>0$, there exists an integer $n_0(\varepsilon,\delta)$ (which depends on the source) such that for all $n\geq n_0(\varepsilon,\delta)$

$$P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})}-H(\boldsymbol{X})\right|\geq\varepsilon\right\}\leq\delta,$$
$$P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})}-H(\boldsymbol{Y})\right|\geq\varepsilon\right\}\leq\delta,$$
$$P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-H(\boldsymbol{X},\boldsymbol{Y})\right|\geq\varepsilon\right\}\leq\delta,$$

simultaneously hold. □

Lemma 2: For arbitrary $\lambda>0,\ \varepsilon>0$, and an arbitrary correlated mixed source $(\boldsymbol{X},\boldsymbol{Y})$,

$$\begin{aligned}
&P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:-\varepsilon+\gamma_n\leq\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\lambda\leq\varepsilon-\frac{c_0}{n}\right\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad-\exp(-n\gamma_n)\\
&\quad\leq P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-\lambda\right|\leq\varepsilon\right\}\leq\\
&P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:-\varepsilon-\gamma_n\leq\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\lambda\leq\varepsilon+\frac{c_0}{n}\right\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad+\exp(-n\gamma_n),
\end{aligned}\tag{4.5}$$

where $c_0=-\log\min(\alpha,1-\alpha)$, and $\{\gamma_n\}_{n=1}^{\infty}$ satisfies $\gamma_1>\gamma_2>\cdots>0$ and $\gamma_n\to 0,\ n\gamma_n\to\infty\ (n\to\infty)$. □

The proof of Lemma 2 is given in Section 4.3.

(Step 1) Preliminaries

First, for an arbitrary subset $A_n\subset\mathcal{X}^n\times\mathcal{Y}^n$, we use the notations

$$\Pr\{(X^n,Y^n)\in A_n\}\ \stackrel{\triangle}{=}\ \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n}P_n(\boldsymbol{x},\boldsymbol{y}),$$
$$\Pr\{(X^n_{(i)},Y^n_{(i)})\in A_n\}\ \stackrel{\triangle}{=}\ \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n}P_n^{(i)}(\boldsymbol{x},\boldsymbol{y}).\qquad(i=1,2)$$

By using these notations, we obtain

$$\begin{aligned}
&\Pr\{(X^n, Y^n) \in A_n\} \\
&= \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n} P_n(\boldsymbol{x},\boldsymbol{y}) \\
&= \alpha \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n} P_n^{(1)}(\boldsymbol{x},\boldsymbol{y}) + (1-\alpha) \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n} P_n^{(2)}(\boldsymbol{x},\boldsymbol{y}) \\
&= \alpha \Pr\{(X_{(1)}^n, Y_{(1)}^n) \in A_n\} + (1-\alpha)\Pr\{(X_{(2)}^n, Y_{(2)}^n) \in A_n\}. \qquad (4.6)
\end{aligned}$$

Next, for an arbitrary $\varepsilon > 0$, define subsets $T_n^{(1)}$ and $T_n^{(2)}$ of $\mathcal{X}^n \times \mathcal{Y}^n$ as

$$\begin{aligned}
T_n^{(1)} \triangleq \Bigg\{ (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \quad & \left| \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) \right| \le \varepsilon, \\
& \left| \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(1)}) \right| \le \varepsilon, \\
& \left| \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(1)}) \right| \le \varepsilon, \Bigg\}, \qquad (4.7) \\
T_n^{(2)} \triangleq \Bigg\{ (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \quad & \left| \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)}) \right| \le \varepsilon, \\
& \left| \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(2)}) \right| \le \varepsilon, \\
& \left| \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(2)}) \right| \le \varepsilon, \Bigg\}, \qquad (4.8)
\end{aligned}$$

and $\overline{T}_n^{(1)}$ (resp. $\overline{T}_n^{(2)}$) denotes a complement of $T_n^{(1)}$ (resp. $T_n^{(2)}$). From (4.6), we immediately obtain

$$\begin{aligned}
&\Pr\{(X^n, Y^n) \notin T_n^{(1)} \cup T_n^{(2)}\} \\
&= \alpha\Pr\{(X_{(1)}^n, Y_{(1)}^n) \notin T_n^{(1)} \cup T_n^{(2)}\} + (1-\alpha)\Pr\{(X_{(2)}^n, Y_{(2)}^n) \notin T_n^{(1)} \cup T_n^{(2)}\} \\
&\le \alpha\Pr\{(X_{(1)}^n, Y_{(1)}^n) \notin T_n^{(1)}\} + (1-\alpha)\Pr\{(X_{(2)}^n, Y_{(2)}^n) \notin T_n^{(2)}\}. \qquad (4.9)
\end{aligned}$$

According to (4.7), the first term in (4.9) can be upperbouded by

$$\begin{aligned}
&\Pr\{(X^n_{(1)}, Y^n_{(1)}) \notin T^{(1)}_n\} \\
&\le P^{(1)}_n\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})\right| > \varepsilon\right\} \\
&\quad + P^{(1)}_n\left\{\boldsymbol{x} \in \mathcal{X}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(1)})\right| > \varepsilon\right\} \\
&\quad + P^{(1)}_n\left\{\boldsymbol{y} \in \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(1)})\right| > \varepsilon\right\}. \qquad (4.10)
\end{aligned}$$

According to Lemma 1 and 2, the first term in (4.10) can be evaluated as

$$\begin{aligned}
&P^{(1)}_n\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})\right| > \varepsilon\right\} \\
&\le P^{(1)}_n\Bigg\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \\
&\qquad \frac{1}{n}\log\frac{1}{P^{(1)}_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) < -\varepsilon + \gamma_n \quad \text{or} \\
&\qquad \frac{1}{n}\log\frac{1}{P^{(1)}_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) > \varepsilon - \frac{c_0}{n}\Bigg\} + \exp(-n\gamma_n) \\
&\to 0 \qquad (n \to \infty).
\end{aligned}$$

In a similar manner, the second and third terms in (4.10) satisfy

$$\begin{aligned}
P^{(1)}_n\left\{\boldsymbol{x} \in \mathcal{X}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(1)})\right| > \varepsilon\right\} &\to 0 \quad (n \to \infty), \\
P^{(1)}_n\left\{\boldsymbol{y} \in \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(1)})\right| > \varepsilon\right\} &\to 0 \quad (n \to \infty).
\end{aligned}$$

Therefore, the first term in (4.9) satisfies

$$\Pr\{(X^n_{(1)}, Y^n_{(1)}) \notin T^{(1)}_n\} \to 0 \qquad (n \to \infty). \qquad (4.11)$$

In a similar manner, the second term in (4.9) also satisfies

$$\Pr\{(X^n_{(2)}, Y^n_{(2)}) \notin T^{(2)}_n\} \to 0 \qquad (n \to \infty). \qquad (4.12)$$

Therefore, we have

$$\Pr\{(X^n, Y^n) \notin T_n^{(1)} \cup T_n^{(2)}\} \to 0 \qquad (n \to \infty). \tag{4.13}$$

On the other hand, according to (4.6), we have

$$\begin{aligned}
&\Pr\{(X^n, Y^n) \in T_n^{(1)} \cap T_n^{(2)}\} \\
&= \alpha \Pr\{(X_{(1)}^n, Y_{(1)}^n) \in T_n^{(1)} \cap T_n^{(2)}\} \\
&\qquad\qquad (1-\alpha)\Pr\{(X_{(2)}^n, Y_{(2)}^n) \in T_n^{(1)} \cap T_n^{(2)}\} \\
&\le \alpha \Pr\{(X_{(1)}^n, Y_{(1)}^n) \in T_n^{(2)}\} + (1-\alpha)\Pr\{(X_{(2)}^n, Y_{(2)}^n) \in T_n^{(1)}\}.
\end{aligned} \tag{4.14}$$

Here, we show $\Pr\{(X_{(2)}^n, Y_{(2)}^n) \in T_n^{(1)}\} \to 0 \quad (n \to \infty)$.

Case i) $H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) \neq H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$

In this case, according to Lemma 1 and Lemma 2, we have

$$\begin{aligned}
&\Pr\{(X_{(1)}^n, Y_{(1)}^n) \in T_n^{(2)}\} \\
&\le P_n^{(1)}\left\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})\right| \le \varepsilon\right\} \\
&\le P_n^{(1)}\Bigg\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \\
&\qquad -\varepsilon - \gamma_n \le \frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x}, \boldsymbol{y})} - H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)}) \le \varepsilon + \frac{c_0}{n}\Bigg\} + \exp(-n\gamma_n) \\
&\to 0 \qquad (n \to \infty).
\end{aligned}$$

Case ii) $H(\boldsymbol{X}_{(1)}) \neq H(\boldsymbol{X}_{(2)})$ or $H(\boldsymbol{Y}_{(1)}) \neq H(\boldsymbol{Y}_{(2)})$

If $H(\boldsymbol{X}_{(1)}) \neq H(\boldsymbol{X}_{(2)})$, according to Lemma 1 and Lemma 2 again, we have

$$\begin{aligned}
&\Pr\{(X_{(1)}^n, Y_{(1)}^n) \in T_n^{(2)}\} \\
&\le P_n^{(1)}\left\{\boldsymbol{x} \in \mathcal{X}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(2)})\right| \le \varepsilon\right\} \\
&\le P_n^{(1)}\left\{\boldsymbol{x} \in \mathcal{X}^n : -\varepsilon - \gamma_n \le \frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x})} - H(\boldsymbol{X}_{(2)}) \le \varepsilon + \frac{c_0}{n}\right\} \\
&\qquad + \exp(-n\gamma_n) \\
&\to 0 \qquad (n \to \infty).
\end{aligned}$$

If $H(\boldsymbol{Y}_{(1)}) \neq H(\boldsymbol{Y}_{(2)})$, we obtain the same result by replacing X with Y. In a similar manner, the second term in (4.14) satisfies

$$\Pr\{(X^n_{(2)}, Y^n_{(2)}) \in T^{(1)}_n\} \quad \to \quad 0 \qquad (n \to \infty).$$

Hence, we have

$$\Pr\{(X^n, Y^n) \in T^{(1)}_n \cap T^{(2)}_n\} \to 0 \qquad (n \to \infty). \tag{4.15}$$

(Step 2) Construction of codes

Suppose that we are given a set of rates (R_1, R_2) which satisfies

$$\begin{aligned} R_1 &\geq \alpha H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) + (1-\alpha) H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}), \\ R_2 &\geq \alpha H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) + (1-\alpha) H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}), \\ R_1 + R_2 &\geq \alpha H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) + (1-\alpha) H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)}). \end{aligned}$$

First, we define a set of rates $(R_{11}, R_{12}, R_{21}, R_{22})$ as

$$\begin{aligned} R_{11} &= H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}), \\ R_{12} &= \min\left(H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) + \frac{\delta_2}{1-\alpha},\ H(\boldsymbol{Y}_{(1)}) + \frac{\delta_3}{\alpha} \right), \\ R_{21} &= H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}) + \frac{\delta_1}{1-\alpha}, \\ R_{22} &= \max\left(H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}),\ H(\boldsymbol{Y}_{(2)}) - \frac{\delta_1}{1-\alpha} \right), \end{aligned}$$

where δ_1, δ_2 and δ_3 are nonnegative numbers defined by

$$\begin{aligned} \delta_1 &\triangleq R_1 - \alpha H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) - (1-\alpha) H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}), \\ \delta_2 &\triangleq R_2 - \alpha H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) - (1-\alpha) H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}), \\ \delta_3 &\triangleq R_1 + R_2 - \alpha H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) - (1-\alpha) H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)}). \end{aligned}$$

Then, it is easy to verify that

$$\left.\begin{aligned} \alpha R_{11} + (1-\alpha) R_{21} &= R_1, \\ \alpha R_{12} + (1-\alpha) R_{22} &= R_2, \end{aligned}\right\} \tag{4.16}$$

and

$$\left.\begin{array}{rr} R_{11} \geq H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}), & R_{21} \geq H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}), \\ R_{12} \geq H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}), & R_{22} \geq H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}), \\ R_{11}+R_{12} \geq H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)}), & R_{21}+R_{22} \geq H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)}). \end{array}\right\} \quad (4.17)$$

According to Corollary 1, we have $(R_{11}, R_{12}) \in \mathcal{R}_{SW}(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ and $(R_{21}, R_{22}) \in \mathcal{R}_{SW}(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$. Hence, there are fixed-length SW codes for the source $(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ and $(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$. Denote the fixed-length SW code for the source $(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ as $\{(f_n^{(1)}, f_n^{(2)}, f_n^{-1})\}_{n=1}^{\infty}$ where

$$\begin{aligned} f_n^{(1)} : \mathcal{X}^n &\to \mathcal{M}_n^{(1)}(f_n), \\ f_n^{(2)} : \mathcal{Y}^n &\to \mathcal{M}_n^{(2)}(f_n), \\ \mathcal{M}_n^{(i)}(f_n) = \{1, 2, \cdots&, M_n^{(i)}(f_n)\} \quad (i = 1, 2). \end{aligned}$$

Similarly, denote the fixed-length SW code for the source $(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$ as $\{(g_n^{(1)}, g_n^{(2)}, g_n^{-1})\}_{n=1}^{\infty}$ where

$$\begin{aligned} g_n^{(1)} : \mathcal{X}^n &\to \mathcal{M}_n^{(1)}(g_n), \\ g_n^{(2)} : \mathcal{Y}^n &\to \mathcal{M}_n^{(2)}(g_n), \\ \mathcal{M}_n^{(i)}(g_n) = \{1, 2, \cdots&, M_n^{(i)}(g_n)\} \quad (i = 1, 2). \end{aligned}$$

These SW codes satisfies

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log M_n^{(1)}(f_n) &\leq R_{11}, \\ \limsup_{n\to\infty} \frac{1}{n} \log M_n^{(2)}(f_n) &\leq R_{12}, \\ \limsup_{n\to\infty} \frac{1}{n} \log M_n^{(1)}(g_n) &\leq R_{21}, \\ \limsup_{n\to\infty} \frac{1}{n} \log M_n^{(2)}(g_n) &\leq R_{22}. \end{aligned}$$

Further, define the sequence of integers $\{N_n\}_{n=1}^{\infty}$ such that

$$0 < N_n \leq n, \quad \lim_{n\to\infty} N_n = \infty, \quad \lim_{n\to\infty} \frac{1}{n} N_n = 0. \quad (4.18)$$

Then, we construct a code $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ as follows:

$\varphi_n^{(11)}$: $\varphi_n^{(11)}(\boldsymbol{x}) = \widehat{\varphi}_{N_n}^{(1)}(\boldsymbol{x}_1)$, where $\boldsymbol{x} = \boldsymbol{x}_1 * \boldsymbol{x}_2 \in \mathcal{X}^n,\ \boldsymbol{x}_1 \in \mathcal{X}^{N_n},\ \boldsymbol{x}_2 \in \mathcal{X}^{n-N_n}$ and $\{(\widehat{\varphi}_n^{(1)}, \widehat{\varphi}_n^{(1)-1})\}_{n=1}^{\infty}$ is a variable-length code for $\boldsymbol{X}$ which satisfies $\limsup_{n\to\infty} \frac{1}{n} E[l(\widehat{\varphi}_n^{(1)}(X^n))] < \infty$.

$\varphi_n^{(21)}$: $\varphi_n^{(21)}(\boldsymbol{y}) = \widehat{\varphi}_{N_n}^{(2)}(\boldsymbol{y}_1)$, where $\boldsymbol{y} = \boldsymbol{y}_1 * \boldsymbol{y}_2 \in \mathcal{Y}^n,\ \boldsymbol{y}_1 \in \mathcal{Y}^{N_n},\ \boldsymbol{y}_2 \in \mathcal{Y}^{n-N_n}$ and $\{(\widehat{\varphi}_n^{(2)}, \widehat{\varphi}_n^{(2)-1})\}_{n=1}^{\infty}$ is a variable-length code for $\boldsymbol{Y}$ which satisfies $\limsup_{n\to\infty} \frac{1}{n} E[l(\widehat{\varphi}_n^{(2)}(Y^n))] < \infty$.

$\varphi_n^{(12)}$: Decode $\boldsymbol{y}_1 (\in \mathcal{Y}^{N_n})$ from a given codeword $\varphi_n^{(21)}(\boldsymbol{y})$, then assign the codeword in the following manner:

If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}$, then $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) = f_n^{(1)}(\boldsymbol{x})$.
If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$, then $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) = g_n^{(1)}(\boldsymbol{x})$.
Otherwise, $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y}))$ may assign any codeword.

$\varphi_n^{(22)}$: Decode $\boldsymbol{x}_1 \in \mathcal{X}^{N_n}$ from a given codeword $\varphi_n^{(11)}(\boldsymbol{x})$, then assign the codeword in the following manner:

If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}$, then $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})) = f_n^{(2)}(\boldsymbol{y})$.
If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$, then $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})) = g_n^{(2)}(\boldsymbol{y})$.
Otherwise, $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x}))$ may assign any codeword.

φ_n^{-1}: For given $s_1 \in \varphi_n^{(11)}(\mathcal{X}^n)$, $s_2 \in \varphi_n^{(12)}(\mathcal{X}^n, \varphi_n^{(21)}(\mathcal{Y}^n))$, $s_3 \in \varphi_n^{(21)}(\mathcal{Y}^n)$, and $s_4 \in \varphi_n^{(22)}(\mathcal{Y}^n, \varphi_n^{(11)}(\mathcal{X}^n))$, we first decode $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in \mathcal{X}^{N_n} \times \mathcal{Y}^{N_n}$ from (s_1, s_3). Then, output an estimate $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) \in \mathcal{X}^n \times \mathcal{Y}^n$ in the following manner:

If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}$, then $(\boldsymbol{x}, \boldsymbol{y}) = f_n^{-1}(s_2, s_4)$.
If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$, then $(\boldsymbol{x}, \boldsymbol{y}) = g_n^{-1}(s_2, s_4)$.
Otherwise, we declare an error.

In the proposed code, we cannot obtain the encoded sequence in the following cases:

(i) $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}$ and $f_n^{-1}(f_n^{(1)}(\boldsymbol{x}), f_n^{(2)}(\boldsymbol{y})) \neq (\boldsymbol{x}, \boldsymbol{y})$

(ii) $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$ and $g_n^{-1}(g_n^{(1)}(\boldsymbol{x}), g_n^{(2)}(\boldsymbol{y})) \neq (\boldsymbol{x}, \boldsymbol{y})$

(iii) $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(1)} \cap T_{N_n}^{(2)}$ or $(\boldsymbol{x}_1, \boldsymbol{y}_1) \notin T_{N_n}^{(1)} \cup T_{N_n}^{(2)}$

According to (2.1), (4.6) and (4.11), the probability of the event (i) can be bounded as

$$\begin{aligned}
&\Pr\{(X^{N_n}, Y^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)} \\
&\qquad \text{and } f_n^{-1}(f_n^{(1)}(X^n), f_n^{(2)}(Y^n)) \neq (X^n, Y^n)\} \\
&= \alpha \Pr\{(X_{(1)}^{N_n}, Y_{(1)}^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)} \\
&\qquad \text{and } f_n^{-1}(f_n^{(1)}(X_{(1)}^n), f_n^{(2)}(Y_{(1)}^n)) \neq (X_{(1)}^n, Y_{(1)}^n)\} \\
&\quad +(1-\alpha)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)} \\
&\qquad \text{and } f_n^{-1}(f_n^{(1)}(X_{(2)}^n), f_n^{(2)}(Y_{(2)}^n)) \neq (X_{(2)}^n, Y_{(2)}^n)\} \\
&\leq \alpha \Pr\{f_n^{-1}(f_n^{(1)}(X_{(1)}^n), f_n^{(2)}(Y_{(1)}^n)) \neq (X_{(1)}^n, Y_{(1)}^n)\} \\
&\qquad +(1-\alpha)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}\} \\
&\leq \alpha \Pr\{f_n^{-1}(f_n^{(1)}(X_{(1)}^n), f_n^{(2)}(Y_{(1)}^n)) \neq (X_{(1)}^n, Y_{(1)}^n)\} \\
&\qquad +(1-\alpha)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \notin T_{N_n}^{(2)}\} \\
&\to 0 \qquad (n \to \infty).
\end{aligned}$$

In a similar manner, according to (2.1), (4.6) and (4.12), the probability of the event (ii) vanishes as $n \to \infty$. Further, the probability of the event (iii) vanishes as $n \to \infty$ due to (4.13) and (4.15). Therefore, the probability of decoding error vanishes as $n \to \infty$.

(Step 3) Evaluation of the average length of codeword

Lastly, we investigate the average length of codeword. For an arbitrary

$\delta > 0$ and a sufficiently large $n_0(\delta)$, if $n \geq n_0(\delta)$, according to (4.6), we have

$$\begin{aligned}
&\frac{1}{n}E[l(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)))] \\
&= \left(\frac{1}{n}\log M_n^{(1)}(f_n)\right) \times \Pr\{(X^{N_n}, Y^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}\} \\
&\qquad + \left(\frac{1}{n}\log M_n^{(1)}(g_n)\right) \times \Pr\{(X^{N_n}, Y^{N_n}) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}\} \\
&\leq (R_{11} + \delta)\Pr\{(X^{N_n}, Y^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}\} \\
&\qquad +(R_{21} + \delta)\Pr\{(X^{N_n}, Y^{N_n}) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}\} \\
&\leq \alpha(R_{11} + \delta)\Pr\{(X_{(1)}^{N_n}, Y_{(1)}^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}\} \\
&\qquad +(1-\alpha)(R_{11} + \delta)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \in T_{N_n}^{(1)} \cap \overline{T}_{N_n}^{(2)}\} \\
&\quad +\alpha(R_{21} + \delta)\Pr\{(X_{(1)}^{N_n}, Y_{(1)}^{N_n}) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}\} \\
&\qquad +(1-\alpha)(R_{21} + \delta)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}\} \\
&\leq \alpha(R_{11} + \delta)\Pr\{(X_{(1)}^{N_n}, Y_{(1)}^{N_n}) \in T_{N_n}^{(1)}\} \\
&\qquad +(1-\alpha)(R_{11} + \delta)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \notin T_{N_n}^{(2)}\} \\
&\quad +\alpha(R_{21} + \delta)\Pr\{(X_{(1)}^{N_n}, Y_{(1)}^{N_n}) \notin T_{N_n}^{(1)}\} \\
&\qquad +(1-\alpha)(R_{21} + \delta)\Pr\{(X_{(2)}^{N_n}, Y_{(2)}^{N_n}) \in T_{N_n}^{(2)}\}.
\end{aligned}$$

Hence, by using (4.11) and (4.12), we obtain

$$\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)))] \leq \alpha R_{11} + (1-\alpha)R_{21} + \delta.$$

In a similar manner, we obtain

$$\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n)))] \leq \alpha R_{12} + (1-\alpha)R_{22} + \delta.$$

Since $\delta > 0$ is arbitrary, (4.16) yields

$$\begin{aligned}
\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)))] &\leq \alpha R_{11} + (1-\alpha)R_{21} = R_1, \\
\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n)))] &\leq \alpha R_{12} + (1-\alpha)R_{22} = R_2.
\end{aligned}$$

Therefore, (R_1, R_2) is the achievable weak variable-length SWL rate pair. Finally, for some constant $K < \infty$, we have

$$\begin{aligned}\frac{1}{n}E[l(\varphi_n^{(11)}(X^n))] &= \frac{1}{n}E[l(\widehat{\varphi}_{N_n}^{(1)}(X^n))] \\ &= \frac{N_n}{n}\cdot\frac{1}{N_n}E[l(\widehat{\varphi}_{N_n}^{(1)}(X^n))] \\ &< \frac{N_n}{n}\cdot K.\end{aligned}$$

Combining the above inequality and (4.18), we immediately obtain

$$\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(11)}(X^n))] = 0.$$

Similarly, we have

$$\limsup_{n\to\infty}\frac{1}{n}E[l(\varphi_n^{(21)}(Y^n))] = 0.$$

This completes the proof of Theorem 2. □

4.2 Proof of Corollary 3

The proof of the converse part is obvious. So, we only gives the achievability part.

For a weak variable-length achievable SWL-I rate pair (R_1, R_2), there exists a weak varible-length SWL-I code $\{(\widetilde{\varphi}_n^{(11)}, \widetilde{\varphi}_n^{(12)}, \widetilde{\varphi}_n^{(21)}, \widetilde{\varphi}_n^{(22)}, \widetilde{\varphi}_n^{-1})\}_{n=1}^{\infty}$. By using this SWL-I code, we construct a weak variable-length SWL-II code $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)},$
$\varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ as follows:

$$\begin{aligned}\varphi_n^{(11)}(\boldsymbol{x}) &\triangleq \widetilde{\varphi}_n^{(11)}(\boldsymbol{x}), \\ \varphi_n^{(21)}(\boldsymbol{y}) &\triangleq \widetilde{\varphi}_n^{(21)}(\boldsymbol{y}), \\ \varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) &\triangleq \varphi_n^{(21)}(\boldsymbol{y}) * \widetilde{\varphi}_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})), \\ &= \widetilde{\varphi}_n^{(21)}(\boldsymbol{y}) * \widetilde{\varphi}_n^{(12)}(\boldsymbol{x}, \widetilde{\varphi}_n^{(21)}(\boldsymbol{y})), \\ \varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})) &\triangleq \varphi_n^{(11)}(\boldsymbol{x}) * \widetilde{\varphi}_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})), \\ &= \widetilde{\varphi}_n^{(11)}(\boldsymbol{x}) * \widetilde{\varphi}_n^{(22)}(\boldsymbol{y}, \widetilde{\varphi}_n^{(11)}(\boldsymbol{x})).\end{aligned}$$

Then, it is easy to show that this code is achievable.

4.3 Proof of Lemma 2

For any $0 < \alpha < 2$, any positive integer n, and any $u > 0,\ v > 0$, it is easy to see

$$\begin{aligned}&\frac{1}{n}\log\min(\alpha, 1-\alpha) + \max(u,v)\\ &\quad\le\ \frac{1}{n}\log\{\alpha\exp(nu) + (1-\alpha)\exp(nv)\} \le \max(u,v).\end{aligned}$$

Substituting $u = \frac{1}{n}\log P_n^{(1)}(\boldsymbol{x},\boldsymbol{y}),\ v = \frac{1}{n}\log P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})$, we obtain

$$\Delta_n(\boldsymbol{x},\boldsymbol{y}) \le \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} \le \Delta_n(\boldsymbol{x},\boldsymbol{y}) + \frac{c_0}{n} \qquad \forall(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n, \tag{4.19}$$

where

$$\Delta_n(\boldsymbol{x},\boldsymbol{y}) \stackrel{\triangle}{=} \min\left(\frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}, \frac{1}{n}\log\frac{1}{P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})}\right).$$

On the other hand, we have

$$\begin{aligned}&P_n^{(1)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log P_n^{(1)}(\boldsymbol{x},\boldsymbol{y}) - \frac{1}{n}\log P_n^{(2)}(\boldsymbol{x},\boldsymbol{y}) \le -\gamma_n\right\}\\
&\quad= \sum_{\substack{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\\ \frac{1}{n}\log\frac{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}{P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})}\le-\gamma_n}} P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})\\
&\quad\le \sum_{\substack{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\\ \frac{1}{n}\log\frac{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}{P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})}\le-\gamma_n}} P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})\exp(-n\gamma_n)\\
&\quad\le \sum_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:} P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})\exp(-n\gamma_n)\\
&\quad\le \exp(-n\gamma_n).\end{aligned} \tag{4.20}$$

Then, by using (4.19) and (4.20), we obtain

$$
\begin{aligned}
&P_n^{(1)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\right.\\
&\qquad\qquad\qquad\qquad\left.\le\frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}+\frac{c_0}{n}\right\}\\
&\overset{ⓐ}{=}\ P_n^{(1)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\right\}\\
&\overset{ⓑ}{\ge}\ P_n^{(1)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\Delta_n(\boldsymbol{x},\boldsymbol{y})\right\}\\
&=\ P_n^{(1)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n :\right.\\
&\qquad\left.\max\left(\frac{1}{n}\log P_n^{(1)}(\boldsymbol{x},\boldsymbol{y}),\frac{1}{n}\log P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})\right)\le\frac{1}{n}\log P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})+\gamma_n\right\}\\
&=\ P_n^{(1)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})\le\frac{1}{n}\log P_n^{(1)}(\boldsymbol{x},\boldsymbol{y})+\gamma_n\right\}\\
&\overset{ⓒ}{\ge}\ 1-\exp(-n\gamma_n),
\end{aligned}
\tag{4.21}
$$

where ⓐ and ⓑ follow from (4.19); ⓒ follows from (4.20). Similarly, we have

$$
\begin{aligned}
&P_n^{(2)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\right.\\
&\qquad\qquad\left.\le\frac{1}{n}\log\frac{1}{P_n^{(2)}(\boldsymbol{x},\boldsymbol{y})}+\frac{c_0}{n}\right\}\ge 1-\exp(-n\gamma_n).
\end{aligned}
\tag{4.22}
$$

Define a subset of $\mathcal{X}^n\times\mathcal{Y}^n$ as

$$
\begin{aligned}
B_n^{(i)}=&\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n :\right.\\
&\quad\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\\
&\qquad\qquad\left.\le\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}+\frac{c_0}{n}\right\}\quad(i=1,2).
\end{aligned}
$$

Then, (4.21) and (4.22) imply

$$\Pr\{(X^n_{(i)}, Y^n_{(i)}) \in B_n^{(i)}\} \geq 1 - \exp(-n\gamma_n) \quad (i = 1, 2). \qquad (4.23)$$

Hence, by using (4.23), we obtain

$$\begin{aligned}
&P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - \lambda\right| \leq \varepsilon\right\} \\
&= P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - \lambda\right| \leq \varepsilon \text{ and } (\boldsymbol{x},\boldsymbol{y}) \in B_n^{(i)}\right\} \\
&\quad + P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - \lambda\right| \leq \varepsilon \text{ and } (\boldsymbol{x},\boldsymbol{y}) \notin B_n^{(i)}\right\} \\
&\leq P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - \lambda\right| \leq \varepsilon \text{ and } (\boldsymbol{x},\boldsymbol{y}) \in B_n^{(i)}\right\} \\
&\qquad + \Pr\left\{(X^n_{(i)}, Y^n_{(i)}) \notin B_n^{(i)}\right\} \\
&\leq P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - \lambda\right| \leq \varepsilon \text{ and } (\boldsymbol{x},\boldsymbol{y}) \in B_n^{(i)}\right\} \\
&\qquad + \exp(-n\gamma_n) \\
&= P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \lambda - \varepsilon \leq \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} \leq \lambda + \varepsilon \text{ and} \right. \\
&\qquad \left. \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})} - \gamma_n \leq \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} \leq \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})} + \frac{c_0}{n}\right\} \\
&\qquad + \exp(-n\gamma_n) \\
&= P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \right. \\
&\qquad \lambda - \varepsilon \leq \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} \leq \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})} + \frac{c_0}{n} \\
&\qquad \left. \text{and } \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})} - \gamma_n \leq \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} \leq \lambda + \varepsilon\right\} \\
&\qquad + \exp(-n\gamma_n) \\
&\leq P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : -\varepsilon - \frac{c_0}{n} \leq \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})} - \lambda \leq \varepsilon + \gamma_n\right\} \\
&\qquad + \exp(-n\gamma_n),
\end{aligned}$$

which is the left-hand side in equality of (4.5). Similarly, by using (4.23) we obtain

$$\begin{aligned}
&P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-\lambda\right|>\varepsilon\right\}\\
\le\;& P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-\lambda\right|>\varepsilon \text{ and } (\boldsymbol{x},\boldsymbol{y})\in B_n^{(i)}\right\}\\
&\qquad +\Pr\left\{(X_{(i)}^n,Y_{(i)}^n)\notin B_n^{(i)}\right\}\\
\le\;& P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-\lambda\right|>\varepsilon \text{ and } (\boldsymbol{x},\boldsymbol{y})\in B_n^{(i)}\right\}\\
&\qquad +\exp(-n\gamma_n)\\
=\;& P_n^{(i)}\Big\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\\
&\quad\left(\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}>\lambda+\varepsilon \text{ or } \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}<\lambda-\varepsilon\right) \text{ and}\\
&\quad\left.\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\le\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}+\frac{c_0}{n}\right\}\\
&\qquad +\exp(-n\gamma_n)\\
\le\;& P_n^{(i)}\Big\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\\
&\quad\lambda+\varepsilon<\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\le\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}+\frac{c_0}{n}\\
&\quad\left.\text{or } \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\gamma_n\le\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}<\lambda-\varepsilon\right\}\\
&\qquad +\exp(-n\gamma_n)\\
\le\;& P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}>\lambda+\varepsilon-\frac{c_0}{n}\right.\\
&\quad\left.\text{or } \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}<\lambda-\varepsilon+\gamma_n\right\}+\exp(-n\gamma_n)\\
=\;& 1-P_n^{(i)}\Big\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\\
&\quad\left.-\varepsilon+\gamma_n\le\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\lambda\le\varepsilon-\frac{c_0}{n}\right\}+\exp(-n\gamma_n).
\end{aligned}$$

This implies that

$$
\begin{aligned}
&P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-\lambda\right|\le\varepsilon\right\}\\
&\quad\ge\; P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : -\varepsilon+\gamma_n\le\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-\lambda\le\varepsilon-\frac{c_0}{n}\right\}\\
&\qquad\qquad -\exp(-n\gamma_n),
\end{aligned}
$$

which shows the right-hand side inequality of (4.5). This completes the proof of Lemma 2. □

Acknowledgement

I am grateful to Prof. Tomohiko Uyematsu for his helpful discussion and useful comments on this thesis. I would like to thank thank Prof. Fumio Kanaya of Shonan Institute of Technology, Prof. Te Sun Han, Prof. Kingo Kobayashi and Prof. Kouji Nagaoka of the University of Electonic Communication for their valuable comments.

Also, I wish to thank colleagues, frineds, and family for their encouragements; most notably, Dr. Daisuke Umehara, Dr. Mitsuru Hamada, Mrs. Junko Goto, Hidehiko Kawakami, Kisra Mohamed Saad, Kouji Maeda, Tomomichi Kimura, Ritsuko Kimura and Michiko Kimura.

Bibliography

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Inform. Theory* , vol.IT-19, pp.471-480, Jul. 1973.

[2] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. on Inform. Theory*, vol.IT-21, pp.226-228, Mar. 1975.

[3] S. Miyake and F. Kanaya, "Coding theorems on correlated general sources," *IEICE Trans. on Fundamentals*, vol.E78-A, No.9, pp.1063-1070, Sep. 1995.

[4] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. on Inform. Theory*, vol.IT-39, pp.752-772, May. 1993.

[5] T. S. Han, *Information-spectrum methods in information theory*, Baifukan, Tokyo, 1998 (in Japanese).

[6] A. Kaspi and T. Berger, "Rate-distortion for correlated sources with partially separated encoders," *IEEE Trans. on Inform. Theory*, vol.IT-28, pp.828-840, Nov. 1982.

[7] T. Ericson and J. Körner, "Successive encoding of correlated sources," *IEEE Trans. on Inform. Theory*, vol.IT-29, pp.390-395, May. 1983.

[8] Y. Oohama, "Universal coding for correlated sources with linked encoders," *IEEE Trans. on Inform. Theory*, vol.42, pp.837-847, May. 1996.

[9] T. S. Han, "Weak variable-length source coding," *submitted to IEEE Trans. on Inform. Theory.*

[10] R. G. Gallager, *Information theory and reliable communication*, New York, Wiley, 1968.