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

Akisato KIMURA* Tomohiko UYEMATSU†

Final version

August 25, 2003

*A. Kimura was with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, Ookayama, Meguro, Tokyo 152-8552, Japan. He is now with NTT Communication Science Laboratories, NTT Corporation, Morinosato Wakamiya 3-1, Atsugi-shi, Kanagawa 243-0198, Japan (E-mail: akisato@eye.brl.ntt.co.jp).

†T. Uyematsu is with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, Ookayama, Meguro, Tokyo 152-8552, Japan (E-mail: uyematsu@ieee.org).

Abstract: Coding problems for correlated information sources were first investigated by Slepian and Wolf. They considered the data compression system, called the SW system, where two sequences emitted from correlated sources are separately encoded to codewords, and sent to a single decoder which has to output original sequence pairs with small probability of error. In this paper, we investigate the coding problem of a modified SW system allowing two encoders to communicate with zero rate. First, we consider the fixed-length coding and clarify that the admissible rate region for general sources is equal to that of the original SW system. Next, we investigate the variable-length coding having the asymptotically vanishing probability of error. We clarify the admissible rate region for mixed sources characterized by two ergodic sources and show that this region is strictly wider than that for fixed-length codes. Further, we investigate the universal coding problem for memoryless sources in the system and show that the SW system with linked encoders has much more flexibility than the original SW system.

Index terms: admissible rate region, average length of codewords, linked encoders, mixed source, weak variable-length code,

I. Introduction

Coding problems for correlated information sources were first investigated by Slepian and Wolf [1]. They considered the data compression system, where two sequences of length n emitted from correlated sources are separately encoded to nR_1 and nR_2 bit codewords, and sent to a single decoder which has to output original sequence pairs with small probability of error. Slepian and Wolf established the admissible rate region (called *the SW region*), namely the closure of the set which consists of the rate (R_1, R_2) such that the error probability of decoding can be made arbitrarily small by letting n to be large. Their coding theorem may be regarded as a substantial starting point of multiterminal information theory, and many variations of their data compression system 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 coding theorem to the class of non-ergodic or non-stationary sources called *general sources* by using the method developed by Han and Verdú [4, 5].

In the system of Slepian and Wolf (called *the SW system*) neither of the encoders can observe the codeword generated by the other encoder. 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 codeword generated by the other encoder. Recently, Oohama [8] has investigated a more general case where there are some mutual linkages between two encoders of the SW system. He called this system the SWL system in the sense of an SW system having the *linkage* of two encoders. Especially, Oohama considered the case where two encoders can observe the codeword generated by the other encoder, and determined the admissible rate region. However, this coding problem allows encoders always to see the codeword of the other encoder, and is rather different from the original coding problem of the SW system, i.e. separate encoding and joint decoding problem. This motivates us to study the other aspect of the SWL system.

In this paper, we investigate the coding problems for the SWL system, where the coding rate for the mutual linkage between two encoders is *negligible*. This SWL system can be regarded as a generalization of the original SW system allowing two encoders to communicate with *zero rate*. First, we consider the fixed-length coding for general sources, and clarify that the admissible rate region is equal to that of the SW region. This shows that the linkage does not reduce the rate of fixed-length codes. Next, we investigate the variable-length coding having the asymptotically vanishing probability of error, and call it the *weak variable-length coding* [9]. We clarify the admissible rate region for mixed sources characterized by two ergodic sources, and show that this region is strictly wider than that for fixed-length codes. This result contrasts with that for the fixed-length coding. Even though the rate of the mutual linkage is zero, this linkage is enough to distinguish which ergodic source the input sequence is typical for, and effectively reduces the coding rate. Further, we investigate the universal coding for memoryless sources in the SWL system, and show that the arbitrary coding rate in the admission region depending on the source can be attained by the weak variable-length code. In case of universal coding, the linkage is used to estimate the probability distribution of the source, and gives drastic flexibility to variable-length coding. The organization of this paper is as follows: In Section II, we describe some coding systems for correlated sources and the formulation of the problem. Then, we clarify the admissible rate for the fixed-length coding. In Section III, we show main results without proofs, and give their proofs in Section IV.

II. Coding Systems for Correlated Sources

(a) Basic Definitions

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets and $\mathcal{B}$ be a binary set. Without loss of generality, we assume that $\mathcal{X} = \mathcal{Y} = \{1, 2, \cdots, M\}$ and $\mathcal{B} = \{0, 1\}$. We denote a set of all sequences of finite length by $\mathcal{B}^*$. 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 $\boldsymbol{X}$ and $\boldsymbol{Y}$ *ergodic sources*, and $(\boldsymbol{X}, \boldsymbol{Y})$ *correlated ergodic source*. The *entropy rate* of an ergodic source $\boldsymbol{X}$ is defined 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 entropy as defined in [10]. Similarly we define the *joint entropy rate* and the *conditional entropy rate* for a correlated ergodic source $(\boldsymbol{X},\boldsymbol{Y})$ by

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

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

Next, we show the formal definition of *general* correlated sources [4, 5]. The *general* correlated source is defined as an infinite sequence

$$(\boldsymbol{X},\boldsymbol{Y}) = \{(X^n,Y^n) = ((X_1^{(n)},Y_1^{(n)}),\cdots,(X_n^{(n)},Y_n^{(n)}))\}_{n=1}^{\infty}$$

of n-dimensional random variables, where each component random variable $(X_i^{(n)},Y_i^{(n)})$ $(1 \le i \le n)$ takes values in $\mathcal{X}\times\mathcal{Y}$. It should be noted here that each component of (X^n,Y^n) may change depending on the block length n. This implies that the sequence $(\boldsymbol{X},\boldsymbol{Y})$ is quite general in the sense that it may not satisfy even the consistency condition, where the consistency condition means that for any integers m,n such that $m<n$ it holds that $(X_i^{(m)},Y_i^{(m)}) \equiv (X_i^{(n)},Y_i^{(n)})$ for all $i=1,2,\cdots,m$. The class of sources thus defined covers a very wide range of sources including all nonstationary and/or nonergodic sources. A typical example of the general correlated source is a correlated mixed source described below.

A *correlated mixed source* $(\boldsymbol{X},\boldsymbol{Y})$ is defined by the following distribution:

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

where $0<\alpha<1$ and $P_n^{(i)}$ $(i=1,2)$ are 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

$$(\boldsymbol{X}_{(i)},\boldsymbol{Y}_{(i)}) \stackrel{\triangle}{=} \left\{\left(X_{(i)j},Y_{(i)j}\right)\right\}_{j=1}^{\infty} \quad (i=1,2).$$

The correlated mixed source is an example of a general source which satisfies the consistency condition.

(b) Slepian-Wolf Coding System

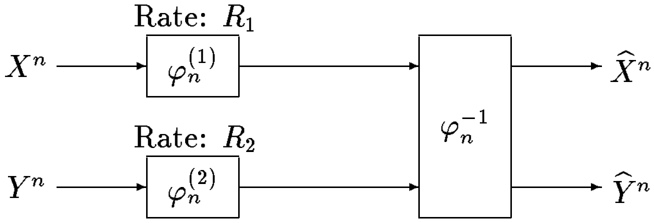

Figure 1: Slepian-Wolf coding system

Slepian and Wolf [1] studied the coding problem for two correlated sources, where two sequences from correlated sources are separately encoded, sent to a single decoder which has to output original sequence pairs (Figure 1). We call this data compression system *the Slepian-Wolf system (the SW system).*

Definition 1: A sequence $\{(\varphi_n^{(1)}, \varphi_n^{(2)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ of codes $(\varphi_n^{(1)}, \varphi_n^{(2)}, \varphi_n^{-1})$ 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

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

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

Definition 2: A rate pair (R_1, R_2) is *admissible for the SW system*, if there exists a SW code which satisfies

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

□

Definition 3 (The SW region): *The SW region* $\mathcal{R}_{SW}(\boldsymbol{X}, \boldsymbol{Y})$ is defined as

$$\mathcal{R}_{SW}(\boldsymbol{X}, \boldsymbol{Y}) = \{(R_1, R_2) : (R_1, R_2) \text{ is admissible for the SW system}\}.$$

□

Miyake and Kanaya [3] investigated the SW system for two correlated *general sources* and clarified the SW region as follows:

Theorem 1 [3]: For any correlated general source $(\boldsymbol{X},\boldsymbol{Y})$,

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

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

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

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

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

respectively. □

The next corollary can be obtained immediately from the definition of sup-entropy rate [5].

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

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

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

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

□

(c) Slepian-Wolf Coding System with Linked Encoders

Oohama [8] considered the coding problem for correlated sources, where two separate encoders of the SW code are mutually linked as shown in Figure 2. We call this compression system *the SWL system* in the sense of an SW system having the *linkage* of two encoders. First, we define the fixed-length coding for the SWL system (called *f-SWL system*).

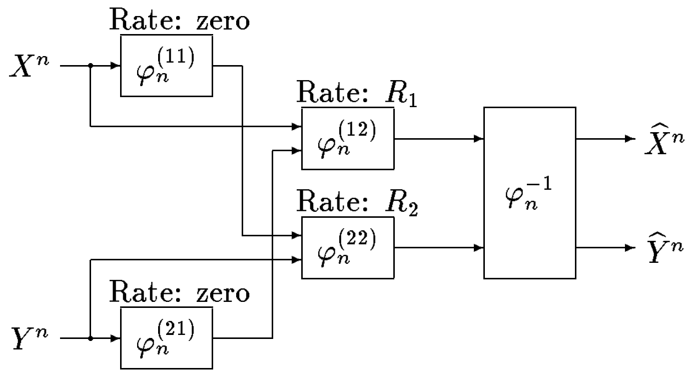

Figure 2: Slepian-Wolf coding system with linked encoders

Definition 5: A sequence $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ of codes $(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})$ is called a *fixed-length SWL code*, if the encoders

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

and the decoder $\varphi_n^{-1} : \mathcal{M}_n^{(12)} \times \mathcal{M}_n^{(22)} \to \mathcal{X}^n \times \mathcal{Y}^n$ satisfy

$$\lim_{n\to\infty} \Pr\{\varphi_n^{-1}(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)), \varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n))) \neq (X^n, Y^n)\} = 0. \tag{3}$$

□

Definition 6: A rate pair (R_1, R_2) is *admissible for the f-SWL system*, if there exists a f-SWL code which satisfies

$$\left.\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log M_n^{(11)} &= 0, \\
\limsup_{n\to\infty} \frac{1}{n} \log M_n^{(21)} &= 0, \\
\limsup_{n\to\infty} \frac{1}{n} \log M_n^{(12)} &\leq R_1, \\
\limsup_{n\to\infty} \frac{1}{n} \log M_n^{(22)} &\leq R_2.
\end{aligned}\right\} \tag{4}$$

□

It should be noted here that we consider the case where both rates of the first encoders $\varphi_n^{(11)}$ and $\varphi_n^{(21)}$ are zero. This implies that the outputs of the first encoders can be provided to the decoder with zero-rate as suitable prefixes in the outputs of the second encoders $\varphi_n^{(12)}$ and $\varphi_n^{(22)}$. Hence, even if the outputs of the first encoders are directly provided to the decoder, the admissible rate region remains the same. On the other hand, Oohama [8] considered the opposite case where both rates of the second encoders $\varphi_n^{(12)}$ and $\varphi_n^{(22)}$ are zero, and the decoder can see the outputs of the first encoders.

Definition 7: A rate pair (R_1, R_2) is *admissible for the f-SWL system*, if there exists a f-SWL code which satisfies (4).

□

Definition 8 (The f-SWL region): *The f-SWL region* is defined as

$$\mathcal{R}_{SWL}(\boldsymbol{X},\boldsymbol{Y}) = \{(R_1,R_2) : (R_1,R_2) \text{ is admissible for the f-SWL system}\}.$$

□

The next theorem clarifies the f-SWL region.

Theorem 2: For any correlated general source $(\boldsymbol{X},\boldsymbol{Y})$,

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

□

The proof of Theorem 2 is given in Appendix.

Comparing Theorem 1 and Theorem 2, we have

$$\mathcal{R}_{SWL}(\boldsymbol{X},\boldsymbol{Y}) = \mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y})$$

for any general source $(\boldsymbol{X},\boldsymbol{Y})$. This implies that the region of admissible rate pairs does not expand even if there are mutual linkages between two encoders.

The next corollary can be obtained immediately from Theorem 2 and Corollary 1.

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

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

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

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

□

(d) Weak Variable-Length Coding

We define weak variable-length coding for the SWL system (called *the wv-SWL system*).

Definition 9: A sequence $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ of codes $(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})$ is called a *wv-SWL 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. $$\lim_{n\to\infty} \Pr\{\varphi_n^{-1}(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)),\ \varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n))) \neq (X^n, Y^n)\} = 0. \qquad (5)$$

□

Definition 10: A rate pair (R_1, R_2) is *admissible for the wv-SWL system*, if there exists a wv-SWL code which satisfies

$$
\left.
\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(11)}(X^n))] &= 0, \\
\limsup_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(21)}(Y^n))] &= 0, \\
\limsup_{n\to\infty} \frac{1}{n} E\left[l(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)))\right] &\leq R_1, \\
\limsup_{n\to\infty} \frac{1}{n} E\left[l(\varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n)))\right] &\leq R_2,
\end{aligned}
\right\} \qquad (6)
$$

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

□

Definition 11: A rate pair (R_1, R_2) is *admissible for the wv-SWL system*, if there exists a wv-SWL code which satisfies (6). □

Definition 12: (The wv-SWL region) *The wv-SWL region* $\mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$ is defined as

$$\mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y}) = \{(R_1, R_2) : (R_1, R_2) \text{ is admissible for the wv-SWL system}\}.$$

□

III. Main Results

In this section, we shall clarify the wv-SWL rate region for correlated mixed sources. The next theorem is our main result.

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

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

□

According to Theorem 3 and Corollary 2, we conclude that the wv-SWL region strictly includes the f-SWL region, i.e.

$$\mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y}) \supset \mathcal{R}_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$$

for any correlated mixed source $(\boldsymbol{X}, \boldsymbol{Y})$. This implies that wv-SWL code can achieve strictly lower coding rate than f-SWL code.

It is instructive to note here how to construct the wv-SWL code. For a given rate $(R_1, R_2) \in \mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$, we can find two rate pairs $(R_{11}, R_{12}) \in \mathcal{R}_{SW}(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ and $(R_{21}, R_{22}) \in \mathcal{R}_{SW}(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ such that

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

Then, we prepare two SW code. One is $(f_n^{(1)}, f_n^{(2)}, f_n^{-1})$ for the source $(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ with rate (R_{11}, R_{12}), and the other is $(g_n^{(1)}, g_n^{(2)}, g_n^{-1})$ for the source $(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$ with rate (R_{21}, R_{22}).

The encoders $\varphi^{(11)}$ and $\varphi^{(21)}$ send the first N_n (e.g. $N_n = \log n$) symbols of each input sequence of length n to the other encoder. Sharing the pair of sequences of length N_n, we can select one of the SW codes, i.e. $(f_n^{(1)}, f_n^{(2)}, f_n^{-1})$ or $(g_n^{(1)}, g_n^{(2)}, g_n^{-1})$ depending on for which source the shared pair of sequences is typical. Then, the encoders $\varphi^{(12)}$ and $\varphi^{(22)}$ send to the decoder the first N_n symbols of the input sequence and the codewords of the selected SW code. Since the decoder can have the knowledge of which SW code the encoders employ, the estimate of the input pair of sequences can be obtained by using the corresponding decoder.

As a special case of Theorem 3, we immediately obtain the wv-SWL region for correlated ergodic sources.

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

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

□

Comparing Corollary 2 and 3, we have

$$\mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y}) = \mathcal{R}_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$$

for any correlated ergodic source $(\boldsymbol{X}, \boldsymbol{Y})$. Hence, we cannot improve the coding rate for ergodic sources even if we employ wv-SWL codes instead of f-SWL codes.

Next, we show that for a restricted class of ergodic sources, the rate pair in the region $\mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$ can be achieved without any linkage of encoders.

Theorem 4: Assume that a correlated mixed source $(\boldsymbol{X}, \boldsymbol{Y})$ satisfies both $H(\boldsymbol{X}_{(1)}) \neq H(\boldsymbol{X}_{(2)})$ and $H(\boldsymbol{Y}_{(1)}) \neq H(\boldsymbol{Y}_{(2)})$. Then, for any $(R_1, R_2) \in \mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$, we can construct a wv-SWL code $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ such that

$$\varphi_n^{(11)}(\boldsymbol{x}) = \varphi_n^{(21)}(\boldsymbol{y}) = \lambda \text{ (null string)},$$

for any $\boldsymbol{x} \in \mathcal{X}^n$, $\boldsymbol{y} \in \mathcal{Y}^n$ and positive integer n. □

This theorem indicates that for a restricted class of correlated mixed sources, the wv-SWL region can be achieved by the SW system. Further, in such a case, $\mathcal{R}_{SW}(\boldsymbol{X}, \boldsymbol{Y}) \subset \mathcal{R}^*_{SWL}(\boldsymbol{X}, \boldsymbol{Y})$, that is, a weak variable-length code can achieve smaller rate than the fixed-length code for the SW system.

In Theorems 3 and 4, we only considered the mixture of two ergodic sources, but it can be easily extended to the mixture of any finite number of ergodic sources. This implies the next corollary which shows a simple version of universal coding for the SWL system.

Corollary 4: Let $S = \{(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}), \cdots, (\boldsymbol{X}_{(m)}, \boldsymbol{Y}_{(m)})\}$ be a set of finite number of correlated ergodic sources. For any set of rate pairs $\{(R_{11}, R_{12}), (R_{21}, R_{22}), \cdots, (R_{m1}, R_{m2})\}$ which satisfies

$$R_{i1} \geq H(\boldsymbol{X}_{(i)}|\boldsymbol{Y}_{(i)}), \quad R_{i2} \geq H(\boldsymbol{Y}_{(i)}|\boldsymbol{X}_{(i)}), \quad R_{i1} + R_{i2} \geq H(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)}),$$

for $i = 1, 2, \cdots, m$, there exists a wv-SWL code $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$ such that

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(11)}(X_{(i)}^n))] &= 0, \\
\lim_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(21)}(Y_{(i)}^n))] &= 0, \\
\limsup_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(12)}(X_{(i)}^n, \varphi_n^{(21)}(Y_{(i)}^n)))] &\leq R_{i1}, \\
\limsup_{n\to\infty} \frac{1}{n} E[l(\varphi_n^{(22)}(Y_{(i)}^n, \varphi_n^{(11)}(X_{(i)}^n)))] &\leq R_{i2},
\end{aligned}$$

for $i = 1, 2, \cdots, m$.

This corollary shows that only if we know that the correlated source belongs to the given set S, we can encode a pair of sequences from the source $(X_{(i)}, Y_{(i)})$ with a rate pair (R_{i1}, R_{i2}) which is an arbitrary point in the admissible rate region of the source $(X_{(i)}, Y_{(i)})$. Though this corollary is valid for a finite set of sources, this result is rather different from the conventional universal coding for SW systems with fixed length codes [11, 12]. The next theorem shows that this property of the weak variable-length universal coding is strengthened for discrete memoryless sources (DMS's).

Theorem 5: Let S be a set of discrete memoryless correlated sources. Further, for every source $(\boldsymbol{X}, \boldsymbol{Y}) \in S$ with the joint probability Q, we correspond a rate pair $(R_1(Q), R_2(Q))$ which is an inner point of the SW region $\mathcal{R}_{SW}(\boldsymbol{X}, \boldsymbol{Y})$. We assume that $(R_1(Q), R_2(Q))$ is a continuous function of Q. Then, there exists a *universal* wv-SWL code

$$\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$$

such that for any joint source $(\boldsymbol{X}, \boldsymbol{Y}) \in S$ with the joint probability Q,

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

This corollary shows that for each memoryless joint source, we may arbitrarily specify the rate pair in its SW region, and that rate pair is admissible by universal wv-SWL code. Without linkage of encoders, we can only realize the (universal) fixed-length coding. However, using the linkage of encoders, we can realize the variable-length coding depending on the source.

IV. Proof of Theorems

Proof of Theorem 3:

(a) Converse part

For simplicity, we first introduce the following notations:

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

It should be noted that $H^*(\cdot)$ and $H^*(\cdot|\cdot)$ coincides with the entropy rate and the conditional entropy rate of the (correlated) mixed source, respectively. For example, $H^*(\boldsymbol{X})$ is equal to the entropy rate of the mixed source $\boldsymbol{X}$, and $H^*(\boldsymbol{X}|\boldsymbol{Y})$ is equal to the conditional entropy rate of the correlated mixed source $(\boldsymbol{X}, \boldsymbol{Y})$. Further, from the definition of entropy rate for ergodic sources, we have the following chain rules

$$H^*(\boldsymbol{X}, \boldsymbol{Y}) = H^*(\boldsymbol{X}) + H^*(\boldsymbol{Y}|\boldsymbol{X}) \;=\; H^*(\boldsymbol{Y}) + H^*(\boldsymbol{X}|\boldsymbol{Y}). \qquad (7)$$

According to [5, Theorem 1.10] or [9, Theorem 2.3], there exists a variable-length code $\{(\widehat{\varphi}_n, \widehat{\varphi}_n^{-1})\}_{n=1}^{\infty}$ for the mixed source $\boldsymbol{Y}$ 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

$$\limsup_{n\to\infty} \frac{1}{n} E[l(\widehat{\varphi}_n(Y^n))] \leq H^*(\boldsymbol{Y}), \tag{8}$$

$$\widehat{\varphi}_n^{-1}(\widehat{\varphi}_n(\boldsymbol{y})) = \boldsymbol{y} \qquad \forall \boldsymbol{y} \in \mathcal{Y}^n. \tag{9}$$

Then, for a given wv-SWL code $\{(\varphi_n^{(11)}, \varphi_n^{(12)}, \varphi_n^{(21)}, \varphi_n^{(22)}, \varphi_n^{-1})\}_{n=1}^{\infty}$, we construct a sequence of codes $\{(\psi_n, \psi_n^{-1})\}_{n=1}^{\infty}$ ($\psi_n : \mathcal{X}^n \times \mathcal{Y}^n \to \mathcal{B}^*$, $\psi_n^{-1} : \mathcal{B}^* \to \mathcal{X}^n \times \mathcal{Y}^n$) for the correlated mixed source $(\boldsymbol{X}, \boldsymbol{Y})$ as follows (see also Figure 2):

$$\left.\begin{array}{r} \psi_n(\boldsymbol{x}, \boldsymbol{y}) \stackrel{\triangle}{=} \varphi_n^{(11)}(\boldsymbol{x}) * \varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) * \widehat{\varphi}_n(\boldsymbol{y}) \\ \psi_n^{-1}(s_1 * s_2 * s_3) \stackrel{\triangle}{=} \varphi_n^{-1}(s_2, \varphi_n^{(22)}(\widehat{\varphi}_n^{-1}(s_3), s_1)) \end{array}\right\} \tag{10}$$

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

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

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 (5) and (9), the error probability of this code can be bounded as

$$\begin{aligned} &\Pr\{\psi_n^{-1}(\psi_n(X^n, Y^n)) \neq (X^n, Y^n)\} \\ &= \Pr\{\varphi_n^{-1}(\varphi_n^{(12)}(X^n, \varphi_n^{(21)}(Y^n)), \varphi_n^{(22)}(Y^n, \varphi_n^{(11)}(X^n))) \neq (X^n, Y^n)\} \\ &\to 0 \quad (n \to \infty). \end{aligned}$$

This implies that $\{(\psi_n, \psi_n^{-1})\}_{n=1}^{\infty}$ is a weak variable-length code. Hence, according to [5, Theorem 1.12] or [9, Theorem 3.1], it must satisfy

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

Hence, we have

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

where the equality ⓐ comes from (10), ⓑ from (6) and ⓒ from (8). This implies

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

where the last equality follows from (7). Therefore, if a wv-SWL code $\{(\varphi_n^{(11)},\varphi_n^{(12)},\varphi_n^{(21)},\varphi_n^{(22)},\varphi_n^{-1})\}_{n=1}^{\infty}$ satisfies

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

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

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

then $R_2\ge H^*(\boldsymbol{Y}|\boldsymbol{X})$. Further, $R_1+R_2\ge H^*(\boldsymbol{X},\boldsymbol{Y})$ is obvious from [5, Theorem 1.12] or [9, Theorem 3.1]. This completes the proof of the converse part.

(b) Achievability part

First, we consider the case where two ergodic sources $(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)})$ and $(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)})$ cannot be discriminated by the entropy rate, i.e. a correlated mixed source $(\boldsymbol{X},\boldsymbol{Y})$ satisfies the following three conditions: $H(\boldsymbol{X}_{(1)})=H(\boldsymbol{X}_{(2)})$, $H(\boldsymbol{Y}_{(1)})=H(\boldsymbol{Y}_{(2)})$, and $H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)})=H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)})$. In this case, we have $\mathcal{R}_{SW}(\boldsymbol{X},\boldsymbol{Y})=\mathcal{R}^*_{SWL}(\boldsymbol{X},\boldsymbol{Y})$. Hence, for a given mixed source $(\boldsymbol{X},\boldsymbol{Y})$ and any rate pair $(R_1,R_2)\in\mathcal{R}^*_{SWL}(\boldsymbol{X},\boldsymbol{Y})$, there exists an SW code $\{(\varphi_n^{(1)},\varphi_n^{(2)},\varphi_n^{(-1)})\}_{n=1}^{\infty}$ which satisfies

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

This shows the existence of the wv-SWL code for any $(R_1,R_2)\in\mathcal{R}^*_{SWL}(\boldsymbol{X},\boldsymbol{Y})$, since the SW code is a special case of the wv-SWL code.

Next, we consider the case where two ergodic sources $(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)})$ and $(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)})$ can be discriminated by the entropy rate, i.e. a correlated mixed source $(\boldsymbol{X},\boldsymbol{Y})$ satisfies at least one of the following conditions ① – ③:

$$\left.\begin{array}{ll}① & H(\boldsymbol{X}_{(1)})\neq H(\boldsymbol{X}_{(2)})\\ ② & H(\boldsymbol{Y}_{(1)})\neq H(\boldsymbol{Y}_{(2)})\\ ③ & H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)})\neq H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)})\end{array}\right\} \tag{11}$$

Here, we introduce two fundamental lemmas.

Lemma 1 (Asymptotic Equipartition Property (AEP)) [10]: For any $\varepsilon>0,\delta>0$ and ergodic sources $(\boldsymbol{X}_{(i)},\boldsymbol{Y}_{(i)})$ $(i=1,2)$, there exists an integer $n_0(\varepsilon,\delta,\boldsymbol{X}_{(i)},\boldsymbol{Y}_{(i)})$ such that for all $n\ge n_0(\varepsilon,\delta,\boldsymbol{X}_{(i)},\boldsymbol{Y}_{(i)})$

$$\begin{aligned}P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x})}-H(\boldsymbol{X}_{(i)})\right|\ge\varepsilon\right\}&\le\delta,\\ P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{y})}-H(\boldsymbol{Y}_{(i)})\right|\ge\varepsilon\right\}&\le\delta,\\ P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}-H(\boldsymbol{X}_{(i)},\boldsymbol{Y}_{(i)})\right|\ge\varepsilon\right\}&\le\delta,\end{aligned}$$

simultaneously hold. □

Lemma 2: For a correlated mixed source $(\boldsymbol{X},\boldsymbol{Y})$ given by (1), and any $\gamma>0$, we have

$$\begin{aligned}P_n^{(i)}\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}-\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x},\boldsymbol{y})}\right|\le\max(\gamma,c_0/n)\right\}\\ \ge 1-\exp(-n\gamma),\end{aligned} \tag{12}$$

for $i = 1, 2$ and any integer $n > 0$, where $c_0 = -\log \min(\alpha, 1-\alpha)$. $\square$

The proof of Lemma 2 is given in Appendix.

(Step 1) Preliminaries

For any subset $A_n \subset \mathcal{X}^n \times \mathcal{Y}^n$, we introduce the notations

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

By the notations above, we immediately have

$$
\begin{aligned}
\Pr\{(X^n, Y^n) \in A_n\} &= \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n} P_n(\boldsymbol{x},\boldsymbol{y}) \\
&= \alpha \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n} P^{(1)}_n(\boldsymbol{x},\boldsymbol{y}) + (1-\alpha) \sum_{(\boldsymbol{x},\boldsymbol{y})\in A_n} P^{(2)}_n(\boldsymbol{x},\boldsymbol{y}) \\
&= \alpha \Pr\{(X^n_{(1)}, Y^n_{(1)}) \in A_n\} + (1-\alpha)\Pr\{(X^n_{(2)}, Y^n_{(2)}) \in A_n\}. \qquad (13)
\end{aligned}
$$

From (11), we can chose an ε such that

$$
\begin{aligned}
0 < 3\varepsilon < \; & \max\{|H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) - H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})|, \\
& |H(\boldsymbol{X}_{(1)}) - H(\boldsymbol{X}_{(2)})|, \; |H(\boldsymbol{Y}_{(1)}) - H(\boldsymbol{Y}_{(2)})|\}. \qquad (14)
\end{aligned}
$$

Then, we define subsets $T^{(i)}_n$ $(i = 1, 2)$ of $\mathcal{X}^n \times \mathcal{Y}^n$ by

$$
\begin{aligned}
T^{(i)}_n \stackrel{\triangle}{=} \Bigg\{ (\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \quad & \left| \frac{1}{n} \log \frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} - H(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)}) \right| \le \varepsilon, \\
& \left| \frac{1}{n} \log \frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(i)}) \right| \le \varepsilon, \\
& \left| \frac{1}{n} \log \frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(i)}) \right| \le \varepsilon, \Bigg\}. \qquad (15)
\end{aligned}
$$

It should be noted that $T^{(i)}_n$ is characterized not by $P^{(i)}_n(\boldsymbol{x},\boldsymbol{y})$ but by $P_n(\boldsymbol{x},\boldsymbol{y})$. According to (14) and (15), it is easy to see that $T^{(1)}_n \cap T^{(2)}_n = \emptyset$, i.e.

$$
\Pr\{(X^n, Y^n) \in T^{(1)}_n \cap T^{(2)}_n\} = 0 \qquad (16)
$$

Suppose that $\overline{T}_n^{(i)}$ denotes the complement of $T_n^{(i)}$. Then, from (13), we immediately obtain

$$\begin{aligned}
&\Pr\{(X^n, Y^n) \notin T_n^{(1)} \cup T_n^{(2)}\} \\
&= \alpha \Pr\{(X^n_{(1)}, Y^n_{(1)}) \notin T_n^{(1)} \cup T_n^{(2)}\} + (1-\alpha)\Pr\{(X^n_{(2)}, Y^n_{(2)}) \notin T_n^{(1)} \cup T_n^{(2)}\} \\
&\le \alpha \Pr\{(X^n_{(1)}, Y^n_{(1)}) \notin T_n^{(1)}\} + (1-\alpha)\Pr\{(X^n_{(2)}, Y^n_{(2)}) \notin T_n^{(2)}\}. \qquad (17)
\end{aligned}$$

From the definition of $T_n^{(i)}$, we have

$$\begin{aligned}
&\Pr\{(X^n_{(i)}, Y^n_{(i)}) \notin T_n^{(i)}\} \\
&\le P_n^{(i)}\left\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - H(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)})\right| > \varepsilon\right\} \\
&\quad + P_n^{(i)}\left\{\boldsymbol{x} \in \mathcal{X}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(i)})\right| > \varepsilon\right\} \\
&\quad + P_n^{(i)}\left\{\boldsymbol{y} \in \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(i)})\right| > \varepsilon\right\}. \qquad (18)
\end{aligned}$$

According to Lemma 1 and Lemma 2 with $\gamma = \varepsilon/2$, for any $\delta > 0$ and sufficiently large n, the first term in (18) can be bounded by

$$\begin{aligned}
&P_n^{(i)}\left\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - H(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)})\right| > \varepsilon\right\} \\
&\le P_n^{(i)}\Bigg\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - H(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)})\right| > \varepsilon \\
&\qquad \text{and } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})}\right| \le \max(\varepsilon/2, c_0/n)\Bigg\} \\
&\qquad + P_n^{(i)}\left\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - \frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})}\right| > \max(\varepsilon/2, c_0/n)\right\} \\
&\le P_n^{(i)}\left\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} - H(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)})\right| > \varepsilon/2\right\} + \exp(-n\gamma) \\
&\le \delta.
\end{aligned}$$

In a similar manner, the second and third terms in (18) satisfy

$$\begin{aligned}
P_n^{(i)}\left\{\boldsymbol{x} \in \mathcal{X}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(i)})\right| > \varepsilon\right\} &\le \delta, \\
P_n^{(i)}\left\{\boldsymbol{y} \in \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(i)})\right| > \varepsilon\right\} &\le \delta,
\end{aligned}$$

for sufficiently large n. Since $\delta > 0$ can be chosen arbitrarily small, the right hand side of (18) vanishes, that is,

$$\lim_{n\to\infty} \Pr\{(X^n_{(i)}, Y^n_{(i)}) \notin T^{(i)}_n\} = 0 \qquad (i = 1, 2). \tag{19}$$

Substituting (19) into (17), we obtain

$$\lim_{n\to\infty} \Pr\{(X^n, Y^n) \notin T^{(1)}_n \cup T^{(2)}_n\} = 0. \tag{20}$$

(Step 2) Determination of rate pairs (R_{11}, R_{12}) *and* (R_{21}, R_{22})

Suppose that we are given a rate pair (R_1, R_2) which satisfies

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

It is easy to see that there exists a pair $(\tilde{R}_1, \tilde{R}_2)$, $c_1 \geq 0$ and $c_2 \geq 0$ such that

$$\tilde{R}_1 + \tilde{R}_2 = \alpha H(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$$
$$\alpha H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}) \leq \tilde{R}_1 \leq \alpha H(\boldsymbol{X}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}),$$

and (R_1, R_2) can be written as

$$\begin{aligned} R_1 &= \tilde{R}_1 + c_1, \\ R_2 &= \tilde{R}_2 + c_2. \end{aligned}$$

Further, there exists $0 \leq \beta \leq 1$ such that

$$\begin{aligned} \tilde{R}_1 &= \beta(\alpha H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)})) \\ &\quad +(1-\beta)(\alpha H(\boldsymbol{X}_{(1)}) + (1-\alpha)H(\boldsymbol{X}_{(2)})) \\ \tilde{R}_2 &= \beta(\alpha H(\boldsymbol{Y}_{(1)}) + (1-\alpha)H(\boldsymbol{Y}_{(2)})) \\ &\quad +(1-\beta)(\alpha H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) + (1-\alpha)H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)})). \end{aligned}$$

By using c_1 and c_2 and β, define two rate pairs (R_{11}, R_{12}) and (R_{21}, R_{22}) as

$$\begin{aligned} R_{11} &\triangleq \beta H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}) + (1-\beta)H(\boldsymbol{Y}_{(1)}) + c_1, \\ R_{12} &\triangleq \beta H(\boldsymbol{Y}_{(1)}) + (1-\beta)H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}) + c_2, \\ R_{21} &\triangleq \beta H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}) + (1-\beta)H(\boldsymbol{Y}_{(2)}) + c_1, \\ R_{22} &\triangleq \beta H(\boldsymbol{Y}_{(2)}) + (1-\beta)H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}) + c_2. \end{aligned}$$

We can easily confirm that

$$\alpha R_{11} + (1-\alpha)R_{21} = R_1, \quad (21)$$

$$\alpha R_{12} + (1-\alpha)R_{22} = R_2, \quad (22)$$

and

$$\left.\begin{array}{ll} R_{11} \geq H(\boldsymbol{X}_{(1)}|\boldsymbol{Y}_{(1)}), & R_{21} \geq H(\boldsymbol{X}_{(2)}|\boldsymbol{Y}_{(2)}), \\ R_{12} \geq H(\boldsymbol{Y}_{(1)}|\boldsymbol{X}_{(1)}), & R_{22} \geq H(\boldsymbol{Y}_{(2)}|\boldsymbol{X}_{(2)}), \\ R_{11}+R_{12} \geq H(\boldsymbol{X}_{(1)},\boldsymbol{Y}_{(1)}), & R_{21}+R_{22} \geq H(\boldsymbol{X}_{(2)},\boldsymbol{Y}_{(2)}). \end{array}\right\} \quad (23)$$

(Step 3) Construction of wv-SWL code

From (23) and Corollary 1, we can see $(R_{11}, R_{12}) \in \mathcal{R}_{SW}(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$ and $(R_{21}, R_{22}) \in \mathcal{R}_{SW}(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$. Hence, there exists an SW code $\{(f_n^{(1)}, f_n^{(2)}, f_n^{-1})\}_{n=1}^{\infty}$ for the ergodic source $(\boldsymbol{X}_{(1)}, \boldsymbol{Y}_{(1)})$, where

$$f_n^{(1)} : \mathcal{X}^n \to \{1,2,\cdots,M_n^{(1)}(f_n)\},$$
$$f_n^{(2)} : \mathcal{Y}^n \to \{1,2,\cdots,M_n^{(2)}(f_n)\},$$

and $\{(f_n^{(1)}, f_n^{(2)}, f_n^{-1})\}_{n=1}^{\infty}$ satisfies

$$\left.\begin{array}{l} \limsup_{n\to\infty} \frac{1}{n}\log M_n^{(1)}(f_n) \leq R_{11}, \\ \limsup_{n\to\infty} \frac{1}{n}\log M_n^{(2)}(f_n) \leq R_{12}, \\ \lim_{n\to\infty} \Pr\{f_n^{-1}(f_n^{(1)}(X_{(1)}^n), f_n^{(2)}(Y_{(1)}^n)) \neq (X_{(1)}^n, Y_{(1)}^n)\} = 0. \end{array}\right\} \quad (24)$$

Similarly, there exists an SW code $\{(g_n^{(1)}, g_n^{(2)}, g_n^{-1})\}_{n=1}^{\infty}$ for the ergodic source $(\boldsymbol{X}_{(2)}, \boldsymbol{Y}_{(2)})$, where

$$g_n^{(1)} : \mathcal{X}^n \to \{1,2,\cdots,M_n^{(1)}(g_n)\},$$
$$g_n^{(2)} : \mathcal{Y}^n \to \{1,2,\cdots,M_n^{(2)}(g_n)\},$$

and $\{(g_n^{(1)}, g_n^{(2)}, g_n^{-1})\}_{n=1}^{\infty}$ satisfies

$$\left.\begin{array}{l} \limsup_{n\to\infty} \frac{1}{n}\log M_n^{(1)}(g_n) \leq R_{21}, \\ \limsup_{n\to\infty} \frac{1}{n}\log M_n^{(2)}(g_n) \leq R_{22}, \\ \lim_{n\to\infty} \Pr\{g_n^{-1}(g_n^{(1)}(X_{(2)}^n), g_n^{(2)}(Y_{(2)}^n)) \neq (X_{(2)}^n, Y_{(2)}^n)\} = 0. \end{array}\right\} \quad (25)$$

Further, since $\mathcal{X}$ is finite, for any positive integer m, we can easily construct a binary fixed-length code $(\widehat{\varphi}_m^{(1)}, \widehat{\varphi}_m^{(1)-1})$ for the source $\boldsymbol{X}$ such that

$$E[\ell(\widehat{\varphi}_m^{(1)}(\boldsymbol{x}))] < m \log M + 1 \qquad \forall \boldsymbol{x} \in \mathcal{X}^m,$$
$$\widehat{\varphi}_m^{(1)-1}(\widehat{\varphi}_m^{(1)}(\boldsymbol{x})) = \boldsymbol{x} \qquad \forall \boldsymbol{x} \in \mathcal{X}^m.$$

A binary fixed-length code $(\widehat{\varphi}_n^{(2)}, \widehat{\varphi}_n^{(2)-1})$ for the source $\boldsymbol{Y}$ can be constructed similarly.

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

$$0 < N_n \le n, \quad \lim_{n\to\infty} N_n = \infty, \quad \lim_{n\to\infty} \frac{N_n}{n} = 0. \tag{26}$$

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

$\varphi_n^{(11)}$: $\varphi_n^{(11)}(\boldsymbol{x}) \triangleq \widehat{\varphi}_{N_n}^{(1)}(\boldsymbol{x}_1)$, where $\boldsymbol{x}_1 \in \mathcal{X}^{N_n}$ is the first N_n symbols of $\boldsymbol{x}$.

$\varphi_n^{(21)}$: $\varphi_n^{(21)}(\boldsymbol{y}) \triangleq \widehat{\varphi}_{N_n}^{(2)}(\boldsymbol{y}_1)$, where $\boldsymbol{y}_1 \in \mathcal{Y}^{N_n}$ is the first N_n symbols of $\boldsymbol{y}$.

$\varphi_n^{(12)}$: Decode $\boldsymbol{y}_1$ 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 $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) \triangleq \varphi_n^{(21)}(\boldsymbol{y}) * f_n^{(1)}(\boldsymbol{x})$.
If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$, then $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) \triangleq \varphi_n^{(21)}(\boldsymbol{y}) * g_n^{(1)}(\boldsymbol{x})$.
Otherwise, $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})) \triangleq \lambda$.

$\varphi_n^{(22)}$: Decode $\boldsymbol{x}_1$ 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 $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})) \triangleq \varphi_n^{(11)}(\boldsymbol{x}) * f_n^{(2)}(\boldsymbol{y})$.
If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$, then $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})) \triangleq \varphi_n^{(11)}(\boldsymbol{x}) * g_n^{(2)}(\boldsymbol{y})$.
Otherwise, $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x})) \triangleq \lambda$.

φ_n^{-1}: Since both $\varphi_n^{(11)}$ and $\varphi_n^{(21)}$ are prefix codes, for given $\varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y}))$ and $\varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x}))$, we have $\varphi_n^{(11)}(\boldsymbol{x})$, $\varphi_n^{(21)}(\boldsymbol{y})$, and either $(f_n^{(1)}(\boldsymbol{x}), f_n^{(2)}(\boldsymbol{y}))$ or $(g_n^{(1)}(\boldsymbol{x}), g_n^{(2)}(\boldsymbol{y}))$. We first decode $(\boldsymbol{x}_1, \boldsymbol{y}_1)$ from $(\varphi_n^{(11)}(\boldsymbol{x}), \varphi_n^{(21)}(\boldsymbol{y}))$ and, 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 $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) = f_n^{-1}(f_n^{(1)}(\boldsymbol{x}), f_n^{(2)}(\boldsymbol{y}))$.
If $(\boldsymbol{x}_1, \boldsymbol{y}_1) \in T_{N_n}^{(2)} \cap \overline{T}_{N_n}^{(1)}$, then $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) = g_n^{-1}(g_n^{(1)}(\boldsymbol{x}), g_n^{(2)}(\boldsymbol{y}))$.
Otherwise, we declare an error.

In the proposed code, we cannot obtain original sequence pairs 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), (13) and (19), the probability of the event (i) can be bounded as

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

In a similar manner, 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 (16) and (20). Therefore, the probability of decoding error vanishes as $n \to \infty$.

(Step 4) Evaluation of the average length of codeword

Lastly, we investigate the average length of codeword for the proposed code. For any

$\delta > 0$ and sufficiently large n, according to (13), (24) and (25), we have

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

By using (19) and (26), we obtain

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

In a similar manner,

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

Since $\delta > 0$ is arbitrary, we can conclude

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

where the last equalities come from (21) and (22). Hence, (R_1, R_2) is admissible for the wv-SWL system. This completes the proof of Theorem 3.

The extension of the theorem to the mixture of any finite number of ergodic sources is immediate. Further, the proof of Corollary 4 can be done similarly by considering a fixed length SW code for each source $(\boldsymbol{X}_{(i)}, \boldsymbol{Y}_{(i)})$ with a rate pair (R_{i1}, R_{i2}) $(i = 1, 2, \cdots, m)$ instead of $\{(f_n^{(1)}, f_n^{(2)}, f_n^{-1})\}_{n=1}^{\infty}$ and $\{(g_n^{(1)}, g_n^{(2)}, g_n^{-1})\}_{n=1}^{\infty}$ in Step 2 and Step 3.

□

Proof of Theorem 4:

By choosing ε such that

$$0 < 3\varepsilon < \min\{|H(\boldsymbol{X}_{(1)}) - H(\boldsymbol{X}_{(2)})|, |H(\boldsymbol{Y}_{(1)}) - H(\boldsymbol{Y}_{(2)})|\}$$

instead of (14), we have

$$\left\{\boldsymbol{x} \in \mathcal{X}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(1)})\right| \leq \varepsilon \text{ and } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(2)})\right| \leq \varepsilon\right\} = \emptyset$$
$$\left\{\boldsymbol{y} \in \mathcal{Y}^n : \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(1)})\right| \leq \varepsilon \text{ and } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(2)})\right| \leq \varepsilon\right\} = \emptyset$$

This implies that we can determine whether $(\boldsymbol{x}, \boldsymbol{y}) \in T_n^{(1)}$ or $(\boldsymbol{x}, \boldsymbol{y}) \in T_n^{(2)}$ by seeing only $\boldsymbol{x}$ or $\boldsymbol{y}$. Hence, by using the codes $\{(f_n^{(1)}, f_n^{(2)}, f_n^{-1})\}_{n=1}^{\infty}$ and $\{(g_n^{(1)}, g_n^{(2)}, g_n^{-1})\}_{n=1}^{\infty}$ described in the proof of Theorem 3, we construct a wv-SWL code $\{(\tilde{\varphi}_n^{(11)}, \tilde{\varphi}_n^{(12)}, \tilde{\varphi}_n^{(21)}, \tilde{\varphi}_n^{(22)}, \tilde{\varphi}_n^{-1})\}_{n=1}^{\infty}$ as follows:

$$\tilde{\varphi}_n^{(11)}(\boldsymbol{x}) \triangleq \lambda$$

$$\tilde{\varphi}_n^{(21)}(\boldsymbol{y}) \triangleq \lambda$$

$$\tilde{\varphi}_n^{(12)}(\boldsymbol{x}, \tilde{\varphi}_n^{(21)}(\boldsymbol{y})) \triangleq \begin{cases} 0 * f_n^{(1)}(\boldsymbol{x}) & \text{if } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(1)})\right| \le \varepsilon, \\ 1 * g_n^{(1)}(\boldsymbol{x}) & \text{if } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x})} - H(\boldsymbol{X}_{(2)})\right| \le \varepsilon, \\ \lambda & \text{otherwise,} \end{cases}$$

$$\tilde{\varphi}_n^{(22)}(\boldsymbol{y}, \tilde{\varphi}_n^{(11)}(\boldsymbol{x})) \triangleq \begin{cases} 0 * f_n^{(2)}(\boldsymbol{y}) & \text{if } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(1)})\right| \le \varepsilon, \\ 1 * g_n^{(2)}(\boldsymbol{y}) & \text{if } \left|\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y})} - H(\boldsymbol{Y}_{(2)})\right| \le \varepsilon, \\ \lambda & \text{otherwise.} \end{cases}$$

Lastly we describe the decoder $\tilde{\varphi}_n^{-1}$. For a given pair of codewords $s_1 * s_2 = \tilde{\varphi}_n^{(12)}(\boldsymbol{x}, \lambda)$ and $s_3 * s_4 = \tilde{\varphi}_n^{(22)}(\boldsymbol{y}, \lambda)$ with $s_1, s_3 \in \mathcal{B}$ and $s_2, s_4 \in \mathcal{B}^*$, we output an estimate $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) \in \mathcal{X}^n \times \mathcal{Y}^n$ as follows:

If $s_1 = s_3 = 0$ then $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) = f_n^{-1}(s_2, s_4)$.

If $s_1 = s_3 = 1$ then $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}) = g_n^{-1}(s_2, s_4)$.

Otherwise, we declare an error.

By using this code, in a similar manner as the proof of Theorem 3, we can show that the rate pair (R_1, R_2) is admissible for the wv-SWL system. □

Proof of Theorem 5:

We only show the construction of universal code. The proof of the theorem can be done in a similar manner as the proof of Theorem 3.

The encoder $\varphi_n^{(11)}$ and $\varphi_n^{(21)}$ are the same ones which defined in the proof of Theorem 3. After sharing the pair of sequence $\{(x_i, y_i)\}_{i=1}^{N_n}$ between two encoders, each encoder calculates the joint type [12] of the shared sequence i.e.

$$P_{xy}(a, b) \triangleq \frac{1}{N_n}|\{i : x_i = a \text{ and } y_i = b, 1 \le i \le N_n\}|, \qquad \forall (a, b) \in \mathcal{X} \times \mathcal{Y}.$$

Let $\mathcal{P}_{N_n}$ denote the set of possible joint types of length N_n. It is well-known that $|\mathcal{P}_{N_n}| \le (N_n + 1)^{|\mathcal{X}||\mathcal{Y}|}$ (see e.g. [12]). For any $\delta > 0$ and any type $Q \in \mathcal{P}_{N_n}$, there exists *universal*

SW code $(f_Q^{(1)}, f_Q^{(2)}, f_Q^{-1})$ with block length n and a rate pair $(R_1(Q)+\delta, R_2(Q)+\delta)$ [11, 12]. In the above notation, it should be noted that the joint probability Q determines only the rate of the universal SW code. Then, we define the second encoders as

$$\begin{aligned}\varphi_n^{(21)}(\boldsymbol{x}) &= \varphi_n^{(11)} * f_{P_{xy}}^{(1)}(\boldsymbol{x}),\\ \varphi_n^{(22)}(\boldsymbol{y}) &= \varphi_n^{(11)} * f_{P_{xy}}^{(2)}(\boldsymbol{y}),\end{aligned}$$

where P_{xy} denotes the type of the shared sequence. This implies that we choose the universal SW code depending on the joint type P_{xy}. Since the decoder can have the knowledge of the type P_{xy}, it can output the estimate $f_{P_{xy}}^{-1}(f_{P_{xy}}^{(1)}(\boldsymbol{x}), f_{P_{xy}}^{(2)}(\boldsymbol{y}))$.

□

Appendix

Proof of Theorem 2:

The proof of the achievability part is obvious from Theorem 1. So, we shall only prove the converse part for the f-SWL system under the condition

$$\left.\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log M_n^{(11)}M_n^{(12)} &\le R_1\\ \limsup_{n\to\infty}\frac{1}{n}\log M_n^{(21)}M_n^{(22)} &\le R_2\end{aligned}\right\} \tag{27}$$

which is weaker than the condition (4).

First, for any $\gamma > 0$, we define the sets $\tilde{T}_n^{(i)}$ $(i = 1, 2, 3)$ and S_n by

$$\begin{aligned}\tilde{T}_n^{(1)} &\triangleq \left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}|\boldsymbol{y})} \ge \frac{1}{n}\log M_n^{(1)} + \gamma\right\},\\ \tilde{T}_n^{(2)} &\triangleq \left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y}|\boldsymbol{x})} \ge \frac{1}{n}\log M_n^{(2)} + \gamma\right\},\\ \tilde{T}_n^{(3)} &\triangleq \left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})} \ge \frac{1}{n}\log M_n^{(1)}M_n^{(2)} + \gamma\right\},\\ S_n &\triangleq \left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n : \varphi_n^{-1}(\varphi_n^{(1)}(\boldsymbol{x},\boldsymbol{y}),\varphi_n^{(2)}(\boldsymbol{x},\boldsymbol{y})) = (\boldsymbol{x},\boldsymbol{y})\right\},\end{aligned}$$

where

$$\begin{aligned}\varphi_n^{(1)}(\boldsymbol{x},\boldsymbol{y}) &\triangleq (\varphi_n^{(11)}(\boldsymbol{x}), \varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y}))),\\ \varphi_n^{(2)}(\boldsymbol{x},\boldsymbol{y}) &\triangleq (\varphi_n^{(21)}(\boldsymbol{y}), \varphi_n^{(22)}(\boldsymbol{y}, \varphi_n^{(11)}(\boldsymbol{x}))),\\ M_n^{(1)} &\triangleq M_n^{(11)}M_n^{(12)},\\ M_n^{(2)} &\triangleq M_n^{(21)}M_n^{(22)}.\end{aligned}$$

By letting $e_n \stackrel{\triangle}{=} \Pr\{(X^n, Y^n) \notin S_n\}$, we obtain

$$\begin{aligned}
&\Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)}\} \\
&= \Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)} \cap S_n\} + \Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)} \cap \overline{S}_n\} \\
&\leq \Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)} \cap S_n\} + \Pr\{(X^n, Y^n) \notin S_n\} \\
&= \Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)} \cap S_n\} + e_n. \qquad (28)
\end{aligned}$$

Note that if $(\boldsymbol{x}, \boldsymbol{y}) \in \tilde{T}_n^{(1)}$ then $P_n(\boldsymbol{x}|\boldsymbol{y}) \leq \exp(-n\gamma)/M_n^{(1)}$. Hence, the first term of (28) can be evaluated as

$$\begin{aligned}
\Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)} \cap S_n\} &= \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \tilde{T}_n^{(1)} \cap S_n} P_n(\boldsymbol{x}, \boldsymbol{y}) \\
&\leq \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \tilde{T}_n^{(1)} \cap S_n} P_n(\boldsymbol{y}) \frac{\exp(-n\gamma)}{M_n^{(1)}} \\
&\leq \sum_{(\boldsymbol{x},\boldsymbol{y}) \in S_n} P_n(\boldsymbol{y}) \frac{\exp(-n\gamma)}{M_n^{(1)}} \\
&= \sum_{\boldsymbol{y} \in \mathcal{Y}^n} P_n(\boldsymbol{y}) |S_n(\boldsymbol{y})| \frac{\exp(-n\gamma)}{M_n^{(1)}}, \qquad (29)
\end{aligned}$$

where $S_n(\boldsymbol{y}) \stackrel{\triangle}{=} \{\boldsymbol{x} \in \mathcal{X}^n : (\boldsymbol{x}, \boldsymbol{y}) \in S_n\}$. Since $\varphi_n^{(1)}(\boldsymbol{x}, \boldsymbol{y}) = (\varphi_n^{(11)}(\boldsymbol{x}), \varphi_n^{(12)}(\boldsymbol{x}, \varphi_n^{(21)}(\boldsymbol{y})))$ and $\varphi_n^{(2)}$ is a function of $\boldsymbol{y}$ and $\varphi^{(11)}$, $\varphi_n^{(2)}$ is a function of $\boldsymbol{y}$ and $\varphi_n^{(1)}$. Thus, for a given $\boldsymbol{y} \in \mathcal{Y}^n$, the number of sequences $\boldsymbol{x} \in \mathcal{X}^n$ which satisfy

$$\varphi_n^{-1}(\varphi_n^{(1)}(\boldsymbol{x}, \boldsymbol{y}), \varphi_n^{(2)}(\boldsymbol{x}, \boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y})$$

is at most the number of codewords of $\varphi_n^{(1)}$. Then, we have

$$|S_n(\boldsymbol{y})| \leq |\varphi_n^{(1)}(\mathcal{X}^n, \mathcal{Y}^n)| \leq M_n^{(1)} \quad \forall \boldsymbol{y} \in \mathcal{Y}^n.$$

Substituting this inequality into (29), we obtain

$$\Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)} \cap S_n\} \leq \exp(-n\gamma).$$

Hence (28) can be rewritten as

$$\Pr\{(X^n, Y^n) \in \tilde{T}_n^{(1)}\} \leq e_n + \exp(-n\gamma). \qquad (30)$$

In a similar manner, we obtain

$$\Pr\{(X^n,Y^n)\in\tilde{T}_n^{(2)}\} \;\le\; e_n+\exp(-n\gamma), \tag{31}$$
$$\Pr\{(X^n,Y^n)\in\tilde{T}_n^{(3)}\} \;\le\; e_n+\exp(-n\gamma). \tag{32}$$

Now, let (R_1,R_2) be admissible for the f-SWL system. Then, there exists a fixed-length SWL code such that for any $\gamma>0$ and sufficiently large n,

$$\frac{1}{n}\log M_n^{(1)} \;\le\; R_1+\gamma, \tag{33}$$
$$\frac{1}{n}\log M_n^{(2)} \;\le\; R_2+\gamma, \tag{34}$$
$$\lim_{n\to\infty} e_n \;=\; 0. \tag{35}$$

Substituting (33)-(34) into (30)-(32), we have

$$\begin{aligned}
e_n &\ge P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}|\boldsymbol{y})}\ge R_1+2\gamma\right\}-\exp(-n\gamma),\\
e_n &\ge P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y}|\boldsymbol{x})}\ge R_2+2\gamma\right\}-\exp(-n\gamma),\\
e_n &\ge P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\ge R_1+R_2+3\gamma\right\}-\exp(-n\gamma).
\end{aligned}$$

Then, according to (35),

$$\begin{aligned}
\lim_{n\to\infty} P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x}|\boldsymbol{y})}\ge R_1+2\gamma\right\} &= 0,\\
\lim_{n\to\infty} P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{y}|\boldsymbol{x})}\ge R_2+2\gamma\right\} &= 0,\\
\lim_{n\to\infty} P_n\left\{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^n\times\mathcal{Y}^n:\frac{1}{n}\log\frac{1}{P_n(\boldsymbol{x},\boldsymbol{y})}\ge R_1+R_2+3\gamma\right\} &= 0.
\end{aligned}$$

From the definitions of the sup-entropy rates, we must have

$$\begin{aligned}
R_1+2\gamma &\ge \overline{H}(\boldsymbol{X}|\boldsymbol{Y}),\\
R_2+2\gamma &\ge \overline{H}(\boldsymbol{Y}|\boldsymbol{X}),\\
R_1+R_2+3\gamma &\ge \overline{H}(\boldsymbol{X},\boldsymbol{Y}).
\end{aligned}$$

Since $\gamma>0$ is arbitrary, we can conclude

$$\begin{aligned}
R_1 &\ge \overline{H}(\boldsymbol{X}|\boldsymbol{Y}),\\
R_2 &\ge \overline{H}(\boldsymbol{Y}|\boldsymbol{X}),\\
R_1+R_2 &\ge \overline{H}(\boldsymbol{X},\boldsymbol{Y}).
\end{aligned}$$

This completes the proof of Theorem 2. □

Proof of Lemma 2:

It is easy to see that for any $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ and $i = 1, 2$

$$\frac{1}{n} \log \frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} \leq \frac{1}{n} \log \frac{1}{\min(\alpha, 1-\alpha) P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} = \frac{1}{n} \log \frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} + \frac{c_0}{n}. \tag{36}$$

On the other hand, for $i = 1, 2$ we have

$$\begin{aligned}
& P_n^{(i)} \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \frac{1}{n} \log \frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - \frac{1}{n} \log \frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} \leq -\gamma \right\} \\
= & \sum_{\substack{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \\ P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y}) \leq P_n(\boldsymbol{x}, \boldsymbol{y}) \exp(-n\gamma)}} P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y}) \\
\leq & \sum_{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n :} P_n(\boldsymbol{x}, \boldsymbol{y}) \exp(-n\gamma) \\
\leq & \exp(-n\gamma).
\end{aligned} \tag{37}$$

Then, by combining (36) and (37), we obtain

$$\begin{aligned}
& P_n^{(i)} \Big\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \\
& \quad \frac{1}{n} \log \frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} - \gamma \leq \frac{1}{n} \log \frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} \leq \frac{1}{n} \log \frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} + \frac{c_0}{n} \Big\} \\
= & P_n^{(i)} \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n : \frac{1}{n} \log \frac{1}{P_n(\boldsymbol{x}, \boldsymbol{y})} - \frac{1}{n} \log \frac{1}{P_n^{(i)}(\boldsymbol{x}, \boldsymbol{y})} \geq -\gamma \right\} \\
\geq & 1 - \exp(-n\gamma).
\end{aligned}$$

This completes the proof of Lemma 2. □

Acknowledgment

We would like to thank Prof. Te Sun Han of the University of Electro-Communications and Prof. Fumio Kanaya of Shonan Institute of Technology for their valuable comments and helpful discussions.

References

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Inform. Theory*, vol.IT-19, pp.471-480, July, 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, March, 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*, Springer, 2003.

[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, 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," *IEEE Trans. on Inform. Theory*, vol.46, pp.1217-1226, July 2000.

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

[11] I. Csiszár, "Linear codes for sources and source networks: error exponents, universal coding," *IEEE Trans. on Inform. Theory*, vol.IT-28, pp.585-592, July 1982.

[12] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.

Biographies of authors

Akisato Kimura (M'00) received the B.E. and M.E. degrees in electrical and electronic engineering from Tokyo Institure of Technology, Tokyo, Japan, in 1998 and 2000, respectively. He is now with NTT Communication Science Laboratories, NTT Corporation, Kanagawa, Japan. He has been working on multimedia information recognition and multimedia database. His research interests include multiterminal information theory and pattern recognition.

Tomohiko Uyematsu (M'95) received the B. E., M. E. and D. E. degrees from Tokyo Institute of Technology in 1982, 1984 and 1988, respectively. From 1984 to 1992, He was with the Department of Electrical and Electronic Engineering of Tokyo Institute of Technology, first as research associate, next as lecturer, and lastly as associate professor. From 1992 to 1997, he was with School of Information Science of Japan Advanced Institute of Science and Technology as associate professor. Since 1997, he returned to Tokyo Institute of Technology as associate professor, and currently he is with the Department of Communications and Integrated Systems as professor. In 1992 and 1996, he was a visiting researcher at the Centre National de la Recherche Scientifique, France and Delft University of Technology, Netherlands, respectively. He received Shinohara Memorial Young Engineer Award in 1989, and the Best Paper Award in 1993, 1996, and 2002 both from IEICE. His current research interests are in the areas of information theory, especially Shannon theory and multi-terminal information theory.