The 32nd Symposium on Information Theory
and its Applications (SITA2009)
Yuda-Onsen, Yamaguchi, Japan, Dec. 1–4, 2009

A Note on the Second Order Separate Source Coding Theorem for Sources with Side Information

Ryo NOMURA * Toshiyasu MATSUSHIMA †

Abstract— The source coding theorem reveals the minimum achievable code length under the condition that the error probability is smaller than or equal to some small constant. For the single source coding problem, the source coding theorem was shown for general sources. Furthermore, there is a study to evaluate the achievable code length more precisely for the restricted class of sources by using the asymptotic normality. In this study, we consider the problem that there exists a side information. This setting is one kind of correlated sources coding problem and show the coding theorem more precisely than the previous result by using the asymptotic normality.

Keywords— Side Information, Asymptotic Normality, Correlated Sources

1 Introduction

The source coding theorem is one of the most basic result in communication systems. It reveals the minimum achievable code length under the condition that the error probability is smaller than or equal to ϵ $(0 \leq \epsilon < 1)$[1][2]. In single-user communication systems, Han et al. and Steinberg et al. shows the source coding theorem for general sources[3][4]. The class of general sources is quite large and their result can be applied for various sources. On the other hand, there are several researches that shows the minimum achievable code length more precisely for the restricted class of sources[5, 6, 7]. These results are based on the asymptotic normality of self-information.

In multi-user communication system there are various types of source coding problems[1][8, 9, 10, 11, 12, 13]. The correlated source coding problem is a typical source coding problem in multi-user communication system. In the correlated source coding problem, there exists several problem settings according to the type of encoder and the decoder. The Slepian-Wolf type problem, Wyner type problem and the source coding problem with side information are included in the class of the correlated source coding problem.

Miyake et al. showed the source coding theorem for Slepian-Wolf type source coding problem and Wyner type source coding problem under the condition that the error probability goes to 0 asymptotically [8]. Han showed the coding theorem for Slepian-Wolf type problem in the case that we allow the small error probability. Please note that their results are very important, since they are valid for general correlated sources. However there was no result to show the coding theorem by using the asymptotic normality for correlated sources.

In this study, we consider the case that the self-information of correlated sources has an asymptotic normality and show the coding theorem for the sources with side information more precisely than the previous result. The analysis is based upon the asymptotic normality.

2 Preliminaries

2.1 Correlated Sources

Let $\mathcal{X}_1^n$ and $\mathcal{X}_2^n$ be alphabets of correlated sources, where $n \in \mathcal{N} \equiv \{1, 2, \cdots\}$. Let

$$(\mathbf{X}_1, \mathbf{X}_2) = \{(X_1^n, X_2^n)\}_{n=1}^{\infty},$$

denote a general correlated sources where

$$(X_1^n, X_2^n) = (X_{11}, X_{21}), (X_{12}, X_{22}), \cdots, (X_n, X_{2n}),$$

be random variables emitted from the source and $\mathbf{x_k} = x_{k1}, x_{k2}, \cdots, x_{kn}$ be a realization of random variable X_k^n. The probability distribution of $(\mathbf{x_1}, \mathbf{x_2})$ is denoted by $P_{X_1^n X_2^n}(\mathbf{x_1}, \mathbf{x_2})$. Please note that each of sources $\mathbf{X}_k$ is general correlated sources. If we assume that correlated sources is stationary memoryless correlated sources then it holds that

$$P_{X_1^n X_2^n}(\mathbf{x_1}, \mathbf{x_2}) = \prod_{i=1}^{n} P_{X_1 X_2}(x_{1i}, x_{2i}).$$

2.2 Problem Settings

In this study, we consider the case that we try to communicate $\mathbf{x}_1$ by using the side information $\mathbf{x}_2$. The fixed-length codes for the source with side information are characterized by a encoder $\phi_n^{(1)}$ and a decoder ψ_n. The encoder is a mapping such as $\phi_n^{(1)} : \mathcal{X}_1^n \to \mathcal{M}_n$, where

$$\mathcal{M}_n = \{1, 2, \cdots, M_n\},$$

denote the codes. The decoder is a mapping defined as $\psi_n : \mathcal{M}_n \times \mathcal{X}_2^n \to \mathcal{X}_1^n$. Please note that the encoder $\phi_n^{(1)}$ does not know the sequence $\mathbf{x}_2$. This setting is called separate coding.

The performance of fixed-length code is evaluated by the error probability and the code length. The code length is given by $\log M_n$ [1].The error probability is given by

$$\epsilon_n = \Pr\{X_1^n \neq \psi_n(\phi_n^{(1)}(X_1^n), X_2^n)\}.$$

We call the a pair of the encoder $\phi_n^{(1)}$ and the decoder ψ_n with the error probability ϵ_n an (n, M_n, ϵ_n)

* School of Science and Engineering, Aoyama Gakuin University. 5–10–1, Fuchinobe, Sagamihara-shi, Kanagawa 229–8558, Japan.
† School of Fundamental Science and Engineering, Waseda University. 3–4–1, Ohkubo, Shinjuku-ku, Tokyo 169–8555, Japan.

[1] The base of logarithm is taken by e.

code. Then we are interested that how the code length can be short under the condition that the error probability is smaller than or equal to ϵ $(0 \le \epsilon < 1)$.

Definition 2.1 *The rate R is called an ϵ-achievable rate if there exists an (n, M_n, ϵ_n) code satisfying*

$$\limsup_{n\to\infty} \frac{1}{n}\log M_n \le R,$$

$$\limsup_{n\to\infty} \epsilon_n \le \epsilon.$$

We consider the infimum of ϵ-achievable rate as follows.

Definition 2.2

$$R(\epsilon|\mathbf{X}_1) = \inf\{R_1 | R_1 \text{ is } \epsilon\text{-achievable rate }\}.$$

The above quantities was not shown directly. However, for general correlated sources, Miyake et al. showed the achievable rate region for Slepian-Wolf type problem and Wyner type problem[8] and Han showed the ϵ-achievable rate region for Slepian-Wolf type problem[2]. Since settings in their results includes the setting defined in this study, we can derive $R(\epsilon|\mathbf{X}_1)$ immediately from their results.

2.3 Generalization of Achievability

Second order source coding theorems give us the precise achievable code length. Kontyiannis showed the second order fixed-to-variable source coding theorem for stationary ergodic sources[5]. Hayashi showed the second order fixed-length source coding theorem for general sources[6]. Furthermore for an i.i.d. source his results can be described by using the asymptotic normality.

Their results implies that for restricted source class, such as an stationary memoryless source, we can obtain the achievable code length more precisely.

In this study, we shall show the second order coding theorem for the source with side information.

The infimum of ϵ-achievable rate denotes the shortest code length under the condition that error probability is smaller than or equal to ϵ. In this subsection we define the achievability to evaluate the code length more precisely. In this study, we consider the following achievability.

Definition 2.3 *A sequence $\{\eta_n\}_{n=1}^{\infty}$ is called an ϵ-achievable sequence if there exists an (n, M_n, ϵ_n) code satisfying*

$$\limsup_{n\to\infty}\left(\frac{1}{\sqrt{n}}\log M_n - \eta_n\right) \le 0,$$

$$\limsup_{n\to\infty} \epsilon_n \le \epsilon.$$

Then we are interested in the condition that $\{\eta_n\}_{n=1}^{\infty}$ is the ϵ-achievable sequence.

Please note that the difference between the previous achievability condition and our achievability condition is the difference of conditions for the code length $\log M_n$. Actually if we use the ϵ-achievable rate, then $o(n)$ term in the condition is neglected. Instead, if we use our definition, we can evaluate the condition of ϵ-achievability more precisely.

Remark 2.1 *Assuming that $\eta_n = \sqrt{n}R$ for each $n = 1, 2, \cdots$ and we divide both sides by $\sqrt{n}$, then the condition in our definition coincides with the condition in Def. 2.1. So our achievability is a generalization of the previous.* □

3 Necessary and Sufficient Condition for ϵ-Achievable Sequence

The infimum of ϵ-achievable rate is considered as the necessary and sufficient condition for ϵ-achievable rate. In this section we show the necessary and sufficient condition that $\{\eta_n\}_{n=1}^{\infty}$ is the ϵ-achievable sequence.

At first we show two lemmas that have important roles in our result.

Lemma 3.1 *Let M_n be an arbitrarily given positive integer and $\{a_n\}_{n=1}^{\infty}$ be a sequence of an arbitrary number satisfying $a_i > 0$ $(\forall i = 1, 2, \cdots)$. Then, for all $n = 1, 2, \cdots$ there exists an (n, M_n, ϵ_n) code that satisfies*

$$\epsilon_n \le \Pr\left\{a_n P_{X_1^n|X_2^n}(X_1^n|X_2^n) \le \frac{1}{M_n}\right\} + a_n. \quad (1)$$

(Proof) At first, we shall define the encoder and the decoder. We use the random coding technique.

Encoder For each $x_1^n \in \mathcal{X}_1^n$, we generate $i \in \mathcal{M}_n$ randomly subject to the uniform distribution and define $\phi_n^{(1)}(\mathbf{x}_1) = i$.

Decoder After receiving i, the decoder $\psi : \mathcal{M}_n \times \mathcal{X}_2^n \to \mathcal{X}_1^n$ decodes $\mathbf{x}_1$ if there exists a unique $\mathbf{x}_1$ such that $\phi_n^{(1)}(\mathbf{x}_1) = i$ and $(\mathbf{x}_1, \mathbf{x}_2) \in B_n$ where

$$B_n \stackrel{\text{def}}{=} \left\{(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X}_1^n \times \mathcal{X}_2^n \middle| a_n P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) > \frac{1}{M_n}\right\}.$$

If there exists no such i or more than one, the error is occurred.

We shall evaluate the above encoder and decoder. Let the event E_n as

$$E_n \stackrel{\text{def}}{=} \left\{(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X}_1^n \times \mathcal{X}_2^n \,|\exists \mathbf{x}_1' \ne \mathbf{x}_1, \phi_n^{(1)}(\mathbf{x}_1') = \phi_n^{(1)}(\mathbf{x}_1), (\mathbf{x}_1', \mathbf{x}_2) \in B_n\right\}.$$

Then the error probability is given by

$$\begin{aligned}\epsilon_n &= \Pr\{X_1^n X_2^n \in E_n \cup (X_1^n, X_2^n) \notin B_n\} \\ &\le \Pr\{X_1^n X_2^n \in E_n\} + \Pr\{(X_1^n, X_2^n) \notin B_n\}. \quad (2)\end{aligned}$$

We shall evaluate the first term of the right hand side

(RHS) of (2) as follows.

$$\begin{aligned}
&\Pr\{X_1^n X_2^n \in E_n\} \\
&\quad = \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\qquad \cdot \sum_{\mathbf{x}_1'\neq\mathbf{x}_1,(\mathbf{x}_1',\mathbf{x}_2)\in B_n} \Pr\left\{\phi_n^{(1)}(\mathbf{x}_1') = \phi_n^{(1)}(\mathbf{x}_1)\right\} \\
&\quad = \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\qquad \cdot \sum_{\mathbf{x}_1'\neq\mathbf{x}_1,(\mathbf{x}_1',\mathbf{x}_2)\in B_n} \frac{1}{M_n} \\
&\quad \le \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \sum_{\mathbf{x}_1',(\mathbf{x}_1',\mathbf{x}_2)\in B_n} \frac{1}{M_n},
\end{aligned}$$

where the second equality holds since we use the random coding. Furthermore set $S_n(\mathbf{x}_2)$ as follows

$$S_n(\mathbf{x}_2) \quad \stackrel{\mathrm{def}}{=} \quad \{\mathbf{x}_1' \in \mathcal{X}_1^n | (\mathbf{x}_1',\mathbf{x}_2) \in B_n\}.$$

Then we have

$$\begin{aligned}
&\Pr\{X_1^n X_2^n \in E_n\} \\
&\quad \le \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \sum_{\mathbf{x}_1',(\mathbf{x}_1',\mathbf{x}_2)\in B_n} \frac{1}{M_n} \\
&\quad = \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2)\, |S_n(\mathbf{x}_2)|\, \frac{1}{M_n}. \quad (3)
\end{aligned}$$

From the definition of B_n, for $(\mathbf{x}_1',\mathbf{x}_2) \in B_n$, we have

$$P_{X_1^n|X_2^n}(\mathbf{x}_1'|\mathbf{x}_2) > \frac{1}{a_n M_n}.$$

Thus we have

$$1 \ge \sum_{\mathbf{x}_1'\in S_n(\mathbf{x}_2)} P_{X_1^n|X_2^n}(\mathbf{x}_1'|\mathbf{x}_2) > \frac{|S_n(\mathbf{x}_2)|}{a_n M_n}.$$

Hence we obtain

$$|S_n(\mathbf{x}_2)| < a_n M_n.$$

Substituting the above inequality into (3) we have

$$\begin{aligned}
&\Pr\{X_1^n X_2^n \in E_n\} \\
&\quad \le \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) a_n \le a_n.
\end{aligned}$$

Substituting the above inequality into (2), we obtain

$$\begin{aligned}
\epsilon_n &\le a_n + \Pr\{(X_1^n, X_2^n) \notin B_n\} \\
&= \Pr\{a_n P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) \le \frac{1}{M_n}\},
\end{aligned}$$

where the last equality is derived from the definition of B_n. Therefore we deduce the lemma. □

Lemma 3.2 *For any (n, M_n, ϵ_n) code, it holds that*

$$\epsilon_n \quad \ge \quad \Pr\left\{P_{X_1^n|X_2^n}(X_1^n|X_2^n) \le \frac{a_n}{M_n}\right\} - a_n,$$

for all $n = 1, 2, \cdots$, where $\{a_n\}_{n=1}^{\infty}$ is a sequence of an arbitrary number satisfying $a_i > 0$ ($\forall i = 1, 2, \cdots$).

(Proof) We shall define the following sets.

$$\begin{aligned}
&C_n \\
&\quad \stackrel{\mathrm{def}}{=} \left\{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n \,\middle|\, P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) \le \frac{a_n}{M_n}\right\},
\end{aligned}$$

$$\begin{aligned}
&D_n \\
&\quad \stackrel{\mathrm{def}}{=} \left\{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{X}_1^n\times\mathcal{X}_2^n \,\middle|\, \psi_n(\phi_n^{(1)}(\mathbf{x}_1),\mathbf{x}_2) = \mathbf{x}_1\right\},
\end{aligned}$$

and

$$C_n(\mathbf{x}_2) \stackrel{\mathrm{def}}{=} \{\mathbf{x}_1 \in \mathcal{X}_1^n | (\mathbf{x}_1,\mathbf{x}_2) \in C_n\},$$

for each $\mathbf{x}_2$. Then we have

$$\begin{aligned}
&\Pr\left\{P_{X_1^n|X_2^n}(X_1^n|X_2^n) \le \frac{a_n}{M_n}\right\} \\
&\quad = \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in C_n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\quad = \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in C_n\cap D_n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\qquad + \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in C_n\cap D_n^c} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\quad \le \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in C_n\cap D_n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\qquad + \sum_{(\mathbf{x}_1,\mathbf{x}_2)\notin D_n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) \\
&\quad = \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in C_n\cap D_n} P_{X_1^n X_2^n}(\mathbf{x}_1,\mathbf{x}_2) + \epsilon_n \\
&\quad \le \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in C_n} P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) P_{X_2^n}(\mathbf{x}_2) + \epsilon_n \\
&\quad = \sum_{\mathbf{x}_2\in\mathcal{X}_2^n} \sum_{\mathbf{x}_1\in C_n(\mathbf{x}_2)} P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) P_{X_2^n}(\mathbf{x}_2) + \epsilon_n \\
&\quad = \sum_{\mathbf{x}_2\in\mathcal{X}_2^n} P_{X_2^n}(\mathbf{x}_2) \sum_{\mathbf{x}_1\in C_n(\mathbf{x}_2)} P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) + \epsilon_n.
\end{aligned}$$

On the other hand for $\forall(\mathbf{x}_1,\mathbf{x}_2) \in C_n$, it holds that

$$P_{X_1^n|X_2^n}(\mathbf{x}_1|\mathbf{x}_2) \le \frac{a_n}{M_n}.$$

Thus we have

$$\begin{aligned}
&\Pr\left\{P_{X_1^n|X_2^n}(X_1^n|X_2^n) \le \frac{a_n}{M_n}\right\} \\
&\quad \le \sum_{\mathbf{x}_2\in\mathcal{X}_2^n} P_{X_2^n}(\mathbf{x}_2) \sum_{\mathbf{x}_1\in C_n(\mathbf{x}_2)} \frac{a_n}{M_n} + \epsilon_n \\
&\quad \le \sum_{\mathbf{x}_2\in\mathcal{X}_2^n} P_{X_2^n}(\mathbf{x}_2)\, |C_n(\mathbf{x}_2)|\, \frac{a_n}{M_n} + \epsilon_n.
\end{aligned}$$

Here, noting that $|C_n(\mathbf{x}_2)| \leq M_n$ we obtain

$$\Pr\left\{P_{X_1^n|X_2^n}(X_1^n|X_2^n) \leq \frac{a_n}{M_n}\right\} \leq \sum_{\mathbf{x}_2 \in \mathcal{X}_2^n} P_{X_2^n}(\mathbf{x}_2)a_n + \epsilon_n = a_n + \epsilon_n.$$

Therefore we deduce the lemma. $\square$

Please note that these lemmas are valid for general correlated sources.

We assume that the following condition holds for correlated sources.

Assumption 3.1 *The conditional self-information has an asymptotic normality, that is*

$$\lim_{n\to\infty} \Pr\left\{\frac{-\log P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} \leq U\right\} = \int_{-\infty}^{U} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz,$$

holds where $H(X_1^n|X_2^n) = E[-\log \frac{1}{P(X_1^n|X_2^n)}]$ *denotes the conditional entropy of the source and* $\sigma(\mathbf{X}_1|\mathbf{X}_2)^2$ *denotes the variance of the random variable* $-\log P_{X_1^n|X_2^n}(X_1^n|X_2^n)$ *that is,*

$$\sigma(\mathbf{X}_1|\mathbf{X}_2)^2 \stackrel{\mathrm{def}}{=} \lim_{n\to\infty} \frac{1}{n} Var\left(\log \frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)}\right).$$

We also assume that $\sigma(\mathbf{X}_1|\mathbf{X}_2)^2$ exists and

$$0 < \sigma(\mathbf{X}_1|\mathbf{X}_2)^2 < \infty,$$

holds. This holds for the case that $|\mathcal{X}_1| < \infty$ and $|\mathcal{X}_2| < \infty$ holds.

Please note that the asymptotic normality of conditional self-information holds for the stationary memoryless correlated sources.

By using the above lemmas and the asymptotic normality, we obtain the necessary and sufficient condition for the ϵ-achievable sequence. The following theorem shows the condition in the case that $0 < \epsilon < 1$ holds.

Theorem 3.1 *Under Assumption 3.1, given* $0 < \epsilon < 1$*, the necessary and sufficient condition of* ϵ*-achievable sequence for sources with side information is as follows*

$$\liminf_{n\to\infty}\left(\eta_n - \frac{H(X_1^n|X_2^n)}{\sqrt{n}}\right) \geq T\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}, \qquad (4)$$

where T *satisfies*

$$\epsilon = \int_{T}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy.$$

(Proof) The proof consists of two parts. The first part is that any $\{\eta_n\}_{n=1}^{\infty}$ satisfying (4) is an ϵ-achievable sequence and the second part is that if $\{\eta_n\}_{n=1}^{\infty}$ is the ϵ-achievable sequence, then (4) holds.

At first we shall show the first part, that is, if (4) holds then $\{\eta_n\}_{n=1}^{\infty}$ is a ϵ-achievable sequence.

From Lemma 3.1, for all $n = 1, 2, \cdots$ there exists an (n, M_n, ϵ_n) code satisfying

$$\epsilon_n \leq \Pr\left\{a_n P_{X_1^n|X_2^n}(X_1^n|X_2^n) \leq \frac{1}{M_n}\right\} + a_n. \quad (5)$$

Then set $a_n = e^{-\frac{\sqrt{n}}{\log n}}$ and substituting it into (5), there exists an (n, M_n, ϵ_n) code satisfying

$$\begin{aligned}\epsilon_n &\leq \Pr\left\{e^{-\frac{\sqrt{n}}{\log n}} P_{X_1^n|X_2^n}(X_1^n|X_2^n) \leq \frac{1}{M_n}\right\} + e^{-\frac{\sqrt{n}}{\log n}} \\ &\leq \Pr\left\{\frac{1}{\sqrt{n}} \log \frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \frac{1}{\sqrt{n}} \log M_n - \frac{1}{\log n}\right\} + e^{-\frac{\sqrt{n}}{\log n}}. \quad (6)\end{aligned}$$

Here, we consider $\{\eta_n\}_{n=1}^{\infty}$ satisfying (4) and define

$$M_n = e^{\sqrt{n}\eta_n}.$$

Then (6) guarantees that, there exists an (n, M_n, ϵ_n) code satisfying

$$\epsilon_n \leq \Pr\left\{\frac{1}{\sqrt{n}} \log \frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \eta_n - \frac{1}{\log n}\right\} + e^{-\frac{\sqrt{n}}{\log n}}. \quad (7)$$

On the other hand from (4) for any small $\nu > 0$ it holds that

$$\eta_n \geq \frac{H(X_1^n|X_2^n)}{\sqrt{n}} + T\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2} - \nu,$$

for sufficiently large n. Thus the first term of the right hand side(RHS) of (7) is evaluated as follows.

$$\begin{aligned}&\Pr\left\{\frac{1}{\sqrt{n}} \log \frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \eta_n - \frac{1}{\log n}\right\} \\ &\leq \Pr\left\{\frac{1}{\sqrt{n}} \log \frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \frac{H(X_1^n|X_2^n)}{\sqrt{n}} + T\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2} - \nu - \frac{1}{\log n}\right\} \\ &\leq \Pr\left\{\frac{-P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} \geq T - \frac{\nu}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} - \frac{1}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}\log n}\right\} \\ &< \Pr\left\{\frac{-P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} \geq T - \frac{2\nu}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}\right\}\end{aligned}$$

for sufficiently large n, since

$$\frac{\nu}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} > \frac{1}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}\log n},$$

holds for sufficiently large n. Here, we denote $\nu' = \frac{\nu}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}$ for simplicity. Then from Assumption 3.1 for $\forall\delta > 0$ there exists N_0 such that for $\forall n > N_0$ it holds that

$$\begin{aligned}
&\Pr\left\{\frac{-P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} \geq T - 2\nu'\right\} \\
&< \int_{T-2\nu'}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy + \delta \\
&= \int_{T}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy \\
&\quad + \int_{T-2\nu'}^{T} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy + \delta. \qquad (8)
\end{aligned}$$

Please note that from the property of continuity of normal distribution, the second term of the RHS of (8) goes to 0 as $\nu' \to 0$. Noting that $\nu' > 0$ is an arbitrary small number, this means that for $\forall\delta > 0$ there exists N_0 such that for $\forall n > N_0$ it holds that

$$\begin{aligned}
&\Pr\left\{\frac{-P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} \geq T - 2\nu\right\} \\
&\leq \int_{T}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy + 2\delta \\
&= \epsilon + 2\delta, \qquad (9)
\end{aligned}$$

where the last equality is due to the definition of T. Hence substituting (9) into (7) we have

$$\limsup_{n\to\infty} \epsilon_n \leq \epsilon.$$

On the other hand from the construction of $\log M_n$

$$\limsup_{n\to\infty}\left(\frac{1}{\sqrt{n}}\log M_n - \eta_n\right) \leq 0,$$

holds obviously. These two inequalities shows that the first part of the theorem holds.

The second part is proved if $\{\eta_n\}_{n=1}^{\infty}$ is an ϵ-achievable sequence, then (4) holds.

We assume that $\{\eta_n\}_{n=1}^{\infty}$ not satisfying (4), is an ϵ-achievable sequence. Then we shall lead a contradiction. Since we assume that $\{\eta_n\}_{n=1}^{\infty}$ is an ϵ-achievable sequence, there exists an (n, M_n, ϵ_n) code satisfying

$$\limsup_{n\to\infty} \epsilon_n \leq \epsilon,$$

and

$$\limsup_{n\to\infty}\left(\frac{1}{\sqrt{n}}\log M_n - \eta_n\right) \leq 0. \qquad (10)$$

(10) implies that for arbitrary small $\gamma > 0$, there exists N_0 such that for $\forall n > N_0$

$$\frac{1}{\sqrt{n}}\log M_n \leq \eta_n + \gamma,$$

holds. Substituting the above inequality into Lemma 3.2 and set $a_n = e^{-\sqrt{n}\gamma}$ we have

$$\begin{aligned}
\epsilon_n &\geq \Pr\left\{P_{X_1^n|X_2^n}(X_1^n|X_2^n) \leq \frac{e^{-\sqrt{n}\gamma}}{M_n}\right\} - e^{-\sqrt{n}\gamma} \\
&\geq \Pr\left\{\frac{1}{\sqrt{n}}\log\frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \frac{1}{\sqrt{n}}\log M_n + \gamma\right\} \\
&\quad - e^{-\sqrt{n}\gamma} \\
&\geq \Pr\left\{\frac{1}{\sqrt{n}}\log\frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \eta_n + \gamma + \gamma\right\} \\
&\quad - e^{-\sqrt{n}\gamma} \\
&= \Pr\left\{\frac{1}{\sqrt{n}}\log\frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \eta_n + 2\gamma\right\} \\
&\quad - e^{-\sqrt{n}\gamma}. \qquad (11)
\end{aligned}$$

Since we assume that $\{\eta_n\}_{n=1}^{\infty}$ does not satisfy (4), there exists a constant $\lambda > 0$ such that

$$\eta_n \leq \frac{H(X_1^n|X_2^n)}{\sqrt{n}} + T\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2} - \lambda, \qquad (12)$$

holds for countably infinite n. Substituting (12) into (11), it holds that

$$\begin{aligned}
\epsilon_n &> \Pr\left\{\frac{1}{\sqrt{n}}\log\frac{1}{P_{X_1^n|X_2^n}(X_1^n|X_2^n)} \geq \frac{H(X_1^n|X_2^n)}{\sqrt{n}}\right. \\
&\quad \left. + T\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2} - \lambda + 2\gamma\right\} - e^{-\sqrt{n}\gamma} \\
&> \Pr\left\{\frac{-\log P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}\right. \\
&\quad \left. \geq T + \frac{2\gamma - \lambda}{\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}\right\} - e^{-\sqrt{n}\gamma} \\
&> \Pr\left\{\frac{-\log P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1^n|X_2^n)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}\right. \\
&\quad \left. \geq T - \frac{\lambda}{2\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}\right\} - e^{-\sqrt{n}\gamma},
\end{aligned}$$

for countably infinite n where the last inequality is derived since $\lambda > 0$ is a constant, $\frac{\lambda}{2} > 2\gamma$ holds for sufficiently small $\gamma > 0$. Here, we denote $\lambda' = \frac{\lambda}{2\sqrt{\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}}$ for short. Then from Assumption 3.1 for $\forall\delta > 0$, we have

$$\begin{aligned}
\epsilon_n &> \Pr\left\{\frac{-\log P_{X_1^n|X_2^n}(X_1^n|X_2^n) - H(X_1|X_2)}{\sqrt{n\sigma(\mathbf{X}_1|\mathbf{X}_2)^2}} \geq\right. \\
&\quad \left. T - \lambda'\right\} - e^{-\sqrt{n}\gamma} \\
&> \int_{T-\lambda'}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy - \delta - e^{-\sqrt{n}\gamma} \\
&= \epsilon + \int_{T-\lambda'}^{T} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy - \delta - e^{-\sqrt{n}\gamma},
\end{aligned}$$

for countably infinite n. Noting that $\lambda' > 0$ is a constant there exists a constant $\alpha > 0$ such that

$$\int_{T-\lambda'}^{T} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right] dy \geq \alpha > 0,$$

holds. Thus we have

$$\epsilon_n > \epsilon + \alpha - \delta - e^{-\sqrt{n}\gamma},$$

for countably infinite n. Thus noting that $\delta > 0$ is an arbitrarily small number and $\alpha > 0$ is a constant, it holds that

$$\liminf_{n\to\infty} \epsilon_n > \epsilon.$$

This means that $\{\eta_n\}_{n=1}^{\infty}$ not satisfying (4), is not an ϵ-achievable sequence. This is a contradiction. Therefore we deduce the second part of the theorem. □

4 Conclusion

In this study, we showed the coding theorem for the sources with side information by using the asymptotic normality. When there does not exist side information, the several researchers showed coding theorems by using the asymptotic normality of self-information[5, 6, 7]. Our result can be considered as a natural extension of these results to the case that we can use the side information.

To show the second order source coding theorems for Slepian-Wolf type problem and Wyner type problem are future work.

Acknowledgment

This work was supported in part by MEXT Grant-in-Aid for Young Scientists (B) No. 21760298.

References

[1] T.M. Cover and J.A.Thomas, "Elements of information theory," Wiley, 1991.

[2] T.S. Han, "Information-Spectrum Methods in Information Theory," Springer, New York, 2002.

[3] T.S. Han and S. Verdú "Approximation Theory of Output Statistics," IEEE Trans. Information Theory, Vol.39, No.3, pp.752-772, 1993.

[4] Y. Steinberg and S. Verdú "Simulation of Random Processes and Rate-Distortion Theory," IEEE Trans. Information Theory, Vol.42, No.1, pp.63-86, 1996.

[5] I. Kontoyiannis, "Second-Order Noiseless Source Coding Theorems," IEEE Trans. Information Theory, vol.43, no.4, pp.1339-1341,1997.

[6] M. Hayashi "Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness," IEEE Trans. Information Theory, Vol.54, No.10, pp.4619-4637, 2008.

[7] R. Nomura, T. Matsushima and S. Hirasawa, "A Note on the ϵ-Overflow probability of Lossless Codes," IEICE Trans. Fundamentals, Vol.E90-A, No.12, pp.2965-2970, 2007.

[8] S. Miyake and F. Kanaya, "Coding Theorems on Correlated General Sources," IEICE Trans. Fundamentals, Vol.E78-A, No9, pp.1063-1070, 1995.

[9] D. Slepian and J.K. Wolf, "Noiseless Coding of Correlated Information Sources," IEEE Trans. Information Theory, Vol.IT-19, No.4, pp.471-480, 1973.

[10] A.D. Wyner "On Source Coding with Side Information at the Decoder," IEEE Trans. Information Theory, Vol.IT-21, No.3, pp.294-300, 1975.

[11] T. Uyematsu and K. Maeda, "Asymptotic Optimality for Universal Data Compression Algorithm with Side Information Based on Incremental Parsing," IEICE Trans. Fundamentals, Vol.J85-A, No.1, pp.95-102, 2002. (in Japanese)

[12] S. Ohta, T. Uyematsu and E. Okamoto, "Coding Theorem of Wyner's Source Network for General Sources," Proceedings of the 19th Symposium on Information Theory and Its Applications(SITA96), pp.109-112, 1996 (in Japanese).

[13] H. Cai, S.R. Kulkarni and S. Verdú, "An Algorithm for Universal Lossless Compression With Side Information," IEEE Trans. Information Theory, Vol.52, No.9, pp.4008-4016 , 2006.