International Symposium on Information Theory and its Applications, ISITA2006
Seoul, Korea, October 29–November 1, 2006

# Multiterminal source coding with complementary delivery

Akisato Kimura[†‡] and Tomohiko Uyematsu[†]

[†] Graduate School of Science and Engineering,
Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro,
Tokyo, 152-8552 Japan.

[‡] NTT Communication Science Laboratories,
NTT Corporation,
3-1 Morinosato Wakamiya, Atsugi,
Kanagawa, 243-0198 Japan.

**Abstract**

A coding problem where messages from two correlated sources are jointly encoded and separately decoded is investigated. Each decoder has access to one of the two messages to enable it to reproduce the other message. The rate-distortion function for lossy coding is clarified. Some related coding problems are also examined.

## 1. Introduction

Coding problems for correlated information sources were originally investigated by Slepian and Wolf [1]. Corresponding rate-distortion coding problems (e.g. [2], [3]) and various coding problems (e.g. [4], [5]) derived from Slepian-Wolf coding have been considered. Including the above studies, the main focus in the 1970's was on coding problems with *separate encoding* (each message is separately encoded) and *joint decoding* (some codewords are sent to a decoder and decoded simultaneously). However, since the 1980's, coding problems have been explored that involve *joint encoding* (messages from several sources are encoded simultaneously) and *separate decoding* (each message is separately decoded). Separate decoding processes have been mainly considered in relation to multiple description (e.g. [6]), while joint encoding processes have been dealt with in some previous papers (e.g. [7, 8, 9, 10]).

The situation we investigate here, which we call *complementary delivery*, involves joint encoding and separate decoding. Figure 1 shows a block diagram of the complementary delivery problem. The encoder observes messages emitted from two correlated sources, and delivers these messages to other locations (i.e. decoders). Each decoder has access to one of two messages, and therefore needs to reproduce the other message. For a lossless configurations, the minimum achievable rate for the coding problem has been clarified by Willems, Wolf and Wyner [11, 8]. Here, we investigate a lossy configuration, and clarify the rate-distortion function and some interesting properties. We also examine an extended version of complementary delivery coding (Fig. 1 below), where only an encoded sequence of a message is available as side information at each decoder. The inner and outer bounds of the rate region for given distortion criteria will be clarified.

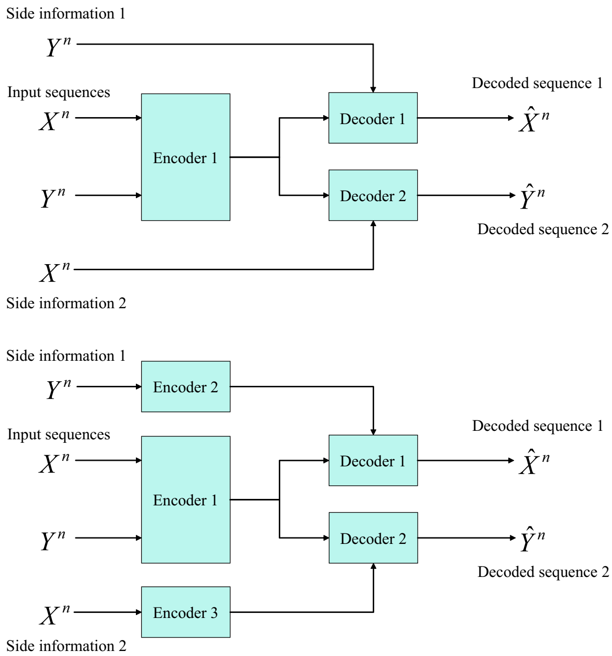


Figure 1: Complementary delivery coding (Above: Full side information is available at decoders, Below: Partial side information is available at decoders)

## 2. Preliminaries

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets, $|\mathcal{X}|$ be the cardinality of $\mathcal{X}$ and $\mathcal{I}_M = \{1, 2, \cdots, M\}$. A member of $\mathcal{X}^n$ is written as $x^n = (x_1, x_2, \cdots, x_n)$, and substrings of $x^n$ are written as $x_i^j = (x_i, x_{i+1}, \cdots, x_j)$ for $i \leq j$. When the dimension is clear from the context, vectors will be denoted by boldface letters. $\mathcal{M}(\mathcal{X})$ denotes a set of all the probability distributions on $\mathcal{X}$. A discrete memoryless source $(\mathcal{X}, P_X)$ is an infinite sequence of independent copies of a random variable $X$ taking values in $\mathcal{X}$ with a generic distribution $P_X \in \mathcal{M}(\mathcal{X})$. We will denote a source $(\mathcal{X}, P_X)$ by referring to its random variable $X$. $\mathcal{M}(\mathcal{X}|P_Y)$ denotes a set of all the conditional probability distributions on $\mathcal{X}$ given a distribution $P_Y \in \mathcal{M}(\mathcal{Y})$, namely each member of $\mathcal{M}(\mathcal{X}|P_Y)$ is characterized by a joint distribution $P_{XY} \in \mathcal{M}(\mathcal{X} \times \mathcal{Y})$ as $P_{XY} = P_{X|Y}P_Y$. For a correlated source $(X, Y, Z)$, $H(X)$, $H(X|Y)$, $I(X;Y)$ and $I(X;Y|Z)$ denote the entropy of $X$, the conditional entropy of $X$ given $Y$, the mutual information of $X$ and $Y$, and the conditional mutual information of $X$ and $Y$ given $Z$, respectively. A similar convention is used for other random variables and vectors. In the following, all bases of exponentials and logarithms are set at $e$ (the base of the natural logarithm). Let $\widehat{\mathcal{X}}$ (resp. $\widehat{\mathcal{Y}}$) stand for a reconstruction alphabet, and let $\Delta_X : \mathcal{X} \times \widehat{\mathcal{X}} \to [0, \overline{\Delta}_X)$ (resp. $\Delta_Y : \mathcal{Y} \times \widehat{\mathcal{Y}} \to [0, \overline{\Delta}_Y)$) be a single-letter distortion function, where $\overline{\Delta}_X, \overline{\Delta}_X < \infty$. The vector distortion function is defined in the usual way, i.e. $\Delta_X^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = 1/n \sum_{k=1}^n \Delta_X(x_k, \widehat{x}_k)$.

## 3. Problem formulation

Here, we provide a problem formulation for complementary delivery.

**Definition 1.** (Lossy FCD (Fully-informed Complementary Delivery) code)
*A set $(\varphi^n, \widehat{\varphi}^n_{(1)}, \widehat{\varphi}^n_{(2)})$ of an encoder and decoders is a lossy FCD code $(n, M_n, \rho_n^{(X)}, \rho_n^{(Y)})$ for the source $(X, Y)$ if and only if*

$$\begin{aligned}
&\varphi^n : \mathcal{X}^n \times \mathcal{Y}^n \to \mathcal{I}_{M_n} \\
&\widehat{\varphi}^n_{(1)} : \mathcal{I}_{M_n} \times \mathcal{Y}^n \to \widehat{\mathcal{X}}^n, \quad \widehat{\varphi}^n_{(2)} : \mathcal{I}_{M_n} \times \mathcal{X}^n \to \widehat{\mathcal{Y}}^n, \\
&\rho_n^{(X)} = E\left[\Delta_X^n(X^n, \widehat{\varphi}^n_{(1)}(\varphi^n(X^n, Y^n), Y^n))\right], \\
&\rho_n^{(Y)} = E\left[\Delta_Y^n(Y^n, \widehat{\varphi}^n_{(2)}(\varphi^n(X^n, Y^n), X^n))\right].
\end{aligned}$$

**Definition 2.** (Lossy FCD-achievable rate)
*$R$ is a lossy FCD-achievable rate of the source $(X, Y)$ for a given distortion pair $(D_1, D_2)$ if and only if there exists a sequence of lossy FCD codes $\{(n, M_n, \rho_n^{(X)}, \rho_n^{(Y)})\}_{n=1}^\infty$ for the source $(X, Y)$ such that*

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log M_n \le R, \\
&\limsup_{n\to\infty} \rho_n^{(X)} \le D_1, \quad \limsup_{n\to\infty} \rho_n^{(Y)} \le D_2.
\end{aligned}$$

**Definition 3.** (Inf lossy FCD-achievable rate)

$$\begin{aligned}
R_F(X, Y|D_1, D_2) = \inf\bigl\{&R : R \textit{ is a lossy} \\
&\textit{FCD-achievable rate of } (X, Y) \textit{ for } (D_1, D_2)\bigr\}.
\end{aligned}$$

## 4. Main results

**Theorem 1.** (Coding theorem of lossy FCD code)

$$\begin{aligned}
&R_F(X, Y|D_1, D_2) \\
&\quad = \min\bigl[\max\{I(X;U|Y), I(Y;U|X)\}\bigr],
\end{aligned}$$

*where the auxiliary random variable $U$ takes a value over an alphabet $\mathcal{U}$ satisfying $|\mathcal{U}| \le |\mathcal{X} \times \mathcal{Y}| + 2$, and the minimum is taken over all $P_{U|XY} \in \mathcal{M}(\mathcal{U}|P_{XY})$ such that there exist functions $\phi_{(1)} : \mathcal{U} \times \mathcal{Y} \to \widehat{\mathcal{X}}$ and $\phi_{(2)} : \mathcal{U} \times \mathcal{X} \to \widehat{\mathcal{Y}}$ that satisfy*

$$\begin{aligned}
D_1 &\ge E\left[\Delta_X(X, \phi_{(1)}(U, Y))\right], \\
D_2 &\ge E\left[\Delta_Y(Y, \phi_{(2)}(U, X))\right].
\end{aligned}$$

From Theorem 1, we can immediately obtain the following properties:

**Corollary 1.** (Compatibility with the known results for a lossless configuration)
*If $\Delta_X(x, \widehat{x}) = 0 \Leftrightarrow x = \widehat{x}$ and $\Delta_Y(y, \widehat{y}) = 0 \Leftrightarrow y = \widehat{y}$,*

$$\begin{aligned}
R_F(X, Y) &\stackrel{\text{def.}}{=} R_F(X, Y|D_1 = 0, D_2 = 0) \\
&= \max\{H(X|Y), H(Y|X)\},
\end{aligned}$$

*which coincides with the results reported by Willems, Wolf and Wyner [11, 8].*

**Corollary 2.** (Relationships to the known results for a lossy configuration)

$$\begin{aligned}
&\max\{R_{C1}(X|Y, D_1), R_{C1}(Y|X, D_2)\} \\
&\quad \le R_F(X, Y|D_1, D_2) \\
&\quad \le R_{C1}(X|Y, D_1) + R_{C1}(Y|X, D_2),
\end{aligned}$$

*where $R_{C1}(X|Y, D)$ is the conditional rate-distortion function [12] for the source $X$ given the source $Y$ and the distortion criterion $D$. On the other hand, for a lossless configuration we have*

$$\begin{aligned}
&\max\{H(X|Y), H(Y|X)\} \\
&\quad = R_F(X, Y) \le H(X|Y) + H(Y|X).
\end{aligned}$$

**Corollary 3.** (Compatibility with known results for special cases)

$$\begin{aligned}
R_F(X, Y|D_1 = d_1, D_2) &= R_{C1}(Y|X, D_2), \\
R_F(X, Y|D_1, D_2 = d_2) &= R_{C1}(X|Y, D_1)
\end{aligned}$$

*if $d_1 \ge \overline{\Delta}_X$ and $d_2 \ge \overline{\Delta}_Y$.*

The most notable property is Corollary 2, which means that there are some rate losses when sending common codewords to two decoders only for a lossy configuration. This property results from the auxiliary random variable $U$, which characterizes the rate-distortion function $R_F(X, Y\ |D_1, D_2)$, although messages used as side information are available at both the encoder and decoder.

In Theorem 1, we have considered only two correlated sources. However, these propositions can be easily extended to any finite number of correlated sources.

Let $\boldsymbol{X} = \{X^{(1)}, X^{(2)}, \cdots, X^{(N)}\}$ be a set of $N$ discrete memoryless sources, each of which takes values in a finite set $\mathcal{X}^{(i)}$ $(i \in \mathcal{I}_N)$. For some $\mathcal{S} \subseteq \mathcal{I}_N$, the corresponding subsets of sources and alphabets will be written as $\boldsymbol{X}^{(\mathcal{S})} = \{X^{(i)} : i \in \mathcal{S}\}$ and $\mathcal{X}^{(\mathcal{S})} = \prod_{i\in\mathcal{S}} \mathcal{X}^{(i)}$, respectively. Let $\boldsymbol{D} = \{D_{j,i}\}_{j\in\mathcal{I}_M, i\in\mathcal{S}_j}$ be a set of distortion criteria. We define the inf lossy FCD-achievable rate $R_F(\boldsymbol{X}|\boldsymbol{D})$ of the source $\boldsymbol{X}$ for given $\boldsymbol{D}$ in the same way as for two correlated sources, where there are $M$ decoders, each of which has access to $\boldsymbol{X}^{(\mathcal{I}_N - \mathcal{S}_j)}$ as side information and therefore has to recover $\boldsymbol{X}^{(\mathcal{S}_j)}$ $(j \in \mathcal{I}_M, \mathcal{S}_j \subseteq \mathcal{I}_N)$ such that for $i \in \mathcal{I}_N$ and $j \in \mathcal{I}_M$ we have

$$D_{j,i} \geq E\left[\Delta_{X^{(i)}}(X^{(i)}, \widehat{X}^{(i)})\right].$$

**Corollary 4.** (Coding theorem of lossy FCD code for any finite number of sources)
*If $\mathcal{S}_{j_1} \not\subseteq \mathcal{S}_{j_2}$ and $\mathcal{S}_{j_1} \not\supseteq \mathcal{S}_{j_2}$ for at least one pair of $(j_1, j_2)$, we have*

$$R_F(\boldsymbol{X}|\boldsymbol{D}) = \min \max_{j\in\mathcal{I}_M} I\left(\boldsymbol{X}^{(\mathcal{S}_j)}; U \middle| \boldsymbol{X}^{(\mathcal{I}_N-\mathcal{S}_j)}\right),$$

*where the auxiliary random variable $U$ takes a value over an alphabet $\mathcal{U}$ satisfying $|\mathcal{U}| \leq |\mathcal{X}^{(\mathcal{I}_N)}| + M$, and the minimum is taken over all $P_{U|\boldsymbol{X}} \in \mathcal{M}(\mathcal{U}|P_{\boldsymbol{X}})$ such that there exist functions $\phi_{(j,i)} : \mathcal{U} \times \mathcal{X}^{(\mathcal{I}_N-\mathcal{S}_j)} \to \widehat{\mathcal{X}}^{(i)}$ that satisfy*

$$D_{j,i} \geq E\left[\Delta_{X^{(i)}}\left(X^{(i)}, \phi_{(j,i)}\left(U, \boldsymbol{X}^{(\mathcal{I}_N-\mathcal{S}_j)}\right)\right)\right].$$

## 5. Extended coding problem

In this section, we examine an extended version of complementary delivery coding, where only an encoded sequence of a message is available as side information at each decoder.

**Definition 4.** (Lossy PCD (Partially-informed Complementary Delivery) code)
*A set $(\varphi^n_{(1)}, \varphi^n_{(2)}, \varphi^n_{(3)}, \widehat{\varphi}^n_{(1)}, \widehat{\varphi}^n_{(2)})$ of encoders and decoders is a lossy PCD code $(n, M^{(1)}_n, M^{(2)}_n, M^{(3)}_n, \rho^{(X)}_n, \rho^{(Y)}_n)$ for the source $(X, Y)$ if and only if*

$$\varphi^n_{(1)} : \mathcal{X}^n \times \mathcal{Y}^n \to \mathcal{I}_{M^{(1)}_n},$$
$$\varphi^n_{(2)} : \mathcal{Y}^n \to \mathcal{I}_{M^{(2)}_n}, \quad \varphi^n_{(3)} : \mathcal{X}^n \to \mathcal{I}_{M^{(3)}_n},$$
$$\widehat{\varphi}^n_{(1)} : \mathcal{I}_{M^{(1)}_n} \times \mathcal{I}_{M^{(2)}_n} \to \widehat{\mathcal{X}}^n,$$
$$\widehat{\varphi}^n_{(2)} : \mathcal{I}_{M^{(1)}_n} \times \mathcal{I}_{M^{(3)}_n} \to \widehat{\mathcal{Y}}^n,$$
$$\rho^{(X)}_n = E\left[\Delta^n_X(X^n, \widehat{\varphi}^n_{(1)}(\varphi^n_{(1)}(X^n, Y^n), \varphi^n_{(2)}(Y^n)))\right],$$
$$\rho^{(Y)}_n = E\left[\Delta^n_Y(Y^n, \widehat{\varphi}^n_{(2)}(\varphi^n_{(1)}(X^n, Y^n), \varphi^n_{(3)}(X^n)))\right].$$

**Definition 5.** (Lossy PCD-achievable rate triad)
*$(R_1, R_2, R_3)$ is a lossy PCD-achievable rate triad of the source $(X, Y)$ for a given distortion pair $(D_1, D_2)$ if and only if there exists a sequence of lossy PCD codes $\left\{(n, M^{(1)}_n, M^{(2)}_n, M^{(3)}_n, \rho^{(X)}_n, \rho^{(Y)}_n)\right\}_{n=1}^{\infty}$ for the source $(X, Y)$ such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log M^{(i)}_n \leq R_i, \ (i = 1, 2, 3)$$
$$\limsup_{n\to\infty} \rho^{(X)}_n \leq D_1, \ \limsup_{n\to\infty} \rho^{(Y)}_n \leq D_2.$$

**Definition 6.** (Lossy PCD-achievable rate region)

$$\mathcal{R}_P(X, Y|D_1, D_2) = \{(R_1, R_2, R_3) : (R_1, R_2, R_3) \text{ is a lossy PCD-achievable rate triad of } (X, Y) \text{ for } (D_1, D_2)\}.$$

**Theorem 2.** (Coding theorem of lossy PCD code)

$$\begin{aligned}
&\mathcal{R}_P(X, Y|D_1, D_2) \subseteq \\
&\quad \{ (R_1, R_2, R_3) : \\
&\qquad R_1 \geq \max\{I(X; U|V), I(Y; U|W)\}, \\
&\qquad R_2 \geq I(Y; V), \ R_3 \geq I(X; W)\} \quad (\textit{converse}) \\
&\mathcal{R}_P(X, Y|D_1, D_2) \supseteq \\
&\quad \{ (R_1, R_2, R_3) : \\
&\qquad R_1 \geq \max\{I(VX; U), I(WY; U)\} \\
&\qquad\qquad - \min\{I(V; U), I(W; U)\}, \\
&\qquad R_2 \geq I(Y; V), \ R_3 \geq I(X; W)\}, \quad (\textit{direct})
\end{aligned}$$

*where the auxiliary random variables $U$, $V$ and $W$ take values over alphabets $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$, respectively, satisfying $|\mathcal{V}| \leq |\mathcal{Y}| + 2$, $|\mathcal{W}| \leq |\mathcal{X}| + 2$, $|\mathcal{U}| \leq |\mathcal{X} \times \mathcal{Y} \times \mathcal{V} \times \mathcal{W}| + 3$, $V \to Y \to X$ and $W \to X \to Y$ form Markov chains, and there exist functions $\phi_{(1)} : \mathcal{U} \times \mathcal{V} \to \widehat{\mathcal{X}}$ and $\phi_{(2)} : \mathcal{U} \times \mathcal{W} \to \widehat{\mathcal{Y}}$ satisfying*

$$D_1 \geq E\left[\Delta_X(X, \phi_{(1)}(U, V))\right],$$
$$D_2 \geq E\left[\Delta_Y(Y, \phi_{(2)}(U, W))\right].$$

From Theorem 2, we immediately obtain the following properties:

**Corollary 5.** (Compatibility with the coding theorem of FCD code)

$$\begin{aligned} & R_P(X,Y|D_1,D_2,R_2=r_2,R_3=r_3) \\ & \overset{\text{def.}}{=} \min\{R:(R,r_2,r_3)\in\mathcal{R}_P(X,Y|D_1,D_2)\} \\ & \geq R_F(X,Y|D_1,D_2) \end{aligned}$$

*for any $r_2, r_3 \geq 0$. The inequality satisfies the equality requirement if $r_2 \geq H(Y)$ and $r_3 \geq H(X)$.*

**Corollary 6.** (Compatibility with known results for special cases)

$$\begin{aligned} \mathcal{R}_P(X,Y|D_1=d_1,D_2) &= \mathcal{R}_{C3}(Y|X,D_2), \\ \mathcal{R}_P(X,Y|D_1,D_2=d_2) &= \mathcal{R}_{C3}(X|Y,D_1), \end{aligned}$$

*if $d_1 \geq \overline{\Delta}_X$ and $d_2 \geq \overline{\Delta}_Y$, where $\mathcal{R}_{C3}(X|Y,D)$ is the minimum achievable rate region when $X$ is encoded and reproduced with encoded sequences from $Y$ to guarantee the distortion level $D$ [7].*

## 6. Proof of theorems

It suffices to prove Theorem 2 because Theorem 1 can be obtained from Corollary 5. Due to space limitations, we would like to show a sketch of the proof. Details will be provided in [14].

### 6.1. Theorem 2: converse part

**Proof.**
Let a sequence $\left\{(\varphi^n_{(1)},\varphi^n_{(2)},\varphi^n_{(3)},\widehat{\varphi}^n_{(1)},\widehat{\varphi}^n_{(2)})\right\}_{n=1}^{\infty}$ of lossy PCD codes be given that satisfy the conditions of Definitions 4 and 5. From Definition 5, for $\delta > 0$ there exists an integer $n_1 = n_1(\delta)$ and then for all $n \geq n_1(\delta)$, we can obtain

$$\frac{1}{n}\log M_n^{(i)} \leq R_i+\delta \quad (i=1,2,3). \tag{1}$$

First, we evaluate Eq. (1) for $i=1$. Let $A_n^{(1)} = \varphi^n_{(1)}(X^n,Y^n)$, $A_n^{(2)} = \varphi^n_{(2)}(Y^n)$ and $A_n^{(3)} = \varphi^n_{(3)}(X^n)$. Then, we obtain

$$n(R_1+\delta) \geq \sum_{k=1}^n I(X_k;A_n^{(1)}|A_n^{(2)}X^{k-1}Y^{k-1}).$$

Let us define the random variables $U_k = A_n^{(1)}$ and $V_k = A_n^{(2)}X^{k-1}Y^{k-1}$. With these definitions, we have the Markov structure $V_k \to Y_k \to X_k$, and therefore we obtain

$$n(R_1+\delta) \geq \sum_{k=1}^n I(X_k;U_k|V_k).$$

In a similar manner, by letting $W_k = A_n^{(3)}X^{k-1}Y^{k-1}$ we obtain $n(R_1+\delta) \geq \sum_{k=1}^n I(Y_k;U_k|W_k)$ and the Markov structure $W_k \to X_k \to Y_k$. Here, let $J$ be a random variable, independent of $X$ and $Y$, and uniformly distributed over the set $\mathcal{I}_n$. We define the random variables $U=(J,U_J)$, $V=(J,V_J)$ and $W=(J,W_J)$. The Markov structures $V\to Y\to X$ and $W\to X\to Y$ still hold, and we have $R_1+\delta \geq I(X;U|V)$ and $R_1+\delta \geq I(Y;U|W)$. Since $\delta>0$ is arbitrary, we obtain $R_1 \geq \max\{I(X;U|V), I(Y;U|W)\}$. Next, we evaluate Eq. (1) for $i=2$.

$$\begin{aligned} n(R_2+\delta) &\geq \sum_{k=1}^n I(Y_k;A_n^{(2)}X^{k-1}Y^{k-1}) \\ &= \sum_{k=1}^n I(Y_k;V_k). \end{aligned}$$

In the same way as with the above discussion, we have $R_2 \geq I(Y;V)$. In a similar manner, we also obtain $R_3 \geq I(X;W)$.

We next show the existence of functions $\phi_{(1)}$ and $\phi_{(2)}$ that satisfy the conditions of Theorem 2. Now, we denote the output of $\widehat{\varphi}^n_{(i)}$ at time $k$ ($k\in\mathcal{I}_n$) by $\widehat{\varphi}_{(i)k}$ ($i=1,2$). From Definition 5, for any $\gamma>0$, there exists an integer $n_2=n_2(\gamma)$, and for all $n\geq n_2(\gamma)$, we have

$$\begin{aligned} & D_1+\gamma \\ & \geq \frac{1}{n}\sum_{k=1}^n E\left[\Delta_X(X_k,\widehat{\varphi}_{(1)k}(\varphi^n_{(1)}(X_k,Y_k),\varphi^n_{(2)}(Y_k)))\right], \\ & D_2+\gamma \\ & \geq \frac{1}{n}\sum_{k=1}^n E\left[\Delta_Y(Y_k,\widehat{\varphi}_{(2)k}(\varphi^n_{(1)}(X_k,Y_k),\varphi^n_{(3)}(X_k)))\right]. \end{aligned}$$

Here, we choose the functions $\phi_{(1)}$ and $\phi_{(2)}$ as follows:

$$\begin{aligned} \phi_{(1)k}(U_k,V_k) &\overset{\text{def.}}{=} \widehat{\varphi}_{(1)k}(A_n^{(1)},A_n^{(2)}), \\ \phi_{(2)k}(U_k,W_k) &\overset{\text{def.}}{=} \widehat{\varphi}_{(2)k}(A_n^{(1)},A_n^{(3)}), \\ \phi_{(1)}(U,V) &\overset{\text{def.}}{=} \phi_{(1)J}(U_J,V_J), \\ \phi_{(2)}(U,W) &\overset{\text{def.}}{=} \phi_{(2)J}(U_J,W_J). \end{aligned}$$

This implies that

$$\begin{aligned} D_1 &\geq E\left[\Delta_X(X,\phi_{(1)}(U,V))\right], \\ D_2 &\geq E\left[\Delta_Y(Y,\phi_{(2)}(U,W))\right]. \end{aligned}$$

It remains to establish that the bounds on $|\mathcal{U}|$, $|\mathcal{V}|$ and $|\mathcal{W}|$ specified in Theorem 2 do not affect the minimization of $R$. To do this, we introduce the support lemma [15, Lemma 3.3.4]. We first reduce the alphabet size of $V$. Here, we have $|\mathcal{Y}|-1$ constraints to preserve the distribution $P_Y$ just defined, and three more constraints for preserving $I(X;U|V)$, $I(Y;V)$ and $E[\Delta_X(X,\phi_{(1)}(U,V))]$. We can reduce the alphabet sizes of $W$ and $U$ similarly.

This completes the proof of the converse part. $\square$

### 6.2. Theorem 2: direct part

**Proof.**
Let a distortion pair $(D_1, D_2)$ be given, and let $U$, $V$ and $W$ satisfy the conditions that define $\mathcal{R}_P(X,Y|D_1,D_2)$. Fix arbitrary $\gamma, \delta > 0$. Let us denote a set of (strong) typical sequences over $\mathcal{X}^n$ as $T_X^n(\delta)$.

*Codeword selection: $\varphi_{(2)}^n$ ($\varphi_{(3)}^n$)*

Randomly generate $M_V$ (resp. $M_W$) independent codewords $\mathcal{A}_V = \{\boldsymbol{v}_i\}_{i=1}^{M_V}$, $\boldsymbol{v}_i \in \mathcal{V}^n$ (resp. $\mathcal{A}_W = \{\boldsymbol{w}_i\}_{i=1}^{M_W}$, $\boldsymbol{w}_i \in \mathcal{W}^n$), each of length $n$, according to $P_V$ (resp. $P_W$).

*Codeword selection: $\varphi_{(1)}^n$*

(1) Randomly generate $M_U$ independent codewords $\mathcal{A}_U = \{\boldsymbol{u}_i\}_{i=1}^{M_U}$, $\boldsymbol{u}_i \in \mathcal{U}^n$, each of length $n$, according to $P_U$.
(2) Partition the codebook $\mathcal{A}_U$ into $N_U$ bins, each containing $L_U = M_U/N_U$ members of $\mathcal{A}_U$. Let $\mathcal{A}_U(j)$ denote the elements $\boldsymbol{u} \in \mathcal{A}_U$ assigned to bin $j$ ($j \in \mathcal{I}_{N_U}$).

*Encoding: $\varphi_{(2)}^n$*

(1) For a given $\boldsymbol{y} \in \mathcal{Y}^n$, the encoder seeks a vector $\boldsymbol{v}_i \in \mathcal{A}_V$ such that $(\boldsymbol{v}_i, \boldsymbol{y}) \in T_{VY}(k_1\delta)$, where $\exists k_1 > 0$. If there is more than one such vector in $\mathcal{A}_V$, the first one is chosen. If there is no such vector in $\mathcal{A}_V$, an error is declared. Denote the selected vector by $\boldsymbol{v}(\boldsymbol{y})$.
(2) The value assigned to the encoder $\varphi_{(2)}^n(\cdot)$ is the index of the selected vector, that is, $\varphi_{(2)}^n(\boldsymbol{y}) = i$, $\boldsymbol{v}(\boldsymbol{y}) = \boldsymbol{v}_i$.

*Encoding: $\varphi_{(3)}^n$*

In a similar manner to $\varphi_{(2)}^n$, the encoder seeks a vector $\boldsymbol{w}(\boldsymbol{x}) = \boldsymbol{w}_i \in \mathcal{A}_W$ such that $(\boldsymbol{w}_i, \boldsymbol{x}) \in T_{WX}(k_2\delta)$, where $\exists k_2 > 0$, and the value assigned to the encoder $\varphi_{(3)}^n(\cdot)$ is the index of the selected vector.

*Encoding: $\varphi_{(1)}^n$*

(1) For a given $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, the encoder seeks a vector $\boldsymbol{u}_i \in \mathcal{A}_U$ such that $(\boldsymbol{u}_i, \boldsymbol{v}(\boldsymbol{y}), \boldsymbol{x}) \in T_{UVX}(k_3\delta)$ and $(\boldsymbol{u}_i, \boldsymbol{w}(\boldsymbol{x}), \boldsymbol{y}) \in T_{UWY}(k_4\delta)$, where $\exists k_3, k_4 > 0$. If there is more than one such vector in $\mathcal{A}_U$, the first one is chosen. If there is no such vector in $\mathcal{A}_U$, an error is declared. The selected vector is denoted by $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y})$.
(2) The value assigned to the encoder $\varphi_{(1)}^n(\cdot)$ is the bin index to which $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y})$ belongs, that is, $\varphi_{(1)}^n(\boldsymbol{x}, \boldsymbol{y}) = j$, $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{A}_U(j)$.

*Decoding: $\widehat{\varphi}_{(1)}^n$*

(1) The decoder has access to the indexes $j_U = \varphi_{(1)}^n(\boldsymbol{x}, \boldsymbol{y})$ and $j_V = \varphi_{(2)}^n(\boldsymbol{y})$.
(2) We can recover the unique vector $\widehat{\boldsymbol{v}} = \boldsymbol{v}(\boldsymbol{y}) = \boldsymbol{v}_{j_V} \in \mathcal{A}_V(j_V)$ from the index $j_V$.
(3) The decoder seeks a unique vector $\boldsymbol{u} \in \mathcal{A}_U(j_U)$ such that $(\boldsymbol{u}, \widehat{\boldsymbol{v}}) \in T_{UV}(k_5\delta)$, where $\exists k_5 > 0$. Denote this vector $\widehat{\boldsymbol{u}}(\widehat{\boldsymbol{v}})$. If there is no or more than one such vector $\boldsymbol{u} \in \mathcal{A}_U(j_U)$, an error is declared.
(4) The reconstruction vector $\widehat{\boldsymbol{x}} = (\widehat{x}_1, \widehat{x}_2, \cdots, \widehat{x}_n)$ is given by

$$\widehat{x}_k \quad = \quad \phi_{(1)}(\widehat{u}_k(\widehat{\boldsymbol{v}}), \widehat{v}_k) \quad (k \in \mathcal{I}_n).$$

*Decoding: $\widehat{\varphi}_{(2)}^n$*

(1) The decoder has access to the indexes $j_U = \varphi_{(1)}^n(\boldsymbol{x}, \boldsymbol{y})$ and $j_W = \varphi_{(3)}^n(\boldsymbol{x})$.
(2) In a similar manner to $\widehat{\varphi}_{(1)}^n$, it seeks a unique vector $\boldsymbol{u} \in \mathcal{A}_U(j_U)$ such that $(\boldsymbol{u}, \widehat{\boldsymbol{w}} = \boldsymbol{w}_{j_W}) \in T_{UW}(k_6\delta)$, where $\exists k_6 > 0$, and the reconstruction vector $\widehat{\boldsymbol{y}}$ is given by

$$\widehat{y}_k \quad = \quad \phi_{(2)}(\widehat{u}_k(\widehat{\boldsymbol{w}}), \widehat{w}_k) \quad (k \in \mathcal{I}_n).$$

*Distortion evaluation: $\widehat{\varphi}_{(1)}^n$, $\widehat{\varphi}_{(2)}^n$*

Noting that $(\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{v}(\boldsymbol{y}), \boldsymbol{x}) \in T_{UVX}(k_3\delta)$, if no error occurs in the encoding/decoding process, we have $(\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{v}(\boldsymbol{y}), \boldsymbol{x}) = (\widehat{\boldsymbol{u}}(\widehat{\boldsymbol{v}}), \widehat{\boldsymbol{v}}, \boldsymbol{x})$, and therefore

$$\Delta_X^n(\boldsymbol{x}, \widehat{\boldsymbol{x}}) \quad \leq \quad D_1 + k_3\delta\overline{\Delta}_X|\mathcal{U} \times \mathcal{V} \times \mathcal{X}|.$$

Since $\delta > 0$ is arbitrarily small for a sufficiently large $n$, if error probabilities in the encoding/decoding processes vanish as $n \to \infty$, we can obtain

$$\limsup_{n \to \infty} E\left[\Delta_X^n(X^n, \widehat{X}^n)\right] \quad \leq \quad D_1.$$

We can obtain the distortion bound for $\widehat{\varphi}_{(2)}^n$ similarly.

*Error evaluation: $\varphi_{(2)}^n$, $\varphi_{(3)}^n$*

If there is no $\boldsymbol{v}_i \in \mathcal{A}_V$ such that $(\boldsymbol{v}_i, \boldsymbol{y}) \in T_{VY}(k_1\delta)$, an encoding error has occurred at $\varphi_{(2)}^n$. The probability of this event vanishes as $n \to \infty$ by setting $M_V$ as

$$M_V \geq \exp\{n(I(Y;V) + m_1\gamma)\}, \ \exists m_1 > 0.$$

We can evaluate the encoding error of $\varphi_{(3)}^n$ similarly.

*Error evaluation: $\varphi_{(1)}^n$*

If there is no $\boldsymbol{u}_i \in \mathcal{A}_U$ such that $(\boldsymbol{u}_i, \boldsymbol{v}(\boldsymbol{y}), \boldsymbol{x}) \in T_{UVX}(k_3\delta)$ and $(\boldsymbol{u}_i, \boldsymbol{w}(\boldsymbol{x}), \boldsymbol{y}) \in T_{UWY}(k_4\delta)$, an encoding error has occurred. This event is denoted as

$$E_2 \quad \stackrel{\text{def.}}{=} \quad E_{21} \cup E_{22},$$

$$E_{21} \stackrel{\text{def.}}{=} \bigcap_{i=1}^{M_U} \{(\boldsymbol{u}_i, \boldsymbol{v}(\boldsymbol{y}), \boldsymbol{x}) \notin T_{UVX}(k_3\delta)\},$$

$$E_{22} \stackrel{\text{def.}}{=} \bigcap_{i=1}^{M_U} \{(\boldsymbol{u}_i, \boldsymbol{w}(\boldsymbol{x}), \boldsymbol{y}) \notin T_{UWY}(k_4\delta)\}.$$

By setting $M_U$ as

$$M_U \geq \exp\{n(I(VX;U) + m_{21}\gamma)\}, \ \exists m_{21} > 0,$$

we have $\lim_{n\to\infty} \Pr\{E_{21}\} = 0$. In a similar manner, we have $\lim_{n\to\infty} \Pr\{E_{22}\} = 0$ by setting $M_U$ as

$$M_U \geq \exp\{n(I(WY;U) + m_{22}\gamma)\}, \ \exists m_{22} > 0.$$

*Error evaluation:* $\widehat{\varphi}^n_{(1)}$, $\widehat{\varphi}^n_{(2)}$

If there is no or more than one $\boldsymbol{u}_i \in \mathcal{A}_U(j_U)$ such that $(\boldsymbol{u}_i, \widehat{\boldsymbol{v}}) \in T_{UV}(k_5\delta)$, a decoding error has occurred at $\widehat{\varphi}^n_{(1)}$. This event is classified into two cases.
(1) The first case: $(\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{v}(\boldsymbol{y})) \notin T_{UV}(k_5\delta)$. However, this error does not occur because $(\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{v}(\boldsymbol{y}), \boldsymbol{x}) \in T_{UVX}(k_3\delta)$.
(2) The second case: If there exists $\boldsymbol{u} \in \mathcal{A}_U(j_U)$, $\boldsymbol{u} \neq \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y})$ such that $(\boldsymbol{u}, \widehat{\boldsymbol{v}}) \in T_{UV}(k_5\delta)$. The probability of this event vanishes as $n \to \infty$ by setting $L_U$ as

$$L_U \leq \exp\{n(I(V;U) - l_{21}\gamma)\}, \ \exists l_{21} > 0.$$

We can evaluate the decoding error of $\widehat{\varphi}^n_{(2)}$ similarly.

*Rate evaluation:* $\varphi^n_{(2)}$, $\varphi^n_{(3)}$

The encoder $\varphi^n_{(2)}$ sends the indexes of the bin using

$$R_2 = \frac{1}{n}\log M_V \quad \geq \quad I(Y;V) + m_1\gamma$$

bits per letter. Since $\gamma > 0$ is arbitrary, we obtain the intended coding rate. We can evaluate the coding rate for $\varphi^n_{(3)}$ similarly.

*Rate evaluation:* $\varphi^n_{(1)}$

The encoder sends the indexes of the bin using

$$\begin{aligned} R_1 &= \frac{1}{n}\log N_U = \frac{1}{n}\log\frac{M_U}{L_U} \\ &\geq \max\{I(VX;U) + m_{21}\gamma, I(WY;U) + m_{22}\gamma\} \\ &\quad - \min\{I(V;U) - l_{21}\gamma, I(W;U) - l_{22}\gamma\} \end{aligned}$$

bits per letter, where $\exists l_{22} > 0$. Since $\gamma > 0$ is arbitrary, we obtain the intended coding rate.

□

## References


[1] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. Inform. Theory, vol.19, no.4, pp.471–480, July 1973.

[2] A.D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Trans. Inform. Theory, vol.22, no.1, pp.1–10, January 1976.

[3] S.Y. Tung, Multiterminal source coding, Ph.D. thesis, School of Electrical Engineering, Cornell University, May 1978.

[4] A.D. Wyner, "On source coding with side information at the decoder," IEEE Trans. Inform. Theory, vol.21, no.3, pp.294–300, May 1975.

[5] A. Sgarro, "Source coding with side information at several decoders," IEEE Trans. Inform. Theory, vol.23, no.2, pp.179–182, March 1977.

[6] A.A. ElGamal and T.M. Cover, "Achievable rates for multiple descriptions," IEEE Trans. Inform. Theory, vol.28, no.6, pp.851–857, November 1982.

[7] R.M. Gray and A.D. Wyner, "Source coding for a simple network," Bell Syst. Tech. J., vol.53, no.9, pp.1681–1721, November 1974.

[8] A.D. Wyner, J.K. Wolf, and F.M.J. Willems, "Communicating via a processing broadcast satellite," IEEE Trans. Inform. Theory, vol.48, no.6, pp.1243–1249, June 2002.

[9] H. Yamamoto, "Source coding theory for a triangular communication system," IEEE Trans. Inform. Theory, vol.42, no.3, pp.848–853, May 1996.

[10] H. Yamamoto, "Source coding theory for cascade and branching communication systems," IEEE Trans. Inform. Theory, vol.27, no.3, pp.299–308, May 1981.

[11] F.M.J. Willems, J.K. Wolf, and A.D. Wyner, "Communicating via a processing broadcast satellite," Proc. the 1989 IEEE/CAM Information Theory Workshop, June 1989.

[12] T. Berger, Rate distortion theory: A mathematical basis for data compression, Prentice Hall, Englewood Cliffs, NJ, 1971.

[13] T. Cover and J. Thomas, Elements of information theory, Wiley, New York, 1991.

[14] A. Kimura and T. Uyematsu, "Multiterminal source coding with complementary delivery," IEICE Trans. Fundamentals, 2006. submitted.

[15] I. Csiszar and J. Körner, Information theory: Coding theorems for discrete memoryless systems, Academic Press, New York, 1981.