

On Cascade System Coding Rates with Respect to Distortion Criteria and Reliability*

E. A. Haroutunian and A. R. Kazarian

Institute for Informatics and Automation Problems
of NAS of RA and YSU

Abstract

A generalization of the multiterminal communication system first studied by Yamamoto is considered. Messages of two correlated sources $\{X\}$ and $\{Y\}$ coded by a common encoder and a separate encoder must be transmitted to two receivers by two common decoders within prescribed distortion levels Δ_x^1, Δ_y^1 and Δ_x^2, Δ_y^2 , respectively.

The region $\mathcal{R}(E_1, E_2, \Delta_x^1, \Delta_y^1, \Delta_x^2, \Delta_y^2)$ of all achievable rates of the best codes ensuring reconstruction of messages of the sources $\{X\}$ and $\{Y\}$ within given distortion levels with error probabilities exponents E_1 and E_2 at the first and second decoders, respectively, called "rate-reliability, distortion" region, is found. A number of important consequent cases are noted.

1 Introduction and Problem Statement

We investigate the problem of two correlated sources common encoding and decoding subject to fidelity criteria at two receivers (see Fig. 1.). Related problems were studied in [1]–[17]. Our problem is a direct generalization of the problem considered by Yamamoto [7].

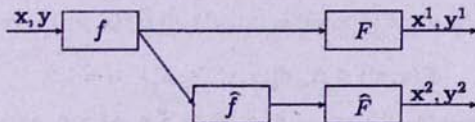


Fig. 1. Cascade communication system.

Let $\{X_i, Y_i\}_{i=1}^{\infty}$ be a sequence of discrete, independent, identically distributed pairs of random variables RV, taking values in finite sets \mathcal{X} and \mathcal{Y} , which are the sets of all messages of the sources $\{X\}$ and $\{Y\}$, respectively. Let generic joint probability distribution (PD) of messages of two sources is

$$P^* = \{P^*(x, y), x \in \mathcal{X}, y \in \mathcal{Y}\}.$$

For memoryless sources the probability $P^{*n}(x, y)$ of a pair of n -sequences of messages $x = (x_1, x_2, \dots, x_n) \in \mathcal{X}^n, y = (y_1, y_2, \dots, y_n) \in \mathcal{Y}^n$ is defined by

*This work was supported by INTAS under Grant 94-469.

$$P^{*n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^n P^{*n}(x_i, y_i).$$

The sets $\mathcal{X}^1, \mathcal{X}^2$ and $\mathcal{Y}^1, \mathcal{Y}^2$ are reconstruction alphabets and in general they are different from \mathcal{X} and \mathcal{Y} , respectively. Let

$$d_x^1: \mathcal{X} \times \mathcal{X}^1 \rightarrow [0, \infty), \quad d_x^2: \mathcal{X} \times \mathcal{X}^2 \rightarrow [0, \infty),$$

$$d_y^1: \mathcal{Y} \times \mathcal{Y}^1 \rightarrow [0, \infty), \quad d_y^2: \mathcal{Y} \times \mathcal{Y}^2 \rightarrow [0, \infty),$$

be corresponding distortion measures. If $\mathcal{X} \equiv \mathcal{X}^1 \equiv \mathcal{X}^2$, $\mathcal{Y} \equiv \mathcal{Y}^1 \equiv \mathcal{Y}^2$, we assume that $d_x^1(x, x) = d_x^2(x, x) = d_y^1(y, y) = d_y^2(y, y) = 0$. The distortion measures for n -sequences are defined by the respective average of per letter distortions

$$d_x^1(\mathbf{x}, \mathbf{x}^1) = n^{-1} \sum_{i=1}^n d_x^1(x_i, x_i^1), \quad d_x^2(\mathbf{x}, \mathbf{x}^2) = n^{-1} \sum_{i=1}^n d_x^2(x_i, x_i^2),$$

$$d_y^1(\mathbf{y}, \mathbf{y}^1) = n^{-1} \sum_{i=1}^n d_y^1(y_i, y_i^1), \quad d_y^2(\mathbf{y}, \mathbf{y}^2) = n^{-1} \sum_{i=1}^n d_y^2(y_i, y_i^2),$$

where $\mathbf{x} \in \mathcal{X}^n$, $\mathbf{y} \in \mathcal{Y}^n$, $\mathbf{x}^1 \in \mathcal{X}^{1n}$, $\mathbf{x}^2 \in \mathcal{X}^{2n}$, $\mathbf{y}^1 \in \mathcal{Y}^{1n}$, $\mathbf{y}^2 \in \mathcal{Y}^{2n}$.

For considered system we name a code and note (f, \hat{f}, F, \hat{F}) the family of four mappings:

$$f: \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, 2, \dots, K(n)\},$$

$$\hat{f}: \{1, 2, \dots, K(n)\} \rightarrow \{1, 2, \dots, L(n)\},$$

$$F: \{1, 2, \dots, K(n)\} \rightarrow \mathcal{X}^{1n} \times \mathcal{Y}^{1n},$$

$$\hat{F}: \{1, 2, \dots, L(n)\} \rightarrow \mathcal{X}^{2n} \times \mathcal{Y}^{2n}.$$

Consider the sets

$$\mathcal{A}_i = \{(\mathbf{x}, \mathbf{y}) : F(f(\mathbf{x}, \mathbf{y})) = (\mathbf{x}^1, \mathbf{y}^1), \hat{F}(\hat{f}(f(\mathbf{x}, \mathbf{y}))) = (\mathbf{x}^2, \mathbf{y}^2),$$

$$d_x^i(\mathbf{x}, \mathbf{x}^i) \leq \Delta_x^i, d_y^i(\mathbf{y}, \mathbf{y}^i) \leq \Delta_y^i\}, \quad i = 1, 2.$$

For given levels of admissible distortions $\Delta_x^1 \geq 0$, $\Delta_x^2 \geq 0$, $\Delta_y^1 \geq 0$, $\Delta_y^2 \geq 0$ error probabilities of the code are:

$$e_i(f, \hat{f}, F, \hat{F}, \Delta_x^i, \Delta_y^i) = 1 - P^{*n}(\mathcal{A}_i), \quad i = 1, 2.$$

For brevity we denote $(E_1, E_2) = E$ and $(\Delta_x^1, \Delta_y^1, \Delta_x^2, \Delta_y^2) = \Delta$. A pair of two nonnegative numbers (R, \hat{R}) is said to be (E, Δ) -achievable rates pair for $E_1 > 0$, $E_2 > 0$, $\Delta_x^1 \geq 0$, $\Delta_x^2 \geq 0$, $\Delta_y^1 \geq 0$, $\Delta_y^2 \geq 0$, if for arbitrary $\varepsilon > 0$ and n sufficiently large there exists a code (f, \hat{f}, F, \hat{F}) such that

$$n^{-1} \log K(n) \leq R + \varepsilon, \quad n^{-1} \log L(n) \leq \hat{R} + \varepsilon,$$

and

$$e_i(f, \hat{f}, F, \hat{F}, \Delta_x^i, \Delta_y^i) \leq \exp(-nE_i), \quad i = 1, 2. \quad (1)$$

Following Shannon we call exponents E_1, E_2 "reliabilities" at the first and second decoders, respectively. Let us denote by $\mathcal{R}(E, \Delta)$ the set of all pairs of (E, Δ) -achievable rates.

$\mathcal{R}(E, \Delta)$ is a generalization of the rates-distortions region $\mathcal{R}(\Delta)$ (corresponding to the case $E_1 \rightarrow 0, E_2 \rightarrow 0$). If at the first decoder only the messages of the source $\{X\}$ and at the second decoder only the messages of the source $\{Y\}$ are reconstructed, then $\mathcal{R}(\Delta)$ becomes the set of all (Δ_x^1, Δ_y^2) -achievable rates pairs $\mathcal{R}_1(\Delta_x^1, \Delta_y^2)$ studied by Yamamoto in [7]. If $\hat{R} = 0$, then $\mathcal{R}(\Delta)$ give us the set of all (Δ_x^1, Δ_y^1) -achievable rates $\mathcal{R}_2(\Delta_x^1, \Delta_y^1)$ for the system presented in Fig. 2.

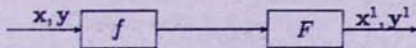


Fig. 2. Two-source with common encoder and decoder.

In the present paper $\mathcal{R}(E, \Delta)$ is specified. For some particular communication systems rates-distortions regions are derived. The results are formulated in the next section, the proofs are given in section 3.

2 Formulation of Results

Let $P = \{P(x, y), x \in \mathcal{X}, y \in \mathcal{Y}\}$ be some PD on $\mathcal{X} \times \mathcal{Y}$ and

$$Q = \{Q(x^1, y^1, x^2, y^2 | x, y), x \in \mathcal{X}, y \in \mathcal{Y}, x^1 \in \mathcal{X}^1, y^1 \in \mathcal{Y}^1, x^2 \in \mathcal{X}^2, y^2 \in \mathcal{Y}^2\}$$

be conditional PD on $\mathcal{X}^1 \times \mathcal{Y}^1 \times \mathcal{X}^2 \times \mathcal{Y}^2$ for given x and y . Let

$$\alpha(E) = \{P : D(P \| P^*) \leq E\}.$$

Denote by $\Phi(P) = Q_P$ the function, which puts into correspondence to PD P some PD Q_P such that for the given Δ , if $E_1 < E_2$ and $P \in \alpha(E_1)$ or $E_2 < E_1$ and $P \in \alpha(E_2)$, then following four inequalities are valid

$$\begin{aligned} E_{P, Q_P} d_x^1(X, X^1) = \\ = \sum_{x, x^1} \sum_{y, y^1} \sum_{x^2, y^2} P(x, y) Q_P(x^1, y^1, x^2, y^2 | x, y) d_x^1(x, x^1) \leq \Delta_x^1, \end{aligned} \quad (2)$$

$$E_{P, Q_P} d_y^1(Y, Y^1) \leq \Delta_y^1, \quad (3)$$

$$E_{P, Q_P} d_x^2(X, X^2) \leq \Delta_x^2, \quad (4)$$

$$E_{P, Q_P} d_y^2(Y, Y^2) \leq \Delta_y^2, \quad (5)$$

and if $E_1 < E_2$ and $P \in \alpha(E_2) - \alpha(E_1)$, then the conditions (4) and (5) take place, or if $E_2 < E_1$ and $P \in \alpha(E_1) - \alpha(E_2)$, then (2) and (3) take place. Let $\mathcal{M}(E, \Delta)$ be the set of all such functions $\Phi(P)$ for given Δ and E . When $E_1 \rightarrow 0, E_2 \rightarrow 0$ we note $\mathcal{M}(\Delta)$ the set of all $\Phi(P)$ for which (2)-(5) are valid. Let us define the region:

$$\mathcal{R}^*(E, \Phi) = \{(R, \hat{R}) :$$

$$\text{for } 0 < E_1 \leq E_2,$$

$$\begin{aligned} R &\geq \max_{P \in \alpha(E_1)} I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1, X^2, Y^2), \hat{R}, \\ \hat{R} &\geq \max_{P \in \alpha(E_2)} I_{P, \Phi(P)}(X, Y \wedge X^2, Y^2), \end{aligned}$$

for $0 < E_2 \leq E_1$,

$$R \geq \max \left(\max_{P \in \alpha(E_2)} I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1, X^2, Y^2), \right.$$

$$\left. \max_{P \in \alpha(E_1) - \alpha(E_2)} I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1) \right),$$

$$\hat{R} \geq \max_{P \in \alpha(E_2)} I_{P, \Phi(P)}(X, Y \wedge X^2, Y^2) \}.$$

Let

$$\mathcal{R}^*(E, \Delta) = \bigcup_{\Phi \in \mathcal{M}(E, \Delta)} \mathcal{R}^*(E, \Phi).$$

Theorem: For $E_1 > 0$, $E_2 > 0$, $\Delta_x^1 \geq 0$, $\Delta_x^2 \geq 0$, $\Delta_y^1 \geq 0$, $\Delta_y^2 \geq 0$,

$$\mathcal{R}(E, \Delta) = \mathcal{R}^*(E, \Delta).$$

Many particular cases are worth to be specified.

Corollary 1: With $E_1 \rightarrow 0$, $E_2 \rightarrow 0$, we obtain "rate-distortion region":

$$\mathcal{R}(\Delta) = \bigcup_{\Phi \in \mathcal{M}(\Delta)} \{(R, \hat{R})\}:$$

$$R \geq I_{P^*, \Phi(P^*)}(X, Y \wedge X^1, Y^1, X^2, Y^2),$$

$$\hat{R} \geq I_{P^*, \Phi(P^*)}(X, Y \wedge X^2, Y^2) \}.$$

Corollary 2: When at the first decoder only the messages of the source $\{X\}$ and at the second decoder only the messages of the source $\{Y\}$ are reconstructed, we arrive to the result of Yamamoto [7]:

$$\mathcal{R}(\Delta) \equiv \mathcal{R}_1(\Delta_x^1, \Delta_y^2),$$

with

$$\mathcal{R}_1(\Delta_x^1, \Delta_y^2) = \bigcup_{\Phi \in \mathcal{M}(\Delta)} \{(R, \hat{R})\}:$$

$$R \geq I_{P^*, \Phi(P^*)}(X, Y \wedge X^1, Y^2), \quad \hat{R} \geq I_{P^*, \Phi(P^*)}(X, Y \wedge Y^2) \}.$$

Corollary 3: If $\hat{R} = 0$, $\mathcal{R}(\Delta)$ becomes

$$\mathcal{R}_2(\Delta_x^1, \Delta_y^1) = \{R : R \geq \min_{\Phi \in \mathcal{M}(\Delta)} I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1) \}.$$

Corollary 4: When only at the first decoder messages of the source $\{X\}$ are reconstructed, then we denote

$$\mathcal{R}_2(\Delta_x^1, \Delta_y^1) \equiv \mathcal{R}_3(\Delta_x^1),$$

with

$$\mathcal{R}_3(\Delta_x^1) = \{R : R \geq \min_{\Phi \in \mathcal{M}(\Delta)} I_{P, \Phi(P)}(X, Y \wedge X^1) \}.$$

Corollary 5: If $\mathcal{X} \equiv \mathcal{X}^1 \equiv \mathcal{X}^2$, $\mathcal{Y} \equiv \mathcal{Y}^1 \equiv \mathcal{Y}^2$, $\Delta_x^1 = \Delta_x^2 = \Delta_y^1 = \Delta_y^2 = 0$, and the distortion measures are the Hamming distances

$$d_x^i(x, x^i) = \begin{cases} 0, & \text{for } x = x^i, \\ 1, & \text{for } x \neq x^i, \end{cases} \quad d_y^i(y, y^i) = \begin{cases} 0, & \text{for } y = y^i, \\ 1, & \text{for } y \neq y^i, \end{cases} \quad i = 1, 2,$$

the region of E -achievable rates pairs will be

$$\mathcal{R}(E) = \{(R, \hat{R}) :$$

$$R \geq \max(\max_{P \in \alpha(E_1)} H_P(X, Y), \hat{R}), \quad \hat{R} \geq \max_{P \in \alpha(E_2)} H_P(X, Y)\}.$$

Corollary 6: If $\mathcal{X} \equiv \mathcal{Y}$, $\mathcal{Y}^1 = \mathcal{X}^3$, $\mathcal{Y}^2 = \mathcal{X}^4$, our system (see Fig. 1.) is a multiple description system. Some cases of this system are studied in [10]-[13]. For our system the region of (E, Δ) -achievable rates pairs $\mathcal{R}_4(E, \Delta)$ may be derived from the theorem as follows:

$$\mathcal{R}_4(E, \Delta) = \bigcup_{\Phi \in \mathcal{M}(E, \Delta)} \{(R, \hat{R}) :$$

$$\text{for } 0 < E_1 \leq E_2,$$

$$R \geq \max(\max_{P \in \alpha(E_1)} I_{P, \Phi(P)}(X \wedge X^1, X^2, X^3, X^4), \hat{R}),$$

$$\hat{R} \geq \max_{P \in \alpha(E_2)} I_{P, \Phi(P)}(X \wedge X^2, X^4),$$

$$\text{for } 0 < E_2 \leq E_1,$$

$$R \geq \max(\max_{P \in \alpha(E_2)} I_{P, \Phi(P)}(X \wedge X^1, X^2, X^3, X^4),$$

$$\max_{P \in \alpha(E_1) - \alpha(E_2)} I_{P, \Phi(P)}(X \wedge X^1, X^3)),$$

$$\hat{R} \geq \max_{P \in \alpha(E_2)} I_{P, \Phi(P)}(X \wedge X^2, X^4)\}.$$

Moreover, if $\hat{R} = R$, we obtain the following (E, Δ) -achievable rates region

$$\mathcal{R}'_4(E, \Delta) = \{R :$$

$$\text{for } 0 < E_1 \leq E_2,$$

$$R \geq \max(\max_{P \in \alpha(E_1)} \min_{\Phi \in \mathcal{M}(E, \Delta)} I_{P, \Phi(P)}(X \wedge X^1, X^2, X^3, X^4),$$

$$\max_{P \in \alpha(E_2)} \min_{\Phi \in \mathcal{M}(E, \Delta)} I_{P, \Phi(P)}(X \wedge X^2, X^4)),$$

$$\text{for } 0 < E_2 \leq E_1,$$

$$R \geq \max(\max_{P \in \alpha(E_2)} \min_{\Phi \in \mathcal{M}(E, \Delta)} I_{P, \Phi(P)}(X \wedge X^1, X^2, X^3, X^4),$$

$$\max_{P \in \alpha(E_1) - \alpha(E_2)} \min_{\Phi \in \mathcal{M}(E, \Delta)} I_{P, \Phi(P)}(X \wedge X^1, X^3))\}.$$

Corollary 7: When $E_1 \rightarrow 0$, $E_2 \rightarrow 0$, then $\mathcal{R}_4(E, \Delta)$ becomes

$$\mathcal{R}_4(\Delta) = \bigcup_{\Phi \in \mathcal{M}(\Delta)} \{(R, \hat{R}) :$$

$$R \geq I_{P^*, \Phi(P^*)}(X \wedge X^1, X^2, X^3, X^4), \quad R \geq I_{P^*, \Phi(P^*)}(X \wedge X^2, X^4),$$

and $\mathcal{R}'_4(E, \Delta)$ becomes

$$\mathcal{R}'_4(\Delta) = \{R : R \geq \min_{\Phi \in \mathcal{M}(\Delta)} I_{P, \Phi(P)}(X \wedge X^1, X^2, X^3, X^4)\}.$$

When $X^3 \equiv X^1$, $X^4 \equiv X^2$, then $\mathcal{R}'_4(\Delta)$ coincides with the set of all Δ -achievable rates obtained by El Gamal and Cover [10].

Remarks: 1. For ordinary one-way source the region $\mathcal{R}(E, \Delta)$ was found by Haroutunian and Mekoush in [8]. The system presented in Fig. 2, will be an ordinary one-way source $\{X, Y\}$ if one common distortion measure for the messages of source $\{X, Y\}$ is considered instead of two separate distortion measures for the messages of each sources $\{X\}$ and $\{Y\}$.

2. For the system presented in Fig. 2 one can consider two separate error probabilities of the code with given exponents E_x and E_y for the messages of sources $\{X\}$ and $\{Y\}$, respectively. We can obtain only the following outer and inner bounds of $(E_x, E_y, \Delta_x^1, \Delta_y^1)$ -achievable rates region

$$\mathcal{R}_5(E_x, E_y, \Delta_x^1, \Delta_y^1) = \{R : R \geq \max_{P \in \alpha(\min(E_x, E_y))} \min_{\Phi \in \mathcal{M}(\Delta)} I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1)\},$$

$$\mathcal{R}_6(E_x, E_y, \Delta_x^1, \Delta_y^1) = \{R : R \geq \max_{P \in \alpha(\max(E_x, E_y))} \min_{\Phi \in \mathcal{M}(\Delta)} I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1)\}.$$

The demonstrations are similar to the proofs of the converse and positive parts of the theorem, respectively.

3 Proof of Theorem

We use the typical sequences technique [2]. In the proof of positive part of the theorem we apply the following modification of covering lemma from [2], [13], [14], [17]:

Lemma: Let for fixed type P , conditional type Q and $\varepsilon > 0$

$$K(P, Q) = \exp\{n(I_{P, Q}(X, Y \wedge X^1, Y^1) + \varepsilon)\},$$

$$L(P, Q) = \exp\{n(I_{P, Q}(X, Y \wedge X^2, Y^2) + \varepsilon)\},$$

$$M(P, Q) = \exp\{n(I_{P, Q}(X, Y \wedge X^1, Y^1 | X^2, Y^2) + \varepsilon)\}.$$

Then for n large enough there exist collections of conditional types

$$\{\mathcal{T}_{P, Q}(X, Y | x_i^1, y_l^1), l = \overline{1, L(P, Q)}\}, \quad \{\mathcal{T}_{P, Q}(X, Y | x_k^1, y_k^1), k = \overline{1, K(P, Q)}\},$$

which are coverings for $\mathcal{T}_P(X, Y)$. Moreover, for any $l = \overline{1, L(P, Q)}$ there exists a collection $\{\mathcal{T}_{P, Q}(X, Y | x_i^1, y_l^1, x_{i, m}^1, y_{l, m}^1), m = \overline{1, M(P, Q)}\}$, which is a covering for $\mathcal{T}_{P, Q}(X, Y | x_i^1, y_l^1)$.

The proof of lemma is similar to the proof of lemma from [17].

We begin the proof of the theorem with the ascertainment of the inclusion

$$\mathcal{R}^*(E, \Delta) \subseteq \mathcal{R}(E, \Delta). \quad (6)$$

Denote by $\mathcal{P}(X, Y, n)$ the set of all types P . Let us present $\mathcal{X}^n \times \mathcal{Y}^n$ as a union of all disjoint sets $\mathcal{T}_P(X, Y)$ (for brevity also called types), which are the sets of all pairs of vectors of joint type P :

$$\mathcal{X}^n \times \mathcal{Y}^n = \bigcup_{P \in \mathcal{P}(X, Y, n)} \mathcal{T}_P(X, Y).$$

Let some $\varepsilon > 0$ be given. A union of all $\mathcal{T}_P(X, Y)$ for $P \notin \alpha(E_i + \varepsilon)$, $i = 1, 2$, has probability small enough

$$\begin{aligned} P^{*n} \left(\bigcup_{P \notin \alpha(E_i + \varepsilon)} \mathcal{T}_P(X, Y) \right) &\leq \sum_{P \notin \alpha(E_i + \varepsilon)} P^{*n}(\mathcal{T}_P(X, Y)) \leq \\ &\leq (n+1)^{|\mathcal{X}||\mathcal{Y}|} \exp\{-n(\min_{P \notin \alpha(E_i + \varepsilon)} D(P \| P^*)\} \leq \\ &\leq \exp\{-nE_i - n\varepsilon + |\mathcal{X}||\mathcal{Y}| \log(n+1)\}. \end{aligned} \quad (7)$$

Hence, to obtain the error probabilities small enough it is sufficient to construct encoding functions f and \hat{f} such that the vectors of the types $P \in \alpha(E_1 + \varepsilon)$ and $P \in \alpha(E_2 + \varepsilon)$, respectively, will be decoded with distortions small enough.

For any type P let us fix $\Phi \in \mathcal{M}(E, \Delta)$ and denote $\Phi(P) = Q_P$. According to the lemma there exist the coverings $\{\mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_l^1, \mathbf{y}_l^2), l = \overline{1, L(P, Q_P)}\}$ and $\{\mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_k^1, \mathbf{y}_k^1), k = \overline{1, K(P, Q_P)}\}$ for $\mathcal{T}_P(X, Y)$. Let

$$\begin{aligned} \mathcal{C}_k(P, Q_P) &= \mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_k^1, \mathbf{y}_k^1) - \\ &- \bigcup_{k' < k} \mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_{k'}^1, \mathbf{y}_{k'}^1), \quad k = \overline{1, K(P, Q_P)}, \\ \mathcal{D}_l(P, Q_P) &= \mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_l^2, \mathbf{y}_l^2) - \\ &- \bigcup_{l' < l} \mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_{l'}^2, \mathbf{y}_{l'}^2), \quad l = \overline{1, L(P, Q_P)}. \end{aligned}$$

For every $l = \overline{1, L(P, Q_P)}$ a covering

$$\{\mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_l^2, \mathbf{y}_l^2, \mathbf{x}_{l, m}^1, \mathbf{y}_{l, m}^1), m = \overline{1, M(P, Q_P)}\}$$

for $\mathcal{T}_{P, Q}(X, Y | \mathbf{x}_l^2, \mathbf{y}_l^2)$ exists. Let

$$\begin{aligned} \mathcal{S}_{l, m}(P, Q_P) &= \mathcal{D}_l(P, Q_P) \cap \{\mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_l^2, \mathbf{y}_l^2, \mathbf{x}_{l, m}^1, \mathbf{y}_{l, m}^1) - \\ &- \bigcup_{m' < m} \mathcal{T}_{P, Q_P}(X, Y | \mathbf{x}_l^2, \mathbf{y}_l^2, \mathbf{x}_{l, m'}^1, \mathbf{y}_{l, m'}^1)\}, \quad m = \overline{1, M(P, Q_P)}. \end{aligned}$$

For $E_1 \leq E_2$ we define a code (f, \hat{f}, F, \hat{F}) as follows: the first encoding is

$$f(\mathbf{x}, \mathbf{y}) = \begin{cases} (l, m), & \text{when } (\mathbf{x}, \mathbf{y}) \in \mathcal{S}_{l, m}(P, Q_P), \quad P \in \alpha(E_1 + \varepsilon), \\ l, & \text{when } (\mathbf{x}, \mathbf{y}) \in \mathcal{D}_l(P, Q_P), \quad P \in \alpha(E_2 + \varepsilon) - \alpha(E_1 + \varepsilon), \\ l', & \text{when } (\mathbf{x}, \mathbf{y}) \in \mathcal{T}_P(X, Y), \quad P \notin \alpha(E_2 + \varepsilon), \end{cases}$$

the second encoding is

$$f(l, m) = l, \quad f(l) = l, \quad f(l') = l'',$$

decodings at the first and second outputs, respectively, are

$$F(l, m) = (x_{l,m}^1, y_{l,m}^1), \quad F(l) = l'', \quad F(l') = l'',$$

$$\hat{F}(l) = (x_l^2, y_l^2), \quad \hat{F}(l'') = l''.$$

According to the definition of the code (f, \hat{f}, F, \hat{F}) , the lemma and the inequalities (2) and (3) for $P \in \alpha(E_1 + \varepsilon)$ we have

$$\begin{aligned} d_x^1(x, x^1) &= n^{-1} \sum_{x, x^1} n(x, x^1 | x, x^1) d_x^1(x, x^1) = \\ &= \sum_{x, x^1} \sum_{y, y^1} \sum_{x^2, y^2} P(x, y) Q_P(x^1, y^1, x^2, y^2 | x, y) d_x^1(x, x^1) = \\ &= E_{P, Q_P} d_x^1(X, X^1) \leq \Delta_x^1, \end{aligned} \quad (8)$$

and, similarly,

$$d_y^1(y, y^1) = E_{P, Q_P} d_y^1(Y, Y^1) \leq \Delta_y^1. \quad (9)$$

For $P \in \alpha(E_2 + \varepsilon)$ from (4) and (5) we have

$$d_x^2(x, x^2) = E_{P, Q_P} d_x^2(X, X^2) \leq \Delta_x^2, \quad (10)$$

$$d_y^2(y, y^2) = E_{P, Q_P} d_y^2(Y, Y^2) \leq \Delta_y^2. \quad (11)$$

We see that when

$$\begin{aligned} K(n) &\geq \exp\{n(\max_{P \in \alpha(E_1 + \varepsilon)} (I_{P, Q_P}(X, Y \wedge X^2, Y^2) + \\ &\quad + I_{P, Q_P}(X, Y \wedge X^1, Y^1 | X^2, Y^2) + 2\varepsilon), L(n))\} = \\ &= \exp\{n(\max_{P \in \alpha(E_1 + \varepsilon)} I_{P, Q_P}(X, Y \wedge X^1, Y^1, X^2, Y^2) + 2\varepsilon, L(n))\}, \\ L(n) &\geq \exp\{n(\max_{P \in \alpha(E_2 + \varepsilon)} I_{P, Q_P}(X, Y \wedge X^2, Y^2) + \varepsilon)\}, \end{aligned}$$

the error probabilities are small enough.

For $E_2 \leq E_1$ we define code (f, \hat{f}, F, \hat{F}) as follows:

$$f(x, y) = \begin{cases} (l, m), & \text{when } (x, y) \in \mathcal{S}_{l, m}(P, Q_P), P \in \alpha(E_2 + \varepsilon), \\ k, & \text{when } (x, y) \in \mathcal{C}_k(P, Q_P), P \in \alpha(E_1 + \varepsilon) - \alpha(E_2 + \varepsilon), \\ l', & \text{when } (x, y) \in \mathcal{T}_P(X, Y), P \notin \alpha(E_1 + \varepsilon), \end{cases}$$

$$\hat{f}(l, m) = l, \quad \hat{f}(k) = l'', \quad \hat{f}(l') = l'',$$

$$F(l, m) = (x_{l, m}^1, y_{l, m}^1), \quad F(k) = (x_k^1, y_k^1), \quad F(l') = l'',$$

$$\hat{F}(l) = (x_l^2, y_l^2), \quad \hat{F}(l'') = l''.$$

According to this definition of the code (f, \hat{f}, F, \hat{F}) , the lemma and the inequalities (2)-(5) we obtain that (8)-(11) take place for $E_2 \leq E_1$ and

$$K(n) \geq \exp\{n(\max_{P \in \alpha(E_1 + \varepsilon)} \max_{P \in \alpha(E_2 + \varepsilon)} I_{P,Q_P}(X, Y \wedge X^1, Y^1, X^2, Y^2) + 2\varepsilon,$$

$$\max_{P \in \alpha(E_1 + \varepsilon) - \alpha(E_2 + \varepsilon)} I_{P,Q_P}(X, Y \wedge X^1, Y^1) + \varepsilon)\},$$

$$L(n) \geq \exp\{n(\max_{P \in \alpha(E_2 + \varepsilon)} I_{P,Q_P}(X, Y \wedge X^2, Y^2) + \varepsilon)\}.$$

The error will be possible at the first and second decoders only if the type P of (x, y) is not in $\alpha(E_1 + \varepsilon)$ and $\alpha(E_2 + \varepsilon)$, respectively. According to (7) the error probabilities are small as it is specified in (1).

Taking into account arbitrariness of ε , continuity of obtained expressions with respect to E_1 and E_2 , we obtain (6).

Now we shall prove the inclusion

$$\mathcal{R}(E, \Delta) \subseteq \mathcal{R}^*(E, \Delta). \quad (12)$$

Let $\varepsilon > 0$ is fixed and a given code (f, \hat{f}, F, \hat{F}) has (E, Δ) -achievable pair (R, \hat{R}) of rates. For n large enough, $E_1 \leq E_2$ and P from $\alpha(E_1 - \varepsilon)$ the following inequality takes place

$$|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| \geq \exp\{n(H_P(X, Y) - \varepsilon)\} \quad (13)$$

(for the proof of (13) see Appendix).

To each pair of vectors $(x, y) \in \mathcal{A}_1 \cap \mathcal{A}_2$ the unique quadruple of vectors (x^1, y^1, x^2, y^2) corresponds such that $(x^1, y^1) = F(f(x, y))$, $(x^2, y^2) = \hat{F}(\hat{f}(f(x, y)))$. These six vectors determine types P and Q , for which

$$(x^1, y^1, x^2, y^2) \in \mathcal{T}_{P,Q}(X^1, Y^1, X^2, Y^2 | x, y),$$

or, which is equivalent,

$$(x, y) \in \mathcal{T}_{P,Q}(X, Y | x^1, y^1, x^2, y^2).$$

The set of all vectors $(x, y) \in \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y)$ is divided into classes corresponding to these conditional types Q . Let us select from them the class, which for given P contains the greatest number of pairs (x, y) . Corresponding conditional type Q we denote by $Q_P = \Phi(P)$, and the class itself we denote by

$$(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P)).$$

Using polynomial upper estimate [2] of the number of conditional types Q we have

$$\begin{aligned} |\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| &\leq \\ &\leq (n+1)^{|x^1||y^1||x^2||y^2|} |(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))|. \end{aligned} \quad (14)$$

Let $\mathcal{D}_{1,2}(P, \Phi(P))$ be the set of all quadruples of vectors (x^1, y^1, x^2, y^2) for which for given $\Phi(P) \in \mathcal{M}(E, \Delta)$ there exist vectors

$$(x, y) \in \tilde{\mathcal{I}}_{P,Q}(X, Y | x^1, y^1, x^2, y^2)$$

such that $F(f(x, y)) = (x^1, y^1)$, $\hat{F}(\hat{f}(f(x, y))) = (x^2, y^2)$. According to the definition of the code

$$|\mathcal{D}_{1,2}(P, \Phi(P))| \leq K(n).$$

We see that

$$\begin{aligned} & |(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))| \leq \\ & \leq \sum_{(x^1, y^1, x^2, y^2) \in \mathcal{D}_{1,2}(P, \Phi(P))} |\mathcal{T}_{P, \Phi(P)}(X, Y | x^1, y^1, x^2, y^2)| \leq \\ & \leq K(n) \exp\{n H_{P, \Phi(P)}(X, Y | X^1, Y^1, X^2, Y^2)\}. \end{aligned} \quad (15)$$

From the last inequality and from (14) we obtain that

$$\begin{aligned} & |\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| \leq \\ & \leq K(n) \exp\{n(H_{P, \Phi(P)}(X, Y | X^1, Y^1, X^2, Y^2) + \varepsilon)\}. \end{aligned}$$

Taking into account the last inequality and (13) we receive that

$$K(n) \geq \exp\{n(I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1, X^2, Y^2) - 2\varepsilon)\}.$$

Since $K(n) \geq L(n)$, then using (15) we obtain that for $E_1 \leq E_2$ and P from $\alpha(E_1 - \varepsilon)$

$$R \geq \max(I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1, X^2, Y^2) - 2\varepsilon, \hat{R}).$$

Remark that

$$\left| \bigcup_{(x^1, y^1) \in \exists(x^2, y^2), (x^1, y^1, x^2, y^2) \in \mathcal{D}_{1,2}(P, \Phi(P))} \mathcal{D}_{1,2}(P, \Phi(P)) \right| \leq L(n).$$

For the same $\Phi(P)$ as in (14) similarly to (15) we have

$$\begin{aligned} & |(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))| \leq \\ & \leq L(n) \exp\{n H_{P, \Phi(P)}(X, Y | X^2, Y^2)\}. \end{aligned}$$

With (13) we obtain

$$L(n) \geq \exp\{n(I_{P, \Phi(P)}(X, Y \wedge X^2, Y^2) - 2\varepsilon)\}.$$

Therefore for $E_1 \leq E_2$ and P from $\alpha(E_1 - \varepsilon)$

$$\hat{R} \geq I_{P, \Phi(P)}(X, Y \wedge X^2, Y^2) - 2\varepsilon. \quad (16)$$

Now note that for n large enough, $E_1 \leq E_2$ and P from $\alpha(E_2 - \varepsilon) - \alpha(E_1 - \varepsilon)$ we have

$$|\mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| \geq \exp\{n(H_P(X, Y) - \varepsilon)\} \quad (17)$$

(for the proof of (17) see Appendix).

Similarly to choice of the class $(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))$ in (14) we shall choose $(\mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))$. Then

$$\begin{aligned} & |\mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| \leq \\ & \leq (n+1)^{|\mathcal{X}||\mathcal{Y}||\mathcal{X}^2||\mathcal{Y}^2|} |(\mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))|. \end{aligned} \quad (18)$$

Let $\mathcal{D}_{2/1}(P, \Phi(P))$ be the set of all pairs of vectors (x^2, y^2) for which for given P and $\Phi(P)$ there exists vectors $(x^1, y^1) \in \mathcal{X}^{1n} \times \mathcal{Y}^{1n}$ and

$$(x, y) \in \mathcal{T}_{P, \Phi(P)}(X, Y | x^2, y^2)$$

such that $F(f(x, y)) = (x^1, y^1)$, $\hat{F}(\hat{f}(f(x, y))) = (x^2, y^2)$. It is clear that

$$|\mathcal{D}_{2/1}(P, \Phi(P))| \leq L(n).$$

We see that

$$\begin{aligned} & |(\mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))| \leq \\ & \leq \sum_{(x^2, y^2) \in \mathcal{D}_{2/1}(P, \Phi(P))} |\mathcal{T}_{P, \Phi(P)}(X, Y | x^2, y^2)| \leq \\ & \leq L(n) \exp\{n(H_{P, \Phi(P)}(X, Y | X^2, Y^2))\}. \end{aligned}$$

Taking into account (17), (18) and the last inequality we receive that (16) takes place also for $P \in \alpha(E_2 - \varepsilon) - \alpha(E_1 - \varepsilon)$. By analogy with the case $E_1 \leq E_2$, $P \in \alpha(E_1 - \varepsilon)$ we can show that for n large enough, $E_2 \leq E_1$ and $P \in \alpha(E_2 - \varepsilon)$, the (15) takes place. Hence for $P \in \alpha(E_2 - \varepsilon)$

$$R \geq I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1, X^2, Y^2) - 2\varepsilon. \quad (19)$$

For n large enough, $E_2 \leq E_1$ and $P \in \alpha(E_1 - \varepsilon) - \alpha(E_2 - \varepsilon)$ the following inequality takes place

$$|\mathcal{A}_1 \cap \mathcal{T}_P(X, Y)| \geq \exp\{n(H_P(X, Y) - \varepsilon)\} \quad (20)$$

(the proof of (20) is similar with the proof of (17)).

Similarly choice of the class $(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y))(\Phi(P))$ in (14) we shall choose $(\mathcal{A}_1 \cap \mathcal{T}_P(X, Y))(\Phi(P))$. Then

$$\begin{aligned} & |\mathcal{A}_1 \cap \mathcal{T}_P(X, Y)| \leq \\ & \leq (n+1)^{|\mathcal{X}^1||\mathcal{Y}^1||\mathcal{X}^2||\mathcal{Y}^2|} |(\mathcal{A}_1 \cap \mathcal{T}_P(X, Y))(\Phi(P))|. \end{aligned} \quad (21)$$

Let $\mathcal{D}_{1/2}(P, \Phi(P))$ be the set of all pairs of vectors (x^1, y^1) for which for given P and $\Phi(P)$ there exists vectors $(x^2, y^2) \in \mathcal{X}^{2n} \times \mathcal{Y}^{2n}$ and

$$(x, y) \in \mathcal{T}_{P, \Phi(P)}(X, Y | x^1, y^1)$$

such that $F(f(x, y)) = (x^1, y^1)$, $\hat{F}(\hat{f}(f(x, y))) = (x^2, y^2)$. It is clear that

$$|\mathcal{D}_{1/2}(P, \Phi(P))| \leq K(n).$$

We can show that

$$\begin{aligned} & |(\mathcal{A}_1 \cap \mathcal{T}_P(X, Y))(\Phi(P))| \leq \\ & \leq \sum_{(x^1, y^1) \in \mathcal{D}_{1/2}(P, \Phi(P))} |\mathcal{T}_{P, \Phi(P)}(X, Y | x^1, y^1)| \leq \\ & \leq K(n) \exp\{n(H_{P, \Phi(P)}(X, Y | X^1, Y^1))\}. \end{aligned}$$

Using the last inequality, (20) and (21) we receive that

$$K(n) \geq \exp\{n(I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1) - 2\varepsilon)\}.$$

Therefore for $P \in \alpha(E_1 - \varepsilon) - \alpha(E_2 - \varepsilon)$

$$R \geq I_{P, \Phi(P)}(X, Y \wedge X^1, Y^1) - 2\varepsilon.$$

By analogy with the case $P \in \alpha(E_1 - \varepsilon)$, $E_1 \leq E_2$ we can show that for the same $\Phi(P)$ as in (19), for $E_2 \leq E_1$ and $P \in \alpha(E_2 - \varepsilon)$ the (16) takes place.

Since all used functions are continuous with respect to E_1 and E_2 , the union of obtained sets of rates pairs (R, \hat{R}) by mappings $\Phi(P)$ belonging to $\mathcal{M}(E, \Delta)$, give us inclusion (12).

3.1 Appendix

The proof of the inequality (13). We can write

$$|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| = |\mathcal{T}_P(X, Y)| - |\overline{\mathcal{A}_1} \cup \overline{\mathcal{A}_2} \cap \mathcal{T}_P(X, Y)|.$$

For $P \in \alpha(E_1 - \varepsilon)$, $E_1 \leq E_2$ we have

$$\begin{aligned} |\overline{\mathcal{A}_1} \cup \overline{\mathcal{A}_2} \cap \mathcal{T}_P(X, Y)| & \leq \frac{P^{*n}(\overline{\mathcal{A}_1} \cup \overline{\mathcal{A}_2} \cap \mathcal{T}_P(X, Y))}{P^{*n}(x, y)} \leq \\ & \leq (\exp(-nE_1) + \exp(-nE_2)) \exp\{n(H_P(X, Y) + D(P \| P^*))\} \leq \\ & \leq \exp\{n(H_P(X, Y) - \varepsilon)\}. \end{aligned}$$

Then for n large enough

$$|\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| \geq \exp\{n(H_P(X, Y) - \varepsilon)\}.$$

The proof of the inequality (17). It is clear that

$$|\mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| = |\mathcal{T}_P(X, Y)| - |\overline{\mathcal{A}_2} \cap \mathcal{T}_P(X, Y)|.$$

For $P \in \alpha(E_2 - \varepsilon) - \alpha(E_1 - \varepsilon)$, $E_1 \leq E_2$ we have

$$\begin{aligned} |\overline{\mathcal{A}_2} \cap \mathcal{T}_P(X, Y)| & \leq \frac{P^{*n}(\overline{\mathcal{A}_2} \cap \mathcal{T}_P(X, Y))}{P^{*n}(x, y)} \leq \\ & \leq \exp(-nE_2) \exp\{n(H_P(X, Y) + D(P \| P^*))\} \leq \\ & \leq \exp\{n(H_P(X, Y) - \varepsilon)\}. \end{aligned}$$

And then

$$|\mathcal{A}_2 \cap \mathcal{T}_P(X, Y)| \geq \exp\{n(H_P(X, Y) - \varepsilon)\}.$$

References

- [1] T. Berger, *Rate Distortion Theory, A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [2] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic, 1981.
- [3] R. M. Gray, "A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions", *IEEE Trans. Inform. Theory*, vol. IT-19, no. 4, pp. 480-489, 1973.
- [4] R. M. Gray and A. D. Wyner, "Source coding for a simple network", *Bell System Tech. J.*, vol. 53, no. 9, pp. 1681-1721, 1974.
- [5] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder", *IEEE Trans. Inform. Theory*, vol. IT-22, no. 1, pp. 1-10, 1976.
- [6] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources", *Inform. Contr.*, vol. 38, pp. 60-80, 1978.
- [7] H. Yamamoto, "Source coding theory for cascade and branching communication systems", *IEEE Trans. Inform. Theory*, vol. IT-27, no. 3, pp. 299-308, 1981..
- [8] E. A. Haroutunian and B. Mekoush, "Estimates of optimal rates of codes with given error probability exponent for certain sources" (in Russian), *Abstr. of VI Intern. Symposium of Inform. Theory*, Tashkent, vol. 1, pp. 22-23, 1984.
- [9] T. Ericson and J. Körner, "Successive encoding of correlated sources", *IEEE Trans. Inform. Theory*, vol. IT-29, no. 3, pp. 390-395, 1983.
- [10] A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions", *IEEE Trans. Inform. Theory*, vol. IT-28, no. 6, pp. 851-857, 1982.
- [11] R. Ahlswede, "The rate-distortion region for multiple descriptions without excess rate", *IEEE Trans. Inform. Theory*, vol. IT-31, no. 6, pp. 721-726, 1985.
- [12] R. Sh. Maroutian, "Achievable rates for multiple descriptions with given exponent and distortion levels" (in Russian), *Problems Inform. Transm.*, vol. 26, no. 1, pp. 83-89, 1990.
- [13] E. A. Haroutunian and R. Sh. Maroutian, " (E, Δ) -achievable rates for multiple descriptions of random varying source", *Problems of Control and Inform. Theory*, vol. 20, no. 2, pp. 165-178, 1991.
- [14] R. Ahlswede, "Coloring hypergraphs. A new approach to multi-user source coding", I, II, *J. Combin. Inform. and Syst. Sci.*, vol. 4, no. 1, pp. 75-115, 1979; vol. 5, no. 2, pp. 220-268, 1980
- [15] A. H. Kaspi, "Rate-distortion function when side-information may be present at the decoder", *IEEE Trans. Inform. Theory*, vol. 40, no. 6, 1994.

- [16] H. Yamamoto, "Triangular communication system", IEEE Trans. Inform. Theory, vol. 42, no. 3, pp. 248-253, 1996.
- [17] E. A. Haroutunian, "Multiterminal Source Coding Achievable Rates and Reliability", IEEE Trans. Inform. Theory, vol. 42, no. 6, 1996.