Mathematical Problems of Computer Science 20, 1998, 98–110.

Rate-reliabilities-distortions Dependence for Source Coding with Many Receivers

E. A. Haroutunian, A. N. Haroutunian and A. R. Kazarian

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

Abstract

We study an information transmission system, where messages of a discrete memoryless source are encoded by one encoder and decoded by many decoders. El Gamal and Cover named it robust descriptions system and found the rate-distortions function, that is the minimal achievable coding rate with respect to distortion criteria at the receivers.

We specify more general function – rate-reliabilities-distortions, that is the minimal achievable coding rate of the best codes ensuring reconstruction of original messages within given distortion levels and error probabilities exponents (reliabilities) at each decoder. Example is given showing that in contrast with the case $E = 0$ (that is the case of rate-distortions function) for rate-reliabilities-distortions function decisive role may play the demand of receiver with greatest reliability, although its distortion level may be great.

1 Problem Statement

The purpose of the present work is to develop a generalization of Shannon rate-distortion concept for the robust description system of El Gamal and Cover [7]. The idea is to consider achievability of coding rate R in relation with demands of receivers not only to distortion levels, but also to error probability exponents (reliabilities). Earlier some models of multiple description systems were studied in [1], [3], [4], [7], [8], [10], [11], [14].

We consider the following information transmission system (see Fig. 1.). The messages of a discrete memoryless source encoded by one encoder must be transmitted to K different receivers. Receivers knowing the same codeword try to recover the original message under the condition of different demanded distortions levels and reliabilities.

Let $\{X_i\}_{i=1}^{\infty}$ be the sequence of discrete independent, identically distributed random variables taking values in the finite alphabet $\mathcal{X}$, which is the set of all messages of the source $\{X\}$. The finite sets $\mathcal{X}^k$, $k = \overline{1,K}$, are the reproduction alphabets of corresponding receivers, in general they are different from $\mathcal{X}$.

Let the probability distribution (PD) of the source messages is

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

Since we study the memoryless source, then

$$P^{*N}(\mathbf{x}) = \prod_{n=1}^{N} P^{*N}(x_n).$$

$$d^k : \mathcal{X} \times \mathcal{X}^k \to [0;\infty), \, k = \overline{1,K},$$

distortion measures between source and corresponding reconstruction alphabets. If the $\mathcal{X}$ coincides with the set $\mathcal{X}^k$, $k = \overline{1,K}$, we suppose, that $d^k(x, x^k) = 0$, $k = \overline{1,K}$. The distortion measures for N-sequences are averages of corresponding per letter distortions:

$$d^k(\mathbf{x}, \mathbf{x}^k) = \frac{1}{N}\sum_{n=1}^{N} d^k(x_n, x_n^k), \, k = \overline{1,K}.$$

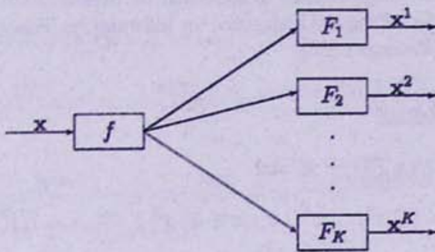

Fig. 1. Source with many receivers.

For the considered system a code $(f, F) = (f, F_1, F_2, ..., F_K)$ is a family of $K+1$ mappings: one encoder:

$$f : (\mathcal{X})^N \to \{1, 2, ..., L(N)\},$$

and K independent decoders:

$$F_k : \{1, 2, ..., L(N)\} \to (\mathcal{X}^k)^N, \, k = \overline{1,K}.$$

Let us consider the following sets "of correct reconstructions":

$$\mathcal{A}_k = \{\mathbf{x} \in (\mathcal{X})^N : F_k(f(\mathbf{x})) = \mathbf{x}^k, \, d^k(\mathbf{x}, \mathbf{x}^k) \le \Delta^k\}, \, k = \overline{1,K}.$$

Error probabilities $e_k(f, F, \Delta^k, N)$ of the code (f, F) at the corresponding receivers are defined as follows:

$$e_k(f, F, \Delta^k, N) = 1 - P^{*N}(\mathcal{A}_k), \, k = \overline{1,K}.$$

For brevity we will write

$$(E^1, ..., E^k) = \mathbf{E}, \quad (\Delta^1, ..., \Delta^K) = \boldsymbol{\Delta}.$$

A number $R \ge 0$ is said to be $(\mathbf{E}, \boldsymbol{\Delta})-$ achievable rate for $E^k > 0$, $\Delta^k \ge 0$, $k = \overline{1,K}$, if for every $\varepsilon > 0$ and sufficiently large N there exists a code (f, F), such that (log-s and exp-s are taken to the base 2)

$$\frac{1}{N}\log L(N) \le R + \varepsilon,$$

and the error probabilities at all receivers are exponentially small with given exponents E^k:

$$e_k(f, F, \Delta^k, N) \le \exp\{-NE^k\}, \, k = \overline{1,K}.$$

Following Shannon we call E^k "reliability" at the k-th decoder, $k = \overline{1,K}$. We call dependence of minimal achievable rate on given $\mathbf{E}$ and Δ rate-reliabilities-distortions function and note it $R(\mathbf{E}, \Delta)$. When $E^k \to 0$, $k = \overline{1,K}$, the function $R(\mathbf{E}, \Delta)$ becomes the corresponding rate-distortions function $R(\Delta)$, which was specified for considered case by El Gamal and Cover in [7].

The result of the paper is formulated in the next section. Proofs are given in section 3. In the Appendix for the case $K = 2$ an example of calculation of the rate-reliabilities-distortions function for binary source and Hamming distortion measures is scrutinized.

The preliminary result was presented in the Scientific Session of Armenian Mathematical Union [12] and at the 13-th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes [13].

2 Formulation of Result

Let $P = \{P(x), x \in \mathcal{X}\}$ be a PD on $\mathcal{X}$ and

$$Q = \{Q(x^1, ..., x^K \mid x),\ x \in \mathcal{X},\ x^k \in \mathcal{X}^k,\ k = \overline{1,K}\}$$

be a conditional PD on $\mathcal{X}^1 \times ... \times \mathcal{X}^K$ for a given x.

We also need

$$Q(x^k \mid x) = \sum_{x^j \in \mathcal{X}^j,\ j=\overline{1,K},\ j \neq k} Q(x^1, ..., x^K \mid x),\ k = \overline{1,K}.$$

Let

$$D(P \parallel P^*) = \sum_x P(x) \log \frac{P(x)}{P^*(x)}$$

be divergence of PD P and P^*.

Consider the sets:

$$\alpha\left(E^k\right) = \{P : D(P \parallel P^*) \leq E^k\},\ k = \overline{1,K}.$$

Since the system is symmetrical, we assume without loss of generality that

$$0 < E^1 \leq E^2 \leq ... \leq E^k.$$

Denote by $\Phi(P, E^k, \Delta) = Q_P^k$ the function, which puts into the correspondence to PD P some conditional PD Q_P^k such that for given Δ, the following conditions take place:
for $\Phi(P, E^1, \Delta) = Q_P^1$ if $P \in \alpha(E^1)$, then

$$\mathbf{E}_{P,Q_P^1} d^k(X, X^k) = \sum_{x, x^k} P(x) Q_P^1(x^k \mid x) d^k(x, x^k) \leq \Delta^k,\ k = \overline{1,K}, \tag{1}$$

and for $\Phi(P, E^j, \Delta) = Q_P^j$ if $P \in \alpha(E^j) - \alpha(E^{j-1})$, $j = \overline{2,K}$, then

$$\mathbf{E}_{P,Q_P^j} d^k(X, X^k) = \sum_{x, x^k} P(x) Q_P^j(x^k \mid x) d^k(x, x^k) \leq \Delta^k,\ k = \overline{j,K}. \tag{2}$$

Denote by $\mathcal{M}(P, E^k, \Delta)$ the set of all such functions $\Phi(P, E^k, \Delta)$ for given Δ, P and E^k, $k = \overline{1,K}$, and by $\mathcal{M}(\Delta)$ the set of all functions $\Phi(P^*, \Delta)$ for the case, when $E^k \to 0$ for all

$k=\overline{1,K}$. Below for brevity we shall write simply $\Phi(P,E^k)$ and $\Phi(P^*)$. We use the following notations for entropy and informations:

$$H_P(X) = -\sum_x P(x)\log P(x),$$

$$I_{P,Q_P}(X\wedge X^1,...,X^K) = \sum_{x,x^1,...,x^K} P(x)Q_P(x^1,...,x^K \mid x)\log\frac{Q_P(x^1,...,x^K\mid x)}{\sum_x P(x)Q_P(x^1,...,x^K\mid x)}.$$

Let us introduce a function

$$R^*(\mathbf{E},\Delta) = \max[\max_{P\in\alpha(E^1)}\min_{\Phi(P,E^1)\in\mathcal{M}(P,E^1,\Delta)} I_{P,\Phi(P,E^1)}(X\wedge X^1,...,X^K),$$

$$\max_{P\in\alpha(E^2)-\alpha(E^1)}\min_{\Phi(P,E^2)\in\mathcal{M}(P,E^2,\Delta)} I_{P,\Phi(P,E^2)}(X\wedge X^2,...,X^K),$$

$$\cdots$$

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

Theorem: For $E^k>0$, $\Delta^k\geq 0$, $k=\overline{1,K}$,

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

Corollary: When $E^k=E$, $k=\overline{1,K}$, we receive

$$R(E,\Delta) = \max_{P\in\alpha(E)}\min_{\Phi(P,E)\in\mathcal{M}(P,E,\Delta)} I_{P,\Phi(P,E)}(X\wedge X^1,...,X^K).$$

When $E^k\to 0$, $k=\overline{1,K}$, we arrive to the result of El Gamal and Cover [7] on rate-distortions function:

$$R(\Delta) = \min_{\Phi(P^*)\in\mathcal{M}(\Delta)} I_{P^*,\Phi(P^*)}(X\wedge X^1,...,X^K).$$

Remark: We shall show by an example that in contradistinction to the case of calculation of the rate-distortions function (where the decisive role has the rate demanded by the receiver with smallest distortion level) for the rate-reliabilities-distortions function decisive role may play the rate of transmission to the receiver, which requires the greatest reliability.

3 Proof of the Theorem

We apply the typical sequences technique [5], [6] and the following random coding lemma, which is a modification of covering lemmas from [2], [6], [8], [11].

Lemma: Let for $\varepsilon>0$,

$$J_k(P,Q) = \exp\{N(I_{P,Q}(X\wedge X^k,...,X^K)+\varepsilon)\},\ k=\overline{1,K}.$$

Then for every type P and conditional distribution Q there exist the collections of vectors

$$\{(\mathbf{x}^k_{j_k},...,\mathbf{x}^K_{j_k})\in\mathcal{T}_{P,Q}(X^k,...,X^K),\ j_k=\overline{1,J_k(P,Q)}\},\ k=\overline{1,K},$$

such that for N large enough the sets

$$\{\mathcal{T}_{P,Q}(X\mid \mathbf{x}^k_{j_k},...,\mathbf{x}^K_{j_k}),\ j=\overline{1,J_k(P,Q)}\},\ k=\overline{1,K},$$

are coverings for $\mathcal{T}_P(X)$.

The proof of the theorem we begin with the inequality

$$R^*(\mathbf{E},\Delta) \geq R(\mathbf{E},\Delta). \tag{3}$$

Let us denote by $\mathcal{P}(X,N)$ the set of all types P. We can present $(\mathcal{X})^N$ as a union of all disjoint types of vectors:

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

For given $\delta > 0$ and every $k = \overline{1,K}$,

$$P^{*N}\Big(\bigcup_{P\notin\alpha(E^k+\delta)} \mathcal{T}_P(X)\Big) = \sum_{P\notin\alpha(E^k+\delta)} P^{*N}(\mathcal{T}_P(X)) \leq$$

$$\leq (N+1)^{|\mathcal{X}|}\exp\{-N \min_{P\notin\alpha(E^k+\delta)} D(P \parallel P^*)\} \leq$$

$$\leq \exp\{-NE^k - N\delta + |\mathcal{X}|\log(N+1)\} \leq \exp\{-N(E^k + \frac{\delta}{2})\}$$

for N large enough.

Consequently, to obtain the desired levels of error probabilities, it is sufficient to construct an errorless encoding procedure only for vectors of the types P from $\alpha(E^K+\delta)$.

Let us fix some type $P \in \alpha\left(E^K+\delta\right)$. If $P \in \alpha(E^1+\delta)$, let us fix some $\Phi(P,E^1) \in \mathcal{M}(P,E^1,\Delta)$. If $P \in \alpha\left(E^k+\delta\right) - \alpha\left(E^{k-1}+\delta\right)$, where $2 \leq k \leq K$, let us fix some $\Phi(P,E^k) \in \mathcal{M}(P,E^k,\Delta)$. Denote $\Phi(P,E^k) = Q_P^k$.

According to the lemma for any $k = \overline{1,K}$, there exists a covering

$$\{\mathcal{T}_{P,Q_P^k}(X \mid \mathbf{x}_{j_k}^k,\dots,\mathbf{x}_{j_k}^K),\ j_k = \overline{1, J_k(P,Q_P^k)}\}$$

for $\mathcal{T}_P(X)$. Let us consider

$$\mathcal{C}(P,Q_P^k,j_k) =$$

$$= \mathcal{T}_{P,Q_P^k}(X \mid \mathbf{x}_{j_k}^k,\dots,\mathbf{x}_{j_k}^K) - \bigcup_{j_k'<j_k} \mathcal{T}_{P,Q_P^k}(X \mid \mathbf{x}_{j_k'}^k,\dots,\mathbf{x}_{j_k'}^K),\ j_k = \overline{1, J_k(P,Q_P^k)},\ k = \overline{1,K}.$$

We define a code $(f,F) = (f,F_1,\dots,F_K)$ as follows:
encoding function:

$$f(\mathbf{x}) = \begin{cases} j_1, & \text{when } \mathbf{x} \in \mathcal{C}(P,Q_P^1,j_1),\ P \in \alpha(E^1+\delta), \\ j_k, & \text{when } \mathbf{x} \in \mathcal{C}(P,Q_P^k,j_k),\ P \in \alpha(E^k+\delta) - \alpha(E^{k-1}+\delta),\ k = \overline{2,K}, \\ j_0, & \text{when } \mathbf{x} \in \mathcal{T}_P(X),\ P \notin \alpha(E^K+\delta), \end{cases}$$

and decoding functions:

$$F_k(j_s) = \begin{cases} \mathbf{x}_{j_k}^k, & \text{for } s = k, \\ \overline{\mathbf{x}}, & \text{for } s \neq k, \end{cases}$$

$$F_k(j_0) = \overline{\mathbf{x}},\ k = \overline{1, K}.$$

According to the definition of the code (f, F), to the lemma and the inequalities (1) and (2) we have for $P \in \alpha(E^k + \delta)$

$$d^j(\mathbf{x}, \mathbf{x}^j) = N^{-1} \sum_{x, x^j} n(x, x^j \mid \mathbf{x}, \mathbf{x}^j) d^j(x, x^j) =$$

$$= \sum_{x, x^j} P(x) Q_P^k(x^j \mid x) d^j(x, x^j) = \mathbf{E}_{P, Q_P^k} d^j(X, X^j) \leq \Delta^j,\ j = \overline{k, K}.$$

For fixed type P and its conditional type $\Phi(P, E^k)$ the number of symbols used in encoding (denote it by $L_{P, \Phi(P, E^k)}(N)$) is

$$L_{P, \Phi(P, E^k)}(N) = \exp\{N(I_{P, \Phi(P, E^k)}(X \wedge X^k, ..., X^K) + \varepsilon)\}.$$

The "worst" types among $\alpha(E^k + \delta)$, (which number have polynomial estimate) and their optimal (among $\mathcal{M}(P, E^k, \Delta)$) conditional distributions $\Phi(P, E^k)$, $k = \overline{1, K}$, determine corresponding bound for transmission rate:

$$\frac{1}{N} \log L(N) - \varepsilon - \frac{|\mathcal{X}| \log(N + 1)}{N} \leq$$

$$\leq \max[\max_{P \in \alpha(E^1 + \delta)} \min_{\Phi(P, E^1) \in \mathcal{M}(P, E^1, \Delta)} I_{P, \Phi(P, E^1)}(X \wedge X^1, ..., X^K),$$

$$\max_{P \in \alpha(E^2 + \delta) - \alpha(E^1 + \delta)} \min_{\Phi(P, E^2) \in \mathcal{M}(P, E^2, \Delta)} I_{P, \Phi(P, E^2)}(X \wedge X^2, ..., X^K), \tag{4}$$

$$\cdots$$

$$\max_{P \in \alpha(E^K + \delta) - \alpha(E^{K-1} + \delta)} \min_{\Phi(P, E^K) \in \mathcal{M}(P, E^K, \Delta)} I_{P, \Phi(P, E^K)}(X \wedge X^K)].$$

Taking into account arbitrariness of ε and δ, continuity of expressions of information in (4) with respect to E^k, $k = \overline{1, K}$, we receive (3).

Now we shall prove the inequality

$$R^*(\mathbf{E}, \Delta) \leq R(\mathbf{E}, \Delta). \tag{5}$$

Let a given code (f, F) of blocklength N has $(\mathbf{E}, \Delta)$-achievable rate R. It is enough to show that for some $\Phi(P, E^k) = Q_P^k(x^1, ..., x^K \mid x) \in \mathcal{M}(P, E^k, \Delta)$, $k = \overline{1, K}$, for N large enough the following inequality takes place

$$\frac{1}{N} \log L(N) \geq \max[\max_{P \in \alpha(E^1)} I_{P, \Phi(P, E^1)}(X \wedge X^1, ..., X^K),$$

$$\max_{P \in \alpha(E^2) - \alpha(E^1)} I_{P, \Phi(P, E^2)}(X \wedge X^2, ..., X^K), \tag{6}$$

$$\cdots$$

$$\max_{P \in \alpha(E^K) - \alpha(E^{K-1})} I_{P, \Phi(P, E^K)}(X \wedge X^K)].$$

We can write:

$$\left|\bigcap_{k=1}^{K} \mathcal{A}_k \bigcap \mathcal{T}_P(X)\right| = |\mathcal{T}_P(X)| - \left|\bigcup_{k=1}^{K} \overline{\mathcal{A}_k} \bigcap \mathcal{T}_P(X)\right|.$$

Let $\varepsilon > 0$ be fixed. For $P \in \alpha(E^1 - \varepsilon)$ the following estimates are valid:

$$\left|\bigcup_{k=1}^{K} \overline{\mathcal{A}_k} \bigcap \mathcal{T}_P(X)\right| = \frac{P^{*N}(\bigcup_{k=1}^{K} \overline{\mathcal{A}_k} \cap \mathcal{T}_P(X))}{P^{*N}(\mathbf{x})} \leq$$

$$\leq \exp\{N(H_P(X) + D(P \parallel P^*))\} \sum_{k=1}^{K} \exp\{-NE^k\} \leq$$

$$\leq K \exp\{-NE^1\} \cdot \exp\{N(H_P(X) + E^1 - \varepsilon)\} =$$

$$= \exp\{N(H_P(X) + \frac{\log K}{N} - \varepsilon)\} \leq \exp\{N(H_P(X) - \varepsilon/2)\}$$

for N large enough.

Hence

$$\begin{aligned}\left|\bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)\right| &\geq (N+1)^{-|\mathcal{X}|} \exp\{NH_P(X)\} - \exp\{N(H_P(X) - \varepsilon/2)\} = \\ &= \exp\{N(H_P(X) - \varepsilon/2)\} \left(\frac{\exp\{N\varepsilon/2\}}{(N+1)^{|\mathcal{X}|}} - 1\right) \geq \exp\{N(H_P(X) - \varepsilon/2)\}\end{aligned} \tag{7}$$

for N large enough.

To each $\mathbf{x} \in \bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)$ a unique collection of vectors $(\mathbf{x}^1, ..., \mathbf{x}^K)$ corresponds, such that $\mathbf{x}^k = F_k(f(\mathbf{x}))$, $k = \overline{1,K}$. This collection of K vectors determines a conditional type Q, for which

$$(\mathbf{x}^1, ..., \mathbf{x}^K) \in \mathcal{T}_{P,Q}(X^1, ..., X^K \mid \mathbf{x}).$$

Since $\mathbf{x} \in \bigcap_{k=1}^{K} \mathcal{A}_k$, then $\mathbf{E}_{P,Q} d^k(X, X^k) = d^k(\mathbf{x}, \mathbf{x}^k) \leq \Delta^k$, $k = \overline{1,K}$. So, $Q \in M(P, E^1, \Delta)$.

The set of all vectors $\mathbf{x} \in \bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)$ is divided into classes corresponding these conditional types Q. Let us select from them the class (let it be $Q_P^1 = \Phi(P, E^1)$), which for given P contains the greatest number of $\mathbf{x}$, and denote it by $\left(\bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)\right)(\Phi(P, E^1))$. Using the polynomial upper estimate [5], [6] for the number of conditional types Q, we have

$$\begin{aligned}&\left|\bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)\right| \leq \\ &\leq (N+1)^{|\mathcal{X}| \prod_{k=1}^{K} |\mathcal{X}^k|} \left|\left(\bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)\right)(\Phi(P, E^1))\right| \leq \\ &\leq \exp\{N\varepsilon/2\} \left|\left(\bigcap_{k=1}^{K} \mathcal{A}_k \cap \mathcal{T}_P(X)\right)(\Phi(P, E^1))\right|\end{aligned} \tag{8}$$

for N large enough.

Let $\mathcal{D}_{1,\dots,K}$ be the set of all $(\mathbf{x}^1,\dots,\mathbf{x}^K)$, which satisfy $F_k(f(\mathbf{x}))=\mathbf{x}^k$, $k=\overline{1,K}$, for some $\mathbf{x}\in\bigcap_{k=1}^{K}\mathcal{A}_k\cap\mathcal{T}_P(X)$, $\mathbf{x}\in\mathcal{T}_{P,\Phi(P,E^1)}(X\mid\mathbf{x}^1,\dots,\mathbf{x}^K)$. According to the definition of the code remark that $|\mathcal{D}_{1,\dots,K}|\le L(N)$. Then

$$\left|\left(\bigcap_{k=1}^{K}(\mathcal{A}_k\bigcap\mathcal{T}_P(X)\right)(\Phi(P,E^1))\right|\le$$

$$\le\sum_{(\mathbf{x}^1,\dots,\mathbf{x}^K)\in\mathcal{D}_{1,\dots,K}}\left|\mathcal{T}_{P,\Phi(P,E^1)}(X\mid\mathbf{x}^1,\dots,\mathbf{x}^K)\right|\le L(N)\exp\{NH_{P,\Phi(P,E^1)}(X\mid X^1,\dots,X^K)\}.$$

From the last inequality, (7) and (8) we receive that for $P\in\alpha(E^1-\varepsilon)$

$$L(N)\ge\exp\{N(I_{P,\Phi(P,E^1)}(X\wedge X^1,\dots,X^K)-\varepsilon)\}.$$

Hence

$$\frac{1}{N}\log L(N)\ge\max_{P\in\alpha(E^1-\varepsilon)}I_{P,\Phi(P,E^1)}(X\wedge X^1,\dots,X^K)-\varepsilon.$$

Similarly we can show that for N large enough and $P\in\alpha(E^k-\varepsilon)-\alpha(E^{k-1}-\varepsilon)$, for $k=\overline{2,K}$, the following inequality takes place

$$\left|\bigcap_{j=k}^{K}\mathcal{A}_j\bigcap\mathcal{T}_P(X)\right|\ge\exp\{N(H_P(X)-\varepsilon/2)\}.\tag{9}$$

By analogy with the selection of the class in (8), for each $k=\overline{2,K}$ we can choose the classes

$$\left(\bigcap_{j=k}^{K}\mathcal{A}_j\bigcap\mathcal{T}_P(X)\right)(\Phi(P,E^k)),\ \ k=\overline{2,K}.$$

Then for any $k=\overline{2,K}$ for N large enough

$$\left|\bigcap_{j=k}^{K}\mathcal{A}_j\cap\mathcal{T}_P(X)\right|\le$$

$$\le(N+1)^{|\mathcal{X}|\prod_{j=k}^{K}|\mathcal{X}^j|}\left|\left(\bigcap_{j=k}^{K}\mathcal{A}_j\cap\mathcal{T}_P(X)\right)(\Phi(P,E^k))\right|\le\tag{10}$$

$$\le\exp\{N\varepsilon/2\}\left|\left(\bigcap_{j=k}^{K}\mathcal{A}_j\cap\mathcal{T}_P(X)\right)(\Phi(P,E^k))\right|.$$

For any $k=\overline{2,K}$ let us denote by $\mathcal{D}_{k,\dots,K}$ the set of all $(\mathbf{x}^k,\dots,\mathbf{x}^K)$, for which there exists $\mathbf{x}\in\bigcap_{j=k}^{K}\mathcal{A}_j\cap\mathcal{T}_P(X)$, $\mathbf{x}\in\mathcal{T}_{P,\Phi(P,E^k)}(X\mid\mathbf{x}^k,\dots,\mathbf{x}^K)$, such that $F_j(f(\mathbf{x}))=\mathbf{x}^j$, $j=\overline{k,K}$. Remark that $|\mathcal{D}_{k,\dots,K}|\le L(N)$. Therefore

$$\left|\left(\bigcap_{j=k}^{K}\mathcal{A}_j\bigcap\mathcal{T}_P(X)\right)(\Phi(P,E^k))\right|\le$$

$$\leq \sum_{(\mathbf{x}^k,\dots,\mathbf{x}^K)\in\mathcal{D}_{k,\dots,K}} \left|\mathcal{T}_{P,\Phi(P,E^k)}(X \mid \mathbf{x}^k,\dots,\mathbf{x}^K)\right| \leq L(N)\exp\{NH_{P,\Phi(P,E^k)}(X \mid X^k,\dots,X^K\}.$$

Taking into account (7), (10) and the last inequality for $P \in \alpha(E^k-\varepsilon)-\alpha(E^{k-1}-\varepsilon)$, $k=\overline{2,K}$, we receive that

$$L(N) \geq \exp\{N(I_{P,\Phi(P,E^k)}(X \wedge X^k,\dots,X^K)-\varepsilon)\}.$$

Therefore

$$\frac{1}{N}\log L(N) \geq \max_{P\in\alpha(E^k-\varepsilon)-\alpha(E^{k-1}-\varepsilon)} I_{P,\Phi(P,E^k)}(X \wedge X^k,\dots,X^K)-\varepsilon.$$

Taking into account arbitrariness of ε, continuity by E^k, $k=\overline{1,K}$ of all functions in above expressions, we complete the proof of the inclusion (6), hence of (5) too.

3.1 Example

Let us consider a binary source, characterized by the alphabets $\mathcal{X}=\mathcal{X}^1=\mathcal{X}^2=\{0,1\}$, with generic probability distribution $P^*=(p^*,1-p^*)$ and Hamming distortion measures

$$d^1(x,x^1)=\begin{cases}0, & x=x^1,\\ 1, & x\neq x^1,\end{cases} \qquad d^2(x,x^2)=\begin{cases}0, & x=x^2,\\ 1, & x\neq x^2.\end{cases}$$

Denote by $R_{\mathrm{BH}}(\mathbf{E},\mathbf{\Delta})$ the binary Hamming rate-reliabilities-distortions function, by $R_{\mathrm{BH}}(\mathbf{\Delta})$- the binary Hamming rate-distortions function. Let us assume that $\Delta^1<\Delta^2$ and $E^1<E^2$. By analogy with the calculation of the binary Hamming rate-distortion function (see [5]) it is not difficult to show that for this system the binary Hamming rate-distortions function is determined by the demands of the output with smallest distortion level:

$$R_{\mathrm{BH}}(\mathbf{\Delta})=\begin{cases}H_{P^*}(X)-H(\Delta^1), & \text{if } \Delta^1\leq\min(p^*,1-p^*)\\ 0, & \text{if } \Delta^1>\min(p^*,1-p^*),\end{cases}$$

where $H(\Delta^1)=-\Delta^1\log\Delta^1-(1-\Delta^1)\log(1-\Delta^1)$.

From the theorem we have

$$R_{\mathrm{BH}}(\mathbf{E},\mathbf{\Delta})=\max[\max_{P\in\alpha(E^1)}\min_{\Phi(P,E^1)\in\mathcal{M}(P,E^1,\mathbf{\Delta})} I_{P,\Phi(P,E^1)}(X\wedge X^1,X^2),$$

$$\max_{P\in\alpha(E^2)-\alpha(E^1)}\min_{\Phi(P,E^2)\in\mathcal{M}(P,E^2,\mathbf{\Delta})} I_{P,\Phi(P,E^2)}(X\wedge X^2)].$$

Similarly to the calculation of the binary Hamming rate-reliability-distortion function (see [9], [11]) we can show the following equalities.

For E^1 such that $(1/2,1/2)\in\alpha(E^1)$

$$\max_{P\in\alpha(E^1)}\min_{\Phi(P,E^1)\in\mathcal{M}(P,E^1,\mathbf{\Delta})} I_{P,\Phi(P,E^1)}(X\wedge X^1,X^2)=\begin{cases}1-H(\Delta^1), & \text{when } \Delta^1\leq 1/2,\\ 0, & \text{when } \Delta^1>1/2.\end{cases}$$

For E^1 such that $(1/2,1/2)\notin\alpha(E^1)$

$$\max_{P\in\alpha(E^1)}\min_{\Phi(P,E^1)\in\mathcal{M}(P,E^1,\mathbf{\Delta})} I_{P,\Phi(P,E^1)}(X\wedge X^1,X^2)=\begin{cases}H_{P_{E^1}}(X)-H(\Delta^1), & \text{when } \Delta^1\leq p_{E^1},\\ 0, & \text{when } \Delta^1>p_{E^1}.\end{cases}$$

For E^2 such that $(1/2,1/2) \in \alpha(E^2) - \alpha(E^1)$

$$\max_{P\in\alpha(E^2)-\alpha(E^1)} \min_{\Phi(P,E^2)\in\mathcal{M}(P,E^2,\Delta)} I_{P,\Phi(P,E^2)}(X \wedge X^2) = \begin{cases} 1-H(\Delta^2), & \text{when } \Delta^2 \le 1/2, \\ 0, & \text{when } \Delta^2 > 1/2. \end{cases}$$

For E^2 such that $(1/2,1/2) \notin \alpha(E^2) - \alpha(E^1)$

$$\max_{P\in\alpha(E^2)-\alpha(E^1)} \min_{\Phi(P,E^2)\in\mathcal{M}(P,E^2,\Delta)} I_{P,\Phi(P,E^2)}(X \wedge X^2) = \begin{cases} H_{P_{E^2}}(X)-H(\Delta^2), & \text{when } \Delta^2 \le p_{E^2}, \\ 0, & \text{when } \Delta^2 > p_{E^2}. \end{cases}$$

In above expressions $P_{E^i} = (p_{E^i}, 1-p_{E^i})$ and p_{E^i} is the nearest to $\frac{1}{2}$ solution of the equation

$$D(P_{E^i} \parallel P^*) = E^i, \ i = 1,2.$$

It is not difficult to show that $R_{\text{BH}}(\mathbf{E},\Delta)$ as a function of E^2 when $(1/2,1/2) \in \alpha(E^1)$ is constant and equals to its possible maximal value $1 - H(\Delta^1)$.

Let us assume that

$$\Delta^1 < \min(p^*, 1-p^*). \tag{11}$$

Then $R_{\text{BH}}(\Delta) = H_{P^*}(X) - H(\Delta^1)$.

Let us consider the following two cases for fixed Δ^1, satisfying (11) and E^1 such that $(1/2,1/2) \notin \alpha(E^1)$, $\Delta^1 < p_{E^1}$.

For E^2 such that

$$(1/2,1/2) \in \alpha\left(E^2\right), \ \Delta^2 < \frac{1}{2}, \tag{12}$$

we receive that

$$R_{\text{BH}}(\mathbf{E},\Delta) = \max[H_{P_{E^1}}(X) - H(\Delta^1), 1 - H(\Delta^2)],$$

and when

$$H_{P_{E^1}}(X) - H(\Delta^1) < 1 - H(\Delta^2), \tag{13}$$

then $R_{\text{BH}}(\mathbf{E},\Delta) = 1 - H(\Delta^2)$.

For E^2 such that

$$(1/2,1/2) \notin \alpha\left(E^2\right), \ \Delta^2 < p_{E^2}, \tag{14}$$

we obtain that

$$R_{\text{BH}}(\mathbf{E},\Delta) = \max[H_{P_{E^1}}(X) - H(\Delta^1), H_{P_{E^2}}(X) - H(\Delta^2)],$$

and when

$$H_{P_{E^1}}(X) - H(\Delta^1) < H_{P_{E^2}}(X) - H(\Delta^2), \tag{15}$$

then $R_{\text{BH}}(\mathbf{E},\Delta) = H_{P_{E^2}}(X) - H(\Delta^2)$.

Remark that $R_{\text{BH}}(\mathbf{E},\Delta)$ as a function of E^2 when $(1/2,1/2) \notin \alpha(E^1)$ and (13) does not takes place is constant and equals to $H_{P_{E^1}}(X) - H(\Delta^1)$.

$R_{\text{BH}}(\mathbf{E},\Delta)$ as a function of E^2 when $(1/2,1/2) \notin \alpha(E^1)$ and (13) is valid, is presented in Fig. 2.

In the Table I some values of parameters P^*, Δ^1, Δ^2, E^1, E^2 meeting the condition (12) are choosen such, that inequality (13) takes place. In this table the calculated values for p_{E^1}, $H_{P_{E^1}}(X) - H(\Delta^1)$, $1 - H(\Delta^2)$, $R_{\text{BH}}(\Delta)$ and $R_{\text{BH}}(\mathbf{E},\Delta)$ are given.

In the Table II some variants of values of parameters P^*, Δ^1, Δ^2, E^1, E^2 meeting the condition (14) are choosen so, that (15) takes place. In this table the calculated values for p_{E^1}, $H_{P_{E^1}}(X) - H(\Delta^1)$, $H_{P_{E^1}}(X) - H(\Delta^2)$, $R_{\mathrm{BH}}(\boldsymbol{\Delta})$ and $R_{\mathrm{BH}}(\mathbf{E}, \boldsymbol{\Delta})$ are presented.

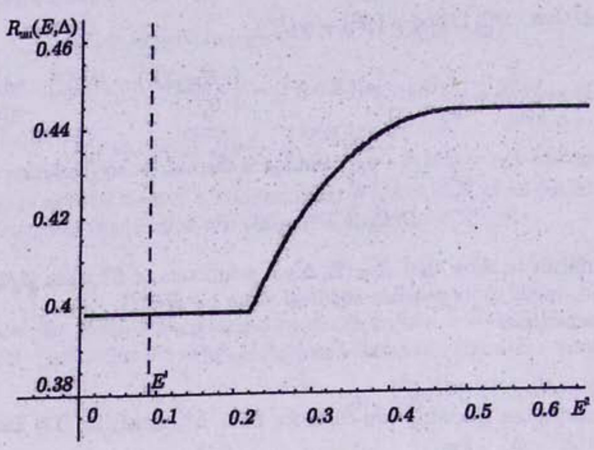

Fig. 2. $R_{\mathrm{BH}}(\mathbf{E}, \boldsymbol{\Delta})$ for $P^* = (0.15, 0.85)$, $\Delta^1 = 0.1$, $\Delta^2 = 0.13$, $E^1 = 0.09$.

Parameters	Variants of values		
p^*	0.15	0.2	0.25
Δ^1	0.1	0.16	0.2
Δ^2	0.13	0.2	0.22
E^1	0.09	0.05	0.05
E^2	0.49	0.4	0.24
	Corresponding calculated values		
p_{E^1}	0.288782	0.311464	0.369015
$H_{P_{E^1}}(X) - H(\Delta^1)$	0.398147	0.260547	0.227985
$1 - H(\Delta^2)$	0.442562	0.278072	0.239832
$R_{\mathrm{BH}}(\boldsymbol{\Delta})$	0.140845	0.0876185	0.08935
$R_{\mathrm{BH}}(\mathbf{E}, \boldsymbol{\Delta})$	0.442562	0.278072	0.239832

Table I. Some variants of parameters P^*, Δ^1, Δ^2, E^1, E^2 and calculated values of p_{E^1}, $H_{P_{E^1}}(X) - H(\Delta^1)$, $1 - H(\Delta^2)$, $R_{\mathrm{BH}}(\boldsymbol{\Delta})$ and $R_{\mathrm{BH}}(\mathbf{E}, \boldsymbol{\Delta})$ in the case (12).

That is the binary Hamming rate-reliabilities-distortions function $R_{\mathrm{BH}}(\mathbf{E}, \boldsymbol{\Delta})$ in the case a) when (13) and in the case b) when (15) take place is determined by the demands of receiver with greatest distortion level, although the binary Hamming rate-distortions function $R_{\mathrm{BH}}(\boldsymbol{\Delta})$ is determined by the demands of the receiver with smallest distortion level.

Parameters	Variants of values		
p^*	0.15	0.2	0.25
Δ^1	0.13	0.13	0.22
Δ^2	0.14	0.14	0.23
E^1	0.08	0.12	0.07
E^2	0.13	0.2	0.13
	Corresponding calculated values		
p_{E^1}	0.28025	0.376975	0.391722
p_{E^2}	0.319374	0.432429	0.445716
$H_{P_{E^1}}(X) - H(\Delta^1)$	0.298353	0.398439	0.205734
$H_{P_{E^2}}(X) - H(\Delta^2)$	0.319461	0.402547	0.213469
$R_{\mathrm{BH}}(\mathbf{\Delta})$	0.0524021	0.16449	0.0511106
$R_{\mathrm{BH}}(\mathbf{E}, \mathbf{\Delta})$	0.319461	0.402547	0.213469

Table II. Some variants of parameters P^*, Δ^1, Δ^2, E^1, E^2 and calculated values of p_{E^1}, p_{E^2}, $H_{P_{E^1}}(X) - H(\Delta^1)$, $H_{P_{E^2}}(X) - H(\Delta^2)$, $R_{\mathrm{BH}}(\mathbf{\Delta})$ and $R_{\mathrm{BH}}(\mathbf{E}, \mathbf{\Delta})$ in the case (14).

References

[1] 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.

[2] 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.

[3] T. Berger and Z. Zhang, "New results in binary multiple descriptions", IEEE Trans. on Inform. Theory, vol. 33, no. 4, pp. 502-521, July 1987.

[4] T. Berger and Z. Zhang, "Multiple description source coding with no excess marginal rate", IEEE Trans. on Inform. Theory, vol. 41, no. 2, pp. 349-357, March 1995.

[5] T. M. Cover and J. A. Thomas, Elements of Information Theory, New York: Wiley, 1991.

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

[7] 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.

[8] 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.

[9] E. A. Haroutunian and A. N. Haroutunian, "The binary Hamming rate-reliability, distortion function", Transactions of the Institute for Informatics and Automation Problems of the NAS of RA and of the Yerevan State University, Mathematical problems of computer science, vol. XVIII, pp. 40-45, 1997.

[10] A. N. Haroutunian and E. A. Haroutunian, "An achievable rates-reliabilities-distortions dependence for source coding with three descriptions", Transactions of the Institute for Informatics and Automation Problems of the NAS of RA and of the Yerevan State University, Mathematical problems of computer science, vol. XVII, pp. 70-75, 1997.

[11] A. N. Haroutunian, Investigation of achievable interdependence between coding rates and reliability for several classes of sources, (in Russian), Kandidat of Science Dissertation, Institute for Informatics and Automation Problems of the NAS of RA and of the YSU, Yerevan, Nov. 1997.

[12] A. N. Haroutunian and A. R. Kazarian, "On optimal rate of source robust description", Abstracts of the Scientific Session of Armenian Mathematical Union, pp. 5-6, Yerevan, 1997.

[13] E. A. Haroutunian, A. N. Haroutunian and A. R. Kazarian, "On rate-reliabilities-distortions function of source with many receivers", Proceedings of the Joint Session of 6-th Prague Symposium on Asymptotic Statistics and 13-th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, vol. 1, pp. 217-220, Prague, 1998.

[14] R. Sh. Maroutian, "Achievable rates for multiple descriptions with given exponent and distortion levels", Problemy peredachi informatsii, vol. 26, no. 1, (in Russian), 1990.