Mathematical Problems of Computer Science 22, 2001, 31–43.

Derivation by Graph Decomposition of the Random Coding and Expurgated Bounds for E-capacity

Evgueni A. Haroutunian

Institute for Informatics and Automation Problems of NAS RA and YSU
E-mail: evhar@ipia.sci.am

Abstract

Random coding and expurgated bounds obtaining by method of graph decomposition for E-capacity of discrete memoryless channel (DMC) is presented. Three decoding rules are considered, the random coding bound is attainable by each of three rules, but the expurgated bound is achievable only by maximum-likelihood decoding.

1 Introduction

Let $\mathcal{X}, \mathcal{Y}$ be finite sets and $W = \{W(y|x), x \in \mathcal{X}, y \in \mathcal{Y}\}$ be a stochastic matrix.

Definition 1: *A discrete channel W with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by stochastic matrix of transition probabilities*

$$W : \mathcal{X} \to \mathcal{Y}.$$

An element $W(y|x)$ of the matrix is a conditional probability of receiving the symbol $y \in \mathcal{Y}$ on the channel's output if the symbol $x \in \mathcal{X}$ was transmitted from the input.

The model for N actions of the channel W is described by the stochastic matrix

$$W^N : \mathcal{X}^N \to \mathcal{Y}^N,$$

the element of which $W^N(\mathbf{y}|\mathbf{x})$ is a conditional probability of receiving vector $\mathbf{y} \in \mathcal{Y}^N$, when $\mathbf{x} \in \mathcal{X}^N$ was transmitted. Here we consider memoryless channels, which operates in each moment of time independently of the previous or next transmitted or received symbols, so for all $\mathbf{x} \in \mathcal{X}^N$ and $\mathbf{y} \in \mathcal{Y}^N$

$$W^N(\mathbf{y}|\mathbf{x}) = \prod_{n=1}^{N} W(y_n|x_n). \tag{1}$$

Let $\mathcal{M}$ denotes the set of messages to be transmitted and M – the number of messages.

Definition 2: *A code (f, g) for the channel W is a pair of mappings, where $f : \mathcal{M} \to \mathcal{X}^N$ is encoding and $g : \mathcal{Y}^N \to \mathcal{M}$ is decoding. N is called code length, and M is called code volume.*

Definition 3: *The probability of erroneous transmission of the message $m \in \mathcal{M}$ by the channel using code (f,g) is defined as*

$$e(m,f,g,N,W) \triangleq W^N(\mathcal{Y}^N - g^{-1}(m)|f(m)) = \\ = 1 - W^N(g^{-1}(m)|f(m)). \tag{2}$$

We shall consider the maximal probability of error of the code (f,g):

$$e(f,g,N,W) \triangleq \max_{m\in\mathcal{M}} e(m,f,g,N,W), \tag{3}$$

$$e(M,N,W) \triangleq \min_{(f,g)} e(f,g,N,W),$$

where minimum is taken among codes (f,g) of volume M.

Definition 4: *The transmission rate of a code (f,g) of volume M is*

$$R(f,g,N) \triangleq \frac{1}{N}\log M. \tag{4}$$

Note that in this paper all exp-s and log-s are to the base two.

We consider the codes, error probability of which exponentially decrease with given exponent E:

$$e(f,g,N,W) \leq \exp\{-NE\}. \tag{5}$$

Denote the best volume of the code of length N for channel W satisfying the condition (5) for given reliability $E > 0$ by $M(E,N,W)$.

Definition 5: *The rate-reliability function, which by analogy with the capacity we call* **E-capacity,** *is*

$$R(E,W) = C(E,W) \triangleq \overline{\lim_{N\to\infty}} \frac{1}{N}\log M(E,N,W). \tag{6}$$

The concept of E-capacity was first considered by author in [1], there was presented derivation of the upper bound $R_{sp}(E,W)$. The simple combinatorial proof of $R_{sp}(E,W)$ was obtained in [2].

Alternative methods for existence part of coding theorems demonstration are Shannon's random coding and Wolfowitz's maximal code methods. In [3] Csiszár and Körner introduced a new original method, based on the lemma of Lovász on graph decomposition. Different methods of error exponent investigation were presented in [4]–[9] and in many other works. Here we shall derive upper bounds for $R(E,W)$ using method of graph decomposition.

2 Formulation of results

For beginning we remind our notations for necessary characteristics of Shannon's entropy and mutual information and Kullback-Leibler's divergence.

The size of the set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. Let P be a PD of RV X

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

V be a conditional PD of RV Y for given value x of RV Y

$$V = \{V(y|x), x \in \mathcal{X}, y \in \mathcal{Y}\}.$$

The joint PD of RV X and Y is

$$P \circ V = \{P \circ V(x,y) = P(x)V(y|x), x \in \mathcal{X}, y \in \mathcal{Y}\},$$

and PD of RV Y is

$$PV = \{PV(y) = \sum_{x \in \mathcal{X}} P(x)V(y|x), y \in \mathcal{Y}\}.$$

Let $\overline{V} : \mathcal{X} \to \mathcal{X}$ is stocastic matrix of conditional probabilities $\overline{V} = \{\overline{V}(x|\bar{x}), x \in \mathcal{X}, \bar{x} \in \mathcal{X}\}$.

We use the following notations: for entropy of RV X with PD P:

$$H_P(X) \triangleq -\sum_{x \in \mathcal{X}} P(x) \log P(x),$$

for joint entropy of RV X and Y:

$$H_{P,V}(X,Y) \triangleq -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)V(y|x) \log P(x)V(y|x),$$

for conditional entropy of RV Y relative to RV X:

$$H_{P,V}(Y|X) \triangleq -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)V(y|x) \log V(y|x),$$

for mutual information of RV X and Y:

$$I_{P,V}(X \wedge Y) \triangleq -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)V(y|x) \log \frac{V(y|x)}{PV(y)},$$

for informational divergence of PD P and PD Q on $\mathcal{X}$:

$$D(P\|Q) \triangleq \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)},$$

and for informational conditional divergence of PD $P \circ V$ and PD $P \circ W$ on $\mathcal{X} \times \mathcal{Y}$:

$$D(V\|W|P) \triangleq \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)V(y|x) \log \frac{V(y|x)}{W(y|x)}.$$

The following identities are often useful

$$D(P \circ V \| Q \circ W) = D(P\|Q) + D(V\|W|P),$$

$$H_{P,V}(X,Y) = H_P(X) + H_{P,V}(Y|X) = H_{PV}(Y) + H_{P,V}(X|Y),$$

$$I_{P,V}(X \wedge Y) = H_{PV}(Y) - H_{P,V}(Y|X) = H_P(X) + H_{PV}(Y) - H_{P,V}(X,Y).$$

The proofs in this paper will be based on the method of types [8], [10]. The type P of a sequence (or vector) $\mathbf{x} = (x_1, \dots, x_N) \in \mathcal{X}^N$ is a PD $P = \{P(x) = N(x|\mathbf{x})/N, x \in \mathcal{X}\}$, where $N(x|\mathbf{x})$ is the number of repetitions of symbol x among $x_1, \dots, x_N$.

The joint type of $\mathbf{x}$ and $\mathbf{y} \in \mathcal{Y}^N$ is the PD $P = \{P(x,y) = N(x,y|\mathbf{x},\mathbf{y})/N,\ x \in \mathcal{X},\ y \in \mathcal{Y}\}$, where $N(x,y|\mathbf{x},\mathbf{y})$ is the number of occurrences of symbols pair (x,y) in the pair of vectors $(\mathbf{x},\mathbf{y})$.

We say that the conditional type of $\mathbf{y}$ for given $\mathbf{x}$ is PD $V = \{V(y|x),\ x \in \mathcal{X},\ y \in \mathcal{Y}\}$ if $N(x,y|\mathbf{x},\mathbf{y}) = N(x|\mathbf{x})V(y|x)$ for all $x \in \mathcal{X},\ y \in \mathcal{Y}$.

The set of all PD on $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$ and the subset of $\mathcal{P}(\mathcal{X})$ consisting of the possible types of sequences $\mathbf{x} \in \mathcal{X}^N$ is denoted by $\mathcal{P}_N(\mathcal{X})$.

The set of vectors $\mathbf{x}$ of type P is denoted by $\mathcal{T}_P^N(X)$, $\mathcal{T}_P^N(X) = \emptyset$ for PD $P \notin \mathcal{P}_N(\mathcal{X})$. The set of all sequences $\mathbf{y} \in \mathcal{Y}^N$ of conditional type V for given $\mathbf{x} \in \mathcal{T}_P^N(X)$ is denoted by $\mathcal{T}_{P,V}^N(Y|\mathbf{x})$ and called V-shell of $\mathbf{x}$. The set of all possible V-shells for $\mathbf{x}$ of type P is denoted $\mathcal{V}_N(\mathcal{Y},P)$.

In the following lemmas very useful properties of types are formulated, for proofs see [8], [10].

Lemma 1: *(Type counting)*

$$|\mathcal{P}_N(\mathcal{X})| \leq (N+1)^{|\mathcal{X}|}, \tag{7}$$

$$|\mathcal{V}_N(\mathcal{Y},P)| \leq (N+1)^{|\mathcal{X}||\mathcal{Y}|}. \tag{8}$$

Lemma 2: *For any type $P \in \mathcal{P}_N(\mathcal{X})$*

$$(N+1)^{-|\mathcal{X}|}\exp\{NH_P(X)\} \leq |\mathcal{T}_P^N(X)| \leq \exp\{NH_P(X)\}, \tag{9}$$

and for any conditional type V and $\mathbf{x} \in \mathcal{T}_P^N(X)$

$$(N+1)^{-|\mathcal{X}||\mathcal{Y}|}\exp\{NH_{P,V}(Y|X)\} \leq |\mathcal{T}_{P,V}^N(Y|\mathbf{x})| \leq \exp\{NH_{P,V}(Y|X)\}. \tag{10}$$

Lemma 3: *If $\mathbf{x} \in \mathcal{T}_P^N(X)$, then for every PD Q on $\mathcal{X}$*

$$Q^N(\mathbf{x}) = \prod_{n=1}^{N} Q(x_n) = \exp\{-N(H_P(X) + D(P\|Q))\}. \tag{11}$$

If $\mathbf{y} \in \mathcal{T}_{P,V}^N(Y|\mathbf{x})$, then for every conditional PD V on $\mathcal{Y}$ for given $\mathbf{x}$

$$W^N(\mathbf{y}|\mathbf{x}) = \exp\{-N(H_{P,V}(Y|X) + D(V\|W|P))\}. \tag{12}$$

Consider the random coding exponent function $R_r(E,W)$, which is an inner estimate for $C(E,W) = R(E,W)$

$$R_r(P,E,W) \triangleq \min_{V:D(V\|W|P)\leq E} |I_{P,V}(X \wedge Y) + D(V\|W|P) - E|^+, \tag{13}$$

$$R_r(E,W) \triangleq \max_P R_r(P,E,W).$$

Consider the expurgated exponent function $R_x(E,W)$, which is another inner estimate for $R(E,W)$

$$R_x(P,E,W) = \min_{\widetilde{V}}\{I_{P,\widetilde{V}}(X \wedge \widetilde{X}) + |\mathbf{E}_{P,\widetilde{V}} d_B(X,\widetilde{X}) - E|^+\},$$

where $d_B(x, \tilde{x})$ is the Bhattacharyya distance

$$d_B(x, \tilde{x}) \triangleq -\log \sum_{y \in \mathcal{Y}} \sqrt{W(y|x)W(y|\tilde{x})},$$

and

$$R_x(E, W) \triangleq \max_P R_x(P, E, W).$$

As an upper bound of $R(E, W)$ serves sphere packing exponent function

$$\mathcal{R}_{sp}(P, E, W) = \min_{V: D(V \| W|P) \leq E} I_{P,V}(X \wedge Y), \tag{14}$$

$$R_{sp}(E, W) = \max_P R_{sp}(P, E, W).$$

It was first considered in [1].

Theorem 4: *For DMC W and for any $E > 0$ the following bound holds*

$$R(E, W) \geq \max(R_r(E, W), R_x(E, W)).$$

Theorem 5: *For $0 < E \leq E_{cr}(P, W)$, where*

$$E_{cr}(P, W) = \min \left\{ E : \frac{\partial R_{sp}(P, E, W)}{\partial E} \geq -1 \right\},$$

$$R(E, W) = R_{sp}(E, W) = R_r(E, W).$$

Remark: For $E \to 0$

$$\lim_{E \to 0} R_{sp}(P, 0, W) = \lim_{E \to 0} R_r(P, 0, W) = I_{P,W}(X \wedge Y).$$

3 Proof of theorem 1

Lemma 6: *Consider a finite set $\mathcal{A}$ and a nonnegative valued function ν on $\mathcal{A} \times \mathcal{A}$ such that for every $a, b \in \mathcal{A}$*

$$\nu(a, b) = \nu(b, a), \ \nu(a, a) = 0.$$

If for some t, for each $a \in \mathcal{A}$

$$\sum_{b \in \mathcal{A}} \nu(a, b) \leq t,$$

and $t_1, t_2, \ldots, t_S$ are nonnegative numbers such that

$$\sum_{s=1}^{S} t_s \geq t,$$

then $\mathcal{A}$ can be partitioned into S disjoint subsets $\mathcal{A}_1, \ldots, \mathcal{A}_S$ such that for every $a \in \mathcal{A}_s$

$$\sum_{b \in \mathcal{A}_s} \nu(a, b) \leq t_s.$$

For proof of the lemma see [3].

In [3] lower bounds for reliability function $E(R,W)$ of DMC and of sources with side information were obtained using this lemma 4. We now present similar derivation of random coding and expurgated bounds for E-capacity $R(E,W)$ of DMC.

Theorem 1 formulated above is a consequence of the following

Theorem 7: *For DMC $W:\mathcal{X}\to\mathcal{Y}$, any $E>0$, $\delta>0$ and type $P\in\mathcal{P}_N(\mathcal{X})$ for sufficiently large N codes (f,g) exist such that*

$$\exp\{-N(E+\delta)\}\le e(f,g,N,W)\le\exp\{-N(E-\delta)\} \tag{15}$$

and

$$R(f,g,N)\ge\max(R_r(P,E+\delta,W),R_x(P,E+\delta,W)).$$

The proof of the theorem 3 consists of several steps. First we shall prove

Lemma 8: *For given type $P\in\mathcal{P}_N(\mathcal{X})$, for any $0<r<|\mathcal{T}_P^N(X)|$ a set $\mathcal{C}$ exists, such that $\mathcal{C}\subset\mathcal{T}_P^N(X)$, $|\mathcal{C}|\ge r$ and for any $\tilde{\mathbf{x}}\in\mathcal{C}$ and matrix $\mathcal{V}:\mathcal{X}\to\mathcal{X}$ different from the identity matrix the following inequality holds*

$$|\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})\bigcap\mathcal{C}|\le r|\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})|\exp\{-N(H_P(X)-\delta_N)\}, \tag{16}$$

where

$$\delta_N=N^{-1}[(|\mathcal{X}|^2+|\mathcal{X}|)\log(N+1)+1].$$

Proof: Using lemma 4 let us take $\mathcal{A}\triangleq\mathcal{T}_P^N(X)$ and

$$\nu(\mathbf{x},\tilde{\mathbf{x}})\triangleq\begin{cases}|\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})|^{-1}, & \text{if } \mathbf{x}\ne\tilde{\mathbf{x}} \text{ and } \mathbf{x}\in\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}}),\\ 0, & \text{if } \mathbf{x}=\tilde{\mathbf{x}}.\end{cases}$$

Because $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are of the same type P, when $\mathbf{x}\in\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})$, then $\tilde{\mathbf{x}}\in\mathcal{T}_{P,\mathcal{V}'}^N(X|\mathbf{x})$ (where $\mathcal{V}'$ is the matrix transposed to $\mathcal{V}$) and therefore $\nu(\mathbf{x},\tilde{\mathbf{x}})=\nu(\tilde{\mathbf{x}},\mathbf{x})$. We have also from (8)

$$\sum_{\mathbf{x}\in\mathcal{T}_P^N(X)}\nu(\mathbf{x},\tilde{\mathbf{x}})=\sum_{\mathcal{V}}\sum_{\mathbf{x}\in\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})}\nu(\mathbf{x},\tilde{\mathbf{x}})\le(N+1)^{|\mathcal{X}|^2}.$$

If we take $t\triangleq(N+1)^{|\mathcal{X}|^2}$, $t_s\triangleq t/S$, $s=\overline{1,S}$, then according to Lemma 4 there exists a partition of $\mathcal{T}_P^N(X)$ into subsets $\mathcal{A}_s$, $s=\overline{1,S}$, such that for each $\tilde{\mathbf{x}}$ from $\mathcal{A}_s$

$$|\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})\bigcap\mathcal{A}_s|\le\frac{1}{S}|\mathcal{T}_{P,\mathcal{V}}^N(X|\tilde{\mathbf{x}})|(N+1)^{|\mathcal{X}|^2},\ s=\overline{1,S}. \tag{17}$$

Taking $\mathcal{C}$ equal to greatest $\mathcal{A}_s$ and S equal to integer part of $|\mathcal{T}_P^N(X)|/r$ we receive $|\mathcal{C}|\ge S^{-1}|\mathcal{T}_P^N(X)|\ge r$, and inequality (16), which follows from (17) and (9), because

$$\frac{1}{S}=\frac{1}{\lfloor|\mathcal{T}_P^N(X)|/r\rfloor}\le\frac{1}{|\mathcal{T}_P^N(X)|/2r}=\frac{2r}{|\mathcal{T}_P^N(X)|}.$$

Lemma 5 is proved. ■

For existence theorems demonstration it is possible to consider various "good" decoding rules. For definition of those rules following [3] apply different real-valued functions α defined on $\mathcal{X}^N \times \mathcal{Y}^N$. One says that g_α decoding is used if to each $\mathbf{y}$ from $\mathcal{Y}^N$ on the output of the channel the message m is accepted when codeword $\mathbf{x}(m)$ minimizes $\alpha(\mathbf{x}(m), \mathbf{y})$. One uses such functions α which depend only on type P of $\mathbf{x}$ and conditional type V of $\mathbf{y}$ for given $\mathbf{x}$. Such functions α can be written in the form $\alpha(P, V)$ and at respective decoding

$$g_\alpha : \mathcal{Y}^N \to \mathcal{M},$$

the message m corresponds to the vector $\mathbf{y}$, if

$$\alpha(P, V) = \min_{\widetilde{V}} \alpha(P, \widetilde{V}), \ \mathbf{y} \in \mathcal{T}^N_{P,V}(Y|\mathbf{x}(m)) \bigcap \mathcal{T}^N_{P,\widetilde{V}}(Y|\widetilde{\mathbf{x}}(m)).$$

Previously the following two rules were used [3]: maximum-likelihood decoding, when accepted codeword $\mathbf{x}(m)$ maximizes transition probability $W^N(\mathbf{x}(m)|\mathbf{y})$, in this case according to (12)

$$\alpha(P, V) = D(V\|W|P) + H_{P,V}(Y|X), \tag{18}$$

and the second decoding rule, called minimum-entropy decoding, according to which the codeword $\mathbf{x}(m)$ minimizing $H_{P,V}(Y|X)$ is accepted, that is

$$\alpha(P, V) = H_{P,V}(Y|X). \tag{19}$$

In [11] and [12] it was proposed another decoding rule by minimization of

$$\alpha(P, V) = D(V\|W|P), \tag{20}$$

which can be called minimum - divergence decoding.

Let $\hat{V} = \{\hat{V}(y|x, \widetilde{x}), x \in \mathcal{X}, \widetilde{x} \in \mathcal{X}, y \in \mathcal{Y}\}$, be conditional distribution of Y given values of X and $\widetilde{X}$ such, that

$$\sum_{\widetilde{x}} P(\widetilde{x})\overline{V}(x|\widetilde{x})\hat{V}(y|x, \widetilde{x}) = P(x)V(y|x), \tag{21}$$

$$\sum_{x} P(x)\overline{V}(\widetilde{x}|x)\hat{V}(y|x, \widetilde{x}) = P(\widetilde{x})\widetilde{V}(y|\widetilde{x}). \tag{22}$$

Following [3] we write $\widetilde{V} \prec_\alpha V$ if $\alpha(P, \widetilde{V}) \leq \alpha(P, V)$ and $P\widetilde{V} = PV$.

Let us denote

$$R_\alpha(P, E, W) \triangleq \min_{V: D(V\|W|P) \leq E} \ \min_{\hat{V}: \widetilde{V} \prec_\alpha V} \{I_{P,\overline{V}}(X \wedge \widetilde{X}) + |I_{P,\overline{V},\hat{V}}(Y \wedge \widetilde{X}|X) + D(V\|W|P) - E|^+\}, \tag{23}$$

where RV $X, \widetilde{X}, Y$ have values, correspondingly, on $\mathcal{X}, \mathcal{X}, \mathcal{Y}$ such that the following is valid:

both X and $\widetilde{X}$ have distribution P and $P\overline{V} = P$, $\overline{V}$ is conditional distribution of $\widetilde{X}$ for given value of X,

$\hat{V}$ is the conditional distribution of RV Y given X and $\widetilde{X}$ satisfying (21) and (22).

The main stage of the theorem demonstration is

Proposition 1: *For any DMC W, any type $P \in \mathcal{P}_N(\mathcal{X})$, any $E > 0$, $\delta_N > 0$, $\delta'_N > 0$, for all sufficiently large N codes (f, g_α) exist such, that*

$$\exp\{-N(E + \delta'_N/2)\} \leq e(f, g_\alpha, N, W) \leq \exp\{-N(E - \delta'_N/2)\}, \tag{24}$$

and

$$R(f, g_\alpha, N) \geq R_\alpha(P, E + \delta'_N, W). \tag{25}$$

Proof: The inequality

$$R_\alpha(P,E,W) \geq R_{\alpha,x}(P,E,W)$$

follows from definitions (23) and (33). For the proof of the inequality

$$R_\alpha(P,E,W) \geq R_{\alpha,r}(P,E,W)$$

remark that

$$I_{P,\overline{V}}(X \wedge \widetilde{X}) + I_{P,\overline{V},\widehat{V}}(Y \wedge \widetilde{X}|X) = I_{P,\overline{V},\widehat{V}}(XY \wedge \widetilde{X}) \geq I_{P,\widehat{V}}(Y \wedge \widetilde{X})$$

and then compare (23) and (32). ■

Lemma 10: *A point $E^*_\alpha(P,W)$ exists, such that*

$$\max[R_{\alpha,x}(P,E,W), R_{\alpha,r}(P,E,W)] = \begin{cases} R_{\alpha,r}(P,E,W), \text{ when } E \leq E^*_\alpha(P,W) \\ R_{\alpha,x}(P,E,W), \text{ when } E \geq E^*_\alpha(P,W). \end{cases}$$

Proof: Note that functions $R_{\alpha,r}(P,E,W)$ and $R_{\alpha,x}(P,E,W)$ are nonnegative and decreasing by E. Let us first prove that for any $E \geq E' \geq 0$

$$R_{\alpha,x}(P,E',W) \leq R_{\alpha,x}(P,E,W) + E - E'. \tag{34}$$

In accordance with (33), taking into account that for any a and b the inequality $|a+b|^+ \leq |a|^+ + |b|^+$ holds, we receive

$$R_{\alpha,x}(P,E',W) = \min_{\overline{V},\widehat{V}:\widehat{V} \prec_\alpha V} \{I_{P,\overline{V}}(X \wedge \widetilde{X}) + |I_{P,\overline{V},\widehat{V}}(Y \wedge \widetilde{X}|X) + D(V\|W|P) - E' + E - E|^+\} \leq$$

$$\leq R_{\alpha,x}(P,E,W) + E - E'.$$

Denote by $E^0_{\alpha,x}(P,W)$ the least value of E, for which $R_{\alpha,x}(E,P,W) = 0$. Let us show that for any E and E', such that

$$0 \leq E' \leq E \leq E^0_{\alpha,x}(P,W),$$

the inequality

$$R_{\alpha,r}(P,E,W) + E - E' \leq R_{\alpha,r}(P,E',W) \tag{35}$$

holds. Really in the interval $[0, E^0_{\alpha,x}(P,W)]$ function $R_{\alpha,r}(E,P,W)$ is strictly positive, then

$$R_{\alpha,r}(P,E',W) = \min_{V:D(V\|W|P)\leq E'} \min_{\widetilde{V}:\widetilde{V} \prec_\alpha V} (I_{P,\widetilde{V}}X \wedge \widetilde{Y}) + D(V\|W|P) - E) + E - E' \geq$$

$$\geq \min_{V:D(V\|W|P)\leq E} \min_{\widetilde{V}:\widetilde{V} \prec_\alpha V} (I_{P,\widetilde{V}}(X \wedge \widetilde{Y}) + D(V\|W|P) - E) + E - E' =$$

$$= R_{\alpha,r}(P,E,W) + E - E'.$$

Denote $E^*_\alpha(P,W)$ the smallest E, for which

$$R_{\alpha,r}(P,E,W) \leq R_{\alpha,x}(P,E,W).$$

Let us show that this inequality holds for all E greater than $E^*_\alpha(P,W)$. Consider two cases.

If $0 \leq E^*_\alpha(P,W) \leq E^0_{\alpha,x}(P,W)$, then it follows from (32) and (33) that for all E from interval $(E^*_\alpha(P,W), E^0_{\alpha,x}(P,W))$

$$R_{\alpha,r}(P,E,W) + E - E^*_\alpha(P,W) \leq R_{\alpha,r}(P, E^*_\alpha(P,W), W) \leq$$

$$\leq R_{\alpha,x}(P, E^*_\alpha(P,W), W) \leq R_{\alpha,x}(P,E,W) + E - E^*_\alpha(P,W).$$

If $E^0_{\alpha,x}(P,W) \leq E^*_\alpha(P,W)$, then for all E greater than $E^*_\alpha(P,W)$ we have

$$R_{\alpha,x}(P,E,W) = 0 = R_{\alpha,r}(P,E,W).$$

In this case $E^*_\alpha(P,W) = E^0_{\alpha,x}(P,W)$. ■

Lemma 11: *For each α-decoding*

$$R_{\alpha,x}(P,E,W) \leq R_x(P,E,W), \tag{36}$$

moreover for maximum likelihood decoding the equality holds.

Proof: First we prove the inequality (36). As

$$D(V\|W|P) + I_{P,\overline{V},\hat{V}}(Y \wedge \widetilde{X}|X) = \sum_{x,\widetilde{x},y} P(x)\overline{V}(\widetilde{x}|x)\hat{V}(y|x,\widetilde{x}) \log \frac{\hat{V}(y|x,\widetilde{x})}{W(y|x)}, \tag{37}$$

and

$$D(\widetilde{V}\|W|P) + I_{P,\overline{V}}(Y \wedge X|\widetilde{X}) = \sum_{x,\widetilde{x},y} P(x)\overline{V}(\widetilde{x}|x)\hat{V}(y|x,\widetilde{x}) \log \frac{\hat{V}(y|x,\widetilde{x})}{W(y|\widetilde{x})}, \tag{38}$$

then

$$R_{\alpha,x}(P,E,W) = \min_{\overline{V},\hat{V}:\widetilde{V} \prec_\alpha V} |I_{\hat{P}}(X \wedge \widetilde{X}) + I_{P,\overline{V},\hat{V}}(Y \wedge \widetilde{X}|X) + D(V\|W|P) - E|^+ \leq$$

$$\leq \min_{\overline{V},V,\widetilde{V}:\widetilde{V}=V} |I_{\hat{P}}(X \wedge \widetilde{X}) +$$

$$+1/2(I_{P,\overline{V},\hat{V}}(Y \wedge \widetilde{X}|X) + D(V\|W|P)) + 1/2(I_{P,\overline{V},\hat{V}}(Y \wedge \widetilde{X}|X) + D(\widetilde{V}\|W|P)) - E|^+.$$

From (37) and (38) denoting

$$\hat{V}_0(y|x,\widetilde{x}) \triangleq \exp d_B(x,\widetilde{x})\sqrt{W(y|x)W(y|\widetilde{x})} \tag{39}$$

we have

$$R_{\alpha,x}(P,E,W) \leq \min_{\overline{V},\hat{V}} |I_{P,\overline{V}}(X \wedge \widetilde{X}) + D(\widetilde{V}\|\widetilde{V}_0|P) + \mathbf{E}_{P,\overline{V}} d_B(X;\widetilde{X}) - E|^+ =$$

$$= \min_{\overline{V}} |I_{P,\overline{V}}(X \wedge \widetilde{X}) + \mathbf{E}_{P,\overline{V}} d_B(X,\widetilde{X}) - E|^+ = R_x(P,E,W).$$

Let us now prove that in the case of maximum likelihood decoding

$$R_x(P,E,W) = R_{\alpha,x}(P,E,W). \tag{40}$$

From the condition $\tilde{V} \prec_\alpha V$ we have

$$D(\tilde{V}\|W|P) + H_{P,V}(Y|X) \leq D(V\|W|P) + H_{P,V}(Y|X).$$

In accordance with (32), (33) and (37) we obtain

$$D(V\|W|P) + I_{P,\overline{V},\hat{V}}(Y \wedge \hat{X}|X) \geq D(\tilde{V}\|\tilde{V}_0|P) + \mathbf{E}_{P,\overline{V}} d_B(X, \widetilde{X}) \geq \mathbf{E}_{P,\overline{V}} d_B(X, \widetilde{X}).$$

Hence

$$|I_{P,\overline{V}}(X \wedge \widetilde{X}) + I_{P,\overline{V},\hat{V}}(Y \wedge \widetilde{X}|X) + D(V\|W|P) - E|^+ \geq$$
$$\geq |I_{P,\overline{V}}(X \wedge \widetilde{X}) + \mathbf{E}_{P,\overline{V}} d_B(X, \widetilde{X}) - E|^+,$$

and then (40) holds. ■

Lemma 12: *For each α-decoding*

$$R_{\alpha,r}(P, E, W) \leq R_r(P, E, W), \tag{41}$$

moreover, for
-maximum likelihood decoding,
-minimum entropy decoding,
-minimum divergence decoding
the equality holds.

Proof: The inequality (41) is valid because

$$R_{\alpha,r}(P, E, W) \leq \min_{V: D(V\|W|P) \leq E} \min_{\tilde{V}: \tilde{V} = V} |I_{P,\tilde{V}}(Y \wedge \widetilde{X}) + D(V\|W|P) - E|^+ = R_r(P, E, W).$$

The inverce inequality is not difficult to receive for all three noted α-decodings using condition $\tilde{V} \prec_\alpha V$. ■

Thus the proof of the theorem 3 is completed. ■

References

[1] E. A. Haroutunian, "Upper estimate of transmission rate for memoryless channel with countable number of output signals under given error probability exponent", (in Russian), *3th Conf. on Coding theory and Information Transmission. Uzhgorod–Tashkent.* pp. 83–86, 1967.

[2] E. A. Haroutunian, "Combinatorial method of construction of the upper bound for E-capacity", (in Russian), *Mezhvuz. Sb. Nouch. Trudov. Matematika.* Yerevan, vol. 1. pp. 213–220, 1982.

[3] I. Csiszár and J. Körner, "Graph decomposition: A new key to coding theorems" *IEEE Trans. on Inform. Theory*, vol. 27, no. 1, pp. 5–12, 1981.

[4] R. L. Dobrushin, "Asymptotic bounds of the probability of error for the transmission of messages over a memoryless channel with a symmetric transition probability matrix", (in Russian), *Teor. Veroyatnost. i Primenen.*, vol. 7, no. 3, pp. 283–311, 1962.

[5] E. A. Haroutunian, "Estimates of the error probability exponent for a semicontinuous memoryless channel", (in Russian), *Problemy Peredachi Informatcii*, vol. 4, no. 4, pp. 37–48, 1968.

[6] R. E. Blahut, "Hypothesis testing and information theory", *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 405–417, 1974.

[7] A. G. Dyachkov, "Bounds to average error probability for the fixed-composition code ensemble", (in Russian), *Problemy Peredachi Informatcii*, vol. 16, no. 4, pp. 3–8, 1980.

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

[9] R. E. Blahut, *Principles and Practice of Information Theory*. Addison-Wesley, Reading, MA, 1987.

[10] I. Csiszár, "The method of types", *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, pp. 2505–2523, October 1998.

[11] E. A. Haroutunian and H. Belbashir, *Lower baund of the optimal transmission rate depending on given error probability exponent for discret memoryless channel and for asymmetric broadcast channel*, (in Russian), Abstracts of papers of Sixth International Symposium on Information Theory, Tashkent, USSR, vol. 1, pp. 19-21, 1984.

[12] E. A. Haroutunian and M. E. Haroutunian, *Information Theory*, (in Armenian), Yerevan State University, 104 p., 1987.

E-ունակության պատահական կոդավորման և արտաքսման գնահատականների արտածումը գրաֆների մասնատմամբ

Ե. Ա. Հարությունյան

Ամփոփում

Աշխատանքում ներկայացվում է ընդհատ առանց հիշողության կապուղու E-ունակության համար պատահական կոդավորման և արտաքսման սահմանների գրաֆների մասնատման եղանակի միջոցով ստացումը: Դիտարկվում են ապակոդավորման երեք կանոններ: Պատահական կոդավորման սահմանը հասանելի է այդ բոլոր կանոնների կիրառմամբ, բայց արտաքսման սահմանը հասանելի է միայն առավելագույն ճշմարտանմանության ապակոդավորմամբ: