Mathematical Problems of Computer Science 31, 28–39, 2008.

Lower Bound for E Capacity of Discrete Memoryless Channel with Two-Sided State Information

Mariam E. Haroutunian and Arthur R. Muradyan

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

Abstract

We study the channel with two-sided state information, a discrete memoryless channel with finite input and output alphabets and random state sequence. Partial information about the state sequence is available to the encoder and decoder. Applications of this study include watermarking, data hiding, communication in presence of partially known interferers. The capacity of this model was obtained by Cover and Chiang in [1]. In this paper the random coding bound of E-capacity is derived for considered model which can be called also generalized channel with state information, as it includes four possible situations of the channel with random parameter.

1 Introduction

The problem of coding for the channel with random parameter, where the random state of the channel is observed by the encoder but not by the decoder was studied in [2, 3]. Applications of this model include computer memories with defects [4], writing on dirty paper [5], information hiding and watermarking [6].

The generalization of the channel with state information, where the sender and the receiver have correlated but different state information (Figure 1), was studied in [1] and the capacity of this channel was obtained. It was proved that capacities of the channel with random parameter in four possible situations are special cases of this capacity.

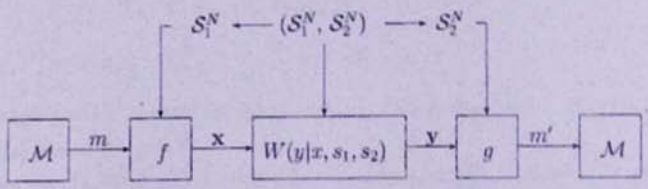


Figure 1. Channel with two-sided state information

In [7] the capacity formulas and random-coding exponents are derived for the model, where partial information about the state sequence is available to the encoder, adversary

and decoder. The investigation of such models is motivated by data-hiding applications in which the decoder has partial or no knowledge of the cover signal.

We investigate the E-capacity, which was first introduced by E. Haroutunian [8, 9] and developed for various channels [10, 11, 12, 13]. This approach is equivalent to studying of error exponents but sometimes is more expedient.

In this paper the random coding bound of E-capacity is derived for the channel with two-sided state information. When $E \to 0$, the limit of this bound coincides with the capacity of the channel, obtained in [1]. We also show, that the lower bounds of E-capacity for four possible situations of the channel with random parameter [9] are special cases of the bound, obtained in this paper.

The paper is organized as follows. In Section II some notations and definitions are given. The notion of E-capacity, its random coding bound with special cases are stated in Section III. The proof of the theorem is presented in Section IV. The proof of Packing lemma, which is used in proof of the theorem, is given in the Appendix.

2 Notations and Definitions

Following conventions are applied within the paper. Capital letters are used for random variables (RV) S_1, S_2, U, X, Y taking values in the finite sets $\mathcal{S}_1, \mathcal{S}_2, \mathcal{U}, \mathcal{X}, \mathcal{Y}$, correspondingly, and lower case letters s_1, s_2, u, x, y for their realizations. Small bold letters are used for N-length vectors $\mathbf{x} = (x_1, .., x_N) \in \mathcal{X}^N$. The cardinality of the set $\mathcal{X}$ we denote by $|\mathcal{X}|$. The notation $|a|^+$ will be used for $\max(a, 0)$.

The channel with two-sided state information is presented in Figure 1. It is defined by a transition probability matrix $W(y|x, s_1, s_2)$, where $x \in \mathcal{X}, y \in \mathcal{Y}, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$. $\mathcal{X}$ is the finite input alphabet, $\mathcal{Y}$ is the output alphabet. The state information is described by the pair of RVs (S_1, S_2) with given joint probability distribution (PD) $Q^* = Q_1^* \circ Q_2^* = \{Q^*(s_1, s_2) = Q_1^*(s_1)Q_2^*(s_2|s_1), s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2\}$. N-length sequences $\mathbf{s}_1$ are available to the encoder and $\mathbf{s}_2$ - to the decoder.

The considered channel is memoryless, it means that for input word $\mathbf{x} \in \mathcal{X}^N$, output word $\mathbf{y} \in \mathcal{Y}^N$ and state sequences $\mathbf{s}_1 \in \mathcal{S}_1^N$, $\mathbf{s}_2 \in \mathcal{S}_2^N$

$$W^N(\mathbf{y}|\mathbf{x}, \mathbf{s}_1, \mathbf{s}_2) = \prod_{n=1}^{N} W(y_n|x_n, s_{1n}, s_{2n}).$$

It is assumed that:

$$Q^{*N}(\mathbf{s}_1, \mathbf{s}_2) = \prod_{n=1}^{N} Q^*(s_{1n}, s_{2n}).$$

We shall use the following PDs:

$$Q = Q_1 \circ Q_2 = \{Q(s_1, s_2|u, x) = Q_1(s_1)Q_2(s_2|u, x, s_1), s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, u \in \mathcal{U}, x \in \mathcal{X}\},$$

$$P = P_0 \circ P_1 = \{P(u, x|s_1) = P_0(u|s_1)P_1(x|u, s_1), s_1 \in \mathcal{S}_1, u \in \mathcal{U}, x \in \mathcal{X}, \},$$

$$V = \{V(y|u, x, s_1, s_2), s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, u \in \mathcal{U}, x \in \mathcal{X}, y \in \mathcal{Y}\},$$

$$Q \circ P \circ V = \{Q(s_1, s_2|u, x)P(u, x|s_1)V(y|u, x, s_1, s_2) =$$

$$= Q_1(s_1)P_0(u|s_1)P_1(x|u, s_1), Q_2(s_2|u, x, s_1)V(y|u, x, s_1, s_2),$$

$$s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, u \in \mathcal{U}, x \in \mathcal{X}, y \in \mathcal{Y}\},$$

$$Q_2^* \circ W = \{Q_2^* \circ W(y, s_2|x, s_1) = Q_2^*(s_2|s_1)W(y|x, s_1, s_2), s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, x \in \mathcal{X}, y \in \mathcal{Y}\}.$$

For the information-theoretic quantities, such as entropy $H_{Q_1}(S_1)$, mutual information $I_{Q_1,P_0}(U \wedge S_1)$, divergence $D(Q_1\|Q_1^*)$ and for the notion of type we refer to [9, 14, 15, 16].

The following properties [14, 15] are used in proofs:

$$D(Q \circ P \circ V\|Q^* \circ P \circ W) = D(Q_1\|Q_1^*) + D(Q_2 \circ V\|Q_2^* \circ W|Q_1, P), \tag{1}$$

for $\mathbf{s}_1 \in \mathcal{T}_{Q_1}^N(S_1)$,

$$Q_1^{*N}(\mathbf{s}_1) = \exp\{-N(H_{Q_1}(S_1) + D(Q_1\|Q_1^*)\}, \tag{2}$$

for $\mathbf{x} \in \mathcal{T}_{Q_1,P}(X|\mathbf{s}_1)$, $(\mathbf{y}, \mathbf{s}_2) \in \mathcal{T}_{Q,P,V}(Y, S_2|\mathbf{x}, \mathbf{s}_1)$,

$$Q_2^{*N} \circ W^N(\mathbf{y}, \mathbf{s}_2|\mathbf{x}, \mathbf{s}_1) = \exp\{-N(H_{Q,P,V}(Y, S_2|X, S_1) + D(Q_2 \circ V\|Q_2^* \circ W|Q_1, P)\}, \tag{3}$$

$$(N+1)^{-|\mathcal{S}_1|}\exp\{NH_{Q_1}(S_1)\} \le |\mathcal{T}_{Q_1}^N(S_1)| \le \exp\{NH_{Q_1}(S_1)\}, \tag{4}$$

$$(N+1)^{-|\mathcal{U}||\mathcal{X}||\mathcal{S}_1||\mathcal{S}_2||\mathcal{Y}|}\exp\{NH_{Q,P,V}(Y, S_2|U, X, S_1)\} \le$$

$$\le |\mathcal{T}_{Q,P,V}^N(Y, S_2|\mathbf{u}, \mathbf{x}, \mathbf{s}_1)| \le \exp\{NH_{Q,P,V}(Y, S_2|U, X, S_1)\}, \tag{5}$$

$$H_{Q,P,V}(Y, S_2|U, X, S_1) \le H_{Q,P,V}(Y, S_2|X, S_1). \tag{6}$$

All logarithms and exponents in the paper are of the base 2.

Let $\mathcal{M}$ be the message set. The N-length code is a pair of mappings (f_N, g_N), where $f_N : \mathcal{M} \times \mathcal{S}_1^N \to \mathcal{X}^N$ is the encoding function and $g_N : \mathcal{Y}^N \times \mathcal{S}_2^N \to \mathcal{M}$ is the decoding function. The nonnegative number $R = \frac{1}{N}\log|\mathcal{M}|$ is called code rate.

The probability of erroneous transmission of the message $m \in \mathcal{M}$ by the channel for the code (f_N, g_N) is:

$$e(m) = e(f_N, g_N, W, Q^*, m) = \sum_{\mathbf{s}_1 \in \mathcal{S}_1^N} Q_1^{*N}(\mathbf{s}_1) * Q_2^{*N} \circ W^N(\mathcal{Y}^N \times \mathcal{S}_2^N \backslash g_N^{-1}(m)|f(m, \mathbf{s}_1), \mathbf{s}_1),$$

where $g_N^{-1}(m) = \{(\mathbf{y}, \mathbf{s}_2) : g_N(\mathbf{y}, \mathbf{s}_2) = m\}$.

The maximal error probability of the code equals:

$$e = e(f_N, g_N, W, Q^*) = \max_{m \in \mathcal{M}} e(m)$$

and the average error probability of the code is (the messages are supposed to be equiprobable):

$$\bar{e} = \bar{e}(f_N, g_N, W, Q^*) = \frac{1}{|\mathcal{M}|}\sum_{m \in \mathcal{M}} e(m).$$

3 Formulation of Results

We investigate the E-capacity function which is defined as:

$$C(E,Q^*,W) = \overline{\lim_{N\to\infty}} \frac{1}{N} \log M(E,Q^*,W,N),$$

where

$$M(E,Q^*,W,N) = \sup_{f_N,\mathcal{A}_N} \{|\mathcal{M}| : e \le \exp(-NE)\}.$$

It is the generalization of the capacity, as it reduces to the latter when $E \to 0$. We denote by $\overline{C}(E,Q^*,W)$ the E-capacity for the average error probability. In this paper the lower bound of E-capacity for maximal and average error probabilities is constructed.

To formulate the lower bound of the E-capacity let us denote:

$$R(E,Q^*,W,Q,P,V) = I_{Q,P,V}(U \wedge S_2, Y) - I_{Q_1,P_0}(U \wedge S_1) + D(Q \circ P \circ V || Q^* \circ P \circ W) - E,$$

$$R_\tau(E,Q^*,W) = \min_{Q_1} \max_{P} \min_{Q_2,V:D(Q\circ P\circ V||Q^*\circ P\circ W)\le E} \left| R(E,Q^*,W,Q,P,V) \right|^+. \tag{7}$$

Theorem. *For the channel with two-sided state information W with given Q^*, for all $E > 0$*

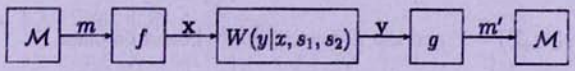
$$R_\tau(E,Q^*,W) \le C(E,Q^*,W) \le \overline{C}(E,Q^*,W).$$

The proof of the theorem is given in Section IV.

Corollary. *When $E \to 0$ we derive the lower bound of channel capacity, which coincides with the capacity obtained in [1].*

We shall show, that the random coding bounds of E-capacity for the four possible situations of the channel with random parameter [9] corresponds to the special cases of Theorem.

Case 1: No state information at the encoder and decoder: $S_1 = \emptyset, S_2 = \emptyset$.



This channel is equivalent to a DMC with

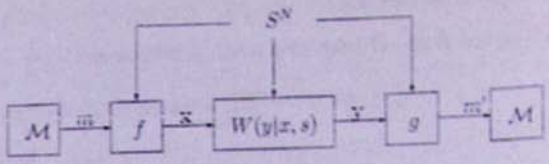
$$W^*(y|x) = \sum_{s_1\in\mathcal{S}_1, s_2\in\mathcal{S}_2} W(y|x,s_1,s_2) Q^*(s_1,s_2).$$

Here $(S_2, Y) = Y$, $I_{Q_1,P_0}(U \wedge S_1) = 0$, $U \to X \to Y$ forms a Markov chain, hence $\max_{P(u,x)} I(U \wedge Y) \le \max_{P(x)} I(X \wedge Y)$ with equality iff $U = X$. Then

$$R_\tau(E,Q^*,W^*) = \max_{P(u,x)} \min_{V:D(P\circ V||P\circ W^*)\le E} \left| I_{P,V}(U \wedge Y) + D(P \circ V || P \circ W^*) - E \right|^+ =$$

$$= \max_{P(x)} \min_{V:D(V||W^*|P)\le E} \left| I_{P,V}(X \wedge Y) + D(V||W^*|P) - E \right|^+.$$

Case 2: State information on both sides is the same: $S_1 = S_2 = S$.



Here $U \to X \to Y$ forms a Markov chain conditioned on S, hence $\max_{P(u,x|s)} I(U \wedge Y|S) \le \max_{P(x|s)} I(X \wedge Y|S)$ with equality iff $U = X$. Then

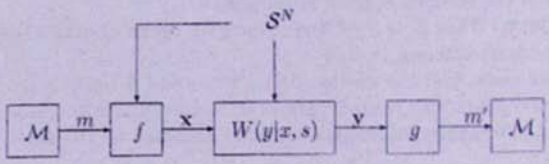
$$R_r(E, Q^*, W) = \min_{Q} \max_{P(u,x|s)} \min_{V: D(Q \circ P \circ V \| Q^* \circ P \circ W) \le E} \left| I_{Q,P,V}(U \wedge S, Y) - I_{Q,P_0}(U \wedge S) + \right.$$

$$\left. + D(Q \circ P \circ V \| Q^* \circ P \circ W) - E \right|^+ =$$

$$= \min_{Q} \max_{P(u,x|s)} \min_{V: D(Q \circ P \circ V \| Q^* \circ P \circ W) \le E} \left| I_{Q,P,V}(U \wedge Y|S) + D(Q \circ P \circ V \| Q^* \circ P \circ W) - E \right|^+ =$$

$$= \min_{Q} \max_{P(x|s)} \min_{V: D(Q \circ P \circ V \| Q^* \circ P \circ W) \le E} \left| I_{Q,P,V}(X \wedge Y|S) + D(Q \circ P \circ V \| Q^* \circ P \circ W) - E \right|^+.$$

Case 3: State information at the encoder: $S_1 = S, S_2 = \emptyset$. This is the channel with random parameter with informed encoder, considered by Gelfand and Pinsker [2].



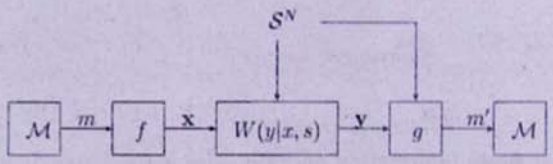
Here $(S_2, Y) = Y$ and the E-capacity lower bound becomes

$$R_r(E, Q^*, W) = \min_{Q} \max_{P(u,x|s)} \min_{V: D(Q \circ P \circ V \| Q^* \circ P \circ W) \le E} \left| I_{Q,P,V}(U \wedge Y) - I_{Q,P}(S \wedge U) + \right.$$

$$\left. + D(Q \circ P \circ V \| Q^* \circ P \circ W) - E \right|^+,$$

which coincides with the bound derived in [10].

Case 4: State information at the decoder: $S_1 = \emptyset, S_2 = S$.



Now $I_{Q,P_0}(U \wedge S_1) = 0$ and $U \to X \to (S,Y)$ forms a Markov chain, hence inequality $\max_{P(u,x)} I(U \wedge S,Y) \le \max_{P(x)} I(X \wedge S,Y)$ with equality iff $U = X$. Then

$$R_r(E,Q^*,W) =$$

$$= \max_{P(u,x)} \min_{Q,V:D(Q\circ P\circ V\|Q^*\circ P\circ W)\le E} \left| I_{Q,P,V}(U \wedge S,Y) + D(Q\circ P\circ V\|Q^*\circ P\circ W) - E \right|^+ =$$

$$= \max_{P(x)} \min_{Q,V:D(Q\circ P\circ V\|Q^*\circ P\circ W)\le E} \left| I_{Q,P,V}(X \wedge S,Y) + D(Q\circ P\circ V\|Q^*\circ P\circ W) - E \right|^+.$$

In the same way from the theorem we can derive the results for the special cases of the information-hiding system [12],[13], i.e. when the channel is independent of state information.

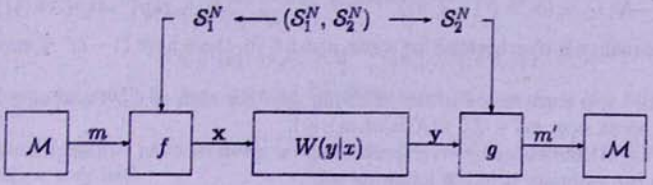


Figure 2. Information-hiding system

4 Proof of Theorem

To prove the theorem we must show the existence of a code with R satisfying (3) and maximal error probability not greater than $\exp\{-N(E-\varepsilon)\}$ for any $0 < \varepsilon < E$.

We will construct encoding and decoding and compute errors caused by each. We use random bin coding technique [2] for encoding and minimum divergence method [8] for decoding.

Denote $\mathcal{Q}(Q_1^*,E) = \{Q_1 : D(Q_1\|Q_1^*) \le E\}$ and

$$\mathcal{T}^N_{Q_1^*,E}(S_1) = \bigcup_{Q_1\in\mathcal{Q}(Q_1^*,E)} \mathcal{T}^N_{Q_1}(S_1).$$

We will construct the code only for $\mathbf{s}_1 \in \mathcal{T}^N_{Q_1^*,E}(S_1)$, because for sufficiently large N, the probability of $\mathbf{s}_1 \notin \mathcal{T}^N_{Q_1^*,E}(S_1)$ is exponentially small:

$$Q_1^{*N}\Big\{ \bigcup_{Q_1\notin\mathcal{Q}(Q_1^*,E)} \mathcal{T}^N_{Q_1}(S_1)\Big\} = \sum_{Q_1\notin\mathcal{Q}(Q_1^*,E)} Q_1^{*N}\{\mathcal{T}^N_{Q_1}(S_1)\} \le$$

$$\le \sum_{Q_1\notin\mathcal{Q}(Q_1^*,E)} \exp\{-ND(Q_1\|Q_1^*)\} < (N+1)^{|\mathcal{S}_1|}\exp\{-NE\} \le \exp\{-N(E-\varepsilon_1)\}, \qquad (8)$$

where ε_1 is positive and small enough.

Encoding. For small $\delta > 0$, any type $Q_1 \in \mathcal{Q}(Q_1^*,E)$, for fixed $P = P_0 \circ P_1$ and E we choose randomly $|\mathcal{M}|$ collections $\mathcal{J}(m)$, $m \in \mathcal{M}$ of vectors $\mathbf{u}_j(m)$, $j = \overline{1,J}$ from $\mathcal{T}^N_{Q_1,P_0}(U)$. where $J = \exp\{N(I_{Q_1,P_0}(S_1 \wedge U) + \delta/2)\}$.

Then for each $\mathbf{s}_1 \in \mathcal{T}^N_{Q_1}(S_1)$ we choose such $\mathbf{u}_j(m)$ from $\mathcal{J}(m)$, that $\mathbf{u}_j(m) \in \mathcal{T}^N_{Q_1,P_0}(U|\mathbf{s}_1)$. Denote this vector by $\mathbf{u}(m,\mathbf{s}_1)$.

If for some $\mathbf{s}_1 \in \mathcal{T}^N_{Q_1}(S_1)$ there is no such $\mathbf{u}_j(m)$ in $\mathcal{J}(m)$, we randomly choose $\mathbf{u}(m,\mathbf{s}_1)$ from $\mathcal{T}^N_{Q_1,P_0}(U|\mathbf{s}_1)$. Denote by $B_{Q_1,P_0}(m,\mathbf{s}_1)$ this event. Its probability can be estimated in the following way.

$$\Pr\{B_{Q_1,P_0}(m,\mathbf{s}_1)\} = \Pr\Big\{\bigcap_{j=1}^{J} \mathbf{u}_j(m) \notin \mathcal{T}^N_{Q_1,P_0}(U|\mathbf{s}_1)\Big\} \le$$

$$\le \prod_{j=1}^{J}\Big(1 - \Pr\Big\{\mathbf{u}_j(m) \in \mathcal{T}^N_{Q_1,P_0}(U|\mathbf{s}_1)\Big\}\Big) \le \Big(1 - \frac{|\mathcal{T}^N_{Q_1,P_0}(U|\mathbf{s}_1)|}{|\mathcal{T}^N_{Q_1,P_0}(U)|}\Big)^J \le$$

$$\le (1 - \exp\{-N(I_{Q_1,P_0}(S_1 \wedge U) + \delta/4)\})^{\exp\{N(I_{Q_1,P_0}(S_1\wedge U)+\delta/2)\}} \le \exp\{-\exp\{N\delta/4\}\}. \quad (9)$$

The last inequality is true because for any n and $t \in (0,1)$ we have $(1-t)^n \le \exp\{-nt\}$.

The codeword $\mathbf{x}$ is constructed in the following way. For each $m \in \mathcal{M}$ and $\mathbf{s}_1 \in \mathcal{T}^N_{Q_1}(S_1)$ we randomly choose $\mathbf{x}(m,\mathbf{s}_1) \in \mathcal{T}^N_{Q_1,P}(X|\mathbf{u}(m,\mathbf{s}_1),\mathbf{s}_1)$.

Denote by $e_E(m)$ the *encoding error probability* for given $m \in \mathcal{M}$. Taking into account (8) and (9) we can estimate it in the following way:

$$e_E(m) = \sum_{\mathbf{s}_1 \in \mathcal{T}^N_{Q_1^*,E}(S_1)} Q_1^{*N}(\mathbf{s}_1)\Pr\{B_{Q_1,P_0}(m,\mathbf{s}_1)\} + \sum_{\mathbf{s}_1 \notin \mathcal{T}^N_{Q_1^*,E}(S_1)} Q_1^{*N}(\mathbf{s}_1) \le$$

$$\le \sum_{Q_1 \in \mathcal{Q}(Q_1^*,E)} Q_1^{*N}(\mathcal{T}^N_{Q_1}(S_1))\exp\{-\exp\{N\delta/4\}\} + \exp\{-N(E-\varepsilon_1)\}.$$

As the number of types $Q_1 \in \mathcal{Q}(Q_1^*,E)$ does not exceed $(N+1)^{|\mathcal{S}_1|}$

$$e_E(m) \le (N+1)^{|\mathcal{S}_1|}\exp\{-\exp\{N\delta/4\}\} + \exp\{-N(E-\varepsilon_1)\} \le$$

$$\le \exp\{-\exp\{N\delta/4\} + \varepsilon_2\} + \exp\{-N(E-\varepsilon_1)\}, \quad (10)$$

for N large enough.

Decoding. For brevity the pair of vectors $\mathbf{u}(m,\mathbf{s}_1), \mathbf{x}(m,\mathbf{s}_1)$ we denote by $\mathbf{u},\mathbf{x}(m,\mathbf{s}_1)$. According to minimum divergence method each $\mathbf{y}$ and $\mathbf{s}_2$ are decoded to such m for which: $(\mathbf{y},\mathbf{s}_2) \in \mathcal{T}^N_{Q,P,V}(Y,S_2|\mathbf{u},\mathbf{x}(m,\mathbf{s}_1),\mathbf{s}_1)$ and Q,P,V are such that $D(Q\circ P\circ V||Q^*\circ P\circ W)$ is minimal.

The decoder g can make an error, when $m \in \mathcal{M}$ is transmitted in the case of state information $\mathbf{s}_1 \in \mathcal{T}^E_{Q_1}(S_1)$, but there exists $m' \ne m$, vector $\mathbf{s}_1'$, types Q', P', V' such that

$$(\mathbf{y},\mathbf{s}_2) \in \mathcal{T}^N_{Q,P,V}(Y,S_2|\mathbf{u},\mathbf{x}(m,\mathbf{s}_1),\mathbf{s}_1) \bigcap \mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathbf{u}',\mathbf{x}'(m',\mathbf{s}_1'),\mathbf{s}_1')$$

and

$$D(Q'\circ P'\circ V'||Q^*\circ P'\circ W) \le D(Q\circ P\circ V||Q^*\circ P\circ W). \quad (11)$$

Denote by $\mathcal{D} = \{Q,P,V,Q',P',V' : (11) \text{ is valid}\}$ and by $e_D(m)$ the *decoding error probability* for given $m \in \mathcal{M}$, then

$$e_D(m) \le \sum_{\mathbf{s}_1 \in \mathcal{T}^N_{Q_1^*,E}(S_1)} Q_1^{*N}(\mathbf{s}_1) * Q_2^{*N} \circ W^N \Big\{ \bigcup_{\mathcal{D}} \mathcal{T}^N_{Q,P,V}(Y, S_2|\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \bigcap$$

$$\bigcup_{m' \ne m} \bigcup_{\mathbf{s}'_1 \in \mathcal{T}^N_{Q'_1}(S_1)} \mathcal{T}^N_{Q',P',V'}(Y, S_2|\mathbf{u}', \mathbf{x}'(m', \mathbf{s}'_1), \mathbf{s}'_1) \Big| \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1 \Big\} \le$$

$$\le \sum_{\mathbf{s}_1 \in \mathcal{T}^N_{Q_1^*,E}(S_1)} \sum_{\mathcal{D}} \Big| \mathcal{T}^N_{Q,P,V}(Y, S_2|\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \bigcap$$

$$\bigcap \bigcup_{m' \ne m} \bigcup_{\mathbf{s}'_1 \in \mathcal{T}^N_{Q'_1}(S_1)} \mathcal{T}^N_{Q',P',V'}(Y, S_2|\mathbf{u}', \mathbf{x}'(m', \mathbf{s}'_1), \mathbf{s}'_1) \Big| \times$$

$$\times Q_1^{*N}(\mathbf{s}_1) * Q_2^{*N} \circ W^N(\mathbf{y}, \mathbf{s}_2|\mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1).$$

The last inequality is true, because for fixed types of $\mathbf{x}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{y}$ the probabilities are constant.

Packing Lemma. *For given W, Q^*, for any $E > \delta \ge 0$, types $Q_1 \in \mathcal{Q}(Q_1^*, E)$ and P there exists a code with*

$$|\mathcal{M}| \ge \exp\Big\{ N \min_{Q_2, V: D(Q_2 \circ P \circ V || Q^* \circ P \circ W) \le E} \Big| R(E, Q^*, W, Q, P, V) - \delta \Big|^+ \Big\} \tag{12}$$

such that

1. for each $\mathbf{s}_1 \in \mathcal{T}^N_{Q_1}(S_1)$ vector pairs $\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1)$ are distinct for different $m \in \mathcal{M}$,

2. for sufficiently large N the following inequality holds for any $Q'_1 \in \mathcal{Q}(Q_1^*, E)$, conditional types $P', Q_2 \circ V, Q'_2 \circ V'$ and for all $m \in \mathcal{M}$ and $\mathbf{s}_1 \in \mathcal{T}^N_{Q_1}(S_1)$ the following inequality holds

$$\Big| \mathcal{T}^N_{Q,P,V}(Y, S_2|\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \bigcap \bigcup_{m' \ne m} \bigcup_{\mathbf{s}'_1 \in \mathcal{T}^N_{Q'_1}(S_1)} \mathcal{T}^N_{Q',P',V'}(Y, S_2|\mathbf{u}', \mathbf{x}'(m', \mathbf{s}'_1), \mathbf{s}'_1) \Big| \le$$

$$\le \Big| \mathcal{T}^N_{Q,P,V}(Y, S_2|\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \Big| \exp\Big\{ -N \Big| E - D(Q' \circ P' \circ V' || Q^* \circ P' \circ W') \Big|^+ \Big\}. \tag{13}$$

Proof of the Lemma is given in the appendix.

Taking into account (1), (2), (3), (5), (6), (11) and (13) we can upper estimate the *decoding error probability*:

$$e_D(m) \le \sum_{\mathbf{s}_1 \in \mathcal{T}^N_{Q_1^*,E}(S_1)} \sum_{\mathcal{D}} \Big| \mathcal{T}^N_{Q,P,V}(Y, S_2|\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \Big| \exp\Big\{ -N(E - D(Q' \circ P' \circ V' || Q^* \circ P' \circ W')) \Big\} \times$$

$$\times Q_1^{*N}(\mathbf{s}_1) * Q_2^{*N} \circ W^N(\mathbf{y}, \mathbf{s}_2|\mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \le$$

$$\le \sum_{Q_1 \in \mathcal{Q}(Q_1^*, E)} \exp\{N H_{Q_1}(S_1)\} \sum_{\mathcal{D}} \exp\{N H_{Q,P,V}(Y, S_2|U, X, S_1)\} \times$$

$$\times \exp\{-N(E - D(Q' \circ P' \circ V' || Q^* \circ P' \circ W'))\} \times$$

$$\times \exp\{-N(H_{Q_1}(S_1) + D(Q_1 || Q_1^*) + H_{Q,P,V}(Y, S_2|X, S_1) + D(Q_2 \circ V || Q_2^* \circ W | Q_1, P))\} \le$$

$$\le (N+1)^{|\mathcal{S}_1|} \sum_{Q,P,V,Q',P',V'} \exp\{-N(E + H_{Q,P,V}(Y, S_2|X, S_1) - H_{Q,P,V}(Y, S_2|U, X, S_1))\} \le$$

$$\leq (N+1)^{|\mathcal{S}_1|} \sum_{Q,P,V,Q',P',V'} \exp\{-NE\} \leq \exp\{-N(E-\varepsilon_3)\}. \tag{14}$$

The error probability of the message $m \in \mathcal{M}$ is $e(m) \leq e_E(m) + e_D(m)$.

Taking into account (10) and (14) we get

$$e(m) \leq \exp\{-\exp\{N\delta/4\}+\varepsilon_2\}+\exp\{-N(E-\varepsilon_1)\}+\exp\{-N(E-\varepsilon_3)\} \leq \exp\{-N(E-\varepsilon)\}.$$

Considering the continuity of all expressions, when $N \to \infty$, arbitrary probability distributions can be considered instead of types.

The theorem is proved.

Appendix. Proof of the Packing Lemma

We can see that if some code satisfies (13) for any $Q_1', P', Q_2 \circ V, Q_2' \circ V'$ then point 1 of the lemma is true. To prove that, it is enough to choose $Q_1 = Q_1', P = P', Q_2 \circ V = Q_2' \circ V'$ and $D(Q' \circ P' \circ V' \| Q^* \circ P' \circ W) < E$.

Now to prove the lemma first suppose that V', P', Q' are such that $D(Q' \circ P' \circ V' \| Q^* \circ P' \circ W) \geq E$. Then $\exp\{-N|E - D(Q' \circ P' \circ V' \| Q^* \circ P' \circ W)|^+\} = 1$ and (13) is valid for any $|\mathcal{M}|$.

It is left to prove the inequality (13) for Q', P', V' such that $D(Q' \circ P' \circ V' \| Q^* \circ P' \circ W) < E$. For any code let us denote

$$A_m(Q_2, V, Q', P', V') = \exp\left\{N(E - D(Q' \circ P' \circ V' \| Q^* \circ P' \circ W) - H_{Q,P,V}(Y, S_2|U, X, S_1)\right\} \times$$

$$\times \max_{\mathbf{s}_1 \in T^N_{Q_1}(S_1)} \left| T^N_{Q,P,V}(Y, S_2|\mathbf{u}, \mathbf{x}(m, \mathbf{s}_1), \mathbf{s}_1) \bigcap \bigcup_{m' \neq m} \bigcup_{\mathbf{s}_1' \in T^N_{Q_1'}(S_1)} T^N_{Q',P',V'}(Y, S_2|\mathbf{u}', \mathbf{x}'(m', \mathbf{s}_1'), \mathbf{s}_1) \right|$$

and

$$A_m = (N+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Y}||\mathcal{S}_1||\mathcal{S}_2|} \sum_{Q_2 \circ V} \sum_{Q',P',V': D(Q' \circ P' \circ V' \| Q^* \circ P' \circ W) < E} A_m(Q_2, V, Q', P', V').$$

It is clear that if $A_m \leq 1$ for all $m \in \mathcal{M}$, then the point 2 of lemma is true. So to prove the lemma it is enough to prove that $A_m \leq 1$ for all $m \in \mathcal{M}$.

Now notice that if for some code the following inequality holds

$$\frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} A_m \leq \frac{1}{2}, \tag{15}$$

then $A_m \leq 1$ for at least $|\mathcal{M}|/2$ indices m. Furthermore, if we denote such indices by m^*, then $A_{m^*} \leq A_m \leq 1$ for every m^*. Therefore the lemma will be proved if we find code satisfying (15) with

$$2\exp\{N \min_{Q_2,V: D(Q \circ P \circ V \| Q^* \circ P \circ W) \leq E} |R(E, Q^*, W, Q, P, V) - \delta|^+\} \leq |\mathcal{M}| \leq$$

$$\leq \exp\{N \min_{Q_2,V: D(Q \circ P \circ V \| Q^* \circ P \circ W) \leq E} |R(E, Q^*, W, Q, P, V) - \delta/2|^+\}. \tag{16}$$

To prove that (15) holds for some code it suffices to show that for random code

$$\mathbf{E} A_m \leq \frac{1}{2}, \qquad m \in \mathcal{M}. \tag{17}$$

To this end we observe that

$$\mathbf{E}\Big|\mathcal{T}^N_{Q,P,V}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathcal{S}_1^N),\mathbf{s}_1)\bigcap\bigcup_{m'\neq m}\bigcup_{\mathbf{s}'_1\in\mathcal{T}^N_{Q'_1}(S_1)}\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathbf{s}'_1)\Big|\leq$$

$$\leq\sum_{(\mathbf{y},\mathbf{s}_2)\in\mathcal{Y}^N\times\mathcal{S}_2^N}\Pr\Big\{(\mathbf{y},\mathbf{s}_2)\in\mathcal{T}^N_{Q,P,V}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathcal{S}_1^N)\bigcap$$

$$\bigcap\bigcup_{m'\neq m}\bigcup_{\mathbf{s}'_1\in\mathcal{T}^N_{Q'_1}(S_1)}\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathbf{s}'_1)\Big\}\leq$$

$$\leq\sum_{m'\neq m}\sum_{(\mathbf{y},\mathbf{s}_2)\in\mathcal{Y}^N\times\mathcal{S}_2^N}\Pr\Big\{(\mathbf{y},\mathbf{s}_2)\in\mathcal{T}^N_{Q,P,V}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathcal{S}_1^N)\Big\}\times$$

$$\times\Pr\Big\{(\mathbf{y},\mathbf{s}_2)\in\bigcup_{\mathbf{s}'_1\in\mathcal{T}^N_{Q'_1}(S_1)}\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathbf{s}'_1)\Big\},$$

as the events in the brackets are independent.

Note that the first probability is different from zero iff $(\mathbf{y},\mathbf{s}_2)\in\mathcal{T}^N_{Q,P,V}(Y,S_2)$. In this case for N large enough

$$\Pr\Big\{(\mathbf{y},\mathbf{s}_2)\in\mathcal{T}^N_{Q,P,V}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathcal{S}_1^N)\Big\}=\frac{|\mathcal{T}^N_{Q,P,V}(U,X,S_1|\mathbf{y},\mathbf{s}_2)|}{|\mathcal{T}^N_{Q,P}(U,X,S_1)|}\leq$$

$$\leq(N+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{S}_1|}\exp\{-NI_{Q,P,V}(Y,S_2\wedge U,X,S_1)\}.$$

The second probability can be estimated in the following way:

$$\Pr\Big\{(\mathbf{y},\mathbf{s}_2)\in\bigcup_{\mathbf{s}'_1\in\mathcal{T}^N_{Q'_1}(S_1)}\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathbf{s}'_1)\Big\}\leq$$

$$\leq\Pr\Big\{(\mathbf{y},\mathbf{s}_2)\in\bigcup_{j}\bigcup_{\mathbf{s}'_1\in\mathcal{T}^N_{Q'_1,P'_0}(S_1|\mathbf{u}_j(m'))}\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathcal{U}_j^N,\mathbf{s}'_1)\Big\}\leq$$

$$\leq\sum_j\Pr\{(\mathbf{y},\mathbf{s}_2)\in\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathbf{u}_j(m'))\}\leq J\frac{|\mathcal{T}^N_{Q',P',V'}(U|\mathbf{y},\mathbf{s}_2)|}{|\mathcal{T}^N_{Q',P'}(U)|}\leq$$

$$\leq(N+1)^{|\mathcal{U}|}\exp\{-N(I_{Q',P',V'}(Y,S_2\wedge U)-I_{Q'_1,P'_0}(S_1\wedge U)-\delta/2)\}.$$

At last we get:

$$\mathbf{E}\Big|\mathcal{T}^N_{Q,P,V}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathcal{S}_1^N)\bigcap\bigcup_{m'\neq m}\bigcup_{\mathbf{s}'_1\in\mathcal{T}^N_{Q'_1}(S_1)}\mathcal{T}^N_{Q',P',V'}(Y,S_2|\mathcal{U}^N,\mathcal{X}^N,\mathbf{s}'_1)\Big|\leq$$

$$\leq(N+1)^{|\mathcal{U}|(|\mathcal{X}||\mathcal{S}_1|+1)}(|\mathcal{M}|-1)|\mathcal{T}^N_{Q,P,V}(Y,S_2)|\times$$

$$\times\exp\Big\{-N\Big(I_{Q,P,V}(Y,S_2\wedge U,X,S_1)+I_{Q',P',V'}(Y,S_2\wedge U)-I_{Q'_1,P'_0}(S_1\wedge U)-\delta/2\Big)\Big\}.$$

It follows from (16) that for any Q', P', V'

$$|\mathcal{M}|-1 \le \exp\Big\{N\Big(I_{Q',P',V'}(U\wedge S_2,Y)-I_{Q',P'_0}(U\wedge S_1)+D(Q'\circ P'\circ V'\|Q^*\circ P'\circ W)-E-\delta/2\Big)\Big\}$$

and we can write

$$\mathbf{E}A_m \le (N+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Y}||\mathcal{S}_1||\mathcal{S}_2|+|\mathcal{X}||\mathcal{S}_1|+1)} \sum_{Q_2\circ V} \sum_{Q',P',V':D(Q'\circ P'\circ V'\|Q^*\circ P'\circ W)<E}$$

$$\exp\Big\{N(E-D(Q'\circ P'\circ V'\|Q^*\circ P'\circ W)-H_{Q,P,V}(Y,S_2|U,X,S_1)\Big\}\exp\{NH_{Q,P,V}(Y,S_2)\}\times$$

$$\times\exp\Big\{N\Big(I_{Q',P',V'}(U\wedge S_2,Y)-I_{Q',P'_0}(U\wedge S_1)+D(Q'\circ P'\circ V'\|Q^*\circ P'\circ W)-E-\delta\Big)\Big\}\times$$

$$\times\exp\Big\{-N\Big(I_{Q,P,V}(Y,S_2\wedge U,X,S_1)+I_{Q',P',V'}(Y,S_2\wedge U)-I_{Q',P'_0}(S_1\wedge U)-\delta/2\Big)\Big\}\le$$

$$\le (N+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{Y}||\mathcal{S}_1||\mathcal{S}_2|+|\mathcal{X}||\mathcal{S}_1|+1)} \sum_{Q_2\circ V} \sum_{Q',P',V'} \exp\{-N\delta/2\}$$

which for N large enough proves (17) and hence the lemma.

References

[1] T. M. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information", *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1629-1638, 2002.

[2] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters", *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19-31, 1980.

[3] C. E. Shannon, "Channels with side information at the transmitter", *IBM J. Res. Develop*, vol. 2, pp. 289-293, 1958.

[4] C. Heegard and A. A. El Gamal, "On the capacity of computer memory with defects", *IEEE Transactions on Information Theory*, vol. IT-29, no. 5, pp. 731-739, 1983.

[5] M. Costa, "Writing on dirty paper", *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439-441, 1983.

[6] P. Moulin and J.A. O'Sullivan, "Information-theoretic analysis of information hiding", *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 563-593, Mar. 2003.

[7] P. Moulin and Y. Wang, "Capacity and random-coding exponents for channel coding with side information", *IEEE Transactions on Information Theory*, vol. 53, no. 4, pp. 1326- 1347, 2007.

[8] E. A. Haroutunian, "On bounds for E-capacity of DMC", *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4210-4220, 2007.

[9] E. A. Haroutunian, M. E. Haroutunian and A. N. Harutyunyan, "Reliability criteria in information theory and in statistical hypothesis testing", *Foundations and Trends in Communications and Information Theory*, vol. 4, no 2-3, pp. 97-263, 2008.

[10] M. E. Haroutunian, "New bounds for E-capacities of arbitrarily varying channel and channel with random parameter", *Transactions of the Institute for Informatics and Automation Problems of the NAS of RA, Mathematical Problems of Computer Sciences*, vol. 22, pp. 44-59, 2001.

[11] M. E. Haroutunian, "Estimates of E-capacity and capacity regions for multiple-access channel with random parameter", *Lecture Notes in Computer Science*, vol. 4123. Springer Verlag, pp. 196-217, 2006.

[12] M. E. Haroutunian and S. A. Tonoyan, "Random coding bound of information hiding *E*-capacity", *Proc. of IEEE International Symposium on Information Theory*. p. 536. USA, Chicago, 2004.

[13] M. E. Haroutunian and S. A. Tonoyan "On estimates of rate-reliability distortion function for information hiding system", *Transactions of the Institute for Informatics and Automation Problems of the NAS of RA, Mathematical Problems of Computer Science*. vol. 23, pp. 20-31, 2004.

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

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

[16] I. Csiszár, "The method of types", *IEEE Transactions on Information Theory*. vol. 44. no. 6, pp. 2505-2523, 1998.

Երկկողմանի վիճակներով ընդհատ առանց հիշողության կապուղու E-ունակության ստորին գնահատականը

Մ. Հարությունյան և Ա. Մուրադյան

Ամփոփում

Մենք ուսումնասիրում ենք երկկողմանի վիճակներով կապուղի, այն է վերջավոր մուտքի և ելքի այբուբեններով ընդհատ առանց հիշողության կապուղի, որը կախված է վիճակների պատահական հաջորդականությունից: Վիճակների հաջորդականության մասնակի ինֆորմացիան հասանելի է կոդավորիչին և ապակոդավորիչին: Այս ուսումնասիրությունները կիրառվում են ինֆորմացիան թաքցնող, ջրանշող համակարգերում: Այս մոդելի ունակությունը ստացվել է Կովերի և Չիանգի կողմից: Այս հոդվածում դուրս է բերվում E-ունակության պատահական կեղավորման գնահատականը դիտարկված մոդելի համար, որը կարելի է անվանել նաև վիճակներով ընդհանրացված կապուղի, քանի որ այն ներառում է պատահական պարամետրերով կապուղում չորս հնարավոր իրավիճակները: