Mathematical Problems of Computer Science 28, 2007, 18-33.

Random Coding Bound of Reversible Information Hiding E-capacity*

Mariam Haroutunian†, Smbat Tonoyan†, Oleksiy Koval‡, Svyatoslav Voloshynovskiy‡

† Institue for Informatics and Automation Problems of NAS of RA
e-mail: {armar, smbatt}@ipia.sci.am
‡University of Geneva, Switzerland
e-mail: {Oleksiy.Koval, svolos}@cui.unige.ch

Abstract

In this paper we consider the problem of reversible information hiding in the case when the attacker uses only discrete memoryless channels (DMC), the decoder knows only the class of channels, but not the DMC chosen by the attacker, the attacker knows the information-hiding strategy, probability distributions of all random variables, but not the side information.

We introduce the notion of reversible information hiding E-capacity, which expresses the dependence of the information hiding rate on the error probability exponent and the distortion levels for information hider, for attacker and for the host data approximation. The random coding bound for reversible information hiding E-capacity is found. When $E \to 0$ we obtain the lower bound for reversibility information hiding capacity.

In particular, we have analyzed two special cases of the general problem formulation, pure reversibility and pure message communications.

1 Introduction

Historically, the main goal of the information-theoretic analysis of communications over the state dependent channels was to establish the optimal information transmission conditions in terms of maximal achievable rates that can be attained in such protocols [1, 2, 3]. However, recently it was realized that in some practical applications, recovery of the channel state at the decoder might be required instead of a pure information transmission [4, 5]. A special attention this line of research has attracted in information hiding or watermarking community where it is known as reversible information hiding [6, 7, 8]. The information theoretic analysis in these referred publications was dedicated to the definition of the optimal achievable region of pure information transmission rates and attainable distortions of channel state recovery at decoder. In particular, in [6, 7] similarly to [9] an optimal protocol of joint pure information and channel interference communication is designed based on a modification of the Gel'fand-Pinsker coding and the corresponding rate-distortion pairs are established

*The work was partially supported by the CRDF and NFSAT grant GRSP 09/06.

by the proof of a source-channel coding theorem. Oppositely, the analysis of [8] considers a formulation where the communication protocol is specifically optimized to a particular pure information communication regime while reversibility is analyzed as a by-product of this design. Similarly to the previous cases, the rate-distortion region of pure information transmission and channel state recovery is defined.

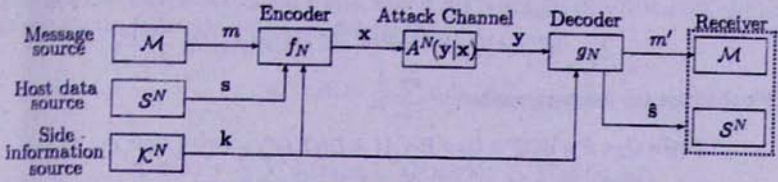

Figure 1. Reversible information hiding system

In this paper we would like to make one step forward with respect to the existing results and to establish the error exponents that can be attained in the reversible information hiding protocols in terms of E-capacity [10]-[14].

2 Notations and Definitions

We use capital letters $X, Y, S, \hat{S}, K, U$ for the random variables (RV) distributed over the discrete finite sets $\mathcal{X}, \mathcal{Y}, \mathcal{S}, \mathcal{K}, \mathcal{U}$ according to the probability distributions (PD) $P, V, Q, \ldots$ and lower case letters $x, y, s, \hat{s}, k, u$ for their realizations. For the N-length vectors of RVs we use the capital letters with superscript N (for example X^N) and we use the boldface letters for their realizations (for example $\mathbf{x} = (x_1, ..., x_N) \in \mathcal{X}^N$). The size of the set $\mathcal{X}$ we denote by $|\mathcal{X}|$.

The notation $|a|^+$ will be used for $\max(a, 0)$.

We shall use the following PD:

$$Q = Q_0 \circ Q_1 = \{Q(s,k) = Q_0(k)Q_1(s|k), s \in \mathcal{S}, k \in \mathcal{K}\},$$

$$Q_2 = \{Q_2(\hat{s}|s,k), \quad \hat{s}, s \in \mathcal{S}, k \in \mathcal{K}\}$$

$$P = P_0 \circ P_1 = \{P(u,x|\hat{s},s,k) = P_0(u|\hat{s},s,k)P_1(x|u,\hat{s},s,k), u \in \mathcal{U}, x \in \mathcal{X}, \hat{s}, s \in \mathcal{S}, k \in \mathcal{K}\},$$

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

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

The joint PD of the RV $U, X, \hat{S}, S, K$ will be denoted

$$Q \circ Q_2 \circ P = \{Q(s,k)Q_2(\hat{s}, |s,k)P(u,x|\hat{s},s,k), x \in \mathcal{X}, u \in \mathcal{U}, \hat{s}, s \in \mathcal{S}, k \in \mathcal{K}\}.$$

Information-theoretic quantities, such as conditional entropy of RV Y relative to RV X will be denoted $H_{P_0,V}(Y|X)$; the conditional mutual information of the RV S and $\hat{S}$ relative to RV K is

$$I_{Q,Q_2}(S \wedge \hat{S}|K) = \sum_{\hat{s},s,k} Q(s,k)Q_2(\hat{s}|s,k) \log \frac{Q_2(\hat{s}|s,k)}{\sum_s Q_2(\hat{s}|s,k)Q_1(s|k)} =$$

$$= H_{Q,Q_2}(\hat{S}|K) - H_{Q,Q_2}(\hat{S}|S,K) = H_Q(S|K) - H_{Q,Q_2}(S|\hat{S},K).$$

The informational divergence of the PD $Q^* = \{Q^*(s,k),\ s \in \mathcal{S},\ k \in \mathcal{K}\}$ and $Q = \{Q(s,k),\ s \in \mathcal{S},\ k \in \mathcal{K}\}$ on $\mathcal{S} \times \mathcal{K}$ is denoted by $D(Q\|Q^*)$ and the conditional informational divergence of the PD $Q^* \circ Q_2 \circ P \circ V$ and $Q^* \circ Q_2 \circ P \circ A$ by

$$D(Q^* \circ Q_2 \circ P \circ V \| Q^* \circ Q_2 \circ P \circ A) = D(V\|A|Q^*, Q_2, P) =$$

$$= \sum_{u,x,y,\hat{s},s,k} Q^*(s,k) Q_2(\hat{s}|s,k) P(u,x|\hat{s},s,k) A(y|x) \log \frac{V(y|x)}{A(y|x)}.$$

We shall use the following relation

$$D(Q \circ Q_2 \circ P \circ V \| Q^* \circ Q_2 \circ P \circ A) = D(Q\|Q^*) + D(V\|A|Q^*, Q_2, P). \qquad (1)$$

The type q_0 of a vector $\mathbf{k} \in \mathcal{K}^N$ is the PD $q_0 = \{Q_0(k) = \frac{n(k|\mathbf{k})}{N}, k \in \mathcal{K}\}$, where $N(k|\mathbf{k})$ is the number of repetitions of symbol k among $k_1, k_2, ..., k_N$.

The conditional type q_1 of $\mathbf{s}$ for given $\mathbf{k}$ is the PD $q_1 = \{Q_1(s|k),\ s \in \mathcal{S},\ k \in \mathcal{K}\}$ if $N(s,k|\mathbf{s},\mathbf{k}) = N(k|\mathbf{k})Q_1(s|k)$ for all $s \in \mathcal{S}$, $k \in \mathcal{K}$. The set of all vectors $\mathbf{k}$ of type q_0 we denote by $\mathcal{T}_{q_0}^N(K)$. The set of all vectors $\mathbf{s} \in \mathcal{S}^N$ of conditional type q_1 for given $\mathbf{k} \in \mathcal{T}_{q_0}^N(K)$ we denote by $\mathcal{T}_q^N(S|\mathbf{k})$. It is called also q-shell of vector $\mathbf{k}$.

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

The following properties will be used [15, 16]:

$$(N+1)^{-|\mathcal{Y}||\mathcal{U}||\mathcal{X}||\mathcal{S}|^2|\mathcal{K}|} \exp\{N H_{q,q_2,p,v}(Y|U,X,S,\hat{S},K)\} \leq$$

$$\leq |\mathcal{T}_{q,q_2,p,v}^N(Y|\mathbf{u}, \mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}), \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k}| \leq$$

$$\leq \exp\{N H_{q_2,q,p,v}(Y|U,X,S,\hat{S},K)\} \leq \exp\{N H_{q_2,q,p,v}(Y|X)\}, \qquad (2)$$

for $(\mathbf{s},\mathbf{k}) \in \mathcal{T}_q^N(S,K)$, $\hat{\mathbf{s}} \in \mathcal{T}_{q,q_2}^N(\hat{S}|\mathbf{s},\mathbf{k})$, $\mathbf{x} \in \mathcal{T}_{q,q_2,p}^N(X)$, $\mathbf{y} \in \mathcal{T}_{q,q_2,p,v}^N(Y|\mathbf{x})$,

$$Q^N(\mathbf{s},\mathbf{k}) = \exp\{-N(H_q(S,K) + D(q\|Q))\}. \qquad (3)$$

$$A^N(\mathbf{y}|\mathbf{x}) = \exp\{-N(H_{q,q_2,p,v}(Y|X) + D(v\|A|q,q_2,p))\}. \qquad (4)$$

Host data source (figure 1) is described by the RV S, which takes values in the discrete finite set $\mathcal{S}$ and generates N-length sequences S^N of independent and identically distributed (i.i.d.) components. The message source produces equiprobable and independent messages m from the message set $\mathcal{M}$, which must be transmitted to the receiver. The side information source is described by the RV K, which takes values in the discrete finite set $\mathcal{K}$, and in the most general case has the given joint PD $Q^* = \{Q^*(s,k),\ s \in \mathcal{S},\ k \in \mathcal{K}\}$ with the RV S. In particular case, when the side information is a cryptographic key, S and K are distributed independently. The side information in the form of N-length sequences K^N of i.i.d. components is available to the encoder and decoder. It is assumed that $Q^{*N}(\mathbf{s},\mathbf{k}) = \prod_{n=1}^{N} Q^*(s,k)$.

The *information hider* (encoder) embeds the message $m \in \mathcal{M}$ in the host data blocks $\mathbf{s} \in \mathcal{S}^N$ using the side information $\mathbf{k} \in \mathcal{K}^N$. The resulting codeword $\mathbf{x} \in \mathcal{X}^N$ is transmitted via attack channel A with the finite input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$. The *attacker* trying to change or remove the message m, transforms the data blocks $\mathbf{x} \in \mathcal{X}^N$ into corrupted blocks $\mathbf{y} \in \mathcal{Y}^N$. The decoder, possessing side information, decodes the data block $\mathbf{y} \in \mathcal{Y}^N$ and derives the message m and the approximation $\hat{\mathbf{s}}$ of the original data block, within the fixed distortion level.

Let the mappings $d_0 : \mathcal{S} \times \mathcal{S} \to [0, \infty)$, $d_1 : \mathcal{S} \times \mathcal{X} \to [0, \infty)$, $d_2 : \mathcal{X} \times \mathcal{Y} \to [0, \infty)$, be single-letter distortion functions. The distortion functions are supposed to be symmetric: $d_0(s, \hat{s}) = d_0(\hat{s}, s), d_1(s, x) = d_1(x, s), d_2(x, y) = d_2(y, x)$ for all $s \in \mathcal{S}, \hat{s} \in \mathcal{S}, x \in \mathcal{X}, y \in \mathcal{Y}$ and assume that $d_0(s, \hat{s}) = 0$, if $s = \hat{s}, d_1(s, x) = 0$, if $s = x, d_2(x, y) = 0$, if $x = y$. Distortion functions for the N-length vectors are defined as

$$d_0^N(\mathbf{s}, \hat{\mathbf{s}}) = \frac{1}{N} \sum_{n=1}^{N} d_0(s_n, \hat{s}_n),$$

$$d_1^N(\mathbf{s}, \mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} d_1(s_n, x_n), \; d_2^N(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{n=1}^{N} d_2(x_n, y_n).$$

Definition 1. The information hiding N-length code is a pair of mappings (f_N, g_N) subject to distortions Δ_0, Δ_1, where

$$f_N : \mathcal{M} \times \mathcal{S}^N \times \mathcal{K}^N \to \mathcal{X}^N$$

is the encoder, mapping host data block $\mathbf{s}$, a message m and side information $\mathbf{k}$ to a sequence $\mathbf{x} = f_N(\mathbf{s}, m, \mathbf{k})$, which satisfies to the following distortion constraint:

$$d_1^N(\mathbf{s}, f_N(\mathbf{s}, m, \mathbf{k})) \leq \Delta_1, \tag{5}$$

$$g : \mathcal{Y}^N \times \mathcal{K}^N \to \mathcal{M} \times \mathcal{S}^N$$

is the decoding, mapping the received sequence $\mathbf{y}$ and the side information $\mathbf{k}$ to a decoded message m' and sequence $\hat{\mathbf{s}}$, which satisfies to the following distortion constraint:

$$d_0^N(\mathbf{s}, \hat{\mathbf{s}}) \leq \Delta_0. \tag{6}$$

Note that the definition of the distortion constraint (5) means that the maximum distortion constraint with respect to $\mathbf{s}, \mathbf{k}$ and m is used, as distinct from [17], where the average distortion constraint is considered, and the maximum distortion constraint is mentioned as more difficult case.

The N-length memoryless expression for the attack channel A is:

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

An attack channel, subject to distortion Δ_2, satisfies to the following condition:

$$\sum_{\mathbf{x} \in \mathcal{X}^N} \sum_{\mathbf{y} \in \mathcal{Y}^N} d_2^N(\mathbf{x}, \mathbf{y}) A^N(\mathbf{y}|\mathbf{x}) p^N(\mathbf{x}) \leq \Delta_2.$$

Definition 2. The nonnegative number

$$R = \frac{1}{N} \log |\mathcal{M}|$$

is called *the information hiding code rate.*

We use the following notation $m = \overline{1,|\mathcal{M}|}$ for $m = 1, 2, \ldots, |\mathcal{M}|$.

For any Q and Δ_0 denote by $\mathcal{Q}_2(Q, \Delta_0)$ the set of all conditional PDs Q_2, for which the following inequality takes place

$$\sum_{s,\hat{s},k} Q(s,k)Q_2(\hat{s}|s,k)d_0(s,\hat{s}) \leq \Delta_0. \tag{7}$$

We use an auxiliary RV U, taking values in the finite set $\mathcal{U}$ and forming the Markov chain $(K, S, \hat{S}, U) \rightarrow X \rightarrow Y$.

Definition 3. A memoryless covert channel P, subject to distortion level Δ_1, is a PD $P = \{P(u,x|s,\hat{s},k),\ u \in \mathcal{U},\ x \in \mathcal{X},\ s,\hat{s} \in \mathcal{S},\ k \in \mathcal{K}\}$ such that for any Q and $Q_2 \in \mathcal{Q}_2(Q, \Delta_0)$

$$\sum_{u,x,s,\hat{s},k} Q(s,k)Q_2(\hat{s}|s,k)P(u,x|s,\hat{s},k)d_1(s,x) \leq \Delta_1. \tag{8}$$

Denote by $\mathcal{P}(Q, Q_2, \Delta_1)$ the set of all covert channels, subject to distortion level Δ_1. The N-length memoryless expression for the covert channel P is:

$$P^N(\mathbf{u},\mathbf{x}|\mathbf{s},\hat{\mathbf{s}},\mathbf{k}) = \prod_{n=1}^{N} P(u_n, x_n|s_n, \hat{s}_n, k_n).$$

Definition 4. A memoryless attack channel A, subject to distortion level Δ_2, under the condition of covert channel $P \in \mathcal{P}(Q, Q_2, \Delta_1)$, is defined by a PD $A = \{A(y|x),\ y \in \mathcal{Y},\ x \in \mathcal{X}\}$, for which for Q and $Q_2 \in \mathcal{Q}_2(Q, \Delta_0)$

$$\sum_{u,x,s,\hat{s},k,y} Q(s,k)Q_2(\hat{s}|s,k)P(u,x|s,\hat{s},k)(x)A(y|x)d_2(x,y) \leq \Delta_2.$$

Denote by $\mathcal{A}(Q, Q_2, P, \Delta_2)$ the set of all attack channels, under the condition of covert channel $P \in \mathcal{P}(Q, Q_2, \Delta_1)$ and subject to distortion level Δ_2. The sets $\mathcal{Q}_2(Q, \Delta_0), \mathcal{P}(Q, Q_2, \Delta_1)$ and $\mathcal{A}(Q, Q_2, P, \Delta_2)$ are defined by linear inequality constraints and hence are convex.

Denote by $g^{-1}_{N,\mathbf{k}}(m, \hat{\mathbf{s}})$ the set of all $\mathbf{y}$ which are decoded into $(m, \hat{\mathbf{s}})$:

$$g^{-1}_{N,\mathbf{k}}(m, \hat{\mathbf{s}}) = \{\mathbf{y} :\ g_N(\mathbf{y}, \mathbf{k}) = (m, \hat{\mathbf{s}})\}.$$

Definition 5. The probability of erroneous reconstruction of the message $m \in \mathcal{M}$ and the approximation of data block $\mathbf{s} \in \mathcal{S}^N$ for $\mathbf{k} \in \mathcal{K}^N$ via channel A is:

$$e(m, \mathbf{s}, \mathbf{k}) = e(f_N, g_N, N, m, \mathbf{s}, \mathbf{k}, A, \Delta_0) = 1 - A^N \left\{ \bigcup_{\hat{\mathbf{s}}:\ d(\mathbf{s},\hat{\mathbf{s}}) \leq \Delta_0} g^{-1}_{N,\mathbf{k}}(m, \hat{\mathbf{s}}) | f_N(m, \mathbf{s}, \mathbf{k}) \right\}. \tag{9}$$

The error probability of the message m averaged over all $(\mathbf{s}, \mathbf{k}) \in \mathcal{S}^N \times \mathcal{K}^N$ equals to

$$e(m, A) = e(f_N, g_N, N, m, Q^*, A, \Delta_0) = \sum_{(\mathbf{s},\mathbf{k}) \in\ \mathcal{S}^N \times \mathcal{K}^N} Q^{*N}(\mathbf{s}, \mathbf{k}) e(m, \mathbf{s}, \mathbf{k}).$$

Denote by $\mathbf{\Delta} = [\Delta_0, \Delta_1, \Delta_2]$ the collection of distortion levels, fixed for the current system.

The error probability of the code, for any message $m \in \mathcal{M}$, maximal over all attack channels from $\mathcal{A}(Q, Q_2, P, \Delta_2)$ is denoted by:

$$e(m) = e(f_N, g_N, N, m, Q^*, \Delta) = \max_{A \in \mathcal{A}(Q,Q_2,P,\Delta_2)} e(m, A).$$

The *maximal error probability* of the code, maximal over all attack channels from $\mathcal{A}(Q, Q_2, P, \Delta_2)$ equals to:

$$e = e(f_N, g_N, N, Q^*, \Delta) = \max_{m \in \mathcal{M}} e(m), \tag{10}$$

and the *average error probability* of the code, maximal over all attack channels from $\mathcal{A}(Q, Q_2, P, \Delta_2)$ equals to:

$$\overline{e} = \overline{e}(f_N, g_N, N, Q^*, \Delta) = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} e(m). \tag{11}$$

3 Reversible Information Hiding E-Capacity

Consider codes the maximal error probability of which exponentially decreases with the given exponent $E > 0$, (called *the reliability*)

$$e \leq \exp\{-NE\}. \tag{12}$$

Denote by $M(Q^*, E, N, \Delta)$ the highest volume of the code, satisfying the condition (12) for the given reliability E and the distortion levels Δ.

The rate-reliability-distortion function, which we call *reversible information hiding E-capacity* by analogy with the E-capacity of ordinary channel [10], is defined as:

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

By $C(Q^*, E, \Delta)$ and $\overline{C}(Q^*, E, \Delta)$ we denote the reversible information hiding E-capacity for maximal and average error probabilities respectively.

In this paper the lower bound of reversible information hiding E-capacity for maximal and average error probabilities are constructed.

Consider the following function, which we call *the random coding bound*

$$R_r(Q^*, E, \Delta) =$$

$$= \min_{Q} \max_{Q_2 \in \mathcal{Q}_2(Q,\Delta_0)} \max_{P \in \mathcal{P}(Q,Q_2,\Delta_1)} \min_{A \in \mathcal{A}(Q,Q_2,P,\Delta_2)} \min_{V: D(Q \circ Q_2 \circ P \circ V \| Q^* \circ Q_2 \circ P \circ A) \leq E} \Big| I_{Q,Q_2,P,V}(Y \wedge U|K) -$$

$$- I_{Q,Q_2,P_0}(S \wedge U, \hat{S}|K) + D(Q \circ Q_2 \circ P \circ V \| Q^* \circ Q_2 \circ P \circ A) - E \Big|^{+}. \tag{13}$$

Theorem. *For all $E > 0$, for reversible information hiding system with distortion levels Δ*

$$R_r(Q^*, E, \Delta) \leq C(Q^*, E, \Delta) \leq \overline{C}(Q^*, E, \Delta).$$

Corollary 1. *When $E \to 0$, from (13) we obtain the lower bound of reversible information hiding capacity:*

$$R_r(Q^*, E, \Delta) =$$

$$= \max_{Q_2\in\mathcal{Q}_2(Q^*,\Delta_0)} \max_{P\in\mathcal{P}(Q^*,Q_2,\Delta_1)} \min_{A\in\mathcal{A}(Q,Q_2,P,\Delta_2)} \left\{I_{Q^*,Q_2,P,A}(Y\wedge U|K) - I_{Q^*,Q_2,P}(S\wedge U,\hat{S}|K)\right\}. \tag{14}$$

Corollary 2: pure reversibility. *If* $\Delta_0 = 0$ *from (13) we have*

$$\begin{aligned} R_r(Q^*,E,\mathbf{\Delta}) = \min_{Q} \max_{P\in\mathcal{P}(Q,\Delta_1)} \min_{A\in\mathcal{A}(Q,P,\Delta_2)} \min_{V:D(Q\circ P\circ V\|Q^*\circ P\circ A)\le E} |I_{Q,P,V}(Y\wedge U|K)- \\ -H_Q(S|K) + D(Q\circ P\circ V\|Q^*\circ P\circ A) - E|^+ . \end{aligned} \tag{15}$$

Corollary 3: pure message communications. *If* $\Delta_0 \to \infty$ *then*

$$\begin{aligned} R_r(Q^*,E,\mathbf{\Delta}) = \min_{Q} \max_{P\in\mathcal{P}(Q,\Delta_1)} \min_{A\in\mathcal{A}(Q,P,\Delta_2)} \min_{V:D(Q\circ P\circ V\|Q^*\circ P\circ A)\le E} |I_{Q,P,V}(Y\wedge U|K)- \\ -I_{Q,P}(S\wedge U|K) + D(Q\circ P\circ V\|Q^*\circ P\circ A) - E|^+ . \end{aligned} \tag{16}$$

In (15) and (16) $P = \{P(u,x|s,k),\ u\in\mathcal{U}, x\in\mathcal{X}, s\in\mathcal{S}, k\in\mathcal{K}\}$ and $\mathbf{\Delta} = (\Delta_1,\Delta_2)$.

4 Proof of the Theorem

The theorem is proved by the Shannon's random coding arguments, using the method of types, covering lemma and demonstration of a generalization of packing lemma [13, 15, 16].

To prove the random coding bound, we must show the existence of a code with R satisfying (13) and

$$e \le \exp\{-N(E-\varepsilon)\},$$

for any $0<\varepsilon<E$.

We will construct the encoding and the decoding and explore the errors caused by each.

For encoding we use the idea of Gelfand-Pinsker [2], which is known as random bin coding technique [18].

The decoding is based on *minimum divergence* method, first introduced by E. Haroutunian [19] and developed by M. Haroutunian [13, 14]. The expansion of this method, adopted for data hiding systems, can be considered as *semi-universal decoding*, as the decoder needs to know only the specific class of channels, instead of the channel, used by attacker.

Encoding

Step 1. Denote by $\mathcal{Q}(Q^*,E) = \{q:\ D(q\|Q^*)\le E\}$ and

$$\mathcal{T}^E_{Q^*}(S,K) = \bigcup_{q\in\mathcal{Q}(Q^*,E)} \mathcal{T}^N_q(S,K). \tag{17}$$

We will construct the code only for $(\mathbf{s},\mathbf{k})$ from $\mathcal{T}^E_{Q^*}(S,K)$, because for sufficiently large N, the probability of $(\mathbf{s},\mathbf{k})\notin\mathcal{T}^E_{Q^*}(S,K)$ is exponentially small:

$$Q^{*N}\left\{\bigcup_{q\notin\mathcal{Q}(Q^*,E)} \mathcal{T}^N_q(S,K)\right\} =$$

$$= \sum_{q\notin\mathcal{Q}(Q^*,E)} Q^{*N}\{\mathcal{T}^N_q(S,K)\} \le \sum_{q\notin\mathcal{Q}(Q^*,E)} \exp\{-ND(q\|Q^*)\} <$$

$$< (N+1)^{|\mathcal{S}||\mathcal{K}|}\exp\{-NE\} \leq \exp\{-N(E-\varepsilon_1)\}, \tag{18}$$

where $\varepsilon_1 > 0$.

Step 2. Denote

$$\hat{q}_1(\hat{s}|k) = \sum_{s} q_2(\hat{s}|s,k) q_1(s|k).$$

Covering lemma. *For every type q, conditional type q_2, vector $\mathbf{k} \in \mathcal{K}^N$ there exists a collection of vectors*

$$\{\hat{\mathbf{s}}_j \in \mathcal{T}^N_{\hat{q}_1}(\hat{S}|\mathbf{k}),\ j = \overline{1, J_1}\},$$

where

$$J_1 = \exp\left\{N\left(I_{q,q_2}(S \wedge \hat{S}|K) + \delta/2\right)\right\},\ \delta > 0$$

such that the set $\left\{\mathcal{T}^N_{q,q_2}(S|\hat{\mathbf{s}}_j, \mathbf{k})\ j = \overline{1, J_1}\right\}$ covers $\mathcal{T}^N_{q_1}(S|\mathbf{k})$ for N large enough:

$$\mathcal{T}^N_{q_1}(S|\mathbf{k}) \subset \bigcup_{j=1}^{J} \mathcal{T}^N_{q,q_2}(S|\hat{\mathbf{s}}_j, \mathbf{k}).$$

For the proof of covering lemma see [15].

For type $q \in \mathcal{Q}(Q^*, E)$ and conditional type $q_2 \in \mathcal{Q}_2(q, \Delta_0)$ denote

$$\mathcal{S}(q, q_2, j) = \mathcal{T}^N_{q,q_2}(S|\hat{\mathbf{s}}_j, \mathbf{k}) \setminus \bigcup_{j'<j} \mathcal{T}^N_{q,q_2}(S|\hat{\mathbf{s}}_{j'}, \mathbf{k}),$$

therefore for the vectors $\mathbf{s}$ from $\mathcal{S}(q, q_2, j)$, we put into correspondence the vector $\hat{\mathbf{s}}_j, j \in [1, J_1]$.

Taking into account the inequality (7), we can show that for types $q \in \mathcal{Q}(Q^*, E), q_2 \in \mathcal{Q}_2(q, \Delta_0)$ and any $j = \overline{1, J_1}$, $\mathbf{s} \in \mathcal{S}(q, q_2, j)$, $\hat{\mathbf{s}}_j$

$$d_0(\mathbf{s}, \hat{\mathbf{s}}_j) = N^{-1} \sum_{s,\hat{s}} n(s, \hat{s}|\mathbf{s}, \hat{\mathbf{s}}_j) d_0(s, \hat{s}) =$$

$$= \sum_{s,\hat{s},k} q(s,k) q_2(\hat{s}|s,k) d_0(s,\hat{s}) \leq \Delta_0,\ j = \overline{1, J_1}.$$

Step 3. Fix the type $p = p_0 \circ p_1 \in \mathcal{P}(q, q_2, \Delta_1)$. For fixed p_0, E, for each types $q \in \mathcal{Q}(Q^*, E)$, $q_2 \in \mathcal{Q}(q, \Delta_0)$ and vectors $\hat{\mathbf{s}}$, $\mathbf{k}$ we choose independently, at random from $\mathcal{T}^N_{q_0,p_0}(U|\hat{\mathbf{s}}, \mathbf{k})$ $|\mathcal{M}|$ collections $\mathcal{J}_2(m), m = \overline{1, |\mathcal{M}|}$ of vectors $\mathbf{u}_j(m)$, $j = \overline{1, J_2}$, where

$$J_2 = \exp\left\{N\left(I_{q,q_2,p_0}(S \wedge U|\hat{S}, K) + \delta/2\right)\right\}\ (\delta > 0).$$

Then for each $\mathbf{s} \in \mathcal{T}^N_q(S|\mathbf{k})$ we choose such $\mathbf{u}_j(m)$ from $\mathcal{J}_2(m)$, that $\mathbf{u}_j(m) \in \mathcal{T}^N_{q,q_2,p_0}(U|\mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})$. Denote this vector by $\mathbf{u}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})$.

If for some $\mathbf{s}$ there is no such vector in $\mathcal{J}_2(m)$, we choose $\mathbf{u}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})$ at random from the $\mathcal{T}^N_{q,q_2,p_0}(U|\mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})$. Denote the probability of such event by $\Pr\{b_{q,q_2,p_0}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})\}$.

$$\Pr\{b_{q,q_2,p_0}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})\} = \Pr\left\{\bigcap_{j=1}^{J_2} \mathbf{u}_j(m) \notin \mathcal{T}^N_{q,q_2,p_0}(U|\mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})\right\} \leq$$

$$\leq \prod_{j=1}^{J_2} [1 - \Pr\{\mathbf{u}_j(m) \in \mathcal{T}^N_{q,q_2,p_0}(U|\mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})\}] \leq \left[1 - \frac{|\mathcal{T}^N_{q,q_2,p_0}(U|\mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})|}{|\mathcal{T}^N_{q_0,p_0}(U|\hat{\mathbf{s}}, \mathbf{k})|}\right]^{J_2} \leq$$

$$\leq [1-\exp\{-N(I_{q,q_2,p_0}(S\wedge U|\hat{S},K)+\delta/4)\}]^{\exp\{N(I_{q,q_2,p_0}(S\wedge U|\hat{S},K)+\delta/2)\}}.$$

Using the inequality $(1-t)^n \leq \exp\{-nt\}$, which is true for any n and $t \in (0,1)$, we can see that

$$\Pr\{b_{q,q_2,p_0}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\} \leq \exp\{-\exp\{N\delta/4\}\}. \tag{19}$$

Notice that for each m the code contains

$$J = J_1 \times J_2 = \exp\left\{N\left(I_{q,q_2}(S\wedge\hat{S}|K)+I_{q,q_2,p_0}(S\wedge U|\hat{S},K)+\delta/2+\delta/2\right)\right\} =$$

$$= \exp\left\{N\left(I_{q,q_2,p_0}(S\wedge U,\hat{S}|K)+\delta\right)\right\}\ (\delta>0)$$

vectors $\mathbf{u}$.

Step 4. The codeword $\mathbf{x}$ is constructed in the following way. For each $m=\overline{1,|\mathcal{M}|}$, $\hat{\mathbf{s}}$, $\mathbf{s}$ and $\mathbf{k}$ we choose at random a vector $\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})$ from $\mathcal{T}^N_{q,q_2,p}(X|\mathbf{u}(m,\mathbf{s},\hat{\mathbf{s}}_j,\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})$.

It is easy to show that the constructed encoding satisfies the distortion constraint. Really, for types $p\in\mathcal{P}(q,q_2,\Delta_1)$, $q\in\mathcal{Q}(Q^*,E)$, $q_2\in\mathcal{Q}_2(q,\Delta_0)$, taking into account the inequality (8), we have

$$d_1^N(\mathbf{s},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})) = N^{-1}\sum_{s,x} n(s,x|\mathbf{s},\mathbf{x})d_1(s,x) =$$

$$= \sum_{u,x,s,\hat{s},k} q(s,k)q_2(\hat{s}|s,k)p(u,x|s,\hat{s},k)d_1(s,x) \leq \Delta_1.$$

Denote by $e_E(m)$ *the encoding error probability* for any $m\in\mathcal{M}$:

$$e_E(m) \leq \sum_{(\mathbf{s},\mathbf{k})\notin\mathcal{T}^E_{Q^*}(S,K)} Q^{*N}(\mathbf{s},\mathbf{k}) + \sum_{(\mathbf{s},\mathbf{k})\in\mathcal{T}^E_{Q^*}(S,K)} Q^{*N}(\mathbf{s},\mathbf{k})\Pr\{b_{q,q_2,p_0}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\}.$$

Now taking into account (17), (18) and (19):

$$e_E(m) \leq \exp\{-N(E-\varepsilon_1)\} + \sum_{q\in\mathcal{Q}(Q^*,E)} Q^{*N}\left\{\mathcal{T}^N_q(S,K)\right\}\exp\{-\exp\{N\delta/4\}\}.$$

As the number of types q in $\mathcal{Q}(Q^*,E)$ does not exceed $(N+1)^{|S||K|}$ according to type counting lemma [15, 16] and $Q^{*N}\left\{\mathcal{T}^N_q(S,K)\right\}\leq 1$, we can write

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

$$\leq \exp\{-N(E-\varepsilon_1)\} + \exp\{-\exp\{N\delta/4\}+\varepsilon_1\}, \tag{20}$$

for N large enough.

The attacker chooses the attack channel A from the set $\mathcal{A}(q,q_2,p,\Delta_2)$ as he knows probability distributions of all random variables. It is clear, that in this case the average distortion constraint is satisfied:

$$\sum_{\mathbf{x}\in\mathcal{X}^N}\sum_{\mathbf{y}\in\mathcal{Y}^N} d_2^N(\mathbf{x},\mathbf{y})A^N(\mathbf{y}|\mathbf{x})p^N(\mathbf{x}) =$$

$$= \mathbf{E}d_2^N(X^N,Y^N) = \frac{1}{N}\sum_{n=1}^{N}\mathbf{E}d_2(x_n,y_n) =$$

$$= \sum_{u,x,s,\hat{s},k,y} q(s,k)q_2(\hat{s}|s,k)p(u,x|s,\hat{s},k)A(y|x)d_2(x,y) \leq \Delta_2.$$

Decoding

For brevity the vector pair $\mathbf{u}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})$ we denote by $\mathbf{u}.\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})$.

Decoding rule is the following: each $\mathbf{y}$ and $\mathbf{k}$ are decoded to such m and $\hat{\mathbf{s}}$, for which $\mathbf{y}\in\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})$, where q,q_2,p,v are such that $\min\limits_{A\in\mathcal{A}(q,q_2,p,\Delta_2)} D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A)$ is minimal.

The decoder can make an error when the message $m\in\mathcal{M}$ is transmitted in the case of $(\mathbf{s},\mathbf{k})\in\mathcal{T}^E_{Q^*}(S,K)$, but there exist the types q',q_2',p',v', vector $\mathbf{s}'$ and pair $(m',\hat{\mathbf{s}}')$, for which $m'\neq m$ or $m'=m$, $d_0(\mathbf{s},\hat{\mathbf{s}}')>\Delta_0$, with

$$\mathbf{y}\in\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\bigcap\mathcal{T}_{q',q_2',p',v'}(Y|\mathbf{u}',\mathbf{x}'(m',\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k}),\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})$$

and

$$\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)} D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)\leq\min_{A\in\mathcal{A}(q,q_2,p,\Delta_2)} D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A). \quad (21)$$

Denote by

$$\mathcal{D}=\{q,q',p,p',q_2,q_2',v,v':(21)\text{ is valid}\}.$$

The decoding error probability $e_D(m)$ of message $m\in\mathcal{M}$, maximal over all attack channels $A\in\mathcal{A}(q,q_2,p,\Delta_2)$ can be estimated in the following way:

$$e_D(m)\leq\max_{A\in\mathcal{A}(q,q_2,p,\Delta_2)}\sum_{(\mathbf{s},\mathbf{k})\in\mathcal{T}^E_{Q^*}(S,K)} Q^{*N}(\mathbf{s},\mathbf{k})\times$$

$$\times A^N\Bigg\{\bigcup_{\mathcal{D}}\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\bigcap$$

$$\bigcap\bigcup_{(m',\hat{\mathbf{s}}'):\left\{\substack{m'\neq m\text{ or}\\ m'=m,\ d_0(\mathbf{s},\hat{\mathbf{s}}')>\Delta_0}\right\}}\ \bigcup_{\mathbf{s}'\in\mathcal{T}^N_{q',q_2'}(S|\hat{\mathbf{s}}',\mathbf{k})}\mathcal{T}^N_{q',q_2',p',v'}(Y|\mathbf{u}',\mathbf{x}'(m',\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k}),\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\Big|\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\Bigg\}\leq$$

$$\leq\sum_{\mathcal{D}}\Big|\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\bigcap$$

$$\bigcap\bigcup_{(m',\hat{\mathbf{s}}'):\left\{\substack{m'\neq m\text{ or}\\ m'=m,\ d_0(\mathbf{s},\hat{\mathbf{s}}')>\Delta_0}\right\}}\ \bigcup_{\mathbf{s}'\in\mathcal{T}^N_{q',q_2'}(S|\hat{\mathbf{s}}',\mathbf{k})}\mathcal{T}^N_{q',q_2',p',v'}(Y|\mathbf{u}',\mathbf{x}'(m',\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k}),\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\Big|\times$$

$$\times\max_{A\in\mathcal{A}(q,q_2,p,\Delta_2)}\sum_{(\mathbf{s},\mathbf{k})\in\mathcal{T}^E_{Q^*}(S,K)} Q^{*N}(\mathbf{s},\mathbf{k})A^N(\mathbf{y}|\mathbf{x}).$$

The last inequality follows from (4), because for fixed types of $\mathbf{x}$ and $\mathbf{y}$ the probability $A^N(\mathbf{y}|\mathbf{x})$ is constant.

For the estimation of decoding error probability we use the statement of the following lemma.

Packing Lemma. *For any $E>2\delta\geq 0$, fixed $q\in\mathcal{Q}(Q^*,E)$, $q_2\in\mathcal{Q}_2(q,\Delta_0)$ and covert channel $p\in\mathcal{P}(q,q_2,\Delta_1)$, there exists a code with*

$$|\mathcal{M}|=\exp\Bigg\{N\min_{A\in\mathcal{A}(q,q_2,p,\Delta_2)}\ \min_{v:D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A)\leq E}|I_{q,q_2,p,v}(Y\wedge U|K)-$$

$$-I_{q,q_2,p,v}(S\wedge U,\hat{S}|K)+D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A)-E+2\delta|^+\Big\}, \tag{22}$$

such that

1) *for each* $\mathbf{k}$, $\mathbf{s}$ *and* $\hat{\mathbf{s}}$, *the vector pairs* $\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k})$ *are distinct for different* $m\in\mathcal{M}$,
2) *for sufficiently large* N *the following inequality holds* for any *types* $q'\in\mathcal{Q}(Q^*,E)$, $q_2'\in\mathcal{Q}_2(q',\Delta_0)$, $p'\in\mathcal{P}(q',q_2',\Delta_1)$, v,v', *and for all* $m=\overline{1,|\mathcal{M}|}$, $(\mathbf{s},\mathbf{k})\in\mathcal{T}_q^N(S,K)$, $\hat{\mathbf{s}}\in\mathcal{T}_{q_2}^N(\hat{S}|\mathbf{s},\mathbf{k})$

$$\Bigg|\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})\bigcap$$

$$\bigcap\bigcup_{(m',\hat{s}'):\left\{\substack{m'\neq m,\ \text{or}\\ m'=m,\ d_0(\hat{\mathbf{s}},\hat{\mathbf{s}}')>\Delta_0}\right\}}\ \bigcup_{\mathbf{s}'\in\mathcal{T}^N_{q',q_2'}(S|\hat{\mathbf{s}}',\mathbf{k})}\mathcal{T}^N_{q',q_2',p',v'}(Y|\mathbf{u}',\mathbf{x}'(m',\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k}),\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\Bigg|\leq$$

$$\leq|\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u},\mathbf{x}(m,\mathbf{s},\hat{\mathbf{s}},\mathbf{k}),\mathbf{s},\hat{\mathbf{s}},\mathbf{k})|\times$$

$$\times\exp\left\{-N\left|E-\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)}D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)\right|^+\right\}. \tag{23}$$

In other words, the lemma guarantees the existence of a good code, the codewords of which must be far from each other in a sense that all q,v-shells have possibly small intersections.

Proof of the lemma is presented in the next section. Now we can bound the decoding error probability using (3), (21) and (23):

$$\epsilon_D(m)\leq\sum_{\mathcal{D}}\exp\{NH_{q,q_2,p,v}(Y|X)\}\times$$

$$\times\exp\left\{-N\left(E-\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)}D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)\right)\right\}\times$$

$$\times\max_{A\in\mathcal{A}(q,q_2,p,\Delta_2)}[\exp\{-N(H_{q,q_2,p,v}(Y|X)+D(q\|Q^*)+D(v\|A|q,q_2,p)\}]=$$

$$=\sum_{\mathcal{D}}\exp\{N(H_{q,q_2,p,v}(Y|X)-E-H_{q,q_2,p,v}(Y|X)-$$

$$-\min_{A\in\mathcal{A}(q,q_2,p,\Delta_2)}D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A)+$$

$$+\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)}D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)\Big)\Big\}\leq$$

$$\leq(N+1)^{2|\mathcal{S}||\mathcal{K}|+2|\mathcal{S}|^2|\mathcal{K}|+2|\mathcal{X}||\mathcal{Y}|+2|\mathcal{U}||\mathcal{X}||\mathcal{S}|^2|\mathcal{K}|+|\mathcal{U}||\mathcal{X}||\mathcal{Y}||\mathcal{S}|^2|\mathcal{K}|}\exp\{-NE\}\leq$$

$$\leq\exp\{-N(E-\varepsilon_2)\}, \tag{24}$$

for N large enough, where $\varepsilon_2>0$.

Taking into account (20) and (24), the error probability of the message $m\in\mathcal{M}$ can be bounded in the following way:

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

$$\leq \exp\{-N(E-\varepsilon)\}, \quad \varepsilon > 0.$$

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

The theorem is proved.

5 Proof of the Packing Lemma

First we shall show that if for any code (23) is satisfied for any q', q_2', p', v, v', then point 1) of lemma is true. To prove that, it is enough to choose $q' = q$, $q_2' = q_2$, $p' = p$, $v' = v$, and $\min_{A \in \mathcal{A}(q', q_2', p', \Delta_2)} D(q' \circ q_2' \circ p' \circ v' \| Q^* \circ q_2' \circ p' \circ A) < E$. Really, if we suppose that for some $m \neq m'$ $\mathbf{u}, \mathbf{x}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k}) = \mathbf{u}', \mathbf{x}'(m', \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})$ then in (23) we shall have

$$\left|\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u}, \mathbf{x}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k}), \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})\right| \leq$$

$$\leq |\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u}, \mathbf{x}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k}), \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k})| \times$$

$$\times \exp\left\{-N(E - \min_{A \in \mathcal{A}(q', q_2', p', \Delta_2)} D(q' \circ q_2' \circ p' \circ v' \| Q^* \circ q_2' \circ p' \circ A))\right\},$$

which is impossible.

Now we shall show that (23) is true for any q', q_2', p' and v'. First consider the case when q', q_2', p' and v' are such that $\min_{A \in \mathcal{A}(q', q_2', p', \Delta_2)} D(q' \circ q_2' \circ p' \circ v' \| Q^* \circ q_2' \circ p' \circ A) \geq E$. Then

$$\exp\left\{-N\left|E - \min_{A \in \mathcal{A}(q', q_2', p', \Delta_2)} D(q' \circ q_2' \circ p' \circ v' \| Q^* \circ q_2' \circ p' \circ A)\right|^+\right\} = 1$$

and (23) is valid for any $|\mathcal{M}|$. It is left to prove the inequality (23) for

$$\mathcal{D}'(Q^*, E) = \{q', q_2', p', v' : \min_{A \in \mathcal{A}(q', q_2', p', \Delta_2)} D(q' \circ q_2' \circ p' \circ v' \| Q^* \circ q_2' \circ p' \circ A) < E\}.$$

Denote

$$B_m(q', q_2', p', v, v') = (N+1)^{|\mathcal{Y}||\mathcal{U}||\mathcal{X}||\mathcal{S}|^2|\mathcal{K}|} \times$$

$$\times \exp\left\{N\left(E - \min_{A \in \mathcal{A}(q', q_2', p', \Delta_2)} D(q' \circ q_2' \circ p' \circ v' \| Q^* \circ q_2' \circ p' \circ A) - H_{q,q_2,p,v}(Y|U, X, S, \hat{S}, K)\right)\right\} \times$$

$$\times \max_{(\mathbf{s},\mathbf{k}) \in \mathcal{T}^N_q(S,K)} \left|\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{u}, \mathbf{x}(m, \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k}), \mathbf{s}, \hat{\mathbf{s}}, \mathbf{k}) \bigcap\right.$$

$$\left.\bigcap \bigcup_{(m', \hat{\mathbf{s}}'): \left\{ {m' \neq m \text{ or} \atop m' = m,\ d_0(\mathbf{s}, \hat{\mathbf{s}}') > \Delta_0} \right\}} \bigcup_{\mathbf{s}' \in \mathcal{T}^N_{q', q_2'}(S|\hat{\mathbf{s}}', \mathbf{k})} \mathcal{T}^N_{q', q_2', p', v'}(Y|\mathbf{u}', \mathbf{x}'(m', \mathbf{s}', \hat{\mathbf{s}}', \mathbf{k}), \mathbf{s}', \hat{\mathbf{s}}', \mathbf{k})\right|$$

and

$$B_m = \sum_{v} \sum_{q', q_2', p', v' \in \mathcal{D}'(Q^*, E)} B_m(q', q_2', p', v, v').$$

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

Notice that if for some code the following inequality holds

$$\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} B_m \leq \frac{1}{2}, \tag{25}$$

then $B_m \leq 1$ for at least $|\mathcal{M}|/2$ indices m. Further, if we denote such indices by m^* then $B_{m^*} \leq B_m \leq 1$ for every index m^*. Hence the lemma will be proved if for any $|\mathcal{M}|$ with

$$\begin{aligned}
&2\exp\left\{N \min_{A\in\mathcal{A}(q,q_2,p,\Delta_2)} \min_{v: D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A)\leq E} |I_{q,q_2,p,v}(Y\wedge U|K)-\right.\\
&\left. -I_{q,q_2,p,v}(S\wedge U,\hat{S}|K) + D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A) - E - 2\delta|^+\right\} \leq |\mathcal{M}| \leq\\
&\leq \exp\left\{N \min_{A\in\mathcal{A}(q,q_2,p,\Delta_2)} \min_{v: D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A)\leq E} |I_{q,q_2,p,v}(Y\wedge U|K)-\right.\\
&\left. -I_{q,q_2,p,v}(S\wedge U,\hat{S}|K) + D(q\circ q_2\circ p\circ v\|Q^*\circ q_2\circ p\circ A) - E - \delta|^+\right\},
\end{aligned} \tag{26}$$

(25) will be satisfied.

To prove that (25) holds for some code it suffices to show that for random code

$$\mathbf{E}B_m \leq \frac{1}{2}, \quad m=\overline{1,|\mathcal{M}|}. \tag{27}$$

To this end we observe that

$$\begin{aligned}
&\mathbf{E}\left|\mathcal{T}^N_{q,q_2,p,v}(Y|U^N,X^N,S^N,\hat{S}^N,\mathbf{k}) \bigcap\right.\\
&\left.\bigcap \bigcup_{(m',\hat{\mathbf{s}}'):\left\{\substack{m'\neq m \text{ or}\\ m'=m,\ d_0(\mathbf{s},\hat{\mathbf{s}}')>\Delta_0}\right\}} \bigcup_{\mathbf{s}'\in\mathcal{T}^N_{q',q_2'}(S|\hat{\mathbf{s}}',\mathbf{k})} \mathcal{T}^N_{q',q_2',p',v'}(Y|U^N,X^N,\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\right| \leq\\
&\leq \sum_{\mathbf{y}\in\mathcal{Y}^N} \Pr\left\{\mathbf{y}\in\mathcal{T}^N_{q,q_2,p,v}(Y|U^N,X^N,S^N,\hat{S}^N,\mathbf{k}) \bigcap\right.\\
&\left.\bigcap \bigcup_{m'} \bigcup_{\hat{\mathbf{s}}'\in\mathcal{T}^N_{q',q_2'}(\hat{S}|\mathbf{k})} \bigcup_{\mathbf{s}'\in\mathcal{T}^N_{q',q_2'}(S|\hat{\mathbf{s}}',\mathbf{k})} \mathcal{T}^N_{q',q_2',p',v'}(Y|U^N,X^N,\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\right\} \leq\\
&\leq \sum_{m'}\sum_{\mathbf{y}\in\mathcal{Y}^N} \Pr\{\mathbf{y}\in\mathcal{T}^N_{q,q_2,p,v}(Y|U^N,X^N,S^N,\hat{S}^N,\mathbf{k})\}\times\\
&\times \Pr\left\{\mathbf{y}\in \bigcup_{(\hat{\mathbf{s}},\mathbf{s}')\in\mathcal{T}^N_{q',q_2'}(\hat{S},S|\mathbf{k})} \mathcal{T}^N_{q',q_2',p',v'}(Y|U^N,X^N,\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\right\},
\end{aligned}$$

because the events in the brackets are independent. Let us note that the first probability is different from zero if and only if $\mathbf{y}\in\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{k})$, in this case for sufficiently large N we have

$$\Pr\{\mathbf{y}\in\mathcal{T}^N_{q,q_2,p,v}(Y|U^N,X^N,S^N,\hat{S}^N,\mathbf{k})\} = \frac{|\mathcal{T}^N_{q,q_2,p,v}(U,X,S,\hat{S}|\mathbf{y},\mathbf{k})|}{|\mathcal{T}^N_{q,q_2,p}(U,X,S,\hat{S}|\mathbf{k})|} \leq$$

$$\leq (N+1)^{|\mathcal{U}||\mathcal{X}||\mathcal{S}|^2|\mathcal{K}|}\exp\{-NI_{q,q_2,p,v}(Y\wedge U,X,S,\hat{S}|K)\}.$$

Taking into account the construction arguments of the code, the second probability can be estimated in the following way:

$$\Pr\left\{\mathbf{y}\in\bigcup_{(\hat{\mathbf{s}},\mathbf{s}')\in\mathcal{T}^N_{q',q_2'}(\hat{S},S|\mathbf{k})}\mathcal{T}^N_{q',q_2',p',v'}(Y|U^N_;X^N,\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\right\}\leq$$

$$\leq\Pr\left\{\mathbf{y}\in\bigcup_j\bigcup_{(\hat{\mathbf{s}}',\mathbf{s}')\in\mathcal{T}^N_{q',q_2',p_0'}(S,\hat{S}|U^N_j,\mathbf{k})}\mathcal{T}^N_{q',q_2',p',v'}(Y|U^N_j,\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\right\}\leq$$

$$\leq\sum_{j=1}^{J}\Pr\{\mathbf{y}\in\mathcal{T}^N_{q',q_2',p',v'}(Y|U^N_j,\mathbf{k})\}\leq J\frac{|\mathcal{T}^N_{q',q_2',p',v'}(U|\mathbf{y},\mathbf{k})|}{|\mathcal{T}^N_{q',q_2',p'}(U|\mathbf{k})|}\leq$$

$$\leq (N+1)^{|\mathcal{U}||\mathcal{K}|}\exp\{-N(I_{q',q_2',p',v'}(Y\wedge U|K)-I_{q',q_2',p_0'}(S\wedge U,\hat{S}|K)-\delta/2)\}.$$

At last we obtain:

$$\mathbf{E}\Big|\mathcal{T}^N_{q,q_2,p,v}(Y|U^N_;X^N,S^N,\hat{S}^N,\mathbf{k})\bigcap$$

$$\bigcap\bigcup_{m'}\bigcup_{\hat{\mathbf{s}}'\in\mathcal{T}^N_{q',q_2'}(\hat{S}|\mathbf{k})}\bigcup_{\mathbf{s}'\in\mathcal{T}^N_{q',q_2'}(S|\hat{\mathbf{s}}',\mathbf{k})}\mathcal{T}^N_{q',q_2',p',v'}(Y|U^N_;X^N,\mathbf{s}',\hat{\mathbf{s}}',\mathbf{k})\Big|\leq$$

$$\leq (N+1)^{|\mathcal{U}||\mathcal{K}|(|\mathcal{X}||\mathcal{S}|^2+1)}|\mathcal{M}||\mathcal{T}^N_{q,q_2,p,v}(Y|\mathbf{k})|\times$$

$$\times\exp\{-N(I_{q,q_2,p,v}(Y\wedge U,X,S,\hat{S}|K)+I_{q',q_2',p',v'}(Y\wedge U|K)-I_{q',q_2',p_0'}(S\wedge U,\hat{S}|K)-\delta/2)\}.$$

From (26) it follows that for any $(q',q_2',p',v')\in\mathcal{D}'(Q^*,E)$

$$|\mathcal{M}|\leq\exp\Big\{N\Big(I_{q',q_2',p',v'}(Y\wedge U|K)-I_{q',q_2',p_0'}(S\wedge U,\hat{S}|K)+$$

$$+\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)}D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)-E-\delta\Big)\Big\}$$

and we obtain

$$\mathbf{E}B_m\leq(N+1)^{|\mathcal{U}||\mathcal{K}|(|\mathcal{S}|^2|\mathcal{X}||\mathcal{Y}|+|\mathcal{S}|^2|\mathcal{X}|+1)}\times$$

$$\times\sum_v\sum_{q',q_2',p',v'\in\mathcal{D}'(Q^*,E)}\exp\Big\{N\Big(I_{q',q_2',p',v'}(Y\wedge U|K)-I_{q',q_2',p_0'}(S\wedge U,\hat{S}|K)+$$

$$+\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)}D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)-E-\delta\Big)\Big\}\times$$

$$\times\exp\{NH_{q,q_2,p,v}(Y|K)\}\times$$

$$\times\exp\{-N(I_{q,q_2,p,v}(Y\wedge U,X,S,\hat{S}|K)+I_{q',q_2',p',v'}(Y\wedge U|K)-I_{q',q_2',p_0'}(S\wedge U,\hat{S}|K)-\delta/2)\}\times$$

$$\times\exp\Big\{N\Big(E-\min_{A\in\mathcal{A}(q',q_2',p',\Delta_2)}D(q'\circ q_2'\circ p'\circ v'\|Q^*\circ q_2'\circ p'\circ A)-$$

$$-H_{q,q_2,p,v}(Y|U,X,S,\hat{S},K)\Big)\Big\}\leq$$

$$\leq(N+1)^{|\mathcal{U}||\mathcal{K}|(|\mathcal{S}|^2|\mathcal{X}||\mathcal{Y}|+|\mathcal{S}|^2|\mathcal{X}|+1)}\sum_v\sum_{q',q_2',p',v'\in\mathcal{D}'(Q^*,E)}\exp\{-N\delta/2\}\leq$$

$$\leq(N+1)^{|\mathcal{U}||\mathcal{K}|(|\mathcal{S}|^2|\mathcal{X}||\mathcal{Y}|+|\mathcal{S}|^2|\mathcal{X}|+1)+|\mathcal{S}||\mathcal{K}|(|\mathcal{U}||\mathcal{X}||\mathcal{S}|+|\mathcal{S}|+1)+2|\mathcal{X}||\mathcal{Y}|}\exp\{-N\delta/2\},$$

which for large enough N proves (27) and hence lemma.

References

[1] A. V. Kusnetsov and B. S. Tsybakov, "Coding in a memory with defective cells", (in Russian), *Probl. Peredachi Informacii*, vol. 10, num. 2, p. 52-60, 1974.

[2] S. I. Gel'fand and M.S. Pinsker, "Coding for channel with random parameters", *Probl. Control and Inf. Theory*, vol. 9, num. 1, p. 19-31, 1980.

[3] M. Costa, "Writing on dirty paper", *IEEE Trans. on Information Theory*, vol. 29, num. 3, p. 439-441, 1983.

[4] H. C. Papadopoulos and C.-E. W. Sundberg, "Simultaneous broadcasting of analog FM and digital audio signals by means of adaptive precanceling techniques", *IEEE Trans. on Communicationsy*, vol. 46, num. 9, p. 1233-1242, 1998.

[5] F. M. J. Willems and T. Kalker, "Methods for reversible embedding," *Proc. 40th Annual Allerton Conference on Communication, Control, and Computing*, Allerton House, Monticello, Illinois, Oct. 2-4, 2002.

[6] T. Kalker and F. M. Willems, "Coding Theorems for Reversible Embedding," *DIMACS Workshop on Network Information Theory*, Rutgers University, Piscataway, NJ, vol. 66, March 2003.

[7] E. Martinian, G. W. Wornell and B. Chen, "Authentication with Distortion Criteria", *IEEE Trans. on Information Theory*, vol. 51, num. 7, p. 2523-2542, 2005.

[8] S. Voloshynovskiy, O. Koval, E. Topak, J. Vila and T. Pun, "Partially reversible data hiding with pure message communications", *IEEE Trans. on Information Forensics and Security*, submitted for publications.

[9] A. Sutivong, M. Chiang, T.M. Cover and Y.-H. Kim, "Channel Capacity and State Estimation for State-Dependent Gaussian Channels", *IEEE Trans. on Information Theory*, vol. 51, num. 4, p. 1486-1495, 2005.

[10] E. A. Haroutunian, "Upper estimate of transmission rate for memoryless channel with countable number of output signals under given error probability exponent" (in Russian), *III All-Union Conf. on Theory of Information Transmission and Coding, Uzhgorod, Publication House of Uzbek Academy of Sciences, Tashkent*, pp. 83-86, 1967.

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

[12] 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 Scence 23*, pp. 20-31, 2004.

[13] M. E. Haroutunian, "New bounds for E-capacities of arbitrarily varying channel and channel with random parameter", *Trans. IIAP NAS RA, Mathematical Problems of Computer sciences*, vol. 22, p. 44-59, 2001. Available at http://ipia.sci.am/itas.

[14] M. E. Haroutunian, "Estimates of E-capacity and capacity regions for multiple-access channel with random parameter", *Electronic Notes in Discrete Mathematics*, v.21, General Theory of Information Transfer and Combinatorics, pp. 303-308, 2005. Available at http://www.sciencedirect.com/science.

[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 Trans. Inform. Theory*, vol. 44, no. 6, pp. 2505–2523, 1998.

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

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

[19] E. A. Haroutunian and H. Belbashir, "Lower bound 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.

Տեղեկություններ թաքցնող շրջելի համակարգի E-ունակության պատահական կոդավորման գնահատականը

Մ. Հարությունյան, Ս. Տոնոյան, Օ. Կովալ, Ս. Վոլոշինովսկի

Ամփոփում

Աշխատանքում հետազոտված է տեղեկությունների շրջելի թաքցման ինֆորմացիոն-տեսական խնդիրը: Դիտարկված համակարգի համար հետազոտվել է E-ունակություն ֆունկցիան: Այն իրենից ներկայացնում է տեղեկությունների թաքցման արագության կախվածությունը հուսալիությունից, տեղեկություններ թաքցնողի ու հարձակվողի համար թույլատրելի շեղման մակարդակներից և նախնական տվյալների վերականգնման համար թույլատրելի շեղման մակարդակից: E-ունակության համար կառուցվել է պատահական կոդավորման գնահատական: Համապատասխան գնահատական է ստացվել համակարգի ունակության համար:

Ուսումնասիրվել են նաև երկու մասնավոր դեպքեր՝ մաքուր շրջելիության և հաղորդագրությունների մաքուր հաղորդման դեպքերը: