Random Coding Bound of Rate-Reliability-Distortion Function for Generalized Channel With Side Information

Mariam, Haroutunian
Institute for Informatics and
Automation Problems of the NAS, RA
e-mail: armar@ipia.sci.am

Arthur, Muradyan
Institute for Informatics and
Automation Problems of the NAS, RA
e-mail: arthur@proximusda.com

ABSTRACT

In this paper we study generalized model of discrete memoryless channel (DMC) with finite input and output alphabets and random state sequence (side information) partially known to the encoder, channel and decoder. The study includes family of Gel'fand-Pinsker and information hiding coding problems as special cases. Information is to be reliably transmitted through the noisy channel selected by adversary. Reasoning from applications the actions of encoder and adversary are limited by distortion constraints. The encoder and decoder depend on a random variable (RV) which can be treated as cryptographic key. Two cases are considered, when the joint distribution of this RV and side information is given or this RV is independent from side information and it's distribution can be chosen for the best code generation. We investigate the rate-reliability-distortion function for the mentioned model and derive the lower bound for it.

1. INTRODUCTION

The DMC with random state information available to the encoder was studied by Gel'fand and Pinsker [1], they derived the capacity of this channel. The capacity of arbitrary varying channel with side information at the encoder was derived by Ahlswede [2]. Error exponents of single-user, multi-user and varying channels with side information were studied in [3, 4, 5, 6, 7].

It was discovered that embedding and hiding [8] is closely related to the channel with random parameter, where the cover signal plays the role of the state information. The difference between the two problems is that in various formulations of data-hiding and watermarking there are distortion constraints for the transmitter and a memoryless adversary and the channel is not fixed as it is chosen by adversary. Motivated by data-hiding applications several models are studied, where partial or no information of the state sequence is available to the encoder, channel designer and decoder. Results on capacity and error exponents problems have been obtained in [8, 9, 10, 11, 12, 13].

A unified framework for studying such problems was first suggested by Cover and Chiang [14], who considered the channel with two-sided state information, where the sender and the receiver have correlated but different state information. This model includes four possible situations of the channel with random parameter as special cases. They obtained the capacity of this channel and explored the duality with source coding problems. The random coding bound of E-capacity for this model was derived in [15].

Later Moulin and Wang [16] studied the generalized model with side information, where the degraded versions of side information are distributed among encoder, adversary and decoder. This model includes also the various cases of information hiding. They derived the capacity formulas and random coding exponents for compound discrete memoryless channels and channels with arbitrary memory.

In this paper we study a similar generalized model of discrete memoryless channel (DMC) with finite input and output alphabets and random state sequence (side information) partially known to the encoder, channel and decoder (fig. 1). Information is to be reliably transmitted through the noisy channel selected by adversary.

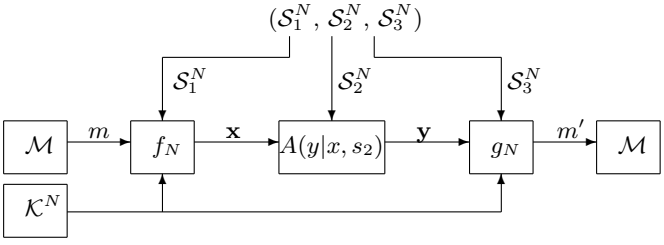

Figure l. Generalized model of a channel with side information

Distortion constraints are imposed on the encoder called *transparency* requirement and on the attacker called *robustness* requirement.

The encoder and decoder depend on a random variable (RV) which can be treated as cryptographic key. Two cases are considered, when the joint distribution of this RV and side information is given or this RV is independent from side information and it's distribution can be chosen for the best code generation.

We investigate the rate-reliability-distortion function for the mentioned model and derive the lower bound for it. This function expresses the dependence of the information hiding rate on reliability and distortion levels for information hider and attacker. This investigation is equivalent to studying of error exponents but sometimes is more expedient. This approach was first introduced by E. Haroutunian [17, 18, 19] and developed for various channels [4, 5, 7, 9, 10, 13, 15]. In this paper we derive the lower bound (random coding bound) of rate-reliability-distortion function .

The paper is organized as follows. Definitions of terms and notations used throughout the paper are described in section 2. The formulation of the main result and its special cases are stated in the section 3.

2. NOTATIONS AND DEFINITIONS

Capital letters are used for RV $K, S_1, S_2, S_3, U, X, Y$ taking values in the finite sets $\mathcal{K}, \mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{U}, \mathcal{X}, \mathcal{Y}$, correspondingly, and lower case letters $k, s_1, s_2, s_3, u, x, y$ 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 generalized model of a channel with side information is depicted in Figure 1. A message m to be transmitted through an attack channel to the receiver is uniformly distributed over the message set $\mathcal{M}$. The joint state sequence is described by random variable $S = (S_1, S_2, S_3)$ the components of which represent the partial information known to the encoder, adversary and decoder, correspondingly. Random variable K represents separate information known only to the encoder and decoder.

Two cases are considered, when the joint PD

$$Q^* = Q_0^* \circ Q_1^* \circ Q_2^* \circ Q_3^* = \{Q^*(k, s_1, s_2, s_3) =$$

$$= Q_0^*(k) Q_1^*(s_1|k) Q_2^*(s_2|k, s_1) Q_3^*(s_3|k, s_1, s_2),$$

$$k \in \mathcal{K}, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, s_3 \in \mathcal{S}_3\}$$

is given or K is independent from side information and it's distribution can be chosen for the best code generation.

It is assumed that:

$$Q^{*N}(\mathbf{k}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3) = \prod_{n=1}^{N} Q^*(k_n, s_{1n}, s_{2n}, s_{3n}).$$

The transmitter encodes the message m using $\mathbf{s}_1$ and $\mathbf{k}$. The resulting codeword $\mathbf{x} \in \mathcal{X}^N$ is transmitted via attack channel $A(y|x, s_2)$. The attacker produces corrupted blocks $\mathbf{y} \in \mathcal{Y}^N$. The decoder does not know $A(y|x, s_2)$ selected by adversary and possessing $\mathbf{s}_3$ derives the message m'.

Following probability distributions are used in the paper:

$$Q = Q_0 \circ Q_1 \circ Q_2 \circ Q_3 = \{Q(k, s_1, s_2, s_3) =$$

$$= Q_0(k) Q_1(s_1|k) Q_2(s_2|k, s_1) Q_3(s_3|k, s_1, s_2),$$

$$k \in \mathcal{K}, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, s_3 \in \mathcal{S}_3\},$$

$$P = P_0 \circ P_1 = \{P(x, u|k, s_1) = P_0(u|k, s_1) P_1(x|u, k, s_1),$$

$$x \in \mathcal{X}, u \in \mathcal{U}, k \in \mathcal{K}, s_1 \in \mathcal{S}_1\},$$

$$V = \{V(y|k, s_1, s_2, s_3, u, x),$$

$$y \in \mathcal{Y}, k \in \mathcal{K}, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, s_3 \in \mathcal{S}_3, u \in \mathcal{U}, x \in \mathcal{X}\},$$

$$Q_3^* \circ A = \{Q_3^* \circ A(y, s_3|x, k, s_1, s_2) = Q_3^*(s_3|k, s_1, s_2) A(y|x, s_2),$$

$$y \in \mathcal{Y}, k \in \mathcal{K}, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, s_3 \in \mathcal{S}_3, x \in \mathcal{X}\},$$

$$QP(x, s_2) = \sum_{k, s_1, u} Q_0(k) Q_1(s_1|k) Q_2(s_2|k, s_1) P(x, u|k, s_1).$$

For brevity we will write indices of Q separated by comma, when mentioning product of respective probability distributions (or types). E.g. $Q_{0,1,2} = Q_0 \circ Q_1 \circ Q_2$.

For the information-theoretic quantities, such as entropy $H_{Q_0,Q_1,Q_2}(K, S_1, S_2)$, mutual information $I_{Q_0,Q_1,P_0}(U \wedge S_1)$, divergence $D(Q_0||Q_0^*)$ and for the notion of type we refer to [19, 20, 21, 22].

The following properties [20, 21] are used in proofs:

for $\mathbf{k} \in \mathcal{T}_{Q_0}(K)$, $\mathbf{s}_1 \in \mathcal{T}_{Q_0,Q_1}(S_1|\mathbf{k})$, $\mathbf{s}_2 \in \mathcal{T}_{Q_0,Q_1,Q_2}(S_2|\mathbf{k}, \mathbf{s}_1)$, $\mathbf{x} \in \mathcal{T}_{Q_0,Q_1,P}(X|\mathbf{k}, \mathbf{s}_1)$, $(\mathbf{y}, \mathbf{s}_3) \in \mathcal{T}_{Q,P,V}(Y, S_3|\mathbf{x}, \mathbf{k}, \mathbf{s}_1, \mathbf{s}_2)$,

$$Q_3^{*N} \circ A^N(\mathbf{y}, \mathbf{s}_3|\mathbf{k}, \mathbf{x}, \mathbf{s}_1, \mathbf{s}_2) =$$

$$= \exp\{-N(H_{Q,P,V}(Y, S_3|X, K, S_1, S_2) +$$

$$+ D(Q_3 \circ V||Q_3^* \circ A|Q_0, Q_1, Q_2, P)\}, \quad (1)$$

$$D(Q \circ P \circ V||Q^* \circ P \circ A) = D(Q_{0,1,2}||Q_{0,1,2}^*) +$$

$$+ D(Q_3 \circ V||Q_3^* \circ A|Q_0, Q_1, Q_2, P), \quad (2)$$

$$D(Q||Q^*) = D(Q_0||Q_0^*) + D(Q_1||Q_1^*|Q_0) + D(Q_2||Q_2^*|Q_0, Q_1) +$$

$$+ D(Q_3||Q_3^*|Q_0, Q_1, Q_2), \quad (3)$$

$$H_{Q,P,V}(Y, S_3|U, X, K, S_1, S_2) \le H_{Q,P,V}(Y, S_3|X, K, S_1, S_2). \quad (4)$$

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

The mappings $d_1 : \mathcal{S}_1 \times \mathcal{X} \to \mathbb{R}^+$ and $d_2 : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+$ are distortion functions over the encoder and attacker correspondingly. They are supposed to be symmetric ($d_1(s_1, x) = d_1(x, s_1)$ and $d_2(x, y) = d_2(y, x), s_1 \in \mathcal{S}_1, x \in \mathcal{X}, y \in \mathcal{Y}$) and become 0 if $s_1 = x$ and $x = y$. Distortion functions for N-length vectors are defined as:

$$d_1^N(\mathbf{s}_1, \mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} d_1(s_{1n}, x_n), d_2^N(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{n=1}^{N} d_2(x_n, y_n).$$

Let $\Delta_1 \ge 0$ be the number indicating the allowed distortion level for the encoder and $\Delta_2 \ge 0$ for the attacker.

The N-length *code* is a pair of mappings (f_N, g_N), where

$$f_N : \mathcal{M} \times \mathcal{K}^N \times \mathcal{S}_1^N \to \mathcal{X}^N,$$

is the encoding function which satisfies the following distortion constraint:

$$d_1^N(\mathbf{s}_1, f_N(m, \mathbf{k}, \mathbf{s}_1)) \le \Delta_1, \quad (5)$$

for all $m, \mathbf{k}, \mathbf{s}_1$ and

$$g_N : \mathcal{Y}^N \times \mathcal{K}^N \times \mathcal{S}_3^N \to \mathcal{M}.$$

is the decoding function.

Note that definition of the distortion constraint (5) means that maximum distortion constraint is used, which is stronger condition than average distortion constraint over $m \in \mathcal{M}$, $\mathbf{k} \in \mathcal{K}^N$ and $\mathbf{s}_1 \in \mathcal{S}_1^N$ i.e. if we find f_N satisfying (5) it will also satisfy average distortion constraint.

N is called *code length* and $|\mathcal{M}|$ is called *code volume*. The nonnegative number $R = \frac{1}{N} \log |\mathcal{M}|$ is called code rate.

The selected channel is memoryless, it means that for $\mathbf{x} \in \mathcal{X}^N$, $\mathbf{y} \in \mathcal{Y}^N$ and $\mathbf{s}_2 \in \mathcal{S}_2^N$:

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

and satisfies the following distortion constraint for QP^N:

$$\sum_{\mathbf{s}_2, \mathbf{x}, \mathbf{y}} QP^N(\mathbf{x}, \mathbf{s}_2) A^N(\mathbf{y}|\mathbf{x}, \mathbf{s}_2) d_2(\mathbf{x}, \mathbf{y}) \le \Delta_2. \quad (6)$$

A memoryless covert channel P, subject to distortion Δ_1, is probability distribution P such that for any $Q_{0,1}$:

$$\sum_{k,s_1,u,x} Q_{0,1}(s_1,k)P(x,u|k,s_1)d_1(s_1,x) \le \Delta_1. \qquad (7)$$

The set of probability distributions P satisfying condition (7) is denoted by $\mathcal{P}(Q_{0,1},\Delta_1)$.

A memoryless attack channel A, subject to distortion Δ_2 is defined by probability distribution A such that for any $Q_{0,1,2}$ and P:

$$\sum_{k,s_1,s_2,u,x,y} Q_{0,1,2}(k,s_1,s_2)P(x,u|k,s_1)A(y|x,s_2)d_2(x,y) \le \Delta_2. \qquad (8)$$

The set of channels A satisfying condition (8) is denoted by $\mathcal{A}(Q_{0,1,2},P,\Delta_2)$.

We will consider cases when the distribution of $\mathbf{k}$ is either given or it is independent of state sequences and it is not given but rather selected in a way to achieve minimal error probability.

In the first case the probability of erroneous reconstruction of message m for $P\in\mathcal{P}(Q^*_{0,1},\Delta_1)$, $A\in\mathcal{A}(Q^*_{0,1,2},P,\Delta_2)$ is calculated in the following way:

$$e^1(m,A)=e(f_N,g_N,A,Q^*,\Delta_1,\Delta_2,m)=$$

$$=\sum_{\mathbf{k},\mathbf{s}_1,\mathbf{s}_2} Q^{*N}_{0,1,2}(\mathbf{k},\mathbf{s}_1,\mathbf{s}_2)\times$$

$$\times Q^{*N}_3\circ A^N(\mathcal{Y}^N\times\mathcal{S}^N_3\backslash g^{-1}_{N,\mathbf{k}}(m)|f_N(m,\mathbf{k},\mathbf{s}_1),\mathbf{k},\mathbf{s}_1,\mathbf{s}_2),$$

where $g^{-1}_{N,\mathbf{k}}(m)=\{\mathbf{y},\mathbf{s}_3: g_N(\mathbf{y},\mathbf{k},\mathbf{s}_3)=m\}$.

In the second case the erroneous reconstruction probability can be calculated in the following way:

$$e^2(m,A)=e(f_N,g_N,A,Q^*,\Delta_1,\Delta_2,m)=$$

$$=\min_{Q_0}\sum_{\mathbf{k},\mathbf{s}_1,\mathbf{s}_2} Q^N_0(\mathbf{k})Q^{*N}_{1,2}(\mathbf{s}_1,\mathbf{s}_2)\times$$

$$\times Q^{*N}_3\circ A^N(\mathcal{Y}^N\times\mathcal{S}^N_3\backslash g^{-1}_{N,\mathbf{k}}(m)|f_N(m,\mathbf{k},\mathbf{s}_1),\mathbf{s}_1,\mathbf{s}_2).$$

The maximal value of error probability of the code over all A for given message m is denoted by:

$$e^i(m)=e^i(f_N,g_N,Q^*,\Delta_1,\Delta_2,m)=\max_A e^i(m,A),\ \mathrm{i}=1,2.$$

The maximal error probability of the code over all $m\in\mathcal{M}$ is equal to:

$$e^i=e^i(f_N,g_N,Q^*,\Delta_1,\Delta_2)=\max_{m\in\mathcal{M}} e^i(m),\ \mathrm{i}=1,2$$

and the average error probability of the code over all $m\in\mathcal{M}$ is:

$$\overline{e^i}=\overline{e^i}(f_N,g_N,Q^*,\Delta_1,\Delta_2)=\frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}} e^i(m),\ \mathrm{i}=1,2.$$

3. FORMULATION OF RESULTS

The E-capacity for maximal error probability is denoted by $C^i(E,Q^*,\Delta_1,\Delta_2)$ and is defined in the following way:

$$C^i(E,Q^*,\Delta_1,\Delta_2)=\overline{\lim_{N\to\infty}}\frac{1}{N}\log M^i(E,Q^*,\Delta_1,\Delta_2,N),$$

where

$$M^i(E,Q^*,\Delta_1,\Delta_2,N)=\sup_{f_N,g_N}\Big\{|\mathcal{M}|: e^i\le\exp(-NE)\Big\},\ \mathrm{i}=1,2.$$

The E-capacity for the average error probability is denoted by $\overline{C^i}(E,Q^*,\Delta_1,\Delta_2)$.

We can observe that E-capacity is the generalization of the capacity because it converges to channel capacity when $E\to 0$. To introduce the main theorem denote:

$$R(E,Q^*,A,Q,P,V)=I_{Q,P,V}(U\wedge S_3,Y|K)-I_{Q_0,Q_1,P_0}(U\wedge S_1|K)+$$

$$+D(Q\circ P\circ V||Q^*\circ P\circ A)-E,$$

$$R^1_r(E,Q^*,\Delta_1,\Delta_2)=\min_{Q_0,Q_1,Q_2}\max_{P\in\mathcal{P}(Q_{0,1},\Delta_1)}\min_{A\in\mathcal{A}(Q_{0,1,2},P,\Delta_2)}$$

$$\min_{Q_3,V:D(Q\circ P\circ V||Q^*\circ P\circ A)\le E}\Big|R(E,Q^*,A,Q,P,V)\Big|^+ \qquad (9)$$

and

$$R^2_r(E,Q^*,\Delta_1,\Delta_2)=\max_{Q_0}\min_{Q_1,Q_2}\max_{P\in\mathcal{P}(Q_{0,1},\Delta_1)}\min_{A\in\mathcal{A}(Q_{0,1,2},P,\Delta_2)}$$

$$\min_{Q_3,V:D(Q_{1,2,3}\circ P\circ V||Q^*_{1,2,3}\circ P\circ A|Q_0)\le E}\Big|R(E,Q^*,A,Q,P,V)\Big|^+. \qquad (10)$$

Theorem. *For generalized channel with distortion constraints imposed on the encoder and channel, for given Q^* and for all $E>0$, $i=1,2$*

$$R^i_r(E,Q^*,\Delta_1,\Delta_2)\le C^i(E,Q^*,\Delta_1,\Delta_2)\le\overline{C^i}(E,Q^*,\Delta_1,\Delta_2).$$

Corollary 1. *When $E\to 0$, $i=2$ we derive the capacity for both compound discrete memoryless channel and channel with arbitrary memory established in [16].*

Corollary 2. *When $S_2=(S_1,S_3), K=\emptyset$ we derive the E-capacity obtained in [15], which in its turn is generalization of the channels for four possible situations with random parameter.*

Corollary 3. *When $S_2=\emptyset, S_3=\emptyset, i=1$ we get the lower bound of E-capacity obtained in [9].*

REFERENCES

[1] 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.

[2] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender", *IEEE Transactions on Information Theory*, vol. IT-32, no. 5, pp. 621-629, 1986.

[3] E. A. Haroutunian, M. E. Haroutunian, "Channel with random parameter", *Proc. of XXII Prague conf. on Inform. Theory, Statistical Decision Functions, Random Processes*, pp. 99-101, 1994.

[4] 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.

[5] M. E. Haroutunian, "On multiple-access channel with random parameter", *Proc. of Int. conf. on Computer Science and Inform. Technologies*, Armenia, Yerevan, pp. 174 - 178, 2003.

[6] A. Somekh-Baruch and N. Merhav, "On the random coding error exponenets of the single-user and the multiple-access Gel'fand-Pinsker channels", *Proc. of IEEE International Symposium on Information Theory*, USA, Chicago, p. 448, 2004.

[7] 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.

[8] 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.

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

[10] M. E. Haroutunian, 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.

[11] A. Somekh-Baruch, N. Merhav, "On the error exponent and capacity games of private watermarking systems", *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 537-562, 2003.

[12] A. Somekh-Baruch, N. Merhav, "On the capacity game of public watermarking systems", *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 511-524, 2004.

[13] M. Haroutunian, S. Tonoyan, O. Koval, S. Voloshynovskiy, "On reversible information hiding system", *Proc. of IEEE International Symposium on Information Theory*, Canada, Toronto, pp. 940-944, 2008.

[14] 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.

[15] M. Haroutunian, A. Muradyan, "Lower bound for E capacity of discrete memoryless channel with two-sided state information", *Transactions of the Institute for Informatics and Automation Problems of the NAS of RA, Mathematical Problems of Computer Sciences*, vol. 31, pp. 28-39, 2008.

[16] 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, 2007.

[17] E. A. Haroutunian, "Upper estimate of transmission rate for memoryless channel with countable number of output signals under given error probability exponent", (in Russian), *3rd All-Union Conf. on Theory of Information Transmission and Coding, Uzhgorod, Publication house of Uzbek Academy of Sciences, Tashkent*, pp. 83-86, 1967.

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

[19] 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.

[20] T. M. Cover and J. A. Thomas, "Elements of information theory", Wiley, New York, 1991.

[21] I. Csiszár and J. Körner, "Information Theory: Coding theorems for discrete memoryless systems", Academic Press, New York, 1981.

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