Identification under Random Processes *

Rudolf Ahlswede and Vladimir B. Balakirsky

*The work was supported in part by SFB-343, Universität Bielefeld, Germany

1 Introduction

Identification via channels was introduced in [1,2] as a problem of a reliable transmission of one of M messages to M receivers in such a way that each receiver decides whether it is his message, which was sent, or not. If a receiver misses his message then the decoding error of the first kind takes place, and if a receiver accepts a message which was sent for a different receiver, then the decoding error of the second kind takes place. An identification scheme should be constructed in such a way that the decoding error probabilities of the both kinds are small for all receivers.

A key idea of constructing good identification codes is an assignement of the probability distributions (PDs) to the messages instead of fixed codewords. Suppose we a given a discrete memoryless channel W having the input alphabet $\mathcal{X}$ and the output alphabet $\mathcal{Y}$, i.e., the probability to receive a vector $\mathbf{y} = (y^{(1)}, ..., y^{(n)}) \in \mathcal{Y}^n$, when a vector $\mathbf{x} = (x^{(1)}, ..., x^{(n)}) \in \mathcal{X}^n$ was transmitted, is defined as

$$W(\mathbf{y}|\mathbf{x}) = \prod_{t=1}^{n} W(y^{(t)}|x^{(t)}).$$

Definition 1 An $(n, M, \Lambda_1, \Lambda_2)$ ID code is a collection

$$\{ \ (Q_i(\cdot), \mathcal{D}_i), \ i = 1, ..., M \ \}$$

such that, for all $i, j = 1, ..., M$,

1) $Q_i(\cdot) = \{ \ Q_i(\mathbf{x}), \mathbf{x} \in \mathcal{X}^n \ \}$ is a PD;

2) $\mathcal{D}_j \subset Y^n$;

3) $\sum_{\mathbf{x}} Q_j(\mathbf{x}) \cdot W(\mathcal{D}_j|\mathbf{x}) \geq 1 - \Lambda_1$;

4) $\sum_{\mathbf{x}} Q_i(\mathbf{x}) \cdot W(\mathcal{D}_j|\mathbf{x}) \leq \Lambda_2$ if $i \neq j$.

The input distributions of an ID code can be realized if the sender has M collections of not necessarily distinct codewords. To send the i-th message, he selects one of the codewords of the i-th collection in accordance with the uniform distribution. Each receiver knows his collection and checks if the received vectors is likely to be generated by one of the codewords of this collection or not. The known results show that the number of messages that can be reliably transmitted using an ID code turns out to be doubly exponential in the blocklength n [1].

We will consider the system of information transmission given in Fig.1, where the sender and the receivers have an access to a public random process and get correlated binary sequences $\mathbf{x}^*$ and $\mathbf{y}^*$ of length k before the sender generates a codeword, depending of the message i, and transmits it over a memoryless channel W. The probability to receive a pair $(\mathbf{x}^*, \mathbf{y}^*)$ is defined as

$$P(\mathbf{x}^*, \mathbf{y}^*) \quad = \quad 2^{-k} p^{d_H(\mathbf{x}^*,\mathbf{y}^*)} (1-p)^{k-d_H(\mathbf{x}^*,\mathbf{y}^*)}, \tag{1.1}$$

where d_H denotes the Hamming distance and $p \leq 1/2$. Our intension is to evaluate the decoding error probabilities of the first and of the second kind as functions of the number of possible messages,

$$\rho = k/n, \tag{1.2}$$

the channel capacity C, and the probability p.

A randomization of the transmission can be represented in such a way that the sender observes a vector $\mathbf{z}$ of length l, whose components are independent identically distributed (i.i.d.) binary random variables taking values 0 and 1 with the probability 1/2, and the use of this vector as a pointer to a collection corresponding to the message to be sent. Such a procedure can be also realized if the sender uses the vector $\mathbf{x}^* \in \{0,1\}^k$ for this purpose. The difference of these two possibilities is concluded in the note that in the second case the receivers have a side information on the pointer, which is $\mathbf{y}^* \in \{0,1\}^k$, and that the length k is assumed to be given, while the length l can be chosen in an arbitrarily way.

Let us denote by R_ρ the log log of the number of messages [1] to be identified with an arbitrarily small decoding error probabilities divided by n (more explicitely, R_ρ is defined in Section 2) and consider two cases when $p = 0$ and $p = 1/2$. If $p = 1/2$, then we have a classical identification model where the channel W is the only link, connecting the sender and the receivers. Therefore, we refer to the known results [1,3,4] and claim that $R_\rho = C$, that the length l should be chosen approximetely equaled nC, and that the observations of a random process are useless (Fig.2).

A common randomness is valuable for constructing effective communication systems [5]. In particular, asymptotically optimal identification scheme with a noiseless feedback consists of two parts [2]. In the first part, the sender and the receivers distribute the result of a random experiment, which is a realization of a noise sequence during transmission of a given or randomly chosen sequence over the channel W. This result becomes common because of feedback. The distribution of a common randomness occupies almost all transmission time and completely determines the asymptotical characteristics of the identification system. Our model is close to the identification with the feedback [2,6]. The main difference is that the random experiment is public and it occures somewhere outside the system. Therefore, we cannot control it, but we also do not include its duration into the transmission time. However, using the results of [2] we can claim that $R_\rho = C + \rho$ for all $\rho \geq 0$ when $p = 0$ (Fig.2).

The cases considered above show that there is a limit on the length $k = \rho n$ when $p = 1/2$ and that this limit is absent when $p = 0$. In Section 2 we examine the general case, $p \in [0, 1/2]$, and come to the conclusion that, from a point of view of the direct coding theorems, this limit does not exist for all $p < 1/2$, but the additional quantity in the maximal identification rate, which we gain while observing the process, is limited at a threshold value for all $p > 0$ and give an expression for that value.

[1] All the logarithms in the paper are based modulo 2. Furthermore, we denote $\exp_2 z = 2^z$ for all z.

During the analysis, we distinguish between two cases when $\rho H(p) < C$ and when $\rho H(p) \geq C$, where

$$H(p) = -p \log p - (1-p) \log(1-p)$$

is the entropy of the noise sequence which makes the observations of the process different at the sender's and at the receiver's sides. In the first case the encoder can improve the behaviour using an additional randomization, and in the second case there is more than enough randomness in these observations.

There was an attempt to represent identification via channels as a model for a human communication (”in a stormy night one sailor drowns in the ocean...” [1,p.27]). We can also continue this direction saying that the observation of a process is a kind of a pray, which allows the sender to predict a way of thinking of the receivers before he transmits the messages to these receivers.

The paper is organized as follows. In subsection 2.1 we give a notational background and some auxiliary results needed in analysis. These results concern different ways of partitioning of the space, because we meet the problem of an upper-bounding the expectation of a product of dependent random variables, and the partitions give an opportunity to select the subsets of independent variables. Two identification schemes, with and without additional randomization, are described in subsection 2.2, and corresponding code ensembles are introduced in subsection 2.3. The direct coding theorems for these cases are given in subsections 2.4,2.5. Subsection 2.6 is devoted to the calculation of an additional quantity in the maximal indentification rate, ΔR_∞, which comes with the infinite observations.

2 Direct Coding Theorems for Identification under a Binary Symmetric Random Process

2.1 Basic ideas and auxiliary results

Let us start with the definitions.

Definition 1.1 An $(n, M, \Lambda_1, \Lambda_2)$ *ID code for identification under a random process*

$$\mathcal{P}^k = \{ \ P(\mathbf{x}^*, \mathbf{y}^*), \ \mathbf{x}^*, \mathbf{y}^* \in \{0,1\}^k \ \}$$

is a collection

$$\{ \ (Q_i(\cdot|\mathbf{x}^*), \mathcal{D}_i(\mathbf{y}^*)), \ \mathbf{x}^*, \mathbf{y}^* \in \{0,1\}^k, \ i = 1, ..., M \ \}$$

such that, for all $\mathbf{x}^*, \mathbf{y}^* \in \{0,1\}^k$ and $i, j = 1, ..., M$,

1) $Q_i(\cdot|\mathbf{x}^*) = \{ \ Q_i(\mathbf{x}|\mathbf{x}^*), \mathbf{x} \in \mathcal{X}^n \ \}$ is a PD for all $\mathbf{x}^* \in \{0,1\}^k$;

2) $\mathcal{D}_j(\mathbf{y}^*) \subset Y^n$;

3) $\sum_{\mathbf{x}^*,\mathbf{y}^*} P(\mathbf{x}^*,\mathbf{y}^*) \cdot \sum_{\mathbf{x}} Q_j(\mathbf{x}|\mathbf{x}^*) \cdot W(\mathcal{D}_j(\mathbf{y}^*)|\mathbf{x}) \geq 1 - \Lambda_1$;

4) $\sum_{\mathbf{x}^*,\mathbf{y}^*} P(\mathbf{x}^*,\mathbf{y}^*) \cdot \sum_{\mathbf{x}} Q_i(\mathbf{x}|\mathbf{x}^*) \cdot W(\mathcal{D}_j(\mathbf{y}^*)|\mathbf{x}) \leq \Lambda_2$ if $i \neq j$.

Definition 1.2 The triple (R, e_1, e_2) will be referred to as a *ρ-achievable ID triple under a random process* $\mathcal{P}^k$, $k = \rho n$, if, for all $\delta > 0$ and $n \geq n(\delta, |\mathcal{X}|, |\mathcal{Y}|)$, an $(n, M, \Lambda_1, \Lambda_2)$ ID code for identification under the process $\mathcal{P}^k$ with

$$\begin{aligned} \frac{1}{n}\log\log M &\geq R - \delta, \\ \Lambda_1 &\leq 2^{-n(e_1-\delta)}, \\ \Lambda_2 &\leq 2^{-n(e_2-\delta)} \end{aligned} \tag{2.1}$$

exists.

Definition 1.3 The parameter R will be referred to as a *ρ-achievable ID rate under a random process* $\mathcal{P}^k$, $k = \rho n$, if there exist $e_1, e_2 > 0$ such that (R, e_1, e_2) is a ρ-achievable ID triple under that process. The ρ-achievable ID rate will be denoted by R_ρ.

To realize an identification, we will use an error-correcting code G consisting of

$$m = 2^{nr} \tag{2.2}$$

codewords $\mathbf{x} \in \mathcal{X}^n$. We suppose that the use of this code to transmit data over a memoryless channel W provides the exponent of the maximal decoding error probability at level $2^{-ne(r)}$ and denote the decoding decision sets by $\mathcal{C}(\mathbf{x}), \mathbf{x} \in G$, i.e.,

$$W(\mathcal{C}(\mathbf{x})|\mathbf{x}) \geq 1 - 2^{-ne(r)} \text{ for all } \mathbf{x} \in G. \tag{2.3}$$

We assign the encoding function for the i-th message, f_i, in such a way that it takes values in the code G. Hence, the transmitted codeword, $\mathbf{x}(i)$, always belongs to G.

Each receiver can construct a list of the vectors $\mathbf{x}^*$ in such a way that, with the high probability, the list contains the vector observed by the sender. Therefore, the decoding error probabilities of the first kind are estimated by the sum consisting of the probability that the sender's observation does not belong to the list and the maximal decoding error probability for the code G.

To estimate the decoding error probabilities of the second kind we introduce the code ensembles. These ensembles are different for the cases $\rho H(p) < C$ and $\rho H(p) \geq C$. However, in every case we prove that, for any fixed pair (i,j), $i \neq j$, an upper bound of the following form :

$$Pr\{\, \lambda_{ij} > \Lambda \,\} \leq \exp_2\{\, -2^{ne} \,\} \tag{2.4}$$

is valid, where λ_{ij} is the decoding error probability of the j-th receiver when the message i was sent, Λ and e are constant chosen in a special way, and $Pr\{\ \}$ is the probability in the code ensemble. Using (1.4) we claim that if

$$M < \exp_2\{\, -2^{ne}/2 \,\}, \tag{2.5}$$

then there exists an ID code such that

$$\lambda_{ij} \leq \Lambda \text{ for all } i \neq j. \tag{2.6}$$

Really, if (1.5) holds, then

$$\begin{aligned} & Pr\{\, \lambda_{ij} > \Lambda \text{ for at least one } (i,j),\ i \neq j \,\} \qquad (2.7) \\ \leq\ & \sum_{i=1}^{M} \sum_{j \neq i} Pr\{\, \lambda_{ij} > \Lambda \,\} < M^2 \cdot \exp_2\{\, -2^{ne} \,\} < 1. \end{aligned}$$

Therefore, the statement (1.6) is true with a positive probability and we get the result of an existence type.

To prove (1.5) we construct an upper bound on the expectation of a product of dependent random variables via selecting the maximal independent subsets of these variables and obtaining the product of expectations based on Hölder's inequality. Such a possibility is realized using the known constructions of error-correcting codes.

Definition 1.4 Given $d, a \in \{1, ..., k/2\}$ and $\mathbf{b} \in \{0,1\}^k$, let

$$\begin{aligned} \{0,1\}^k_d &= \{\, \mathbf{c} \in \{0,1\}^k :\ w_H(\mathbf{c}) = d \,\}, \qquad (2.8) \\ \mathbf{S}^k_d(\mathbf{b}) &= \mathbf{b} + \{0,1\}^k_d, \\ \mathbf{S}^k_{d,a}(\mathbf{b}) &= \mathbf{b} + \{0,1\}^k_d + \sum_{\tau=0}^{a} \{0,1\}^k_\tau. \end{aligned}$$

where w_H denotes the Hamming weight. Furthermore, for any given collection of sets $\{\mathcal{B}_\nu\}$ and any index t, we write

$$\varepsilon_t(\{\mathcal{B}_\nu\}) = \max_\nu \left| \mathbf{S}^k_t(\mathbf{b}) \bigcap \left(\bigcup_{\mathbf{b}' \in \mathcal{B}_\nu \setminus \{\mathbf{b}\}} \mathbf{S}^k_t(\mathbf{b}') \right) \right|^{-1} \cdot \left|\, \mathbf{S}^k_t(\mathbf{b}) \,\right|.$$

(1D) A system of ν_τ pairwise disjoint subsets $\mathcal{B}_\nu, \nu = 1, ..., \nu_d$, consisting of μ_d vectors $\mathbf{b} \in \{0,1\}^k$, will be referred to as a $(d, \mu_d \times \nu_d, \varepsilon_d)$-*pairwise disjoint decomposition of the set* $\{0,1\}^k$ if

$$\nu_d = \frac{|\{0,1\}^k|}{\mu_d} = \frac{2^k}{\mu_d} \tag{2.9}$$

and

$$\varepsilon_d(\{\mathcal{B}_\nu\}) = \varepsilon_d.$$

(2D) A system of ν_d pairwise disjoint subsets $\mathcal{B}_\nu, \nu = 1, ..., \nu_{d,a}$, consisting of $\mu_{d,a}$ vectors $\mathbf{b} \in \{0,1\}^k$, will be referred to as a $((d,a), \mu_{d,a} \times \nu_{d,a}, \varepsilon_{d,a})$-*pairwise disjoint decomposition of the set* $\{0,1\}^k$ if

$$\nu_d = \frac{|\{0,1\}^k|}{\mu_{d,a}} = \frac{2^k}{\mu_{d,a}} \tag{2.10}$$

and

$$\varepsilon_{d,a}(\{\mathcal{B}_\nu\}) = \varepsilon_{d,a}.$$

(3D) A system of $\nu_d^{(a)}$ subsets $\mathcal{B}_\nu^0, \nu = 1, ..., \nu_d^{(a)}$, consisting of $\mu_d^{(a)}$ vectors $\mathbf{b} \in \{0,1\}_d^k$, will be referred to as a $(d|a, \mu_d^{(a)} \times \nu_d^{(a)}, \varepsilon_d^{(a)})$-*decomposition of the set* $\{0,1\}_d^k$ if

$$\bigcup_{\nu=1}^{\nu_d^{(a)}} \mathcal{B}_\nu^0 = \{0,1\}_d^k \tag{2.11}$$

and

$$\varepsilon_{d^{(a)}}(\{\mathcal{B}_\nu^0\}) = \varepsilon_d^{(a)}.$$

Lemma 1.5 Let $d = k\gamma$ and $a = k\alpha$.

(1L) One can find $\varepsilon_k, \delta_k \to 0$, as $k \to \infty$, such that there exists a $(d, \mu_d \times \nu_d, \varepsilon_d)$-pairwise disjoint decomposition of the set $\{0,1\}^k$ with

$$\begin{aligned} \mu_d &= 2^{k(1-H(\gamma)+\delta_k)}, \\ \varepsilon_d &= \varepsilon_k. \end{aligned} \tag{2.12}$$

(2L) One can find $\varepsilon_k, \delta_k \to 0$, as $k \to \infty$, such that there exists a $((d,a), \mu_{d,a} \times \nu_{d,a}, \varepsilon_{d,a})$-pairwise disjoint decomposition of the set $\{0,1\}^k$ with

$$\begin{aligned} \mu_{d,a} &= 2^{k(1-H(\alpha*\gamma)+\delta_k)}, \\ \varepsilon_{d,a} &= \varepsilon_k, \end{aligned} \tag{2.13}$$

where

$$\alpha * \gamma = \alpha(1-\gamma) + (1-\alpha)\gamma.$$

(3L) One can find $\varepsilon_k^0, \delta_k^0 \to 0$, as $k \to \infty$, such that there exists a $(d|a, \mu_d^{(a)} \times \nu_d^{(a)}, \varepsilon_d^{(a)})$-decomposition of the set $\{0,1\}_d^k \backslash \{0,1\}_{kq}^k$, where $q > 0$, with

$$\begin{aligned} \mu_d^{(a)} &= 2^{k(H(\alpha*\gamma)-H(\alpha)+\delta_k^0)}, \\ \nu_d^{(a)} &= k!, \\ \varepsilon_d^{(a)} &= \varepsilon_k^0, \end{aligned} \tag{2.14}$$

Proof To obtain (1.12) we take a best error-correcting code for a binary symmetric channel (BSC) with crossover probability $\gamma + \delta_k'$ as $\mathcal{B}_1$ and define all the other sets by the shifts $\mathbf{c} + \mathcal{B}_1$, $\mathbf{c} \in \{0,1\}_t^k$, where $t = 1, ..., d$. The construction, which provides (1.13) is similar. To obtain (1.14) we construct a best error-correcting code for a BSC with crossover probability $\alpha + \delta_k^{0\prime}$, whose codewords belong to $\{0,1\}_d^k$, define it as $\mathcal{B}_1^0$, and generate all the other sets by all possible permutations of the components of this code. An existence of codes satisfying (1.12)-(1.14) follows from the well-known results of coding theory [7]. Q.E.D.

Convention 1.6 In the further considerations we will use the relation $' \sim '$ and write

$$a_n \sim b_n$$

if

$$\lim_{n\to\infty} \frac{\log a_n}{n} = \lim_{n\to\infty} \frac{\log b_n}{n},$$

where we also assume that the limits exist. Furthermore, we write :

$$c_n \lesssim b_n$$

if $c_n \leq a_n$ and $a_n \sim b_n$. In particular, the inequalities (1.12)-(1.14) will be rewritten as

$$\begin{aligned} m_d &\sim 2^{k(1-H(\gamma))}, \\ m_{d,a} &\sim 2^{k(1-H(\alpha*\gamma))}, \\ m_d^{(a)} &\sim 2^{k(H(\alpha*\gamma)-H(\alpha))}. \end{aligned} \tag{2.15}$$

2.2 Identification schemes

Given observation $\mathbf{y}^* \in \{0,1\}^k$, let us introduce the set

$$\mathbf{X}_\tau^k(\mathbf{y}^*) = \mathbf{y}^* + \sum_{d=kq}^{\tau} \{0,1\}_d^k, \tag{2.16}$$

where the set $\{0,1\}_d^k$ is defined in (1.8) and the values of τ and q will be specified later. Let

$$\begin{aligned} c_d &= \binom{k}{d}, \\ p_d &= p^d(1-p)^{k-d}, \quad d = 0, ..., \tau, \\ P_\tau &= P(\mathbf{X}_\tau^k(\mathbf{y}^*)|\mathbf{y}^*), \\ C_\tau &= |\mathbf{X}_\tau^k(\mathbf{y}^*)|, \end{aligned} \tag{2.17}$$

where we have used the fact that the probability $P(\mathbf{X}_\tau^k(\mathbf{y}^*)|\mathbf{y}^*)$ and the cardinality $|\mathbf{X}_\tau^k(\mathbf{y}^*)|$ do not depend on $\mathbf{y}^*$. Note that if $\tau = k\beta > kp$ and $d = k\gamma$ then, as it is well-known [7],

$$\begin{aligned} c_d &\sim 2^{kH(\gamma)}, \\ C_\tau &\sim 2^{kH(\beta)}, \\ 1 - P_\tau &\sim 2^{-kD(\beta\|p)}, \end{aligned} \tag{2.18}$$

where

$$D(\beta \parallel p) = \beta \log \frac{\beta}{p} + (1-\beta) \log \frac{1-\beta}{1-p}$$

denotes the I-divergence between $(\beta, 1-\beta)$ and $(p, 1-p)$.

We assume that one of the codewords of the code G is transmitted and assign the decoding decision regions of the j-th receiver, $\mathcal{D}_j(\mathbf{y}^*)$, as follows :

$$\mathcal{D}_j(\mathbf{y}^*) \;=\; \bigcup_{\mathbf{x} \in \mathcal{F}_j(\mathbf{y}^*)} \{\mathcal{C}_j(\mathbf{x})\}\,, \tag{2.19}$$

where

$$\mathcal{F}_j(\mathbf{y}^*) \;=\; \bigcup_{\mathbf{x}^* \in \mathbf{X}_\tau^k(\mathbf{y}^*)} \bigcup_{\mathbf{z}} \{f_j(\mathbf{x}^*, \mathbf{z})\}\,. \tag{2.20}$$

Then

$$\lambda_j \;=\; 2^{-l} \sum_{\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}} P(\mathbf{x}^*, \mathbf{y}^*) \sum_{\mathbf{y}} W(\mathbf{y} | f_j(\mathbf{x}^*, \mathbf{z})) \cdot \chi\{\, \mathbf{y} \notin \mathcal{D}_j(\mathbf{y}^*)\, \} \tag{2.21}$$

and

$$\lambda_{ij} \;=\; 2^{-l} \sum_{\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}} P(\mathbf{x}^*, \mathbf{y}^*) \sum_{\mathbf{y}} W(\mathbf{y} | f_i(\mathbf{x}^*, \mathbf{z})) \cdot \chi\{\, \mathbf{y} \in \mathcal{D}_j(\mathbf{y}^*)\, \} \tag{2.22}$$

are the decoding error probabilities of the first and of the second kind at the output of the j-th receiver, provided that the message i was generated. If $\rho H(p) \geq C$, then $l = 0$ and we omit dependence on $\mathbf{z}$ in (2.5)-(2.7).

2.3 Ensembles of ID codes

If $\rho H(p) < C$, we introduce an ensemble of ID codes in such a way that the codewords, assigned to the message i, are selected from a fixed code G in accordance with the uniform distribution, and such a selection is realized independently for all i, $\mathbf{x}^* \in \{0,1\}^k$, and $\mathbf{z} \in \{0,1\}^l$. We denote the probability in this ensemble by $Pr\{\ \}$ and write :

$$Pr\{\ f_i(\mathbf{x}^*, \mathbf{z}) = \mathbf{c}\ \} \;=\; \begin{cases} 1/m, & \text{if } \mathbf{c} \in G, \\ 0, & \text{if } \mathbf{c} \notin G. \end{cases} \tag{2.23}$$

If $\rho H(p) \geq C$, we will use a different code ensemble. Its construction depends on a parameter $\alpha \in (0, 1/2)$, and we will denote the probability in this ensemble by $Pr_\alpha\{\ \}$. Let $a = k\alpha$ and $\mathcal{A} = \{\mathbf{x}_1^*, ..., \mathbf{x}_{m_a}^*\}$, where

$$m_a \;\sim\; 2^{k(1-H(\alpha))}, \tag{2.24}$$

be a 'good' code for a BSC with crossover probability α, i.e., the maximal decoding error probability tends to zero when k tends to infinity. Let

$$\mathbf{A}_a^k(\mathbf{x}_t^*) = \mathbf{x}_t^* + \sum_{d=0}^{a} \{0,1\}_d^k \tag{2.25}$$

and let $\mathbf{A}_a(\mathbf{x}_t^*)$ be the subset of $\mathbf{A}_a^k(\mathbf{x}_t^*)$ consisting of the vectors, which do not belong to $\mathbf{A}_a^k(\mathbf{x}_{t'}^*,)$ $t' \neq t$; $t = 1, ..., m_a$. Let

$$\mathbf{A}'_a = \{0,1\}^k \backslash \bigcup_{t=1}^{m_a} \mathbf{A}_a(\mathbf{x}_t^*). \tag{2.26}$$

Since the sets $\mathbf{A}_a^k(\mathbf{x}_t^*)$, $t = 1, ..., m_a$, are disjoint, for any $\mathbf{x}^* \in \{0,1\}^k$ we can either find a unique h such that $\mathbf{x}^* \in \mathbf{A}_a(\mathbf{x}_h^*)$, or claim that $\mathbf{x}^* \in \mathbf{A}'_a$. For all $\mathbf{x}^* \in \mathcal{A}$, let

$$Pr_\alpha\{\ f_i(\mathbf{x}^*) = \mathbf{c}\ |\mathcal{A}\ \} \ = \ \begin{cases} 1/m, & \text{if } \mathbf{c} \in G, \\ 0, & \text{if } \mathbf{c} \notin G, \end{cases} \tag{2.27}$$

and for all $\mathbf{x}^* \notin \mathcal{A}$, let

$$Pr_\alpha\{\ f_i(\mathbf{x}^*) = \mathbf{c}\ |\mathcal{A}\ \} \ = \ \begin{cases} f_i(\mathbf{x}_h^*), & \text{if } \mathbf{x}^* \notin \mathbf{A}'_\alpha, \\ \mathbf{c}_0, & \text{if } \mathbf{x}^* \in \mathbf{A}'_\alpha, \end{cases} \tag{2.28}$$

where $\mathbf{c}_0 \in G$ is some codeword, assigned in advance. We also need a randomization all over possible shifts of the set $\mathcal{A}$. Therefore, we define

$$Pr_\alpha\{\ f_i(\mathbf{x}^*) = \mathbf{c}\ \} = |\mathbf{A}_a^k|^{-1} \cdot \sum_{\Delta\mathbf{x}^* \in \mathbf{A}_a^k} Pr_\alpha\{\ f_i(\mathbf{x}^*) = \mathbf{c}\ |\mathcal{A} + \Delta\mathbf{x}^*\ \}, \tag{2.29}$$

where

$$\mathbf{A}_a^k = \sum_{d=0}^{a} \{0,1\}_d^k. \tag{2.30}$$

We also denote by

$$\varepsilon_\alpha = Pr_\alpha\{\ f_i(\mathbf{x}^*) = \mathbf{c}_0\ \} \tag{2.31}$$

the probability that a certain vector $\mathbf{x}^* \in \{0,1\}^k$ belongs to the set $\mathbf{A}'_\alpha$ and note that (3.7), (3.8) lead to ε_α, which does not depend on $\mathbf{x}^*$.

2.4 Direct coding theorem for the case $\rho H(p) < C$

Theorem 4.1 For all $\beta \in (p, 1/2]$ and $\sigma > 0$, the triple (R, e_1, e_2) such that

$$\begin{aligned} e_1 &= \min\{\ e(r),\ \rho D(\beta \,\|\, p)\ \}, \\ e_2 &= \min\{\ e(r),\ r - \rho H(\beta) - \sigma\ \}, \\ R &= \rho + \sigma - e_2 \end{aligned} \tag{2.32}$$

is a ρ-achievable ID triple.

Corollary 4.2

$$R_\rho \;\geq\; C+\rho(1-H(p)) \;\text{ for all }\; \rho<\rho_0, \tag{2.33}$$

where

$$\rho_0 = \frac{C}{H(p)}. \tag{2.34}$$

Proof We may assign

$$\begin{aligned} r &= C-\varepsilon_n, \\ \beta &= p+\delta_n, \\ \sigma &= C-\rho H(p)+\delta_n', \end{aligned} \tag{2.35}$$

where $\varepsilon_n, \delta_n, \delta_n' \to 0$, as $n\to\infty$. Then the exponents of the decoding error probabilities, given in (4.1), are still positive, and the rate R asymptotically coincides with the expression at the right hand side of (4.2). This expression is a lower bound on R_ρ since a more explicit analysis can give more tight results. Q.E.D.

The inequality (4.2) provides a straight line, given in Fig.3.

Proof of the theorem 4.1

Let us fix

$$\tau = k\beta > p$$

and assign a positive $q<p$ such that $D(q \parallel p) = D(\beta \parallel p)$. Then using (1.3) and (2.4)-(2.6) we write :

$$\begin{aligned} \lambda_j &\leq \sum_{\mathbf{y}^*} \sum_{\mathbf{x}^* \notin \mathbf{X}_\tau^k(\mathbf{y}^*)} P(\mathbf{x}^*,\mathbf{y}^*) \\ &+ 2^{-l} \sum_{\mathbf{x}^*,\mathbf{y}^*,\mathbf{z}} P(\mathbf{x}^*,\mathbf{y}^*) \sum_{\mathbf{y}\in\mathcal{C}(f_i(\mathbf{x}^*,\mathbf{z}))} W(\mathbf{y}|f_i(\mathbf{x}^*,\mathbf{z})) \\ &\lesssim 2\cdot 2^{-\rho n D(\beta\|p)} + 2^{-ne(r)}. \end{aligned} \tag{2.36}$$

Therefore, e_1, given in (4.1), is an asymptotically achievable exponent of the decoding error probability of the first kind.

Using (1.3) and (2.4)-(2.7) we can also estimate λ_{ij} as follows :

$$\begin{aligned}
\lambda_{ij} &= 2^{-l}\sum_{\mathbf{x}^*,\mathbf{y}^*,\mathbf{z}} P(\mathbf{x}^*,\mathbf{y}^*)\sum_{\mathbf{y}\notin\mathcal{C}(f_i(\mathbf{x}^*,\mathbf{z}))} W(\mathbf{y}|f_i(\mathbf{x}^*,\mathbf{z}))\cdot\chi\{\ \mathbf{y}\in\mathcal{D}_j(\mathbf{y}^*)\ \} \qquad (2.37)\\
&+ 2^{-l}\sum_{\mathbf{x}^*,\mathbf{y}^*,\mathbf{z}} P(\mathbf{x}^*,\mathbf{y}^*)\sum_{\mathbf{y}\in\mathcal{C}(f_i(\mathbf{x}^*,\mathbf{z}))} W(\mathbf{y}|f_i(\mathbf{x}^*,\mathbf{z}))\cdot\chi\{\ \mathbf{y}\in\mathcal{D}_j(\mathbf{y}^*)\ \}\\
&\leq 2^{-ne(r)}\\
&+ 2^{-l}\sum_{\mathbf{x}^*,\mathbf{y}^*,\mathbf{z}} P(\mathbf{y}^*)P(\mathbf{x}^*|\mathbf{y}^*)\sum_{\mathbf{y}\in\mathcal{C}(f_i(\mathbf{x}^*,\mathbf{z}))} W(\mathbf{y}|f_i(\mathbf{x}^*,\mathbf{z}))\cdot\chi\{\ \mathbf{y}\in\mathcal{D}_j(\mathbf{y}^*)\ \}\\
&= 2^{-ne(r)} + 2^{-(k+l)}\sum_{d=kq}^{\tau} p_d \sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}),
\end{aligned}$$

where

$$\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) \ = \sum_{\mathbf{x}^*\in\mathbf{S}_d^k(\mathbf{y}^*)} \chi\{\ f_i(\mathbf{x}^*,\mathbf{z})\in\mathcal{F}_j(\mathbf{y}^*)\ \} \qquad (2.38)$$

and notations (1.8), (2.2) are used.

Let $d = k\gamma$. Then, given $\Lambda > 0$, we can use Chernov's inequality and write :

$$\log Pr\left\{\sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) > 2^{k+l}c_d\Lambda\right\} \leq -s\cdot 2^{k+l}c_d\Lambda + \log G^{(d)}(s), \qquad (2.39)$$

where $s \geq 0$,

$$G^{(d)}(s) \ = \ E\left[\prod_{\mathbf{y}^*,\mathbf{z}} 2^{s\cdot\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z})}\right], \qquad (2.40)$$

and $E\,[\,]$ denotes the expectation in the code ensemble.

The result below is proved in Appendix.

Lemma 4.3 If there exists a $(d, \mu_d \times \nu_d, \varepsilon_d)$-pairwise disjoint decomposition of the set $\{0,1\}^k$, then

$$\log G^{(d)}(s) \ \leq \ \mu_d c_d 2^l \cdot (\ \log g(s\nu_d) + s\varepsilon_d\ ), \qquad (2.41)$$

where

$$g(s) = 1 - \Pi + \Pi\cdot 2^s, \qquad (2.42)$$
$$\Pi = \frac{C_\tau 2^l}{m} \sim 2^{-n(r-\rho H(\beta)-\sigma)},$$

where $l = n\sigma$ and notations (2.2) are used.

Using (1.10), (1.12), and (2.3) in (4.10) we have :

$$\log G^{(d)}(s) \quad \lesssim \quad 2^{k+l} \cdot \left(\log g \left(s \cdot 2^{kH(\gamma)} \right) + s\varepsilon_d \right). \tag{2.43}$$

Hence, (4.8) and (4.12) lead to the following asymptotic inequality :

$$\begin{aligned} &\log Pr \left\{ \sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) > 2^{k+l} c_d \Lambda \right\} \\ \lesssim \quad & 2^{k+l} \left(-s \cdot 2^{kH(\gamma)} \Lambda + \log g \left(s \cdot 2^{kH(\gamma)} \right) + s\varepsilon_d \right). \end{aligned} \tag{2.44}$$

Let

$$\Lambda > \Pi. \tag{2.45}$$

Then we can set

$$s = 2^{-kH(\gamma)} \cdot \log \frac{\Lambda}{1-\Lambda} \frac{1-\Pi}{\Pi} \tag{2.46}$$

and

$$\log Pr \left\{ \sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) > 2^{k+l} c_d \Lambda \right\} \lesssim -2^{k+l} \cdot D(\Lambda \parallel \Pi). \tag{2.47}$$

The estimate (4.16) does not depend on d. Therefore,

$$\begin{aligned} &Pr \left\{ \sum_{d=kq}^{\tau} p_d \sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) > P_\tau \cdot \Lambda \right\} \\ = \quad & Pr \left\{ \sum_{d=kq}^{\tau} p_d \left(\sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) - c_d \Lambda \right) > 0 \right\} \\ \le \quad & \sum_{d=kq}^{\tau} Pr \left\{ \sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) > c_d \Lambda \right\} \\ \lesssim \quad & \tau \exp_2 \left\{ -2^{k+l} \cdot D(\Lambda \parallel \Pi) \right\}. \end{aligned} \tag{2.48}$$

If $\Lambda \to 0$, as $n \to \infty$, then

$$D(\Lambda \parallel \Pi) \sim \Lambda, \tag{2.49}$$

and (4.17), (4.18) lead to the inequality :

$$Pr \left\{ \sum_{d=kq}^{\tau} p_d \sum_{\mathbf{y}^*,\mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) > P_\tau \cdot \Lambda \right\} \lesssim \exp_2 \left\{ -2^{k+l} \cdot \Lambda \right\}. \tag{2.50}$$

Let us assign

$$\Lambda = P_\tau \cdot (2^{-ne_2} + 2^{-ne(r)}), \tag{2.51}$$

where e_2 is given in (4.1). Then we can refer to the considerations of subsection 2.1 and using (4.6), (4.19) conclude that the rate R_ρ, given in (4.1), is asymptotically achievable. Q.E.D.

2.5 Direct coding theorem for the case $\rho H(p) \geq C$

Theorem 5.1 For all $\beta \in (p, 1/2]$ and $\alpha > 0$, the triple (R, e_1, e_2) such that

$$\begin{aligned} e_1 &= \min\{\, e(r),\ \rho D(\beta \,\|\, p)\,\}, \\ e_2 &= \min\{\, e(r),\ r - \rho(H(\alpha * \beta) - H(\alpha))\,\}, \\ R &= \rho(1 - H(\alpha)) - e_2 \end{aligned} \tag{2.52}$$

is a ρ-achievable ID triple.

Corollary 5.2

$$\begin{aligned} R_\rho &\geq \rho(1 - H(\alpha)) \\ &= C + \rho(1 - H(\alpha * p)) \ \text{ for all } \ \rho \geq \rho_0, \end{aligned} \tag{2.53}$$

where the parameter ρ_0 is defined in (4.3), and α is chosen in such a way that

$$C = \rho(H(\alpha * p) - H(\alpha)). \tag{2.54}$$

Proof We substitute the expressions at the right hand side of (4.4) for r and β into (5.1) and complete the proof. Q.E.D.

The inequality (5.2) provides a convex up curve, given in Fig.3.

Proof of the theorem 5.1 The same considerations as in the proof of theorem 4.1 lead to the inequalities :

$$\begin{aligned} \lambda_j &\leq 2 \cdot 2^{-\rho n D(\beta \| p)} + 2^{-ne(r)}, \\ \lambda_{ij} &\leq 2^{-ne(r)} + 2^{-k} \sum_{d=kq}^{\tau} p_d \sum_{\mathbf{y}^*} \eta_{ij}^{(d)}(\mathbf{y}^*), \end{aligned} \tag{2.55}$$

where the variables $\eta_{ij}^{(d)}(\mathbf{y}^*)$ are defined similarly to (4.7), i.e.,

$$\eta_{ij}^{(d)}(\mathbf{y}^*) = \sum_{\mathbf{x}^* \in \mathbf{S}_d^k(\mathbf{y}^*)} \chi\{\, f_i(\mathbf{x}^*) \in \mathcal{F}_j(\mathbf{y}^*)\,\}. \tag{2.56}$$

Furthermore,

$$\log Pr\left\{ \sum_{\mathbf{y}^*} \eta_{ij}^{(d)}(\mathbf{y}^*) > 2^k c_d \Lambda \right\} \leq -s \cdot 2^k c_d \Lambda + \log G_\alpha^{(d)}(s), \tag{2.57}$$

where $\Lambda > 0$ is a given constant, $s \geq 0$,

$$G_\alpha^{(d)}(s) = E_\alpha\left[\prod_{\mathbf{y}^*} 2^{s \cdot \eta_{ij}^{(d)}(\mathbf{y}^*)} \right], \tag{2.58}$$

and $E_\alpha[\,]$ denotes the expectation in the code ensemble.

All the peculiarities of the analysis are concentraned in a new upper bound on $G_\alpha^{(d)}(s)$. The result below is proved in Appendix.

Lemma 5.3 Let $a = k\alpha$, $\tau = k\beta$, and $d = k\gamma$.

1. If there exists a $((d,a), \mu_{d,a} \times \nu_{d,a}, \varepsilon_{d,a})$-pairwise disjoint decomposition of the set $\{0,1\}^k$, then
$$\log G_\alpha^{(d)}(s) \ \le \ \mu_{d,a} \cdot \left(\log \hat{g}^{(d)}(s\nu_{d,a}) + s\varepsilon_{d,a} \right), \tag{2.59}$$
where
$$\hat{g}_\alpha^{(d)}(s) \ = \ E_\alpha\left[2^{s\cdot\eta_{ij}^{(d)}(\mathbf{y}^*)} \right] \tag{2.60}$$
for some $\mathbf{y}^* \in \{0,1\}^k$.

2. Let $\mathcal{B}_\nu^0, \nu = 1, ..., \nu_d^{(a)}$, be a $(d|a, \mu_d^{(a)} \times \nu_d^{(a)}, \varepsilon_d^{(a)})$-decomposition of the set $\{0,1\}_d^k$ such that $n(\mathbf{x}^*)$ is the number of occurences of $\mathbf{x}^*$ in the collection $\mathcal{B}_\nu^0, \nu = 1, ..., \nu_d^{(a)}$, i.e.,
$$n(\mathbf{x}^*) = \sum_{\nu=1}^{\nu_d^{(a)}} \chi\{\ \mathbf{x}^* \in \mathcal{B}_\nu^0\ \}. \tag{2.61}$$
Then
$$\hat{g}_\alpha^{(d)}(s) \ \le \ \prod_{\nu=1}^{\nu_d^{(a)}} \left(\prod_{\mathbf{x}^*\in\mathcal{B}_\nu^0} g_\alpha\left(s\nu_d^{(a)}/n(\mathbf{x}^*) \right) \right)^{1/\nu_d^{(a)}}, \tag{2.62}$$
where
$$\begin{aligned} g_\alpha(s) &= (1-\varepsilon_d^{(a)} - \varepsilon_\alpha)\cdot(1-\Pi_\alpha + \Pi_\alpha\cdot 2^s) + (\varepsilon_d^{(a)} + \varepsilon_\alpha)\cdot 2^s \\ &\sim 1 - \Pi_\alpha + \Pi_\alpha \cdot 2^s, \\ \Pi_\alpha &\sim 2^{-n(r-\rho(H(\alpha*\beta)-H(\alpha)))}. \end{aligned} \tag{2.63}$$

Let us use the construction of a $(d|a, \mu_d^{(a)} \times \nu_d^{(a)}, \varepsilon_d^{(a)})$-decomposition of the set $\{0,1\}_d^k$, described in the proof of lemma 1.5. Then, as it easy to see,
$$n(\mathbf{x}^*) = \frac{k!\cdot\mu_d^{(a)}}{c_d} \tag{2.64}$$
Thus, substituting (1.14), (5.13) to (5.11) we obtain :
$$\begin{aligned} \log \hat{g}_\alpha^{(d)}(s) &\lesssim \mu_d^{(a)} \cdot \log g_\alpha\left(sc_d 2^{-k(H(\alpha*\gamma)-H(\alpha))} \right) \\ &\sim 2^{k(H(\alpha*\gamma)-H(\alpha))} \log g_\alpha\left(sc_d 2^{-k(H(\alpha*\gamma)-H(\alpha))} \right). \end{aligned} \tag{2.65}$$

Hence (1.13), (1.15), (5.8), and (5.14) lead to the inequality :

$$\log G_\alpha^{(d)}(s) \;\leq\; 2^{k(1-H(\alpha))} \cdot \left(\log g_\alpha \left(s c_d \cdot 2^{kH(\alpha)} \right) + s\varepsilon_{d,a} \right) \tag{2.66}$$

and

$$\log Pr \left\{ \sum_{\mathbf{y}^*} \eta_{ij}^{(d)}(\mathbf{y}^*) > 2^k c_d \Lambda \right\} \tag{2.67}$$
$$\lesssim\; 2^{k(1-H(\alpha))} \left(-s c_d \cdot 2^{kH(\alpha)} \Lambda + \log g_\alpha \left(s c_d \cdot 2^{kH(\alpha)} \right) + s\varepsilon_{d,a} \right).$$

Let

$$\Lambda > \Pi_\alpha. \tag{2.68}$$

Then we can set

$$s = c_d^{-1} \cdot 2^{-kH(\alpha)} \cdot \log \frac{\Lambda}{1-\Lambda} \frac{1-\Pi_\alpha}{\Pi_\alpha} \tag{2.69}$$

and

$$\log Pr \left\{ \sum_{\mathbf{y}^*, \mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*, \mathbf{z}) > 2^k c_d \Lambda \right\} \lesssim -2^{k(1-H(\alpha))} \cdot D(\Lambda \parallel \Pi_\alpha). \tag{2.70}$$

We set $\Lambda \to 0$, as $n \to \infty$, in such a way that (5.17) is valid, note that

$$D(\Lambda \parallel \Pi_\alpha) \sim \Lambda, \tag{2.71}$$

and repeat the considerations of subsection 2.4 to complete the proof. Q.E.D.

2.6 Direct coding theorem for the case $\rho H(p) \geq C$

Theorem 5.1 For all $\beta \in (p, 1/2]$ and $\alpha > 0$, the triple (R, e_1, e_2) such that

$$\begin{aligned} e_1 &= \min\{\, e(r),\ \rho D(\beta \parallel p) \,\}, \\ e_2 &= \min\{\, e(r),\ r - \rho(H(\alpha * \beta) - H(\alpha)) \,\}, \\ R &= \rho(1 - H(\alpha)) - e_2 \end{aligned} \tag{2.72}$$

is a ρ-achievable ID triple.

Corollary 5.2

$$\begin{aligned} R_\rho &\geq \rho(1 - H(\alpha)) \\ &= C + \rho(1 - H(\alpha * p)) \ \text{ for all } \ \rho \geq \rho_0, \end{aligned} \tag{2.73}$$

where the parameter ρ_0 is defined in (4.3), and α is chosen in such a way that

$$C = \rho(H(\alpha * p) - H(\alpha)). \tag{2.74}$$

Proof We substitute the expressions at the right hand side of (4.4) for r and β into (5.1) and complete the proof. Q.E.D.

The inequality (5.2) provides a convex up curve, given in Fig.3.

Proof of the theorem 5.1 The same considerations as in the proof of theorem 4.1 lead to the inequalities :

$$\begin{aligned} \lambda_j &\leq 2 \cdot 2^{-\rho n D(\beta \| p)} + 2^{-ne(r)}, \\ \lambda_{ij} &\leq 2^{-ne(r)} + 2^{-k} \sum_{d=kq}^{\tau} p_d \sum_{\mathbf{y}^*} \eta_{ij}^{(d)}(\mathbf{y}^*), \end{aligned} \tag{2.75}$$

where the variables $\eta_{ij}^{(d)}(\mathbf{y}^*)$ are defined similarly to (4.7), i.e.,

$$\eta_{ij}^{(d)}(\mathbf{y}^*) = \sum_{\mathbf{x}^* \in \mathbf{S}_d^k(\mathbf{y}^*)} \chi\{ f_i(\mathbf{x}^*) \in \mathcal{F}_j(\mathbf{y}^*) \}. \tag{2.76}$$

Furthermore,

$$\log Pr \left\{ \sum_{\mathbf{y}^*} \eta_{ij}^{(d)}(\mathbf{y}^*) > 2^k c_d \Lambda \right\} \leq -s \cdot 2^k c_d \Lambda + \log G_\alpha^{(d)}(s), \tag{2.77}$$

where $\Lambda > 0$ is a given constant, $s \geq 0$,

$$G_\alpha^{(d)}(s) = E_\alpha \left[\prod_{\mathbf{y}^*} 2^{s \cdot \eta_{ij}^{(d)}(\mathbf{y}^*)} \right], \tag{2.78}$$

and $E_\alpha[\,]$ denotes the expectation in the code ensemble.

All the peculiarities of the analysis are concentraned in a new upper bound on $G_\alpha^{(d)}(s)$. The result below is proved in Appendix.

Lemma 5.3 Let $a = k\alpha$, $\tau = k\beta$, and $d = k\gamma$.

1. If there exists a $((d,a), \mu_{d,a} \times \nu_{d,a}, \varepsilon_{d,a})$-pairwise disjoint decomposition of the set $\{0,1\}^k$, then

   $$\log G_\alpha^{(d)}(s) \leq \mu_{d,a} \cdot \left(\log \hat{g}^{(d)}(s\nu_{d,a}) + s\varepsilon_{d,a} \right), \tag{2.79}$$

   where

   $$\hat{g}_\alpha^{(d)}(s) = E_\alpha \left[2^{s \cdot \eta_{ij}^{(d)}(\mathbf{y}^*)} \right] \tag{2.80}$$

   for some $\mathbf{y}^* \in \{0,1\}^k$.

2. Let $\mathcal{B}^0_\nu, \nu = 1, ..., \nu^{(a)}_d$, be a $(d|a, \mu^{(a)}_d \times \nu^{(a)}_d, \varepsilon^{(a)}_d)$-decomposition of the set $\{0,1\}^k_d$ such that $n(\mathbf{x}^*)$ is the number of occurences of $\mathbf{x}^*$ in the collection $\mathcal{B}^0_\nu, \nu = 1, ..., \nu^{(a)}_d$, i.e.,

$$n(\mathbf{x}^*) = \sum_{\nu=1}^{\nu^{(a)}_d} \chi\{ \mathbf{x}^* \in \mathcal{B}^0_\nu \}. \tag{2.81}$$

Then

$$\hat{g}^{(d)}_\alpha(s) \ \le \ \prod_{\nu=1}^{\nu^{(a)}_d} \left(\prod_{\mathbf{x}^* \in \mathcal{B}^0_\nu} g_\alpha \left(s\nu^{(a)}_d / n(\mathbf{x}^*) \right) \right)^{1/\nu^{(a)}_d}, \tag{2.82}$$

where

$$\begin{aligned} g_\alpha(s) &= (1 - \varepsilon^{(a)}_d - \varepsilon_\alpha) \cdot (1 - \Pi_\alpha + \Pi_\alpha \cdot 2^s) + (\varepsilon^{(a)}_d + \varepsilon_\alpha) \cdot 2^s \\ &\sim 1 - \Pi_\alpha + \Pi_\alpha \cdot 2^s, \\ \Pi_\alpha &\sim 2^{-n(r - \rho(H(\alpha*\beta) - H(\alpha)))}. \end{aligned} \tag{2.83}$$

Let us use the construction of a $(d|a, \mu^{(a)}_d \times \nu^{(a)}_d, \varepsilon^{(a)}_d)$-decomposition of the set $\{0,1\}^k_d$, described in the proof of lemma 1.5. Then, as it easy to see,

$$n(\mathbf{x}^*) = \frac{k! \cdot \mu^{(a)}_d}{c_d} \tag{2.84}$$

Thus, substituting (1.14), (5.13) to (5.11) we obtain :

$$\begin{aligned} \log \hat{g}^{(d)}_\alpha(s) &\lesssim \mu^{(a)}_d \cdot \log g_\alpha \left(s c_d 2^{-k(H(\alpha*\gamma) - H(\alpha))} \right) \\ &\sim 2^{k(H(\alpha*\gamma) - H(\alpha))} \log g_\alpha \left(s c_d 2^{-k(H(\alpha*\gamma) - H(\alpha))} \right). \end{aligned} \tag{2.85}$$

Hence (1.13), (1.15), (5.8), and (5.14) lead to the inequality :

$$\log G^{(d)}_\alpha(s) \ \le \ 2^{k(1 - H(\alpha))} \cdot \left(\log g_\alpha \left(s c_d \cdot 2^{kH(\alpha)} \right) + s\varepsilon_{d,a} \right) \tag{2.86}$$

and

$$\begin{aligned} &\log Pr \left\{ \sum_{\mathbf{y}^*} \eta^{(d)}_{ij}(\mathbf{y}^*) > 2^k c_d \Lambda \right\} \\ &\lesssim \ 2^{k(1 - H(\alpha))} \left(-s c_d \cdot 2^{kH(\alpha)} \Lambda + \log g_\alpha \left(s c_d \cdot 2^{kH(\alpha)} \right) + s\varepsilon_{d,a} \right). \end{aligned} \tag{2.87}$$

Let

$$\Lambda > \Pi_\alpha. \tag{2.88}$$

Then we can set

$$s = c_d^{-1} \cdot 2^{-kH(\alpha)} \cdot \log \frac{\Lambda}{1-\Lambda} \frac{1-\Pi_\alpha}{\Pi_\alpha} \tag{2.89}$$

and

$$\log Pr \left\{ \sum_{\mathbf{y}^*, \mathbf{z}} \eta_{ij}^{(d)}(\mathbf{y}^*, \mathbf{z}) > 2^k c_d \Lambda \right\} \lesssim -2^{k(1-H(\alpha))} \cdot D(\Lambda \parallel \Pi_\alpha). \tag{2.90}$$

We set $\Lambda \to 0$, as $n \to \infty$, in such a way that (5.17) is valid, note that

$$D(\Lambda \parallel \Pi_\alpha) \sim \Lambda, \tag{2.91}$$

and repeat the considerations of subsection 2.4 to complete the proof. Q.E.D.

2.7 Calculation of the function ΔR_∞

Using (5.2) we conclude that additional quantity in the identification rate, which we gain because of the observations of the process, is measured as $\rho(1 - H(\alpha * p))$, where the parameter α is defined by the equation (5.3). Let us consider the function

$$\Delta R_\infty = \lim_{\rho \to \infty} \rho \cdot (1 - H(\alpha * p)), \tag{2.92}$$

which derives this additional quantity conditioned on the infinite length of the vectors $\mathbf{x}^*, \mathbf{y}^*$. This function can be calculated exactly, as it follows from the result below.

Lemma 6.1

$$\Delta R_\infty = \frac{(1-2p)^2}{1-(1-2p)^2} \cdot C. \tag{2.93}$$

Proof The parameter α satisfying (5.3) is a function of ρ, and we may write :

$$\begin{aligned} \alpha &= \frac{1}{2} - \varepsilon_\rho, \\ \alpha * p &= \frac{1}{2} - (1-2p) \cdot \varepsilon_\rho, \end{aligned}$$

where

$$\varepsilon_\rho \to 0, \quad \text{as} \quad \rho \to \infty. \tag{2.94}$$

Since, for all $\varepsilon \in (0, 1/2)$,

$$\begin{aligned} \log \left(\frac{1}{2} - \varepsilon \right) &= -1 - \frac{1}{\ln 2} \cdot \sum_{i \geq 1} \frac{(2\varepsilon)^i}{i}, \\ \log \left(\frac{1}{2} + \varepsilon \right) &= -1 - \frac{1}{\ln 2} \cdot \sum_{i \geq 1} (-1)^i \cdot \frac{(2\varepsilon)^i}{i}, \end{aligned}$$

we have

$$
\begin{aligned}
H(\alpha) &= 1-\frac{1}{\ln 2}\cdot\sum_{k\geq 1}(2\varepsilon_\rho)^{2k}\cdot\left(\frac{1}{2k-1}-\frac{1}{2k}\right),\\
H(\alpha * p) &= 1-\frac{1}{\ln 2}\cdot\sum_{k\geq 1}(1-2p)^{2k}\cdot(2\varepsilon_\rho)^{2k}\cdot\left(\frac{1}{2k-1}-\frac{1}{2k}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\Delta R_\infty &= \lim_{\rho\to\infty}\frac{\rho}{\ln 2}\cdot\sum_{k\geq 1}(1-2p)^{2k}\cdot(2\varepsilon_\rho)^{2k}\cdot\left(\frac{1}{2k-1}-\frac{1}{2k}\right),\\
C &= \frac{\rho}{\ln 2}\cdot\sum_{k\geq 1}(1-(1-2p)^{2k})\cdot(2\varepsilon_\rho)^{2k}\cdot\left(\frac{1}{2k-1}-\frac{1}{2k}\right).
\end{aligned}
$$

Therefore,

$$
\Delta R_\infty = \lim_{\rho\to\infty}\frac{2((1-2p)^2)+\delta_\rho^{(R)}}{2(1-(1-2p)^2)+\delta_\rho^{(C)}}\cdot C, \tag{2.95}
$$

where

$$
\begin{aligned}
\delta_\rho^{(R)} &= \sum_{k\geq 2}(1-2p)^{2k}\cdot(2\varepsilon_\rho)^{2k-2}\cdot\left(\frac{1}{2k-1}-\frac{1}{2k}\right),\\
\delta_\rho^{(C)} &= \sum_{k\geq 2}(1-(1-2p)^{2k})\cdot(2\varepsilon_\rho)^{2k-2}\cdot\left(\frac{1}{2k-1}-\frac{1}{2k}\right).
\end{aligned}
$$

Using (6.3) we conclude that

$$
\delta_\rho^{(R)}, \delta_\rho^{(C)} \to 0, \quad \text{as} \quad \rho\to\infty, \tag{2.96}
$$

and (6.2) follows from (6.4), (6.5). Q.E.D.

3 Acknowledgement

The authors are greatful to Prof. Imre Csiszár for interesting discussions.

Appendix

1 Proof of Lemma 4.3

Let $\mathcal{B}_\nu, \nu = 1, ..., \nu_d$ be a $(d, \mu_d \times \nu_d, \varepsilon_d)$-pairwise disjoint decomposition of the set $\{0,1\}^k$. Then

$$\begin{aligned} G^{(d)}(s) &= E\left[\prod_{\mathbf{y}^*,\mathbf{z}} 2^{s\cdot\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z})}\right] \\ &= \prod_{\mathbf{z}} E\left[\prod_{\mathbf{y}^*} 2^{s\cdot\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z})}\right] \\ &= \prod_{\mathbf{z}} E\left[\prod_{\nu=1}^{\nu_d}\prod_{\mathbf{y}^*\in\mathcal{B}_\nu} 2^{s\cdot\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z})}\right] \\ &\leq \prod_{\mathbf{z}}\prod_{\nu=1}^{\nu_d} E^{1/\nu_\tau}\left[\prod_{\mathbf{y}^*\in\mathcal{B}_\nu} 2^{s\nu_d\cdot\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z})}\right], \end{aligned} \tag{A.1}$$

where Hölder's inequality was used.

Given ν, let $\mathbf{S}_d(\mathbf{y}^*)$, $\mathbf{y} \in \mathcal{B}_\nu$, be a set consisting of the vectors $\mathbf{x}^*$, which do not belong to the sets $\mathbf{S}_d^k(\mathbf{y}^{*\prime})$, $\mathbf{y}^{*\prime} \in \mathcal{B}_\nu\backslash\{\mathbf{y}^*\}$. Then

$$\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z}) \;\leq\; c_d\varepsilon_d + \sum_{\mathbf{x}^*\in\mathbf{S}_d(\mathbf{y}^*)} \chi\{\, f_i(\mathbf{x}^*,\mathbf{z}) \in \mathcal{F}_j(\mathbf{y}^*)\,\}$$

and

$$\begin{aligned} E\left[\prod_{\mathbf{y}^*\in\mathcal{B}_\nu} 2^{s\nu_d\cdot\eta_{ij}^{(d)}(\mathbf{y}^*,\mathbf{z})}\right] &\leq 2^{sc_d\mu_d\nu_d\varepsilon_d}\cdot E\left[\prod_{\mathbf{y}^*\in\mathcal{B}_\nu}\prod_{\mathbf{x}^*\in\mathbf{S}_d(\mathbf{y}^*)} 2^{s\nu_d\cdot\eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*,\mathbf{z})}\right] \\ &= 2^{sc_d\mu_d\nu_d\varepsilon_d}\cdot \prod_{\mathbf{y}^*\in\mathcal{B}_\nu}\prod_{\mathbf{x}^*\in\mathbf{S}_d(\mathbf{y}^*)} E\left[2^{s\nu_d\cdot\eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*,\mathbf{z})}\right], \end{aligned} \tag{A.2}$$

where

$$\eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*,\mathbf{z}) = \chi\{\, f_i(\mathbf{x}^*,\mathbf{z}) \in \mathcal{F}_j(\mathbf{y}^*)\,\}.$$

Let us represent the expectation $E[\;]$ as a concatenation of the expectation $E_i[\;]$ taken on all assignements of the codewords for the message i and the expectation $E_j[\;]$ taken on all assignements of the codewords for the message j, i.e.,

$$E\,[\,(*)\,] = E_j\,[E_i\,[\,(*)\,]]\,.$$

The codewords, which are used to encode the i-th message, are independent random vectors for different $\mathbf{x}^*$. Therefore,

$$
\begin{aligned}
E\left[\, 2^{s\nu_d \cdot \eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*,\mathbf{z})} \,\right] &= E_j\left[\, E_i\left[\, 2^{s\nu_d \cdot \eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*,\mathbf{z})} \,\right]\,\right] \\
&= E_j\left[\, 1 - \frac{|\mathcal{F}_j(\mathbf{y}^*)|}{m} + \frac{|\mathcal{F}_j(\mathbf{y}^*)|}{m} 2^{s\nu_d} \,\right] \\
&\leq g(s\nu_d),
\end{aligned} \tag{A.3}
$$

where the inequality $|\mathcal{F}_j(\mathbf{y}^*)| \leq C_\tau 2^l$ and notations (4.11) were used. Combining (A.1)-(A.3) we complete the proof. Q.E.D.

2 Proof of Lemma 5.3

The proof of the first part of lemma 5.3 repeats the steps (A.1), (A.2) with the accuracy to notations.

Let $\mathcal{B}^0_\nu, \nu = 1, ..., \nu_d^{(a)}$, be a $(d|a, \mu_d^{(a)} \times \nu_d^{(a)}, \varepsilon_d^{(a)})$-decomposition of the set $\{0,1\}^k_d$. Then

$$
\mathcal{B}_\nu = \mathbf{y}^* + \mathcal{B}^0_\nu, \tag{A.4}
$$

$\nu = 1, ..., \nu_d^{(a)}$, is a $(d|a, \mu_d^{(a)} \times \nu_d^{(a)}, \varepsilon_d^{(a)})$-design of the set $\mathbf{y}^* + \{0,1\}^k_d$. For all $\mathbf{x}^* \in \mathbf{X}^k_d(\mathbf{y}^*)$, let $n'(\mathbf{x}^*)$, be the number of occurences of $\mathbf{x}^*$ in the collection $\mathcal{B}_\nu, \nu = 1, ..., \nu_d^{(a)}$, i.e.,

$$
n'(\mathbf{x}^*) = \sum_{\nu=1}^{\nu_d^{(a)}} \chi\{\ \mathbf{x}^* \in \mathcal{B}_\nu\ \}. \tag{A.5}
$$

Then

$$
\begin{aligned}
\hat{g}^{(d)}_\alpha(s) &= E_\alpha\left[\, 2^{s \cdot \eta_{ij}^{(d)}(\mathbf{y}^*)} \right] \\
&= E_\alpha\left[\, \prod_{\mathbf{x}^* \in \mathbf{S}^k_d(\mathbf{y}^*)} 2^{s \cdot \eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*)} \,\right] \\
&= E_\alpha\left[\, \prod_{\nu=1}^{\nu_d^{(a)}} \prod_{\mathbf{x}^* \in \mathcal{B}_\nu} 2^{s \cdot \eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*)/n'(\mathbf{x}^*)} \,\right] \\
&\leq \prod_{\nu=1}^{\nu_d^{(a)}} E_\alpha^{1/\nu_d^{(a)}}\left[\, \prod_{\mathbf{x}^* \in \mathcal{B}_\nu} 2^{s\nu_d^{(a)} \cdot \eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*)/n'(\mathbf{x}^*)} \,\right],
\end{aligned} \tag{A.6}
$$

where

$$
\eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*) = \chi\{\ f_i(\mathbf{x}^*) \in \mathcal{F}_j(\mathbf{y}^*)\ \}.
$$

Given shift $\Delta\mathbf{x}^* \in \mathbf{A}_a^k$, the codewords of the j-th message are assigned for all vectors, belonging to the set

$$\hat{\mathcal{F}}_j(\mathbf{y}^*) = \left\{ \mathbf{x}^* \in \mathcal{A} : \; (\mathbf{x}^* + \Delta\mathbf{x}^* + \mathbf{A}_{a)}^k) \bigcap (\mathbf{y}^* + \{0,1\}_{d)}^k) \neq \emptyset \right\}. \tag{A.7}$$

It is known that

$$|\hat{\mathcal{F}}_j(\mathbf{y}^*)| \sim m_{\alpha,\beta}, \tag{A.8}$$

where

$$m_{\alpha,\beta} = 2^{k(H(\alpha*\beta)-H(\alpha))}. \tag{A.9}$$

Since we are dealing with an upper bound, we can set that all these codewords are different, and since the codewords of the i-th message are independent random variables, fix these codewords as $\mathbf{c}_1, ..., \mathbf{c}_{m_{\alpha,\beta}}$. Thus,

$$\begin{aligned} & E_\alpha \left[\prod_{\mathbf{x}^* \in \mathcal{B}_\nu} 2^{s\nu_d^{(a)} \cdot \eta_{ij}^{(d)}(\mathbf{x}^*|\mathbf{y}^*)/n'(\mathbf{x}^*)} \right] \\ \leq \;\; & E_\alpha \left[\prod_{\mathbf{x}^* \in \mathcal{B}_\nu} 2^{s\nu_d^{(a)}/n'(\mathbf{x}^*)} \cdot \chi\{ \; f_i(\mathbf{x}^*) \in \{\mathbf{c}_1, ..., \mathbf{c}_{m_{\alpha,\beta}}\} \; \} \right]. \end{aligned} \tag{A.10}$$

Using the symmetry properties of the code ensemble, defined by (3.5)-(3.8), we note that, with the probability $\varepsilon_\alpha + \varepsilon_d^{(a)}$, a codeword assigned to a vector $\mathbf{x}^* \in \mathcal{B}_\nu$, is either fixed or depends on the codewords assigned to the vectors belonging to the set $\mathcal{B}_\nu \backslash \{\mathbf{x}^*\}$. Otherwise, this codeword is an independent random vector, and using (5.12) and (A.6)-(A.10) we complete the proof of (5.11), where $n(\mathbf{x}^*)$ are replaced with $n'(\mathbf{x}^*)$. However, the shifts (A.4) keep $\{n(\mathbf{x}^*)\}$ as the set of all possible values of $n'(\mathbf{x}^*)$, defined in (A.5). Therefore, (5.11) is also valid with the values of $n(\mathbf{x}^*)$. Q.E.D.

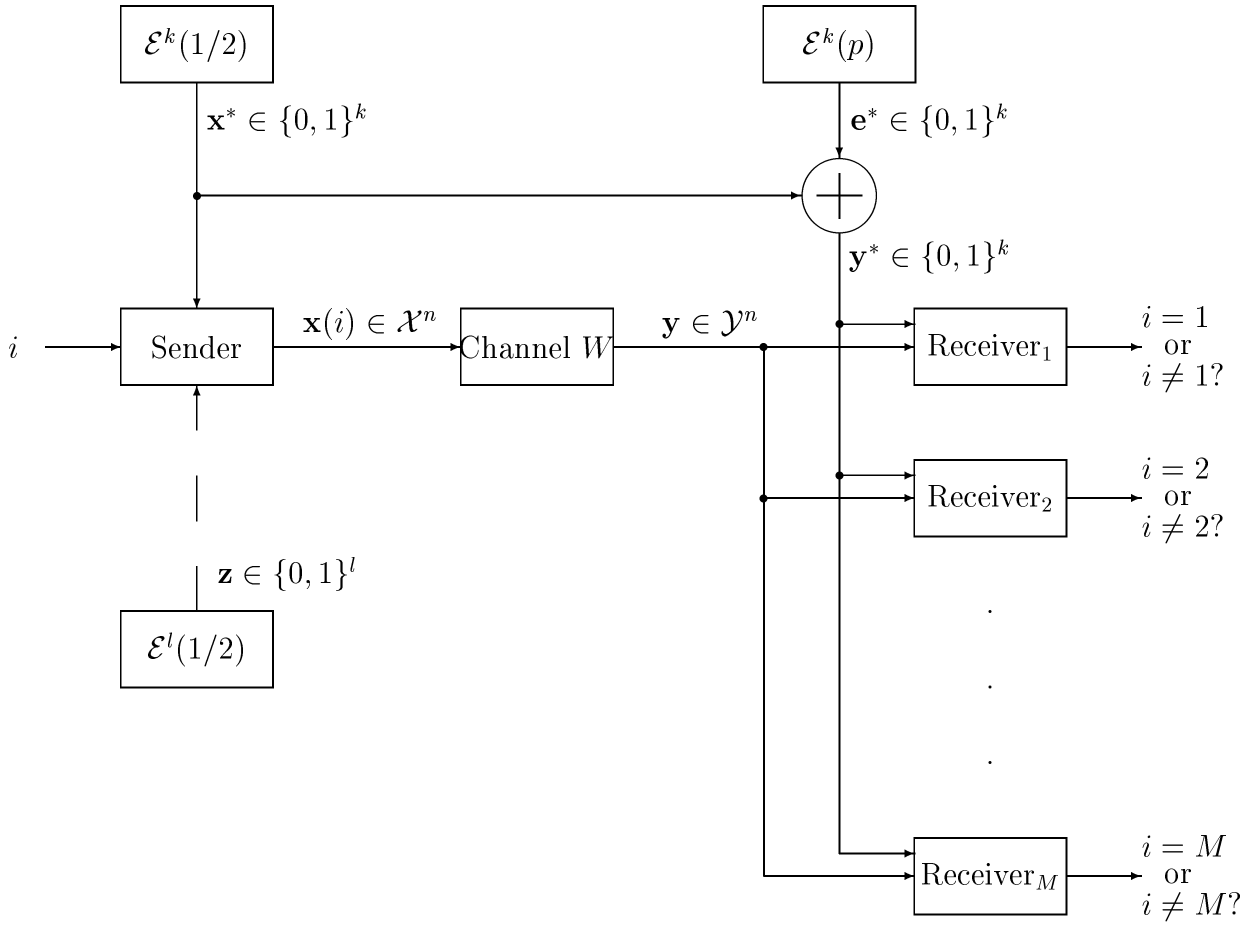

Figure 1: A model of identification over a channel W under a random process generated by a binary symmetric source $(\mathcal{E}^k(1/2), \mathcal{E}^k(p))$.

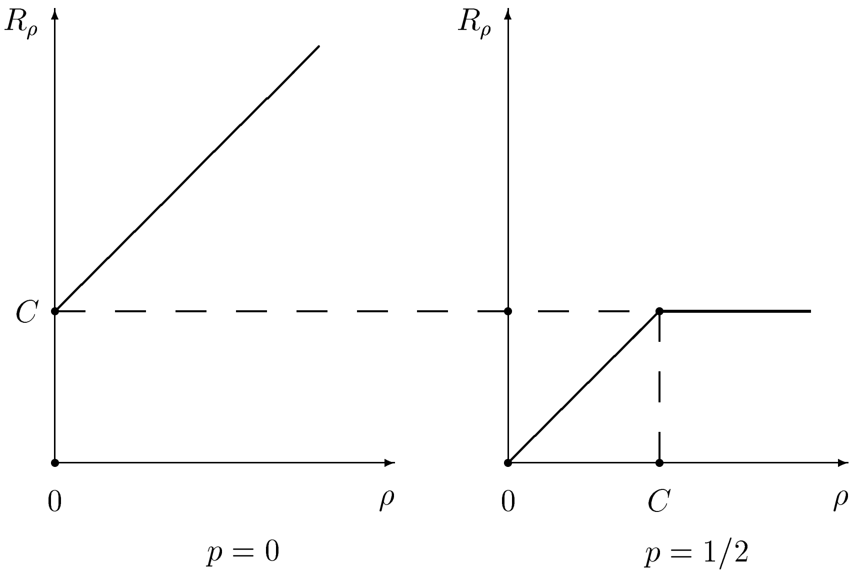

Figure 2: The ρ-achievable ID rate R_ρ as a function of ρ for $p = 0$ and $p = 1/2$.

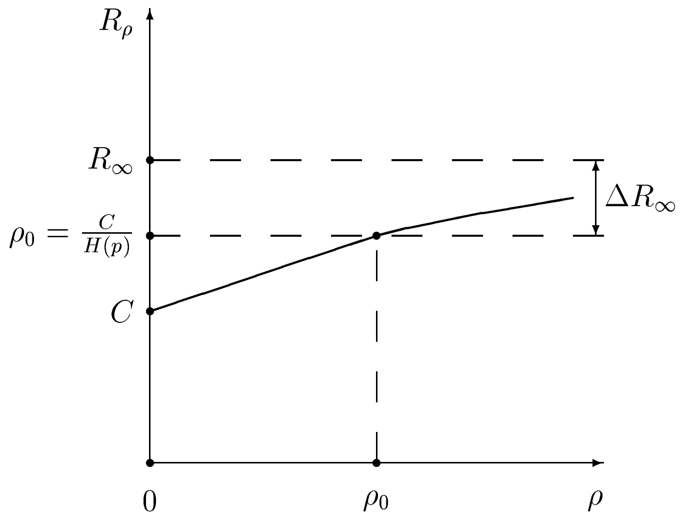

Figure 3: A lower bound on the ρ-achievable ID rate R_ρ as a function of ρ for $p \in (0, 1/2)$.

References

[1] R.Ahlswede and G.Dueck, "Identification via channels," *IEEE Trans.Inform.Theory,* vol.35, pp.15-29, Jan. 1989.

[2] R.Ahlswede and G.Dueck, "Identification in the presence of feedback - A discovery of new capacity formulas," *IEEE Trans.Inform.Theory,* vol.35, pp.30-36, Jan. 1989.

[3] T.S.Han and S.Verdu, "New results in the theory of identification via channels," *IEEE Trans.Inform.Theory,* vol.38, pp.14-25, Jan. 1992.

[4] T.S.Han and S.Verdu, "Approximation theory of output statistics," *IEEE Trans.Inform.Theory,* vol.39, pp.752-772, May 1993.

[5] R.Ahlswede and I.Csiczár, "Common randomness in information theory and cryptography - Part 1: Secret sharing," *IEEE Trans.Inform.Theory,* vol.39, pp.1121-1131, July 1993.

[6] R.Ahlswede and Z.Zhang, "New directions in the theory of identification via channels," Preprint 94 -010.

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