RANDOM CODING BOUND FOR CHANNELS WITH MEMORY — DECODING FUNCTION WITH PARTIAL OVERLAPPING

Part 2. Examples and Discussion

A. N. Trofimov[a], *PhD, Tech., Associate Professor, andrei.trofimov@vu.spb.ru*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: *Suboptimal random coding exponent $E_r^\star(R; \psi)$ for a wide class of finite-state channel models using a mismatched decoding function ψ was obtained and presented in the first part of this work. We used ψ function represented as a product of a posteriori probabilities of non-overlapped input subblocks of length $2B+1$ relative to the overlapped output subblocks of length $2W+1$. It has been shown that the computation of function $E_r^\star(R; \psi)$ is reduced to the calculation of the largest eigenvalue of a square non-negative matrix of an order depending on the B and W values.* **Purpose:** *To illustrate the approach developed in the first part of this study with its application to various channel modelled as a probabilistic finite-state machine.* **Results:** *We consider channels with state transitions not depending on the input symbol (channels with freely evolving states), and channels with deterministic state transitions, in particular, intersymbol interference channels. We present and discuss numerical results of calculating this random coding exponent in a full range of code rates for some of channel models for which similar results were not obtained before. Practical computations were carried out for relatively small values of B and W. Nevertheless, even for small values of these parameters a good correspondence with some known results for optimal decoding was shown.*
Keywords — *Random Coding Bound Exponent, Finite-State Channel Model, Mismatched Decoding, Perron — Frobenius Theorem, Intersymbol Interference Channel.*

Introduction

This paper is the second part of the work, the first part of which is published earlier [1]. In the previous part, a random coding bound is presented for a wide class of channels with memory, including those for which this bound could not be managed to obtain. The basic idea is to apply a *suboptimal* decoding rule that is different from the maximum likelihood (ML) decoding. In this part, we give numerical examples of the application of this boundary and their discussion.

For the connectedness of the exposition, we give the definitions and the main result of the previous part of the paper. Let $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ be transition probability of the discrete-time channel; for the continuous-output channel it is instead a probability density function (p. d. f.); $\mathbf{x} \in X^N$, where X be a discrete input channel alphabet and $q_x = |X| < \infty$; $\mathbf{y} \in Y^N$, where Y is the channel output alphabet, $|Y| = q$ and N is the length of a block code.

To indicate a segment of an arbitrary vector $\mathbf{z}$ we use the notation $\mathbf{z}_a^b = (z^{(\max(1,a))}, z^{(\max(1,a)+1)}, \ldots, z^{(\min(b,L))})$, where L is length of the vector $\mathbf{z}$.

For subvectors, or segments of vectors, $\mathbf{x}$ and $\mathbf{y}$ the notation $\mathbf{x}$ and $\mathbf{y}$ is used. The difference between them is noted due to the use of ordinary and *sans serif* font.

The decoding rule is given as

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \psi(\mathbf{y};\, \mathbf{x}),$$

where $\psi(\mathbf{y};\, \mathbf{x})$ is a real-valued positive decoding function, and the maximization is performed over all code words.

For simplicity we assume that code ensemble is generated by using of independent uniformly distributed (i. u. d.) code symbols. This assumption leads to *loss of optimality* but simplifies further consideration. Using the classic approach [2] one can derive the suboptimal exponent of the random coding bound in *asymptotic form*

$$E_r^*(R;\, \psi) = \max_{1 \geq \rho \geq 0}\left(\max_{\lambda > 0} E_0^*(\psi,\, \rho,\, \lambda) - \rho R \right),$$

where

$$E_0^*(\psi,\, \rho,\, \lambda) = (1+\rho)\log q_x -$$

$$- \lim_{N \to \infty} \frac{1}{N} \log \sum_{\mathbf{y}} \sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})\psi(\mathbf{y};\, \mathbf{x})^{-\lambda\rho} \left(\sum_{\mathbf{x}'} \psi(\mathbf{y};\, \mathbf{x}')^{\lambda} \right)^{\rho}.$$

Hereafter $\log(\cdot)$ denotes the binary logarithm, the asterisk in the superscript hereafter means that the code symbols are chosen as i. u. d. random variables.

Similarly one can get the random coding exponent for ML decoding for fixed code length N

$$E_r^*(N, R) = \max_{1 \geq \rho \geq 0} \left(E_0^*(N, \rho) - \rho R \right), \qquad (1)$$

where

$$E_0^*(N, \rho) = (1+\rho)\log q_x - \\ -\frac{1}{N}\log \sum_{\mathbf{y}} \left(\sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})^{\frac{1}{1+\rho}} \right)^{1+\rho} \qquad (2)$$

with the bound for maximum information rate

$$R_{\max}^*(N) = \partial E_0^*(N, \rho) / \partial \rho \big|_{\rho=0},$$

where $R_{\max}^*(N) \leq C$, where C is channel capacity. Evidently, the inequality $E_r^*(R; \psi) \leq E_r^*(R)$ is valid. The asymptotic random coding exponent for ML decoding and the code ensemble with i. u. d. symbols is $E_r^*(R) = \lim_{N \to \infty} E_r^*(N, R)$.

By analogy with the channel capacity C and maximal information rate $R_{\max}^*(N)$ let us define the lower bound on maximum achievable code rate for mismatched decoding $C^*(\psi)$ as

$$C^*(\psi) = \frac{\partial \max\limits_{\lambda>0} E_0^*(\psi, \rho, \lambda)}{\partial \rho}\Bigg|_{\rho=0} \leq C \qquad (3)$$

and value

$$R_0^*(\psi) = \max_{\lambda>0} E_0^*(\psi, 1, \lambda),$$

giving a bound on the cut-off rate R_0; evidently, the inequalities $R_0^*(\psi) \leq R_0^* \leq R_0$ are valid.

In this study we assume that the channel model is given as a *probabilistic finite-state machine* [2], i. e. conditional probabilities characterizing this model are given as follows

$$p_{\mathbf{y}|\mathbf{xs}}(\mathbf{y}\,|\,\mathbf{x}, \mathbf{s}) = \prod_{n=1}^N p_{y|xs}(y^{(n)}\,|\,x^{(n)}, s^{(n-1)});$$

$$p_{\mathbf{s}|\mathbf{x}}(\mathbf{s}\,|\,\mathbf{x}) = p_s(s^{(0)})\prod_{n=1}^N p_{s|xs}(s^{(n)}\,|\,x^{(n)}, s^{(n-1)}),$$

where $\mathbf{s} = (s^{(0)}, s^{(1)}, \ldots, s^{(n)}, \ldots)$ is the sequence of the channel states, $s^{(n)} \in S$, S is a set of the channel states, and $|S| < \infty$; $p_{y|xs}(y^{(n)}|x^{(n)}, s^{(n-1)})$ and $p_{y|xs}(y^{(n)}|x^{(n)}, s^{(n-1)})$ are conditional probabilities of the channel output and channel state transition, respectively, $p_s(\cdot)$ is an unconditional (stationary) distribution on the set of the channel states. Also, we assume that the input channel symbol $x^{(n)}$ and the current channel state $s^{(n-1)}$ are independent. It has been shown in [1] that the probabilities $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ can be represented in form of matrix product as

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x}) = \mathbf{p}_s\left(\prod_{n=1}^N \mathbf{P}(y^{(n)}\,|\,x^{(n)}) \right)\mathbf{1}^{\mathrm{T}}, \qquad (4)$$

where

$$\mathbf{P}(y\,|\,x) = [p_{y|xs}(y\,|\,x, s)\,p_{s|xs}(s'\,|\,x, s)] \qquad (5)$$

is a matrix of size $|S| \times |S|$; $\mathbf{p}_s = [p_s(1), \ldots, p_s(|S|)]$ is the vector of the unconditional state probabilities at $n = 0$, and $\mathbf{1} = (1, \ldots, 1)$ is vector of 1's of dimensions $1 \times |S|$.

Next, we specify type of decoding function. The appropriate choice of decoding function, which allows obtaining a result in the final form, was one of the main problems of this study. In this paper, we proposed a decoding function $\psi(\mathbf{y}; \mathbf{x})$ with partial overlap, which depends on two integer parameters W and B, $W \geq B \geq 0$. For i. u. d. segments $\mathbf{x}_{k(n)-B}^{k(n)+B}$ the decoding function $\psi(\mathbf{y}; \mathbf{x})$ can be written as

$$\psi(\mathbf{y}; \mathbf{x}) = \prod_{n=0}^{N(B)-1} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k(n)-W}^{k(n)+W}\,|\,\mathbf{x}_{k(n)-B}^{k(n)+B}), \qquad (6)$$

where $p_{\mathbf{y}|\mathbf{x}}(\cdot|\cdot)$ is the conditional probability for segments of *different*, in general, lengths $2W+1$ and $2B+1$ respectively, and $k(n) = n(2B+1)+1$.

Denote square matrices of order $|S|$ as

$$\mathbf{P}_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x}) = \prod_{l=1}^{2B+1} \mathbf{P}(y^{(l)}\,|\,x^{(l)}), \\ \mathbf{y} \in Y^{2B+1}, \ \mathbf{x} \in X^{2B+1}, \qquad (7)$$

where $\mathbf{P}(y|x)$ is matrix defined in (5). Let $D_1(\mathbf{y}; \lambda)$ be a scalar quantity, and $\mathbf{D}_2(\mathbf{y}; \lambda\rho)$ and $\mathbf{D}(\mathbf{y}; \lambda,\rho)$ be square matrices of order $|S|$ defined as follows

$$D_1(\mathbf{y}; \lambda) = \sum_{\mathbf{x} \in X^{2B+1}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})^{\lambda},$$

$$\mathbf{D}_2(\mathbf{y}; \lambda\rho) = \sum_{\mathbf{x} \in X^{2B+1}} \mathbf{P}_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{W-B+1}^{W+B+1}\,|\,\mathbf{x})\,p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})^{-\lambda\rho},$$

$$\mathbf{D}(\mathbf{y}; \lambda, \rho) = D_1(\mathbf{y}; \lambda)^{\rho}\mathbf{D}_2(\mathbf{y}; \lambda\rho), \ \mathbf{y} \in Y^{2W+1}.$$

Let us also define the square matrices $\mathbf{K}_{ij}(\lambda, \rho)$ of order $|S|$ as

$$\mathbf{K}_{ij}(\lambda, \rho) = \begin{cases} \mathbf{D}(\mathbf{y}; \lambda, \rho), & W \geq 2B+1; \\ \sum\limits_{\mathbf{y}_{2(W-B)+1}^{2B+1}} \mathbf{D}(\mathbf{y}; \lambda, \rho), & W < 2B+1. \end{cases} \qquad (8)$$

The correspondence of the indices i, j and the vector $\mathbf{y}$ in the expression (8) is given as $i \leftrightarrow \mathbf{y}_1^{2(W-B)}$

and $j \leftrightarrow \mathbf{y}_{2B+2}^{2W+1}$. Finally, we define a square block matrix of order $|S|q^{2(W-B)}$

$$\mathbf{K}(\lambda,\, \rho) = [\mathbf{K}_{ij}(\lambda,\, \rho)] =$$

$$= \begin{bmatrix} \mathbf{K}_{11}(\lambda,\, \rho) & \dots & \mathbf{K}_{1\,q^{2(W-B)}}(\lambda,\, \rho) \\ \dots & \dots & \dots \\ \mathbf{K}_{q^{2(W-B)}\,1}(\lambda,\, \rho) & \dots & \mathbf{K}_{q^{2(W-B)}\,q^{2(W-B)}}(\lambda,\, \rho) \end{bmatrix}. \quad (9)$$

The main result obtained in the first part of this paper [1] is formulated as the following assertion.

Theorem. Let channel be specified by conditional probabilities (4), where the matrices (5) are irreducible, and let the decoding function ψ be given by equation (6) with integer parameters W and B, where $W \geq B \geq 0$. Then the random coding exponent $E_r^*(R;\, \psi)$ for the code ensemble with i. u. d. code symbols is

$$E_r^*(R,\, \psi) = \max_{0 \leq \rho \leq 1}\left(E_0^*(\psi,\, \rho) - \rho R\right), \qquad (10)$$

where

$$E_0^*(\psi,\, \rho) = \max_{\lambda > 0} E_0^*(\psi,\, \rho,\, \lambda) =$$

$$= (1+\rho)\log q_x - (2B+1)^{-1}\log\left(\min_{\lambda>0} r\left(\mathbf{K}(\lambda,\, \rho)\right)\right),$$

and $r(\mathbf{K}(\lambda,\, \rho))$ is the maximum eigenvalue (spectral radius) of matrix $\mathbf{K}(\lambda,\, \rho)$, given in equation (9).

The computational complexity of obtaining the values of the function $E_r^*(R;\, \psi)$ depends on the dimensions of $p_{\mathbf{x}|\mathbf{y}}(\mathbf{y}|\mathbf{x})$ equal to $q^{2W+1} \times q_x^{2B+1}$ [see (7)], and on the order of the square matrix $\mathbf{K}(\lambda,\, \rho)$ equal to $|S|q^{2(W-B)}$.

Determination of analytical dependence of asymptotic random coding exponent $E_r^*(R;\, \psi)$ on the values of W and B is equivalent to the description of the dependency of spectral radius of the matrix $\mathbf{K}(\lambda,\, \rho)$ on these parameters. This dependency cannot be expressed in exact and closed analytical form. From general considerations, it follows that the greater the values of W and B are, then better approximation of the ML exponent can be achieved in principle. Moreover, for various values of code rate R different combinations of the values W and B may be preferable. Unfortunately, the increase of the parameters W and B causes the great growth of computational complexity. The common approach consists of testing some combinations and selecting one that gives acceptable results for a given coding rate at a reasonable computational complexity. In next section, we present some results of calculating the random coding exponent for several channel models with memory and comparison with some known results.

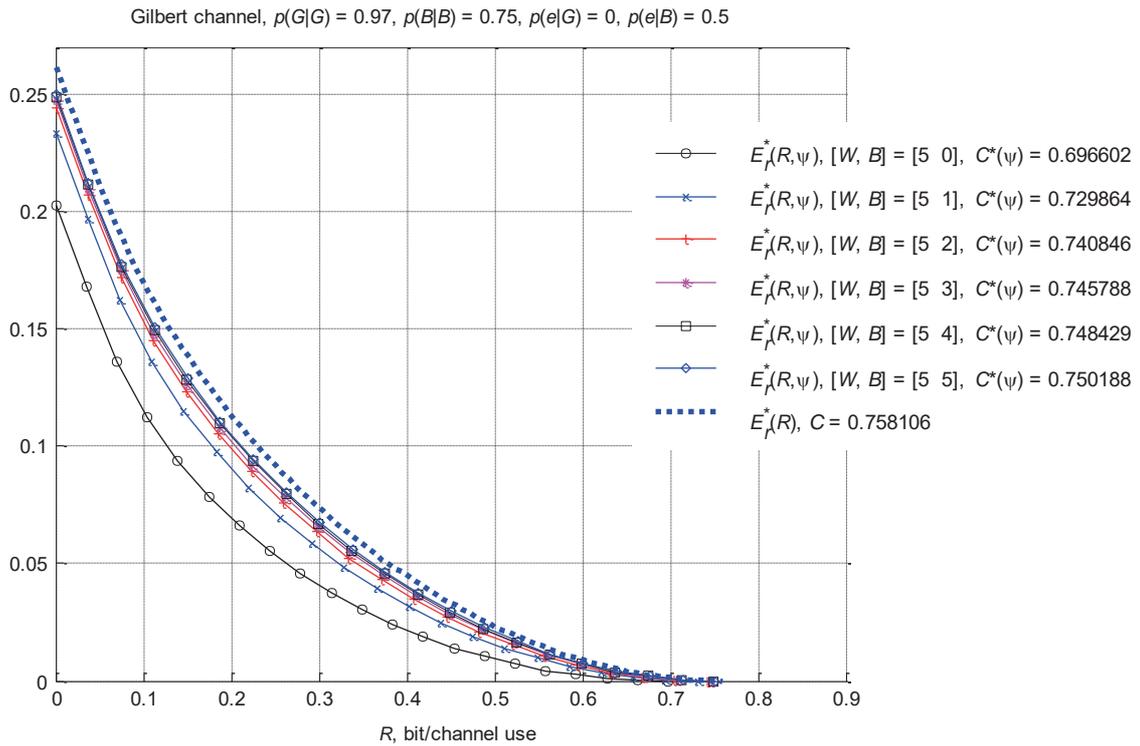## Numerical Examples and Discussion

To illustrate application of the suggested approach let us consider some examples. The first example is classical Gilbert model and its generalization — Gilbert — Elliott model. The second example presents a simple model for a fading channel with nonbinary Frequency Shift Keying (FSK). These models give examples of symmetric channels with binary and nonbinary inputs and freely evolving states. The third example is for the channel model that is defined as a deterministic finite-state machine, and the last example is for channel with linear intersymbol interference and q-level quantized output (deterministic finite-state machine model as well).

Example 1. *Gilbert channel and Gilbert — Elliott channel.* Consider the well-known Gilbert channel model with two states 1 ("good") and 2 ("bad"). In this case $q_x = q = 2$ and $|S| = 2$. Let the channel state transition probabilities be given as $p_{s|s}(1|1) = 0.75$ and $p_{s|s}(2|2) = 0.97$, and let the symbol crossover probabilities be equal to 0 and 1/2 for states 1 and 2 respectively (here we follow example in [3]). For this model, the complexity of calculation of the exponent of the random coding bound is not too large and the function $E_r^*(R;\, \psi)$ can be computed for comparatively large values of the parameters W and B. Results of the computations are shown in Fig. 4 for $W = 5$ and $B = 0, 1, ..., 5$. The values of $C^*(\psi)$, shown in Fig. 1 and further, are computed as

$$C^*(\psi) \approx \frac{\max\limits_{\lambda>0} E_0^*(\psi,\, \delta,\, \lambda)}{\delta},\ \delta << 1, \qquad (11)$$

giving an approximation for formula (3). The error exponent $E_r^*(R)$ for ML decoding computed for this example using the Egarmin algorithm [4] is also shown in Fig. 4. The channel capacity can be found using the original Gilbert approach [3] and for this example the capacity is $C = 0.758$ bit/channel use. It can be seen from the Fig. 1 that the curves $E_r^*(R;\, \psi)$ are approaching the curve $E_r^*(R)$ from below with the increase of the parameter B. For $B = 5$ the function $E_r^*(R;\, \psi)$ and value of $C^*(\psi)$ give good approximations for the ML random coding exponent $E_r^*(R)$ and the channel capacity C respectively .

It is interesting to compare the function $E_r^*(R;\, \psi)$ with the random coding exponent $E_r^*(N,\, R)$ for the ML decoding for some (small) values of the code length N. The function $E_r^*(N,\, R)$ can be computed by formulas (1) and (2). In equation (2) the channel conditional probability $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ is computed according to the equation (4). Evidently, to compute a single value of $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ we have to perform approximately $2N|S|^2$ operations (multiplications and additions). To compute the value $E_0^*(N,\, \rho)$ we need

Gilbert channel, $p(G|G) = 0.97$, $p(B|B) = 0.75$, $p(e|G) = 0$, $p(e|B) = 0.5$

Legend:
- $E_r^*(R, \psi)$, $[W, B] = [5\ 0]$, $C^*(\psi) = 0.696602$
- $E_r^*(R, \psi)$, $[W, B] = [5\ 1]$, $C^*(\psi) = 0.729864$
- $E_r^*(R, \psi)$, $[W, B] = [5\ 2]$, $C^*(\psi) = 0.740846$
- $E_r^*(R, \psi)$, $[W, B] = [5\ 3]$, $C^*(\psi) = 0.745788$
- $E_r^*(R, \psi)$, $[W, B] = [5\ 4]$, $C^*(\psi) = 0.748429$
- $E_r^*(R, \psi)$, $[W, B] = [5\ 5]$, $C^*(\psi) = 0.750188$
- $E_r^*(R)$, $C = 0.758106$

R, bit/channel use

■ *Fig. 1.* Functions $E_r^*(R; \psi)$ and $E_r^*(R)$, Gilbert channel

to calculate the values of $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ for all $\mathbf{y} \in Y^N$ and for all $\mathbf{x} \in X^N$. Thus the total number of operations in general case is about $2N|S|^2 q^N q_x^N$. For this example $|S| = 2$ and $q = q_x = 2$. Therefore, the total number of operations is about $N2^{2N+3}$ and it can be too large even for small N. For example, for $N = 16$ it is equal to $2^{39} \approx 5.5 \cdot 10^{11}$. But this channel can be considered as a channel with binary additive *modulo* 2 noise. Therefore, the channel output vector is $\mathbf{y} = \mathbf{x} \oplus \mathbf{e}$, where $\mathbf{e}$ is a binary error vector, and $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = p_{\mathbf{e}}(\mathbf{y} \oplus \mathbf{x})$, where $p_{\mathbf{e}}(\cdot)$ is a distribution on the set of error vectors. Then for the sum over $\mathbf{x}$ in the right hand side of (2) we can write

$$\sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}} = \sum_{\mathbf{x}} p_{\mathbf{e}}(\mathbf{y} \oplus \mathbf{x})^{\frac{1}{1+\rho}} = \sum_{\mathbf{x}} p_{\mathbf{e}}(\mathbf{x})^{\frac{1}{1+\rho}}$$

and hence

$$E_0^*(N, \rho) = \rho - \frac{1+\rho}{N} \log \left(\sum_{\mathbf{x}} p_{\mathbf{e}}(\mathbf{x})^{\frac{1}{1+\rho}} \right). \qquad (12)$$
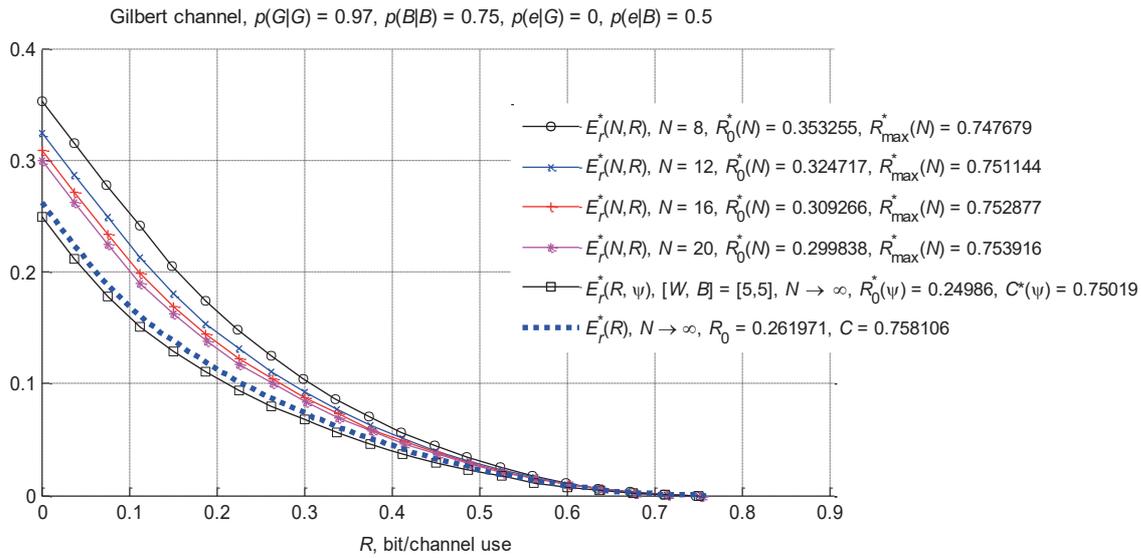
Evidently, $p_{\mathbf{e}}(\mathbf{x}) = p_{\mathbf{y}|\mathbf{x}}(\mathbf{0}|\mathbf{x})$, where the probabilities $p_{\mathbf{y}|\mathbf{x}}(\cdot|\cdot)$ are given by equation (4). In this case, to compute all values $p_{\mathbf{e}}(\mathbf{x})$ it is required to perform about $2N|S|^2 q_x^N = N2^{N+3}$ operations. For example, for $N = 16$ the number of operations is equal to an acceptable value $2^{23} \approx 8.4 \cdot 10^6$. Fig. 2 presents re-

sults of comparison of the function $E_r^*(R; \psi)$ with the random coding exponent $E_r^*(N, R)$ for ML decoding for some fixed values of the code length N.
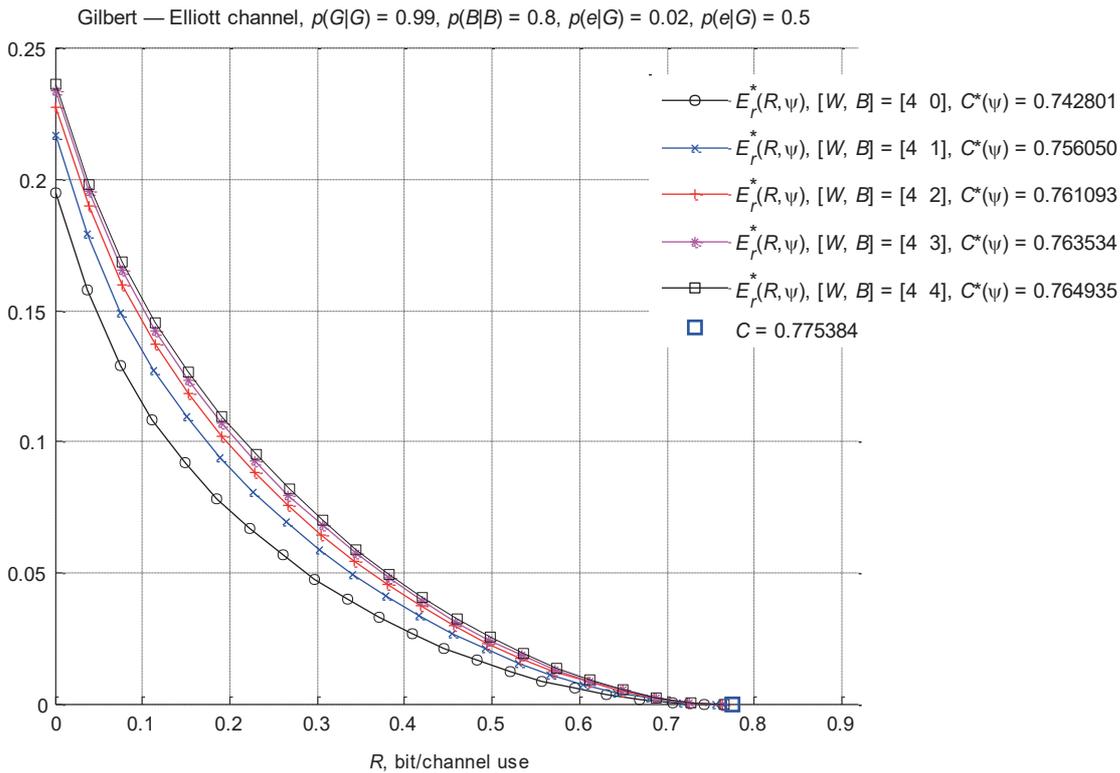
The asymptotic random coding exponent $E_r^*(R)$ for ML decoding computed by the Egarmin algorithm [4] is also shown in Fig. 2. It can be seen from the Fig. 2 that functions $E_r^*(N, R)$ approach the asymptotic function $E_r^*(R)$ from above with increasing of N, and the function $E_r^*(R; \psi)$ gives quite good approximation of the asymptotic function $E_r^*(R)$ from below.

Let us consider the next example — Gilbert — Elliot model [5]. Here again $q = q_x = 2$ and $|S| = 2$. Let the channel state transition probabilities be given as $p_{s|s}(1|1) = 0.99$ and $p_{s|s}(2|2) = 0.8$, and let the symbol crossover probabilities be equal to 0.02 and 1/2 for states 1 ("good") and 2 ("bad") respectively. For this model the ML random coding exponent $E_r^*(R)$ is unknown, but the channel capacity can be computed as it is shown in [6, 7]. The plots of function $E_r^*(R; \psi)$ are presented in Fig. 3.

It can be seen that the curves are being shifted upwards with the increase of the parameter B. For this case, we do not have a curve for the ML random coding exponent for comparison, but we can compute the capacity C for this channel using *statistical version* [7] of the algorithm presented in [6]. For this example the true capacity $C \approx 0.775$ bit/channel use. For $W = 4$, $B = 4$ the value of $C^*(\psi) = 0.765$ bit/channel use, that is very close to the true capacity.
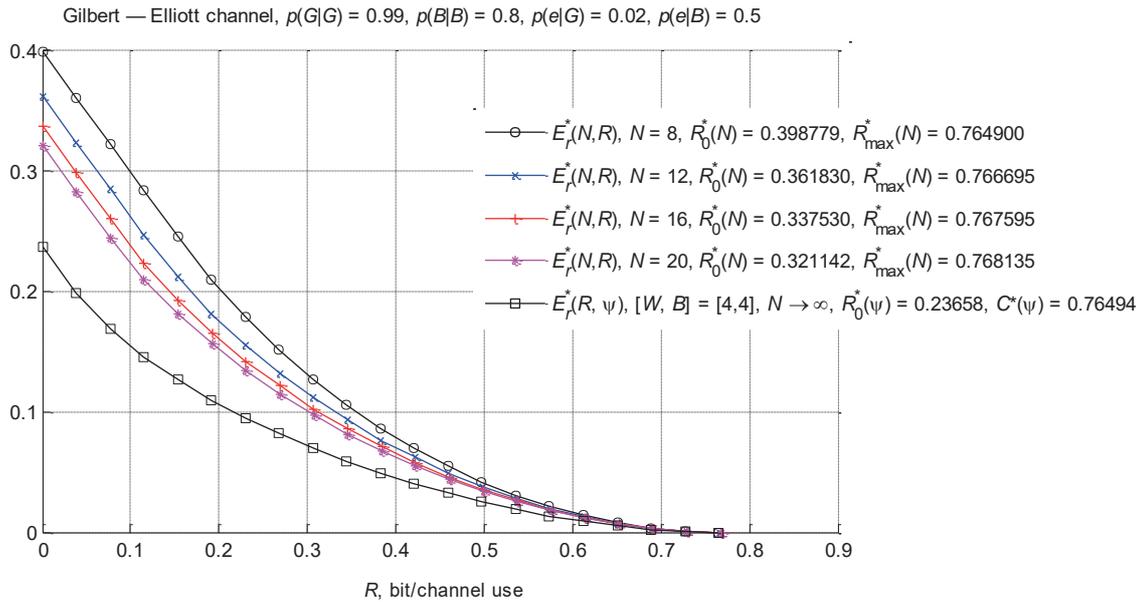
Gilbert channel, $p(G|G) = 0.97$, $p(B|B) = 0.75$, $p(e|G) = 0$, $p(e|B) = 0.5$

- $E_r^*(N,R)$, $N = 8$, $R_0^*(N) = 0.353255$, $R_{max}^*(N) = 0.747679$
- $E_r^*(N,R)$, $N = 12$, $R_0^*(N) = 0.324717$, $R_{max}^*(N) = 0.751144$
- $E_r^*(N,R)$, $N = 16$, $R_0^*(N) = 0.309266$, $R_{max}^*(N) = 0.752877$
- $E_r^*(N,R)$, $N = 20$, $R_0^*(N) = 0.299838$, $R_{max}^*(N) = 0.753916$
- $E_r^*(R, \psi)$, $[W, B] = [5,5]$, $N \to \infty$, $R_0^*(\psi) = 0.24986$, $C^*(\psi) = 0.75019$
- $E_r^*(R)$, $N \to \infty$, $R_0 = 0.261971$, $C = 0.758106$

■ *Fig. 2.* Functions $E_r^*(N, R)$, $E_r^*(R; \psi)$ and $E_r^*(R)$, Gilbert channel

Gilbert — Elliott channel, $p(G|G) = 0.99$, $p(B|B) = 0.8$, $p(e|G) = 0.02$, $p(e|G) = 0.5$

- $E_r^*(R, \psi)$, $[W, B] = [4\ \ 0]$, $C^*(\psi) = 0.742801$
- $E_r^*(R, \psi)$, $[W, B] = [4\ \ 1]$, $C^*(\psi) = 0.756050$
- $E_r^*(R, \psi)$, $[W, B] = [4\ \ 2]$, $C^*(\psi) = 0.761093$
- $E_r^*(R, \psi)$, $[W, B] = [4\ \ 3]$, $C^*(\psi) = 0.763534$
- $E_r^*(R, \psi)$, $[W, B] = [4\ \ 4]$, $C^*(\psi) = 0.764935$
- $C = 0.775384$

■ *Fig. 3.* Functions $E_r^*(R; \psi)$, Gilbert — Elliott channel

The plots of the functions $E_r^*(N, R)$ for ML decoding for some fixed N values calculated using equations (12), (4) and (1) are shown in Fig. 4 for comparison with the suboptimal asymptotic exponent $E_r^*(R; \psi)$. It can be seen that the curves $E_r^*(N, R)$ are shifting downward, and the achievable code rate $R_{max}^*(N)$ increases with increasing N. As $N \to \infty$, we have $R_{max}^*(N) \to C$ (note that in this example $C^* = C$), and the suboptimal random coding exponent $E_r^*(R; \psi)$ can serve as a lower bound for the random coding exponent for the Gilbert — Elliott channel.

Gilbert — Elliott channel, $p(G|G) = 0.99$, $p(B|B) = 0.8$, $p(e|G) = 0.02$, $p(e|B) = 0.5$

- ○ $E_r^*(N, R)$, $N = 8$, $R_0^*(N) = 0.398779$, $R_{max}^*(N) = 0.764900$
- × $E_r^*(N, R)$, $N = 12$, $R_0^*(N) = 0.361830$, $R_{max}^*(N) = 0.766695$
- ＋ $E_r^*(N, R)$, $N = 16$, $R_0^*(N) = 0.337530$, $R_{max}^*(N) = 0.767595$
- ✳ $E_r^*(N, R)$, $N = 20$, $R_0^*(N) = 0.321142$, $R_{max}^*(N) = 0.768135$
- □ $E_r^*(R, \psi)$, $[W, B] = [4,4]$, $N \to \infty$, $R_0^*(\psi) = 0.23658$, $C^*(\psi) = 0.76494$

R, bit/channel use

■ **Fig. 4.** Functions $E_r^*(N, R)$ and $E_r^*(R; \psi)$, Gilbert — Elliot channel
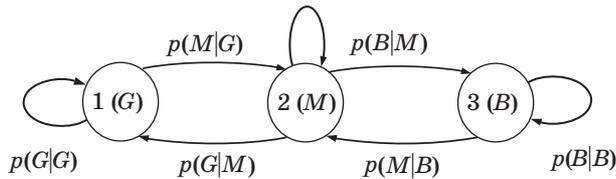
■ **Fig. 5.** State transitions diagram

Example 2. *Simple model for fading channel with nonbinary FSK.* Let us define the channel states as an additive white Gaussian noise channels with different noise power. Consider transmission of q_x-ary orthogonal FSK signals over this channel and optimal noncoherent reception. For this model $q_x = q$, and the symbol crossover probabilities are given as

$$p(y^{(n)} \mid x^{(n)}, s^{(n-1)}) = \begin{cases} 1 - \varepsilon(s^{n-1}), & y^{(n)} = x^{(n)}; \\ \dfrac{\varepsilon(s^{(n-1)})}{q-1}, & y^{(n)} \neq x^{(n)}, \end{cases}$$

where $\varepsilon(s)$ is symbol error probability for noncoherent reception of q-ary FSK signal for channel state s. This probability can be found as (see, e. g., [8])

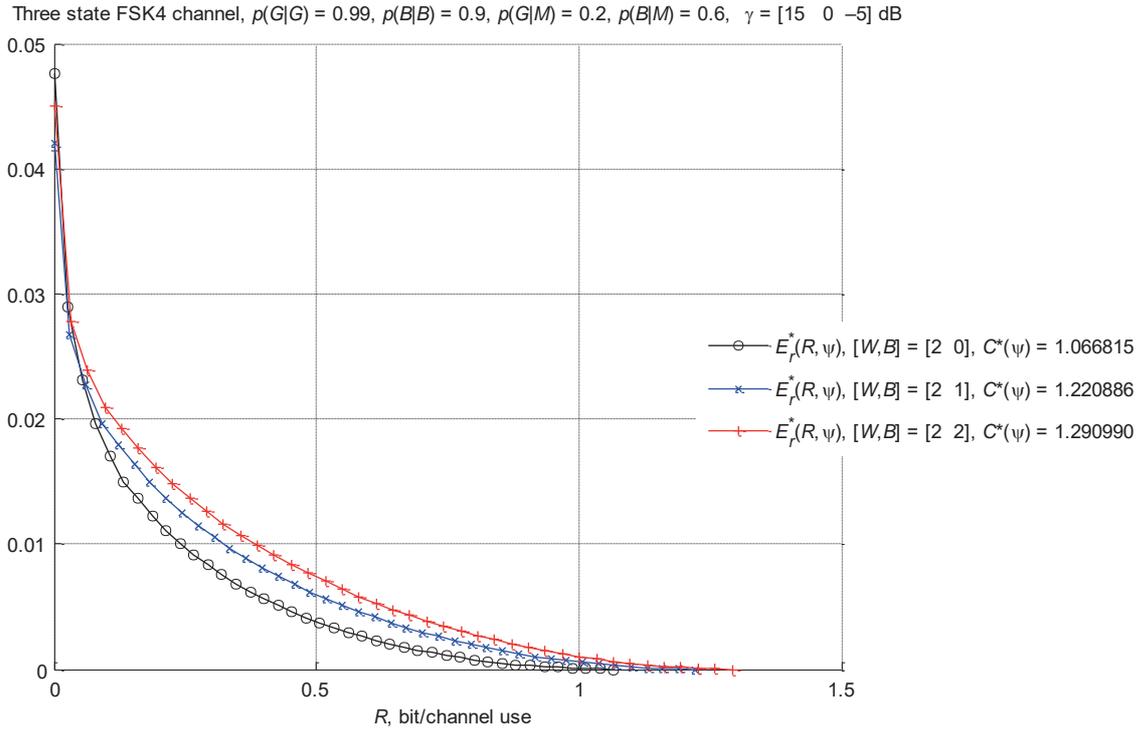$$\varepsilon(s) = \sum_{l=1}^{q-1} \binom{q-1}{l} \frac{(-1)^{l+1}}{l+1} \exp\left(-\frac{l}{l+1}\gamma(s)\right),$$

where $\gamma(s)$ is signal-to-noise ratio (SNR) in channel state s. Let for instance, $q_x = q = 4$ and $S = \{1, 2, 3\}$, i. e. the channel can be in one of three states: 1 ("good", or G), 2 ("medium", or M) and 3 ("bad", or B). Assume that the channel state transitions are given by the diagram shown in Fig. 5 with the following channel state transition probabilities: $p_{s|s}(1|1) = p(G|G) = 0.99$, $\quad p_{s|s}(1|2) = p(G|M) = 0.2$, $p_{s|s}(3|2) = p(B|M) = 0.6$, $p_{s|s}(3|3) = p(B|B) = 0.9$. Let the state SNR be as follows $\gamma(1) = 15$ dB, $\gamma(2) = 0$ dB and $\gamma(3) = -5$ dB. The plots of the function $E_r^*(R; \psi)$ are shown in Fig. 6.

Example 3. *Channel with deterministic state transitions.* Let $X = Y = \{0, 1\}$, $S = \{1, 2\}$, and the probabilities $p_{s|xs}(s^{(n)}|x^{(n)}, s^{(n-1)})$ and $p_{y|xs}(y^{(n)}|x^{(n)}, s^{(n-1)})$ are given in Table 1, where ε is symbol crossover probability, $0 < \varepsilon < 1/2$.

■ **Table 1.** Probabilities $p_{s|xs}(s^{(n)}|x^{(n)}, s^{(n-1)})$ and $p_{y|xs}(y^{(n)}|x^{(n)}, s^{(n-1)})$

$p_{s|xs}(s^{(n)}|x^{(n)}, s^{(n-1)})$					$p_{y|xs}(y^{(n)}|x^{(n)}, s^{(n-1)})$				
$s^{(n)}$	$(x^{(n)}, s^{(n-1)})$				$y^{(n)}$	$(x^{(n)}, s^{(n-1)})$			
	(0,1)	(0,2)	(1,1)	(1,2)		(0,1)	(0,2)	(1,1)	(1,2)
1	1	1	0	0	0	1/2	$1-\varepsilon$	ε	1/2
2	0	0	1	1	1	1/2	ε	$1-\varepsilon$	1/2

Three state FSK4 channel, $p(G|G) = 0.99$, $p(B|B) = 0.9$, $p(G|M) = 0.2$, $p(B|M) = 0.6$, $\gamma = [15 \quad 0 \quad -5]$ dB

■ Fig. 6. Functions $E_r^*(R; \psi)$, three state FSK4 channel

For this model in state 1, the symbol $x = 0$ flips with probability $1/2$, and the symbol $x = 1$ flips with small probability ε. In state 2, on the contrary, the symbol $x = 0$ flips with small probability ε, and the symbol $x = 1$ — with probability $1/2$. In addition, the channel state becomes 1 after symbol 0 comes in, and is equal to 2 after the coming in the symbol 1. The functions $E_r^*(R; \psi)$ for this model are plotted in Fig. 7 for $\varepsilon = 0.01$. In Fig. 7 we present examples for some good combinations of the parameters W and B for $W = 0, 1, ..., 6$, and the best pairs for this example are $[W, B] = [W, W - 1]$, $W > 1$.

To compare the plots of functions $E_r^*(R; \psi)$ with result for ML decoding let us consider the function $R_0^* - R$. This function coincides with the random coding exponent $E_r^*(R)$ in the interval $0 \leq R \leq R_{cr}$, where R_{cr} is the *critical rate* [2], and the plot of the linear function $R_0^* - R$ can be considered in example as a *known part* of the curve for whole random coding exponent $E_r^*(R)$. It can be shown (see Appendix) that R_0^* for ML decoding for this example can be found as $R_0^* = 2\log q_x - \log r(\mathbf{H})$, where $r(\mathbf{H})$ is maximum eigenvalue of the matrix $\mathbf{H}$, and

$$\mathbf{H} = \begin{bmatrix} 1 & a(\varepsilon) & a(\varepsilon) & 1 \\ a(\varepsilon) & 1 & b(\varepsilon) & a(\varepsilon) \\ a(\varepsilon) & b(\varepsilon) & 1 & a(\varepsilon) \\ 1 & a(\varepsilon) & a(\varepsilon) & 1 \end{bmatrix}, \quad (13)$$

where $a(\varepsilon) = \sqrt{(1-\varepsilon)/2} + \sqrt{\varepsilon/2}$ and $b(\varepsilon) = 2\sqrt{\varepsilon(1-\varepsilon)}$. It follows from the Fig. 7, that the random coding exponent $E_r^*(R; \psi)$ for $W = 6$, $B = 5$ is very close to the straight line $R_0^* - R$, so the function $E_r^*(R; \psi)$ can be considered as a quite good approximation for the ML random coding exponent $E_r^*(R)$.

Example 4. *Intersymbol interference channel with q-level quantized output.* A simple interference channel model is defined by the vector of coefficients $\mathbf{g} = [g_0, g_1, ..., g_L]$. The channel input $\mathbf{x} = (x^{(1)}, ..., x^{(N)})$ is a binary sequence and unquantized channel output is

$$y_0^{(n)} = \sum_{l=0}^{L} g_l (-1)^{x^{(n-l)}} + \xi^{(n)}, \quad (14)$$

where $x^{(\cdot)} = 0,1$, and $\xi^{(\cdot)}$ are independent Gaussian random variables with zero mean and variance σ^2. The following modulation mapping is assumed in equation (14): $0 \rightarrow +1$, $1 \rightarrow -1$. The SNR is defined as $\gamma = \|\mathbf{g}\|^2/(2\sigma^2)$. The number of channel states is equal to 2^L. The continuous channel output $y_0^{(n)}$ is subjected by a q-level quantization. In this example we assume that the quantization algorithm is one-dimensional quantization maximizing the Bhattacharya distance between conditional distributions of the quantized values [9]. We consider the simplest case of the model known as the *dicode channel* with parameters $L = 1$, $\mathbf{g} = [+1, -1]$ and $|S| = 2$. For this case $y_0^{(n)} = (-1)^{x^{(n)}} - (-1)^{x^{(n-1)}} + \xi^{(n)}$.
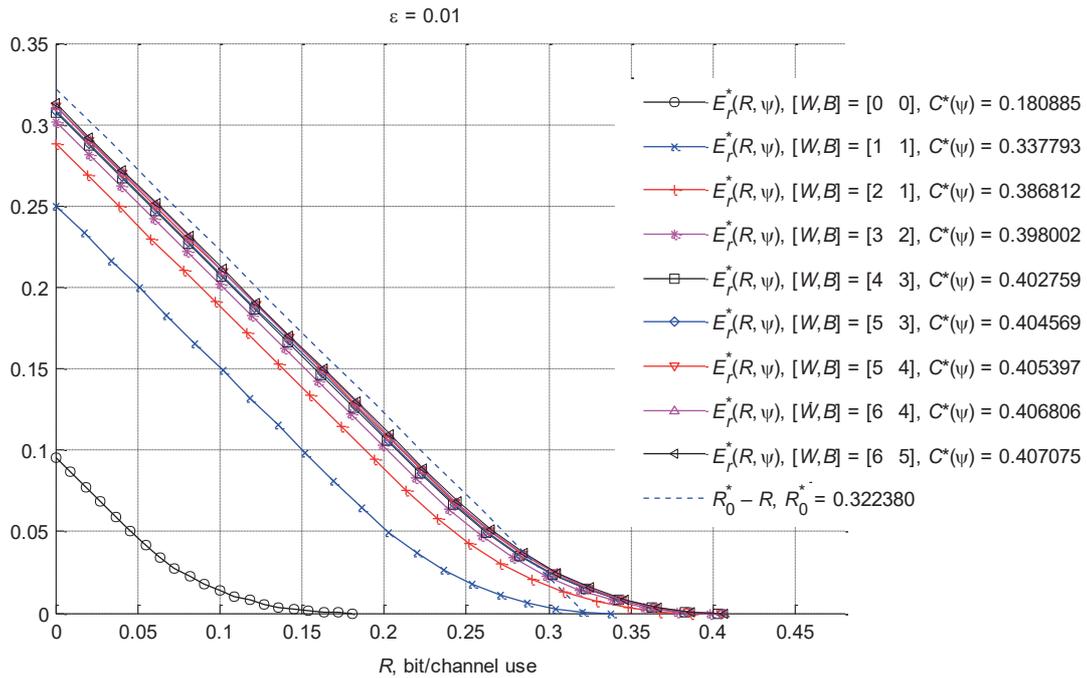
■ *Fig. 7.* Functions $E_r^*(R; \psi)$, channel with deterministic state transitions, $\varepsilon = 0.01$

The dicode channel with two-level quantization of the continuous channel output is equivalent to the channel with deterministic state transitions in Example 3. Plots of the functions $E_r^*(R; \psi)$ computed according to the equation (10) for $\gamma = 4$ dB, $q = 8$ and some combinations of the parameters B and W, are shown in Fig. 8. The values of $C^*(\psi)$, computed by the formula (11), and values of R_0^* for the ML decoding computed using known techniques [10–12] (see also Appendix), are also presented in Fig. 8.
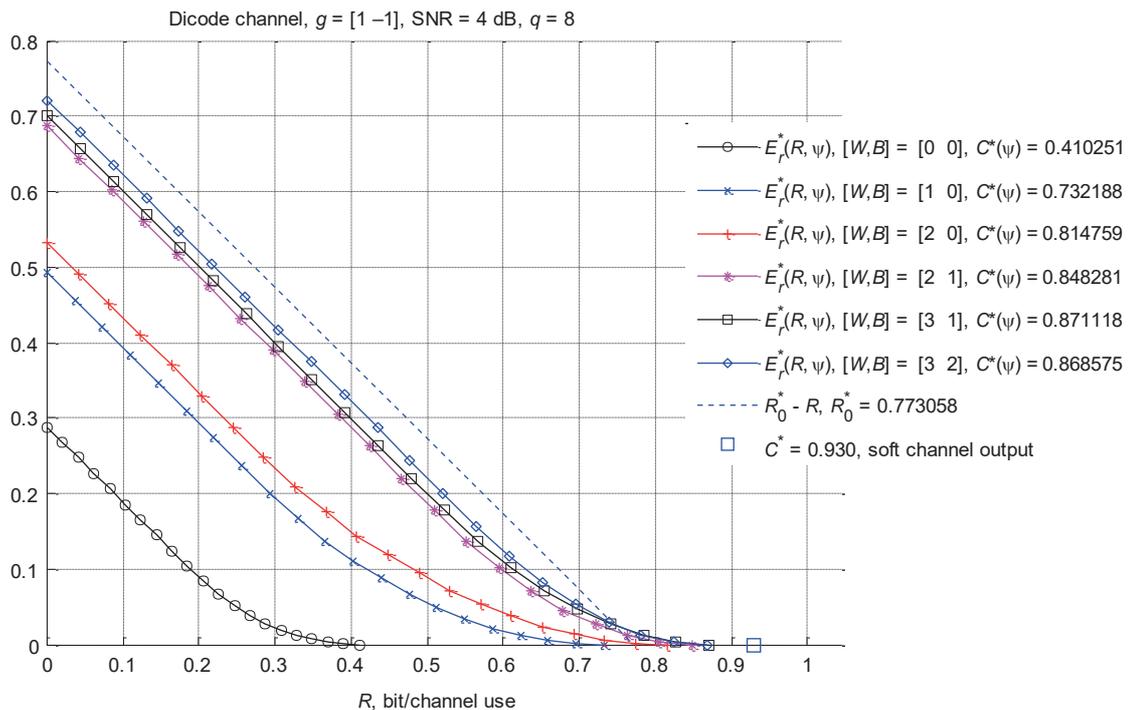
■ *Fig. 8.* Function $E_r^*(R; \psi)$, dicode channel, $\gamma = 4$ dB

As with the previous example we present the plot for the function $R_0^* - R$ in Fig. 8, where R_0^* is computed as it is shown in Appendix. Also in the Fig. 8 we indicate the value of maximum information rate C^* for the dicode channel with soft output computed in [13]. We see that maximum of functions $E_r^*(R; \psi)$ for $[W, B] = [3, 2]$ and $[W, B] = [3, 1]$ are close to the straight line $R_0^* - R$ and to the value of C^*.

Random coding exponent $E_r^*(N, R)$ for the ML decoding computed by formulas (1) and (2) for several small values of the code length N is depicted in Fig. 9. The function $E_r^*(R; \psi)$ and asymptotic ($N \to \infty$) linear function $R_0^* - R$ for the ML decoding are also presented in Fig. 9 for comparison. Clearly, as $N \to \infty$, the function $E_r^*(N, R)$ for $0 \le R \le R_{cr}$ tends to the line $R_0^* - R$. As it follows from the Fig. 9 the suboptimal random coding exponent $E_r^*(R; \psi)$ is very close to the random coding exponent $E_r^*(N, R)$ for the presented examples of the code length N and is not far from the asymptotic linear function $R_0^* - R$. The functions $E_r^*(N, R)$ for larger values of N are not presented due to high complexity of their computation.
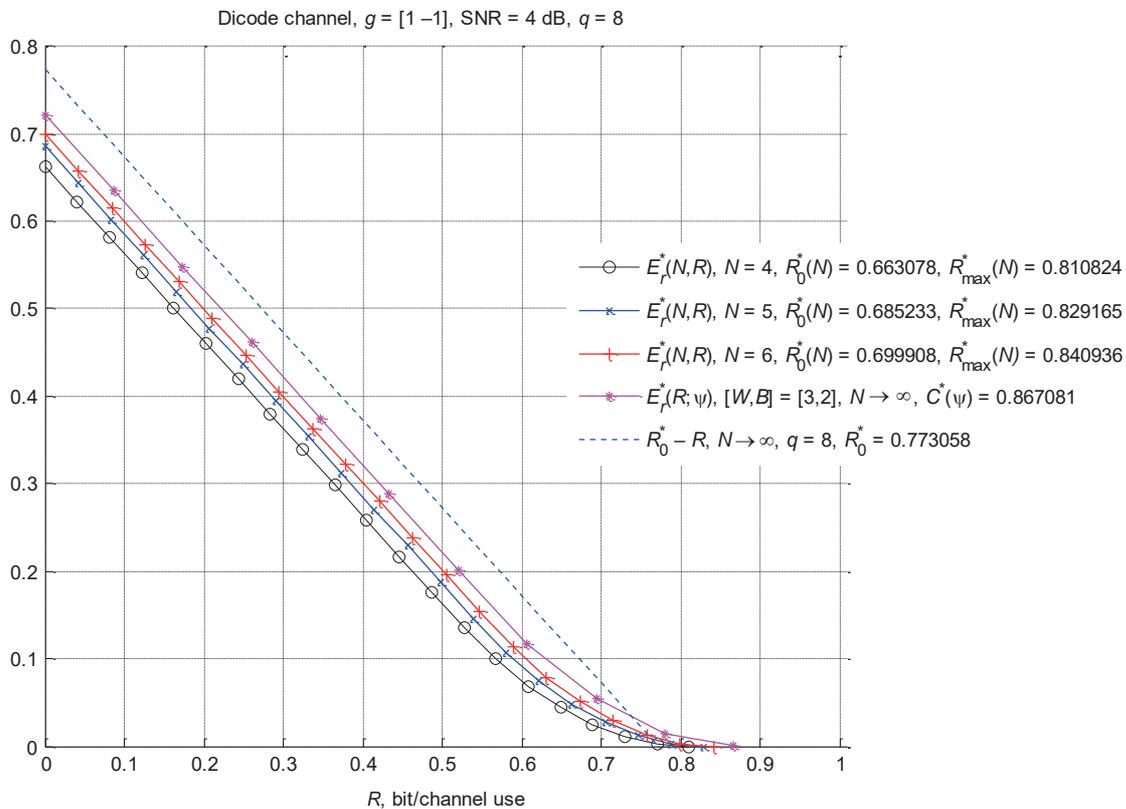
Plots of R_0^* and $R_0^*(\psi)$ as a function of the SNR are presented in Fig. 10. The values of R_0^* are calculated by a known method [14–16]. For this case (dicode channel) the values of R_0^* can be found in closed form as

$$R_0^* = -\log\left(\frac{3}{8} + \frac{1}{8}\left(e^{-2\gamma} + \sqrt{1 + 16e^{-\gamma} - 2e^{-2\gamma} + e^{-4\gamma}}\right)\right).$$
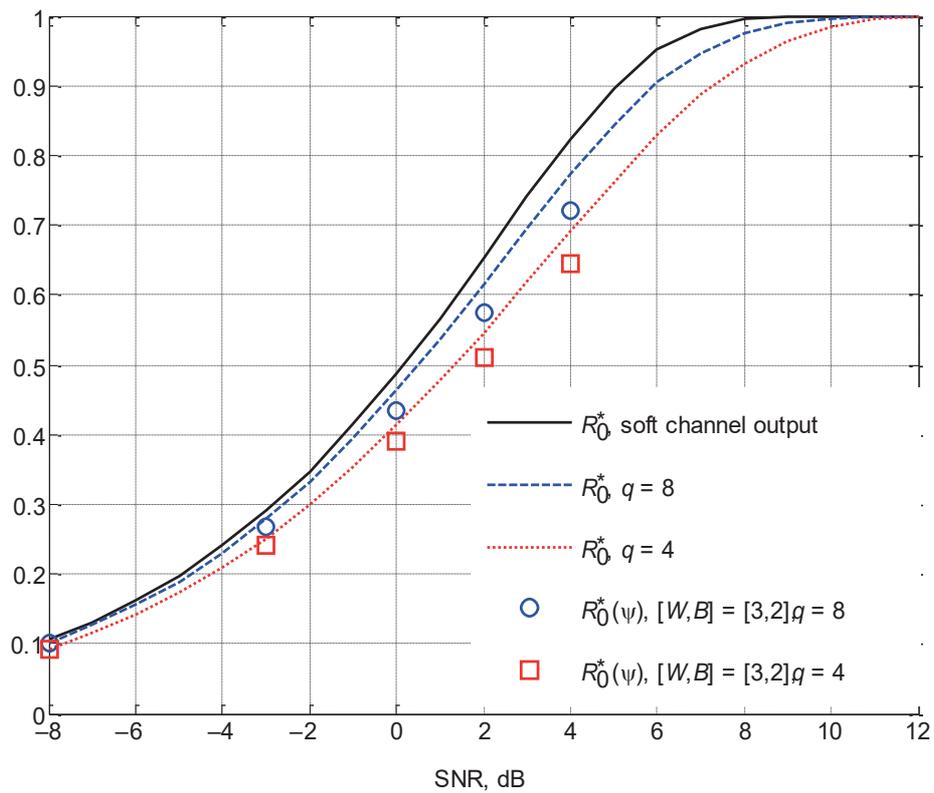
The values of $C^*(\psi)$ and values of *maximum information rate* for the channel with nonquantized output found in [13] by a simulation-based algorithm are shown in Fig. 11. Note that in [9] a different definition of SNR is used, namely $\gamma = \|g\|/\sigma^2$; therefore the plot showing the data of [9] is moved to the left by 3 dB. We see in Fig. 10 and Fig. 11 that the difference is about 10 %, and it seems that the increasing number of quantization levels q leads to a decrease in this difference.

Conclusion

In this work, consisted of two parts, we present the derivation of the exponent of the random coding bound $E_r^*(R; \psi)$ for suboptimal, or mismatched, decoding. We proposed a decoding function in the form of a product of the *a posteriori* probabilities of the non-overlapped input subblocks of length $2B + 1$ relative to the overlapped output subblocks of length $2W + 1$ [see (6)]. The computation of the values of the function $E_r^*(R; \psi)$ is reduced to the calculation of the largest eigenvalue of a square nonnegative matrix $\mathbf{K}(\lambda, \rho)$ of order $|S|q^{2(W-B)}$, where $|S|$
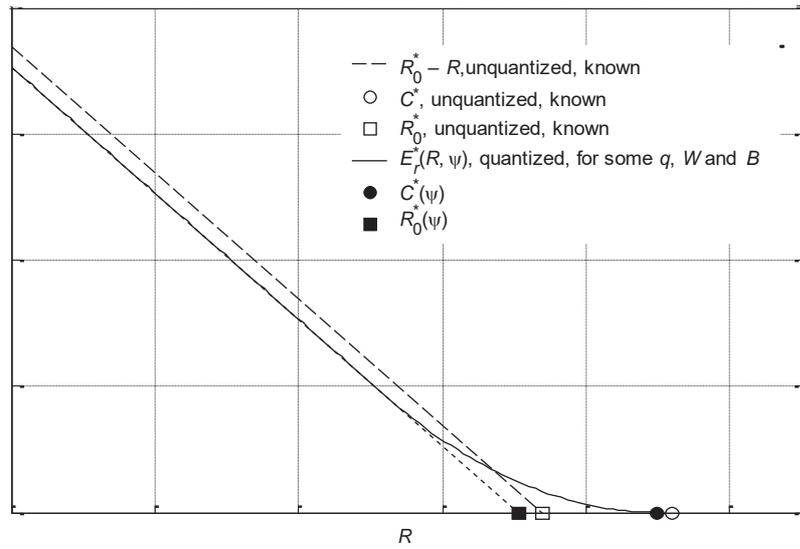

Dicode channel, $g = [1\ {-}1]$, SNR = 4 dB, $q = 8$

- $E_r^*(N,R)$, $N = 4$, $R_0^*(N) = 0.663078$, $R_{max}^*(N) = 0.810824$
- $E_r^*(N,R)$, $N = 5$, $R_0^*(N) = 0.685233$, $R_{max}^*(N) = 0.829165$
- $E_r^*(N,R)$, $N = 6$, $R_0^*(N) = 0.699908$, $R_{max}^*(N) = 0.840936$
- $E_r^*(R; \psi)$, $[W,B] = [3,2]$, $N \to \infty$, $C^*(\psi) = 0.867081$
- $R_0^* - R$, $N \to \infty$, $q = 8$, $R_0^* = 0.773058$

R, bit/channel use

■ *Fig. 9.* Functions $E_r^*(N, R)$ and $E_r^*(R; \psi)$, dicode channel $\gamma = 4$ dB

■ *Fig. 10.* Plots of R_0^* и $R_0^*(\psi)$ for dicode channel
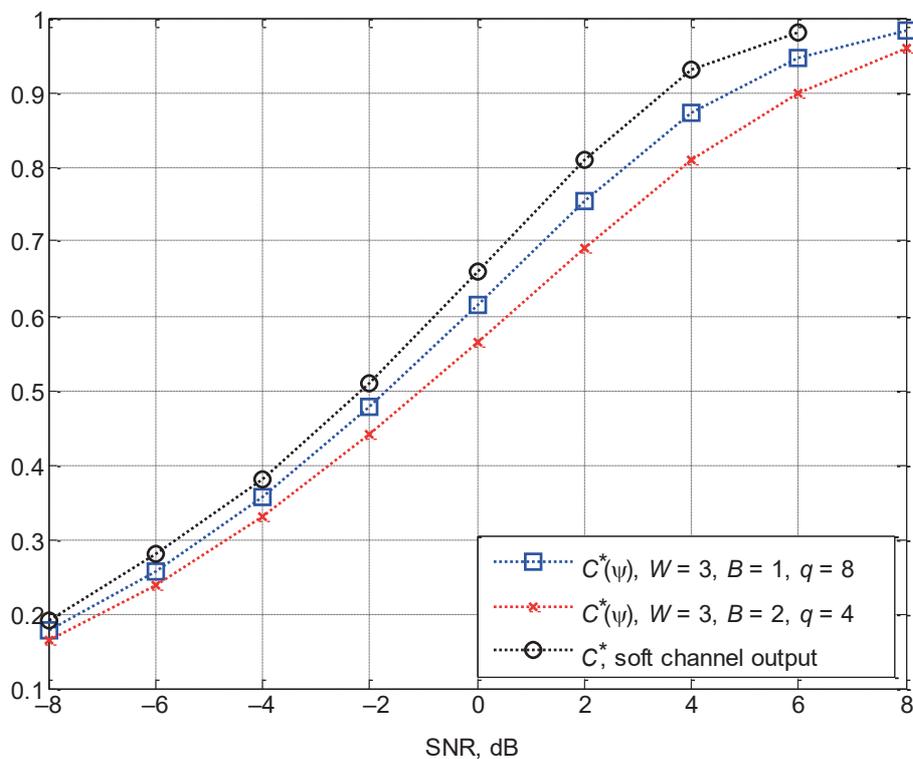
■ *Fig. 11.* Plots C^* and $C^*(\psi)$ for dicode channel

■ *Fig. 12*. General view of the functions $\tilde{E}_r^*(R)$ and $E_r^*(R; \psi)$

is the number of channel states and q is the cardinality of the channel output alphabet.

The computational complexity of obtaining the values of the function $E_r^*(R; \psi)$ depends on the dimension of the values $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$, equal to $q^{2W+1} \times q_x^{2B+1}$ [see (7)], and on the order of the matrix $\mathbf{K}(\lambda, \rho)$, equal to $|S|q^{2(W-B)}$. Therefore, in the examples presented in this part of the work, the practical computations were carried out for relatively small values of B, W and q. Nevertheless even for small values of these parameters, good results were obtained. The values of $R_0^*(\psi)$ and $C^*(\psi)$ found for the suggested suboptimal decoding functions are close to the corresponding values found before for the case of ML decoding for the intersymbol interference channels with soft output. A qualitative picture of the relationship between function $\tilde{E}_r^*(R)$, presented in the introduction in the first part [1] of this work, and function $E_r^*(R; \psi)$ is shown in Fig. 12. The values of R_0^* and C^* for a discrete-time channel with intersymbol interference can be obtained by known techniques [14–16] and [9]. As we see from the Fig. 12 the curve for the function $E_r^*(R; \psi)$ (solid line) goes higher than known bound $\tilde{E}_r^*(R)$ for high code rates.

We see the same in example in Fig. 8, where for quantized channel output with $q = 8$ and $W = 3$, $B = 1$, we have $R_0^*(\psi) = 0.710$ bit/channel use, $C^*(\psi) = 0.871$ bit/channel use; for continuous channel output $R_0^*(\psi) = 0.824$ bit/channel use and $C^*(\psi) = 0.920$ bit/channel use as follows from Fig. 10 and Fig. 11 for SNR = 4 dB. The curve $E_r^*(R; \psi)$ is shifted upwards and to the right with increasing values of the parameters q, W and B. This conclusion follows from the fact that the decoding function ψ becomes the ML decoding function with increasing parameters W and B. Thus, the function $E_r^*(R; \psi)$ can be a *good approximation* for the true

but unknown function $E_r^*(R)$. The problem of the extension of the proposed approach to other channel models and the problem of finding an efficient algorithm for numerical computation for large values of B, W and q remain open for research.

Appendix

In this appendix we present a derivation of the expression for R_0^* for the channel with deterministic state transitions. This derivation is a minor modification of the known results [12, 14–16]. The general expression for R_0^* follows from the equation or $\rho = 1$ and $N \to \infty$ is

$$R_0^* = 2\log q_x - \lim_{N \to \infty} \frac{1}{N} \sum_{\mathbf{y}} \left(\sum_{\mathbf{x}} \sqrt{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})} \right)^2. \quad \text{(A1)}$$

Let us consider the sum over $\mathbf{y}$ in (A1)

$$\sum_{\mathbf{y}} \left(\sum_{\mathbf{x}} \sqrt{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})} \right)^2 =$$

$$= \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \left(\sum_{\mathbf{y}} \sqrt{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}')} \right). \quad \text{(A2)}$$

For finite-state channel model we have the channel conditional probability

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = \sum_{\mathbf{s}} p_{\mathbf{y}|\mathbf{xs}}(\mathbf{y}|\mathbf{x}, \mathbf{s}) p_{\mathbf{s}|\mathbf{x}}(\mathbf{s}|\mathbf{x}),$$

where

$$p_{\mathbf{y}|\mathbf{xs}}(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{n=1}^{N} p_{y|xs}(y^{(n)}|x^{(n)}, s^{(n-1)})$$

■ **Table A1.** Values of $h(s_a, s'_a; s_b, s'_b)$

(s_a, s'_a)	(s_b, s'_b)			
	(1,1)	(1,2)	(2,1)	(2,1)
(1,1)	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,1)p_{y|xs}(y\,|\,0,1)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,1)p_{y|xs}(y\,|\,1,1)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,1)p_{y|xs}(y\,|\,0,1)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,1)p_{y|xs}(y\,|\,1,1)}$
(1,2)	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,1)p_{y|xs}(y\,|\,0,2)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,1)p_{y|xs}(y\,|\,1,2)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,1)p_{y|xs}(y\,|\,0,2)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,1)p_{y|xs}(y\,|\,1,2)}$
(2,1)	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,2)p_{y|xs}(y\,|\,0,1)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,2)p_{y|xs}(y\,|\,1,1)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,2)p_{y|xs}(y\,|\,0,1)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,2)p_{y|xs}(y\,|\,1,1)}$
(2,2)	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,2)p_{y|xs}(y\,|\,0,2)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,0,2)p_{y|xs}(y\,|\,1,2)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,2)p_{y|xs}(y\,|\,0,2)}$	$\sum\limits_y \sqrt{p_{y|xs}(y\,|\,1,2)p_{y|xs}(y\,|\,1,2)}$

and

$$p_{\mathbf{s}|\mathbf{x}}(\mathbf{s}\,|\,\mathbf{x}) = p_s(s^{(0)})\prod_{n=1}^{N} p_{s|xs}(s^{(n)}\,|\,x^{(n)}, s^{(n-1)}).$$

Hence,

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x}) = \sum_{s^{(0)}} p_s(s^{(0)})\prod_{n=1}^{N}\sum_{s^{(n)}} p_{y|xs}(y^{(n)}\,|\,x^{(n)},$$
$$s^{(n-1)})p_{s|xs}(s^{(n)}\,|\,x^{(n)}, s^{(n-1)}). \qquad \text{(A3)}$$

For channel with deterministic state transitions the pair $(x^{(n)}, s^{(n-1)})$ uniquely defines the next channel state $s^{(n)}$, therefore the sum over $s^{(n)}$ in (A3) contains only one term. Then it can be written

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x}) = \sum_{s^{(0)}} p_s(s^{(0)})p_{y|xs}(y^{(1)}\,|\,x^{(1)},$$
$$s^{(0)})p_{y|xs}(y^{(2)}\,|\,x^{(2)}, s^{(2)})...p_{y|xs}(y^{(N)}\,|\,x^{(N)}, s^{(N-1)}),$$

where $s^{(n)} = f(x^{(n)}, s^{(n-1)})$, $n = 1, 2, ..., N$, and $f(\cdot;\cdot)$ is a function defining deterministic transition $s^{(n-1)} \overset{x^{(n)}}{\to} s^{(n)}$. Hence, for the expression (A2) we have

$$\sum_{\mathbf{x}}\sum_{\mathbf{x}'}\left(\sum_{\mathbf{y}}\sqrt{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x}')}\right) \le$$
$$\le \max_{s^{(0)}, s'^{(0)}}\sum_{\mathbf{x}}\sum_{\mathbf{x}'}\prod_{n=1}^{N}\sum_{y}\sqrt{\begin{array}{l} p_{y|xs}(y\,|\,x^{(n)}, s^{(n-1)})\times \\ \times p_{y|xs}(y\,|\,x'^{(n)}, s'^{(n-1)}) \end{array}}. \qquad \text{(A4)}$$

Let us introduce for two pairs of states (s_a, s'_a) and (s_b, s'_b) the values

$$h(s_a, s'_a; s_b, s'_b) =$$
$$= \begin{cases} \sum\limits_y \sqrt{p_{y|xs}(y\,|\,x, s_a)p_{y|xs}(y\,|\,x', s_b)}, \\ \text{if } s_b = f(x, s_a) \text{ and } s'_b = f(x', s'_a); \\ 0, \text{ otherwise}, \end{cases} \qquad \text{(A5)}$$

and build matrix $\mathbf{H}$ of size $|S|^2 \times |S|^2$ with the entries $h(s_a, s'_a; s_b, s'_b)$, where the pairs (s_a, s'_a) и (s_b, s'_b) represent first and second indices of the matrix entry respectively. Then the inequality (A4) can be written as

$$\sum_{\mathbf{x}}\sum_{\mathbf{x}'}\left(\sum_{\mathbf{y}}\sqrt{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x}')}\right) \le$$
$$\le \max(\mathbf{H}^N\mathbf{1}^{\mathrm{T}}) \le \mathbf{1}\mathbf{H}^N\mathbf{1}^{\mathrm{T}},$$

where $\mathbf{1} = (1, ..., 1)$ is vector of the dimension $|S|^2$. Further, we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\mathbf{y}}\left(\sum_{\mathbf{x}}\sqrt{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\,|\,\mathbf{x})}\right)^2 = \lim_{N\to\infty}\frac{1}{N}\mathbf{1}\mathbf{H}^N\mathbf{1}^{\mathrm{T}},$$

and using the corollary from the Perron — Frobenius theorem [2, 15] we get finally from (A1) $R_0^* = 2\log q_x - \log r(\mathbf{H})$, where $r(\mathbf{H})$ is maximum eigenvalue of the matrix $\mathbf{H}$.

For the channel model given in the Examples 3 and 4 the values of $h(s_a, s'_a; s_b, s'_b)$ defined in (A5) are listed in Table A1.

Expression (13) is obtained using the data given in Table A1.

References

1. **Trofimov A. N.** Random Coding Bound for Channels with Memory — Decoding Function with Partial Overlapping. Part 1. Derivation of Main Expression. *Informatsionno-upravlyayushhie sistemy* [Information and Control Systems], 2018, no. 3, pp. 79–88. doi:10.15217/issn1684-8853.2018.3.79

2. **Gallager R.** *Information Theory and Reliable Communication.* New York, John Wiley & Sons, 1968. 588 p.

3. **Gilbert E. N.** Capacity of a Burst-Noise Channel. *Bell System Technical Journal*, 1960, vol. 39, no. 5, pp. 1253–1265.

4. **Egarmin V. K.** Lower and Upper Bounds on Decoding Error Probability for Discrete Channels. *Problemy peredachi informatsii* [Problems of Information Transmission], 1969, vol. 5, no. 1, pp. 23–39 (In Russian).

5. **Elliott E. O.** Estimates of Error Rates for Codes on Burst-Noise Channels. *Bell System Technical Journal*, 1963, vol. 42, no. 5, pp. 1977–1997.

6. **Mushkin M., Bar-David I.** Capacity and Coding for the Gilbert — Elliott Channels. *IEEE Transactions on Information Theory*, 1989, vol. 35, no. 6, pp. 1277–1290.

7. **Rezaeian M.** Computation of Capacity for Gilbert — Elliott Channels, using a Statistical Method. *6th Australian Communications Theory Workshop 2005,* Feb. 2–4, 2005, Brisbane, Australia, pp. 56–61.

8. **Proakis J. G.** *Digital Communications.* McGraw-Hill, New York, 1995. 928 p.

9. **Lin S., Costello D.** *Error Control Coding — Fundamentals and Applications.* Prentice Hall, N. J., 2004. 624 p.

10. **Raghavan S., Kaplan G.** Optimum Soft Decision Demodulation for ISI Channels. *IEEE Transactions on Communications*, 1993, vol. 41, no. 1, pp. 83–89.

11. **Shamai (Shitz) S., Raghavan S.** On the Generalized Symmetric Cutoff Rate for Finite-State Channels. *IEEE Transactions on Information Theory*, 1995, vol. 41, no. 9, pp. 1333–1346.

12. **Trofimov A., Chan Keong Sann.** Complexity-Performance Trade-off for Intersymbol Interference Channels – Random Coding Analysis. *IEEE Transactions on Magnetics*, 2010, vol. 46, no. 4, pp. 1077–1091.

13. **Arnold D. M., Loeliger H.-A., Vontobel P. O., Kavcic A., Wei Zeng.** Simulation-Based Computation of Information Rates for Channels With Memory. *IEEE Transactions on Information Theory*, 2006, vol. 52, no. 8, pp. 3498–3508.

14. **Biglieri E.** The Computational Cutoff Rate of Channel Having Memory. *IEEE Transactions on Information Theory*, 1981, vol. 27, pp. 352–357.

15. **Gantmacher F. R.** *Teoriia matrits* [Theory of Matrices]. Moscow, Nauka Publ., 1988. 552 p. (In Russian).

Граница случайного кодирования для каналов с памятью — декодирующая функция с частичным перекрытием. Часть 2. Примеры и обсуждение

Трофимов А. Н.[а], канд. техн. наук, доцент, andrei.trofimov@vu.spb.ru
[а]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: субоптимальная экспонента случайного кодирования $E_r^*(R; \psi)$ для широкого класса моделей канала с конечным числом состояний, использующая несогласованную декодирующую функцию ψ, получена и описана в первой части этой работы. Мы использовали функцию ψ, представленную в виде произведения апостериорных вероятностей неперекрывающихся входных подблоков длины $2B + 1$ относительно перекрывающихся выходных подблоков длины $2W + 1$. Показано, что вычисление значений функции $E_r^*(R; \psi)$ сводится к вычислению наибольшего собственного значения квадратной неотрицательной матрицы, порядок которой зависит от параметров канала и от величин W и B. **Цель:** проиллюстрировать развитый в первой части исследования подход к вычислению экспоненты случайного кодирования в приложении его к различным каналам, модели которых представляют собой вероятностный конечный автомат. **Результаты:** рассмотрены каналы, в которых переходы в множестве состояний не зависят от входного символа, и каналы с детерминированными переходами, в частности каналы с межсимвольной интерференцией. Получены численные результаты вычисления экспоненты случайного кодирования в полном интервале скоростей кода для ряда моделей каналов, для которых подобные результаты не были ранее получены. Практические вычисления выполнены для относительно малых значений B и W. Тем не менее даже при малых значениях этих параметров получено хорошее соответствие с известными результатами для оптимального декодирования.

Ключевые слова — граница случайного кодирования, канал с конечным числом состояний, несогласованное декодирование, теорема Перрона — Фробениуса, канал с межсимвольной интерференцией.