Acta Polytechnica **60**(1):65–72, 2020
DOI:10.14311/AP.2020.60.0065

available online at https://ojs.cvut.cz/ojs/index.php/ap

ENERGY GAIN FROM ERROR-CORRECTING CODING IN CHANNELS WITH GROUPING ERRORS

Alexandr Kuznetsov[a,b,*], Oleg Oleshko[a], Kateryna Kuznetsova[a]

[a] *V. N. Karazin Kharkiv National University, Svobody sq., 6, Kharkiv, 61022, Ukraine*
[b] *JSC "Institute of Information Technologies", Bakulin St., 12, Kharkiv, 61166, Ukraine*
[*] corresponding author: kuznetsov@karazin.ua

Abstract. This article explores a mathematical model of a data transmission channel with errors grouping. We propose an estimating method for energy gain from coding and energy efficiency of binary codes in channels with grouped errors. The proposed method uses a simplified Bennet and Froelich's model and allows leading the research of the energy gain from coding for a wide class of data channels without restricting the way of the length distributing the error bursts. The reliability of the obtained results is confirmed by the information of the known results in the theory of error-correcting coding in the simplified variant.

Keywords: Modelling of data transmission channels, grouping errors, coding theory, energy gain from coding, channels with grouped errors.

1. Introduction

The usage of redundant codes to increase the reliability of the transmitted information requires the designer to take into account various factors, including the nature of the error distribution in the communication channel [1–9]. A detailed research of statistical properties of error sequences in real channels has shown that the errors are dependent and have a tendency to batch groups [1–3]. Most of the time, the information is transferred via communication channels without any distortions. However, at any point of time, error condensations, so-called error bursts, can occur, inside of which the error probability is significantly higher than the average error probability calculated for a considerable transmission time. In such conditions, protection methods that are optimal for the independent error hypothesis are absolutely ineffective in real communication channels [4–9]. It is necessary to use a (scientific and) methodological apparatus to estimate the error-correcting codes efficiency correctly. It allows to describe the error behaviour in the communication channel and to develop practical recommendations on using error correcting codes.

One of the main criteria of the error-correcting code efficiency is the energy gain from coding (EGC), which refers to the reduction of the minimum required ratio of signal energy to spectral power density of noise. It allows to apply a system of an error-correcting code while ensuring the given probability of the erroneous receiver of signs [10–12]. Today, the EGC is calculated in the known manner, for channels with independent errors. The current direction of research is to develop methods of estimating the energy gain from coding in channels with grouping errors.

2. Model of channel with independent errors

Let's consider the process of a code word decoding in conformity with the binary symmetrical channel model with an independent error distribution. Let's suppose that transmission errors occur independently with a probability P_0. If t is the number of corrected errors by (n,k,d)-block code, $t = \left\lfloor \frac{d-1}{2} \right\rfloor$, then the probability of the erroneous decoding is calculated as

$$P_{ed}(n) = 1 - \sum_{i=0}^{t} C_n^i P_0^i (1-P_0)^{n-i} - \\ - \sum_{i=t+1}^{n} u(i) P_0^i (1-P_0)^{n-i}, \qquad (1)$$

where C_n^i is a binomial coefficient; $u(i)$ is the number of error vectors of weight i, errors being fixed by the code.

If the code corrects all errors within the radius of the code batch and does not fix other bugs, then

$$P_{ed}(n) = P(>t,n) = \sum_{i=t+1}^{n} C_n^i P_0^i (1-P_0)^{n-i} = \\ = 1 - \sum_{i=0}^{t} C_n^i P_0^i (1-P_0)^{n-i}, \qquad (2)$$

To recalculate the error probability of the decoding for one character, P_{ed} will use the expression [10–12]:

$$P_{ed} = \frac{d}{n} P_{ed}(n).$$

After the substitution in (2), we will obtain

$$P_{ed} = \frac{d}{n} \sum_{i=t+1}^{n} C_n^i P_0^i (1-P_0)^{n-i}. \qquad (3)$$

To calculate the EGC, we will consider the probability dependence of P_0 from the ratio of the signal energy to spectral power density of noise to [13]:

$$P_0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\sqrt{E(1-b_s)/N_0}} exp\left(-\frac{y^2}{2}\right) dy = \\ = V\left(\sqrt{\frac{E(1-b_s)}{N_0}}\right), \quad (4)$$

where E is the signal energy, N_0 is the spectral power density of white noise; b_s is the coefficient of mutual correlation between signals; $V(x)$ is the error integral.

Thus, in the case of using binary phase-shift keyed (PSK) signal with a 180^0 phase manipulation, the coefficient of correlation is $b_s = -1$. The probability of P_0 for such signals is determined by the expression

$$P_0 = 0,5\left(1 - \Phi\left(\sqrt{\frac{2E}{N_0}}\right)\right) = 1 - \Phi'\left(\sqrt{\frac{2E}{N_0}}\right),$$

where Φ and Φ' are tabulated functions that represent the probability integrals:

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x exp\left(-\frac{y^2}{2}\right) dy = \\ = 1 - \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{-x} exp\left(-\frac{y^2}{2}\right) dy,$$

$$\Phi'(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} exp\left(-\frac{y^2}{2}\right) dy = \\ = 1 - \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} exp\left(-\frac{y^2}{2}\right) dy.$$

The usage of (n, k, d)-block codes detecting and correcting errors leads to increasing the redundancy of the transmitted data. If you fix the message energy transmitted in the channel, then the energy per symbol is reduced proportionally to the introduced redundancy. To calculate the error probability per symbol at the output of the decoder according to the expression (3), considering the introduced redundancy, we will decrease the ratio of the signal energy to spectral power density of noise in the expression (4) in $R = k/n$ times.

The final expression for the error probability per symbol, using the error-correcting (n, k, d)-block code, will be as follows:

$$P_{ed} = \frac{d}{n} \sum_{i=t+1}^{n} C_n^i \left(V\left(\sqrt{\frac{kE(1-b_s)}{nN_0}}\right)\right)^i \times \\ \times \left(1 - V\left(\sqrt{\frac{kE(1-b_s)}{nN_0}}\right)\right)^{n-i}. \quad (5)$$

For binary PSK signals, the last expression we will rewrite as follows:

$$P_{ed} = \frac{d}{n} \sum_{i=t+1}^{n} C_n^i \left(1 - \Phi'\left(\sqrt{\frac{2kE}{nN_0}}\right)\right)^i \times \\ \times \left(\Phi'\left(\sqrt{\frac{2kE}{nN_0}}\right)\right)^{n-i}. \quad (6)$$

Let's fix the required probability of an error on one character P_d and calculate the required ratio $\gamma_1 = E/N_0$ at $P_0 = P_d$ according to the expression (4), and the ratio $\gamma_2 = E/N_0$ at $P_{er} = P_d$, according to the expression (5). The difference $\gamma_2 - \gamma_1$ gives the needed estimation of the EGC. If the EGC is positive, then the usage of the error-correcting code leads to a gain, and, on the contrary, it is inappropriate to use the chosen (n, k, d)-block code with a negative EGC. The values γ_1, γ_2 and EGC are usually on a logarithmic scale.

3. Model of channel with grouping errors

A convenient tool to describe data transmission channels is Bennet and Froelich's mathematical model, which has no restrictions as to the way of length distribution of error bursts [1–9]. The main features of the Bennet and Froelich's model [4] are:

- constant P_n-probability is the probability that the error package will start from a certain position;
- the independence of error packets occurrence;
- the independence of the P_l-probability of error packet occurrence with l-length from the lengths of other packet errors;
- the independence of errors within the package;
- constant P_ε-probability inside the package;
- no errors outside the package;
- the possibility of contiguity and mutual overlapping of the error packets.

To specify the model, it is enough to determine P_n, P_ε probabilities and $P_n(l, n)$ distribution. At the same time, only solid batches with $P_\varepsilon = 1$ are experimentally singled out. In [6–8], a simplified Bennet and Froelich's model is proposed:

- an error can only arise within the error burst with the constant $(P_\varepsilon = 1)$-probability (continuous packages);
- contiguity and mutual overlap of solid packages do not exist;
- the constant P_n-probability is the probability that a solid error burst of any length will start from a certain position;
- $P(l)$ is the probability of continuous l-length packet occurrence;

- $P_n(l)$ is the probability that a continuous error burst of l-length will start from a certain position,

$$P_n(l) = P_n \cdot P(l).$$

To specify a simplified Bennet and Froelich's model, it is sufficient to specify the P_n-probability and $P(l)$-distribution. The P_n probability value and $P(l)$-distribution can be obtained experimentally on a large enough sample size [1, 2, 4].

According to the Bennet and Froelich's model, the distribution of $P_{ner}(m)$ -probabilities of occurrence of error-free m-length intervals between adjacent continuous error bursts has the following geometrical meaning:

$$P_{ner}(m) = P_n(1 - P_n)^{m-1}. \tag{7}$$

If $P(l)$-distribution can also be presented as a geometrical law

$$P(l) = (1 - g)g^{l-1}, \tag{8}$$

then, the average error burst length l_{av}, the average length of error-free interval m_{av}, the error probability for the bit P_0 and the error burst probability P_n are associated with the values

$$m_{av} = 1, P_0 = P_n l_{av}, l_{av}(1 - g) = 1. \tag{9}$$

To define the considered model, it is sufficient to specify only two parameters, for example, P_0 and l_{av}.

In Figure 1, there are $P(l)$ dependencies for cases: 1) $l_{av} = 2$; 2) $l_{av} = 4$; 3) $l_{av} = 8$; 4) $l_{av} = 16$; 5) $l_{av} = 32$; 6) $l_{av} = 64$.

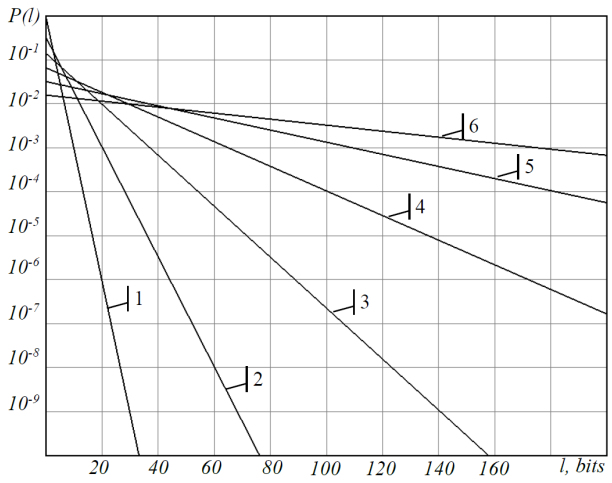

FIGURE 1. The dependencies of the probability of solid package occurrence.

The analysis of dependencies presented in Figure 1 shows that even with a small average error burst length, there is a high probability of a continuous package occurrence. Indeed, with the average error burst length $l_{av} = 8$ bits the probability of a solid package occurrence of 20 errors is $\approx 10^{-2}$. With the same average error burst length, the probability of a solid package occurrence of 40 errors is $\approx 10^{-3}$.

Using the considered simplified model, it is possible to calculate the features of an error-correcting coding efficiency in channels with grouping errors. Let us fix the average error burst length l_{av} and error probability per bit P_0. Then, we obtain:

$$P_n = \frac{P_0}{l_{av}}, g = 1 - \frac{1}{l_{av}},$$

$$P(l) = \frac{1}{l_{av}}\left(1 - \frac{1}{l_{av}}\right)^{l-1},$$

$$P_n(l) = \frac{P_0}{l_{av}^2}\left(1 - \frac{1}{l_{av}}\right)^{l-1}. \tag{10}$$

As an example, Figure 2 shows the dependencies $P_n(l)$ for different values l_{av}: a) $l_{av} = 2$; b) $l_{av} = 4$; c) $l_{av} = 8$, d) $l_{av} = 16$, e) $l_{av} = 32$, f) $l_{av} = 64$. The values $P_n(l)$ were calculated for the case of receiving a binary phase-shift keyed signal and correspond to the following values: 1) $P_0 = P_n(1)$; 2) $P_n(2)$; 3) $P_n(4)$; 4) $P_n(8)$; 5) $P_n(16)$; 6) $P_n(32)$; 7) $P_n(64)$.

The analysis of the dependencies presented in Figure 2 shows that, whith increasing the average error burst length l_{av}, there is a very slight (one to two orders of magnitude) decrease in the probability of the continuous package occurrence of small length ($2 \le l \le 4$ bits). At the same time, there is a significant increase in the probability of the solid package occurrence of great length ($32 \le l \le 64$ bits). So, with $l_{av} = 2$, the probability of the continuous error burst occurrence of the length $l = 64$ bits is

$$P(64) = \frac{1}{2}\left(1 - \frac{1}{2}\right)^{64-1} = \left(\frac{1}{2}\right)^{64} \approx 5.42 \cdot 10^{-20},$$

and the probability of the continuous error burst occurrence of the length $l = 64$ bits from the current position is as follows

$$P_n(64) \approx P_0 \cdot 5.42 \cdot 10^{-20}.$$

With $l_{av} = 64$, the corresponding probabilities are:

$$P(64) = \frac{1}{64}\left(1 - \frac{1}{64}\right)^{64-1} \approx 5.70 \cdot 10^{-3},$$

$$P_n(64) \approx P_0 \cdot 5.70 \cdot 10^{-3},$$

as it can be seen, the probability of occurrence of long solid package increased by seventeen orders of magnitude.

Thus, as the analysis shows, with the increase in the average length of packet errors, a redistribution of the probabilities of occurrence of packet errors takes place: the reduction of the probability of occurrence of packets of short length and the increase in the probability of occurrence of packages of greater length.

Consider the event consisting in the error decoding of linear (n, k, d)-block code, when used in channels with grouping errors. If, for a block of n-symbols, the

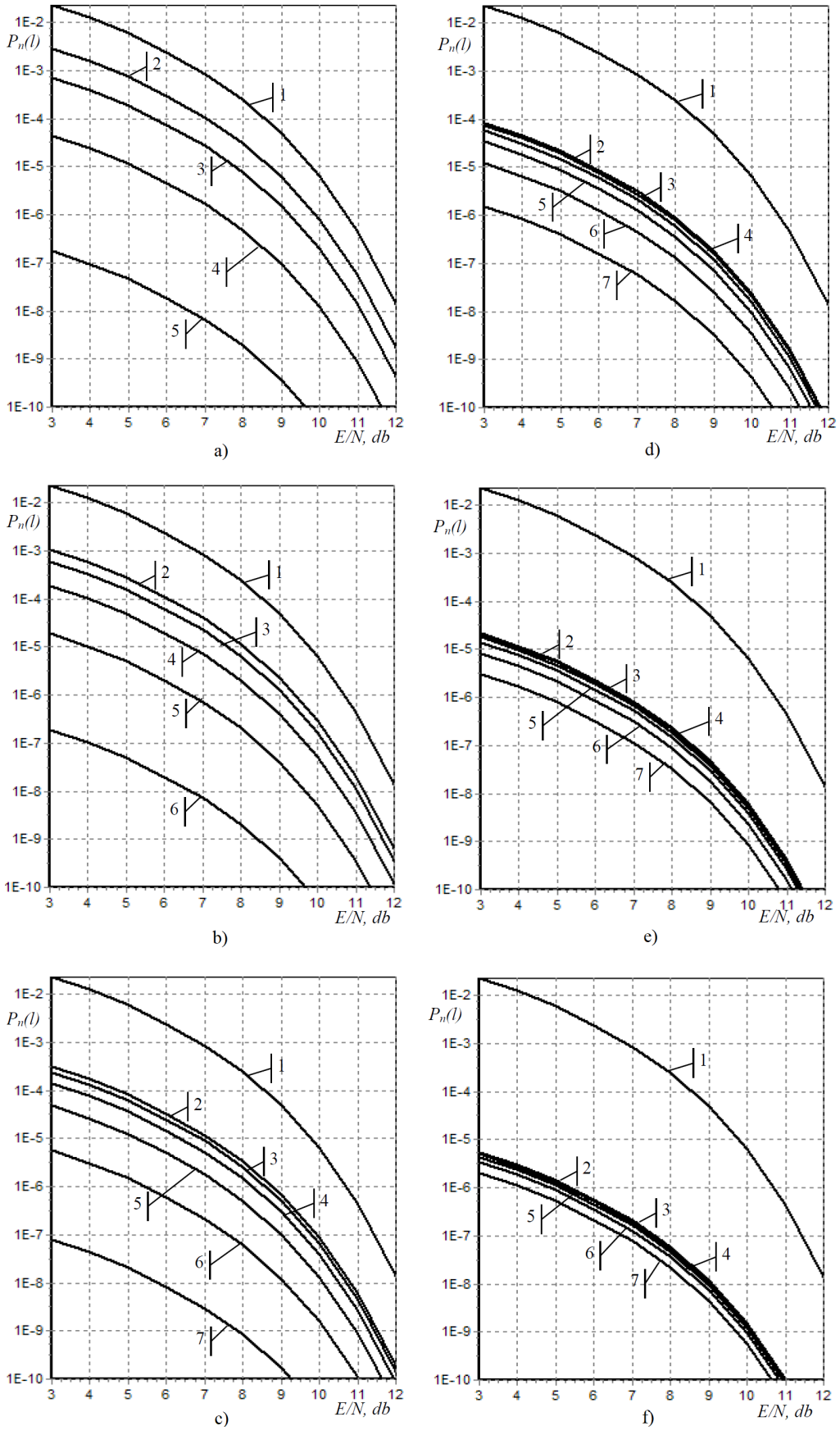

Figure 2. The dependencies $P_n(l)$ for different values l_{av}.

code corrects all errors of t-weight and less, doesn't corrects other errors and errors that occurred in an accordance with the considered model are grouped in packets of l -symbols, then the errors that occurred in an accordance with the considered model are grouped in packets of ξ-packages, so that $\xi l > t$.

Let's consider a simplified Bennet and Froelich's model with a disjoint not adjacent to any other error bursts. In this case, the probability of the error burst occurrence ξ-packages of an l-length on a block with n-symbols is determined by the quantity combination of the package number (with no overlapping parts) on the length of n-symbols. The probability of one error burst occurrence of an l-length on the length of the package of n-symbols is computed as:

$$P_1(l,n) = (n-l+1)P_n(l)(1-P_n(l))^{n-l}.$$

It applies that, on a block length of n-symbols, no more than $\lambda = \left\lfloor \frac{n+1}{l+1} \right\rfloor$ error blocks of an l-length can appear. The combination number of ξ-packages on the length of n-symbols is defined as a value of the binomial coefficient

$$C^{\xi}_{\lambda+n+1-\xi(l+1)-\lambda+\xi} = C^{\xi}_{n-\xi l+1}. \qquad (11)$$

Then, the expression for the probability of ξ-error burst occurrence with l-length on the package of n-symbols is written as:

$$P_{\xi}(l,n) = C^{\xi}_{n-\xi l+1} P_n(l)^{\xi}(1-P_n(l))^{n-\xi l}. \qquad (12)$$

For the probability of $\xi > 1$-packages occurrence from l-errors on the block length of n-symbols, the $\xi l \le t$ is determined by the expression:

$$P_{l<\xi l \le t}(l,n) =$$

$$= \sum_{\substack{\xi=1,\\ l<\xi l \le t}}^{\lambda} C^{\xi}_{n-\xi l+1} P_n(l)^{\xi}(1-P_n(l))^{n-\xi l}. \qquad (13)$$

The probability of the error decoding is defined as

$$P_{ed} = 1-(1-P_n)^n - \sum_{l=1}^{n} P_{l<\xi l \le t}(l,n),$$

where $(1-P_n)^n$ is the probability that, on the block of n-symbols, no error bursts will happen. Then, taking into account (13), we get

$$P_{ed} = 1-(1-P)^n -$$

$$-\sum_{l=1}^{n} \sum_{\substack{\xi=1,\\ l<\xi l \le t}}^{\lambda} C^{\xi}_{n-\xi l+1} P_n(l)^{\xi}(1-P_n(l))^{n-\xi l}. \qquad (14)$$

When compared to the model with independent errors, the disadvantage of the model with no attached error bursts is irreducibility, even in the case of the fixed $l = 1$. Indeed, let us suggest, that $l_{av} = 1$, then $g = 0$, $P(1) = 1$, $P(>1) = 0$, $P_n = P_0 = P_n(l)$, with only single errors, which are not adjacent to each other, appearing. Then, the expression (12-13) will be rewritten as:

$$P_{\xi}(l,n) = C^{\xi}_{n-\xi+1} {P_0}^{\xi} \cdot (1-P_0)^{n-\xi},$$

$$P_{1<\xi \le t}(l,n) = \sum_{\xi=1}^{t} C^{\xi}_{n-\xi+1} {P_0}^{\xi}(1-P_0)^{n-\xi},$$

and the expression for the decoding of an error probability will be written as

$$P_{ed} = 1-(1-P_0)^n - \sum_{\xi=1}^{t} C^{\xi}_{n-\xi+1} {P_0}^{\xi}(1-P_0)^{n-\xi} =$$

$$= 1 - \sum_{\xi=0}^{t} C^{\xi}_{n-\xi+1} {P_0}^{\xi}\,(1-P_0)^{n-\xi}. \qquad (15)$$

that does not correspond to the expressions (1-2) for the model with independent errors. Let's analyse the reasons of this disparity.

The expression (15) conforms to the probability of an erroneous decoding, distorted by such single errors, which cannot adhere to each other. In general, the model of a binary symmetrical channel without memory, described by the expressions (1-2), allows adjoining single errors that cause the discrepancy between the corresponding formulas. Let's consider a simplified Bennet and Froelich's model with disjoint error bursts and their possible adjacency to each other. In this case, on the block length of n-symbols, there could be no more than $\lambda' = \lfloor n/l \rfloor$ error bursts of an l-length. The number of combinations of ξ-packages on the length of n-symbols is defined by the binomial coefficient

$$C^{\xi}_{\lambda'+n-\xi l-\lambda'+\xi} = C^{\xi}_{n-\xi l+\xi}. \qquad (16)$$

Then the expression for the ξ-error bursts occurrence probability of l-length on the block of n-symbols is determined by the expression:

$$P_{\xi}(l,n) = C^{\xi}_{n-\xi l+\xi} P_n(l)^{\xi}(1-P_n(l))^{n-\xi l}. \qquad (17)$$

For any occurrence of $\xi > 1$ of l-errors packages on the block of n-symbols with $\xi l \le t$, $P_{l<\xi l \le t}(l,n)$, the probability is computed by the expression

$$P_{l<\xi l \le t}(l,n) =$$

$$= \sum_{\substack{\xi=1,\\ l<\xi l \le t}}^{\lambda'} C^{\xi}_{n-\xi l+\xi} \cdot P_n(l)^{\xi} \cdot (1-P_n(l))^{n-\xi l}. \qquad (18)$$

The probability of erroneous decoding is computed as

$$P_{ed} = 1 - (1 - P_n)^n - \sum_{l=1}^{n} P_{1<\xi \le t}(l, n),$$

and, taking into account (18), we get

$$P_{ed} = 1 - (1 - P_n)^n - \\ - \sum_{l=1}^{n} \sum_{\substack{\xi=1, \\ l<\xi l \le t}}^{\lambda'} C^{\xi}_{n-\xi l+\xi} P_n(l)^{\xi} (1 - P_n(l))^{n-\xi l}. \quad (19)$$

The disadvantage of the model with not-adjacent error bursts is the irreducibility, even with the fixed $l = 1$, when compared to the model with independent errors. Indeed, let's suggest that $l_{av} = 1$, $g = 0$, $P(1) = 1$, $P(> 1) = 0$, $P_n = P_0 = P_n(l)$ and only single errors occur, which are not adjacent to each other.

Then, the expression (17-18) will be rewritten in the following form

$$P_{\xi}(l, n) = C^{\xi}_n P_0{}^{\xi} (1 - P_0)^{n-\xi},$$

$$P_{1<\xi \le t}(l, n) = \sum_{\xi=1}^{t} C^{\xi}_n P_0{}^{\xi} (1 - P_0)^{n-\xi},$$

and the expression for the error probability decoding will be transformed to

$$P_{ed} = 1 - (1 - P_0)^n - \sum_{\xi=1}^{t} C^{\xi}_n P_0{}^{\xi} (1 - P_0)^{n-\xi} =$$

$$= 1 - \sum_{\xi=0}^{t} C^{\xi}_n P_0{}^{\xi} (1 - P_0)^{n-\xi},$$

that with $i = \xi$, it fully complies with the expressions (1-2) for the model with independent errors.

To calculate the EGC of (n, k, d)-block code in a channel with grouping errors, it is necessary to fix a required error probability on one character P_d and calculate the corresponding value of E/N_0 by expressions (14) and/or (19) (taking into account the introduced redundancy and the multiplier d/n). The difference of γ_2 and γ_1 gives the required estimation of the EGC:

$$EGC = \gamma_2 - \gamma_1,$$

where γ_2 is the ratio E/N_0, the minimum required to achieve a desired probability of an erroneous reception of symbols P_d for a fixed ensemble of signals (without coding); γ_1 is the ratio E/N_0, the minimum required to achieve P_d for a fixed ensemble of signals by using the (n, k, d)-code block.

It should be noted that the considered mathematical model and methodology of evaluating the energy gain are not limited by the distribution of the length of error bursts that allows to explore the EGC for a wide class of data channels.

$GF(2^m)$ (n, k, d)	**The roots the polynomial** $g(x)$
$GF(2^4)$ $(15, 7, 5)$	α^1, α^2, α^4, α^8, α^3, α^6, α^9, α^{12}
$GF(2^5)$ $(31, 16, 7)$	α^1, α^2, α^4, α^8,α^{16}, α^3, α^6, α^{12},α^{24}, α^{17}, α^5, α^{10}, α^{20}, α^9, α^{18}
$GF(2^6)$ $(63, 36, 11)$	α^1, α^2, α^4, α^8, α^{16}, α^{32}, α^3, α^6, α^{12}, α^{24}, α^{48}, α^{33}, α^5, α^{10}, α^{20}, α^{40}, α^{17}, α^{34}, α^7, α^{14}, α^{28}, α^{56}, α^{49}, α^{35}, α^9, α^{18}, α^{36}
$GF(2^7)$ $(127, 64, 21)$	α^1, α^2, α^4, α^8, α^{16}, α^{32}, α^{64}, α^3, α^6, α^{12}, α^{24}, α^{48}, α^{96}, α^{65}, α^5, α^{10}, α^{20}, α^{40}, α^{80}, α^{33}, α^{66}, α^7, α^{14}, α^{28}, α^{56}, α^{112}, α^{97}, α^{67}, α^9, α^{18}, α^{36}, α^{72}, α^{17}, α^{34}, α^{68}, α^{11}, α^{22}, α^{44}, α^{88}, α^{49}, α^{98}, α^{69}, α^{13}, α^{26}, α^{52}, α^{104}, α^{81}, α^{35}, α^{70}, α^{15}, α^{30}, α^{60}, α^{120}, α^{113}, α^{99}, α^{71}, α^{19}, α^{38}, α^{76}, α^{25}, α^{50}, α^{100}, α^{73}

TABLE 1. Primitive binary BCH-code.

4. EGC-evaluation of BCH codes

Consider the binary Bose-Chaudhuri-Hocquenghem (BCH) codes for an assessment of their EGC into channels with independent and grouping errors. To evaluate the EGC, we will use the techniques developed earlier. The theory and methods of constructing BCH codes are best described in the monographs [10–12], in which it is shown that BCH codes yield the highest win at $R \approx 1/2$.

Fix a finite field $GF(2^m)$, $m = 4, 5, 6, 7$ and primitive binary BCH-code with $R \approx 1/2$. In Table 1, the corresponding code parameters and the roots of the generating polynomial $g(x)$ in the form of the degrees of the primitive element of the field are presented.

Figure 3 shows the dependencies of the probability of erroneous receiving of symbols for the cases:

- 1 - the optimum receiving of binary PSK signals (without coding);
- 2 - using the $(15, 7, 5)$ BCH-code;
- 3 - using the $(31, 16, 7)$ BCH-code;
- 4 - using the $(63, 36, 11)$ BCH-code;
- 5 - using the $(127, 64, 21)$ BCH-code.

In Figure 3a), the reduced dependencies correspond to the model considered earlier with an independent occurrence of errors, which is equivalent to the case $l_{av} = 1$. In the rest of the figures, the reduced dependencies correspond to the probability of receiving erroneous channel symbols, which is described by a simplified Bennet and Froelich's model containing disjoint error packets and their possible contiguity to each

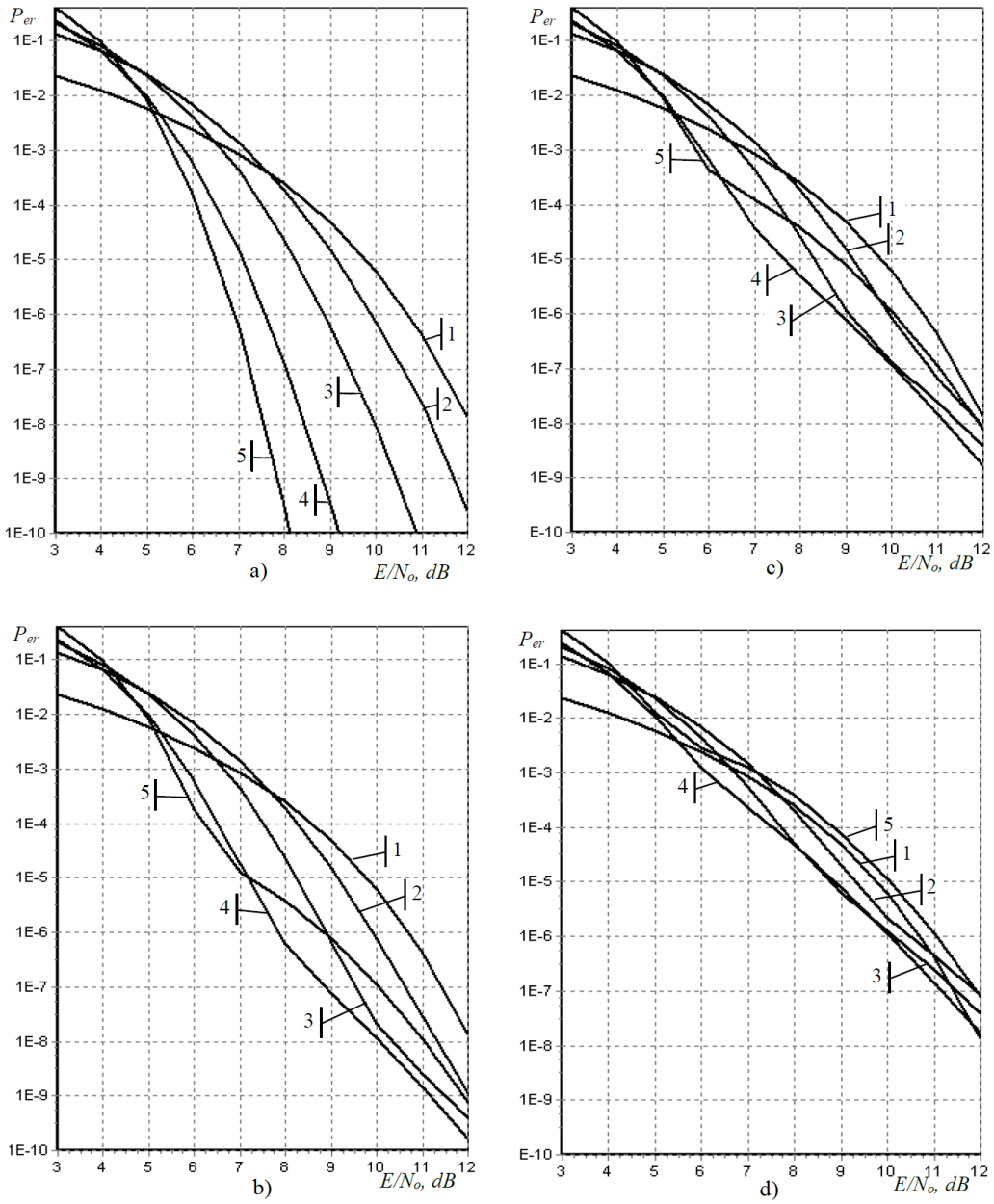

FIGURE 3. The dependencies of the probability of erroneous receiving of a symbol.

other for the cases: b) $l_{av} = 1,00001$; c) $l_{av} = 1,0001$; d) $l_{av} = 1,001$. As follows from the reduced dependencies, even a small clustering of errors leads to a sharp decrease of the EGC.

Indeed, for the above considered model with independent errors at $P_d = 10^{-5}$, the EGC binary $(127, 64, 21)$ BCH-code is equal to $\approx 3.2dB$ (see Figure 3a). But when $l_{av} = 1,0001$, the EGC is sharply reduced and, for the same, value P_d is approximately equal to 0.9 dB (see Figure 3b). With an increasing average length of a package of continuous errors, the EGC continues to decrease and becomes negative (see Figure 3d). Practically, this means that the use of strictly random error-correcting coding in channels with a strong clustering of errors is inefficient and leads to energy loss.

5. CONCLUSIONS

As the result of the research, we have developed a scientific and methodological apparatus description of the error behaviour in discrete channels. For the first time in the open literature, we have developed the methodology of estimating the EGC-binary codes in channels with grouping errors.

The proposed method uses a simplified Bennet and Froelich's model and allows conducting a research of the EGC for a wide class of data channels with different laws of length distribution of the error bursts. The reliability of the obtained results is confirmed by the information in the simplified variant of the known results in the theory of the error-correcting coding. So, the estimating expression of probability of erroneous decoding in channels with grouping errors

with $l_{av} = 1$ (19) is reduced to the known expression (2) concerning the computation of the probability of erroneous decoding in channels with an independent error distribution.

Using the developed methodology, we have conducted the estimation of the EGC for binary BCH-codes, which has shown a significant decrease of the efficiency of error-correcting coding in channels with grouping errors. Even with a negligible error grouping, the EGC of binary BCH-codes is sharply reduced, while further increasing the average length of an error burst, the EGC reaches negative values.

A remark should be added, stating that many binary BCH codes are optimum for correcting burst errors when they satisfy the Reiger bound, i.e. whenever $2b = n - k$, where b denotes the maximum guaranteed correctable burst length and the BCH code has parameters (n, k, d). However, the introduced redundancy with this encoding still requires significant energy costs (for the transmission of each redundant bit). For channels with a strong clustering of errors, the EGC may be small or negative, i.e. coding may lead to an energy loss. Thus, even codes that are optimal for correcting error bursts may not compensate for the energy costs of transmitting redundant symbols. Obviously, in this case, it is better to use protocols with error detection and automatic repeat request (ARQ).

A perspective direction for further research is the development of estimating methods of the EGC for non-binary codes in channels with grouping errors and a direct estimation of the EGC for such codes (e.g., for non-binary BCH-codes, Reed-Solomon codes, algebraic geometric codes).

These results may also be useful in other important practical applications: cryptography, authentication, the theory of complex signals, etc. [14–16].

References

[1] E. N. Gilbert. Capacity of a burst-noise channel. *Bell System Technical Journal* **39**(5):1253–1265, 1960. doi:10.1002/j.1538-7305.1960.tb03959.x.

[2] A. Fontaine, R. Gallager. Error statistics and coding for binary transmission over telephone circuits. *Proceedings of the IRE* **49**(6):1059–1065, 1961. doi:10.1109/jrproc.1961.287890.

[3] M. Zorzi, R. Rao, L. Milstein. Error statistics in data transmission over fading channels. *IEEE Transactions on Communications* **46**(11):1468–1477, 1998. doi:10.1109/26.729391.

[4] W. Bennett, F. Froehlich. Some results on the effectiveness of error-control procedures in digital data transmission. *IRE Transactions on Communications Systems* **9**(1):58–65, 1961. doi:10.1109/TCOM.1961.1097651.

[5] E. Bloch, O. Popov, W. Y. Turin. Models of error sources in channels for digital information transmission. *Moscow, Svyaz* 1971.

[6] W. Turin. Error source models. In *Performance Analysis and Modeling of Digital Transmission Systems*, pp. 9–59. Springer US, 2004. doi:10.1007/978-1-4419-9070-9_2.

[7] O. Melentiev, V. Yatsukov, E. Minina. The estimation technique of parameters of discrete channel with grouping errors. In *2003 Siberian Russian Workshop on Electron Devices and Materials. Proceedings. 4th Annual (IEEE Cat. No.03EX664)*. Novosibirsk Student Branch of IEEE, 2003. doi:10.1109/sredm.2003.1224208.

[8] T. Kløve, V. I. Korzhik. Channel models. In *Error Detecting Codes*, pp. 1–16. Springer US, 1995. doi:10.1007/978-1-4615-2309-3_1.

[9] J. Poikonen, J. Paavola. Error models for the transport stream packet channel in the dvb-h link layer. In *2006 IEEE International Conference on Communications*, vol. 4, pp. 1861–1866. 2006. doi:10.1109/ICC.2006.254991.

[10] G. C. Clark, J. B. Cain. *Error-Correction Coding for Digital Communications*. Springer US, 1981. doi:10.1007/978-1-4899-2174-1.

[11] I. S. Reed, X. Chen. *Error-Control Coding for Data Networks*. Springer US, 1999. doi:10.1007/978-1-4615-5005-1.

[12] F. MacWilliams, N. Sloane. *The Theory of Error-Correcting Codes*. Elsevier, 1977. doi:10.1016/s0924-6509(08)x7030-8.

[13] B. Sklar. *Digital Communications: Fundamentals and Applications (Paperback)*. Prentice Hall Communications Engineering and Emerging Techno. Pearson Education, 2016.

[14] A. Kuznetsov, M. Lutsenko, N. Kiian, et al. Code-based key encapsulation mechanisms for post-quantum standardization. In *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, pp. 276–281. 2018. doi:10.1109/DESSERT.2018.8409144.

[15] A. Kuznetsov, A. Pushkar'ov, N. Kiyan, T. Kuznetsova. Code-based electronic digital signature. In *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, pp. 331–336. 2018. doi:10.1109/DESSERT.2018.8409154.

[16] A. Kuznetsov, I. Svatovskij, N. Kiyan, A. Pushkar'ov. Code-based public-key cryptosystems for the post-quantum period. In *2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S T)*, pp. 125–130. 2017. doi:10.1109/INFOCOMMST.2017.8246365.