ISSN 1064-2269, Journal of Communications Technology and Electronics, 2017, Vol. 62, No. 6, pp. 669–685. © Pleiades Publishing, Inc., 2017.
Original Russian Text © V.B. Afanassiev, A.A. Davydov, D.K. Zigangirov, 2016, published in Informatsionnye Protsessy, 2016, Vol. 16, No. 4, pp. 382–404.

THEORY AND METHODS OF INFORMATION PROCESSING

Estimating the Fraction of Erasure Patterns Correctable by Linear Codes

V. B. Afanassiev*, A. A. Davydov, and D. K. Zigangirov*****

Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, 127051 Russia

**e-mail: afanv@iitp.ru*

***e-mail: adav@iitp.ru*

****e-mail: zig@iitp.ru*

Received December 2, 2016

Abstract—The conditional probability (fraction) of the successful decoding of erasure patterns of high (greater than the code distance) weights is investigated for linear codes with the partially known or unknown weight spectra of code words. The estimated conditional probabilities and the methods used to calculate them refer to arbitrary binary linear codes and binary Hamming, Panchenko, and Bose–Chaudhuri–Hocquenghem (BCH) codes, including their extended and shortened forms. Error detection probabilities are estimated under erasure-correction conditions. The product-code decoding algorithms involving the correction of high weight erasures by means of component Hamming, Panchenko, and BCH codes are proposed, and the upper estimate of decoding failure probability is presented.

Keywords: linear code, erasure correction, product code

DOI: 10.1134/S106422691706002X

1. INTRODUCTION

The problem concerning erasure corrections in the erasure channel has been investigated for a long time. The throughput for the channel with independent errors and erasure is known. The exponential bounds of the error probability are known for codding–decoding schemes involving erasures, lists of variable length, and feedback [1]. The algorithms and bounds on generalized distance decoding and many other results are known. The list of publications is huge! However, for binary linear codes, many problems are still unresolved, e.g., *the fraction of correctable erasure combinations with weights greater than or equal to the code distance for arbitrary and particular classes of codes and the decoder failure probability during correction of only erasures or errors and erasures*. At present, the correcting ability of low-density parity-check (LDPC) codes in the erasure channel is being actively studied. For example, in [2] has been investigated the erasure correction by concatenation of LDPC codes with the Hamming code. In this case, for binary Hamming codes, the exact estimates of the fraction of correctable erasures with large weights are derived and employed. It is worthy to note [3], where the fraction of erasures corrected by nonbinary codes was discussed. The results concerning investigations of minimal codewords in linear codes [4] can be useful for estimates of the fraction of correctable erasure patterns.

The estimate of the fraction of correctable erasures with large weights is necessary to optimize decoding in the channel with erasures and errors and the decoding of product or concatenated codes. It is interesting to consider decreasing of decoder fault and error probability with extension of the decoding area.

The issues concerning the complexity of correction of erasures of large weights have been intensely investigated in 1963–1980 under the condition of the low capacity of computation devices. For example, the author of [5, 6], examined the algorithms and schemes of correction of erasures of large weights and error detection for linear and certain cyclic codes. Nowadays, the complexity is not critical but remains an important problem.

The theoretical part of the work related to the corrections of erasures of large weights contains the exact or lower estimates of the fractions of correctable erasure combinations of specified weights (from the code distance to the number of code checks) for arbitrary and particular block codes with the known weight spectrum and partially known and completely unknown codeword weight spectrum. In a number of cases, these estimates are accurate. It was proved that the fraction of correctable configurations does not decrease upon code shortening and, in principle, can grow. In addition, it was demonstrated that the fraction of correctable configurations remains unchangeable with code extension under the certain (often fulfilled) conditions. In particular, this refers to nonshortened Hamming codes with $d = 3$ and Bose–Chaudhuri–Hocquenghem (BCH) codes with $d = 5$. As the objects for future investigations there was chosen Hamming and

Panchenko codes with code distance 4 and BCH code with distance 6.

In the applied aspects, this work continues work [7] in the direction of increasing the reliability of long-term memory devices (e.g., solid-state disks (SSDs)) or their modifications. In the memory devices of the given type, channel (memory cells) degradation looks like an error accumulation. However, the error-accumulation rate is relative small, and each recording cycle again and again initiates the error-accumulation process. With the aim of optimizing the reading and recording of swap files, data exchange between random-access and long-term memories occurs by blocks of sufficiently large size (tens or even hundreds or thousands of standard words). The natural compromise between the complexity and the reliability is the usage of the concatenated coding structure, ~~preserves~~ encoding for standard words and incorporates them with outer encoding. As such a compromised structure, the product of component code was studied in [7].

In the case of the product of binary codes, an unsolved problem is, e.g., the best decoding with partial correction of dependent error configurations in "neighboring" (in some structural sense) words. Such error "spots" often arise when memory microcircuits are externally irradiated or the temperature conditions are critical. Correction of erasures of large weight make it possible to extend the decoding area of the products of Hamming, Panchenko, and BCH codes whose code distances are 4 and 6. As a result, the size of almost all correctable error spots can be greater than 4×4 and 6×6.

In this work, the fraction of correctable error configurations of large weights is estimated for component codes and the probabilities of correct and faulty decoding are determined for product codes. Strictly speaking, estimation is performed under the assumption of independence of errors and erasures due to the absence of adequate model (to the application area) of dependent events. Actually, this is not a substantial "minus" of the given work because all correctable erasure combinations are located in the certain (restricted) amount of rows and columns. Therefore, any spot configuration is their subset. For "extended" decoding of the product of Panchenko and BCH codes with correction of erasure patterns of a large weight, the failure probability estimate indicates a radical decrease in fault probability in comparison with usual ("nonextended") decoding of the product code even if the extension of the decoding area is only unity.

As a matter of fact, the concept of error spot (or a 2D error burst) has appeared long ago. This concept is related to the theory of Markov processes occurring in the communications channels. In applied problems, the given concept is connected with decoding for hard disks, tape and diskette information carriers, and optical discs (compact disks and digital video disks) [8]. In the decoding range of SSDs, an attainable resource is strictly restricted first of all by the allowable delay. For this reason Hamming and Panchenko codes with distance 4 and BCH codes with distance 6 were chosen as the objects under study.

Let the $[n, n-r, d]$ code be the binary linear code with length n, redundancy r, minimum distance d. A parity check matrix of a code will be called, for short, by "check matrix". The quantities coupled with this code can be designated as follows: n is the code length, r is the number of check symbols, d is the minimum code distance, A_w is the number of codewords of weight w, ρ is the weight of erasures whose correction is analyzed (only the case of $\rho \le r$ is considered), S_ρ is the number of erasure configurations with weight ρ correctable by the code (or, equivalently, the amount of different sets of ρ linearly independent columns of the check matrix of the code or the number of different $r \times \rho$ submatrices of the full rank in the check matrix), and δ_ρ is the fraction of erasures of weight ρ correctable by the code:

$$\delta_\rho = \frac{S_\rho}{\binom{n}{\rho}} \le 1. \quad (1.1)$$

Obviously, for any $[n, n-r, d]$ code, we have

$$S_\rho = \binom{n}{\rho}, \quad \delta_\rho = 1, \text{ if and only if } \rho \le d-1. \quad (1.2)$$

In the case of the binary nonshortened $[2^r-1, 2^r-1-r, 3]$ Hamming code, the following system of notation is used:

(i) $S^H_{\rho,r}$ is the number of correctable erasure configurations with weight ρ;

(ii) $\delta^H_{\rho,r}$ is the fraction of correctable erasures with weight ρ, where $\delta^H_{\rho,r} = \frac{S^H_{\rho,r}}{\binom{2^r-1}{\rho}} \le 1$ in compliance with (1.1).

For the binary nonshortened extended $[2^{r-1}, 2^{r-1}-r, 4]$ Hamming code obtained from the $[2^{r-1}-1, 2^{r-1}-r, 3]$ code by supplement of overall parity check symbol, the following system of notation is employed:

(i) $S^{H\bullet}_{\rho,r}$ is the number of correctable erasure configurations with weight ρ.

(ii) $\delta^{H\bullet}_{\rho,r}$ is the fraction of correctable erasures with weight ρ, where $\delta^{H\bullet}_{\rho,r} = \frac{S^{H\bullet}_{\rho,r}}{\binom{2^{r-1}}{\rho}} \le 1$ in compliance with (1.1).

This paper is organized as follows. Section 2 deals with the method for calculating the number S_ρ of the

sets containing ρ linearly independent columns of the check matrix of the code with the known weight spectrum. The corresponding asymptotic estimates are derived. The recursive approach to the estimation of quantity δ_ρ is proposed. Section 3 provides the recurrent estimates of S_ρ and δ_ρ for arbitrary codes and codes with even weights. The dependence between weight and recurrent approaches and their joint application are demonstrated. For the Hamming code with $d = 3$ and its extension, the exact values of $S^H_{\rho,r}$, $\delta^H_{\rho,r}$, $S^{H\bullet}_{\rho,r}$, and $\delta^{H\bullet}_{\rho,r}$ are obtained. The relationships between S_ρ and δ_ρ derived with the help of shortened and nonshortened codes and punctured and extended ones are investigated in Section 4. The formulas for determining S_ρ and δ_ρ of Hamming, Panchenko, and BCH codes, which are based on the results of previous calculations, are deduced in Section 5, where the corresponding graphs and tables are presented as well. Erasure corrections with error detection are discussed in Section 6. The product-code decoding algorithms relied on the correction of erasures with large weights are proposed in Sections 7 and 8, which also contain formulas for calculating the successful decoding probabilities. Several examples are presented all over the paper text.

2. FRACTION OF ERASURES CORRECTABLE BY LINEAR CODES WITH KNOWN SPECTRAL WEIGHTS

2.1. Number of Sets of Linearly Independent Columns of the Check Matrix for a Code with the Known Weight Spectrum

The necessary condition of the correction of erasure patterns of large weight ρ is the full rank of the submatrix composed of the columns of check matrix of the code, which correspond to the erased positions. In this section, the number and fraction of the such submatrices are estimated.

Below, for the $[n, n-r, d]$ code with weight spectrum $A_0, A_1, \ldots, A_n$, we introduce the function

$$\Psi(n,d,\rho) = \binom{n}{\rho} - \sum_{w=d}^{\rho} A_w \binom{n-w}{\rho-w},\ d \le \rho \le r, \quad (2.1)$$

the value of which is the lower estimate of the number S_ρ of erasure configurations with weights ρ correctable by the code under consideration. (In a number of cases, this estimate is accurate.) To emphasize that function $\Psi(n, d, \rho)$ is calculated for certain $[n, n-r, d]$ code C, this quantity is written as $\Psi(n, d, \rho, C)$.

Theorem 2.1. *The fraction δ_ρ of erasures with weights ρ correctable by an $[n, n-r, d]$ code and the number S_ρ of different sets of ρ linearly independent columns of the check matrix of the given code satisfy the following lower estimates*:

$$\delta_\rho \ge \frac{\Psi(n,d,\rho)}{\binom{n}{\rho}},\quad S_\rho \ge \Psi(n,d,\rho),\ d \le \rho \le r. \quad (2.2)$$

In particular, the equality

$$\delta_\rho = \frac{\Psi(n,d,\rho)}{\binom{n}{\rho}},\ S_\rho = \Psi(n,d,\rho), \quad (2.3)$$

holds under the fulfillment of the condition

$$\rho - d \le \frac{d-1}{2}. \quad (2.4)$$

Proof. Quantity S_ρ is the difference between the total number of sets of ρ columns of check matrix $\binom{n}{\rho}$ and the number of configurations of ρ linearly dependent columns. Any configuration of ρ linearly dependent columns of the check matrix can be obtained if $\rho - w$ columns are added to the set from w columns with the zero sum corresponding to the codeword with weight w (hence, $\rho \ge w$). For the fixed set of w columns, the number of methods for selecting the above $\rho - w$ columns is $\binom{n-w}{\rho-w}$. The foregoing explains the structure of expression (2.1). In principle, the $r \times \rho$ submatrix of the check matrix, where $d \le \rho \le r$, can contain more than one subset from $\ge d$ columns with the zero sum. This leads to the fact that, in expression (2.1), certain linearly dependent configurations of columns are calculated (and subtracted from $\binom{n}{\rho}$) more than onefold. Hence, (2.2) comprises the sign "≥". Under condition (2.4), all linearly dependent configurations of columns are calculated (and subtracted from $\binom{n}{\rho}$) onefold. From this follows equality (2.3). The theorem is proved.

Remark 1. As far as is known to us, O.V. Popov pioneered the application of the estimate analogous to (2.1).

Remark 2. It follows from (2.1), (2.3), and (2.4) that, for all $[n, n-r, d]$ codes, the number S_d of different sets of d linearly independent columns of the check matrix is $S_d = \binom{n}{d} - A_d$. Therefore, $\max\{S_d(n, r)\} = \binom{n}{d} - \min\{A_d(n, r)\}$, where $\max\{S_d(n, r)\}$ is the largest possible value of quantity S_d at fixed n, r, and d and $\min\{A_d(n, r)\}$ is the smallest possible number of words with a minimum weight at fixed n, r, and d. The

bounds corresponding to $\min\{A_d(n, r)\}$ and codes achieving this bounds can be found in [9–13].

Many authors (see, e.g., [7, 9–21] and references therein) have been studied (and continue to study) the weight spectra of codes and their asymptotics.

The estimates of Theorem 2.1 can be improved with the help of the following lemma.

Lemma 2.1. *Any set of* ρ *linearly dependent columns of the check matrix is the combination of the set from w columns with the zero sum corresponding to a codeword of weight w and the set from* $\rho - w$ *linearly independent columns, where* $d \le w \le \rho$.

Using function (2.1) in a recursive manner, let us construct the "inclusion–exclusion" *recursive scheme*, which, in principle, can improve estimate (2.2):

$$\tilde{\Psi}(n,d,\rho) = \binom{n}{\rho} - \sum_{w=d}^{\rho} A_w(n)\,\tilde{\Psi}(n-w,d,\rho-w),$$

where $A_w(n)$ is the number of words with weight w in the code of length n. For one and two steps of recurrence, the estimates of the fraction of erasures are expressed, respectively, as

$$\tilde{\delta}(n,d,\rho) = \frac{\tilde{\Psi}(n,d,\rho)}{\binom{n}{\rho}}$$

$$= 1 - \sum_{w=d}^{\rho} A_w(n)\,\tilde{\delta}(n-w,d,\rho-w)\frac{\binom{n-w}{\rho-w}}{\binom{n}{\rho}},$$

$$\tilde{\delta}(n,d,\rho) = 1 - \sum_{w_1=d}^{\rho} A_{w_1}(n)\frac{\binom{n-w_1}{\rho-w_1}}{\binom{n}{\rho}}$$

$$\times\left[1 - \sum_{w_2=d}^{\rho-w_1} A_{w_2}(n-w_1)\frac{\binom{n-w_1-w_2}{\rho-w_1-w_2}}{\binom{n-w_1}{\rho-w_1}}\right].$$

2.2. Asymptotic Estimation of the Fraction of Correctable Erasures

The following observation can motivate the inclusion of erasures with large weights into the decoding process. Using the know binomial approximation of the weight spectrum of the [15, 17, 18, 20] linear code $A_w \approx 2^{-z}\binom{n}{w}$, ($r - 1 < z \le r$ and $w \ge d$, where z is the real number allowing for (in principle) the correction term in the aforementioned approximations and weight restriction $w \ge d$), we can obtain the following approximate estimate of the behavior of function S_ρ in the interval $d \le \rho < r$:

$$S_\rho \ge \binom{n}{\rho} - \sum_{w=d}^{\rho} A_w\binom{n-w}{\rho-w} \approx \binom{n}{\rho} - 2^{-z}\sum_{w=d}^{\rho}\binom{n}{w}\binom{n-w}{\rho-w}$$

$$= \binom{n}{\rho} - 2^{-z}\binom{n}{\rho}\sum_{w=d}^{\rho}\binom{\rho}{w}.$$

Thus (see [18, Lemma 10.8]), we obtain the estimate of the fraction of correctable erasures with large weights:

$$\delta_\rho \ge \frac{S_\rho}{\binom{n}{\rho}} \approx 1 - 2^{-z}\sum_{w=d}^{\rho}\binom{\rho}{w}$$

$$\approx 1 - 2^{-z}\times 2^{\rho H(d/\rho)} \ge 1 - 2^{\rho - z},$$

$$d \le \rho < z,$$

where $H(d/\rho)$ is the binary entropy. The reference point is $S_d \times \binom{n}{d}^{-1} \approx 1$.

It is seen from the proposed estimate that the fraction of correctable erasures with large weights diminishes exponentially with increasing weight at the fixed number of checks. For this reason, it can be assumed that the weight interval of the combinations of erasures is sufficiently (softly) restricted by the value on the order of $2d$ (or the lesser value).

3. RECURRENT ESTIMATION OF THE FRACTION OF CORRECTABLE ERASURES

Let us introduce the designation

$$\lambda(d,\rho) = \begin{cases} 0 & \text{for} \quad d = 3 \\ \sum_{i=2}^{d-2}\binom{\rho-1}{i} & \text{for} \quad d \ge 4 \end{cases}.$$

Lemma 3.1. *The number* S_ρ *of different sets of* ρ *linearly independent columns of the check matrix of an* $[n, n-r, d]$ *code is recurrently estimated as follows*:

$$S_\rho \ge \frac{1}{\rho}S_{\rho-1}\left(n + 1 - 2^{\rho-1} + \lambda(d,\rho)\right), \qquad (3.1)$$

$$d \le \rho \le r.$$

In particular, the $[2^r - 1, 2^r - 1 - r, 3]$ *Hamming code satisfies the equality*

$$S^H_{\rho,r} = \frac{1}{\rho}S^H_{\rho-1,r}\left(2^r - 2^{\rho-1}\right), \quad 3 \le \rho \le r. \qquad (3.2)$$

Proof. Set Γ_ρ containing the ρ linearly independent columns of the check matrix can be obtained if the certain (correctly chosen) column is added to set $\Gamma^{(b)}_{\rho-1}$

from $\rho - 1$ linearly independent columns. Here, $b = 1, 2, \ldots, S_{\rho-1}$ is the set number. The added column is chosen from the $n - (\rho - 1)$ columns of the matrix not included in the set $\Gamma_{\rho-1}^{(b)}$. This explains term $n - \rho + 1$ in formula (3.3).

Let $\rho - 1 \geq d - 1$ and $2 \leq j \leq \rho - 1$. It is assumed that $\Gamma_{\rho-1,j}^{(b,u)}$ is the subset from the j columns of the set $\Gamma_{\rho-1}^{(b)}$, where $u = 1, 2, \ldots, \binom{\rho-1}{j}$ is the subset number. Any subset $\Gamma_{\rho-1,j}^{(b,u)}$ is linearly independent, and the sum $\sum_j^{(b,u)}$ of all subset columns is not equal to zero. Hence, for all subsets $\Gamma_{\rho-1,j}^{(b,u)}$, the added column cannot be equal to $\sum_j^{(b,u)}$. Otherwise, we obtain the linearly dependent set of ρ columns. If $2 \leq j \leq d - 2$, the check matrix cannot comprise the column equal to $\sum_j^{(b,u)}$: this would lead to the fact that the $j + 1$ columns with the zero sum would exist in the matrix, where $j + 1 \leq d - 1$. Hence, the situations with $d - 1 \leq j \leq \rho - 1$ are considered below. In formula (3.3), the term $\sum_{j=d-1}^{\rho-1} \binom{\rho-1}{j}$ estimates from above the number of columns that cannot be added to set $\Gamma_{\rho-1}^{(b)}$. It can be demonstrated that at fixed b and arbitrary u and j, all sums $\sum_j^{(b,u)}$ are different. Otherwise, $\Gamma_{\rho-1}^{(b)}$ could not be linearly independent. On the other hand, it is possible that the column equal to $\sum_j^{(b,u)}$ is not involved in the check matrix. Hence, the term $\sum_{j=d-1}^{\rho-1} \binom{\rho-1}{j}$ is the upper estimate, explaining the sign "$\geq$" in (3.3).

In the aforementioned construction, each set Γ_ρ will be repeated ρ-fold. Hence, divisor ρ appears in (3.3). Thus, the number S_ρ of different sets Γ_ρ can be estimated from the formula

$$S_\rho \geq \frac{1}{\rho} S_{\rho-1}\left(n - \rho + 1 - \sum_{j=d-1}^{\rho-1} \binom{\rho-1}{j}\right). \quad (3.3)$$

It is readily seen that

$$-\rho + 1 - \sum_{j=d-1}^{\rho-1} \binom{\rho-1}{j} = -\left(2^{\rho-1} - 1 - \lambda(d,\rho)\right),$$

which leads to (3.1).

The check matrix of the $[2^r - 1, 2^r - 1 - r, 3]$ Hamming code incorporates all nonzero columns of size r. Hence, inequalities are replaced by equality in (3.3) and, accordingly, from which we obtain (3.2). The lemma is proved.

Below, for the $[n, n - r, d]$ code, we employ the function

$$\Phi(n,d,\rho) = \frac{1}{d(d+1)\ldots\rho}\binom{n}{d-1} \times \prod_{j=d}^{\rho}\left(n + 1 - 2^{j-1} + \lambda(d,j)\right), \quad d \leq \rho \leq r, \quad (3.4)$$

the value of which is the lower estimate of the number of erasure configurations of weight ρ correctable by the code under consideration. (For the Hamming code with $d = 3$, this estimate is accurate.)

Theorem 3.1. *The number S_ρ of different sets of ρ linearly independent columns of the check matrix of an $[n, n - r, d]$ code is estimated as follows*:

$$S_\rho \geq \Phi(n,d,\rho), \quad d \leq \rho \leq r. \quad (3.5)$$

In particular, for the $[2^r - 1, 2^r - 1 - r, 3]$ Hamming code, the following equality holds [2, 3, 4]:

$$S_{\rho,r}^H = \frac{1}{\rho!}\prod_{j=1}^{\rho}\left(2^r - 2^{j-1}\right), \quad 3 \leq \rho \leq r. \quad (3.6)$$

Proof. Assuming that $\rho = d - 1$ in (1.2) and iteratively using (3.1), we obtain (3.5). In the case of the Hamming code, (3.2) is iteratively employed.

Remark 3. Let q be a power prime. In [3, lemma, p. 64], it was proved that, in the $r \times (q^r - 1)$ matrix containing all possible nonzero q-ary r-size columns, the number of linearly independent sets from ρ columns is defined as

$$\frac{1}{\rho!}\left(q^r - 1\right)\left(q^r - q\right)\left(q^r - q^2\right)\ldots\left(q^r - q^{\rho-1}\right). \quad (3.7)$$

In [4, Theorem 2.7, proof], it was demonstrated that, in the check matrix of the q-ary $\left[\frac{q^r-1}{q-1}, \frac{q^r-1}{q-1} - r, 3\right]$ Hamming code, the number of linearly independent sets from ρ columns is $\frac{1}{\rho!}\prod_{j=1}^{\rho}\frac{q^r - q^{j-1}}{q-1}$. For $q = 2$, relationship (3.6) is the particular case of the foregoing expressions from [3, 4]. In [2, Lemma 2], expression (3.6) was derived for the binary case.

Let us consider binary *codes with even weights*.

Let us introduce the designation

$$\lambda^{\bullet}(d,\rho) = \begin{cases} 0 & \text{for} \quad d = 4 \\ \sum_{i=2}^{d/2-1}\binom{\rho-1}{2i-1} & \text{for} \quad d \geq 6 \end{cases}.$$

Lemma 3.2. *The number S_ρ of different sets of ρ linearly independent columns of the check matrix of an $[n, n-r, d]$ code is recurrently estimated as follows*:

$$S_\rho \geq \frac{1}{\rho} S_{\rho-1}\left(n - 2^{\rho-2} + \lambda^{\bullet}(d,\rho)\right), \quad d \leq \rho \leq r, \ \text{if all code weights are even.} \tag{3.8}$$

In particular, for the $[2^{r-1}, 2^{r-1}-r, 4]$ Hamming code, the following equality is valid:

$$S^{H\bullet}_{\rho,r} = \frac{1}{\rho} S^{H\bullet}_{\rho-1,r}\left(2^{r-1} - 2^{\rho-2}\right), \quad 4 \leq \rho \leq r. \tag{3.9}$$

Proof. The proof is analogous to the proof of Lemma 3.1. In a check matrix of the code with even weights, the sum of the even number of columns cannot be equal to any check-matrix column. Estimate (3.3) takes the form

$$S_\rho \geq \frac{1}{\rho} S_{\rho-1}\left(n - \rho + 1 \sum_{j=d/2}^{\lceil(\rho-1)/2\rceil} \binom{\rho-1}{2j-1}\right).$$

It is easily seen that

$$-\rho + 1 - \sum_{j=d/2}^{\lceil(\rho-1)/2\rceil} \binom{\rho-1}{2j-1} = -\left(2^{\rho-2} - \lambda^{\bullet}(d,\rho)\right),$$

from which relation (3.8) immediately follows.

The extended Hamming code check matrix contains all possible columns of size r with unity in the upper position. The lemma is proved.

Below, for the $[n, n-r, d]$ code with even weights, we employ the function

$$\Phi^{\bullet}(n,d,\rho) = \frac{1}{d(d+1)\ldots\rho}\binom{n}{d-1} \times \prod_{j=d}^{\rho}\left(n - 2^{j-2} + \lambda^{\bullet}(d,\rho)\right), \tag{3.10}$$

the value of which is the lower estimate of the amount of erasure configurations with weight ρ correctable by the code under consideration. (For the Hamming code with $d = 4$, this estimate is accurate.)

Theorem 3.2. *The number S_ρ of different sets of ρ linearly independent columns of the check matrix of an $[n, n-r, d]$ code is estimated as follows*:

$$S_\rho \geq \Phi^{\bullet}(n,d,\rho), \quad d \leq \rho \leq r, \ \text{if all code weights are even.} \tag{3.11}$$

In particular, for the $[2^{r-1}, 2^{r-1}-r, 4]$ extended Hamming code, the following equality holds [4]:

$$S^{H\bullet}_{\rho,r} = \frac{2^{r-1}}{\rho!}\prod_{j=2}^{\rho}\left(2^{r-1} - 2^{j-2}\right), \quad 4 \leq \rho \leq r. \tag{3.12}$$

Proof. Assuming that $\rho = d - 1$ in (1.2) and iteratively using (3.8), we obtain (3.11). In the case of the Hamming code, (3.9) is iteratively employed.

Remark 4. Relationship (3.12) was first derived in [4, Theorem 2.8, proof].

Corollaries 3.1 and 3.2 follow from Theorems 2.1, 3.1, and 3.2.

Corollary 3.1. *For $d \leq \rho \leq r$, the fraction δ_ρ of erasures with weight ρ correctable by an $[n, n-r, d]$ code is estimated as follows*:

$$\delta_\rho \geq \prod_{j=d}^{\rho} \frac{n + 1 - 2^{j-1} + \lambda(d,j)}{n+1-j} \quad \text{for arbitrary code}, \tag{3.13}$$

$$\delta_\rho \geq \prod_{j=d}^{\rho} \frac{n - 2^{j-2} + \lambda^{\bullet}(d,j)}{n+1-j}, \quad \text{if all code weights are even.} \tag{3.14}$$

In particular, for the $[2^r - 1, 2^r - 1 - r, 3]$ and $[2^{r-1}, 2^{r-1} - r, 4]$ Hamming codes, respectively, the following equalities are valid:

$$\delta^{H}_{\rho,r} = \prod_{j=3}^{\rho} \frac{2^r - 2^{j-1}}{2^r - j}, \tag{3.15}$$

$$\delta^{H\bullet}_{\rho,r} = \prod_{j=4}^{\rho} \frac{2^{r-1} - 2^{j-2}}{2^{r-1} + 1 - j} = \delta^{H}_{\rho-1,r-1}. \tag{3.16}$$

It is seen from (3.13) and (3.14), at the fixed n, the fraction δ_ρ diminishes with growing ρ. In addition, it follows from (3.15) and (3.16), at the fixed r, the fractions $\delta^{H}_{\rho,r}$ and $\delta^{H\bullet}_{\rho,r}$ reduces with increasing ρ.

Corollary 3.2. *For $\rho \leq r$ and $\rho - d > \frac{d-1}{2}$, the number S_ρ of different sets of ρ linearly independent columns of the check matrix of an $[n, n-r, d]$ code that is not nonshortened Hamming code is estimated as follows*:

$$S_\rho \geq \begin{cases} \max\{\Psi(n,d,\rho), \Phi(n,d,\rho)\} \\ \text{for arbitrary code.} \\ \max\{\Psi(n,d,\rho), \Phi^{\bullet}(n,d,\rho)\}, \\ \text{if all code weights are even.} \end{cases} \tag{3.17}$$

From Theorem 2.1 and the proof of Theorems 3.1 and 3.2, it follows that, as a rule, the maximum in formula (3.17) is equal to $\Psi(n, d, \rho)$ because the check matrix of the code incorporates not all possible columns. Moreover, when the inequality

$$n \leq \begin{cases} 2^{\rho-1} - 1 - \lambda(d,\rho) & \text{for arbitrary code} \\ 2^{\rho-2} - \lambda^{\bullet}(d,\rho), & \\ \text{if all code weights are even} & \end{cases},$$

takes place, the inequalities $\Phi(n, d, \rho) \leq 0$ and $\Phi^{\bullet}(n, d, \rho) \leq 0$ are valid.

On the other hand, the weight spectrum is not always known. Only data on relatively small weights are often available. In this case, it is reasonable to calculate function $\Psi(n, d, \rho)$ if possible. Afterward, the last calculated value of function $\Psi(n, d, \rho)$ is used to start the recurrent process based on Lemmas 3.1 and 3.2. The given approach was employed in Corollary 3.3, Statement 2, and Examples 2, 5, and 6.

Corollary 3.3. *For $d \le \rho_0 < \rho \le r$, the number S_ρ of different sets of ρ linearly independent columns of the check matrix of an $[n, n-r, d]$ code that is not nonshortened Hamming code is estimated as follows.*

(i) For arbitrary $[n, n-r, d]$ code, we obtain

$$S_\rho \ge \frac{1}{(\rho_0+1)(\rho_0+2)\dots\rho}\Psi(n,d,\rho_0) \times \prod_{j=\rho_0+1}^{\rho}\left(n+1-2^{j-1}+\lambda(d,j)\right). \quad (3.18)$$

(ii) When all weights of the $[n, n-r, d]$ code are even, we have

$$S_\rho \ge \frac{1}{(\rho_0+1)(\rho_0+2)\dots\rho}\Psi(n,d,\rho_0) \times \prod_{j=\rho_0+1}^{\rho}\left(n-2^{j-2}+\lambda^{\bullet}(d,j)\right). \quad (3.19)$$

Remark 5. In [3], erasure correction by the q-ary cyclic $[n, n-r, d]_q$ code was examined, and the lower estimate of the fraction of correctable erasures, which is based on relationship (3.7), is presented. For binary codes with the distance $d > 3$, this estimate is worse than the estimates obtained in this paper.

4. FRACTION OF ERASURES CORRECTABLE BY SHORTENED, EXTENDED, AND PUNCTURED CODES

In this section, it is shown that the fraction of correctable erasure configurations does not diminish and, in principle, can grow if codes are shortened. In addition, it is demonstrated that, under definite (often fulfilled) conditions, the fraction of correctable configurations remains unchanged if the code is extended. In particular, this is applicable to nonshortened Hamming codes with $d = 3$ and BCH codes with $d = 5$.

4.1. Fraction of Erasures Correctable by Shortened Codes

Theorem 4.1. *Let $\rho \le r$ and the check matrix of a binary linear $[n_0, n_0-r, d]$ code C_{n_0} of length n_0 contain $S_\rho(n_0)$ sets of ρ linearly independent columns. Then, there is a shortened $[n, n-r, d]$ code C_n of length $n < n_0$ with the check matrix incorporating $S_\rho(n)$ sets of ρ linearly independent columns, where*

$$S_\rho(n) \ge S_\rho(n_0)\frac{\binom{n_0-\rho}{n-\rho}}{\binom{n_0}{n}} = S_\rho(n_0)\frac{\binom{n}{\rho}}{\binom{n_0}{\rho}}. \quad (4.1)$$

Proof. Let us perform shortening by excluding the columns from the check matrix H_{n_0} of code C_{n_0}. As a result, we obtain the check matrix H_n of code C_n. Each set of ρ linearly independent columns of nonshortened check matrix H_{n_0} remains invariable in $\binom{n_0-\rho}{n-\rho}$ shortened matrices H_n. Therefore, in all shortened codes C_n, the sum of the sets of ρ linearly independent columns of the check matrix is $S_\rho(n_0)\binom{n_0-\rho}{n-\rho}$. The total number of shortened codes is $\binom{n_0}{n}$. Performing averaging over all shortened codes, we obtain $S_\rho(n) \ge S_\rho(n_0)\frac{\binom{n_0-\rho}{n-\rho}}{\binom{n_0}{n}}$. The final form of relationship (4.1) is derived by means of simple transformations. The theorem is proved.

Corollary 4.1. *Let $\rho \le r$ and the fraction of correctable erasures with weight ρ be $\delta_\rho(n_0)$ for the check matrix of a binary $[n_0, n_0-r, d]$ code C_{n_0} with length n_0. Then, there is a shortened $[n, n-r, d]$ code C_n of length $n < n_0$ with the check matrix ensuring the fraction $\delta_\rho(n)$ of correctable erasures with weight ρ, which is defined as*

$$\delta_\rho(n) \ge \delta_\rho(n_0). \quad (4.2)$$

Proof. It follows from (4.1) that

$$\delta_\rho(n) = \frac{S_\rho(n)}{\binom{n}{\rho}} \ge \frac{S_\rho(n_0)}{\binom{n_0}{\rho}} = \delta_\rho(n_0).$$

Example 3 illustrating theorem 4.1 and corollary 4.1 is presented in Section 5.2.

Below, for $t-(\nu, k, \lambda)$ designes, the system of notation and definitions correspond to [18, Chapter 2].

Theorem 4.2. *Let $d \le \rho \le r$ and words of arbitrary weight w in an $[n_0, n_0-r, d]$ code C_{n_0} of length n_0 generate an $1-(n_0, w, \lambda)$ design. It is assumed that the check matrix of an $[n, n-r, d]$ code C_n with the length $n = n_0 - 1$ was obtained by deleting one column from the check*

matrix of the code C_{n_0}. Then, for all ρ, the following equality is valid:

$$\frac{\Psi(n_0, d, \rho, C_{n_0})}{\binom{n_0}{\rho}} = \frac{\Psi(n, d, \rho, C_n)}{\binom{n}{\rho}}, \quad (4.3)$$

where the left- and right-hand sides of the equality were derived for the codes C_{n_0} and C_n, respectively.

In particular, if $\rho - d \le \frac{d-1}{2}$, the fraction of correctable erasures with weight ρ is identical for codes C_{n_0} and C_n.

Proof. The number of blocks of $1-(n_0, w, \lambda)$ design is equal to the number $A_w(C_{n_0})$ of words with weight w in code C_{n_0}. Parameter λ is equal to the number of words with weight w coupled with each column of the check matrix of code C_{n_0}. When some of the columns is removed, these words "are broken down" and only $A_w(C_{n_0}) = A_w(C_{n_0}) - \lambda$ words of weight w are preserved in code C_n. In accordance with [18, Chapter 2, Corollary 10], we obtain $\lambda = wA_w(C_{n_0})/n_0$. Therefore, $A_w(C_n) = A_w(C_{n_0})(n_0 - w)/n_0$, and relationship (4.3) can be deduced from (2.1) by means of simple transformations. The final statement of the theorem follows from (2.3) and (4.3).

We note that words with any weight w generate $t-(n_0, w, \lambda)$ designs with $t \ge 1$, in the Hamming, Panchenko, and BCH codes discussed below.

4.2. Fraction of Erasures Correctable by Extended and Punctured Codes

Statement 1. *Let the check matrix H_0 of a binary $[n, n-r, 2t+1]$ code C_0 with odd code distance contain $S_\rho(H_0)$ sets from ρ ($1 \le \rho \le r$) linearly independent columns. It is assumed that an $[n+1, n-r, 2t+2]$ code C is constructed from the code C_0 by supplement of overall parity check symbol. Then, the check matrix H of the code C incorporates $S_\rho(H)$ sets from ρ linearly independent columns. In this case*,

$$S_\rho(H) \ge S_{\rho-1}(H_0) + S_\rho(H_0), \quad 2 \le \rho \le r. \quad (4.4)$$

Proof. The check matrix H of code C can be obtained if the check matrix H_0 of code C_0 is supplemented by the upper row composed of unities and column $(10\ldots0)^T$. All $S_\rho(H_0)$ sets involving ρ linearly independent columns of matrix H_0 are also linearly independent in matrix H. Introduction of column $(10\ldots0)^T$ into each of the $S_{\rho-1}(H_0)$ sets from $\rho-1$ linearly independent columns, provides the set from ρ linearly independent columns of matrix H. The sign "≥" in (4.4) is explained by the fact that the $r \times \rho$ submatrices of rank $\rho-1$, the rank of which is increased to ρ upon adding the upper row composed of unities to them, can exist in matrix H_0. The statement is proved.

We note that $r \times \rho$ submatrices of rank $\rho-1$ mentioned in the proof of Statement 1 exists, leading to the sign ">" in (4.4)

Let $A_w^{(0)}$ and A_w designate the number of words in codes C_0 and C of Statement 1, respectively. As is known,

$$A_{2j} = A_{2j-1}^{(0)} + A_{2j}^{(0)}. \quad (4.5)$$

Equality (4.5) and inequality (4.4) emphasize distinctions in approaches and estimates related to the weight spectrum and the amount of the sets from linearly independent columns.

In addition, it should be noted that code puncturing is the process inverse to its extension.

Definition. ([18, Section 8.5]). A code C possesses *property P* if the removal of the fixed coordinate from each codeword of the code C provides the punctured code C^* with the same weight spectrum that is independent of the punctured coordinate.

Lemma 4.1. ([18, Section 8.5, Theorem 8.14]). *Let an $[n, n-r, 2t+2]$ code C, all code words of which have an even weight, possess property P. Then, the code C^* obtained via puncturing of a certain coordinate of the code C is an $[n-1, n-r, 2t+1]$ code. In this case, it holds that*

$$A_{2j-1}^* = \frac{2j}{n} A_{2j}, \quad A_{2j}^* = \frac{n-2j}{n} A_{2j}, \quad j = t+1, \ t+2, \ldots, \quad (4.6)$$

where A_w and A_w^* are the number of words with weight w in the codes C and C^*, respectively.

We note that the equality $A_{2j} = A_{2j-1}^* + A_{2j}^*$ resembling the relationship (4.5) follows from (4.6). At the same time, the authors of [18] indicated that the code extended by the parity check does not always possesses property P and provided the following sufficient condition ([18, Section 8.5, Corollary 15]): *the code invariant with respect to the transitive group of substitutions possesses property P.* Many extended codes have transitive groups of substitutions [18]. In particular (see, e.g., [17]), the nonshortened extended BCH code is twice transitive.

Theorem 4.3. *Let an $[n, n-r, 2t+2]$ code C whose code words have an even weight possess property P and the $[n-1, n-r, 2t+1]$ code C^* is constructed via puncturing of the some coordinate of code C. Then, for ρ = $2t+2$, $2t+3$, $2t+4$, and $2t+5$, it is valid that*

$$\frac{\Psi(n, 2t+2, \rho, C)}{\binom{n}{\rho}} = \frac{\Psi(n-1, 2t+1, \rho, C^*)}{\binom{n-1}{\rho-1}}, \quad (4.7)$$

where the left- and right-hand sides of the equality were obtained for the codes C and C^*, respectively. In particular, when $\rho \in \{2t+2, 2t+3, 2t+4, 2t+5\}$, we have $\delta_\rho(C) = \delta_{\rho-1}(C^*)$, where $\delta_\rho(C)$ and $\delta_{\rho-1}(C^*)$ are the

fractions of erasures with weights ρ *and* $\rho - 1$ *correctable by C and C* codes, respectively, under the condition that* $\rho - (2t + 2) \le \frac{2t+1}{2}$.

Proof. We employ relationship (4.6), substitute the corresponding weights into (2.1) with allowance for $A_{2t+3} = A_{2t+5} = 0$ for code C, and perform simple transformations.

It can be assumed that relationship (4.7) and the equality $\delta_\rho(C) = \delta_{\rho-1}(C^*)$ in Theorem 4.3, which is a consequence of the former, hold for all admissible values of ρ.

5. FRACTION OF ERASURES CORRECTABLE BY NONSHORTENED HAMMING, PANCHENKO, AND BCH CODES

5.1. Nonshortened [$2^r - 1$, $2^r - 1 - r$, 3] and [2^{r-1}, $2^{r-1} - r$, 4] Hamming Codes

Example 1. For the nonshortened extended [2^{r-1}, $2^{r-1} - r$, 4] Hamming code, the dependences between the fraction $\delta^{H\bullet}_{\rho,r}$ of correctable erasures with weight ρ and r (see (3.17)), which were calculated at $4 \le \rho \le 8$, $\rho \le r$, and $7 \le r \le 18$, are depicted in Fig. 1. We remind that $\delta^{H\bullet}_{\rho,r} = \delta^{H}_{\rho-1,r-1}$. Therefore, the presented curves are the graphs of the fraction $\delta^{H}_{\rho-1,r-1}$ of correctable erasures with weight $\rho - 1$ as a function of $r - 1$ in the nonshortened [$2^{r-1} - 1$, $2^{r-1} - r$, 3] Hamming code.

The numerical values of $\delta^{H\bullet}_{\rho,r}$ are presented in Table 1.

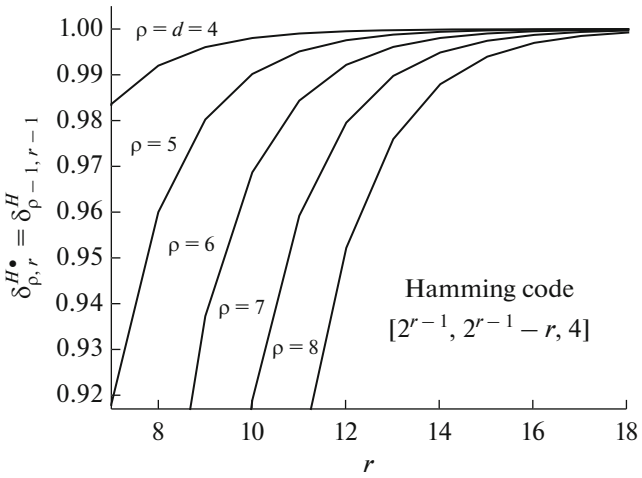

Fig. 1. Variations in the fraction $\delta^{H\bullet}_{\rho,r}$ of correctable erasures with weight ρ vs. r for the nonshortened [2^{r-1}, $2^{r-1} - r$, 4] Hamming code (or, equivalently, variations in the fraction $\delta^{H}_{\rho-1,r-1}$ of correctable erasures with weight $\rho - 1$ vs. $r - 1$ for the nonshortened [$2^{r-1} - 1$, $2^{r-1} - r$, 3] Hamming code) under the condition that $4 \le \rho \le 8$, $\rho \le r$, and $7 \le r \le 18$.

5.2. Nonshortened [$5 \times 2^{r-4}$, $5 \times 2^{r-4} - r$, 4] Panchenko Code with the Distance $d = 4$

Below, we consider the binary [$5 \times 2^{r-4}$, $5 \times 2^{r-4} - r$, 4] Panchenko code suggested by V.I. Panchenko in [13]. This code was analyzed in [7, 12]. For the given code, let us introduce the following system of notation: $A^{\Pi}_{w,r}$ is the number of words with weight w, $S^{\Pi}_{\rho,r}$ is the number of the correctable erasure configurations

Table 1. Fraction $\delta^{H\bullet}_{\rho,r}$ of correctable erasures with weight ρ as a function of r for the nonshortened [2^{r-1}, $2^{r-1} - r$, 4] Hamming code (or, equivalently, the fraction $\delta^{H}_{\rho-1,r-1}$ of correctable erasures with weight $\rho - 1$ as a function of $r - 1$ for the nonshortened [$2^{r-1} - 1$, $2^{r-1} - r$, 3] Hamming code) at $4 \le \rho \le 12$, $\rho \le r$, and $7 \le r \le 20$

r	$\rho = d = 4$	$\rho = 5$	$\rho = 6$	$\rho = 7$	$\rho = 8$	$\rho = 9$	$\rho = 10$	$\rho = 11$	$\rho = 12$
7	0.9836	0.9180	0.7469	0.4121					
8	0.9920	0.9600	0.8741	0.6879	0.3638				
9	0.9960	0.9802	0.9373	0.8398	0.6476	0.3342			
10	0.9980	0.9902	0.9687	0.9189	0.8152	0.6211	0.3161		
11	0.9990	0.9951	0.9844	0.9592	0.9055	0.7985	0.6042	0.3051	
12	0.9995	0.9976	0.9922	0.9796	0.9522	0.8962	0.7876	0.5936	0.2984
13	0.9998	0.9988	0.9961	0.9898	0.9760	0.9473	0.8901	0.7807	0.5871
14	0.9999	0.9994	0.9980	0.9949	0.9879	0.9735	0.9441	0.8862	0.7764
15	0.9999	0.9997	0.9990	0.9974	0.9940	0.9867	0.9718	0.9420	0.8837
16	1.0000	0.9998	0.9995	0.9987	0.9970	0.9933	0.9858	0.9707	0.9407
17	1.0000	0.9999	0.9998	0.9994	0.9985	0.9967	0.9929	0.9853	0.9701
18	1.0000	1.0000	0.9999	0.9997	0.9992	0.9983	0.9964	0.9926	0.9850
19	1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9982	0.9963	0.9925
20	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996	0.9991	0.9982	0.9962

with weight ρ, and $\delta^{\Pi}_{\rho,r}$ is the number of correctable erasures of weight ρ. In conformity with (1.1), we have

$$\delta^{\Pi}_{\rho,r} = \frac{S^{\Pi}_{\rho,r}}{\binom{5\times 2^{r-4}}{\rho}} \le 1. \quad (5.1)$$

It is known [7, 12] that

$$A^{\Pi}_{4,r} = \frac{5\times 2^{r-6}\left(2^{r-4}-1\right)\left(2^{r-2}+5\times 2^{r-5}-1\right)}{3}, \quad (5.2)$$

$$A^{\Pi}_{5,r} = 2^{4r-16}. \quad (5.3)$$

Statement 2. *For the* $[5\times 2^{r-4}, 5\times 2^{r-4}-r, 4]$ *Panchenko code, the number* $S^{\Pi}_{\rho,r}$ *of the correctable erasure configurations with weight* ρ *is estimated as*

$$S^{\Pi}_{4,r} = \binom{5\times 2^{r-4}}{4} - \frac{5\times 2^{r-6}\left(2^{r-4}-1\right)\left(2^{r-2}+5\times 2^{r-5}-1\right)}{3}, \quad (5.4)$$

$$S^{\Pi}_{5,r} = \binom{5\times 2^{r-4}}{5} - \frac{3\times 2^{4r-16}+5\times 2^{r-6}\left(2^{r-4}-1\right)\left(2^{r-2}+5\times 2^{r-5}-1\right)\left(5\times 2^{r-4}-4\right)}{3}, \quad (5.5)$$

$$S^{\Pi}_{\rho,r} \ge \frac{1}{6\times 7\times\ldots\times\rho} S^{\Pi}_{5,r} \times \prod_{j=6}^{\rho}\left(5\times 2^{r-4}+1-2^{j-1}+\lambda(4,j)\right), \quad 6\le\rho\le r. \quad (5.6)$$

Proof. In the cases of ρ = 4 and 5, condition (2.4) is fulfilled. With the use of (5.2) and (5.3), formulas (2.1) and (2.3) provide (5.4) and (5.5). Relationship (5.6) is derived using (3.18) with $\rho_0 = 5$ and $\Psi(n, d, \rho_0) = S^{\Pi}_{5,r}$.

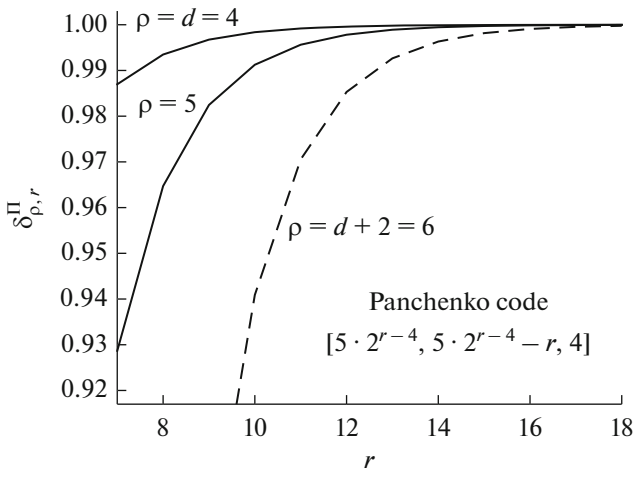

Fig. 2. Variations in the fraction $\delta^{\Pi}_{\rho,r}$ of correctable erasures with weight ρ vs. r for the nonshortened $[5\times 2^{r-4}, 5\times 2^{r-4}-r, 4]$ Panchenko code under the condition that $7\le r\le 18$ and ρ = 4, 5, and 6. The continuous lines are the exact values of $\delta^{\Pi}_{\rho,r}$, and the dashed curve designates the lower estimate of $\delta^{\Pi}_{6,r}$.

Example 2. For the nonshortened $[5\times 2^{r-4}, 5\times 2^{r-4}-r, 4]$ Panchenko code, the dependences between the fraction $\delta^{\Pi}_{\rho,r}$ of correctable erasures with weight ρ and r, which were obtained at ρ = 4, 5, and 6 and $7\le r\le 18$, are presented in Fig. 2. The exact values of $\delta^{\Pi}_{4,r}$ and $\delta^{\Pi}_{5,r}$ were determined according to formulas (5.1), (5.4), and (5.5). The lower estimate of quantity $\delta^{\Pi}_{6,r}$ follows from (5.1) and (5.6).

The numerical values of $\delta^{\Pi}_{\rho,r}$ are summarized in Table 2, where the exact values of $\delta^{\Pi}_{\rho,r}$ correspond to ρ = 4 and 5 and the lower estimate is designated by ρ = 6.

Example 3. In the case of the nonshortened [80, 72, 4] Panchenko code from (5.1), (5.4), and (5.5), we have $\delta^{\Pi}_{4,8} = 0.993488$ and $\delta^{\Pi}_{5,8} = 0.964712$. Let the [76, 64, 8] Panchenko code is shortened by eight symbols according to Algorithm 1 from [12]. Then, with allowance for the algorithm modification in [7], it is valid that $A_4 = 6654$ and $A_5 = 38586$ [7, Table 2]. Hence, using formulas (2.1) and (5.1), we arrive at $\delta^{\Pi(8)}_{4,8} = 0.993532 > \delta^{\Pi}_{4,8}$ and $\delta^{\Pi(8)}_{5,8} = 0.964903 > \delta^{\Pi}_{5,8}$, where $\delta^{\Pi(v)}_{\rho,r}$ designates the fraction of erasures of weight ρ correctable by the $[5\times 2^{r-4}-v, 5\times 2^{r-4}-v-r, 4]$ Panchenko code, which is shortened by v symbols. The aforesaid illustrates Theorem 4.1 and Corollary 4.1.

5.3. $[2^{(r-1)/2}-1, 2^{(r-1)/2}-r, 5]$ and $[2^{(r-1)/2}, 2^{(r-1)/2}-r, 6]$ Nonshortened BCH Codes

Let $B^{\bullet}$ designate the binary extended $[2^{(r-1)/2}, 2^{(r-1)/2}-r, 6]$ BCH code with even weights. For the given code, it is possible to introduce the following

Table 2. Fraction $\delta_{\rho,r}^{\Pi}$ of correctable erasures with weight ρ as a function of r for the nonshortened $[5 \times 2^{r-4}, 5 \times 2^{r-4} - r, 4]$ Panchenko code under the condition that $7 \le r \le 18$ and $\rho = 4$, 5, and 6

r	7	8	9	10	11	12
n	40	80	160	320	640	1280
$\rho = d = 4$	0.9870	0.9935	0.9967	0.9984	0.9992	0.9996
$\rho = d + 1 = 5$	0.9287	0.9647	0.9825	0.9913	0.9956	0.9978
$\rho = d + 2 = 6 \ge$	0.5041	0.7589	0.8810	0.9409	0.9705	0.9853
r	13	14	15	16	17	18
n	2560	5120	10240	20480	40960	81920
$\rho = d = 4$	0.9998	0.9999	0.9999	1.0000	1.0000	1.0000
$\rho = d + 1 = 5$	0.9989	0.9995	0.9997	0.9999	0.9999	1.0000
$\rho = d + 2 = 6 \ge$	0.9927	0.9963	0.9982	0.9991	0.9995	0.9998

system of notation: $A_{w,r}^{B^\bullet}$ is the number of words with weight w, $S_{\rho,r}^{B^\bullet}$ is the number of the correctable erasure configurations with weight ρ, and $\delta_{\rho,r}^{B^\bullet}$ is the fraction of correctable erasures of weight ρ. In compliance with (1.1), we have

$$\delta_{\rho,r}^{B^\bullet} = \frac{S_{\rho,r}^{B^\bullet}}{\binom{2^{(r-1)/2}}{\rho}} \le 1. \tag{5.7}$$

It is known [16, p. 434] that

$$A_{6,r}^{B^\bullet} = \begin{cases} \dfrac{2^{(r-1)/2}\left(2^{(r-1)/2}-1\right)\left(2^{(r-1)/2}-2\right)\left(2^{(r-1)/2}-8\right)}{720}, \\ \text{if } (r-1)/2 \text{ is odd} \\ \dfrac{2^{(r-1)/2}\left(2^{(r-1)/2}-1\right)\left(2^{(r-1)/2}-4\right)^2}{720}, \\ \text{if } (r-1)/2 \text{ is even.} \end{cases} \tag{5.8}$$

Let B designate the $[2^{(r-1)/2} - 1, 2^{(r-1)/2} - r, 5]$ binary BCH code obtained via puncturing of the certain position of code $B^\bullet$. Let $\delta_{\rho,r-1}^{B}$ be the fraction of erasures with weight ρ correctable by code B. The code $B^\bullet$ is twice transitive (see, e.g., [17]) and, consequently, possesses property P from the definition in Section 4.2. Hence, for $\rho = 6$, 7, 8, and 9, Theorem 4.3 provides the equality

$$\frac{\Psi\left(2^{(r-1)/2}, 6, \rho, B^\bullet\right)}{\binom{2^{(r-1)/2}}{\rho}} = \frac{\Psi\left(2^{(r-1)/2} - 1.5, \rho - 1, B\right)}{\binom{2^{(r-1)/2}-1}{\rho-1}}, \tag{5.9}$$

where the left- and right-hand sides of the equality were deduced for codes $B^\bullet$ and B, respectively. Moreover, if condition (2.4) is taken into account, we have

$$\delta_{\rho,r}^{B^\bullet} = \delta_{\rho-1,r-1}^{B}, \quad \rho = 6, 7, 8. \tag{5.10}$$

Using (2.1), (2.3), (2.4), (5.7), (5.8), and (5.10) and allowing for the evenness of all weights in code $B^\bullet$, we obtain

Statement 3. *The fraction* $\delta_{\rho,r}^{B^\bullet} = \delta_{\rho-1,r-1}^{B}$ *of erasures with weights* ρ *and* $\rho - 1$ *correctable by* $[2^{(r-1)/2}, 2^{(r-1)/2} - r, 6]$ *and* $[2^{(r-1)/2} - 1, 2^{(r-1)/2} - r, 5]$ *BCH codes is defined as*

$$\delta_{6,r}^{B^\bullet} = \delta_{5,r-1}^{B} = 1 - \begin{cases} \dfrac{\left(2^{(r-1)/2}-8\right)}{\left(2^{(r-1)/2}-3\right)\left(2^{(r-1)/2}-4\right)\left(2^{(r-1)/2}-5\right)}, \\ \text{if } (r-1)/2 \textit{ is odd} \\ \dfrac{\left(2^{(r-1)/2}-4\right)}{\left(2^{(r-1)/2}-2\right)\left(2^{(r-1)/2}-3\right)\left(2^{(r-1)/2}-5\right)}, \\ \text{if } (r-1)/2 \textit{ is even.} \end{cases} \tag{5.11}$$

$$\delta_{7,r}^{B^\bullet} = \delta_{6,r-1}^{B} = \delta_{6,r}^{B^\bullet} - 6\left(1 - \delta_{6,r}^{B^\bullet}\right). \tag{5.12}$$

We note that code $B^\bullet$ can be interpreted as extended code B.

Example 4. For the $[2^{(r-1)/2}, 2^{(r-1)/2} - r, 6]$ nonshortened extended BCH code $B^\bullet$, the dependences between the fraction $\delta_{\rho,r}^{B^\bullet}$ of correctable erasures with weight ρ and r (see (2.1), (2.3), (2.4), (5.7), (5.11), and (5.12)), which were determined at $\rho = 6$, 7, 8, and 9 and $r = 13$, 15, and 17, are depicted in Fig. 3. In the cases of $\rho = 6$, 7, and 8, condition (2.4) holds. At $\rho = 9$, this condition is not fulfilled and the corresponding

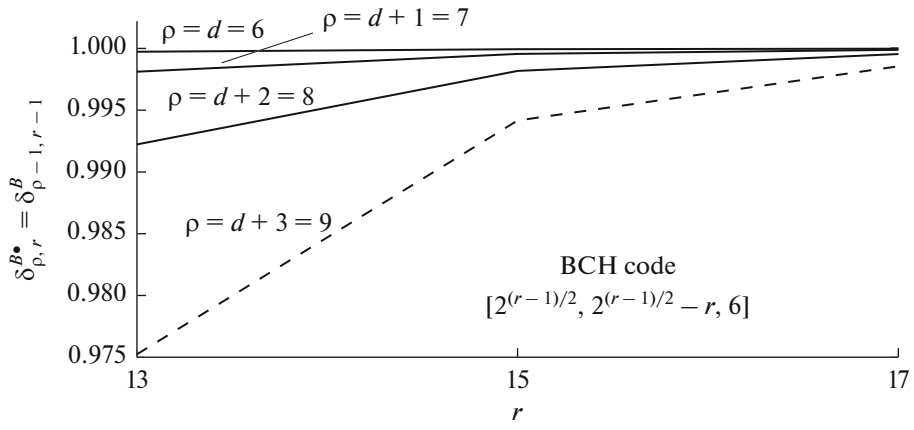

Fig. 3. Variations in the fraction $\delta^{B\bullet}_{\rho,r}$ of correctable erasures with weight ρ vs. r for the nonshortened extended $[2^{(r-1)/2}, 2^{(r-1)/2} - r, 6]$ BCH code (or, equivalently, the fraction $\delta^{B}_{\rho-1,r-1}$ of correctable erasures with weight $\rho - 1$ vs. $r - 1$ for the nonshortened $[2^{(r-1)/2} - 1, 2^{(r-1)/2} - r, 5]$ BCH code) under the condition that ρ = 6, 7, 8, and 9 and r = 13, 15, and 17. The continuous lines are the exact values of $\delta^{B\bullet}_{\rho,r} = \delta^{B}_{\rho-1,r-1}$, and the dashed curve designates the lower estimate of $\delta^{B\bullet}_{9,r} = \delta^{B\bullet}_{8,r-1}$.

graph is the lower estimate of fraction $\delta^{B\bullet}_{9,r}$. In calculations based on ρ = 8 and 9, quantities $A^{B\bullet}_{8,r}$ were taken from [19, Appendix A] and [21]. If (5.9) and (5.10) are taken into account, the aforementioned curves are the graphs of fraction $\delta^{B}_{\rho-1,r-1}$ of correctable erasures with weight $\rho - 1$ as a function of $r - 1$ for the $[2^{(r-1)/2} - 1, 2^{(r-1)/2} - r, 5]$ code B.

The numerical values of $\delta^{B\bullet}_{\rho,r}$ are presented in Table 3, where the exact values of $\delta^{B\bullet}_{\rho,r}$ correspond to ρ = 6, 7, and 8 and the lower estimate is designated by ρ = 9. For r = 19, only the values of $\delta^{B\bullet}_{6,19}$ and $\delta^{B\bullet}_{7,19}$ were calculated.

Table 3. Fraction $\delta^{B\bullet}_{\rho,r}$ of correctable erasures with weight ρ as a function of r for the nonshortened extended $[2^{(r-1)/2}, 2^{(r-1)/2} - r, 6]$ BCH code (or, equivalently, the fraction $\delta^{B}_{\rho-1,r-1}$ of correctable erasures with weight $\rho - 1$ as a function of $r - 1$ for the nonshortened $[2^{(r-1)/2} - 1, 2^{(r-1)/2} - r, 5]$ BCH code) under the condition that ρ = 6, 7, 8, and 9 and r = 13, 15, 17, and 19

r	13	15	17	19
n	64	128	256	512
$\rho = d = 6$	0.9997	0.9999	1.0000	1.0000
$\rho = d + 1 = 7$	0.9981	0.9996	0.9999	1.0000
$\rho = d + 2 = 8$	0.9922	0.9982	0.9995	
$\rho = d + 3 = 9 \geq$	0.9752	0.9942	0.9986	

6. ERASURE CORRECTION WITH DETECTION OF ERRORS

In this section, the case where ρ erasures and ν errors take place ($d \leq \rho < r$ and $\nu > 0$) is considered for an $[n, r, d]$ code. Let us introduce the following set of notation:

(i) $\mathbf{A}^{(\nu)}_{\rho}$ is the $r \times (\rho + \nu)$ submatrix of the check matrix corresponding to erasures and errors.

(ii) $\mathbf{A}_{\rho}$ is the $r \times \rho$ submatrix of the check matrix corresponding only to erasures.

(iii) $\mathcal{L}(\mathbf{A}_{\rho})$ is the overdetermined system of r linear equations with ρ unknowns, the matrix of which is $\mathbf{A}_{\rho}$ and the column of free elements is the column of syndrome.

(iv) $\Delta_{\nu,\rho}$ is the fraction of errors with multiplicity ν that are detected upon attempts to correct ρ erasures corresponding to $r \times \rho$ matrix $\mathbf{A}_{\rho}$ of rank ρ.

It should be noted that matrix $\mathbf{A}_{\rho}$ is known to a decoder and is the submatrix of matrix $\mathbf{A}^{(\nu)}_{\rho}$. The syndrome is the sum of the columns of matrix $\mathbf{A}^{(\nu)}_{\rho}$. In decoding, system $\mathcal{L}(\mathbf{A}_{\rho})$ is solved. The decoding process is not fulfilled if the system $\mathcal{L}(\mathbf{A}_{\rho})$ cannot be solved, i.e., in two cases:

(i) The rank of matrix $\mathbf{A}_{\rho}$ is less than ρ (system $\mathcal{L}(\mathbf{A}_{\rho})$ is degenerate).

(ii) The rank of matrix $\mathbf{A}_{\rho}$ is ρ (or, equivalently, erasures correspond to the ρ linearly independent columns of the check matrix), but the system $\mathcal{L}(\mathbf{A}_{\rho})$ is inconsistent.

Case (ii) is interpreted as error detection during an attempt of correction of ρ erasures corresponding to

the ρ linearly independent columns of the check matrix.

Lemma 6.1. *If* ρ *erasures and* ν *errors occur* ($d \le \rho < r$ *and* $\nu > 0$) *and the erasures correspond to* ρ *linearly independent columns of the check matrix of an* $[n, r, d]$ *code, errors are detected if and only if the rank of the* $r \times (\rho + \nu)$ *matrix* $\mathbf{A}_\rho^{(\nu)}$ *corresponding to the erasures and errors is greater than or equal to* $\rho + 1$.

Proof. If the rank of matrix $\mathbf{A}_\rho^{(\nu)}$ is less than $\rho + 1$, this matrix contains column h not belonging to submatrix $\mathbf{A}_\rho$ and linearly independent of other columns of the given submatrix. The column h is one of the syndrome summands, i.e., the column of free elements of system $\mathcal{L}(\mathbf{A}_\rho)$. Hence, the rank of the extended matrix of system $\mathcal{L}(\mathbf{A}_\rho)$ is greater than or equal to $\rho + 1$ and is not equal to the rank of the system's basic matrix, signifying its inconsistency. If the rank of matrix $\mathbf{A}_\rho^{(\nu)}$ is ρ, such column h does not exist and the system turns out to be consistent.

Let $\lambda(d, \rho)$ and $\lambda^\bullet(d, \rho)$ are defined as in Section 3.

Theorem 6.1. *If* ρ *erasures and* ν *errors take place* ($d \le \rho < r$ *and* $\nu > 0$) *and the erasures correspond to the* ρ *linearly independent columns of the check matrix of an* $[n, r, d]$ *code, the fraction* $\Delta_{\nu,\rho}$ *of detectable errors is estimated as follows*:

$$\Delta_{\nu,\rho} \ge 1 - \frac{\binom{2^\rho - \rho - 1 - \lambda(d, \rho+1)}{\nu}}{\binom{n-\rho}{\nu}} \tag{6.1}$$

for an arbitrary $[n, r, d]$ *code*,

$$\Delta_{\nu,\rho} = 1 - \frac{\binom{2^\rho - \rho - 1}{\nu}}{\binom{2^r - \rho - 1}{\nu}} \tag{6.2}$$

for the $\left[2^r - 1,\ 2^r - 1 - r,\ 3\right]$ *Hamming code*,

$$\Delta_{\nu,\rho} \ge 1 - \frac{\binom{2^{\rho-1} - \rho - \lambda^\bullet(d, \rho+1)}{\nu}}{\binom{n-\rho}{\nu}}, \tag{6.3}$$

if all weights of $[n, r, d]$ *code are even*,

$$\Delta_{\nu,\rho} = 1 - \frac{\binom{2^{\rho-1} - \rho}{\nu}}{\binom{2^{r-1} - \rho}{\nu}} \tag{6.4}$$

for the $\left[2^{r-1},\ 2^{r-1} - r,\ 4\right]$ *Hamming code.*

Proof. By analogy with the proof of Lemma 3.1, it is possible to demonstrate that, in the case of matrix $\mathbf{A}_\rho$ with rank ρ, the remaining part of the check matrix comprises the set of no less than $m = n + 1 - 2^\rho + \lambda(d, \rho + 1)$ columns whose substitution into $\mathbf{A}_\rho$ provides the matrix of rank $\rho + 1$. Let T designate the given set. To found matrix $\mathbf{A}_\rho^{(\nu)}$, matrix $\mathbf{A}_\rho$ must be added with some $r \times \nu$ matrix $\mathbf{D}$. The matrix $\mathbf{A}_\rho^{(\nu)}$ has rank ρ (i.e., errors are not detected) if and only if the matrix $\mathbf{D}$ has no columns from T. The number of such matrices $\mathbf{D}$ is $\binom{n-\rho-m}{\nu}$. The total number of matrices $\mathbf{D}$ is $\binom{n-\rho}{\nu}$. Therefore, there are no less than $\binom{n-\rho}{\nu} - \binom{n-\rho-m}{\nu}$ matrices $\mathbf{A}_\rho^{(\nu)}$ with rank $\ge \rho + 1$. On the other hand, the $\binom{n-\rho}{\nu}$ errors of multiplicity ν exists beyond ρ erasures. Dividing the number of "good" matrices $\mathbf{A}_\rho^{(\nu)}$ by $\binom{n-\rho}{\nu}$, we obtain (6.1). Relations (6.2)–(6.4) can be proved similarly with the help of approaches used to prove Lemmas 3.1 and 3.2.

7. ALGORITHMS FOR EXTENDED DECODING OF BINARY PRODUCT CODES

Let us consider decoding procedure for a product code with identical $[n, n - r, d]$ component codes on its rows and columns. The code word is the $n \times n$ matrix. It is assumed that $d = 4$ or 6.

Let us introduce the following system of notation: H is the $r \times n$ parity check matrix of a binary component code, U is the $n \times n$ matrix of the received word with errors, S_{row} is the $n \times r$ matrix of row syndromes, and S_{col} is the $r \times n$ matrix of column syndromes.

Algorithm 1.

(i) Calculation of the syndromes $S_{\text{row}} = UH^T$ and $S_{\text{col}} = HU$.

(ii) Formation of the lists L_{row} and L_{col} of row and column numbers with detected errors.

(iii) Testing whether submatrices $H(L_{\text{row}})$ and $H(L_{\text{col}})$ are nondegenerate. If at least one of them is nondegenerate, erasures are corrected (with the help of the minimum-size submatrix at the intersection of two lists) and the corrected codeword is produced. Otherwise, the <failure> is formed.

Algorithm 2.

(i) Algorithm 1 is carried out. In the case of <failure>, step (ii).

(ii) List L_{row} with correction of one error (or up to two for the BCH code) is inspected, and the renewed

list L_{row} with corrections kept in matrix U_{res} is complied.

(iii) The syndrome $S_{col} = HU_{res}$ is calculated, and renewed list L_{col} is prepared.

(iv) For the renewed L_{col}, it is revealed whether submatrix $H(L_{col})$ is nondegenerate. If the submatrix is nondegenerate, the erasures are corrected at the intersection of renewed lists and the corrected codeword is generated. Otherwise, <failure> or step (v).

(v) List L_{col} with correction of one error (or up to two for the BCH code) is examined, and the renewed list L_{col} with corrections preserved in matrix U_{res} is complied.

(vi) The syndrome $S_{row} = HU_{res}$ is calculated, and the renewed list L_{row} is prepared.

(vii) For the renewed L_{row}, it is revealed whether submatrix $H(L_{row})$ is nondegenerate. If the submatrix is nondegenerate, the erasures are corrected at the intersection of renewed lists and the corrected codeword is produced. Otherwise, <failure> (or step (iii) until a double failure is fulfilled in steps (iv) and (vii)).

Comment. In the proposed algorithms, rows and column are independently decoded up to the final step (i.e., step (iii) in Algorithm 1 and steps (iv) and (vii) in Algorithm 2). Owing to such a construction of the algorithms, preliminary decoding is performed independently and parallel over rows and columns and, consequently, the delay and complexity of decoding is decreased. However, in this case, the set of decodable error configuration and the correct decoding probability diminish to some extent as compared to the best attainable result. It is significant that the diminution accumulates only improbable events and the main term is preserved.

8. ESTIMATING THE PROBABILITY OF SUCCESSFUL DECODING

The probability of successful decoding is proposed to estimate for somewhat simplified general scheme. The simplification of the decoding algorithm and calculation scheme consists in that the preliminary decoding of rows and columns is interpreted as independent events. Since rows and columns are independently decoded, the complexity and delay of decoding of a product code can be reduced, while independent calculations provide the estimated main term of the probability of successful decoding.

Let Ω designate the probability of the event, namely, successful decoding of rows and, independently, successful decoding of columns for the symmetric structure of a product code. Therefore, the probability of successful decoding of the product code can be estimated (from below) as $1 - (1 - \Omega)^2$ because the fault probability during column decoding performed after the failure arising from row decoding can differ from $1 - \Omega$.

Let p designate the bit-error probability at the decoder input. Evidently, we have

(i) The probability that a row (column) has no errors is

$$P_0 = P(0) = (1-p)^n.$$

(ii) The probability that a row (column) has at least one error is

$$P_1 = P(>0) = 1 - (1-p)^n.$$

(iii) The probability that a row (column) has more than one error is

$$P_2 = P(>1) = 1 - (1-p)^n - np(1-p)^{n-1}.$$

(iv) The probability that a row (column) has more than two errors is

$$P_3 = P(>2) = 1 - (1-p)^n - np(1-p)^{n-1} - \binom{n}{2} p^2 (1-p)^{n-2}.$$

Let us introduce the following system of notation: d^+ is the threshold for the extended correction of erasures, $d^+ \geq d$, and δ_ρ is fraction of erasures with weight ρ correctable by a component code.

8.1. Product of Codes with Distance 4

The probability that just ρ rows (columns) contain more than one error is defined by

$$\binom{n}{\rho} (P_2)^\rho (1-P_2)^{n-p}.$$

The probability of successful decoding of rows (columns) is expressed as

$$\Omega_2 = \sum_{\rho=0}^{d+} \binom{n}{\rho} (P_2)^\rho (1-P_2)^{n-\rho} \delta_\rho. \quad (8.1)$$

For product of codes with distance 4, the probability of failure is defined by

$$(1-\Omega_2)^2. \quad (8.2)$$

The probability of successful decoding for product of codes with distance 4 is expressed as

$$P_{prod,4} = 1 - (1-\Omega_2)^2.$$

Example 5. For the product of shortened [72, 64, 4] and [137, 128, 4] Panchenko codes, the fault probabilities $(1 - \Omega_2)^2$ are presented in Tables 4 and 5, respectively. In this case, the input probabilities are $p = 10^{-1}$, 10^{-2}, 5×10^{-3}, 10^{-3}, 5×10^{-4}, and 10^{-4}, quantities $d^+ = 3$, 4, 5, and 6, and entry e-m designates 10^{-m}.

In [7, Table 2], we constructed the Panchenko code with the weights $A_4 = 6654$, $A_5 = 38586$, and $A_6 = 695799$. Using the given weights and formulas (2.1)–

Table 4. Fault probabilities of the product of [72, 64, 4] Panchenko codes

p	10^{-1}	10^{-2}	5×10^{-3}	10^{-3}	5×10^{-4}	10^{-4}
$d^+ = 3$	1	0.996	0.250	1.1e-09	2.3e-14	1.9e-25
$d^+ = 4$	1	0.988	0.092	1.6e-12	5.1e-18	1.1e-31
$d^+ = 5$	1	0.967	0.027	7.0e-14	1.045e-18	4.931e-32
$d^+ = 6$	1	0.926	0.008	5.8e-14	1.029e-18	4.931e-32

Table 5. Fault probabilities of the product of [137, 128, 4] Panchenko codes

p	10^{-1}	10^{-2}	5×10^{-3}	10^{-3}	5×10^{-4}	10^{-4}
$d^+ = 3$	1	1	0.9999989	9,2e-4	7,4e-8	1,021e-18
$d^+ = 4$	1	1	0.9999933	4,5e-5	2,8e-10	3,304e-23
$d^+ = 5$	1	1	0.9999681	2,3e-6	5,0e-12	1,138e-23
$d^+ = 6$	1	1	0.9998819	4,2e-7	2,5e-12	1,135e-23

(2.4), we obtain the exact values of δ_4 and δ_5 and the lower estimate of δ_6. Afterward, Table 4 is filled in according to (8.1) and (8.2).

In [12, p. 899], the authors created the [137, 128, 4] Panchenko code with the weight $A_4 = 45443$. Using (2.1) and (2.3), we get $\Psi(137, 4, 4)$ and the exact value of δ_4. Next, with the help of (3.18) and (1.1), we find the lower estimates of δ_5 and δ_6. The probability $(1 - \Omega_2)^2$ from table 5 is calculated in compliance with (8.1) and (8.2).

8.2. Product of Codes with Distance 6

The probability that just ρ rows (columns) contain more than two error is defined by

$$\binom{n}{\rho}(P_3)^{\rho}(1-P_3)^{n-\rho}.$$

The probability of successful decoding of rows or columns is expressed as

$$\Omega_3 = \sum_{\rho=0}^{d+}\binom{n}{\rho}(P_3)^{\rho}(1-P_3)^{n-\rho}\delta_{\rho}. \qquad (8.3)$$

For product of codes with distance 6, the probability of failure is defined by

$$(1-\Omega_3)^2. \qquad (8.4)$$

The probability of successful decoding of product of codes with distance 6 is expressed as

$$P_{prod,6} = 1-(1-\Omega_3)^2.$$

Example 6. For the product of extended even-weight [79, 64, 6] and [145 128, 6] BCH codes, the fault probabilities $(1 - \Omega_3)^2$ are presented in Tables 6 and 7, respectively. In this case, the input probabilities are $p = 10^{-1}$, 10^{-2}, 5×10^{-3}, 10^{-3}, 5×10^{-4}, and 10^{-4} and thresholds are $d^+ = 5$, 6, 7, 8, and 9.

In [7, Table 5], the [79, 64, 6] BCH code with $A_6 = 17375$ was constructed. From formulas (2.1), (2.3) and (2.4), we obtain the exact values of δ_6 and δ_7. Next, on the basis of formula (3.19), we find the lower estimates of δ_8 and δ_9. Afterward, Table 6 is filled in according to (8.3) and (8.4).

Using [7, Theorem 3.1] and relationship (5.8), it is possible to demonstrate that the even-weight [145, 128, 6] BCH code with $A_6 = 181611$ exists. Once again, the exact values of δ_6 and δ_7 are determined according to

Table 6. Fault probabilities of the product of [79, 64, 6] BCH codes

p	10^{-1}	10^{-2}	5×10^{-3}	10^{-3}	5×10^{-4}	10^{-4}
$d^+ = 5$	1	0.02149	8.9e-10	0	0	0
$d^+ = 6$	1	0.00435	5.4e-12	0	0	0
$d^+ = 7$	1	0.00069	2.5e-14	0	0	0
$d^+ = 8$	1	0.00021	3.2e-15	0	0	0
$d^+ = 9$	1	0.00019	3.1e-15	0	0	0

Table 7. Fault probabilities of the product of [145, 128, 6] BCH codes

p	10^{-1}	10^{-2}	5×10^{-3}	10^{-3}	5×10^{-4}	10^{-4}
$d^+ = 5$	1	0.9999998	0.1981	7.9e-21	1.43e-29	0
$d^+ = 6$	1	0.9999987	0.0827	6.3e-25	1.11e-29	0
$d^+ = 7$	1	0.9999940	0.0282	1.7e-28	1.11e-29	0
$d^+ = 8$	1	0.9999794	0.0103	5.9e-29	1.11e-29	0
$d^+ = 9$	1	0.9999537	0.0066	5.9e-29	1.11e-29	0

formulas (2.1), (2.3) and (2.4), and the lower estimates of δ_8 and δ_9 are found from formula (3.19). The probability $(1 - \Omega_3)^2$ from Table 7 is calculated in compliance with (8.3) and (8.4).

Tables 4–7 illustrate the expected fact that the probability of failure diminishes with increasing d^+.

CONCLUSIONS

In this work, different methods for estimating the number and fractions of erasures with arbitrary weights correctable by binary linear codes with the known weight spectrum of code words (or with partially known or completely unknown weight spectra) are constructed. The examples of calculations are given for Hamming and Panchenko codes with distances 3 and 4 and Bose–Chaudhuri–Hocquenghem codes with distance 6, including their shortenings. Such examples were chosen according to the anticipated field of application: solid-state memory devices and their modifications. It is pertinent to note that, in conformity with calculations, the product of Panchenko codes with distance 4 or BCH codes with distance 6 and identical code sizes of 64 or 128 bit ensure a high reliability: when the error probability in the memory cell is on the order of 0.0001, the fault probability of the [72, 64, 4] Panchenko codes reduces by seven orders of magnitude if the decoding region is extended from three to five erasures. In the case of the product of BCH code of the same size, the analogous effect is observed at an input probability of 0.001. The presented examples indicate that practically maximum effect is attained at a sufficiently small extension of the decoding region (from $d - 1$ to $3/2d$). In essence, this result signifies that, beyond the aforementioned interval, the selection of the practical restriction of the decoding region is determined only by the complexity of erasure corrections.

In the theoretical part of the given work, it is necessary to highlight the different methods of obtainment of estimates: spectral, combinational, and recurrent. It is demonstrated that these methods can be combined depending on the volume of known data on the specified class of codes. As a special case, estimates are deduced for shortened and extended codes. An important result is the derived estimates of the conditional probability of error detection during erasure correction in the extended decoding region. From this estimate (Theorem 6.1), it follows that practically all errors are detected in the extended decoding region under consideration.

ACKNOWLEDGMENTS

This work was performed at the Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, and supported by the Russian Science Foundation, project no. 14-50-00150.

REFERENCES

1. G. D. Forney, "Exponential error bounds for erasure, list, and decision feedback schemes," IEEE Trans. Inf. Theory **14**, 206–220 (1968).
2. V. V. Zyablov and P. S. Rybin, "Erasure correction by low-density codes," Probl. Inf. Transm. **45** (3), 204–220 (2009).
3. M. N. Nazarov and S. P. Mishin, "Almost optimum codes used to correct erasures," in Vestn. Volgograd. Gos. Univ., Ser. 1: Mat., Fiz., No. 4, 59–69 (1999) [in Russian].
4. A. Ashikhmin and A. Barg, "Minimal vectors in linear codes," IEEE Trans. Inf. Theory **44**, 2010–2017 (1998).
5. O. V. Popov, "On assessment of linear code ability to correct erasures and find errors in the presence of erasures," Elektrosvyaz', No. 10, 1 (1967) [in Russian].
6. O. V. Popov, *On Correction of Erasures by Cyclic Codes. Digital Data Transmission over Channels with Memory* (Nauka, Moscow, 1970), pp. 111–124 [in Russian].
7. V. B. Afanassiev, A. A. Davydov, and D. K. Zigangirov, "Design and analysis of codes with distance 4 and 6 minimizing the probability of decoder error," J. Commun. Technol. Electron. **61**, 1440–1455 (2016).
8. M. Bossert, M. Braitbakh, V. V. Zyablov, and V. P. Sidorenko, "Codes correcting the set of error spots or erasures," Probl. Inf. Transmis. **33** (4), 297–306 (1997).
9. A. A. Davydov, A. Yu. Drozhzhina-Labinskaya, and L. M. Tombak, "Supplementary correcting possibilities of BCH codes correcting double errors and finding triple errors," in *Cybernetics questions. Complex design of element and design base of the supercomputer*, by Ed. V. A. Mel'nikov, and Yu. I. Mitropol'skii (VINITI, Moscow, 1988), pp. 86–112 [in Russian].

10. A. A. Davydov, L. P. Kaplan, Yu. V. Smerkis, and G. L. Tauglikh, "Optimization of shortened Hamming codes," Probl. Inf. Transmis. **17** (4), 261–267 (1981).
11. A. A. Davydov and L. M. Tombak, "Number of words with minimum weights in block codes," Probl. Inf. Transmis. **24** (1), 11–24 (1988).
12. A. A. Davydov and L. M. Tombak, "An alternative to the Hamming code in the class of SEC-DED codes in semiconductor memory," IEEE Trans. Inf. Theory **37**, 897–902 (1991).
13. V. I. Panchenko, "On optimization of a linear code with distance 4," in *Proc. VIII All-Union Conf. on According to the Theory of Coding and Information Transfer, Kuibyshev, 1981,* Part 2: *The Theory of Coding* (Akad. Nauk SSSR, Moscow, 1981), pp. 132–134 [in Russian].
14. A. Barg and I. Dumer, "On computing the weight spectrum of cyclic codes," IEEE Trans. Inf. Theory **38**, 1382–1386 (1992).
15. K. M. Cheung, "The weight distribution and randomness of linear codes," in *TDA Progress Report 42-97, Jet Propulsion Lab., Pasadena, CA, California, USA, 1989* (Inst. of Tech., Pasadena, 1989), pp. 208–215. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19890018521.pdf. Accessed 22.11.2016.
16. E. R. Berlekamp, *Algebraic Coding Theory* (McGrow-Hill, New York, 1968).
17. I. Krasikov and S. Litsyn, "On spectra of BCH codes," IEEE Trans. Inf. Theory **41**, 786–788 (1995).
18. F. J. MacWilliams and N. J. A. Sloane, *The Theory Error-Correcting Codes* (North-Holland, Amsterdam, 1977; Svyaz', Moscow, 1979).
19. R. H. Morelos-Zaragoza, *The Art of Error Correcting Coding* (Wiley, Chichester, 2002; Tekhnosfera, Moscow, 2005).
20. V. M. Sidel'nikov, "On spectrum of weights of binary Bose–Chaudhuri–Hocquenghem codes," Probl. Inf. Transmis. **7** (1), 11–17 (1971).
21. Weight Distribution. http://www.ec.okayama-u.ac.jp//~infsys/kusaka/wd/index.html (Accessed 27.11.2016).

Translated by S. Rodikov