 doi:10.20944/preprints202110.0315.v1

ABOUT ACCESSIBLE MAXIMUM EFFICIENCY CORRECTING BINARY CODES

Adzhemov Artem Sergeevich,

MTUCI, President - Chairman of the MTUCI Board of Trustees, Doctor of Technical Sciences, Professor, Moscow, Russia asa@mtuci.ru

Albov Nikolay Yurievich

MTUCI, postgraduate

Annotation

The digital representation of various signals allows, at the subsequent stages of their transmission, to apply correction codes that provide protection against possible errors arising from the action of interference in the communication channel. At the same time, it is important that, with the required correcting ability, these codes have the maximum possible speed. The article presents the results of calculations for linear codes, showing their really achievable limiting capabilities.

Keywords

Code construction, minimum code distance, noise immunity, coding efficiency, theoretically achievable boundary, correction codes.

Introduction

The theory and practice of correcting codes continues to develop [1 - 11]. Various coding constructions and theoretically achievable limits on the coding rate are known. For example, the Hamming border:

$$2^k \le \frac{2^n}{\sum_{i=0}^{\frac{d_{min}-1}{2}} C_n^i} \quad \text{or} \quad r \ge \log_2\left(\sum_{i=0}^{\frac{d_{min}-1}{2}} C_n^i\right), \qquad (1)$$

The Plotkin boundary:

$$d_{min} \le \frac{n\cdot 2^{k-1}}{2^k-1} \quad \text{or} \quad r \ge 2\cdot(d_{min}-1) - \log_2 d_{min} \text{ for } n \le 2\cdot d_{min}-1, \qquad (2)$$

Varshamov - Hilbert boundary:

$$r \ge \log_2\left(\sum_{i=0}^{d_{min}-2} C_{n-1}^i\right), \qquad (3)$$

Singleton boundary:

$$d_{min} \le n-k+1. \qquad (4)$$

where d_{min} is the minimum code distance, n is the length of the code word, k is the number of information symbols, r is the number of check symbols. (The k / n ratio determines the encoding rate).

In this case, it is practically important to know about the real possibility of the attainability of these boundaries, which is investigated in this work for linear block codes.

1. Limit and achievable boundaries of the efficiency of error-correcting coding

Using [10] and expressions (1, 4), let us compare the limiting and achievable boundaries of error-correcting coding for $d_{min} = 2, 3, 4$.

For $d_{min} = 2$ from [10]

$$k = n - 1; \tag{5}$$

for $d_{min} = 3$

$$k = n - 1 - \lfloor log_2 n \rfloor \tag{6}$$

for $d_{min} = 4$

$$k = n - 2 - \lfloor log_2(n-1) \rfloor \, . \tag{7}$$

For the upper Hamming bound for $d_{min} = 2$ we have:

$$k \leq n \, ; \tag{8}$$

for $d_{min} = 3$ and 4

$$k \leq n - \log_2(n+1) \, . \tag{9}$$

For the upper Singleton boundary at $d_{min} = 2$ we have:

$$k \leq n - 1 \, ; \tag{10}$$

for $d_{min} = 3$

$$k \leq n - 2 \, ; \tag{11}$$

for $d_{min} = 4$

$$k \leq n - 3 \, . \tag{12}$$

Comparing the corresponding formulas, we are convinced that the Hamming bound gives slightly overestimated estimates, especially when the minimum code distance is an even number.

The Singleton boundary and the achievable estimate coincide for $d_{min} = 2$. However, further with $d_{min} = 3, 4$ the differences become more and more noticeable.

For the purpose of further comparison at other values of d_{min}, computer calculations were carried out. In this case, the algorithm corresponded to the following model.

As follows from [12], all possible code combinations of a binary code, the number of which is equal to $N_n = 2^n$, for a given value of the minimum code distance d_{min}, can be decomposed into a set of allowed code combinations, the number of which is $K_n^{d_{min}} = 2^k$, and the remaining set of prohibited code combinations, which in turn can be structured into $R_n^{d_{min}}$ groups. Moreover

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2021 doi:10.20944/preprints202110.0315.v1

$$N_n = K_n^{d_{min}} \cdot R_n^{d_{min}} \tag{13}$$

The set of allowed code combinations, among which there will be a zero code combination, will be called the main group and written in the form of a matrix A_n. The dimension of A_n is obviously $K_n^{d_{min}} \times n$. The remaining groups included in the set of forbidden code combinations will be called adjacent groups. The dimension of these groups is the same, but we will denote them by B_n^i, where $i = \overline{1,\ (d_{min}-1)}$ is the code distance between the main group and the adjacent one.

Since there is a zero code word in the main group, then, consequently, the weight of the remaining code words in this group is $w \geq d_{min}$. This condition is not met for adjacent groups.

For $d_{min} = 1$, it is obvious that $K_n^{d_{min}} = N_n$, and there is only one main group. When $d_{min} = 2$, there are two groups. One main and one adjacent. Moreover $2 \cdot K_n^{d_{min}} = N_n$.

With a larger value of the minimum code distance, the number of adjacent groups also changes with increasing n. Moreover, this process is complex, depending on the value of the minimum code distance.

It is easy to see that for a given d_{min} the value $K_n^{d_{min}} = 2$ with an increase in the length of the code word will be for the first time at $n = d_{min}$. And further, for all increasing values of the code word length $n = \overline{d_{min},\ (d_{min} - 1 + \lceil d_{min}/2 \rceil)}$, the number of allowed code words will remain the same. And, therefore, for these values of length n, the number of adjacent groups will increase for the fulfillment of equality (13).

According to [12], we can write that the number of allowed code combinations and the number of groups, respectively, are equal to $K_n^{d_{min}} = 2^k$ and $R_n^{d_{min}} = 2^r$, where k and r are integers, and equality $n = k + r$.

The first main group consists of a zero code word and a combination of ones, the number of which is equal to d_{min}. According to [12], it is possible to construct a recurrence relation for constructing the matrix A_{n+i}, which allows them to be calculated programmatically.

$$A_{n+i} = \begin{Bmatrix} 0_1 0_2 & \cdots & 0_i & A_n \\ 1_1 1_2 & & 1_i & B_n^{d_{min}-i} \end{Bmatrix}, \tag{14}$$

where $i = \overline{1,\ (d_{min}-1)}$, and (0_i) are columns of zeros and (1_i) are columns of ones of the corresponding dimension.

The main problem of forming the matrix A_{n+i} with known A_n is to determine the matrix $B_n^{d_{min}-1}$.

The matrix $B_n^{d_{min}-1}$ is constructed from the matrix A_n by summing all its rows with the generating code combination of weight $w = d_{min} - i$. In this case, the search for such a combination is carried out starting with $i = 1$. After detecting the generating codeword, the adjacent matrix $B_n^{d_{min}-i}$ is formed and then, according to (14), the main matrix A_{n+i}.

Let's give an example. Since the zero code word is included in the main group, then $A_0 = (0)$. For $n = d_{min}$, it is obvious that expression (14) will take the form:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2021 doi:10.20944/preprints202110.0315.v1

$$A_{n=d_{min}} = \begin{Bmatrix} 0_1 0_2 & \dots & 0_{n-1} & A_0 \\ 1_1 1_2 & \dots & 1_{n-1} & B_0^1 \end{Bmatrix} = \begin{Bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{Bmatrix}, \quad (15)$$

where the first row of the matrix consists of $n = d_{min}$ zeros, and the second of $n = d_{min}$ ones. For illustration purposes, let $d_{min} = 5$. Then (15) becomes:

$$A_{n=d_{min}=5} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{Bmatrix}. \quad (16)$$

Continuing the construction of the next basic matrix, it is easy to determine that the generating combination can be, for example, the combination $b_n^{d_{min}-i} = b_8^2 = (11000)$, which in relation to the combinations from (16) has a minimum code distance equal to 2. Therefore, according to (14), we have

$$A_{n=8} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \end{Bmatrix}. \quad (16)$$

It should be noted that when searching for the generating code word $b_n^{d_{min}-i}$, there may be several solutions that allow constructing the basic matrix A_{n+i}. So in the example considered, any code combination with weight $w = 2$ could be chosen as a generating code word, for example $b_8^2 = (00011)$ or $b_8^2 = (10001)$. However, with a further increase in the length of the code word n, the number of such solutions decreases.

Continuing the construction of the main matrix, we are convinced of the need to use the capabilities of computer search for the generating code combination $b_n^{d_{min}-i}$. For this example, this combination in the next step becomes the combination $b_8^3 = (10010100)$. This combination is used to construct the adjacent matrix $B_n^{d_{min}-i} = B_8^3$, equal to

$$B_8^3 = \begin{Bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{Bmatrix}. \quad (17)$$

Further, according to (14), the main matrix is formed

$$A_{n=10} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{Bmatrix}. \quad (18)$$

In [10], the dependences of k on n are presented for $d_{min} = 3$ and 4 in comparison with the Hamming boundary. These dependences are stepwise and only at separate points at $d_{min} = 3$ coincide with the potential Hamming boundary.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2021 doi:10.20944/preprints202110.0315.v1

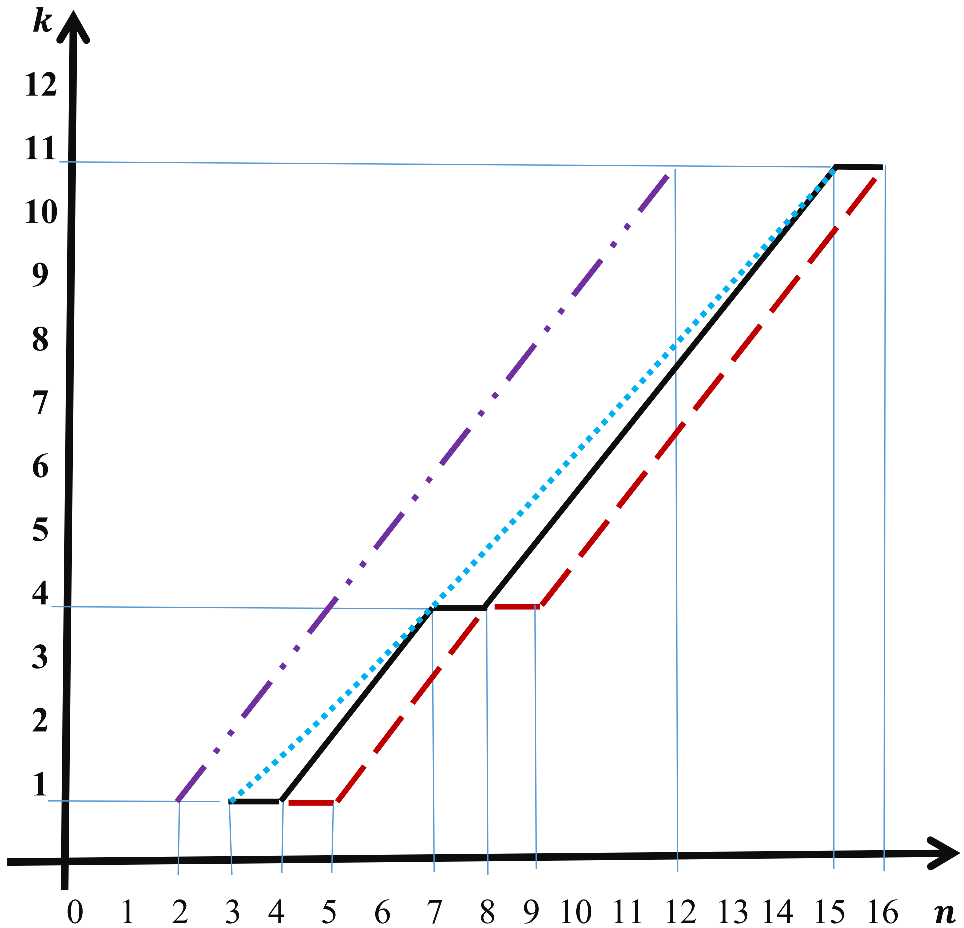

Fig. 1 Dependence of k on n

(Figure 1 shows: black for $d_{min} = 3$, red for $d_{min} = 4$, blue for $d_{min} = 3$ and 4 the upper Hamming limit, purple for $d_{min} = 2$)

Further calculations made it possible to obtain similar data shown in Fig. 2 for $d_{min} = \overline{1, 20}$ and $n = \overline{1, 34}$ (The value of k is shown on the vertical axis in Fig. 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2021 doi:10.20944/preprints202110.0315.v1

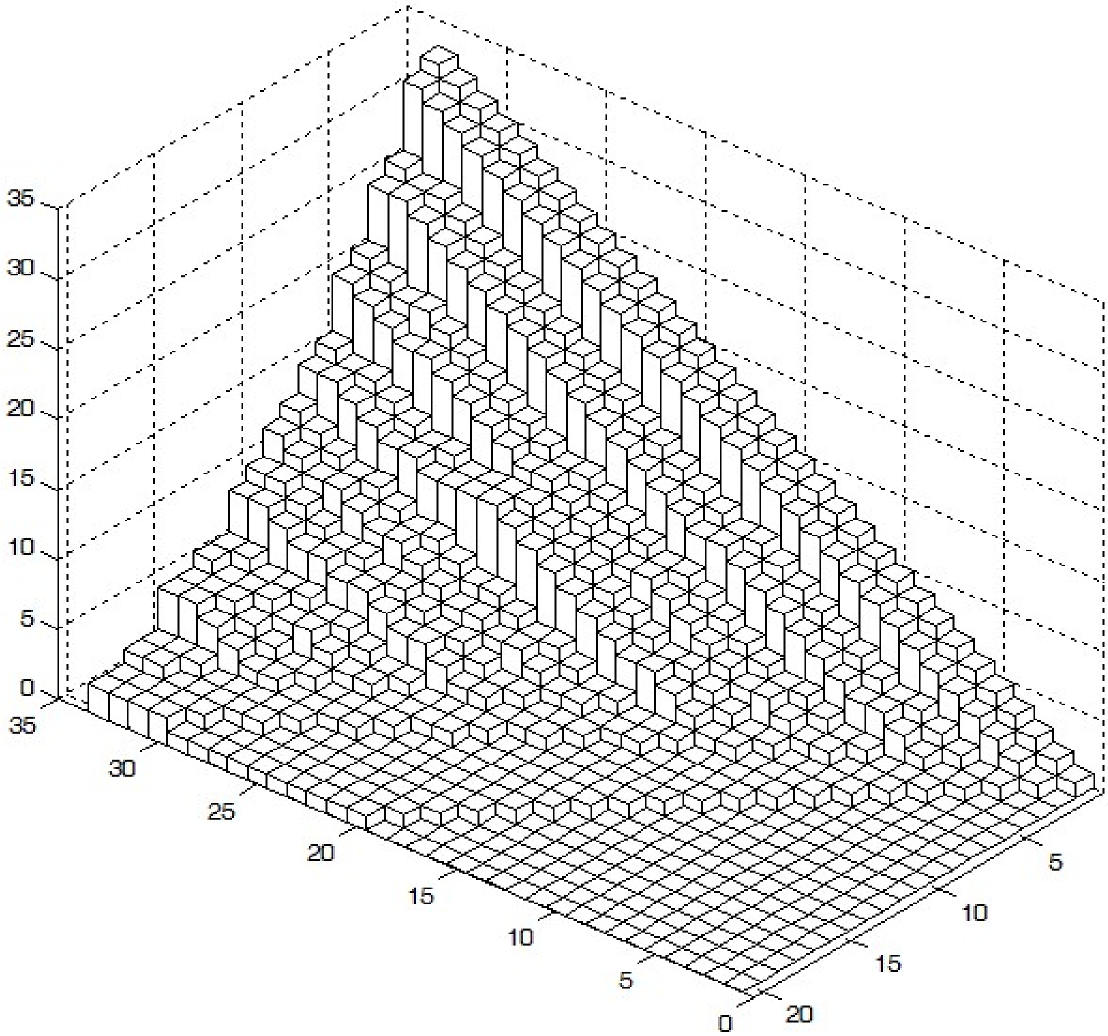

Fig. 2 The number of information symbols k for different values $d_{min} = \overline{1,20}$ and $n = \overline{1,33}$

2. Efficiency of matrix iterated coding

The matrix construction of the correcting code makes it possible to obtain a simpler practical implementation of the code with the required correcting properties. Moreover, as shown in [8], the construction itself can be both two-dimensional and multidimensional.

The total length of the code word is determined by the product of the lengths of the codes used in this design, as well as the number of information symbols.

The minimum code distance of the matrix code is equal to the product of the minimum code distances of the applied correction codes.

In practice, due to the requirements for simplicity of implementation, a two-dimensional construction using a parity-checked code is usually used.

At the same time, it is of interest to evaluate the effectiveness of such a solution, provided that codes that reach theoretically possible boundaries are used for coding.

Suppose that block codes are used according to the Hamming bound.

In this case, the parameters of these codes will be (n_1, k_1) and ((n_2, k_2), respectively. Considering the above, the length of the code word is $n_1 \cdot n_2$, and the number of information symbols is $k_1 \cdot k_2$.

Since it was assumed that the codes implement the Hamming boundary, using (1), we can estimate the corresponding number of information symbols k_{tot}, based on the total length of the codeword, and then, assuming that $n_1 = n_2 = n$ and $k_1 = k_2 = k$, we obtain:

$$E_H = \frac{k_{tot}}{k^2} = \frac{n^2 - \log_2 \sum_{i=0}^{\frac{d_{min}^2-1}{2}} C_n^i}{\left(n - \mathrm{l}\ {}_2\sum_{i=0}^{\frac{d_{min}-1}{2}} C_n^i\right)^2}. \qquad (19)$$

In fig. 3 shows the results of calculations proving that the efficiency of the optimal choice of a code of the corresponding length $n^2$is higher than the efficiency of a matrix construction consisting of two codes of shorter length n. At the same time, with an increase in the length of the code word, this gain decreases and tends to 1.

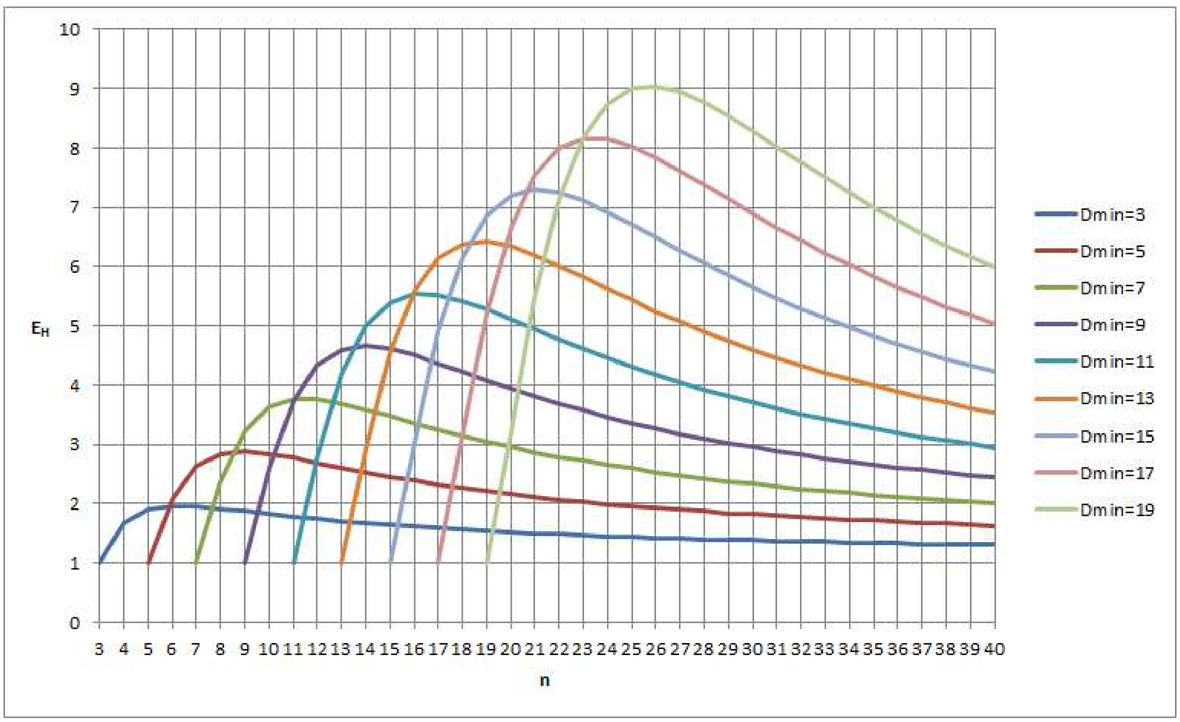

Fig. 3 Efficiency of optimization as a whole in relation to optimization by parts (Hamming bounds

For the Singleton boundary, a similar expression can be obtained

$$E_S = \frac{k_{tot}}{k^2} = \frac{n^2 - d_{min}^2 + 1}{(n - d_{min} + 1)^2}. \qquad (20)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2021 doi:10.20944/preprints202110.0315.v1

The illustration is shown in Fig. 4.

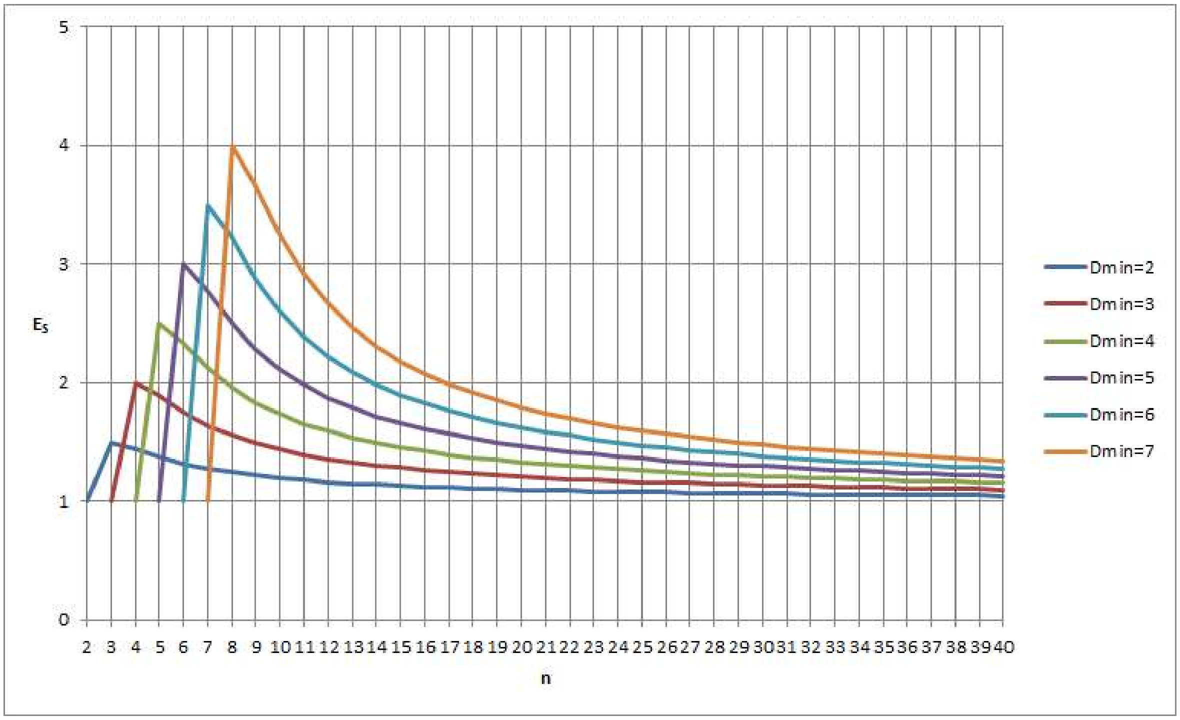

Fig. 4 Efficiency of optimization as a whole in relation to optimization by parts (Singleton bounds)

The results obtained show that for relatively small values of the length of the codeword, preference should be given to the search for the optimal code of length n^2, unless, of course, the requirements for the simplicity of the implementation of the correcting code are not decisive.

Conclusion

The obtained practically achievable boundaries as close as possible to the potential boundaries show the existing possibilities of implementing error-correcting codes with maximum efficiency. At the same time, they differ from the theoretically achievable boundaries by 5-10% and this difference cannot be reduced.

Matrix constructions of error-correcting codes, which have a relatively simple implementation, provide lower efficiency (coding rate) compared to a block code of length n^2 with the same correcting properties. Moreover, this effect is more noticeable at large values of the minimum code distance d_{min}.

Literature.

1. Peterson W., Weldon E. Error-correcting codes / Transl. from English // Ed. R.L. Dobrushin and S.I. Samoilenko. - M .: Ed. "MIR", 1976. - 594 p.
2. Zolotarev V.V., Ovechkin G.V. Noise-resistant coding. Methods and algorithms: Handbook / Ed. Corresponding Member RAS Yu.B. Zubarev. - M .: Hot line - Telecom, 2004 .-- 126 p.

 doi:10.20944/preprints202110.0315.v1

3. Zolotarev V.V., Zubarev Yu.B., Ovechkin G.V. Multi-threshold decoders and coding optimization theory / Ed. Academician V.K. Levin. - M .: Hot line - Telecom, 2013 .-- 232 p.
4. Elias P. Encoding and decoding. - In the book: Lectures on the theory of communication systems. Ed. E.J. Baghdadi. Per. from English M., 1964, p. 289-317.
5. Adzhemov A.S., Sannikov V.G. General communication theory: Textbook for universities - M .: Hot line - Telecom, 2018, 624 p.
6. D. Forney. Concatenated codes; Per. from English V. V. Zyablova and O. V. Popova; Ed. S. I. Samoilenko. - Moscow: Mir, 1970 .-- 205 p.
7. Blokh E.L., Zyablov V.V. Linear concatenated codes. - M .: Nauka, 1982
8. Adzhemov A.S. Possibilities of error-correcting coding when using an iterative code // Electrosvyaz, 2017, No. 11
9. Berrou C., Glavieux A., Thitmajshima P. Near Shannon Limit error-correcting coding and decoding: Turbo-codes Proceedings of ICC '93 - IEEE International Conference on Communications 23-26 May 1993, Geneva, Switzerland
10. Adzhemov A.S., Adzhemov S.A. Estimation of the maximum possible number of information symbols in a noise-immune block code with $d_{min} = 2, 3$ and 4 // Electrosvyaz, 2019, No. 11
11. Robinson, J.P., Bernstein A.J. A class of binary recurrent codes with limited error propagation. IEEE Trans. Inf. Theory, IT-13, January, 2010. – C.106– 113.
12. Adzhemov A.S., Adzhemov S.A. On some features of binary code combinations // IEEE International Scientific Conference Systems of Signals Generating and Processing in the Field of on Board Communications (ON BOARD 2019). Moscow, 20-21 March 2019.pp.1-6
13. Adzhemov A.S. Possibilities of error-correcting coding when using an iterative code // Electrosvyaz, 2017, No. 11