BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **9**, No 3

Sofia • 2009

On Error Detection with Block Codes

Rositza Dodunekova

Chalmers University of Technology and the University of Gothenburg, 41296 Gothenburg, Sweden

Abstract: *In error detection with block codes over symmetric memoryless channels, the code performance is measured by the probability of undetected error. This probability depends on code characteristics and on ε, the symbol error probability of the channel. When the undetected error probability behaves irregularly with respect to ε, difficulties arise in finding a code, appropriate for error detection over a channel with not exactly known symbol error probability (which is most often the case).* ***Good*** *and* ***proper*** *codes are to be preferred in such cases. We present a survey of known methods and techniques for the study of block codes with respect to properness and goodness, together with applications to families of block codes, and some open problems.*

Keywords*: Error detection, block code, proper code, good code.*

1. Introduction: Briefly about channels and codes

Codes are used to control errors, when information is transmitted over noisy communication systems. The system channel may be a telephone line, a high frequency radio link, or a satellite communication link. The noise may be caused by human errors, lightnings, thermal fluctuations, imperfection in equipment, etc. In error control, the original message is encoded in the beginning of the channel by using codewords. An encoded message contains redundant information, used at the end of the channel for better recovering of the original message. The general model of a communication system is sketched below.

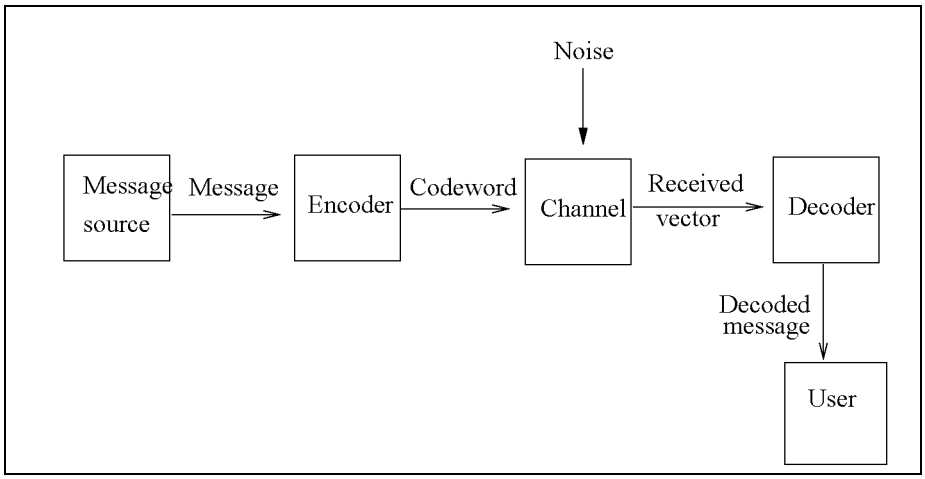

Fig. 1. Basic model of a general communication system

The simple example, where the only messages we want to transmit, are "YES" or "NO", is illustrated below. "YES" is encoded as 00000 and "NO" as 11111. Suppose "YES" was sent and 01001 was obtained. There are two basic methods of error control used by the decoder, both in agreement with the Maximum Likelihood Principle: error detection and error correction. In error detection the decoder would ask for retransmission since the vector obtained is not a codeword. In error correction the decoder decodes the vector into the "nearest" codeword, which is 00000, "YES".

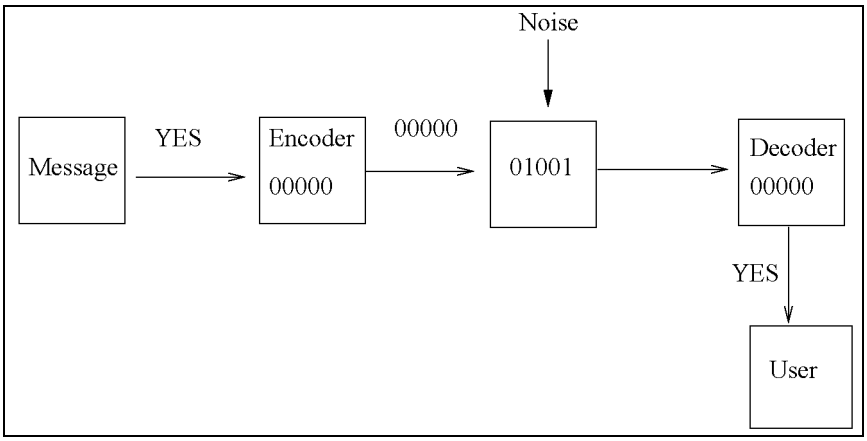

Fig. 2. A simple example where the message source consists of "YES" and "NO" and "YES" is transmitted

Further on we consider the transmission over a q-ary symmetric memoryless channel (SMC). Such a channel has an alphabet with q symbols, each of them remaining unchanged during the transmission with probability $1 - \varepsilon$ and it may change into any of the other $q - 1$ symbols with the same probability $\varepsilon/(q - 1)$. A natural assumption for the channel is that it is more likely for a symbol to remain unchanged during the transmission, than to be changed into some other symbol,

which results in the restriction $0 < \varepsilon < (q - 1)/q$. The q-ary symmetric channel is memoryless, if the errors which occur in separate uses of the channel, are independent.

The next picture describes the mathematical model of a ternary SMC with alphabet {0, 1, 2}.

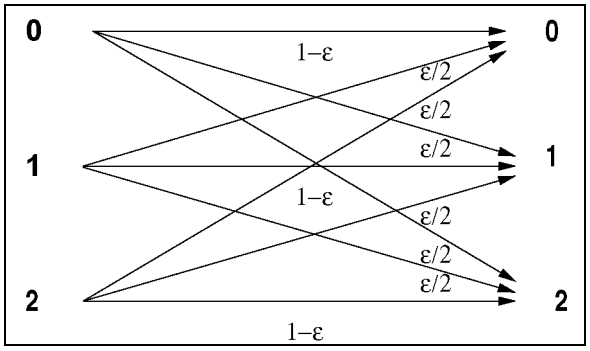

Fig. 3. A ternary SMC with symbol error probability ε

2. Block codes

A q-ary block code is a set of sequences of the same length (codewords) with elements from a finite set F_q with q elements. It is well known, that there exists a field GF(q) over F_q (the Galois Field) if and only if q is a prime power. Let F_q^n denote the n-dimensional vector space over GF(q). The *Hamming distance* between two vectors x and y from F_q^n is the number of non-zero elements in x–y. The *Hamming weight* of a vector is its distance to the zero vector. The *distance distribution* of a block code $C \subset F_q^n$ is a collection of numbers $\{A_0, \ldots, A_n\}$, where A_i equals the number of pairs of codewords in C at distance i, divided by the number of all codewords. The smallest positive distance between two codewords in C is denoted by d and called the *code distance.* A *linear code* of dimension k is a k-dimensional subspace of F_q^n. In this case d equals the minimum positive weight in the code and A_i equals the number of codewords of weight i, $0 \le i \le n$. The *dual code* of a linear code C is defined as the subspace $C^{\perp\cup} \subset F_q^n$ orthogonal to C.

Remark 1. Linear q-ary codes are not defined unless q is a prime power. However, reasonable q-ary codes can be obtained from linear codes in different ways, for example by omitting all codewords containing a given fixed symbol.

Remark 2. The restriction to linear codes is not a sign of weakness. It turns out that codes that are optimal in some way, very frequently are linear.

3. Error detection with block codes

Let $C \subset F_q^n$ be a block code with M codewords. Encoding with C is carried out as follows. After data compression, the initial source information is presented as a series of symbols from F_q, which is divided into blocks of length $k < n$. Each block is a *message,* which is encoded in a codeword from C. In this way a message, written in k symbols, is after encoding written in n symbols, so that $n - k$ symbols are redundant.

Suppose the codeword x was sent and vector y was received. In error detection, the decoder accepts y as the codeword sent, when it is a codeword, or asks for a retransmission, when it is not. Thus transmission errors remain undetected only if the codeword sent changes during the transmission into another codeword. The probability of undetected error of C is given by [19, Ch. 2]:

$$P_{\text{ue}}(C,\varepsilon) = \sum_{i=1}^{n} A_i \left(\frac{\varepsilon}{q-1}\right)^i (1-\varepsilon)^{n-i}, \quad 0 \le \varepsilon \le \frac{q-1}{q}. \tag{3.1}$$

The formula is derived under the assumption of equally likely messages. This assumption is based on the Law of Large Numbers and is basic in C. Shannon's fundamental paper "Mathematical theory of communication" from 1948. Another expression of the probability of undetected error of C is

$$P_{\text{ue}}(C,\varepsilon) = \frac{M}{q^n} A_C^{\text{MW}}\left(1 - \frac{q\varepsilon}{q-1}\right) - (1-\varepsilon)^n \tag{3.2}$$

where

$$A_C^{\text{MW}}(z) = \frac{1}{M}(1-(q-1)z)^n \sum_0^n A_i \left(\frac{1-z}{1+(q-1)zq-1}\right)^i$$

is the MacWilliams transformation [19, Ch. 2]. When C is linear (3.2) leads to

$$P_{\text{ue}}(C,\varepsilon) = q^{-(n-k)} \sum_{i=0}^{n} B_i \left(1 - \frac{q\varepsilon}{q-1}\right)^i - (1-\varepsilon)^n, \quad 0 \le \varepsilon \le \frac{q-1}{q}. \tag{3.3}$$

Here $\{B_i, 0 \le i \le n\}$ is the weight distribution of the dual code $C^\perp$.

Suppose C is a linear code of dimension k. Any $k \times n$ matrix, the rows of which form a basis in C, is called a *generator matrix* of C. In error detection with C, one standard way to check if the received vector is a codeword or not, is to compute its scalar product with a specified generator matrix of $C^\perp$. This product equals zero if and only if the vector is a codeword.

Remark 3. In recent years the problem of making a fast decision if a vector is a codeword or not, has become still more important in connection with large data bases encoded and stored in computers. For different reasons we have to make a fast decision if the information stored has not been substantially corrupted. To check for every single vector if it is a codeword or not, might be expensive. For this reason in some situations we are content with answers of the type "75% of the information is not destroyed" with sufficiently high probability of being true, if only

quick efficient algorithms exist for such answers. These randomized algorithms and codes, for which efficient randomized algorithms exist, are called locally testable codes. Nowadays just a few codes are known to be locally testable. Among these are the shortened first-order Reed-Muller codes and the Reed-Muller codes of constant order.

4. Good and proper codes

For a channel with symbol error probability ε, the most appropriate for error detection would be code C with the smallest possible value of $P_{\text{ue}}(C, \varepsilon)$. It is difficult, however, to find such a code, since no efficient method for such search exists. Furthermore, the symbol error probability ε of the channel is often not known exactly and a code found to be best for some ε may be completely inappropriate for the real channel. It is reasonable in these situations to use codes, which are good or proper. C is *good* for error detection if

$$P_{\text{ue}}(C, \varepsilon) \le P_{\text{ue}}\left(C, \frac{q-1}{q}\right) = q^{-n}(q^k - 1), \quad 0 \le \varepsilon \le \frac{q-1}{q}, \tag{4.1}$$

and C is *proper* if $P_{\text{ue}}(C, \varepsilon)$ is an increasing function of $\varepsilon \in [0, (q-1)/q]$. Thus a good code performs in any channel at least as well as it does in the worst channel with $\varepsilon = (q-1)/q$, and a proper code is just a good code with the advantage that it performs better in better channels. In fact, in the first decades of the foundation of Coding Theory (4.1) this was believed to be true for any linear code, but examples later disproved this.

W o l f, M i c h e l s o n and L e v e s q u e [20] found the average of $P_{\text{ue}}(C, \varepsilon)$ over all q-ary linear codes of length n and dimension k. The result is an increasing function

$$P_{\text{ue}}(\varepsilon) = q^{-(n-k)}[1 - (1 - \varepsilon)^k].$$

Thus a hypothetical "average" $[n, k]_q$ code would be proper, and in this sense a proper code just imitates an "average" error detecting code. This is another strong reason to prefer a proper error detecting code to a non-proper one in situations, where it is impossible to find an optimal code (following the rule to keep to the average, if nothing better can be done).

The concepts of a good and a proper code have been introduced in 1979 [18]. After this, until 1995, when the first monograph on error detecting codes appeared [17], just a few codes have been studied regarding properness and goodness. Among them are the so called Maximum Distance Separable (MDS) codes, shown to be proper in 1984 [15].

5. Study of block codes with respect to goodness and properness

To find out if a single code is proper, good, or non-good, we can use computer graphs or numerical methods for the study of the polynomial representing the undetected error probability. Below the graph of the normed probability of code C

is shown, which is dual to $[819, 12, 384]_2$ Delsarte-Goethals cyclic code. The code is long and has high dimension, so one can expect nice properties. Indeed, considering the picture, this code is proper. However, a proper scaling reveals another picture, see Fig. 2. In fact C is a member of a parametric subclass of non-good Delsarte-Goethals cyclic codes, as shown in [14]:

$$P_1(C, \varepsilon) = \frac{P(C, \varepsilon)}{2^{-n}(2^k - 1)}..$$

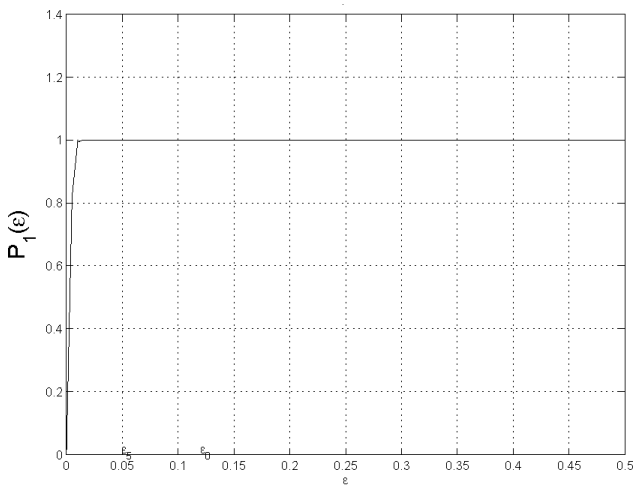

Fig. 4. The normed function $P(C, \varepsilon)$

It becomes much more complicated when we want to investigate parametric families of codes regarding properness and goodness. Below we will present sufficient conditions for goodness or properness, which have shown to be efficient in the study of parametric families of block codes, together with some applications. Some conditions are expressed in terms of basic code parameters and may involve the so called extended binomial moments of the code, others are analytic. It should be mentioned, however, that so far the analytic study of the undetected error probability function for families of codes has shown to be efficient only in a small number of cases.

The *extended binomial moments* of a linear $[n, k, d]_q$ code C are synonymously related to its weight distribution $\{A_0, \ldots, A_n\}$ and are defined as [2]

$$A_0^* = 0, \; A_l^* = \sum_{i=1}^{l} \frac{l_{(i)}}{n_{(i)}} A_i, \quad l = 1, 2, \ldots, n,$$

where $j_{(i)}$ denotes the i-th factorial moment $j(j-1) \ldots (j-i+1)$ of j. In [2, 4, 5], the extended binomial moments have been used to study the undetected error probability function, in particular, to obtain discrete sufficient conditions for properness and goodness.

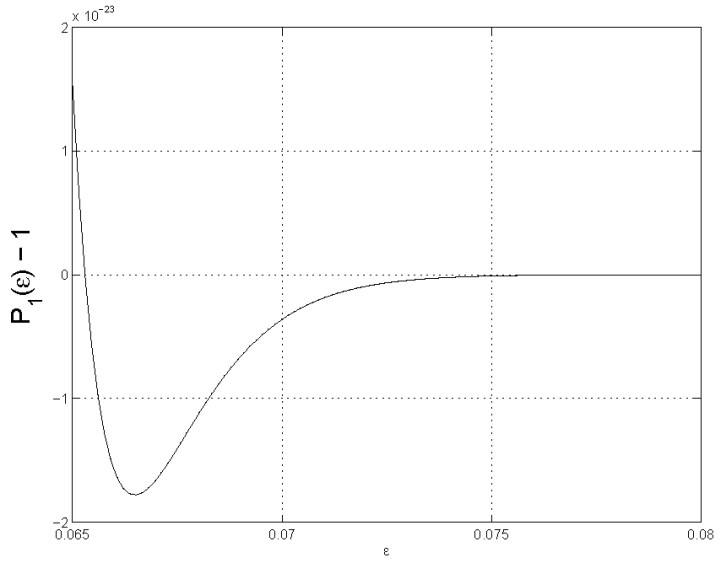

Fig. 5. A scaled picture shows that C is not good

6. Sufficient conditions for properness and goodness and in terms of the weight distribution and basic code parameters

As mentioned in Section 4, the MDS codes have been shown to be proper in [15]. A MDS code is a linear code which is distance optimal, i.e., it has the largest code distance among the linear codes with the same length and dimension. Its dual code is MDS as well. Moreover, a linear code of length n is MDS, if and only if its code distance d and the dual code distance $d^{\perp}$ satisfy $d + d^{\perp} = n + 2$. Also, for any non-MDS code, we have $d + d^{\perp} \le n$. The next two theorems present sufficient conditions for properness and goodness of linear non MDS codes [4, 5, 2].

Theorem 6.1. Let C be a $[n, k, d]_q$ linear code with $d + d^{\perp} \le n$. Then:
(i) if the extended binomial moments of C satisfy

(6.1) $$A_l^* \ge qA_{l-1}^*, \; l = d+1, \ldots, n-d^{\perp}+1,$$

then C is proper;
(ii) if the extended binomial moments of the dual code satisfy

(6.2) $$B_{n-l}^* \ge qB_{n-l+1}^* - q^{n-k-1}(q-1), \; l = d+1, \ldots, n-d^{\perp}+1,$$

then C is proper.

The extended binomial moments of a linear code are strictly increasing so that (6.1) is just a condition on the rate of increase. Though (6.1) and (6.2) are equivalent, (6.2) is more efficient in situations where the dual code distance or the number of non-zero weights in the dual code are small.

Theorem 6.2. Let C be an $[n, k, d]_q$ code with $d + d^{\perp} \le n$. Then:
(i) if the extended binomial moments of C satisfy

(6.3) $$A_l^* q^{-l} \le q^{-n}(q^k - 1),\ l = d, \ldots,\ n - d^\perp,$$

then C is good;

(ii) if the extended binomial moments of the dual code satisfy

(6.4) $$q^{-n+l} B_{n-l}^* \le q^{-k} - q^{-n-k+l},\ l = d, \ldots,\ n - d^\perp,$$

then C is good.

As above, the dual conditions in (6.4) are more efficient in situations, where the dual code distance or the number of non-zero weights in the dual code are small.

Applications. The MDS codes are distance optimal and proper. Next in optimality are Near MDS (NMDS) codes, then the Maximum Minimum Distance (MMD) codes. The dual of an NMDS code is NMDS as well. Many NMDS codes turn out to be proper, by Theorem 6.1; some are good – by Theorem 6.2 [3], [8]. The MMD codes and their duals turn out to be proper [6], [1]. All unique optimal binary linear codes of dimension at most seven and their dual codes are proper [11]. Also, many Cyclic Redundancy-Check codes (CRC) are proper or good, by the above theorems, but some standardized such codes are non-good [16].

7. Sufficient conditions for properness in terms of basic parameters

Computation of the weight distribution of a liner code is an NP hard problem. As a result, relatively few codes are known with their weight distribution. For this reason, to have sufficient conditions for properness, not involving the code weight distribution, would be very useful.

Theorem 7.1. Suppose C is a q-ary linear code of length n, code distance d and dual code distance $d^\perp$. If

$$\max(dd^\perp) \ge [n(q-1) + 1]/q,$$

then both C and its dual code are proper [13].

Applications. Parametric families of Griesmer codes turn out to satisfy the above theorem [12, 13]. A Griesmer code is a linear code which is length optimal, i.e., it has the smallest length among the linear codes with the same dimension and code distance.

8. Properness and goodness in intervals

In our work we have often encountered codes, for which the probability of undetected error has extrema in a relatively small interval $[0, a]$, and then becomes an increasing function up to the endpoint $\varepsilon = (q - 1)/q$. We call such a code proper in $[a, (q-1)/q]$. The following results have been shown in [13].

Theorem 8.1. Let C be a q-ary linear code of length n and dual code distance $d^\perp$. If

(8.1) $$\frac{n(q-1)+2}{q+1} \le d^\perp \le \frac{n(q-1)+1}{q},$$

then C is proper in the interval

$$\left[\frac{n(q-1) - d^{\perp}q + 1}{n(q-1) - d^{\perp}q + 1 + \dfrac{d^{\perp} - 1}{q-1}}, \quad \frac{q-1}{q} \right]. \tag{8.2}$$

Corollary. Suppose the condition of Theorem (8.1) holds. If also

$$\frac{n(q-1) - d^{\perp}q + 1}{n(q-1) - d^{\perp}q + 1 + \dfrac{d^{\perp} - 1}{q-1}} \le \frac{d}{n}. \tag{8.3}$$

then C is proper.

Applications. The above Theorem 8.1 and its Corollary turn out to work well for parametric families of Griesmer codes, see [12, 13]. In all examples of interval then C is proper in the interval properness considered in these works the codes are also asymptotically proper and have small redundancy. Since codes with small redundancy are intensely used in error detection, such examples might be of practical interest.

9. Analytic methods

Unfortunately, at the present time we are not aware of routine analytic methods for the study of parametric polynomials representing the undetected error probability of parametric families of codes, and development of such methods would of course be a challenge.

Below we present two theorems, which are obtained by analytic study of the undetected error probability function.

Theorem 9.1 [10]. A binary block code of length n and code distance d with $d \ge \dfrac{n - \sqrt{n}}{2}$ and symmetric distance distribution ($A_i = A_{n-i}$) is proper.

Theorem 9.2 [7]. Let C be a linear $[n, k, d]_q$ code and assume that for some $\varepsilon_0 \in (0, (q-1)/q)$ we have

$$P_{\text{ue}}(C, \varepsilon_0) \ge q^{-(n-k)}. \tag{9.1}$$

Then $C^{\perp}$ is not good.

Applications. Theorem 9.1 was used in [10] and to snow that some Kerdock codes and the Preparata codes are proper. These are perhaps the first examples of proper non-linear block codes. The Kerdock and the Preparata codes seem to be of a permanent theoretical appeal because of their interesting algebraic-combinatorial properties. Also, it follows from Theorem 9.1, that binary self-complementary block codes, linear and non-linear, which satisfy the so called Grey Rankin bound are proper.

A code for which (9.1) holds, is called ugly. Obviously, an ugly code is non-good. According to theorem 9.2 when a code is ugly, its dual is non-good. We made use of this result in [7], where a full classification is given with respect to properness and goodness of a parametric class of q-ary cyclic codes and their dual codes.

10. List of proper codes

More details regarding the list below can be found in [9]:

All Perfect codes over finite fields;
Some Reed-Muller codes;
Some BCH codes;
The MDS codes;
The MMD codes and their duals;
Some NMDS codes;
Some CRC codes;
Some Griesmer codes;
The unique optimal binary codes of dimension at most seven and their dual codes.

11. Open problems

There are many interesting questions related to good and proper codes. We will mention two of them.

Claude Shannon proved that codes exist for reliable transmission of information at any rate below the channel capacity, but did not provide a construction of an optimal code. Even today, it is not known what an optimal code looks like. Instead, the efforts are devoted to the search for codes, for which the performance in error control is efficient in one sense or another. As a result, codes may be optimal in many different ways. Of greatest interest are the codes, whose parameters are in some sense extremal, like the MDS and Griesmer codes.

Our studies have shown that many linear codes, which are optimal in some sense, or close to optimal, are also proper, and most often their dual codes are proper, too. It is natural to ask if properness and optimality are closely related properties. If so, what kind of relationship would this be?

Another interesting question is to compare the error detecting performance of a proper code with the performance of an “average” code. In the case of binary linear codes our experience shows that a proper code is never worse than an “average” code. If this was general, it would have a strong impact on the theory and practice in communications.

References

1. Dodunekova, R. The Duals of the MMD Codes are Proper for Error Detection. – IEEE Trans. Inform. Theory, **49**, 2003, 2034-2038.
2. Dodunekova, R. The Extended Binomial Moments of a Linear Code and the Undetected Error Probability. – Problemy Peredachi Informatsii, **39**, 2003, 28-39. English translation in Problems Inform. Transmission, **39**, 2003, 255-265.
3. Dodunekova, R., S. Dodunekov. On the Probability of Undetected Error for Near MDS Codes. Dept. Math. Göteborg University, 1995.
4. Dodunekova, R., S. Dodunekov. Sufficient Conditions for Good and Proper Error Detecting Codes. – IEEE Trans. Inform. Theory, **43**, 1997, 2023-2026.
5. Dodunekova, R., S. Dodunekov. Sufficient Conditions for Good and Proper Error Detecting Codes via Their Duals. – Math. Balkanica (NS), **11**, 1997, 375-381.
6. Dodunekova, R., S. Dodunekov. The MMD Codes are Proper for Error Detection. – IEEE Trans. Inform. Theory, **48**, 2002, 3109-3111.
7. Dodunekova, R., S. Dodunekov. Error Detection with a Class of Cyclic Codes. – Math. Balkanica (NS), **21**, 2007, Fasc. 3-4, 361-376.
8. Dodunekova, R., S. Dodunekov, T. Kløve. Almost MDS and Near MDS Codes for Error Detection. – IEEE Trans. Inform. Theory, **43**, 1997, 285-290.
9. Dodunekova, R., S. Dodunekov, E. Nikolova. A Survey on Proper Codes. – Disc. Appl. Math., **156**, 2008, No 9, 1499-1509.
10. Dodunekova, R., S. Dodunekov, E. Nikolova. On the Error-Detecting Performance of Some Classes of Block Codes. – Problemy Peredachi Informatsii, **40**, 2004, No 4, 68-78 (in Russian). English translation in Problems Inform. Transmission, **40**, 2004, No 4, 356-364.
11. Dodunekova, R., S. M. Xiaolei Hu. On the Properness of Some Binary Linear Codes and their Dual Codes. – In: Proc. 11th Intern. Workshop on Algebraic and Combinatorial Coding Theory, Pamporovo, 2008, 76-81.
12. Dodunekova, R., E. Nikolova. Properness of Binary Linear Error-Detecting Codes in Terms of Basic Parameters. – In: Proc. 4th Intern. Workshop on Optimal Codes and Related Topics, Pamporovo, 2005, 133-138.
13. Dodunekova, R., Li Weng. Sufficient Conditions for Interval Properness of Linear Error Detecting Codes. – Math. Balkanica (NS), **21**, 2007, Fasc. 3-4, 245-258.
14. Dodunekova, R., O. Rabaste, J. L. Vega Páez. Error Detection with a Class of Irreducible Binary Cyclic Codes and their Dual Codes. – IEEE Trans. Inform. Theory, **51**, 2005, No 3, 1206-1209.
15. Kasami, T., S. Lin. On the Probability of Undetected Error for the Maximum Distance Separable Codes. – IEEE Trans. Commun., **32**, 1984, No 9, 998-1006.
16. Kazakov, P. Application of Polynomials to CRC and Spherical Codes. PhD Thesis, Technishe Universitet, Delft, 2000.
17. Kløve, T., V. Korzhik. Error Detecting Codes, General Theory and their Application in Feedback Communication Systems. Boston, MA, Kluwer, 1995.
18. Leung-Yan-Cheong, S. K., E. R. Barnes, D. U. Friedman. On Some Properties of the Undetected Error Probability of Linear Codes. – IEEE Trans. Inform. Theory, **25**, 1979, No 1, 110-112.
19. MacWilliams, F. J., N. J. A. Sloane. The Theory of Error-Correcting Codes. Amsterdam, North-Holland, 1977.
20. Wolf, J. K., A. M. Michelson, A. H. Levesque. On the Probability of Undetected Error For Linear Block Codes. – IEEE Trans. Commun., **COM-30**, 1982, No 2, 317-324.