On the properness of some optimal binary linear codes and their dual codes

R. Dodunekova[*]
Mathematical Sciences
Chalmers University of Technology
and the University of Gothenburg
412 96 Gothenburg, Sweden

S. M.Xiaolei Hu
Signals and Systems
Chalmers University of Technology
412 96 Gothenburg, Sweden

1 Introduction

A linear code is said to be proper in error detection over a symmetric memoryless channel if its undetected error probability is an increasing function of the channel symbol error probability. A proper code performs well in error detection in the sense that the better the channel, the better the performance, which makes the code appropriate for use in channels where the symbol error probability is not known exactly.

A q-ary linear code may be optimal in different ways. Of most interest are codes whose parameters are in some sense extremal. For example, Maximum Distance Separable (MDS) codes are distance-optimal among the q-ary linear codes of the same length and dimension. Codes may be also length-optimal and size-optimal.

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. For example, proper are the Perfect codes over finite fields, MDS codes and some Near MDS codes, many Griesmer codes, and Maximum Minimum Distance codes and their duals. Could it be the case that properness and optimality are closely related? What kind of relation would this be?

It is most natural to start the study of these questions by looking for optimal codes which are not proper. In this work we present some preliminary results in this direction. We have studied some binary linear codes of optimal length which cannot be obtained by shortening or puncturing other binary linear codes. The codes turn out to be proper, together with their dual codes. Moreover, like most of the codes listed above, these binary codes satisfy certain conditions that imply properness. These conditions are expressed in terms of the so called extended binomial moments, which are just linear combinations of the elements of the weight distribution of the codes. One interesting observation based on

[*] Partially supported by Research Platform MP2 of the University of Gothenburg.

computer graphs is that the extended binomial moments of these binary proper codes are rather close to a certain general lower bound.

2 Preliminaries

Error detection with linear codes. Let C be a linear $[n, k, d]_q$ code over the finite field $GF(q)$ of q elements, i.e., a k-dimensional subspace of the n-dimensional vector space $GF(q)^n$ over $GF(q)$, with minimum Hamming weight d. Suppose C is used to detect transmission errors on a q-ary discrete memoryless channel. In such a channel, any symbol transmitted has a probability $1 - \varepsilon$ of being received correctly and a probability $\frac{\varepsilon}{q-1}$ of being transformed into each of the $q - 1$ other symbols. Naturally, it should be more likely for a symbol to remain unchanged during the transmission than to change into another symbol, which leads to the restriction $0 \leq \varepsilon \leq \frac{q-1}{q}$.

Let $x \in C$ be the code word transmitted and $y \in GF(q)^n$ be the vector received. In error detection, when y is not a codeword the decoder makes the correct decision that a transmission error has occurred, and asks for a retransmission. When y is a codeword, the decoder decides that y was sent. Such a decision is of course incorrect when y and x are different, thus a transmission error for which the error vector $y - x$ is a non-zero codeword remains undetected. The probability $P_{ue}(C, \varepsilon)$ that an undetected error occurs depends on ε, the basic parameters n, k, d, and q of C, and its weight distribution $\{A_i,\ 0 \leq i \leq n\}$, where A_i is the number of code words in C with weight i. The formula is given by [7]

$$P_{ue}(C, \varepsilon) = \sum_{i=1}^{n} A_i \Big(\frac{\varepsilon}{q-1}\Big)^i (1 - \varepsilon)^{n-i}, \quad 0 \leq \varepsilon \leq \frac{q-1}{q}. \tag{2.1}$$

Proper error detecting codes. In error detection over a particular channel, codes with the smallest probability of undetected error would be the best. However, in order to find such a code, one has to use exhaustive search since presently we don't have any efficient general method for such a search. But even if we would have such a method, this would not solve the problem, since most often ε is not known exactly, and a best code for some ε' may be very inappropriate for the channel, even if its symbol error probability is close to ε'. For this reason the concept of a proper code has been introduced [8, 6, 7].

A linear code is *proper*, if its undetected error probability is an increasing function of ε. Thus the error detecting performance of a proper code is better on better channels, i.e., channels with smaller symbol error probability, which makes the code appropriate for channels where ε is not known exactly.

Another view to properness is gained by comparing the function $P_{ue}(C, \varepsilon)$ of a proper $[n, k, d]_q$ code C to the function $P_{ue}(\varepsilon)$ obtained by averaging the undetected error probability in some set of $[n, k]_q$ codes. In the set of systematic $[n, k]_q$ codes, the averaging procedure gives [10, 11]

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

which is an increasing function of ε. Also in the set of binary $[n,k]$ codes the average undetected error probability is an increasing function [8]:

$$P_{ue}(\varepsilon) = \frac{2^k - 1}{2^n - 1}[1 - (1-\varepsilon)^n].$$

Hence a hypothetical "average" code in the class would be proper. In this sense a proper code is similar to an "average" code, which makes the code a reasonable choice in situations where we cannot do better.

Codes, which are optimal or close to optimal in some sense, are prevailing in the list of proper codes [4]. The question we want to address is if properness and optimality are closely related. As a first step, we have studied some length-optimal binary codes from [1].

Discrete sufficient conditions for properness. Let C be an $[n,k,d]_q$ linear code with weight distribution $\{A_0, A_1, \dots, A_n\}$. The *extended binomial moments* A^*_ℓ of C are defined as [2]

$$\begin{aligned} A^*_\ell &= \sum_{i=d}^{\ell} \frac{\ell(\ell-1)\dots(\ell-i+1)}{n(n-1)\dots(n-i+1)} A_i, \quad d \le \ell \le n, \\ A^*_0 &= 0, \quad 0 \le \ell \le d-1. \end{aligned} \tag{2.2}$$

Let B^*_ℓ be the extended binomial moments of the dual code. It holds [2]

$$B^*_\ell + 1 = q^{\ell-k}(A^*_{n-\ell} + 1), \quad \ell = 0, \dots, n. \tag{2.3}$$

Denote by $d^\perp$ the minimum Hamming distance of the dual code. The following results have been derived in [3, 5, 2].

Theorem 1 *If*

$$A^*_\ell \ge q A^*_{\ell-1}, \quad \ell = d+1, \dots n - d^\perp + 1, \tag{2.4}$$

then C is proper.

Theorem 2 *Suppose C is a binary code. If*

$$d \ge \left\lceil \frac{n}{2} \right\rceil$$

or

$$\left\lceil \frac{n}{3} \right\rceil + 1 \le d^\perp \le \left\lfloor \frac{n}{2} \right\rfloor \quad \textit{and} \quad n(n+1-2d^\perp) \le d(n-d^\perp),$$

then C is proper.

Theorem 3 *The extended binomial moments satisfy*

$$\max\{0,\ q^{\ell-n+k} - 1\} < A^*_\ell < q^{\min(\ell+1-d,\ k+1-d^\perp)} - 1, \quad \ell = d, \dots, n - d^\perp$$
$$A^*_\ell = q^{\ell-n+k} - 1, \quad \ell = n - d^\perp + 1, \dots, n.$$

Linear binary codes of dimension at most 7. Following [1], we say that an $[n, k, d]$ code is *distance-optimal* if no $[n, k, d-1]$ code exists; it is *length-optimal* (which is a stronger condition) if no $[n-1, k, d]$ code exists, and *optimal*, if no $[n+1, k+1, d]$ or $[n+1, k, d+1]$ code exists. An optimal code cannot be obtained by shortening or puncturing other binary linear codes.

Summary of optimal binary codes with $k \leq 7$, $n \leq 2^k$ [1].

$[n, k, d]$	# codes $(form.equiv.)$	$[n, k, d]$	# codes $(form.equiv.)$	$[n, k, d]$	# codes $(form.equiv.)$
$[8, 4, 4]$	1	$[12, 4, 6]$	1	$[16, 5, 8]$	1
$[21, 5, 10]^*$	2	$[24, 5, 12]$	1	$[28, 5, 14]$	1
$[32, 6, 16]$	1	$[38, 6, 18]$	1	$[45, 6, 22]$	1
$[48, 6, 24]$	1	$[53, 6, 26]$	2	$[56, 6, 28]$	1
$[60, 6, 30]$	1	$[24, 7, 10]^*$	6(5)	$[27, 7, 12]$	1
$[40, 7, 18]$	172(46)	$[43, 7, 20]$	7(3)	$[56, 7, 26]^*$	> 19000
$[59, 7, 28]$	143(38)	$[64, 7, 32]$	1	$[71, 7, 34]$	1
$[75, 7, 36]^*$	3603	$[79, 7, 38]$	216(22)	$[82, 7, 40]$	11(7)
$[87, 7, 42]$	55(36)	$[90, 7, 44]$	6(6)	$[93, 7, 46]$	1
$[96, 7, 48]$	1	$[102, 7, 50]^*$	3	$[105, 7, 52]$	1
$[109, 7, 54]$	1	$[112, 7, 56]$	1	$[117, 7, 58]$	2
$[120, 7, 60]$	1	$[124, 7, 62]$	1		

Recall that two codes are formally equivalent if they have the same basic parameters and weight distribution. Clearly, the undetected error probability function (2.1) of such codes is the same. In the above table, the even columns show the number of non-isomorphic codes with the given basic parameters and, in parentheses, the number of classes of formal equivalence.

3 The result

Theorem 4 *All codes in the above table and their duals are proper, except those marked by an asterisk.*

The proof is based on theorems 1 and 2 above. We have used Matlab for computing the extended binomial moments of the codes and their duals, given in (2.2) and (2.3), and for checking the conditions of the theorems. Information about weight enumerators and dual code distances has been taken from [1] and also from the Internet based data bases http://www.codetables.de/ and

http://www.math.unl.edu/∼djaffe2/codes/webcodes/binary/codes.cgi?n=28&k=5.

The codes $[8, 4, 4]$, $[12, 4, 6]$, $[16, 5, 81]$, $[24, 5, 12]$, $[28, 5, 14]$, $[60, 6, 30]$, and $[56, 6, 28]$ have minimum distance $n/2$ and are proper by the first part of Theorem 2. The dual codes have minimum distance at least 3, and are proper by the second part of the theorem. In fact the codes achieve the Grisemer bound. It has been noticed earlier [5] that Theorem 2 is quite efficient for the study of such codes.

We end by noting the following. The extended binomial moments have shown to be a useful tool in the study of the undetected error probability function. We plotted the extended binomial moments of the above codes together with their bounds from Theorem 3. It turns out that the extended binomial moments of these optimal proper codes almost lie on the lower bound.

References

[1] Iliya Bouykliev, David B. Jaffe, Optimal binary linear codes of dimension at most seven. *Disc. Math.* 226, 51-70, 2001.

[2] R. Dodunekova, The extended binomial moments of a linear code and the undetected error probability. *Problemy Peredachi Informatsii* vol. 39, no. 3, 28–39, 2003, English translation in *Problems Inform. Transmission* vol. 39, no. 3, 255–265, 2003.

[3] R. Dodunekova and S. M. Dodunekov, Sufficient conditions for good and proper error detecting codes, *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 2023-2026, 1997.

[4] R. Dodunekova, S. Dodunekov, and E. Nikolova, A survey on proper codes. *Disc. Appl. Math.* 156, no. 9, 1499-1509, 2008.

[5] R. Dodunekova, E. Nikolova, Sufficient conditions for the monotonicity of the undetected error probability for large channel error probabilities. *Problemy Peredachi Informatsii* 41, no. 3, 3–16, 2005.

[6] T. Kasami, S. Lin, On the probability of undetected error for the maximum distance separable codes. *IEEE Trans. Commun.*, vol. 32, no. 9, pp. 998-1006, 1984.

[7] T. Kløve, V. Korzhik, *Error detecting codes, General Theory and their Application in Feedback Communication Systems.* Kluwer, Boston, MA 1995.

[8] S. K. Leung-Yan-Cheong, E. R. Barnes, D. U. Friedman, On some properties of the undetected error probability of linear codes. *IEEE Trans. Inform. Theory*, vol. 25, no. 1, pp. 110-112, 1979.

[9] S. K. Leung-Yan-Cheong, M. E. Hellman, Concerning a bound on undetected error probability. *IEEE Trans. Inform. Theory*, vol. 22, no. 2, pp. 235-237, 1976.

[10] J. Massey, Coding techniques for digital data networks. In *Proc. Int. Conf. Inform. Theory and Syst.*, NTG-Fachberichte, Sept. 18-20, Berlin, Germany, vol. 65, 1978.

[11] J.K. Wolf, A.M. Michelson, A.H. Levesque, On the probability of undetected error for linear block codes. *IEEE Trans. Commun.*, vol. COM-30, no. 2, pp. 317-324, 1982.