t-Good and t-Proper Linear Error Correcting Codes

R. Dodunekova
Department of Mathematics
Chalmers University of Technology
and the University of Gothenburg
412 96 Gothenburg, Sweden

S. M. Dodunekov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
8 G. Bontchev Street
1113 Sofia, Bulgaria

Abstract

The probability of undetected error after using a linear code to correct errors is investigated. Sufficient conditions for a code to be t-*good* or t-*proper* for error correction are derived. Applications to various classes of codes are discussed.

Index terms: error correcting codes, probability of undetected error

I Introduction

Let C be a linear $[n, k, d; q]$ code which is used to correct t or less errors, where $d \geq 2t + 1$. We shall consider a discrete memoryless channel with q inputs and q outputs. Any transmitted symbol has a probability $1 - \varepsilon$ of being received

[0]The work of R. Dodunekova was supported in part by Alice Wallenberg Foundation. The work of S. M. Dodunekov was supported in part by Bulgarian NSF under contracts MM502/95 and MM901/99. The results of this paper were presented in part at the Second International Workshop on Optimal Codes and Related Topics in Sozopol, Bulgaria, June 9-15, 1998.

correctly and a probability $\varepsilon/(q-1)$ of being transformed into each of the $q-1$ other symbols. We assume that $0 \leq \varepsilon \leq \frac{q-1}{q}$.

Let $P_{ud}^{(t)}(C,\varepsilon)$ denote the probability of undetected error after t-error correction and $P_h(\varepsilon)$ denote the probability that an undetectable error pattern in a coset of weight h occurs, $0 \leq h \leq t$. Let $Q_{h,\ell}$ be the number of vectors of weight ℓ in the cosets of weight h, excluding the coset leaders. Then (see [1] and [2])

$$P_h(\varepsilon) = \sum_{\ell=o}^{n} Q_{h,\ell} \left(\frac{\varepsilon}{q-1}\right)^{\ell} (1-\varepsilon)^{n-\ell} \tag{1}$$

and

$$P_{ud}^{(t)}(C,\varepsilon) = \sum_{h=0}^{t} P_h(\varepsilon). \tag{2}$$

The code C is called *t-proper* if $P_{ud}^{(t)}(C,\varepsilon)$ is monotonous and *t-good* if

$$P_{ud}^{(t)}(C,\varepsilon) \leq P_{ud}^{(t)}(C, \frac{q-1}{q})$$

for all $\varepsilon \in \left[0, \frac{q-1}{q}\right]$. It is easy to check that

$$P_{ud}^{(t)}(C, \frac{q-1}{q}) = (q^{-(n-k)} - q^{-n}) V_q(t), \tag{3}$$

where $V_q(t)$ is the volume of the q-nary sphere of radius t in the n-dimensional vector space over $GF(q)$.

In this paper we first derive unified representation of $P_{ud}^{(t)}(C,\varepsilon)$ as a function of $z = \frac{\varepsilon q}{q-1}$, $0 \leq z \leq 1$. Using this representation we obtain then sufficient conditions for a code to be *t-good* or *t-proper*. In the last section of the paper we list some applications of our sufficient conditions, leading to examples of *t-good* and *t-proper* error-correcting codes. For all notions which are not defined here we refer to [3].

II Unified representation of $P_{ud}^{(t)}(C,\varepsilon)$

For $z \in [0,1]$ introduce the functions

$$R_\ell(z) = \binom{n}{\ell} z^\ell (1-z)^{n-\ell},\ \ell = 1, 2, \dots, n \tag{4}$$

and

$$L_\ell(z) = \sum_{j=\ell}^{n} R_j(z),\ \ell = 1, 2, \dots, n. \tag{5}$$

Let C be a linear $[n,k,d;q]$ block code with weight distribution $\{A_i : 0 \le i \le n\}$. We will express the probability of undetected error after error correction $P_{ud}^{(t)}(C,\varepsilon)$ in (2) in terms of either the functions (4) or the functions (5) and the weight distribution

$$\{A_i^{(t)} : A_i^{(t)} = \sum_{h=0}^{t} Q_{h,i},\ i = t+1, \dots, n\} \tag{6}$$

of the vectors in the cosets of weight at most t excluding the leaders. For brevity, denote for $\ell = t+1, \dots, n$

$$A_{\ell,t}^{*} = \sum_{i=t+1}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} A_i^{(t)},\ \ A_{\ell,0}^{*} = A_{\ell}^{*}, \tag{7}$$

where

$$m_{(i)} = m(m-1)\dots(m-i+1) \quad \text{for any integer } m \ge 1.$$

Lemma 1. *The probability of undetected error $P_{ud}^{(t)}(C,\varepsilon)$ has the following representations:*

$$P_{ud}^{(t)}(C,\varepsilon) = P_{ud}^{(t)}(C,z),\ \ z = \frac{\varepsilon q}{q-1} \tag{8}$$

where

$$P_{ud}^{(t)}(C,z) = \sum_{\ell=t+1}^{n} q^{-\ell} A_{\ell,t}^{*} R_{\ell}(z) \tag{9}$$

$$= q^{-(t+1)} A_{t+1,t}^{*} L_{t+1}(z) + \sum_{\ell=t+2}^{n} q^{-\ell}(A_{\ell,t}^{*} - qA_{\ell-1,t}^{*}) L_{\ell}(z). \tag{10}$$

Proof. Let $0 \le h \le t$. Then $Q_{h,\ell} = 0$ for $h \le \ell < t+1$. The functions $P_h(\varepsilon)$ in (1) can be written as

$$\begin{aligned} P_h(\varepsilon) &= \sum_{i=0}^{n} Q_{h,i} q^{-i} \big(\frac{q\varepsilon}{q-1}\big)^{i} (1-\varepsilon)^{n-i} \\ &= \sum_{i=t+1}^{n} Q_{h,i} q^{-i} z^{i} (1 - z + z/q)^{n-i} \\ &= \sum_{i=t+1}^{n} Q_{h,i} q^{-i} z^{i} \sum_{j=0}^{n-i} \binom{n-i}{j} \left(\frac{z}{q}\right)^{j} (1-z)^{n-i-j} \\ &= \sum_{i=t+1}^{n} Q_{h,i} \sum_{j=0}^{n-i} q^{-(i+j)} \binom{n-i}{j} z^{i+j} (1-z)^{n-(i+j)}. \end{aligned}$$

Put $\ell = i + j$ above and use the identity

$$\binom{n-i}{\ell-i} = \binom{n}{\ell}\frac{\ell_{(i)}}{n_{(i)}}$$

to get

$$\begin{aligned} P_h(\varepsilon) &= \sum_{i=t+1}^{n} Q_{h,i} \sum_{\ell=i}^{n} q^{-\ell} \frac{\ell_{(i)}}{n_{(i)}} R_\ell(z) \\ &= \sum_{\ell=t+1}^{n} q^{-\ell} \left[\sum_{i=t+1}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} Q_{n,i} \right] R_\ell(z). \end{aligned}$$

Then by (2)

$$\begin{aligned} P_{ud}^{(t)}(C,\varepsilon) &= \sum_{h=0}^{t} P_h(\varepsilon) \\ &= \sum_{\ell=t+1}^{n} q^{-\ell} \sum_{i=t+1}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} \left[\sum_{h=0}^{t} Q_{h,i} \right] R_\ell(z) \\ &= \sum_{\ell=t+1}^{n} q^{-\ell} \left[\sum_{i=t+1}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} A_i^{(t)} \right] R_\ell(z) = \sum_{\ell=t+1}^{n} q^{-\ell} A_{\ell,t}^{*} R_\ell(z) \end{aligned}$$

which shows (8) with $P_{ud}^{(t)}(C,z)$ as in (9). We show now (10):

$$\begin{aligned} P_{ud}^{(t)}(C,z) &= \sum_{\ell=t+1}^{n-1} q^{-\ell} A_{\ell,t}^{*}[L_\ell(z) - L_{\ell+1}(z)] + q^{-n} A_{n,t}^{*} L_n(z) \\ &= \sum_{\ell=t+1}^{n} q^{-\ell} A_{\ell,t}^{*} L_\ell(z) - \sum_{\ell=t+2}^{n} q^{-(\ell-1)} A_{\ell-1,t}^{*} L_\ell(z) \\ &= q^{-(t+1)} A_{t+1,t}^{*} L_{t+1}(z) + \sum_{\ell=t+2}^{n} q^{-\ell}(A_{\ell,t}^{*} - qA_{\ell-1,t}^{*}) L_\ell(z). \end{aligned}$$

Remark. In the case of $t = 0$, $P_{ud}^{(0)}(C,\varepsilon) = P_{ud}(C,\varepsilon)$, the probability of undetected error when C is used for error detection only. The unified representation of $P_{ud}(C,\varepsilon)$ in terms of the functions $R_\ell(z)$ and $L_\ell(z)$ were found earlier in [4].

Lemma 2. *The functions* $L_\ell(z)$, $\ell = 1, 2, \dots, n$ *are strictly increasing in* $z \in [0,1]$.

Proof. For the proof see [4].

III t-good error correcting codes

Let C be an $[n, k, d; q]$ code over a finite field of q elements $GF(q)$ with weight distribution $\{A_i : 0 \le i \le n\}$. As before, let $V_q(t)$ denote the volume of the q-nary sphere of radius t in the n-dimensional vector space over $GF(q)$. Next theorem gives sufficient conditions for the code C to be t-*good*.

Theorem 1. *If for* $\ell = t+1, \ldots, n$

$$(q^{-(n-k)} - q^{-n})V_q(t) \ge q^{-\ell} \sum_{i=t+1}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} A_i^{(t)} \tag{11}$$

then C *is* t-*good.*

Proof. Note first that

$$A^*_{n,t} = \sum_{i=t+1}^{n} A_i^{(t)} = \sum_{h=0}^{t} \sum_{i=t+1}^{n} Q_{h,i},$$

which is the number of all vectors in the cosets of weight at most t, excluding the leaders. The number of these cosets is $\sum_{h=0}^{t} \binom{n}{h} (q-1)^h$ and every such a coset has q^k elements with one leader among them. Then

$$A^*_{n,t} = (q^k - 1) \sum_{h=0}^{t} \binom{n}{h} (q-1)^h = (q^k - 1) V_q(t) \tag{12}$$

and thus the left-hand side of (11) is equal to $q^{-n} A^*_{n,t}$. Then (11) can be written as

$$q^{-n} A^*_{n,t} \ge q^{-\ell} A^*_{\ell,t}. \tag{13}$$

The theorem now follows from (8)-(9) and the chain of simple relations

$$\begin{aligned} P_{ud}^{(t)}(C, \varepsilon) &= \sum_{\ell=t+1}^{n} q^{-\ell} A^*_{\ell,t} R_\ell(z) \\ &\le q^{-n} A^*_{n,t} \sum_{\ell=t+1}^{n} R_\ell(z) = \\ &= q^{-n} A^*_{n,t} L_{t+1}(z) \le q^{-n} A^*_{n,t} L_{t+1}(1) \\ &= (q^{-(n-k)} - q^{-n}) \sum_{h=0}^{t} \binom{n}{h} (q-1)^h \\ &= P_{ud}^{(t)}(C, \frac{q-1}{q}), \end{aligned}$$

where we have used (13), (5), Lemma 2 and the fact that $L_{t+1}(1) = 1$, (12), and finally (3).

Remark. If $t = 0$, (11) becomes

$$q^{-(n-k)} - q^{-n} \geq q^{-\ell} \sum_{i=d}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} A_i,\ \ell = d, \dots, n$$

and by Theorem 1 the above conditions must be sufficient for the code C to be *good* for error detection. This result was obtained earlier in [4].

IV t-proper error correcting codes

Again, let C be an $[n, k, d; q]$ code with weight distribution $\{A_i, 0 \leq i \leq n\}$. Next theorem gives sufficient conditions for the code to be *t-proper.*

Theorem 2. *If for $i = t+2, \dots, n$*

$$\sum_{i=t+1}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} A_i^{(t)} \geq q \sum_{i=t+1}^{\ell-1} \frac{(\ell-1)_{(i)}}{n_{(i)}} A_i^{(t)} \tag{14}$$

then C is t-proper.

Proof. In terms of (7), (14) is written as

$$A^*_{\ell,t} - qA^*_{\ell-1,t} \geq 0,\ \ell = t+2, \dots, n.$$

Using the above and Lemma 2 in the representation (10) of the probability of undetected error we see that $P_{ud}^{(t)}(C, z)$ is non-decreasing in $z \in [0, 1]$. Since $P_{ud}^{(t)}(C, z)$ is a polynomial, it must strictly increase in z. Thus $P_{ud}^{(t)}(C, \varepsilon)$ is strictly increasing in ε, too.

Remark. If $t = 0$, (14) becomes

$$\sum_{i=d}^{\ell} \frac{\ell_{(i)}}{n_{(i)}} A_i \geq q \sum_{i=d}^{\ell-1} \frac{(\ell-1)_{(i)}}{n_{(i)}} A_i,\ \ell = d+1, \dots, n$$

and by Theorem 2 the above conditions must be sufficient for C to be *proper* for error detection. This result was obtained earlier in [4].

V Applications

Although the problem of finding the weight distribution of a code is known to be NP hard (see [7]), it is often solvable for codes with relatively small parameters. It turns out that for such codes Theorems 1 and 2 are quite effective. Below we refer to some applications.

(i) In [5] the performance of the ternary [13, 7, 5] quadratic-residue code was investigated. Using Theorem 2 it was shown that this code is *t-proper* for error correction, $t = 0, 1, 2$.

(ii) In [6] the performance of all binary cyclic codes of lengths up to 31 and ternary cyclic and negacyclic codes of length up to 20 were systematically investigated. Applying Theorems 1 and 2 a large amount of *t-good* and *t-proper* codes have been found. For more details we refer to [6].

(iii) In [8-10] the corresponding versions of Theorems 1 and 2 for the case of error detection, presented in [4], were used to analyze the performance of CRC-codes of 8-bit and 16-bit redundancy. Many examples of CRC-codes which perform better than the standardized ones were found.

For complete information we refer to [11].

References

1. T. Kasami and S. Lin, "On the probability of undetected error for the Maximum Distance Separable codes", *IEEE Trans. Commun.,* vol. COM-32, no. 9, pp. 998-1006, Sept. 1984.

2. F.J.MacWilliams, "A theorem on the distribution of weights in a systematic code," The Bell System Technical Journal, vol. 42 (Jan. 1963), 79-94.

3. T. Kløve and V. Korzhik, *Error Detecting Codes.* Boston: Kluwer Academic Publishers, 1995.

4. R. Dodunekova and S. M. Dodunekov, "Sufficient conditions for good and proper error detecting codes", *IEEE Trans. Inform. Theory*, vol. 43, pp. 2023-2026. Nov. 1997.

5. Ts. Baicheva, S. Dodunekov, R. Kötter, "On the performance of the ternary [13, 7, 5] quadratic-residue code", in *Proc. Sixth International Workshop on Algebraic and Combinatorial Coding Theory*, Pskov, Russia, September 6-12, 1998, pp. 93-97.

6. Ts. Baicheva, "Binary and ternary linear codes which are good and proper for error correction", in *Proc. Seventh International Workshop on Algebraic and Combinatorial Coding Theory*, Bansko, Bulgaria, June 18-24, 2000, pp. 55-60.

7. E. R. Berlekamp, R. J. McEliece, and H. C. A, van Tilborg, "On the inherent intractability of certain coding problems", *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 384-386, 1978.

8. Ts. Baicheva, S. Dodunekov, and P. Kazakov, "On the cyclic redundancy-check codes with 8-bit redundansy", *Computer Communications*, vol. 21, no. 11, 1998, 1030-1033.

9. Ts. Baicheva, S. Dodunekov, and P. Kazakov, "On the cyclic redundancy-check codes with 16-bit redundansy", in *Proc. Sixth International Workshop on Algebraic and Combinatorial Coding Theory,*Pskov, Russia, September 6-12, 1998, pp. 17-21.

10. Ts. Baicheva, S. Dodunekov, and P. Kazakov, "Undetected error probability performance of cyclic redundancy-check codes of 16-bit redundancy", *IEE Proc. Commun.*, vol. 147, no. 5, Oct. 2000, pp. 253 -256.

11. P. Kazakov, *Application of polynomials to CRC and spherical codes.* PhD Thesis, Delft University of Technology, Delft, 1999.

12. R. Dodunekova and S. Dodunekov, "Sufficient conditions for good and proper linear error correcting codes", in *Proc. of the Second International Workshop on Optimal Codes and Related Topics*, Sozopol, Bulgaria, June 9-15, 1998, pp. 62-67.