DIMACS Technical Report 99-05
January 1999

A new upper bound on codes decodable into size-2 lists

by

A.Ashikhmin
Los Alamos National Laboratory
Group CIC-3
Mail Stop P990
Los Alamos, NM 87545
alexei@c3serve.c3.lanl.gov.

A. Barg *
Bell Labs, Lucent Technologies
600 Mountain avenue 2C-375
Murray Hill, NJ 07974-0636
abarg@research.bell-labs.com

S. Litsyn †
Department of Electrical Engineering–Systems
Tel Aviv University
Ramat Aviv 69978
Israel
litsyn@eng.tau.ac.il.

*Permanent Member
†Research done while visiting DIMACS

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, Bellcore and NEC Research Institute.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

A new asymptotic upper bound on the size of binary codes with the property described in the title is derived. The proof relies on the properties of the distance distribution of binary codes established in earlier related works of the authors.

1 Introduction

Let $C \in \mathbb{Z}_2^n$ be a binary block code. One says that C corrects r errors if every sphere of radius r in $\mathbb{Z}_2^n$ contains at most one codevector and r is the maximal number with such property. Relaxing this definition, one may require that every such sphere contain at most m vectors from the code. Then if r or fewer errors occur in the channel, the transmitted vector can be isolated by compiling a list of m codevectors closest to the received vector. If these conditions hold true, one says that C corrects r errors under list decoding. For brevity, we call such a code C an (m, r) code. The number r will be called the size-m list radius of C.

Let C be an (m, r) code of rate $R(C) = \log_2 |C|/n$. We assume that $r = \rho n$, i.e., the number of errors depends linearly on n, and m is a constant. The main asymptotic problem for (m, r) list codes is to determine the value of

$$R(m, \rho) = \limsup_{n\to\infty} R(C),$$

where the limit is computed over all sequences of codes whose size-m list radius converges to ρ.

The concept of list decoding was introduced by Elias [6] and Wozencraft [12]. Ahlswede [1] showed that it enables one to determine capacity of a wide class of communication channels. Some 30 years after (m, r) codes had been introduced, Blinovsky [4] (see also [5]) derived lower and upper asymptotic bounds on their size for any given value of m. Since in this paper we deal only with the case of $m = 2$, in the theorem below we quote only the relevant bounds from [4].

Let $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ be the entropy function and $H^{-1}(x)$ its inverse.

Theorem 1 [4] *We have* $\overline{R}_2(\rho) \geqslant R(2, \rho) \geqslant \underline{R}_2(\rho)$, *where for* $0 \leqslant h < \infty$, $\overline{R}_2(\rho)$ *is defined parametrically as follows:*

$$\underline{R}_2 = 1 - \frac{1}{2}\Big[h\rho + \log_2\Big(1 + 3\cdot 2^{-h/3}\Big)\Big], \tag{1}$$

$$\rho = \frac{2^{-h/3}}{1 + 3\cdot 2^{-h/3}}. \tag{2}$$

Further,

$$\overline{R}_2 \leqslant 1 - H\big(\frac{1}{2} - \frac{1}{2}\sqrt{1-4\rho}\big) \tag{3}$$

Lower bounds on (m, r) codes for finite n were derived in [7].

Note that formally the upper bound (3) coincides with the well-known Bassalygo-Elias bound on the size of error-correcting codes. Technically it will be more convenient to us to study the function

$$\rho(m, R) = \limsup_{n\to\infty} \frac{1}{n} r(m, C),$$

where $r(m, C)$ is the size-m list radius of the code C. In this paper we are concerned with upper bounds on $\rho(2, R)$ (typically, any such bound also gives an upper bound on $R(2, \rho)$). Eq. (3) implies the bound

$$\rho(2, R) \leqslant H^{-1}(1 - R)(1 - H^{-1}(1 - R)). \tag{4}$$

In this paper we derive an improvement of this bound (and so also of (3)).

The principal technical tool of obtaining the new bound is an application of Delsarte's linear programming method to deriving lower bounds on code invariants, found recently in [10], [2]. In particular, in [10] it is proved that in every code of rate $R > 0$ and sufficiently large length n, there necessarily exists an exponentially large component of the weight distribution. This theorem was used in [3] together with bounds on constant-weight codes to prove sharp estimates of the distance distribution of codes meeting the MRRW upper bound [9] (provided that such exist). These results are also used below. More details and notation are given in Section 2. Section 3 is devoted to the new bound.

2 Notation and preliminaries

Let

$$\delta(R) = \limsup_{n\to\infty} \operatorname{dist}(C),$$

where $\operatorname{dist} C$ is the distance of the code C and the limsup is computed over all sequences of codes of rate R. In other words, $\delta(R) = 2\rho(1, R)$. We shall use the upper (linear programming) bound on $\delta(R)$ [9], which has the form

$$\delta(R) \leqslant \delta^{lp}(R) = \min_{\substack{0\leqslant\beta\leqslant\alpha\leqslant 1/2 \\ H(\alpha)-H(\beta)=1-R}} 2\frac{\alpha(1-\alpha) - \beta(1-\beta)}{1 + 2\sqrt{\beta(1-\beta)}} \tag{5}$$

Likewise, let C be a binary code of distance $d = \delta n$ and constant weight $w = \alpha n$. Define

$$\begin{aligned} \delta(R, \alpha) &= \limsup_{n\to\infty} \frac{d(Rn, \alpha n)}{n}, \\ R(\delta, \alpha) &= \limsup_{n\to\infty} R(C). \end{aligned}$$

By [9], we have

$$\delta(R, \alpha) \leqslant \delta^{lp}(R, \alpha) = 2\frac{\alpha(1-\alpha) - H^{-1}(R)(1 - H^{-1}(R))}{1 + 2\sqrt{H^{-1}(R)(1 - H^{-1}(R))}}, \quad H^{-1}(R) \leqslant \alpha \leqslant \frac{1}{2}$$

In a certain range of parameters this bound can be improved. The improvement is based on a result in [8] and appears in an explicit form in [11]. In the form convenient to us it is given in [3]. Let $\alpha_m(R)$ be the value of α that furnishes the minimum to the right-hand side of (5)*. Then

$$\delta(R,\alpha) \leqslant \delta^{lp}(1+R-H(\alpha)), \quad \alpha_m(R) \leqslant \alpha \leqslant \frac{1}{2}.$$

Let us summarize these results in the following theorem.

Theorem 2

$$\delta(R,\alpha) \leqslant \delta^{up}(R,\alpha) = \begin{cases} \delta^{lp}(R,\alpha), & H^{-1}(R) \leqslant \alpha \leqslant \alpha_m(R), \\ \delta^{lp}(1+R-H(\alpha)), & \alpha_m(R) \leqslant \alpha \leqslant \frac{1}{2}. \end{cases} \tag{6}$$

The second ingredient that we need is the following theorem, which gives a lower bound on the components of the distance distribution of the code. Let

$$A_i = \frac{1}{|C|}\{(c', c'') \in C^2 : \mathrm{dist}\,(c', c'') = i\}.$$

Theorem 3 [10] *For every code of rate R and sufficiently large length n there exists a number ξ,*

$$\xi \in \Big(0, 2\frac{\alpha(1-\alpha)-\beta(1-\beta)}{1+2\sqrt{\beta(1-\beta)}}\Big], \tag{7}$$

such that

$$\frac{1}{n}\log_2 A_{\xi n} \geqslant R-1+H(\beta)+2H(\alpha)-2q(\alpha,\beta,\xi/2)-\xi-(1-\xi)H\left(\frac{\alpha-\xi/2}{1-\xi}\right), \tag{8}$$

where α and β are arbitrary numbers satisfying

$$0 \leqslant \beta \leqslant \alpha \leqslant 1/2, \quad H(\alpha)-H(\beta) \geqslant 1-R, \tag{9}$$

and

$$\begin{aligned} q(\alpha,\beta,\gamma) = H(\beta) + \int_0^\gamma \log_2\Bigg(&\frac{\alpha(1-\alpha)-y(1-2y)-\beta(1-\beta)}{2(\alpha-y)(1-\alpha-y)} \\ &+ \frac{\sqrt{(\alpha(1-\alpha)-y(1-2y)-\beta(1-\beta))^2-4(\alpha-y)(1-\alpha-y)y^2}}{2(\alpha-y)(1-\alpha-y)}\Bigg)dy. \end{aligned} \tag{10}$$

*Note that β in (5) is a dummy variable whose value is determined uniquely given α and R.

3 The new bound

Theorem 4 *Let C be a $(2, \rho n)$ code of rate R, $0 \leqslant R \leqslant 1$. Then*

$$\rho \leqslant \rho(2, R) \leqslant \frac{1}{2} \min_{\alpha,\beta} \max_{\xi} \delta^{up}(R'(\alpha, \beta, \xi), \delta^{lp}(R)),$$

where

$$R'(\alpha, \beta, \xi) = R - 1 + H(\beta) + 2H(\alpha) - 2q(\alpha, \beta, \xi/2) - \xi - (1 - \xi)H\left(\frac{\alpha - \xi/2}{1 - \xi}\right),$$

ξ, α, β satisfy (7)-(9), and $q(\alpha, \beta, \gamma)$ is defined in (10).

Proof. By Theorem 3, there exists ξ in the interval (7) such that the number of codevectors on the sphere of radius ξn centered at a certain codevector a satisfies (8). We can translate the space $\mathbb{Z}_2^n$ by a; then this claim is equivalent to the existence of a constant-weight code of rate R'. This code has relative minimum distance at most $\delta^{up}(R, \xi)$ (cf. (6)). Take two codevectors c', c'' at a distance $n\delta^{up}(R, \xi)$ and consider them together with the center of the sphere (0, for that matter). It is easy to see that the center of the sphere of minimal radius that contains c', c'' and 0 has weight $\frac{1}{2} n\delta^{up}(R'(\alpha, \beta, \xi), \xi)$. ∎

Optimization carried out in [3] leads to the following corollary.

Corollary 5

$$\rho(2, R) \leqslant \begin{cases} \frac{1}{2}\delta^{lp}(R - 1 + H(\delta^{lp}(R)), \delta^{lp}(R)), & 0 \leqslant R \leqslant R_0, \\ \frac{1}{2}\delta^{lp}(R), & R_0 \leqslant R \leqslant 1, \end{cases} \tag{11}$$

where $R_0 = 0.421\ldots$ is the root of the equation $\alpha_m(R) = \delta^{lp}(R)$.

Bound (11) is plotted in Fig. 1 together with bounds (1)-(2) and (3). Computations show that it is better than (3) for all $R \in (0, 1)$. Note that the second segment in (11) coincides with the best known upper bound (5) on $\rho(1, R) = \delta(R)$. We wish to stress a difference between this result and the upper bound in Theorem 1. The upper bound in Theorem 1 is the same for the cases of $m = 1$ and $m = 2$ simply because the way of counting the contribution to the weight of the center of the sphere in [4] cannot tell between $m = 2i - 1$ and $m = 2i$. Corollary 5, in contrast, indicates a geometric property of (hypothetical) codes meeting the MRRW upper bound (5), namely, that for some pairs of vectors at a distance $n\delta^{lp}(R)$ apart, there is a third vector at the same distance from each of them.

For reference purposes we also give a short table of values of the bounds.

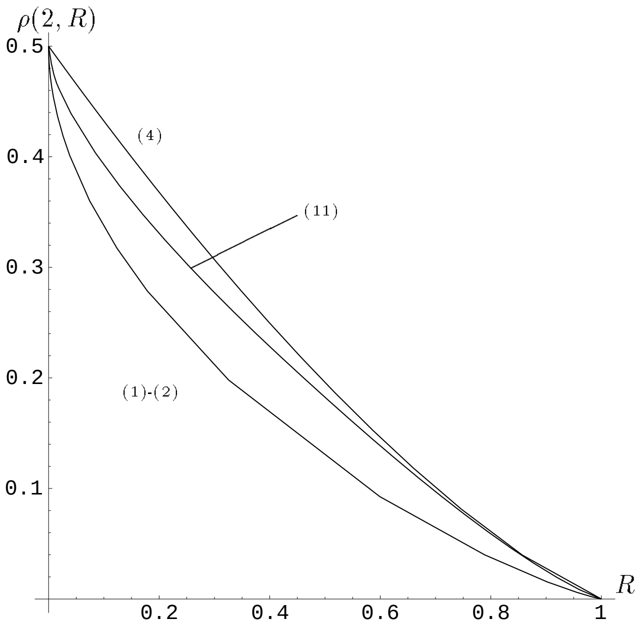

Figure 1: Bounds on the size-2 list radius of a code of rate R

Table 1: Bounds on $\rho(2, R)$.

R	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Lower bound (1)-(2)	0.168	0.133	0.105	0.082	0.063	0.046	0.031	0.018	0.0079
Elias bound (4)	0.216	0.184	0.153	0.125	0.098	0.073	0.050	0.030	0.0128
New bound (11)	0.196	0.165	0.138	0.114	0.091	0.069	0.048	0.029	0.0127

References

[1] R. Ahlswede (1973). Channel capacities for list codes. *J. Appl. Probability*, 10:824–836.

[2] A. Ashikhmin and A. Barg (1998). Binomial moments of the distance distribution: Bounds and applications. *IEEE Trans. Inform. Theory*, to appear.

[3] A. Ashikhmin, A. Barg, and S. Litsyn (1998). New upper bounds on generalized distances. *IEEE Trans. Inform. Theory*, submitted for publication.

[4] V. Blinovsky (1986). Bounds for codes decodable in a list of finite size. *Problems of Information Transmission*, 22(1):11–25.

[5] V. Blinovsky (1997). *Asymptotic Combinatorial Coding Theory.* Kluwer Academic Publishers, Boston.

[6] P. Elias (1957). List decoding for noisy channels. Rep. No. 335 Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass. MR 20 #5702.

[7] P. Elias (1991). Error correcting codes for list decoding. *IEEE Trans. Inform. Theory*, 37:5–12.

[8] V. I. Levenshtein (1971). Upper-bound estimates for fixed-weight codes. *Problemy Peredachi Informatsii*, 7(4):3–12, in Russian. English translation in *Probl. Inform. Trans.* 7:281–287.

[9] R. J. McEliece, E. R. Rodemich, H. Rumsey, and L. R. Welch (1977). New upper bound on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Trans. Inform. Theory*, 23:157–166.

[10] S. Litsyn (1998). New bounds on error exponents. *IEEE Trans. Inform. Theory*, to appear.

[11] A. Samorodnitsky (1998). On the optimum of Delsarte's linear program. *J. Combinatorial Theory, Ser. A*, to appear.

[12] J. M. Wozencraft (1958) List decoding. Quarterly Progr. Rep., Res. Lab. Electronics, MIT, 48:90–95.