Correspondence

Estimates for the Range of Binomiality in Codes' Spectra

Ilia Krasikov and Simon Litsyn, *Member, IEEE*

***Abstract*—We derive new estimates for the range of binomiality in a code's spectra, where the distance distribution of a code is upperbounded by the corresponding normalized binomial distribution. The estimates depend on the code's dual distance.**

***Index Terms*—Krawtchouk polynomials, spectra of codes.**

I. Introduction

It is well-known (see, e.g. [12, p. 287]) that the expected number of codewords of weight $i > 0$ in a random linear code C of length n is $\binom{n}{i}|C|/2^n$, i.e., is just the normalized binomial distribution. For the codes with the spectrum (distance distribution) upper-bounded by the corresponding binomial distribution, the error probability exponent for the maximum-likelihood decoding coincides with the random coding exponent, see, e.g., [1], [2]. The accuracy of the binomial approximation to the distance distribution increases with the dual distance. Such estimates for the case of the dual distance being close to $n/2$ (for instance, this is the case for BCH codes) were derived in [5]–[7], [14], [15]. These results are based on estimates for values of Krawtchouk polynomials, and for d', the dual distance, growing linearly in the code length n, only weak bounds exist for several spectrum elements in the neighborhood of $n/2$.

In [8] we demonstrated that every even code with relative dual distance $\delta' = d'/n$ has an asymptotically (in n) binomial distance distribution in the interval

$$\left(\frac{n}{2}\left(1-\frac{\delta'}{1-\delta'}\right), \frac{n}{2}\left(1+\frac{\delta'}{1-\delta'}\right)\right) \tag{1}$$

in the following sense: all the components in this range are upper-bounded by

$$\mathrm{const}\cdot\sqrt{n}\binom{n}{i}|C|/2^n$$

and for every subinterval of size about $\sqrt{n\,\ln\,n}$ there exists a component asymptotically achieving the upper bound. For even codes with even dual codes, the interval can be further extended to

$$\left(\frac{n}{2}\left(1-\sqrt{\frac{\delta'}{1-\delta'}}\right), \frac{n}{2}\left(1+\sqrt{\frac{\delta'}{1-\delta'}}\right)\right). \tag{2}$$

Even more can be said about self-dual codes. In particular, we proved in [8] that if there exist self-dual codes with

$$\delta = d/n \ge 1/2 - \sqrt{2}/4 \approx 0.146\cdots$$

then their spectrum is binomial in the interval

$$(n(1/2-\sqrt{3}/6), n(1/2+\sqrt{3}/6)).$$

In this correspondence we use the linear programming approach for further extending the range of "binomiality." It is shown is Section II that for d' growing linearly in n, the binomial upper bound is valid in the interval of indices of spectrum elements

$$\left(\frac{n}{2}\left(1-\sqrt{\delta'(2-\delta')}\right), \frac{n}{2}(1-\sqrt{\delta'(2-\delta')})\right).$$

Notice, that this bound is reminiscent of Tietäväinen's bound [16] on the covering radius of codes with known dual distance. In Section III we develop a dual approach yielding binomiality of spectra of even codes in the range

$$\left(\frac{n}{2}(1-2\delta')^2, n-\frac{n}{2}(1-2\delta')^2\right)$$

which is better than the previous bound for $\delta' \ge 0.164\cdots$.

In what follows, $H(x)$ stands, as usual, for the entropy function

$$H(x) = -x\log x - (1-x)\log(1-x)$$

all logarithms are base e.

The Krawtchouk polynomials are defined as follows:

$$P_k^n(x) = P_k(x) = \sum_{j=0}^{k}(-1)^j\binom{x}{j}\binom{n-x}{k-j}.$$

For their properties consult [9]–[12]. In the sequel, n always stands for the length of the code, and is omitted in the notation for Krawtchouk polynomials.

II. The First Estimate

Let C be a code of length n and size $|C|$. Let $\boldsymbol{B} = (B_0, B_1, \cdots, B_n)$ stand for the distance distribution of C, and $\boldsymbol{B}' = (B_0', B_1', \cdots, B_n')$ be the dual distance distribution. $\boldsymbol{B}$ and $\boldsymbol{B}'$ are related by the MacWilliams transform

$$|C|B_i' = \sum_{j=0}^{n} B_j P_i(j) \tag{3}$$

where P_i are the corresponding Krawtchouk polynomials. Let d', the dual distance of C, be the minimal positive index of the nonzero component of $\boldsymbol{B}'$. Denote $|C'| = 2^n/|C|$. The following lemma is apparently due to Delsarte [3].

Lemma 1: Let

$$\alpha(x) = \sum_{i=0}^{s}\alpha_i P_i(x), \qquad 0 \le s \le n$$

then

$$\alpha_0|C| + |C|\sum_{i=d'}^{s}\alpha_i B_i' = \sum_{j=0}^{n}\alpha(j)B_j. \tag{4}$$

Proof: Calculating $|C|\Sigma_{i=0}^{r}\alpha_i B_i'$, we get the claim from (3). □

We will also need a bound on r_t, the smallest root of $P_t(j)$. For t growing linearly in n and $t = \tau n$ (see, e.g., [12])

$$\xi_t = \frac{r_t}{n} = \frac{1}{2} - \sqrt{\tau(1-\tau)} + o(1). \tag{5}$$

Denote $\xi = j/n$. The next observation is due to Kalai and Linial [4].

Manuscript received March 29, 1995; revised October 22, 1996. This work was supported in part by the Guastallo Fellowship.

I. Krasikov is with the Tel-Aviv University, School of Mathematical Sciences, Ramat-Aviv 69978 Tel-Aviv, Israel, and with Beit-Berl College, Kfar-Sava, Israel.

S. Litsyn is with the Tel-Aviv University, Department of Electrical Engineering–Systems, Ramat-Aviv 69978, Tel-Aviv, Israel.

Publisher Item Identifier S 0018-9448(97)02316-X.

0018–9448/97$10.00 © 1997 IEEE

Lemma 2: For $\xi < \xi_t$

$$\begin{aligned}\frac{1}{n}\log P_t(j) = {}&\frac{1}{n}\log\binom{n}{t}\\ &+\int_0^{\xi}\log\left(1-2\tau+\frac{\sqrt{(1-2\tau)^2-4z(1-z)}}{2(1-z)}\right)dz\\ &+O\left(\frac{1}{n}\right).\end{aligned}\tag{6}$$

Proof: It is known (see, e.g., [12, Lemma 36]) that for $\xi < \xi_t$

$$\frac{P_t(j+1)}{P_t(j)} = \frac{n-2t+\sqrt{(n-2t)^2-4j(n-j)}}{2(n-j)} \cdot \left(1+O\left(\frac{1}{n}\right)\right).$$

Taking the logarithm on both sides, applying this recursively to $P_t(0) = \binom{n}{t}$, and approximating the sum by the integral we get the claim. □

We will make use of the following multiplication rule for Krawtchouk polynomials, see, e.g., [11, p. 17] and [9].

Lemma 3: For $0 \le x \le n$

$$P_i(x)P_j(x) = \sum_{k=\max(0,i+j-n)}^{\min(i,j)} \binom{n-i-j+2k}{k} \cdot \binom{i+j-2k}{j-k} P_{i+j-2k}.$$ □

Now we are in the position to prove the next

Theorem 1: For n growing and

$$j/n \in (1/2 - 1/2\sqrt{\delta'(2-\delta')}, 1/2 + 1/2\sqrt{\delta'(2-\delta')})\tag{7}$$

we have

$$B_j = O\left(n\frac{\binom{n}{j}}{|C'|}\right).$$

Proof: Let us choose in Lemma 1 $\alpha(j) = P_t^2(j)$ where $2t+1 \le d'$. The following expansion is due to Lemma 3:

$$P_t^2(j) = \sum_{i=0}^{t}\binom{2i}{i}\binom{n-2i}{t-i}P_{2i}(j)$$

that yields

$$\alpha_0 = \binom{n}{t}.$$

Now, we get

$$|C|\binom{n}{t} + |C|\sum_{i=d'}^{n}\alpha_i B_i' = \sum_{j=0}^{n} P_t^2(j)B_j.\tag{8}$$

Since $2t < d'$ the sum in the left-hand side vanishes, so

$$B_j \le \frac{|C|\binom{n}{t}}{P_t^2(j)} = \frac{\binom{n}{j}}{|C'|}\,\frac{2^n\binom{n}{t}}{\binom{n}{j}P_t^2(j)}.\tag{9}$$

Consider $j < r_t$, but such that

$$\xi = 1/2 - \sqrt{\tau(1-\tau)} - o(1).$$

By virtue of (5), this choice is possible (that is, the appropriate $t < d'/2$ does exist) if

$$\xi \in (1/2 - 1/2\sqrt{\delta'(2-\delta')}, 1/2 + 1/2\sqrt{\delta'(2-\delta')}).$$

To prove the theorem we will show that for such t and j

$$R = \frac{2^n\binom{n}{t}}{\binom{n}{j}P_t^2(j)} = O(n).\tag{10}$$

Notice, that the integral in Lemma 2 can be expressed explicitly [4], namely,

$$\begin{aligned}&\int\log\left(\frac{1-2\tau+\sqrt{(1-2\tau)^2-4z(1-z)}}{2(1-z)}\right)dz\\ &\quad= \log(1-z) + \frac{c}{2}\log(1-2z-\sqrt{c^2-4z+4z^2})\\ &\qquad+ z\log\left(\frac{c+\sqrt{c^2-4z+4z^2}}{2-2z}\right)\\ &\qquad- \frac{1}{2}\log(2-c^2-2z-c\sqrt{c^2-4z+4z^2})\end{aligned}$$

where $c = 1-2\tau$. For $\tau = 1/2 - \sqrt{\xi(1-\xi)}$ we have

$$\begin{aligned}I &= \int_0^{\xi}\log\left(\frac{1-2\tau+\sqrt{(1-2\tau)^2-4z(1-z)}}{2(1-z)}\right)dz\\ &= \frac{1}{2}\log(1-2\xi) + 2\sqrt{(1-\xi)}\log(1-2\xi)\\ &\quad+ \log(1-\xi) - \xi\log(1-\xi)\\ &\quad+ \xi\log\xi - 2\sqrt{(1-\xi)\xi}\log(1-2\sqrt{(1-\xi)\xi}).\end{aligned}$$

Taking into account that

$$\log\binom{n}{k} = nH\left(\frac{k}{n}\right) + \frac{1}{2}\log\frac{n}{k(n-k)} + O(1)$$

we get

$$\begin{aligned}\frac{1}{n}\log R = {}&\log 2 + H(\tau) - H(\xi_t) - 2(H(\tau)+I)\\ &+ \frac{1}{2n}\log\frac{jt(n-t)(n-j)}{n^2} + O\left(\frac{1}{n}\right).\end{aligned}$$

Plugging in the value of I, after direct calculations we get (10). Moreover, observing that $P_t^2(j)$, the coefficient of B_j in (8), is equal to $P_t^2(n-j)$, the coefficient of B_{n-j}, we conclude that (9) is valid also for B_{n-j}. Therefore, the claim of the theorem holds also for $j > n/2$. □

Notice, that the above proof actually shows that, although $P_t(j)$ tends to zero while j tends to r_t, the right-hand side of (6) for $\xi \to \xi_t$ is

$$1/2(\log 2 + H(\tau) - H(\xi_t)) \ne 0.$$

This surprizing phenomenon can be interpreted as follows. Consider the function $f(j) = P_t(j)/(r_t - j)$. Clearly, (6) is still valid for $f(j)$. However, $f(r_t)$ is (asymptotically) greater than the absolute value of the first local extremum of $P_t(j)$ (i.e., the extremum between the first and the second roots of $P_t(j)$). This extremum is obviously far from being zero. This is actually the reason why we are allowed to substitute the limit value $\xi = 1/2 - \sqrt{\tau(1-\tau)}$ in the proof of the above theorem.

III. The Dual Estimate

Here we will employ polynomials which are, roughly speaking, dual to $P_t^2(x)$. The following lemma is complementary to Lemma 1.

Lemma 4: Let

$$\beta(x)=\sum_{i=0}^{s}\beta_i P_i(x), 0<s<n.$$

Then

$$\frac{|C|}{2^n}\left(\beta(0)+\sum_{i=d'}^{n}\beta(i)B_i'\right)=\sum_{j=0}^{s}\beta_j B_j. \qquad \square$$

Notice, that if we choose $\beta(x)$ such that $\beta(i)\le 0$ for such $i\ge d'$, $\beta(0)>0$, and all $\beta_j\ge 0$, then indeed

$$|C|\ge\frac{2^n\beta_0}{\beta(0)}. \tag{11}$$

Choosing the polynomial $\beta(x)$ as in [13] (see also [12, ch. 17, Theorem 35]) we get the following lemma (see also [10] for other estimates).

Lemma 5: For a code C of length n with the given dual distance $d'=\delta n$ and n tending to infinity

$$\frac{1}{n}\log|C|\ge 1-H(1/2-\sqrt{\delta'(1-\delta')}). \qquad \square$$

We will give the estimate for even codes, i.e., when the dual spectrum is symmetric with respect to $n/2$. We choose in Lemma 4

$$\beta(x)=(x-n/2+\sqrt{t(n-t)})(x-n/2-\sqrt{t(n-t)})P_t^2(x) \tag{12}$$

where $r_t\le d'$. Then

$$\frac{|C|}{2^n}\beta(0)\le\sum_{j=0}^{s}\beta_j B_j$$

is valid if

$$\tau=t/n\in(1/2-\sqrt{\delta'(1-\delta')},1/2+\sqrt{\delta'(1-\delta')}). \tag{13}$$

Lemma 6:

$$\beta(x)=1/2(\Sigma_1+\Sigma_2+\Sigma_3)$$

where

$$\Sigma_1=\sum_{i=1}^{t+1}\frac{i^2(n-2i+2)(n-2i+1)}{2(t-i+1)(n-t-i+1)}\binom{2i}{i}\binom{n-2i}{t-i}P_{2i}(x)$$

$$\Sigma_2=\sum_{i=0}^{t}\left(2i(n-2i)+\frac{n-4tn+4t^2}{2}\right)\binom{2i}{i}\binom{n-2i}{t-i}P_{2i}(x)$$

$$\Sigma_3=\sum_{i=0}^{t-1}\frac{(2i+1)(t-i)(n-t-i)}{i+1}\binom{2i}{i}\binom{n-2i}{t-i}P_{2i}(x).$$

Proof: Observing that

$$(x-n/2+\sqrt{t(n-t)})(x+n/2-\sqrt{t(n-t)})=\frac{1}{2}P_2(x)+\frac{n-4tn+4t^2}{4}$$

and applying Lemma 3 after straightforward calculations we get the claim. $\square$

Lemma 7: For n sufficiently large, t linear in n, $t/n\le\text{const}<1/2$

$$i>\left(\frac{2t^2(n-t)^2}{(n-2t)^2}\right)^{1/3}=O(n^{2/3})$$

β_{2i} are positive.

Proof: Clearly, β_{t+1} is positive. For $i=t, t<n/2$, β_t is also positive since

$$\frac{t^2(n-2t+2)(n-2t+1)}{2(n-2t+1)}+\left(2t(n-2t)+\frac{n-4tn+4t^2}{2}\right)=\frac{n+t^2(n-2t-2)}{2}>0.$$

Now, by Lemma 6 it is left to show that under our assumptions

$$S(i)=\frac{i^2(n-2i+2)(n-2i+1)}{2(t-i+1)(n-t-i+1)}+\left(2i(n-2i)+\frac{n-4tn+4t^2}{2}\right)+\frac{(2i+1)(t-i)(n-t-i)}{i+1}>0 \tag{14}$$

for $i<t$.

Put, for some $c>1$,

$$i=c\left(\frac{2t^2(n-t)^2}{(n-2t)^2}\right)^{1/3}$$

then we calculate

$$2(1+i)(1-i+n-t)(1-i+t)S(i)=2(c-1)(1+c+c^2)(n-2t)^{10/3}(n-t)^2t^2n^{-10/3}+O(n^{11/3}).$$

Notice, that the coefficient of $S(i)$ in the left-hand side is positive. The main term in the right-hand side is of order at least n^4 and is also positive, completing the proof. $\square$

Theorem 2: For even codes, n growing, and

$$2j/n\in\left(\frac{(1-2\delta')^2}{2},1-\frac{(1-2\delta')^2}{2}\right) \tag{15}$$

we have

$$\log B_{2j}=\log\frac{\binom{n}{2j}}{|C'|}+O(\log n).$$

Proof: Denote

$$a=\left(\frac{2t^2(n-t)^2}{(n-2t)^2}\right)^{1/3}.$$

For t linear in n, $a=O(n^{2/3})$. Choose $\beta(x)$ as given by (12). Then by Lemmas 4 and 7 we have

$$\beta_{2j}B_{2j}<\frac{|C|}{2^n}\beta(0)+\sum_{i=0}^{a}|\beta_i|B_i=I_1+I_2. \tag{16}$$

Assume $j\le n/4$ (for $j>n/4$ the proof is similar, see the end of the proof of Theorem 1). Given j we choose t to be the nearest integer to $n/2-\sqrt{n(n-4j)}/2$, that is, $j\approx t-t^2/n$. Of course, for sufficiently large n we may just put $j=t-t^2/n$. For such the choice we have

$$\beta(0)=\frac{(n-2t)^2}{4}\binom{n}{t}^2$$

$$\frac{I_1}{\beta_{2j}}=\frac{|C|\binom{n}{t}^2}{2^n\binom{2j}{j}\binom{n-2j}{t-j}}\,\frac{(n-2t)^2}{4S(j)}=\frac{\binom{n}{2j}}{|C'|}\,\frac{\binom{n}{t}^2}{\binom{2j}{j}\binom{n-2j}{t-j}\binom{n}{2j}}\,\frac{(n-2t)^2}{4S(j)}$$

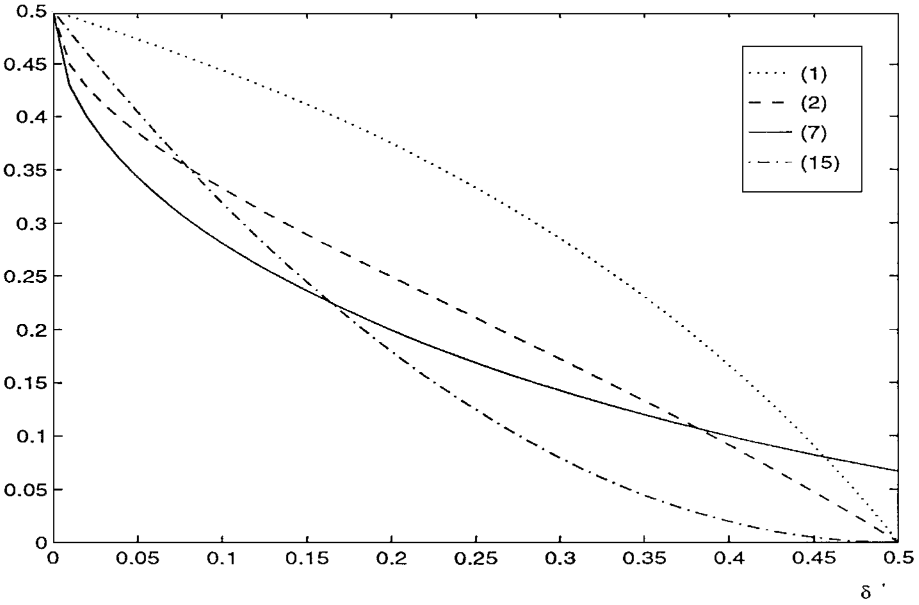

Fig. 1. Bounds for the range of binomiality.

where $S(j)$ is defined by (14). Now, we show that

$$\frac{1}{n}\log\left(\frac{\binom{n}{t}^2}{\binom{2j}{j}\binom{n-2j}{t-j}\binom{n}{2j}}\,\frac{(n-2t)^2}{4S(j)}\right)=O\left(\frac{\log n}{n}\right). \tag{17}$$

Indeed, denoting $\tau = t/n, \xi = j/n,$ we have for the last expression

$$2H(\tau)-2\xi\log 2-(1-2\xi)H\left(\frac{\tau-\xi}{1-2\xi}\right)-H(2\xi)+O\left(\frac{\log n}{n}\right).$$

Substituting $\xi = \tau - \tau^2$, after straightforward calculations we get (17).

Let us estimate I_2. Using trivial

$$B_{2i}<\binom{n}{2i}$$

we get

$$\frac{1}{n}\log I_2<\max_{0\le 2i\le a}\log\left(\binom{n}{2i}\binom{2i}{i}\binom{n-2i}{t-i}\right)+O\left(\frac{\log n}{n}\right).$$

For t linear in n the maximum is achieved at $i = a/2$, and it yields

$$\frac{1}{n}\log I_2<H(\tau)+O\left(\frac{\log n}{n}\right).$$

Now, calculations give

$$\begin{aligned}\frac{1}{n}\log\left(\frac{I_2}{\beta_{2j}}\right)&<H(\tau)-2\xi\log 2-(1-2\xi)\\&\quad\cdot H\left(\frac{\tau-\xi}{1-2\xi}\right)+O\left(\frac{\log n}{n}\right)\\&=H(2\xi)-H(\tau)+O\left(\frac{\log n}{n}\right).\end{aligned}$$

By Lemma 5, (13), and since $H(x)$ is increasing on $[0, 0.5)$

$$\begin{aligned}H(\tau)&\ge H\left(\frac{1}{2}-\sqrt{\delta'(1-\delta')}\right)\\&\ge 1-\left(1-H\left(\frac{1}{2}-\sqrt{\delta'(1-\delta')}\right)\right)\\&\ge 1-\frac{1}{n}\log|C|.\end{aligned}$$

Hence

$$\frac{1}{n}\log\left(\frac{I_2}{\beta_{2j}}\right)\le\frac{1}{n}\log\left(\frac{I_1}{\beta_{2j}}\right)+O\left(\frac{\log n}{n}\right).$$

Thus

$$\frac{1}{n}\log B_{2j}\le\log\frac{\binom{n}{j}}{|C'|}+O\left(\frac{\log n}{n}\right).$$

Now, from (13)

$$\xi\ge\left(\tfrac{1}{2}-\delta'\right)^2$$

which completes the proof. □

Comparison of the bounds in Theorems 1 and 2 shows that the first bound is better for $\delta' < 0.164\cdots$. Fig. 1 presents graphs for dependence of the left-end point of the binomiality interval as a function of δ' (recall that the right-end point is symmetric with respect to $1/2$). Here (1), (2), and (15) are valid for even codes, and (7) requires the dual code to be even as well.

IV. Conclusion

In this correspondence we have derived new upper bounds for spectral components. This approach provides a better estimate for the interval of binomiality in the spectra of codes. Notice that similar results can be obtained for the case of d' close to $n/2$, using a more accurate analysis. On the other hand, the proposed approach does not seem (at least to the authors) to give an easy way to establish lower bounds for spectral components. In contrast, in our previous paper [8], such estimates were derived, showing that, in a certain sense, the binomial upper bound is tight. Such estimates can be of importance in analysis of self-dual codes. For instance, if one could establish an existence of components achieving (asymptotically) the binomial upper bound in the interval guaranteed by Theorem 2, it would imply that for the self-dual codes

$$\delta\le\frac{(1-2\delta)^2}{2}$$

giving $\delta \le (3-\sqrt{5})/4 < 0.191$.

REFERENCES

[1] Th. Beth, H. Kalouti, and D. E. Lazic, "Which families of long binary linear codes have a binomial weight distribution?," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes* (*Proc. 11th Int. Symp.*), G. Cohen, M. Giusti, and T. Mora, Eds. (Paris, France, 1995), pp. 120–130.

[2] E. L. Blokh and V. V. Zyablov, *Linear Concatenated Codes*. Moscow, USSR: Nauka, 1982 (in Russian).

[3] Ph. Delsarte, "An algebraic approach to the association schemes of coding theory," *Philips Res. Rep. Suppl.*, no. 10, 1973.

[4] G. Kalai and N. Linial, "On the distance distribution of codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1467–1472, 1995.

[5] T. Kasami, T. Fujiwara, and S. Lin, "An approximation to the weight distribution of binary linear codes," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 769–780, 1985.

[6] I. Krasikov and S. Litsyn, "On spectra of BCH codes," *IEEE Trans. Inform. Theory*, vol. 41, pp. 786–788, 1995.

[7] ——, "On accuracy of binomial approximation to the distance distribution of codes," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1472–1474, 1995.

[8] ——, "Bounds on spectra of codes with known dual distance," submitted for publication.

[9] ——, "On integral zeroes of Krawtchouk polynomials," *J. Comb. Theory*, ser. A, vol. 74, no. 1, pp. 71–99.

[10] V. Levenshtein, "Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1303–1321, 1995.

[11] J. H. van Lint, *Introduction to Coding Theory*. New York: Springer-Verlag, 1992.

[12] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. New York: North-Holland, 1977.

[13] R. J. McEliece, E. R. Rodemich, H. Rumsey, Jr., and L. R. Welch, "New upper bounds on the rate of a code via the Delsarte–McWilliams inequalities," *IEEE Trans. Inform.Theory*, vol. IT-23, pp. 157–166, 1977.

[14] V. M. Sidelnikov, "Weight spectrum of binary Bose–Chaudhuri–Hocquenghem codes," *Probl. Pered. Inform.*, vol. 7, no. 1, pp. 14–22, 1971 (in Russian).

[15] P. Solé, "A limit law on the distance distribution of binary codes," *IEEE Trans. Inform. Theory*, vol. 36, pp. 229–232, 1990.

[16] A.Tietäväinen, "An upper bound on the covering radius as a function of the dual distance," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1472–1474, 1990.

Lengthening and the Gilbert–Varshamov Bound

Yves Edel and Jürgen Bierbrauer

***Abstract*— We use lengthening and an enhanced version of the Gilbert–Varshamov lower bound for linear codes to construct a large number of record-breaking codes. Our main theorem may be seen as a closure operation on databases.**

***Index Terms*—Gilbert–Varshamov bound, lengthening, linear codes.**

I. INTRODUCTION

Let q be a prime power, which will be fixed throughout the discussion. Denote by $\mathbb{F}_q$ the field of q elements and by $V(n,i)$ the number of vectors of weight at most i in $\mathbb{F}_q^n$. It is clear that

$$V(n,i) = \sum_{j=0}^{i} \binom{n}{j} (q-1)^j. \qquad (1)$$

Let $\mathcal{C}$ be a q-ary code with parameters $[n, k-1, d]$. As $\mathcal{C}$ has q^{k-1} elements it follows that if $q^{k-1}V(n, d-1) < q^n$, there is a vector $v \in \mathbb{F}_q^n$, that has distance $\geq d$ from every codeword $\in \mathcal{C}$. This leads to the Gilbert–Varshamov bound:

Theorem 1 (Gilbert–Varshamov Bound): If

$$V(n, d-1) < q^{n-k+1}$$

then a q-ary linear code with parameters $[n, k, d]$ exists.

Using orthogonal arrays the following can be proved.

Theorem 2: If $V(n-1, d-2) < q^{n-k}$, then a q-ary linear code with parameters $[n, k, d]$ exists. Moreover, every code $[n-1, k-1, d]$ can be embedded in a code $[n, k, d]$.

This can be found in MacWilliams and Sloane [3, p. 34]. For the sake of completeness we shall give a proof in the final section. It is easy to see that this is always stronger than the Gilbert–Varshamov bound. Combining Theorem 2 with the method of lengthening yields new codes:

Theorem 3: Assume $V(n–1, d–2) < q^{n-k}$. If codes $[n–i, k–i, d+\delta]$ and $[e, i, \delta]$ exist, then a code $[n+e, k, d+\delta]$ can be constructed.

A proof of Theorem 3 will be given in the following section. It should be noted that Theorem 3 uses only the code parameters. No information on subcodes is needed. We like to think of it as of a closure operation on databases. In order to illustrate its use we give a binary example: a code $\mathcal{D}$ with parameters $[126, 36, 34]$ is known to exist. It can be derived from a $[128, 36, 36]$ constructed in [4]. As $V(126, 26) < 2^{90}$ it follows from Theorem 2 that $\mathcal{D}$ can be embedded in a code $\mathcal{C}$ with parameters $[127, 37, 28]$. Applying construction X to the pair $\mathcal{C} \supset \mathcal{D}$ with $[6, 1, 6]$ as auxiliary code yields the new code $[133, 37, 34]$.

In Table I we list additional applications of Theorem 3. In all cases $i = 1$, so that the auxiliary code is the repetition code $[e, i, \delta] = [\delta, 1, \delta]$. The following parameters are given:

- $q \in \{2, 3, 4\}$,
- the parameters $[n-1, k-1, d+\delta]$ of the known code $\mathcal{D}$,
- δ,
- the parameters $[n, k, d+\delta]$ of the resulting code $\mathcal{E}$.

Manuscript received June 23, 1995; revised July 25, 1996.

Y. Edel is with the Mathematisches Institut der Universität, 69120 Heidelberg, Germany.

J. Bierbrauer is with the Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931 USA.

Publisher Item Identifier S 0018-9448(97)02317-1.