BOUNDS ON DISTANCE DISTRIBUTIONS FOR CODES OF KNOWN SIZE

Alexei Ashikhmin

- Distance distribution of block codes
- Previously known estimates on distance distribution
- Beckner/Nelson inequality
- New estimates on distance distribution
- Applications

This is a joint work with G.Cohen, M.Krivelevich, and S.Litsyn

Block Codes

Let C be an $(n, M, d, d^{\perp})$ binary block code.

- length n

- size $|C| = 2^k$ or rate $R = \frac{\log_2 M}{n}$

- minimum distance d

- dual distance $d^{\perp}(C)$

Distance Distribution of Block Codes

Def. 1 *The distance distribution of* C*:*

$$A_i = \frac{1}{|C|}|\{(\mathbf{v}, \mathbf{w}) \in C^2 : \text{dist}\,(\mathbf{v}, \mathbf{w}) = i\}|$$

Obviously

$$A_i \geq 0, \;\; A_0 = 1,$$

and

$$A_0 + A_1 + \ldots + A_n = |C|$$

Denote

$$a_\xi = \frac{1}{n} \log_2 A_i, \;\; \xi = \frac{i}{n}, \;\; \delta = \frac{d}{n}$$

Traditionally bounds on the minimum distance were a subject of investigation

$$\delta \leq g(R) \text{ or } R \leq h(\delta)$$

The famous McEliece-Rodemich-Rumsey-Welch Linear Programming bound:

$$R \leq R_{LP}(\delta) = H\left(\frac{1}{2} - \sqrt{\delta\,(1-\delta)}\right)$$

Here $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$

Our goal is to construct asymptotic bounds on a_ξ as function of R and δ as $n \to \infty$

$$A_i \leq f(n, M, d), \text{ and } a_\xi \leq f(R, \delta)$$

Trivial bounds:

$$\begin{aligned} A_i &\leq |C| \quad \text{or} \quad a_\xi \leq R \\ A_i &\leq \binom{n}{i} \quad \text{or} \quad a_\xi \leq H(\xi) \end{aligned}$$

We will show that a_ξ is strictly lower than $H(\xi)$

(Linear programming approach (previous results)

Delsarte Inequalities:

$$\frac{1}{|C|}\sum_{i=0}^{n} A_i K_j(i) \ \geq 0, \quad j = 0, \ldots, n$$

where

$$K_j(i) = \sum_{t=0}^{j} (-1)^t \binom{i}{t}\binom{n-i}{j-t}$$

are Krawtchouk polynomials.

C is a code with minimum distance d and dual minimum distance $d^{\perp}$

Let $Z_w(i) = \sum_{j=0}^{n} z_j K_j(i)$ be s.t.

- $Z_w(i) \geq g_w(i), i = d, \ldots, n$
- $z_j \leq 0, j = d^{\perp}, \ldots, n$

Theorem 1 *(Ashikhmin, Barg, Litsyn, 2001)*

$$\sum_{i=d}^{n} g_w(i) A_i \leq z_0 |C| - Z_w(0)$$

Theorem 2 *(Ashikhmin, Barg, Litsyn, 2001)*

$$a_\xi(\delta) \leq H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right) + H(\xi) - 1 + o(1), \quad (1)$$

$$a_\xi(\delta) \leq H(\xi) - H\left(\frac{\delta}{2}\right) + o(1), \quad (2)$$

A number of other estimates for A_i are known: G.Kalai and N.Linial (1995), N. Linial and A.Samorodnit 2002.

Good bounds for good codes

Theorem 3 *(Ashikhmin, Barg, Litsyn, 2001) Let C be such that its minimum distance meets either*

- *McEliece-Rodemich-Rumsey-Welch Linear Programming bound or*
- *the Hamming bound.*

Then

$$a_\xi = H(\xi) + R - 1, \ \xi \geq \delta(C),$$

that is a_ξ are the same as in random codes.

Ex. (Ashikhmin, Barg, Litsyn)

Let C have length $n = 47$,
size $|C| = 9 \cdot 2^{38}$, and minimum distance $d = 3$.

i	3	4	5	6	7
$\log \overline{A}_i$	8.483	11.973	14.971	17.815	20.345
$\log \underline{A}_i$	7.938	11.675	14.312	17.468	19.643

8	14	18	22	24
22.714	32.874	36.722	38.616	38.904
22.333	32.306	35.881	37.146	36.650

These bounds are so good because C almost meets the Hamming bound on minimum distance

$$\sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} \leq \frac{2^n}{|C|}$$

Beckner/Nelson Inequality

Let $f\colon \{0,1\}^n \to \mathbb{R}$. For an $s \in \mathbb{R}, s > 0$ the s-norm of f is

$$\|f\|_s = \left(\frac{1}{2^n} \sum_{\mathbf{v} \in \{0,1\}^n} |f(\mathbf{v})|^s \right)^{1/s}$$

Define for $0 < \varepsilon < 1$

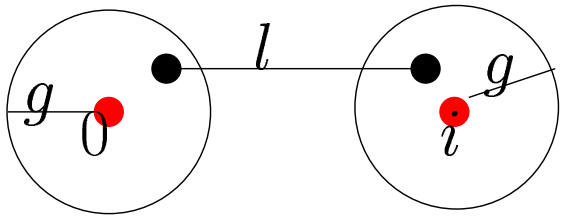
$$T_\varepsilon(\mathbf{v}) = \sum_{\mathbf{u} \in \{0,1\}^n} f(\mathbf{u}) \left(\frac{1+\varepsilon}{2} \right)^{n-d(\mathbf{u},\mathbf{v})} \left(\frac{1-\varepsilon}{2} \right)^{d(\mathbf{u},\mathbf{v})}$$

Theorem 4 (Beckner/Nelson)

$$\|T_\varepsilon\|_2 \leq \|f\|_{1+\varepsilon^2}$$

Let

$$h(i,l) = |\{(\mathbf{u}_1,\mathbf{u}_2) : \mathbf{u}_1 \in B(0^n,g), \mathbf{u}_2 \in B(1^i0^{n-i},g) : d(\mathbf{u}_1,\mathbf{u}_2) = l\}|$$



$V(g) = \sum_{i=0}^{i\leq g} \binom{n}{i}$ is the volume of the binary ball of radius g

$A(n,d,\leq g)$ is an upper bound on the number of n-vectors of weight g of less being at Hamming distance at least d

Theorem 5 *For any code C of length n, minimum distance d, and parameters $p \in [0,1/2]$, $0 \leq g \leq \frac{1}{2}$:*

$$\sum_{i=0}^{n} A_i \sum_{l=0}^{n} h(i,l) p^l (1-p)^{n-l}$$

$$\leq \left(V(g) A^{1-2p}(n,d,\leq g)\right)^{\frac{1}{1-p}} \left(\frac{|C|}{2^n}\right)^{\frac{p}{1-p}},$$

Corollary 6

$$\begin{aligned} & a_\mu(R,\delta) \\ \leq\ & \min\Big\{-\frac{p}{1-p}(1-R)+\frac{1}{1-p}H(\gamma)+\frac{1-2p}{1-p}R(\delta,\gamma) \\ & \quad -\hbar(\mu,\upsilon)-\upsilon\log p-(1-\upsilon)\log(1-p)\},\end{aligned}$$

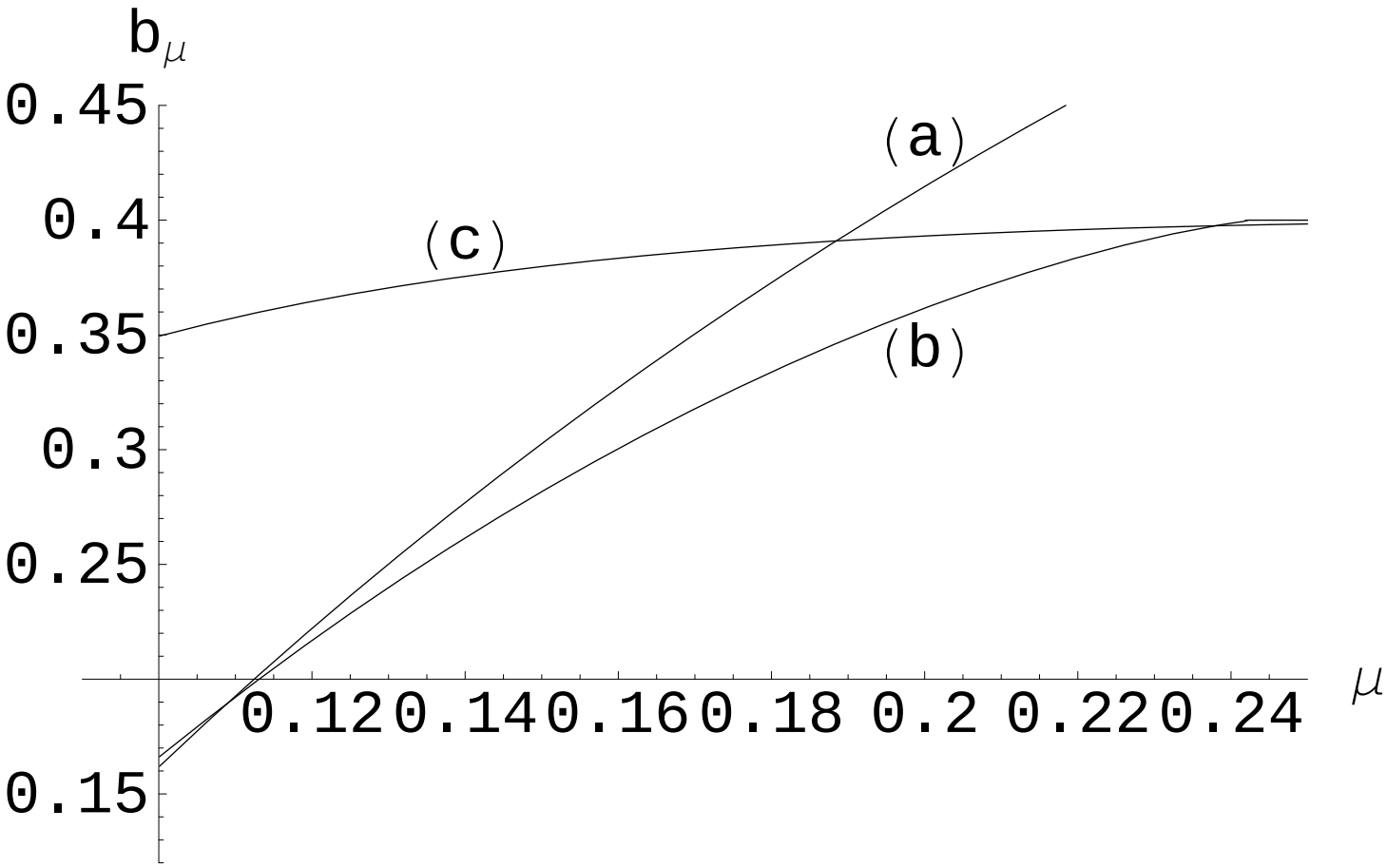
where the minimum is taken for $p \in [0,1/2], \gamma \in [\frac{1}{2}-\sqrt{\frac{1}{4}-\frac{\delta}{2}}, 1/2]$, *and* $\upsilon = [0,1]$ *and*

$$\begin{aligned} R(\delta,\gamma) &= \frac{1}{n}\log A(n,\delta n,\leq \gamma n) \\ \hbar(\xi,\upsilon) &= \frac{1}{n}\log h(\xi n,\upsilon n)\end{aligned}$$

Example 1 $R = 0.4,\ \delta = 0.1$

LP bound: $R \leq R_{LP}(0.1) = 0.693$

Gilbert-Varshamov (existence bound): there exists a code with $R \geq R_{VG}(0.1) = 0.531$



- a) is the best LP bound
- b) is the new bound
- c) Extension of Linial, Samorodnitsky bound for linear codes

Setting $g = 0$ in Theorem 5, we obtain

Theorem 7 *If* $0 < R < 1$, *and*

$$0 < \mu \leq 1 - 2\sqrt{(1-R)\ln 2} + \ln 2 - R\ln 2$$

then

$$a_\mu(R) = a_{1-\mu}(R) \leq \tag{3}$$

$$\leq -\frac{p^*}{1-p^*}(1-R) - \mu \log p^* - (1-\mu)\log(1-p^*) + o(1)$$

where

$$\begin{aligned} p^* &= \frac{1}{2}\Big(1 + \mu - \ln 2 + R\ln 2 \\ &\quad -\sqrt{-4\mu + (1+\mu-\ln 2 + R\ln 2)^2}\Big) \end{aligned}$$

Setting $p^\star = \mu$ in Theorem 7 we obtain

Corollary 8

$$a_\mu(R) = a_{1-\mu}(R) \leq -\frac{\mu}{1-\mu}(1-R) + H(\mu)$$

This shows that in any code of rate $R < 1$

$$A_i < \binom{n}{i}$$

Let us now consider the code C consisting of all vectors of weight w and their complements, and let $2\binom{n}{w} = |C|$.

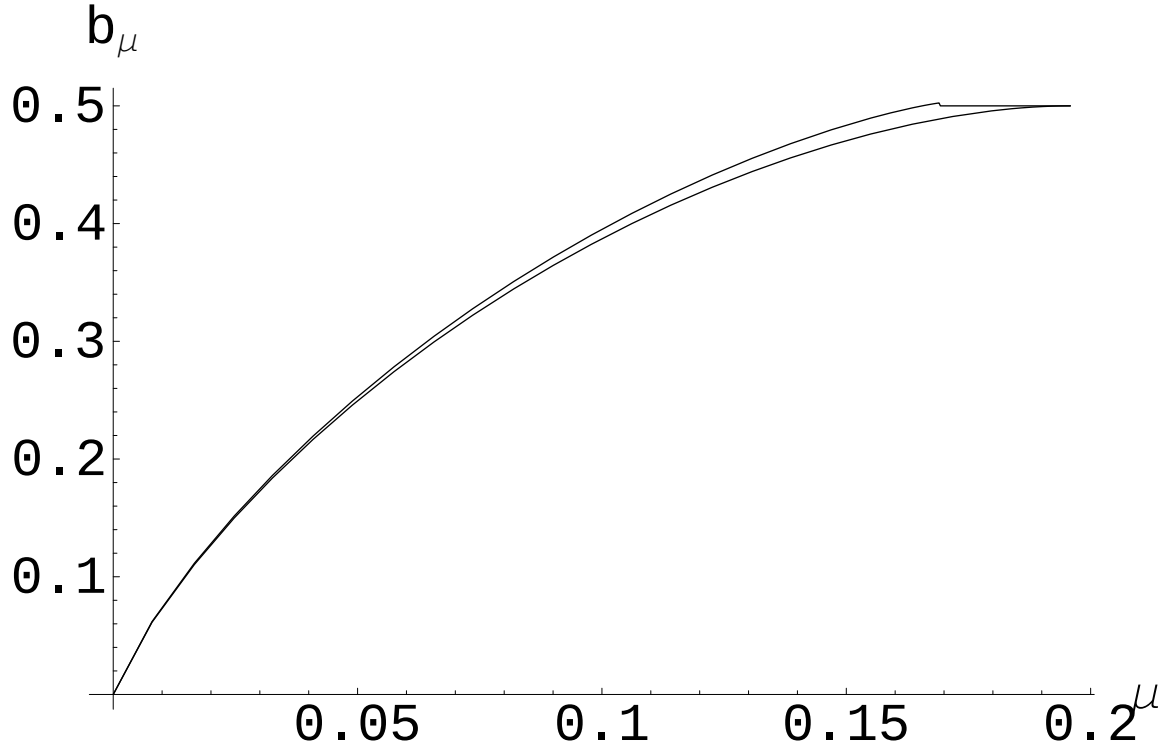
Theorem 9 *(Low bound) For any* $0 \leq \mu \leq 2\omega(1-\omega)$

$$a_\mu(R) = a_{1-\mu}(R) \geq \omega H\left(\frac{\mu}{2\omega}\right) + (1-\omega)H\left(\frac{\mu}{2(1-\omega)}\right)$$

where $\omega = H^{-1}(R)$. *Otherwise, for* $\mu \leq 0.5$,

$$a_\mu(R) = a_{1-\mu}(R) = R$$

Example 2 $R = 0.5$



Applications

Undetected Error Probability

Theorem 10 *Every code C used on a BSC channel with transition probability $0 \leq \rho \leq \frac{1}{2}$ has*

$$P_{ue}(C, \rho) \leq \left(\frac{|C|}{2^n}\right)^{\frac{\rho}{1-\rho}} - (1-\rho)^n$$

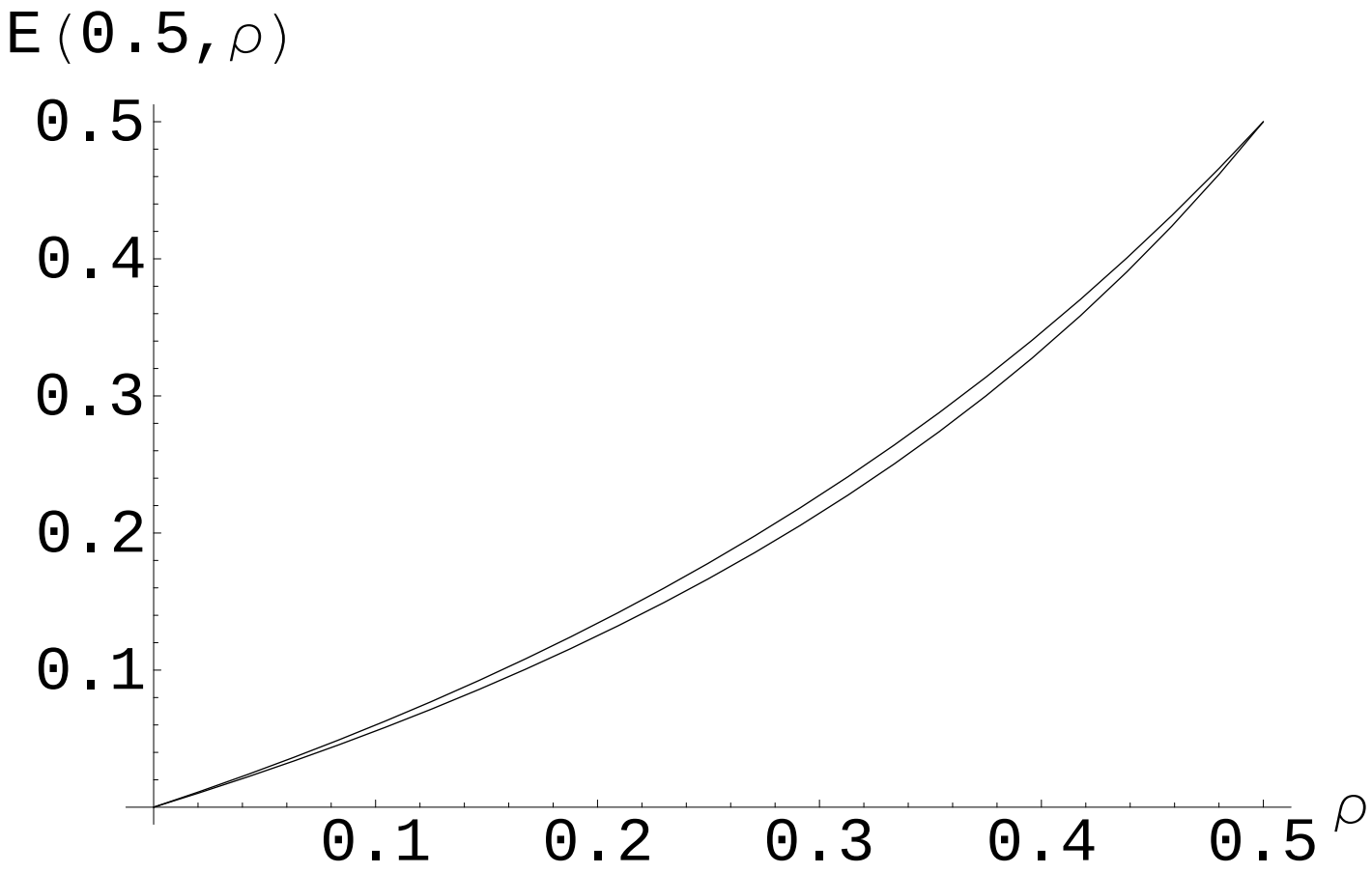
Theorem 11 *For the constant weight code consisting of all vectors of weight w and $0 \leq \rho \leq 1/2$,*

$$\frac{1}{n}\log_2 p_{ue}(R, \rho) \geq -\left(\omega H\left(\frac{\xi}{\omega}\right) + (1-\omega)H\left(\frac{\xi}{1-\omega}\right) + 2\xi \log \rho + (1-2\xi)\log(1-\rho)\right)$$

where

$$\omega = H^{-1}(R),$$

$$\xi = \frac{\rho^2 - \rho\sqrt{\rho^2 - 4\omega(1-\omega)(2\rho-1)}}{2(2\rho-1)}$$



$$E(R,\rho) = -\frac{1}{n}\log_2 P_{ue}(C,\rho)$$

$$R = 0.5$$

The Threshold Probability

Let C be a code of rate R and relative minimum distance δ.

Theorem 12 *(Tillich and Zemor, 2000) Let ϑ be such that*

$$H(\vartheta) < \max_{(\delta/\vartheta)<\alpha\leq 2} \left\{ a_{\alpha\vartheta}(R,\delta) + \alpha\vartheta + (1-\alpha\vartheta) H\left(\frac{(1-\alpha/2)\vartheta}{1-\alpha\vartheta}\right) \right\}$$

Then ML decoding of C in BSC fails to decode only vanishing fraction of error vectors of weight ϑn.

The previous bounds on ϑ were obtained with the help of upper estimates on a_ξ as functions of δ.

Example 3 *Let $\delta = 0.25$ then the threshold $\vartheta =$* 0.165.

The new bounds are functions of R and δ:

R	0.1	0.12	0.14	0.16	0.18
ϑ	0.2204	0.2019	0.1866	0.1736	0.1626

Covering Radius

Let C be a code of rate R and relative dual minimum distance $\delta^{\perp}$.

Theorem 13 *Let ρ be the minimal number such that*

$$\max_{\delta^{\perp} \leq \xi \leq 2\rho} \{(1-\xi) H\left(\frac{\rho - \xi/2}{1-\xi}\right) - H(\rho) + \xi + b_{\xi}(R, \delta^{\perp}).$$

Then the covering radius of C satisfies $\mathcal{R}(C)/n \leq \rho$.

Traditionally covering radius was estimated as a function of the code dual distance

We estimate it as a function of both size and dual distance

Example 4 *Let $\delta^{\perp} = 0.25$ then the best bound is $\mathcal{R}/n \leq 0.1291$*

New bounds as function of R and $\delta^{\perp} = 0.25$:

$1-R$	0.1	0.12	0.14	0.16	0.18
$\mathcal{R}/n$	0.0854	0.0985	0.110	0.1213	0.1288