Bounds on Distance Distributions in Codes of Known Size

Alexei Ashikhmin
Bell Laboratories, Lucent Technologies, 600 Mountain Avenue, Murray Hill, NJ 07974, USA, e-mail: aea@research.bell-labs.comi.

Gérard Cohen
Département Informatique et Réseaux, ENST, 46 Rue Barrault, Paris, France; e-mail: cohen@inf.enst.fr.

Michael Krivelevich Simon Litsyn
Tel Aviv University, Ramat-Aviv, 69978 Israel; e-mail: krivelev@math.tau.ac.il litsyn@eng.tau.ac.il

Abstract — **We treat the problem of bounding components of the possible distance distributions of codes given the knowledge of their size and possibly minimum distance. Using the Beckner inequality from Harmonic Analysis we derive upper bounds on distance distribution components which are sometimes better than earlier ones due to Ashikhmin, Barg and Litsyn. We use an alternative approach to derive upper bounds on distance distributions in linear codes. As an application of the suggested estimates we get an upper bound on the undetected error probability for an arbitrary code of given size. We also use the new bounds to derive better upper estimates on the covering radius, as well as a lower bound on the error-probability threshold, as a function of the code's size and minimum distance.**

Let F^n be the space of binary vectors of length n endowed with the Hamming metric $d(\cdot,\cdot)$, for $\mathbf{x}=(x_1,...,x_n), \mathbf{y}=(y_1,...,y_n), \mathbf{x},\mathbf{y}\in F^n$, $d(\mathbf{x},\mathbf{y})=|\{i: x_i\neq y_i\}|$. Let the (Hamming) weight of $\mathbf{x}\in F^n$ be $wt(\mathbf{x})=|\{i: x_i=1\}|$, i.e. the number of ones in $\mathbf{x}$. Let $B(\mathbf{x},r)\subseteq F^n$ stand for the ball of radius r centered at $\mathbf{x}$, $V(r)=\sum_{i=0}^{r}\binom{n}{i}$, being its volume. Let $C\subseteq F^n$ be a code of rate $R(C)=R=\frac{1}{n}\log_2|C|$. Assume that the minimum distance $d(C)$ of the code, $d(C)=d=\min_{\mathbf{c}_1,\mathbf{c}_2\in C} d(\mathbf{c}_1,\mathbf{c}_2)$, is $d=\delta(C)n$, where $\delta=\delta(C)$ is the code's relative distance. Denote by $\mathbf{B}(C)=(B_0(C)=1,B_1(C),\ldots,B_n(C))$ the distance distribution of the code, i.e. $B_i(C)=B_i=\frac{1}{|C|}|\{\mathbf{c}_1,\mathbf{c}_2:\mathbf{c}_1,\mathbf{c}_2\in C; d(\mathbf{c}_1,\mathbf{c}_2)=i\}|$. For $\mathbf{c}\in C$ let $A_i^C(\mathbf{c})=A_i(\mathbf{c})=|\{\mathbf{c}_1\in C: d(\mathbf{c},\mathbf{c}_1)=i\}|$. Notice that $B_i=\frac{1}{|C|}\sum_{\mathbf{c}\in C}A_i(\mathbf{c})$. Whenever we deal with a linear code (i.e. a code closed under component-wise modulo 2 sum) then for every $\mathbf{c}\in C$ we have $A_i(\mathbf{c})=B_i$. We will also be using the exponents $b_\xi(C)$ of $B_{\lfloor\xi n\rfloor}(C)$, namely, $b_\xi(C)=b_\xi=\frac{1}{n}\log_2 B_{\lfloor\xi n\rfloor}$.

Problem statement: *Given R, δ, and ξ, such that $0\le R\le 1$, $0\le\delta\le 1/2$, and $\delta\le\xi\le 1$, we wish to estimate*

$$b_\xi(R,\delta):=\limsup_{n\to\infty}\max_C b_\xi(C) \tag{1}$$

where the maximum is taken over all codes C of length n, rate at most R and minimum distance at least δn.

The main inequality we prove is

Theorem 1 *For any code C of length n, minimum distance d, and parameters $p\in[0,1/2]$, $g\in[0,\frac{n}{2}]$, g being an integer, the following inequality holds:*

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

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

where $h(i,\ell)=$

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

Corollary 1 *With the above notation, for any $p\in[0,1/2]$, and $0\le\gamma<\gamma_E:=\frac{1}{2}-\frac{1}{2}\sqrt{1-2\delta}$,*

$$\sum_{i=0}^{n}B_i\sum_{\ell=0}^{n}h(i,\ell)p^\ell(1-p)^{n-\ell}\le V^{\frac{1}{1-p}}(\gamma n)n^\alpha\left(\frac{|C|}{2^n}\right)^{\frac{p}{1-p}}.$$

Here α is a non-negative constant depending on γ. ◇

This yields the following bound on distance components.

Theorem 2 *If $0<R<1$, and*

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

then $b_\mu(R)\le$

$$\le-\frac{p^*}{1-p^*}(1-R)-\mu\log p^*-(1-\mu)\log(1-p^*)+o(1), \tag{2}$$

where $p^*=$

$$\frac{1}{2}\left(1+\mu-\ln 2+R\ln 2-\sqrt{-4\mu+(1+\mu-\ln 2+R\ln 2)^2}\right).$$

Otherwise, the trivial bound holds:

$$b_\mu(R)\le R.$$

The same bounds are valid also for $b_{1-\mu}(R)$.

The results can be applied to the problem of estimation of undetected error probability.

Theorem 3 *Every code C used on a BSC channel with transition probability $\rho\in[0,\frac{1}{2}]$ satisfies:*

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

In the case of linear codes we use an alternative approach.

Theorem 4 *The distance distribution of a linear code C is bounded from above as follows*

$$b_\mu^L(R,\delta)\le\begin{cases}\min_{0\le\alpha\le 1}\left\{\alpha RH\left(\frac{w}{\beta^*(\alpha)}\right)+(1-\alpha)R\right\},\\ \text{if } \beta^*(\alpha)>2\mu;\\ R,\ \text{otherwise.}\end{cases} \tag{3}$$

where $\beta^*(\alpha)$ is the root of the equation

$$\frac{(1-\alpha)R}{\beta}=R^*\left(\frac{\delta}{\beta}\right). \tag{4}$$