[7] E. N. Gilbert, "A comparison of signaling alphabet," *Bell Syst. Tech. J.*, vol. 31, pp. 504–522, 1952.
[8] J. Kahn, G. Kalai, and N. Linial, "The influence of variables on Boolean functions," in *Proc. 29th Ann. Symp. on Foundations of Computer Science*. Los Alamitos, CA: Computer Soc. Press, 1988, pp. 68-80.
[9] G. A. Kabatiansky and V. I. Levenshtein, "Bounds on packing on a sphere and in space," *Probl. Inform. Transmission*, vol. 14, pp. 1–17, 1978.
[10] R. J. McEliece and H. C. Rumsey, "Sphere-packing in the Hamming metric," *Bull. Amer. Math. Soc.*, vol. 75, pp. 32–34, 1969.
[11] R. J. McEliece, E. R. Rodemich, H. C. Rumsey, and L. R. Welch, "New upper bounds on the rate of codes via the Delsarte-MacWilliams inequalities," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 157–166, 1977.
[12] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.
[13] N. J. A. Sloane, "Recent bounds for codes, sphere packings and related problems obtained by linear programming and other methods," *Contemp. Math.*, vol. 9, pp. 153–185, 1982.
[14] J. H. van Lint, *Introduction to Coding Theory*. New York: Springer-Verlag, 1982.

On the Accuracy of the Binomial Approximation to the Distance Distribution of Codes

Ilia Krasikov and Simon Litsyn

***Abstract*—The binomial distribution is a well-known approximation to the distance spectra of many classes of codes. We derive a lower estimate for the deviation from the binomial approximation.**

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

I. Introduction

The binomial distribution is a well-known approximation to the distance spectra of many classes of codes. For example, it is known to be tight for the weights of BCH codes (see, e.g. [7, sec. 9.10]). Several upper bounds for the error term of such approximation have been derived in [1], [2], [4], [8], [9]. These estimates show that, provided the dual distance is large enough, the spectrum of the code rapidly converges to the binomial distribution. How close can the real distribution be to the binomial one? In this correspondence we give a lower estimate for the deviation from the binomial approximation thus showing that it cannot be too sharp. We also establish an identity relating the error terms to the dual spectrum of a code.

II. Results

We start with the following auxiliary lemma [3]. The proof is presented for self-completeness.

Manuscript received July 19, 1994; revised February 2, 1995. This research was partially supported by the Guastallo Fellowship and a Grant from the Israeli Ministry of Science and Technology.

I. Krasikov is with Tel-Aviv University, School of Mathematical Sciences, Ramat-Aviv 69978, Tel-Aviv, Israel, and Beit-Berl College, Kfar-Sava, Israel.
S. Litsyn is with Tel-Aviv University, Department of Electrical Engineering–Systems, Ramat-Aviv 69978, Tel-Aviv, Israel.

IEEE Log Number 9413879.

Lemma 1: Let $\mathcal{F}$ be the set of real monic polynomials of degree c. Define

$$E(a,b,c) = \min_{f\in\mathcal{F}} \max_{x\in[-1,1]} |(1-x)^a\,(1+x)^b f(x)|.$$

Then

$$\frac{4^{a+b+c}}{2a+2b+2c+1} \le \binom{2a+2b+2c}{c}\binom{2a+2b+2c}{2a+c}(E(a,b,c))^2$$

provided $a, b > -1/2$, real.

Proof: Let f be an optimal polynomial. Expand f in the series of Jacobi polynomials (see, e.g., [10])

$$f(x) = \sum_{j=0}^{c} q_j P_j^{(2a,2b)}(x)$$

where

$$(1-x)^\alpha(1+x)^\beta P_j^{(\alpha,\beta)}(x) = \frac{(-1)^j}{2^j j!}\frac{d^j}{dx^j}((1-x)^{\alpha+j}(1+x)^{\beta+j}).$$

The leading coefficient of $P_j^{(\alpha,\beta)}(x)$ is $2^{-j}\binom{\alpha+\beta+2j}{j}$, and so

$$q_c = \frac{2^c}{\binom{2a+2b+2c}{c}}.$$

The orthogonality relation for Jacobi polynomials is given by

$$\begin{aligned} g_{jl}(\alpha,\beta) &= \int_{-1}^{1}(1-x)^\alpha(1+x)^\beta P_j^{(\alpha,\beta)}(x)P_l^{(\alpha,\beta)}(x)\,dx \\ &= \frac{2^{\alpha+\beta+1}\Gamma(j+\alpha+1)\Gamma(j+\beta+1)}{(2j+\alpha+\beta+1)\Gamma(j+1)\Gamma(j+\alpha+\beta+1)}\delta_{jl} \end{aligned}$$

where δ_{jl} is the Kronecker delta.

Now we get

$$\begin{aligned} &\max_{x\in[-1,1]}(1-x)^a(1-x)^b(f(x))^2 \\ &\ge \frac{1}{2}\int_{-1}^{1}(1-x)^{2a}(1+x)^{2b}(f(x))^2\,dx \\ &\ge \frac{1}{2}\sum_{j=0}^{c}\sum_{l=0}^{c} q_j q_l \int_{-1}^{1}(1-x)^{2a}(1+x)^{2b}P_j^{(2a,2b)}(x) \\ &\quad\cdot P_l^{(2a,2b)}(x)\,dx \\ &\ge \frac{1}{2}q_c^2\; g_{cc}(2a,2b) \end{aligned}$$

and we are done. □

The binary Krawtchouk polynomial $P_k^n(x)$ (of degree k in x) is defined by the following generating function:

$$\sum_{k=0}^{\infty} P_k^n(x)z^k = (1-z)^x(1+z)^{n-x}. \tag{1}$$

When it does not lead to confusion n is omitted, i.e., $P_k(x) = P_k^n(x)$. The following values are of importance for us:

$$P_i(0) = \binom{n}{i} \qquad P_n(i) = (-1)^i \qquad P_i(n) = (-1)^i\binom{n}{i}.$$

Let the distance distribution of a code C be $\underline{B} = (B_0, \cdots, B_n)$, and $\underline{B}' = (B'_0, \cdots, B'_n)$ stand for the the dual spectrum, that is, $\underline{B}'$ is determined by the MacWilliams transform of $\underline{B}$

$$B'_k = \frac{1}{|C|}\sum_{i=0}^{n} B_i P_k(i). \tag{2}$$

The inverse is given by

$$B_i = \frac{|C|}{2^n}\sum_{k=0}^{n} B'_k P_i(k). \tag{3}$$

Hence

$$B_i = \frac{|C|\binom{n}{i}}{2^n}(1+(-1)^i B'_n) + \sum_{k=1}^{n-1} B'_k P_i(k).$$

Quite often the first term turns out to be dominating. Note also that $B'_n \in [0,1]$.

Define

$$r_i = B_i - \frac{|C|\binom{n}{i}}{2^n}(1+(-1)^i B'_n).$$

This is evidently the deviation of the ith spectrum element from the "expected" value given by the binomial distribution.

Theorem 1: Let $B'_i = 0$, for $i \in [1, d'_1 - 1] \cup [d'_2 + 1, n-1]$. Then

$$\sum_{i=0}^{n} |r_i| \geq \frac{2^n - 2|C|}{\sqrt{(n+1)}} \cdot \left(\max\left\{\binom{n}{[\frac{d'_2}{2}] - [\frac{d'_1+1}{2}]}\binom{n}{[\frac{d'_2}{2}] + [\frac{d'_1+1}{2}]}, \binom{n}{[\frac{d'_2-1}{2}] - [\frac{d'_1}{2}]}\binom{n}{[\frac{d'_2-1}{2}] + [\frac{d'_1}{2}]+1}\right\}\right)^{-\frac{1}{2}}.$$

Proof: Let $\iota = \sqrt{-1}$, $\varphi \in [0, \pi/2]$, and put $x = \cos 2\varphi$. Denote also

$$a_1 = [(d'_1+1)/2],\ b_1 = [d'_2/2],\ a_2 = [d'_1/2],\ b_2 = [(d'_2-1)/2].$$

From the definition of r_i and (3)

$$r_i = \frac{|C|}{2^n}\sum_{j=d'_1}^{d'_2} B'_j P_i(j).$$

Denote by $\sum_j^{(e)}$ and $\sum_j^{(o)}$ the sums over all even (odd) $j \in [d'_1, d'_2]$.

Using (1) with $z = e^{2\iota\varphi}$, we get

$$\begin{aligned}
\frac{2^n}{|C|}\sum_{i=0}^{n}|r_i| &= \sum_{i=0}^{n}\left|\sum_j B'_j P_i(j)\right| \\
&= \left|\sum_j\sum_{i=0}^{n} e^{2i\iota\varphi} B'_j P_i(j)\right| \\
&= \left|\sum_j B'_j (1-e^{2\iota\varphi})^j (1+e^{2\iota\varphi})^{n-j}\right| \\
&= 2^n\left(\left|\sum_j B'_j(-\iota)^j \sin^j\varphi \cos^{n-j}\varphi\right|\right) \\
&= 2^{n/2}\left(\left|\sum_j^{(e)} B'_j (-1)^{j/2}(1-x)^{j/2}(1+x)^{(n-j)/2}\right.\right. \\
&\qquad \left.\left. - \iota\sum_j^{(o)} B'_j(-1)^{(j+1)/2}(1-x)^{j/2}(1+x)^{(n-j)/2}\right|\right) \\
&\geq 2^{n/2}\max\left\{\left|\sum_j^{(e)} B'_j(-1)^{j/2}(1-x)^{j/2} \cdot (1+x)^{(n-j)/2}\right|,\right. \\
&\qquad \left.\left|\sum_j^{(o)} B'_j(-1)^{(j+1)/2}(1-x)^{j/2}(1+x)^{(n-j)/2}\right|\right\} \\
&= 2^{n/2}\max\left\{(1-x)^{a_1}(1+x)^{n/2-b_1} \cdot \left|\sum_{j=a_1}^{b_1} B'_{2j}(-1)^j(1-x)^{j-a_1}(1+x)^{b_1-j}\right|,\right. \\
&\qquad (1-x)^{a_2+1/2}(1+x)^{(n-1)/2-b_2} \\
&\qquad \left.\cdot\left|\sum_{j=a_2}^{b_2} B'_{2j+1}(-1)^{j+1}(1-x)^{j-a_2}(1+x)^{b_2-j}\right|\right\} \\
&= 2^{n/2}\max\left\{(1-x)^{a_1}(1+x)^{n/2-b_1}|f_1(x)|,\right. \\
&\qquad \left.(1-x)^{a_2+1/2}(1+x)^{(n-1)/2-b_2}|f_2(x)|\right\}.
\end{aligned} \tag{4}$$

Now we apply Lemma 1 to both terms of the last expression. Observe that $f_1(x)$ and $f_2(x)$ are just polynomials in x of degrees $(b_1 - a_1)$ and $(b_2 - a_2)$, respectively. The absolute value of the leading coefficients of $f_1(x)$ and $f_2(x)$ are $\sum_j^{(e)} B'_j$ and $\sum_j^{(o)} B'_j$.

Taking into account that

$$\sum_{j=d'_1}^{d'_2} B'_j = \frac{2^n}{|C|} - 1 - B'_n \geq \frac{2^n}{|C|} - 2$$

and applying Lemma 1 with $a = a_1$, $b = n/2 - b_1$, $c = b_1 - a_1$, to the first term of (4), and with $a = a_2 + 1/2$, $b = (n-1)/2 - b_2$, $c = b_2 - a_2$, to the second one, we get the result. □

For wide classes of codes, $d'_1 = n - d'_2$. For example, it is the case when the code contains only even weight vectors. For even n and d'_1 the estimate gets the form

$$\sum_{i=0}^{n}|r_i| \geq \frac{2^n - 2|C|}{\sqrt{(n+1)\binom{n}{n/2-d'_1}\binom{n}{n/2}}}. \tag{5}$$

Consider BCH codes of distance $d = 2t+1 < \sqrt{n}$. Upper estimates for the distance of the code, obtained by extending the code dual to the BCH code, may be deduced from the lower bound on exponential sums (see, e.g. [5])

$$d'_1 \leq n/2 - c_1\sqrt{tn}$$

for some constant c_1. Then

$$\log_2 \sum_{i=0}^{n}|r_i| \geq \frac{n - c_1\sqrt{n\,t}\log n}{2}. \tag{6}$$

For constant t this estimate turns out to be asymptotically tight. This follows from results of [1], [4] where it was shown that

$$\frac{1}{n}\lim_{n\to\infty}\log_2|r_i| \leq \frac{1}{2}H\left(\frac{i}{n}\right),$$

where H is the binary entropy function.

In what follows we will derive an identity relating the deviations to the dual distance distribution. This is achieved by refining some arguments due to Gashkov and Sidelnikov [1].

We need (see, e.g., [6]) the following properties of Krawtchouk polinomials (for integer $i, j, l, k \in [0, n]$):

$$\binom{n}{j}P_i(j) = \binom{n}{i}P_j(i) \tag{7}$$

$$\sum_{i=0}^{n} P_l(i)P_i(k) = \delta_{lk}\,2^n \tag{8}$$

$$P_i(j) = (-1)^i P_i(n-j) = (-1)^j P_{n-i}(j). \tag{9}$$

Lemma 2:

$$\frac{1}{|C|}\sum_{i=0}^{n}\frac{B_i^2}{\binom{n}{i}} = \frac{|C|}{2^n}\sum_{j=0}^{n}\frac{B_j'^2}{\binom{n}{j}}.$$

Proof: We have

$$\begin{aligned}\sum_{i=0}^{n}\frac{B_i^2}{\binom{n}{i}} &= \sum_{i=0}^{n}\frac{|C|^2}{4^n\binom{n}{i}}\left(\sum_{j=0}^{n}B_j'P_j(i)\right)^2\\ &= \sum_{i=0}^{n}\frac{|C|^2}{4^n\binom{n}{i}}\sum_{j=0}^{n}\sum_{l=0}^{n}B_j'B_l'P_j(i)P_l(i)\\ &= \frac{|C|^2}{4^n}\sum_{j=0}^{n}\sum_{l=0}^{n}\frac{B_j'B_l'}{\binom{n}{j}}\sum_{i=0}^{n}P_i(j)P_l(i)\\ &= \frac{|C|^2}{2^n}\sum_{j=0}^{n}\frac{B_j'^2}{\binom{n}{j}}.\end{aligned}$$

□

Let $d' = \min\{d_1', n - d_2'\}$.

Corollary 1:

$$\sum_{i=0}^{n}\frac{B_i^2}{\binom{n}{i}} - \frac{|C|^2}{2^n} < \frac{2^n}{\binom{n}{d'}}.$$

Proof: Just follows from the evident

$$\sum_{j=d_1'}^{d_2'} B_j' \le 2^n/|C| - 1.$$

A similar bound was obtained in [1] by more complicated arguments. □

Theorem 2:

$$\sum_{i=0}^{n}\frac{r_i^2}{\binom{n}{i}} = \frac{|C|^2}{2^n}\sum_{i=d_1'}^{d_2'}\frac{B_i'^2}{\binom{n}{i}}.$$

Proof: Using (2) observe that

$$\sum_{i=0}^{n} r_i = 0.$$

by

$$\sum_{i=0}^{n} B_i = |C|$$

and

$$\begin{aligned}\sum_{i=0}^{n}(-1)^i r_i &= \sum_{i=0}^{n}(-1)^i B_i - |C|B_n'\\ &= \sum_{i=0}^{n} B_i P_n(i) - |C|B_n' = 0.\end{aligned}$$

Hence

$$\begin{aligned}\sum_{i=0}^{n}\frac{B_i^2}{\binom{n}{i}} &= \sum_{i=0}^{n}\frac{\left(r_i + \frac{|C|\binom{n}{i}}{2^n}(1+(-1)^i B_n')\right)^2}{\binom{n}{i}}\\ &= \sum_{i=0}^{n}\frac{r_i^2}{\binom{n}{i}} + \frac{2|C|}{2^n}\sum_{i=0}^{n} r_i + \frac{2|C|B_n'}{2^n}\sum_{i=0}^{n}(-1)^i r_i\\ &\quad + \frac{|C|^2}{4^n}\sum_{i=0}^{n}\binom{n}{i} + \frac{2|C|^2B_n'}{4^n}\sum_{i=0}^{n}(-1)^i\binom{n}{i}\\ &\quad + \frac{|C|^2B_n'^2}{4^n}\sum_{i=0}^{n}\binom{n}{i}\\ &= \sum_{i=0}^{n}\frac{r_i^2}{\binom{n}{i}} + \frac{|C|^2(1+B_n'^2)}{2^n}.\end{aligned}$$

By the previous lemma this is also

$$\begin{aligned}\sum_{i=0}^{n}\frac{B_i^2}{\binom{n}{i}} &= \frac{|C|^2}{2^n}\sum_{i=0}^{n}\frac{B_i'^2}{\binom{n}{i}}\\ &= \frac{|C|^2(1+B_n'^2)}{2^n} + \frac{|C|^2}{2^n}\sum_{i=d_1'}^{d_2'}\frac{B_i'^2}{\binom{n}{i}}\end{aligned}$$

and we are done. □

Acknowledgment

The authors wish to thank V. Sidelnikov and P. Solé for helpful suggestions.

References

[1] I. Gashkov and V. Sidelnikov, "Linear ternary quasiperfect codes correcting double errors," *Probl. Peredachi Inform.*, vol. 22, no. 4, pp. 43–48, 1986.

[2] T. Kasami, T. Fujiwara, and S. Lin, "An approximation to the weight distribution of binary linear codes," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 6, pp. 769–780, 1985.

[3] I. Krasikov, "Degree conditions for vertex switching reconstruction," submitted.

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

[5] V. Levenshtein, "Bounds for packings of metric spaces and some their applications," in *Problemy Kibernetiki*, vol. 40. Moscow, USSR: Nauka, 1983, pp. 43–110 (in Russian).

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

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

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

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

[10] G. Szegö, "Orthogonal polynomials," *Amer. Math. Soc. Colloq. Publ.*, vol. 23. Providence, RI: Amer. Math. Soc., 1975.