

AN UPPER BOUND ON THE SIZE OF A BINARY CODE WITH s DISTANCES

Ivan Landjev[✉], Konstantin Vorobev

Received on March 4, 2025

Presented by P. Boyvalenkov, Corresponding Member of BAS, on March 25, 2025

Abstract

Let C be a binary code of length n with distances $0 < d_1 < \dots < d_s \leq n$. In this note we prove a general upper bound on the size of C without any restriction on the distances d_i . The bound is asymptotically optimal.

Key words: s -weight codes, non-linear codes, main problem of coding theory

2020 Mathematics Subject Classification: 05A05, 05A20, 94B25, 94B65

1. Introduction. A binary (nonlinear) code $C \subset \mathbb{F}_2^n$ is called an s -distance code if the possible distances between two different words of the code take on s different values, i.e.

$$\{d(u, v) \mid u, v \in C, u \neq v\} = \{d_1, d_2, \dots, d_s\},$$

where $0 < d_1 < d_2 < \dots < d_s \leq n$. Here $d(u, v)$ denotes the Hamming distance between u and v . A binary code of length n , cardinality M , and with distances from the set $\mathcal{D}$ is called an $(n, M, \mathcal{D})$ -code. If the cardinality is not specified, we speak of an $(n, \mathcal{D})$ -code. A natural problem is to determine the maximal cardinality, denoted by $A_2(n, \mathcal{D})$, of a binary code of fixed length n and distances in the set $\mathcal{D}$.

The first author was supported by the Bulgarian National Science Fund under Contract KP-06-N72/6 – 2023. The second author was supported by the NSP P. Beron project CP-MACT.

<https://doi.org/10.7546/CRABS.2025.05.02>

In a recent paper by BARG et al. [3] it was proved that

$$(1) \quad A_2(n, \{d_1, d_2\}) \leq \binom{n}{2} + 1,$$

for $n \geq 6$. There is no restriction on the distances d_1 and d_2 . This bound is achieved, e.g. we have $A_2(n, \{2, 4\}) = \binom{n}{2} + 1$ for all $n \geq 6$ [1, 2]. The proof of [3] is based on the idea to embed a binary code into a sphere S^{n-1} and relate its size to spherical two distance sets and spherical equiangular sets, using some recently proved bounds.

For codes with more than two weights, we have the following general results.

Theorem 1 (BARG and MUSIN [4]). *Let $\mathcal{D} = \{d_1, \dots, d_s\}$, where $\sum_i d_i \leq$*

$\frac{1}{2}sn$. Then

$$A_2(n, \mathcal{D}) \leq \sum_{i=0}^{s-2} \binom{n}{i} + \binom{n}{s}.$$

Theorem 2 (NOZAKI and SHINOHARA [5]). *Let $\mathcal{D} = \{d_1, \dots, d_s\}$ and let*

$$f(t) = \prod_i \frac{d_i - t}{d_i} = \sum_{k=0}^s f_k \phi_k(t),$$

where

$$\phi_k(x) = \sum_{j=0}^k (-1)^j \binom{x}{j} \binom{n-x}{k-j}, k = 0, \dots, n$$

are the Krawtchouk polynomials. Then

$$A_2(n, \mathcal{D}) \leq \sum_{k: f_k > 0} \binom{n}{k}.$$

Both results require some restrictions on the distances in the $\mathcal{D}$. In this note, we prove a bound on $A_2(n, \mathcal{D})$ without restriction on the distances in $\mathcal{D}$.

2. The upper bound. Let C be a binary $(n, M, \mathcal{D})$ -code with distances $0 < d_1 < d_2 < \dots < d_s \leq n$. Consider the ring of polynomials $\mathbb{R}[x_1, x_2, \dots, x_n]$ over the field of the real numbers. Let us denote $\mathbf{x} = (x_1, x_2, \dots, x_n)$. For a vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$ over $\mathbb{R}$, define the polynomial

$$P_{\mathbf{v}}(\mathbf{x}) = \prod_{i=1}^s (d_i - (\mathbf{x}, \mathbf{1}) - (\mathbf{v}, \mathbf{1}) + 2(\mathbf{v}, \mathbf{x})).$$

Here $(\mathbf{x}, \mathbf{y}) = \sum_i x_i y_i$ is the usual dot product (over $\mathbb{R}$) of $\mathbf{x}$ and $\mathbf{y}$.

Let us note that

$$2(\mathbf{v}, \mathbf{x}) - (\mathbf{x}, \mathbf{1}) = \pm x_1 \pm x_2 \pm \cdots \pm x_n,$$

where the sign of x_i is “+” if $v_i = 1$, and “−” if $v_i = 0$. If we set

$$\sigma_{\mathbf{v}}(j) = \begin{cases} 0 & \text{if } v_j = 1, \\ 1 & \text{if } v_j = 0, \end{cases}$$

we get

$$2(\mathbf{v}, \mathbf{x}) - (\mathbf{x}, \mathbf{1}) = \sum_{j=1}^n (-1)^{\sigma_{\mathbf{v}}(j)} x_j.$$

For a vector $\mathbf{v}$ of Hamming weight d_k , i.e. $\text{wt}(\mathbf{v}) = d_k$:

$$(2) \quad P_{\mathbf{v}}(\mathbf{x}) = \prod_{i=1}^s \left(d_i - d_k + \sum_j (-1)^{\sigma_{\mathbf{v}}(j)} x_j \right).$$

Lemma 3. *Let $\mathbf{u}, \mathbf{v} \in C$. Then*

$$P_{\mathbf{u}}(\mathbf{v}) = \begin{cases} \prod_{i=1}^s d_i & \text{if } \mathbf{u} = \mathbf{v}, \\ 0 & \text{if } \mathbf{u} \neq \mathbf{v}. \end{cases}$$

Proof. As usual, for the vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$, we define $\mathbf{u} * \mathbf{v} = (u_1 v_1, \dots, u_n v_n)$. Then we have

$$\begin{aligned} P_{\mathbf{u}}(\mathbf{u}) &= \prod_{i=1}^s (d_i - (\mathbf{u}, \mathbf{1}) - (\mathbf{u}, \mathbf{1}) + 2(\mathbf{u}, \mathbf{u})) \\ &= \prod_{i=1}^s (d_i - \text{wt}(\mathbf{u}) - \text{wt}(\mathbf{u}) + 2\text{wt}(\mathbf{u})) \\ &= \prod_{i=1}^s d_i, \end{aligned}$$

and

$$\begin{aligned} P_{\mathbf{u}}(\mathbf{v}) &= \prod_{i=1}^s (d_i - (\mathbf{u}, \mathbf{1}) - (\mathbf{v}, \mathbf{1}) + 2(\mathbf{u}, \mathbf{v})) \\ &= \prod_{i=1}^s (d_i - \text{wt}(\mathbf{u}) - \text{wt}(\mathbf{v}) + 2\text{wt}(\mathbf{u} * \mathbf{v})) \\ &= \prod_{i=1}^s (d_i - d(\mathbf{u}, \mathbf{v})). \end{aligned}$$

Since $d(\mathbf{u}, \mathbf{v}) = d_i$ for some $i \in \{1, \dots, s\}$ we have $P_{\mathbf{u}}(\mathbf{v}) = 0$. □

Lemma 4. Let C be a binary code with distances $0 < d_1 < \dots < d_s \leq n$. The polynomials in the set

$$\{P_{\mathbf{u}}(\mathbf{x}) \mid \mathbf{u} \in C\}$$

are linearly independent.

Proof. Assume that $\sum_{\mathbf{u} \in C} c_{\mathbf{u}} P_{\mathbf{u}}(\mathbf{x}) = 0$ for some real numbers $c_{\mathbf{u}}$, $\mathbf{u} \in C$.

Plug in $\mathbf{x} = \mathbf{v}$ for an arbitrary codeword $\mathbf{v} \in C$. By Lemma 4, we get

$$0 = \sum_{\mathbf{u} \in C} c_{\mathbf{u}} P_{\mathbf{u}}(\mathbf{v}) = c_{\mathbf{v}} P_{\mathbf{v}}(\mathbf{v}) = c_{\mathbf{v}} \prod_{i=1}^s d_i.$$

This implies $c_{\mathbf{v}} = 0$ for all $\mathbf{v} \in C$, which proves the Lemma. $\square$

Theorem 5. Let $C \subset \{0, 1\}^n$ be a binary code with distances $0 < d_1 < \dots < d_s \leq n$. Then

$$|C| \leq \binom{n+s}{s}.$$

Proof. Let V be the subspace of $\mathbb{R}[x_1, \dots, x_n]$ of all polynomials of degree at most s , and U – the subspace spanned by $\{P_{\mathbf{u}}(\mathbf{x}) \mid \mathbf{u} \in C\}$. Clearly $U \subseteq V$, whence $\dim U \leq \dim V$.

By Lemma 4, $\dim U = |C|$. On the other hand, V is spanned by the monomials $x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$, where $\alpha_1 + \alpha_2 + \dots + \alpha_n \leq s$. The number of the monomials of degree j is equal to the number of the solutions of $y_1 + y_2 + \dots + y_n = j$ in non-negative integers which is $\binom{n+j-1}{j}$. Thus

$$\dim V = \sum_{j=0}^s \binom{n+j-1}{j} = \binom{n+s}{s}.$$

Now

$$|C| = \dim U \leq \dim V = \binom{n+s}{s}.$$

$\square$

It should be noted that this bound has been proved in [6–8] for s -distance sets on the Euclidean sphere. This result implies then Theorem 5. In the next section, we give an improvement on the bound from Theorem 5.

Theorem 5 says that if C is a code with s distances, then $|C| = O(n^s)$. It has to be noted that this bound is asymptotically tight in the sense that for every s there exist codes of size $\approx cn^s$ for some constant c . An example is provided by the code C consisting of all words of Hamming weight s . If $\mathbf{u}$ and $\mathbf{v}$ are two words from C with $wt(\mathbf{u} * \mathbf{v}) = i$, then $d(\mathbf{u}, \mathbf{v}) = 2(s - i)$, where $i \in \{0, \dots, n - 1\}$. Hence C is a code with s distances: $2, 4, \dots, 2s$. If s is even, the zero word can

be taken in C . This is an example of a code with size $\binom{n}{s} + \varepsilon$, $\varepsilon \in \{0, 1\}$, which is even optimal in many cases, e.g. for $s = 2$ (cf. [1–3, 9]).

Note that the blocks of every design with s intersection numbers form a code with s distances. In particular, quasi-symmetric designs give rise to codes with two distances. If a 2-distance code has the additional property that there is a word at equal distance from all codewords, then the construction described above is the best possible. The following theorem is well-known [10] (Proposition 3.5).

Theorem 6. *If a quasi-symmetric design does not contain repeated blocks, then*

$$b \leq \frac{v(v-1)}{2}.$$

If we exclude the trivial examples, we have the following strengthening [10, Proposition 3.6].

Theorem 7. *Let $\mathcal{D}$ be a quasi-symmetric design without repeated blocks and with $4 \leq k \leq v-4$. Then $b \leq \binom{v}{2}$ with equality if and only if $\mathcal{D}$ is a 4-design.*

In fact, up to taking the complement, there exists exactly one design for which we have equality in the above theorem. This is the classical 4-(23, 7, 1)-design with intersection numbers 1 and 3. Its blocks form a (23, 253, {8, 12})-code.

3. An improvement on the upper bound. The bound from Theorem 8 can be improved by inspecting more closely the dimension of the vector space U . In what follows, let C be a binary code of length n with s distances: $0 < d_1 < \dots < d_s \leq n$.

Consider the coefficient of the monomial $x_1^{2\alpha_1} x_2^{2\alpha_2} \dots x_n^{2\alpha_n}$ in the polynomial $P_v(\mathbf{x})$, where $\text{wt}(\mathbf{v}) = d_k$ for some k . If $t = \sum_i \alpha_i$, then $2t \leq s$. From (2) we can deduce that this coefficient is equal to

$$\begin{aligned} & \binom{2t}{2\alpha_1} \binom{2t-2\alpha_1}{2\alpha_2} \dots \binom{2t-2\alpha_1-\dots-2\alpha_{n-1}}{2\alpha_n} \\ & \times \sum_{\{i_1, \dots, i_{s-2t}\}} (d_{i_1} - d_k)(d_{i_2} - d_k) \dots (d_{i_{s-2t}} - d_k), \end{aligned}$$

where the sum is taken over all $(s-2t)$ -tuples of indices $i_1, \dots, i_{s-2t}$ contained in $\{1, \dots, s\}$. Thus the coefficients of $x_{\sigma(1)}^{2\alpha_1} x_{\sigma(2)}^{2\alpha_2} \dots x_{\sigma(n)}^{2\alpha_n}$ are the same for all permutations σ of $\{1, \dots, n\}$.

This implies the following improvement for the bound of Theorem 5.

Theorem 8. *For every binary code C of length n with s distances, it holds*

$$|C| \leq \binom{n+s}{s} - \sum_{t: 2t \leq s} \left(\binom{t+n-1}{t} - p(t) \right),$$

where $p(t)$ is the number of the partitions of t .

Proof. Two solutions (α_i) and (β_i) of the equation

$$(3) \quad y_1 + \cdots + y_n = t, \quad y_i \geq 0$$

are said to be equivalent if one of them is a permutation of the other, i.e. $\beta_i = \alpha_{\sigma(i)}$ for some permutation σ . The number of the equivalent classes is equal to $p(t)$, the number of partitions of t .

From our preliminary observation before the theorem we have that for the equivalent solutions (α_i) and (β_i) the coefficients of $x_1^{2\alpha_1} x_2^{2\alpha_2} \cdots x_n^{2\alpha_n}$ and $x_1^{2\beta_1} x_2^{2\beta_2} \cdots x_n^{2\beta_n}$ are equal in every polynomial $P_v(\mathbf{x})$.

Let U and V be the same as in Theorem 8. We have $V = V_1 \oplus V_2$, where

$$\begin{aligned} V_1 &= \langle x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha_i \text{ even for all } i \rangle, \\ V_2 &= \langle x_1^{\beta_1} \cdots x_n^{\beta_n} \mid \text{at least one } \beta_i \text{ odd} \rangle. \end{aligned}$$

Let $U_i = V_i \cap U$. We have proved that we have that

$$U_1 \subseteq \left\langle \sum_{\alpha} \mathbf{x}^{\alpha} \mid \alpha \text{ runs over all classes of equivalent solutions of (3)} \right\rangle.$$

Now we have

$$\begin{aligned} \dim V_2 &= \dim V - \dim V_1 \\ &= \binom{n+s}{s} - \sum_{t:t \leq s/2} \#(\text{solutions of (3)}) \\ &= \binom{n+s}{s} - \sum_{t:t \leq s/2} \binom{t+n-1}{t}. \end{aligned}$$

On the other hand,

$$\dim V_1 \leq \sum_{t:t \leq s/2} \#(\text{equivalence classes of solutions of (3)}) = \sum_{t:t \leq s/2} p(t).$$

Now

$$\begin{aligned} \dim U &= \dim U_1 + \dim U_2 \\ &\leq \dim V_1 + \dim V_2 \\ &= \sum_{t:t \leq s/2} p(t) + \binom{n+s}{s} - \sum_{t:t \leq s/2} \binom{t+n-1}{t}, \end{aligned}$$

which was to be proved. □

REFERENCES

- [1] BOYVALENKOV P., K. DELCHEV, D. V. ZINOVIEV, V. A. ZINOVIEV (2021) On two-weight codes, *Discrete Math.*, **344**(5), Paper No. 112318, 16 pp.
- [2] LANDJEV I., A. ROUSSEVA, K. VOROBEOV (2023) Constructions of binary codes with two distances, *Discrete Math.*, **346**(6), Paper No. 113337, 10 pp.
- [3] BARG A., A. GLAZYRIN, W.-J. KAO et al. (2024) On the size of maximal codes with 2, 3, and 4 distances, *Combinatorial Theory*, **4**(1), Paper No. 7, 28 pp.
- [4] BARG A., O. R. MUSIN (2011) Bounds on sets with few distances, *J. Comb. Theory Ser. A*, **118**(4), 1465–1474.
- [5] NOZAKI H., M. SHINOHARA (2010) On a generalization of distance sets, *J. Comb. Theory Ser. A*, **117**(7), 810–826.
- [6] BANNAI E., E. BANNAI, D. STANTON (1982) An upper bound for the cardinality of an s -distance subset in real Euclidean space, II, *Combinatorica*, **3**(2), 147–152.
- [7] BLOKHUIS A. (1982) An upper bound for the cardinality of s -distance sets in E^d and H^d , Eindhoven Univ. Techn. Memorandum 1982-68, May 1982, <https://pure.tue.nl/ws/portalfiles/portal/4373585/696963.pdf>.
- [8] BLOKHUIS A. (1983) Few distance sets, PhD Thesis, Eindhoven University of Technology.
- [9] BOYVALENKOV P., K. DELCHEV, V. A. ZINOVIEV, D. V. ZINOVIEV (2022) On codes with distances d and n , *Probl. of Inform. Transmission*, **58**(4), 352–371.
- [10] CAMERON P., J. H. VAN LINT (1980) *Graphs, Codes and Designs*, Revised edition of Graph theory, coding theory and block designs (in Russian), London Math. Soc. Lecture Note Ser., vol. **43**, Cambridge-New York, Cambridge University Press.

*Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
Akad. G. Bonchev St, Bl. 8, 1113 Sofia, Bulgaria
e-mails: ivan@math.bas.bg, konstantin.vorobev@gmail.com*