An Elementary Proof of the First LP Bound on the Rate of Binary Codes

Nati Linial
School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel

Elyassaf Loyfer
School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel

Abstract

The asymptotic rate vs. distance problem is a long-standing fundamental problem in coding theory. The best upper bound to date was given in 1977, and has since received numerous proofs and interpretations. Here we provide a new, elementary proof of this bound that is based on counting walks in the Hamming cube.

2012 ACM Subject Classification Theory of computation → Error-correcting codes

Keywords and phrases Coding theory, code bounds, convex optimization, linear progamming

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2026.48

Category RANDOM

Related Version *Previous Version*: `https://arxiv.org/abs/2303.16619`

Funding *Nati Linial*: Supported in part by a NSF-BSF research grant "Global Geometry of Graphs".

Acknowledgements The authors are thankful to Leonardo Nagami Coregliano, Fernando Granha Jeronimo and Chris Jones for insightful discussions. We also thank Alex Samorodnitsky for helpful advice and inspiration.

1 Introduction

A binary code of length n is a subset $C \subset \mathbb{F}_2^n$. Its distance is the smallest Hamming distance among all pairs of code words, $\mathrm{dist}(C) := \min_{x \neq y \in C} |x + y|$. A fundamental problem in coding theory is to find the largest possible size, $A(n, d)$, of a code of length n and distance d:

$$A(n, d) := \max\{|C| : C \subset \mathbb{F}_2^n,\ \mathrm{dist}(C) \geq d\}$$

The asymptotic version of this question, the *rate vs. distance problem*, seeks the largest possible rate $R(C) := n^{-1} \log_2(|C|)$, where the distance is linear in n:

$$\mathcal{R}(\delta) := \limsup_{n \to \infty} \{R(C) \colon C \subset \mathbb{F}_2^n,\ \mathrm{dist}(C) \geq \lfloor \delta n \rfloor\}$$

It is known that $\mathcal{R}(0) = 1$, and $\mathcal{R}(\delta) = 0$ for all $\delta \geq 1/2$. The exact value of $\mathcal{R}(\delta)$ remains unknown for all $\delta \in (0, 1/2)$.

The best and only lower bound, $\mathcal{R}(\delta) \geq 1 - H(\delta)$, is due to Gilbert [7] and Varshamov [13]. Here and throughout, $H(\cdot)$ is the binary entropy function.

The best upper bound that we have on $\mathcal{R}(\delta)$ for a large range of δ is the *first linear programming (LP) bound*.[1] Due to McEliece, Rodemich, Rumsey and Welch (MRRW, [9]), this bound states that:

[1] MRRW [9] also proved a second LP bound which improves upon the first for $\delta \in (0, 0.273)$. Because this paper focuses exclusively on the first LP bound, we omit the statement of the second.

▶ **Theorem 1** (The first LP bound [9]). *$\mathcal{R}(\delta) \le H(1/2 - \sqrt{\delta(1-\delta)})$ for every $0 \le \delta \le 1/2$.*

These bounds are based on a linear program due to Delsarte [4], the optimum of which is an upper bound on $A(n,d)$. Although this maximum is still unknown, upper bounds on $A(n,d)$ can be derived by passing to the LP dual and constructing a dual feasible solution.

Numerous works have constructed dual feasible solutions [9, 6, 10, 11, 12, 2, 1], yet all lead to the same asymptotic bound. Interestingly, we observe that not only do all these works reach the same bound, but also construct the same dual feasible solution. They differ, however, in proof techniques, and introduce new insight or generalizations to other metric spaces.

Can we achieve a better bound by designing a different solution? This question motivated the current work. It led us to identify the common pattern underlying previous approaches and develop a slightly more general template for dual solutions. Based on this template, we build a new family of solutions which, while different from previous constructions, yields the same asymptotic upper bound stated in Theorem 1.

Although our new construction does not improve upon Theorem 1, it offers potential insights into the nature of the bound. First, exploring alternate constructions may either lead to improved solutions or clarify the inherent limitations of the Delsarte–MRRW framework. Furthermore, while recent LP/SDP hierarchies offer stronger relaxations to the rate-distance problem ([3, 8]), they have struggled to yield concrete bounds due to their complexity. By constructing an explicit dual solution that relies on different machinery than its predecessors, our walk-based combinatorial approach may prove more adaptable to these advanced hierarchies. Second, our proof technique establishes a connection between bounds on binary codes and small sets with large expansion properties, which can possibly lead to mutual benefits. Third, the main part of our proof is a simple counting argument, which allows for a complete and transparent exposition of the bound, with no technical details omitted. Nonetheless, the proof necessarily includes a non-elementary component common to all previous approaches, detailed in Section 3.

Organization of the paper: Section 2 introduces notation, provides necessary background in Fourier analysis, and recalls Delsarte's LP for binary codes. In Section 3, we introduce the general template for dual feasible solutions and demonstrate how previous works can be seen as implementations of this template. In Section 4, we construct the new solutions. In Appendix A, we state and prove some standard results for completeness.

2 Preliminaries

The Hamming cube. As usual, the Hamming cube is the graph whose vertex set is $\mathbb{F}_2^n$, where adjacency means having Hamming distance one. We denote the adjacency matrix of the cube by $A_H \in \{0,1\}^{\mathbb{F}_2^n \times \mathbb{F}_2^n}$. The j-th *level* of the cube, for $0 \le j \le n$, is the set of vertices of Hamming weight j. We denote by $L_j \colon \mathbb{F}_2^n \to \{0,1\}$ the indicator of the j-th level.

Binary codes. A binary code is a subset $C \subset \mathbb{F}_2^n$. The minimal distance of a code, denoted $\mathrm{dist}(C)$, is determined by the minimum Hamming distance between distinct codewords:

$$\mathrm{dist}(C) \coloneqq \min\{|x-y| \colon x,y \in C \wedge x \neq y\}$$

Analysis of Boolean functions. Let $f, g : \mathbb{F}_2^n \to \mathbb{R}$ be real functions on the cube. We define their inner product w.r.t. the uniform measure,

$$\langle f, g \rangle := 2^{-n} \sum_{x \in \mathbb{F}_2^n} f(x)g(x)$$

Convolution between f and g is defined by

$$(f * g)(x) := 2^{-n} \sum_{y \in \mathbb{F}_2^n} f(y)g(x+y)$$

To every $x \in \mathbb{F}_2^n$ there corresponds a *Fourier character* $\chi_x : \mathbb{F}_2^n \to \mathbb{R}$ that is defined via $\chi_x(y) = (-1)^{\langle x,y \rangle}$. The set of characters $\{\chi_x\}_{x \in \mathbb{F}_2^n}$ forms an orthonormal basis of the space of real functions on the Hamming cube. The Fourier transform of f is denoted $\mathcal{F}[f]$ or $\widehat{f}$ in short, and defined by

$$\widehat{f}(x) := \langle \chi_x, f \rangle = 2^{-n} \sum_{y \in \mathbb{F}_2^n} (-1)^{\langle x,y \rangle} f(y)$$

In Fourier space, operators are defined with normalization factor of 1, and we denote them explicitly with a subscript $\mathcal{F}$, e.g.

$$\langle \widehat{f}, \widehat{g} \rangle_{\mathcal{F}} = \sum_{x \in \mathbb{F}_2^n} \widehat{f}(x) \cdot \widehat{g}(x)$$
$$(\widehat{f} *_{\mathcal{F}} \widehat{g})(x) = \sum_{y \in \mathbb{F}_2^n} \widehat{f}(y)\widehat{g}(y+x)$$

The Fourier transform is its own inverse, up to normalization:

$$f(x) = \mathcal{F}^{-1}[\widehat{f}](x) = \langle \chi_x, \widehat{f} \rangle_{\mathcal{F}} = \sum_{y \in \mathbb{F}_2^n} \chi_x(y)\widehat{f}(y)$$

Parseval's theorem states that $\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle_{\mathcal{F}}$.
The convolution theorem states that $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$.

Krawtchouk polynomials. The Krawtchouk polynomials $\{K_0, \ldots, K_n\}$ are a family of polynomials orthogonal with respect to the binomial probability measure,

$$\sum_{i=0}^{n} 2^{-n} \binom{n}{i} K_j(i)K_k(i) = \begin{cases} \binom{n}{j} & j = k \\ 0 & \text{otherwise} \end{cases}$$

The j-th Krawtchouk can be defined by the sum

$$K_j(t) = \sum_{k=0}^{n} (-1)^k \binom{t}{k} \binom{n-t}{j-k}, \qquad t \in \mathbb{R}$$

In particular, $K_0(t) = 1$ and $K_1(t) = n - 2t$.

2.1 Delsarte's LP

Here we recall Delsarte's LP for binary codes. Although we only require the dual LP, we start with the primal as it better suits the exposition. In his original work [4], Delsarte formulated the LP via association schemes, which readily generalizes to other spaces, e.g. fixed-weight codes or spherical codes. For binary codes with the Hamming distance, the LP can be developed completely through Fourier analysis, as we choose to do here.

Delsarte's LP for binary codes can be stated as follows. Inequalities on functions are defined pointwise.

$$\begin{aligned} \max \quad & \sum_{x \in \mathbb{F}_2^n} f(x) \\ \text{s.t.} \quad & f \colon \mathbb{F}_2^n \to \mathbb{R} \\ & f(0) = 1 && \text{(normalization)} \\ & f \geq 0 && \text{(non-negativity)} \\ & \widehat{f} \geq 0 && \text{(Fourier)} \\ & f(x) = 0 \text{ if } 1 \leq |x| \leq d-1 && \text{(distance)} \end{aligned} \tag{1}$$

▶ **Proposition 2.** *The optimum of Delsarte's LP is an upper bound to $A(n,d)$.*

Proof. The proof is by constructing a feasible solution to (1) whose value is $A(n,d)$. Consider a binary code $C \subset \mathbb{F}_2^n$ of length n, minimal distance $\operatorname{dist}(C) \geq d$, and maximum cardinality, $|C| = A(n,d)$. Let $\mathbb{1}_C : \mathbb{F}_2^n \to \mathbb{R}$ be its indicator function, and let $f_C \coloneqq \frac{2^n}{|C|} \mathbb{1}_C * \mathbb{1}_C$. It is not hard to verify that f_C satisfies the normalization and non-negativity constraints. The Fourier constraints are a consequence of the convolution theorem, as $\mathcal{F}[\mathbb{1}_C * \mathbb{1}_C] = \widehat{\mathbb{1}}_C^2 \geq 0$. The distance constraints follow from the distance property of the code, $\operatorname{dist}(C) \geq d$. Finally, the value of this solution is $\sum_{x \in \mathbb{F}_2^n} f_C(x) = |C|$ which is by assumption $A(n,d)$. ◀

Passing to the LP dual,

$$\begin{aligned} \text{min.} \quad & g(0)/\widehat{g}(0) \\ \text{s.t.} \quad & g \colon \mathbb{F}_2^n \to \mathbb{R} \\ & \widehat{g}(0) > 0 && \text{(normalization)} \\ & \widehat{g} \geq 0 && \text{(Fourier)} \\ & g(x) \leq 0 \text{ if } |x| \geq d && \text{(distance)} \end{aligned} \tag{2}$$

Note that to turn the minimization problem (2) into an honest LP we can just require that $\widehat{g}(0) = 1$.

The dual can be derived by standard LP methods. For completeness we include a proof of weak duality.

▶ **Proposition 3.** *If $g : \mathbb{F}_2^n \to \mathbb{R}$ is feasible for (2), then*

$$A(n,d) \leq g(0)/\widehat{g}(0) \tag{3}$$

Proof. Let $g \colon \mathbb{F}_2^n \to \mathbb{R}$ be feasible as in the proposition's statement, and let $f \colon \mathbb{F}_2^n \to \mathbb{R}$ be an optimal solution to (1). By Proposition 2,

$$\widehat{g}(0) \cdot A(n,d) \leq \widehat{g}(0) \cdot \sum_x f(x) = 2^n \widehat{g}(0) \cdot \widehat{f}(0)$$

Since both $\widehat{f}$ and $\widehat{g}$ are non-negative,

$$\widehat{g}(0) \cdot \widehat{f}(0) \leq \sum_{x \in \mathbb{F}_2^n} \widehat{g}(x) \cdot \widehat{f}(x) = \langle \widehat{g}, \widehat{f} \rangle_{\mathcal{F}}$$

Using Parseval's inequality, we have

$$\langle \widehat{g}, \widehat{f} \rangle_{\mathcal{F}} = \langle g, f \rangle = 2^{-n} \sum_x f(x)g(x)$$

If $1 \leq |x| \leq d-1$ then $f(x) = 0$. If $|x| \geq d$ then $f(x) \geq 0$ while $g(x) \leq 0$. Hence,

$$2^{-n} \sum_x f(x)g(x) \leq 2^{-n} f(0)g(0)$$

Recalling that $f(0) = 1$ and $\widehat{g}(0) > 0$, we conclude that $A(n,d) \leq g(0)/\widehat{g}(0)$. ◀

3 A recipe for dual solutions

We next describe a general framework that yields feasible solutions to (2). The idea is to decouple the two families of constraints: first handle the distance constraints and then the Fourier constraints. We then demonstrate how previous works fit into this framework by implementing a dual solution similar to [9, 6, 10, 11, 12, 2, 1]. Finally, we remark shortly on the proof techniques used in prior works.

▶ **Proposition 4.** *Let $\phi, \Gamma : \mathbb{F}_2^n \to \mathbb{R}$ such that*

1. *$\phi(0) > 0$ and $\phi(x) \leq 0$ whenever $|x| \geq d$,*
2. *The nonzero function Γ satisfies*

$$\widehat{\Gamma} \geq 0, \ \widehat{\phi} *_{\mathcal{F}} \widehat{\Gamma} \geq \widehat{\Gamma}$$

Take $g := \phi \cdot \Gamma^2$. Then, g is feasible for (2), and

$$g(0)/\widehat{g}(0) \leq \phi(0) \cdot |\mathrm{supp}(\widehat{\Gamma})|$$

where $\mathrm{supp}(\widehat{\Gamma})$ is the support of $\widehat{\Gamma}$.

Combined with Proposition 3, Proposition 4 gives an upper bound to $A(n,d)$ in terms of $\phi(0)$ and the support size of $\widehat{\Gamma}$.

Note that this method involves only $\widehat{\Gamma}$ with no direct reference to Γ.

Proof. Let ϕ, Γ and g as above. Indeed, such g is non-positive on levels $d, d+1, \ldots, n$ of the cube, and, by the convolution theorem, $\widehat{g} \geq \widehat{\Gamma} *_{\mathcal{F}} \widehat{\Gamma} \geq 0$. In particular, $\widehat{g}(0) \geq \|\widehat{\Gamma}\|_2^2 > 0$. Hence, g is feasible for (2).

In spirit of [10], we derive an upper bound on $A(n,d)$ in terms of the support size of $\widehat{\Gamma}$. Start with an upper bound on $g(0)$:

$$g(0) = \phi(0) \cdot \Gamma^2(0) = \phi(0) \cdot \left(\sum_{x \in \mathbb{F}_2^n} \widehat{\Gamma}(x) \right)^2$$

Let $S \subset \mathbb{F}_2^n$ be the support of $\widehat{\Gamma}$ and let $\mathbb{1}_S$ be its indicator. Then $\sum_{x \in \mathbb{F}_2^n} \widehat{\Gamma}(x) = \langle \widehat{\Gamma}, \mathbb{1}_S \rangle_{\mathcal{F}}$. By Cauchy-Schwarz,

$$\langle \widehat{\Gamma}, \mathbb{1}_S \rangle_{\mathcal{F}} \leq \left(\langle \widehat{\Gamma}, \widehat{\Gamma} \rangle_{\mathcal{F}} \cdot \langle \mathbb{1}_S, \mathbb{1}_S \rangle_{\mathcal{F}} \right)^{1/2} = \left(\langle \widehat{\Gamma}, \widehat{\Gamma} \rangle_{\mathcal{F}} \cdot |\mathrm{supp}(\widehat{\Gamma})| \right)^{1/2}$$

On the other hand,

$$\widehat{g}(0) \geq (\widehat{\Gamma} *_{\mathcal{F}} \widehat{\Gamma})(0) = \sum_{x \in \mathbb{F}_2^n} \widehat{\Gamma}^2(x) = \langle \widehat{\Gamma}, \widehat{\Gamma} \rangle_{\mathcal{F}}$$

Dividing $g(0)/\widehat{g}(0)$, we get the desired result. ◀

3.1 The MRRW solution

The MRRW construction, and subsequent works, can be described by choosing

- $\phi(x) = 2(d - |x|)$, and
- $\widehat{\Gamma} = f_r$, where f_r is the principal eigenfunction of the adjacency matrix $A_{B(r)}$ of the Hamming ball with radius $r = n/2 - \sqrt{d(n-d)}$ (a more detailed definition is given below).

We turn to give the details in a language close to that of [10]. At the end of this section we remark on the different techniques used by previous authors to obtain this solution.

Let $\phi(x) = 2(d - |x|)$. It can be expressed with Krawtchouk polynomials,

$$\phi(x) = 2(d - |x|) = (n - 2|x|) - (n - 2d) = K_1(|x|) - K_1(d)$$

The Fourier transform of the linear Krawtchouk K_1 is the level-1 indicator L_1, (Proposition 8),

$$\widehat{K_1}(|x|) = L_1(x) = \begin{cases} 1 & |x| = 1 \\ 0 & \text{otherwise} \end{cases}$$

By the convolution theorem,

$$\widehat{\phi \cdot \Gamma}(x) = \widehat{\phi} *_{\mathcal{F}} \widehat{\Gamma} = L_1 *_{\mathcal{F}} \widehat{\Gamma} - K_1(d) \cdot \widehat{\Gamma}$$

The operator "$L_1 *_{\mathcal{F}}$" has a nice interpretation: it is equivalent to multiplication by the adjacency matrix A_H of the Hamming cube (see Proposition 9). The conditions on Γ from Proposition 4 can be expressed as

$$\begin{aligned} &\text{minimize} && |\mathrm{supp}(\widehat{\Gamma})| \\ &\text{s.t.} && \Gamma \neq 0, \quad \widehat{\Gamma} \geq 0, \quad A_H \widehat{\Gamma} \geq (1 + K_1(d)) \cdot \widehat{\Gamma} \end{aligned} \tag{4}$$

The problem is now reduced to finding an approximate eigenfunction of A_H with eigenvalue at least $n - 2d + 1$ and small support. The best known choice is to take $\widehat{\Gamma}$ as the principal eigenfunction of the adjacency matrix restricted to the Hamming ball of radius r.

Let $A_{B(r)} \in \{0,1\}^{\mathbb{F}_2^n \times \mathbb{F}_2^n}$ be the adjacency matrix of the Hamming ball of radius $0 \leq r \leq n$,

$$A_{B(r)}(x,y) = \begin{cases} 1 & |x|, |y| \leq r \wedge |x + y| = 1 \\ 0 & \text{otherwise} \end{cases}$$

We now state facts regarding the operator $A_{B(r)}$, without proof. These facts are compiled from [9, 6, 10, 11, 12, 2, 1], where proofs can be found. The largest eigenvalue of $A_{B(r)}$ is equal to the first root of the r-th Krawtchouk polynomial, and its value is $2\sqrt{r(n-r)} + o(n)$. Let f_r be the corresponding eigenfunction. Then f_r is non-negative and supported on $B(r)$, thus its support size is at most $\sum_{j=0}^{r} \binom{n}{r} = 2^{nH(r/n)+o(n)}$.

Choose $\widehat{\Gamma} = f_r$, for some r to be determined later. Since A_H has only non-negative entries, $A_H f_r \geq 2\sqrt{r(n-r)} \cdot f_r$, element-wise. Consequently, the choice $\widehat{\Gamma} = f_r$ satisfies (4) if

$$2\sqrt{r(n-r)} \geq n - 2d + 1 \tag{5}$$

hence

$$A(n,d) \leq \phi(0) \cdot |\mathrm{supp}(f_r)| \leq 2^{nH(r/n)+o(n)}$$

The first LP bound, Theorem 1, is obtained by optimizing r.

In [9], f_r is obtained using Christoffel-Darboux formula. The problem (4) is equivalent to the Faber-Krahn maximizer in [6]. In [10], f_r is obtained directly by solving a recurrence relation. In [2, 1], a function close to f_r is found by spectral and functional tools. In [5], they use the Rayleigh quotient to prove existence of a feasible solution to (4) supported on levels $[r - \sqrt{n}, r]$ of the Boolean hypercube, for r satisfying (5).

4 A new family of solutions to Delsarte's dual LP

In this section we present our proof, which is cast in the pattern described in the previous section.

The main idea is as follows. By choosing a slightly more involved ϕ than in prior work, we can dramatically simplify $\widehat{\Gamma}$, reducing it to the indicator of a set. The sets of interest are small, and have the property that a random walk, started inside the set, has a high probability of returning to it. This is captured in Proposition 5. The symmetry of the Hamming cube suggests the level sets as natural candidates for such subsets. In Proposition 6 we count the walks between levels of the cube. Finally, we derive the first LP bound (Theorem 1) from Proposition 5 and Proposition 6.

A walk of length m on the cube is a sequence of vertices $(x_1, x_2, \ldots, x_m) \in (\mathbb{F}_2^n)^m$, where every two consecutive vertices are neighbors, $|x_i + x_{i+1}| = 1$, for $i = 1, \ldots, m-1$.

▶ **Proposition 5.** *Let $n, d \in \mathbb{N}$ such that $d \leq n/2$. Let $m \in \mathbb{N}$ and let $S \subset \mathbb{F}_2^n$ be a set of vertices such that for any $x \in S$, the number of length-m walks starting at x and ending in S is at least $(n-2d)^m + 1$. Then,*

$$A(n,d) \leq n^m \cdot |S|$$

Proof. We choose ϕ and Γ to satisfy Proposition 4 and derive the bound. Define

$$\phi_m(x) = (n - 2|x|)^m - (n-2d)^m$$

and let $\widehat{\Gamma} = \mathbb{1}_S$, the indicator function of the set S. Observe that

1. $\phi_m(0) > 0$ and $\phi_m(x) \leq 0$ for $|x| \geq d$.
2. $\widehat{\Gamma} \neq 0$, $\widehat{\Gamma} \geq 0$.

It remains to show that $\widehat{\phi}_m *_{\mathcal{F}} \widehat{\Gamma} \geq \widehat{\Gamma}$. The Fourier transform of $(n-2|x|)^m = K_1^m(x)$ is the m-wise convolution of L_1, the level-1 indictor (see Proposition 8),

$$\mathcal{F}[K_1^m] = \widehat{K}_1 *_{\mathcal{F}} \ldots *_{\mathcal{F}} \widehat{K}_1 = L_1 *_{\mathcal{F}} \ldots *_{\mathcal{F}} L_1$$

Convolution with the level-1 indictor is equivalent to multiplying by the adjacency matrix of the Hamming cube A_H (see Proposition 9). Namely,

$$\widehat{\phi}_m *_{\mathcal{F}} \widehat{\Gamma} = A_H^m \cdot \widehat{\Gamma} - (n-2d)^m \widehat{\Gamma}$$

Hence the remaining condition $\widehat{\phi}_m *_{\mathcal{F}} \widehat{\Gamma} \geq \widehat{\Gamma}$ is equivalent to

$$(A_H^m \cdot \mathbb{1}_S)(x) \geq ((n-2d)^m + 1)\mathbb{1}_S(x) \qquad \forall x \in \mathbb{F}_2^n \tag{6}$$

Let $x \in \mathbb{F}_2^n$. The left-hand side counts walks that start at x and end in S,

$$(A_H^m \mathbb{1}_S)(x) = \sum_{y \in \mathbb{F}_2^n} A_H^m(x,y)\mathbb{1}_S(y) = \sum_{y \in S} A_H^m(x,y)$$

the last expression sums the number of length-m paths between x and y, for every $y \in S$.

If $x \notin S$, then the left-hand side of (6) is non-negative while the right-hand side is 0. If $x \in S$, the number of walks from x to S exceeds $((n-2d)^m+1)$, by assumption. Hence (6) holds for every $x \in \mathbb{F}_2^n$, and ϕ_m, Γ satisfy Proposition 4.

Finally, $\phi_m(0) = n^m - (n-2d)^m \leq n^m$, and the support size of $\mathbb{1}_S$ is $|S|$. By Proposition 4, $A(n,d) \leq n^m \cdot |S|$. ◀

▶ **Proposition 6.** *Let $j, r, m \in \mathbb{N}$ such that $m = o(\sqrt{r})$ and $r \leq n/2$. Let $P_{r,m,j}$ be the number of walks that start at some vertex x in level r and end at any vertex in level j. Then,*

1. *If $m + r + j = 1 \pmod 2$ or $m < |r - j|$ then $P_{r,m,j} = 0$.*
2. *Otherwise,*

$$P_{r,m,j} = \binom{m}{(m-r+j)/2}(n-r)^{(m+r-j)/2} r^{(m-r+j)/2}(1+O(m/r))^m$$

Proof. A vertex in level k has $n-k$ neighbors in level $k+1$, and k neighbors at level $k-1$. The walk takes m steps and thus contained in levels $r-m, \ldots, r+m$, so from every vertex encountered during the walk there are $n - r + O(m)$ ways to go up one level and $r + O(m)$ ways to go down one level.

If the walk terminates at level j, it takes in total $(m+r-j)/2$ steps up and $(m-r+j)/2$ steps down. There are $\binom{m}{(m-r+j)/2}$ ways to choose the order of the up/down steps, $n-r+O(m)$ options for each "up" step and $r + O(m)$ options for each "down" step, so in total

$$\binom{m}{(m-r+j)/2}(n-r+O(m))^{(m+r-j)/2}(r+O(m))^{(m-r+j)/2}$$ ◀

▶ **Corollary 7.** *Let $1 \leq d \leq n/2$ be integers. Then $A(n,d) \leq 2^{nH(1/2-\sqrt{d/n(1-d/n)})+o(n)}$.*

Proof. Let m be odd, $m = \omega_n(1)$ and $m = o(\sqrt{n})$. Let $r = n/2 - \sqrt{d(n-d)}$ and let $S = \{r-1 \leq |x| \leq r\}$. It suffices to show that S has sufficiently many returning walks. Let $x \in S$. If $|x| = r$, by Proposition 6, the number of length-m walks from x to level r is 0, and to level $r-1$ it is

$$P_{r,m,r-1} = \binom{m}{(m-1)/2}\sqrt{((n-r)r)^m \frac{n-r}{r}}(1+O(m/r))^m$$

Using the well known approximation for binomial coefficients, $\binom{n}{k} = 2^{nH(k/n)-O(\log n)}$, we obtain

$$((A_H^m \mathbb{1}_S)(x))^{1/m} \geq 2\sqrt{r(n-r)} + O(m)$$

By our choice of r, this is greater than $((n-2d)^m+1)^{1/m}$. The proof for $x \in S$ with $|x| = r-1$ is similar. ◀

Theorem 1 follows directly from the above corollary.

References

1 Alexander Barg and Dmitry Nogin. A functional view of upper bounds on codes. In *Coding and cryptology*, pages 15–24. World Scientific, 2008.

2 Alexander M Barg and D Yu Nogin. Spectral approach to linear programming bounds on codes. *Problems of Information Transmission*, 42(2):77–89, 2006. doi:10.1134/S0032946006020025.

3 Leonardo Nagami Coregliano, Fernando Granha Jeronimo, and Chris Jones. A complete linear programming hierarchy for linear codes. *arXiv preprint arXiv:2112.09221*, 2021. arXiv:2112.09221.

4 Philippe Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.*, 10:vi+–97, 1973.

5 Dean Doron and Amnon Ta-Shma. Lecture notes on error correcting codes, 2017. Accessed: 2026-05-04. URL: `https://www.math.tau.ac.il/~amnon/Classes/2017-ECC/Lectures/LPBound.pdf`.

6 Joel Friedman and Jean-Pierre Tillich. Generalized Alon–Boppana theorems and error-correcting codes. *SIAM Journal on Discrete Mathematics*, 19(3):700–718, 2005. `doi:10.1137/S0895480102408353`.

7 Edgar N Gilbert. A comparison of signalling alphabets. *The Bell system technical journal*, 31(3):504–522, 1952.

8 Elyassaf Loyfer and Nati Linial. New lp-based upper bounds in the rate-vs.-distance problem for binary linear codes. *IEEE Transactions on Information Theory*, 69(5):2886–2899, 2023. `doi:10.1109/TIT.2023.3236660`.

9 Robert McEliece, Eugene Rodemich, Howard Rumsey, and Lloyd Welch. New upper bounds on the rate of a code via the Delsarte-Macwilliams inequalities. *IEEE transactions on Information Theory*, 23(2):157–166, 1977. `doi:10.1109/TIT.1977.1055688`.

10 Michael Navon and Alex Samorodnitsky. On Delsarte's linear programming bounds for binary codes. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 327–336. IEEE, 2005.

11 Michael Navon and Alex Samorodnitsky. Linear programming bounds for codes via a covering argument. *Discrete & Computational Geometry*, 41(2):199–207, 2009. `doi:10.1007/S00454-008-9128-0`.

12 Alex Samorodnitsky. One more proof of the first linear programming bound for binary codes and two conjectures. *arXiv preprint arXiv:2104.14587*, 2021. `arXiv:2104.14587`.

13 Rom Rubenovich Varshamov. Estimate of the number of signals in error correcting codes. *Docklady Akad. Nauk, SSSR*, 117:739–741, 1957.

A Supplementary proofs

▶ **Proposition 8.** *Let K_j be the j-th Krawtchouk polynomial,*

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

and let $L_j(x) = \mathbb{1}_{|x|=j}$ be the indicator function of the j-level of the Hamming cube. Then,

$$\widehat{K_j} = L_j$$

Proof. It is easier to apply the Fourier transform to L_j, and recall that Fourier transform is its own inverse up to normalization, $\mathcal{F}^{-1}[\mathcal{F}[f]] = 2^n \mathcal{F}[\mathcal{F}[f]] = f$, for every $f\colon \mathbb{F}_2^n \to \mathbb{R}$.

Indeed, let $x \in \mathbb{F}_2^n$ with $|x| = t$. Then,

$$2^n \widehat{L}_j(x) = \sum_{y\in\mathbb{F}_2^n} L_j(y)\chi_x(y) = \sum_{|y|=j} (-1)^{\langle x,y\rangle}$$

where $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$. We group elements by the value of $\langle x, y\rangle$:

$$\sum_{|y|=j} (-1)^{\langle x,y\rangle} = \sum_{k=0}^{n} \sum_{\substack{y\in\mathbb{F}_2^n:\\ |y|=j\\ \langle y,x\rangle=k}} (-1)^k$$

The term $(-1)^k$ does not depend on y anymore, so it remains to count $|\{y \in \mathbb{F}_2^n \colon |y| = j \langle y, x\rangle = k\}|$. If $\langle x, y\rangle = k$, there are $\binom{|x|}{k} = \binom{t}{k}$ options to choose the bits of y that intersect with x, and $\binom{n-|x|}{|y|-k} = \binom{n-t}{j-k}$ options to choose the bits of y outside the support of x. Hence

$$2^n \widehat{L}_j(x) = \sum_{k=0}^{n} (-1)^k \binom{t}{k} \binom{n-t}{j-k}$$ ◀

▶ **Proposition 9.** *Let $f \colon \mathbb{F}_2^n \to \mathbb{R}$, then $\widehat{K_1} *_{\mathcal{F}} f = A_H \cdot f$, where A_H is the adjacency matrix of the Hamming cube.*

Proof. By Proposition 8, $\widehat{K}_1 *_{\mathcal{F}} f = L_1 *_{\mathcal{F}} f$. By definition,

$$(\widehat{K}_1 *_{\mathcal{F}} f)(x) = (L_1 *_{\mathcal{F}} f)(x) = \sum_{y \in \mathbb{F}_2^n} L_1(x+y) f(y) = \sum_{y \in \mathbb{F}_2^n \colon |x+y|=1} f(y)$$

The condition $|x + y| = 1$ is the adjacency condition between vertices x, y in the Hamming cube, hence the last sum is equal $(A_H f)(x)$. ◀