New Bounds on the Size of Binary Codes with Large Minimum Distance

James (Chin-Jen) Pang, Hessam Mahdavifar, and S. Sandeep Pradhan
Department of Electrical Engineering and Computer Science, University of Michgan, Ann Arbor, MI 48109, USA
Email: cjpang, hessam, pradhanv@umich.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834432

***Abstract*—Let $A(n,d)$ denote the maximum number of codewords in a binary code of length n and minimum Hamming distance d. Deriving upper and lower bounds on $A(n,d)$ has been a subject for extensive research in coding theory. In this paper, we examine upper and lower bounds on $A(n,d)$ in the high-minimum distance regime, in particular, when $d = n/2 - \Theta(\sqrt{n})$. We will first provide a lower bound based on a cyclic construction for codes of length $n = 2^m - 1$ and show that $A(n, d = n/2 - 2^{c-1}\sqrt{n}) \geqslant n^c$, where c is an integer with $1 \leqslant c \leqslant m/2 - 1$. With a Fourier-analytic view of Delsarte's linear program, novel upper bounds on $A(n, n/2 - \sqrt{n})$ and $A(n, n/2 - 2\sqrt{n})$ are obtained, and, to the best of the authors' knowledge, are the first upper bounds scaling polynomially in n for the regime with $d = n/2 - \Theta(\sqrt{n})$.**

I. Introduction

Low-capacity scenarios are of increasingly greater importance with the advent of the Internet of Things (IoT) and the next generation of mobile networks. In general, IoT devices operate under extreme power constraints and often need to communicate at very low signal-to-noise ratio (SNR), e.g., -13 dB or 0.03 bits per transmission (in terms of channel capacity) in Narrow-Band-IoT protocols [1]. In the standard, legacy turbo codes or convolutional codes at moderate rates together with many repetitions are adopted. This implies effective code rates as low as 1.6×10^{-4} are supported in such protocols. It is expected, however, that repeating a moderate-rate code to enable low-rate communication will result in rate loss and suboptimal performance. As a result, studying ultra-low-rate error-correcting codes for reliable communications in such low-capacity regimes becomes necessary [2]–[6].

In this paper we focus on minimum distance properties of codes in the ultra-low-rate regime, which can be also described as the large minimum distance regime, to be specified later. Let C be a binary (n, M, d) code of length n, size M, minimum distance $d = (n - j)/2$. The *dimension* of C is given by $k = \log_2 M$, the rate by $R = k/n$, and relative distance by $\delta = d/n$. Given positive integers n and d, $A(n,d)$ denote the maximal value of M such that a (n, M, d) code exists. $A(n,d)$ is a fundamental quantity in coding theory subject to extensive studies in the past. Lower bounds on $A(n,d)$ are usually obtained by constructions. For a survey on the known bounds with finite n and d, the reader is referred to [7] and the websites [8], [9]. For asymptotic lower bounds and a survey of prior asymptotic results the reader is referred to [10], [11]. In this paper, we focus on studying bounds on $A(n,d)$ in the high-minimum distance regime, in particular, when $d = n/2 - \Theta(\sqrt{n})$.

This work was supported by the National Science Foundation under grants CCF–2132815 and CCF–1909771.

For $j = n - 2d \leqslant 0$, provided that a sufficient number of Hadamard matrices exist, a widely accepted conjecture, Plotkin and Levenshtein (see [12, Chapter 2, Theorem 8]) have essentially settled the problem and showed that $A(2d, d) = 4d$, $A(n,d) = 2\lfloor d/(2d-n) \rfloor$ for even $d > n/2$, and $A(n,d) = 2\left\lfloor \frac{d+1}{2d+1-n} \right\rfloor$ for odd $d > (n-1)/2$.

In what follows, we consider the scenario with $j > 0$. When j scales linearly with n, asymptotic results can be found in [10], [13]. In particular, the conjecture is that there does not exist any binary code exceeding the Gilbert-Varshamov lower bound (Theorem 1).

There are very few studies in the literature targeting the regime where j is sub-linear in n. In 1973, McEliece (see [12, Chapter 17, Theorem 38]), using the linear programming approach, established the following bound that is valid for $j = o(\sqrt{n})$:

$$A(n,d) \leqslant n(j+2). \tag{1}$$

For $j \approx n^{1/3}$, codes have been constructed [14] to meet McEliece's upper bound, hence showing the tightness of this bound in this regime. A few improvements [15], [16] have been derived in the literature in the regime $j = o(n^{1/3})$. However, to the best of the authors' knowledge, no explicit upper bounds (or lower bounds with general enough parameters) on $A(n,d)$ are derived in the regime $j = \Theta(\sqrt{n})$. In this paper, we attempt to answer the following question: If the term $j = n - 2d$ scales as $j = \Theta(\sqrt{n})$, what is the best size M one can achieve?

The rest of this paper is organized as follows. In Section II we review some well-known asymptotic bounds on $A(n,d)$ and examine their scaling behaviour when $j = \Theta(\sqrt{n})$. In Section III-A, a BCH-like cyclic code construction is given to establish a non-trivial new lower bound. In Section III-B, we review an alternative proof of a well-known *first linear programming bound* on $A(n,d)$ (formally decribed in Theorem 7) through a covering argument using Fourier analysis on the group $\mathbb{F}_2^n$. Then two upper bounds on $A(n,d)$ with $d \geqslant n/2 - \sqrt{n}$ and $d \geqslant n/2 - 2\sqrt{n}$, that are strictly tighter than all prior results, are derived. Finally, the paper is concluded in Section IV.

II. Preliminaries

Let $H_2(\cdot)$ denotes the binary entropy function. Let $n \geqslant r$ be positive integers. Let $B_r(\mathbf{0}, n) \in \{0,1\}^n$ denote the Hamming ball of radius r centered at $\mathbf{0} = (0, 0, \ldots, 0)$, and its volume by

$Vol(r,n) \stackrel{\text{def}}{=} |B_r(\mathbf{0},n)| = \sum_{i=0}^{r} \binom{n}{i}$. We recall the following bounds

1) $Vol(r,n) \leqslant 2^{H_2(r/n)n}$; and
2) $Vol(r,n) \geqslant 2^{H_2(r/n)n-o(n)}$ for sufficiently large n.

A. Known Bounds on Code Sizes

The following bounds on the size of binary codes can be found in standard coding theory textbooks, e.g. [12], [13]. Bounds for the regime $j = \Theta(\sqrt{n})$ are derived and given following the general bounds, e.g. inequalities (3), (5), (7), (8), (9), and (11). When the scaling behaviour of j matters, we choose $j = 2a\sqrt{n}$, i.e., $d = n/2 - a\sqrt{n}$, for ease of comparison between bounds, as in (3), (9), and (11).

Theorem 1 (Gilbert-Varshamov or GV lower bound, [17]): Let positive integers n and $d \leqslant n/2$ be given. Then

$$A(n,d) \geqslant \frac{2^n}{Vol(d-1,n)}. \tag{2}$$

Asymptotically, suppose $0 \leqslant \delta < 1/2$, then there exists an infinite sequence of (n,M,d) binary linear codes with $d/n > \delta$ and rate $R = k/n$ satisfying $R \geqslant 1 - H_2(d/n)$. To evaluate Theorem 1 when $j = \Theta(\sqrt{n})$, consider $j = 2a\sqrt{n}$. Central limit theorem, coupled with the Berry–Esseen theorem, provides an upper bound

$$Vol(d-1,n) \leqslant 2^n \left[Q(2a) + O(1/\sqrt{n})\right],$$

where $Q(\cdot)$ denotes the tail distribution function of the standard normal distribution. Hence (2) becomes

$$A(n, n/2 - a\sqrt{n}) \geqslant \left[Q(2a) + O(1/\sqrt{n})\right]^{-1}, \tag{3}$$

which is loose compared with the Plokin-Levenshtein bound $A(2d,d) = 4d$.

Theorem 2 (Hamming Bound): For every (n,M,d) code $C \subset \{0,1\}^n$,

$$M \leqslant 2^n / Vol(e,n), \tag{4}$$

where $e = \lfloor (d-1)/2 \rfloor$.

In the asymptotics, Theorem 2 bounds the rate from above, in terms of the relative distance δ, by $R \leqslant 1 - H_2(\delta/2)$. For $j = \Theta(\sqrt{n})$, the term $e = n/4 - \Theta(\sqrt{n})$, and $Vol(e,n) \geqslant 2^{H_2(1/4)n - o(n)}$. Hence Theorem 2 becomes

$$M \leqslant 2^{(1-H_2(1/4))n + o(n)} \leqslant 2^{0.189n}, \tag{5}$$

for all sufficiently large n.

Theorem 3 (Singleton Bound): The following holds for any code $C \subset \{0,1\}^n$ with distance d and dimension k.

$$k \leqslant n - d + 1. \tag{6}$$

Under the regime $j = \Theta(\sqrt{n})$, Theorem 3 becomes

$$M \leqslant 2^{n/2 + o(n)}, \tag{7}$$

which is weak compared to (5).

Theorem 4 (Plotkin Bound, [18]): The following holds for any code $C \subset \{0,1\}^n$ with distance d

1) If $d = n/2$, $|C| \leqslant 2n$.
2) If $d > n/2$, $|C| < 2\left\lfloor \frac{d}{2d-n} \right\rfloor$.

One may use a combinatorial argument and Theorem 4 to derive the following corollary.

Corollary 5: If a (n,M,d) binary code C has distance $d < n/2$, then the size $M \leqslant d \cdot 2^{n-2d+2}$.

Using Corollary 5, one may bound the size of any code with $d = (n-j)/2 < n/2$ by

$$M \leqslant d \cdot 2^{j+2} < 2n \cdot 2^j. \tag{8}$$

When j scales as $j = \Theta(\sqrt{n})$, the size M is bounded from above by a polynomial scaling sub-exponentially in n. In particular, set $j = 2a\sqrt{n}$, i.e. $d = n/2 - a\sqrt{n}$, (8) becomes

$$M \leqslant 2n \cdot 2^{2a\sqrt{n}}. \tag{9}$$

Theorem 6 (Elias-Bassalygo Bound): For sufficiently large n, every code $C \subset \{0,1\}^n$ with relative distance δ and rate R satisfies the following:

$$R \leqslant 1 - H_2(J_2(\delta)) + o(1), \tag{10}$$

where $J_2(\delta) \stackrel{\text{def}}{=} \frac{1}{2}(1 - \sqrt{1-2\delta})$.

Assuming $d = n/2 - a\sqrt{n}$, one may adopt steps similar to the proof of Theorem 6 as in [13, p.147] to show an upper bound:

$$M \leqslant n^3 \cdot 2^{\frac{a}{\ln 2}\sqrt{n} + O(1)}. \tag{11}$$

The last upper bound we introduce is known as the *first linear programming bound* or the *MRRW bound* on binary error correcting codes, or, alternatively, on optimal packing of Hamming balls in a Hamming cube. The bound was originally proved by McEliece, Rodemich, Rumsey, and Welch [19], following Delsarte's linear programming approach [20]. Delsarte viewed the distance distribution of a code C of length n as an $(n+1)$-dimensional vector $a = (a_0, a_1, \ldots, a_n)$, where a_i is given by the number of pairs of codewords at distant i, and discovered a system of linear inequalities satisfied by every a associated with a length-n, minimumm distance d code. The coefficients of the linear constraints can be viewed as values of a family of orthogonal polynomials called the Krawchouk polynomials. Based on the duality theorem of linear programming, one may find a feasible solution to the dual program, and view the obtained linear program as an extremal problem. (See [12, Chapter 17] for details.) Good feasible solutions of the dual program were constructed in [19]. The resulting bound is the best known asymptotic upper bound on the cardinality of a code with a given minimal distance scaling linearly in n, for a significant range of the relative distance.

Theorem 7 (MRRW Bound, [19]): For sufficiently large n, every code $C \subset \{0,1\}^n$ with relative distance δ and rate R satisfies the following:

$$R \leqslant H_2\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right). \tag{12}$$

Remark 1: Another bound, known as the *second linear programming bound*, is also given in [19] in the form

$$R \leqslant \min_{0 \leqslant u \leqslant 1-2\delta} 1 + g(u^2) - g(u^2 + 2\delta u + 2\delta), \tag{13}$$

where the function $g(x) \stackrel{\text{def}}{=} H_2((1 - \sqrt{1-x})/2)$. For $0.273 \leqslant \delta \leqslant 0.5$, the bound (13) simplifies to that of (12). For $\delta < 0.273$, the inequality (13) is strictly tighter than (12).

Plugging in $\delta = d/n$ into (12), we have the following bound:

$$M \leqslant 2^{nH_2\left(1/2-\sqrt{d/n(1-d/n)}\right)+o(n)}. \tag{14}$$

Note that, due to the $o(n)$ term, the bound (14) is not tighter than (8) when $j = \Theta(\sqrt{n})$. This appears to the contrary of the fact the MRRW bound is tighter than all the other bounds for relative distance $\delta > 0.273$. However, a tailored treatment of the proof technique may lead to a nontrivial bound as in the derivation of (11) from Theorem 6. In Section III-B, one such bound is given through an alternative proof of the Theorem 7 by working with the maximal eigenfunctions of Hamming balls.

III. Main Results

A. Cyclic Code with High Minimum Distance

Let $n = 2^m - 1$, and $m \in \mathbb{N}$ be an even integer with $m \geqslant 4$. Let c be an integer with $1 \leqslant c \leqslant m/2 - 1$. We construct a binary cyclic code C of length n with high minimum distance as follows.

Theorem 8: There exists a binary cyclic code C of length n, dimension cm, and minimum distance

$$d \geqslant 2^{m-1} - 2^{m/2+c-1} \geqslant \frac{n}{2} - 2^{c-1}\sqrt{n}. \tag{15}$$

Proof: Consider the finite field $F = \mathbb{F}_{2^m}$ and a subfield $K = \mathbb{F}_2 < F$. Let α be a primitive root of unity in F, and set $\alpha_i = \alpha^{1+2^{m/2+i}}$ for $i = 1, 2, \ldots, c$. Consider the binary cyclic code with the generator polynomial

$$g(x) = \frac{x^n - 1}{\prod_{i=1}^{c} M_{\alpha_i}(x)},$$

where $M_\beta(\cdot)$ is the minimal polynomial of β over K. Note that the α_i's belong to different conjugacy classes, i.e,

$$\begin{aligned} A_i &\stackrel{\text{def}}{=} \left\{\alpha_i^{2^j} | j = 0, 1, 2, \ldots, m-1\right\} \\ &= \left\{\alpha^{2^j + 2^{m/2+i+j}} | j = 0, 1, 2, \ldots, m-1\right\} \end{aligned}$$

are disjoint subsets of $F \setminus \{0\}$, and $|A_i| = m$ for each i. This is ensured by the particular choice of α_i's. More specifically, let $P_i = \left\{2^j + 2^{m/2+i+j} \bmod 2^m - 1 | j = 0, 1, 2, \ldots, m-1\right\}$ be the set of the powers over α for elements in A_i. Each P_i is a cyclotomic coset mod 2 in F and the length-m binary representation for each p, p' in P_i are cyclic shifts of each other. Let $p_i = 1 + 2^{m/2+i}$ be the coset representative of P_i. The claim on the size of $|A_i|$ holds by noting that $|A_i| = |P_i| = m$. To claim that A_i's are disjoint, it suffice to show that the cyclotomic cosets P_i's are disjoint. First note that for two cyclotomic cosets P_i and P_k, they are either disjoint or identical. Assume for some $i \neq k$, cosets P_i and P_k are identical. Then $p_i = 1 + 2^{m/2+i}$ is an element in P_k, that is, there is a $p' = 2^l + 2^{m/2+k+l} \in P_k$ for which $p_i = p'$ modulo $2^m - 1$. As both p_i and p' are sums of two powers of 2, we note that neither $m|l$ and $m|(l + k - i)$, nor $m|(l - m/2 - i)$ and $m|(m/2 + k - l)$, can happen. Hence $p_i \notin P_k$, and thus P_i and P_k are disjoint. Thus the degree of the polynomial $g(x)$ is $n - cm$.

For the minimum distance, let $t = 2^{m-1} + 2^{m/2+c-1} + 1$. We show next that for $j = t, t+1, \ldots, 2^m - 1$, α^j is a root for the generator polynomial $g(x)$. In other words, $A_i \cap \{t, t+1, \ldots, 2^m - 1\} = \emptyset$, for $i = 1, 2, \ldots, c$. This is by noting that the powers of the elements in A_i, after taking modulo $2^m - 1$, can be written as the sum of two powers of two, i.e., $2^l + 2^j$, where the difference between l and j is at least $m/2 - c$, and that such a number does not belong to $\{t, t+1, \ldots, 2^m - 1\}$. Hence, the minimum distance of the code d is at least $2^m - t + 1 = 2^{m-1} - 2^{m/2+c-1}$ by BCH bound. ∎

Note that the parameters of the codes constructed in Theorem 8 are, in a sense, sitting between those of the first order and the second order Reed-Muller (RM) codes of length $n = 2^m$. More specifically, RM$(m, 1)$ has minimum distance equal to $n/2$ and dimension equal to $m + 1$, while RM$(m, 2)$ has minimum distance $n/4$, and dimension $1 + m + \binom{m}{2}$. Hence, there is a wide gap in the minimum distance between the first order and the second order RM codes and, intuitively speaking the *BCH-like* codes constructed in Theorem 8 can be used to fill this gap. In particular, the parameters of the constructed code in the extreme cases of c, i.e., the cases of $c = 1$ and $c = m/2 - 1$, are close to those of RM$(m, 1)$ and RM$(m, 2)$, respectively.

B. Improved Code Size Upper Bound

We will follow the covering argument by Navon and Samorodnitsky [21] and show two upper bounds on the size of any code C with length n and minimum distance d, the first holds for any $d \geqslant n/2 - \sqrt{n}$, and the second for any $d \geqslant n/2 - 2\sqrt{n}$. The viewpoint presented in [21], providing an alternative proof to Theorem 7, is different from that in [19], which relies on analytical properties of the Krawchouk polynomials, and instead adopts Fourier analysis on the group $\mathbb{F}_2^n$ as their main tool.

In particular, the authors of [21] exploit the expediency of working with the maximal eigenfunctions of Hamming balls. One key finding was that, given any real-valued function f on $\{0,1\}^n$ with a small support $B \subset \{0,1\}^n$, such that the adjacency matrix of the Hamming cube acts on f by multiplying it pointwise by a large factor, the cardinality of error-correcting codes with minimum distance d can be upper bounded by $n|B|$. The applicability will depend on the value of the multiplying factor. By proposing functions f supported on Hamming balls $B = B_r(\mathbf{0}, n)$ of different radii r, one may derive a lower bound of the multiplying factor, formally called the *maximal eigenvalue of adjacency matrix of the subgraph incduced by* B. This made possible a simple proof of the first linear programming bound.

Let us now state the definition of the *maximal eigenvalue* of a graph. Let $G = (V, E)$ be a (finite, undirected, simple) graph. Let $A_G = (A_{ij})$ denote the $|V| \times |V|$ adjacency matrix of G, defined by $A_{ij} = 1$ if $(i, j) \in E$ and $A_{ij} = 0$ otherwise for vertices $i, j \in V$. Note that A_G is symmetric, so its eigenvalues are real, and can be ordered as $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$. For any function f on $\mathbb{F}_2^n$, the function Af sums at each point of $\{0,1\}^n$ the values of f at its neighbours. That is, the value taken by the function Af at a vertex $x \in \mathbb{F}_2^n$, denoted by $(Af)(x)$ or $Af(x)$, is given by $Af(x) = \sum_{y \in \mathbb{F}_2^n : w_H(x,y)=1} f(y)$. When the graph is a subset of the cube, $B \subseteq \{0,1\}^n$, set

$$\lambda_B \stackrel{\text{def}}{=} \max\left\{ \frac{\langle Af, f\rangle}{\langle f, f\rangle} \,\middle|\, f : \mathbb{F}_2^n \to \mathbb{R}, \operatorname{supp}(f) \subseteq B \right\}. \tag{16}$$

That is, λ_B is the maximal eigenvalue of adjacency matrix of the subgraph of $\{0,1\}^n$ induced by B.

Two lemmas were shown in [21] to show (12).

Lemma 9 ([21] Prop 1.1): Let C be a code with block length n and minimal distance d. Let B be a subset of $\{0,1\}^n$ with $\lambda_B \geqslant n-2d+1$. Then $|C| = M \leqslant n\,|B|$.

Lemma 10 ([21] Lemma 1.4): Let $B = B_r(\mathbf{0}, n) \subseteq \{0,1\}^n$. The maximal eigenvalue associated with B is $\lambda_B \geqslant 2\sqrt{r(n-r)} - o(n)$.

To prove (12), we note that Lemma 10 implies that a radius $r^* = n/2 - \sqrt{d(n-d)} + o(n)$ exists such that $\lambda_{B(r^*)} \geqslant n-2d+1$. Lemma 9 in turn shows that any code of length n and minimal distance d has at most $n\,|B(r^*)| = nVol(r^*, n)$ codewords. The cardinality of a Hamming ball of radius r is $Vol(r,n) = 2^{H_2(r/n)n+o(n)}$. Equation (14) follows the above argument, hence yielding equation (12).

We note that the above argument can not be used directly to show an upper bound when $d = n/2 - \Theta(\sqrt{n})$. In particular, the $o(n)$ term in Lemma 10 renders the search for a meaningful r^* impossible, as we would ideally require a subset B with λ_B close to $n-2d+1 = \Theta(\sqrt{n})$.

First we provide a proposition in place of Lemma 10 when the radius of Hamming ball does not scale linearly in n.

Proposition 11: Let $B = B_3(\mathbf{0}, n) \subseteq \{0,1\}^n$ be the Hamming ball of radius 3. The maximal eigenvalue associated with B is $\lambda_B \geqslant \sqrt{3+\sqrt{6}}\sqrt{n} \approx 2.334\sqrt{n}$.

Proof: Recall the definition of the maximal eigenvalue in (16). We prove the proposition by constructing a function f with support in B, and for which $\langle Af, f\rangle / \langle f, f\rangle = \sqrt{3+\sqrt{6}}\sqrt{n}$. The function f will be symmetric, namely its value at a point will depend only on the Hamming weight of the point. With a slight abuse of notation, such a function is fully defined by its values $f(0), f(1), \ldots, f(n)$ at Hamming weights $0, 1, \ldots, n$.

Set $f(0) = 1$, $f(j) = 0$ for $j \geqslant 4$, and let $\lambda f(i) = Af(i) = if(i-1) + (n-i)f(i+1)$ for $i = 0, 1, 2$ (assuming $f(-1) = 0$), where $\lambda = t\sqrt{n}$. We have

$$f(1) = \frac{\lambda f(0)}{n} = \frac{t}{\sqrt{n}},\ f(2) = \frac{\lambda f(1) - 1f(0)}{n-1} = \frac{t^2-1}{n-1},$$
$$f(3) = \frac{\lambda f(2) - 2f(1)}{n-2} = \frac{1}{n-2}\left(\frac{t^2-1}{n-1}t\sqrt{n} - 2\frac{t}{\sqrt{n}}\right).$$

We may use the values $f(i)$ and calculate

$$\langle Af, f\rangle = 2t\sqrt{n} + t(t^2-1)^2\frac{n\sqrt{n}}{n-1},$$
$$\langle f, f\rangle = 1 + t^2 + (t^2-1)^2/2 + t^2(t^2-3)^2/6 + o(1).$$

We are now ready to optimize the value

$$\frac{\langle Af, f\rangle}{\langle f, f\rangle} = \left[\frac{2t + t(t^2-1)^2}{(t^6 - 3t^4 + 9t^2 + 9)/6} + o(1)\right]\sqrt{n} \qquad (17)$$

over $t > 0$. Taking $t = \sqrt{3+\sqrt{6}}$, the square bracket term in (17) becomes $\sqrt{3+\sqrt{6}} + o(1)$. ∎

In order to provide a bound as tight as possible, we improve upon Lemma 9 and show the following proposition.

Proposition 12: Let C be a code with block length n and minimal distance d. Let B be a subset of $\{0,1\}^n$ with $\lambda_B > n-2d$. Then $|C| = M \leqslant \frac{n}{\lambda_B - (n-2d)}\,|B|$.

The proof can be shown using a similar argument as in the proof of Lemma 9 in [21], and is provided in Appendix-B for reference.

With Proposition 11 and Proposition 12, we are ready to state the upper bound on $A(n, n/2 - \sqrt{n})$.

Theorem 13: If a (n, M, d) binary code C has minimum distance $d \geqslant n/2 - \sqrt{n}$, then $M \leqslant \frac{\sqrt{n}}{\sqrt{3+\sqrt{6}}-2}Vol(3,n) = O(n^{3.5})$.

Proof: Let $B = B_3(\mathbf{0}, n)$ be the radius-3 Hamming ball. The maximal eigenvalue $\lambda_B \geqslant \sqrt{3+\sqrt{6}}\sqrt{n}$ according to Proposition 11. Since $n - 2d \leqslant 2\sqrt{n} < \lambda_B$, the cardinality of C can be upper bounded using Proposition 12 as

$$M \leqslant \frac{n}{\lambda_B - (n-2d)}\,|B| \leqslant \frac{\sqrt{n}}{\sqrt{3+\sqrt{6}}-2}Vol(3,n).$$

∎

Remark 2: We note that the argument above can upper bound the size as $M = O(n^{3.5})$ as long as $(n-2d)/\sqrt{n}$ is strictly smaller than $\sqrt{3+\sqrt{6}}$. That is, for any $d \geqslant n/2 - t\sqrt{n}$, for some constant $t < \sqrt{3+\sqrt{6}}/2 \approx 1.167$, we have $A(n,d) = O(n^{3.5})$.

Using similar technique as in the proof of Proposition 11, one may show lower bounds of the maximal eigenvalues associated with Hamming balls of different radii. For example, we can list bounds of $\lambda_{B_r}/\sqrt{n}$ for $2 \leqslant r \leqslant 8, r \neq 3$:

$$\lambda_{B_2}/\sqrt{n} \geqslant \sqrt{3}, \qquad \lambda_{B_4}/\sqrt{n} \geqslant (5+\sqrt{10})^{0.5} \approx 2.857,$$
$$\lambda_{B_5}/\sqrt{n} \geqslant 3.324, \qquad \lambda_{B_6}/\sqrt{n} \geqslant 3.75,$$
$$\lambda_{B_7}/\sqrt{n} \geqslant 4.14, \qquad \lambda_{B_8}/\sqrt{n} \geqslant 4.51.$$

This implies $A(n, n/2 - t_1\sqrt{n}) = O(n^{2.5})$, $A(n, n/2 - t_2\sqrt{n}) = O(n^{4.5})$, and so on, for all $t_1 > \sqrt{3}/2 \approx 0.866$ and $t_2 > 1.428$. In particular, we have the following bound

Theorem 14: If a (n, M, d) binary code C has minimum distance $d \geqslant n/2 - 2\sqrt{n}$, then $M \leqslant \frac{\sqrt{n}}{4.14-4}Vol(7,n) = O(n^{7.5})$.

Proof: The theorem can be proved using a similar argument as Theorem 13 by taking $B = B_7(\mathbf{0}, n)$. ∎

In general, by finding the values of, or the lower bounds thereof, the maximal eigenvalues λ_{B_r} for $r = 1, 2, \ldots$, one may obtain a sequence of upper bounds on $A(n, d = n/2 - t\sqrt{n})$ for various t.

IV. Concluding Remarks

In this paper, we study the asymptotic performance bounds of the cardinality of codes with minimum distance $d = n/2 - \Theta(\sqrt{n})$. The codes in this regime have vanishing rate and thus renders ineffective most bounds that dictate the tradeoff between code rate and relative distance. We obtain a sequence of lower bounds based on a cyclic code construction, and two upper bound for $d \geqslant n/2 - \sqrt{n}$ and $d \geqslant n/2 - 2\sqrt{n}$, respectively.

The proposed cyclic linear code is the first construction observed in the specified distance regime, and allows one to construct codes of sizes polynomial in n. The proof of the upper

bound makes extensive use of Fourier analysis on the Hamming cube as a group, and the calculation of the maximal eigenvalue associated with Hamming balls of small radii.

An interesting problem for future work is to study the relationship between the maximal eigenvalue λ_B associated with $B = B_r(\mathbf{0}, n)$ and the radius r, for $r \ll n$. The solution can be used to provide a sequence of upper bounds on $A(n, d)$ for the regime $d = n/2 - \Theta(\sqrt{n})$.

APPENDIX

A. Harmonic Analysis

We compile in this section harmonic analysis preliminaries as in [21], [22]. See [22] for a more detailed treatment. Here we list several necessary definitions and simple facts.

Consider the abelian group structure $\mathbb{F}_2^n = (\mathbb{Z}/2\mathbb{Z})^n$ on the hypercube $\{0,1\}^n$. The characters of the abelian group $\mathbb{F}_2^n$ are $\{\chi_z\}_{z\in\mathbb{F}_2^n}$, where $\chi_z : \{0,1\}^n \to \{-1,1\}$ is given by $\chi_z(x) = (-1)^{\langle x,z\rangle}$ and $\langle x, z\rangle = \sum_{i=1}^n x_i z_i$.

Consider the $\mathbb{R}$-vector space $\mathcal{L}(\mathbb{F}_2^n) = \{f : \mathbb{F}_2^n \to \mathbb{R}\}$ endowed with the inner product $\langle\cdot,\cdot\rangle$, associated with the uniform distribution on $\{0,1\}^n$:

$$\langle f, g\rangle = \mathbb{E}_{U_n} fg = \frac{1}{2^n}\sum_{x\in\mathbb{F}_2^n} f(x)g(x). \tag{18}$$

The set of 2^n characters $\{\chi_z\}_{z\in\mathbb{F}_2^n}$ form an orthonormal basis in the space $\mathcal{L}(\mathbb{F}_2^n)$, equipped with uniform probability distribution. That is, for each $z, z' \in \{0,1\}^n$, $\langle\chi_z, \chi_{z'}\rangle = \delta_{z,z'}$, where δ is the Kronecker delta function. The *Fourier transform* of a function $f \in \mathcal{L}(\mathbb{F}_2^n)$ is the function $\hat{f} \in \mathcal{L}(\mathbb{F}_2^n)$ given by the coefficients of the unique expansion of f in terms of the characters:

$$f(x) = \sum_z \hat{f}(z)\chi_z(x) \text{ or equivalently, } \hat{f}(z) = \langle f, \chi_z\rangle. \tag{19}$$

One may show that $\hat{\hat{f}} = 2^n f$, and $\mathbb{E} f = \hat{f}(0)$. If $f, g \in \mathcal{L}(\mathbb{F}_2^n)$, we have Parseval's identity: $\langle f, g\rangle = \sum_z \hat{f}(z)\hat{g}(z) = 2^n\left\langle \hat{f}, \hat{g}\right\rangle$. A special case of the above equality is the following equality: $\mathbb{E} f^2 = \sum_z \hat{f}(z)^2 = \left\|\hat{f}\right\|_2^2$.

The convolution of f and g is defined by $(f * g)(x) = \mathbb{E}_y f(y)g(x+y)$. The convolution transforms to dot product: $\widehat{f * g} = \hat{f}\cdot\hat{g}$. The convolution operator is commutative and associative. One may also show that for arbitrary functions $f, g, h \in \mathcal{L}(\mathbb{F}_2^n)$, the following equality holds:

$$\langle f * g, h\rangle = \langle f, g * h\rangle. \tag{20}$$

In this section and in Appendix-B, let $L(x) = 2^n$ for $x \in \{0,1\}^n$ with $w_H(x) = 1$ and $L(x) = 0$ otherwise. For any $f \in \mathcal{L}(\mathbb{F}_2^n)$ holds $Af = f * L$ because for $x \in \mathbb{F}_2^n$, $Af(x) = \sum_{y:d_H(x,y)=1} f(y) = \sum_{y:w_H(y)=1} f(x+y) = \mathbb{E}_y L(y)f(x+y) = (f * L)(x)$. The Fourier transform of L is the function $\hat{L}$ given by $\hat{L}(z) = \langle L, \chi_z\rangle = \sum_{x:w_H(x)=1}(-1)^{\langle x,z\rangle} = n - 2\cdot w_H(z)$.

For $C \subset \mathbb{F}_2^n$, let $1_C \in \mathcal{L}(\mathbb{F}_2^n)$ be the indicator function of C. It can be shown that a code C has minimum distance d if and only if $(1_C * 1_C)(x) = 0$ for all $0 < w_H(x) < d$.

B. Proof of Proposition 12

Let f_B be an eigenfunction supported on B corresponding to its maximal eigenvalue λ_B. That is $\lambda_B = \langle Af_B, f_B\rangle/\langle f_B, f_B\rangle$. It is known that the maximum can be attained with an non-negative function f_B, and further we have $Af_B \geqslant \lambda_B f_B$ (see [23, p.13-15 and appendix C]) for details). We write $f = f_B$ and $\lambda = \lambda_B$ interchangeably, and denote the Hamming weight of $x \in \mathbb{F}_2^n$ by $|x| = w_H(x)$, in this proof. As f is supported on B, Cauchy-Schwarz inequality yields the following:

$$\mathbb{E}^2 f = \langle f, 1_B\rangle^2 \leqslant \mathbb{E} f^2 \cdot \mathbb{E}(1_B)^2 = \mathbb{E} f^2 \cdot |B|/2^n. \tag{21}$$

Let $\phi \in \mathcal{L}(\mathbb{F}_2^n)$ be a function such that $(\hat{\phi})^2 = \widehat{\phi * \phi} = 1_C * 1_C$. Equivalently, $\phi * \phi = 2^n \widehat{1_C * 1_C} = 2^n \widehat{1_C}^2$. Therefore we have

$$\phi * \phi \geqslant 0 \text{ and } \frac{\mathbb{E}(\phi^2)}{\mathbb{E}^2(\phi)} = \frac{(\phi * \phi)(0)}{\hat{\phi}^2(0)} = |C|. \tag{22}$$

Now let $F = \phi * f$. We estimate the product $\langle AF, F\rangle$ in two ways. First,

$$\begin{aligned}\langle AF, F\rangle &= \langle (\phi * f) * L, \phi * f\rangle = \langle \phi * \phi * f, f * L\rangle \\ &= \langle \phi * \phi * f, Af\rangle \geqslant \langle \phi * \phi * f, \lambda f\rangle \\ &= \lambda\langle \phi * f, \phi * f\rangle = \lambda\langle F, F\rangle = \lambda\,\mathbb{E} F^2.\end{aligned}$$

Second, by Parseval's identity,

$$\begin{aligned}\langle AF, F\rangle &= 2^n\left\langle \widehat{AF}, \hat{F}\right\rangle = 2^n\left\langle \hat{L}\cdot\hat{F}, \hat{F}\right\rangle \\ &= \sum_z (n - 2|z|)\,\hat{F}^2(z).\end{aligned}$$

Since $\hat{F} = \hat{\phi}\cdot\hat{f}$ and $(\hat{\phi})^2(z) = (1_C * 1_C)(z)$, $\hat{F}(z) = 0$ for all $0 < |z| < d$. We can estimate $\langle AF, F\rangle$ by

$$\begin{aligned}&\sum_z (n - 2|z|)\,\hat{F}^2(z) = n\hat{F}^2(0) + \sum_{z:|z|\geqslant d}(n - 2|z|)\,\hat{F}^2(z) \\ &\leqslant n\hat{F}^2(0) + (n - 2d)\sum_z \hat{F}^2(z) = n\,\mathbb{E}^2 F + (n - 2d)\,\mathbb{E} F^2.\end{aligned}$$

Combining the two estimates, we have the following inequality: $n\,\mathbb{E}^2 F \geqslant (\lambda - (n - 2d))\,\mathbb{E} F^2$. Since

$$\mathbb{E}^2 F = \mathbb{E}^2(\phi * f) = [\widehat{\phi * f}(0)]^2 = [\hat{\phi}(0)\hat{f}(0)]^2 = \mathbb{E}^2\,\phi\,\mathbb{E}^2 f,$$

$$\begin{aligned}\mathbb{E} F^2 &= \langle F, F\rangle = \langle \phi * f, \phi * f\rangle = \langle \phi * \phi, f * f\rangle \\ &\geqslant 1/2^n(\phi * \phi)(0)(f * f)(0) = 1/2^n\,\mathbb{E}\,\phi^2\,\mathbb{E} f^2,\end{aligned}$$

as $\phi * \phi = 2^n \cdot \widehat{1_C}^2 \geqslant 0$, we now have

$$n\,\mathbb{E}^2\,\phi\,\mathbb{E}^2 f \geqslant (\lambda - (n - 2d))\,\frac{1}{2^n}\,\mathbb{E}\,\phi^2\,\mathbb{E} f^2. \tag{23}$$

Leveraging equations (21), (22), and (23), the size of any code C with minimum distance d is

$$|C| = \frac{\mathbb{E}\,\phi^2}{\mathbb{E}^2\,\phi} \leqslant \frac{n}{\lambda - (n - 2d)}\cdot 2^n\frac{\mathbb{E}^2 f}{\mathbb{E} f^2} \leqslant \frac{n}{\lambda - (n - 2d)}|B|.$$

■

REFERENCES

[1] R. Ratasuk, N. Mangalvedhe, Y. Zhang, M. Robert, and J.-P. Koskinen, "Overview of narrowband IoT in LTE Rel-13," in *2016 IEEE conference on standards for communications and networking (CSCN)*. IEEE, 2016, pp. 1–7.

[2] M. Fereydounian, M. V. Jamali, H. Hassani, and H. Mahdavifar, "Channel coding at low capacity," in *2019 IEEE Information Theory Workshop (ITW)*. IEEE, 2019, pp. 1–5.

[3] F. Abbasi, H. Mahdavifar, and E. Viterbo, "Hybrid non-binary repeated polar codes for low-SNR regime," in *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2021, pp. 1742–1747.

[4] I. Dumer and N. Gharavi, "Codes for high-noise memoryless channels," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*. IEEE, 2020, pp. 101–105.

[5] F. Abbasi, H. Mahdavifar, and E. Viterbo, "Polar coded repetition for low-capacity channels," in *2020 IEEE Information Theory Workshop (ITW)*. IEEE, 2021, pp. 1–5.

[6] I. Dumer and N. Gharavi, "Combined polar-LDPC design for channels with high noise," in *2021 IEEE Information Theory Workshop (ITW)*. IEEE, 2021, pp. 1–6.

[7] S. Litsyn, "An update table of the best binary codes known," *Handbook of Coding Theory*, 1998.

[8] E. Agrell, "Bounds for unrestricted binary codes." [Online]. Available: https://codes.se/bounds/unr.html

[9] A. Brouwer. [Online]. Available: https://www.win.tue.nl/\%7Eaeb/codes/binary-1.html

[10] T. Jiang and A. Vardy, "Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes," *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1655–1664, 2004.

[11] P. Gaborit and G. Zemor, "Asymptotic improvement of the gilbert–varshamov bound for linear codes," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 3865–3872, 2008.

[12] F. J. MacWilliams and N. J. A. Sloane, *The theory of error correcting codes*. Elsevier, 1977, vol. 16.

[13] V. Guruswami, A. Rudra, and M. Sudan, "Essential coding theory."

[14] V. M. Sidel'nikov, "On mutual correlation of sequences," in *Doklady Akademii Nauk*, vol. 196, no. 3. Russian Academy of Sciences, 1971, pp. 531–534.

[15] A. Tietäväinen, "Bounds for binary codes just outside the Plotkin range," *Information and Control*, vol. 47, no. 2, pp. 85–93, 1980.

[16] I. Krasikov and S. Litsyn, "On upper bounds for the distance of codes of small size," in *Proceedings of IEEE International Symposium on Information Theory*. IEEE, 1997, p. 84.

[17] E. N. Gilbert, "A comparison of signalling alphabets," *The Bell system technical journal*, vol. 31, no. 3, pp. 504–522, 1952.

[18] M. Plotkin, "Binary codes with specified minimum distance," *IRE Transactions on Information Theory*, vol. 6, no. 4, pp. 445–450, 1960.

[19] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities," *IEEE Transactions on Information Theory*, vol. 23, no. 2, pp. 157–166, 1977.

[20] P. Delsarte, "An algebraic approach to the association schemes of coding theory," *Philips Res. Rep. Suppl.*, vol. 10, pp. vi+–97, 1973.

[21] M. Navon and A. Samorodnitsky, "Linear programming bounds for codes via a covering argument," *Discrete & Computational Geometry*, vol. 41, no. 2, p. 199, 2009.

[22] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on Boolean functions*. Citeseer, 1989.

[23] J. Friedman and J.-P. Tillich, "Generalized Alon–Boppana theorems and error-correcting codes," *SIAM Journal on Discrete Mathematics*, vol. 19, no. 3, pp. 700–718, 2005.