Upper bound on list-decoding radius of binary codes

Yury Polyanskiy

***Abstract*—Consider the problem of packing Hamming balls of a given relative radius subject to the constraint that they cover any point of the ambient Hamming space with multiplicity at most L. For odd $L \geq 3$ an asymptotic upper bound on the rate of any such packing is proven. The resulting bound improves the best known bound (due to Blinovsky'1986) for rates below a certain threshold. The method is a superposition of the linear-programming idea of Ashikhmin, Barg and Litsyn (that was used previously to improve the estimates of Blinovsky for $L = 2$) and a Ramsey-theoretic technique of Blinovsky. As an application it is shown that for all odd L the slope of the rate-radius tradeoff is zero at zero rate.**

***Index Terms*—Combinatorial coding theory, list-decoding, converse bounds**

I. Main result and discussion

One of the most well-studied problems in information theory asks to find the maximal rate at which codewords can be packed in binary space with a given minimum distance between codewords. Operationally, this (still unknown) rate gives the capacity of the binary input-output channel subject to adversarial noise of a given level. A natural generalization was considered by Elias and Wozencraft [1], [2], who allowed the decoder to output a list of size L. In this paper we provide improved upper bounds on the latter question.

Our interest in bounding the asymptotic tradeoff for the list-decoding problem is motivated by our study of fundamental limits of joint source-channel communication [3]. The best known converse bound for that problem – a straightforward extension of [3, Theorem 7] to lists of size > 1 – reduces to bounding rate for the list-decoding problem, cf. [4, Theorem 6].

We proceed to formal definitions and brief overview of known results. For a binary code $\mathcal{C} \subset \mathbb{F}_2^n$ we define its list-size L decoding radius as

$$\tau_L(\mathcal{C}) \triangleq \frac{1}{n} \max\{r : \forall x \in \mathbb{F}_2^n \ |\mathcal{C} \cap \{x + B_r^n\}| \leq L\},$$

where Hamming ball B_r^n and Hamming sphere S_r^n are defined as

$$B_r^n \triangleq \{x \in \mathbb{F}_2^n : |x| \leq r\}, \tag{1}$$

$$S_r^n \triangleq \{x \in \mathbb{F}_2^n : |x| = r\} \tag{2}$$

with $|x| = |\{i : x_i = 1\}|$ denoting the Hamming weight of x. Alternatively, we may define τ_L as follows:[1]

$$\tau_L(\mathcal{C}) = \frac{1}{n}\left(\min\left\{\mathrm{rad}(S) : S \in \binom{\mathcal{C}}{L+1}\right\} - 1\right),$$

where $\mathrm{rad}(S)$ denotes radius of the smallest ball containing S (known as Chebyshev radius):

$$\mathrm{rad}(S) \triangleq \min_{y \in \mathbb{F}_2^n} \max_{x \in S} |y - x|.$$

The asymptotic tradeoff between rate and list-decoding radius τ_L is defined as usual:

$$\tau_L^*(R) \triangleq \limsup_{n\to\infty} \max_{\mathcal{C}:|\mathcal{C}|\geq 2^{nR}} \tau_L(\mathcal{C}) \tag{3}$$

$$R_L^*(\tau) \triangleq \limsup_{n\to\infty} \max_{\mathcal{C}:\tau_L(\mathcal{C})\geq\tau} \frac{1}{n}\log|\mathcal{C}| \tag{4}$$

The best known upper (converse) bounds on this tradeoff are as follows:

- List size $L = 1$: The best bound to date was found by McEliece, Rodemich, Rumsey and Welch [5]:

$$R_1^*(\tau) \leq R_{LP2}(2\tau), \tag{5}$$

$$R_{LP2}(\delta) \triangleq \min \log 2 - h(\alpha) + h(\beta), \tag{6}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ and minimum is taken over all $0 \leq \beta \leq \alpha \leq 1/2$ satisfying

$$2\frac{\alpha(1-\alpha) - \beta(1-\beta)}{1 + 2\sqrt{\beta(1-\beta)}} \leq \delta$$

For rates $R < 0.305$ this bound coincides with the simpler bound:

$$\tau_1^*(R) \leq \frac{1}{2}\delta_{LP1}(R), \tag{7}$$

$$\delta_{LP1}(R) \triangleq \frac{1}{2} - \sqrt{\beta(1-\beta)}, \tag{8}$$

$$R = \log 2 - h(\beta), \quad \beta \in [0, 1/2] \tag{9}$$

- List size $L = 2$: The bound found by Ashikhmin, Barg and Litsyn [6] is given as[2]

$$R_2^*(\tau) \leq \log 2 - h(2\tau) + R_{up}(2\tau, 2\tau),$$

where $R_{up}(\delta, \alpha)$ is the best known upper bound on rate of codes with minimal distance δn constrained to live on Hamming spheres $S_{\alpha n}^n$. The expression for $R_{up}(\delta, \alpha)$ can be obtained by using the linear programming bound

YP is with the Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139 USA. e-mail: yp@mit.edu.

The research was supported by the NSF grant CCF-13-18620 and NSF Center for Science of Information (CSoI) under grant agreement CCF-09-39370.

[1] $\binom{\mathcal{C}}{j}$ denotes the set of all subsets of $\mathcal{C}$ of size j.

[2] This result follows from optimizing [6, Theorem 4]. It is slightly stronger than what is given in [6, Corollary 5].

from [5] and applying Levenshtein's monotonicity, cf. [7, Lemma 4.2(6)]. The resulting expression is

$$R_2^*(\tau) \le \begin{cases} R_{LP2}(2\tau)\,, & \tau \le \tau_0 \\ \log 2 - h(2\tau) + h(u(\tau)), & \tau > \tau_0\,, \end{cases} \tag{10}$$

where $\tau_0 \approx 0.1093$ and

$$u(\tau) = \frac{1}{2} - \sqrt{\frac{1}{4} - (\sqrt{\tau - 3\tau^2} - \tau)^2}$$

(cf. [7, (9)]).

- For list sizes $L \ge 3$: The original bound of Blinovsky [8] appears to be the best (before this work):

$$\tau_L^*(R) \le \sum_{i=1}^{\lceil L/2 \rceil} \frac{\binom{2i-2}{i-1}}{i} (\lambda(1-\lambda))^i\,, \tag{11}$$

$$R = 1 - h(\lambda)\,, \lambda \in [0, 1/2] \tag{12}$$

Note that [8] also gives a non-constructive lower bound on $\tau_L^*(R)$. Results on list-decoding over non-binary alphabets are also known, see [9], [10].

In this paper we improve the bound of Blinovsky for lists of odd size and rates below a certain threshold. To that end we will mix the ideas of Ashikhmin, Barg and Litsyn (namely, extraction of a large spectrum component from the code) and those of Blinovsky (namely, a Ramsey-theoretic reduction to study of symmetric subcodes).

To present our main result, we need to define exponent of Krawtchouk polynomial $K_{\beta n}(\xi n) = \exp\{nE_\beta(\xi) + o(n)\}$. For $\xi \in [0, \frac{1}{2} - \sqrt{\beta(1-\beta)}]$ the value of $E_\beta(\xi)$ was found in [11]. Here we give it in the following parametric form, cf. [12] or [13, Lemma 4]:

$$E_\beta(\xi) = \xi \log(1-\omega) + (1-\xi)\log(1+\omega) - \beta \log \omega \tag{13}$$

$$\xi = \frac{1}{2}(1 - (1-\beta)\omega - \beta\omega^{-1})\,, \tag{14}$$

where

$$\omega \in \left[\frac{\beta}{1-\beta}, \sqrt{\frac{\beta}{1-\beta}}\right].$$

Our main result is the following:

Theorem 1. *Fix list size $L \ge 2$, rate R and an arbitrary $\beta \in [0, 1/2]$ with $h(\beta) \le R$. Then any sequence of codes $\mathcal{C}_n \subset \{0,1\}^n$ of rate R satisfies*

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \le \max_{j,\xi_0} \xi_0 g_j\left(1 - \frac{\xi_1}{2\xi_0}\right) + (1-\xi_0) g_j\left(\frac{\xi_1}{2(1-\xi_0)}\right)\,, \tag{15}$$

where maximization is over ξ_0 satisfying

$$0 \le \xi_0 \le \frac{1}{2} - \sqrt{\beta(1-\beta)} \tag{16}$$

and j ranging over $\{0, 1, 3, \dots, 2k+1, \dots, L\}$ if L is odd and over $\{0, 2, \dots, 2k, \dots L\}$ if L is even. Quantity $\xi_1 = \xi_1(\xi_0, \delta, R)$ is a unique solution of

$$R + h(\beta) - 2E_\beta(\xi_0) = h(\xi_0) - \xi_0 h\left(\frac{\xi_1}{2\xi_0}\right) - (1-\xi_0) h\left(\frac{\xi_1}{2(1-\xi_0)}\right)\,, \tag{17}$$

on the interval $[0, 2\xi_0(1-\xi_0)]$ and functions $g_j(\nu)$ are defined as

$$g_j(\nu) \triangleq \frac{L\nu - \mathbb{E}\left[|2W - L - j|^+\right]}{L+j}\,, \quad W \sim \mathrm{Bino}(L, \nu) \tag{18}$$

TABLE I

RATES FOR WHICH NEW BOUND IMPROVES STATE OF THE ART

List size L	Range of rates
$L = 3$	$0 < R \le 0.361$
$L = 5$	$0 < R \le 0.248$
$L = 7$	$0 < R \le 0.184$
$L = 9$	$0 < R \le 0.144$
$L = 11$	$0 < R \le 0.108$

As usual with bounds of this type, cf. [14], it appears that taking $h(\beta) = R$ can be done without loss. Under such choice, our bound outperforms Blinovsky's for all odd L and all rates small enough (see Corollary 3 below). The bound for $L = 3$ is compared in Fig. 1 with the result of Blinovsky numerically. For larger odd L the comparison is similar, but the range of rates where our bound outperforms Blinovsky's becomes smaller, see Table I.

Evaluation of Theorem 1 is computationally possible, but is somewhat tedious.[3] Fortunately, for small L the maximum over ξ_0 and j is attained at $\xi_0 = \frac{1}{2} - \sqrt{\beta(1-\beta)}$ and $j = 1$. We rigorously prove this for $L = 3$:

Corollary 2. *For list-size $L = 3$ we have*

$$\tau_L^*(R) \le \frac{3}{4}\delta - \frac{1}{16}\left(\frac{(2\delta - \xi_1)^3}{\delta^2} + \frac{\xi_1^3}{(1-\delta)^2}\right)\,, \tag{19}$$

where $\delta \in (0, 1/2]$ and $\xi_1 \in [0, 2\delta(1-\delta)]$ are functions of R determined from

$$R = h\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)\,, \tag{20}$$

$$R = \log 2 - \delta h\left(\frac{\xi_1}{2\delta}\right) - (1-\delta) h\left(\frac{\xi_1}{2(1-\delta)}\right) \tag{21}$$

Another interesting implication of Theorem 1 is that it allows us to settle the question of slope of the curve $R_L^*(\tau)$ at zero rate. Notice that Blinovsky's converse bound (11) has a negative slope, while his achievability bound has a zero slope. Our bound always has a zero slope for odd L (but not for even L, see [15] for details):

[3] Notice that proofs of each of the two Corollaries below contain a different relaxation of the bound (15), which may appear useful separately.

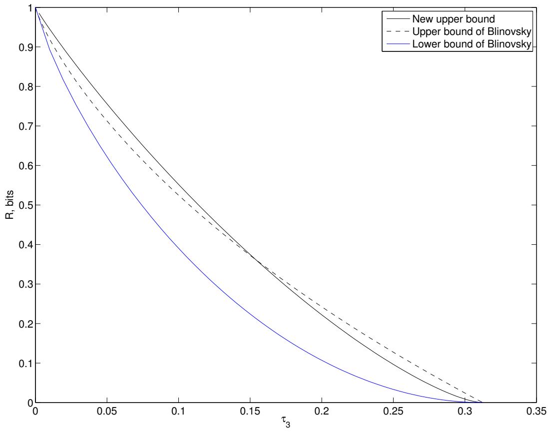

Fig. 1. Comparison of bounds on $R_L^*(\tau)$ for list size $L = 3$

Corollary 3. *Fix arbitrary odd $L \geq 3$. There exists $R_0 = R_0(L) > 0$ such that for all rates $R < R_0$ we have*

$$\tau_L^*(R) \leq g_1(\delta_{LP1}(R))\,, \tag{22}$$

where $g_1(\cdot)$ is a degree-L polynomial defined in (18)*. In particular,*

$$\left.\frac{d}{d\tau}\right|_{\tau=\tau_L^*(0)} R_L^*(\tau) = 0\,, \tag{23}$$

where the zero-rate radius is $\tau_L^*(0) = \frac{1}{2} - 2^{-L-1}\binom{L}{\frac{L-1}{2}}$.

We close our discussion with some additional remarks:

1) The bound in Theorem 1 can be slightly improved by replacing $\delta_{LP1}(R)$, that appears in the right-hand side of (16), with a better bound, a so-called second linear-programming bound $\delta_{LP2}(R)$ from [5]. This would enforce the usage of the more advanced estimate of Litsyn [16, Theorem 5] and complicate analysis significantly. Notice that $\delta_{LP2}(R) \neq \delta_{LP1}(R)$ only for rates $R \geq 0.305$. If we focus attention only on rates where new bound is better than Blinovsky's, such a strengthening only affects the case of $L = 3$ and results in a rather minuscule improvement (for example, for rate $R = 0.33$ the improvement is $\approx 3 \cdot 10^{-5}$).
2) For even L it appears that $h(\beta) = R$ is no longer optimal. However, the resulting bound does not appear to improve upon Blinovsky's.
3) When L is large (e.g. 35) the maximum in (15) is not always attained by either $j = 1$ or $\xi_0 = \delta_{LP1}(R)$. It is not clear whether such anomalies only happen in the region of rates where our bound is inferior to Blinovsky's.
4) The result of Corollary 3 follows by weakening (15) to

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \leq \max_{j,\xi_0} g_j\,(\xi_0) = \max_j g_j(\delta_{LP1}(R))\,.$$

The $R < R_0(L)$ condition is only used to show that the maximum is attained at $j = 1$.

II. Proofs

Several key Lemmas are omitted for space constraints, those can be found in [15].

A. Proof of Theorem 1

Consider an arbitrary sequence of codes $\mathcal{C}_n$ of rate R. As in [6] we start by using Delsarte's linear programming to select a large component of the distance distribution of the code. Namely, we apply result of Kalai and Linial [11, Proposition 3.2]: For every β with $h(\beta) \leq R$ there exists a sequence $\epsilon_n \to 0$ such that for every code $\mathcal{C}$ of rate R there is a ξ_0 satisfying (16) such that

$$A_{\xi_0 n}(\mathcal{C}) \triangleq \frac{1}{|\mathcal{C}|} \sum_{x,x'\in\mathcal{C}} 1\{|x-x'| = \xi_0 n\} \tag{24}$$

$$\geq \exp\{n(R + h(\beta) - 2E_\beta(\xi_0) + \epsilon_n)\}\,. \tag{25}$$

Without loss of generality (by compactness of the interval $[0, 1/2 - \sqrt{\beta(1-\beta)}]$ and passing to a proper subsequence of codes $\mathcal{C}_{n_k}$) we may assume that ξ_0 selected in (25) is the same for all blocklengths n. Then there is a sequence of subcodes $\mathcal{C}'_n$ of asymptotic rate

$$R' \geq R + h(\beta) - 2E_\beta(\xi_0)$$

such that each $\mathcal{C}'_n$ is situated on a sphere $c_0 + S_{\xi_0}$ surrounding another codeword $c_0 \in \mathcal{C}$. Our key geometric result is: If there are too many codewords on a sphere $c_0 + S_{\xi_0}$ then it is possible to find L of them that are includable in a small ball that also contains c_0. Precisely, we have:

Lemma 4. *Fix $\xi_0 \in (0, 1)$ and positive integer L. There exist a sequence $\epsilon_n \to 0$ such that for any code $\mathcal{C}'_n \subset S_{\xi_0 n}$ of rate $R' > 0$ there exist L codewords $c_1, \ldots, c_L \in \mathcal{C}'_n$ such that*

$$\frac{1}{n}\,\mathrm{rad}(0, c_1, \ldots, c_L) \leq \theta(\xi_0, R', L) + \epsilon_n\,, \tag{26}$$

where

$$\theta(\xi_0, R', L) \triangleq \max_j \theta_j(\xi_0, R', L) \tag{27}$$

$$\theta_j(\xi_0, R', L) \triangleq \xi_0 g_j\left(1 - \frac{\xi_1}{2\xi_0}\right) + (1-\xi_0) g_j\left(\frac{\xi_1}{2(1-\xi_0)}\right), \tag{28}$$

with $\xi_1 = \xi_1(\xi_0)$ found as unique solution on interval $[0, 2\xi_0(1-\xi_0)]$ of

$$R' = h(\xi_0) - \xi_0 h\left(\frac{\xi_1}{2\xi_0}\right) - (1-\xi_0) h\left(\frac{\xi_1}{2(1-\xi_0)}\right), \tag{29}$$

functions g_j are defined in (18) *and j in maximization* (27) *ranging over the same set as in Theorem 1.*

Equipped with Lemma 4 we immediately conclude that

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \leq \max_{\xi_0\in[0,\delta]} \theta(\xi_0, R + h(\beta) - 2E_\beta(\xi_0), L)\,. \tag{30}$$

Clearly, (30) coincides with (15). So it suffices to prove Lemma 4.

B. Proof of Lemma 4

Let $\mathcal{T}_L$ be the (2^L-1)-dimensional space of probability distributions on $\mathbb{F}_2^L$. If $T \in \mathcal{T}_L$ then we have

$$T = (t_v, v \in \mathbb{F}_2^L) \qquad t_v \ge 0, \sum_v t_v = 1\,.$$

We define distance on $\mathcal{T}_L$ to be the L_∞ one:

$$\|T - T'\| \triangleq \max_{v \in \mathbb{F}_2^L} |t_v - t'_v|\,.$$

Permutation group S_L acts naturally on $\mathbb{F}_2^L$ and this action descends to probability distributions $\mathcal{T}_L$. We will say that T is symmetric if

$$T = \sigma(T) \quad \iff \quad t_v = t_{\sigma(v)} \quad \forall v \in \mathbb{F}_2^L$$

for any permutation $\sigma : [L] \to [L]$. Note that symmetric T is completely specified by $L+1$ numbers (weights of Hamming spheres in $\mathbb{F}_2^L$):

$$\sum_{v:|v|=j} t_v\,, \qquad j = 0, \ldots, L\,.$$

Next, fix some total ordering of $\mathbb{F}_2^n$ (for example, lexicographic). Given a subset $S \subset \mathbb{F}_2^n$ we will say that S is given in ordered form if $S = \{x_1, \ldots, x_{|S|}\}$ and $x_1 < x_2 \cdots < x_{|S|}$ under the fixed ordering on $\mathbb{F}_2^n$. For any subset of codewords $S = \{x_1, \ldots, x_L\}$ given in ordered form we define its *joint type* $T(S)$ as an element of $\mathcal{T}_L$ with

$$t_v \triangleq \frac{1}{n}|j : x_1(j) = v_1, \ldots, x_L(j) = v_j|\,,$$

where here and below $y(j)$ denotes the j-th coordinate of binary vector $y \in \mathbb{F}_2^n$. In this way every subset S is associated to an element of $\mathcal{T}_L$. Note that $T(S)$ is symmetric if and only if the $L \times n$ binary matrix representing S (by combining row-vectors x_j) has the property that the number of columns equal to $[1, 0, \ldots, 0]^T$ is the same as the number of columns $[0, 1, \ldots, 0]^T$ etc. For any code $\mathcal{C} \subset \mathbb{F}_2^n$ we define its average joint type:

$$\bar{T}_L(\mathcal{C}) = \frac{1}{L! \cdot \binom{|\mathcal{C}|}{L}} \sum_\sigma \sum_{S \in \binom{\mathcal{C}}{L}} \sigma(T(S))\,.$$

Evidently, $\bar{T}_L(\mathcal{C})$ is symmetric.

Our proof crucially depends on a (slight extension of the) brilliant idea of Blinovsky [8]:

Lemma 5. *For every $L \ge 1$, $K \ge L$ and $\delta > 0$ there exist a constant $K_1 = K_1(L, K, \delta)$ such that for all $n \ge 1$ and all codes $\mathcal{C} \subset \mathbb{F}_2^n$ of size $|\mathcal{C}| \ge K_1$ there exists a subcode $\mathcal{C}' \subset \mathcal{C}$ of size at least K such that for any $S \in \binom{\mathcal{C}'}{L}$ we have*

$$\|T(S) - \bar{T}_L(\mathcal{C}')\| \le \delta\,. \tag{31}$$

Remark 1. *Note that if $S' \subset S$ then every element of $T(S')$ is a sum of $\le 2^L$ elements of $T(S)$. Hence, joint types $T(S')$ are approximately symmetric also for smaller subsets $|S'| < L$.*

Proof. See [15]. □

Before proceeding further we need to define the concept of an average radius (or a moment of inertia):

$$\overline{\mathrm{rad}}(x_1, \ldots, x_m) \triangleq \min_y \frac{1}{m} \sum_{i=1}^m |x_i - y|\,.$$

Note that the minimizing y can be computed via a per-coordinate majority vote (with arbitrary tie-breaking for even m). Consider now an arbitrary subset $S = \{c_1, \ldots, c_L\}$ and define for each $j \ge 0$ the following functions

$$h_j(S) \triangleq \frac{1}{n}\,\overline{\mathrm{rad}}(\underbrace{0, \ldots, 0}_{j \text{ times}}, c_1, \ldots, c_L)\,.$$

It is easy to find an expression for $h_j(S)$ in terms of the joint-type of S:

$$h_j(S) = \frac{1}{L+j}\left(\mathbb{E}[W] - \mathbb{E}[|2W - L - j|^+]\right) \tag{32}$$

$$\mathbb{P}[W = w] = \sum_{v:|v|=w} t_v\,, \tag{33}$$

where t_v are components of the joint-type $T(S) = \{t_v, v \in \mathbb{F}_2^L\}$. To check (32) simply observe that if one arranges L codewords of S in an $L \times n$ matrix and also adds j rows of zeros, then computation of $h_j(S)$ can be done per-column: each column of weight w contributes

$$\min(w, L + j - w) = w - |2w - L - j|^+$$

to the sum. In view of expression (32) we will abuse notation and write

$$h_j(T(S)) \triangleq h_j(S)\,.$$

We now observe that for symmetric codes satisfying (31) average-radii $h_j(S)$ in fact determine the regular radius:

Lemma 6. *Consider an arbitrary code $\mathcal{C}$ satisfying conclusion (31) of Lemma 5. Then for any subset $S = \{c_1, \ldots, c_L\} \subset \mathcal{C}$ we have*

$$\left|\mathrm{rad}(0, c_1, \ldots, c_L) - n \cdot \max_j h_j(\bar{T}_L(\mathcal{C}))\right| \le 2^L(1 + \delta n)\,, \tag{34}$$

where j in maximization (34) ranges over $\{0, 1, 3, \ldots, 2k+1, \ldots, L\}$ if L is odd and over $\{0, 2, \ldots, 2k, \ldots L\}$ if L is even.

Proof. See [15]. □

Lemma 7. *There exist constants C_1, C_2 depending only on L such that for any $\mathcal{C} \subset \mathbb{F}_2^n$ the joint-type $\bar{T}_L(\mathcal{C})$ is approximately a mixture of product Bernoulli distributions*[4]*, namely:*

$$\left\|\bar{T}_L(\mathcal{C}) - \frac{1}{n}\sum_{i=1}^n \mathrm{Bern}^{\otimes L}(\lambda_i)\right\| \le \frac{C_1}{|\mathcal{C}|}\,, \tag{35}$$

[4]Distribution $\mathrm{Bern}^{\otimes L}(\lambda)$ assigns probability $\lambda^{|v|}(1-\lambda)^{L-|v|}$ to element $v \in \mathbb{F}_2^L$.

where $\lambda_i = \frac{1}{|\mathcal{C}|}\sum_{c\in\mathcal{C}} 1\{c(i)=1\}$ be the density of ones in the j-th column of a $|\mathcal{C}|\times n$ matrix representing the code. In particular,

$$\left|h_j(\bar{T}_L(\mathcal{C})) - \frac{1}{n}\sum_j g_j(\lambda_j)\right| \le \frac{C_2}{|\mathcal{C}|}\,, \qquad (36)$$

where functions g_j were defined in (18)*.*

Proof. See [15] □

Lemma 8. *Functions g_j defined in* (18) *are concave on* $[0,1]$.

Proof. See [15] □

Proof of Lemma 4. Our plan is the following:

1) Apply Elias-Bassalygo reduction to pass from $\mathcal{C}'_n$ to a subcode $\mathcal{C}''_n$ on an intersection of two spheres $S_{\xi_0 n}$ and $y+S_{\xi_1 n}$.
2) Use Lemma 5 to pass to a symmetric subcode $\mathcal{C}'''_n \subset \mathcal{C}''_n$
3) Use Lemmas 7-8 to estimate maxima of average radii h_j over $\mathcal{C}'''_n$.
4) Use Lemma 6 to transport statement about h_j to a statement on $\tau_L(\mathcal{C}'''_n)$.

We proceed to details. It is sufficient to show that for some constant $C=C(L)$ and arbitrary $\delta>0$ estimate (26) holds with $\epsilon_n = C\delta$ whenever $n \ge n_0(\delta)$. So we fix $\delta>0$ and consider a code $\mathcal{C}' \subset S_{\xi_0 n} \subset \mathbb{F}_2^n$ with $|\mathcal{C}'| \ge \exp\{nR' + o(n)\}$. Note that for any r , even m with $m/2 \le \min(r, n-r)$ and arbitrary $y \in S_r^n$ intersection $\{y+S_m^n\}\cap S_r^n$ is isometric to the product of two lower-dimensional spheres:

$$\{y+S_m^n\}\cap S_r^n \cong S_{r-m/2}^r \times S_{m/2}^{n-r}\,. \qquad (37)$$

Therefore, we have for $r=\xi_0 n$ and valid m:

$$\sum_{y\in S_r^n} |\{y+S_m^n\}\cap \mathcal{C}'| = |\mathcal{C}'|\binom{\xi_0 n}{\xi_0 n - m/2}\binom{n(1-\xi_0)}{m/2}\,.$$

Consequently, we can select $m = \xi_1 n - o(n)$, where ξ_1 defined in (29), so that for some $y\in S_r^n$:

$$|\{y+S_{\rho n}^n\}\cap \mathcal{C}'| > n\,.$$

Note that we focus on solution of (29) satisfying $\xi_1 < 2\xi_0(1-\xi_0)$. For some choices of R,δ and ξ_0 choosing $\xi_1 > 2\xi_0(1-\xi_0)$ is also possible, but such a choice appears to result in a weaker bound.

Next, we let $\mathcal{C}'' = \{y+S_{\rho n}^n\}\cap\mathcal{C}'$. For sufficiently large n the code $\mathcal{C}''$ will satisfy assumptions of Lemma 5 with $K\ge\frac{1}{\delta}$. Denote the resulting large symmetric subcode $\mathcal{C}'''$.

Note that because of (37) column-densities λ_i's of $\mathcal{C}'''$, defined in Lemma 7, satisfy (after possibly reordering coordinates):

$$\sum_{i=1}^{\xi_0 n}\lambda_i = \xi_1 n/2 + o(n), \qquad \sum_{i>\xi_0 n}\lambda_i = \xi_1 n/2 + o(n)\,.$$

Therefore, from Lemmas 7-8 we have

$$h_j(\bar{T}_L(\mathcal{C}''')) \le \xi_0 g_j\left(1-\frac{\xi_1}{2\xi_0}\right) + (1-\xi_0)g_j\left(\frac{\xi_1}{2(1-\xi_0)}\right) + \epsilon'_n + \frac{C_1}{|\mathcal{C}'''|}\,, \qquad (38)$$

where $\epsilon'_n \to 0$. Note that by construction the last term in (38) is $O(\delta)$. Also note that the first two terms in (38) equal θ_j defined in (27).

Finally, by Lemma 6 we get that for any codewords $c_1,\ldots,c_L \in \mathcal{C}'''$, some constant C and some sequence $\epsilon''_n \to 0$ the following holds:

$$\frac{1}{n}\operatorname{rad}(0,c_1,\ldots,c_L) \le \theta(\xi_0,R',L) + \epsilon''_n + C\delta\,.$$

By the initial remark, this concludes the proof of Lemma 4. □

Acknowledgement

We thank Prof. A. Barg for reading and commenting on an earlier draft.

References

[1] P. Elias, "List decoding for noisy channels," MIT, Cambridge, MA, Tech. Rep. RLE-TR-335, 1957.

[2] J. Wozencraft, "List decoding," MIT, Cambridge, MA, Tech. Rep. RLE Quart. Progr., vol. 48, 1958.

[3] Y. Kochman, A. Mazumdar, and Y. Polyanskiy, "The adversarial joint source-channel problem," in *Proc. 2012 IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, Jul. 2012.

[4] A. J. Young and Y. Polyanskiy, "Converse and duality results for combinatorial source-channel coding in binary Hamming spaces," in *Proc. 2015 IEEE Int. Symp. Inf. Theory (ISIT)*, Hong Kong, China, Jun. 2015. [Online]. Available: http://people.lids.mit.edu/yp/homepage/data/cjscc_convdual.pdf

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

[6] A. Ashikhmin, A. Barg, and S. Litsyn, "A new upper bound on codes decodable into size-2 lists," in *Numbers, Information and Complexity*. Springer, 2000, pp. 239–244.

[7] A. Samorodnitsky, "On the optimum of Delsarte's linear program," *J. Comb. Th., Ser. A*, vol. 96, pp. 261–287, 2001.

[8] V. Blinovsky, "Bounds for codes in the case of list decoding of finite volume," *Prob. Peredachi Inform.*, vol. 22, no. 1, pp. 7–19, 1986.

[9] ——, "Code bounds for multiple packings over a nonbinary finite alphabet," *Prob. Peredachi Inform.*, vol. 41, no. 1, pp. 23–32, 2005.

[10] V. Guruswami and S. Vadhan, "A lower bound on list size for list decoding," in *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*. Springer, 2005, pp. 318–329.

[11] G. Kalai and N. Linial, "On the distance distribution of codes," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1467–1472, 1995.

[12] M. E. H. Ismail and P. Simeonov, "Strong asymptotics for Krawtchouk polynomials," *J. Comp. and Appl. Math.*, vol. 100, pp. 121–144, 1998.

[13] Y. Polyanskiy, "Hypercontractivity of spherical averages in Hamming space," *Arxiv preprint arXiv:1309.3014*, 2013.

[14] A. Barg and A. McGregor, "Distance distribution of binary codes and the error probability of decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4237–4246, 2005.

[15] Y. Polyanskiy, "Upper bound on list-decoding radius of binary codes," *arXiv preprint arXiv:1409.7765*, 2014. [Online]. Available: http://people.lids.mit.edu/yp/homepage/data/listconv.pdf

[16] S. Litsyn, "New upper bounds on error exponents," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 385–398, 1999.