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]. Namely, in [4, Theorem 6] we proposed an extension of the previous result in [3, Theorem 7] that required bounding rate for the list-decoding problem.

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:

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

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

where $\beta \in [0, \frac{1}{2}]$.

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

$$R_2^*(\tau) \leq \begin{cases} R_{LP2}(2\tau), & \tau \leq \tau_0 \\ \log 2 - h(2\tau) + h(u(\tau)), & \tau > \tau_0, \end{cases} \tag{9}$$

where $\tau_0 \approx 0.1093$ and

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

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. This work was presented at 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, CN, Jun 2015.

[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].

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

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

$$\tau_L^*(R) \leq \sum_{i=1}^{\lceil L/2 \rceil} \frac{\binom{2i-2}{i-1}}{i} (\lambda(1-\lambda))^i \,, \qquad R = 1 - h(\lambda)\,, \tag{10}$$

where $\lambda \in [0, \frac{1}{2}]$. 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]:

$$E_\beta(\xi) = \xi \log(1-\omega) + (1-\xi)\log(1+\omega) - \beta \log \omega \tag{11}$$

$$\xi = \frac{1}{2}(1 - (1-\beta)\omega - \beta\omega^{-1})\,, \tag{12}$$

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

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \leq \max_{j,\xi_0} \xi_0 g_j\left(1 - \frac{\xi_1}{2\xi_0}\right) + (1-\xi_0) g_j\left(\frac{\xi_1}{2(1-\xi_0)}\right)\,, \tag{13}$$

where maximization is over ξ_0 satisfying

$$0 \leq \xi_0 \leq \frac{1}{2} - \sqrt{\beta(1-\beta)} \tag{14}$$

and j ranging 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. Quantity $\xi_1 = \xi_1(\xi_0, \delta, R)$ is a unique solution of

$$R + h(\beta) - 2E_\beta(\xi_0) = h(\xi_0) - \xi_0 h\left(\frac{\xi_1}{2\xi_0}\right) - (1-\xi_0) h\left(\frac{\xi_1}{2(1-\xi_0)}\right)\,, \tag{15}$$

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

$$g_j(\nu) \triangleq \frac{1}{L+j}\left(L\nu - \mathbb{E}\left[|2W - L - j|^+\right]\right)\,, W \sim \mathrm{Bino}(L, \nu) \tag{16}$$

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.

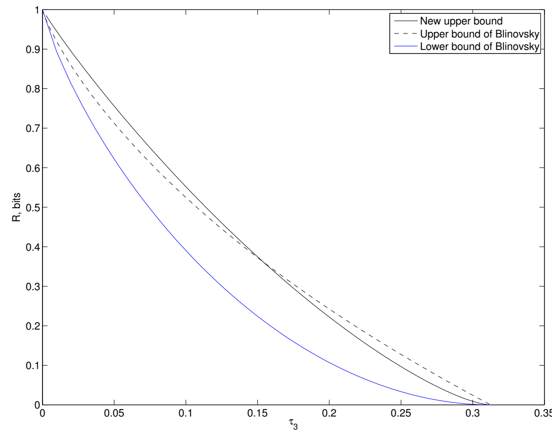

Fig. 1. Comparison of bounds on $R_L^*(\tau)$ for list size $L = 3$

TABLE I
RATES FOR WHICH NEW BOUND* IMPROVES STATE OF THE ART

List size L	Range of rates
$L = 3$	$0 < R \leq 0.361$
$L = 5$	$0 < R \leq 0.248$
$L = 7$	$0 < R \leq 0.184$
$L = 9$	$0 < R \leq 0.136$
$L = 11$	$0 < R \leq 0.100$

* This is computation of (13) with $h(\beta) = R$.

Evaluation of Theorem 1 is computationally possible, but is somewhat tedious. 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 $L = 3$:[3]

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

$$\tau_L^*(R) \leq \frac{3}{4}\delta - \frac{1}{16}\left(\frac{(2\delta - \xi_1)^3}{\delta^2} + \frac{\xi_1^3}{(1-\delta)^2}\right)\,, \tag{17}$$

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

$$R = h\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)\,, \tag{18}$$

$$R = \log 2 - \delta h\left(\frac{\xi_1}{2\delta}\right) - (1-\delta) h\left(\frac{\xi_1}{2(1-\delta)}\right) \tag{19}$$

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 (10) 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 Remark 2 in Section II-C):

[3]Notice that proofs of each of the two Corollaries below contain different relaxations of the bound (13), e.g. (22), which are easier to evaluate. Notice also that in Table I for the last two entries ($L = 9, 11$) at the high endpoint of rate the maximum over ξ_0 is attained *not* at $\frac{1}{2} - \sqrt{\beta(1-\beta)}$.

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*

$$\tau_L^*(R) \leq g_1(\delta_{LP1}(R)), \tag{20}$$

where $g_1(\cdot)$ is a degree-L polynomial defined in (16)*. In particular,*

$$\frac{d}{d\tau}\bigg|_{\tau=\tau_L^*(0)} R_L^*(\tau) = 0, \tag{21}$$

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

Before closing our discussion we make 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 (14), 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 [15, 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 (13) 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 (13) (via concavity of g_j, Lemma 8) to

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \leq \max_{j,\xi_0} g_j(\xi_0) = \max_j g_j(\delta_{LP1}(R)). \tag{22}$$

The $R < R_0(L)$ condition is only used to show that the maximum is attained at $j = 1$. Note also that weakening (22) corresponds to omitting the extra Elias-Bassalygo type reduction, which is responsible for the extra optimization over ξ_1 in (13).

Finally, at the invitation of anonymous reviewer we give our intuition about why our bound outperforms Blinovsky's for odd L. It is easiest to compare with the weakening (22) of our bound. Now compare the two proofs:

1) Blinovsky [8] first uses Elias-Bassalygo reduction to restrict attention to a subcode $\mathcal{C}'$ situated on a Hamming sphere of radius $\approx \delta_{GV}(R) = h^{-1}(1-R)$. Then he proves an upper bound for $\tau_L(\mathcal{C}')$ valid as long as $|\mathcal{C}'| \gg 1$ via a Plotkin-type argument together with a great symmetrization idea.
2) Our bound (following Ashikhmin, Barg and Litsyn [6]) instead uses a Kalai-Linial [11] reduction to select a subcode $\mathcal{C}''$ situated on a Hamming sphere of radius $\approx \delta_{LP1}(R)$. We then proceeded to prove a (Plotkin-type) upper bound on a strange quantity:

$$\tau_L^o(\mathcal{C}'') = \frac{1}{n}\left(\min\left\{\mathrm{rad}(\{0\}\cup S) : S \in \binom{\mathcal{C}}{L}\right\} - 1\right),$$

which corresponds to a requirement that the code contain not more than $L-1$ codewords in any ball of radius τ_L^o, but only for those balls that happen to also contain the origin.

Notice that the sphere returned by Kalai-Linial is bigger than that of Elias-Bassalygo (which is the reason our bound deteriorates at large rates), but the good thing is that the subcode $\mathcal{C}''$ has another codeword c_0 at the center of the Hamming sphere. Now, intuitively τ_L^o is roughly equivalent to τ_{L-1}. The zero-rate (Plotkin) radius for a list-L decoding of binary codes on Hamming sphere $S^n_{\xi n}$ is given by

$$p_L(\xi) = \frac{\mathbb{E}\left[\min(W_\xi, L+1-W_\xi)\right]}{L+1}, W_\xi \sim \mathrm{Bino}(L+1, \xi).$$

So intuitively, we expect that Blinovsky's bound should give

$$\tau_L^*(R) \lesssim p_L(\delta_{GV}(R))$$

while our bound should give

$$\tau_L^*(R) \lesssim p_{L-1}(\delta_{LP1}(R)).$$

Finally, it is easy to check that for even L we have $p_L = p_{L-1}$, while for odd L, $p_L > p_{L-1}$. This is the main intuitive reason why our bound succeeds in improving Blinovsky's, but only for odd L.

II. Proofs

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 (14) such that

$$A_{\xi_0 n}(\mathcal{C}) \triangleq \frac{1}{|\mathcal{C}|}\sum_{x,x'\in\mathcal{C}} 1\{|x-x'| = \xi_0 n\} \geq \exp\{n(R + h(\beta) - 2E_\beta(\xi_0) + \epsilon_n)\}. \tag{23}$$

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 (23) 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*

$$\frac{1}{n}\operatorname{rad}(0, c_1, \ldots, c_L) \le \theta(\xi_0, R', L) + \epsilon_n\,, \qquad (24)$$

where

$$\theta(\xi_0, R', L) \triangleq \max_j \theta_j(\xi_0, R', L) \qquad (25)$$

$$\theta_j(\xi_0, R', L) \triangleq \xi_0 g_j\left(1 - \frac{\xi_1}{2\xi_0}\right) + (1-\xi_0) g_j\left(\frac{\xi_1}{2(1-\xi_0)}\right), \qquad (26)$$

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

$$R' = h(\xi_0) - \xi_0 h\left(\frac{\xi_1}{2\xi_0}\right) - (1-\xi_0) h\left(\frac{\xi_1}{2(1-\xi_0)}\right), \qquad (27)$$

functions g_j are defined in (16) *and j in maximization* (25) *ranging over the same set as in Theorem 1.*

Equipped with Lemma 4 we immediately conclude that

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \le \max_{\xi_0 \in [0,\delta]} \theta(\xi_0, R + h(\beta) - 2E_\beta(\xi_0), L)\,. \qquad (28)$$

Clearly, (28) coincides with (13). 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

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

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*

$$\|T(S) - \bar{T}_L(\mathcal{C}')\| \le \delta\,. \qquad (29)$$

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. We first will show that for any $\delta_1 > 0$ and sufficiently large $|\mathcal{C}|$ we may select a subcode $\mathcal{C}'$ so that the following holds: For any pair of subsets $S, S' \subset \mathcal{C}'$ s.t. $|S| = |S'| \le L$ we have:

$$\|T(S) - T(S')\| \le \delta_1 \qquad (30)$$

Consider any code $\mathcal{C}_1 \subset \mathbb{F}_2^n$ and define a hypergraph with vertices indexed by elements of $\mathcal{C}$ and hyper-edges corresponding to each of the subsets of size L. Now define a $\delta_1/2$-net on the space $\mathcal{T}_L$ and label each edge according to the closest element of the $\delta_1/2$-net. By a theorem of Ramsey there exists K_L such that if $|\mathcal{C}_1| \ge K_L$ then there is a subset $\mathcal{C}'_1 \subset \mathcal{C}$ such that $|\mathcal{C}'_1| \ge K$ and each of the internal edges, indexed by $\binom{\mathcal{C}'_1}{L}$, is assigned the same label. Thus, by triangle inequality (30) follows for all $S, S' \in \binom{\mathcal{C}'_1}{L}$.

Next, apply the previous argument to show that there is a constant K_{L-1} such that for any $\mathcal{C}_2 \subset \mathbb{F}_2^n$ of size $|\mathcal{C}_2| \ge K_{L-1}$ there exists a subcode $\mathcal{C}'_2$ of size $|\mathcal{C}'_2| \ge K_L$ satisfying (30) for all $S, S' \in \binom{\mathcal{C}'_2}{L-1}$. Since $\mathcal{C}'_2$ satisfies the size assumption on $\mathcal{C}_1$ made in previous paragraph, we can select a further subcode $\mathcal{C}''_2 \subset \mathcal{C}'_2$ of size $\ge K_L$ so that for $\mathcal{C}''_2$ property (30) holds for all S, S' of size L or $L-1$.

Continuing similarly, we may select a subcode $\mathcal{C}'$ of arbitrary $\mathcal{C}$ such that (30) holds for all $|S| = |S'| \le L$ provided that $|\mathcal{C}| \ge K_1$.

Next, we show that (30) implies

$$\|T(S_0) - \sigma(T(S_0))\| \le C\delta_1\,, \qquad (31)$$

where $S_0 \in \binom{\mathcal{C}'}{L}$ is arbitrary and $C = C(L)$ is a constant depending on L only.

Now to prove (31) let $T(S_0) = \{t_v, v \in \mathbb{F}_2^L\}$ and consider an arbitrary transposition $\sigma : [L] \to [L]$. It will be clear that our proof does not depend on what transposition is chosen, so for simplicity we take $\sigma = \{(L-1) \leftrightarrow L\}$. We want to show that (30) implies

$$|t_v - t_{\sigma(v)}| \le \delta_1 \,. \qquad \forall v \in \mathbb{F}_2^L \tag{32}$$

Since transpositions generate permutation group S_L, (31) then follows. Notice that (32) is only informative for v whose last two digits are not equal, say $v = [v_0, 0, 1]$. Suppose that $S_0 = \{c_1, \ldots, c_L\}$ given in the ordered form. Let

$$S = \{c_1, \ldots c_{L-1}\}\,, \tag{33}$$

$$S' = \{c_1, \ldots, c_{L-2}, c_L\} \tag{34}$$

Joint types $T(S)$ and $T(S')$ are expressible as functions of $T(S_0)$ in particular, the number of occurrences of element $[v_0, 0]$ in S is $t_{[v_0,0,1]} + t_{[v_0,0,0]}$ and in S' is $t_{[v_0,0,0]} + t_{[v_0,1,0]}$. Thus, from (30) we obtain:

$$|(t_{[v_0,0,1]} + t_{[v_0,0,0]}) - (t_{[v_0,0,0]} + t_{[v_0,1,0]})| \le \delta$$

implying (32) and thus (31).

Finally, we show that (31) implies (29). Indeed, consider the chain

$$\|T(S) - \bar{T}_L(\mathcal{C}')\|$$

$$= \left\| T(S) - \frac{1}{L! \cdot \binom{|\mathcal{C}'|}{L}} \sum_{\sigma} \sum_{S' \in \binom{\mathcal{C}'}{L}} \sigma(T(S')) \right\| \tag{35}$$

$$\le \frac{1}{L! \cdot \binom{|\mathcal{C}'|}{L}} \sum_{\sigma} \sum_{S' \in \binom{\mathcal{C}'}{L}} \|T(S) - \sigma(T(S'))\| \tag{36}$$

$$\le \frac{1}{L! \cdot \binom{|\mathcal{C}'|}{L}} \sum_{\sigma} \sum_{S' \in \binom{\mathcal{C}'}{L}} \|T(S) - T(S')\|$$

$$+ \|T(S') - \sigma(T(S'))\| \tag{37}$$

$$\le (1 + C)\delta_1\,, \tag{38}$$

where (36) is by convexity of the norm, (37) is by triangle inequality and (38) is by (30) and (31). Consequently, setting $\delta_1 = \frac{\delta}{1+C}$ we have shown (29). □

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:

$$h_j(S) = \frac{1}{L+j} \left(\mathbb{E}\left[W\right] - \mathbb{E}\left[|2W - L - j|^+\right] \right) \tag{39}$$

$$\mathbb{P}[W = w] = \sum_{v: |v| = w} t_v\,, \tag{40}$$

where t_v are components of the joint-type $T(S) = \{t_v, v \in \mathbb{F}_2^L\}$. To check (39) 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 (39) we will abuse notation and write

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

We now observe that for symmetric codes satisfying (29) average-radii $h_j(S)$ in fact determine the regular radius:

Lemma 6. *Consider an arbitrary code $\mathcal{C}$ satisfying conclusion (29) of Lemma 5. Then for any subset $S = \{c_1, \ldots, c_L\} \subset \mathcal{C}$ we have*

$$\left| \mathrm{rad}(0, c_1, \ldots, c_L) - n \cdot \max_j h_j(\bar{T}_L(\mathcal{C})) \right| \le 2^L(1 + \delta n)\,, \tag{41}$$

where j in maximization (41) 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. For joint-types of size L and all $j \ge 0$ we clearly have (cf. expression (39))

$$|h_j(T_1) - h_j(T_2)| \le 2^{L-1}\|T_1 - T_2\|\,, \qquad \forall T_1, T_2 \in \mathcal{T}_L\,. \tag{42}$$

We also trivially have

$$\frac{1}{n} \mathrm{rad}(0, c_1, \ldots, c_L) \ge h_j(S) \qquad \forall j \ge 0\,. \tag{43}$$

Thus from (29) and (42) we already get

$$\frac{1}{n} \mathrm{rad}(0, c_1, \ldots, c_L) \ge \max_j h_j(\bar{T}_L(\mathcal{C})) - 2^{L-1}\delta\,.$$

It remains to show

$$\frac{1}{n} \mathrm{rad}(0, c_1, \ldots, c_L) \le \max_j h_j(\bar{T}_L(\mathcal{C})) + \delta + \frac{2^L}{n}\,. \tag{44}$$

This evidently requires constructing a good center y for the set $\{0, c_1, \ldots, c_L\}$. To that end fix arbitrary numbers $q = (q_0, \ldots, q_L) \in [0, 1]^L$. Next, for each $v \in \mathbb{F}_2^L$ let $E_v \subset [n]$ be all coordinates on which restriction of $\{c_1, \ldots, c_L\}$ equals v. On E_v put y to have a fraction $q_{|v|}$ of ones and remaining set to zeros (rounding to integers arbitrarily). Proceed for all $v \in \mathbb{F}_2^L$. Call resulting vector $y(q) \in \mathbb{F}_2^n$.

Denote for convenience $c_0 = 0$. We clearly have

$$\mathrm{rad}(c_0, c_1, \ldots, c_L) \le \min_q \max_p \sum_{i=0}^L p_i |c_i - y(q)|\,, \tag{45}$$

where $p = (p_0, \ldots, p_L)$ is a probability distribution.

Denote

$$T(S) = \{t_v, v \in \mathbb{F}_2^L\} \tag{46}$$

$$\bar{T}_L(\mathcal{C}) = \{\bar{t}_v, v \in \mathbb{F}_2^L\} \tag{47}$$

We proceed to computing $|c_i - y(q)|$.

$$|c_i - y(q)| \le n \sum_{v\in\mathbb{F}_2^L} t_v(q_{|v|}1\{v(i)=0\} + (1-q_{|v|})1\{v(i)=1\}) + 2^L\,, \quad (48)$$

where 2^L comes upper-bounding the integer rounding issues and we abuse notation slightly by setting $v(0)=0$ for all v (recall that $v(i)$ is the i-th coordinate of $v\in\mathbb{F}_2^L$).

By (29) we may replace t_v with $\bar t_v$ at the expense of introducing $2^L\delta n$ error, so we have:

$$|c_i - y(q)| \le n \sum_{v\in\mathbb{F}_2^L} \bar t_v(q_{|v|}1\{v(i)=0\} + (1-q_{|v|})1\{v(i)=1\}) + 2^L(1+\delta n)\,. \quad (49)$$

Next notice that the sum over v only depends on whether $i=0$ or $i\neq 0$ (by symmetry of $\bar t_v$). Furthermore, for any given weight w and $i\neq 0$ we have

$$\sum_{v:|v|=w} 1\{v(i)=1\} = \binom{L}{w}\frac{w}{L}\,.$$

Thus, introducing the random variable $\bar W$, cf. (39),

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

we can rewrite:

$$\sum_{v\in\mathbb{F}_2^L} \bar t_v(q_{|v|}1\{v(i)=0\} + (1-q_{|v|})1\{v(i)=1\}) = \frac{1}{L}\mathbb{E}\,[\bar W + (L-2\bar W)q_{\bar W}]\,. \quad (50)$$

For $i=0$ the expression is even simpler:

$$\sum_{v\in\mathbb{F}_2^L} \bar t_v(q_{|v|}1\{v(0)=0\} + (1-q_{|v|})1\{v(0)=1\}) = \mathbb{E}\,[q_{\bar W}]\,.$$

Substituting derived upper bound on $|c_i - y(q)|$ into (45) we can see that without loss of generality we may assume $p_1 = \cdots = p_L$, so our upper bound (modulo $O(\delta)$ terms) becomes:

$$\min_q \max_{p_1\in[0,L^{-1}]} (1-Lp_1)\mathbb{E}\,[q_{\bar W}] + p_1\mathbb{E}\,[\bar W + (L-2\bar W)q_{\bar W}]$$
$$= \min_q \max_{p_1\in[0,L^{-1}]} p_1\mathbb{E}\,[\bar W] + \mathbb{E}\,[q_{\bar W}(1-2\bar W p_1)]$$

By von Neumann's minimax theorem we may interchange min and max, thus continuing as follows:

$$= \max_{p_1\in[0,L^{-1}]} \min_q p_1\mathbb{E}\,[\bar W] + \mathbb{E}\,[q_{\bar W}(1-2\bar W p_1)] \quad (51)$$
$$= \max_{p_1\in[0,L^{-1}]} p_1\mathbb{E}\,[\bar W] - \mathbb{E}\,[|2\bar W p_1 - 1|^+]\,. \quad (52)$$

The optimized function of p_1 is piecewise-linear, so optimization can be reduced to comparing values at slope-discontinuities and boundaries. The point $p_1 = 0$ is easily excluded, while the rest of the points are given by $p_1 = \frac{1}{L+j}$ with j ranging over the set specified in the statement of Lemma[4]. So we continue (52) getting

$$= \max_j \frac{1}{L+j}\left(\mathbb{E}\,[\bar W] - \mathbb{E}\,[|2\bar W - L - j|^+]\right) \quad (53)$$

We can see that expression under maximization is exactly $h_j(\bar T_L(\mathcal{C}))$ and hence (44) is proved. □

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*[5]*, namely:*

$$\left\|\bar T_L(\mathcal{C}) - \frac{1}{n}\sum_{i=1}^n \mathrm{Bern}^{\otimes L}(\lambda_i)\right\| \le \frac{C_1}{|\mathcal{C}|}\,, \quad (54)$$

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,

$$\left|h_j(\bar T_L(\mathcal{C})) - \frac{1}{n}\sum_j g_j(\lambda_j)\right| \le \frac{C_2}{|\mathcal{C}|}\,, \quad (55)$$

where functions g_j were defined in (16)*.*

Proof. Second statement (55) follows from the first via (42) and linearity of $h_j(T)$ in the type T, cf. (39). To show the first statement, let $M = |\mathcal{C}|$, $M_i = \lambda_i M$ and p_w – total probability assigned to vectors v of weight w by $\bar T_L(\mathcal{C})$. Then by computing p_w over columns of $M\times n$ matrix we obtain

$$p_w = \frac{1}{n}\sum_{i=1}^n \frac{\binom{M_i}{w}\binom{M-M_i}{L-w}}{\binom{M}{L}}\,.$$

By a standard estimate we have for all $w = \{0,\dots,L\}$:

$$\frac{\binom{M_i}{w}\binom{M-M_i}{L-w}}{\binom{M}{L}} = \binom{L}{w}\lambda_i^w(1-\lambda_i)^{L-w} + O(\frac{1}{M})\,,$$

with $O(\cdot)$ term uniform in w and λ_i. By symmetry of the type $\bar T_L(\mathcal{C})$ the result (54) follows. □

Lemma 8. *Functions g_j defined in* (16) *are concave on* $[0,1]$.

Proof. Let $W_\lambda \sim \mathrm{Bino}(L,\lambda)$ and $V_\lambda \sim \mathrm{Bino}(L-1,\lambda)$. Denote for convenience $\bar\lambda = 1-\lambda$ and take j_0 to be an integer

[4] The difference between odd and even L occurs due to the boundary point $p_1 = \frac{1}{L}$ not being a slope-discontinuity when L is odd, so we needed to add it separately.

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

between 0 and L. We have then

$$\frac{\partial}{\partial\lambda}\mathbb{E}\left[|W_\lambda - j_0|^+\right]$$
$$= \sum_{w=j_0+1}^{L}\binom{L}{w}(w-j_0)\lambda^w\bar\lambda^{L-w}\left\{w\lambda^{-1}-(L-w)\bar\lambda^{-1}\right\} \tag{56}$$
$$= \binom{L}{j_0+1}(j_0+1)\lambda^{j_0}\bar\lambda^{L-j_0-1}$$
$$+\sum_{w=j_0+1}^{L-1}\left[\binom{L}{w+1}(w+1-j_0)(w+1)\right.$$
$$\left.-\binom{L}{w}(w-j_0)(L-w)\right]\lambda^w\bar\lambda^{L-w-1} \tag{57}$$
$$= L\binom{L-1}{j_0}\lambda^{j_0}\bar\lambda^{L-1-j_0}+L\sum_{w=j_0+1}^{L-1}\binom{L-1}{w}\lambda^w\bar\lambda^{L-1-w} \tag{58}$$
$$= L\mathbb{P}[V_\lambda \ge j_0]\,, \tag{59}$$

where in (57) we shifted the summation by one for the first term under the sum in (56), and in (58) applied identities $\binom{L}{w+1} = \binom{L}{w}\frac{L-w}{w+1} = \binom{L-1}{w}\frac{L}{w+1}$. Similarly, if $\theta \in [0,1)$ we have

$$\frac{\partial}{\partial\lambda}\mathbb{E}\left[|W_\lambda - j_0-\theta|^+\right] = L\mathbb{P}[V_\lambda \ge j_0+1]+L(1-\theta)\mathbb{P}[V_\lambda = j_0]\,. \tag{60}$$

Similarly, one shows (we will need it later in Lemma 9):

$$\frac{\partial}{\partial\lambda}\mathbb{P}[W_\lambda \ge j_0] = L\mathbb{P}[V_\lambda = j_0-1]\,. \tag{61}$$

Since clearly the function in (60) is strictly increasing in λ for any j_0 and θ we conclude that

$$\lambda \mapsto \mathbb{E}\left[|W_\lambda - j_0-\theta|^+\right]$$

is convex. This concludes the proof of concavity of g_j. □

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 (24) 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:

$$\{y+S_m^n\}\cap S_r^n \cong S^r_{r-m/2}\times S^{n-r}_{m/2}\,. \tag{62}$$

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 (27), so that for some $y \in S_r^n$:

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

Note that we focus on solution of (27) 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 $\mathcal{C}'' = \{y+S^n_{\rho n}\}\cap\mathcal{C}'$. 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 (62) 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

$$h_j(\bar T_L(\mathcal{C}''')) \le \xi_0 g_j\left(1-\frac{\xi_1}{2\xi_0}\right) + (1-\xi_0)g_j\left(\frac{\xi_1}{2(1-\xi_0)}\right)+\epsilon'_n+\frac{C_1}{|\mathcal{C}'''|}\,, \tag{63}$$

where $\epsilon'_n \to 0$. Note that by construction the last term in (63) is $O(\delta)$. Also note that the first two terms in (63) equal θ_j defined in (25).

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

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

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

C. Proof of Corollary 3

Lemma 9. *For any odd $L = 2a+1$ there exists a neighborhood of $x = \frac{1}{2}$ such that*

$$\max_j g_j(x) = g_1(x)\,, \tag{64}$$

maximum taken over j equal all the odd numbers not exceeding L and $j = 0$. We also have for some $c > 0$

$$g_1(x) = \frac{1}{2} - 2^{-L-1}\binom{L}{\frac{L-1}{2}}+cx+O((2x-1)^2), \qquad x \to \frac{1}{2}\,. \tag{65}$$

Proof. First, the value $g_1(1/2)$ is computed trivially. Then from (60) we have

$$\frac{d}{dx}g_j(x) = \frac{L}{L+j}\left(1-2\mathbb{P}\left[V_x \ge \frac{L+j}{2}\right]\right)\,, \tag{66}$$

where $j \ge 1$ and $V_x \sim \mathrm{Bino}(x, L-1)$. This implies (65). For future reference we note that (69) (below) and (61) imply

$$\frac{d}{dx}g_0(x) = 1-2\mathbb{P}[V_x \ge \frac{L+1}{2}] - \mathbb{P}[V_x = \frac{L-1}{2}], \qquad V_x \sim \mathrm{Bino}(x, L-1)\,. \tag{67}$$

By continuity, (64) follows from showing

$$g_1(1/2) > \max_{j\in\{0,3,5,\dots L\}} g_j(1/2)\,. \tag{68}$$

Next, consider $W_x \sim \mathrm{Bino}(x, L)$ and notice the upper-bound

$$g_j(x) \le \frac{1}{L+j}\mathbb{E}\,[W_x 1\{W_x \le a\} + (L+j-W_x)1\{W_x \ge a+1\}]\,.$$

Then, substituting expression for $g_1(x)$ we get

$$g_1(x) - g_0(x) = \frac{1}{L}\left(\mathbb{P}[W_x \ge a+1] - g_1(x)\right) \tag{69}$$

$$g_1(x) - g_j(x) \ge \frac{j-1}{L+j}\left(g_1(x) - \mathbb{P}[W_x > a+1]\right)\,. \tag{70}$$

Thus, to show (68) it is sufficient to prove that for $x = 1/2$ we have

$$\mathbb{P}[W_{\frac12} > a+1] < g_1(1/2) < \mathbb{P}[W_{\frac12} \ge a+1]\,. \tag{71}$$

The right-hand inequality is trivial since $\mathbb{P}[W_{\frac12} \ge a+1] = 1/2$ while from (65) we know $g_1(1/2) < 1/2$. The left-hand inequality, after simple algebra, reduces to showing

$$\sum_{u=0}^{a-1}(2a+1-2u)\binom{2a+1}{u} < (2a+1)\binom{2a+1}{a}\,. \tag{72}$$

Notice, that

$$(n-2u)\binom{n}{u} = n\left[\binom{n-1}{u} - \binom{n-1}{u-1}\right]\,\forall u \ge 0$$

and therefore

$$\sum_{u\le \ell}(n-2u)\binom{n}{u} = n\binom{n-1}{\ell}\,.$$

Plugging this identity into the right-hand side of (72) we get

$$\sum_{u=0}^{a-1}(2a+1-2u)\binom{2a+1}{u} = (2a+1)\binom{2a}{a-1}$$

$$< (2a+1)\binom{2a}{a} < (2a+1)\binom{2a+1}{a} \tag{73}$$

completing the proof of (72). ☐

Proof of Corollary 3. We first show that (20) implies (21). To that end, fix a small $\epsilon >$ so that $\frac12 - \epsilon$ belongs to the neighborhood existence of which is claimed in Lemma 9. Choose rate so that $\delta_{LP1}(R) = 1/2 - \epsilon$ and notice that this implies

$$R = h(\epsilon^2 + o(\epsilon^2))\,, \tag{74}$$

By Lemma 9, the right-hand side of (20) is

$$\tau_L^*(0) - \mathrm{const}\cdot\epsilon + o(\epsilon)\,,$$

which together with (74) implies (21).

To prove (20) we use Theorem 1 with $\delta = \delta_{LP1}(R)$. Next, use concavity of g_j's (Lemma 8) to relax (13) to

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \le \max_{j,\xi_0} g_j(\xi_0)\,.$$

From (66) and (67) it is clear that $\xi_0 \mapsto g_j(\xi_0)$ is monotonically increasing for all $j \ge 0$ on the interval $[0, 1/2]$. Thus, we further have

$$\limsup_{n\to\infty} \tau_L(\mathcal{C}_n) \le \max_j g_j(\delta_{LP1}(R))\,. \tag{75}$$

Bound (75) is valid for all $R \in [0,1]$ and arbitrary (odd/even L). However, when R is small (say, $R < R_0$) and L is odd, $\delta_{LP1}(R)$ belongs to the neighborhood of $1/2$ in Lemma 9 and thus (20) follows from (75) and (64). ☐

Remark 2. *It is, perhaps, instructive to explain why Corollary 3 cannot be shown for even L (via Theorem 1). For even L the maximum over j of $g_j(1/2-\epsilon)$ is attained at $j = 0$ and*

$$g_0(\frac12 - \epsilon) = \tau_L^*(0) + c\epsilon^2 + O(\epsilon^3)\,, \epsilon \to 0 \tag{76}$$

Therefore, for $\delta_{LP1}(R) = \frac12 - \epsilon$ we get from (76) *that the right-hand side of* (75) *evaluates to*

$$\tau_L^*(0) - \mathrm{const}\cdot\epsilon^2\log\frac{1}{\epsilon}\,. \tag{77}$$

Thus, comparing (77) *with* (74) *we conclude that for even L our bound on $R_L^*(\tau)$ has negative slope at zero rate. Note that Blinovsky's bound* (10) *has negative slope at zero rate for both odd and even L.*

D. Proof of Corollary 2

Proof. Instead of working with parameter δ we introduce $\beta \in [0, 1/2]$ such that

$$\delta = \frac12 - \sqrt{\beta(1-\beta)}\,.$$

We then apply Theorem 1 with $h(\beta) = R$. Notice that the bound on ξ_0 in (14) becomes

$$0 \le \xi_0 \le \delta\,.$$

By a simple substitution $\omega = \sqrt{\frac{\beta}{1-\beta}}$ we get from (11)

$$E_\beta(\delta) = \frac12(\log 2 - h(\delta) + h(\beta))\,.$$

Therefore, when $\xi_0 = \delta$ we notice that

$$R + h(\beta) - 2E_\beta(\xi_0) = R - \log 2 + h(\delta)$$

implying that defining equation for ξ_1, i.e. (15), coincides with (19).

Next for $L = 3$ we compute

$$g_0(\nu) = \nu(1-\nu)\,, \tag{78}$$

$$g_1(\nu) = \frac34\nu - \frac12\nu^3\,, \tag{79}$$

$$g_3(\nu) = \frac12\nu\,. \tag{80}$$

Note that the right-hand side of (17) is precisely equal to

$$\delta g_1\left(1 - \frac{\xi_1}{2\delta}\right) + (1-\delta)g_1\left(\frac{\xi_1}{2(1-\delta)}\right)\,.$$

So this corollary simply states that for $L = 3$ the maximum in (13) is achieved at $j = 1, \xi_0 = \delta$. Let us restate this last statement rigorously: The maximum

$$\max_{j\in\{0,1,3\}} \max_{\xi_0\in\delta} \xi_0 g_j\left(1 - \frac{x}{2\xi_0}\right) + (1-\xi_0) g_j\left(\frac{x}{2(1-\xi_0)}\right) \tag{81}$$

is achieved at $j = 1, \xi_0 = \delta$. Here $x = x(\xi_0, \beta)$ is a solution of

$$\begin{aligned} &2(h(\beta) - E_\beta(\xi_0)) \\ &\quad = h(\xi_0) - \xi_0 h\left(\frac{x}{2\xi_0}\right) - (1-\xi_0) h\left(\frac{x}{2(1-\xi_0)}\right). \end{aligned} \tag{82}$$

For notational convenience we will denote the function under maximization in (81) by $g_j(\xi_0, x)$.

We proceed in two steps:

- First, we estimate the maximum over ξ_0 for $j = 0$ as follows:

$$\begin{aligned} \max_{\xi_0} g_0(\xi_0, x) \le \frac{\log 2 - R}{4\log 2} \cdot \left(1 - \frac{1-\delta}{a_{max}(1-a_{max})}\right) \\ + (1-\delta) g_0(a_{min}), \end{aligned} \tag{83}$$

  where $a_{max}, a_{min} \le \frac{1}{2}$ are given by

$$a_{max} = h^{-1}(\log 2 - R), \tag{84}$$

$$a_{min} = h^{-1}\left(\log 2 - \frac{R}{1-\delta}\right). \tag{85}$$

- Second, we prove that for $j = 1$ function

$$\xi_0 \mapsto g_j(\xi_0, x(\xi_0))$$

  is monotonically increasing.

Once these two steps are shown, it is easy to verify (for example, numerically) that $g_1(\delta, x(\delta))$ exceeds both $\frac{1}{2}\delta$ (term corresponding to $j = 3$ in (81)) and the right-hand side of (83) (term corresponding to $j = 0$). Notice that this relation holds for all rates. Therefore, maximum in (81) is indeed attained at $j = 1, \xi_0 = \delta$.

One trick that will be common to both steps is the following. From the proof of Lemma 4 it is clear that the estimate (24) is monotonic in R'. Therefore, in equation (82) we may replace $E_\beta(\xi)$ with any upper-bound of it. We will use the well-known upper-bound, which leads to binomial estimates of spectrum components [15, (46)]:

$$E_\beta(\xi_0) \le \frac{1}{2}(\log 2 + h(\beta) - h(\xi_0)). \tag{86}$$

Furthermore, it can also be argued that maximum cannot be attained by ξ_0 so small that

$$h(\beta) - \frac{1}{2}(\log 2 + h(\beta) - h(\xi_0)) < 0.$$

So from now on, we assume that

$$h^{-1}(\log 2 - h(\beta)) \le \xi_0 \le \delta,$$

and that $x = x(\xi_0) \le 2\xi_0(1-\xi_0)$ is determined from the equation:

$$\log 2 - R = \xi_0 h\left(\frac{x}{2\xi_0}\right) + (1-\xi_0) h\left(\frac{x}{2(1-\xi_0)}\right) \tag{87}$$

(we remind $R = h(\beta)$).

We proceed to demonstrating (83). For convenience, we introduce

$$a_1 \triangleq 1 - \frac{x}{2\xi_0}, \tag{88}$$

$$a_2 \triangleq \frac{x}{2 - 2\xi_0}. \tag{89}$$

By constraints on x it is easy to see that

$$0 \le a_2 \le \min(a_1, 1 - a_1).$$

Therefore, we have

$$\log 2 - R = \xi_0 h(a_1) + (1-\xi_0) h(a_2) \ge h(a_2)$$

and thus $a_2 \le a_{max}$ defined in (84). Similarly, we have

$$\log 2 - R = \xi_0 h(a_1) + (1-\xi_0) h(a_2) \le \xi_0 \log 2 + (1-\xi_0) h(a_2),$$

and since $\xi_0 \le \delta$ we get that $a_2 \ge a_{min}$ defined in (85).

Next, notice that $\frac{h(x)}{x(1-x)}$ is decreasing on $(0, 1/2]$. Thus, we have

$$h(a_1) \ge g_0(a_1) 4\log 2 \tag{90}$$

$$\begin{aligned} h(a_2) &\ge h(a_{max}) \frac{g_0(a_2)}{g_0(a_{max})} \\ &= \frac{\log 2 - R}{a_{max}(1-a_{max})} g_0(a_2) \triangleq c \cdot g_0(a_2), \end{aligned} \tag{91}$$

where in the last step we introduced $c > 4\log 2$ for convenience. Consequently, we get

$$\begin{aligned} &\log 2 - R \\ &= \xi_0 h(a_1) + (1-\xi_0) h(a_2) && (92) \\ &\ge 4\log 2 \cdot \xi_0 g_0(a_1) + (1-\xi_0) c \cdot g_0(a_2) && (93) \\ &= 4\log 2 \cdot g_0(\xi_0, x) + (1-\xi_0)(c - 4\log 2) \cdot g_0(a_2) && (94) \\ &\ge 4\log 2 \cdot g_0(\xi_0, x) + (1-\delta)(c - 4\log 2) \cdot g_0(a_{min}). && (95) \end{aligned}$$

Rearranging terms yield (83).

We proceed to proving monotonicity of (82). The technique we will use is general (can be applied to $L > 3$ and $j > 1$), so we will avoid particulars of $L = 3, j = 1$ case until the final step.

Notice that regardless of the function $g(\nu)$ we have the equivalence:

$$\begin{aligned} &\frac{d}{d\xi_0} \xi_0 g(a_1) + (1-\xi_0) g(a_2) \ge 0 \qquad \iff \\ &\frac{1}{2}\frac{dx}{d\xi_0}(g'(a_2) - g'(a_1)) \ge \int_{a_2}^{a_1} (1-x)(-g''(x))dx - g'(a_2), \end{aligned} \tag{96}$$

where we recall definition of a_1, a_2 in (88)-(89). Differentiating (87) in ξ_0 (and recalling that R is fixed, while $x = x(\xi_0)$ is an implicit function of ξ_0) we find

$$\frac{dx}{d\xi_0} = -2 \frac{\log \frac{1-a_2}{a_1}}{\log \frac{1-a_2}{a_2} \frac{a_1}{1-a_1}} < 0.$$

Next, one can notice that the map $(\xi_0, x, R) \mapsto (a_1, a_2)$ is a bijection onto the region

$$\{(a_1, a_2) : 0 \le a_1 \le 1, 0 \le a_2 \le a_1(1-a_1)\}. \tag{97}$$

With the inverse map given by

$$\xi_0 = \frac{a_2}{1-a_1+a_2}, x = \frac{2a_2^2}{1-a_1+a_2}, \qquad R = \log 2 - \xi_0 h(a_1) - (1-\xi_0)h(a_2)\,.$$

Thus, verifying (96) can as well be done for all a_1, a_2 inside the region (97). Substituting $g = g_1$ into (96) we get that monotonicity in (82) is equivalent to a two-dimensional inequality:

$$-2\log\frac{1-a_2}{a_1}\cdot(a_1^2-a_2^2) \geq (2a_1^2 - \frac{4}{3}(a_1^3-a_2^3)-1)\log\frac{1-a_2}{a-2}\frac{a_1}{1-a_1}\,. \quad (98)$$

It is possible to verify numerically that indeed (98) holds on the set (97). For example, one may first demonstrate that it is sufficient to restrict to $a_2 = 0$ and then verify a corresponding inequality in a_1 only. We omit mechanical details. □

Acknowledgement

We thank Prof. A. Barg for reading and commenting on an earlier draft and anonymous reviewers for pointing out a mistake in the previous version of Table I and for simplifying proof of (72).

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.

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

Yury Polyanskiy (S'08-M'10-SM'14) is an Associate Professor of Electrical Engineering and Computer Science and a member of LIDS at MIT. Yury received M.S. degree in applied mathematics and physics from the Moscow Institute of Physics and Technology, Moscow, Russia in 2005 and Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ in 2010. In 2000-2005 he lead the development of the embedded software in the Department of Oilfield Surface Equipment, Borets Company LLC (Moscow). Currently, his research focuses on basic questions in information theory, error-correcting codes, wireless communication and fault-tolerant and defect-tolerant circuits. Dr. Polyanskiy won the 2013 NSF CAREER award and 2011 IEEE Information Theory Society Paper Award.