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. Resulting bound improves the best known bound (due to Blinovsky’1986) for rates below a certain threshold. 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.

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.

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

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

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

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

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

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

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

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

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

The asymptotic tradeoff between rate and list-decoding radius $\tau_L$ is defined as usual:

$$\tau^*_L(R) \triangleq \limsup_{n \to \infty} \max_{C : |C| \geq 2^n R} \tau_L(C)$$

$$R^*_L(\tau) \triangleq \limsup_{n \to \infty} \max_{C : \tau_L(C) \geq \tau} \frac{1}{n} \log |C|$$

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 [4]:

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

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

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:

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

(7)

$$
\delta_{LP1}(R) \triangleq \frac{1}{2} - \sqrt{\beta(1-\beta)}, \quad R = \log 2 - h(\beta), \quad \beta \in [0, 1/2]
$$

(8)

$$
\tau_2^*(\tau) \leq \log 2 - h(2\tau) + R_{up}(2\tau, 2\tau),
$$

where $R_{up}(\delta, \alpha)$ is the best known upper bound on rate of codes with minimal distance $\delta n$ 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 [4] and applying Levenshtein’s monotonicity, cf. [6, Lemma 4.2(6)]. The resulting expression is

$$
\tau_2^*(\tau) \leq \begin{cases} 
R_{LP2}(2\tau), & \tau \leq \tau_0 \\
\log 2 - h(2\tau) + h(u(\tau)), & \tau > \tau_0,
\end{cases}
$$

(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. [6, (9)]).

- For list sizes $L \geq 3$: The original bound of Blinovsky [7] appears to be the best (before this work):

$$
\tau_L^*(R) \leq \sum_{i=1}^{[L/2]} \frac{(2i-2)!(i-1)!}{i!} (\lambda(1-\lambda))^i, \quad R = 1 - h(\lambda), \lambda \in [0, 1/2]
$$

(11)

Note that [7] also gives a non-constructive lower bound on $\tau_L^*(R)$. Results on list-decoding over non-binary alphabets are also known, see [8], [9].

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).

---

1This result follows from (fixing typos and) optimizing [5, Theorem 4]. It is slightly stronger than what is given in [5, Corollary 5].
To present our main result, we need to define exponent of Krawtchouk polynomial $K_{\beta n}(\xi n) = \exp\{nE_\beta(\xi) + o(n)\}$. We have the following parametric expression [10] (see also [11, Lemma 4]):

$$E_\beta(\xi) = \xi \log |1 - \omega| + (1 - \xi) \log |1 + \omega| - \beta \log |\omega|$$  \hspace{1cm} (12)

$$\xi = \frac{1}{2} (1 - (1 - \beta) \omega - \beta \omega^{-1})$$  \hspace{1cm} (13)

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 $C_n \subset \{0,1\}^n$ of rate $R$ satisfies

$$\limsup_{n \to \infty} \tau_L(C_n) \leq \max_{j,\xi_0} \xi_0 \xi_1 \left(1 - \frac{\xi_1}{2\xi_0}\right) + (1 - \xi_0) \xi_1 \left(1 - \frac{\xi_1}{2(1 - \xi_0)}\right),$$  \hspace{1cm} (14)

where maximization is over $\xi_0$ satisfying

$$0 \leq \xi_0 \leq \frac{1}{2} - \sqrt{\beta(1 - \beta)}$$  \hspace{1cm} (15)

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

$$R + h(\beta) - 2E_\beta(\xi_0) = h(\xi_0) - \xi_0 h \left(\frac{\xi_1}{2\xi_0}\right) - (1 - \xi_0) h \left(\frac{\xi_1}{2(1 - \xi_0)}\right),$$  \hspace{1cm} (16)

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

$$g_j(\nu) \triangleq \frac{1}{L + j} \left(L\nu - \mathbb{E}[|2W - L - j|^+]\right), \hspace{1cm} W \sim \text{Bino}(L, \nu)$$  \hspace{1cm} (17)

As usual with bounds of this type, cf. [12], 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. Fortunately, for small $L$ the maximum over $\xi_0$ 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

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

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

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

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

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 non-zero 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):

**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

$$\tau^*_L(R) \leq g_1(\delta_{LP1}(R)).$$

Notice that proofs of each of the two Corollaries below contain a different relaxation of the bound (14), which may appear useful separately.
TABLE I

RATES FOR WHICH NEW BOUND IMPROVES STATE OF THE ART

<table>
<thead>
<tr>
<th>List size $L$</th>
<th>Range of rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L = 3$</td>
<td>$0 &lt; R \leq 0.361$</td>
</tr>
<tr>
<td>$L = 5$</td>
<td>$0 &lt; R \leq 0.248$</td>
</tr>
<tr>
<td>$L = 7$</td>
<td>$0 &lt; R \leq 0.184$</td>
</tr>
<tr>
<td>$L = 9$</td>
<td>$0 &lt; R \leq 0.144$</td>
</tr>
<tr>
<td>$L = 11$</td>
<td>$0 &lt; R \leq 0.108$</td>
</tr>
</tbody>
</table>

In particular,

$$\left. \frac{d}{d\tau} \right|_{\tau = \tau^*_L(0)} R^*_L(\tau) = 0,$$

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

We close our discussion with some additional remarks on Theorem 1:

1) The bound in Theorem 1 can be slightly improved by replacing $\delta_{LP1}(R)$, that appears in the right-hand side of (15), with a better bound, a so-called second linear-programming bound $\delta_{LP2}(R)$ from [4]. This would enforce the usage of the more advanced estimate of Litsyn [13, 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 (14) 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.

II. PROOFS

A. Proof of Theorem 1

Consider an arbitrary sequence of codes $C_n$ of rate $R$. As in [5] 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 [14, Proposition 3.2]: For every $\beta$ with $h(\beta) \leq R$ there exists a sequence $\epsilon_n \to 0$ such that for every code $C$ of rate $R$ there is a $\xi_0$ satisfying (15) such that

$$A_{\xi_0n}(C) \triangleq \frac{1}{|C|} \sum_{x,x' \in C} 1\{|x - x'| = \xi_0n\} \geq \exp\{n(R + h(\beta) - 2E_\beta(\xi_0) + \epsilon_n)\}. \quad (23)$$

Without loss of generality (by compactness of the interval $[0, 1/2 - \sqrt{\beta(1 - \beta)}]$ and passing to a proper subsequence of codes $C_{n_k}$) we may assume that $\xi_0$ selected in (23) is the same for all blocklengths $n$. Then there is a sequence of subcodes $C'_n$ of asymptotic rate $R' \geq R + h(\beta) - 2E_\beta(\xi_0)$ such that each $C'_n$ is situated on a sphere $c_0 + S_{\xi_0}$ surrounding another codeword $c_0 \in C$. 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 $C'_n \subset S_{\xi_0n}$ of rate $R' > 0$ there exist $L$ codewords $c_1, \ldots, c_L \in C'_n$ such that

$$\frac{1}{n} \text{rad}(0, c_1, \ldots, c_L) \leq \theta(\xi_0, R', L) + \epsilon_n, \quad (24)$$

where

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

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

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

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

functions $g_j$ are defined in (17) and $j$ in maximization (25) ranging over the same set as in Theorem 1.

Equipped with Lemma 4 we immediately conclude that

$$\limsup_{n \to \infty} \tau_L(C_n) \leq \max_{\xi_0 \in [0, \delta]} \theta(\xi_0, R + h(\beta) - 2E_\beta(\xi_0), L). \quad (28)$$

Clearly, (28) coincides with (14). So it suffices to prove 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

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

We define distance on $\mathcal{T}_L$ to be the $L_\infty$ 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

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

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

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

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 $C \subset \mathbb{F}_2^n$ we define its average joint type:

$$\bar{T}_L(C) = \frac{1}{L! \cdot \binom{|C|}{L}} \sum_{\sigma} \sum_{S \subset\subset_{\ell} C} \sigma(T(S)).$$

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

Our proof crucially depends on a (slight extension of the) brilliant idea of Blinovsky [7]:
Lemma 5: For every \( L \geq 1, K \geq L \) and \( \delta > 0 \) there exist a constant \( K_1 = K_1(L, K, \delta) \) such that for all \( n \geq 1 \) and all codes \( C \subset \mathbb{F}_2^n \) of size \( |C| \geq K_1 \) there exists a subcode \( C' \subset C \) of size at least \( K \) and a symmetric \( T_0 \in T_L \) such that for any \( S \in \binom{C'}{L} \) we have

\[
\|T(S) - T_L(C')\| \leq \delta.
\]  

(29)

Remark 1: Note that if \( S' \subset S \) then every element of \( T(S') \) is a sum of \( \leq 2L \) 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 \( |C| \) we may select a subcode \( C' \) so that the following holds: For any pair of subsets \( S, S' \subset C' \) s.t. \( |S| = |S'| \leq L \) we have:

\[
\|T(S) - T(S')\| \leq \delta_1
\]  

(30)

Consider any code \( C_1 \subset \mathbb{F}_2^n \) and define a hypergraph with vertices indexed by elements of \( C \) and hyper-edges corresponding to each of the subsets of size \( L \). Now define a \( \delta_1/2 \)-net on the space \( T_L \) 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 \( |C_1| \geq K_L \) then there is a subset \( C'_1 \subset C \) such that \( |C'_1| \geq K \) and each of the internal edges, indexed by \( \binom{C'_1}{L} \), is assigned the same label. Thus, by triangle inequality (30) follows for all \( S, S' \in \binom{C'_1}{L} \).

Next, apply the previous argument to show that there is a constant \( K_{L-1} \) such that for any \( C_2 \subset \mathbb{F}_2^n \) of size \( |C_2| \geq K_{L-1} \) there exists a subcode \( C'_2 \) of size \( |C'_2| \geq K_L \) satisfying (30) for all \( S, S' \in \binom{C'_2}{L-1} \). Since \( C'_2 \) satisfies the size assumption on \( C_1 \) made in previous paragraph, we can select a further subcode \( C''_2 \subset C'_2 \) of size \( \geq K_L \) so that for \( C''_2 \) property (30) holds for all \( S, S' \) of size \( L \) or \( L-1 \).

Continuing similarly, we may select a subcode \( C' \) of arbitrary \( C \) such that (30) holds for all \( |S| = |S'| \leq L \) provided that \( |C| \geq K_1 \).

Next, we show that (30) implies

\[
\|T(S_0) - \sigma(T(S_0))\| \leq C\delta_1,
\]  

(31)

where \( S_0 \in \binom{C'}{L} \) 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

$$|t_v - t_{\sigma(v)}| \leq \delta_1. \quad \forall v \in \mathbb{F}_2^L$$

(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

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

(33)

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

(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:

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

implying (32) and thus (31).

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

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

(35)

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

(36)

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

(37)

$$\leq (1 + C)\delta_1,$$

(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):

$$\overline{\text{rad}}(x_1, \ldots, x_m) \triangleq \min_y \frac{1}{m} \sum_{i=1}^{m} |x_i - m|.$$
Consider now an arbitrary subset \( S = \{c_1, \ldots, c_L\} \) and define for each \( j \geq 0 \) the following functions

\[
h_j(S) \triangleq \frac{1}{n} \frac{1}{\text{rad}(0, \ldots, 0, c_1, \ldots, c_L) \text{, \( j \) times}}.
\]

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

\[
h_j(S) = \frac{1}{L+j} \left( \mathbb{E}[W] - \mathbb{E}[|2W - L - j|] \right), \quad \mathbb{P}[W = w] = \sum_{v:|v|=w} t_v, \tag{39}
\]

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 \( C \) satisfying conclusion (29) of Lemma 5. Then for any subset \( S = \{c_1, \ldots, c_L\} \subset C \) we have

\[
\left| \frac{1}{n} \text{rad}(0, c_1, \ldots, c_L) - n \cdot \max_j h_j(T_L(C)) \right| \leq 2^L(1 + \delta n), \tag{40}
\]

where \( j \) in maximization (40) 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 \( j \geq 0 \) we clearly have (cf. expression (39))

\[
|h_j(T_1) - h_j(T_2)| \leq 2^{L-1}||T_1 - T_2||, \quad \forall T_1, T_2 \in T_L. \tag{41}
\]

We also trivially have

\[
\frac{1}{n} \text{rad}(0, c_1, \ldots, c_L) \geq h_j(S) \quad \forall j \geq 0. \tag{42}
\]

Thus from (29) and (41) we already get

\[
\frac{1}{n} \text{rad}(0, c_1, \ldots, c_L) \geq \max_j h_j(T_L(C)) - 2^{L-1}\delta.
\]
It remains to show

\[ \frac{1}{n} \text{rad}(0, c_1, \ldots, c_L) \leq \max_j h_j(\bar{T}_L(C)) + \delta + \frac{2L}{n}. \]  

(43)

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

\[ \text{rad}(c_0, c_1, \ldots, c_L) \leq \min_q \max_p \sum_{i=0}^L p_i |c_i - y(q)|, \]  

(44)

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

Denote

\[ T(S) = \{ t_v, v \in \mathbb{F}_2^L \} \]  

(45)

\[ \bar{T}_L(C) = \{ \bar{t}_v, v \in \mathbb{F}_2^L \} \]  

(46)

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

\[ |c_i - y(q)| \leq n \sum_{v \in \mathbb{F}_2^L} t_v(q_v) 1\{v(i) = 0\} + (1 - q_v) 1\{v(i) = 1\} + 2L, \]  

(47)

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 \( 2\delta n \) error, so we have:

\[ |c_i - y(q)| \leq n \sum_{v \in \mathbb{F}_2^L} \bar{t}_v(q_v) 1\{v(i) = 0\} + (1 - q_v) 1\{v(i) = 1\} + 2L(1 + \delta n). \]

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:
\[
\sum_{v \in F_2^L} \bar{t}_v(q|v|)1\{v(i) = 0\} + (1 - q|v|)1\{v(i) = 1\}) = \frac{1}{L} \mathbb{E} [\bar{W} + (L - 2\bar{W})q_{\bar{W}}].
\]

For \( i = 0 \) the expression is even simpler:
\[
\sum_{v \in F_2^L} \bar{t}_v(q|v|)1\{v(0) = 0\} + (1 - q|v|)1\{v(0) = 1\}) = \mathbb{E} [q_{\bar{W}}].
\]

Substituting derived upper bound on \(|c_i - y(q)|\) into (44) 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:
\[
= \max_{p_1 \in [0, L^{-1}]} \min_{q} p_1 \mathbb{E} [\bar{W}] + \mathbb{E} [q_{\bar{W}}(1 - 2\bar{W}p_1)]
\]
\[
= \max_{p_1 \in [0, L^{-1}]} p_1 \mathbb{E} [\bar{W}] - \mathbb{E} [|2\bar{W}p_1 - 1|^+].
\]

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 3. So we continue (49) getting
\[
= \max_{j} \frac{1}{L + j} (\mathbb{E} [\bar{W}] - \mathbb{E} [|2\bar{W} - L - j|^+])
\]

We can see that expression under maximization is exactly \( h_j(T_L(C)) \) and hence (43) is proved.

\]

Lemma 7: There exist constants \( C_1, C_2 \) depending only on \( L \) such that for any \( C \subset \mathbb{F}_2^n \) the joint-type \( T_L(C) \) is approximately a mixture of product Bernoulli distributions, namely:
\[
\left\| T_L(C) - \frac{1}{n} \sum_{i=1}^{n} \text{Bern}^\otimes L(\lambda_i) \right\| \leq \frac{C_1}{|C|},
\]

3The 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.

4Distribution \( \text{Bern}^\otimes L(\lambda) \) assigns probability \( \lambda^{|v|}(1 - \lambda)^{|L - |v|} \) to element \( v \in F_2^L \).
where \( \lambda_i = \frac{1}{|C|} \sum_{c \in C} 1\{c(i) = 1\} \) be the density of ones in the \( j \)-th column of a \( |C| \times n \) matrix representing the code. In particular,

\[
\left| h_j(\tilde{T}_L(C)) - \frac{1}{n} \sum_j g_j(\lambda_j) \right| \leq \frac{C_2}{|C|},
\]

where functions \( g_j \) were defined in (17).

**Proof:** Second statement (52) follows from the first via (41) and linearity of \( h_j(T) \) in the type \( T \), cf. (39). To show the first statement, let \( M = |C|, M_i = \lambda_i M \) and \( p_w \) – total probability assigned to vectors \( v \) of weight \( w \) by \( \tilde{T}_L(C) \). Then by computing \( p_w \) over columns of \( M \times n \) matrix we obtain

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

By a standard estimate we have for all \( w = \{0, \ldots, L\} \):

\[
\frac{\binom{M_i}{w} \binom{M - M_i}{L - w}}{\binom{M}{w}} = \left( \frac{L}{w} \right) \lambda^w_i (1 - \lambda_i)^{L-w} + O\left( \frac{1}{M} \right),
\]

with \( O(\cdot) \) term uniform in \( w \) and \( \lambda_i \). By symmetry of the type \( \tilde{T}_L(C) \) the result (51) follows. ■

**Lemma 8:** Functions \( g_j \) defined in (17) are concave on \([0, 1]\).

**Proof:** Let \( W_\lambda \sim \text{Bino}(L, \lambda) \) and \( V_\lambda \sim \text{Bino}(L - 1, \lambda) \). Denote for convenience \( \bar{\lambda} = 1 - \lambda \) and take \( j_0 \) to be an integer between 0 and \( L \). We have then

\[
\frac{\partial}{\partial \lambda} \mathbb{E} [ |W_\lambda - j_0| ] = \sum_{w = j_0 + 1}^L \binom{L}{w} (w - j_0) \{ w \lambda^{w-1} \bar{\lambda}^{L-w} - (L - w) \lambda^w \bar{\lambda}^{L-w-1} \}
\]

(53)

\[
= \binom{L}{j_0 + 1} (j_0 + 1) \lambda^{j_0} \bar{\lambda}^{L-j_0-1} + \sum_{w = j_0 + 1}^{L-1} \binom{L}{w+1} (w + 1 - j_0)(w + 1) - \binom{L}{w} (w - j_0)(L - w) \lambda^w \bar{\lambda}^{L-w-1}
\]

(54)

\[
= L \binom{L-1}{j_0} \lambda^{j_0} \bar{\lambda}^{L-1-j_0} + L \sum_{w = j_0 + 1}^{L-1} \binom{L-1}{w} \lambda^w \bar{\lambda}^{L-1-w}
\]

(55)

\[
= LP[V_\lambda \geq j_0],
\]

(56)

where in (54) we shifted the summation by one for the first term under the sum in (53), and in (55) 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

\[
\frac{\partial}{\partial \lambda} \mathbb{E} [ |W_\lambda - j_0 - \theta| ] = LP[V_\lambda \geq j_0 + 1] + L(1 - \theta)P[V_\lambda = j_0].
\]

(57)
Similarly, one shows (we will need it later in Lemma 9):

$$\frac{\partial}{\partial \lambda} \mathbb{P}[W_\lambda \geq j_0] = L \mathbb{P}[V_\lambda = j_0 - 1].$$

(58)

Since clearly the function in (57) is strictly increasing in \( \lambda \) for any \( j_0 \) and \( \theta \) we conclude that

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

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 \( C'_n \) to a subcode \( C''_n \) 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 \( C'''_n \subset C''_n \)
3) Use Lemmas 7-8 to estimate maxima of average radii \( h_j \) over \( C'''_n \).
4) Use Lemma 6 to transport statement about \( h_j \) to a statement on \( \tau_L(C'''_n) \).

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 \( n \geq n_0(\delta) \). So we fix \( \delta > 0 \) and consider a code \( C' \subset S_{\xi_0 n} \subset \mathbb{F}_2^n \) with \( |C'| \geq \exp\{n R' + o(n)\} \). Note that for any \( r \), even \( m \) with \( m/2 \leq \min(r, n-r) \) 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:

$$\{y + S_m^n\} \cap S_r^n \simeq S_{r-m/2}^r \times S_{m/2}^{n-r}.$$  
(59)

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

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

Consequently, we can select \( m = \xi_1 n - o(n) \), where \( \xi_1 \) defined in (27), so that for some \( y \in S_r^n \):

$$|\{y + S_m^n\} \cap C'| > n.$$  

Note that we focus on solution of (27) satisfying \( \xi_1 < 2\xi_0(1 - \xi_0) \). For some choices of \( R, \delta \) and \( \xi_0 \) 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 \( C'' = \{y + S_m^n\} \cap C' \). For sufficiently large \( n \) the code \( C'' \) will satisfy assumptions of Lemma 5 with \( K \geq \frac{1}{\epsilon} \). Denote the resulting large symmetric subcode \( C''' \).
Note that because of (59) column-densities \( \lambda_i \)'s of \( C''' \), defined in Lemma 7, satisfy (after possibly reordering coordinates):

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

Therefore, from Lemmas 7-8 we have

\[
h_j(T_L(C''')) \leq \xi_0 g_j \left( 1 - \frac{\xi_1}{2\xi_0} \right) + (1 - \xi_0)g_j \left( \frac{\xi_1}{2(1 - \xi_0)} \right) + \epsilon'_n + \frac{C_1}{|C'''|},
\]

where \( \epsilon'_n \to 0 \). Note that by construction the last term in (60) is \( O(\delta) \). Also note that the first two terms in (60) equal \( \theta_j \) defined in (25).

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

\[
\frac{1}{n} \text{rad}(0, c_1, \ldots, c_L) \leq \theta(\xi_0, R', L) + \epsilon''_n + C\delta.
\]

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

\[
\max_j g_j(x) = g_1(x),
\]

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

\[
g_1(x) = \frac{1}{2} - 2^{-L-1} \left( \frac{L}{L-1} \right) + cx + O((2x-1)^2), \quad x \to \frac{1}{2}.
\]

**Proof:** First, the value \( g_1(1/2) \) is computed trivially. Then from (57) we have

\[
\frac{d}{dx} g_j(x) = \frac{L}{L+j} \left( 1 - 2^P \left[ V_x \geq \frac{L+j}{2} \right] \right), \quad j \geq 1, V_x \sim \text{Bino}(x, L - 1)
\]

and this implies (62). For future reference we note that (66) (below) and (58) imply

\[
\frac{d}{dx} g_0(x) = 1 - 2^P(V_x \geq \frac{L+1}{2}) - P(V_x = \frac{L-1}{2}), \quad V_x \sim \text{Bino}(x, L - 1).
\]

By continuity, (61) follows from showing

\[
g_1(1/2) > \max_{j \in \{0,3,5,...,L\}} g_j(1/2).
\]
Next, consider \( W_x \sim \text{Bino}(x, L) \) and notice the upper-bound
\[
g_j(x) \leq \frac{1}{L+j} \mathbb{E} \left[ W_x \mathbb{1}\{W_x \leq a\} + (L+j-W_x)\mathbb{1}\{W_x \geq a+1\} \right].
\]

Then, substituting expression for \( g_1(x) \) we get
\[
g_1(x) - g_0(x) = \frac{1}{L} (\mathbb{P}[W_x \geq a + 1] - g_1(x)) \tag{66}
\]
\[
g_1(x) - g_j(x) \geq \frac{j-1}{L+j} (g_1(x) - \mathbb{P}[W_x > a + 1]) \tag{67}
\]

Thus, to show (65) it is sufficient to prove that for \( x = \frac{1}{2} \) we have
\[
\mathbb{P}[W_{\frac{1}{2}} > a + 1] < g_1(1/2) < \mathbb{P}[W_{\frac{1}{2}} \geq a + 1]. \tag{68}
\]

The right-hand inequality is trivial since \( \mathbb{P}[W_{\frac{1}{2}} \geq a + 1] = \frac{1}{2} \) while from (62) we know \( g_1(1/2) < 1/2 \). Proof of the left-hand inequality is more involved. First, after simple algebra it reduces to showing
\[
\sum_{v=1}^{a} (2v+1) \binom{2a+1}{a-v} < (2a+1) \binom{2a+1}{a} \tag{69}
\]

To show the latter inequality, we consider the bounds
\[
1 - x \leq e^{-x}, \quad \frac{1}{1+x} \leq e^{-x \ln 2}, \quad x \in [0,1].
\]
Then we have the chain
\[
\binom{2a+1}{a-v} = \frac{a}{a+2} \frac{a-1}{a+3} \cdots \frac{a-v}{a+v+1} \binom{2a+1}{a} \tag{70}
\]
\[
\leq \binom{2a+1}{a} e^{-\frac{1}{a} \ln^2(v+1)} e^{-\frac{1}{2} \ln^2(v+3)} \tag{71}
\]
\[
= \binom{2a+1}{a} e^{-\frac{v^2}{a} \ln \frac{a+2}{a} - \frac{v}{a} \ln \frac{a+2}{2} - \frac{1}{2}} \tag{72}
\]
\[
\leq \binom{2a+1}{a} e^{-\frac{v^2}{a} - \frac{c_2}{a}}, \tag{73}
\]
where in the last step we denoted \( c = \frac{\ln^2 2+1}{2}, c_2 = \frac{3\ln^2 2-1}{2} \) and used the fact that \( v \geq 1 \).

With estimate (73) we have due to monotonicity of the summand:
\[
\sum_{v=1}^{a} \binom{2a+1}{a-v} \leq \binom{2a+1}{a} e^{-\frac{c_2}{a}} \int_0^{\infty} e^{-\frac{v^2}{a}} dv = \binom{2a+1}{a} e^{-\frac{c_2}{a}} \sqrt{\pi a} \frac{4}{c}. \tag{74}
\]
Next, notice that for any non-negative continuous function that has only one extremum on the interval \([b_0, b_1]\) we have
\[
\sum_{n=b_0}^{b_1} f(n) \leq \int_{b_0}^{b_1} f(x) dx + \max_{x \in [b_0, b_1]} f(x) .
\] (75)

The maximum of \(2ve^{-cv^2/a}\) is given by
\[
\max_{v \geq 0} 2ve^{-cv^2/a} = \sqrt{2a/c}.
\] (76)

Meanwhile, the integral of this very function is
\[
\int_{0}^{\infty} 2ve^{-cv^2/a} dv = \frac{a}{c}.
\] (77)

Then we can estimate, via (73), (75), (76) and (77):
\[
\sum_{v=1}^{a} \left( \frac{2a + 1}{a - v} \right) \leq \left( \frac{2a + 1}{a} \right) e^{-\frac{c^2}{a}} \left\{ \frac{a}{c} + \sqrt{2a/c} \right\}.
\] (78)

Together (74) and (78) imply
\[
\sum_{v=1}^{a} (2v + 1) \left( \frac{2a + 1}{a - v} \right) \leq \left( \frac{2a + 1}{a} \right) e^{-\frac{c^2}{a}} \left\{ \frac{a}{c} + \sqrt{2a/c} + \sqrt{\frac{\pi a}{4c}} \right\}.
\] (79)

To further upper-bound this sum, notice that for all \(a \geq 1\) we have
\[
e^{-\frac{c^2}{a}} \left\{ \frac{a}{c} + \sqrt{2a/c} + \sqrt{\frac{\pi a}{4c}} \right\} < 2a + 1
\] (80)

For \(a \gg 1\) the inequality (80) holds since \(\frac{1}{c} < 2\). For small \(a\) its validity is readily verified numerically. Overall, (79) and (80) imply (69).

\textbf{Proof of Corollary 3:} We first show that (21) implies (22). To that end, fix a small \(\epsilon > 0\) so that \(\frac{1}{2} - \epsilon\) 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
\[
R = h(e^2 + o(e^2)) ,
\] (81)

By Lemma 9, the right-hand side of (21) is
\[
\tau_L^*(0) - \text{const} \cdot \epsilon + o(\epsilon),
\]
which together with (81) implies (22).
To prove (21) we use Theorem 1 with $\delta = \delta_{LP_1}(R)$. Next, use concavity of $g_j$’s (Lemma 8) to relax (14) to
\[
\limsup_{n \to \infty} \tau_L(C_n) \leq \max_j g_j(\xi_0).
\]
From (63) and (64) it is clear that $\xi_0 \mapsto g_j(\xi_0)$ is monotonically increasing for all $j \geq 0$ on the interval $[0, 1/2]$. Thus, we further have
\[
\limsup_{n \to \infty} \tau_L(C_n) \leq \max_j g_j(\delta_{LP_1}(R)). \tag{82}
\]
Bound (82) 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, $\delta_{LP_1}(R)$ belongs to the neighborhood of $1/2$ in Lemma 9 and thus (21) follows from (82) and (61).

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
\[
g_0(\frac{1}{2} - \epsilon) = \tau^*_L(0) + c\epsilon^2 + O(\epsilon^3), \epsilon \to 0 \tag{83}
\]
Therefore, for $\delta_{LP_1}(R) = \frac{1}{2} - \epsilon$ we get from (83) that the right-hand side of (82) evaluates to
\[
\tau^*_L(0) - \text{const} \cdot \epsilon^2 \log \frac{1}{\epsilon}. \tag{84}
\]
Thus, comparing (84) with (81) we conclude that for even $L$ our bound on $R^*_L(\tau)$ has positive slope at zero rate. Note that Blinovsky’s bound (11) has non-zero slope at zero rate for both odd and even $L$.

D. Proof of Corollary 2

Proof: Instead of working with parameter $\delta$ we introduce $\beta \in [0, 1/2]$ such that
\[
\delta = \frac{1}{2} - \sqrt{\beta(1 - \beta)}.
\]
We then apply Theorem 1 with $h(\beta) = R$. Notice that the bound on $\xi_0$ in (15) becomes
\[
0 \leq \xi_0 \leq \delta.
\]
By a simple substitution $\omega = \sqrt{\frac{\beta}{1 - \beta}}$ we get from (12)
\[
E_{\beta}(\delta) = \frac{1}{2}(\log 2 - h(\delta) + h(\beta)).
\]

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 $\xi_1$, i.e. (16), coincides with (20).

Next for $L = 3$ we compute

$$g_0(\nu) = \nu(1 - \nu), \quad (85)$$
$$g_1(\nu) = \frac{3}{4} \nu - \frac{1}{2} \nu^3, \quad (86)$$
$$g_3(\nu) = \frac{1}{2} \nu. \quad (87)$$

Note that the right-hand side of (18) is precisely equal to

$$\delta g_1 \left(1 - \frac{\xi_0}{2\delta}\right) + (1 - \delta)g_1 \left(\frac{\xi_0}{2(1 - \delta)}\right).$$

So this corollary simply states that for $L = 3$ the maximum in (14) is achieved at $j = 1, \xi_0 = \delta$.

Let us restate this last statement rigorously: The maximum

$$\max_{j \in \{0, 1, 3\}} \max_{\xi_0 \in \delta} \max g_j \left(1 - \frac{x}{2\xi_0}\right) + (1 - \xi_0)g_j \left(\frac{x}{2(1 - \xi_0)}\right) \quad (88)$$

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

$$2(h(\beta) - E_\beta(\xi_0)) = h(\xi_0) - \xi_0h \left(\frac{x}{2\xi_0}\right) - (1 - \xi_0)h \left(\frac{x}{2(1 - \xi_0)}\right). \quad (89)$$

For notational convenience we will denote the function under maximization in (88) by $g_j(\xi_0, x)$.

We proceed in two steps:

- First, we estimate the maximum over $\xi_0$ for $j = 0$ as follows:

$$\max_{\xi_0} g_0(\xi_0, x) \leq \frac{\log 2 - R}{4 \log 2} \cdot \left(1 - \frac{1 - \delta}{a_{\max}(1 - a_{\max})}\right) + (1 - \delta)g_0(a_{\min}), \quad (90)$$

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

$$a_{\max} = h^{-1}(\log 2 - R), \quad (91)$$
$$a_{\min} = h^{-1}\left(\log 2 - \frac{R}{1 - \delta}\right). \quad (92)$$

- 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 (88)) and the right-hand side of (90) (term corresponding to \( j = 0 \)). Notice that this relation holds for all rates. Therefore, maximum in (88) 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 (89) 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 \([13, (46)]\):

\[
E_{\beta}(\xi_0) \leq \frac{1}{2}(\log 2 + h(\beta) - h(\xi_0)).
\]

(93)

Furthermore, it can also be argued that maximum cannot be attained by \( \xi_0 \) so small that

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

So from now on, we assume that

\[
h^{-1}(\log 2 - h(\beta)) \leq \xi_0 \leq \delta,
\]

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

\[
\log 2 - R = \xi_0 h\left(\frac{x}{2\xi_0}\right) + (1 - \xi_0)h\left(\frac{x}{2(1 - \xi_0)}\right)
\]

(94)

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

We proceed to demonstrating (90). For convenience, we introduce

\[
a_1 \triangleq 1 - \frac{x}{2\xi_0},
\]

(95)

\[
a_2 \triangleq \frac{x}{2 - 2\xi_0}.
\]

(96)

By constraints on \( x \) it is easy to see that

\[
0 \leq a_2 \leq \min(a_1, 1 - a_1).
\]

Therefore, we have

\[
\log 2 - R = \xi_0 h(a_1) + (1 - \xi_0)h(a_2) \geq h(a_2)
\]

and thus \( a_2 \leq a_{\text{max}} \) defined in (91). Similarly, we have

\[
\log 2 - R = \xi_0 h(a_1) + (1 - \xi_0)h(a_2) \leq \xi_0 \log 2 + (1 - \xi_0)h(a_2),
\]
and since \( \xi_0 \leq \delta \) we get that \( a_2 \geq a_{\min} \) defined in (92).

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

\[
h(a_1) \geq g_0(a_1)4\log 2 \tag{97}
\]

\[
h(a_2) \geq \frac{g_0(a_2)}{g_0(a_{\max})} = \frac{\log 2 - R}{a_{\max}(1-a_{\max})}g_0(a_2) \triangleq c \cdot g_0(a_2), \tag{98}
\]

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

\[
\log 2 - R = \xi_0 h(a_1) + (1 - \xi_0) h(a_2) \tag{99}
\]

\[
\geq 4\log 2 \cdot \xi_0 g_0(a_1) + (1 - \xi_0)c \cdot g_0(a_2) \tag{100}
\]

\[
= 4\log 2 \cdot g_0(\xi_0, x) + (1 - \xi_0)(c - 4\log 2) \cdot g_0(a_2) \tag{101}
\]

\[
\geq 4\log 2 \cdot g_0(\xi_0, x) + (1 - \delta)(c - 4\log 2) \cdot g_0(a_{\min}). \tag{102}
\]

Rearranging terms yield (90).

We proceed to proving monotonicity of (89). 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:

\[
\frac{d}{d\xi_0} \xi_0 g(a_1) + (1 - \xi_0) g(a_2) \geq 0 \iff \int_{a_2}^{a_1} \frac{1}{2} \frac{dx}{d\xi_0} (g'(a_2) - g'(a_1)) \geq \int_{a_2}^{a_1} (1-x)(-g''(x))dx - g'(a_2), \tag{103}
\]

where we recall definition of \( a_1, a_2 \) in (95)-(96). Differentiating (94) in \( \xi_0 \) (and recalling that \( R \) is fixed, while \( x = x(\xi_0) \) is an implicit function of \( \xi_0 \)) we find

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

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

\[
\{(a_1, a_2) : 0 \leq a_1 \leq 1, 0 \leq a_2 \leq a_1(1-a_1)\}. \tag{105}
\]

With the inverse map given by

\[
\xi_0 = \frac{a_2}{1-a_1+a_2}, x = \frac{2a_2}{1-a_1+a_2}, R = \log 2 - \xi_0 h(a_1) - (1-\xi_0)h(a_2).
\]
Thus, verifying (104) can as well be done for all $a_1, a_2$ inside the region (105). Substituting $g = g_1$ into (104) we get that monotonicity in (89) is equivalent to a two-dimensional inequality:

$$-2 \log \frac{1 - a_2}{a_1} \cdot (a_1^2 - a_2^2) \geq \left(2a_1^2 - \frac{4}{3}(a_1^3 - a_2^3) - 1\right) \log \frac{1 - a_2}{a - 2} \cdot \frac{a_1}{1 - a_1}.$$  \hspace{1cm} (106)

It is possible to verify numerically that indeed (106) holds on the set (105). 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.

**REFERENCES**