Improved List-Decodability of Random Linear Binary Codes

Ray Li[1]
Department of Computer Science, Stanford University, USA
rayyli@stanford.edu

Mary Wootters
Departments of Computer Science and Electrical Engineering, Stanford University, USA
marykw@stanford.edu

Abstract

There has been a great deal of work establishing that random linear codes are as list-decodable as uniformly random codes, in the sense that a random linear binary code of rate $1 - H(p) - \epsilon$ is $(p, O(1/\epsilon))$-list-decodable with high probability. In this work, we show that such codes are $(p, H(p)/\epsilon + 2)$-list-decodable with high probability, for any $p \in (0, 1/2)$ and $\epsilon > 0$. In addition to improving the constant in known list-size bounds, our argument – which is quite simple – works simultaneously for all values of p, while previous works obtaining $L = O(1/\epsilon)$ patched together different arguments to cover different parameter regimes.

Our approach is to strengthen an existential argument of (Guruswami, Håstad, Sudan and Zuckerman, IEEE Trans. IT, 2002) to hold with high probability. To complement our upper bound for random linear binary codes, we also improve an argument of (Guruswami, Narayanan, IEEE Trans. IT, 2014) to obtain a tight lower bound of $1/\epsilon$ on the list size of uniformly random binary codes; this implies that random linear binary codes are in fact *more* list-decodable than uniformly random binary codes, in the sense that the list sizes are strictly smaller.

To demonstrate the applicability of these techniques, we use them to (a) obtain more information about the distribution of list sizes of random linear binary codes and (b) to prove a similar result for random linear rank-metric codes.

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

Keywords and phrases List-decoding, Random linear codes, Rank-metric codes

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2018.50

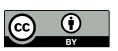
Related Version [26], `https://arxiv.org/abs/1801.07839`

Acknowledgements We thank anonymous reviewers for helpful comments on a previous version of this manuscript.

1 Introduction

An *error correcting code* is a subset $\mathcal{C} \subseteq \mathbb{F}_2^n$, which is ideally "spread out." In this paper, we focus on one notion of "spread out" known as *list-decodability.* We say that a code $\mathcal{C}$ is (p, L)-list-decodable if any Hamming ball of radius pn in $\mathbb{F}_2^n$ contains at most L points of $\mathcal{C}$: that is, if for all $x \in \mathbb{F}_2^n$, $|\mathcal{B}(x, pn) \cap \mathcal{C}| \leq L$, where $\mathcal{B}(x, pn)$ is the Hamming ball of radius pn centered at x. Since list-decoding was introduced in the 1950's [6, 44], it has found

[1] Research supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE - 1656518.



applications beyond communication, for example in pseudorandomness [41] and complexity theory [40].

A classical result in list-decoding is known as the *list-decoding capacity theorem:*

▶ **Theorem 1** (List-decoding capacity theorem). *Let $p \in (0, 1/2)$ and $\varepsilon > 0$.*

1. *There exist binary codes of rate $1 - H(p) - \varepsilon$ that are $(p, \lceil 1/\varepsilon \rceil)$-list-decodable.*
2. *Any binary code of rate $1 - H(p) + \varepsilon$ that is (p, L)-list-decodable up to distance p must have $L \geq 2^{\Omega(\varepsilon n)}$.*

Above, $H(p) = -(1-p)\log_2(1-p) - p\log_2(p)$ is the binary entropy function. We say that a family of binary codes with rate approaching $1 - H(p)$ which are (p, L)-list-decodable for $L = O(1)$ *achieves list-decoding capacity.*[2]

Theorem 1 is remarkable because it means than even when pn is much larger than half the minimum distance of the code – the radius at which at most one codeword $c \in \mathcal{C}$ lies in any Hamming ball – it still can be the case that only a constant number of $c \in \mathcal{C}$ lie in any Hamming ball of radius pn. Because of this, there has been a great deal of work attempting to understand what codes achieve the bound in Theorem 1.

The existential part of Theorem 1 is proved by showing that a uniformly random subset of $\mathbb{F}_2^n$ is $(p, 1/\varepsilon)$-list-decodable with high probability. For a long time, uniformly random codes were the only example of binary codes known to come close to this bound, and today we still do not have many other options. There are explicit constructions of capacity-achieving list-decodable codes over large alphabets (either growing with n or else large-but-constant) [5, 21, 22], but over binary alphabets we still do not have any explicit constructions; we refer the reader to the survey [11] for an overview of progress in this area.

Because it is a major open problem to construct explicit binary codes of rate $1 - H(p) - \varepsilon$ with constant (or even $\text{poly}(n)$) list-sizes, one natural line of work has been to study structured random approaches, in particular *random linear codes.* A random linear code $\mathcal{C} \subset \mathbb{F}_2^n$ is simply a random subspace of $\mathbb{F}_2^n$, and the list-decodability of these codes has been well-studied [45, 13, 12, 2, 43, 34, 36]. There are several reasons to study the list-decodability of random linear codes. Not only is it a natural question in its own right as well as a natural stepping stone in the quest to obtain explicit binary list-decodable codes, but also the list-decodability of random linear codes is useful in other coding-theoretic applications. One example of this is in concatenated codes and related constructions [14, 17, 24, 23], where a random linear code is used as a short inner code. Here, the linearity is useful because (a) a linear code can be efficiently described; (b) it is sometimes desirable to obtain a linear code at the end of the day, hence all components of the construction must be linear; and (c) as in [24] sometimes the linearity is required for the construction to work.

To this end, the line of work mentioned above has aimed to establish that random linear codes are "as list-decodable" as uniformly random codes. That is, uniformly random codes are viewed (as is often the case in coding theory) as the optimal construction, and we try to approximate this optimality with random linear codes, despite the additional structure.

Our contributions

In this paper, we give an improved analysis of the list-decodability of random linear binary codes. More precisely, our contributions are as follows.

[2] Sometimes the phrase "achieves list-decoding capacity" is also used when $L = \text{poly}(n)$; since this paper focuses on the exact constant in the $O(1)$ term however, we use it to mean that $L = O(1)$.

- **A unified analysis.** As we discuss more below, previous work on the list-decodability of random linear binary codes either work only in certain (non-overlapping) parameter regimes [12, 43], or else get substantially sub-optimal bounds on the list-size [36]. Our argument obtains improved list size bounds over all these results and works in all parameter regimes.
  Our approach is surprisingly simple: we adapt an existential argument of Guruswami, Håstad, Sudan and Zuckerman [13] to hold with high probability. Extending the argument in this way was asked as an open question in [13] and had been open until now.
- **Improved list-size for random linear codes.** Not only does our result imply that random linear codes of rate $1-H(p)-\varepsilon$ are (p, L)-list-decodable with list-size $L = O(1/\varepsilon)$, in fact we show that $L \leq H(p)/\varepsilon + 2$. In particular, the leading constant is small and – to the best of our knowledge – is the best known, even existentially, for any list-decodable code.
- **Finer-grained information about the combinatorial structure of random linear codes.** We extend our argument to obtain more information about the distribution of list sizes of random linear codes. More precisely, we obtain high-probability bounds on the number of points x so that the *list size at* x, $L_{\mathcal{C}}(x) := |\mathcal{B}(x, pn) \cap \mathcal{C}|$, is at least ℓ.
- **Tight list-size lower bound for uniformly random codes.** To complement our upper bound, we strengthen an argument of Guruswami and Narayanan [15] to show that a uniformly random binary code of rate $1 - H(p) - \varepsilon$ requires $L \geq (1-\gamma)/\varepsilon$ for any constant $\gamma > 0$ and sufficiently small ε. In other words, the list size of $1/\varepsilon$ in Theorem 1 is tight even in the leading constant. Thus, random linear codes are, with high probability, list-decodable with smaller list sizes than completely random codes.[3]
- **Results for rank-metric codes.** Finally, we adapt our argument for random linear codes to apply to random linear rank-metric codes. As with standard (Hamming-metric) codes, recent work aimed to show that random linear rank-metric codes are nearly as list-decodable as uniformly random codes [4, 16]. Our approach establishes that in fact, random linear binary rank-metric codes are more list-decodable than their uniformly random counterparts in certain parameter regimes, in the sense that the list sizes near capacity are strictly smaller. Along the way, we show that low-rate random linear binary rank-metric codes are list-decodable to capacity, answering a question of [16].

On the downside, we note that our arguments only work for binary codes and do not extend to larger alphabets; additionally, our positive results do not establish *average-radius* list-decodability with list size $O(1/\varepsilon)$, a stronger notion which was established in some of the previous works [2, 43, 36]. It would be very interesting to extend our results to these settings.

1.1 Outline of paper

After a brief overview of the notation in §1.2, we proceed in §2 with a survey of related work for both random linear codes and rank-metric codes, and we formally state our results in this context. In §3, we prove Theorem 5, which establishes our upper bound for random linear binary codes. In Appendix A, we prove Theorem 6, characterizing the list size distribution of random linear codes. We refer the interested reader to the full version of the paper [26] for proofs of our other results, Theorems 7, 11, and 12.

[3] In retrospect, this may not be surprising: for example, it is well-known that random linear codes have better distance than completely random codes. However, the fact that we are able to prove this is surprising to the authors, since it requires taking advantage of the dependence between the codewords, rather than trying to get around it.

1.2 Notation

Throughout most of the paper, we are interested in binary codes $\mathcal{C} \subseteq \mathbb{F}_2^n$ of *block length* n. The *dimension* of a code $\mathcal{C}$ is defined as $k = \log_2 |\mathcal{C}|$, and the *rate* is the ratio k/n. We define a *uniformly random binary code of rate* R to be a set $\mathcal{C}$ of 2^{Rn} elements chosen independently and uniformly at random from $\mathbb{F}_2^n$. We say that a binary code is *linear* if it forms a linear subspace of $\mathbb{F}_2^n$. We define a *random linear binary code of rate* R to be the span of $k = Rn$ independently random vectors $b_1, \ldots, b_k \in \mathbb{F}_2^n$.[4]

For two points $x, y \in \mathbb{F}_2^n$, we use $\Delta(x, y) = \sum_{i=1}^n \mathbb{I}[x_i = y_i]$ to denote the Hamming distance between x and y, where, for an event $\mathcal{E}$, $\mathbb{I}[\mathcal{E}]$ is 1 if $\mathcal{E}$ occurs and 0 otherwise. For $x \in \mathbb{F}_2^n, r \in [0, n]$, we define the Hamming ball $\mathcal{B}(x, r)$ of radius r centered at x to be $\mathcal{B}(x, r) = \{y \in \mathbb{F}_2^n : \Delta(x, y) \leq r\}$, and the volume of $\mathcal{B}(x, r)$ to be $\mathrm{Vol}(n, r) := |\mathcal{B}(0^n, r)| = \sum_{i=0}^r \binom{n}{i}$. We use the well known bound that, for any $p \in [0, 1]$, $\mathrm{Vol}(n, pn) \leq 2^{H(p)n}$, where $H(p) = -(1-p)\log_2(1-p) - p\log_2(p)$ is the binary entropy function. One of our main technical results is about the distribution of *list sizes* of points $x \in \mathbb{F}_2^n$: given a code $\mathcal{C}$ and $p \in (0, 1/2)$, we define the list size of a point $x \in \mathbb{F}_2^n$ to be $L_{\mathcal{C}}(x) := |\mathcal{B}(x, pn) \cap \mathcal{C}|$.

For $\alpha > 0, \beta \in \mathbb{R}$, let $\exp_\alpha(\beta) := \alpha^\beta$ and assume $\alpha = e$ when it is omitted. For two sets $A, B \subseteq \mathbb{F}_2^n$, define the sumset $A + B = \{a + b : a \in A, b \in B\}$. When $b \in \mathbb{F}_2^n$, let $A + b$ denote $A + \{b\}$.

2 Previous Work and Our Results

In §2.1 below, we survey related work on the list-decodability of random linear binary codes. In §2.2, we state our positive results for random linear codes. In §2.3, we state our negative result for uniformly random codes. In §2.4, we introduce and survey existing work on random rank-metric codes. In §2.5, we state our results on the list-decodability of random linear and uniformly random binary rank-metric codes.

2.1 Prior work: uniformly random and random linear codes

The list-decodability of random linear binary codes has been well studied. Here we survey the results that are most relevant for this work. As this work focuses on binary codes, we focus this survey on results for binary codes, even though many of the works mentioned also apply to general q-ary codes. We additionally remark that, in contrast to the large alphabet setting [18], capacity achieving binary codes have no known explicit constructions.

A modification of the proof of the list-decoding capacity theorem shows that a random linear code of rate $1 - H(p) - \varepsilon$ is $(p, \exp(O(\frac{1}{\varepsilon})))$-list-decodable [45]. However, whether or not random linear codes of this rate have list-sizes that do not depend exponentially on ε remained open for decades: this question was explicitly asked in [7].

A first step was given in the work of Guruswami, Håstad, Sudan and Zuckerman [13], who proved via a beautiful potential-function-based-argument that there *exist* binary linear codes or rate $1 - H(p) - \varepsilon$ which are $(p, 1/\varepsilon)$-list-decodable. However, this result did not hold with high probability. Our approach relies heavily on the approach of [13], and we return to their argument in §3.

[4] Our definitions of uniformly random binary code and random linear binary code are slightly different than the standard definitions, which are a uniformly random set of size 2^{Rn} and a uniformly random subspace of dimension k, respectively. However, our definitions are easier to work with. Furthermore, for this paper, the difference is negligible. For example, a random linear code has rank strictly less than k with probability at most $2^{-(n-k)}$, and each dimension k code is represented in the same number of ways, so our results hold also for the more standard definition.

Over the next 15 years, a line of work [12, 2, 43, 34, 35, 36] has focused on the list-decodability (and related properties) of random linear codes, which should hold with high probability. The works most relevant to ours are [12, 43], which together more or less settle the question. We state these results here for binary alphabets, although both works address larger alphabets as well.

The first result, of [12], establishes a result for a constant p, bounded away from $1/2$.

▶ **Theorem 2** (Theorem 2 of [12]). *Let $p \in (0, 1/2)$. Then there exist constants $C_p, \delta > 0$ such that for all $\varepsilon > 0$ and sufficiently large n, for all $R \leq 1 - H(p) - \varepsilon$, if $\mathcal{C} \subseteq \mathbb{F}_2^n$ is a random linear code of rate R, then $\mathcal{C}$ is $(p, C_p/\varepsilon)$-list-decodable with probability at least $1 - 2^{-\delta n}$.*

However, C_p is not small and tends to ∞ as p approaches $1/2$. The following result of [43] fills in the gap when p is quite close to $1/2$.

▶ **Theorem 3** (Theorem 2 of [43]). *There exist constants C_1, C_2 so that for all sufficiently small $\varepsilon > 0$ and sufficiently large n, for $p = 1/2 - C_1\sqrt{\varepsilon}$ and for all $R \leq 1 - H(p) - \varepsilon$, if $\mathcal{C} \subseteq \mathbb{F}_2^n$ is a random linear code of rate R, then $\mathcal{C}$ is $(p, C_2/\varepsilon)$-list-decodable with probability at least $1 - o(1)$.*

The list-decoding capacity theorem implies that we cannot hope to take the rate R substantially larger than $1 - H(p) - \varepsilon$ and obtain a constant list size. Moreover, the list size $\Theta(1/\varepsilon)$ is optimal for both random linear codes and uniformly random codes [32, 15]. More precisely, Guruswami and Narayanan show the following theorem (which we have specialized to binary codes).

▶ **Theorem 4** (Theorem 20 of [15]). *Let $\varepsilon > 0$. A uniformly random binary code of rate $1 - H(p) - \varepsilon$ is $(p, (1 - H(p))/\varepsilon)$-list-decodable with probability at most $\exp(-\Omega_{p,\varepsilon}(n))$.*[5]

We note that for general codes (not uniformly random or random linear) it is still unknown what the "correct" list size L is in terms of ε, although there are results in particular parameter regimes [1, 19] and for stronger notions of list-decodability [15].

2.2 Our main results: random linear codes

We show that, with high probability, a random linear binary code of rate $1 - H(p) - \varepsilon$ is (p, L)-list-decodable with $L \sim H(p)/\varepsilon$. More precisely, the upper bound is as follows (proved in §3).

▶ **Theorem 5.** *Let $p \in (0, 1/2)$, let $\varepsilon > 0$, and let $R = 1 - H(p) - \varepsilon$. Let $\mathcal{C} \subseteq \mathbb{F}_2^n$ be a random linear code of rate R. Then with probability $1 - \exp(-\Omega_\varepsilon(n))$, the code $\mathcal{C}$ is $(p, H(p)/\varepsilon + 2)$-list-decodable.*

Theorem 5 improves upon the picture given by Theorems 2 and 3 in two ways. First, the leading constant on the list size, which is $H(p)$, improves over both the constant C_p from Theorem 2 (which blows up as $p \to 1/2$) and on the constant C_2 from Theorem 3 (which the authors do not see how to make less than 2). Moreover, when $p \to 1/2$, Theorem 5 improves on Theorem 3 in that it decouples p from ε: in Theorem 3, we must take $p = 1/2 - O(\sqrt{\varepsilon})$

[5] In fact, in [15], Theorem 20 is stated with a list size of $(1 - H(p))/2\varepsilon$ as a lower bound. However, the constant can be improved to $1 - H(p)$, because the factor of 2 is introduced to handle an additive constant term. Thus, for sufficiently large n their argument proves the stronger statement stated above.

and $R = 1 - H(p) - \varepsilon$, while in Theorem 5, p and ε may be chosen independently. Thus, Theorem 5 offers the first true "list-decoding capacity theorem for binary linear codes," in that it precisely mirrors the quantifiers in Theorem 1.

The techniques that we use to prove Theorem 5 can be refined to give more combinatorial information about random linear codes. It is our hope that such information will help in further derandomizing constructions of binary codes approaching list-decoding capacity. In Appendix A, we prove the following structural result about random linear binary codes.

▶ **Theorem 6.** *Let $\varepsilon, \gamma \in (0,1)$ be constants, $p \in (0,1/2)$, L be a positive integer, and let $R = 1 - H(p) - \varepsilon$. Let $\mathcal{C} \subset \mathbb{F}_2^n$ be a random linear code of rate R. Then with probability $1 - \exp(-\Omega_L(\gamma\varepsilon n))$, the code $\mathcal{C}$ satisfies, for all $0 \le \ell \le L$,*

$$|\{x \in \mathbb{F}_2^n : L_{\mathcal{C}}(x) \ge \ell\}| \;\le\; 2^n \cdot 2^{-n\ell\varepsilon(1-\gamma)}.$$

To interpret this result, it is helpful to think of γ as close to 0. Intuitively, it says that, with high probability over the choice of a random linear code $\mathcal{C}$, the number of words $x \in \mathbb{F}_2^n$ with "list size ℓ" decays approximately exponentially as $2^{-n\ell\varepsilon}$. As we show in Appendix A, Theorem 5 follows as a corollary of Theorem 6 (see Corollary 19), but as a warm-up to Theorem 6 we present a proof of Theorem 5 independent of Theorem 6 in §3.

▶ Remark. Theorem 6 implies that, with high probability over the choice of the code, for any codeword $c \in \mathcal{C}$, $\mathbf{Pr}_{x \in \mathcal{B}(c,pn)}[L_{\mathcal{C}}(x) = 1] \ge 1 - 2^{-n(\varepsilon(1-\gamma)-o(1))}$. That is, "most" points $x \in \{0,1\}^n$ within pn of a codeword $c \in \mathcal{C}$ are not within pn of any other codeword $c' \ne c$. This is in line with the conventional wisdom from the Shannon model: with high probability, random linear codes achieve capacity on the BSC, so for a random linear code, a random center x obtained by sending a codeword $c \in \mathcal{C}$ through the binary symmetric channel $\mathrm{BSC}(p)$ has $L_{\mathcal{C}}(x) \le 1$ with high probability. (See also [33]). Thus, Theorem 6 recovers this intuition for list size 1, and quantitatively extends it to list sizes larger than 1.

2.3 Our results: uniformly random codes

In Theorem 5, the list size of $H(p)/\varepsilon + 2$ is smaller than the list size of $1/\varepsilon$ given by the classical list-decoding capacity theorem for uniformly random codes. Further, the following negative result shows that the list size of $1/\varepsilon$ given by uniformly random binary codes in the list-decoding capacity theorem is tight, even in the leading constant of 1.

▶ **Theorem 7.** *For any $p \in (0,1/2)$ and $\varepsilon > 0$, there exists a $\gamma_{p,\varepsilon} = \exp(-\Omega_p(\frac{1}{\varepsilon}))$ and $n_{p,\varepsilon} \in \mathbb{N}$ such that for all $n \ge n_{p,\varepsilon}$, a random linear code $\mathcal{C} \subseteq \mathbb{F}_2^n$ of rate $R = 1 - H(p) - \varepsilon$ is with probability $1 - \exp(-\Omega_{p,\varepsilon}(n))$ not $(p, \frac{1-\gamma_{p,\varepsilon}}{\varepsilon})$-list-decodable.*

The proof of Theorem 7 is obtained by tightening the second moment method proof of [15], and can be found in Appendix A of [26]. Theorem 7, combined with Theorem 5, implies that, for all $p \in (0,1/2)$ and sufficiently small ε, random linear codes of rate $1 - H(p) - \varepsilon$ with high probability can be list-decoded up to distance p with smaller list sizes than uniformly random codes. Perhaps surprisingly, the difference between the list size upper bound in Theorem 1 and the lower bound in Theorem 7 is bounded by 1 as $\varepsilon \to 0$, implying that the "correct" list size of a uniformly random code is tightly concentrated between $\lfloor 1/\varepsilon \rfloor \pm 1$ for small ε.

We are unaware of results in the literature that give even the existence of binary codes list-decodable with list size *better* than $H(p)/\varepsilon$. We remark that the Lovasz Local Lemma also gives the *existence* of $(p, H(p)/\varepsilon)$-list-decodable codes, matching our high probability result for random linear codes. We refer the reader to Appendix C of [26] for the details.

2.4 Prior work: rank metric codes

As an application of our techniques for random linear codes, we turn our attention to *rank metric codes*. Rank metric codes, introduced by Delsarte in [3], are codes $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$; that is, the codewords are $m \times n$ matrices, where typically $m \geq n$. The distance between two codewords X and Y is given by the rank of their difference: $\Delta_R(X,Y) := \frac{1}{n}\operatorname{rank}(X-Y)$, where Δ_R is called the *rank metric*. We denote the *rank ball* by $\mathcal{B}_{q,R}(X,p) := \{Y \in \mathbb{F}_q^{m\times n} : \Delta_R(X<Y) \leq pn\}$, and say that a rank metric code $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ is (p,L)-list-decodable if $|\mathcal{B}_{q,R}(X,p) \cap \mathcal{C}| \leq L$ for all $X \in \mathbb{F}_q^{m\times n}$. The *rate* R of a rank metric code $\mathcal{C} \subset \mathbb{F}_q^{m\times n}$ is $R := \log_q(|\mathcal{C}|)/(mn)$.

Rank metric codes generalize standard (Hamming metric) codes, which are simply diagonal rank metric codes. The study of rank metric codes has been motivated by a wide range of applications, including magnetic storage [31], cryptography [8, 27, 28], space-time coding [30, 29], and network coding [25, 39], and distributed storage [38, 37].

The natural "list-decoding capacity" for rank metric codes is $R = (1-p)(1-(n/m)p)$, which is the analog of the Gilbert-Varshamov bound [10]. It was shown in [4, 16] that this is achievable by a uniformly random rank metric code.

▶ **Theorem 8** ([16], Proposition A.1.[6]). *Let $\varepsilon > 0$ and $p \in (0,1)$ and suppose that m, n are sufficiently large compared to $1/\varepsilon$. A uniformly random code $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ of rate $R = (1-p)(1-bp) - \varepsilon$ is $(p, \lceil 1/\varepsilon \rceil)$-list-decodable with probability at least $1 - O(q^{-\varepsilon mn})$, where $b = n/m$.*

Moreover, it is shown in [4] that no linear code can beat this bound.

▶ **Theorem 9.** *Let $b = \lim_{n\to\infty} \frac{n}{m}$ be a constant. Then for any $R \in (0,1)$ and $p \in (0,1)$, a (p,L)-list-decodable $\mathbb{F}_q$-linear rank metric code $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ with rate R satisfies $R \leq (1-p)(1-bp)$.*

There has been a great deal of work aimed at establishing (or refuting) the list-decodability of explicit rank metric codes. It is shown in [42] that Gabidulin codes [9] – the rank-metric analog of Reed-Solomon codes – are *not* list-decodable to capacity, or even much beyond half their minimum distance. However, there have been works [22, 20] designing explicit codes meeting the lower bound of Theorem 9.

As with standard (Hamming-metric) codes, it is interesting to study the list-decodability of random linear rank-metric codes; we say that $\mathcal{C} \subseteq \mathbb{F}_q^{n\times m}$ is linear if it forms an $\mathbb{F}_q$-linear space. Following the approach of [45] for Hamming metric codes, [4] shows that random linear rank metric codes of rate $R = (1-p)(1-bp) - \varepsilon$ are $(p, \exp(O(1/\varepsilon))$-list-decodable, where as above $b = n/m$. In a recent paper of Guruswami and Resch [16], this result was strengthened to give a list size of $O(1/\varepsilon)$.

▶ **Theorem 10** ([16]). *Let $p \in (0,1)$ and $q \geq 2$. There is some constant $C_{p,q}$ so that the following holds. For all sufficiently large n, m with $b = n/m$, a random linear rank metric code $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ of rate $R = (1-p)(1-bp) - \varepsilon$ is $(p, C_{p,q}/\varepsilon)$-list-decodable with high probability.*

The proof of Theorem 10 uses ideas from the approach of [12] to prove Theorem 2. This is a beautiful argument, but as with the results of [12], the result of [16] suffers from the fact that $C_{p,q}$ tends to ∞ as p approaches 1. It is shown in [16], Proposition A.2, that

[6] In [16], the result is stated with $L = O(1/\varepsilon)$, but an inspection of the proof shows that we may take the leading constant to be 1.

when $p = 1 - \eta$, a uniformly random rank metric code of rate $R = (\eta - \eta b + \eta^2 b)/2$ is $(p, 4/(\eta - \eta b + \eta^2 b))$-list-decodable, and that work poses the question of whether or not a random linear rank metric code can achieve this. Our results, described in the next section, show that the answer is "yes" for $q = 2$.

2.5 Our results: rank metric codes

By applying the techniques in the proof of Theorem 5, we prove the following upper bound on the list size of random linear binary rank-metric codes.

▶ **Theorem 11.** *Let $p \in (0,1)$ and $\varepsilon > 0$. There is a constant C_ε so that the following holds. Let m and n be sufficiently large positive integers with $n < m$ and let $b = n/m$. A random linear rank metric code $\mathcal{C} \subseteq \mathbb{F}_2^{m \times n}$ of rate $R = (1-p)(1-bp) - \varepsilon$ is $(p, \frac{p+bp-bp^2}{\varepsilon} + 2)$-list-decodable with probability at least $1 - \exp(-C_\varepsilon mn)$.*

Notice that Theorem 11 works for all p, improving upon Theorem 10. In particular, when $p = 1 - \eta$, then setting $\varepsilon = (1-p)(1-bp)/2$ and applying Theorem 11 implies that a random linear binary rank metric code of rate $R = (\eta - \eta b + \eta^2 b)/2$ is (p, L) list-decodable for $L \leq \frac{2}{\eta - \eta b + \eta^2 b}$, answering the aforementioned open question of [16] in the affirmative.

We also prove a new lower bound on the list size of uniformly random rank-metric codes.

▶ **Theorem 12.** *Let $p \in (0,1)$ and $\varepsilon > 0$. Suppose m, n are sufficiently large so that $b = n/m$. Let $\mathcal{C} \subseteq \mathbb{F}_q^{m \times n}$ be a uniformly random rank metric code of rate $R = (1-p)(1-bp) - \varepsilon$. Then $\mathcal{C}$ is $(p, (1-p)(1-bp)/\varepsilon - 1)$-list-decodable with probability at most $\exp(-\Omega_{p,\varepsilon}(n))$.*

Theorem 12 again uses the method of [15]. The proofs of Theorems 11 and 12 can be found in Section 5 and Appendix B, respectively, of [26]. Together, Theorems 11 and 12 show that for some values of p, random linear binary rank metric codes have a strictly smaller list size than uniformly random rank metric codes with the same parameters. In particular, the upper bound of Theorem 11 is strictly smaller than the lower bound of Theorem 12 whenever $p < \frac{1-b}{2}$. For larger values of p, we remark that the list size obtained by Theorem 11 is still strictly smaller than the $1/\varepsilon$ list size given by uniformly random codes in Theorem 8, even though in this case we don't have a lower bound which proves that this is tight.

3 Simplified result for random linear binary codes

In this section, we prove Theorem 5, which we restate here.

▶ **Theorem 13** (Theorem 5, restated)**.** *Let $p \in (0, 1/2)$, let $\varepsilon > 0$, and let $R = 1 - H(p) - \varepsilon$. Let $\mathcal{C} \subseteq \mathbb{F}_2^n$ be a random linear code of rate R. Then with probability $1 - \exp(-\Omega_\varepsilon(n))$, the code $\mathcal{C}$ is $(p, H(p)/\varepsilon + 2)$-list-decodable.*

Theorem 5 also follows from our more refined result, Theorem 6. However, since our techniques give a very simple proof of Theorem 5 on its own, we begin with just this simple proof. We start by reviewing the approach of [13], which is the basis of our proof.

3.1 The approach of [13]

Before anything was known about the list-decodability of a typical random linear code, Guruswami, Håstad, Sudan and Zuckerman [13] proved the *existence* of binary linear codes of rate $1 - H(p) - \varepsilon$ that are $(p, 1/\varepsilon)$-list-decodable. Their result followed from a beautiful potential-function argument, which is the basis of our approach and which we describe here.

Let $k := Rn = (1 - H(p) - \varepsilon)n$. We choose vectors $b_1, \ldots, b_k$ one at a time, so that the code $\mathcal{C}_i := \text{span}(b_1, \ldots, b_i)$ remains "nice": formally, so that a potential function $\tilde{S}_{\mathcal{C}_i}$ remains small. Once we have picked all k vectors, we set $\mathcal{C} = \mathcal{C}_k$, and the fact that $\tilde{S}_{\mathcal{C}_k}$ is small implies list-decodability.

Recall that for a code $\mathcal{C}$ and $x \in \mathbb{F}_2^n$, we set $L_{\mathcal{C}}(x) = |\mathcal{B}(x, pn) \cap \mathcal{C}|$. Define

$$\tilde{S}_{\mathcal{C}} := \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} 2^{\varepsilon n L_{\mathcal{C}}(x)}.$$

It is not hard to show that for any vectors $b_1, \ldots, b_i \in \mathbb{F}_2^n$,

$$\mathop{\mathbf{E}}_{b_{i+1} \sim \mathbb{F}_2^n} \left[\tilde{S}_{\mathcal{C}_i + \{0, b_{i+1}\}} | b_1, \ldots, b_i \right] \leq \tilde{S}_{\mathcal{C}_i}^2. \tag{1}$$

That is, when a uniformly random vector b_{i+1} is added to the basis $\{b_1, \ldots, b_i\}$, we expect the potential function not to grow too much. Hence, there exists a choice of vectors $b_1, \ldots, b_k$ so that $\tilde{S}_{\mathcal{C}_{i+1}} \leq \tilde{S}_{\mathcal{C}_i}^2$ for $i = 0, 1, \ldots, k-1$.[7]

As $\mathcal{C}_0 = \{0\}$, we have $\tilde{S}_{\mathcal{C}_0} \leq 1 + 2^{-n(1-H(p)-\varepsilon)}$. Setting $\mathcal{C} = \mathcal{C}_k = \text{span}(b_1, \ldots, b_k)$, we have

$$\tilde{S}_{\mathcal{C}} \leq \tilde{S}_{\mathcal{C}_0}^{2^k} \leq \left(1 + 2^{-n(1-H(p)-\varepsilon)}\right)^{2^k} \leq \exp\left(2^{k-n(1-H(p)-\varepsilon)}\right) \leq e$$

by our choice of k. This implies that $\sum_x 2^{\varepsilon n L_{\mathcal{C}}(x)} \leq e \cdot 2^n$, and in particular, for all $x \in \mathbb{F}_2^n$, we have $2^{\varepsilon n L_{\mathcal{C}}(x)} \leq e \cdot 2^n$. Thus, for all x, $L_{\mathcal{C}}(x) \leq \frac{1}{\varepsilon} + o(1)$, as desired.

The approach of [13] is extremely clever, but these ideas have not, to the best of our knowledge, been used in subsequent work on the list-decodability of random linear codes. One reason is that the crux of the argument, which is (1), holds in expectation, and it was not clear how to show that it holds with high probability; thus, the result remained existential, and other techniques were introduced to study typical random codes [12, 2, 43, 34, 36].

3.2 Proof of Theorem 5

We improve the argument of [13] in two ways. First, we show that in fact, (1) essentially holds with high probability over the choice of b_{i+1}, which allows us to use the approach sketched above for random linear codes. Second, we introduce one additional trick which takes advantage of the linearity of the code in order to reduce the constant in the list size from 1 to $H(p)$. Before diving into the details, we briefly describe the main ideas.

The first improvement follows from looking at the potential function in the right way. In this paragraph, all $o(1)$ terms are exponentially small in n. Our goal is $\tilde{S}_{\mathcal{C}_k} \leq O(1)$. Write $\tilde{S}_{\mathcal{C}_i} = 1 + \tilde{T}_{\mathcal{C}_i}$. By above, $\tilde{T}_{\mathcal{C}_0} = \tilde{S}_{\mathcal{C}_0} - 1 = o(1)$. We show that with high probability, for all $i \leq k$, we have $\tilde{T}_{\mathcal{C}_i} = o(1)$. In the [13] argument we have

$$\mathbf{E}\, \tilde{S}_{\mathcal{C}_{i+1}} \leq \tilde{S}_{\mathcal{C}_i}^2 = (1 + \tilde{T}_{\mathcal{C}_i})^2 = 1 + 2\tilde{T}_{\mathcal{C}_i}(1 + o(1)),$$

and so $\mathbf{E}\, \tilde{T}_{\mathcal{C}_{i+1}} = 2\tilde{T}_{\mathcal{C}_i}(1 + o(1))$. One can show that, always, $2\tilde{T}_{\mathcal{C}_i} \leq \tilde{T}_{\mathcal{C}_{i+1}}$. By Markov's inequality, $\tilde{T}_{\mathcal{C}_{i+1}} = 2\tilde{T}_{\mathcal{C}_i}(1 + o(1))$ with probability $1 - o(1)$, for appropriately chosen $o(1)$ terms. Union bounding over the $o(1)$ failure probabilities in the k steps, we conclude that $\tilde{T}_{\mathcal{C}_i}$ grows roughly as slowly as in the existential argument, giving the desired list-decodability.

[7] As a technical detail, one needs to be careful that $b_{i+1} \notin \mathcal{C}_i$. One can guarantee $b_{i+1} \notin \mathcal{C}_i$ by carefully examining the proof of (1), or use (1) to get a similar equation where we additionally condition $b_{i+1} \notin \mathcal{C}_i$.

The second improvement follows from the linearity of the code. In the last step of the [13] argument, we replace the summation "$\sum_x$" in $\sum_x 2^{\varepsilon n L_{\mathcal{C}}(x)} \le e \cdot 2^n$ with a "$\forall x$." We can save a bit because, by linearity, the contribution $2^{\varepsilon n L_{\mathcal{C}}(x)}$ is the same for all x in a coset $y + \mathcal{C}$.

Now we go through the details. It is convenient to change the definition of the potential function very slightly: losing the tilde, define, for a code $\mathcal{C} \subset \mathbb{F}_2^n$,

$$A_{\mathcal{C}}(x) \;:=\; 2^{\frac{\varepsilon n L_{\mathcal{C}}(x)}{1+\varepsilon}} \qquad \text{and} \qquad S_{\mathcal{C}} \;:=\; \mathop{\mathbf{E}}_{x \sim \mathbb{F}_2^n} \left[A_{\mathcal{C}}(x)\right].$$

The term $S_{\mathcal{C}}$ differs from the term $\tilde{S}_{\mathcal{C}}$ above in that $A_{\mathcal{C}}(x)$ has an extra factor of $\frac{1}{1+\varepsilon}$ in the exponent. This is an extra "slack" term that helps guarantee a high probability result under the same parameters. However, this definition does not change how the potential function behaves. In particular, we still have the following lemma:

▶ **Lemma 14** (Following [13]). *For all linear $\mathcal{C} \subseteq \mathbb{F}_2^n$ and all $b \in \mathbb{F}_2^n$,*

$$L_{\mathcal{C}+\{0,b\}}(x) \;\le\; L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b) \tag{2}$$
$$A_{\mathcal{C}+\{0,b\}}(x) \;\le\; A_{\mathcal{C}}(x) \cdot A_{\mathcal{C}}(x+b), \tag{3}$$

with equality if and only if $b \notin \mathcal{C}$.

Proof. To see (2), notice that

$$\begin{aligned} L_{\mathcal{C}+\{0,b\}}(x) \;&=\; |\mathcal{B}(x, pn) \cap (\mathcal{C} \cup (\mathcal{C}+b))| \\ &\le\; |\mathcal{B}(x, pn) \cap \mathcal{C}| + |\mathcal{B}(x, pn) \cap (\mathcal{C}+b)| \\ &=\; |\mathcal{B}(x, pn) \cap \mathcal{C}| + |\mathcal{B}(x+b, pn) \cap \mathcal{C}| \\ &=\; L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b), \end{aligned}$$

with equality in the second line if and only if $b \notin \mathcal{C}$. Inequality (3) follows as a consequence of (2), and this proves the lemma. ◀

We additionally define

$$B_{\mathcal{C}}(x) \;:=\; A_{\mathcal{C}}(x) - 1 \qquad \text{and} \qquad T_{\mathcal{C}} \;:=\; S_{\mathcal{C}} - 1.$$

As noted above, it is helpful to think of $T_{\mathcal{C}}$ as a very small term; we would like to show – in accordance with (1) – that $T_{\mathcal{C}}$ approximately doubles each time we add a basis vector. Note that

$$S_{\{0\}} \;=\; 1 + \left(2^{\frac{\varepsilon n}{1+\varepsilon}} - 1\right) \cdot \frac{\mathrm{Vol}(n, pn)}{2^n} \;<\; 1 + 2^{\frac{\varepsilon n}{1+\varepsilon}} \cdot \frac{2^{H(p)n}}{2^n} \;=\; 1 + 2^{-n(1-H(p)-\frac{\varepsilon}{1+\varepsilon})}. \tag{4}$$

With these definitions, we can prove the concentration result we need.

▶ **Lemma 15.** *Let p, ε, and $R = 1 - H(p) - \varepsilon$ be as in the statement of Theorem 5. Let $\mathcal{C}_{Rn} \subset \mathbb{F}_2^n$ be a random linear code of rate R. Then with probability $1 - \exp(-\Omega_\varepsilon(n))$, the code $\mathcal{C}_{Rn}$ satisfies $S_{\mathcal{C}_{Rn}} \le 2$.*

Before we prove Lemma 15 we show that this implies Theorem 5.

Proof of Theorem 5 given Lemma 15. We show that, for a binary linear code $\mathcal{C}$ of rate $R = 1 - H(p) - \varepsilon$, $S_{\mathcal{C}} \le 2$ implies $(p, \frac{H(p)}{\varepsilon} + 2)$-list-decodability. Suppose for sake of contradiction

that a code $\mathcal{C}$ satisfies $S_{\mathcal{C}} \leq 2$ and there exists $x^* \in \mathbb{F}_2^n$ such that $|\mathcal{B}(x^*, pn) \cap \mathcal{C}| > H(p)/\varepsilon + 2$. For all $x \in \mathbb{F}_2^n$ and $c \in \mathcal{C}$, we have

$$|\mathcal{B}(x+c, pn) \cap \mathcal{C}| = |\mathcal{B}(x, pn) \cap (\mathcal{C} - c)| = |\mathcal{B}(x, pn) \cap \mathcal{C}|,$$

so $|\mathcal{B}(x^* + c, pn) \cap \mathcal{C}| > H(p)/\varepsilon$ for all $c \in \mathcal{C}$. If $S_{\mathcal{C}} \leq 2$, then we have

$$\begin{aligned}
2^{n+1} \geq 2^n S_{\mathcal{C}} &= \sum_{x \in \mathbb{F}_2^n} \exp_2\left(n \cdot \frac{\varepsilon}{1+\varepsilon} \cdot |\mathcal{B}(x, pn) \cap \mathcal{C}|\right) \\
&> \sum_{c \in \mathcal{C}} \exp_2\left(n \cdot \frac{\varepsilon}{1+\varepsilon} \cdot |\mathcal{B}(x^* + c, pn) \cap \mathcal{C}|\right) \\
&\geq \sum_{c \in \mathcal{C}} \exp_2\left(n \cdot \frac{\varepsilon}{1+\varepsilon} \cdot (H(p)/\varepsilon + 2)\right) \\
&= |\mathcal{C}| \cdot \exp_2\left(n \cdot \frac{H(p) + 2\varepsilon}{1+\varepsilon}\right) \\
&= \exp_2\left(n\left(1 + \frac{\varepsilon}{1+\varepsilon}(1 - H(p) - \varepsilon)\right)\right),
\end{aligned}$$

where the first inequality is because we sum over strictly fewer terms, and the second inequality is by definition of x^*. This is a contradiction for large enough n. ◀

Finally, we prove Lemma 15.

Proof of Lemma 15. As in §3.1, let $b_1, b_2, \ldots, b_k \in \mathbb{F}_2^n$ be independently and uniformly chosen, and let $\mathcal{C}_i = \operatorname{span}\{b_1, \ldots, b_i\}$.

▶ **Lemma 16.** *Suppose that $\mathcal{C}$ is fixed and satisfies $T_{\mathcal{C}} < 1$, so that $S_{\mathcal{C}} < 2$. Then*

$$\Pr_{b \sim \mathbb{F}_2^n}\left[S_{\mathcal{C}+\{0,b\}} > 1 + 2T_{\mathcal{C}} + T_{\mathcal{C}}^{1.5}\right] < T_{\mathcal{C}}^{0.5}.$$

Proof. By Lemma 14, for all b,

$$\begin{aligned}
S_{\mathcal{C}+\{0,b\}} &= \mathop{\mathbf{E}}_x\left[A_{\mathcal{C}+\{0,b\}}(x)\right] \\
&\leq \mathop{\mathbf{E}}_x\left[A_{\mathcal{C}}(x) A_{\mathcal{C}}(x+b)\right] \\
&= \mathop{\mathbf{E}}_x\left[(1 + B_{\mathcal{C}}(x)) \cdot (1 + B_{\mathcal{C}}(x+b))\right] \\
&= \mathop{\mathbf{E}}_x\left[1 + B_{\mathcal{C}}(x) + B_{\mathcal{C}}(x+b) + B_{\mathcal{C}}(x) B_{\mathcal{C}}(x+b)\right] \\
&= 1 + 2T_{\mathcal{C}} + \mathop{\mathbf{E}}_x\left[B_{\mathcal{C}}(x) B_{\mathcal{C}}(x+b)\right].
\end{aligned}$$

Over the randomness of b and x, we have x and $x+b$ are statistically independent and uniform over $\mathbb{F}_2^n$, so we have

$$\mathop{\mathbf{E}}_b \mathop{\mathbf{E}}_x\left[B_{\mathcal{C}}(x) B_{\mathcal{C}}(x+b)\right] = \mathop{\mathbf{E}}_{b,x}\left[B_{\mathcal{C}}(x)\right] \cdot \mathop{\mathbf{E}}_{b,x}\left[B_{\mathcal{C}}(x+b)\right] = T_{\mathcal{C}}^2. \tag{5}$$

As $B_{\mathcal{C}}$ is always nonnegative, we have, by Markov's inequality,

$$\Pr_b\left[S_{\mathcal{C}+\{0,b\}} > 1 + 2T_{\mathcal{C}} + T_{\mathcal{C}}^{1.5}\right] \leq \Pr_b\left[\mathop{\mathbf{E}}_x\left[B_{\mathcal{C}}(x) B_{\mathcal{C}}(x+b)\right] > T_{\mathcal{C}}^{1.5}\right] < \frac{T_{\mathcal{C}}^2}{T_{\mathcal{C}}^{1.5}} = T_{\mathcal{C}}^{0.5}. \blacktriangleleft$$

Returning to the proof of Lemma 15, consider the sequence

$$\begin{aligned}\delta_0 &:= 2^{-n(1-H(p)-\frac{\varepsilon}{1+\varepsilon})}\\ \delta_i &:= 2\delta_{i-1} + \delta_{i-1}^{1.5}.\end{aligned}$$

We prove by induction that, for $0 \le i \le n(1-H(p)-\varepsilon)$, we have $\delta_i < 2^{i+1}\delta_0$, which is at most $2^{-\frac{\varepsilon^2 n}{2}}$ for n sufficiently large. The base case $i = 0$ is straightforward. If $\delta_j < 2^{j+1}\delta_0$ for $j < i$, then

$$\delta_i \;=\; 2\delta_{i-1}(1+\delta_{i-1}^{0.5}) \;=\; 2^i\delta_0 \cdot \prod_{j=0}^{i-1}(1+\delta_j^{0.5}) \;\le\; 2^i\delta_0 \cdot \exp\left(\sum_{j=0}^{i-1}\delta_j^{0.5}\right) \;<\; 2^{i+1}\delta_0.$$

In the first two equalities, we applied the definitions of δ_i and $\delta_{i-1}, \dots, \delta_1$, respectively. In the first inequality, we used the estimate $1 + z \le e^z$, and in the second we used the inductive hypothesis $\delta_j < 2^{-\frac{\varepsilon^2 n}{2}}$ for $j < i$. By this induction, we conclude that, if $k = n(1-H(p)-\varepsilon)$, then $\delta_k < 2^{-\frac{\varepsilon^2 n}{2}}$.

Let $b_1, \dots, b_k \in \mathbb{F}_2^n$ be randomly chosen vectors, and let $\mathcal{C}_i = \operatorname{span}(b_1, \dots, b_i)$ with $\mathcal{C}_k = \mathcal{C}$. By Lemma 16, conditioned on a fixed $\mathcal{C}_i$ satisfying $T_{\mathcal{C}_i} \le \delta_i$, we have, with probability at most $T_{\mathcal{C}_i}^{0.5}$, which is at most $\delta_i^{0.5}$, that $T_{\mathcal{C}_{i+1}} > \delta_{i+1}$. Furthermore, $T_{\mathcal{C}_0} \le \delta_0$ by the initial condition (4). Thus, with probability at least

$$1 - \left(\delta_0^{0.5} + \delta_1^{0.5} + \cdots + \delta_k^{0.5}\right) \;>\; 1 - k2^{-\varepsilon^2 n/2} \;\ge\; 1 - 2^{-\Omega_\varepsilon(n)}$$

we have $T_{\mathcal{C}_i} \le \delta_i$ for all i. In particular, $T_{\mathcal{C}} = T_{\mathcal{C}_k} < \delta_k < 2^{-\frac{\varepsilon^2 n}{2}}$. Thus, $S_{\mathcal{C}} = 1 + T_{\mathcal{C}} \le 2$ with probability $1 - \exp(-\Omega_\varepsilon(n))$, completing the proof of Lemma 15. ◀

▶ Remark. We do not see how to extend this proof to larger alphabets. If, for example, $q = 3$, then Lemma 16 would need to say $\mathbf{Pr}[S_{C+\{0,b,2b\}} > 1 + 3T_C + o(T_C)] < o(1)$. However, the same proof would fail to establish this, as we can no longer separate the expectation in (5); that is we cannot say

$$\mathop{\mathbf{E}}_{b}\mathop{\mathbf{E}}_{x}\left[B_{\mathcal{C}}(x)B_{\mathcal{C}}(x+b)B_{\mathcal{C}}(x+2b)\right] = \mathop{\mathbf{E}}_{b,x}\left[B_{\mathcal{C}}(x)\right] \cdot \mathop{\mathbf{E}}_{b,x}\left[B_{\mathcal{C}}(x+b)\right] \cdot \mathop{\mathbf{E}}_{b,x}\left[B_{\mathcal{C}}(x+2b)\right] = T_{\mathcal{C}}^3.$$

4 Conclusion

In this work, we have given an improved analysis of the list-decodability of random linear binary codes. Our analysis works for all values of p, and also obtains improved bounds on the list size as the rate approaches list-decoding capacity. In particular, not only do our bounds improve on previous work for random linear codes, but they show that random linear codes are more list-decodable than completely random codes, in the sense that the list size is strictly smaller. Our techniques are quite simple, and strengthen an argument of [13] to hold with high probability. In order to demonstrate the applicability of these techniques, we use them to (a) obtain more information about the distribution of list sizes of random linear codes and (b) to prove a similar result for random linear rank-metric codes, improving a recent result of [16].

We end with some open questions raised by our work.

1. With the exception of Theorem 12, our results – both our upper bounds and our lower bounds – hold only for binary alphabets. We conjecture that analogous results, and in particular list-decoding random linear codes with list size C/ε for $C < 1$, hold over larger alphabets.

2. We showed that random linear binary codes of rate $1 - H(p) - \varepsilon$ are with high probability (p, L) list-decodable with $L \leq H(p)/\varepsilon$. The lower bounds of [32, 15] show that we must have $L \geq C/\varepsilon$, but the constant C is much smaller than $H(p)$. Thus, we still do not know what the correct leading constant is for random linear codes.
3. Finally, there are currently no known explicit constructions of capacity-achieving binary list-decodable codes for general p. It is our hope that this work – which gives more information about the structure of linear codes which achieve list-decoding capacity – could lead to progress on this front. Given that we don't know how to efficiently check if a given code is (p, L)-list-decodable, even an efficient Las Vegas construction (as opposed to a Monte Carlo construction) would be interesting.

References

1 Volodia M. Blinovsky. Bounds for codes in the case of list decoding of finite volume. *Problems of Information Transmission*, 22(1):7–19, 1986.

2 Mahdi Cheraghchi, Venkatesan Guruswami, and Ameya Velingker. Restricted isometry of Fourier matrices and list decodability of random linear codes. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 432–442. ACM-SIAM, 2013. doi:10.1137/1.9781611973105.31.

3 Philippe Delsarte. Bilinear forms over a finite field, with applications to coding theory. *Journal of Combinatorial Theory, Series A*, 25(3):226–241, 1978.

4 Yang Ding. On list-decodability of random rank metric codes and subspace codes. *IEEE Trans. Information Theory*, 61(1):51–59, 2015. doi:10.1109/TIT.2014.2371915.

5 Zeev Dvir and Shachar Lovett. Subspace evasive sets. In *Proceedings of the Forty-Fourth annual ACM Symposium on Theory of Computing (STOC)*, pages 351–358. ACM, 2012.

6 Peter Elias. List decoding for noisy channels. *Wescon Convention Record, Part 2*, pages 94–104, 1957.

7 Peter Elias. Error-correcting codes for list decoding. *IEEE Trans. Information Theory*, 37(1):5–12, 1991.

8 Ernst M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov. Ideals over a non-commutative ring and their applications in cryptology. In *Proceedings of Advances in Cryptology - EUROCRYPT '91, Workshop on the Theory and Application of of Cryptographic Techniques*, pages 482–489, 1991. doi:10.1007/3-540-46416-6_41.

9 Ernst Mukhamedovich Gabidulin. Theory of codes with maximum rank distance. *Problemy Peredachi Informatsii*, 21(1):3–16, 1985.

10 Maximilien Gadouleau and Zhiyuan Yan. On the decoder error probability of bounded rank-distance decoders for maximum rank-distance codes. *IEEE Trans. Information Theory*, 54(7):3202–3206, 2008. doi:10.1109/TIT.2008.924697.

11 Venkatesan Guruswami. List decoding of binary codes-a brief survey of some recent results. In *Proceedings of Coding and Cryptology, Second International Workshop (IWCC)*, pages 97–106, 2009. doi:10.1007/978-3-642-01877-0_10.

12 Venkatesan Guruswami, Johan Håstad, and Swastik Kopparty. On the list-decodability of random linear codes. *IEEE Trans. Information Theory*, 57(2):718–725, 2011. doi:10.1109/TIT.2010.2095170.

13 Venkatesan Guruswami, Johan Håstad, Madhu Sudan, and David Zuckerman. Combinatorial bounds for list decoding. *IEEE Trans. Information Theory*, 48(5):1021–1034, 2002. doi:10.1109/18.995539.

14 Venkatesan Guruswami and Piotr Indyk. Efficiently decodable codes meeting gilbert-varshamov bound for low rates. In *Proceedings of the fifteenth annual ACM-SIAM sym-*

posium on Discrete algorithms, pages 756–757. Society for Industrial and Applied Mathematics, 2004.

15 Venkatesan Guruswami and Srivatsan Narayanan. Combinatorial limitations of average-radius list-decoding. *IEEE Trans. Information Theory*, 60(10):5827–5842, 2014. doi:10.1109/TIT.2014.2343224.

16 Venkatesan Guruswami and Nicolas Resch. On the list-decodability of random linear rank-metric codes. *arXiv preprint arXiv:1710.11516*, 2017.

17 Venkatesan Guruswami and Atri Rudra. Concatenated codes can achieve list-decoding capacity. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*, pages 258–267, 2008. URL: http://dl.acm.org/citation.cfm?id=1347082.1347111.

18 Venkatesan Guruswami and Atri Rudra. Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy. *IEEE Trans. Information Theory*, 54(1):135–150, 2008.

19 Venkatesan Guruswami and Salil Vadhan. A lower bound on list size for list decoding. In Chandra Chekuri, Klaus Jansen, José D. P. Rolim, and Luca Trevisan, editors, *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 318–329, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

20 Venkatesan Guruswami, Carol Wang, and Chaoping Xing. Explicit list-decodable rank-metric and subspace codes via subspace designs. *IEEE Trans. Information Theory*, 62(5):2707–2718, 2016.

21 Venkatesan Guruswami and Chaoping Xing. Folded codes from function field towers and improved optimal rate list decoding. In *Proceedings of the Forty-Fourth annual ACM Symposium on Theory of Computing (STOC)*, pages 339–350. ACM, 2012.

22 Venkatesan Guruswami and Chaoping Xing. List decoding Reed-Solomon, Algebraic-Geometric, and Gabidulin subcodes up to the Singleton bound. In *Proceedings of the Forty-Fifth annual ACM Symposium on Theory of Computing (STOC)*, pages 843–852. ACM, 2013.

23 Brett Hemenway, Noga Ron-Zewi, and Mary Wootters. Local list recovery of high-rate tensor codes & applications. In *58th Annual IEEE Symposium on Foundations of Computer Science*, 2017.

24 Brett Hemenway and Mary Wootters. Linear-time list recovery of high-rate expander codes. In *International Colloquium on Automata, Languages, and Programming*, pages 701–712. Springer, 2015.

25 Ralf Koetter and Frank R Kschischang. Coding for errors and erasures in random network coding. *IEEE Trans. Information Theory*, 54(8):3579–3591, 2008.

26 Ray Li and Mary Wootters. Improve list-decodability of random linear binary code. *CoRR*, abs/1801.07839, 2018. arXiv:1801.07839.

27 Pierre Loidreau. Designing a rank metric based mceliece cryptosystem. In *Proceedings of the Post-Quantum Cryptography, Third International Workshop on Post-Quantum Cryptography, (PQCrypto)*, pages 142–152, 2010. doi:10.1007/978-3-642-12929-2_11.

28 Pierre Loidreau. A new rank metric codes based encryption scheme. In *Proceedings of the Post-Quantum Cryptography, 8th International Workshop on Post-Quantum Cryptography, (PQCrypto)*, pages 3–17, 2017. doi:10.1007/978-3-319-59879-6_1.

29 Hsiao-feng Lu and P. Vijay Kumar. A unified construction of space-time codes with optimal rate-diversity tradeoff. *IEEE Trans. Information Theory*, 51(5):1709–1730, 2005. doi:10.1109/TIT.2005.846403.

30 P. Lusina, Ernst M. Gabidulin, and Martin Bossert. Maximum rank distance codes as space-time codes. *IEEE Trans. Information Theory*, 49(10):2757–2760, 2003. doi:10.1109/TIT.2003.818023.

31 Ron M. Roth. Maximum-rank array codes and their application to crisscross error correction. *IEEE Trans. Information Theory*, 37(2):328–336, 1991. doi:10.1109/18.75248.

32 Atri Rudra. Limits to list decoding of random codes. *IEEE Trans. Information Theory*, 57(3):1398–1408, 2011. doi:10.1109/TIT.2010.2054750.

33 Atri Rudra and Steve Uurtamo. Two theorems on list decoding. In Maria Serna, Ronen Shaltiel, Klaus Jansen, and José Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 696–709, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

34 Atri Rudra and Mary Wootters. Every list-decodable code for high noise has abundant near-optimal rate puncturings. In *Proceedings of the Forty-Sixth annual ACM Symposium on Theory of Computing (STOC)*, pages 764–773. ACM, 2014.

35 Atri Rudra and Mary Wootters. It'll probably work out: Improved list-decoding through random operations. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 287–296. ACM, 2015. doi:10.1145/2688073.2688092.

36 Atri Rudra and Mary Wootters. Average-radius list-recovery of random linear codes. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. ACM-SIAM, 2018.

37 Natalia Silberstein, Ankit Singh Rawat, O Ozan Koyluoglu, and Sriram Vishwanath. Optimal locally repairable codes via rank-metric codes. In *Proceedings of the 2013 IEEE International Symposium on Information Theory (ISIT)*, pages 1819–1823. IEEE, 2013.

38 Natalia Silberstein, Ankit Singh Rawat, and Sriram Vishwanath. Error resilience in distributed storage via rank-metric codes. In *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1150–1157. IEEE, 2012.

39 Danilo Silva, Frank R Kschischang, and Ralf Koetter. A rank-metric approach to error control in random network coding. *IEEE Trans. Information Theory*, 54(9):3951–3967, 2008.

40 Madhu Sudan. List decoding: algorithms and applications. *SIGACT News*, 31(1):16–27, 2000. doi:10.1145/346048.346049.

41 Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1-3):1–336, 2012. doi:10.1561/0400000010.

42 Antonia Wachter-Zeh. Bounds on list decoding of rank-metric codes. *IEEE Trans. Information Theory*, 59(11):7268–7277, 2013.

43 Mary Wootters. On the list decodability of random linear codes with large error rates. In *Proceedings of the Forty-Fifth Symposium on Theory of Computing Conference (STOC)*, pages 853–860, 2013. doi:10.1145/2488608.2488716.

44 Jack Wozencraft. List decoding. *Quarter Progress Report*, 48:90–95, 1958.

45 Victor Vasilievich Zyablov and Mark Semenovich Pinsker. List concatenated decoding. *Problemy Peredachi Informatsii*, 17(4):29–33, 1981.

A Characterizing the list size distribution

In this section we establish Theorem 6. Define

$$\begin{aligned} P_{\mathcal{C}}^{(\ell)} &:= 2^{-n}|\{x : L_{\mathcal{C}}(x) = \ell\}| \\ P_{\mathcal{C}}^{(\geq \ell)} &:= \sum_{i=\ell}^{\infty} P_{\mathcal{C}}^{(i)} \\ Q_{\mathcal{C}}^{(\geq \ell)} &:= \sum_{i=\ell}^{\infty} i \cdot P_{\mathcal{C}}^{(i)} \end{aligned}$$

The goal (per Theorem 6) is to bound $P_{\mathcal{C}}^{(\geq \ell)}$; in our argument, it is more convenient to work with $Q_{\mathcal{C}}^{(\geq \ell)}$, which is a proxy for $P_{\mathcal{C}}^{(\geq \ell)}$. We note a few useful properties of these definitions.

▶ **Proposition 17.** *Suppose that $\mathcal{C} \subseteq \mathbb{F}_2^n$ is a linear code. Then the following hold for $\ell \geq 1$.*
1. $Q_{\mathcal{C}}^{(\geq \ell)} = 2^{-n} \cdot \sum_{x \in \mathbb{F}_2^n : L_{\mathcal{C}}(x) \geq \ell} L_{\mathcal{C}}(x) = \mathbf{E}_x \left[\mathbb{I}\left[L_{\mathcal{C}}(x) \geq \ell \right] \cdot L_{\mathcal{C}}(x) \right]$
2. $Q_{\mathcal{C}}^{(\geq 1)} = 2^{-n} \cdot \mathrm{Vol}(n, pn) \cdot 2^k \leq 2^{-n(1-H(p))+k}$

Proof. To see Item 1, notice that

$$\begin{aligned} Q_{\mathcal{C}}^{(\geq \ell)} &= 2^{-n} \sum_{i=\ell}^{\infty} i \cdot |\{x \mid L_{\mathcal{C}}(x) = i\}| \\ &= 2^{-n} \sum_{i=\ell}^{\infty} \sum_{x \in \mathbb{F}_2^n} i \cdot \mathbb{I}\left[L_{\mathcal{C}}(x) = i \right] \\ &= 2^{-n} \sum_{x : L_{\mathcal{C}}(x) \geq \ell} L_{\mathcal{C}}(x). \end{aligned}$$

To see Item 2, we begin with Item 1 and derive

$$\begin{aligned} Q_{\mathcal{C}}^{(\geq 1)} &= 2^{-n} \sum_{x : L_{\mathcal{C}}(x) \geq 1} L_{\mathcal{C}}(x) \\ &= 2^{-n} \sum_{x \in \mathbb{F}_2^n} L_{\mathcal{C}}(x) \\ &= 2^{-n} \sum_{x \in \mathbb{F}_2^n} \sum_{v \in \mathcal{B}(x, pn)} \mathbb{I}[x + v \in \mathcal{C}] \\ &= 2^{-n} \sum_{v \in \mathcal{B}(x, pn)} \sum_{x \in \mathbb{F}_2^n} \mathbb{I}[x + v \in \mathcal{C}] \\ &= 2^{-n} \sum_{v \in \mathcal{B}(x, pn)} |\mathcal{C}| \\ &= 2^{-n} \cdot \mathrm{Vol}(n, pn) \cdot 2^k. \end{aligned}$$

◀

Using these properties, we are now can state and prove Theorem 18, which implies Theorem 6.

▶ **Theorem 18.** *Fix $L \geq 0$. There exists a constant C_L depending only on L so that, for all $\gamma \in (0,1)$ and sufficiently large n, the following holds. Suppose that $k \leq (1-\gamma)n$. If $\mathcal{C} \subseteq \mathbb{F}_2^n$ is a random linear code of dimension k, then with probability $1 - \exp(-C_L \gamma n)$, for all $1 \leq \ell \leq L$,*

$$Q_{\mathcal{C}}^{(\geq \ell)} \leq \left(2^{-n(1-H(p))} \cdot 2^k \right)^{\ell} \cdot 2^{\gamma \ell^2 n}. \tag{6}$$

Before we prove Theorem 18, we explain why it implies Theorem 6. By setting $k = n(1 - H(p) - \varepsilon)$ and $\gamma = \varepsilon\gamma'/L$ in Theorem 18, we obtain that with high probability, for all $1 \leq \ell \leq L$,

$$P_{\mathcal{C}}^{(\geq \ell)} \leq Q_{\mathcal{C}}^{(\geq \ell)} \leq 2^{-n\varepsilon\ell} \cdot 2^{\frac{\varepsilon\gamma'}{L}\ell^2 n} \leq 2^{-n\varepsilon\ell(1-\gamma')},$$

which is Theorem 6. Theorem 18 also implies Theorem 5:

▶ **Corollary 19** (Theorem 5). *Let $p \in (0, 1/2)$, let $\varepsilon > 0$, and let $R = 1 - H(p) - \varepsilon$. Let $\mathcal{C} \subseteq \mathbb{F}_2^n$ be a random linear code of rate R. Then with probability $1 - \exp(-\Omega_\varepsilon(n))$, the code $\mathcal{C}$ is $(p, H(p)/\varepsilon + 2)$-list-decodable.*

Proof. Let $L = H(p)/\varepsilon + 2$ and choose $\gamma \ll L^{-3}$, and let $\mathcal{C} \subseteq \mathbb{F}_2^n$ be a random linear code of dimension $k = n(1 - H(p) - \varepsilon)$. Then $k \leq (1-\gamma)n$, so, by Theorem 18, with probability at least $1 - \exp(-C_L \gamma n)$, code $\mathcal{C}$ satisfies (6) with $k = n(1 - H(p) - \varepsilon)$. In particular, choosing $\ell = L$, this means that

$$Q_{\mathcal{C}}^{(\geq L)} \;\leq\; 2^{L(-n(1-H(p))+k)+\gamma L^2 n} \;=\; 2^{-nL\varepsilon+\gamma L^2 n} \;<\; 2^{-n(1-R)}$$

Suppose there exists $x \in \mathbb{F}_2^n$ such that $L_{\mathcal{C}}(x) \geq L$. Then $L_{\mathcal{C}}(x+c) \geq L$ for all $c \in C$, so that $Q_{\mathcal{C}}^{(\geq L)} > P_{\mathcal{C}}^{(\ell)} \geq 2^{-n(1-R)}$. This is a contradiction, so there are no $x \in \mathbb{F}_2^n$ such that $L_{\mathcal{C}}(x) \geq L$. ◀

Proof of Theorem 18. The proof of Theorem 18 proceeds by induction on k, with the inductive hypothesis that (6) holds for all $1 \leq \ell \leq L$. We begin with the base case by noting that (6) is satisfied for $k = 0$, for all $1 \leq \ell \leq L$. To see this, notice that for $\ell = 1$, we have $Q_{\{0\}}^{(\geq 1)} \leq 2^{-n(1-H(p))}$ by Proposition 17, Item 2, which satisfies (6) for $k = 0, \ell = 1$. For $\ell \geq 2$, $Q_{\{0\}}^{(\geq \ell)} = 0$, and so again (6) holds.

Now that we have established the base case of $k = 0$, we proceed by induction. Lemma 20 provides the inductive step; similar to the approach in §3, it shows that at every step the "expected behavior" holds with high probability.

▶ **Lemma 20.** *Let $\gamma > 0$, and suppose that $\mathcal{C} \subseteq \mathbb{F}_2^n$ is a linear code of dimension $k \leq (1-\gamma)n$ such that for all $1 \leq \ell \leq L$, we have*

$$Q_{\mathcal{C}}^{(\geq \ell)} \;\leq\; \left(2^{-n(1-H(p))} \cdot 2^k\right)^{\ell} \cdot 2^{\gamma \ell^2 n}. \tag{7}$$

Then, for a uniformly chosen $b \in \mathbb{F}_2^n$, with probability at least $1 - 4 \cdot L^3 \cdot 2^{-\gamma n}$ over the choice of b, we have, for all $1 \leq \ell \leq L$,

$$Q_{\mathcal{C}+\{0,b\}}^{(\geq \ell)} \;\leq\; \left(2^{-n(1-H(p))} \cdot 2^{k+1}\right)^{\ell} \cdot 2^{\gamma \ell^2 n}. \tag{8}$$

Proof. By Proposition 17, Item 2, (8) always holds for $\ell = 1$, so suppose that $\ell \geq 2$. We have, for any $b \in \mathbb{F}_2^n$,

$$\begin{aligned}
Q_{\mathcal{C}+\{0,b\}}^{(\geq \ell)} &= \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}+\{0,b\}}(x) \geq \ell] \cdot L_{\mathcal{C}+\{0,b\}}(x)\right] && \text{(By Prop. 17, Item 1)} \\
&\leq \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b) \geq \ell] \cdot (L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b))\right] && \text{(By Lemma 14)} \\
&= \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b) \geq \ell] \cdot L_{\mathcal{C}}(x) + \mathbb{I}[L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b) \geq \ell] \cdot L_{\mathcal{C}}(x+b)\right] \\
&= 2 \cdot \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}}(x) + L_{\mathcal{C}}(x+b) \geq \ell] \cdot L_{\mathcal{C}}(x)\right] \\
&= 2 \cdot \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}}(x) \geq \ell] \cdot L_{\mathcal{C}}(x) + \sum_{i=0}^{\ell-1} \mathbb{I}[L_{\mathcal{C}}(x) = i, L_{\mathcal{C}}(x+b) \geq \ell - i] \cdot L_{\mathcal{C}}(x)\right] \\
&= 2 \cdot Q_{\mathcal{C}}^{(\geq \ell)} + 2 \cdot \sum_{i=0}^{\ell-1} i \cdot \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}}(x) = i] \cdot \mathbb{I}[L_{\mathcal{C}}(x+b) \geq \ell - i]\right], && (9)
\end{aligned}$$

where we have used Proposition 17, Item 1 in the final line. Since the only inequality above is an application of Lemma 14, which is an equality when $b \notin \mathcal{C}$, the derivation above is an equality when $b \notin \mathcal{C}$. Thus, when $b \notin \mathcal{C}$, we have

$$Q_{\mathcal{C}+\{0,b\}}^{(\geq \ell)} = 2 \cdot Q_{\mathcal{C}}^{(\geq \ell)} + 2 \cdot \sum_{i=0}^{\ell-1} i \cdot \mathop{\mathbf{E}}_x \left[\mathbb{I}[L_{\mathcal{C}}(x) = i] \cdot \mathbb{I}[L_{\mathcal{C}}(x+b) \geq \ell - i]\right] \geq 2 \cdot Q_{\mathcal{C}}^{(\geq \ell)}. \tag{10}$$

This fact (10) is useful later. For now, we move on with no assumption on b. Taking expectations on both sides of (9), we see

$$\begin{aligned}
\mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}}\right] &\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot \sum_{i=0}^{\ell-1} i \cdot \mathop{\mathbf{E}}_{b,x}\left[\mathbb{I}[L_{\mathcal{C}}(x)=i]\cdot \mathbb{I}[L_{\mathcal{C}}(x+b)\geq \ell - i]\right] \\
&= 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot \sum_{i=0}^{\ell-1} i \cdot P^{(i)}_{\mathcal{C}}\cdot P^{(\geq \ell-i)}_{\mathcal{C}} \\
&\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot \sum_{i=0}^{\ell-1} i \cdot Q^{(\geq i)}_{\mathcal{C}}\cdot Q^{(\geq \ell-i)}_{\mathcal{C}}
\end{aligned}$$

Continuing, we bound

$$\begin{aligned}
\mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}}\right] &\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot \sum_{i=0}^{\ell-1} i \cdot Q^{(\geq i)}_{\mathcal{C}}\cdot Q^{(\geq \ell-i)}_{\mathcal{C}} \\
&\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot L\cdot \sum_{i=1}^{\ell-1} Q^{(\geq i)}_{\mathcal{C}}\cdot Q^{(\geq \ell-i)}_{\mathcal{C}} \\
&\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot L\cdot \sum_{i=1}^{\ell-1} \left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma(\ell^2-2i\ell+2i^2)n} \qquad \text{(By (7))} \\
&\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot L\cdot \sum_{i=1}^{\ell-1} \left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma\ell^2 n}\cdot 2^{-\gamma n} \\
&\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot L^2\cdot 2^{-\gamma n}\cdot \left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma\ell^2 n} \\
&=: (\star)
\end{aligned}$$

The above derivation holds whether or not $b\in\mathcal{C}$. Thus,

$$\begin{aligned}
(\star) \geq \mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}}\right] &= \mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} | b\notin\mathcal{C}\right]\mathop{\mathbf{Pr}}_{b}\left[b\notin\mathcal{C}\right] + \mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} | b\in\mathcal{C}\right]\mathop{\mathbf{Pr}}_{b}\left[b\in\mathcal{C}\right] \\
&\geq \mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} | b\notin\mathcal{C}\right]\mathop{\mathbf{Pr}}_{b}\left[b\notin\mathcal{C}\right] \\
&= \mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} | b\notin\mathcal{C}\right]\cdot\left(1-\frac{|\mathcal{C}|}{2^n}\right),
\end{aligned}$$

and so

$$\begin{aligned}
&\mathop{\mathbf{E}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} | b\notin\mathcal{C}\right] \\
&\quad\leq \frac{2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot L^2\cdot 2^{-\gamma n}\cdot\left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma\ell^2 n}}{1-\frac{|C|}{2^n}} \\
&\quad\leq \left(1+\frac{2|C|}{2^n}\right)\left(2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 2\cdot L^2\cdot 2^{-\gamma n}\cdot\left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma\ell^2 n}\right) \\
&\quad\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + \left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma\ell^2 n}\left(\frac{4|C|}{2^n} + 3\cdot L^2\cdot 2^{-\gamma n}\right) \qquad \text{(By (7))} \\
&\quad\leq 2\cdot Q^{(\geq \ell)}_{\mathcal{C}} + 4\cdot L^2\cdot 2^{-\gamma n}\cdot\left(2^{-n(1-H(p))}\cdot 2^{k}\right)^{\ell}\cdot 2^{\gamma\ell^2 n}.
\end{aligned}$$

In the third line, we used the fact that $1/(1-x)\leq 1+2x$ for all $0\leq x\leq 1/2$, along with the fact that, since $k<n$, we have $|\mathcal{C}|/2^n = 2^{k-n}\leq 1/2$. In the last line, we used that

$|\mathcal{C}| = 2^k \leq 2^{n(1-\gamma)}$ and $2 \leq \ell \leq L$. By (10), when $b \notin \mathcal{C}$, the quantity $Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} - 2Q^{(\geq \ell)}_{\mathcal{C}}$ is nonnegative. Hence, we may apply Markov's inequality to obtain

$$\mathop{\mathbf{Pr}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} - 2Q^{(\geq \ell)}_{\mathcal{C}} \geq \left(2^{-n(1-H(p))} \cdot 2^k\right)^{\ell} \cdot 2^{\gamma \ell^2 n} | b \notin \mathcal{C}\right] \;\leq\; 4 \cdot L^2 \cdot 2^{-\gamma n}.$$

When $b \in \mathcal{C}$, we have $Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} = Q^{(\geq \ell)}_{\mathcal{C}}$. Thus,

$$\mathop{\mathbf{Pr}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} - 2Q^{(\geq \ell)}_{\mathcal{C}} \geq \left(2^{-n(1-H(p))} \cdot 2^k\right)^{\ell} \cdot 2^{\gamma \ell^2 n} | b \in \mathcal{C}\right] = 0.$$

Together these imply

$$\mathop{\mathbf{Pr}}_{b}\left[Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} - 2Q^{(\geq \ell)}_{\mathcal{C}} \geq \left(2^{-n(1-H(p))} \cdot 2^k\right)^{\ell} \cdot 2^{\gamma \ell^2 n}\right] \;\leq\; 4 \cdot L^2 \cdot 2^{-\gamma n}. \tag{11}$$

Thus, with probability at least $1 - 4L^2 2^{-\gamma n}$, we have

$$\begin{aligned} Q^{(\geq \ell)}_{\mathcal{C}+\{0,b\}} &\leq 2Q^{(\geq \ell)}_{\mathcal{C}} + \left(2^{-n(1-H(p))} \cdot 2^k\right)^{\ell} \cdot 2^{\gamma \ell^2 n} \\ &\leq 3\left(2^{-n(1-H(p))} \cdot 2^k\right)^{\ell} \cdot 2^{\gamma \ell^2 n} \\ &\leq \left(2^{-n(1-H(p))} \cdot 2^{k+1}\right)^{\ell} \cdot 2^{\gamma \ell^2 n} \end{aligned}$$

where the first inequality is from (11), the second inequality is by the assumption (7), and the final inequality is because $\ell \geq 2$. This completes the proof of Lemma 20. ◀

Returning to the proof of Theorem 18, call a code $\mathcal{C}$ of dimension k *good* if (6) holds for all $1 \leq \ell \leq L$. Lemma 20 states that if $\mathcal{C}$ is good, then $\mathcal{C} + \{0, b\}$ fails to be good with probability at most $4L^3 2^{-\gamma n}$ over the choice of a uniformly random $b \in \mathbb{F}_2^n$.

Since we have already shown that $\{0\}$ is good at the beginning of the proof, it follows from the union bound that a random linear binary code $\mathcal{C} = \mathrm{span}(b_1, \ldots, b_k)$ fails to be good with probability at most $k \cdot 4L^3 2^{-\gamma n} = 2^{-\Omega(\gamma n)}$. This completes the proof of Theorem 18. ◀