Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 139 (2021)

Punctured Large Distance Codes, and Many Reed-Solomon Codes, Achieve List-Decoding Capacity

Venkatesan Guruswami Jonathan Mosheiff

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
{venkatg,jmosheif}@cs.cmu.edu

Abstract

We prove the existence of Reed-Solomon codes of any desired rate $R \in (0,1)$ that are combinatorially list-decodable up to a radius approaching $1-R$, which is the information-theoretic limit. This is established by starting with the full-length $[q,k]_q$ Reed-Solomon code over a field $\mathbb{F}_q$ that is polynomially larger than the desired dimension k, and "puncturing" it by including k/R randomly chosen codeword positions.

Our puncturing result is more general and applies to any code with large minimum distance: we show that a random rate R puncturing of an $\mathbb{F}_q$-linear "mother" code whose relative distance is close enough to $1-1/q$ is list-decodable up to a radius approaching the q-ary list-decoding capacity bound $h_q^{-1}(1-R)$. In fact, for large q, or under a stronger assumption of low-bias of the mother-code, we prove that the threshold rate for list-decodability with a specific list-size (and more generally, any "local" property) of the random puncturing approaches that of fully random linear codes. Thus, all current (and future) list-decodability bounds shown for random linear codes extend automatically to random puncturings of any low-bias (or large alphabet) code. This can be viewed as a general derandomization result applicable to random linear codes.

To obtain our conclusion about Reed-Solomon codes, we establish some hashing properties of field trace maps that allow us to reduce the list-decodability of RS codes to its associated trace (dual-BCH) code, and then apply our puncturing theorem to the latter. Our approach implies, essentially for free, optimal rate *list-recoverability* of punctured RS codes as well.

Research supported in part by NSF grants CCF-1814603 and CCF-1908125, and a Simons Investigator Award. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

ISSN 1433-8092

Contents

1 Introduction

The two main subjects of this work are *Reed-Solomon codes* and *randomly punctured codes*. We discuss each of them in turn, and then explain how our main result on the latter helps us achieve the main result on the former.

1.1 List-decoding of Reed-Solomon codes

Reed-Solomon (RS) codes are a classical family of error-correcting codes. To define an RS code, fix a prime power q and let $\mathbb{F}_q$ denote the finite field of size q. Let $1 \leq k < n \leq q$, and pick an *evaluation set* $S \subseteq \mathbb{F}_q$ of size n. The Reed-Solomon (RS) code of dimension k over the set S is defined as

$$\mathrm{RS}_{\mathbb{F}_q}(S; k) = \{(f(\alpha_1), \ldots, f(\alpha_n)) \mid f \in \mathbb{F}_q[x],\ \deg(f) < k\} \ ,$$

where $\alpha_1, \ldots, \alpha_n$ are the elements of S in some fixed order. We omit the subscript $\mathbb{F}_q$ when the underlying field is clear from context. When $S = \mathbb{F}_q$, one has the "full" Reed-Solomon code.

The code $\mathrm{RS}(S; k)$ has rate $R = \frac{k}{n}$, and every two distinct codewords differ in at least $n-k+1$ positions. For any $z \in \mathbb{F}_q^n$, there can be at most one codeword of the RS code within Hamming distance $(n-k)/2$ from z, meaning that this code is (combinatorially) *uniquely-decodable* up to a normalized radius $\frac{1-R}{2}$. There are also many classical algorithms for efficient unique decoding of RS codes up to this radius. By the Singleton bound, this is the optimal trade-off between rate and unique-decoding radius.

List-decoding is a powerful model when one wishes to decode beyond the unique-decoding radius. A code is said to be (combinatorially[1]) list-decodable up to radius ρ if every Hamming ball of radius ρn in $\mathbb{F}_q^n$ has, at most, a small (i.e., $\mathrm{poly}(n)$, or even constant) intersection with the code. Note that such a code, when accompanied by a decoding algorithm, will allow the correction of a ρ fraction of errors up to some bounded ambiguity in the worst-case. We refer the reader to [Gur06] for a detailed discussion of the motivation, usefulness, and potential of the list-decoding model.

The *List-Decoding Capacity Theorem* [GRS, Thm. 7.4.1] asserts the existence of codes that are list-decodable up to radius $1 - R - o(1)$, and no such codes exist for radius exceeding $1 - R$. The $1 - R$ bound is a basic information-theoretic limit on the error-fraction (or even erasure-fraction) that can be decoded, since in order to recover Rn messages symbols with small ambiguity, we need to receive at least $(R - o(1))n$ symbols without error. The List-Decoding Capacity Theorem means that list-decoding offers a potential twofold improvement in decoding radius over unique-decoding. Furthermore, certain explicit code constructions based on RS codes, namely *folded Reed-Solomon* and *multiplicity codes*, are known to be efficiently list-decodable up to this optimal radius [GR08; GW13; KRSW18], known as the list-decoding capacity.

In this paper we study the list-decodability of RS codes themselves. The *Johnson Bound* [Joh62; GRS, Thm. 7.3.1] is a general lower bound on the list-decoding radius as a function of the code distance, which, when applied to RS codes, shows that any RS code of rate R is list-decodable at least up to the Johnson Radius $1 - \sqrt{R} - o(1)$. Furthermore, an efficient algorithm for list-decoding RS codes up to the Johnson radius is given in [GS99]. The question of whether this algorithm can be improved, perhaps for RS codes with carefully structured evaluation sets S, has remained open.

[1] All of the results in this paper are *combinatorial*, rather than *algorithmic*.

A pre-requisite for efficient decoding beyond the Johnson radius is a combinatorial guarantee of small list-size: if some Hamming ball of radius ρ has too many codewords, then efficient list-decoding up to radius ρ is also not possible. The only such limitations known (for rates R bounded away from 0) apply for $\rho \geq 1 - R - o(1)$ [JH01] (For vanishing rates $R \to 0$, stronger limitations are shown in [BKR10] for full Reed-Solomon codes over extension fields; specifically for $R = n^{\delta-1}$, $\rho \approx 1-n^{\sqrt{\delta}-1}$). These bounds pin the list-decoding radius of RS codes in the range $1 - \sqrt{R} \leq \rho \leq 1 - R - o(1)$, resulting in a quadratic gap.

On the positive side, there has been a recent surge of results [RW14a; ST20; GLSTW20; FKS20; GST21], using techniques ranging from high-dimensional probability to tree packings to extremal combinatorics, showing the existence of RS codes that are (combinatorially) list-decodable beyond the Johnson Radius in certain regimes. These results are all probabilistic, and show that a *random* RS code in these regimes (that is, an RS code over a random evaluation set of the appropriate size) is likely to beat the Johnson Bound. For example, a beautiful recent result of [FKS20] shows the existence of RS codes of rate $\Omega(\delta)$ with list-decoding radius $1 - \delta$ for any $\delta > 0$.

Despite these advances, until the present work it was unknown whether there exist RS codes which are list-decodable all the way up to the capacity radius of $1 - R - o(1)$. In this work, we show that such codes in fact exist for every rate, over fields of arbitrary characteristic.

Theorem A (Main result about RS codes)**.** *For every $R \in (0, 1)$, $\varepsilon > 0$ and prime p, there exists a family of Reed-Solomon codes of rate R over fields of characteristic p, which are list-decodable with constant list-size up to radius $1 - R - \varepsilon$.*

The full details of this result are given in Theorem 3. For now, we mention that like previous results, Theorem A is probabilistic. Concretely, given a suitable underlying field size q, we show that an RS code over a random evaluation set of size n is very likely to be list-decodable up to capacity. The required field size is $q = n^{O_{R,\varepsilon}(1)}$, with $q = O(n^2)$ sufficing for a wide range of parameters. We note that our methods for proving Theorem A depart from these previous works. While we achieve the optimal rate vs. list-decoding radius trade-off, Theorem A does not entirely subsume the recent results [FKS20; GST21] as our field size is a larger polynomial, and the list-size is worse when the decoding radius $\rho \to 1$.

List-recovery up to capacity. We can also extend Theorem A to the model of *list-recovery* where the decoder is given not one but a subset of ℓ symbols per position, and the goal is to list all codewords which "miss" at most a fraction ρ of these subsets. Here again we show the existence of RS codes list-recoverable up to a radius approaching the optimal $1 - R$ capacity limit, *independent of ℓ* (see Theorem 5). In comparison, the performance of previous existence results for RS codes degraded with the parameter ℓ—e.g., the bounds on ρ were $1 - O(\sqrt{\ell}R\log(1/R))$ in [GLSTW20] and $1 - O(\ell R)$ in [GST21]. Also, it was shown in [GR06] that when ℓ is a prime power, certain full RS codes of rate $R = 1/\ell$ cannot be list-recovered even for $\rho = 0$ with polynomial output lists.

1.2 List-decoding of randomly punctured codes

A puncturing of a code $\mathcal{D} \subseteq \mathbb{F}_q^m$ is a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ (where we usually think of m as being much large than n) whose coordinates are taken from those of $\mathcal{D}$. More formally, $\mathcal{C}$ is a puncturing of $\mathcal{D}$ if $\mathcal{C} = \{(x_{i_1}, \ldots, x_{i_n}) \mid x = (x_1 \ldots x_m) \in \mathcal{D}\}$, for some integers $i_1, \ldots, i_n \in [m]$. We sometimes refer to $\mathcal{D}$ as the *mother-code*. When $i_1, \ldots, i_n$ are sampled uniformly and independently from

$[m]$, we say that $\mathcal{C}$ is a random n-puncturing of $\mathcal{D}$. Thus, the code $\mathrm{RS}_{\mathbb{F}_q}(S;k)$ (where $S \subseteq \mathbb{F}_q$ and $|S| = n$) is a puncturing of the full RS code $\mathrm{RS}_{\mathbb{F}_q}(\mathbb{F}_q;k)$, where we think of the coordinates of the latter as indexed by $\mathbb{F}_q$. Furthermore, the code $\mathrm{RS}_{\mathbb{F}_q}(S;k)$ with S a random set of size n, almost[2] corresponds to a random n-puncturing of $\mathrm{RS}(\mathbb{F}_q;k)$.

Many of the aforementioned works about list-decodability of RS codes, beginning with [RW14a], consider these objects within the more general framework of randomly punctured codes initiated in [Woo13]. Notably, in the low-rate regime, Ferber, Kwan, and Sauermann [FKS20] recently proved that a certain ensemble of random RS codes are, with high probability, list-decodable up to radius $1-cR$ (and thus within an additive $O(R)$ term from capacity). Their codes are random puncturings of $\mathrm{RS}(\mathbb{F}_q;k)$ for $q \geq n^{1.1}$. Their result in fact works for puncturing *arbitrary* codes of large enough distance and is *not about RS codes per se.* [3] A follow-up to [FKS20] exploited linearity of the code to show that one can take $c \approx \frac{2}{1+R}$, implying the existence of RS codes of rate R that are list-decodable up to radius $1 - \frac{2R}{1+R} - o(1)$ [GST21].

Our Theorem A is also derived from a general statement about randomly punctured codes, stated below as Theorem B. In contrast to these prior works, however, Theorem A is not a special case of Theorem B, but rather follows from it by means of a reduction, as explained in Section 1.3.

Theorem B (Main result about puncturing of large-distance codes)**.** *Fix a prime power q, $R \in (0,1)$ and $\varepsilon > 0$. Let $\mathcal{D}$ be a linear code over $\mathbb{F}_q$ of minimal distance at least $\left(1-\frac{1}{q}\right)\cdot(1-\eta)$ for $\eta(q,R,\varepsilon) > 0$. Then, a random n-puncturing of $\mathcal{D}$ of rate R is list-decodable with constant list-size up to radius $\rho^{(q)}_{\mathsf{capacity}}(R) - \varepsilon$, with high probability as $n \to \infty$. Here $\rho^{(q)}_{\mathsf{capacity}}(R)$ stands for the optimal list-decoding radius of a code of rate R over $\mathbb{F}_q$.*[4]

The full details of this result are given in Theorems 1 and 2. As Theorem B is significant even independently of Theorem A (due to the generality of the hypothesis and the fact that we achieve list-decoding capacity, and since code puncturing is such a fundamental notion), we discuss the former on its own before connecting the two results.

Theorem B essentially means that random puncturings of linear codes of sufficiently large distance are list-decodable up to capacity. This can be seen as generalizing previous works about *list-decodability of random linear codes.* A random linear code (RLC) of rate R and length n over $\mathbb{F}_q$ is the kernel[5] of a uniformly random matrix in $\mathbb{F}_q^{(1-R)n\times n}$. Almost equivalently, an RLC can be seen as a random n-puncturing of the *Hadamard Code* of length q^{Rn} over $\mathbb{F}_q$, the latter having minimum distance $1-\frac{1}{q}$. Random linear codes are well known to achieve list-decoding capacity [ZP81], and Theorem B generalizes this fact from puncturings of Hadamard codes (which have

[2] Note that our random puncturing model allows the same coordinate of $\mathcal{D}$ to be chosen several times for inclusion in $\mathcal{C}$. In contrast, the evaluation set of an RS code may not be a multiset. Hence, a random puncturing of $\mathrm{RS}_{\mathbb{F}_q}(\mathbb{F}_q;k)$ is an RS code if and only if no coordinate was sampled more than once. This quirk of the model is only significant at one point in this work—during the proof of Theorem 3—where we deal with it via conditioning.

[3] The results in [GLSTW20; ST20] are specific to RS codes. The existence of RS codes over exponentially large fields that have rate $\Omega(\varepsilon/\log(1/\varepsilon))$ and are list-decodable up to radius $1-\varepsilon$ is shown in [GLSTW20]. RS codes over exponential fields list-decodable up to radius $\frac{L}{L+1}(1-R)$ for list sizes $L = 2,3$ are shown in [ST20].

[4] The term $\rho^{(q)}_{\mathsf{capacity}}(R)$ is explicitly given by $h_q^{-1}(1-R)$, where h_q^{-1} is the inverse of the q-ary entropy function $h_q(x) = -x\log_q x - (1-x)\log_q(1-x) + x\log_q(q-1)$. In particular, $\rho^{(q)}_{\mathsf{capacity}}(R) \approx 1-R$ for large q.

[5] Another reasonable RLC model is to directly sample a code uniformly from the set of all rate R linear codes in $\mathbb{F}_q^n$. A third model is to take the image of a uniformly random matrix in $\mathbb{F}_q^{n\times Rn}$. All three models are within total-variation distance exponentially small in n of each other, and thus they are essentially identical.

distance $1 - 1/q$), to puncturings of any code of large distance (close enough to $1 - 1/q$).

In this sense, Theorem B yields a certain *derandomization* of RLCs. To illustrate this, suppose that one wishes to obtain, with high probability, a binary code $\mathcal{C}$ of rate R that achieves list-decoding capacity. By [ZP81], one may take $\mathcal{C}$ to be an RLC, namely, a random n-puncturing of the Hadamard code of length $m = 2^{Rn}$. This requires $n \log_2 m = Rn^2$ random bits. By Theorem B, one may instead take as the mother-code any explicit code of large enough distance. In Section 10 we show how, by puncturing a suitable mother-code, one can construct binary codes that achieve list-decoding capacity using just $O(n)$ random bits. The previous best randomness bound for sampling such codes was $\Omega(n \log^2 n)$ [GR10].

Connections to list-decodability of random linear codes. The list-decodability of RLCs has been studied in different regimes in many previous works [ZP81; GHK11; CGV13; Woo13; RW14b; RW18; LW21; GLMRSW20]. Since [ZP81] already establishes that RLCs are list-decodable up to capacity with list-size constant in n, the focus of the later works is pinpointing the exact dependence of the list-size on the field size q, rate R and gap to capacity ε.

Our proof of Theorem B works by *reducing* the list-decodability of a random code $\mathcal{C}$, obtained as a random puncturing of some large-distance linear code, to that of an RLC. Specifically, Theorem 1 shows that $\mathcal{C}$ is likely to be list-decodable up to capacity, with a similar list-size to that given for an RLC in [GHK11]. Theorem 2 shows that, under a wide range of conditions, we can do better. Namely, under these conditions, the list-decodability parameters of $\mathcal{C}$ are similar to those of an RLC, *independently of any specific RLC bound.* In particular, any known bound on RLC list-decodability can then immediately be applied to $\mathcal{C}$. Moreover, the same would also be true for any positive RLC list-decoding bound discovered in the future. The latter may be relevant since there are still some gaps in our knowledge of RLC list-decodability, especially for the large q regime. Fortunately, q being large is a sufficient condition for Theorem 2 to apply.

Broader pseudorandomness perspective. The idea of relating the list-decodability of a more structured code $\mathcal{C}$ to that of an RLC figures in a different context in [MRRSW20], where a *Gallagher LDPC Code* is cast in the role of $\mathcal{C}$. While our methods in the present work are very different, our proof of Theorem 2 does use the framework of [MRRSW20], as well as the *RLC Threshold Theorem* [MRRSW20, Thm. 2.8] proven there. [MRRSW20] (and its follow-up works [GMRSW21; GLMRSW20]) treat list-decodability as a special case of a *monotone, local and row-symmetric property of codes* (list-decodability is a *local property* because it can be characterized as not containing a set of L bad (i.e., clustered) words, where, crucially, L is small). Under this viewpoint, the main result of [MRRSW20] can be loosely stated as "With high probability, a Gallagher code has the the same monotone-decreasing, local and row-symmetric properties as an RLC".

Similarly, Theorem 2 also has a more general formulation in terms of code properties (see Section 8). Loosely put, this generalization states that the code $\mathcal{C}$ is "locally similar" to an RLC (from a derandomization point of view, this means that the derandomized code $\mathcal{C}$ is similar to an RLC not just in terms of list-decodability, but in many other ways as well). In particular, the latter generalization applies to *list-recoverability.* Thus, our proof of Theorem 2 immediately yields a positive result (Theorem 4) about the list-recoverability of $\mathcal{C}$, via reduction to established results about the list-recoverability of an RLC.

We see Theorem B as part of a broader theme in coding theory, where a very random code (in this case, an RLC) has excellent *combinatorial* properties, but is *resistant to algorithms* due to its randomness. One thus seeks to *derandomize* this code—preserving the desired combinatorial features and simultaneously opening the door to efficient algorithms by adding structure. Theorem B states that the punctured code $\mathcal{C}$ is combinatorially similar to an RLC, and offers a decent amount of freedom in choosing the mother-code, leading to many possible structures that can be enforced on $\mathcal{C}$. It is our hope that some of these choices will lead to algorithmic results (see also Remark 3.2).

1.3 From Theorem B to Theorem A via trace codes

In each of the previous works [RW14a; FKS20; GST21], the main result about list-decodability of RS codes is directly obtained as a special case of a more general theorem about randomly punctured codes. Our work differs, since Theorem B cannot be directly applied to RS codes. The issue is that, in order to be able to satisfy the distance requirement, the mother-code $\mathcal{D}$ must be of length significantly larger than q. Unfortunately, the length of the full RS code $\mathrm{RS}(\mathbb{F}_q; k)$ is exactly q.

We are thus forced to take a different approach. Rather than applying Theorem A to the full RS code $\mathrm{RS}(\mathbb{F}_q; k)$, we apply a small variation of Theorem A to a certain code $\mathcal{D}$, known as the *trace code* of $\mathrm{RS}(\mathbb{F}_q; k)$, to show that a random puncturing of $\mathcal{D}$ is likely to be "quasi-list-decodable," (see Sections 2.2 and 6.2.2 for details). Crucially, $\mathcal{D}$ has a much smaller underlying field. To finish the argument, we reduce the list-decodability of a random puncturing of $\mathrm{RS}(\mathbb{F}_q; k)$ to the quasi-list-decodability of a random puncturing of $\mathcal{D}$. Our result about list-recoverability of RS codes is obtained in a similar manner.

2 Main Results

Before stating our main results, we formally define some of the relevant notions.

Definition 2.1 (Random puncturing)**.** *Fix some prime power q. Let $m, n \in \mathbb{N}$. An $(m \to n)$* puncturing map *is a function $\varphi : \mathbb{F}_q^m \to \mathbb{F}_q^n$ of the form $\varphi(u = (u_1, \dots, u_m)) = (u_{i_1}, u_{i_2}, \cdots, u_{i_n})$ for some $i_1, \dots, i_n \in [m]$. If $i_1, \dots, i_n$ are sampled i.i.d. and uniformly from $[m]$, we say that φ is a* random $(m \to n)$ puncturing map.

A random n-puncturing *of a code $\mathcal{D} \subseteq \mathbb{F}_q^m$ is a random code $\mathcal{C} = \varphi(\mathcal{D}) = \{\varphi(u) \mid u \in \mathcal{D}\}$, where $\varphi : \mathbb{F}_q^m \to \mathbb{F}_q^n$ is a random puncturing map. The* design rate *of $\mathcal{C}$ is $\frac{\log_q |\mathcal{D}|}{n}$.*

Definition 2.2. *Let $\mathcal{D} \subseteq \mathbb{F}_q^m$, where q is a power of some prime p, be a linear code and let $\eta > 0$.*

1. *If every $u \in \mathcal{D} \setminus \{0\}$ has $\mathrm{wt}(u) \geq \frac{(q-1)(1-\eta)}{q}$, we say that $\mathcal{D}$ has η-*optimal distance*. Here, $\mathrm{wt}(u) = \frac{|\{i \in [m] | u_i \neq 0\}|}{m}$ denotes the* normalized Hamming weight *of $u \in \mathbb{F}_q^m$.*

2. *A vector $u \in \mathbb{F}_q^m$ is said to be η-*biased *if $\left|\sum_{i=1}^m \omega^{\mathrm{tr}(a \cdot u_i)}\right| \leq m\eta$ for all $a \in \mathbb{F}_q^*$. Here, $\omega = e^{\frac{2\pi i}{p}}$ and $\mathrm{tr} : \mathbb{F}_q \to \mathbb{F}_p$ is the field trace map (see Section 4.4). The code $\mathcal{D}$ is said to be η-*biased *if every $u \in \mathcal{D} \setminus \{0\}$ is η-biased.*

As shown in Lemma 4.14, an η-biased code also has η-optimal distance, so the former is a stronger notion. For intuition, note that in the binary case η-bias implies $\frac{1-\eta}{2} \leq \mathrm{wt}(u) \leq \frac{1+\eta}{2}$ for any $u \in \mathcal{D} \setminus \{0\}$, whereas η-optimal distance only implies the lower bound on $\mathrm{wt}(u)$.

Up to and including Section 5, it may be simpler for the reader to focus on the case where q is a prime, i.e., $q = p$. In this case, tr is merely the identity map.

If $\mathcal{C}$ is a random n-puncturing of a code $\mathcal{D}$, the rate of $\mathcal{C}$ is clearly bounded from above by its design rate. The following lemma shows that when $\mathcal{D}$ is of almost optimal distance, these two terms are very likely to coincide. In light of this lemma, we blur the distinction between design rate and actual rate.

Lemma 2.3 (Actual rate equals design rate whp)**.** *Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a code of η-optimal distance, and let $\mathcal{C}$ be a length-n random puncturing of $\mathcal{C}$, of design rate $R \leq 1 - \log_q(1+\eta q) - \varepsilon$. Then, with probability at least $1 - q^{-n\varepsilon}$, the rate of $\mathcal{C}$ is equal to its design rate.*

Proof. The rate of $\mathcal{C}$ is smaller than R if and only if there exists a non-zero word $u \in \mathcal{D}$ such that only coordinates $i \in [m]$ for which $u_i = 0$ are sampled for inclusion in $\mathcal{C}$. For a given u, this happens with probability

$$(1 - \mathrm{wt}(u))^n \leq \left(\frac{1}{q} + \frac{q-1}{q}\eta\right)^n \leq \left(\frac{1}{q} + \eta\right)^n = q^{-n(1-\log(1+q\eta))} \ .$$

The claim follows by a union bound over the non-zero words of $\mathcal{D}$, of which there are $q^{Rn} - 1$, and the assumed upper bound on R. □

Definition 2.4 (clustered sets and list-decodability)**.** *Fix $\rho \in [0,1]$. A set of vectors $W \subseteq \mathbb{F}_q^n$ is called ρ-*clustered *if there exists some $z \in \mathbb{F}_q^n$ such that $\mathrm{wt}(u - z) \leq \rho$ for each $u \in W$. A code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is said to be (ρ, L)-*list-decodable *if it does not contain any ρ-clustered set of size $L+1$.*

2.1 List-decoding of punctured codes

In our discussion of list-decoding capacity in Section 1 we treated the rate R as fixed, and the list-decoding capacity radius ρ as a function of R and the field size. Henceforth we will prefer to think of R as depending on some fixed ρ.

The *List-Decoding Capacity Theorem* [GRS, Thm. 7.4.1] states that the *optimal rate* for radius ρ list-decoding over the field $\mathbb{F}_q$ is $R^* = 1 - h_q(\rho)$, where

$$h_q(\rho) = -\rho \log_q \rho - (1-\rho)\log_q(1-\rho) + \rho \log_q(q-1)$$

is the q-ary entropy function (see Section 4.5). In other words, there exist infinite families of codes of rate $R^* - \varepsilon$ that are list-decodable up to radius ρ, but no such families exist for rate $R^* + \varepsilon$. Note that $R^* \approx 1 - \rho$ when q is large.

A GHK-style bound for randomly punctured codes. Theorem 1 is a concrete version of Theorem B. In particular, it states that a random puncturing of a code of near-optimal distance almost surely achieves list-decoding capacity with constant list-size, i.e., list-size independent of n.

Theorem 1 (A puncturing of a near-optimal distance code is whp list-decodable up to capacity)**.** *Fix a prime power q. Let $L, n \in \mathbb{N}$ and $0 < \rho < \frac{q-1}{q}$, such that $\frac{n}{\log_q n} \geq \omega\left(q^{L+1}\right)$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a linear code with η-optimal distance, where $\eta = q^{-L+1}$. Let $\mathcal{C}$ be a random n-puncturing of $\mathcal{D}$ of design rate R, where $R \leq 1 - h_q(\rho) - \frac{K}{L}$ for some constant $K = K_{\rho,q}$. Then,*

$$\Pr\left[\mathcal{C} \textit{ is } (\rho, L)\textit{-list-decodable}\right] \geq 1 - q^{-\Omega(n)} \ .$$

Furthermore, one can take

$$K_{\rho,q} \leq \exp\left(O\left(\frac{(\log q)^2}{\min\left\{(1-1/q-\rho)^2, \rho\right\}}\right)\right) \tag{1}$$

and, in particular, $K_{\rho,q} \leq \mathrm{poly}(q)$ *whenever* ρ *is bounded away from* 0 *and* $1-\frac{1}{q}$.

By the List-Decoding Capacity Theorem, Theorem 1 achieves the optimal trade-off between q, ρ and R. We thus turn to discuss the secondary trade-off, which involves the former three parameters and the *list-size* L. As mentioned in Section 1.2, Theorem 1 is derived by reduction to the result of [GHK11] on list-decodability of RLCs. The main theorem of [GHK11] states that a RLC of rate $R = 1 - h_q(\rho) - \frac{K'_{\rho,q}}{L}$ is with high probability is (ρ, L)-list-decodable, where $K'_{\rho,q} \leq \exp\left(O\left(\frac{(\log q)^2}{\min\{(1-1/q-\rho)^2,\rho\}}\right)\right)$ is proportional to the constant $K_{\rho,q}$ that appears in Theorem 1. Denoting the *gap-to-capacity of the rate* by $\varepsilon = 1 - h_q(\rho) - R$, [GHK11] shows that an RLC of rate R is almost surely (ρ, L)-list-decodable with $L \approx \frac{K'_{\rho,q}}{\varepsilon}$. In Theorem 1, we have $\varepsilon = \frac{K_{\rho,q}}{L}$, so $L = \frac{K_{\rho,q}}{\varepsilon} = O\left(\frac{K'_{\rho,q}}{\varepsilon}\right)$. Thus, we can informally state Theorem 1 as "A random puncturing of a code of near-optimal distance is very likely to be list-decodable up to capacity, with a similar list-size trade-off to that guaranteed by [GHK11] for RLCs".

The list-size L guaranteed by Theorem 1 inherits some desirable properties from [GHK11]: it is constant in terms of n, and has linear dependence on $\frac{1}{\varepsilon}$, which is tight for RLCs [GN14, Thm. 16]. As for the dependence on q and ρ, we get good list-size bounds when q is not too large and ρ is bounded away from 0 and $1-\frac{1}{q}$, but, unfortunately, the constant $K_{\rho,q}$ grows exponentially as $\rho \to 1-\frac{1}{q}$. In comparsion with [GHK11], other works on RLC list-decodability are more specialized, and give tighter upper bounds on the list-size in specific regimes. Notably, [Woo13] does well when ρ is large and ε is of similar magnitude to R, and [LW21] gives an extremely tight upper bound (see [GLMRSW20]) on the list-size for every ρ and ε, when $q = 2$.

Punctured codes are as list-decodable as RLCs. To state our next general result (Theorem 2), we require some notation about RLCs. For $n \in \mathbb{N}$, $R \in [0,1]$ such that $Rn \in \mathbb{Z}$, and a prime power q, let $C^{n,q}_{\mathrm{RLC}}(R)$ denote a *random linear code* of rate R and length n over $\mathbb{F}_q$. More precisely, $C^{n,q}_{\mathrm{RLC}}(R)$ is the kernel of a uniformly random matrix in $\mathbb{F}_q^{(1-R)n\times n}$. For $L, n \in \mathbb{N}$, $0 < \rho < \frac{q-1}{q}$ and a prime power q, define the *RLC threshold rate for* (ρ, L)*-list-decodability* by

$$\mathrm{RLC}^{n,q}(\rho, L) = \max\left\{R \in [0,1] \mid \Pr\left[C^{n,q}_{\mathrm{RLC}}(R) \text{ is } (\rho, L)\text{-list-decodable}\right] \geq \tfrac{1}{2}\right\} \quad .$$

This terminology is motivated by the following theorem, which states that the probability of an RLC of rate R being (ρ, L)-list-decodable, as a function of R, rapidly drops near the threshold from $1 - o(1)$ to $o(1)$.

Theorem 2.5 (Threshold behavior of RLC list-decodability [MRRSW20, Thm. 2.8])**.** *Let* $R^* = \mathrm{RLC}^{n,q}(\rho, L)$ *and fix* $\varepsilon > 0$. *Then,*

$$\Pr\left[C^{n,q}_{\mathrm{RLC}}(R^*+\varepsilon) \textit{ is } (\rho, L)\textit{-list-decodable}\right] \leq q^{-(\varepsilon - o(1))n}, \quad \textit{and}$$
$$\Pr\left[C^{n,q}_{\mathrm{RLC}}(R^*-\varepsilon) \textit{ is } (\rho, L)\textit{-list-decodable}\right] \geq 1 - q^{-(\varepsilon - o(1))n} \quad .$$

Motivated by the myriad results concerning the list-decodability of RLCs, Theorem 2 stated below is a strengthening of Theorem 1. It directly reduces the list-decodability parameter trade-off of $\mathcal{C}$ to that of an RLC, rather than going through a specific RLC bound such as [GHK11]. When applying Theorem 2, one can thus apply the most suitable RLC list-decoding bound for a given situation, e.g., the one that gives the tightest bound on the list-size for the given set of parameters. Another benefit of Theorem 2 is that any newly discovered positive RLC list-decoding result would immediately be applicable to the punctured code $\mathcal{C}$ as well.

More broadly, Theorem 2 is interesting because it hints that a random puncturing of any large-distance code "resembles" an RLC, in a formal sense. As we discuss in Section 8, this resemblance is rather deep, going beyond just similarity in list-decoding parameters. Theorem 2 applies in two situations: 1. When q is large. 2. When the mother-code $\mathcal{D}$ has small bias.

Theorem 2 (Puncturings of certain linear codes are as list-decodable as RLCs)**.** *Let q be a prime power, $0 < \rho < \frac{q-1}{q}$, $L \in \mathbb{N}$ and $n \in \mathbb{N}$ such that $\frac{n}{\log_q n} \geq \omega\left(q^{-2(L+1)}\right)$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a linear code. Let $\mathcal{C}$ be a random n-puncturing of $\mathcal{D}$ of design rate $R \leq \mathrm{RLC}^{n,q}(\rho, L) - \varepsilon$ for some $\varepsilon > 0$. Suppose that at least one of the following conditions holds:*

1. *$\mathcal{D}$ has $q^{-(L+1)}$-optimal distance and $q \geq 2^{\frac{2}{\varepsilon}}$.*
2. *$\mathcal{D}$ is $\left(\frac{\varepsilon(L+1)\ln q}{q^{L+1}}\right)$-biased.*

Then, $\Pr\left[\mathcal{C} \text{ is } (\rho, L)\text{-list-decodable}\right] \geq 1 - q^{-(\varepsilon - o(1))n}$.

The reasons for the conditions in Theorem 2 are discussed in Remark 8.14.

2.2 List-decoding of Reed-Solomon codes

Our main result about RS codes is the following more detailed version of Theorem A.

Theorem 3 (RS codes list-decodable up to capacity)**.** *For every $\rho \in (0,1)$ and $0 < \varepsilon < \min\{\rho, 1-\rho\}$, any prime p and any $n \in \mathbb{N}$ large enough in terms of p and ε, there exist $L = L(p, \rho, \varepsilon) \in \mathbb{N}$ and $q \leq O_{p,\varepsilon}\left(n^{\max\left\{2, \frac{1}{\rho-\varepsilon}\right\}}\right)$, which is a power of p, for which the following holds: Let $S \subseteq \mathbb{F}_q$ be a uniformly sampled subset of size n. Then, $\mathrm{RS}_{\mathbb{F}_q}(S; (1-\rho-\varepsilon)n)$ is (ρ, L)-list-decodable with probability $1 - p^{-\Omega(n)}$.*

Furthermore, one can take

$$L \leq \exp\left(O\left(\frac{\left(\log p + \frac{1}{\varepsilon}\right)^2}{\min\left\{(1-\rho)^2, \rho\right\}}\right)\right) . \tag{2}$$

Note that $\mathrm{RS}_{\mathbb{F}_q}(S; (1-\rho-\varepsilon)n)$ is essentially a random puncturing of the full Reed-Solomon code $\mathrm{RS}_{\mathbb{F}_q}(\mathbb{F}_q; (1-\rho-\varepsilon)n)$. Despite this, Theorem 3 cannot be obtained as a special case of Theorem 1 or Theorem 2. This is because the η-optimal-distance condition of these theorems becomes too restrictive for a mother-code of equal length and alphabet size, such as the full RS code.

Hence, to prove Theorem 3, we pick some suitable $Q = Q(p, \rho, \varepsilon)$ such that $\mathbb{F}_Q \subseteq \mathbb{F}_q$, and consider the code $\mathcal{D}$, defined as the image of $\mathrm{RS}_{\mathbb{F}_q}(\mathbb{F}_q; k = (1-\rho-\varepsilon)n)$ under the $\mathbb{F}_q \to \mathbb{F}_Q$ trace map (see Section 6.1 for details). The code $\mathcal{D}$ is known as a *trace code* or a *dual-BCH code*. As explained in Section 1.3, we first prove that a random puncturing of $\mathcal{D}$, denoted $\varphi(\mathcal{D})$, is likely to be *quasi-list-decodable* (see Definition 6.8) up to capacity. We then establish certain pseudorandom properties of the field trace map in order to reduce the list-decodability of $\mathrm{RS}_{\mathbb{F}_q}(S;k)$ to the quasi-list-decodability of $\varphi(\mathcal{D})$.

We show that $\varphi(\mathcal{D})$ is almost surely quasi-list-decodable, via methods similar to those used to prove Theorems 1 and 2. We note that these theorems cannot be directly applied to $\varphi(\mathcal{D})$ because $\mathcal{D}$ is not generally guaranteed to have near-optimal distance. For the sake of simplicity, our proof of the quasi-list-decodability of $\varphi(\mathcal{D})$ follows the framework of Theorem 1 rather than Theorem 2, that is, we reduce to the bound given by [GHK11], rather than to the actual behavior of the RLC rate threshold.

2.3 List-recovery and row-symmetric local properties of codes

We formally define the notion of (combinatorial) list-recovery.

Definition 2.6. *Fix $1 \le \ell \le q$ and let $\rho \in (0, 1 - \ell/q)$. The set W is said to be (ρ,ℓ)-*recovery-clustered *if there exist sets $Z_1, \dots, Z_n \subseteq \mathbb{F}_q$, each of which is of size at most ℓ, such that $|\{i \in [n] \mid u_i \notin Z_i\}| \le \rho n$ for all $u \in W$. A code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is called (ρ,ℓ,L)*-list-recoverable *if it does not contain any (ρ,ℓ)-recovery-clustered set of size $L+1$.*

Note that list-recovery generalizes list-decodability (Definition 2.4), i.e., a set W is ρ-clustered if and only if it is $(\rho,1)$-recovery-clustered. Likewise, a code is (ρ,L)-list-decodable if and only if it is $(\rho,1,L)$-list-recoverable.

The *List-Recovery Capacity Theorem* [Res20, Thm. 2.4.12] gives the threshold rate for list-recoverability as $R^* = 1 - h_{q,\ell}(\rho)$, where $h_{q,\ell}(\rho) = \rho \log_q\left(\frac{q-\ell}{\rho}\right) + (1-\rho)\log_q\left(\frac{\ell}{1-\rho}\right)$. Namely, for every $\varepsilon > 0$ there exists a family of $(\rho,\ell,O_{\rho,\ell,\varepsilon}(1))$-list-recoverable codes of rate at least $R^* - \varepsilon$ but, on the other hand, every (ρ,ℓ,L)-list-recoverable family of codes of rate $\ge R^* + \varepsilon$ has L exponentially large in εn.

RLCs are known to achieve list-recovery capacity. Concretely, let

$$\mathrm{RLC}^{n,q}(\rho,\ell,L) = \max\left\{R \in [0,1] \mid \Pr\left[C^{n,q}_{\mathrm{RLC}}(R) \text{ is } (\rho,\ell,L)\text{-list-recoverable}\right] \ge \tfrac{1}{2}\right\} .$$

As explained in Section 8, the threshold behavior of Theorem 2.5 applies to list-recovery as well (see Theorem 8.10), so an RLC of rate $\mathrm{RLC}^{n,q}(\rho,\ell,L) - \varepsilon$ is very likely to be (ρ,ℓ,L)-list-recoverable. One simple bound (see Section 8.3 for a proof) is[6]

$$\mathrm{RLC}^{n,q}(\rho,\ell,L) \ge 1 - h_{q,\ell}(\rho) - \frac{\ell}{\log_q L} - o_{n\to\infty}(1) \tag{3}$$

for any fixed q, ρ, ℓ and L. Eq. (3) means that RLCs get to within ε of the capacity rate for list-recovery with list-size $L \approx q^{\frac{\ell}{\varepsilon}}$.

Theorem 4—an analog of Theorem 2—reduces the list-recoverability of random puncturings of near-optimal distance codes to that of RLCs. Together with Eq. (3), the theorem implies, in particular, that these punctured codes achieve list-recovery capacity. List-recoverability beyond the Johnson bound of random puncturings of large distance codes was shown in [LP20; GST21].

[6] Better lower bounds on $\mathrm{RLC}^{n,q}(\rho,\ell,L)$ are known. See, e.g., [RW18].

Theorem 4 (List-recovery of random puncturings of certain linear codes)**.** *Let q be a prime power, $0 < \rho < \frac{q-1}{q}$, $L \in \mathbb{N}$, $1 \le \ell \le q$ and $n \in \mathbb{N}$ such that $\frac{n}{\log_q n} \ge \omega\left(q^{-2(L+1)}\right)$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a linear code. Let $\mathcal{C}$ be a random n-puncturing of $\mathcal{D}$ of design rate $R \le 1 - h_{q,\ell} - \varepsilon$ for some $\varepsilon > 0$. Suppose that at least one of the following conditions holds:*

1. *$\mathcal{D}$ has $q^{-(L+1)}$-optimal distance and $q \ge 2^{\frac{2}{\varepsilon}}$.*
2. *$\mathcal{D}$ is $\left(\frac{\varepsilon(L+1)\ln q}{q^{L+1}}\right)$-biased.*

Then, $\Pr\left[\mathcal{C} \text{ is } (\rho, \ell, L)\text{-list-recoverable}\right] \ge 1 - q^{-(\varepsilon - o(1))n}$.

As we show in Section 8, Theorem 4 follows essentially for free from the proof of Theorem 2 via the general framework of local and row-symmetric properties of codes.

For RS codes, we have the following analog of Theorem 3. The notable feature is that the parameter ℓ has no effect on the rate rate vs. decoding radius trade-off, and in fact we achieve capacity $\rho \approx 1 - R$ for any desired rate. Previous results only applied for rates $R \le 1/\ell^{\Omega(1)}$, and did not achieve capacity for any rate.

Theorem 5 (RS codes list-recoverable up to capacity)**.** *For every $\rho \in (0,1)$, $\ell \in \mathbb{N}$ and $0 < \varepsilon < \min\{\rho, 1-\rho\}$, any prime p and any $n \in \mathbb{N}$ large enough in terms of p and ε, there exist $L = L(p, \rho, \varepsilon) \in \mathbb{N}$ and $q \le O_{p,\varepsilon}\left(n^{\max\left\{2, \frac{1}{\rho-\varepsilon}\right\}}\right)$, which is a power of p, for which the following holds: Let $S \subseteq \mathbb{F}_q$ be a uniformly sampled subset of size n. Then, $\mathrm{RS}_{\mathbb{F}_q}(S; (1-\rho-\varepsilon)n)$ is (ρ, ℓ, L)-list-recoverable with probability $1 - p^{-\Omega(n)}$.*

Furthermore, one can take

$$L \le q^{\frac{2\ell}{\varepsilon}} + 1 \ . \tag{4}$$

2.4 Derandomization of RLCs

As discussed in Section 1.2, Theorems 1 and 2 can be invoked to derandomize RLCs by casting a short code of low bias in the role of the mother-code $\mathcal{D}$. One result that can be achieved via this method is the following theorem. For simplicity, we focus on the binary case.

Theorem 6 (Codes achieving list-decoding capacity with $O(n)$ randomness)**.** *There exists a randomized algorithm that, given $\rho \in \left(0, \frac{1}{2}\right)$, $L \in \mathbb{N}$, $\varepsilon > 0$ and $n \in \mathbb{N}$ where $\frac{n}{\log_2 n} \ge \omega\left(2^{2L}\right)$ and $n \ge \omega(1/\varepsilon)$, samples a generating matrix for a linear code $\mathcal{C}$ of rate $R \ge \mathrm{RLC}^{n,2}(\rho, L) - \varepsilon$ such that*

$$\Pr\left[\mathcal{C} \text{ is } (\rho, L)\text{-list-decodable}\right] \ge 1 - 2^{-\Omega(\varepsilon n)}.$$

This algorithm uses $O\left(n\left(L + \log_2 \frac{1}{\varepsilon}\right)\right)$ random bits, and works in time polynomial in n.

In Section 10, we use the aforementioned framework of local and row-symmetric properties to prove a stronger statement than Theorem 6, namely, that the code $\mathcal{C}$ is likely to “locally resemble” an RLC.

2.5 Organization

The rest of the paper is organized as follows. In Section 3 we survey our techniques by sketching proofs for simplified versions of the main theorems. Section 4 establishes some general definitions and lemmas used in the main proof. In Section 5 we prove Theorem 1 about list-decodability of randomly punctured codes. In Section 6 we provide the necessary background about finite fields and the trace map, and then prove Theorem 3 about list-decodable RS codes modulo a certain algebraic theorem about hash properties of trace maps, which we prove in Section 7. In Section 8 we recall the framework for properties of codes from [MRRSW20], and prove a generalization of Theorems 2 and 4. Theorem 5 about list-recoverability of RS codes is proven in Section 9. Finally, a more general version of Theorem 6, dealing with derandomization of RLCs, is proven in Section 10.

3 Technical overview

For the sake of exposition, we begin by sketching a proof for a weaker version of Theorem B. This version assumes that the mother-code has small bias, rather than just near-optimal distance. For simplicity, we also restrict ourselves to the binary field.

Theorem B'. *Let $\rho \in \left(0, \frac{1}{2}\right)$ and $L \in \mathbb{N}$. Then, there exist $\eta(L) > 0$ and $\varepsilon(L) > 0$ with $\varepsilon(L) \xrightarrow[L\to\infty]{} 0$, such that the following holds. Let $\mathcal{D} \subseteq \mathbb{F}_2^m$ be a linear η-biased code, where $\eta(L)$ is positive and small enough, and let $\mathcal{C}$ be a random n-puncturing of $\mathcal{D}$ of design rate $R \le 1 - h_2(\rho) - \varepsilon$. Then $\mathcal{C}$ is (ρ, L)-list-decodable with high probability as $n \to \infty$.*

Proof sketch. Let $\varphi : \mathbb{F}_2^m \to \mathbb{F}_2^n$ be the random puncturing map by which $\mathcal{C}$ is generated from $\mathcal{D}$. Write $b = \lceil \log_2(L+1) \rceil$. Now any set of $L+1$ vectors in $\mathbb{F}_2^n$ must contain a subset of b linearly-independent vectors. In particular, for $\mathcal{C}$ to contain a ρ-clustered set of size $L+1$, it must contain a ρ-clustered set of b linearly-independent vectors (this argument originated in [ZP81]). Thus, the probability, taken over the random puncturing φ, that $\mathcal{C}$ is **not** (ρ, L)-list-decodable is at most

$$\begin{aligned}
&\Pr\left[\exists v_1, \dots, v_b \in \mathcal{C} \text{ which are } \rho\text{-clustered and linearly-independent}\right] \\
&= \Pr\left[\exists u_1, \dots, u_b \in \mathcal{D} \text{ s.t. } \varphi(u_1), \dots, \varphi(u_b) \text{ are } \rho\text{-clustered and linearly-independent}\right] \\
&\le \Pr\left[\exists u_1, \dots, u_b \in \mathcal{D} \text{ which are linearly independent, s.t. } \varphi(u_1), \dots, \varphi(u_b) \text{ are } \rho\text{-clustered}\right] \\
&\le \sum_{\substack{u_1,\dots u_b \in \mathcal{D} \\ \text{linearly independent}}} \Pr\left[\varphi(u_1), \dots, \varphi(u_b) \text{ are } \rho\text{-clustered}\right] \quad ,
\end{aligned}$$

where the penultimate inequality is because linear-independence of $u_1, \dots, u_b$ is a necessary condition for linear-independence of $\varphi(u_1), \dots, \varphi(u_b)$. The sum on the right hand side has at most $|\mathcal{D}|^b = 2^{bRn}$ terms, so it suffices to show that

$$\Pr\left[\varphi(u_1), \dots, \varphi(u_b) \text{ are } \rho\text{-clustered}\right] \le 2^{-bRn-\omega(1)} \tag{5}$$

whenever $u_1, \dots, u_b \in \mathcal{D}$ are linearly independent.

Let $B \in \mathbb{F}_2^{m\times b}$ be the matrix whose columns are $u_1, \dots, u_b$, and let σ denote the distribution, over $\mathbb{F}_2^b$, of a uniformly random row of B. Let $A \in \mathbb{F}_2^{n\times b}$ be the matrix whose columns are $\varphi(u_1), \dots, \varphi(u_b)$. A crucial observation is that A is a random matrix whose rows are sampled independently from σ. At this point, if σ were the uniform distribution over $\mathbb{F}_2^b$, we would be done. Indeed, σ being uniform means that the columns of A, call them $c_1, c_2, \ldots, c_b$, are sampled independently and uniformly from $\mathbb{F}_2^n$. This establishes Eq. (5) since

$$\begin{aligned}
\Pr_{c_1,\ldots,c_b \sim \mathsf{U}(\mathbb{F}_2^n)}[c_1,\ldots,c_b \text{ are } \rho\text{-clustered}] &\leq \sum_{z\in\mathbb{F}_2^n}\sum_{y_1,\ldots,y_b\in B(z,\rho n)}\Pr_{c_1,\ldots,c_b\sim\mathsf{U}(\mathbb{F}_2^n)}\left[\bigwedge_{i=1}^{b}(c_i=y_i)\right] \\
&= \sum_{z\in\mathbb{F}_2^n}\sum_{y_1,\ldots,y_b\in B(z,\rho n)}(2^{-b})^n \qquad (6)\\
&\leq \sum_{z\in\mathbb{F}_2^n} 2^{bh_2(\rho)n}(2^{-b})^n = 2^{n(bh_2(\rho)-b+1)} \\
&\leq 2^{-bRn-n}\ , \qquad (7)
\end{aligned}$$

where $B(z, \rho n)$ denotes the Hamming ball of radius ρn around z, and the last inequality Eq. (7) holds for, say, $\varepsilon = \frac{2}{b}$. Note that $\varepsilon \leq O\left(\frac{1}{\log L}\right)$.

We now use a certain formulation of the *Vazirani XOR-Lemma* (see, e.g., [Gol11]) to show that σ is in fact arbitrarily close to the uniform distribution over $\mathbb{F}_2^b$. This allows us to finish the theorem by extending the above argument from uniform σ to almost-uniform σ.

Lemma 3.1 (Vazirani XOR-Lemma)**.** *Let σ be a distribution over $\mathbb{F}_2^b$ such that for every $y \in \mathbb{F}_q^b \setminus \{0\}$, we have $\frac{1-\eta}{2} \leq \Pr_{x\sim\sigma}[\langle x, y\rangle = 1] \leq \frac{1+\eta}{2}$. Then, σ is $\left(2^b \cdot \eta\right)$-close in* total-variation distance *to the uniform distribution over $\mathbb{F}_2^b$.*

In our case, $\Pr_{x\sim\sigma}[\langle x, y\rangle = 1] = \mathrm{wt}(By)$. Since the columns of B belong to $\mathcal{D}$ and are linearly-independent, By is a non-zero codeword of $\mathcal{D}$. Our assumption about $\mathcal{D}$ having small bias means that $\mathrm{wt}(By)$ is very close to $\frac{1}{2}$, so the hypothesis of Lemma 3.1 is satisfied. Thus, in the above calculation the rows of A are sampled i.i.d from a distribution $\sigma \sim \mathbb{F}_2^b$ which has statistical distance at most $2^b\eta$ from uniform. Therefore, we can replace the 2^{-b} term in Eq. (6) by an upper bound $(2^{-b} + 2^b\eta)$. By taking η small enough, say at most 2^{-2b}, the bound in Eq. (7) remains valid by slightly adjusting parameters (e.g., taking $\varepsilon = \frac{3}{b}$). □

We next sketch a proof for a weaker version of Theorem A.

Theorem A'. *For every $\rho \in (0, 1)$, $\varepsilon > 0$, there is some $L = L(\rho, \varepsilon) \in \mathbb{N}$ such that there exist (ρ, L) -list-decodable Reed-Solomon codes of rate at least $1 - \rho - \varepsilon$, and arbitrarily large length n.*

The above statement is weaker than Theorem A since it has no universal quantifier over the field characteristic p. As explained in Section 2.2, our proof of Theorem A works by reducing the list-decodability of the punctured RS code over $\mathbb{F}_q$ to that of a punctured trace code over $\mathbb{F}_Q$, for some suitable Q for which $\mathbb{F}_Q$ is a subfield of $\mathbb{F}_q$. In the proof of Theorem A', we take the characteristic to be a prime p such that $Q = p$ is a suitable choice for the above reduction. This greatly simplifies the proof for two reasons. First, when Q is prime, the punctured trace code can be shown to be almost surely list-decodable, rather than just quasi-list-decodable. The plain list-decodability of the punctured trace code then makes the reduction step simpler as well. In particular, Theorem 7, the subject of Section 7, is needed in the general case but not when Q is prime.

Proof sketch for Theorem A'. Let p be a prime in the range $[2^{\frac{\varepsilon}{2}}, 2^{\frac{\varepsilon}{2}+1}]$. Let q be a prime power of order of magnitude $\Theta(n^c)$, for some $c(\varepsilon) > 1$. Let φ be a random $(q \to n)$ puncturing map. Consider the code $\mathcal{C} = \varphi\left(\mathrm{RS}_{\mathbb{F}_q}(\mathbb{F}_q; k)\right)$, namely, $\mathcal{C}$ is a random puncturing of the full Reed-Solomon code of dimension k over $\mathbb{F}_q$. Suppose no coordinate of $\mathrm{RS}(\mathbb{F}_q; k)$ is sampled twice by φ for inclusion in $\mathcal{C}$. Then, $\mathcal{C}$ itself is of the form $\mathrm{RS}(S; k)$ for some set $S \subseteq \mathbb{F}_q$ of size n. For this sketch we assume that this is indeed the case. Under this assumption, it suffices to show that $\mathcal{C}$ has some positive probability of being (ρ, L)-list-decodable, for a large enough list-size $L = L(\rho, \varepsilon)$.

Let $\mathrm{tr} : \mathbb{F}_q \to \mathbb{F}_p$ denote the *trace map* (see Section 6.1 for details). Given a vector $u \in \mathbb{F}_q^m$ (for some $m \in \mathbb{N}$) let $\mathrm{tr}(u) = \langle \mathrm{tr}(u_i) \rangle_{i \in [m]} \in \mathbb{F}_p^m$. Let $\mathcal{D} = \mathrm{tr}(\mathrm{RS}(\mathbb{F}_q; k)) = \{\mathrm{tr}(u) \mid u \in \mathrm{RS}(\mathbb{F}_q; k)\} \subseteq \mathbb{F}_p^q$. It is well known that the *trace code* $\mathcal{D}$ is **almost** a small-bias code. That is, $\mathcal{D}$ contains the all-a's vector for every $a \in \mathbb{F}_p^*$, but outside of these $p-1$ vectors, $\mathcal{D}$ has bias at most $\frac{k-2}{\sqrt{q}}$ (see Corollary 6.3). Using this property, an argument in the spirit of Theorem 1 shows that $\varphi(\mathcal{D})$ is (ρ, L')-list-decodable with high probability, where we can take $L' = cL$ for some small $0 < c < 1$. The following diagram illustrates the relations between the four relevant codes:

$$\begin{tikzcd}
\mathrm{RS}(\mathbb{F}_q; k) \arrow[r, "\mathrm{tr}"] \arrow[d, "\varphi"] & \mathcal{D} \arrow[d, "\varphi"] \\
\mathcal{C} \arrow[r, "\mathrm{tr}"] & \varphi(\mathcal{D})
\end{tikzcd}$$

The remaining part of the proof reduces the list-decodability of $\mathcal{C}$ to that of $\varphi(\mathcal{D})$ (with a slightly smaller list-size). Now, suppose that $\mathcal{C}$ is not (ρ, L)-list-decodable. Then, there exist $L+1$ distinct ρ-clustered codewords $U = \{u_1, \ldots, u_{L+1}\} \subseteq \mathcal{C}$. It is not hard to see that the vectors $\mathrm{tr}(U) = \{\mathrm{tr}(u_1), \ldots \mathrm{tr}(u_{L+1})\} \subseteq \mathbb{F}_p^n$ are also ρ-clustered (this follows immediately from $\mathbb{F}_p$-linearity of the trace map; see Observation 6.11). The tr and φ maps in the above diagram commute and thus $\mathrm{tr}(U) \subseteq \varphi(\mathcal{D})$. If $|\mathrm{tr}(U)| > L'$, i.e., if applying tr to U does not result in too many collisions, then $\mathrm{tr}(U)$ is a counterexample to the (ρ, L')-list-decodability of $\varphi(\mathcal{D})$, which suffices to show that $\mathcal{C}$ is (ρ, L)-list-decodable with high probability. Unfortunately, it is possible that $\mathrm{tr}(U)$ is smaller than L'.

To overcome this final challenge we consider applying other maps to U. Concretely, for each $a \in \mathbb{F}_q$, define $f_a : \mathcal{C} \to \varphi(\mathcal{D})$ by $f_a(x) = \mathrm{tr}(a \cdot x)$. Each of these functions maps ρ-clustered sets to ρ-clustered sets. Furthermore, it is highly likely (over the choice of φ) that the family of functions $\{f_a\}_{a \in \mathbb{F}_q}$ enjoys some good hash properties (Claim 6.14 and Lemma 6.12), and, in particular, that for every set U as above there exists some $a \in \mathbb{F}_q$ such that $|f_a(U)| > L'$. Applying the above argument with f_a instead of tr yields the theorem. □

Remark 3.2 (An algorithmic reduction from list-decodability of $\mathcal{C}$ to that of $\varphi(\mathcal{D})$)**.** *Let $\alpha_1, \ldots, \alpha_d$ be a p-linear basis for $\mathbb{F}_q$ as a vector space over $\mathbb{F}_p$, where $d = \log_p q$. Let $f : \mathbb{F}_q \to \mathbb{F}_p^d$ denote the bijection $f(x) = (\mathrm{tr}(\alpha_1 x), \ldots, \mathrm{tr}(\alpha_d x))$. This map f induces a weight-preserving injection of $\mathcal{C}$ into the* interleaved code *$\varphi(\mathcal{D})^{\odot d}$, so the list-decodability of the former is reduced to that of the latter. The list-decodability of $\varphi(\mathcal{D})^{\odot d}$ can be further reduced to that of $\varphi(\mathcal{D})$ as a special case of [GGR11, Thm. 2.5], which deals with list-decodability of interleaved codes. The result of [GGR11] is algorithmic. Hence, given an algorithm to list-decode $\varphi(\mathcal{D})$, this reduction yields such an algorithm for $\mathcal{C}$ as well.*

We note that this approach only works when the underlying field of $\varphi(\mathcal{D})$ is of prime size. Hence, it cannot be used to prove the full Theorem 3. Also, even when the field size is prime, this approach *yields a worse list-size bound than the one in Theorem 3.*

Remark 3.3 (Traces are good hashes)**.** *Theorem 7 shows that $\{x \mapsto \mathrm{tr}(a \cdot x)\}_{a \in \mathbb{F}_q}$ has good properties as a hash family when applied to bounded degree polynomials over $\mathbb{F}_q$. This is a rather natural statement which may be of independent interest, so we isolate its proof in Section 7. As noted above, Theorem 7 is used to prove the general Theorem 3, but is not needed for the weaker Theorem A'.*

4 Preliminaries

4.1 General notation

We denote the uniform distribution over a finite nonempty set S by $\mathsf{U}(S)$.

For $a, b \in \mathbb{R}$, we denote $\exp_a(b) = a^b$.

The constants implied by asymptotic notation are universal unless stated otherwise. To indicate that the hidden constant may depend on, e.g., the parameter p, we write "$O_p(\cdot)$".

If $A \in \mathbb{F}_q^{m \times b}$ and $\mathcal{C} \subseteq \mathbb{F}_q^m$, we write $A \subseteq \mathcal{C}$ to mean that each column of A is a codeword in $\mathcal{C}$. Given a puncturing map $\varphi : \mathbb{F}_q^m \to \mathbb{F}_q^n$, let $\varphi(A)$ denote the matrix obtained from A by applying φ to each column.

4.2 A characterization of linear list-decodable linear codes

Recall the notion of a ρ-clustered set (Definition 2.4.)

Definition 4.1. *Fix $\rho \in [0,1]$, $L \in \mathbb{N}$. A matrix $A \in \mathbb{F}_q^{n \times b}$ ($1 \le b \le L$) with* $\operatorname{rank} A = b$ *is* (ρ, L)-span-clustered *if the column-span of A contains a ρ-clustered set of size L.*

Note that for a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ we have

$$\mathcal{C} \text{ is } \textit{not } (\rho, L)\text{-list-decodable} \iff \exists A \in \mathbb{F}_q^{n \times b} \text{such that } A \text{ is } (\rho, L+1)\text{-span-clustered and } A \subseteq \mathbb{C} \ .$$

Furthermore, we can always take b to be in the range $[\log_q(L+1), L+1]$. Indeed, a matrix of rank smaller than $\log_q(L+1)$ cannot be $(\rho, (L+1))$-span-clustered since its span has cardinality smaller than $L+1$. On the other hand, a rank larger than $L+1$ is never needed since, given a ρ-clustered set $W \subseteq \mathcal{C}$ with $|W| = L+1$, one can take A to be a matrix whose columns are a maximal linearly independent subset of W.

4.3 The scalar-multiplied code $\Lambda\mathcal{C}$ and scalar-expanded code $\mathcal{D}^*$

Let

$$\Gamma_n = \left\{\Lambda \in \mathbb{F}_q^{n \times n} \mid \Lambda \text{ is diagonal and of full-rank}\right\} \ .$$

The following is immediate.

Observation 4.2. [7] *Let $\mathcal{C} \in \mathbb{F}_q^n$ be a code. Fix a matrix $\Lambda \in \Gamma_n$ and let $\Lambda\mathcal{C} = \{\Lambda u \mid u \in \mathcal{C}\}$. Then, for any $\rho \in [0,1]$ and $L \in \mathbb{N}$, we have*

$$\mathcal{C} \textit{ is } (\rho, L)\textit{-list-decodable} \iff \Lambda\mathcal{C} \textit{ is } (\rho, L)\textit{-list-decodable} \ .$$

[7] Definition 8.3 more generally discusses the class of *scalar-invariant* code properties, namely, these are the code properties for which Observation 4.2 holds.

The question of the list-decodability of $\mathcal{C}$ thus reduces to that of any code of the form $\Lambda\mathcal{C}$. To take advantage of this reduction, we shall study the list-decodability of $\Lambda\mathcal{C}$ where $\Lambda \sim \mathsf{U}(\Gamma_n)$.

If $\mathcal{C}$ is a random puncturing of some code $\mathcal{D}$, we can realize the code $\Lambda\mathcal{C}$ as a puncturing of the code $\mathcal{D}^*$, which we now define.

Definition 4.3. *Given $u \in \mathbb{F}_q^m$, let $u^* \in \mathbb{F}_q^{m(q-1)}$ denote the vector*

$$u^* = \bigodot_{a \in \mathbb{F}_q^*} (au) \ ,$$

where $\bigodot$ stands for concatenation of vectors. Given a matrix $B \in \mathbb{F}_q^{m \times b}$ with columns $a_1, \dots, a_b$, let $B^* \in \mathbb{F}_q^{m(q-1)\times b}$ be the matrix whose columns are $a_1^*, \dots, a_b^*$. Denote $\mathcal{D}^* = \{u^* \mid u \in \mathcal{D}\} \subseteq \mathbb{F}_q^{m(q-1)}$.

Observation 4.4. *The code $\Lambda\mathcal{C}$, where $\Lambda \sim \mathsf{U}(\Gamma_n)$ and $\mathcal{C}$ is a random n-puncturing of $\mathcal{D}$, is distributed identically to a random n-puncturing of $\mathcal{D}^*$.*

4.4 Fourier transform

We recall the following elementary facts about the Fourier transform[8] of a function $f : \mathbb{F}_q^b \to \mathbb{C}$.

Definition 4.5 (Fourier (and inverse Fourier) transform)**.** *Suppose that $q = p^r$ for some prime p, and let $\omega = e^{\frac{2\pi i}{p}}$. Let $b \in \mathbb{N}$ and let $f : \mathbb{F}_q^b \to \mathbb{C}$. Then $\widehat{f} : \mathbb{F}_q^b \to \mathbb{C}$ is defined by*

$$\widehat{f}(y) = \sum_{x \in \mathbb{F}_q^b} f(x) \cdot \overline{\chi_y(x)}, \quad \textit{where} \quad \chi_y(x) = \omega^{\operatorname{tr}\langle x,y\rangle} \ .$$

Here, $\operatorname{tr} : \mathbb{F}_q \to \mathbb{F}_p$ *stands for the field trace function* $\operatorname{tr}(x) = \sum_{i=0}^{r-1} x^{p^i}$. *We also have the Fourier inversion formula:*

$$f(x) = q^{-b} \sum_{y \in \mathbb{F}_q^b} \widehat{f}(y)\chi_y(x) \ .$$

Fact 4.6 (Parseval's identity)**.** *Let $f, g : \mathbb{F}_q^b \to \mathbb{C}$. Then,*

$$\sum_{x \in \mathbb{F}_q^b} f(x)\overline{g(x)} = \mathbb{E}_{y \sim \mathsf{U}(\mathbb{F}_q^b)} \left[\widehat{f}(y)\overline{\widehat{g}(y)}\right] \ .$$

In particular, $\sum_{x \in \mathbb{F}_q^b} |f(x)|^2 = \mathbb{E}_{y \sim \mathsf{U}(\mathbb{F}_q^b)} \left[\left|\widehat{f}(y)\right|^2\right]$.

4.5 Entropy and KL-divergence

Let τ be a distribution over a finite set. The base-q entropy of τ is

$$H_q(\tau) = -\sum_{x \in \operatorname{supp}(\tau)} \tau(x) \log_q \tau(x) \ .$$

[8] Our convention is to use counting norm for f and expectation norm for $\hat{f}$.

Given $x \in [0,1]$, we write

$$h_q(x) = -x \log_q x - (1-x)\log_q(1-x) + x\log_q(q-1)$$

for the base-q entropy of a random variable over $\{0,\dots,q-1\}$, which takes 0 with probability $1-x$ and each $i \in \{1,\dots,q-1\}$ with probability $\frac{x}{q-1}$.

The q-ary Kullback-Leibler Divergence of two distributions τ, σ over a finite set S is

$$D_{\mathrm{KL}q}(\tau \parallel \sigma) = \sum_{s \in S} \tau(s) \log_q \frac{\tau(s)}{\sigma(s)} \quad .$$

4.6 The empirical distribution of the rows of a matrix

Definition 4.7. *Given a vector $a \in \mathbb{F}_q^n$ we define its* empirical distribution Emp_a *over $\mathbb{F}_q$ by*

$$\mathsf{Emp}_a(x) = \Pr_{i \in [n]}[a_i = x] \quad .$$

More generally, given $A \in \mathbb{F}_q^{n \times b}$, let Emp_A denote its empirical row distribution*, that is, the distribution over $\mathbb{F}_q^b$ defined by*

$$\mathsf{Emp}_A(x) = \Pr_{i \in [n]}[A_i = x] \quad ,$$

where A_i denotes the i'th row of A.

Fact 4.8 ([CT06, Thm. 11.1.4])**.** *Let $X \in \mathbb{F}_q^{n \times b}$ have rows sampled identically and independently from some distribution σ over $\mathbb{F}_q^b$. Then, for any distribution τ over $\mathbb{F}_q^b$,*

$$\Pr[\mathsf{Emp}_A = \tau] \le q^{-D_{\mathrm{KL}q}(\tau \| \sigma) \cdot n} \quad .$$

Definition 4.9. *Let τ be a distribution over $\mathbb{F}_q^b$. We denote* $\dim(\tau) = \dim \operatorname{supp}(\tau)$*. If* $\dim(\tau) = b$*, we say that τ is a* full-rank *distribution.*

Definition 4.10 (Matrix of a particular distribution)**.** *Let τ be a distribution over $\mathbb{F}_q^b$ (where $b \in \mathbb{N}$). For $n \in \mathbb{N}$, we denote*

$$\mathcal{M}_{n,\tau} = \left\{ A \in \mathbb{F}_q^{n \times b} \mid \mathsf{Emp}_A = \tau \right\} \quad .$$

A distribution τ over $\mathbb{F}_q^b$ is said to be n-feasible if $\tau(x) \cdot n$ is an integer for all $x \in \mathbb{F}_q^b$. Observe that any n-feasible distribution over $\mathbb{F}_q^b$ corresponds to a partition of n identical balls into q^b buckets. The bound below thus follows immediately.

Fact 4.11. *The number of n-feasible distributions over $\mathbb{F}_q^b$ is at most $(n+1)^{q^b}$.*

Clearly, n-feasibility of τ is a necessary condition for $\mathcal{M}_{n,\tau}$ to be nonempty. When this condition holds, $|\mathcal{M}_{n,\tau}|$ is equal to the multinomial coefficient $\frac{n!}{\prod_{x \in \mathbb{F}_q^b}(\tau(x)n)!}$. By standard bounds on multinomial coefficients, we have

$$n^{-O(q^b)} \cdot q^{n \cdot H_q(\tau)} \le |\mathcal{M}_{n,\tau}| \le q^{n \cdot H_q(\tau)} \quad . \tag{8}$$

4.6.1 The Fourier transform of an empirical distribution

We record several useful properties of the function $\widehat{\mathsf{Emp}_A}$ for a given matrix A. The following is immediate.

Fact 4.12. *A vector $u \in \mathbb{F}_q^n$ is η-biased ($\eta > 0$) if and only if $\left|\widehat{\mathsf{Emp}_u}(a)\right| \le \eta$ for all $a \in \mathbb{F}_q^*$.*

The following identity shows that the Fourier transform of Emp_A (where $A \in \mathbb{F}_q^{n\times b}$) is in fact composed of the Fourier transforms of Emp_{Ay} over $y \in \mathbb{F}_q^b$. Let $a \in \mathbb{F}_q$. Then,

$$\widehat{\mathsf{Emp}_A}(ay) = \sum_{x\in\mathbb{F}_q^b} \mathsf{Emp}_A(x)\omega^{-\mathrm{tr}(a\langle x,y\rangle)} = \mathbb{E}_{x\sim\mathsf{Emp}_A}\left[\omega^{-\mathrm{tr}(a\langle x,y\rangle)}\right] = \mathbb{E}_{z\sim\mathsf{Emp}_{Ay}}\left[\omega^{-\mathrm{tr}(az)}\right] = \widehat{\mathsf{Emp}_{Ay}}(a)\ . \tag{9}$$

By Fact 4.6, the normalized Hamming Weight of a vector $u \in \mathbb{F}_q^n$ can be conveniently expressed in terms of the Fourier transform of Emp_u.

$$\mathrm{wt}(u) = \sum_{x\in\mathbb{F}_q} \mathbf{1}_{x\neq 0}\cdot\mathsf{Emp}_u(x) = \frac{q-1}{q}\cdot\widehat{\mathsf{Emp}_u}(0) - \frac{1}{q}\cdot\sum_{a\in\mathbb{F}_q^*}\widehat{\mathsf{Emp}_u}(a) = \frac{q-1}{q} - \frac{1}{q}\cdot\sum_{a\in\mathbb{F}_q^*}\widehat{\mathsf{Emp}_u}(a). \tag{10}$$

This yields the following relation between bias and weight.

Lemma 4.13. *Let $u \in \mathbb{F}_q^n$ be η-biased for some $\eta > 0$. Then*

$$\frac{q-1}{q}(1-\eta) \le \mathrm{wt}(u) \le \frac{q-1}{q}(1+\eta)\ .$$

Proof. By Eq. (10) and Fact 4.12,

$$\left|\mathrm{wt}(u) - \frac{q-1}{q}\right| = \left|\frac{1}{q}\cdot\sum_{a\in\mathbb{F}_q^*}\widehat{\mathsf{Emp}_u}(a)\right| \le \frac{q-1}{q}\cdot\eta\ . \quad\square$$

We have the following immediate conclusion.

Lemma 4.14. *For any $\eta \ge 0$, an η-biased code also has η-optimal distance.*

5 A random puncturing of a near-optimal-distance code is likely to be list-decodable

5.1 GHK list-decodability bound for random linear codes revisited

The main result of [GHK11] gives bounds on the list-size for list-decoding of RLCs up to capacity. Here, we go deeper and slightly reformulate[9] the main technical claim of that paper.

[9] See Remark 5.2 for the differences in our formulation.

Theorem 5.1 ([GHK11, Thm. 6.1])**.** *Let q be a prime power and let $\rho \in (0, 1-1/q)$. Then, there is a constant $K' = K'_{\rho,q} \geq 1$ such that, for all $b, L \in \mathbb{N}$, we have*

$$\left|\left\{A \in \mathbb{F}_q^{n\times b} \mid A \textit{ is } (\rho, L+1)\textit{-span-clustered}\right\}\right| \leq q^{(bh_q(\rho)-4)\cdot n}$$

whenever $L \geq K' \cdot b$ and n is large enough, and

$$\left|\left\{A \in \mathbb{F}_q^{n\times b} \mid A \textit{ is } (\rho, L+1)\textit{-span-clustered}\right\}\right| \leq q^{(bh_q(\rho)+1)\cdot n} \tag{11}$$

in general.

Furthermore, one can take

$$K' \leq \exp\left(O\left(\frac{(\log_2 q)^2}{\min\left\{(1-1/q-\rho)^2, \rho\right\}}\right)\right) . \tag{12}$$

Remark 5.2. *There are several differences between our formulation of the theorem and the one that appears in [GHK11]. We list and justify them here.*

(i) The random vectors $X_1, \dots, X_\ell$ from the original formulation have become the columns of the matrix A, and we changed the name ℓ to b.

(ii) The original statement of [GHK11, Thm. 6.1] only deals with matrices whose span contains a large set clustered around 0. In our statement we already apply the reduction to a ball with arbitrary center, which appears in [GHK11, Thm. 2.1].

(iii) Eq. (11) is a rather naive bound, originally derived as part of the proof of [GHK11, Thm. 2.1].

(iv) The asymptotic statement about $K'_{\rho,q}$ comes from inspecting the proof in [GHK11]. Specifically, in the notation of that paper, [GHK11, Lem. 6.3] yields a 2-increasing chain of length $d = \Omega(\log_q K')$ whenever $L \geq K' \cdot b$. The exponent in the q-ary analog of [GHK11, Lem 4.1] satisfies $\delta_p = \Theta\left(\frac{\min\left\{\rho, \left(1-\frac{1}{q}-\rho\right)^2\right\}}{\log_2 q}\right)$. Finally, the requirement in [GHK11, Thm. 6.1] is that K' be large enough so that $d \cdot \delta_p \geq \Omega(1)$.

It will be convenient to formulate a corollary from Theorem 5.1 in terms of the row-distributions of certain matrices.

Definition 5.3. *Fix a prime power q. Let $b, n \in N$ and let τ be an n-feasible distribution over $\mathbb{F}_q^b$. If a matrix $A \in \mathcal{M}_{n,\tau}$ is $(\rho, L+1)$-span-clustered, we say that τ is $(\rho, L+1)$*-span-clustered (with regard to n).

Remark 5.4. *Observe that the notion of τ being $(\rho, L+1)$-span clustered is well defined, and in particular does not depend on the choice of A in Definition 5.3. In other words, either every matrix in $\mathcal{M}_{n,\tau}$ is $(\rho, L+1)$-span-clustered, or non of them are. Indeed, suppose that $A \in \mathbb{F}_q^{n\times b}$ is $(\rho, L+1)$-span-clustered with regard to some center $z \in \mathbb{F}_q^n$, and let B be a matrix obtained from A by permuting the rows of the latter according to some permutation π over $[n]$. Then, B is $(\rho, L+1)$-span-clustered with regard to the center vector resulting from applying π to z.*

This idea is generalized in Definition 8.3, with the concept of a row-symmetric *code property.*

Corollary 5.5. *In the setting of Theorem 5.1, every $(\rho, L+1)$-span-clustered (with regard to n), n-feasible distribution τ over $\mathbb{F}_q^b$ satisfies*

$$H_q(\tau) \le b \cdot \left(h_q(\rho) + \frac{5K'_{\rho,q}}{L}\right) - 3 \ .$$

for every b and n such that $\frac{n}{\log_q n} \ge \omega\left(q^{L+1}\right)$.

Proof. By Remark 5.4, $\mathcal{M}_{n,\tau} \subseteq \left\{A \in \mathbb{F}_q^{n\times b} \mid A \text{ is } (\rho, L+1)\text{-span-clustered}\right\}$. Thus, Eq. (8) and Theorem 5.1 yield the following:

If $L \ge K'_{\rho,q} \cdot b$:

$$H_q(\tau) \le \log_q |\mathcal{M}_{n,\tau}| + O\left(\frac{q^b \cdot \log_q n}{n}\right) \le bh_q(\rho) - 4 + O\left(\frac{q^b \cdot \log_q n}{n}\right)$$

If $L < K'_{\rho,q} \cdot b$:

$$\begin{aligned} H_q(\tau) &\le \log_q |\mathcal{M}_{n,\tau}| + O\left(\frac{q^b \cdot \log_q n}{n}\right) \le bh_q(\rho) + 1 + O\left(\frac{q^b \cdot \log_q n}{n}\right) \\ &\le bh_q(\rho) + \frac{5bK'_{\rho,q}}{L} - 4 + O\left(\frac{q^b \cdot \log_q n}{n}\right) \ . \end{aligned}$$

The claim now follows from our assumption that $\frac{n}{\log_q n} \ge \omega\left(q^{L+1}\right)$. □

5.2 Proof of Theorem 1

Let us restate Theorem 1 before proving it.

Theorem 1 (A puncturing of a near-optimal distance code is whp list-decodable up to capacity)**.** *Fix a prime power q. Let $L, n \in \mathbb{N}$ and $0 < \rho < \frac{q-1}{q}$, such that $\frac{n}{\log_q n} \ge \omega\left(q^{L+1}\right)$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a linear code with η-optimal distance, where $\eta = q^{-L+1}$. Let $\mathcal{C}$ be a random n-puncturing of $\mathcal{D}$ of design rate R, where $R \le 1 - h_q(\rho) - \frac{K}{L}$ for some constant $K = K_{\rho,q}$. Then,*

$$\Pr\left[\mathcal{C} \text{ is } (\rho, L)\text{-list-decodable}\right] \ge 1 - q^{-\Omega(n)} \ .$$

Furthermore, one can take

$$K_{\rho,q} \le \exp\left(O\left(\frac{(\log q)^2}{\min\left\{(1-1/q-\rho)^2, \rho\right\}}\right)\right) \tag{1}$$

and, in particular, $K_{\rho,q} \le \operatorname{poly}(q)$ whenever ρ is bounded away from 0 and $1 - \frac{1}{q}$.

Our proof of Theorem 1 relies on the following lemma, which we prove below in Section 5.3.

Lemma 5.6 (Puncturings of large-distance codes are locally similar to random linear codes)**.** *Fix $b \in \mathbb{N}$ and a full-rank distribution τ over $\mathbb{F}_q^b$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a linear code of η-optimal distance ($\eta \geq 0$). Let $\Lambda \sim \mathsf{U}(\Gamma_n)$ and, independently, let φ be a random ($m \to n$) puncturing map. Denote $R = \frac{\log_q |\mathcal{D}|}{n}$. Then,*

$$\mathbb{E}\left[\left|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \Lambda \cdot \varphi(\mathcal{D})\}\right|\right] \leq \exp_q\Big(n \cdot \Big(H_q(\tau) - (1-R)b + \log_q \Big(1 + \eta q^b\Big) + \log_q 2\Big)\Big)$$

Remark 5.7. *Lemma 5.6 bounds the expected number of τ-distributed matrices in the code $\Lambda \cdot \varphi(\mathcal{D})$. The lemma says that this number is not much larger than the expected number of τ-distributed matrices in a random linear code of similar rate. Indeed, for a given matrix A, the probability of A being contained in the random linear code $C_{\mathrm{RLC}}^{n,q}(R)$ is $q^{-n(1-R)\cdot \mathrm{rank}(A)}$. Thus, by Eq. (8),*

$$\mathbb{E}\left[\left|\left\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq C_{\mathrm{RLC}}^{n,q}(R)\right\}\right|\right] = |\mathcal{M}_{n,\tau}| \cdot q^{n(R-1)\cdot b} \approx q^{n(H_q(\tau)-(1-R)b)} \ .$$

Using Lemma 5.6, we conclude Theorem 1 from Theorem 5.1.

Proof of Theorem 1. Take $K_{\rho,q} = 5K'$, where $K'_{\rho,q}$ is as in Theorem 5.1. We need to show that

$$\Pr\left[\mathcal{C} \text{ is not } (\rho, L)\text{-list-decodable}\right] \leq q^{-\Omega(\varepsilon n)}.$$

By Observation 4.2, it suffices to show instead that

$$\Pr\left[\Lambda\mathcal{C} \text{ is not } (\rho, L)\text{-list-decodable}\right] \leq q^{-\Omega(\varepsilon n)} \ , \tag{13}$$

where the matrix Λ is sampled uniformly from Γ_n.

Now, if $\Lambda\mathcal{C}$ is *not* (ρ, L)-list-decodable, then $\Lambda\mathcal{C}$ contains some $(\rho, L+1)$-span-clustered matrix $A \in \mathbb{F}_q^{n\times b}$ for some b, $\log_q(L+1) \leq b \leq L+1$. Hence,

$$\begin{aligned}
&\Pr\left[\Lambda\mathcal{C} \text{ is not } (\rho, L)\text{-list-decodable}\right] \\
&\leq \sum_{b=\lceil \log_q(L+1)\rceil}^{L+1} \Pr\left[\exists A \in \mathbb{F}_q^{n\times b} \text{ s.t. } A \text{ is } (\rho, L+1)\text{-span-clustered and } A \subseteq \Lambda\mathcal{C}\right] \\
&\leq \sum_{b=\lceil \log_q(L+1)\rceil}^{L+1} \mathbb{E}\left[\left|\left\{A \in \mathbb{F}_q^{n\times b} \mid A \text{ is } (\rho, L+1)\text{-span-clustered and } A \subseteq \Lambda\mathcal{C}\right\}\right|\right] \ .
\end{aligned}$$

By Remark 5.4, we can write

$$\left\{A \in \mathbb{F}_q^{n\times b} \mid A \text{ is } (\rho, L+1)\text{-span-clustered}\right\} = \bigcup_{\tau \in T_b} \mathcal{M}_{n,\tau}$$

where T_b is a set of n-feasible distributions over $\mathbb{F}_q^b$. Therefore, by Lemma 5.6 and our assumption that $\eta \le q^{-L+1} \le q^{-b}$, the probability that $\Lambda\mathcal{C}$ is not (ρ, L)-list-decodable is at most

$$\begin{aligned}
&\sum_{b=\lceil \log_q(L+1)\rceil}^{L+1} \sum_{\tau\in T_b} \mathbb{E}\left[|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \Lambda\mathcal{C}\}|\right] \\
&\le \sum_{b=\lceil \log_q(L+1)\rceil}^{L+1} \sum_{\tau\in T_b} \exp_q\Big(n \cdot \Big(H_q(\tau) - (1-R)b + \log_q(1+\eta q^b) + \log_q 2\Big)\Big) \\
&\le \sum_{b=\lceil \log_q(L+1)\rceil}^{L+1} \sum_{\tau\in T_b} \exp_q(n\cdot (H_q(\tau) - (1-R)b + 2)) \ .
\end{aligned}$$

By Corollary 5.5, each term of the inner sum is at most q^{-n}. Therefore, by Fact 4.11,

$$\Pr\left[\Lambda\mathcal{C} \text{ is not } (\rho, L)\text{-list-decodable}\right] \le \sum_{b=\lceil \log_q(L+1)\rceil}^{L+1} \sum_{\tau\in T_b} q^{-n} \le q^{-n}\sum_{b=1}^{L+1}(n+1)^{q^b} \le q^{-n}(L+1)(n+1)^{q^{L+1}} \ ,$$

and the theorem follows due to our assumption that $\frac{n}{\log_q n} \ge \omega\left(q^{L+1}\right)$. □

5.3 Proof of Lemma 5.6: random puncturings of large-distance codes are similar to RLCs

Lemma 5.6 follows from Lemmas 5.8 to 5.10, stated and proven below. The proof of Lemma 5.6 is then completed at the end of this section.

Lemma 5.8 is a variation of the *Vazirani XOR-Lemma* (see [Gol11], and Lemma 3.1 for a special case). Given a distribution σ over $\mathbb{F}_q^b$, the XOR-Lemma relates the *total-variation distance* of σ from the uniform distribution over $\mathbb{F}_q^b$, to the maximum of $|\widehat{\sigma}(y)|$ over all $y \neq 0$. In Lemma 5.8, rather than taking a maximum, we consider the ℓ_1 norm of $\widehat{\sigma}$, which yields a tighter bound when only a small number of entries of $\widehat{\sigma}$ are large in absolute value.

Lemma 5.8. *Fix a prime power q, and $b \in \mathbb{N}$. Let σ be a distribution over $\mathbb{F}_q^L$ and let $f : \mathbb{F}_q^b \to \mathbb{R}$ be a non-negative function. Then,*

$$\mathbb{E}_{x\sim\sigma}\left[f(x)\right] \le \left(\sum_{y\in\mathbb{F}_q^b} |\widehat{\sigma}(y)|\right) \cdot \mathbb{E}_{x\sim \mathsf{U}(\mathbb{F}_q^b)}\left[f(x)\right] \ .$$

Proof. We have

$$\sigma(x) = q^{-b}\sum_{y\in\mathbb{F}_q^b} \widehat{\sigma}(y)\omega^{-\mathrm{tr}(\langle x,y\rangle)} \le q^{-b}\sum_{y\in\mathbb{F}_q^b}|\widehat{\sigma}(y)|$$

for all $x \in \mathbb{F}_q^b$. So

$$\mathbb{E}_{x\sim\sigma}\left[f(x)\right] = \sum_{x\in\mathbb{F}_q^b}\sigma(x)f(x) \le q^{-b}\left(\sum_{y\in\mathbb{F}_q^b}|\widehat{\sigma}(y)|\right)\cdot\left(\sum_{x\in\mathbb{F}_q^b} f(x)\right) = \left(\sum_{y\in\mathbb{F}_q^b}|\widehat{\sigma}(y)|\right)\cdot\mathbb{E}_{x\sim\mathsf{U}(\mathbb{F}_q^b)}\left[f(x)\right] \ . \quad \square$$

We next bound the expectation of an arbitrary non-negative test function over the empirical row-distribution of a given matrix B, assuming that the column-span of B has good bias or distance. The bias based bound is an immediate application of Lemma 5.8. The weight based bound requires an additional trick, and only yields a result relating to the row-distribution of B^* rather than B itself (recall Definition 4.3 for a reminder about B^*). One reason for the difference between the two cases is that under the weight-based hypothesis we have an upper bound only on the entries of the Fourier transform (Eq. (16)), rather than on their absolute value.

Lemma 5.9. *Let $B \in \mathbb{F}_q^{m\times b}$ have $\operatorname{rank} B = b$, and let $f : \mathbb{F}_q^b \to \mathbb{R}$ be a non-negative function. Then, the following holds for all $\eta \geq 0$:*

1. *Suppose that the column-span of B (as a code in $\mathbb{F}_q^m$) is η-biased. Then,*
$$\mathbb{E}_{x\sim \mathsf{Emp}_B}\left[f(x)\right] \leq (1+q^b\eta)\cdot \mathbb{E}_{x\sim \mathsf{U}(\mathbb{F}_q^b)}\left[f(x)\right] \ .$$
2. *Suppose that the column-span of B has η-optimal distance. Then,*
$$\mathbb{E}_{x\sim \mathsf{Emp}_{B^*}}\left[f(x)\right] \leq 2(1+q^b\eta)\cdot \mathbb{E}_{x\sim \mathsf{U}(\mathbb{F}_q^b)}\left[f(x)\right] \ .$$

Proof. We first prove Item 1. Hence, by Lemma 5.8 it suffices to show that
$$\sum_{y\in\mathbb{F}_q^b}\left|\widehat{\mathsf{Emp}_B}(y)\right| \leq 1+q^b\eta \ .$$

By Eq. (9), the above is equivalent to
$$\sum_{y\in\mathbb{F}_q^b}\left|\widehat{\mathsf{Emp}_{By}}(1)\right| \leq 1+q^b\eta \ . \tag{14}$$

For $y=0$ we have $\widehat{\mathsf{Emp}_{By}}(1) = \widehat{\mathsf{Emp}_0}(1) = 1$. For any $y \in \mathbb{F}_q^b\setminus\{0\}$, since B has full column-rank, By is a non-zero codeword of $\mathcal{D}$. By hypothesis, By is η-biased, so Fact 4.12 yields $\left|\widehat{\mathsf{Emp}_{By}}(1)\right| \leq \eta$, establishing Eq. (14).

We now turn to Item 2. Let σ denote the distribution, over $\mathbb{F}_q^b$, of the random variable $a\cdot x$, where $a \sim \mathsf{U}(\mathbb{F}_q^*)$ and x is independently sampled from Emp_B. By Lemma 5.8, to prove Item 2 it suffices to show that
$$\sum_{y\in\mathbb{F}_q^b}\left|\widehat{\sigma}(y)\right| \leq 2\left(1+q^b\eta\right) \ . \tag{15}$$

By Eq. (10) followed by Eq. (9),
$$\operatorname{wt}(By) = \frac{q-1}{q} - \frac{1}{q}\cdot\sum_{a\in\mathbb{F}_q^*}\widehat{\mathsf{Emp}_{By}}(a) = \frac{q-1}{q} - \frac{1}{q}\cdot\sum_{a\in\mathbb{F}_q^*}\widehat{\mathsf{Emp}_B}(ay) = \frac{q-1}{q}\cdot\left(1-\widehat{\sigma}(y)\right) \ ,$$

so $\widehat{\sigma}(y) = 1-\frac{q}{q-1}\cdot\operatorname{wt}(By)$.

In particular, if $y \neq 0$ then By is a non-zero element in the column-span of B. Hence, by hypothesis,

$$\widehat{\sigma}(y) = 1 - \frac{q}{q-1} \cdot \mathrm{wt}(By) \leq \eta \ . \tag{16}$$

Let $P = \{y \in \mathbb{F}_q^b \mid \widehat{\sigma}(y) \geq 0\}$ and $N = \mathbb{F}_q^b \setminus P$. By Eq. (16),

$$\sum_{y \in P \setminus \{0\}} \widehat{\sigma}(y) \leq q^b \eta \ .$$

Note that $\widehat{\sigma}(0) = \sum_{x \in \mathbb{F}_q^b} \sigma(x) = 1$, and thus,

$$\sum_{y \in P} \widehat{\sigma}(y) = 1 + \sum_{y \in P \setminus \{0\}} \widehat{\sigma}(y) \leq 1 + q^b \eta \ .$$

Consequently,

$$0 \leq q^b \cdot \sigma(0) = \sum_{y \in \mathbb{F}_q^b} \widehat{\sigma}(y) = \sum_{y \in P} \widehat{\sigma}(y) + \sum_{y \in N} \widehat{\sigma}(y) \leq 1 + q^b \eta + \sum_{y \in \mathbb{N}} \widehat{\sigma}(y)$$

and so,

$$\sum_{y \in N} |\widehat{\sigma}(y)| = -\sum_{y \in N} \widehat{\sigma}(y) \leq 1 + q^b \eta \ .$$

Eq. (15) now follows since

$$\sum_{y \in \mathbb{F}_q^b} |\widehat{\sigma}(y)| = \sum_{y \in P} |\widehat{\sigma}(y)| + \sum_{y \in N} |\widehat{\sigma}(y)| \leq 2(1 + q^b \eta) \ . \quad \square$$

Lemma 5.10 bounds the probability of a random puncturing of a given matrix B having a certain empirical distribution τ. Due to the concavity argument in Eq. (18), this lemma gives tighter bounds when Emp_B is close to the uniform distribution over $\mathbb{F}_q^b$. Notably, as Lemma 5.9 shows, good bias or similar properties of the column-span of B ensure that Emp_B is indeed close to uniform.

Lemma 5.10. *Fix some distribution τ over $\mathbb{F}_q^b$. Let $B \in \mathbb{F}_q^{m \times b}$ have $\operatorname{rank} B = b$. Let $\varphi : \mathbb{F}_q^m \to \mathbb{F}_q^n$ be a random puncturing map. Then,*

$$\Pr\left[\varphi(B) \in \mathcal{M}_{n,\tau}\right] \leq \exp_q\!\big(n\left(\log_q \mathbb{E}_{x \sim \mathsf{Emp}_B}\left[\tau(x)\right] + H_q(\tau)\right)\big) \ .$$

Proof. By Fact 4.8,

$$\Pr\left[\varphi(B) \in \mathcal{M}_{n,\tau}\right] = \Pr\left[\mathsf{Emp}_{\varphi(B)} = \tau\right] \leq q^{-n \cdot D_{\mathrm{KL}_q}(\tau \| \mathsf{Emp}_B)} \ . \tag{17}$$

By concavity of log,

$$\begin{aligned} D_{\mathrm{KL}_q}\left(\tau \parallel \mathsf{Emp}_B\right) = \sum_{x \in \mathbb{F}_q^b} \tau(x) \log_q \frac{\tau(x)}{\mathsf{Emp}_B(x)} &= -H_q(\tau) - \sum_{x \in \mathbb{F}_q^b} \tau(x) \log_q \mathsf{Emp}_B(x) \\ &\geq -H_q(\tau) - \log_q \mathbb{E}_{x \sim \mathsf{Emp}_B}\left[\tau(x)\right] \ . \end{aligned} \tag{18}$$

The claim follows from Eq. (17) and Eq. (18). $\square$

Proof of Lemma 5.6. By Observation 4.4, $\Lambda \cdot \varphi(\mathcal{D})$ is distributed identically to $\varphi^*(\mathcal{D}^*)$, where φ^* is a random $((q-1)m \to n)$ puncturing map. Thus,

$$\begin{aligned}\mathbb{E}\left[\left|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \Lambda \cdot \varphi(\mathcal{D})\}\right|\right] &= \mathbb{E}\left[\left|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \varphi^*(\mathcal{D}^*)\}\right|\right] \\ &\le \mathbb{E}\left[\left|\left\{B \in \mathbb{F}_q^{m\times b} \mid B \subseteq \mathcal{D} \text{ and } \varphi^*(B^*) \in \mathcal{M}_{n,\tau}\right\}\right|\right] \ . \qquad (19)\end{aligned}$$

We proceed to bound the expectation of the right-hand side.

Suppose that $\varphi^*(B^*) \in \mathcal{M}_{n,\tau}$. Because τ is of full-rank, we have $\operatorname{rank} B = \operatorname{rank} B^* \ge \operatorname{rank} \varphi^*(B^*) = b$, so $\operatorname{rank} B = b$.

Let $B \in \mathbb{F}_q^{m\times b}$ such that $\operatorname{rank} B = b$ and $B \subseteq \mathcal{D}$. Since the column-span of B is contained in $\mathcal{D}$, it is of η-optimal distance. Hence, by Item 2 of Lemma 5.9,

$$\mathbb{E}_{x\sim \mathsf{Emp}_{B^*}}\left[\tau(x)\right] \le \Upsilon \ , \qquad (20)$$

where $\Upsilon = 2q^{-b}\left(1+q^b\eta\right)$. Lemma 5.10 yields

$$\begin{aligned}\mathbb{E}\left[\left|\left\{B \in \mathbb{F}_q^{m\times b} \mid B \subseteq \mathcal{D} \text{ and } \varphi^*(B^*) \in \mathcal{M}_{n,\tau}\right\}\right|\right] &= \sum_{\substack{B\in\mathbb{F}_q^{m\times b} \\ B\subseteq\mathcal{D} \\ \operatorname{rank} B = b}} \Pr_{\Lambda,\varphi}\left[\varphi^*(B^*) \in \mathcal{M}_{n,\tau}\right] \\ &\le q^{bRn} \cdot \exp_q\!\left(n\left(\log_q \Upsilon + H_q(\tau)\right)\right) \ , \qquad (21)\end{aligned}$$

and the claim follows from Eqs. (19), (20) and (21). □

6 List-decodable Reed-Solomon codes

In this section we prove Theorem 3, restated below.

Theorem 3 (RS codes list-decodable up to capacity)**.** *For every $\rho \in (0,1)$ and $0 < \varepsilon < \min\{\rho, 1-\rho\}$, any prime p and any $n \in \mathbb{N}$ large enough in terms of p and ε, there exist $L = L(p,\rho,\varepsilon) \in \mathbb{N}$ and $q \le O_{p,\varepsilon}\left(n^{\max\left\{2, \frac{1}{\rho-\varepsilon}\right\}}\right)$, which is a power of p, for which the following holds: Let $S \subseteq \mathbb{F}_q$ be a uniformly sampled subset of size n. Then, $\mathrm{RS}_{\mathbb{F}_q}(S; (1-\rho-\varepsilon)n)$ is (ρ, L)-list-decodable with probability $1 - p^{-\Omega(n)}$.*

Furthermore, one can take

$$L \le \exp\left(O\left(\frac{\left(\log p + \frac{1}{\varepsilon}\right)^2}{\min\{(1-\rho)^2, \rho\}}\right)\right) \ . \qquad (2)$$

As explained in Section 3, the algebraic part of the proof can be significantly simplified if one is content with a version of Theorem 3 which works only for certain choices of the field characteristic p. Below, we prove the full version of the theorem.

6.1 Preliminaries for Theorem 3: The field trace map and the trace code

We recall some facts about RS codes and the field trace map. Fix a prime p. Let $q = p^t$ and $Q = p^r$ for some $r, t \in \mathbb{N}$ such that r divides t. Then $\mathbb{F}_Q$ is a subfield of $\mathbb{F}_q$. We use $\mathrm{tr}_{q\to Q} : \mathbb{F}_q \to \mathbb{F}_Q$ to denote the $\mathbb{F}_Q$-linear trace map $\mathrm{tr}_{q\to Q}(x) = \sum_{i=0}^{\frac{t}{r}-1} x^{Q^i}$. We also allow $\mathrm{tr}_{q\to Q}$ to operate element-wise on vectors and matrices over $\mathbb{F}_q$.

Recall that $\mathbb{F}_q$ can be viewed as a t/r-dimensional linear space over the field $\mathbb{F}_Q$. It is well known that every $\mathbb{F}_Q$-linear functional $\mathbb{F}_q \to \mathbb{F}_Q$ is of the form $x \mapsto \mathrm{tr}_{q\to Q}(\alpha x)$ for some $\alpha \in \mathbb{F}_q$.

Trace maps also behave well with respect to towers of fields, namely,

$$\mathrm{tr}_{q\to p} = \mathrm{tr}_{Q\to p} \circ \mathrm{tr}_{q\to Q} \ . \tag{22}$$

This yields the following Fourier identity about the trace map on vectors.

Lemma 6.1. *Suppose that $\mathbb{F}_Q$ is a subfield of $\mathbb{F}_q$. Let $u \in \mathbb{F}_q^n$, $a \in \mathbb{F}_Q$, and denote $v = \mathrm{tr}_{q\to Q}(u)$. Then,*

$$\widehat{\mathsf{Emp}_v}(a) = \widehat{\mathsf{Emp}_u}(a) \ .$$

Proof. By Eq. (22),

$$\widehat{\mathsf{Emp}_v}(a) = \frac{1}{n}\sum_{i=1}^{n} \omega^{-\mathrm{tr}_{Q\to p}\left(a\cdot \mathrm{tr}_{q\to Q}(u_i)\right)} = \frac{1}{n}\sum_{i=1}^{n} \omega^{-\mathrm{tr}_{Q\to p}\left(\mathrm{tr}_{q\to Q}(a\cdot u_i)\right)} = \frac{1}{n}\sum_{i=1}^{n} \omega^{-\mathrm{tr}_{q\to p}(a\cdot u_i)} = \widehat{\mathsf{Emp}_u}(a).$$

□

Say that a vector is constant if all of its entries are identical. We recall the famous Weil-Carlitz-Uchiyama bound [CU57] (see also [LN96, Thm. 5.38] and subsequent discussion).

Theorem 6.2 (Weil-Carlitz-Uchiyama)**.** *Let q be a prime power. For every nonzero word $u \in \mathbb{F}_q^q$ in the full Reed-Solomon code $\mathrm{RS}\left(\mathbb{F}_q; k\right)$, and every $a \in \mathbb{F}_q$, either the vector $\mathrm{tr}_{q\to p}(a\cdot u) = \langle \mathrm{tr}(au_i)\rangle_{i\in[q]}$ is constant, or*

$$\left|\widehat{\mathsf{Emp}_u}(a)\right| = \frac{1}{q}\left|\sum_{i=1}^{q} \omega^{-\mathrm{tr}_{q\to p}(au_i)}\right| \le \frac{k-2}{\sqrt{q}} \ ,$$

where $\omega = e^{\frac{2\pi i}{p}}$.

We have the following corollary on the distance of the trace code. By Delsarte's connection [Del75], the dual of the trace code is the subfield subcode of the dual RS code, so the following is related to the distance bound for dual-BCH codes.

Corollary 6.3 (Small bias of the full $q \to p$ trace code)**.** *Let q be a power of some prime p and fix $1 \le k \le \sqrt{q}+2$. Then, every non-constant vector in the trace code $\mathrm{tr}_{q\to p}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q; k\right)\right) \subseteq \mathbb{F}_p^q$ is η-biased for $\eta = \frac{k-2}{\sqrt{q}}$. Furthermore, the code $\mathrm{tr}_{q\to p}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q; k\right)\right)$ has η-optimal distance.*

Proof. Let $u \in \mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q; k\right)$ such that $v := \mathrm{tr}_{q\to p}(u)$ is non-constant, and let $a \in \mathbb{F}_p^*$. By Lemma 6.1, $\widehat{\mathsf{Emp}_v}(a) = \widehat{\mathsf{Emp}_u}(a)$. Note that $\mathrm{tr}_{q\to p}(a\cdot u) = a\cdot \mathrm{tr}_{q\to p}(u)$ is also non-constant, so $\left|\widehat{\mathsf{Emp}_u}(a)\right| \le \eta$ by Theorem 6.2. Consequently, Fact 4.12 implies that v is η-biased.

Due to the above and Lemma 4.13, every non-constant vector in $\mathrm{tr}_{q\to p}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q; k\right)\right)$ has η-optimal distance, establishing the distance claim as well. □

Remark 6.4. *Corollary 6.3 does not necessarily hold for the trace code* $\mathrm{tr}_{q\to Q}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q;k\right)\right)$ *if* Q *is not prime. The above proof fails because it is possible, e.g., to have a non-constant vector* $v \in \mathrm{tr}_{q\to Q}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q;k\right)\right)$ *such that* $\mathrm{tr}_{Q\to p}(av)$ *is constant for some specific* $a \in \mathbb{F}_Q^*$.

For example, suppose that $Q = p^2$ *and* $k > p$*. Consider the vector* $v = \mathrm{tr}_{q\to Q}(u)$ *where* $u \in \mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q;k\right)$ *is defined by* $u_x = x - x^p$*. Then,* $v_x = \sum_{i=0}^{\log_p q-1} x^{p^i}(-1)^i$*, so* v *is non-constant, but* $\mathrm{tr}_{Q\to p}(v) = \mathrm{tr}_{q\to p}(u) = 0$ *is constant.*

The following definitions will help us control the set of vectors for which Corollary 6.3 fails with regard to the $\mathrm{tr}_{q\to Q}$ map, when Q is not necessarily prime.

Definition 6.5 (Constant traces)**.** *Let* u *be a vector over* $\mathbb{F}_q$*, where* q *is a power of some prime* p*. Define*

$$\mathsf{Const}_u = \{a \in \mathbb{F}_q \mid \mathrm{tr}_{q\to p}(au) \text{ is a constant vector}\} \ . \tag{23}$$

Observe that Const_u *is an* $\mathbb{F}_p$*-linear subspace of* $\mathbb{F}_q$.

Definition 6.6 (Trace-friendliness)**.** *Fix a prime* p *and let* Q *be a power of* p*. A vector* $v \in \mathbb{F}_Q^n$ *is called* trace-friendly *if* $\mathrm{tr}_{Q\to p}(av)$ *is non-constant for all* $a \in \mathbb{F}_Q^*$.

A matrix $A \in \mathbb{F}_Q^{n\times b}$ *is said to be* trace-friendly *if* $\operatorname{rank} A = b$*, and every non-constant vector in the column-span of* A *is trace-friendly.*

The following lemma is a straightforward generalization of Corollary 6.3 to the case of not necessarily prime Q.

Lemma 6.7 (Small bias of the full $q \to Q$ trace code, subject to trace-friendliness)**.** *Let* q, Q *be powers of some prime* p *such that* $\mathbb{F}_Q$ *is a subfield of* $\mathbb{F}_q$*. Fix* $1 \le k \le \sqrt{q}+2$*. Then, every non-constant trace-friendly vector in the trace code* $\mathrm{tr}_{q\to Q}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q;k\right)\right) \subseteq \mathbb{F}_Q^q$ *is* η*-biased for* $\eta = \frac{k-2}{\sqrt{q}}$*. Furthermore, if* $A \in \mathbb{F}_Q^{q\times b}$ *is a trace-friendly matrix with* $A \subseteq \mathrm{tr}_{q\to Q}\left(\mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q;k\right)\right)$*, then the column-span of* A *has* η*-optimal distance.*

Proof. Let $u \in \mathrm{RS}_{\mathbb{F}_q}\left(\mathbb{F}_q;k\right)$ such that $v := \mathrm{tr}_{q\to Q}(u)$ is non-constant and trace-friendly, and let $a \in \mathbb{F}_Q^*$. Since, v is trace-friendly, $\mathrm{tr}_{Q\to p}(av)$ is non-constant. Lemma 6.1 and Theorem 6.2 now yield $\widehat{\mathsf{Emp}_v}(a) = \widehat{\mathsf{Emp}_u}(a) \le \eta$. By virtue of Fact 4.12, v is η-biased.

For the distance claim, the above implies that every vector in the column-span of A is either constant or η-biased. In the latter case, the vector is also of η-optimal weight, due to Lemma 4.13. The claim follows. □

6.2 Proof of Theorem 3

6.2.1 Notation and parameters

Write $R = 1-\rho-\varepsilon$ and $k = Rn$. Let Q be a power of p such that $2^{\frac{2}{\varepsilon}} \le Q \le p \cdot 2^{\frac{2}{\varepsilon}}$. Let $L = \frac{20K'}{\varepsilon(1-\rho)}$, where $K' = K'_{\rho,Q}$ is as in Theorem 5.1. Let

$$L' = \frac{20K'}{\varepsilon} = (1-\rho)\cdot L \ .$$

Suppose that n is large enough so that $\frac{\varepsilon n}{\log_p n} \geq \omega\left(Q^L\right)$. Let q be the minimum power of p such that $\mathbb{F}_Q$ is a subfield of F_q (equivalently, $\log_p Q$ should divide $\log_p q$) and

$$q \geq \max\left\{n^2, \left(2kQ^{L+1}\right)^2, k^{\frac{1}{1-R-2\varepsilon}}, 3 \cdot \binom{L+1}{2}^{\frac{1}{\varepsilon}}\right\} \ . \tag{24}$$

Note that

$$q \leq O_{p,\varepsilon}\left(n^{\max\left\{2, \frac{1}{\rho-\varepsilon}\right\}}\right) \ . \tag{25}$$

We claim that the upper bound on L in Eq. (2) holds. Since $\varepsilon < 1-\rho$, we have

$$1 - 1/Q - \rho \geq 1 - 2^{-\frac{2}{\varepsilon}} - \rho \geq 1 - \rho - 2^{-\frac{2}{1-\rho}} \geq \Omega(1-\rho) \ .$$

Hence, by Eq. (12),

$$K'_{\rho,Q} \leq \exp\left(O\left(\frac{(\log Q)^2}{\min\left\{(1-1/q-\rho)^2, \rho\right\}}\right)\right) \leq \exp\left(O\left(\frac{\left(\log p + \frac{1}{\varepsilon}\right)^2}{\min\left\{(1-\rho)^2, \rho\right\}}\right)\right) \ ,$$

establishing Eq. (2).

Let φ be a random $q \to n$ puncturing map. Write $\mathcal{C} = \varphi(\mathrm{RS}\left(\mathbb{F}_q; k\right))$ and $\mathcal{D} = \mathrm{tr}_{q\to Q}(\mathrm{RS}\left(\mathbb{F}_q; k\right))$. Note that the maps φ and $\mathrm{tr}_{q\to Q}$ commute. In other words, we have the following commuting diagram.[10]

$$\begin{tikzcd}
\mathrm{RS}\left(\mathbb{F}_q; k\right) \arrow[r, "\mathrm{tr}_{q\to Q}"] \arrow[d, "\varphi"] & \mathcal{D} \arrow[d, "\varphi"] \\
\mathcal{C} \arrow[r, "\mathrm{tr}_{q\to Q}"] & \varphi(\mathcal{D})
\end{tikzcd} \tag{26}$$

6.2.2 Main technical lemmas

Definition 6.8 (Quasi-list-decodability of $\varphi(D)$)**.** *The code $\varphi(\mathcal{D})$ is said to be (ρ, L')*-quasi-list-decodable *if there does not exist any trace-friendly matrix A such that $A \subseteq \mathcal{D}$ and $\varphi(A)$ is $(\rho, L'+1)$ -span-clustered. In other words, quasi-list-decodability means that, while $\varphi(\mathcal{D})$ may contain a ρ-clustered set of size $L'+1$, such a set cannot originate from a trace-friendly matrix.*

The proof of Theorem 3 relies on Lemmas 6.9 and 6.10, proven in Sections 6.3 and 6.4, respectively.

Lemma 6.9. *In the setting of Section 6.2.1, consider the following events:*

1. *$\mathcal{C}$ is (ρ, L) -list-decodable.*
2. *$\varphi(\mathcal{D})$ is not (ρ, L') -quasi-list-decodable.*

The probability that ***at least one of*** *the above events occurs is $1 - q^{-\Omega(\varepsilon n)}$.*

Lemma 6.10. *In the setting of Section 6.2.1, with probability at least $1 - Q^{-\Omega(n)}$, the code $\varphi(\mathcal{D})$ is (ρ, L') -quasi-list-decodable.*

In both lemmas, the probability is taken over the random choice of the puncturing map φ.

[10] We overload the puncturing map notation φ to be from $\mathbb{F}_q^q \to \mathbb{F}_q^n$ on the left and from $\mathbb{F}_Q^q \to \mathbb{F}_Q^n$ on the right; this should cause no confusion.

6.2.3 Concluding Theorem 3

Let T denote the event that $\mathcal{C}$ is (ρ, L)-list-decodable. It follows immediately from Lemmas 6.9 and 6.10 that $\Pr[T] \geq 1 - Q^{-\Omega(n)} - q^{-\Omega(\varepsilon n)} \geq 1 - Q^{-\Omega(n)}$.

Let J denote the event that each coordinate of $\mathrm{RS}(\mathbb{F}_q; k)$ is sampled at most once for inclusion in $\mathcal{C}$. Since $q \geq n^2$, the event J occurs with some positive probability bounded away from 0. Since J has probability bounded away from 0, we also have $\Pr[T \mid J] \geq 1 - \frac{1-\Pr[T]}{\Pr[J]} \geq 1 - Q^{-\Omega(n)}$. Observe that, conditioned on J, the distribution of $\mathcal{C}$ is uniform on the set of all q-ary Reed-Solomon codes of degree k and length n. Hence, such a Reed-Solomon code, chosen uniformly at random, is (ρ, L)-list-decodable with probability $1 - Q^{-\Omega(n)}$. Theorem 3 follows.

6.3 Proof of Lemma 6.9: quasi-list-decodability of $\varphi(\mathcal{D})$ almost surely implies list-decodability of $\mathcal{C}$

We begin with the following simple observation.

Observation 6.11 (The trace map preserves clustering)**.** *Fix $a \in \mathbb{F}_q$. Let $U = \{u_1, \dots, u_L\}$ be a set of ρ-clustered vectors $\mathbb{F}_q^n$. Then, $U' = \{\mathrm{tr}_{q\to Q}(a \cdot u_1), \dots, \mathrm{tr}_{q \to Q}(a \cdot u_L)\} \subseteq \mathbb{F}_Q^n$ is also a ρ-clustered set. Note, however, that U' may be smaller than U.*

Proof. Define $f : \mathbb{F}_q^n \to \mathbb{F}_Q^n$ by $f(u) = \mathrm{tr}_{q\to Q}(a \cdot u)$. Note that $\mathrm{wt}(u) \geq \mathrm{wt}(f(u))$ for all u. Let $z \in \mathbb{F}_q^n$ such that $\mathrm{wt}(u_i - z) \leq \rho$ for all $i \in [L]$. Then,

$$\mathrm{wt}\left(f(u_i) - f(z)\right) = \mathrm{wt}\left(f(u_i - z)\right) \leq \mathrm{wt}(u_i - z) \leq \rho. \quad \square$$

To prove Lemma 6.9 we will assume that $\mathcal{C}$ is not list-decodable, so that it contains some large ρ-clustered set U. We will use Observation 6.11 to map U to a ρ-clustered set $U' = \mathrm{tr}_{q\to Q}(a \cdot U)$. Provided that U' is large enough, it can serve as a witness for the non quasi-list-decodability of $\varphi(\mathcal{D})$. The main challenge is showing that, conditioned on some very likely event, we can always choose some $a \in \mathbb{F}_q$ for which U' is large. For this last part, we require Lemma 6.12 and Theorem 7, stated below.

Lemma 6.12 (Whp, $\{u \mapsto \mathrm{tr}_{q\to p}(au)\}_{a\in\mathbb{F}_q}$ is a good hash family on $\mathcal{C}$)**.** *Fix a prime p and let q be a power of p. Fix $\varepsilon > 0$. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a random n-puncturing of $\mathrm{RS}(\mathbb{F}_q; k)$, of rate $R \in (0, 1)$. Suppose that $k \leq q^{1-R-2\varepsilon}$ and that $n \geq \omega\left(\log_p q + \frac{1}{\varepsilon}\right)$. Then, with probability $1 - q^{-\Omega(\varepsilon n)}$, every non-constant $u \in \mathcal{C}$ has*

$$\dim \mathsf{Const}_u < (1 - \varepsilon) \cdot \log_p q \ .$$

Proof. Let $t = \log_p q$. Let $\mathcal{C} = \varphi(\mathrm{RS}(\mathbb{F}_q; k))$, where φ is a random $q \to n$ puncturing map. We claim that

$$\Pr_{\varphi}\left[\dim \mathsf{Const}_{\varphi(v)} \geq (1-\varepsilon)t\right] \leq q^{-(R+\varepsilon)n+\varepsilon t+1} \ . \tag{27}$$

for each non-constant $v \in \mathrm{RS}(\mathbb{F}_q; k)$. Eq. (27) implies the lemma by the union bound on all non-constant vectors in $\mathrm{RS}(\mathbb{F}_q; k)$, of which there are at most q^{Rn}.

Note that

$$\Pr\left[\dim \mathsf{Const}_{\varphi(v)} \geq (1-\varepsilon)t\right] \leq \sum_{\substack{U \subseteq \mathbb{F}_q \\ U \text{ is } \mathbb{F}_p\text{-linear} \\ \dim U = (1-\varepsilon)t}} \Pr\left[U \subseteq \mathsf{Const}_{\varphi(u)}\right] \ .$$

The sum on the right-hand side has at most $q^{\varepsilon t}$ terms. Thus, to prove Eq. (27), it suffices to show that

$$\Pr\left[U \subseteq \mathsf{Const}_{\varphi(v)}\right] \le q^{-(R+\varepsilon)n+1} \tag{28}$$

whenever U is a p-linear subspace of $\mathbb{F}_q$ with $\dim U = (1-\varepsilon)t$.

Fix a basis $a_1, \dots, a_d$ for U over $\mathbb{F}_p$, where $d = (1-\varepsilon)t$. Now,

$$\Pr\left[U \subseteq \mathsf{Const}_{\varphi(v)}\right] = \sum_{z \in \mathbb{F}_p^d} \Pr\left[\forall i \in [d] \;\; \mathrm{tr}(a_i \varphi(v)) \text{ is the constant } z_i \text{ vector}\right] \tag{29}$$

Observe that there are at most kq^{ε} entries $1 \le j \le q$ such that $\mathrm{tr}(a_i \cdot v_j) = z_i$ for all $1 \le i \le d$. Indeed, there are $p^{\varepsilon t} = q^{\varepsilon}$ elements $x \in \mathbb{F}_q$ such that $\mathrm{tr}(a_i x) = z_i$ for all $1 \le i \le d$. Due to the distance property of Reed-Solomon codes, each of these q^{ε} elements can appear in at most k entries of v.

Now, the event inside the sum in Eq. (29) occurs if and only if the above $\le kq^{\varepsilon}$ entries of v are the *only* ones sampled for inclusion in $\varphi(v)$. This happens with probability at most

$$\left(\frac{kq^{\varepsilon}}{q}\right)^n = q^{\left(\log_q k + \varepsilon - 1\right)n} \le q^{-(R+\varepsilon)n} \enspace .$$

Thus, by Eq. (29),

$$\Pr\left[U \subseteq \mathsf{Const}_{\varphi(v)}\right] \le q^{1-\varepsilon} \cdot q^{-(R+\varepsilon)n} \le q^{-(R+\varepsilon)n+1} \enspace .$$

Eq. (28) follows, and consequently, so do Eq. (27) and the lemma. □

Theorem 7 (The trace maps $\{u \mapsto \mathrm{tr}_{q\to p}(au)\}_{a\in\mathbb{F}_q}$ form a good hash family on the full RS code)**.** *Let q be a power of some prime p. Let $1 \le k \le \sqrt{q}$. Then, every non-constant vector $u \in \mathrm{RS}\left(\mathbb{F}_q; k\right)$ has*

$$\dim \mathsf{Const}_u < \log_p(k-1) \enspace .$$

Remark 6.13. *Theorem 7 is proven in Section 7. In fact, we prove a stronger version of the theorem than the one stated above, namely, we also give a bound on the number of codewords $u \in \mathrm{RS}\left(\mathbb{F}_q; k\right)$ such that $\mathsf{Const}_u = d$, for any $d \le \log_p(k-1)$. This stronger claim is not needed in the present work but may be of independent interest.*

We now show that Lemma 6.9 follows from Observation 6.11, Lemma 6.12, and Theorem 7. It is straightforward to verify that the parameter settings in Section 6.2.1 satisfy the hypothesis of Lemma 6.12. Suppose that every non-constant $u \in \mathcal{C}$ has

$$\dim \mathsf{Const}_u < (1-\varepsilon)\log_p q \enspace . \tag{30}$$

By Lemma 6.12, to prove Lemma 6.9 it suffices to show that the conclusion of the latter holds deterministically under the assumption of Eq. (30). Hence, we now assume that $\mathcal{C}$ is not (ρ, L)-list-decodable and prove that $\varphi(\mathcal{D})$ is nor (ρ, L')-quasi-list-decodable.

Since $\mathcal{C}$ is not (ρ, L)-list-decodable, there exists some matrix $B \in \mathbb{F}_q^{q\times b}$ $(b \le L+1)$ with $B \subseteq \mathrm{RS}\left(\mathbb{F}_q; k\right)$ and $\operatorname{rank} B = b$, such that $\varphi(B)$ is $(\rho, L+1)$-span-clustered. Let $\{u_1, \dots, u_{L+1}\} \in \mathbb{F}_q^n$ denote a ρ-clustered set of $L+1$ distinct vectors in the column-span of $\varphi(B)$.

We need the following claim, proven below.

Claim 6.14. *There exists some $a \in \mathbb{F}_q$ with the following properties:*

1. *The set $\{\mathrm{tr}_{q\to Q}(a \cdot u_i) \mid 1 \le i \le L+1\}$ is of cardinality at least $L' + 1 = (1-\rho)L + 1$.*
2. *The matrix $\mathrm{tr}_{q\to Q}(a \cdot B)$ is trace-friendly.*

Let $a \in \mathbb{F}_q$ be as in Claim 6.14. Let $A = \mathrm{tr}_{q\to Q}(a \cdot B)$. By assumption, A is trace-friendly. By Eq. (26), $\varphi(A) = \mathrm{tr}_{q\to Q}(\varphi(B))$. In particular, the column-span of $\varphi(A)$ contains the vectors $\mathrm{tr}_{q\to Q}(a \cdot u_1), \dots, \mathrm{tr}_{q\to Q}(a \cdot u_{L+1})$. By assumption, this list contains at least $L'+1$ distinct vectors, and by Observation 6.11, they are ρ-clustered. Therefore, $\varphi(A)$ is $(\rho, L'+1)$-span-clustered, and so $\varphi(\mathcal{D})$ is not (ρ, L')-quasi-list-decodable, proving Lemma 6.9 modulo Theorem 7 and Claim 6.14.

Proof of Claim 6.14. Pick a uniformly at random from $\mathbb{F}_q$. We bound from below the probability, over the choice of a, of a having each of the properties stated in the claim.

Property 1. Write $F = \{u_i \mid 1 \le i \le L+1\}$. Let $G_1, \dots, G_s \subseteq F$ be the equivalence classes of F with regard to the relation

$$u \sim v \iff u - v \text{ is a constant vector} \ .$$

We claim that each set G_i $(1 \le i \le s)$ is of size at most $\frac{1}{1-\rho}$. Indeed, since F is ρ-clustered, there exists some $z \in \mathbb{F}_q^n$ such that $\mathrm{wt}(u - z) \le \rho$ for all $u \in F$. In particular,

$$\sum_{u \in G_i} \mathrm{wt}(u - z) \le |G_i| \cdot \rho \ .$$

On the other hand, since every distinct pair of vectors $u, v \in G_i$ disagree on every coordinate, at most one vector in G_i can agree with z on a given coordinate. Hence,

$$\sum_{u \in G_i} \mathrm{wt}(u - z) \ge |G_i| - 1 \ .$$

Therefore, $|G_i| \le \frac{1}{1-\rho}$. In particular, it follows that that $s \ge (1-\rho) \cdot (L+1) > L'$.

Hence, to achieve Property 1, it suffices that $\mathrm{tr}_{q\to Q}(a \cdot u) \ne \mathrm{tr}_{q\to Q}(a \cdot v)$ for every $u, v \in F$ such that $u \nsim v$. Equivalently, we need $\mathrm{tr}_{q\to Q}(a \cdot (u - v)) \ne 0$. By our assumption that $u \nsim v$, the vector $u - v$ is non-constant. Hence, by Eq. (30), there are at least $q - q^{1-\varepsilon}$ choices of a for which $\mathrm{tr}_{q\to p}(a \cdot (u - v))$, and hence $\mathrm{tr}_{q\to Q}(a \cdot (u - v))$ is non-constant, and, a fortiori, non-zero. Therefore, by the union bound over all pairs $u, v \in F$ with $u \nsim v$, Property 1 holds with probability at least

$$1 - \binom{L+1}{2} \cdot q^{-\varepsilon} \ge \frac{2}{3} \ ,$$

due to our assumed bound on q in Eq. (24).

Property 2. Write $A = \mathrm{tr}_{q\to Q}(aB)$. Let $Y = \left\{y \in \mathbb{F}_Q^b \mid By \text{ is non-constant}\right\}$. For each $y \in Y$, let X_y be an indicator for the event that $\mathrm{tr}_{Q\to p}(Ay)$ is constant. Note that $Ay = \mathrm{tr}_{q\to Q}(aBy)$ is constant for all $y \in \mathbb{F}_Q^b \setminus Y$. Hence, for A to be trace-friendly, it suffices that $X_y = 0$ for all $y \in Y$.

For $y \in Y$, we have

$$\mathrm{tr}_{Q\to p}(Ay) = \mathrm{tr}_{Q\to p}\left(\mathrm{tr}_{q\to Q}(aB)y\right) = \mathrm{tr}_{Q\to p}\left(\mathrm{tr}_{q\to Q}(aBy)\right) = \mathrm{tr}_{q\to p}\left(a(By)\right) \quad ,$$

where the last equality is due to Eq. (22). By assumption, By is a non-constant vector. Applying Theorem 7 to By yields

$$\Pr\left[X_y = 1\right] = \frac{p^{\dim \mathsf{Const}_{By}}}{q} \le \frac{k}{q} \quad .$$

Hence, by the union bound on all $y \in Y$

$$\Pr\left[A \text{ is trace-friendly}\right] \ge 1 - |Y| \cdot \frac{k}{q} \ge 1 - Q^b \cdot \frac{k}{q} \ge 1 - Q^{L+1}\frac{k}{q} \ge \frac{2}{3} \quad ,$$

where the last inequality is due to the assumption about field size q from Eq. (24).

We conclude that a uniformly random a simultaneously satisfies Properties 1 and 2 with probability at least $\frac{1}{3}$. In particular, some a satisfying both properties exists. □

6.4 Proof of Lemma 6.10: quasi-list-decodability of $\varphi(\mathcal{D})$

Lemma 6.10 follows from similar principles to those by which we proved Theorem 1. By Observations 4.2 and 4.4 and Markov's bound, it suffices to show that

$$\mathbb{E}\left[\left|\left\{A \in \mathbb{F}_Q^{q\times b} \mid A \text{ is trace-friendly, } A \subseteq \mathcal{D} \text{ and } \varphi^*(A^*) \text{ is } \left(\rho, L'+1\right)\text{-span-clustered}\right\}\right|\right] \le Q^{-\Omega(n)} \quad ,$$

where φ^* is a random $((Q-1)q \to n)$ puncturing-map.

The left-hand side of the above is equal to

$$\sum_{b \le L'+1} \sum_{\tau} \mathbb{E}\left[\left|\left\{A \in \mathbb{F}_Q^{q\times b} \mid A \text{ is trace-friendly, } A \subseteq \mathcal{D} \text{ and } \varphi^*(A^*) \in \mathcal{M}_{n,\tau}\right\}\right|\right] \quad ,$$

where the inner sum goes over all n-feasible $(\rho, L'+1)$-span-clustered distributions τ over $\mathbb{F}_Q^b$. By Fact 4.11, the total number of terms (including both the inner and outer sums) is at most $p^{O\left(\log_p n \cdot Q^{L'+1}\right)} \le Q^{o(n)}$, so it suffices to bound each term separately by $Q^{-\Omega(n)}$. Namely, the lemma would follow if we show that

$$\mathbb{E}\left[\left|\left\{A \in \mathbb{F}_Q^{q\times b} \mid A \text{ is trace-friendly, } A \subseteq \mathcal{D} \text{ and } \varphi^*(A^*) \in \mathcal{M}_{n,\tau}\right\}\right|\right] \le Q^{-\Omega(n)} \quad . \tag{31}$$

for every $b \le L'+1$ and each $(\rho, L'+1)$-span-clustered distribution τ over $\mathbb{F}_Q^b$.

Let $\eta = \frac{k-1}{\sqrt{q}}$. By Eq. (24), $\eta \le Q^{-b}$. By Lemma 6.7, the column-span of A has η-optimal distance. Let $Y = \left\{A \in \mathbb{F}_Q^{q \times b} \mid A \text{ is trace-friendly, } A \subseteq \mathcal{D}\right\}$ and note that $|Y| \le Q^{bRn}$. Now,

$$\begin{aligned}
&\mathbb{E}\left[|\{A \in Y \mid \varphi^*(A^*) \in \mathcal{M}_{n,\tau}\}|\right] \\
&\le \sum_{A \in Y} \Pr\left[\varphi^*(A^*) \in \mathcal{M}_{n,\tau}\right] \\
&\le \sum_{A \in Y} \exp_Q\!\left(n\left(\log_Q \mathbb{E}_{x \sim \mathsf{Emp}_{A^*}}\left[\tau(x)\right] + H_Q(\tau)\right)\right) && \text{by Lemma 5.10} \\
&\le \sum_{A \in Y} \exp_Q\!\left(n\left(\log_Q\left(2Q^{-b} \cdot \left(1 + Q^b \eta\right)\right) + H_Q(\tau)\right)\right) && \text{by Lemma 5.9} \\
&\le \exp_Q\!\left(n\left(bR + \log_Q\left(2Q^{-b} \cdot \left(1 + Q^b \eta\right)\right) + H_Q(\tau)\right)\right) \\
&\le \exp_Q\!\left(n\left(bR - b + 2\log_Q 2 + H_Q(\tau)\right)\right) \\
&\le \exp_Q\!\left(n\left(-b(1-R) + 2\log_Q 2 + bh_Q(\rho) + \frac{5bK'}{L'} - 3\right)\right) && \text{by Corollary 5.5} \\
&\le \exp_Q\!\left(n\left(-b(1-R) - 1 + bh_Q(\rho) + \frac{5bK'}{L'}\right)\right) \\
&= \exp_Q\!\left(n\left(-b(\rho+\varepsilon) - 1 + bh_Q(\rho) + \frac{\varepsilon b}{4}\right)\right) && \text{since } L' = \frac{20K'}{\varepsilon} \\
&= \exp_Q\!\left(n\left(-b\varepsilon - 1 + \frac{\varepsilon}{2} + \frac{\varepsilon b}{4}\right)\right) && \text{since } h_Q(\rho) - \rho \le \frac{1}{\log_2 Q} \\
&\le Q^{-n} \ ,
\end{aligned}$$

establishing Eq. (31) and the lemma.

7 Proof of Theorem 7: trace maps form a good hash family for the full RS code

We now prove the following expanded version of Theorem 7.

Theorem 7 (expanded) (Codewords of the full RS code have many non-constant traces)**.** *Fix $q = p^t$ for some prime p and $t \in \mathbb{N}$. Let $k \le \sqrt{q}$ and let $0 \le d \le t$. Then, every non-constant vector $u \in \mathrm{RS}(\mathbb{F}_q; k)$ has $\dim \mathsf{Const}_u < \log_p(k-1)$. Moreover, for $d < \log_p(k-1)$, we have*

$$|\{u \in \mathrm{RS}(\mathbb{F}_q; k) \mid \dim \mathsf{Const}_u \ge d\}| \le \exp_q\!\left(d + k - \frac{(k-1)d}{\log_p(k-1)}\right) \ . \tag{32}$$

In this proof we use the shorthand $\mathrm{tr} = \mathrm{tr}_{q \to p}$. Fix some $\mathbb{F}_p$-linear subspace $V \subseteq \mathbb{F}_q$. Note that $\{u \in \mathrm{RS}(\mathbb{F}_q; k) \mid V \subseteq \mathsf{Const}_u\}$ is an $\mathbb{F}_p$-linear space. We claim that

$$\dim_{\mathbb{F}_p}\{u \in \mathrm{RS}(\mathbb{F}_q; k) \mid V \subseteq \mathsf{Const}_u\} \le t \cdot \max\left\{k - \frac{(k-1)\dim V}{\log_p(k-1)}, 1\right\} \ . \tag{33}$$

We now show that Eq. (33) implies the lemma. Note that $u \in \mathrm{RS}\left(\mathbb{F}_q; k\right)$ is constant if and only if $\mathsf{Const}_u = \mathbb{F}_q$. Now, if $\log_p(k-1) \leq \dim V < t$, then

$$\begin{aligned}|\{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid \mathsf{Const}_u = V\}| &\leq |\{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid V \subseteq \mathsf{Const}_u\}| - |\{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid u \text{ is constant}\}| \\ &= |\{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid V \subseteq \mathsf{Const}_u\}| - q \leq p^t - q = 0 \ ,\end{aligned}$$

where the last inequality is due to Eq. (33). Hence, every con-constant $u \in \mathrm{RS}\left(\mathbb{F}_q; k\right)$ has $\dim u < \log_p(k-1)$. Next, let $d < \log_p(k-1)$. Then,

$$\begin{aligned}|\{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid \dim \mathsf{Const}_u \geq d\}| &\leq \sum_{\substack{V \subseteq \mathbb{F}_q \text{ is } \mathbb{F}_p\text{-linear} \\ \dim_{\mathbb{F}_p} V = d}} |\{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid V \subseteq \mathsf{Const}_u\}| \\ &\leq q^d \cdot \exp_p\left(t\left(k - \frac{(k-1)d}{\log_p(k-1)}\right)\right) \\ &= \exp_q\left(d + k - \frac{(k-1)d}{\log_p(k-1)}\right) \ ,\end{aligned}$$

establishing Eq. (32).

We turn to proving Eq. (33). Let $u \in \mathrm{RS}\left(\mathbb{F}_q; k\right)$ and let $F(x) = \sum_{i=0}^{k-1} a_i x^i$ be the polynomial associated with u (here, $a_0, \ldots, a_{k-1} \in \mathbb{F}_q$).

Let $O_1, \ldots, O_s \subseteq \{0, \ldots, q-2\}$ be the equivalence classes of the relation

$$r \sim s \iff \exists j \in \mathbb{N} \ \ rp^j \equiv a \mod (q-1) \ .$$

Note that $|O_i| \leq t$ for all $1 \leq i \leq s$. Assume without loss of generality, that $O_s = \{0\}$. We need the following claim, proven at the end of this section.

Claim 7.1. *For each $1 \leq i \leq s$, either $O_i \cap \{0, 1, \ldots, k-1\} = \emptyset$, or there exist some $\gamma_i \in O_i$ and $1 \leq \mu_i \leq \log_p(k-1)$ such that $|O_i \cap \{0, 1 \ldots, (k-1)\}| = \mu_i$ and*

$$O_i \cap \{1, \ldots, (k-1)\} = \{\gamma_i \cdot p^j \mid 0 \leq j \leq \mu_i - 1\} \ . \tag{34}$$

Now, for any $\beta \in \mathbb{F}_q$,

$$\begin{aligned}\mathrm{tr}\left(\beta F(x)\right) = \sum_{i=0}^{k-1}\sum_{d=0}^{t-1}(\beta a_i x^i)^{p^d} &= \sum_{i=1}^{s}\sum_{j=0}^{\mu_i-1}\sum_{d=0}^{t-1}\left(\beta \cdot a_{\gamma_i \cdot p^j} \cdot x^{\gamma_i \cdot p^j}\right)^{p^d} \\ &= \sum_{i=1}^{s}\sum_{j=0}^{\mu_i-1}\sum_{d=0}^{t-1}\left(\beta \cdot a_{\gamma_i \cdot p^j}\right)^{p^d} \cdot x^{\gamma_i p^{j+d}} \ .\end{aligned}$$

Re-indexing $(j+d, j) \to (j, d)$, and using the convention that if r is a negative integer we take p^r to mean p^s for some non-negative s satisfying $s \equiv r \mod (q-1)$, we can rewrite the above expression as

$$\sum_{i=1}^{s}\sum_{j=0}^{\mu_i+t-2} x^{\gamma_i \cdot p^j}\left(\sum_{d=0}^{\mu_i-1}\left(\beta \cdot a_{\gamma_i \cdot p^d}\right)^{p^{j-d}}\right) = \sum_{i=1}^{s}\sum_{j=0}^{\mu_i+t-2} x^{\gamma_i \cdot p^j}\lambda_{i,\beta,j} \ , \tag{35}$$

where $\lambda_{i,\beta,j} = \sum_{d=0}^{\mu_i - 1} \left(\beta \cdot a_{\gamma_i \cdot p^d}\right)^{p^{j-d}}$.

By Eq. (35), the polynomial $\mathrm{tr}(\beta F(x))$ is constant in x if and only if $\lambda_{i,\beta,j} = 0$ for all $1 \leq i \leq s-1$ and $0 \leq j \leq \mu_i + t - 2$. Since $\lambda_{i,\beta,j} = \lambda_{i,\beta,0}^{p^j}$ for all $0 \leq j \leq |O_i| - 1$, we have the equivalent statement

$$\mathrm{tr}\left(\beta \cdot F(x)\right) \text{ is constant in } x \quad \Longleftrightarrow \quad \lambda_{i,\beta} = 0 \ \ \forall 1 \leq i \leq s-1 \tag{36}$$

where $\lambda_{i,\beta} = \lambda_{i,\beta,0}$. Simplifying this further, we can write

$$b_{i,d} = a_{\gamma_i \cdot p^d}^{p^{-d}} \tag{37}$$

and

$$\lambda_{i,\beta} = \lambda_{i,\beta,0} = \sum_{d=0}^{\mu_i - 1} \left(\beta \cdot a_{\gamma_i \cdot p^d}\right)^{p^{-d}} = \sum_{d=0}^{\mu_i - 1} \beta^{p^{-d}} \cdot b_{i,d} \ .$$

Let $\beta_1, \ldots, \beta_{\dim V} \in \mathbb{F}_q$ be a basis for V. Then,

$$V \subseteq \mathsf{Const}_u \iff \mathrm{tr}(\beta_\ell F(x)) \text{ is constant in } x \ \forall \ell \in [\dim V] \iff \lambda_{i,\beta_\ell} = 0 \ \forall i \in [s-1], \ell \in [\dim V] \ .$$

Thus, $\dim_{\mathbb{F}_p} \{u \in \mathrm{RS}\left(\mathbb{F}_q; k\right) \mid V \subseteq \mathsf{Const}_u\}$ is the $\mathbb{F}_p$-dimension of the solution space $M \subseteq \mathbb{F}_q^k$ of the system of equations

$$\sum_{d=0}^{\mu_i - 1} \beta_\ell^{p^{-d}} \cdot b_{i,d} = 0 \qquad \forall 1 \leq i \leq s-1, \quad 1 \leq \ell \leq \dim V \tag{38}$$

in the variables

$$\{b_{i,d} \mid 1 \leq i \leq s, \quad 0 \leq d \leq \mu_i - 1\} \ .$$

Note that there are indeed $\sum_{i=1}^{s} \mu_i = k$ variables. In this reduction, we used the fact that the change of variables in Eq. (37) is $\mathbb{F}_p$-linear and invertible, so it does not affect the dimension over $\mathbb{F}_p$.

Note that M is not only $\mathbb{F}_p$-linear but also $\mathbb{F}_q$-linear. The equations corresponding to different values of i are disjoint, so in fact, for each i we have a separate system given by the matrix $B^{(i)} \in \mathbb{F}_q^{\dim V \times \mu_i}$ where $B_{\ell,d}^{(i)} = \beta_\ell^{p^{-d}}$. Below, we argue that

$$\mathrm{rank}_{\mathbb{F}_q}\left(B^{(i)}\right) = \min\left\{\dim V, \mu_i\right\} \ . \tag{39}$$

We first show that this implies Eq. (33). Indeed,

$$\dim_{\mathbb{F}_q} M = k - \sum_{i=1}^{s-1} \mathrm{rank}\left(B^{(i)}\right) = k - \sum_{i=1}^{s-1} \min\left\{\dim V, \mu_i\right\} \ .$$

Observe that, under the constraints $\sum_{i=1}^{s-1} \mu_i = k-1$ and $\mu_i \leq \log_p(k-1)$ for all $1 \leq i \leq s-1$, the right-hand side above is maximized when $\frac{(k-1)}{\log_p(k-1)}$ of the μ_i's are $\log_p(k-1)$, and the rest are 0. Hence,

$$\dim_{\mathbb{F}_q} M \leq \max\left\{k - \frac{k-1}{\log_p(k-1)} \cdot \dim V, 1\right\} \ .$$

Eq. (33) follows since $\dim_{\mathbb{F}_p} M = t \cdot \dim_{\mathbb{F}_q} M$.

We turn to prove Eq. (39). Note that it suffices to prove the case where $B^{(i)}$ is a square matrix, i.e., when $\mu_i = \dim V$. Suppose that $B^{(i)} \cdot \lambda = 0$ for some $\lambda \in \mathbb{F}_q^{\mu_i}$. Then, $\beta_1, \ldots, \beta_{\dim V}$ are roots of the polynomial $G(x) = \sum_{d=0}^{\mu_i - 1} \lambda_d x^{p^{-d}}$. Consequently, they are also roots of $G'(x) := G(x) \cdot x^{p^{\mu_i - 1}} = \sum_{d=0}^{\mu_i - 1} \lambda_{\mu_i - 1 - d} x^{p^d}$. Since $G'(x)$ is $\mathbb{F}_p$-linear, all $\mathbb{F}_p$-linear combinations of $\beta_1, \ldots, \beta_{\dim V}$ are also roots, so $G'(x)$ has at least $p^{|V|}$ roots, yet its degree is at most $p^{\mu_i - 1} = p^{\dim V - 1}$. Thus, $\lambda_0 = \ldots = \lambda_{\mu_i - 1} = 0$.[11] We conclude that $B^{(i)}$ has full rank, establishing the theorem modulo Claim 7.1.

Proof of Claim 7.1. The claim is immediate for the cases $i = s$ and $O_i \cap \{0, 1, \ldots, k-1\} = \emptyset$. Suppose that $1 \le i \le s - 1$ so that $0 \notin O_i$, and that $O_i \cap \{1, \ldots k-1\} \neq \emptyset$. Let γ_i be the minimum of $O_i \cap \{1, \ldots, k-1\}$. Note that

$$O_i = \{\gamma_i, \gamma_i \cdot p, \gamma_i \cdot p^2, \ldots, \gamma_i \cdot p^{t-1}\} \ . \tag{40}$$

For an integer $1 \le a \le q - 1$, let $\mathsf{repr}(a) = (\mathsf{repr}(a)_{t-1}, \ldots, \mathsf{repr}(a)_0)$ denote its base p representation. Denote $S_a = \{0 \le i \le t-1 \mid \mathsf{repr}(a)_i \neq 0\}$. The proof now follows from the following facts:

1. Since $k \le \sqrt{q} = p^{\frac{t}{2}}$, if $1 \le a < k$ then S_a must be contained in the interval $I := \{0, \ldots, \lfloor \frac{t-1}{2} \rfloor\}$.
2. $\mathsf{repr}(p \cdot a)$ is a cyclic-shift of $\mathsf{repr}(a)$, namely, $\mathsf{repr}(p \cdot a) = (\mathsf{repr}(a)_{t-2}, \ldots, \mathsf{repr}(a)_0, \mathsf{repr}(a)_{t-1})$.

In particular, by the minimality of γ_i, we have $S_{\gamma_i} \subseteq I$ and $0 \in S_{\gamma_i}$. Let $w = \max S_{\gamma_i} \le \lfloor \frac{t-1}{2} \rfloor$. It is not hard to see that $\{\gamma_i, \gamma_i \cdot p, \ldots, \gamma_i \cdot p^{t-1-w}\}$ is an increasing sequence, so the part of that sequence contained in $\{1, \ldots, k-1\}$ is a prefix. We denote this prefix by $\{\gamma_i, \ldots \gamma_i \cdot p^{\mu_i - 1}\}$.

It remains to show that $\gamma_i \cdot p^i \ge k$ for every $t - w \le i \le t - 1$. This is indeed the case since $a := \gamma_i \cdot p^i$ is represented by a cyclic i-shift of $\mathsf{repr}(\gamma_i)$. In particular, because $0 \in S_{\gamma_i}$, we have $i \in S_a$. Consequently, $i \ge t - w > \lfloor \frac{t-1}{2} \rfloor$ implies $a \ge k$. □

8 Local row-symmetric properties of punctured codes

8.1 Properties of codes

In this section we recall some of the framework for studying *local and row-symmetric properties of linear codes* [12], established in [MRRSW20; GMRSW21].[13]

A property $\mathcal{P}$ of length-n linear codes over $\mathbb{F}_q$ is a collection of linear codes in $\mathbb{F}_q^n$. A linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ such that $\mathcal{C} \in \mathcal{P}$ is said to satisfy $\mathcal{P}$. If $\mathcal{P}$ is upwards closed with regard to containment, it is said to be monotone-increasing. A monotone-increasing property $\mathcal{P}$ of linear codes has a unique minimal-set $\mathcal{M}_{\mathcal{P}}$, namely, a matrix $A \subseteq \mathbb{F}_q^{n \times b}$ with distinct columns belongs to $\mathcal{M}_{\mathcal{P}}$ if the code consisting of the column-span of A satisfies $\mathcal{P}$, but no proper linear subspace of that code does so.

[11] Equivalently, unless $G'(x) = 0$, the dimension of the space of roots of the linearized polynomial $G'(x)$ is at most its p-degree, which is $\mu_i - 1$.

[12] This framework makes sense for linear as well as non-linear codes. In this work we restrict ourselves to the linear case.

[13] The notion of a *local property* from [MRRSW20] was later refined and split into two parts in [GMRSW21], where it appears as a *row-symmetric* and *local* property. We follow the latter convention.

Example 8.1 (List-decodability as a property of codes)**.** *Fix a prime power q, $n, L \in \mathbb{N}$ and $\rho \in [0, 1]$. Consider the monotone increasing property $\mathcal{P}$ consisting of all linear codes in $\mathbb{F}_q^n$ that are **not** (ρ, L) -list-decodable. Then,*

$$\mathcal{M}_{\mathcal{P}} = \left\{ A \in \mathbb{F}_q^{n \times b} \mid \log_q(L+1) \le b \le L+1, \textit{ and } A \textit{ is } (\rho, L+1)\textit{-span-clustered and } \text{minimal} \right\} .$$

By minimal, *we mean that there is no $(\rho, L+1)$ -span-clustered matrix whose column-span is strictly contained in that of A.*

Say that a matrix $A \in \mathbb{F}_q^{n \times b}$ is (ρ, ℓ, L) -recovery-span-clustered if the column span of A contains a (ρ, ℓ)-clustered set of size L. Example 8.1 can now be readily generalized to list-recoverability.

Example 8.2 (List-recoverablity as a property of codes)**.** *Fix a prime power q, $n, L \in \mathbb{N}$, $1 \le \ell \le q$ and $\rho \in [0, 1]$. Consider the monotone increasing property $\mathcal{P}$ consisting of all linear codes in $\mathbb{F}_q^n$ that are **not** (ρ, ℓ, L) -list-recoverable. Then, $\mathcal{M}_{\mathcal{P}}$ consists of all the matrices $A \in \mathbb{F}_q^{n \times b}$ ($\log_q(L+1) \le b \le L+1$) that are $(\rho, \ell, L+1)$ -recovery-span-clustered and* minimal.

$$\mathcal{M}_{\mathcal{P}} = \left\{ A \in \mathbb{F}_q^{n \times b} \mid \log_q(L+1) \le b \le L+1, \ \&\ A \textit{ is } (\rho, \ell, L+1)\textit{-recovery-span-clustered and minimal} \right\} .$$

Definition 8.3 (Local, row-symmetric and scalar-invariant properties)**.** *Let $\mathcal{P}$ be a monotone-increasing property of codes in $\mathbb{F}_q^n$. We define the following notions. The first two are from [GMRSW21] and the third is specific to this work.*

1. *If every matrix in $\mathcal{M}_{\mathcal{P}}$ has at most b columns ($b \in \mathbb{N}$), we say that $\mathcal{P}$ is b-*local*.*
2. *If, for each $A \in \mathcal{M}_{\mathcal{P}}$, it holds that every matrix obtained by permuting the rows of A also belongs to $\mathcal{M}_{\mathcal{P}}$, we say that $\mathcal{P}$ is* row-symmetric*.*
3. *If, for each $A \in \mathcal{M}_{\mathcal{P}}$ and every full-rank diagonal matrix $\Lambda \in \mathbb{F}_q^{n \times n}$ it holds that $\Lambda A \in \mathcal{M}_{\mathcal{P}}$, then $\mathcal{P}$ is called* scalar-invariant*.*

The following is immediate.

Observation 8.4. *Let q be a prime power and $n \in \mathbb{N}$. Fix $\rho \in (0, 1)$, $L \in \mathbb{N}$, and let $\mathcal{P}$ be the monotone increasing property consisting of codes in $\mathbb{F}_q^n$ that are **not** (ρ, L) -list-decodable. Then, $\mathcal{P}$ is $(L+1)$-local, row-symmetric and scalar-invariant. Moreover, given $1 \le \ell \le q$, the same holds for the property consisting of codes that are **not** (ρ, ℓ, L) -list-recoverable.*

Remark 8.5 (Average versions of list-decodability and list-recoverability)**.** *Average-radius list-decodability is a stronger property where we demand that for every $L+1$ codewords their average distance to any center exceeds ρ (as opposed to maximum distance for list-decodability). A code not being (ρ, L)-average-radius list-decodable is also an $(L+1)$-local, row-symmetric and scalar-invariant property.*

For list-recovery, we can define a stronger variant where in Definition 2.6 we allow input sets Z_i such that the average size $|Z_i|$ over all $i \in [n]$ is at most ℓ. A violation of this stronger property is also a local, row-symmetric and scalar-invariant property.

The generality of our framework thus means that we can get results for these variants also automatically. We note that certain results for list-decodability for RLCs, e.g., [GHK11; LW21], do not extend to average-radius list-decoding (or list-recovery).

Let $\mathcal{P}$ be a monotone-increasing property over $\mathbb{F}_q^n$. Suppose that $\mathcal{P}$ is nonempty, namely, that it is satisfied by $\mathbb{F}_q^n$. We denote its threshold by

$$\mathrm{RLC}(\mathcal{P}) = \min\left\{R \in [0,1] \mid \Pr\left[C_{\mathrm{RLC}}^{n,q}(R) \text{ satisfies } \mathcal{P}\right] \geq \frac{1}{2}\right\} .$$

We have the following generalization of Theorem 2.5.

Theorem 8.6 (Thresholds for local and row-symmetric properties [MRRSW20, Thm. 2.8])**.** [14] *Let $\mathcal{P} \subseteq \mathbb{F}_q^n$ be a random linear code of radius R and Let $\mathcal{P}$ be a monotone-increasing, b-local and row-symmetric property over $\mathbb{F}_q^n$, where $\frac{n}{\log_q n} \geq \omega\left(q^{2b}\right)$. The following now holds for every $\varepsilon > 0$.*

1. *If $R \leq \mathrm{RLC}(\mathcal{P}) - \varepsilon$ then*
$$\Pr\left[\mathcal{C} \textit{ satisfies } \mathcal{P}\right] \leq q^{-(\varepsilon - o(1))n} .$$

2. *If $R \geq \mathrm{RLC}(\mathcal{P}) + \varepsilon$ then*
$$\Pr\left[\mathcal{C} \textit{ satisfies } \mathcal{P}\right] \geq 1 - q^{-(\varepsilon - o(1))n} .$$

8.2 Theorems 2 and 4 generalized to local row-symmetric properties

We now generalize the statements of Theorems 2 and 4 to the language of properties of codes.

Theorem 8 (Puncturings of certain linear codes are locally similar to random linear codes)**.** *Let q be a prime power, and let $\mathcal{P}$ be a monotone-increasing, row-symmetric and b-local property over $\mathbb{F}_q^n$, where $\frac{n}{\log n} \geq \omega_q\left(q^{2b}\right)$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be a linear code. Let $\mathcal{C}$ be a random n-puncturing of $\mathcal{D}$ of design rate $R \leq \mathrm{RLC}(\mathcal{P}) - \varepsilon$ for some $\varepsilon > 0$. Suppose that at least one of the following two conditions holds:*

1. *$\mathcal{P}$ is scalar-invariant, $\mathcal{D}$ has q^{-b}-optimal distance, and $q \geq 2^{\frac{2}{\varepsilon}}$.*
2. *$\mathcal{D}$ is $\left(\frac{\varepsilon b \ln q}{q^b}\right)$-biased.*

Then,
$$\Pr\left[\mathcal{C} \textit{ satisfies } \mathcal{P}\right] \leq q^{-(\varepsilon - o(1))n} .$$

By virtue of Observation 8.4, Theorem 8 yields Theorem 2 (resp., Theorem 4) by taking $\mathcal{P}$ to be the family of codes in $\mathbb{F}_q^n$ that are **not** (ρ, L)-list-decodable (resp., **not** (ρ, ℓ, L)-list-recoverable). Hence, to prove Theorems 2 and 4 it suffices to prove Theorem 8. We do so in Section 8.4.

8.3 Row-symmetric b-local properties in terms of distributions over $\mathbb{F}_q^b$

Thresholds for row-symmetric and local properties can be characterized in terms of empirical distributions of certain matrices. We recall this connection.

[14] The theorem as stated in [MRRSW20] deals only with the regime of constant q and b. The current statement, which allows q and b to depend on n, follows by inspecting the proof in [MRRSW20].

Fact 8.7 ([GMRSW21, Fact 2.15]). *Let $\mathcal{P}$ be a monotone-increasing, b-local, row-symmetric property over $\mathbb{F}_q^n$. Then, there exists a set $\mathcal{T}_\mathcal{P}$ of distributions over $\mathbb{F}_q^b$ such that $|\mathcal{T}_\mathcal{P}| \le (n+1)^{q^b}$, and $\mathcal{M}_\mathcal{P} = \bigcup_{\tau \in \mathcal{T}_\mathcal{P}} \mathcal{M}_{n,\tau}$.*

Definition 8.8 (Implied distribution [MRRSW20, Def. 2.6]). *Let τ be a distribution over $\mathbb{F}_q^b$ and let $D \in \mathbb{F}_q^{b \times a}$ such that $\operatorname{rank} D = a$ for some $a \le b$. The distribution (over $\mathbb{F}_q^a$) of the random vector xD, where $x \sim \tau$ (note that x is a row vector), is said to be τ-implied. We denote the set of τ-implied distributions by $\mathcal{I}_\tau$.*

The motivation for Definition 8.8 is the following observation, which follows immediately from the linearity of the code.

Observation 8.9. *Let τ be a distribution over $\mathbb{F}_q^b$, and let $\tau' \in \mathcal{I}_\tau$. Then, any linear code containing a matrix in $\mathcal{M}_{n,\tau}$ must also contain some matrix in $\mathcal{M}_{n,\tau'}$.*

We now have the following characterization of the threshold.

Theorem 8.10 ([MRRSW20, Thm. 2.8]). [15] *Let $\mathcal{P}$ be a monotone-increasing, b-local, row-symmetric property over $\mathbb{F}_q^n$, and let $\mathcal{T}_\mathcal{P}$ be as in Fact 8.7. Then,*

$$\mathrm{RLC}(\mathcal{P}) = \min_{\tau \in \mathcal{T}_\mathcal{P}} \max_{\tau' \in \mathcal{I}_\tau} \left(1 - \frac{H_q(\tau')}{\dim(\tau')}\right) \pm \frac{2q^{2b}\log_q n}{n} \quad .$$

To illustrate the notions that appear in Theorem 8.10, we use this theorem to prove the lower bound on the RLC list-recovery threshold rate stated in Eq. (3). We need the following claim.

Claim 8.11. *Let $B \in \mathbb{F}_q^{n \times b}$ be a matrix whose columns form a (ρ, ℓ)-recovery-clustered set. Denote $\tau = \mathsf{Emp}_B$. Then, $H_q(\tau) \le b \cdot h_q(\ell, \rho) + \ell$, where $h_{q,\ell}(\rho) = \rho \log_q\left(\frac{q-\ell}{\rho}\right) + (1-\rho)\log_q\left(\frac{\ell}{1-\rho}\right)$.*

Proof. Let $Z_1, \ldots, Z_n$ be subsets of $\mathbb{F}_q$, each of size ℓ, such that for all $j \in [b]$, we have $|\{i \in [n] \mid B_{i,j} \notin Z_i\}| \le \rho n$. Let i be sampled uniformly from $[n]$. We now have

$$H_q(\tau) = H_q(B_i) \le H_q(B_i, Z_i) = H_q(Z_i) + H_q(B_i \mid Z_i) \le H_q(Z_i) + \sum_{j=1}^{b} H_q(B_{i,j} \mid Z_i).$$

The number of different options for Z_i is $\binom{q}{\ell}$ so $H_q(Z_i) \le \log_q(\binom{q}{\ell}) \le \ell$. Let ρ_j' ($j \in [b]$) denote the probability that $B_{i,j} \notin Z_i$, and note that $\rho_j' \le \rho$. Then,

$$H_q(B_{i,j} \mid Z_i) \le h_{q,\ell}(\rho_j') \le h_{q,\ell}(\rho).$$

Consequently, $H_q(\tau) \le b \cdot h_{q,\ell}(\rho) + \ell$, establishing the claim. □

Proof of Eq. (3). Let $\mathcal{P}$ denote the property consisting of codes over $\mathbb{F}_q^n$ that are **not** (ρ, ℓ, L)-list-recoverable. Let $\tau \in \mathcal{T}_\mathcal{P}$ and let $A \in \mathcal{M}_{n,\tau}$ be a matrix in $\mathbb{F}_q^{n \times a}$ ($a \in \mathbb{N}$). By Example 8.2, A is $(\rho, \ell, L+1)$-recovery-span-clustered. Let W be a (ρ, ℓ)-recovery-clustered set of

[15] The precise error term does not appear in the statement of this theorem in [MRRSW20], but follows by inspecting the proof there.

size $L+1$, contained in the column-span of A. Note that W must contain a linearly-independent subset U of size $b := \lceil \log_q |W| \rceil = \lceil \log_q(L+1) \rceil$. Let $D \in \mathbb{F}_q^{a\times b}$ such that $B := AD$ is the matrix whose columns are the elements of U, and note that U is also (ρ,ℓ)-recovery-clustered. Let $\tau' = \mathsf{Emp}_B$. By Claim 8.11, $H_q(\tau') \le b \cdot h_{q,\ell}(\rho) + \ell$. Furthermore, we can express τ' as the distribution of the random vector xD, where $x \sim \tau$. Consequently, $\tau' \in \mathcal{I}_\tau$. Therefore,

$$\max_{\tau'' \in \mathcal{I}_\tau} \left(1 - \frac{H_q(\tau'')}{\dim(\tau')}\right) \ge 1 - \frac{H_q(\tau')}{b} \ge 1 - h_{q,\ell}(\rho) - \frac{\ell}{b}.$$

The claim now follows by Theorem 8.10. □

We note that the above derivation of Eq. (3) could also be achieved via more standard arguments, which do not require Theorem 8.10. The actual power Theorem 8.10 is that it enables reductions from other random code models to the RLC model, as demonstrated in Lemma 8.12 below, and later, in the proof of Theorem 8. This sort of argument involves an application of Theorem 8.10 in *its less intuitive direction*: rather than starting from an upper bound on $H_q(\tau)$ for some set of distributions and using Theorem 8.10 to obtain a lower bound on $\mathrm{RLC}(\mathcal{P})$, we start from some known lower bound on $\mathrm{RLC}(\mathcal{P})$ and use the theorem to get an upper bound on the entropy of certain "bad distributions." The latter entropy bound is then typically used in a union-bound argument to obtain a lower bound on the threshold rate for some non-RLC model. This type of argument was used in [MRRSW20] to prove that *Gallagher LDPC codes* are as list-decodable (and list-recoverable) as RLCs.

Lemma 8.12 (A generic reduction to random linear codes)**.** *Let $n \in \mathbb{N}$, q a prime power and $b \in \mathbb{N}$ such that $\frac{n}{\log_q n} \ge \omega\left(q^{2b}\right)$. Let $\mathcal{C} \in \mathbb{F}_q^n$ be a linear code of rate $R \in [0,1]$, sampled at random from some ensemble. Suppose that, for every $1 \le a \le b$, every distribution τ over $\mathbb{F}_q^a$ and every matrix $B \in \mathbb{F}_q^{Rn\times a}$ with $\operatorname{rank} B = a$, we have*

$$\mathbb{E}_{\mathcal{C}}\left[\left|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \mathcal{C}\}\right|\right] \le q^{(H_q(\tau) - a(1-R) + a\varepsilon)n} \quad , \tag{41}$$

for some fixed $\varepsilon > 0$. Then, for any row-symmetric and b-local property $\mathcal{P}$ over $\mathbb{F}_q^n$ such that $R \le \mathrm{RLC}(\mathcal{P}) - 2\varepsilon$, it holds that

$$\Pr_{\mathcal{C}}\left[\mathcal{C} \textit{ satisfies } \mathcal{P}\right] \le q^{-n(\varepsilon - o(1))} \ .$$

Proof. Let $\tau \in \mathcal{T}_{\mathcal{P}}$. By Theorem 8.10, there is some distribution $\tau' \in \mathcal{T}_{\mathcal{P}}$ over $\mathbb{F}_q^a$ (where $1 \le a \le b$) such that

$$\frac{H_q(\tau')}{a} \le 1 - \mathrm{RLC}(\mathcal{P}) + o(1) \ .$$

Now, by Observation 8.9, followed by Markov's bound,

$$\begin{aligned}\Pr\left[\exists A \in \mathcal{M}_{n,\tau} \ \ A \subseteq \mathcal{C}\right] &\le \Pr\left[\exists A \in \mathcal{M}_{n,\tau'} \ \ A \subseteq \mathcal{C}\right] \le \mathbb{E}\left[\left|\left\{A \in \mathcal{M}_{n,\tau'} \mid A \subseteq \mathcal{C}\right\}\right|\right] \\ &\le \exp_q\left(\left(H_q(\tau') - a(1-R) + a\varepsilon\right) n\right) \le \exp_q(an\left(R - \mathrm{RLC}(\mathcal{P}) + \varepsilon + o(1)\right)) \\ &\le \exp_q(-na(\varepsilon - o(1))) \ .\end{aligned}$$

Therefore, by Fact 8.7,

$$\begin{aligned}\Pr\left[\mathcal{C} \text{ satisfies } \mathcal{P}\right] &\le \sum_{\tau \in \mathcal{T}_{\mathcal{P}}} \Pr\left[\exists A \in \mathcal{M}_{n,\tau} \ \ A \subseteq \mathcal{C}\right] \le |\mathcal{T}_{\mathcal{P}}|\, q^{-n(\varepsilon - o(1))} \\ &\le (n+1)^{q^b} q^{-n(\varepsilon - o(1))} \le q^{-n(\varepsilon - o(1))} \ .\end{aligned}$$

□

8.4 Proof of Theorems 2, 4 and 8

We now prove Theorem 8, which generalizes Theorems 2 and 4.

Proof of Condition 1 of Theorem 8. Let τ be a distribution over $\mathbb{F}_q^a$, where $a \le b$. By Lemma 5.6,

$$\begin{aligned}\mathbb{E}_{\mathcal{C}}\left[|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \Lambda\mathcal{C}\}|\right] &\le \exp_q\!\big(n \cdot \big(H_q(\tau) - (1-R)a + \log_q\left(1+\eta q^a\right) + \log_q 2\big)\big)\\ &\le \exp_q\!\big(n \cdot \big(H_q(\tau) - (1-R)a + 2\log_q 2\big)\big)\\ &\le \exp_q\!\left(n \cdot \left(H_q(\tau) - (1-R)a + \varepsilon\right)\right)\\ &\le \exp_q\!\left(n \cdot \left(H_q(\tau) - (1-R)a + a\varepsilon\right)\right)\ ,\end{aligned}$$

where $\Lambda \sim \mathsf{U}(\Gamma_n)$. By Lemma 8.12, $\Lambda\mathcal{C}$ satisfies $\mathcal{P}$ with probability at most $q^{-n(\varepsilon - o(1))}$. Since $\mathcal{P}$ is scalar-invariant, the same holds for $\mathcal{C}$. □

Condition 2 of Theorem 8 requires the following lemma, which is analogous to Lemma 5.6.

Lemma 8.13 (Puncturings of low-bias codes are locally similar to random linear codes)**.** *Fix $b \in \mathbb{N}$ and a full-rank distribution τ over $\mathbb{F}_q^b$. Let $\mathcal{D} \subseteq \mathbb{F}_q^m$ be an η-biased linear code ($\eta \ge 0$). Let φ be a random ($m \to n$) puncturing map. Denote $R = \frac{\log_q|\mathcal{D}|}{n}$. Then,*

$$\mathbb{E}_{\mathcal{C}}\left[|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \mathcal{C}\}|\right] \le \exp_q\!\left(n \cdot \left(H_q(\tau) - (1-R)b + \log_q\left(1+\eta q^b\right)\right)\right)$$

Proof. Let τ be a full-rank distribution over $\mathbb{F}_q^b$. Item 1 of Lemma 5.9 yields B

$$\mathbb{E}_{x\sim \mathsf{Emp}_B}\left[\tau(x)\right] \le q^{-b}\left(1+q^b\eta\right)\ ,$$

for all $B \in \mathbb{F}_q^{m\times b}$ such that $\operatorname{rank} B = b$ and $B \subseteq \mathcal{D}$. By Lemma 5.10,

$$\Pr\left[\varphi(B) \in \mathcal{M}_{n,\tau}\right] \le \exp_q\!\left(n\left(-b + H_q(\tau) + \log_q\left(1+q^b\eta\right)\right)\right)\ .$$

The claim now follows by the union bound over the $\le q^{Rnb}$ choices of . □

Proof of Condition 2 of Theorem 8. Let τ be a distribution over $\mathbb{F}_q^a$ with $a \le b$. Lemma 8.13 yields

$$\begin{aligned}\mathbb{E}_{\mathcal{C}}\left[|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \mathcal{C}\}|\right] &\le \exp_q\!\big(n \cdot \big(H_q(\tau) - (1-R)a + \log_q\left(1+\eta q^a\right)\big)\big)\\ &\le \exp_q\!\left(n \cdot \left(H_q(\tau) - (1-R)a + \frac{\eta q^a}{\ln q}\right)\right)\\ &= \exp_q\!\left(n \cdot \left(H_q(\tau) - (1-R)a + \frac{\varepsilon b}{q^{b-a}}\right)\right)\\ &\le \exp_q\!\left(n \cdot \left(H_q(\tau) - (1-R)a + a\varepsilon\right)\right)\ ,\end{aligned}$$

which implies the claim by virtue of Lemma 8.12. □

Remark 8.14 (On the conditions in Theorem 8, and comparison to Theorem 1)**.** *In the above proof of Theorem 8, as in the proof of Theorem 1, the core of the proof is obtaining an upper bound on terms of the form* $\mathbb{E}\left[\left|\{A \in \mathcal{M}_{n,\tau} \mid A \subseteq \Lambda\mathcal{C}\}\right|\right]$ *for certain distributions* τ*, where* $\mathcal{C}$ *is a random puncturing of a mother-code* $\mathcal{D}$*.*

When our assumption about $\mathcal{D}$ *is that of near-optimal distance, we bound this expectation via Lemma 5.6, which includes a bothersome* $\log_q 2$ *term. This term needs to be bounded from above by* $a\varepsilon$*. One way to overcome this term is to take* q *large enough to make* $\log_q 2$ *negligibly small, as we do in Condition 1 of Theorem 8. In Theorem 1 where we use [GHK11], the problem is handled differently: Corollary 5.5 provides us with a "slack" that dominates the* $\log_q 2$ *term whenever* a *is small, whereas for large* a *the* $a\varepsilon$ *upper bound is not too restrictive. Finally, in Condition 2 we avoid the bothersome term altogether via Lemma 8.13, due to our assumption that* $\mathcal{D}$ *has small bias.*

9 List-recovery for Reed-Solomon codes

The proof of Theorem 5 closely follows that of Theorem 3. We allow ourselves to skip parts of the proof that remain essentially identical.

9.1 Parameter settings

Write $R = 1 - \rho - \varepsilon$ and $k = Rn$, and take Q be a power of p such that $\max\{\ell, 2\}^{\frac{4}{\varepsilon}} \le Q \le p \cdot \max\{\ell, 2\}^{\frac{4}{\varepsilon}}$. Take $L = \frac{\ell}{1-\rho} Q^{\frac{2\ell+3}{\varepsilon}}$ and $L' = Q^{\frac{2\ell+3}{\varepsilon}}$.

As in Theorem 3, we suppose that n is large enough so that $\frac{\varepsilon n}{\log_p n} \ge \omega\left(Q^L\right)$, and take q to be the minimum power of p such that $\mathbb{F}_Q$ is a subfield of F_q and

$$q \ge \max\left\{n^2, \left(2kQ^{L+1}\right)^2, k^{\frac{1}{1-R-2\varepsilon}}, 3 \cdot \binom{L+1}{2}^{\frac{1}{\varepsilon}}\right\} .$$

In particular, we have

$$q \le O_{p,\varepsilon}\left(n^{\max\left\{2, \frac{1}{\rho-\varepsilon}\right\}}\right) .$$

Let φ be a random $(q \to n)$ puncturing map. Write $\mathcal{C} = \varphi(\mathrm{RS}\left(\mathbb{F}_q; k\right))$ and $\mathcal{D} = \mathrm{tr}_{q \to Q}(\mathrm{RS}\left(\mathbb{F}_q; k\right))$. We say that the code $\varphi(\mathcal{D})$ is (ρ, ℓ, L')-quasi-list-recoverable if there does not exist any trace-friendly matrix $A \in \mathbb{F}_Q^{q \times b}$ such that $A \subseteq \mathcal{D}$ and $\varphi(A)$ is $(\rho, \ell, L' + 1)$-recovery-span-clustered.

9.2 Proof of Theorem 5

The following lemmas, proven below, are respective analogs of Lemmas 6.9 and 6.10

Lemma 9.1. *In the setting of Section 9.1, consider the following events:*

1. $\mathcal{C}$ *is* (ρ, ℓ, L)*-list-recoverable*

2. $\varphi(\mathcal{D})$ *is not* (ρ, ℓ, L')*-quasi-list-recoverable.*

The probability that ***at least one of*** *the above events occurs is* $1 - q^{-\Omega(\varepsilon n)}$.

Lemma 9.2. *In the setting of Section 9.1, with probability at least* $1 - Q^{-\Omega(n)}$*, the code* $\varphi(\mathcal{D})$ *is* (ρ, ℓ, L') *-quasi-list-recoverable.*

Lemmas 9.1 and 9.2 yield the theorem in exactly the same manner that their analogs imply Theorem 3. It thus remains to prove Lemmas 9.1 and 9.2.

Proof of Lemma 9.1. The proof of this lemma is virtually identical to that of Lemma 6.9, except for one difference. In the proof of Claim 6.14, we show that a ρ-clustered set G_i such that $u - v$ is a constant vector for all $u, v \in G_i$, must satisfy $|G_i| \le \frac{1}{1-\rho}$. Here, rather than assuming that G_i is ρ-clustered, we assume that it is (ρ, ℓ)-recovery-clustered. Accordingly, the relevant bound is now $|G_i| \le \frac{\ell}{1-\rho}$, and it is obtained in a similar fashion.

To account for this difference, our parameter setting now has $L' = \frac{1-\rho}{\ell} \cdot L$, rather than $L' = (1-\rho)L$ as in the proof of Theorem 3. □

Proof of Lemma 9.2. Let $r = \log_Q(L'+1)$. Say that a matrix $G \in \mathbb{F}_Q^{n\times b}$ is (ρ, ℓ)-recovery-clustered if the columns of G are distinct and form a (ρ, ℓ)-recovery-clustered set.

If $\varphi(\mathcal{D})$ is not (ρ, ℓ, L')-quasi-list-recoverable, then there exists a trace-friendly matrix $A \in \mathbb{F}_Q^{q\times b}$ (for $b \in \mathbb{N}$) such that $A \subseteq \mathcal{D}$ and $\varphi(A)$ is $(\ell, \rho, L'+1)$-recovery-span-clustered. Let W be a (ρ, ℓ)-recovery-clustered set of size $L'+1$, contained in the column-span of $\varphi(A)$. This set W must contain a linearly-independent subset U of size $\log_Q |W| = r$. Since U is linearly independent, there exists some $P \in \mathbb{F}_Q^{b\times r}$ of rank r such that $G := \varphi(A)P$ is a matrix whose columns are the elements of U. Note that G is (ρ, ℓ)-recovery-clustered. Moreover, $G = \varphi(AP)$, where $F := AP$ is trace-friendly and contained in $\mathcal{D}$.

We conclude that a necessary condition for $\varphi(\mathcal{D})$ to not be (ρ, ℓ, L')-quasi-list-recoverable is the existence of some trace-friendly matrix $F \in \mathbb{F}_q^{q\times r}$ such that $F \subseteq \mathcal{D}$, and $\varphi(F)$ is (ρ, ℓ)-recovery-clustered. By Observations 8.4 and 4.4 and Markov's bound, it suffices to prove that

$$\mathbb{E}\left[\left|\left\{F \in \mathbb{F}_Q^{q\times r} \mid F \text{ is trace friendly, } F \subseteq \mathcal{D}, \text{ and } \varphi^*(F) \text{ is } (\rho, \ell)\text{-recovery-clustered}\right\}\right|\right] \le Q^{-\Omega(n)} \enspace ,$$

where φ^* is a random $((Q-1)q \to n)$ puncturing-map.

Similarly to the proof of Lemma 6.10, Lemma 9.2 would now follow if we show that

$$\mathbb{E}\left[\left|\left\{F \in \mathbb{F}_Q^{q\times r} \mid F \text{ is trace-friendly, } F \subseteq \mathcal{D} \text{ and } \varphi^*(F^*) \in \mathcal{M}_{n,\tau}\right\}\right|\right] \le Q^{-\Omega(n)} \tag{42}$$

for every distribution τ over $\mathbb{F}_q^r$ of the form $\tau = \mathsf{Emp}_G$, where $G \in \mathbb{F}_q^{n\times r}$ is (ρ, ℓ)-recovery-clustered. As in the proof of Lemma 6.9, we can show that the left-hand side of Eq. (42) is at most

$$\exp_Q\left(n\left(rR + \log_Q\left(2Q^{-b}\cdot\left(1 + Q^b\eta\right)\right) + H_Q(\tau)\right)\right) \enspace ,$$

where $\eta = \frac{k-1}{\sqrt{q}} \le Q^{-b}$. By Claim 8.11, $H_Q(\tau) \le \ell + r \cdot h_{Q,\ell}(\rho)$. Note that

$$h_{Q,\ell}(\rho) \le \rho + \log_Q \ell + \frac{1}{\log_2 Q} \le \rho + \frac{\varepsilon}{2} \enspace .$$

Thus, we can further bound the left-hand side of Eq. (42) by

$$\begin{aligned}
&\exp_Q\Big(n\left(rR+\log_Q\left(4Q^{-r}\right)+\ell+r\cdot\rho+\frac{r}{2}\varepsilon\right)\Big)\\
&=\exp_Q\Big(n\left(\log_Q 4+\ell-\frac{r\varepsilon}{2}\right)\Big)\\
&\le\exp_Q\Big(n\left(\ell-\frac{(r-1)\varepsilon}{2}\right)\Big)\\
&=\exp_Q\Big(n\left(\ell-\frac{(\log_Q(L'+1)-1)\varepsilon}{2}\right)\Big)\\
&\le Q^{-n}\ .
\end{aligned}$$

□

10 Derandomization of RLCs

Here we prove the following theorem, which generalizes Theorem 6 by virtue of Observation 8.4.

Theorem 9 (Codes locally similar to an RLC with linear randomness)**.** *There exists a randomized algorithm that, given $b \in \mathbb{N}$, $\varepsilon > 0$, $R^* \in [\varepsilon, 1]$ and $n \in \mathbb{N}$, where $\frac{n}{\log_2 n} \ge \omega\left(2^{2b}\right)$ and $n \ge \omega(1/\varepsilon)$, samples a generating matrix for a code $\mathcal{C}$ of rate $R = R^* - \varepsilon$ such that*

$$\Pr\left[\mathcal{C} \textit{ satisfies some property } \mathcal{P} \in \mathcal{K}\right] \le 2^{-\Omega(\varepsilon n)}. \tag{43}$$

Here, $\mathcal{K}$ is the family of all monotone-increasing, b-local and row-symmetric properties $\mathcal{P}$ over $\mathbb{F}_2^n$ for which the threshold $\mathrm{RLC}(\mathcal{P})$ is at least R^*. This algorithm uses $O\left(n\left(b+\log_2\frac{1}{\varepsilon}\right)\right)$ random bits, and works in time polynomial in n.

Proof. Fix a property $\mathcal{P} \in \mathcal{K}$. Fix $\eta = \frac{\varepsilon b \ln 2}{2^b}$. Let $\mathcal{D} \subseteq \mathbb{F}_2^m$ be an η-biased linear code of dimension Rn, where $m \le O(n \cdot \eta^{-c})$ for some universal $c \ge 2$. Explicit constructions of such a code $\mathcal{D}$ are given in [ABNNR92; Ta-17]. We also assume that

$$m \ge \frac{n}{1-2^{-\frac{\varepsilon}{2}}}\ , \tag{44}$$

noting that $\frac{m}{n}$ can be taken to be as large as desired.

Sample a random increasing sequence of n integers $1 \le i_1 < i_2 < \cdots < i_n \le m$ uniformly from among all such sequences. Note that such a sequence can be encoded by

$$\log_2\binom{m}{n} + O(1) \le n\left(\log_2\frac{m}{n} + O(1)\right) \le O\left(n\left(b+\log_2\frac{1}{\varepsilon}\right)\right)$$

random bits, whose decoding can be done in poly(m) time.

Let $\mathcal{C}$ be the code defined by the random sequence $i_1, \dots, i_n$ via $\mathcal{C} = \{(u_{i_1} \dots u_{i_n}) \mid u \in \mathcal{D}\} \subseteq \mathbb{F}_2^n$. Clearly, a generating matrix for $\mathcal{C}$ can be obtained from that of $\mathcal{D}$ in poly(m) = poly(n) time. Hence, to prove the theorem it suffices to show that $\mathcal{C}$ satisfies Eq. (43). Let $\mathcal{C}' \subseteq \mathbb{F}_2^n$ be a random n-puncturing of $\mathcal{D}$. Let T be the event that $\mathcal{C}'$ satisfies $\mathcal{P}$. Let J denote the event that no coordinate of $\mathcal{D}$ is sampled more than once for inclusion in $\mathcal{C}'$. Note that $\Pr[J] \ge \left(1-\frac{n}{m}\right)^n$. By Theorem 8, $\Pr[T] \le 2^{-(\varepsilon - o(1))n}$. Thus,

$$\Pr[T \mid J] \le \frac{\Pr[T]}{\Pr[J]} \le \exp_2\Big(\Big(-\varepsilon - \log_2\left(1-\frac{n}{m}\right) + o(1)\Big)n\Big) \le 2^{-\Omega(\varepsilon n)}\ ,$$

where the last inequality follows from Eq. (44).

By row-symmetry, $\mathcal{P}$ is invariant to coordinate permutations of $\mathcal{C}$. Observe that a uniformly random coordinate permutation of $\mathcal{C}$ yields a code distributed identically to the distribution of $\mathcal{C}'$ conditioned on the event J. Therefore,

$$\Pr\left[\mathcal{C} \text{ satisfies } \mathcal{P}\right] = \Pr\left[T \mid J\right] \le 2^{-\Omega(\varepsilon n)} \tag{45}$$

for every $\mathcal{P} \in \mathcal{K}$.

It remains to show that Eq. (45) implies Eq. (43). Let $\mathcal{K}' = (\mathcal{P} \in \mathcal{K} \mid |\mathcal{T}_{\mathcal{P}}| = 1)$ (recall Fact 8.7 for the definition of $\mathcal{T}_{\mathcal{P}}$). Observe that a necessary condition for the event in Eq. (43) is that $\mathcal{C}$ satisfies some property in $\mathcal{K}'$. Indeed, suppose that $\mathcal{C}$ satisfies a property $\mathcal{P} \in \mathcal{K}$ and let $\tau \in \mathcal{T}_{\mathcal{P}}$ such that $\mathcal{C}$ contains a matrix in $\mathcal{M}_{n,\tau}$. Let $\mathcal{P}'$ denote the b-local, row-symmetric and monotone-increasing property for which $\mathcal{T}_{\mathcal{P}'} = \{\tau\}$. Clearly, $\mathcal{C}$ satisfies $\mathcal{P}'$. Since $\mathcal{P}'$ implies $\mathcal{P}$, we have $\mathrm{RLC}(\mathcal{P}') \ge \mathrm{RLC}(\mathcal{P}) \ge R^*$ and so $\mathcal{P}' \in \mathcal{K}'$. Thus, to prove the theorem, it suffices to show that

$$\Pr\left[\mathcal{C} \text{ satisfies some property } \mathcal{P}' \in \mathcal{K}'\right] \le 2^{-\Omega(\varepsilon n)} \ . \tag{46}$$

Now, by Fact 4.11, $|\mathcal{K}'| \le (n+1)^{2^b} \le 2^{o(n)}$. Thus, Eq. (46) follows from Eq. (45) by a union bound on $\mathcal{K}'$, noting that $\mathcal{K}' \subseteq \mathcal{K}$. □

Acknowledgments

The authors thank João Ribeiro for pointing out an improvement to Theorem 6.

References

[ABNNR92] Noga Alon, Jehoshua Bruck, Joseph Naor, Moni Naor, and Ron M. Roth. "Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs". In: *IEEE Trans. Inf. Theory* 38.2 (1992), pp. 509–516. DOI: 10.1109/18.119713. URL: https://doi.org/10.1109/18.119713.

[BKR10] Eli Ben-Sasson, Swastik Kopparty, and Jaikumar Radhakrishnan. "Subspace polynomials and limits to list decoding of Reed-Solomon codes". In: *IEEE Trans. Inf. Theory* 56.1 (2010), pp. 113–120. DOI: 10.1109/TIT.2009.2034780. URL: https://doi.org/10.1109/TIT.2009.2034780.

[CGV13] Mahdi Cheraghchi, Venkatesan Guruswami, and Ameya Velingker. "Restricted Isometry of Fourier Matrices and List Decodability of Random Linear Codes". In: *SIAM J. Comput.* 42.5 (2013), pp. 1888–1914. DOI: 10.1137/120896773. URL: https://doi.org/10.1137/120896773.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of information theory (2. ed.)* Wiley, 2006. ISBN: 978-0-471-24195-9. URL: http://www.elementsofinformationtheory.com/.

[CU57] L. Carlitz and S. Uchiyama. "Bounds for exponential sums". In: *Duke Math. J.* 24 (1957), pp. 37–41. ISSN: 0012-7094. URL: http://projecteuclid.org/euclid.dmj/1077467207.

[Del75] P. Delsarte. “On subfield subcodes of modified Reed–Solomon codes”. In: *IEEE Trans. Inform. Theory* 21.5 (1975), pp. 575–576.

[FKS20] Asaf Ferber, Matthew Kwan, and Lisa Sauermann. “List-decodability with large radius for Reed-Solomon codes”. In: *CoRR* abs/2012.10584 (2020). To appear in FOCS 2021. arXiv: 2012.10584. URL: https://arxiv.org/abs/2012.10584.

[GGR11] Parikshit Gopalan, Venkatesan Guruswami, and Prasad Raghavendra. “List Decoding Tensor Products and Interleaved Codes”. In: *SIAM J. Comput.* 40.5 (2011), pp. 1432–1462. DOI: 10.1137/090778274. URL: https://doi.org/10.1137/090778274.

[GHK11] Venkatesan Guruswami, Johan Håstad, and Swastik Kopparty. “On the List-Decodability of Random Linear Codes”. In: *IEEE Trans. Inf. Theory* 57.2 (2011), pp. 718–725. DOI: 10.1109/TIT.2010.2095170. URL: https://doi.org/10.1109/TIT.2010.2095170.

[GLMRSW20] Venkatesan Guruswami, Ray Li, Jonathan Mosheiff, Nicolas Resch, Shashwat Silas, and Mary Wootters. “Bounds for List-Decoding and List-Recovery of Random Linear Codes”. In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2020, August 17-19, 2020, Virtual Conference*. Ed. by Jaroslaw Byrka and Raghu Meka. Vol. 176. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020, 9:1–9:21. DOI: 10.4230/LIPIcs.APPROX/RANDOM.2020.9. URL: https://doi.org/10.4230/LIPIcs.APPROX/RANDOM.2020.9.

[GLSTW20] Zeyu Guo, Ray Li, Chong Shangguan, Itzhak Tamo, and Mary Wootters. “Improved List-Decodability of Reed-Solomon Codes via Tree Packings”. In: *CoRR* abs/2011.04453 (2020). To appear in FOCS 2021. arXiv: 2011.04453. URL: https://arxiv.org/abs/2011.04453.

[GMRSW21] Venkatesan Guruswami, Jonathan Mosheiff, Nicolas Resch, Shashwat Silas, and Mary Wootters. “Sharp Threshold Rates for Random Codes”. In: *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference*. Ed. by James R. Lee. Vol. 185. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, 5:1–5:20. DOI: 10.4230/LIPIcs.ITCS.2021.5. URL: https://doi.org/10.4230/LIPIcs.ITCS.2021.5.

[GN14] Venkatesan Guruswami and Srivatsan Narayanan. “Combinatorial Limitations of Average-Radius List-Decoding”. In: *IEEE Trans. Inf. Theory* 60.10 (2014), pp. 5827–5842. DOI: 10.1109/TIT.2014.2343224. URL: https://doi.org/10.1109/TIT.2014.2343224.

[Gol11] Oded Goldreich. “Three XOR-Lemmas - An Exposition”. In: *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*. Ed. by Oded Goldreich. Vol. 6650. Lecture Notes in Computer Science. Springer, 2011, pp. 248–272. DOI: 10.1007/978-3-642-22670-0_22. URL: https://doi.org/10.1007/978-3-642-22670-0%5C_22.

[GR06] Venkatesan Guruswami and Atri Rudra. “Limits to List Decoding Reed-Solomon Codes”. In: *IEEE Trans. Inf. Theory* 52.8 (2006), pp. 3642–3649. DOI: 10.1109/TIT.2006.878164. URL: https://doi.org/10.1109/TIT.2006.878164.

[GR08] Venkatesan Guruswami and Atri Rudra. “Explicit Codes Achieving List Decoding Capacity: Error-Correction With Optimal Redundancy”. In: *IEEE Trans. Inf. Theory* 54.1 (2008), pp. 135–150. DOI: 10.1109/TIT.2007.911222. URL: https://doi.org/10.1109/TIT.2007.911222.

[GR10] Venkatesan Guruswami and Atri Rudra. “The existence of concatenated codes list-decodable up to the hamming bound”. In: *IEEE Trans. Inf. Theory* 56.10 (2010), pp. 5195–5206. DOI: 10.1109/TIT.2010.2059572. URL: https://doi.org/10.1109/TIT.2010.2059572.

[GRS] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential coding theory.* Draft available at http://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/.

[GS99] Venkatesan Guruswami and Madhu Sudan. “Improved decoding of Reed-Solomon and algebraic-geometry codes”. In: *IEEE Trans. Inf. Theory* 45.6 (1999), pp. 1757–1767. DOI: 10.1109/18.782097. URL: https://doi.org/10.1109/18.782097.

[GST21] Eitan Goldberg, Chong Shangguan, and Itzhak Tamo. “List-decoding and list-recovery of Reed-Solomon codes beyond the Johnson radius for any rate”. In: *CoRR* abs/2105.14754 (2021). arXiv: 2105.14754. URL: https://arxiv.org/abs/2105.14754.

[Gur06] Venkatesan Guruswami. “Algorithmic Results in List Decoding”. In: *Found. Trends Theor. Comput. Sci.* 2.2 (2006). DOI: 10.1561/0400000007. URL: https://doi.org/10.1561/0400000007.

[GW13] Venkatesan Guruswami and Carol Wang. “Linear-Algebraic List Decoding for Variants of Reed-Solomon Codes”. In: *IEEE Trans. Inf. Theory* 59.6 (2013), pp. 3257–3268. DOI: 10.1109/TIT.2013.2246813. URL: https://doi.org/10.1109/TIT.2013.2246813.

[JH01] Jørn Justesen and Tom Høholdt. “Bounds on list decoding of MDS codes”. In: *IEEE Trans. Inf. Theory* 47.4 (2001), pp. 1604–1609. DOI: 10.1109/18.923744. URL: https://doi.org/10.1109/18.923744.

[Joh62] Selmer M. Johnson. “A new upper bound for error-correcting codes”. In: *IRE Trans. Inf. Theory* 8.3 (1962), pp. 203–207. DOI: 10.1109/TIT.1962.1057714. URL: https://doi.org/10.1109/TIT.1962.1057714.

[KRSW18] Swastik Kopparty, Noga Ron-Zewi, Shubhangi Saraf, and Mary Wootters. “Improved Decoding of Folded Reed-Solomon and Multiplicity Codes”. In: *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018.* Ed. by Mikkel Thorup. IEEE Computer Society, 2018, pp. 212–223. DOI: 10.1109/FOCS.2018.00029. URL: https://doi.org/10.1109/FOCS.2018.00029.

[LN96] Rudolf Lidl and Harald Niederreiter. *Finite Fields.* 2nd ed. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996. DOI: 10.1017/CBO9780511525926.

[LP20] Ben Lund and Aditya Potukuchi. "On the List Recoverability of Randomly Punctured Codes". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2020, August 17-19, 2020, Virtual Conference*. Ed. by Jaroslaw Byrka and Raghu Meka. Vol. 176. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020, 30:1–30:11. DOI: 10.4230/LIPIcs.APPROX/RANDOM.2020.30. URL: https://doi.org/10.4230/LIPIcs.APPROX/RANDOM.2020.30.

[LW21] Ray Li and Mary Wootters. "Improved List-Decodability of Random Linear Binary Codes". In: *IEEE Trans. Inf. Theory* 67.3 (2021), pp. 1522–1536. DOI: 10.1109/TIT.2020.3041650. URL: https://doi.org/10.1109/TIT.2020.3041650.

[MRRSW20] Jonathan Mosheiff, Nicolas Resch, Noga Ron-Zewi, Shashwat Silas, and Mary Wootters. "LDPC Codes Achieve List Decoding Capacity". In: *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*. IEEE, 2020, pp. 458–469. DOI: 10.1109/FOCS46700.2020.00050. URL: https://doi.org/10.1109/FOCS46700.2020.00050.

[Res20] Nicolas Resch. "List-Decodable Codes: (Randomized) Constructions and Applications". PhD thesis. 2020. URL: http://reports-archive.adm.cs.cmu.edu/anon/2020/CMU-CS-20-113.pdf.

[RW14a] Atri Rudra and Mary Wootters. "Every list-decodable code for high noise has abundant near-optimal rate puncturings". In: *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*. Ed. by David B. Shmoys. ACM, 2014, pp. 764–773. DOI: 10.1145/2591796.2591797. URL: https://doi.org/10.1145/2591796.2591797.

[RW14b] Atri Rudra and Mary Wootters. "It'll probably work out: improved list-decoding through random operations". In: *Electronic Colloquium on Computational Complexity (ECCC)* 21 (2014), p. 104. URL: http://eccc.hpi-web.de/report/2014/104.

[RW18] Atri Rudra and Mary Wootters. "Average-radius list-recovery of random linear codes". In: *Proceedings of the 2018 ACM-SIAM Symposium on Discrete Algorithms, SODA*. 2018.

[ST20] Chong Shangguan and Itzhak Tamo. "Combinatorial list-decoding of Reed-Solomon codes beyond the Johnson radius". In: *Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*. Ed. by Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy. ACM, 2020, pp. 538–551. DOI: 10.1145/3357713.3384295. URL: https://doi.org/10.1145/3357713.3384295.

[Ta-17] Amnon Ta-Shma. "Explicit, almost optimal, epsilon-balanced codes". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*. Ed. by Hamed Hatami, Pierre McKenzie, and Valerie King. ACM, 2017, pp. 238–251. DOI: 10.1145/3055399.3055408. URL: https://doi.org/10.1145/3055399.3055408.

[Woo13] Mary Wootters. “On the list decodability of random linear codes with large error rates”. In: *Symposium on Theory of Computing Conference, STOC’13, Palo Alto, CA, USA, June 1-4, 2013*. 2013, pp. 853–860. DOI: 10.1145/2488608.2488716. URL: http://doi.acm.org/10.1145/2488608.2488716.

[ZP81] Victor Vasilievich Zyablov and Mark Semenovich Pinsker. “List Concatenated Decoding”. In: *Problemy Peredachi Informatsii* 17.4 (1981), pp. 29–33.