Randomly Punctured Reed–Solomon Codes Achieve the List Decoding Capacity over Polynomial-Size Alphabets

Zeyu Guo
CSE Department
The Ohio State University
zguotcs@gmail.com

Zihan Zhang
CSE Department
The Ohio State University
zhang.13691@osu.edu

Abstract

This paper shows that, with high probability, randomly punctured Reed–Solomon codes over fields of polynomial size achieve the list decoding capacity. More specifically, we prove that for any $\varepsilon > 0$ and $R \in (0, 1)$, with high probability, randomly punctured Reed–Solomon codes of block length n and rate R are $(1 - R - \varepsilon, O(1/\varepsilon))$ list decodable over alphabets of size at least $2^{\mathrm{poly}(1/\varepsilon)} n^2$. This extends the recent breakthrough of Brakensiek, Gopi, and Makam (STOC 2023) that randomly punctured Reed–Solomon codes over fields of exponential size attain the generalized Singleton bound of Shangguan and Tamo (STOC 2020).

1 Introduction

Reed–Solomon (RS) codes [RS60] are a classical family of error-correcting codes that have found numerous applications both in theory and in practice. They are obtained by evaluating low-degree univariate polynomials over a finite field $\mathbb{F}_q$ at a set of evaluation points. Formally, given distinct elements $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}_q$, the $[n, k]$ RS code over $\mathbb{F}_q$ with evaluation points $\alpha_1, \ldots, \alpha_n$ is defined to be the linear code

$$\mathsf{RS}_{n,k}(\alpha_1, \ldots, \alpha_n) := \left\{ \big(f(\alpha_1), \ldots, f(\alpha_n)\big) \,:\, f(X) \in \mathbb{F}_q[X],\ \deg(f) < k \right\} \subseteq \mathbb{F}_q^n.$$

It has rate $R = k/n$ and relative minimum distance $\delta = (n - k + 1)/n$, attaining the Singleton bound [Sin64]. Thus, an RS code of rate R has the *unique decoding radius* $(1 - R)/2$, which is optimal by the Singleton bound. In this paper, we consider the more challenging problem of determining the *list decoding radius* of RS codes.

List decoding. The notion of list decoding was introduced independently by Elias [Eli57] and Wozencraft [Woz58] in the 1950s as a natural generalization of unique decoding, where the decoder is allowed to output $L \geq 1$ codewords and can potentially correct more than $\delta/2$ fraction of errors, δ being the relative minimum distance of the code. Since its introduction, list decoding has found many applications in theoretical computer science [Sud00, Vad12, GL89, CPS99, GRS00] and information theory [Eli91, Ahl73, Bli86, Bli97].

Formally, a code $C \subseteq \Sigma^n$ over an alphabet Σ is said to be (*combinatorially*) (ρ, L) *list decodable* if for every $y \in \Sigma^n$, the Hamming ball centered at y with relative radius $\rho \in [0, 1]$ contains at most L codewords in C. By the *list decoding capacity theorem* [GRS19, Theorem 7.4.1], for $q \geq 2$, $0 \leq \rho < 1 - \frac{1}{q}$, $\varepsilon > 0$, and sufficiently large n, there exist (ρ, L) list decodable codes of block length n, rate R, alphabet size q, and list size $L = O(1/\varepsilon)$ such that

$$R \geq 1 - H_q(\rho) - O(\varepsilon) \tag{1.1}$$

where $H_q(\cdot)$ denotes the q-ary entropy function. Codes satisfying (1.1) are said to achieve the *list decoding capacity*. When $q \geq 2^{\Omega(1/\varepsilon)}$, Condition (1.1) can be rewritten as $\rho \geq 1 - R - O(\varepsilon)$.

In a seminal paper [GR08], Guruswami and Rudra constructed the first explicit list decodable codes achieving the list decoding capacity, known as *folded Reed–Solomon codes*. Other explicit capacity-achieving codes have been discovered since then, which are based on the same or similar techniques [Gur09, GW13, Kop15, GRZ21, GX22]. These codes are not only combinatorially list decodable, but also *efficiently* list decodable, meaning that they admit efficient list decoding algorithms. On the other hand, the known bounds on the list size in these constructions are substantially worse than the bound $O(1/\varepsilon)$ in the list decoding capacity theorem, being at least exponential in $1/\varepsilon$.[1]

List decodability of RS codes. While *folded* RS codes have been shown to achieve the list decoding capacity in [GR08], understanding the list decodability of RS codes remains an important problem and has attracted a lot of attention. See, e.g., [Gur04, Rud07, Vad12]. By the classical *Johnson bound* [Joh62, GS01], an RS code of rate R is list decodable up to the radius around $1 - \sqrt{R}$. Guruswami and Sudan [GS06], built on the earlier work of Sudan [Sud97], gave an efficient algorithm that list decodes RS codes up to the Johnson bound.

Going beyond the Johnson bound is much more challenging. Ben-Sasson, Kopparty, and Radhakrishnan [BSKR09] proved that over certain (non-prime) finite fields $\mathbb{F}_q$, full-length RS codes are not list decodable substantially beyond the Johnson bound. However, this result does not rule out the possibility that *randomly punctured RS codes* (i.e., RS codes with a random subset of evaluation points in $\mathbb{F}_q$) are, with high probability, list decodable beyond the Johnson bound. Proving that these codes can indeed outperform the Johnson bound is, however, highly nontrivial due to the strong algebraic structure of RS codes. In particular, the lack of independence between the codewords prevents one from simply applying the probabilistic method.

As a side remark, some other natural families of random codes with structures have been shown to be list decodable well beyond the Johnson bound and in fact achieve the list decoding capacity, including random linear codes [GHK11, CGV13, Woo13, RW18, GLM$^+$21] and random LDPC codes [MRRZ$^+$20]. In [GM22], it was shown that random puncturings of low-bias linear codes over $\mathbb{F}_q$ of rate R are, with high probability, (ρ, L) list decodable if $R < 1 - H_q(\rho) - \frac{C_{\rho,q}}{L} - \varepsilon$, where $C_{\rho,q}$ depends only on ρ and q.

Rudra and Wootters [RW14] were the first to show that randomly punctured RS codes are list decodable beyond the Johnson bound for a certain range of parameters. Specifically, they proved that for small enough ε and large enough q, randomly punctured RS codes of rate $\frac{\varepsilon}{\log^5(1/\varepsilon)\log q}$ over $\mathbb{F}_q$ are, with high probability, (ρ, L) list decodable with $\rho = 1 - O(\varepsilon)$ and $L = O(1/\varepsilon)$. In [ST20], Shangguan and Tamo proved that if C is a linear code of rate R that is (ρ, L) list decodable, then

$$\rho \leq \frac{L}{L+1}(1 - R). \tag{1.2}$$

Bound (1.2) is called the *generalized Singleton bound* in [ST20]. Indeed, it generalizes the bound $\rho \leq (1-R)/2$ for unique decoding (i.e., the case where $L = 1$) that follows from the Singleton bound.

For $L = 2, 3$, Shangguan and Tamo proved in the same paper [ST20] that their generalized Singleton bound is (with high probability) attained by randomly punctured RS codes over alphabets of exponential size. They further conjectured that this also holds for arbitrary L. In a follow-up paper [GLS$^+$22], Guo, Li, Shangguan, Tamo, and Wootters proved that randomly punctured RS codes of rate R over alphabets of exponential size are $(1-\varepsilon, O(1/\varepsilon))$ list decodable for some $R = \Omega(\frac{\varepsilon}{\log(1/\varepsilon)})$. Their result was greatly improved by Ferber, Kwan, and Sauermann [FKS22], who used a short and clever proof to show that, over a large enough alphabet of (at least) polynomial size, a code of rate R obtained by randomly puncturing *any* code is, with high probability, $(1-\varepsilon, O(1/\varepsilon))$ list decodable for some $R = \Omega(\varepsilon)$. Using the proof of [FKS22],

[1] For folded RS codes [GR08], the best known upper bound for the list size is $(1/\varepsilon)^{O(1/\varepsilon)}$, proved in [KRZSW18].

Goldberg, Shangguan, and Tamo [GST22] showed that a randomly punctured RS code of rate R and block length n over a large enough field of size $n^{O_{R,1/\varepsilon}(1)}$ is, with high probability, $(1-\frac{2}{R+1}R-\varepsilon, O(1/\varepsilon))$ list decodable. This follows from a more general result that they proved about the list decodability of randomly punctured linear codes (see [GST22, Theorem 5]).

In a recent breakthrough [BGM22a], Brakensiek, Gopi, and Makam resolved the conjecture of Shangguan and Tamo in the affirmative by showing that generic RS codes achieve the generalized Singleton bound. This means that over a large enough alphabet, randomly punctured RS codes of rate R are, with high probability, $(\frac{L}{L+1}(1-R), L)$ list decodable.[2] Brakensiek et al. proved their result by establishing connections among three notions of linear codes that strengthen the classical notion of *maximum distance separable (MDS) codes*. These are *MDS(ℓ) codes* studied in [BGM22b, BDG22], *GZP(ℓ) codes*, which are linear MDS codes whose generating matrices attain generic zero patterns [DSY14, BGM22a], and *LD-MDS(ℓ) codes* (introduced and called ℓ-MDS codes in [Rot22]), which are linear codes that attain the bound (1.2) in a strong sense. Surprisingly, Brakensiek et al. showed that these notions are all equivalent. More precisely, they proved that for $\ell \geq 2$, a linear code C is MDS(ℓ) iff it is GZP(ℓ), which holds iff the dual code of C is LD-MDS($\leq \ell-1$) (i.e., LD-MDS(ℓ') for all $\ell' \leq \ell-1$). It is known that a generic RS code is GZP(ℓ), which was proved independently by Lovett [Lov18] and Yildiz and Hassibi [YH19] in their resolutions of the *GM-MDS conjecture* [DSY14]. Combining this fact with the above equivalence and the duality of (generalized) RS codes, Brakensiek et al. proved that generic RS codes are LD-MDS(ℓ) for all ℓ, which implies the conjecture of Shangguan and Tamo.

The alphabet size. While Brakensiek et al. [BGM22a] showed that randomly punctured RS codes over large enough alphabets attain the generalized Singleton bound $\rho \leq \frac{L}{L+1}(1-R)$, the alphabet size they need is quite large, which is at least exponential in nL when the rate $R=k/n$ is a bounded away from zero and one. Moreover, Brakensiek, Dhar, and Gopi [BDG22] recently proved an exponential lower bound on the alphabet size for such R and $L=2$ (see [BDG22, Corollary 1.7 and Theorem 1.8]). Also see [BGM22b] for an earlier lower bound.

However, as noted in [BDG22], the exponential lower bound applies only if we want to *exactly* achieve the generalized Singleton bound. In particular, it does not rule out the possibility that a randomly punctured RS code is, with high probability, $(1-R-\varepsilon, O(1/\varepsilon))$ list decodable over an alphabet of polynomial size, which is consistent with the list decoding capacity theorem and known lower bounds.

We remark that both in theory and in practice, codes over smaller alphabets tend to have more applications. Whether or not there exist RS codes over polynomial-size alphabets that still achieve the list decoding capacity is thus a very important question. In this paper, we answer this question in the affirmative. See Table 1 below for a summary of known results on the list decodability of randomly punctured RS codes over $\mathbb{F}_q$.

1.1 Our Results

We now state our main results. Recall that a linear code $C \subseteq \mathbb{F}_q^n$ is (ρ, L) *list decodable* (resp. (ρ, L) *average-radius list decodable*) if there do not exist $y \in \mathbb{F}_q^n$ and distinct codewords $x_1, \ldots, x_{L+1} \in C$ such that the maximum relative distance (resp. average relative distance) between x_i and y over $i \in [L+1]$ is bounded by ρ. And a *randomly punctured* $[n,k]$ RS code over $\mathbb{F}_q$ is just $\mathsf{RS}_{n,k}(\alpha_1, \ldots, \alpha_n)$ where $(\alpha_1, \ldots, \alpha_n)$ is uniformly distributed over the set of vectors in $\mathbb{F}_q^n$ with *distinct* coordinates.

Our main theorem states that with high probability, randomly punctured RS codes over alphabets of polynomial size are (average-radius) list decodable up to a radius that almost attains the generalized Singleton bound.

[2] In fact, Brakensiek, Gopi, and Makam [BGM22a] proved the stronger statement that randomly punctured RS codes of rate R are, with high probability, $(\frac{L}{L+1}(1-R), L)$ *average-radius* list decodable. See Definition 2.1.

Table 1: Adapted from [GST22] and [BGM22a]. Known results on the combinatorial list decodability of randomly punctured RS codes over $\mathbb{F}_q$. We use c and $c_{R,\varepsilon}$ to denote an absolute constant and a constant depending on R and ε, respectively.

	Radius ρ	List size L	Rate R	Field size q
Johnson bound [GS01]	$1-\varepsilon$	$O(1/\varepsilon^2)$	$\Omega(\varepsilon^2)$	$q \geq n$
[RW14]	$1-\varepsilon$	$O(1/\varepsilon)$	$\Omega\left(\frac{\varepsilon}{\log^5(1/\varepsilon)\log q}\right)$	$q \geq cn\log^c(n/\varepsilon)/\varepsilon$
[ST20]	$\frac{L(1-R)}{L+1}$	$2,3$	R	$q \geq 2^{cn}$
[GLS+22]	$1-\varepsilon$	$O(1/\varepsilon)$	$\Omega\left(\frac{\varepsilon}{\log(1/\varepsilon)}\right)$	$q \geq \left(\frac{1}{\varepsilon}\right)^{cn}$
[FKS22]	$1-\varepsilon$	$O(1/\varepsilon)$	$\Omega(\varepsilon)$	$q \geq \mathrm{poly}(n)$
[GST22]	$1-\frac{2R}{R+1}-\varepsilon$	$O(1/\varepsilon)$	R	$q \geq n^{c_{R,\varepsilon}}$
[BGM22a]	$1-R-\varepsilon$	$\frac{1-R-\varepsilon}{\varepsilon}$	R	$q \geq \exp(\widetilde{\Theta}(n/\varepsilon))$
Our work (Thm. 1.1)	$\frac{L(1-R-\varepsilon)}{L+1}$	L	R	$q \geq 2^{\mathrm{poly}(L)/\varepsilon}n^2$
Our work (Cor. 1.2)	$1-R-\varepsilon$	$\frac{2(1-R-\varepsilon)}{\varepsilon}$	R	$q \geq 2^{\mathrm{poly}(1/\varepsilon)}n^2$

Theorem 1.1. *For $\varepsilon > 0$, positive integers n, k, L with $k \leq n$, and a prime power $q \geq 2^{\mathrm{poly}(L)/\varepsilon}nk$, a randomly punctured $[n,k]$ RS code of rate $R = k/n$ over $\mathbb{F}_q$ is, with high probability, $\left(\frac{L}{L+1}(1-R-\varepsilon), L\right)$ average-radius list decodable (and hence also $\left(\frac{L}{L+1}(1-R-\varepsilon), L\right)$ list decodable).*

As a corollary, we prove that randomly punctured RS codes over alphabets of polynomial size achieve the list decoding capacity with high probability.

Corollary 1.2. *For $\varepsilon > 0$, positive integer n, k with $k \leq n$, and a prime power $q \geq 2^{\mathrm{poly}(1/\varepsilon)}nk$, a randomly punctured $[n,k]$ RS code of rate $R = k/n$ over $\mathbb{F}_q$ is, with high probability, $(1-R-\varepsilon, L)$ average-radius list decodable (and hence also $(1-R-\varepsilon, L)$ list decodable) with $L = \max\left\{\left\lceil\frac{2(1-R)}{\varepsilon}\right\rceil - 1, 1\right\}$.*

See Theorem 4.7 and Corollary 4.8 for the more detailed versions of Theorem 1.1 and Corollary 1.2, respectively.

Recall that Brakensiek et al. [BGM22a] proved that with high probability, randomly punctured RS codes are $\left(1-R-\varepsilon, \frac{1-R-\varepsilon}{\varepsilon}\right)$ list decodable over alphabets of exponential size. Compared with their result, Corollary 1.2 reduces the required alphabet size to $O_\varepsilon(n^2)$.

On the other hand, the list size in Corollary 1.2 is worse than that in [BGM22a] by a constant factor. This constant factor can be brought arbitrarily close to one at the cost of increasing the field size. See Corollary 4.8 for details.

1.2 Proof Overview

The ideas in our proof are quite natural and intuitive. To explain these ideas, we first take a look at why previous results in [ST20, GLS+22, BGM22a] require an exponentially large alphabet. In [ST20], Shangguan and Tamo showed that proving the list decodability of randomly punctured RS codes reduces to proving that certain matrices, which they call *intersection matrices*, have full column rank. In this paper, we use an equivalent variant called *reduced intersection matrices*, but the basic idea is the same. Namely, one can show that randomly punctured RS codes of block length n over a large enough finite field $\mathbb{F}_q$ are, with high probability, (average-radius) list decodable if a collection of reduced intersection matrices all have full column rank. See Section 3, and in particular, Lemma 3.5 for details.

It follows from the analysis in [BGM22a] that the reduced intersection matrices, as symbolic matrices in variables $X_1, \ldots, X_n$, do have full column rank. Then by the Schwartz–Zippel lemma [Sch80, Zip79], these matrices still have full column rank with high probability under a random assignment $X_1 = \alpha_1, \ldots, X_n = \alpha_n$, where $\alpha_1, \ldots, \alpha_n$ are the random evaluation points of the RS code. More specifically, the probability that each of the reduced intersection matrices fails to have full column rank under a random assignment $X_1 = \alpha_1, \ldots, X_n = \alpha_n$ is bounded by a function inverse linear in q. However, we also need to apply a union bound over the set of reduced intersection matrices to prove that there exist evaluation points $\alpha_1, \ldots, \alpha_n$ for which all the reduced intersection matrices have full column rank *simultaneously*. As there are exponentially many reduced intersection matrices, applying the union bound requires the alphabet size q to be exponentially large in n.

It is not clear to us whether it is possible to use much fewer (e.g., polynomially many) reduced intersection matrices. Nevertheless, to improve the alphabet size q, one may also try to reduce the probability that each reduced intersection matrix fails to have full column rank under a random assignment. If this probability can be brought down to $\exp(-\Omega(n))$ even if q is only polynomially large, then we would be able to afford the union bound over polynomially large alphabets.

Our key observation is that reducing the failure probability to $\exp(-\Omega(n))$ is indeed possible if we introduce a little "slackness" in the parameters, which corresponds to slightly worsening the list decoding radius of the code. To see this, consider the toy problem of independently picking m random row vectors $v_1, \ldots, v_m \in \mathbb{F}_q^n$ to form an $m \times n$ matrix M, which we want to have full column rank. If we choose $m = n$, which is the optimal choice of m, then the probability that M has full column rank is bounded by a function inverse linear in q, and this happens only if each v_i is not in the span of $v_1, \ldots, v_{i-1}$. However, suppose we choose $m = (1 + \lambda)n$ for some small $\lambda > 0$. In this case, we could afford λn "faulty" vectors v_i, i.e., v_i may be in the span of previous vectors, in which case we just skip it and consider the next vector. The probability that the matrix M has full column rank is then exponentially small in λn even if q is only polynomially large.

Our actual analysis is somewhat more complicated than the one sketched above, but the intuition remains the same. In our analysis, we consider a reduced intersection matrix A of full column rank, which is a symbolic matrix in the variables $X_1, \ldots, X_n$. Then we fix a nonsingular maximal square submatrix M of A, and consider if its nonsingularity changes under a partial random assignment $X_1 = \alpha_1, \ldots, X_i = \alpha_i$, where i goes from zero to n. If M remains nonsingular after assigning all of the n variables, then we have certified that A continues to have full column rank under the assignment. On the other hand, if M becomes singular after assigning some variable X_i, then we call i a *faulty index*. In this case, we update A by deleting all the rows that depend on X_i, pick a new nonsingular maximal square submatrix M of A, and start all over again. By repeating this process up to r times, where r is some parameter linear in n, we either certify that A has full column rank under the randomly chosen assignment, or obtain a sequence of faulty indices $(i_1, \ldots, i_r)$. Moreover, our analysis shows that the latter case occurs with exponentially small probability even after taking the union bound over all the possible sequences $(i_1, \ldots, i_r)$. Further taking the union bound over all the reduced intersection matrices establishes the average-radius list decodability of randomly punctured RS codes.

The above analysis requires the reduced intersection matrices to have the property that they have full column rank even after deleting a small number of rows, namely, those whose associated variables have faulty indices. We prove this property (formally stated as Lemma 3.11) by following and extending the proof in [BGM22a]. We also remark that this full-rank property is used in a black-box manner in our analysis, which makes the analysis quite general.

2 Notations and Preliminaries

Let $\mathbb{N} = \{0, 1, \dots\}$, $\mathbb{N}^+ = \{1, 2, \dots\}$, and $[n] = \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$. Denote by $|S|$ the cardinality of a set S. We write $S = P_1 \sqcup P_2 \sqcup \dots \sqcup P_k$ if the sets $P_1, \dots, P_k$ are nonempty and form a partition of the set S. All logarithms are to the base 2. Denote by $\mathbb{F}_q$ the finite field of size q. For convenience, we use I_J to denote a set system $(I_j : j \in J)$ indexed by a set J.

Write $\mathbb{F}^{n\times m}$ for the vector space of $n \times m$ matrices over a field $\mathbb{F}$. For $M \in \mathbb{F}^{n\times m}$, $S \subseteq [n]$, and $T \subseteq [m]$, denote by $M_{S,T}$ the $|S| \times |T|$ submatrix of M where the rows are selected by S and the columns are selected by T, and the order of rows and that of columns are preserved. We write $M_{i,T}$ instead of $M_{\{i\},T}$ if $S = \{i\}$, and similarly write $M_{S,j}$ instead of $M_{S,\{j\}}$ if $T = \{j\}$.

Unless stated otherwise, all vectors are column vectors. For a (column) vector $v \in \mathbb{F}^n$ and $S \subseteq [n]$, define $v|_S \in \mathbb{F}^{|S|}$ to be the vector obtained by restricting v to the subset S of coordinates, where the order of coordinates is preserved.

For a matrix $M \in \mathbb{F}^{n\times m}$, define $\mathrm{Im}(M) := \{Mx : x \in \mathbb{F}^m\}$ and $\mathrm{Ker}(M) := \{x \in \mathbb{F}^m : Mx = 0\}$, i.e., $\mathrm{Im}(M)$ and $\mathrm{Ker}(M)$ are the image and the kernel of the linear map $\mathbb{F}^n \to \mathbb{F}^m$ represented by M, respectively.

Codes. A *code* over an alphabet Σ is simply a set $C \subseteq \Sigma^n$, where n is called the *block length* of C. The elements of C are called the *codewords* of C. The *rate* of C is $r(C) := \frac{\log|C|}{n \log|\Sigma|} \in [0, 1]$. The *(relative) Hamming distance* $\delta(x, y)$ between $x, y \in \Sigma^n$ is the fraction of coordinates where x and y differ. The *(relative) minimum distance* of C is $\delta(C) := \min_{x,y\in C, x\neq y} \delta(x, y)$.

When Σ is a finite field $\mathbb{F}_q$ and $C \subseteq \mathbb{F}_q^n$ is an $\mathbb{F}_q$-subspace, we say C is a *linear code* over $\mathbb{F}_q$. A linear code of block length n and dimension k is also called an $[n, k]$ *code*, and its rate is simply k/n. For a linear code $C \subseteq \mathbb{F}_q^n$ of dimension k, a *generating matrix* of C is a matrix $G \in \mathbb{F}_q^{n\times k}$ such that $C = \mathrm{Im}(G)$, and a *parity-check matrix* of C is a matrix $H \in \mathbb{F}_q^{(n-k)\times n}$ such that $C = \mathrm{Ker}(H)$. Generating matrices always have full column rank and parity check matrices always have full row rank.

Next, we define list decodable codes and average-radius list decodable codes.

Definition 2.1 (List decodable code)**.** For $\rho \in [0, 1]$ and $L \in \mathbb{N}^+$, a code $C \subseteq \Sigma^n$ is said to be (ρ, L) *list decodable* if for every $y \in \Sigma^n$, the Hamming ball $B_{y,\rho} := \{x \in \Sigma^n : \delta(x, y) \leq \rho\}$ contains at most L codewords of C. And C is said to be (ρ, L) *average-radius list decodable* if there do not exist $y \in \Sigma^n$ and distinct codewords $x_1, \dots, x_{L+1} \in C$ such that $\frac{1}{L+1} \sum_{i=1}^{L+1} \delta(x_i, y) \leq \rho$.

Note that a code is (ρ, L) list decodable if it is (ρ, L) average-radius list decodable. In [BGM22a], a linear code of rate R that is $(\frac{L}{L+1}(1 - R), L)$ average-radius list decodable is also called an *LD-MDS*(L) code.

Reed–Solomon codes. Fix a finite field $\mathbb{F}_q$. Given distinct $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{F}_q$, the $[n, k]$ *Reed–Solomon (RS) code* over $\mathbb{F}_q$ with evaluation points $\alpha_1, \dots, \alpha_n$ is the linear code

$$\mathsf{RS}_{n,k}(\alpha_1, \dots, \alpha_n) := \left\{ \big(f(\alpha_1), \dots, f(\alpha_n)\big) : f(X) \in \mathbb{F}_q[X],\ \deg(f) < k \right\} \subseteq \mathbb{F}_q^n$$

which has dimension k and minimum relative distance $(n - k + 1)/n$.

The *Vandermonde matrix*

$$V_{n,k}(\alpha_1, \dots, \alpha_n) := (\alpha_i^{j-1})_{i\in[n], j\in[k]} \in \mathbb{F}_q^{n\times k}$$

is a generating matrix of $\mathsf{RS}_{n,k}(\alpha_1, \alpha_2, \dots, \alpha_n)$. We also use $V_{n,k}$ to denote the *symbolic Vandermonde matrix*

$$V_{n,k} := (X_i^{j-1})_{i\in[n], j\in[k]} \in \mathbb{F}_q(X_1, \dots, X_n)^{n\times k}.$$

Both $V_{n,k}(\alpha_1, \ldots, \alpha_n)$ and $V_{n,k}$ are *MDS matrices*, i.e., their maximal minors are all nonzero.

Duality of (generalized) RS codes. It is well known that if C is an RS code, then its dual code $C^\perp$ is a *generalized RS code*. This means $C^\perp$ has a generating matrix that is, up to scaling the rows, a Vandermonde matrix. We record this fact in terms of symbolic Vandermonde matrices as follows.

Lemma 2.2. *$V_{n,n-k}^\top D V_{n,k} = 0$, where $D = \operatorname{diag}(v_1, \ldots, v_n)$ and $v_i = \prod_{j\in[n]\setminus\{i\}} \frac{1}{X_i - X_j}$ for $i \in [n]$.*

Proof. For $j \in [n-k]$ and $\ell \in [k]$, the (j, ℓ)-th entry of $V_{n,n-k}^\top D V_{n,k}$ is $\sum_{i\in[n]} v_i X_i^{j+\ell-2}$, and we want to prove that it equals zero. We will show that in fact, for any polynomial $f(X) \in \mathbb{F}_q[X]$ of degree at most $n-2$, we have $\sum_{i\in[n]} v_i f(X_i) = 0$.

By Lagrange's interpolation formula, we have $f(X) = \sum_{i\in[n]} \left(\prod_{j\in[n]\setminus\{i\}} \frac{X-X_j}{X_i-X_j}\right) f(X_i)$. Thus, the coefficient of X^{n-1} in $f(X)$ is $\sum_{i\in[n]} \left(\prod_{j\in[n]\setminus\{i\}} \frac{1}{X_i-X_j}\right) f(X_i) = \sum_{i\in[n]} v_i f(X_i)$, which we know is zero as $\deg(f) \le n-2$. $\square$

3 Reduced Intersection Matrices

We introduce the notion of *reduced intersection matrices*. They are essentially equivalent[3] to the notion of *intersection matrices* introduced in [ST20], but are somewhat more lightweight to use.

3.1 Definitions and Basic Properties

Following [ST20], we first define the weight function $\operatorname{wt}(\cdot)$ for finite set systems on $[n]$.

Definition 3.1 (Weight function)**.** Let $n \in \mathbb{N}^+$. For a collection of subsets $I_j \subseteq [n]$ indexed by a finite set J, define

$$\operatorname{wt}(I_J) := \sum_{j\in J} |I_j| - \Big|\bigcup_{j\in J} I_j\Big|.$$

We now give the formal definition of reduced intersection matrices.

Definition 3.2 (Reduced intersection matrix)**.** Let $n, k, t \in \mathbb{N}^+$ and $I_j \subseteq [n]$ for $j \in [t]$. Let $G \in \mathbb{F}^{n\times k}$ be a matrix over a field $\mathbb{F}$. For $i \in [n]$, let $J_i := \{j \in [t] : i \in I_j\} \subseteq [t]$. In other words, the sets $J_i \subseteq [t]$ are chosen such that for $(i, j) \in [n] \times [t]$, we have $i \in I_j$ iff $j \in I_i$.

Construct a matrix $R_{G,I_{[t]}}$ over $\mathbb{F}$ as follows. Enumerate all $i \in [n]$ such that $|J_i| \ge 2$. For each such i, write $J_i = \{j_1, \ldots, j_{|J_i|}\}$ with $j_1 < \cdots < j_{|J_i|}$, and for $u = 2, \ldots, |J_i|$, add to $R_{G,I_{[t]}}$ a row $r_{i,u} = (r^{(1)}, \ldots, r^{(t-1)})$ of length $(t-1)k$ that is determined as follows:

- $r^{(j_1)} = G_{i,[k]}$ (i.e., the i-th row of G).
- $r^{(j_u)} = -G_{i,[k]}$ if $j_u \ne t$.
- All the other $r^{(j)}$ are zero row vectors of length k.

[3] In [ST20], an intersection matrix is used to represent a system of linear equations satisfied by the pairwise differences $f_{ij} = f_j - f_i$ between a list of codewords f_i. We define a reduced intersection matrix such that it represents an equivalent system of linear equations obtained by eliminating some variables using the cycle relations $f_{ij} + f_{jk} = f_{ik}$ and removing some redundant linear equations.

Order the rows $r_{i,u}$ in the lexicographic order of (i,u). This yields the matrix $R_{G,I_{[t]}}$ with $(t-1)k$ columns. Its number of rows is $\sum_{i\in[n]:|J_i|\geq 2}(|J_i|-1) = \mathrm{wt}(I_{[t]})$. So $R_{G,I_{[t]}} \in \mathbb{F}^{\mathrm{wt}(I_{[t]})\times(t-1)k}$. We call $R_{G,I_{[t]}}$ the *reduced intersection matrix* associated with G and $I_{[t]}$.

We now explain the motivation behind Definition 3.2 above. Suppose $C \subseteq \mathbb{F}_q^n$ is a linear code with a generating matrix $G \in \mathbb{F}_q^{n\times k}$, $y \in \mathbb{F}_q^n$ is a transmitted word, and $c_1,\ldots,c_t \in C$ are t codewords. We may choose (unique) messages $f_1,\ldots,f_t \in \mathbb{F}_q^k$ such that $c_j = Gf_j$ for $j \in [t]$. Let $I_1,\ldots,I_t \subseteq [n]$ such that each I_j is the set of indices of the coordinates where y and c_j agree. For $j \in [t]$, let $f_{j,t} := f_j - f_t$. We are interested in these differences $f_{j,t}$ as knowing them and some f_{j_0} allows us to determine all f_j via $f_j = f_{j_0} - f_{j_0,t} + f_{j,t}$.

Consider $i \in [n]$. The set J_i consists of the set of $j \in [t]$ such that $c_j = Gf_j$ and y agree at the i-th coordinate. In particular, for every pair $j, j' \in J_i$, we have a linear equation $G_{i,[k]}f_{j,t} - G_{i,[k]}f_{j',t} = 0$ that the differences $f_{j,t}$ must satisfy. But note that these equations are not linearly independent. In fact, let j_1 be the smallest element in J_i. Then these linear equations are generated by the subset of equations $G_{i,[k]}f_{j_1,t} - G_{i,[k]}f_{j,t} = 0$, where j ranges over $J_i \setminus \{j_1\}$.

View the coordinates of the vectors $f_{j,t}$ as unknowns. Then consider the system of linear equations $G_{i,[k]}f_{j_1,t} - G_{i,[k]}f_{j,t} = 0$ in these unknowns, where $i \in [n]$, $j \in J_i$, and $j_1 = j_1(u)$ is the smallest element of J_i. The reduced intersection matrix $R_{G,I_{[t]}}$ is defined to exactly represent this system of linear equations, except for one technicality: A we already know $f_{t,t} = 0$, we exclude the coordinates of $f_{t,t}$ from the unknowns. This is reflected in Definition 3.2 above, where we let $r^{(j_u)} = -G_{i,[k]}$ only if $j_u \neq t$.

Finally, we remark that while the above explanation assumes G and $R_{G,I_{[t]}}$ are matrices over a finite field $\mathbb{F}_q$, in our applications, they will be matrices over a function field $\mathbb{F} = \mathbb{F}_q(X_1,\ldots,X_n)$. In particular, we will choose G to be a symbolic Vandermonde matrix $V_{n,k}$, which corresponds to a "generic RS code."

Example 3.3. Let $n = 6$, $k = 3$, $t = 4$, and $G = V_{n,k} = \begin{pmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ \vdots & \vdots & \vdots \\ 1 & X_6 & X_6^2 \end{pmatrix}$. Also let $I_1 = \{1,3,4\}, I_2 = \{1,4,5\}, I_3 = \{2,3,4,5\}, I_4 = \{1,2,4,6\}$. Then $J_1 = \{1,2,4\}, J_2 = \{3,4\}, J_3 = \{1,3\}, J_4 = \{1,2,3,4\}, J_5 = \{2,3\}, J_6 = \{4\}$, and $\mathrm{wt}(I_{[4]}) = 8$. The reduced intersection matrix $R_{G,I_{[t]}} \in \mathbb{F}(X_1,\ldots,X_6)^{8\times 9}$ is given as follows.

$$R_{G,I_{[t]}} = \begin{pmatrix} 1 & X_1 & X_1^2 & -1 & -X_1 & -X_1^2 & & & \\ 1 & X_1 & X_1^2 & & & & & & \\ & & & & & & 1 & X_2 & X_2^2 \\ 1 & X_3 & X_3^2 & & & & -1 & -X_3 & -X_3^2 \\ 1 & X_4 & X_4^2 & -1 & -X_4 & -X_4^2 & & & \\ 1 & X_4 & X_4^2 & & & & -1 & -X_4 & -X_4^2 \\ 1 & X_4 & X_4^2 & & & & & & \\ & & & 1 & X_5 & X_5^2 & -1 & -X_5 & -X_5^2 \end{pmatrix}.$$

When G is a symbolic Vandermonde matrix, we have the following easy observation regarding the degree of the maximal minors of $R_{G,I_{[t]}}$, which will be used later.

Lemma 3.4. *If $G = V_{n,k}$ and M is a $(t-1)k \times (t-1)k$ submatrix of $R_{G,I_{[t]}}$, then $\det(M)$ is a polynomial in $\mathbb{F}[X_1,\ldots,X_n]$ whose degree in each variable X_i is at most $(t-1)(k-1)$.*

Proof. This follows from the definition. The degree bound holds since for each $i \in [n]$, there are at most $t-1$ rows of $R_{G,I_{[t]}}$ that depend on X_i, and the degree of each entry of $R_{G,I_{[t]}}$ in X_i is at most $k-1$. $\square$

The following lemma states that if a linear code with a generating matrix G is not average-radius list decodable, then we can identify a reduced intersection matrix $R_{G,I_{[t]}}$ that does not have full column rank, where $I_{[t]}$ is a set system satisfying certain conditions.

Lemma 3.5. *Let $\rho \in [0,1]$, $\lambda \geq 0$, and $L \in \mathbb{N}^+$. Let C be an $[n,k]$ linear code over a finite field $\mathbb{F}_q$ with a generating matrix $G \in \mathbb{F}_q^{n\times k}$. Suppose C is not (ρ, L) average-radius list decodable and $\rho \leq \frac{L}{L+1}\frac{(n-(1+\lambda)k)}{n}$. Then there exist $t \in \{2,3,\ldots,L+1\}$ and sets $I_1,\ldots,I_t \subseteq [n]$ such that*

(1) $\mathrm{Ker}(R_{G,I_{[t]}}) \neq 0$,

(2) $\mathrm{wt}(I_{[t]}) \geq (1+\lambda)(t-1)k$, *and*

(3) $\mathrm{wt}(I_J) \leq (1+\lambda)(|J|-1)k$ *for all nonempty* $J \subsetneq [t]$.

Proof. As C is not (ρ, L) average-radius list decodable, there exist $y = (y_1,\ldots,y_n) \in \mathbb{F}_q^n$ and distinct $c_1,\ldots,c_{L+1} \in C$ such that $\sum_{j=1}^{L+1}\delta(c_j,y) \leq (L+1)\rho$. For each $j \in [L+1]$, write $c_j = (c_{j,1},\ldots,c_{j,n})$ and let I_j be the set of indices $i \in [n]$ where c_j and y agree, i.e., $c_{j,i} = y_i$. As $n - |I_j| = n\cdot\delta(c_j,y)$ for $j \in [L+1]$, we have

$$\sum_{j=1}^{L+1}|I_j| = n(L+1) - n\cdot\sum_{j=1}^{L+1}\delta(c_j,y) \geq n(L+1)(1-\rho).$$

Therefore,

$$\mathrm{wt}(I_{[L+1]}) = \sum_{j=1}^{L+1}|I_j| - \left|\bigcup_{j=1}^{L+1} I_j\right| \geq n(L+1)(1-\rho) - n \geq (1+\lambda)Lk \tag{3.1}$$

where the last inequality uses the assumption $\rho \leq \frac{L}{L+1}\frac{(n-(1+\lambda)k)}{n}$. Choose a minimal set $S \subseteq [L+1]$ with respect to inclusion such that $|S| \geq 2$ and $\mathrm{wt}(I_S) \geq (1+\lambda)(|S|-1)k$. By (3.1), such a set S exists. Let $t = |S|$. By permuting the codewords c_j and the corresponding sets I_j, we may assume $S = [t]$. So (2) in the lemma holds.

By definition, we have $\mathrm{wt}(I_J) = 0$ for any $J \subseteq [t]$ of size one. Therefore, (3) in the lemma holds by the minimality of S.

Finally, we show that there exists a nonzero vector $v \in \mathbb{F}_q^{(t-1)k}$ such that $R_{G,I_{[t]}} \cdot v = 0$. For $j \in [t]$, let $x_j \in \mathbb{F}_q^k$ be the unique vector satisfying $Gx_j = c_j$, and let $x_{j,t} = x_j - x_t$. In particular, $x_{t,t} = 0$.

Let $v = (x_{1,t},\ldots,x_{t-1,t}) \in \mathbb{F}_q^{(t-1)k}$. Now consider an arbitrary row r of $R_{G,I_{[t]}}$. By the definition of $R_{G,I_{[t]}}$ (Definition 3.2), there exist $\ell,\ell' \in [t]$ with $\ell < \ell'$ and $i \in I_\ell \cap I_{\ell'}$ such that $r = (r^{(1)},\ldots,r^{(t-1)})$, where $r^{(\ell)} = G_{i,[k]}$, $r^{(\ell')} = -G_{i,[k]}$ if $\ell' \neq t$, and all the other $r^{(j)}$ are zero row vectors of length k.

If $\ell' \neq t$, then $r\cdot v = G_{i,[k]}x_{\ell,t} - G_{i,[k]}x_{\ell',t}$. If $\ell' = t$, then we still have $r\cdot v = G_{i,[k]}x_{\ell,t} = G_{i,[k]}x_{\ell,t} - G_{i,[k]}x_{\ell',t}$ since $x_{t,t} = 0$. Therefore,

$$\begin{aligned} r\cdot v = G_{i,[k]}x_{\ell,t} - G_{i,[k]}x_{\ell',t} &= G_{i,[k]}(x_\ell - x_t) - G_{i,[k]}(x_{\ell',t} - x_t) \\ &= G_{i,[k]}x_\ell - G_{i,[k]}x_{\ell'} = c_{\ell,i} - c_{\ell',i} = 0 \end{aligned}$$

where the last equality holds since $c_{\ell,i}, c_{\ell',i} = y_i$ due to the fact $i \in I_\ell \cap I_{\ell'}$. As r is an arbitrary row of $R_{G,I_{[t]}}$, we have $R_{G,I_{[t]}} \cdot v = 0$. So $\mathrm{Ker}(R_{G,I_{[t]}}) \neq 0$. $\square$

3.2 Fundamental Results Established by Brakensiek–Gopi–Makam

We review several results proved by Brakensiek, Gopi, and Makam [BGM22a, BGM22b], which will be used in our analysis.

Recall that for $M \in \mathbb{F}^{k\times n}$, $S \subseteq [k]$, and $T \subseteq [n]$, we denote by $M_{S,T}$ the $|S| \times |T|$ submatrix of M where the rows are selected by S and the columns are selected by T. In particular, $M_{[k],T}$ is the $k \times |T|$ matrix consisting of the columns of M with indices in T, and $\mathrm{Im}(M_{[k],T})$ is its column span.

In the following, we say an $n\times m$ matrix W is *generic* (with the base field $\mathbb{F}$) if $W \in \mathbb{F}(X_{1,1}, \dots, X_{n,m})^{n\times m}$ and the (i, j)-th entry of W is the indeterminate $X_{i,j}$ for $i \in [n]$ and $j \in [m]$. The next theorem, proved in [BGM22a], gives a formula for computing the dimension of the intersection of several column spans $\mathrm{Im}(W_{[k],A_i})$ for a generic $k \times n$ matrix W.

Theorem 3.6 ([BGM22a, Theorem 1.15])**.** *For a generic $k \times n$ matrix W and sets $A_1, \dots, A_\ell \subseteq [n]$, each of size at most k, it holds that*

$$\dim\left(\bigcap_{i=1}^{\ell} \mathrm{Im}(W_{[k],A_i})\right) = \max_{P_1\sqcup P_2\sqcup\cdots\sqcup P_s=[\ell]} \left(\sum_{i=1}^{s}\left|\bigcap_{j\in P_i} A_j\right| - (s-1)k\right), \tag{3.2}$$

where the maximum is taken over all partitions of $[\ell]$.

In [BGM22a], an $[n, k]$ linear code with a generating matrix G is said to be *MDS*(ℓ) if the LHS of (3.2) equals $\dim\left(\bigcap_{i=1}^{\ell} \mathrm{Im}((G^\top)_{[k],A_i})\right)$ for all $A_1, \dots, A_\ell \subseteq [n]$ of size at most k. One of the main results of [BGM22a] is that generic RS codes are MDS(ℓ) for all ℓ. In other words, the following theorem holds.

Theorem 3.7 ([BGM22a, Corollary 1.14])**.** *Let $G = V_{n,k} = (X_i^{j-1})_{i\in[n],j\in[k]}$. Let W be a generic $k \times n$ matrix. Then for all $\ell \in \mathbb{N}^+$ and sets $A_1, \dots, A_\ell \subseteq [n]$, each of size at most k,*

$$\dim\left(\bigcap_{i=1}^{\ell} \mathrm{Im}((G^\top)_{[k],A_i})\right) = \dim\left(\bigcap_{i=1}^{\ell} \mathrm{Im}(W_{[k],A_i})\right).$$

Finally, we need the following lemma in [BGM22b], which relates the dimension of the intersection of several column spans $\mathrm{Im}(H_{[k],A_i})$ to the sum of the dimensions of these column spans and the rank of a certain matrix.

Lemma 3.8 ([BGM22b, Claim B.1]. See also [Tia19])**.** *For $H \in \mathbb{F}^{k\times n}$, $\ell \geq 2$, and $A_1, A_2, \dots, A_\ell \subseteq [n]$,*

$$\dim\left(\bigcap_{i=1}^{\ell} \mathrm{Im}(H_{[k],A_i})\right) = \sum_{i=1}^{\ell} \dim(\mathrm{Im}(H_{[k],A_i})) - \mathrm{rank}\begin{pmatrix} H_{[k],A_1} & H_{[k],A_2} & & & \\ H_{[k],A_1} & & H_{[k],A_3} & & \\ \vdots & & & \ddots & \\ H_{[k],A_1} & & & & H_{[k],A_\ell} \end{pmatrix}.$$

3.3 Full Rankness of Reduced Intersection Matrices

In this subsection, we prove a crucial statement, Lemma 3.11, which states that under certain conditions, a reduced intersection matrix $R_{G,I_{[t]}}$ with $G = V_{n,k}$ has full column rank even if we ignore all the rows associated with a small subset B of variables. Here the size of B is controlled by a parameter $\lambda \geq 0$. The lossless (i.e. $\lambda = 0$) case of this statement was (essentially) proved by Brakensiek, Gopi, and Makam [BGM22a].[4] Our proof of Lemma 3.11 follows and extends their proof.

[4] See [BGM22a, Appendix A] for the proof of the conjecture of Shangguan and Tamo [ST20] about the full rankness of intersection matrices. We also remark that the conjecture concerns list decodability rather than the stronger notion of average-radius list decodability, but the same proof works for the latter.

First, we introduce the following notation, $R^B_{G,I_{[t]}}$, which is a submatrix of a reduced intersection matrix $R_{G,I_{[t]}}$ obtained by deleting some rows.

Definition 3.9 (Deleting rows in $R_{G,I_{[t]}}$)**.** Under the notations in Definition 3.2, for $B \subseteq [n]$, define $R^B_{G,I_{[t]}}$ to be the submatrix of $R_{G,I_{[t]}}$ obtained by deleting all rows $r_{i,u}$ with $i \in B$. Equivalently, $R^B_{G,I_{[t]}} = R_{G,I'_{[t]}}$ where $I'_j = I_j \setminus B$ for $j \in [t]$.

The next lemma states that under some conditions, the kernel of a reduced intersection matrix $R_{G,I_{[t]}}$ can be embedded as a subspace of the kernel of a certain matrix M, which also appears in Lemma 3.8.

Lemma 3.10. *Let $G \in \mathbb{F}^{n\times k}$ and $H \in \mathbb{F}^{(n-k)\times n}$ be matrices over a field $\mathbb{F}$ such that G has full column rank and $HG = 0$. Let $t > 1$ be an integer. Let $I_j \subseteq [n]$ and $A_j = [n] \setminus I_j$ for $j \in [t]$. Finally, let*

$$M = \begin{pmatrix} H_{[n-k],A_1} & H_{[n-k],A_2} & & & \\ H_{[n-k],A_1} & & H_{[n-k],A_3} & & \\ \vdots & & & \ddots & \\ H_{[n-k],A_1} & & & & H_{[n-k],A_t} \end{pmatrix}.$$

Suppose $\bigcup_{j\in[t]} I_j = [n]$. Then there exists a linear map $\psi : \mathbb{F}^{(t-1)k} \to \mathbb{F}^{\sum_{j=1}^t |A_j|}$ that maps $\mathrm{Ker}(R_{G,I_{[t]}})$ *injectively to* $\mathrm{Ker}(M)$.

Proof. For each $i \in [n]$, let $j_i \in [t]$ be the smallest index satisfying $i \in I_{j_i}$. Such an index j_i always exists as $\bigcup_{j\in[t]} I_j = [n]$. Define the linear map $\phi : \mathbb{F}^{(t-1)k} \to \mathbb{F}^n$ sending $x = (x_1, \dots, x_{t-1}) \in (\mathbb{F}^k)^{t-1}$ to $(\phi_1(x), \dots, \phi_n(x)) \in \mathbb{F}^n$ such that $\phi_i(x) := G_{i,[k]}x_{j_i}$ for $i \in [n]$. Also define the linear map $\psi : \mathbb{F}^{(t-1)k} \to \mathbb{F}^{\sum_{j=1}^t |A_j|}$ sending $x = (x_1, \dots, x_{t-1}) \in (\mathbb{F}^k)^{t-1}$ to $(-y_1, y_2, \dots, y_t)$ such that $y_j = (\phi(x) - Gx_j)|_{A_j} \in \mathbb{F}^{|A_j|}$ for $j \in [t]$, where we let $x_t = 0 \in \mathbb{F}^k$. We will show that ψ is the desired linear map.

Consider $x = (x_1, \dots, x_{t-1}) \in \mathrm{Ker}(R_{G,I_{[t]}})$ and let $x_t = 0 \in \mathbb{F}^k$. We claim that $(\phi(x) - Gx_j)|_{I_j} = 0$ for $j \in [t]$. To see this, consider arbitrary $j \in [t]$ and $i \in I_j$. Then the i-th coordinate of $\phi(x) - Gx_j$ is $G_{i,[k]}x_{j_i} - G_{i,[k]}x_j$. By definition, either j is the only index in $[t]$ satisfying $i \in I_j$ (and hence $j = j_i$), or $R_{G,I_{[t]}}$ has a row that expresses the linear equation $G_{i,[k]}x_{j_i} - G_{i,[k]}x_j$ that $(x_1, \dots, x_{t-1}) \in \mathrm{Ker}(R_{G,I_{[t]}})$ and $x_t = 0$ must satisfy. In either case, we have $G_{i,[k]}x_{j_i} - G_{i,[k]}x_j = 0$. This proves the claim that

$$(\phi(x) - Gx_j)|_{I_j} = 0 \text{ for } j \in [t]. \tag{3.3}$$

For $j \in [t]$, let $y_j = (\phi(x) - Gx_j)|_{A_j}$, and we have

$$\begin{aligned} H_{[n-k],A_j} y_j &= H_{[n-k],A_j}(\phi(x) - Gx_j)|_{A_j} \\ &= H(\phi(x) - Gx_j) - H_{[n-k],I_j}(\phi(x) - Gx_j)|_{I_j} \\ &= H\phi(x) \end{aligned} \tag{3.4}$$

where the last equality holds by the fact $HG = 0$ and (3.3).

Now for arbitrary $j \in [t-1]$, the j-th block of $M\psi(x)$ equals $H_{[n-k],A_1}(-y_1) + H_{[n-k],A_{j+1}}y_{j+1}$ by definition, which equals zero by (3.4). So $M\psi(x) = 0$. This proves $\psi(\mathrm{Ker}(R_{G,I_{[t]}})) \subseteq \mathrm{Ker}(M)$.

Now further assume $\psi(x) = 0$, i.e., $(\phi(x) - Gx_j)|_{A_j} = 0$ for $j \in [t]$. Combining this with (3.3) and the fact $x_t = 0$, we see that $Gx_j = \phi(x) = Gx_t = 0$ for $j \in [t]$. As G has full column rank, this implies $x_1 = \dots = x_{t-1} = 0$, i.e., $x = 0$. So ψ maps $\mathrm{Ker}(R_{G,I_{[t]}})$ injectively to $\mathrm{Ker}(M)$. $\square$

We now combine the statements established so far to prove the following key lemma, which states that under certain conditions, a submatrix $R^B_{G,I_{[t]}}$ of a reduced intersection matrix $R_{G,I_{[t]}}$ with $G = V_{n,k}$ has full column rank if the set $B \subseteq [n]$ is small enough. This lemma will be used in our analysis in Section 4.

Lemma 3.11. *Let $G = V_{n,k}$. Let $\lambda \geq 0$ and let $t > 1$ be an integer. Let $I_1, \ldots, I_t \subseteq [n]$ such that $\mathrm{wt}(I_{[t]}) \geq (1+\lambda)(t-1)k$ and $\mathrm{wt}(I_J) \leq (1+\lambda)(|J|-1)k$ for all nonempty $J \subseteq [t]$. Finally, let $B \subseteq [n]$ such that $|B|(t-1) \leq \lambda k$. Then $\mathrm{Ker}(R^B_{G,I_{[t]}}) = 0$, i.e., $R^B_{G,I_{[t]}}$ has full column rank.*

Proof. Let $I'_j = I_j \setminus B$ for $j \in [t]$. Then $R^B_{G,I_{[t]}} = R_{G,I'_{[t]}}$ and we want prove $\mathrm{Ker}(R_{G,I'_{[t]}}) = 0$. Let $S = \bigcup_{j\in[t]} I'_j$ and $n' = |S|$. Note that permuting the indices in $[n]$ and changing $I'_1, \ldots, I'_t$ accordingly corresponds to permuting the variables X_i in $R_{G,I'_{[t]}}$, which does not change the dimension of $\mathrm{Ker}(R_{G,I'_{[t]}})$. So by applying a permutation on $[n]$, we may assume $S = [n']$. Let $G' = G_{[n'],[k]} = V_{n',k}$. Note $R_{G,I'_{[t]}} = R_{G',I'_{[t]}}$ since the definition of $R_{G,I'_{[t]}}$ only uses the rows of G whose indices are in $S = [n']$ (see Definition 3.2).

Also note that

$$\mathrm{wt}(I'_{[t]}) \geq \mathrm{wt}(I_{[t]}) - |B|(t-1) \geq (1+\lambda)(t-1)k - \lambda k \tag{3.5}$$

and

$$\mathrm{wt}(I'_J) \leq \mathrm{wt}(I_J) \leq (1+\lambda)(|J|-1)k \quad \text{for } \emptyset \neq J \subseteq [t]. \tag{3.6}$$

Let $A_j = [n'] \setminus I'_j$ for $j \in [t]$. For $j \in [t]$, we have $|I'_j| \geq \mathrm{wt}(I'_{[t]}) - \mathrm{wt}(I'_{[t]\setminus\{j\}}) \overset{(3.5),(3.6)}{\geq} k$ and hence $|A_j| \leq n' - k$.

Let $H = V^\top_{n',n'-k}$ and

$$M = \begin{pmatrix} H_{[n'-k],A_1} & H_{[n'-k],A_2} & & & \\ H_{[n'-k],A_1} & & H_{[n'-k],A_3} & & \\ \vdots & & & \ddots & \\ H_{[n'-k],A_1} & & & & H_{[n'-k],A_t} \end{pmatrix}.$$

By Theorem 3.6 and Theorem 3.7, we have

$$\dim\left(\bigcap_{i=1}^{t} \mathrm{Im}(H_{[n'-k],A_i})\right) = \max_{P_1\sqcup P_2\sqcup\cdots\sqcup P_s=[t]}\left(\sum_{i=1}^{s}\Big|\bigcap_{j\in P_i} A_j\Big| - (s-1)(n'-k)\right). \tag{3.7}$$

Now we calculate the RHS of (3.7). In the case where $s = 1$ and $P_1 = [t]$, as $[n'] = \bigcup_{j\in[t]} I'_j$, we have

$$\Big|\bigcap_{j\in[t].} A_j\Big| = n' - \Big|\bigcup_{j\in[t]} I'_j\Big| = 0. \tag{3.8}$$

For $s \geq 2$ and nonempty sets $P_1, \ldots, P_s$ that form a partition of $[t]$,

$$\begin{aligned}\sum_{i=1}^{s}\Big|\bigcap_{j\in P_i} A_j\Big| &= \sum_{i=1}^{s}\left(n' - \Big|\bigcup_{j\in P_i} I'_j\Big|\right) = sn' + \sum_{i=1}^{s}\mathrm{wt}(I'_{P_i}) - \sum_{i=1}^{s}\sum_{j\in P_i}|I'_j| \\ &\overset{(3.6)}{\leq} sn' + (1+\lambda)k\sum_{i=1}^{s}(|P_i|-1) - \sum_{j\in[t]}|I'_j| \\ &= sn' + (1+\lambda)(t-s)k - \mathrm{wt}(I'_{[t]}) - n' \\ &\overset{(3.5)}{\leq} (s-1)n' + (1+\lambda)(t-s)k - (1+\lambda)(t-1)k + \lambda k \\ &\leq (s-1)(n'-k)\end{aligned} \tag{3.9}$$

where the last inequality uses the fact $s \geq 2$. Combining (3.7), (3.8), and (3.9) yields $\bigcap_{i=1}^{t} \mathrm{Im}(H_{[n'-k],A_i}) = 0$. Now by Lemma 3.8, we have

$$\mathrm{rank}(M) = \sum_{i=1}^{t} \dim(\mathrm{Im}(H_{[n'-k],A_i})) - \dim\left(\bigcap_{i=1}^{t} \mathrm{Im}(H_{[n'-k],A_i})\right) = \sum_{i=1}^{t} \dim(\mathrm{Im}(H_{[n'-k],A_i})) = \sum_{i=1}^{t} |A_i|$$

where the last equality uses the facts that $H^\top = V_{n',n'-k}$ is an MDS code and $|A_j| \leq n' - k$ for $j \in [t]$. So M has full column rank, or equivalently, $\mathrm{Ker}(M) = 0$.

By Lemma 2.2, we have $H(DG') = V_{n',n'-k}^\top D V_{n',k} = 0$, where $D = \mathrm{diag}(v_1, \ldots, v_{n'})$ and $v_i = \prod_{j \in [n'] \setminus \{i\}} \frac{1}{X_i - X_j}$ for $i \in [n']$. Also note $\bigcup_{j \in [t]} I'_j = S = [n']$. Applying Lemma 3.10 to DG', H, and $I'_1, \ldots I'_t \subseteq [n']$ then yields $\mathrm{Ker}(R_{DG',I'_{[t]}}) = 0$. Observe that by the definition of reduced intersection matrices (Definition 3.2), $R_{DG',I'_{[t]}}$ can be obtained from $R_{G',I'_{[t]}}$ by scaling its rows by the nonzero elements v_i, and this does not change the dimension of its kernel. So $\mathrm{Ker}(R_{G',I'_{[t]}}) = 0$, i.e., $\mathrm{Ker}(R_{G,I'_{[t]}}) = 0$. □

4 Capacity-Achieving RS Codes over Polynomial-Size Alphabets

We prove our main results (Theorem 1.1 and Corollary 1.2) in this section. To achieve this, we design an algorithm to certify that a given reduced intersection matrix has full column rank even after assigning the variables. We then bound the probability that the algorithm fails to do so under a randomly chosen assignment. Our main results follow easily from this bound.

4.1 Full Rankness of Reduced Intersection Matrices under a Random Assignment

Fix $n, k \in \mathbb{N}^+$ and a finite field $\mathbb{F}_q$ in the following. Given a matrix $A \in \mathbb{F}^{m \times \ell}$ with $m \geq \ell$, we can order its $\ell \times \ell$ submatrices according to the lexicographic order of the row indices. Call this the *lexicographic order* on the $\ell \times \ell$ submatrices of A. The formal definition is given as follows.

Definition 4.1 (Lexicographic order)**.** Let $A \in \mathbb{F}^{m \times \ell}$ be a matrix over a field $\mathbb{F}$, where $m \geq \ell$. The *lexicographic order* $\prec$ on the $\ell \times \ell$ submatrices M of A is the lexicographic order determined by the column indices of M. In other words, if M and M' have the column indices $i_1 < i_2 < \cdots < i_\ell$ and $i'_1 < i'_2 < \cdots < i'_\ell$ respectively, then $M \prec M'$ if $(i_1, \ldots, i_\ell)$ is smaller than $(i'_1, \ldots, i'_\ell)$ in the lexicographic order.

We also introduce the following notation to denote a matrix that is obtained from a symbolic matrix by assigning a subset of variables.

Definition 4.2 (Partial assignment)**.** Let $A \in \mathbb{F}_q(X_1, \ldots, X_n)^{m \times \ell}$ be a matrix such that the entries of A are in $\mathbb{F}_q[X_1, \ldots, X_n]$. For $i \in \{0, 1, \ldots, n\}$ and $\alpha_1, \ldots, \alpha_i \in \mathbb{F}_q$, denote by $A|_{X_1=\alpha_1,\ldots,X_i=\alpha_i}$ the matrix obtained from A by substituting α_j for X_j for $j = 1, \ldots, i$. Note that the entries of $A|_{X_1=\alpha_1,\ldots,X_i=\alpha_i}$ are in $\mathbb{F}_q[X_{i+1}, \ldots, X_n]$.

Finally, we need the notion of *faulty indices*. Given a matrix $A \in \mathbb{F}_q(X_1, \ldots, X_n)^{m \times \ell}$ of full column rank, consider the process of gradually assigning $X_i = \alpha_i$ for $i = 1, \ldots, n$. The faulty index of A is the index where the smallest $\ell \times \ell$ nonsingular submatrix of A in the lexicographic order changes from nonsingular to singular. The formal definition is given as follows.

Definition 4.3 (Faulty index)**.** Let $A \in \mathbb{F}_q(X_1, \ldots, X_n)^{m \times \ell}$ be a matrix such that $\mathrm{rank}(A) = \ell$ and the entries of A are in $\mathbb{F}_q[X_1, \ldots, X_n]$. For $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$, we say $i \in [n]$ is the *faulty index* of A (with respect to $\alpha_1, \ldots, \alpha_n$) if $\det(M|_{X_1=\alpha_1,\ldots,X_{i-1}=\alpha_{i-1}}) \neq 0$ but $\det(M|_{X_1=\alpha_1,\ldots,X_i=\alpha_i}) = 0$, where M is the smallest nonsingular $\ell \times \ell$ submatrix of A in the lexicographic order.

Note that by definition, the faulty index of A is uniquely determined by $\alpha_1, \ldots, \alpha_n$ if it exists. Also note that if i is the faulty index of A and M is the smallest nonsingular $\ell \times \ell$ submatrix of A in the lexicographic order, then $\det(M|_{X_1=\alpha_1,\ldots,X_j=\alpha_j})$ is nonzero for $j = 0, 1, \ldots, i-1$ and is zero for $j = i, i+1, \ldots, n$.

Next, we describe an algorithm that given $I_1, \ldots, I_t \subseteq [n]$, $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$, and a parameter $r \in \mathbb{N}^+$, tries to certify that the matrix $R_{G,I_{[t]}}|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}$ has full column rank, where $G = V_{n,k}$. When it fails to do so, it outputs either "FAIL" or a sequence of faulty indices $(i_1, \ldots, i_r) \in [n]^r$. See Algorithm 1 for the pseudocode of this algorithm.

Algorithm 1: `CertifyFullColumnRankness`

Input: Sets $I_1, \ldots, I_t \subseteq [n]$, $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$, and $r \in \mathbb{N}^+$.
Output: "SUCCESS", "FAIL", or $(i_1, \ldots, i_r) \in [n]^r$.

Let $G \leftarrow V_{n,k}$ and $B \leftarrow \emptyset$.
for $j = 1$ **to** r **do**
  if $\operatorname{rank}(R^B_{G,I_{[t]}}) < (t-1)k$ **then**
    Output "FAIL" and halt.
  else if *the faculty index $i \in [n]$ of $R^B_{G,I_{[t]}}$ exists* **then**
    $i_j \leftarrow i$ and $B \leftarrow B \cup \{i\}$.
  else
    Output "SUCCESS" and halt.
  end
end
Output $(i_1, \ldots, i_r)$.

For the input values in which we are interested, the behavior of Algorithm 1 is described by the following lemma.

Lemma 4.4. *Let $\lambda \geq 0$ and let $t > 1$ be an integer. Let $I_1, \ldots, I_t \subseteq [n]$ such that $\operatorname{wt}(I_{[t]}) \geq (1+\lambda)(t-1)k$ and $\operatorname{wt}(I_J) \leq (1+\lambda)(|J|-1)k$ for all nonempty $J \subseteq [t]$. Let r be a positive integer such that $r \leq \lambda k/(t-1)+1$. Then for all $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$, running Algorithm 1 on the input $I_1, \ldots, I_t$, $\alpha_1, \ldots, \alpha_n$, and r yields one of the following two possible scenarios:*

(1) Algorithm 1 outputs "SUCCESS". In this case, $R_{G,I_{[t]}}|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}$ has full column rank.

(2) Algorithm 1 outputs an r-tuple $(i_1, \ldots, i_r) \in [n]^r$ consisting of r distinct indices. In this case, for each $j \in [r]$, i_j is the faulty index of $R^{B_j}_{G,I_{[t]}}$, where $B_j := \{i_1, \ldots, i_{j-1}\}$.

Proof. If the algorithm reaches the j-th round of the loop, where $j \in [r]$, then at the beginning of this round, we have $|B| = j-1 \leq r-1 \leq \lambda k/(t-1)$. Then by Lemma 3.11, the matrix $R^B_{G,I_{[t]}}$ has full column rank, i.e., $\operatorname{rank}(R^B_{G,I_{[t]}}) = (t-1)k$. So the algorithm never outputs "FAIL".

Suppose the algorithm outputs "SUCCESS" and halts in the j-th round for some $j \in [r]$, which means the faculty index of $R^B_{G,I_{[t]}}$ does not exist in that round. Let M be the smallest nonsingular $\ell \times \ell$ submatrix of $R^B_{G,I_{[t]}}$ in the lexicographic order. Then $\det(M|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}) \neq 0$, i.e., $M|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}$ is nonsingular. As $R^B_{G,I_{[t]}}$ is a submatrix of $R_{G,I_{[t]}}$ obtained by deleting rows, we see that $R_{G,I_{[t]}}|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}$ has full column rank.

Now suppose the algorithm does not output "SUCCESS". Then it outputs some $(i_1, \ldots, i_r) \in [n]^r$ where i_j is the faulty index of $R^{B_j}_{G,I_{[t]}}$ and $B_j = \{i_1, \ldots, i_{j-1}\}$ for $j \in [r]$. Note that $i_1, \ldots, i_r$ must be distinct.

This is because if an index i is in B_j, then $R^{B_j}_{G,I_{[t]}}$ does not depend on X_i, and hence i cannot be the faulty index of $R^{B_j}_{G,I_{[t]}}$. □

The next lemma bounds the probability that Algorithm 1 outputs a particular sequence of faulty indices over randomly chosen $(\alpha_1, \ldots, \alpha_n)$.

Lemma 4.5. *Under the notations and conditions in Lemma 4.4, suppose $q \geq n$ and $(\alpha_1, \ldots, \alpha_n)$ is chosen uniformly at random from the set of all n-tuples of distinct elements in $\mathbb{F}_q$. Then for any r-tuple $(i_1, \ldots, i_r) \in [n]^r$ of distinct indices, the probability that Algorithm 1 outputs $(i_1, \ldots, i_r)$ on the input $I_1, \ldots, I_t$, $\alpha_1, \ldots, \alpha_n$, and r is at most $\left(\frac{(t-1)(k-1)}{q-n+1}\right)^r$.*

Proof. For $j \in [r]$, define the following:

1. $B_j := \{i_1, \ldots, i_{j-1}\}$.
2. Let M_j be the smallest nonsingular $(t-1)k \times (t-1)k$ submatrix of $R^{B_j}_{G,I_{[t]}}$ in the lexicographic order. As argued in the proof of Lemma 4.4, since $|B_j| = j-1 \leq r-1 \leq \lambda k/(t-1)$, the matrix $R^{B_j}_{G,I_{[t]}}$ has full column rank and hence M_j exists.
3. Let E_j be the event that $\det(M_j|_{X_1=\alpha_1,\ldots,X_{i_j-1}=\alpha_{i_j-1}}) \neq 0$ but $\det(M_j|_{X_1=\alpha_1,\ldots,X_{i_j}=\alpha_{i_j}}) = 0$.

If Algorithm 1 outputs $(i_1, \ldots, i_r)$, then i_j is the faulty index of $R^{B_j}_{G,I_{[t]}}$ for $j \in [r]$ by Lemma 4.4 and hence the events $E_1, \ldots, E_r$ all occur. So we just need to prove that $\Pr[E_1 \wedge \cdots \wedge E_r] \leq \left(\frac{(t-1)(k-1)}{q-n+1}\right)^r$.

Let $(j_1, j_2, \ldots, j_r)$ be a permutation of $(1, 2, \ldots, r)$ such that $i_{j_1} < \cdots < i_{j_r}$, i.e., i_{j_ℓ} is the ℓ-th smallest index among $i_1, \ldots, i_r$ for $\ell \in [r]$. For $\ell \in \{0, 1, \ldots, r\}$, define $F_\ell := E_{j_1} \wedge \cdots \wedge E_{j_\ell}$, where we let F_0 be the event that always occurs. Then $F_r = E_{j_1} \wedge \cdots \wedge E_{j_r} = E_1 \wedge \cdots \wedge E_r$. If $\Pr[F_r] = 0$ then we are done. So assume $\Pr[F_r] > 0$. By definition, if F_ℓ occurs and $\ell' < \ell$, then $F_{\ell'}$ also occurs. So $\Pr[F_\ell] > 0$ for all $\ell \in \{0, 1, \ldots, r\}$. Note

$$\Pr[E_1 \wedge \cdots \wedge E_r] = \Pr[F_r] = \prod_{\ell=1}^{r} \frac{\Pr[F_\ell]}{\Pr[F_{\ell-1}]}.$$

So it suffices to prove that $\frac{\Pr[F_\ell]}{\Pr[F_{\ell-1}]} \leq \frac{(t-1)(k-1)}{q-n+1}$ for $\ell \in [r]$.

Fix $\ell \in [r]$ and let $j = j_\ell$. Let S be the set of all $\beta = (\beta_1, \ldots, \beta_{i_j-1}) \in \mathbb{F}_q^{i_j-1}$ such that $\Pr\left[\left(\alpha_{<i_j} = \beta\right) \wedge F_{\ell-1}\right] > 0$, where $\alpha_{<i_j} = \beta$ is a shorthand for $(\alpha_1 = \beta_1) \wedge \cdots \wedge (\alpha_{i_j-1} = \beta_{i_j-1})$. Note that for $\beta \in S$, the event $\left(\alpha_{<i_j} = \beta\right) \wedge F_{\ell-1}$ is simply $\alpha_{<i_j} = \beta$ since $F_{\ell-1} = E_{j_1} \wedge \cdots \wedge E_{j_{\ell-1}}$ depends only on $\alpha_1, \ldots, \alpha_{i_{j_{\ell-1}}}$ and is bound to happen conditioned on $\alpha_{<i_j} = \beta$. We then have

$$\begin{aligned}
\frac{\Pr[F_\ell]}{\Pr[F_{\ell-1}]} &= \frac{\sum_{\beta \in S} \Pr\left[\left(\alpha_{<i_j} = \beta\right) \wedge F_\ell\right]}{\sum_{\beta \in S} \Pr\left[\left(\alpha_{<i_j} = \beta\right) \wedge F_{\ell-1}\right]} = \frac{\sum_{\beta \in S} \Pr\left[\left(\alpha_{<i_j} = \beta\right) \wedge E_j\right]}{\sum_{\beta \in S} \Pr\left[\alpha_{<i_j} = \beta\right]} \\
&\leq \max_{\beta \in S} \frac{\Pr\left[\left(\alpha_{<i_j} = \beta\right) \wedge E_j\right]}{\Pr\left[\alpha_{<i_j} = \beta\right]} = \max_{\beta \in S} \Pr\left[E_j \mid \alpha_{<i_j} = \beta\right].
\end{aligned}$$

Fix $\beta = (\beta_1, \ldots, \beta_{i_j-1}) \in S$. We just need to prove that $\Pr\left[E_j \mid \alpha_{<i_j} = \beta\right] \leq \frac{(t-1)(k-1)}{q-n+1}$. Let

$$Q := \det(M_j|_{X_1=\beta_1,\ldots,X_{i_j-1}=\beta_{i_j-1}}) \in \mathbb{F}_q[X_{i_j}, \ldots, X_n].$$

If $Q = 0$, then E_j never occurs conditioned on $\alpha_{<i_j} = \beta$ and we are done. So assume $Q \neq 0$. View Q as a polynomial in $X_{i_j+1}, \ldots, X_n$ over the ring $\mathbb{F}_q[X_{i_j}]$, and let $Q_0 \in \mathbb{F}_q[X_{i_j}]$ be the coefficient of a nonzero

term of Q. Then conditioned on $\alpha_{<i_j} = \beta$, the event E_j occurs only if α_{i_j} is a root of $Q_0 \neq 0$. Note that $\deg Q_0 \leq \deg_{X_{i_j}} Q \leq \deg_{X_{i_j}} (\det(M_j))$, which is bounded by $(t-1)(k-1)$ by Lemma 3.4. Also note that conditioned on $\alpha_{<i_j} = \beta$, the random variable α_{i_j} is uniformly distributed over the set $\mathbb{F}_q \setminus \{\beta_1, \ldots, \beta_{i_j-1}\}$, whose size is at least $q-n+1$. It follows that $\Pr\left[E_j \mid \alpha_{<i_j} = \beta\right] \leq \frac{(t-1)(k-1)}{q-n+1}$, as desired. □

By combining Lemma 4.4 and Lemma 4.5 and taking a union bound over the set of possible outputs $(i_1, \ldots, i_r)$ of the algorithm, we obtain the following corollary.

Corollary 4.6. *Under the notations and conditions in Lemma 4.4, suppose $q \geq n$ and $(\alpha_1, \ldots, \alpha_n)$ is chosen uniformly at random from the set of all n-tuples of distinct elements in $\mathbb{F}_q$. Then*

$$\Pr\left[\operatorname{Ker}(R_{G,I_{[t]}}|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}) \neq 0\right] \leq \left(\frac{(t-1)n(k-1)}{q-n+1}\right)^r.$$

Proof. By Lemma 4.5 and the union bound, the probability that Algorithm 1 outputs some sequence $(i_1, \ldots, i_r) \in [n]^r$ on the input $I_1, \ldots, I_t, \alpha_1, \ldots, \alpha_n$, and r is at most $n^r \cdot \left(\frac{(t-1)(k-1)}{q-n+1}\right)^r = \left(\frac{(t-1)n(k-1)}{q-n+1}\right)^r$. By Lemma 4.4, whenever this does not occur, the matrix $R_{G,I_{[t]}}|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}$ has full column rank, i.e., its kernel is zero. The claim follows. □

4.2 Proof of Theorem 1.1 and Corollary 1.2

We are now ready to prove our main results, Theorem 1.1 and Corollary 1.2.

Theorem 4.7 (Theorem 1.1 restated)**.** *Let $\varepsilon \in (0,1)$, $c > 2$, and $n, k, L \in \mathbb{N}^+$ with $k \leq n$. Let q be a prime power such that $q \geq 2^{\frac{L(L+c)}{\varepsilon}} Ln(k-1) + n$. Then with probability at least $1 - 2^{-(c-2)n} > 0$, a randomly punctured RS code of block length n and rate $R = k/n$ over $\mathbb{F}_q$ is $\left(\frac{L}{L+1}(1-R-\varepsilon), L\right)$ average-radius list decodable.*

Proof. Let $\lambda = \varepsilon/R = \varepsilon n/k$. By Lemma 3.5, if an $[n,k]$ linear code with a generating matrix $\overline{G}$ is not $\left(\frac{L}{L+1}(1-R-\varepsilon), L\right)$ average-radius list decodable, then there exist $t \in \{2, 3, \ldots, L+1\}$ and $I_1, \ldots, I_t \subseteq [n]$ such that

(1) $\operatorname{Ker}(R_{\overline{G},I_{[t]}}) \neq 0$,

(2) $\operatorname{wt}(I_{[t]}) \geq (1+\lambda)(t-1)k$, and

(3) $\operatorname{wt}(I_J) \leq (1+\lambda)(|J|-1)k$ for all nonempty $J \subsetneq [t]$.

Choose $(\alpha_1, \ldots, \alpha_n)$ uniformly at random from the set of n-tuples of distinct elements in $\mathbb{F}_q$. Let $\overline{G} = V_{n,k}(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_q^{n \times k}$. Then to prove the theorem, it suffices to show that the probability that there exist $t \in \{2, 3, \ldots, L+1\}$ and $I_1, \ldots, I_t \subseteq [n]$ satisfying (1)–(3) above is at most $2^{-(c-2)n}$.

Fix $t \in \{2, 3, \ldots, L+1\}$ and $I_1, \ldots, I_t \subseteq [n]$ satisfying (2) and (3) above. Let $r = \lfloor \lambda k/(t-1) + 1 \rfloor \geq \lambda k/(t-1) = \varepsilon n/(t-1)$. Note that $R_{\overline{G},I_{[t]}}$ is exactly $R_{G,I_{[t]}}|_{X_1=\alpha_1,\ldots,X_n=\alpha_n}$, where $G = V_{n,k}$. So by Corollary 4.6, the probability that $\operatorname{Ker}(R_{\overline{G},I_{[t]}}) \neq 0$ holds is at most $\left(\frac{(t-1)n(k-1)}{q-n+1}\right)^r \leq \left(\frac{Ln(k-1)}{q-n+1}\right)^{\varepsilon n/L}$. Here we use the facts $r \geq \varepsilon n/(t-1) \geq \varepsilon n/L$ and $q-n+1 \geq Ln(k-1) \geq (t-1)n(k-1)$.

The number of $(t, I_1, \ldots, I_t)$ with $t \in \{2, 3, \ldots, L+1\}$ and $I_1, \ldots, I_t \subseteq [n]$ is bounded by $\sum_{t=2}^{L+1} 2^{tn} \leq 2^{(L+2)n}$. By the union bound, the probability that (1)–(3) above hold for some $t \in \{2, 3, \ldots, L+1\}$ and $I_1, \ldots, I_t \subseteq [n]$ is at most $2^{(L+2)n}\left(\frac{Ln(k-1)}{q-n+1}\right)^{\varepsilon n/L} \leq 2^{-(c-2)n}$, as desired. □

Remark. While Theorem 4.7 is stated with $\varepsilon > 0$, its proof still makes sense when $\varepsilon = 0$. In this case, $\lambda = \varepsilon/R = 0$ and we choose $r = \lfloor \lambda k/(t-1) + 1 \rfloor = 1$. Then the above proof shows that for sufficiently large q that is (at least) exponential in n, a randomly punctured RS code of block length n over $\mathbb{F}_q$ is $\left(\frac{L}{L+1}(1-R), L\right)$ average-radius list decodable with high probability, recovering one of the main results of [BGM22a].

Similarly, we restate and prove (a generalization of) Corollary 1.2 as follows.

Corollary 4.8 (Corollary 1.2 restated)**.** *Let $\varepsilon, \delta \in (0,1)$, $c > 2$, $n, k \in \mathbb{N}^+$ with $k \le n$, and $R = k/n$. Let $L = \max\left\{\left\lceil \frac{1-R}{(1-\delta)\varepsilon} \right\rceil - 1, 1\right\}$. Let q be a prime power such that $q \ge 2^{\frac{L(L+c)}{\delta\varepsilon}} Ln(k-1) + n$. Then with probability at least $1 - 2^{-(c-2)n} > 0$, a randomly punctured RS code of block length n and rate R over $\mathbb{F}_q$ is $(1 - R - \varepsilon, L)$ average-radius list decodable.*

Proof. Let $\varepsilon' = \delta\varepsilon$. By Theorem 4.7, with probability at least $1 - 2^{-(c-2)n}$, a randomly punctured RS code of block length n and rate $R = k/n$ over $\mathbb{F}_q$ is $\left(\frac{L}{L+1}(1 - R - \varepsilon'), L\right)$ average-radius list decodable. Then note that $\frac{L}{L+1}(1 - R - \varepsilon') \ge 1 - R - \frac{1-R}{L+1} - \varepsilon' \ge 1 - R - \varepsilon$. □

5 Future Work and Open Problems

We list some open problems and directions for future work.

Further improving the alphabet size. For $(1 - R - \varepsilon, O(1/\varepsilon))$ list decodable RS codes, the smallest alphabet size we are able to achieve in this paper is $q = O\left(2^{\mathrm{poly}(1/\varepsilon)} nk\right)$, which is quadratic in the block length n when the rate $R = k/n$ is a positive constant. We conjecture that it can be improved to a function linear in n. Formally, we ask the following question.

Question 5.1. *For $\varepsilon > 0$ and positive integers n, k with $k \le n$, do there exist RS codes of block length n and rate $R = k/n$ that are $(1 - R - \varepsilon, O(1/\varepsilon))$ list decodable over an alphabet of size at most $c_\varepsilon n$, where c_ε is a constant only depending on ε? And how small can c_ε be in terms of ε?*

It is possible that achieving an alphabet size linear in n would require establishing and exploiting other properties of intersection matrices or reduced intersection matrices, such as an appropriate notion of exchangeability. We leave this question for future studies.

Efficient encoding and list decoding algorithms for RS codes achieving the capacity. Given the results in [BGM22a] and in this paper that randomly punctured RS codes achieve the list decoding capacity, the next natural (although challenging) questions are if there exist explicit constructions of capacity-achieving RS codes, and if they admit efficient list decoding algorithms.

Question 5.2. *Are there any explicit constructions of RS codes that achieve the list decoding capacity over alphabets of size at most exponential (or even polynomial) in the block length of the codes?*

Question 5.3. *Are there efficient algorithms, deterministic or randomized, that list decode an $[n, k]$ RS code C up to the radius $1 - k/n - \varepsilon$, assuming that C is$(1 - k/n - \varepsilon, O(1/\varepsilon))$ list decodable and a generating matrix of C is given to the algorithm?*

Remark. Cheng and Wan [CW07] proved that unless the discrete logarithm problem has an efficient algorithm, it is impossible to efficiently list decode RS codes over $\mathbb{F}_q$ up to the relative radius $1 - \widehat{g}/n$, where $\widehat{g} := \min\left\{g \in \mathbb{N} : \binom{n}{g} q^{k-g} \le 1\right\}$. One can estimate that $1 - \widehat{g}/n \ge 1 - k/n - 1/\log q$. Thus, this result does not rule out the possibility of an affirmative answer to Question 5.3 when $\varepsilon > 1/\log q$.

Further applications of our techniques. Finally, we find our techniques to be quite general and versatile. Therefore, we expect that these techniques or their generalizations can find applications in the study of other algebraic constructions of linear codes and other coding-theoretic problems.

References

[Ahl73] Rudolf Ahlswede. Channel capacities for list codes. *Journal of Applied Probability*, 10(4):824–836, 1973.

[BDG22] Joshua Brakensiek, Manik Dhar, and Sivakanth Gopi. Improved field size bounds for higher order MDS codes. *arXiv preprint arXiv:2212.11262*, 2022.

[BGM22a] Joshua Brakensiek, Sivakanth Gopi, and Visu Makam. Generic Reed-Solomon codes achieve list-decoding capacity. *arXiv preprint arXiv:2206.05256*, 2022. To appear in STOC 2023.

[BGM22b] Joshua Brakensiek, Sivakanth Gopi, and Visu Makam. Lower bounds for maximally recoverable tensor codes and higher order MDS codes. *IEEE Transactions on Information Theory*, 68(11):7125–7140, 2022.

[Bli86] Volodia M. Blinovskiĭ. Bounds for codes in decoding by a list of finite length. *Problemy Peredachi Informatsii*, 22(1):11–25, 1986.

[Bli97] Volodia M. Blinovsky. *Asymptotic Combinatorial Coding Theory*. Springer Science & Business Media, 1997.

[BSKR09] Eli Ben-Sasson, Swastik Kopparty, and Jaikumar Radhakrishnan. Subspace polynomials and limits to list decoding of Reed–Solomon codes. *IEEE Transactions on Information Theory*, 56(1):113–120, 2009.

[CGV13] Mahdi Cheraghchi, Venkatesan Guruswami, and Ameya Velingker. Restricted isometry of Fourier matrices and list decodability of random linear codes. *SIAM Journal on Computing*, 42(5):1888–1914, 2013.

[CPS99] Jin-Yi Cai, Aduri Pavan, and D. Sivakumar. On the hardness of permanent. In *Proceedings of the 16th Annual Conference on Theoretical Aspects of Computer Science (STACS)*, pages 90–99. Springer, 1999.

[CW07] Qi Cheng and Daqing Wan. On the list and bounded distance decodability of Reed–Solomon codes. *SIAM Journal on Computing*, 37(1):195–209, 2007.

[DSY14] Son Hoang Dau, Wentu Song, and Chau Yuen. On the existence of MDS codes over small fields with constrained generator matrices. In *2014 IEEE International Symposium on Information Theory (ISIT)*, pages 1787–1791. IEEE, 2014.

[Eli57] Peter Elias. List decoding for noisy channels. *Wescon Convention Record, Part 2, Institute of Radio Engineers*, pages 99–104, 1957.

[Eli91] Peter Elias. Error-correcting codes for list decoding. *IEEE Transactions on Information Theory*, 37(1):5–12, 1991.

[FKS22] Asaf Ferber, Matthew Kwan, and Lisa Sauermann. List-decodability with large radius for Reed-Solomon codes. *IEEE Transactions on Information Theory*, 68(6):3823–3828, 2022.

[GHK11] Venkatesan Guruswami, Johan Håstad, and Swastik Kopparty. On the list-decodability of random linear codes. *IEEE Transactions on Information Theory*, 57(2):718–725, 2011.

[GL89] Oded Goldreich and Leonid A. Levin. A hard-core predicate for all one-way functions. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC)*, pages 25–32, 1989.

[GLM+21] Venkatesan Guruswami, Ray Li, Jonathan Mosheiff, Nicolas Resch, Shashwat Silas, and Mary Wootters. Bounds for list-decoding and list-recovery of random linear codes. *IEEE Transactions on Information Theory*, 68(2):923–939, 2021.

[GLS+22] Zeyu Guo, Ray Li, Chong Shangguan, Itzhak Tamo, and Mary Wootters. Improved list-decodability and list-recoverability of Reed-Solomon codes via tree packings. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 708–719. IEEE, 2022.

[GM22] Venkatesan Guruswami and Jonathan Mosheiff. Punctured low-bias codes behave like random linear codes. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 36–45. IEEE, 2022.

[GR08] Venkatesan Guruswami and Atri Rudra. Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy. *IEEE Transactions on Information Theory*, 54(1):135–150, 2008.

[GRS00] Oded Goldreich, Dana Ron, and Madhu Sudan. Chinese remaindering with errors. *IEEE Transactions on Information Theory*, 46(4):1330–1338, 2000.

[GRS19] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential Coding Theory*. 2019. Draft available at https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/.

[GRZ21] Zeyu Guo and Noga Ron-Zewi. Efficient list-decoding with constant alphabet and list sizes. *IEEE Transactions on Information Theory*, 68(3):1663–1682, 2021.

[GS01] Venkatesan Guruswami and Madhu Sudan. Extensions to the Johnson bound. Unpublished manuscript. Available at https://madhu.seas.harvard.edu/papers/2001/johnson.pdf, 2001.

[GS06] Venkatesan Guruswami and Madhu Sudan. Improved decoding of Reed-Solomon and algebraic-geometry codes. *IEEE Transactions on Information Theory*, 45(6):1757–1767, 2006.

[GST22] Eitan Goldberg, Chong Shangguan, and Itzhak Tamo. List-decoding and list-recovery of Reed-Solomon codes beyond the Johnson radius for every rate. *IEEE Transactions on Information Theory*, 2022.

[Gur04] Venkatesan Guruswami. *List decoding of error-correcting codes: winning thesis of the 2002 ACM doctoral dissertation competition*, volume 3282. Springer Science & Business Media, 2004.

[Gur09] Venkatesan Guruswami. Artin automorphisms, cyclotomic function fields, and folded list-decodable codes. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 23–32, 2009.

[GW13] Venkatesan Guruswami and Carol Wang. Linear-algebraic list decoding for variants of Reed–Solomon codes. *IEEE Transactions on Information Theory*, 59(6):3257–3268, 2013.

[GX22] Venkatesan Guruswami and Chaoping Xing. Optimal rate list decoding over bounded alphabets using algebraic-geometric codes. *Journal of the ACM*, 69(2):1–48, 2022.

[Joh62] Selmer Johnson. A new upper bound for error-correcting codes. *IRE Transactions on Information Theory*, 8(3):203–207, 1962.

[Kop15] Swastik Kopparty. List-decoding multiplicity codes. *Theory of Computing*, 11(1):149–182, 2015.

[KRZSW18] Swastik Kopparty, Noga Ron-Zewi, Shubhangi Saraf, and Mary Wootters. Improved decoding of folded Reed-Solomon and multiplicity codes. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 212–223. IEEE, 2018.

[Lov18] Shachar Lovett. MDS matrices over small fields: A proof of the GM-MDS conjecture. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 194–199. IEEE, 2018.

[MRRZ$^+$20] Jonathan Mosheiff, Nicolas Resch, Noga Ron-Zewi, Shashwat Silas, and Mary Wootters. LDPC codes achieve list decoding capacity. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 458–469. IEEE, 2020.

[Rot22] Ron M. Roth. Higher-order MDS codes. *IEEE Transactions on Information Theory*, 68(12):7798–7816, 2022.

[RS60] Irving S. Reed and Gustave Solomon. Polynomial codes over certain finite fields. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):300–304, 1960.

[Rud07] Atri Rudra. *List decoding and property testing of error-correcting codes*. PhD thesis, University of Washington, 2007.

[RW14] Atri Rudra and Mary Wootters. Every list-decodable code for high noise has abundant near-optimal rate puncturings. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pages 764–773, 2014.

[RW18] Atri Rudra and Mary Wootters. Average-radius list-recoverability of random linear codes. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 644–662. SIAM, 2018.

[Sch80] Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980.

[Sin64] Richard Singleton. Maximum distance q-nary codes. *IEEE Transactions on Information Theory*, 10(2):116–118, 1964.

[ST20] Chong Shangguan and Itzhak Tamo. Combinatorial list-decoding of Reed-Solomon codes beyond the Johnson radius. In *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing (STOC)*, pages 538–551, 2020.

[Sud97] Madhu Sudan. Decoding of Reed-Solomon codes beyond the error-correction bound. *Journal of Complexity*, 13(1):180 – 193, 1997.

[Sud00] Madhu Sudan. List decoding: algorithms and applications. *ACM SIGACT News*, 31(1):16–27, 2000.

[Tia19] Yongge Tian. Formulas for calculating the dimensions of the sums and the intersections of a family of linear subspaces with applications. *Beiträge zur Algebra und Geometrie*, 60:471–485, 2019.

[Vad12] Salil P. Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.

[Woo13] Mary Wootters. On the list decodability of random linear codes with large error rates. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC)*, pages 853–860, 2013.

[Woz58] John M. Wozencraft. List decoding. *Quarterly Progress Report, Research Laboratory of Electronics, MIT*, 48:90–95, 1958.

[YH19] Hikmet Yildiz and Babak Hassibi. Optimum linear codes with support-constrained generator matrices over small fields. *IEEE Transactions on Information Theory*, 65(12):7868–7875, 2019.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, EUROSAM ’79, page 216–226, 1979.