

Average-radius list decoding of random linear matrix codes at rank-metric capacity*

Matěj Kripner Václav Rozhoň Robert Šámal
Martin Šindelář Adrián Zámečník

Abstract

We show that for every fixed prime power q and $0 < p < 1$, a uniformly random $\mathbb{F}_q$ -linear rank-metric code in $\mathbb{F}_q^{n \times n}$ of rate $(1 - p)^2 - \varepsilon$ is, with high probability, average-radius list-decodable at radius pn with list size $O_{p,q}(1/\varepsilon)$. This answers a question of Guruswami, Li and Singhal; the case $q = 2$ was known. The failure probability is explicit and remains small when ε depends on n with $\varepsilon n^2 \rightarrow \infty$. The proof combines a weighted affine extraction, span estimates for matrices from nearby rank shells, and Gaussian summation over rank profiles.

1 Introduction

1.1 Background and the source problem

Rank measures the number of independent error directions, rather than the number of changed matrix entries: a nonzero outer product uv^T has rank one even when every entry is nonzero. This geometry leads to the rank metric, studied in Delsarte's work on bilinear forms [4]. Let q be a prime power and let $V_n = \mathbb{F}_q^{n \times n}$. The rank metric on V_n is $d(X, Y) = \text{rank}(X - Y)$, and a linear matrix code is an $\mathbb{F}_q$ -linear subspace $\mathcal{C} \leq V_n$. Its rate is $\dim_{\mathbb{F}_q} \mathcal{C} / n^2$. Ordinary list decoding controls the number of codewords in a ball about any center. Average-radius list decoding controls the average distance of a list from its center, so it also excludes lists with some members outside the prescribed ball.

Definition 1.1 (Average-radius convention). A code $\mathcal{C} \subseteq V_n$ has no bad average-radius list of size L at radius pn if every $Y \in V_n$ and every L distinct codewords $c_1, \dots, c_L \in \mathcal{C}$ satisfy

$$\sum_{i=1}^L \text{rank}(c_i - Y) > pnL. \quad (1)$$

Here L is the *forbidden-list size*. In the convention of [1, Section 1], this is $(p, L - 1)$ -average-radius list-decodability, with normalized rank distance in place of normalized Hamming distance.

For square matrices, the exponent of a rank ball of relative radius p is $f(p)n^2$, where $f(p) = 2p - p^2$. The corresponding capacity rate is $1 - f(p) = (1 - p)^2$. Guruswami and Resch [2, Theorem 1.1] prove ordinary list-decodability of random linear rank-metric codes at this rate minus ε , with list size $O_{p,q}(1/\varepsilon)$. Their result also treats rectangular matrices. Guruswami, Li and Singhal prove the average-radius counterpart for the Hamming metric in [1, Theorem 1], and ask the following rank-metric question in their final section.

*This work was produced within the Bolzano project (<https://bolzano.app>), using AI systems for mathematical research and manuscript preparation. The listed authors contributed to the project and the preparation of this release. The manuscript received only limited human checking.

Question 1.2 ([1, Section 4, “Extension to the rank metric”]). Can the $O_{p,q}(1/\varepsilon)$ list-size result for random $\mathbb{F}_q$ -linear codes in $\mathbb{F}_q^{n \times n}$ of rate $(1-p)^2 - \varepsilon$ be extended to average-radius list decoding at radius pn ?

The binary case was already observed in [3, Remark 5.6]: its potential argument gives the average-radius guarantee for random linear rank-metric codes over $\mathbb{F}_2$, with a sharper list-size constant. The contribution here is the general finite-field assertion asked for in Question 1.2, together with a finite estimate uniform in the capacity gap. We restrict to square matrices, as in that question. No decoding algorithm or extension to every rectangular aspect ratio is asserted.

1.2 Main results

Our main result answers Question 1.2 affirmatively.

Theorem 1.3. *Fix a prime power q and $0 < p < 1$. There are positive constants B_1, B_2, B_3 , depending only on p, q , such that the following holds for every $n \geq 1$ and $0 < \varepsilon < (1-p)^2$. Set*

$$k = \left\lfloor ((1-p)^2 - \varepsilon)n^2 \right\rfloor,$$

and choose $\mathcal{C}$ uniformly among the k -dimensional subspaces of V_n . There is an integer $L \leq B_3/\varepsilon$ for which

$$\Pr[\mathcal{C} \text{ has a bad list of size } L] \leq q^{-2n^2 + B_1/\varepsilon + B_2 \log_q(n+1)}. \quad (2)$$

An explicit finite estimate and choice of L are given in Theorem 7.1 below.

The proof has three components.

1. Lemma 3.1 extracts either a large independent set of small average weight or a smaller independent set whose span contains many of the original points. Its affine version preserves the single-center condition in a codeword list.
2. Theorem 5.3 bounds the number of low-rank points in every translate of the span of matrices drawn from nearby rank shells. The bound is uniform over translations; no independence of the generators as vectors is required.
3. Proposition 6.1 combines this estimate with a quadratic rank-profile deficit. Theorem 7.1 then proves (2) by counting affine witnesses.

The proof follows the span-counting framework behind [2] and the greedy-basis approach in [1, Section 1.2]. The two-sum and span estimates in [2, Lemmas 4.1–4.2] concern generators drawn from a rank ball. Here those estimates are proved for prescribed nearby rank shells and all translated spans, then combined with a weighted profile count. This is what allows a constraint on the average rank.

The distinction between probabilistic independence of sampled matrices and linear independence of a witness is essential. The shell estimates use the former, while containment in a random code is charged only for the independent differences in an affine witness.

Corollary 1.4. *For fixed p, q , let $0 < \varepsilon_n < (1-p)^2$ and suppose $\varepsilon_n n^2 \rightarrow \infty$. There are forbidden-list sizes $L_n = O_{p,q}(1/\varepsilon_n)$ such that the random codes of Theorem 1.3 have no bad lists of size L_n with probability at least $1 - q^{-2n^2 + o(n^2)}$.*

Proof. The hypothesis gives $1/\varepsilon_n = o(n^2)$, and $\log(n+1) = o(n^2)$. Substitute these estimates in (2). $\square$

1.3 Related work

Rank-metric codes and list decoding. Rank-metric codes were introduced by Delsarte [4] and Gabidulin [5], who constructed the maximum rank distance codes now called Gabidulin codes; Roth [6] independently studied them for crisscross error correction. List decoding goes back to Elias and Wozencraft; see Elias [7]. For Gabidulin codes, Wachter-Zeh [8] showed that list sizes become exponential beyond the Johnson radius, and that no Johnson-type polynomial bound depending only on the length and minimum distance can hold in the rank metric. Capacity-achieving list decodability in the rank metric is therefore sought in other ensembles, such as random linear codes.

Random linear codes. The central difficulty for random linear codes is dependence between codewords. A list can be large while its linear span has small dimension, so estimates based only on independent codewords lose an exponential factor in the list size. Guruswami, Håstad and Kopparty [9, Theorems 6 and 10 of the preprint] overcame this obstacle in the Hamming metric by controlling how many points of a ball can lie in the span of a few randomly chosen ball points. Their method gives list size $O_{p,q}(1/\varepsilon)$ at capacity minus ε for every fixed field and admissible Hamming radius.

Rank metric. In the rank metric, Ding [10, Section IV] established capacity for random $\mathbb{F}_q$ -linear matrix codes with list size exponential in $1/\varepsilon$. Guruswami and Resch [2] reduced this to $O_{p,q}(1/\varepsilon)$ by adapting the span argument. For binary codes, Li and Wootters [11, Section 5] obtained a sharper ordinary-list constant through a potential-function argument. The extension of that argument to average radius is precisely the earlier binary result in [3, Remark 5.6 of the 2020 preprint]. Thus neither the binary theorem nor the ordinary rank-metric capacity theorem is a new claim here.

Average-radius list decoding. Average-radius guarantees also have a substantial Hamming-metric history. Rudra and Wootters [12, Corollary 3.4 of the preprint] gave capacity-rate guarantees with quasipolynomial list sizes in a specified constant-alphabet parameter range, within a framework also covering list recovery. The all-fixed-field linear dependence on $1/\varepsilon$ is supplied by Guruswami, Li and Singhal [1]. Their greedy-basis argument and entropy count motivate the weighted extraction and rank-profile count below. The two metrics have different volume exponents; in particular, the rank radius may be any $p < 1$, whereas their Hamming theorem assumes $p < 1 - 1/q$. For large Hamming alphabets, Goyal and Guruswami [16, Theorems 1.1–1.2] obtain improved alphabet-size bounds for average-radius guarantees approaching the generalized Singleton bound, for both random linear and additive codes. Their coordinatewise constraint count addresses a different metric and alphabet-size question.

Other ensembles and metrics. Other rank-metric results use different ensembles or guarantees. Guo, Xing, Yuan and Zhang [13, Theorem 4] prove average-radius list-decodability for random Gabidulin codes over $\mathbb{F}_{q^m}$. Writing their block length as ν and their gap parameter as δ , the theorem requires $m \geq \mathcal{O}(\ell + 1)\nu/\delta$ and gives radius $\frac{\ell}{\ell+1}(1 - R - \delta)$ with maximum list size ℓ . These are extension-field-linear codes represented by $m \times \nu$ matrices; this theorem does not supply the square, uniformly random $\mathbb{F}_q$ -subspace assertion considered here.

Liu, Baumeister and Wachter-Zeh [14, Theorem 6] extend the ordinary-list span method to the sum-rank metric, where the weight is the sum of ranks of matrix blocks. Their conclusion bounds the number of codewords in each sum-rank ball. Recent work of Li, Xing, Yuan and Zhu [15, Definition 1.1 and Theorems 1.2–1.3] instead controls how often points on an affine line are close to a rank-metric code, with applications to proofs of proximity. Neither statement is the average-distance guarantee for lists of codewords in Question 1.2.

1.4 Why the weighted argument works

A bad list may contain offsets of several different ranks. Removing those with rank above $(p + \eta)n$, for a small fixed $\eta > 0$, retains a fixed positive fraction of the list and leaves its average rank at most pn . The greedy affine extraction then gives two possibilities: many affinely independent offsets, whose simultaneous occurrence is unlikely, or a smaller witness whose affine span contains

many low-rank points. This second possibility is ruled out by estimates on matrices sampled independently from exact-rank shells.

There is a useful distinction within the latter estimate. If many ranks are far from pn , their profile already has few realizations. Indeed, for $a_i = r_i/n$ and $\sum_i a_i \leq tp$, the quadratic rank exponent $f(a) = 2a - a^2$ satisfies

$$\sum_{i=1}^t f(a_i) \leq tf(p) - \sum_{i=1}^t (a_i - p)^2.$$

Otherwise only a bounded number of generators fall below $p - \eta$; the others lie in nearby shells, where the translated-span estimate applies. Retaining half the quadratic deficit allows the remaining profile sum to factor into bounded discrete Gaussian sums. This replaces an overhead $(n+1)^t$ by a constant to the power t , and is what makes the estimate useful when the gap varies with n .

Finally, the center is encoded by offsets. An affinely independent t -tuple $X_i = c_i - Y$ requires only its $t-1$ independent differences to lie in the random code. Counting those tuples already accounts for all possible centers. Charging a further factor q^{n^2} for the center would lose part of the available exponential saving.

Example 1.5. The difference between average and maximum distance can already be seen at one center in $\mathbb{F}_2^{2 \times 2}$. The offsets 0 and I_2 have ranks 0 and 2. Their average rank is 1, although the second offset is outside the radius-1 ball. Thus a proof for average-radius lists must allow generators with different ranks; counting only one ball is not sufficient.

2 Rank profiles and random-subspace containment

Throughout the proof put $N = n^2$, $f(a) = 2a - a^2$, and

$$B(a) = \{X \in V_n : \text{rank}(X) \leq an\}, \quad \Gamma = \prod_{j=1}^{\infty} (1 - q^{-j})^{-1}.$$

The radius in $B(a)$ is automatically rounded down. The infinite product is finite because $\sum_j q^{-j} < \infty$.

Lemma 2.1. *If M_r is the number of rank- r matrices in V_n , then*

$$\Gamma^{-2} q^{2nr-r^2} \leq M_r \leq \Gamma q^{2nr-r^2} \quad (0 \leq r \leq n). \quad (3)$$

For an exact rank profile $r_1, \dots, r_t$ with $\sum_i r_i \leq pnt$, put $a_i = r_i/n$ and $\Delta = \sum_i (a_i - p)^2$. The number of tuples realizing the profile is at most

$$\Gamma^t q^{(tf(p) - \Delta)N}. \quad (4)$$

Also $|B(a)| \leq \Gamma(n+1)q^{f(a)N}$ for $0 \leq a \leq 1$.

Proof. Factor a rank- r matrix into a full-column-rank $n \times r$ matrix and a full-row-rank $r \times n$ matrix. Each matrix has exactly $|\text{GL}_r(\mathbb{F}_q)|$ such factorizations. Hence

$$M_r = q^{2nr-r^2} \frac{\prod_{i=0}^{r-1} (1 - q^{i-n})^2}{\prod_{i=0}^{r-1} (1 - q^{i-r})}.$$

Every finite product of distinct factors $1 - q^{-j}$ lies between Γ^{-1} and 1, proving (3), including the empty product at $r = 0$. For (4), multiply the upper bounds and use the exact identity

$$f(a) = f(p) + 2(1-p)(a-p) - (a-p)^2. \quad (5)$$

The linear terms sum to a nonpositive number. Finally, f is increasing on $[0, 1]$, so summing (3) over ranks at most an gives the ball bound. $\square$

Remark 2.2 (The capacity rate). The rank count also explains the rate in Question 1.2. If a nonempty code has no bad list of size $L \geq 2$, each radius- pn ball contains at most $L - 1$ codewords. Count pairs (c, Y) with $c \in \mathcal{C}$ and $\text{rank}(c - Y) \leq pn$ first by c and then by Y . Translation invariance gives

$$|\mathcal{C}| |B(p)| \leq (L - 1)q^N.$$

With $r = \lfloor pn \rfloor$, the lower bound in (3) therefore gives

$$\frac{\log_q |\mathcal{C}|}{N} \leq 1 - f(r/n) + \frac{\log_q(L - 1) + 2 \log_q \Gamma}{N}.$$

If $\log L = o(n^2)$, then $r/n \rightarrow p$ and the asymptotic upper bound is $(1 - p)^2$. This argument does not require linearity.

The Gaussian binomial coefficient, counting v -dimensional subspaces of $\mathbb{F}_q^u$, satisfies

$$\Gamma^{-1} q^{v(u-v)} \leq \begin{bmatrix} u \\ v \end{bmatrix}_q \leq \Gamma q^{v(u-v)}. \quad (6)$$

Indeed, counting ordered independent bases gives $\begin{bmatrix} u \\ v \end{bmatrix}_q = q^{v(u-v)} \prod_{i=0}^{v-1} (1 - q^{i-u}) / \prod_{i=0}^{v-1} (1 - q^{i-v})$, to which the same product bounds apply.

Lemma 2.3. *For a fixed s -dimensional subspace $U \leq \mathbb{F}_q^N$ and a uniformly random k -dimensional subspace $\mathcal{C}$,*

$$\Pr[U \subseteq \mathcal{C}] \leq \Gamma q^{-(N-k)s}.$$

Proof. The probability is zero if $s > k$. Otherwise invariance under invertible linear maps, applied to ordered independent s -tuples, gives

$$\Pr[U \subseteq \mathcal{C}] = \prod_{i=0}^{s-1} \frac{q^k - q^i}{q^N - q^i}.$$

After factoring out $q^{-(N-k)s}$, the numerator is at most 1 and the denominator at least Γ^{-1} . The formula also covers $s = 0$. $\square$

Lemma 2.4. *Set $G_q = 2 \sum_{j=0}^{\infty} q^{-j^2/2}$. Then $G_q \leq 2/(1 - q^{-1/2})$, and for every real x ,*

$$\sum_{r=0}^n q^{-(r-x)^2/2} \leq G_q.$$

Consequently the number of ordered t -tuples of matrices of average rank at most pn is at most

$$(\Gamma G_q)^t q^{tf(p)N}. \quad (7)$$

Proof. For every integer $j \geq 0$, the integers at distance in $[j, j + 1)$ from x lie in $[x + j, x + j + 1)$ or $(x - j - 1, x - j]$. Each interval contains at most one integer. Their total contribution is at most $2q^{-j^2/2}$. Summing proves the first bound, and $j^2 \geq j$ gives the geometric-series estimate on G_q . For a rank profile, (4) has deficit $\Delta N = \sum_i (r_i - pn)^2$. Sum over the allowed profiles, enlarge to all profiles in $\{0, \dots, n\}^t$, and use $q^{-u^2} \leq q^{-u^2/2}$. The resulting product of t one-dimensional sums proves (7). $\square$

3 Weighted affine extraction

Lemma 3.1. *Let S be a nonempty finite set of nonzero vectors over a field, with $|S| = s$. Give its elements nonnegative weights of average at most μ . For every positive integer D , either*

1. *S contains D linearly independent vectors of average weight at most μ ; or*
2. *for some $1 \leq d < D$, it contains independent vectors $x_1, \dots, x_d$ of average weight at most μ with $|S \cap \text{span}(x_1, \dots, x_d)| > sd/D$.*

Proof. Greedily choose a minimum-weight element outside the previous span, obtaining a basis $x_1, \dots, x_r$ of $\text{span}(S)$. Put $b_j = w(x_j)$, $W_j = \text{span}(x_1, \dots, x_j)$, and $m_j = |S \cap W_j|$. The sequence b_j is nondecreasing, and every point outside W_{j-1} has weight at least b_j . If the first D elements exist and have average weight at most μ , we have the first conclusion. If all greedy prefixes have average at most μ but this conclusion fails, then $r < D$, and $d = r$ gives $m_r = s > sr/D$.

In the remaining case let h be the first index with $\sum_{i=1}^h b_i > h\mu$. Then $2 \leq h \leq D$: the first weight is at most the average over S , and failure of the first conclusion forces such an index by D if the basis has that many elements. All shorter prefixes have average at most μ , and $b_h > \mu$. Suppose $m_j \leq sj/D$ for every $j < h$. Points entering the span at step i have weight at least b_i , and those outside W_{h-1} have weight at least b_h . Thus

$$\begin{aligned} \sum_{x \in S} w(x) &\geq sb_h - \sum_{j=1}^{h-1} m_j(b_{j+1} - b_j) \\ &\geq sb_h - \frac{s}{D} \sum_{j=1}^{h-1} j(b_{j+1} - b_j) \\ &= \frac{s}{D} \left(\sum_{i=1}^h b_i + (D-h)b_h \right) > s\mu. \end{aligned}$$

The last inequality uses $D-h \geq 0$ and $b_h > \mu$. This contradicts the assumed average. Some $d < h$ therefore satisfies $m_d > sd/D$, and its prefix gives the second conclusion. $\square$

Corollary 3.2. *If $S \subseteq V_n$ is nonempty and has average rank at most pn , then either it contains D affinely independent matrices of average rank at most pn , or, for some $1 \leq t < D$, it contains affinely independent matrices $Z_1, \dots, Z_t$ of that average rank such that*

$$|S \cap \text{aff}(Z_1, \dots, Z_t)| > \frac{|S|t}{D}.$$

Proof. Apply Lemma 3.1 to the nonzero lifted vectors $(X, 1)$, weighted by $\text{rank}(X)$. Independence of these lifts is affine independence of the matrices. A point $(X, 1)$ is in their linear span exactly when its coefficients sum to 1, which is precisely the affine span condition. $\square$

4 Separated balls and nearby rank shells

Lemma 4.1. *Let $0 \leq r < n$ be an integer. Suppose $\text{rank}(A) \geq an$ and $n-r \geq bn$ for positive a, b , and put $\kappa = \min\{ab/2, b^2/4\}$. Then*

$$|\{X : \text{rank}(X) \leq r, \text{rank}(X-A) \leq r\}| \leq (n+1)\Gamma^5 q^{2nr-r^2-\kappa N}. \quad (8)$$

Proof. Write $a_A = \text{rank}(A)$. For each counted X , extend the column spaces of X and $X-A$ to r -dimensional subspaces $U, V \leq \mathbb{F}_q^n$. Overcounting is harmless. Put $j = r - \dim(U \cap V)$, so $0 \leq j \leq \min(r, n-r)$ and $\dim(U+V) = r+j$. For fixed U, V , a necessary and sufficient condition

for the column constraints to have a solution is $\text{col}(A) \subseteq U + V$. When it holds, each column of X lies in an affine coset of $U \cap V$, giving exactly $q^{n(r-j)}$ possible matrices.

For a fixed j , the number of ordered pairs (U, V) is

$$\begin{bmatrix} n \\ r \end{bmatrix}_q \begin{bmatrix} r \\ j \end{bmatrix}_q \begin{bmatrix} n-r \\ j \end{bmatrix}_q q^{j^2} \leq \Gamma^3 q^{r(n-r)+jn-j^2}.$$

To obtain the formula, choose U , its $(r-j)$ -dimensional intersection with V , and the j -dimensional image of V in $\mathbb{F}_q^n/U$. The remaining choice is the graph of a map from that image into the j -dimensional quotient of U , giving q^{j^2} possibilities. The inequality follows from (6). By the transitive action of $\text{GL}_n(\mathbb{F}_q)$, the sums $U + V$ are equally distributed over all $(r+j)$ -dimensional subspaces. If $a_A > r+j$, no pair is admissible. Otherwise the fraction containing $\text{col}(A)$ is

$$\begin{bmatrix} n-a_A \\ r+j-a_A \end{bmatrix}_q \begin{bmatrix} n \\ r+j \end{bmatrix}_q^{-1} \leq \Gamma^2 q^{-a_A(n-r-j)}.$$

Multiplication by $q^{n(r-j)}$ gives, for this j , the bound

$$\Gamma^5 q^{2nr-r^2-j^2-a_A(n-r-j)}.$$

Write $h = n - r$. If $j \geq h/2$, then $j^2 \geq b^2 N/4$. If $j < h/2$, then $a_A(h-j) \geq abN/2$. In both cases the exponent loses at least κN . There are at most $n+1$ values of j , proving (8). $\square$

Lemma 4.2. *There are constants $\eta, c > 0$, depending only on p , with $0 < p - \eta < p + \eta < 1$, such that the following holds. Put $\rho = p + \eta$ and $C_0 = 2\Gamma^7$. Let X_1, X_2 be independent and uniform on exact-rank shells of ranks r_1, r_2 , where $p - \eta \leq r_i/n \leq p + \eta$. For every fixed $Z \in V_n$ and nonzero $a_1, a_2 \in \mathbb{F}_q$,*

$$\Pr[a_1 X_1 + a_2 X_2 + Z \in B(\rho)] \leq C_0(n+1)q^{-cN}. \quad (9)$$

Proof. Let $\beta = 1 - p$, $c_2 = \min\{p\beta/8, \beta^2/16\}$, and choose

$$0 < \eta \leq \min\{p/4, \beta/4, c_2/8\}.$$

Set $c_1 = f(3p/4) - f(p/2) > 0$ and $c = \min\{c_1, c_2/2\}$. Nonzero scalar multiplication preserves every exact-rank shell, so it suffices to take $a_1 = a_2 = 1$. Write $A = X_1 + Z$. The number of X_1 with $\text{rank}(A) < pn/2$ is at most $|B(p/2)|$. Using Lemma 2.1 and $p - \eta \geq 3p/4$ gives

$$\Pr[\text{rank}(A) < pn/2] \leq \Gamma^3(n+1)q^{[f(p/2)-f(p-\eta)]N} \leq \Gamma^3(n+1)q^{-c_1 N}.$$

For all other X_1 , put $r = \lfloor \rho n \rfloor$. Then $n - r \geq \beta n/2$. Lemma 4.1 with $a = p/2$ and $b = \beta/2$ yields

$$|B(\rho) \cap (-A + B(\rho))| \leq (n+1)\Gamma^5 q^{[f(\rho)-c_2]N}.$$

The rank- r_2 shell is contained in $B(\rho)$ and has at least $\Gamma^{-2} q^{f(p-\eta)N}$ elements. Conditional on this X_1 , the probability that $X_2 + A$ is in $B(\rho)$ is therefore at most

$$(n+1)\Gamma^7 q^{[f(p+\eta)-f(p-\eta)-c_2]N} \leq (n+1)\Gamma^7 q^{-c_2 N/2},$$

since $f(p+\eta) - f(p-\eta) = 4(1-p)\eta \leq c_2/2$. Add the exceptional probability and use $\Gamma \geq 1$. All bounds are uniform in Z . $\square$

5 Coefficient sets and translated spans

For $v \in \mathbb{F}_q^g$, write $\text{supp}(v) = \{j : v_j \neq 0\}$. Call $w_1, \dots, w_m$ *two-increasing* if each w_i has at least two support coordinates outside the union of the earlier supports. A coordinate rectangle in $T \subseteq \mathbb{F}_q^g$ means that its projection onto two coordinates contains the Cartesian product of two two-element sets. Witnesses to the four corners need not agree elsewhere. Put $A_q = \binom{g}{2}$.

Lemma 5.1. *A set $T \subseteq \mathbb{F}_q^g$ with no coordinate rectangle satisfies $|T| \leq 1 + A_q g$.*

Proof. Induct on g , the case $g = 0$ being immediate. Project onto the first $g - 1$ coordinates, obtaining P . For each $x \in P$, let F_x be its nonempty fiber of last-coordinate values. Then

$$|T| = |P| + \sum_{x \in P} (|F_x| - 1).$$

A fixed pair of distinct symbols can belong to at most one F_x . If it belonged to F_x and F_y for $x \neq y$, a coordinate where x, y differ, together with the last coordinate, would give a rectangle. Since $s - 1 \leq \binom{s}{2}$ for $s \geq 1$, the sum is at most A_q . The projection P has no rectangle either, so induction gives $|P| \leq 1 + A_q(g - 1)$, as required. $\square$

Lemma 5.2. *If $m \geq 1$ and $|T| > q^{2(m-1)}(1 + A_q g)$, then there are $v_0, v_1, \dots, v_m \in T$ such that $v_1 - v_0, \dots, v_m - v_0$ is two-increasing. For $g \geq 1$, the sufficient condition $|T| > Hg$ may be used, where $H = q^{2(m-1)}(1 + A_q)$.*

Proof. Induct on m , simultaneously for every g . For $m = 1$, Lemma 5.1 supplies a rectangle, and two opposite corners give points differing in two coordinates. For $m \geq 2$, choose a rectangle and partition T by the values of its two coordinates. A largest fiber F has size

$$|F| \geq |T|/q^2 > q^{2(m-2)}(1 + A_q g) \geq q^{2(m-2)}(1 + A_q(g - 2)).$$

Delete the two fixed coordinates and apply induction inside F , obtaining $v_0, \dots, v_{m-1}$. Let the fixed coordinate values be a, b . In each of the two sides of the original rectangle, choose a value different from the corresponding a or b . Its corner has a witness $v_m \in T$. The difference $v_m - v_0$ is nonzero in both fixed coordinates, whereas all earlier differences vanish there. Appending it gives the chain. Finally, $(1 + A_q)g \geq 1 + A_q g$ for $g \geq 1$. $\square$

Fix the constants of Lemma 4.2, and henceforth put

$$m = \lceil 4/c \rceil, \quad H = q^{2(m-1)}(1 + A_q), \quad b = \lceil 6/\eta^2 \rceil, \quad K = q^b(H + 1). \quad (10)$$

All depend only on p, q .

Theorem 5.3. *Let $X_1, \dots, X_g$ be independent, each uniform on an exact-rank shell whose normalized rank belongs to $[p - \eta, p + \eta]$, where $g \geq 1$. Then*

$$\Pr [\exists Z : |(Z + \text{span}(X_1, \dots, X_g)) \cap B(\rho)| > Hg] \leq C_0^m (n + 1)^m q^{-3N + mg}. \quad (11)$$

Proof. For a realization, define $T_X(v) = \sum_j v_j X_j$. If a translate $Z + \text{span}(X_1, \dots, X_g)$ contains more than Hg distinct points of $B(\rho)$, then

$$S = \{v \in \mathbb{F}_q^g : Z + T_X(v) \in B(\rho)\}$$

has more than Hg elements: each such point has a preimage. Lemma 5.2 supplies $v_0, \dots, v_m \in S$, with $w_i = v_i - v_0$ two-increasing. Set $A = Z + T_X(v_0)$. Then $A + T_X(w_i) \in B(\rho)$ for every i .

Fix now a *deterministic* matrix A and a deterministic two-increasing sequence $w_1, \dots, w_m$. Let E_i denote the corresponding low-rank event and put $\delta = C_0(n + 1)q^{-cN}$. Choose two coordinates of w_m that do not occur in any earlier support. Conditional on all generators except these two,

the earlier events are determined, while E_m involves two independent shell matrices with nonzero coefficients and a fixed shift. Lemma 4.2 gives

$$\Pr(E_1 \cap \dots \cap E_m) \leq \delta \Pr(E_1 \cap \dots \cap E_{m-1}).$$

Repeat on the shortened chain to obtain a bound of δ^m . There are q^N possible deterministic A and at most q^{mg} coefficient sequences. A union bound yields

$$q^{N+mg}\delta^m = C_0^m(n+1)^m q^{-(cm-1)N+mg} \leq C_0^m(n+1)^m q^{-3N+mg}.$$

This union covers all possible shifts. In particular, the proof never conditions on the random anchor $Z + T_X(v_0)$; doing so would not preserve the independence used in the two-shell estimate. $\square$

6 Weighted spans and Gaussian summation

Proposition 6.1. *For every positive integer t , the number of ordered tuples $(X_1, \dots, X_t)$ satisfying*

$$X_i \in B(\rho), \quad \sum_i \text{rank}(X_i) \leq pnt, \quad |\text{span}(X_1, \dots, X_t) \cap B(\rho)| > Kt \quad (12)$$

is at most

$$(\Gamma G_q)^t C_0^m(n+1)^m q^{(tf(p)-3)N+mt}. \quad (13)$$

No linear or affine independence is required, and t may depend on n .

Proof. Fix an allowed rank profile, put $a_i = r_i/n$, and let $\Delta = \sum_i (a_i - p)^2$. Write $P_{n,t} = C_0^m(n+1)^m q^{mt} \geq 1$. We will bound the number of rich tuples with this profile by

$$\Gamma^t P_{n,t} q^{(tf(p)-3)N} q^{-\Delta N/2}. \quad (14)$$

If $\Delta \geq 6$, the bound for all tuples in (4) suffices, since $\Delta \geq 3 + \Delta/2$.

Suppose $\Delta < 6$. Call i exceptional when $a_i < p - \eta$. Each exceptional index contributes more than η^2 to Δ , so there are at most b of them. All other ranks lie in $[p - \eta, p + \eta]$, because $X_i \in B(\rho)$. Sample the matrices independently and uniformly from their specified shells, and condition on the exceptional matrices. Their span U has at most q^b elements. Let W be the span of the g remaining matrices. If $g = 0$, the full span has size at most $q^b < Kt$, and cannot be rich. If $g \geq 1$ and every translate of W contains at most Hg low-rank points, then

$$|(U + W) \cap B(\rho)| \leq \sum_{Z \in U} |(Z + W) \cap B(\rho)| \leq q^b Hg \leq q^b Ht < Kt.$$

Thus richness implies the event in Theorem 5.3 for the good generators. They remain independent and uniform on their shells after conditioning. The conditional probability of richness is at most $P_{n,t} q^{-3N}$, uniformly in the exceptional values. Multiply by the profile count $\Gamma^t q^{tf(p)N - \Delta N}$ from (4). The result is bounded by (14), since $\Delta \geq 0$. This proves the same profile bound in both cases.

Finally sum (14) over allowed profiles and enlarge the residual nonnegative sum to all $r_i \in \{0, \dots, n\}$. Because $\Delta N = \sum_i (r_i - pn)^2$, this sum factors into t one-dimensional Gaussian sums, each at most G_q by Lemma 2.4. Substituting $P_{n,t}$ gives (13). $\square$

7 Counting bad affine lists

Theorem 7.1. *Use the constants above and set*

$$A_{p,q} = \Gamma G_q q^m, \quad D = \lceil 3/\varepsilon \rceil, \quad L = \left\lceil \frac{\rho K}{\eta} D \right\rceil.$$

For the random code of Theorem 1.3,

$$\Pr[\mathcal{C} \text{ has a bad list of size } L] \leq D \Gamma C_0^m(n+1)^m A_{p,q}^D q^{-2n^2}. \quad (15)$$

Moreover $L \leq (4\rho K/\eta + 1)/\varepsilon$.

Proof. Put $\alpha = f(p) + \varepsilon$. Then $\alpha < 1$ and $N - k \geq \alpha N$. If $X_1, \dots, X_t$ are affinely independent, their differences $X_i - X_1$ for $i \geq 2$ span a $(t - 1)$ -dimensional space. Lemma 2.3 bounds the probability that all these differences lie in $\mathcal{C}$ by $\Gamma q^{-\alpha(t-1)N}$. There is no additional factor for the center: offsets $X_i = c_i - Y$ from a common center have all their differences in $\mathcal{C}$; conversely, if those differences lie in $\mathcal{C}$, then the codewords $c'_i = X_i - X_1$ and the center $Y' = -X_1$ realize the same offsets. Distinct offsets produce distinct codewords.

First consider affinely independent D -tuples of average rank at most pn whose differences lie in $\mathcal{C}$. By (7) and the containment bound, their existence has probability at most

$$\Gamma(\Gamma G_q)^D q^{[Df(p) - \alpha(D-1)]N} = \Gamma(\Gamma G_q)^D q^{(\alpha - \varepsilon D)N} \leq \Gamma(\Gamma G_q)^D q^{-2N}. \quad (16)$$

The last inequality uses $\varepsilon D \geq 3$ and $\alpha < 1$.

For each $1 \leq t < D$, consider affinely independent tuples satisfying (12) whose differences lie in $\mathcal{C}$. Proposition 6.1 and the same containment bound give probability at most

$$\begin{aligned} \Gamma(\Gamma G_q)^t C_0^m (n+1)^m q^{mt + [tf(p) - 3 - \alpha(t-1)]N} \\ \leq \Gamma C_0^m (n+1)^m A_{p,q}^t q^{-2N}, \end{aligned} \quad (17)$$

because $tf(p) - 3 - \alpha(t-1) = \alpha - 3 - \varepsilon t \leq -2$. If affine independence is impossible for dimension reasons, the corresponding event is empty, and these estimates remain valid.

It remains to prove that every bad list produces one of the witnesses just counted. Let $X_i = c_i - Y$ be L distinct offsets with total rank at most pnL . Keep precisely those with rank at most pn , forming S of size s . The removed offsets have rank greater than pn , so

$$pn(L - s) \leq pnL, \quad s \geq \frac{\eta}{\rho} L \geq KD.$$

Their removal leaves average rank at most pn : subtracting numbers greater than pn from a total at most pnL leaves a total at most pns . Apply Corollary 3.2 to S . Its first alternative gives the witness in (16). Otherwise there are $t < D$ affinely independent $Z_1, \dots, Z_t \in S$ of average rank at most pn with

$$|S \cap \text{aff}(Z_1, \dots, Z_t)| > st/D \geq Kt.$$

Since $S \subseteq B(\rho)$ and the affine span is contained in the linear span, this tuple satisfies (12). Its differences lie in $\mathcal{C}$, because its members come from the same original list. It is therefore a witness in (17).

There are D event types: one large type and $D - 1$ small types. As $A_{p,q} > 1$ and $C_0^m (n+1)^m \geq 1$, each bound is at most $\Gamma C_0^m (n+1)^m A_{p,q}^D q^{-2N}$. Summing proves (15). Finally, $0 < \varepsilon < 1$ gives $D \leq 4/\varepsilon$, and hence

$$L \leq \frac{\rho K}{\eta} D + 1 \leq \frac{4\rho K/\eta + 1}{\varepsilon}.$$

□

Proof of Theorem 1.3. Take the list size in Theorem 7.1. The base- q logarithm of its prefactor is

$$\log_q D + \log_q \Gamma + m \log_q C_0 + m \log_q (n+1) + D \log_q A_{p,q}.$$

All constants depend only on p, q , and $D \leq 4/\varepsilon$. Since $q \geq 2$, we have $\log_q D \leq D$. Using also $1/\varepsilon > 1$, one may therefore take

$$B_1 = 4 + \log_q \Gamma + m \log_q C_0 + 4 \log_q A_{p,q}, \quad B_2 = m, \quad B_3 = 4\rho K/\eta + 1.$$

These constants give (2). □

8 Concluding remarks

The constants obtained from translated coefficient chains and exceptional rank profiles are large. Improving their dependence on p and q is a separate question. The proof treats p, q as fixed and gives no uniform control as p approaches 0 or 1 or as the field grows. Its finite estimate does allow the capacity gap to vary with n ; when εn^2 stays bounded, that estimate need not tend to zero. The rectangular model of [2] would require a version of the shell and profile arguments with its two dimension parameters.

References

- [1] V. Guruswami, S. Li and M. Singhal, *Average-Radius List-Decodability of Random Linear Codes*, arXiv:2608.22663v1, 2026. <https://arxiv.org/abs/2608.22663v1>.
- [2] V. Guruswami and N. Resch, *On the list-decodability of random linear rank-metric codes*, ISIT 2018, 1505–1509. <https://doi.org/10.1109/ISIT.2018.8437698>. Theorem and lemma numbering follows arXiv:1710.11516v1 (2017), <https://arxiv.org/abs/1710.11516v1>.
- [3] V. Guruswami, R. Li, J. Mosheiff, N. Resch, S. Silas and M. Wootters, *Bounds for list-decoding and list-recovery of random linear codes*, IEEE Transactions on Information Theory **68** (2022), 923–939. <https://doi.org/10.1109/TIT.2021.3127126>. The cited Remark 5.6 is in the authors’ April 2020 preprint, <https://www.cs.bgu.ac.il/~mosheiff/publications/RLCBounds.pdf>.
- [4] P. Delsarte, *Bilinear forms over a finite field, with applications to coding theory*, Journal of Combinatorial Theory, Series A **25** (1978), 226–241. [https://doi.org/10.1016/0097-3165\(78\)90015-8](https://doi.org/10.1016/0097-3165(78)90015-8).
- [5] E. M. Gabidulin, *Theory of codes with maximum rank distance*, Problems of Information Transmission **21** (1985), no. 1, 1–12.
- [6] R. M. Roth, *Maximum-rank array codes and their application to crisscross error correction*, IEEE Trans. Inform. Theory **37** (1991), no. 2, 328–336. <https://doi.org/10.1109/18.75248>.
- [7] P. Elias, *Error-correcting codes for list decoding*, IEEE Trans. Inform. Theory **37** (1991), no. 1, 5–12. <https://doi.org/10.1109/18.61123>.
- [8] A. Wachter-Zeh, *Bounds on list decoding of rank-metric codes*, IEEE Trans. Inform. Theory **59** (2013), no. 11, 7268–7277. <https://doi.org/10.1109/TIT.2013.2274653>.
- [9] V. Guruswami, J. Håstad and S. Kopparty, *On the list-decodability of random linear codes*, IEEE Transactions on Information Theory **57** (2011), 718–725. <https://doi.org/10.1109/TIT.2010.2095170>. Preprint: <https://arxiv.org/abs/1001.1386>.
- [10] Y. Ding, *On list-decodability of random rank metric codes and subspace codes*, IEEE Transactions on Information Theory **61** (2015), 51–59. <https://doi.org/10.1109/TIT.2014.2371915>. Preprint: <https://arxiv.org/abs/1401.2693>.
- [11] R. Li and M. Wootters, *Improved list-decodability of random linear binary codes*, IEEE Transactions on Information Theory **67** (2021), 1522–1536. <https://doi.org/10.1109/TIT.2020.3041650>. Preprint: <https://arxiv.org/abs/1801.07839>.
- [12] A. Rudra and M. Wootters, *Average-radius list-recoverability of random linear codes*, SODA 2018, 644–662. <https://doi.org/10.1137/1.9781611975031.42>. Preprint: <https://arxiv.org/abs/1704.02420>.

- [13] Z. Guo, C. Xing, C. Yuan and Z. Zhang, *Gabidulin codes achieve list decoding capacity with an order-optimal column-to-row ratio*, APPROX/RANDOM 2025, LIPIcs **353**, 43:1–43:20. <https://doi.org/10.4230/LIPIcs.APPROX/RANDOM.2025.43>.
- [14] Y. Liu, A. Baumeister and A. Wachter-Zeh, *On the list-decodability of random (linear) sum-rank metric codes*, arXiv:2503.10234v1, 2025. <https://arxiv.org/abs/2503.10234v1>.
- [15] S. Li, C. Xing, C. Yuan and R. Zhu, *Proximity gaps for Gabidulin codes and applications*, arXiv:2609.09838v1, September 9, 2026. <https://arxiv.org/abs/2609.09838v1>.
- [16] R. Goyal and V. Guruswami, *Improved analysis of list-decodability of random linear codes: It's all about counting constraints*, Electronic Colloquium on Computational Complexity, Report TR26-074, May 12, 2026. <https://eccc.weizmann.ac.il/report/2026/074/>.