Independent Squares and Average-Radius Obstructions for Linear Codes

Scott Schubr

October 19, 2026
Revision 2

Abstract

We study binary-input average-radius list recovery using a configuration formed by independent affine squares sharing their origin. If an $\mathbb{F}_q$-linear code in $(\mathbb{F}_q^s)^n$ has dimension k, then, for $t \geq 1$, $c \geq 0$, $d = 2t + c$, and $M = 1 + 3t + c$, it contains M distinct words and input lists of size at most two with mean relative error at most $1 - 2/M - \min\{2n, n + 2tu, (4t + c)u\}/(Mn)$, where $k \geq d$ and $u = \lfloor (k - d)/s \rfloor$. Allocating unequal shortening groups can strengthen the equal-group bound; the asymptotic improvement occurs when coloop directions are present. For sixteen words this gives the asymptotic converse $7/8 - (5/4)\min\{R, 1/10\}$, uniformly in the folding parameter and field. We also prove that the associated all-field two-fiber intersection-dimension certificate has exact cost $4t + c$. Finally, standard full-length folded Reed–Solomon codes realize the sixteen-word profile with mean error exactly $3/4$, including families at the fixed rate $5/48$ and arbitrarily large folding and length. These results rule out the average-radius extrapolation of the Boolean-cube curve for output cap fifteen, and analogous extrapolations for larger power-of-two lists. They do not refute the ordinary list-recovery conjecture of Chen and Zhang. The increment is the sparse independent-square profile and its sharp certificate calculation; shortening, independent kernel representatives, and local-profile methods have established predecessors.

1 Introduction

List recovery asks how many codewords can agree with prescribed small sets of alphabet symbols at most coordinates. Its average-radius version bounds the mean error of a whole set of codewords. The latter condition is stronger: a word that agrees everywhere can compensate, in an average, for other words that each have substantially larger error. This distinction matters even for input lists of size two and a fixed small output cap.

For $M = 2^a$, the Boolean-grid construction underlying the converse of Chen and Zhang [4, Theorem 3.1 and Corollary 3.2] gives the curve

$$\rho_{\mathrm{cube},a}(R) = 1 - \frac{2}{M} - \frac{B_a}{M}R, \qquad B_a = \frac{a(M-2)}{a-1}. \tag{1}$$

Their Conjecture 3.6 proposes an ordinary list-recovery guarantee for folded Reed–Solomon codes below the corresponding curve, with positive slack and sufficiently large folding and length. It does not propose an average-radius guarantee. We show that the latter, stronger extrapolation fails. In particular, at $M = 16$, the cube curve is

$$\rho_{\mathrm{cube},4}(R) = \frac{7}{8} - \frac{7}{6}R, \tag{2}$$

whereas every linear code of a fixed positive rate has the sixteen-word average-radius obstruction

$$\frac{7}{8} - \frac{5}{4}\min\left\{R, \frac{1}{10}\right\} + o(1). \tag{3}$$

The new obstruction is smaller than (2) whenever $0 < R < 3/28$.

1.1 The independent-square profile

Take t independent pairs a_h, b_h and c further independent vectors z_j. The configuration

$$S = \{0\} \ \cup \ \bigcup_{h=1}^{t} \{a_h, b_h, a_h + b_h\} \ \cup \ \{z_1, \dots, z_c\} \tag{4}$$

has $M = 1+3t+c$ points in dimension $d = 2t+c$. The term "square" means the affine parallelogram $\{0, a_h, b_h, a_h + b_h\}$; it imposes no metric structure and is valid in characteristic two. The extra independent single points will also be called coloop points.

Quotienting by a square-axis line allows two fibers to capture four points of S. Quotienting by a coloop-axis line allows them to capture three. These local counts lead to two related results. First, shortening an arbitrary linear code on disjoint groups of coordinates produces independent codewords with these prescribed vanishing properties. Assigning coordinates to square directions before the coloop directions gives the finite converse in Theorem 2.1; equal group sizes are unnecessary. The resulting asymptotic excess agreement is $\min\{2, 1 + 2tR, (4t + c)R\}$. This improves the coloop cases while leaving the sixteen-word curve (3) unchanged. Second, for every subspace K of the configuration's direction space, the sum $b_2(S/K)$ of its two largest fiber populations satisfies

$$b_2(S/K) \le 2 + \dim K + \sum_{h=1}^{t} \dim\bigl(K \cap \operatorname{span}(a_h, b_h)\bigr). \tag{5}$$

Among all nonnegative, K-independent certificates of this intersection-dimension form, the exact minimum cost is $C = 4t + c$ (Theorem 3.2). Five independent squares therefore give sixteen points of exact cost 20, exceeding the cube coefficient $56/3$.

The universal shortening argument only guarantees lower bounds on coordinate agreement. A separate product-polynomial construction realizes the profile with exact one-dimensional block kernels in standard full-length folded Reed–Solomon codes. For five squares it gives exactly four agreeing words at every block, hence mean error $3/4$. At rate $R_* = 5/48$, choose $\rho_* = 751/1000$. Then

$$\frac{3}{4} < \rho_* < \frac{217}{288} - \frac{1}{1000} = \rho_{\mathrm{cube},4}(R_*) - \frac{1}{1000}. \tag{6}$$

The construction exists at this exact rate beyond any prescribed folding and length thresholds. Thus the gap persists with fixed parameters and fixed positive slack.

1.2 Predecessors and scope

Chen and Zhang's proof uses a full Cartesian grid and prescribed vanishing groups [4, Section 3.1]. Li and Shagrithaya establish general linear-code list-recovery lower bounds using related grid and support constructions [7, Section 5]. Chen's note [2, Theorem 2.2] explicitly generalizes the grid converse to folded linear codes by shortening on disjoint coordinate groups and choosing independent representatives from the resulting kernels. The proof of Theorem 2.1 uses these established

ideas with a different, sparse message set. Its simple sufficient dimension bound permits a direct greedy choice of representatives.

Average-radius recovery is the average-weight recovery notion of Levi, Mosheiff, and Shagrithaya [6, Definition 2.1]. Their local coordinate-wise linear framework and the full digit-grid converse in Section 5.2 are prior context. Brakensiek, Chen, Dhar, and Zhang use discrete Brascamp–Lieb inequalities and subspace designs to obtain sharp zero-error results and larger-list positive-error bounds; their Remark 4.8 explains the average-radius extension of the latter bounds [1]. Goyal and Guruswami's variational formulation already expresses the optimal large-alphabet random-linear average-radius list-size problem through local profiles [5, Definition 21 and Remark 4]. Our result evaluates a concrete independent-square profile and gives an exact certificate for it; it does not introduce this framework.

The present note follows the author's eight-point certificate paper [8]. That paper supplies selected-fiber rank arguments and an attached-square dual witness, as well as restricted sixteen-point bounds. The union of five independent squares has affine dimension ten and lies outside those restricted hypotheses. The new cost 20 disproves the unrestricted sixteen-point target $56/3$; it leaves the proved eight-point and restricted sixteen-point results unchanged.

The comparison with the literature is deliberately limited to these statements and methods.[1] In particular, neither the certificate theorem nor the explicit FRS example proves that the same codes satisfy ordinary recovery at the parameters where average-radius recovery fails.

2 A finite converse for every linear code

Let q be a prime power, and let n, s be positive integers. We use $\mathbb{F}_q^s$ as a vector alphabet. An $\mathbb{F}_q$-linear code $\mathcal{C} \le (\mathbb{F}_q^s)^n$ of $\mathbb{F}_q$-dimension k has rate

$$R_n = \frac{k}{sn}. \tag{7}$$

Linearity over an extension field of order q^s is not required. For a table $\mathcal{A} = (A_1, \dots, A_n)$ of nonempty symbol sets, put

$$\operatorname{dist}(w, \mathcal{A}) = \frac{1}{n}\big|\{j : w_j \notin A_j\}\big|.$$

For a finite nonempty set $\mathcal{S}$ of codewords, its mean relative error against $\mathcal{A}$ is

$$D(\mathcal{S}, \mathcal{A}) = \frac{1}{|\mathcal{S}|}\sum_{w\in\mathcal{S}} \operatorname{dist}(w, \mathcal{A}) = 1 - \frac{1}{|\mathcal{S}|n}\sum_{w\in\mathcal{S}}\sum_{j=1}^{n} \mathbf{1}[w_j \in A_j]. \tag{8}$$

We call $\mathcal{C}$ average-radius $(\rho, 2, M-1)$ list recoverable if every M distinct codewords and every table with $1 \le |A_j| \le 2$ have mean error strictly greater than ρ. Ordinary recovery requires instead that at most $M-1$ codewords individually have error at most ρ. All counterexamples below use a strict gap, and therefore also apply if the definition uses a strict error threshold for a bad list.

Throughout the note, the profile parameters are

$$t \ge 1, \quad c \ge 0, \qquad d = 2t + c, \quad M = 1 + 3t + c, \quad C = 4t + c, \tag{9}$$

where t, c are integers. The symbol k always denotes code dimension, not the number of squares.

[1] Chen, Cheraghchi, Si, and Zhang's manuscript [3] is reported by Brakensiek, Chen, Dhar, and Zhang [1, Section 1] as proving the binary zero-error case. The counterexamples here concern an average-radius extrapolation and do not refute the ordinary Chen–Zhang conjecture.

Theorem 2.1 (Finite independent-square converse)**.** *Let $\mathcal{C} \leq (\mathbb{F}_q^s)^n$ be $\mathbb{F}_q$-linear of dimension k, and use the parameters* (9)*. Suppose $k \geq d$, and set*

$$u = \left\lfloor \frac{k-d}{s} \right\rfloor, \qquad E = \min\{2n, n+2tu, Cu\}. \tag{10}$$

There are M distinct codewords $\mathcal{S} \subseteq \mathcal{C}$ and lists $A_j \subseteq \mathbb{F}_q^s$, with $1 \leq |A_j| \leq 2$, such that

$$\sum_{w \in \mathcal{S}} \sum_{j=1}^{n} \mathbf{1}[w_j \in A_j] \geq 2n + E. \tag{11}$$

Consequently

$$D(\mathcal{S}, \mathcal{A}) \leq 1 - \frac{2}{M} - \frac{E}{Mn}. \tag{12}$$

No condition $du \leq n$ is required.

Proof. Choose disjoint sets $G_0, \ldots, G_{d-1} \subseteq \{1, \ldots, n\}$, each of size at most u. Fill the first $2t$ sets to their capacity before assigning coordinates to the remaining c sets. If fewer than du coordinates are available, some groups are smaller or empty. The total square-group and coloop-group sizes are respectively

$$A = \min\{n, 2tu\}, \qquad B = \min\{n - A, cu\}. \tag{13}$$

These sizes are realized by assigning successive coordinates to the groups until either their capacity or the coordinate supply is exhausted.

Let N_i be the subspace of codewords vanishing on G_i. The coordinate projection on G_i has codomain dimension $s|G_i|$, so rank-nullity gives

$$\dim_{\mathbb{F}_q} N_i \geq k - s|G_i| \geq k - su \geq d.$$

Inductively choose $v_i \in N_i$ outside $\operatorname{span}(v_0, \ldots, v_{i-1})$. At stage $i < d$ the latter space has dimension $i < d \leq \dim N_i$, so this is possible. Thus $v_0, \ldots, v_{d-1}$ are independent, and v_i vanishes at every coordinate of G_i. No independence assumption on the spaces N_i is used.

Form the codeword set

$$\mathcal{S} = \{0\} \ \cup \ \bigcup_{h=0}^{t-1} \{v_{2h}, v_{2h+1}, v_{2h} + v_{2h+1}\} \ \cup \ \{v_{2t}, \ldots, v_{d-1}\}. \tag{14}$$

It has exactly M distinct elements. Indeed, in the independent v_i coordinates these are zero, distinct unit vectors, and distinct vectors supported on the specified pairs. This support argument also applies in characteristic two.

At a coordinate $j \in G_i$ with $i < 2t$, let i' be the other index in the pair $\{2h, 2h+1\}$ containing i, and set

$$A_j = \{0, (v_{i'})_j\}.$$

The four distinct words $0, v_i, v_{i'}, v_i + v_{i'}$ agree with this table, because their symbols are $0, 0, (v_{i'})_j, (v_{i'})_j$. If $(v_{i'})_j = 0$, the table is a singleton and the four agreements still hold. Additional symbol coincidences only increase the count.

At a coordinate $j \in G_i$ with $i \geq 2t$, set $A_j = \{0, (v_0)_j\}$. This covers the three distinct words $0, v_i, v_0$, including when their symbols all coincide. The hypothesis $t \geq 1$ ensures that v_0 exists and is different from v_i. At every remaining coordinate use $A_j = \{0, (v_0)_j\}$; this covers at least the two distinct words $0, v_0$.

By (13), the total agreement is at least

$$4A + 3B + 2(n - A - B) = 2n + 2A + B.$$

If $n \le 2tu$, then $2A + B = 2n$. If $2tu < n \le du$, then $2A + B = n + 2tu$. If $n \ge du$, then $2A + B = (4t + c)u = Cu$. At shared endpoints these expressions agree; the middle interval is empty when $c = 0$. In all cases $2A + B = \min\{2n, n + 2tu, Cu\} = E$. This proves (11), and (8) proves (12). □

Corollary 2.2 (Equal groups)**.** *If an integer $v \ge 0$ satisfies*

$$dv \le n, \qquad k - sv \ge d, \tag{15}$$

then there are M distinct codewords and binary input lists with mean relative error at most $1 - 2/M - Cv/(Mn)$. In particular, this holds for

$$v = \min\left\{ \left\lfloor \frac{n}{d} \right\rfloor, \left\lfloor \frac{k-d}{s} \right\rfloor \right\} \quad \text{when } k \ge d.$$

Proof. One can take all d groups to have size v in the proof above. There are then $2tv$ square-group, cv coloop-group, and $n - dv$ unassigned coordinates, giving agreement $4(2tv) + 3cv + 2(n - dv) = 2n + Cv$. □

Remark 2.3 (Lists of exactly two symbols). Singleton tables are necessary to allow arbitrary symbol collisions in the universal proof. If a definition requires exactly two symbols, pad each singleton by an unused symbol. Since $q^s \ge 2$, this is always possible and cannot decrease agreement.

2.1 Rate normalization and saturation

Define, for $x \ge 0$,

$$e(x) = \min\{2, 1 + 2tx, Cx\}. \tag{16}$$

For $c > 0$ this is the continuous piecewise-linear function

$$e(x) = \begin{cases} Cx, & 0 \le x \le 1/d, \\ 1 + 2tx, & 1/d \le x \le 1/(2t), \\ 2, & x \ge 1/(2t). \end{cases} \tag{17}$$

Indeed, the first two expressions meet at $x = 1/d$, and the second and third meet at $x = 1/(2t)$. For $c = 0$, these breakpoints coincide and $e(x) = \min\{2, 4tx\}$.

Corollary 2.4. *If $k \ge d$, the code in Theorem 2.1 admits a bad list with*

$$D(\mathcal{S}, \mathcal{A}) \le 1 - \frac{2}{M} - \frac{e(R_n)}{M} + \frac{C(d/s + 1)}{Mn}. \tag{18}$$

For fixed t, c, consider any sequence with $n \to \infty$, arbitrary $s = s_n \ge 1$, arbitrary finite fields, and $R_n \to R > 0$. At every radius strictly greater than

$$D_\infty(R) = 1 - \frac{2}{M} - \frac{\min\{2, 1 + 2tR, CR\}}{M}, \tag{19}$$

these codes eventually fail average-radius $(\rho, 2, M - 1)$ list recovery. The error term in (18) *is at most $C(d + 1)/(Mn)$, uniformly in s and q.*

Proof. The choice (10) satisfies

$$0 \le R_n - \frac{u}{n} \le \frac{d/s+1}{n}.$$

The function e is nondecreasing and C-Lipschitz: each of its three defining affine functions has slope between zero and C, since $2t \le C$, and their minimum has the same two properties. Also $E/n = e(u/n)$ exactly. Consequently

$$\frac{E}{n} = e(u/n) \ge e(R_n) - \frac{C(d/s+1)}{n}.$$

Substitute this in (12). For the sequence, $k = R_n sn \to \infty$, so $k \ge d$ eventually. The displayed error tends to zero and e is continuous. Every strict gap above (19) therefore eventually exceeds the exhibited mean error. □

The saturation in (16) is essential; the low-rate line does not extend indefinitely. Equal groups alone give the weaker excess $C\min\{R, 1/d\}$, which is at most $e(R)$. For $c > 0$, square-first allocation improves that excess at every $R > 1/d$, until its own saturation; for $c = 0$, the two asymptotic bounds agree. If $R = 0$, the limiting statement requires the separate assumption $k \ge d$ eventually. The parameters t, c are fixed here; growing profiles require tracking the remainder anew.

Proposition 2.5 (Strongest profile within this family)**.** *Fix $M \ge 4$. Among profiles with $t \ge 1$ and $c = M - 1 - 3t \ge 0$, the largest allowable $t = \lfloor (M-1)/3 \rfloor$ gives the strongest finite bound* (12) *whenever a smaller-t profile is admissible.*

Proof. With M fixed, $d = M - 1 - t$ decreases and $C = M - 1 + t$ increases with t. Thus $u = \lfloor (k-d)/s \rfloor$ is nondecreasing. Once a profile is admissible, $u \ge 0$, and each of $2n, n + 2tu, Cu$ is nondecreasing as t increases. Their minimum E is therefore nondecreasing, while the denominator Mn is fixed. Larger t may also be admissible when a smaller value is not, because its d is lower. □

This is optimality only within the stated square/coloop family, not over arbitrary affine configurations or possible coding converses. For two finite comparisons, at $M = 8, t = 2, c = 1, n = 100, s = 1, k = 30$, we have $u = 25$ and $E = 200$. The refined mean-error bound is $1/2$, whereas equal groups give $21/40$. At $M = 32, t = 10, c = 1, n = 2100, s = 1, k = 126$, we have $u = 105$ and $E = 4200$; the bound is $7/8$ rather than $589/672$, an improvement of $1/672$. These compare two bounds from the present argument at the specified dimensions.

Remark 2.6 (Exact projection ranks). For a particular code, the proof needs only $\dim N_i \ge d$. The sufficient cardinality condition can therefore be replaced by $\operatorname{rank}(\mathcal{C} \to (\mathbb{F}_q^s)^{G_i}) \le k - d$. Dependent projections may permit larger groups.

Corollary 2.7 (Sixteen words)**.** *For $t = 5, c = 0$, every positive-rate sequence as in Corollary 2.4 has the obstruction*

$$D_\infty(R) = \begin{cases} 7/8 - 5R/4, & 0 < R \le 1/10, \\ 3/4, & R \ge 1/10. \end{cases} \tag{20}$$

It is strictly below the cube curve (2) *for every $0 < R < 3/28$.*

Proof. Here $M = 16, d = 10, C = 20$. Below rate $1/10$, subtracting the two expressions gives the positive gap $R/12$. Above rate $1/10$, the gap is $1/8 - 7R/6$, which is positive precisely when $R < 3/28$. □

At the fixed rate 5/48, the universal converse already gives $D \le 3/4 + o(1)$ uniformly over all folding parameters. For example, if $n/240 \ge 10/s$, then $(k-10)/s = 5n/48 - 10/s \ge n/10$. The equal-group corollary therefore applies with $v = \lfloor n/10 \rfloor$, giving

$$D \le \frac{3}{4} + \frac{5}{4n}. \tag{21}$$

Thus every sufficiently long code of this exact rate fails at $\rho_* = 751/1000$. Section 4 supplies full-length FRS families at the same rate with error exactly 3/4.

3 A sharp all-field two-fiber certificate

This section concerns linear geometry over an arbitrary field F, which need not be finite. Let

$$V = U_1 \oplus \cdots \oplus U_t \oplus W, \qquad U_h = \operatorname{span}(a_h, b_h), \qquad W = \operatorname{span}(z_1, \ldots, z_c),$$

where the displayed $2t + c$ vectors are a basis. Define S by (4). It spans V, contains zero, and has cardinality M.

For $K \le V$, its affine K-fibers on S are the nonempty sets $S \cap (v + K)$. Let $b_2(S/K)$ be the sum of the two largest fiber sizes, with the absent second fiber assigned size zero.

Definition 3.1. A two-fiber certificate for S is a finite collection of nonnegative real weights y_X, indexed by linear subspaces $X \le V$, such that

$$b_2(S/K) \le 2 + \sum_X y_X \dim(K \cap X) \quad \text{for every } K \le V. \tag{22}$$

The weights are fixed independently of K. Its cost is $\operatorname{cost}(y) = \sum_X y_X \dim X$.

Theorem 3.2 (Exact independent-square cost)**.** *For every field F and every $t \ge 1, c \ge 0$, the minimum cost in Definition 3.1 is*

$$\min_y \operatorname{cost}(y) = 4t + c. \tag{23}$$

A rational minimizer places weight one on V and weight one on each U_h. Weights are added when named supports coincide.

3.1 The lower bound

For the axis line $A = Fa_h$, the two fibers $\{0, a_h\}$ and $\{b_h, a_h + b_h\}$ have size two, and all remaining fibers are singletons. Thus $b_2(S/A) = 4$. The same is true for $A = Fb_h$. For a coloop-axis line $A = Fz_j$, the only nonsingleton fiber is $\{0, z_j\}$. A second fiber exists because $t \ge 1$, so $b_2(S/A) = 3$.

Let $\mathcal{L}$ be the collection of these d coordinate-axis lines. For every $X \le V$,

$$\sum_{A \in \mathcal{L}} \dim(X \cap A) \le \dim X. \tag{24}$$

Indeed, each nonzero intersection contains the corresponding basis vector, and those vectors are independent. Apply (22) at each coordinate-axis line, subtract its intercept two, and sum. Using nonnegativity of the weights and (24) gives

$$\begin{aligned} 4t + c &\le \sum_X y_X \sum_{A \in \mathcal{L}} \dim(X \cap A) \\ &\le \sum_X y_X \dim X. \end{aligned}$$

This proves the lower bound directly. It is a finite weak-duality argument and uses neither a solver nor a finite-field assumption.

3.2 The matching upper bound

We prove (5). If K is proper, the spanning property of S ensures at least two fibers. Choose the two largest, denoted A, B, and put $r = |A| + |B|$. In the coefficient space F^S let

$$E = \left\{ \lambda : \operatorname{supp} \lambda \subseteq A \cup B,\ \sum_{v \in A} \lambda_v = \sum_{v \in B} \lambda_v = 0 \right\}.$$

The two equations have disjoint nonempty supports, so $\dim E = r - 2$ over every field. Let

$$H = \left\{ \lambda \in F^S : \sum_{v \in S} \lambda_v = 0 \right\}, \qquad T : H \longrightarrow V, \quad T(\lambda) = \sum_{v \in S} \lambda_v v,$$

and write $\Lambda = \ker T$. Because A, B are separate K-fibers and each coefficient sum is zero, $T(E) \subseteq K$. Rank-nullity therefore yields

$$r - 2 \leq \dim K + \dim(E \cap \Lambda). \tag{25}$$

For $1 \leq h \leq t$, define the square relation $\lambda^{(h)}$ by the four nonzero coefficients

$$\lambda^{(h)}_0 = 1, \quad \lambda^{(h)}_{a_h} = -1, \quad \lambda^{(h)}_{b_h} = -1, \quad \lambda^{(h)}_{a_h + b_h} = 1. \tag{26}$$

These lie in Λ and form a basis of it. To see completeness, the independent a_h, b_h, z_j coordinates in $T(\lambda) = 0$ give

$$\lambda_{a_h} = \lambda_{b_h} = -\lambda_{a_h + b_h}, \qquad \lambda_{z_j} = 0.$$

The equation defining H then fixes λ_0 to be $\sum_h \lambda_{a_h + b_h}$. Thus every dependency is the corresponding linear combination of (26). They are independent because the coordinate $a_h + b_h$ occurs in only its own relation.

If $K \cap U_h = 0$, the quotient map $V \to V/K$ is injective on U_h. Hence $a_h, b_h, a_h + b_h$ have three distinct quotient images. The two selected fibers cannot contain all three. Every vector in E consequently vanishes at at least one of those three coordinates. That coordinate occurs in only $\lambda^{(h)}$ among the dependency basis vectors, with coefficient 1 or -1. Any vector in $E \cap \Lambda$ therefore has zero $\lambda^{(h)}$ coefficient. It follows that

$$\dim(E \cap \Lambda) \leq \left| \{ h : K \cap U_h \neq 0 \} \right| \leq \sum_{h=1}^{t} \dim(K \cap U_h). \tag{27}$$

This argument includes characteristic two: $-1 = 1$ there remains nonzero, and the three quotient images remain distinct.

Combining (25) and (27) proves (5) for every proper K, including $K = 0$. At $K = V$, its left side is $M = 1 + 3t + c$, whereas its right side is $2 + d + 2t = 2 + 4t + c$, so the same inequality holds. The asserted weights are therefore feasible and have cost $d + 2t = 4t + c$. This completes the proof of Theorem 3.2.

Remark 3.3 (Degenerate support and excluded parameters). If $t = 1, c = 0$, then $U_1 = V$, and the prescription means a single weight two on V, of cost four. The condition $t \geq 1$ is substantive. For example, with $t = 0, c = 1$ the two-point set has $b_2 \leq 2$ for every kernel and minimum certificate cost zero, rather than one.

3.3 What the certificate comparison establishes

The first power-of-two instances are

M	t	c	d	C	B_a
4	1	0	2	4	4
8	2	1	5	9	9
16	5	0	10	20	56/3
32	10	1	21	41	75/2

Thus a proposed universal sixteen-point certificate of cost at most 56/3 is false in the precise model of Definition 3.1. Realizing a bad list for a code requires its relevant quotient kernels to occur in that code. Theorem 2.1 and the next section supply separate coding arguments.

4 Exact realization in full-length folded Reed–Solomon codes

Fix a primitive element $\gamma \in \mathbb{F}_q^\times$. For positive integers s, n, k with $sn = q-1$ and $1 \le k \le sn$, define the full-length folded Reed–Solomon encoder by

$$E_r(f) = \big(f(\gamma^{sr}), f(\gamma^{sr+1}), \dots, f(\gamma^{sr+s-1})\big), \qquad 0 \le r < n, \tag{28}$$

for $f \in \mathbb{F}_q[X]$ of degree $< k$. These are consecutive disjoint blocks of all nonzero field elements. The encoder is injective by the root bound, since a nonzero polynomial of degree $< k \le sn$ cannot vanish at all sn distinct evaluation points. Its rate is $k/(sn)$.

Theorem 4.1 (Full-length realization)**.** *Use* (9)*. Suppose $s \ge d-1$ and $ds \mid (q-1)$, and put*

$$m = \frac{q-1}{ds}, \qquad n = dm, \qquad sm+1 \le k \le sn. \tag{29}$$

The code (28) *contains M distinct words and tables of exactly two symbols with mean error*

$$D = 1 - \frac{8t+3c}{Md} = 1 - \frac{2}{M} - \frac{C}{Md}. \tag{30}$$

For $t = 5, c = 0$, this error is exactly 3/4. In that case the zero word has error zero and each other word has error 4/5.

4.1 Product polynomials and independence

Set $\delta = \gamma^m$, which has order ds, and define, for $0 \le i < d$,

$$P_i(Y) = \prod_{\ell=0}^{s-1} (Y - \delta^{si+\ell}), \qquad p_i(X) = P_i(X^m). \tag{31}$$

Each p_i is monic of degree $sm < k$.

We record the coefficient argument to make its field dependence explicit. For $0 \le r \le h \le s$, set

$$Q_{h,r} = \prod_{v=0}^{r-1} \frac{1-\delta^{h-v}}{1-\delta^{r-v}}, \qquad Q_{h,0} = 1.$$

The denominators are nonzero because all their positive exponents are at most $s < ds$. With $Q_{h,r} = 0$ outside $0 \le r \le h$, directly putting the two products over a common denominator gives

$$Q_{h,r} = Q_{h-1,r} + \delta^{h-r} Q_{h-1,r-1}. \tag{32}$$

For $1 \le r < h$, for example, divide the right side by $Q_{h,r}$. The two ratios are $(1-\delta^{h-r})/(1-\delta^h)$ and $\delta^{h-r}(1-\delta^r)/(1-\delta^h)$, whose sum is one. The boundary cases follow from $Q_{h,0} = Q_{h,h} = 1$. Multiplication by the last factor in $\prod_{\ell=0}^{h-1}(1+\delta^\ell T)$, together with (32), proves inductively that

$$\prod_{\ell=0}^{h-1}(1+\delta^\ell T) = \sum_{r=0}^{h} \delta^{r(r-1)/2} Q_{h,r} T^r. \tag{33}$$

Consequently, writing $P_0(Y) = \sum_{r=0}^{s} \beta_r Y^{s-r}$, we have

$$\beta_r = (-1)^r \delta^{r(r-1)/2} \prod_{v=0}^{r-1} \frac{1-\delta^{s-v}}{1-\delta^{r-v}} \neq 0. \tag{34}$$

Every numerator is nonzero for the same order reason. This is an identity using invertible field elements; no assumption about ordinary integer binomial coefficients is involved.

Let $\alpha_i = \delta^{si}$. These d elements are distinct, since $\alpha_i = \alpha_j$ would imply $ds \mid s(i-j)$ and hence $d \mid i-j$. Scaling roots gives

$$P_i(Y) = \alpha_i^s P_0(Y/\alpha_i) = \sum_{r=0}^{s} \beta_r \alpha_i^r Y^{s-r}. \tag{35}$$

Since $s \ge d-1$, the first d coefficient rows form the product of the invertible diagonal matrix with entries $\beta_0, \ldots, \beta_{d-1}$ and the Vandermonde matrix $(\alpha_i^r)_{0 \le r, i < d}$. Its determinant is a nonzero multiple of $\prod_{i<j}(\alpha_j - \alpha_i)$. Thus the P_i are independent, and substitution $Y = X^m$ shows the p_i are independent as well.

The full-length divisibility condition excludes characteristics dividing ds. This restriction belongs to this explicit realization, not to the all-field geometry or the universal converse.

4.2 Exact block kernels and agreement counts

Let $V = \mathrm{span}(p_0, \ldots, p_{d-1})$. Write every block index uniquely as $r = i + dj$, with $0 \le i < d$ and $0 \le j < m$. Raising its evaluation elements to the mth power gives

$$(\gamma^{s(i+dj)+\ell})^m = \delta^{si+\ell}, \qquad 0 \le \ell < s. \tag{36}$$

Every $f \in V$ is $P(X^m)$ for a unique polynomial P of degree at most s. Its evaluation vanishes on the block precisely when P vanishes at the s distinct roots of P_i. Divisibility by their monic product implies that P is a scalar multiple of P_i. Conversely p_i vanishes there. Therefore

$$\ker(E_{i+dj}|_V) = \mathbb{F}_q p_i. \tag{37}$$

The folded evaluations of all polynomials in V depend on the type i but not on the replicate j.

Use the M message polynomials

$$S = \{0\} \ \cup \ \bigcup_{h=0}^{t-1} \{p_{2h}, p_{2h+1}, p_{2h} + p_{2h+1}\} \ \cup \ \{p_{2t}, \ldots, p_{d-1}\}. \tag{38}$$

Independence makes them distinct, and encoder injectivity makes their codewords distinct. At a square type $i < 2t$, with partner i', put

$$A_{i+dj} = \{0, E_{i+dj}(p_{i'})\}. \tag{39}$$

The second symbol is nonzero by (37). The two corresponding fibers on S are exactly

$$\{0, p_i\}, \qquad \{p_{i'}, p_i + p_{i'}\}.$$

To check that no other message lies in them, express the message in the independent p_ℓ basis: equality of folded evaluations is equivalent to equality of all coefficients other than the ith. The supports in (38) give precisely the four listed possibilities. Thus exactly four messages agree at every block of a square type.

At a coloop type $i \geq 2t$, use $A_{i+dj} = \{0, E_{i+dj}(p_0)\}$. These again are distinct symbols. The zero fiber on S is $\{0, p_i\}$, and the other fiber is the singleton $\{p_0\}$: the set (38) contains no message $p_0 + \lambda p_i$ with $\lambda \neq 0$. Hence exactly three messages agree there.

There are m blocks of each type, so total agreement is $(8t + 3c)m$. Division by $Mn = Mdm$ proves (30), using $2d + C = 8t + 3c$. When $c = 0, t = 5$, the zero message agrees at all $10m$ blocks and each nonzero message at precisely the $2m$ blocks of its square. Their individual errors are consequently 0 and $4/5$, while the mean is $3/4$. This completes the proof of Theorem 4.1.

4.3 Fixed rate and arbitrarily large folding and length

At the smallest degree bound $k = sm + 1$, the rate in Theorem 4.1 is $1/d + 1/(dsm)$. Increasing k preserves all the constructed messages and tables. This gives exact fixed-rate families as follows.

Proposition 4.2. *Fix a rational $R_0 \in (1/d, 1)$ and an integer $s \geq d-1$. There are arbitrarily large prime powers q for which the full-length FRS code has rate exactly R_0 and the mean error* (30)*. In particular, for every prescribed folding threshold and length threshold, there is such a code with $t = 5, c = 0$, rate $5/48$, and mean error $3/4$ exceeding both thresholds.*

Proof. Write $R_0 = u_0/v_0$ in lowest terms and let $D = \operatorname{lcm}(ds, v_0)$. Choose a prime p not dividing D. For positive integers h, Euler's theorem gives $D \mid p^{\varphi(D)h} - 1$. Set $q = p^{\varphi(D)h}$, choose a primitive element of $\mathbb{F}_q$, and put

$$m = \frac{q-1}{ds}, \quad n = \frac{q-1}{s}, \quad k = R_0(q-1).$$

All three are integers, and $k/(sn) = R_0$. Since $R_0 > 1/d$, $k - sm = (R_0 - 1/d)(q-1) \geq 1$ for all sufficiently large h. Also $k < sn$, and $n \to \infty$. Theorem 4.1 applies.

For the last assertion, first choose any integer $s \geq 9$ above the prescribed folding threshold, and then choose h to make n large. Here $D = \operatorname{lcm}(10s, 48)$ is divisible by 240, and

$$k - sm = \left(\frac{5}{48} - \frac{1}{10}\right)(q-1) = \frac{q-1}{240} \geq 1.$$

Thus the rate is exactly $5/48$ and the error is exactly $3/4$. □

At $R_0 = 5/48$, the curve (2) equals $217/288$ and

$$\frac{217}{288} - \frac{751}{1000} = \frac{89}{36000} > \frac{1}{1000}.$$

This proves (6). The quantifier order is explicit: fix the rate, radius, and slack; choose any sufficiently large fixed folding; then obtain arbitrarily large lengths. Thus no finite folding or length threshold repairs the proposed average-radius guarantee below the cube curve. The fifteen nonzero words still have individual error 4/5, so the exhibited list gives no ordinary counterexample at this radius.

4.4 Two finite instances

The construction gives the following examples over prime fields. The generators and all message polynomials are specified by (31); no generic choice is required.

q	γ	s	m	n	k	R
541	2	9	6	60	55	11/108
2161	23	9	24	240	225	5/48

Both have mean error 3/4. Their individual agreement counts are, respectively, 60 and fifteen copies of 12, and 240 and fifteen copies of 48. The cube radii are 245/324 and 217/288. For the first instance, $3/4 < 151/200 < 245/324$.

For completeness, $q-1$ has prime divisors $2, 3, 5$ in each case. The residues of $\gamma^{(q-1)/r}$ for $r = 2, 3, 5$ are $(540, 129, 48)$ modulo 541 and $(2160, 593, 953)$ modulo 2161. All are different from one, so the stated elements are primitive. The reproducibility supplement contains the expanded polynomials, folded symbols, tables, and every agreement bit for these instances.

5 Power-of-two lists

For $M = 2^a$, choose as many squares as possible:

$$t = \left\lfloor \frac{M-1}{3} \right\rfloor, \qquad c = (M-1) \bmod 3, \qquad d = 2t + c, \qquad C = 4t + c. \tag{40}$$

Here $c = 0$ for even a and $c = 1$ for odd a.

Corollary 5.1. *For every fixed $a \geq 4$, the independent-square converse* (19) *is strictly below the cube curve* (1) *at every fixed rate*

$$0 < R < \frac{2}{B_a}. \tag{41}$$

In particular this holds throughout $0 < R \leq 1/d$. The earlier equal-group comparison has the possibly narrower range $0 < R < C/(dB_a)$. For every rational

$$R_0 \in \left(\frac{1}{d}, \frac{C}{dB_a} \right), \tag{42}$$

full-length FRS codes at that exact rate realize the constant error $1 - 2/M - C/(Md)$ at arbitrarily large prescribed fixed folding and arbitrarily large length.

Proof. For even a, substitute $t = (M-1)/3, c = 0$; for odd a, substitute $t = (M-2)/3, c = 1$. Direct subtraction from $B_a = a(M-2)/(a-1)$ gives

$$C - B_a = \begin{cases} \dfrac{(a-4)M + 2a + 4}{3(a-1)}, & a \text{ even}, \\ \dfrac{(a-4)M + a + 5}{3(a-1)}, & a \text{ odd}. \end{cases} \tag{43}$$

Both expressions are positive for $a \geq 4$. We also have $4t > B_a$: for even a, $4t = C$; for odd $a \geq 5$,

$$4t - B_a = \frac{(a-4)(M-2)}{3(a-1)} > 0.$$

If $0 < R < 2/B_a$, then $2 > B_a R$ and $CR > B_a R$. The third needed inequality is $1 + 2tR > B_a R$. It is immediate when $B_a \leq 2t$. Otherwise

$$(B_a - 2t)R < (B_a - 2t)\frac{2}{B_a} = 2 - \frac{4t}{B_a} < 1.$$

Thus all three terms in $e(R) = \min\{2, 1 + 2tR, CR\}$ strictly exceed $B_a R$. The difference between the cube radius and (19) is $(e(R) - B_a R)/M > 0$. Since $C = 2d - c \leq 2d$ and $C > B_a$, the earlier endpoint $C/(dB_a)$ lies above $1/d$ and is at most $2/B_a$. For equal groups the excess is $C \min\{R, 1/d\}$, whose comparison with $B_a R$ gives precisely the stated earlier range.

The exact-rate realization follows from Theorem 4.1 and Proposition 4.2. At rates (42), the difference between the cube radius and the explicit FRS mean error is $(C/d - B_a R_0)/M > 0$. Any radius strictly between them gives a counterexample with fixed positive gap. The unequal-group universal converse does not change the agreement counts in this particular FRS construction. □

For $M = 16$, the universal range is still $0 < R < 3/28$. For $M = 32$, it increases from $0 < R < 82/1575$ to $0 < R < 4/75$; the displayed exact FRS construction retains the former upper endpoint.

For $a = 2, 3$, the two coefficients agree, as the table in Section 3 shows. As $a \to \infty$, $C/B_a \to 4/3$. This coefficient comparison concerns the chosen configurations and certificate model.

6 Remaining questions

The sparse profile shows that the full Boolean grid does not govern the fine average-radius tradeoff at these small output caps. The finite converse holds for every linear code in a vector alphabet, while the explicit FRS construction explains the local geometry without the dimension slack needed for greedy shortening. The unequal-group refinement and its optimized square/coloop allocation belong to this same construction and result family.

Unequal allocation removes the avoidable loss from equal groups in the coloop cases, but the saturated part may still admit stronger profiles. Proposition 2.5 gives the optimum within its stated family. The obstruction concerns average-radius recovery. The larger-list average-radius guarantees in [1, 5] are compatible with these fixed-cap obstructions.

A Reproducibility and proof status

All universal statements in the note have analytic proofs above. The theorems follow from their printed analytic proofs. The accompanying `reproducibility/reproduce.py` supplies finite diagnostics and expanded instances. The supplement includes the frozen examples `examples/q541_explicit_certificate.json` and `examples/q2161_fixed_rate_certificate.json`. Coefficients are stored in ascending degree order, and field elements in these prime-field instances are integer residues. From the package directory, the diagnostics can be run with

```
python reproducibility/reproduce.py --output-dir ../reproduction-output
```

Choose an absent or empty output directory outside the sealed package.

For a concrete FRS instance, the checks are the following: the evaluation points are all distinct; the generator has full order; the p_i have the prescribed degree and rank; each restricted block kernel is its named axis line; the message list has the prescribed cardinality; the two table symbols are distinct; and the coordinate and individual agreement counts equal the displayed values. The complete agreement bits allow direct inspection of the difference between average and individual error.

Finite linear-code diagnostics also cover small fields and vector alphabets, including characteristic two, collapsed singleton tables, empty groups, and exhausted coordinate supplies. They check both the equal-group construction and the unequal-group agreement guarantee in Theorem 2.1. A separate exact allocation check compares the square-first formula with exhaustive finite allocations; exact rational checks cover the two finite improvements, the piecewise asymptotic curve, and the power-of-two comparison ranges. The fixed-rate FRS gap and the coefficient formulas are checked as well. These checks support reproducibility; their finite coverage does not replace the quantifiers in the proofs.

AI assistance. AI tools assisted with mathematical exploration, literature retrieval, drafting, and internal proof and code checks.

References

[1] Joshua Brakensiek, Yeyuan Chen, Manik Dhar, and Zihan Zhang. Combinatorial bounds for list recovery via discrete Brascamp–Lieb inequalities. *arXiv:2510.13775v2*, October 11, 2026. An extended abstract appeared in STOC 2026. `https://arxiv.org/abs/2510.13775v2`.

[2] Yeyuan Chen. List-recovery lower bound for folded linear codes. Author-posted manuscript, cited as 2026 in [5]; the PDF has no printed date. Author-posted PDF.

[3] Yeyuan Chen, Mahdi Cheraghchi, King-Siong Si, and Zihan Zhang. Evidence toward a conjecture in list recoverability of folded Reed–Solomon codes. Manuscript, 2025, as reported in [1]. Full text not examined here.

[4] Yeyuan Chen and Zihan Zhang. Explicit folded Reed–Solomon and multiplicity codes achieve relaxed generalized Singleton bounds. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1–12, 2025. doi:10.1145/3717823.3718114. Full version: *arXiv:2408.15925v3*, April 14, 2025. `https://arxiv.org/abs/2408.15925v3`.

[5] Rohan Goyal and Venkatesan 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/`.

[6] Matan Levi, Jonathan Mosheiff, and Nikhil Shagrithaya. Random Reed–Solomon codes and random linear codes are locally equivalent. *arXiv:2406.02238v7*, November 14, 2025. `https://arxiv.org/abs/2406.02238v7`.

[7] Ray Li and Nikhil Shagrithaya. Near-optimal list-recovery of linear code families. In *APPROX/RANDOM 2025*, volume 353 of LIPIcs, pages 53:1–53:14, 2025. doi:10.4230/LIPIcs.APPROX/RANDOM.2025.53. Full version: *arXiv:2502.13877v2*, February 27, 2025. `https://arxiv.org/abs/2502.13877v2`.

[8] Scott Schubr. Eight-point fiber certificates and binary-input list recovery of folded Reed–Solomon codes. Luminary Research Archive, edition 1, October 18, 2026. Public archive record.