Counterexamples to the fine-grained list-recovery conjecture for folded Reed–Solomon and random linear codes

Scott Schubr
Luminary Oddities

Abstract

We give counterexamples to the fine-grained list-recovery conjecture of Chen and Zhang for folded Reed–Solomon codes. For every fixed input-list size $\ell \geq 3$ and integer $a \geq 3$, and for input-list size two with $a \geq 4$, we construct ℓ^a distinct codewords that each meet a common error radius strictly below the conjectured threshold for output cap $\ell^a - 1$. For each fixed pair (ℓ, a) this happens at every rational rate in an explicit interval $(0, R_0^*)$, in every folded Reed–Solomon code with appropriate evaluation points, folding above a fixed bound and sufficiently large length. The construction combines coordinate-square message sets with blocks on which every selected message has a distinct symbol. An elementary degree-scheduling argument assigns the remaining input tables to equalize all individual errors. In consecutive full-length codes, polynomial identities and Vandermonde determinants realize the required evaluation kernels exactly; in general, random codewords vanishing on prescribed blocks realize them on all but a few blocks. The argument therefore applies to every linear code with a block-surjectivity property, and, with high probability, to random linear codes over every fixed field with at least ℓ elements. As a concrete instance, sixteen codewords each have error $55/64$ at exact rate $1/79$, violating the conjectured ordinary cap-fifteen tradeoff with fixed positive slack.

1 Introduction and statements

List recovery asks how many codewords can agree with a small list of allowed symbols at each coordinate. In its ordinary version each recovered codeword must meet the error threshold individually; that is the version considered throughout this paper.

Chen and Zhang [4, Conjecture 3.6] proposed a precise tradeoff for folded Reed–Solomon codes [5] when the output cap plus one is a power of the input-list size, and showed with a hypercube-like message configuration that the conjectured radius cannot be exceeded [4, Theorem 3.1]. Brakensiek, Chen, Dhar and Zhang [1] proved the zero-error case, restated the positive-error question as their Conjecture 1.3, and call it the fine-grained question [1, Remark 1.7]; the case of input-list size two and output cap three is proved in [4, Theorem 1.8] (see [1, Remark 1.7]). For every pair (ℓ, a) in the region of Theorem 1.2 we disprove that tradeoff at all rates in an explicit interval $(0, R_0^*)$, for every folded Reed–Solomon code with appropriate evaluation points and sufficiently large folding and length. An earlier note [7] introduced the configuration of independent coordinate squares sharing an origin, realized it in standard full-length folded Reed–Solomon codes, and used it to rule out the average-radius extrapolation of the conjectured curve at output cap fifteen; it left the ordinary conjecture open. The new ingredient here is the injective padding and degree scheduling of Section 3, which equalizes every individual error and so yields ordinary counterexamples, together with the extension to every (ℓ, a) in (4).

The mechanism has two parts. Several coordinate squares sharing one origin give large local captures, but initially produce unequal individual agreements. We then add evaluation-block types

on which the selected messages have distinct symbols. At those blocks any ℓ selected messages can be chosen by an input table. Prescribed degree sequences in uniform multihypergraphs distribute these choices so that every message reaches the same agreement count. Filling remaining coordinates evenly is already used in the upper-bound construction of Chen–Zhang [4, Section 3.1]; what is new is a schedule over injective blocks that balances unequal deficits from a different message geometry. An exact polynomial construction supplies all the special and injective block types in the ordinary consecutive full-length code. The mechanism uses the code only through this pattern of evaluation kernels, and Section 4 realizes the pattern, up to a few defective blocks, in every linear code with block surjectivity: in every folded Reed–Solomon code with appropriate evaluation points, and with high probability in random linear codes.

1.1 The code and the conjecture

Let q be a prime power, let γ generate $\mathbb{F}_q^\times$, and let s, n, k be positive integers with $k < sn$. The message space consists of the polynomials $f \in \mathbb{F}_q[X]$ of degree less than k. For evaluation points $\alpha_1, \dots, \alpha_n \in \mathbb{F}_q$, the s-folded encoding is

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

The rate over the alphabet $\mathbb{F}_q^s$ is $R = k/(sn)$. Following [4, Definition 1.2], evaluation points $\alpha_1, \dots, \alpha_n$ are *appropriate* if the sn elements $\gamma^i\alpha_j$ $(0 \le i < s,\ 1 \le j \le n)$ are distinct. Then $sn \le q$, and the encoding is injective, since a nonzero polynomial of degree less than $k < sn$ has fewer than sn roots. The *consecutive full-length* code takes $\alpha_{r+1} = \gamma^{sr}$ and $sn = q - 1$:

$$E_r(f) = \big(f(\gamma^{sr}), \dots, f(\gamma^{sr+s-1})\big), \qquad 0 \le r < n. \tag{2}$$

Its unfolded points γ^{sr+i} run through the nonzero field elements exactly once and are appropriate.

An input table is a sequence $\mathcal{T} = (\mathcal{T}_r)_{r=0}^{n-1}$ with $\mathcal{T}_r \subseteq \mathbb{F}_q^s$ and $|\mathcal{T}_r| \le \ell$. The relative error of a message is

$$\operatorname{err}_{\mathcal{T}}(f) = \frac{1}{n}\big|\{r : E_r(f) \notin \mathcal{T}_r\}\big|.$$

The code is (ρ, ℓ, L) list recoverable if, for every such table, at most L messages have error at most ρ. For $M = \ell^a$, define

$$\Phi_{\ell,a}(R) = \frac{M - \ell}{M}\left(1 - \frac{aR}{a-1}\right). \tag{3}$$

Conjecture 1.1 (Chen–Zhang [4, Conjecture 3.6])**.** *Let $\ell \ge 2$ and $a \ge 2$ be integers, put $L+1 = \ell^a$, and fix $\varepsilon > 0$, a rate $R \le (a-1)/a$, and a generator γ. There is a constant C_0 such that, if $s \ge C_0$, $(k-1)/s \ge a$ and n is sufficiently large, then every rate-R folded Reed–Solomon code with appropriate evaluation points is $(\Phi_{\ell,a}(R) - \varepsilon, \ell, L)$ list recoverable.*

The conjecture quantifies over every code with appropriate evaluation points, and C_0 may depend on ε, ℓ, a, R and γ. It is therefore refuted by exhibiting, for one fixed (ℓ, a, R) with $R \le (a-1)/a$ and one fixed $\varepsilon > 0$, codes with arbitrarily large folding s and, for each such s, arbitrarily large length and $(k-1)/s \ge a$, for every choice of generator, that contain $L+1$ messages of error at most $\Phi_{\ell,a}(R) - \varepsilon$ against some table. The reformulation [1, Conjecture 1.3] puts the slack in the rate instead; since $\Phi_{\ell,a}$ is linear and strictly decreasing, a radius slack ε is the same as a rate slack $(a-1)M\varepsilon/(a(M-\ell))$, and all errors below are at most $1 - \ell/M$, so the strict violations below refute both forms.

1.2 Parameters and results

Fix integers ℓ, a with

$$(\ell \geq 3 \text{ and } a \geq 3) \quad \text{or} \quad (\ell = 2 \text{ and } a \geq 4), \tag{4}$$

and put

$$M = \ell^a, \quad N = \frac{M-1}{\ell-1}, \quad c = \begin{cases} 0, & a \text{ even}, \\ 1, & a \text{ odd}, \end{cases} \quad t = \frac{N-c}{\ell+1}, \quad d = 2t + c,$$

$$C = (\ell-1)(2\ell t + c), \qquad B = \frac{a(M-\ell)}{a-1}. \tag{5}$$

Section 3 shows that t is a positive integer and $C > B$.

Theorem 1.2. *Fix integers ℓ, a such that either $\ell \geq 3$, $a \geq 3$, or $\ell = 2$, $a \geq 4$, put $M = \ell^a$, and put*

$$R_0^* = R_0^*(\ell, a) = \frac{\ell C}{B(Md - C)}$$

in the notation of Section 1.2. For every rational rate $R \in (0, R_0^*)$ there are a rational $\varepsilon = \varepsilon(\ell, a, R) > 0$ and an integer $n_0 = n_0(\ell, a, R)$ such that the following holds with $s_0 = s_0(\ell, a) = d+1$. For every integer $s \geq s_0$, every s-folded Reed–Solomon code of rate exactly R and length $n \geq n_0$ with appropriate evaluation points, over any field $\mathbb{F}_q$ and for any generator γ, admits a table of exactly ℓ symbols per block and M distinct messages, each satisfying

$$\mathrm{err}_{\mathcal{T}}(f) < \Phi_{\ell,a}(R) - \varepsilon.$$

These codes satisfy $(k-1)/s \geq a$, and they exist for every s and every n with $Rsn \in \mathbb{Z}$: for every prime power $q > sn$ the consecutive points $\alpha_{r+1} = \gamma^{sr}$ are appropriate. For example $R_0^*(2,4) = 3/196$ and $R_0^*(3,3) = 19/906$.

Thus every indicated fixed parameter instance of the conjecture fails at every sufficiently small fixed rate, each with its own fixed positive slack, and the failure occurs in every code over which the conjecture quantifies once $s \geq d+1$ and $n \geq n_0$. The parameters ℓ, a are fixed before the folding and length quantifiers. In consecutive full-length codes, at the rates covered by Corollary 3.3, the folding may equal d and all M errors are exactly equal. Proposition 3.4 in Section 3 compares the resulting error with the optimal list-recovery radius.

Theorem 1.3 (A numerical cap-fifteen instance)**.** *For every integer $s \geq 10$ there are arbitrarily long consecutive full-length folded Reed–Solomon codes of exact rate $1/79$ and tables of exactly two symbols per block for which sixteen distinct messages each have relative error $55/64$. At this rate,*

$$\Phi_{2,4}(1/79) - \frac{55}{64} = \frac{13}{15168} > \frac{1}{10000},$$

so these codes are not $(55/64 + 1/10000,\ 2,\ 15)$ list recoverable.

The rate $1/79$ lies in $(0, R_0^*(2,4))$, so Theorem 1.2 also covers it, for every code of that rate with appropriate evaluation points, folding $s \geq 11$ and large length; Theorem 1.3 records the instance with a direct pair schedule and explicit constants.

The construction is a statement about linear codes with a prescribed pattern of evaluation kernels. Section 4 isolates a block-surjectivity property that produces this pattern, proves Theorem 1.2 from it, and shows that uniformly random $\mathbb{F}_q$-linear codes, for every fixed prime power $q \geq \ell$ and large folding, fail the same tradeoff with high probability (Theorem 4.6). We first realize prescribed axis kernels and injective blocks by polynomials (Section 2), then build the equal-error configuration (Section 3), extend it to codes with block surjectivity and prove Theorems 1.2 and 4.6 (Section 4), and prove Theorem 1.3 (Section 5).

2 A full-length realization with injective padding blocks

The following elementary polynomial construction supplies both axis-kernel blocks and blocks injective on the entire message configuration, with no defective blocks. Section 4 obtains the same kernel pattern in general codes by a random choice of codewords.

Lemma 2.1 (Polynomial realization)**.** *Fix integers $D > d \geq 1$ and $s \geq d$. Suppose $q - 1 = Dsm$, where m is a positive integer, and let γ generate $\mathbb{F}_q^\times$. Set $n = Dm$, $\delta = \gamma^m$, and, for $0 \leq i < D$, define*

$$\theta_i = \delta^{si}, \tag{6}$$

$$P_i(Y) = \prod_{u=0}^{s-1} (Y - \delta^{si+u}), \qquad p_i(X) = P_i(X^m). \tag{7}$$

Use the consecutive folded evaluations (2)*. The polynomials $p_0, \ldots, p_{d-1}$ are independent, and, on $V = \operatorname{span}_{\mathbb{F}_q}\{p_0, \ldots, p_{d-1}\}$,*

$$\ker(E_{i+Dj}|_V) = \begin{cases} \mathbb{F}_q p_i, & 0 \leq i < d, \\ \{0\}, & d \leq i < D, \end{cases} \qquad 0 \leq j < m. \tag{8}$$

Every p_i has degree sm.

Proof. The element δ has order Ds. To establish independence without any characteristic restriction beyond the existence of this root of unity, write

$$P_0(Y) = \sum_{r=0}^{s} (-1)^r e_{s,r} Y^{s-r}.$$

The finite geometric-product identity gives

$$e_{s,r} = \delta^{r(r-1)/2} \prod_{u=0}^{r-1} \frac{1 - \delta^{s-u}}{1 - \delta^{r-u}}, \qquad 0 \leq r \leq s, \tag{9}$$

where the empty product is one. For completeness, the coefficients in $\prod_{u=0}^{s-1}(1 + \delta^u T)$ obey $e_{s,r} = e_{s-1,r} + \delta^{s-1} e_{s-1,r-1}$, with zero terms outside the coefficient range. The displayed products satisfy this recurrence: after removing $\delta^{r(r-1)/2}$, it is the identity

$$\begin{bmatrix} s \\ r \end{bmatrix}_\delta = \begin{bmatrix} s-1 \\ r \end{bmatrix}_\delta + \delta^{s-r} \begin{bmatrix} s-1 \\ r-1 \end{bmatrix}_\delta,$$

whose product verification reduces to $1 - \delta^s = (1 - \delta^{s-r}) + \delta^{s-r}(1 - \delta^r)$. Here the bracket denotes the product in the preceding coefficient formula without its power of δ. The boundary values and induction prove (9). All numerator and denominator factors are nonzero, since their positive exponents are at most $s < Ds$. In particular every coefficient of P_0 is nonzero in $\mathbb{F}_q$.

The scaling relation $P_i(Y) = \theta_i^s P_0(Y/\theta_i)$ shows that its coefficient of Y^{s-r} is $(-1)^r e_{s,r} \theta_i^r$. The D parameters θ_i are distinct, since δ^s has order D. On any $d + 1$ selected columns, the coefficient rows $r = 0, \ldots, d$ form an invertible diagonal matrix times a Vandermonde matrix. Thus any $d + 1$ of the P_i are independent. Substitution $Y = X^m$ preserves linear independence.

Write $r = i + Dj$. For an unfolded point in that block,

$$(\gamma^{s(i+Dj)+u})^m = \delta^{si+u},$$

because δ has order Ds. Hence every polynomial $P(X^m)$ has the same evaluation on the m replicas of type i. A polynomial P of degree at most s vanishes on the s distinct points $\delta^{si}, \ldots, \delta^{si+s-1}$ if and only if P is a scalar multiple of P_i: its monic root product has degree s and divides P. For $i < d$ this gives the asserted axis kernel on V. For $i \geq d$, the span of $P_0, \ldots, P_{d-1}$ has zero intersection with $\mathbb{F}_q P_i$, by the $d+1$ independence. This proves (8). □

Since the $sn = q-1$ unfolded points are distinct, the encoding is injective for every $k \leq q-1$, and the rate is $k/(sn)$. Whenever $sm < k \leq sn$, all of V lies in the message space. For a finite set of distinct messages in V, every type $i \geq d$ is injective, so a table consisting of the evaluations of any chosen ℓ distinct messages captures exactly those ℓ messages from that set.

3 All higher power caps

Throughout this section (ℓ, a) satisfies (4) and M, N, c, t, d, C, B are as in Section 1.2. All parameters are fixed before taking field-size or length limits. The integer t is positive: reducing $N = 1 + \ell + \cdots + \ell^{a-1}$ modulo $\ell + 1$ gives remainder c. In particular $1 + t(\ell^2 - 1) + c(\ell - 1) = M$. Elementary expansion gives

$$C = \frac{2\ell(M-1) - c(\ell-1)^2}{\ell+1} < 2M, \tag{10}$$

$$(C-B)(\ell+1)(a-1) = \begin{cases} [a(\ell-1) - 2\ell]M + \ell[a(\ell-1)+2], & a \text{ even}, \\ [a(\ell-1) - 2\ell]M + a(\ell-1) + \ell^2 + 1, & a \text{ odd}. \end{cases} \tag{11}$$

In (4), the coefficient $a(\ell-1) - 2\ell$ is nonnegative and the remaining term is positive. Thus $C > B$. For $a = 2$ and for $(\ell, a) = (2, 3)$ the two constants are equal.

Lemma 3.1 (Uniform subset scheduling)**.** *Let $G \geq 0$ and let nonnegative integers $(z_v)_{v \in S}$ satisfy $\sum_v z_v = \ell G$ and $\max_v z_v \leq G$. Then there is a multiset of G subsets of S, each of size exactly ℓ, in which vertex v occurs exactly z_v times.*

Proof. Induct on G. At $G = 0$ all degrees vanish. If $G > 0$, at least ℓ degrees are positive, since fewer could not sum to ℓG while each is at most G. Select the ℓ largest positive degrees and decrease each by one. Every old degree equal to G is selected: otherwise there would be at least $\ell + 1$ such degrees, exceeding the total. The new maximum is at most $G - 1$ and the new sum is $\ell(G-1)$. Apply induction and add the selected ℓ-set. Repeated subsets are permitted; repeated vertices inside a subset are not. □

Proposition 3.2 (Ordinary padding at every higher power cap)**.** *Let $D \geq (Md - C)/\ell$ be an integer, let $s \geq d$, and let $q \geq \ell$ be a prime power with $DsM \mid q - 1$. Put $m = (q-1)/(Ds)$ and $n = Dm$, and let k be any integer with $sm < k \leq sn$. The consecutive full-length s-folded code of dimension k contains M distinct messages and a table of exactly ℓ symbols at every block against which every message has exactly $m(\ell D + C)/M$ agreements, that is, agreement fraction*

$$\tau_D = \frac{\ell D + C}{MD}.$$

Proof. Apply Lemma 2.1. It supplies independent degree-sm polynomials $p_0, \ldots, p_{d-1}$ spanning V, with kernel $\mathbb{F}_q p_i$ on special type $i < d$ and zero kernel on every other type; each type has m blocks. Since $k > sm$, V lies in the message space.

Choose $A \subseteq \mathbb{F}_q$ of size ℓ, containing $0, 1$. No subfield assumption is made. In each of the t coordinate planes put an $A \times A$ square, sharing only the origin, and on the possible remaining axis put an A-line:

$$S = \{0\} \cup \bigcup_{h=0}^{t-1} \{\alpha p_{2h} + \beta p_{2h+1} : (\alpha, \beta) \in A^2 \setminus \{(0,0)\}\} \cup \bigcup_{i=2t}^{d-1} \{\alpha p_i : \alpha \in A \setminus \{0\}\}.$$

Independence makes all these messages distinct, so $|S| = M$.

At special type $2h$, use the ℓ symbols $E_{2h}(\beta p_{2h+1})$, $\beta \in A$; at type $2h+1$, use $E_{2h+1}(\alpha p_{2h})$, $\alpha \in A$. The symbols are distinct because the kernel is the corresponding other coordinate axis. The selected fibers capture precisely the ℓ^2 square messages: its zero fiber is the A-axis, and its other fibers are the parallel A-lines. No message in another square or on the remaining axis line can enter one of these fibers, by coordinate independence.

If $c = 1$, at its special axis use symbol zero and the $\ell - 1$ symbols $E_i(\beta p_0)$, $\beta \in A \setminus \{0\}$. The zero fiber consists of the ℓ messages of the remaining axis line. Each other selected fiber contains the single message βp_0: there are no messages in S combining a square with the remaining axis. These symbols are distinct, again by the exact kernel. This table therefore captures exactly $2\ell - 1$ messages. Here $t \geq 1$ ensures the required square containing p_0.

Let h_v count the special types capturing message v. Then

$$0 \leq h_v \leq d, \qquad \sum_{v \in S} h_v = 2t\ell^2 + c(2\ell - 1) = \ell d + C. \tag{12}$$

After their m replicas, message v has mh_v agreements. There are $G = (D - d)m$ generic blocks. Prescribe additional degrees

$$z_v = m\left(\frac{\ell D + C}{M} - h_v\right). \tag{13}$$

Since $M \mid m$, these are integers. They are nonnegative because $\ell D + C \geq Md$ gives $(\ell D + C)/M \geq d \geq h_v$; the origin has $h_v = d$. For the upper bound, every message has $h_v \geq 1$, so it suffices that $(\ell D + C)/M - 1 \leq D - d$, that is $(M - \ell)D \geq Md + C - M$. In the region (4) one has $d \geq 7$ and $M/\ell \geq 4$. Put $x = Md - C$; then $x \geq M(d-2) \geq 5M$ by (10), and $\ell D \geq x$ gives

$$(M - \ell)D - (Md + C - M) \geq \Big(\frac{M}{\ell} - 2\Big)x - (2C - M) \geq 10M - 3M > 0.$$

Also $D > d$, since $x \geq M(d-2) > \ell d$, because $(M - \ell)d > 2M$ for $d \geq 7$ and $M/\ell \geq 4$. Hence $z_v \leq G$. Equation (12) yields $\sum_v z_v = \ell G$. Lemma 3.1 supplies G subsets of ℓ distinct messages with those degrees. Assign them to the generic blocks. Each generic evaluation is injective on V, so its assigned subset provides exactly ℓ distinct table symbols and captures precisely those messages. Every message now has the same number of agreements,

$$mh_v + z_v = \frac{m(\ell D + C)}{M} = n\tau_D. \qquad \square$$

The displayed estimate uses D only through $\ell D \geq Md - C$, so it holds for every real $D \geq (Md - C)/\ell$; Section 4 uses it in this form.

Corollary 3.3 (The defect-free case)**.** *Let $D \geq (Md - C)/\ell$ be an integer and let R be rational with $1/D < R < C/(BD)$. For every integer $s \geq d$ there are arbitrarily large prime powers q and consecutive full-length s-folded Reed–Solomon codes of rate exactly R, with $q = sn + 1$ and $(k-1)/s \geq a$, containing M distinct messages and a table of exactly ℓ symbols per block against which every message has error exactly $1 - \tau_D < \Phi_{\ell,a}(R)$.*

Proof. Let $R = \nu/\omega$ in lowest terms and $s \geq d$. Pick a prime π not dividing $DsM\omega$ and put $q = \pi^{\varphi(DsM\omega)\lambda}$ for integers $\lambda \geq 1$; Euler's theorem gives $DsM\omega \mid q-1$, and $q \geq \ell$ for large λ. Then $m = (q-1)/(Ds)$ is divisible by M and by ω, so $k = Rsn = \nu sDm/\omega$ is an integer. Moreover $k > sm$ because $RD > 1$, and $k < sn$ because $R < 1$; the code has rate exactly R. By Proposition 3.2, its M messages all have error $1 - \tau_D$, while

$$\Phi_{\ell,a}(R) - (1 - \tau_D) = \frac{1}{M}\Big(\frac{C}{D} - BR\Big) > 0.$$

As $\lambda \to \infty$ with s fixed, $m \to \infty$ and $(k-1)/s > m - 1/s \to \infty$, which gives arbitrarily large length and the eventual condition $(k-1)/s \geq a$. □

Corollary 3.3 is the case of Theorem 1.2 in which the kernel pattern (8) holds with no defective blocks: the folding may equal d, and all M errors coincide. Consecutive windows $(1/D, C/(BD))$ and $(1/(D+1), C/(B(D+1)))$ overlap exactly when $D(C-B) > B$. For general appropriate evaluation points the pattern comes from the block-surjectivity lemma of Section 4, where the proportion $1/D$ of each special type becomes a real parameter ξ and the rate range becomes the whole interval $(0, R_0^*)$.

Proposition 3.4. *Fix (ℓ, a) as in Theorem 1.2. For every rational R with $0 < R \leq 1/(Md-1)$, the codes of Proposition 3.2 at rate exactly R, built with $D = \lfloor 1/R \rfloor + 1$ and the field choice in the proof of Corollary 3.3, contain M messages each of error exactly*

$$1 - \frac{\ell}{M} - \frac{C}{M(\lfloor 1/R \rfloor + 1)} \;\leq\; 1 - \frac{\ell}{M} - \frac{C}{M} \cdot \frac{R}{1+R}.$$

Compared with the optimal large-alphabet radius $(L+1-\ell-LR)/(L+1)$ of [3], with $L+1 = M$, this error is smaller by at least $\big(C/(1+R) - (M-1)\big)R/M$, which is positive for these R.

Proof. With $D = \lfloor 1/R \rfloor + 1$ we have $RD > 1$ and $D \geq Md$, so Proposition 3.2, with the field choice in the proof of Corollary 3.3, applies at rate exactly R. It gives agreement fraction $(\ell D + C)/(MD)$, and $1/D \geq R/(1+R)$. Positivity follows from $C/(M-1) > 1+R$, which holds since $C/(M-1) \geq 4/3 - 1/(3(M-1)) \geq 4/3 - 1/45$ by (10) (for $\ell = 2$ this is the formula with $c \leq 1$; for $\ell \geq 3$ the main term is at least $3/2$ and the subtracted term is at most $(\ell-1)^2/((\ell+1)(\ell^3-1)) < 1/6$), while $Md \geq 160$ in the stated region gives $R \leq 1/159$. □

Remark 3.5. For fixed (ℓ, a), as $R \to 0$, the conjectured curve, this construction and the optimal radius have slopes B/M, C/M and $(M-1)/M$ in R. The upper bound construction of Chen and Zhang [4, Theorem 3.1] already shows that folded Reed–Solomon codes miss the optimal slope by the factor $B/(M-1)$, which tends to 1 as $a \to \infty$. The factor here is

$$\frac{C}{M-1} = \frac{2\ell}{\ell+1} - \frac{c(\ell-1)^2}{(\ell+1)(M-1)},$$

which tends to $2\ell/(\ell+1)$. The range $R \leq 1/(Md-1)$ on which Proposition 3.4 applies depends on a. List-size barriers for linear codes near capacity are treated in [3, Fact 1.4].

4 Codes with block surjectivity and random linear codes

The proof of Proposition 3.2 uses the code only through three features: d independent codewords $p_0, \ldots, p_{d-1}$, equally many special blocks of each type $i < d$ with kernel exactly $\mathbb{F}_q p_i$ on their span

V, and further blocks injective on V. The squares, the fiber captures, the injective padding and Lemma 3.1 are linear algebra and combinatorics. This section obtains the kernel pattern in an arbitrary linear code, at the cost of a few defective blocks, and deduces Theorems 1.2 and 4.6. Throughout, (ℓ, a) satisfies (4) and M, d, C, B are as in Section 1.2.

An $\mathbb{F}_q$*-linear code* of length n over the alphabet $\mathbb{F}_q^s$ is an $\mathbb{F}_q$-subspace $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$; its rate is $\dim_{\mathbb{F}_q} \mathcal{C}/(sn)$, and $c_r \in \mathbb{F}_q^s$ denotes block r of a codeword c. Tables, the error $\mathrm{err}_{\mathcal{T}}(c) = |\{r : c_r \notin \mathcal{T}_r\}|/n$ and (ρ, ℓ, L) list recoverability are defined for codewords exactly as in Section 1.2. A folded Reed–Solomon code with appropriate evaluation points is such a code of dimension k, and, its encoding being injective, messages and codewords correspond.

Definition 4.1 (Block surjectivity)**.** Let $z \geq 1$ with $dz \leq n$, and fix pairwise disjoint sets $Z_0, \ldots, Z_{d-1} \subseteq \{0, \ldots, n-1\}$ of z blocks each. An $\mathbb{F}_q$-linear code $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$ has property BS_z (with respect to $Z_0, \ldots, Z_{d-1}$) if, for every $i < d$ and every block $r \notin Z_i$, the projection of $\mathcal{C}$ onto the blocks $Z_i \cup \{r\}$ is surjective onto $(\mathbb{F}_q^s)^{Z_i \cup \{r\}}$. A block lying in none of the Z_i is *generic*.

Lemma 4.2 (Planted kernels)**.** *Let $\mathcal{C}$ have* BS_z *with respect to $Z_0, \ldots, Z_{d-1}$, and for $i < d$ let $\mathcal{C}_i = \{c \in \mathcal{C} : c_r = 0$ for all $r \in Z_i\}$. There are $p_i \in \mathcal{C}_i$ $(i < d)$ and a set $\mathcal{B}$ of fewer than $q^{d-s}n$ blocks, the* bad *blocks, such that the following holds whenever $n - dz \geq q^{d-s}n$. The codewords $p_0, \ldots, p_{d-1}$ are linearly independent, and on $V = \mathrm{span}_{\mathbb{F}_q}\{p_0, \ldots, p_{d-1}\}$ the map $c \mapsto c_r$ has kernel $\mathbb{F}_q p_i$ at every block $r \in Z_i \setminus \mathcal{B}$ and kernel $\{0\}$ at every generic block $r \notin \mathcal{B}$.*

Proof. Choose p_i uniformly from $\mathcal{C}_i$, independently for $i = 0, \ldots, d-1$. For $r \notin Z_i$ the linear map $\mathcal{C}_i \to \mathbb{F}_q^s$, $c \mapsto c_r$, is onto: given $y \in \mathbb{F}_q^s$, BS_z provides $c \in \mathcal{C}$ vanishing on Z_i with $c_r = y$. All fibers of a surjective linear map are cosets of its kernel and have equal size, so $(p_i)_r$ is uniform on $\mathbb{F}_q^s$.

Fix a block r and let $J_r = \{j < d : r \notin Z_j\}$. Since the Z_j are disjoint, $|J_r| \geq d-1$, and $(p_i)_r = 0$ for the at most one index $i \notin J_r$. The vectors $(p_j)_r$, $j \in J_r$, are independent and uniform. Call r *bad* if they are linearly dependent. Adding them one at a time, the u-th lies in the span of the earlier ones with probability at most q^{u-1-s}, so

$$\Pr[r \text{ bad}] \leq \sum_{u=1}^{|J_r|} q^{u-1-s} = \frac{q^{|J_r|} - 1}{(q-1)q^s} < q^{d-s}.$$

The expected number of bad blocks is less than $q^{d-s}n$; fix an outcome whose set $\mathcal{B}$ of bad blocks has $|\mathcal{B}| < q^{d-s}n$.

Consider the coefficient map $\phi_r : \mathbb{F}_q^d \to \mathbb{F}_q^s$, $(c_j) \mapsto \sum_j c_j (p_j)_r$. At a block $r \in Z_i \setminus \mathcal{B}$ we have $(p_i)_r = 0$ and the other $d-1$ values are independent, so $\ker \phi_r = \mathbb{F}_q e_i$; at a generic block $r \notin \mathcal{B}$ all d values are independent and $\ker \phi_r = 0$. If $n - dz \geq q^{d-s}n$, the $n - dz$ generic blocks are not all bad, and a good generic block r exists. Then $\sum_j c_j p_j = 0$ forces $\phi_r(c) = 0$ and $c = 0$, so the p_j are independent, V is identified with $\mathbb{F}_q^d$ through the coefficients, and the kernels of ϕ_r are the asserted kernels on V. □

Put

$$\tau(\xi) = \frac{\ell + C\xi}{M}, \tag{14}$$

so that $\tau(1/D) = \tau_D$ in Proposition 3.2.

Theorem 4.3 (Square configurations under block surjectivity)**.** *Let $q \geq \ell$ be a prime power and $s \geq 1$, and let $\mathcal{C} \subseteq (\mathbb{F}_q^s)^n$ be an $\mathbb{F}_q$-linear code with* BS_z *with respect to some $Z_0, \ldots, Z_{d-1}$. Put*

$\xi = z/n$ and $\eta = q^{d-s}$, and suppose that

$$(Md - C)\xi + \ell\eta + \frac{M}{n} \le \ell, \tag{15}$$

$$(Md + C - M)\xi + (M - \ell)\eta + \frac{M}{n} \le M - \ell. \tag{16}$$

Then $\mathcal{C}$ contains M distinct codewords and a table with exactly ℓ symbols per block such that each of the M codewords has error less than

$$1 - \tau(\xi) + \frac{(\ell d + C)\eta}{M} + \frac{1}{n}.$$

The error bound is $1 - \tau(\xi) + O(q^{d-s}) + O(1/n)$. Condition (15) keeps the origin's deficit nonnegative, and condition (16) is the real-variable form, with ξ in place of $1/D$, of the inequality $(M - \ell)D \ge Md + C - M$ that bounds the deficits from above in the proof of Proposition 3.2.

Proof. By (12), $\sum_v h_v = \ell d + C$ with every $h_v \ge 1$ and $h_0 = d > 1$, so $\ell d + C > M$; also $Md - C > \ell d$, as shown in the proof of Proposition 3.2. Since $(Md + C - M)\xi = (M - \ell)d\xi + (\ell d + C - M)\xi$, condition (16) gives $(M - \ell)(1 - d\xi - \eta) \ge (\ell d + C - M)\xi + M/n > 0$, that is $n - dz > \eta n$. Lemma 4.2 therefore supplies independent $p_0, \dots, p_{d-1}$, their span V of dimension d, and a set $\mathcal{B}$ of $b < \eta n$ bad blocks with the stated kernels.

Blocks. Let b_0 and b_1 be the numbers of bad generic and bad special blocks, so $b = b_0 + b_1$. Put $m' = \min_i |Z_i \setminus \mathcal{B}|$, so that $m' \ge z - b_1$, and choose $Z_i' \subseteq Z_i \setminus \mathcal{B}$ with $|Z_i'| = m'$ for each i. Let W be the set of good generic blocks and $G' = |W| = n - dz - b_0$. All other blocks, namely the bad blocks and the surplus good special blocks in $Z_i \setminus (\mathcal{B} \cup Z_i')$, are *discarded*. A surplus good special block has kernel $\mathbb{F}_q p_i$ on V, so it is not injective on V and cannot carry padding; discarding it keeps the special counts equal across types. Then

$$\ell G' + (\ell d + C)m' \ge \ell(n - dz - b_0) + (\ell d + C)(z - b_1) \ge \ell n + Cz - (\ell d + C)b. \tag{17}$$

Special blocks. Form the message set $S \subseteq V$ and the special tables of Proposition 3.2 with these p_i, using the blocks of Z_i' as the blocks of special type i. The elements of S have M distinct coefficient vectors, so they are M distinct codewords. The arguments for the square tables and the axis table in that proof use only independence of the p_j and the kernel $\mathbb{F}_q p_i$ on V at type i, both of which hold on Z_i'. Hence each of these blocks receives exactly ℓ distinct symbols, and codeword v has exactly $m' h_v$ agreements on $\bigcup_i Z_i'$, with h_v as in (12).

Deficits. Put $T = \lfloor (\ell G' + (\ell d + C)m')/M \rfloor$ and $z_v = T - m' h_v$. For the lower bound, $Md - \ell d - C > 0$, $m' \le z$ and $G' \ge n - dz - b$ give

$$\ell G' - (Md - \ell d - C)m' \ge \ell n - (Md - C)z - \ell b > n\Big(\ell - (Md - C)\xi - \ell\eta\Big) \ge M$$

by (15). Thus $(\ell G' + (\ell d + C)m')/M > m'd + 1$, so $T \ge m'd + 1$ and $z_v \ge T - m'd > 0$. For the upper bound, $\ell d + C - M > 0$, $m' \le z$ and $G' \ge n - dz - b$ give

$$\begin{aligned}(M - \ell)G' - (\ell d + C - M)m' &\ge (M - \ell)n - (Md + C - M)z - (M - \ell)b \\ &> n\Big(M - \ell - (Md + C - M)\xi - (M - \ell)\eta\Big) \ge M\end{aligned}$$

by (16). Thus $(\ell G' + (\ell d + C)m')/M < G' - 1 + m'$, and since $h_v \ge 1$, $z_v \le T - m' \le G' - 1$. The shortfall $\sigma = \ell G' - \sum_v z_v = \ell G' + (\ell d + C)m' - MT$ is an integer in $[0, M)$. Increase z_v by one for σ distinct codewords v; the new degrees z_v' lie in $[0, G']$ and sum to $\ell G'$.

Padding. Lemma 3.1 supplies G' subsets of S of size ℓ with degrees z'_v; assign them bijectively to the blocks of W. Each block of W is injective on V, so the evaluations of its assigned subset are ℓ distinct symbols, and the table captures exactly that subset of S. At every discarded block use any ℓ distinct symbols of $\mathbb{F}_q^s$, which exist since $q^s \geq q \geq \ell$; agreements there only lower errors.

Errors. Every codeword $v \in S$ has at least $m'h_v + z'_v \geq T$ agreements. By (17) and $b < \eta n$,

$$T > \frac{\ell n + Cz - (\ell d + C)b}{M} - 1 > n\tau(\xi) - \frac{(\ell d + C)\eta n}{M} - 1,$$

so each error is at most $1 - T/n < 1 - \tau(\xi) + (\ell d + C)\eta/M + 1/n$. □

Lemma 4.4 (Margins)**.** *Put $\xi_1 = \ell/(Md - C)$. For $0 \leq \xi \leq \xi_1$,*

$$\mu_-(\xi) = \ell - (Md - C)\xi \geq 0, \qquad \mu_+(\xi) = (M - \ell) - (Md + C - M)\xi > 0,$$

with $\mu_-(\xi) > 0$ for $\xi < \xi_1$. Moreover $d\xi_1 < 1$.

Proof. The claims on μ_- are immediate from its definition, and $d\xi_1 < 1$ follows from $Md - C > \ell d$. The function μ_+ is linear with $\mu_+(0) = M - \ell > 0$, and, with the real number $D_1 = 1/\xi_1 = (Md - C)/\ell$, $\mu_+(\xi_1) = \xi_1\big((M - \ell)D_1 - (Md + C - M)\big) \geq 7M\xi_1 > 0$ by the estimate in the proof of Proposition 3.2, which holds for every real $D \geq (Md - C)/\ell$. Hence $\mu_+ > 0$ on $[0, \xi_1]$. □

Proposition 4.5 (Folded Reed–Solomon codes)**.** *Let $\mathcal{C}$ be an s-folded Reed–Solomon code of length n and dimension $k < sn$ with appropriate evaluation points. If $z \geq 1$, $dz \leq n$ and $s(z + 1) \leq k$, then $\mathcal{C}$ has* BS_z *with respect to every family of pairwise disjoint sets $Z_0, \ldots, Z_{d-1}$ of z blocks each.*

Proof. For $i < d$ and $r \notin Z_i$, the blocks $Z_i \cup \{r\}$ carry $s(z + 1)$ unfolded points $\gamma^u \alpha_j$, which are distinct because the evaluation points are appropriate. Given any values at these points, Lagrange interpolation provides a polynomial of degree less than $s(z+1) \leq k$ taking them, that is, a message whose codeword has the prescribed blocks. □

Proof of Theorem 1.2. Let $R \in (0, R_0^*)$ be rational and put $\xi_1 = \ell/(Md - C)$. Since $C > B$ we have $BR/C < R$, and $R < R_0^*$ is equivalent to $BR/C < \xi_1$. Fix a rational ξ_* with $BR/C < \xi_* < \min\{R, \xi_1\}$, chosen as a function of (ℓ, a, R), and put

$$\varepsilon = \frac{C\xi_* - BR}{3M} > 0.$$

Since $B = a(M - \ell)/(a - 1)$, we have $\Phi_{\ell,a}(R) = 1 - (\ell + BR)/M$, so $1 - \tau(\xi_*) = \Phi_{\ell,a}(R) - 3\varepsilon$. By Lemma 4.4, $\mu_\pm = \mu_\pm(\xi_*) > 0$. Let n_0 be the least integer exceeding all of

$$\ell, \quad \frac{1}{\xi_*}, \quad \frac{1}{R - \xi_*}, \quad \frac{a + 1}{R}, \quad \frac{\ell + M}{\mu_-}, \quad \frac{2M - \ell}{\mu_+}, \quad \frac{2C + \ell d + M}{M\varepsilon};$$

it depends only on (ℓ, a, R).

Let $s \geq d + 1$ and consider a code as in the theorem, of length $n \geq n_0$ and dimension $k = Rsn$. Its sn unfolded points are distinct, so $q \geq sn \geq n \geq \ell$ and $\eta = q^{d-s} \leq q^{-1} \leq 1/n$. Put $z = \lfloor \xi_* n \rfloor$ and $\xi = z/n \in (\xi_* - 1/n, \xi_*]$. Then $z \geq 1$, $dz \leq n$ because $d\xi_* < d\xi_1 < 1$, and $s(z + 1) \leq s(\xi_* n + 1) \leq Rsn = k$ because $n(R - \xi_*) \geq 1$. Take $Z_i = \{iz, \ldots, iz + z - 1\}$; Proposition 4.5 gives BS_z. Condition (15) holds since its left side is at most $\ell - \mu_- + (\ell + M)/n \leq \ell$, and condition (16) holds since its left side is at most $M - \ell - \mu_+ + (2M - \ell)/n \leq M - \ell$. In fact $\eta n < 1$, so in Lemma 4.2

some choice of the p_i has no bad blocks at all. Theorem 4.3 provides M distinct codewords, hence M distinct messages, and a table of exactly ℓ symbols per block, with errors less than

$$1-\tau(\xi_*)+\frac{C}{Mn}+\frac{\ell d+C}{Mn}+\frac{1}{n}=\Phi_{\ell,a}(R)-3\varepsilon+\frac{2C+\ell d+M}{Mn}<\Phi_{\ell,a}(R)-2\varepsilon.$$

Moreover $(k-1)/s = Rn-1/s \geq a$ because $Rn \geq a+1$. Every such R satisfies $R < R_0^* = C\ell/(B(Md-C)) < 2\ell/(5B) \leq 2/15 < (a-1)/a$, by (10), $Md-C \geq 5M$ and $B \geq M-\ell \geq 3\ell$. For every s and every n with $Rsn \in \mathbb{Z}$, any prime power $q > sn$ and the points $\alpha_{r+1} = \gamma^{sr}$ give an appropriate code, since the exponents $sr+u$ with $0 \leq u < s$ are the distinct integers $0, \ldots, sn-1 < q-1$. Finally R_0^* is rational, and $R_0^*(2,4) = 3/196$, $R_0^*(3,3) = 19/906$ follow from $(C, B, Md-C) = (20, 56/3, 140)$ and $(38, 36, 151)$. □

We turn to random linear codes. Following [2], for integers $s, n \geq 1$ and $0 < k < sn$, a *uniformly random $\mathbb{F}_q$-linear code of rate $R = k/(sn)$ over the alphabet $\mathbb{F}_q^s$* is the row space of a uniformly random $k \times sn$ matrix Γ over $\mathbb{F}_q$, whose columns are grouped into n consecutive blocks of s; the asymptotics in n run over the lengths with $Rsn \in \mathbb{Z}$. Conditioned on full rank, Γ has a uniformly random k-dimensional row space, and full rank holds with probability $1-o(1)$; the statement below therefore holds equally for a uniformly random k-dimensional subspace of $(\mathbb{F}_q^s)^n$. Write $\rho_{\mathrm{CZ}}(R) = \Phi_{\ell,a}(R)$ for the radius of Conjecture 1.1.

Theorem 4.6 (Random linear codes)**.** *Let (ℓ, a) satisfy (4), let $q \geq \ell$ be a prime power and let $R \in (0, R_0^*)$ be rational. There are $\varepsilon > 0$ and s_0, depending only on ℓ, a, q, R, such that for every $s \geq s_0$ a uniformly random $\mathbb{F}_q$-linear code of rate R over the alphabet $\mathbb{F}_q^s$ is, with probability $1-o(1)$ as $n \to \infty$, not $(\rho_{\mathrm{CZ}}(R)-\varepsilon, \ell, \ell^a-1)$ list-recoverable.*

Proof. Choose ξ_*, ε and $\mu_\pm = \mu_\pm(\xi_*) > 0$ as in the proof of Theorem 1.2. Let $s_0 \geq d+1$ be the least integer for which $\eta_0 = q^{d-s_0}$ satisfies

$$\ell\eta_0 \leq \frac{\mu_-}{2}, \qquad (M-\ell)\eta_0 \leq \frac{\mu_+}{2}, \qquad \frac{(\ell d+C)\eta_0}{M} \leq \frac{\varepsilon}{2};$$

it depends only on (ℓ, a, q, R). Fix $s \geq s_0$, so that $\eta = q^{d-s} \leq \eta_0$, and let $k = Rsn$. Put $z = \lfloor \xi_* n \rfloor$, $\xi = z/n$, $w = s(z+1)$, and fix the deterministic sets $Z_i = \{iz, \ldots, iz+z-1\}$ for $i < d$; for large n, $z \geq 1$ and $dz \leq n$.

Full rank. The rows of Γ are independent uniform vectors in $\mathbb{F}_q^{sn}$, and the j-th lies in the span of the earlier ones with probability at most q^{j-1-sn}. Hence $\Pr[\operatorname{rank}\Gamma < k] < q^{k-sn}/(q-1) \to 0$.

Block surjectivity. The projection of the row space onto the blocks $Z_i \cup \{r\}$ is onto exactly when the $k \times w$ submatrix of Γ on their w columns has rank w. Here $w-k \leq s(\xi_* n+1) - Rsn = s - sn(R-\xi_*)$, which is negative for large n. The w columns are independent uniform vectors in $\mathbb{F}_q^k$, so, as above, this submatrix is rank deficient with probability less than $q^{w-k}/(q-1)$. A union bound over the at most dn pairs (i, r) gives

$$\Pr[\mathrm{BS}_z \text{ fails}] < dn\, q^{\,s-sn(R-\xi_*)} \to 0,$$

since $\xi_* < R$ and s is fixed.

Conclusion. With probability $1-o(1)$ the code has rate exactly R and property BS_z. For $n \geq 2M/\mu_-$ and $n \geq 2M/\mu_+$, the left sides of (15) and (16) are at most $\ell - \mu_- + \mu_-/2 + M/n \leq \ell$ and $M-\ell-\mu_+ + \mu_+/2 + M/n \leq M-\ell$. Theorem 4.3 gives $M = \ell^a$ distinct codewords and a table of exactly ℓ symbols per block, with errors less than

$$1-\tau(\xi_*)+\frac{C}{Mn}+\frac{\varepsilon}{2}+\frac{1}{n}=\rho_{\mathrm{CZ}}(R)-\frac{5\varepsilon}{2}+\frac{C+M}{Mn}<\rho_{\mathrm{CZ}}(R)-\varepsilon$$

for large n. The code is therefore not $(\rho_{\mathrm{CZ}}(R)-\varepsilon, \ell, \ell^a-1)$ list-recoverable. □

Remark 4.7 (Unfolded codes). The row space of Γ, read as a code of length sn over the alphabet $\mathbb{F}_q$, is a uniformly random $\mathbb{F}_q$-linear code of rate R in the usual unfolded sense. A folded table with ℓ symbols per block induces an unfolded table: at the u-th coordinate of block r allow the u-th entries of the symbols in $\mathcal{T}_r$, at most ℓ values. A folded agreement gives s unfolded agreements, so unfolded errors are at most folded errors. With $s = s_0$, Theorem 4.6 therefore gives the same conclusion for uniformly random $\mathbb{F}_q$-linear codes over $\mathbb{F}_q$ itself, along lengths divisible by s_0 and every fixed prime power $q \geq \ell$; for $q \geq \ell^a$ this range is genuinely new, since smaller alphabets already admit averaging arguments. The bound $q \geq \ell$ is the one needed to choose the ℓ-set $A \subseteq \mathbb{F}_q$.

Remark 4.8 (The black-box route). For large fields, $q = \Theta(sn)$, failure of random linear codes also follows from Theorem 1.2 through known transfer theorems. Brakensiek, Chen, Dhar and Zhang [2, Theorems 1.4 and 3.1] show that if random $\mathbb{F}_q$-linear codes of rate R avoid a b-local profile with high probability, then explicit s-folded Reed–Solomon codes with $q = \Theta(sn)$, $s \geq b(R/\varepsilon' + 1)$ and rate $R - \varepsilon'$ avoid it as well; list recovery with output cap L corresponds to a family of $(L+1)$-local profiles [2, Section 1.1]. Applied in contrapositive form to the configurations of Theorem 1.2, which covers the explicit codes of [2] since it holds for all appropriate evaluation points, this shows that random linear codes of slightly larger rate are not free of them with high probability, and the sharp-threshold theorem of Levi, Mosheiff and Shagrithaya [6, Theorem 4.4], quoted as [2, Theorem 2.4], upgrades this to containment with high probability at rates slightly above. The reverse reduction [1, Theorem 2.5], [2, Corollary 4.3] transfers positive list-recovery guarantees from subspace designs to random linear codes and plays no role here. Theorems 4.3 and 4.6 add an elementary, self-contained proof; a statement for every code with block surjectivity, in particular for folded Reed–Solomon codes with arbitrary appropriate evaluation points; fixed fields of every size $q \geq \ell$; and the explicit threshold R_0^* with explicit slack.

5 Proof of Theorem 1.3

Proposition 3.2 also applies with $D = 80$; here a direct pair schedule with explicit multiplicities replaces Lemma 3.1.

Proof of Theorem 1.3. Apply Lemma 2.1 with $D = 80$, $d = 10$ and m divisible by $84 \cdot 79$. Set $n = 80m$ and

$$k = \frac{80sm}{79}, \qquad V = \operatorname{span}\{p_0, \dots, p_9\}.$$

Then $k - sm = sm/79 \geq 1$, so these polynomials belong to the message space, and $k/(sn) = 1/79$. Use the sixteen messages

$$S = \{0\} \cup \bigcup_{h=0}^{4} \{p_{2h}, p_{2h+1}, p_{2h} + p_{2h+1}\}. \tag{18}$$

They are distinct because the ten axes are independent. Encoding is injective, so their codewords are distinct too.

At a block $r = i + 80j$ with $0 \leq i < 10$, let i' be the other index in its pair $\{2h, 2h+1\}$ and choose the two-symbol table

$$\mathcal{T}_r = \{0, E_r(p_{i'})\}.$$

The second symbol is nonzero by the axis-kernel statement. Within S, its zero fiber consists exactly of $0, p_i$, and its other chosen fiber consists exactly of $p_{i'}, p_i + p_{i'}$. Indeed evaluation equality is equivalence modulo the line $\mathbb{F}_q p_i$, and independent coordinates exclude any other message. Thus

the origin agrees at all $10m$ special blocks. Each of the fifteen other messages agrees at exactly the two types of its own square, contributing $2m$ agreements.

There remain $70m$ blocks, all injective on V. At each such block a table made from the evaluations of two distinct members of S captures exactly those two members. Assign unordered pairs of members to these blocks using the following multiplicities:

pair type	number of pairs	multiplicity per pair
origin and a nonzero message	15	$m/12$
two nonzero messages	105	$55m/84$.

All multiplicities are integral. Their sum is

$$\frac{15m}{12} + 105\frac{55m}{84} = 70m,$$

so assign the pair occurrences bijectively to the generic blocks and use their two evaluations as the tables. The origin gains $5m/4$ agreements. Each nonzero message gains

$$\frac{m}{12} + 14\frac{55m}{84} = \frac{37m}{4}.$$

Every message therefore has exactly

$$10m + \frac{5m}{4} = 2m + \frac{37m}{4} = \frac{45m}{4}$$

agreements, or agreement fraction $9/64$ out of $80m$ blocks. Every individual error is $55/64$.

To obtain the field and length quantifiers, fix $s \geq 10$ and put $H = 80s \cdot 84 \cdot 79$. Choose a prime π not dividing H, and for integers $\lambda \geq 1$ take

$$q = \pi^{\varphi(H)\lambda}.$$

Euler's theorem gives $H \mid q-1$. Thus $m = (q-1)/(80s)$ is divisible by $84 \cdot 79$, and the construction applies. As $\lambda \to \infty$ its lengths tend to infinity, with s fixed. The argument works for any choice of generator. The unfolded points comprise all $q - 1 = sn$ nonzero field elements, so the evaluations are appropriate and $q > sn > k$. Also $(k-1)/s \geq 4$ for sufficiently large length.

At rate $1/79$,

$$\Phi_{2,4}(1/79) - \frac{55}{64} = \frac{7}{8} - \frac{7}{6 \cdot 79} - \frac{55}{64} = \frac{13}{15168} > \frac{2}{10000}.$$

Hence with $\rho_\circ = 55/64 + 1/10000$ all sixteen messages have error below $\rho_\circ < \Phi_{2,4}(1/79) - 1/10000$, and the codes are not $(\rho_\circ, 2, 15)$ list recoverable. Since s may first be chosen above any finite folding threshold and then λ taken arbitrarily large, this contradicts Conjecture 1.1 at fixed positive slack. □

For the rate-slack reformulation [1, Conjecture 1.3], the threshold at radius ρ is $R^*(\rho) = \frac{3}{4}(1 - 8\rho/7)$, and

$$R^*(\rho_\circ) - \frac{1}{79} = \frac{6}{7}\left(\frac{13}{15168} - \frac{1}{10000}\right) > \frac{1}{10000}.$$

Thus the example also lies strictly inside that formulation's conjectured region.

6 Scope of the results

The counterexamples hold at every rational rate in $(0, R_0^*)$, with $R_0^*(2,4) = 3/196$ and $R_0^*(3,3) = 19/906$, in every folded Reed–Solomon code with appropriate evaluation points, folding $s \geq d+1$ and length $n \geq n_0(\ell, a, R)$; in consecutive full-length codes, at the rates covered by Corollary 3.3, the folding $s = d$ suffices and all errors are exactly equal. The same tradeoff fails in every linear code with block surjectivity at the appropriate proportion, and, with high probability, in uniformly random $\mathbb{F}_q$-linear codes for every fixed prime power $q \geq \ell$ and folding $s \geq s_0(\ell, a, q, R)$. The sixteen-message instance has slack $13/15168$. The rate range is set by the condition $(Md - C)\xi < \ell$ on the proportion ξ of each special type, which keeps the origin's deficit nonnegative; together with the requirement $C\xi > BR$ for a violation, it gives exactly $R < R_0^*$. The region (4) consists exactly of the pairs (ℓ, a) with $C > B$; for $a = 2$ and for $(\ell, a) = (2, 3)$ the square configuration gives $C = B$, and the case $(\ell, a) = (2, 2)$ of the conjecture is proved in [4, Theorem 1.8] (see [1, Remark 1.7]).

Assistance and data availability

AI-assisted tools were used in mathematical exploration and manuscript preparation. The mathematical arguments are contained in this article. The constructions use no numerical experiment or external dataset. As an independent check, exact computations carried out the proof of Theorem 4.3 step by step on folded Reed–Solomon codes, with all errors computed from raw polynomial evaluations; for instance, at $(\ell, a) = (3, 3)$, $s = 8$, $n = 1377$, $z = 27$ and $k = 226$, over $q = 11027$ with consecutive points and over $q = 33049$ with random appropriate points, conditions (15) and (16) hold and all 27 codewords had error at most 1186/1377, below $\Phi_{3,3}(k/(sn))$. These computations are not used in any proof.

References

[1] J. Brakensiek, Y. Chen, M. Dhar and Z. Zhang, *Combinatorial Bounds for List Recovery via Discrete Brascamp–Lieb Inequalities*, full version of the STOC 2026 extended abstract, arXiv:2510.13775v2, 11 October 2026. `https://arxiv.org/abs/2510.13775v2`.

[2] J. Brakensiek, Y. Chen, M. Dhar and Z. Zhang, *From Random to Explicit via Subspace Designs With Applications to Local Properties and Matroids*, arXiv:2510.13777v1, 15 October 2025. `https://arxiv.org/abs/2510.13777v1`.

[3] J. Brakensiek, Y. Chen, A. Putterman and Z. Zhang, *Bounds and Limitations on Codes Achieving List Recovery Capacity*, full version of the SODA 2027 extended abstract, arXiv:2607.19576v2, 9 October 2026. `https://arxiv.org/abs/2607.19576v2`.

[4] Y. Chen and Z. Zhang, *Explicit Folded Reed–Solomon and Multiplicity Codes Achieve Relaxed Generalized Singleton Bounds*, STOC 2025; arXiv:2408.15925v3, 14 April 2025. `https://arxiv.org/abs/2408.15925v3`.

[5] V. Guruswami and A. Rudra, *Explicit codes achieving list decoding capacity: error-correction with optimal redundancy*, IEEE Trans. Inform. Theory **54** (2008), 135–150.

[6] M. Levi, J. Mosheiff and N. Shagrithaya, *Random Reed–Solomon Codes and Random Linear Codes are Locally Equivalent*, arXiv:2406.02238v7, 14 November 2025. `https://arxiv.org/abs/2406.02238v7`.

[7] S. Schubr, *Independent Squares and Average-Radius Obstructions for Linear Codes*, Luminary Research Archive, edition 2, 27 October 2026. `https://research.luminaryoddities.org/independent-squares-average-radius-obstructions-linear-codes/v/2`.