The Syndrome-Space Lens

A Complete Resolution of Proximity Gaps for Reed-Solomon Codes

Russell Okamoto
arithmocene@proton.me

September 20, 2025

Abstract

We resolve the Correlated Agreement (CA) problem for Reed-Solomon codes up to the information-theoretic capacity limit by introducing a fundamental change of basis: from the traditional evaluation domain to the *syndrome space*. Viewed through this "Syndrome-Space Lens," the problem of proximity testing transforms into a transparent question of linear-algebraic geometry: a single affine line of syndromes traversing a family of low-dimensional subspaces. This new perspective makes a sharp phase transition at the capacity boundary visible, allowing for a complete characterization of the problem's behavior across all parameter regimes, yielding short, self-contained proofs.

Classification. We establish a precise trichotomy organized by the rank margin $\Delta := t - d$. At the capacity boundary ($\Delta = 0$), the CA premise is information-theoretically *vacuous*, and we prove that no rigidity can be concluded without imposing additional structure. One step beyond capacity ($\Delta = 1$), the problem enters a "knife-edge" regime where unconditional rigidity does not hold; soundness is recovered either through a combinatorial *witness* (such as a repeated error support or a small union of supports) or by adding protocol-level structure (such as independent two-fold MCA checks, DEEP/STIR out-of-domain sampling, or a global error locator). For stricter gaps ($\Delta \geq 2$), *unconditional* rigidity holds under a simple algebraic condition ($(r+1)k < m+1$), with explicit quantitative bounds.

MCA and Practical Implications. Below capacity ($\delta < 1 - \rho$), the strengthened *mutual* correlated agreement (MCA) problem reduces to ordinary correlated agreement. MCA holds under the same hypotheses as CA. When folds are generated with *independent* challenges (e.g., via domain-separated Fiat-Shamir), the per-round security margins add. The model-scoped soundness law is $\Pr[\mathrm{FA}] \leq q^{-(\sum \Delta_i)s}$, providing a clear and complete rulebook for selecting safe and efficient parameters in FRI/STARK systems. This work bypasses the complex machinery of list-decoding algorithms entirely and resolves the long-standing open problem concerning the gap between the Johnson bound and capacity.

Contents

1 Responsible Disclosure & Community Advisory

Why are we publishing now rather than waiting for the Proximity Prize launch? This work studies core assumptions used by hash-based Reed-Solomon proximity testing (CA/MCA) and clarifies proximity gap boundaries, with explicit, reviewable parameter floors for optimal, sound, and non-vacuous operation. We contacted proximityprize@ethereum.org three weeks ago and learned there was no target start date for the prize program at that time. Given the **potential**

benefits and caveats suggested by these results, we are releasing this preprint prior to any prize launch to spur **independent verification** and to provide **potentially useful, immediately actionable guidance** for practitioners. Our aim is to inform the community in good faith and reduce uncertainty—not to speculate about specific deployments. Community corroboration, refutation, and improvement are not only welcome but essential and urgent.

2 Introduction

2.1 The Role of Proximity Testing in Modern Cryptography

The relentless pursuit of computational integrity in decentralized systems has led to the development of sophisticated cryptographic proof systems. Among the most prominent are STARKs (Scalable Transparent Arguments of Knowledge) and the underlying FRI (Fast Reed-Solomon Interactive Oracle Proof of Proximity) protocol [2, 3]. These systems enable a computationally weak verifier to check the correctness of a massive computation performed by a powerful, untrusted prover, with security guarantees that are transparent (requiring no trusted setup) and often post- quantum. At their core, these systems rely on a technique called *arithmetization*, which transforms a claim about a computation into a claim about the properties of a low-degree polynomial. The verifier's task is then reduced to checking that a function provided by the prover is indeed the evaluation of such a polynomial. Since reading the entire function is infeasible, the verifier performs a *proximity test*, querying the function at a small number of random points to gain confidence that it is "close" to a valid low-degree polynomial (i.e., a codeword of a Reed-Solomon code). The soundness of the entire proof system rests on the soundness of this proximity test.

2.2 The Correlated Agreement Problem

The FRI protocol achieves its remarkable efficiency through a recursive process. In each step, a claim about one large low-degree polynomial is reduced to a claim about a smaller low-degree polynomial. This reduction step, known as "folding," is where the central soundness issue arises. The prover provides two functions f and g, and the verifier combines them using a random challenge $z \in \mathbb{F}_q$ to form the test function $h_z := f + z\,g$, which is then checked for proximity to a Reed-Solomon codeword. This procedure is repeated for many independent challenges z.

Correlated Agreement (CA). CA formalizes the soundness of this step: if, for a large fraction of random challenges z, the folded word h_z is close to *some* Reed-Solomon codeword, then f and g themselves must be close to *correlated* codewords—namely, there exist codewords $P, Q \in \mathrm{RS}_{n,d}(D)$ such that f is close to P, g is close to Q, and P and Q agree on a large set of positions. Consequently, a prover that succeeds on many challenges necessarily exhibits global low-degree structure rather than arbitrary, unstructured behavior.

2.3 A Brief History: The Johnson Bound as a Theoretical Barrier

The analysis of proximity tests for Reed-Solomon codes is deeply intertwined with the theory of list decoding. Classical unique decoding corrects errors up to half the minimum distance of the code. *List decoding*, proposed in the 1950s by Elias and Wozencraft, allows the decoder to output a small list of candidate codewords, enabling error correction from a much larger number of errors. A fundamental combinatorial limit in this area is the **Johnson bound** [7]. It specifies a radius of agreement, larger than the unique decoding radius, within which the number of codewords in a Hamming ball around any received word is guaranteed to be small. For decades, this bound represented a theoretical barrier for provable soundness. The landmark work of Guruswami and Sudan provided the first efficient algorithm to list-decode RS codes up to the Johnson bound [5], and subsequent research on FRI soundness, culminating in the work of Ben-Sasson, Kopparty, and Saraf (BKS), established CA soundness up to this same bound [3]. These analyses, while powerful, were fundamentally tied to the list-decoding paradigm. They operated in the function-evaluation domain and involved intricate arguments about the behavior of decoding algorithms on formal polynomials [5, 3]. The question of what happens in the gap between the Johnson bound and the code's information-theoretic capacity remained a major open problem.

2.4 Our Contribution: A Paradigm Shift

This paper resolves the CA and Mutual Correlated Agreement (MCA) behavior up to and at the information-theoretic capacity limit by introducing a fundamental change of basis. We move from the traditional evaluation domain to the *syndrome space* and organize the analysis by the *rank margin*

$$\Delta := t - d\,,$$

where t is the agreement threshold, d is the code dimension, m is the number of parity checks, and k is the error budget. Rather than the traditional decoding/list-size view, we study the affine syndrome line $s(z) = A + zB$ and its incidences with column spans $U_T = \mathrm{Span}(H_T) \subseteq \mathbb{F}_q^m$. This **syndrome-space lens** makes a sharp phase transition visible and yields short, elementary linear-algebraic proofs that bypass list decoding entirely.

Table 1: A New Perspective: Classical View vs. Syndrome-Space View

Concept	**Classical View (Evaluation Domain)**	**Syndrome-Space View (Our Approach)**
Objects	Functions $f, g, h_z \in \mathbb{F}_q^n$	Vectors $A, B, s(z) \in \mathbb{F}_q^m$
Structure	Low-degree polynomials (codewords)	Codewords have *zero syndrome*; structure is subspace incidence in $\mathbb{F}_q^m$
Proximity	Large agreement with a codeword	Line $s(z)$ lies in a low-dimensional subspace U_T
Protocol	Family of folded functions $\{h_z\}_{z\in\mathbb{F}_q}$	Affine line $s(z) = A + zB$ in syndrome space
Analysis Tool	List-decoding algorithms; Johnson bound	Linear algebra; MDS property

At its core, the rank margin Δ provides a "Rosetta Stone," translating the specialized parameters of a proximity protocol into the universal language of linear algebra. A *trichotomy* of behaviors is observed corresponding directly to the well-understood states of a system of linear equations, where m is the number of equations (constraints) and k is the number of variables (unknowns).

- $\Delta = 0 \iff$ **Effectively underconstrained:** A system with an equal number of equations and unknowns ($m = k$). In an adversarial setting, this balance guarantees that a non-trivial solution (a kernel event) can always be found. The system has no constraining power and is thus vacuous.

- $\Delta = 1 \iff$ **Critically constrained:** A system with exactly one more equation than unknowns ($m = k + 1$). The system is on a knife-edge, neither fully constrained nor vacuous.

Soundness is fragile.

- $\Delta \geq 2 \iff$ **Overconstrained:** A system with at least two more equations than unknowns ($m \geq k+2$). The system is rigid, and finding a solution by chance is statistically negligible. This is the regime of robust soundness.

This mapping from Δ to the state of a linear system is the foundational intuition behind the Syndrome-Space Lens.

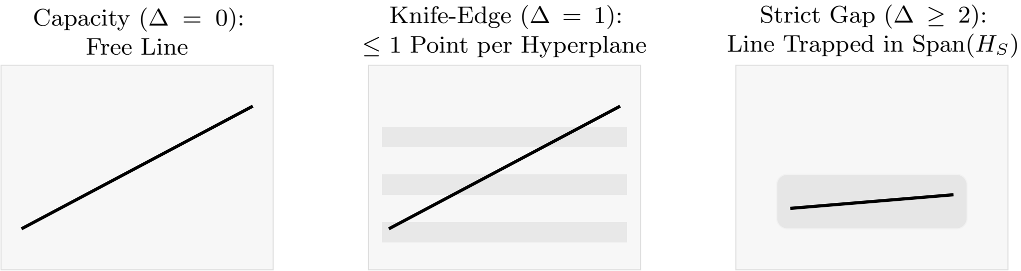
Capacity ($\Delta = 0$): Free Line
Knife-Edge ($\Delta = 1$): ≤ 1 Point per Hyperplane
Strict Gap ($\Delta \geq 2$): Line Trapped in Span(H_S)

Figure 1: Syndrome–space intuition in $\mathbb{F}_q^m$: the affine line $s(z) = A + zB$ is unconstrained at capacity, meets each hyperplane in at most one point when $\Delta = 1$, and is trapped in a fixed low-dimensional span when $\Delta \geq 2$.

The rank-margin (syndrome-space) lens yields a complete, geometric understanding of proximity gaps up to capacity and provides concrete guidance for safely unlocking substantial efficiency gains in deployed proof systems.

Summary of Results (in rank-Δ language).

- **Capacity boundary ($\Delta = 0$; $t = d$).** The CA premise holds for *any* pair of functions (Theorem 5.1), meaning the test carries no soundness information. More strongly, no small-span rigidity conclusion can be drawn at $\Delta = 0$ without additional assumptions (Theorem 5.2); we provide probabilistic and explicit small-field witnesses to this impossibility.

- **Knife-edge ($\Delta = 1$; $k = m - 1$).** A sharp line–hyperplane dichotomy holds (Theorem 6.1): either the syndrome line collapses into a single $(m-1)$-span (rigidity) or a success can occur in a given $(m-1)$-span at most once (no global structure). A combinatorial *witness* (e.g., a repeated error support or a small union of supports where $|T_{z_i} \cup T_{z_j}| \leq m-1$) is sufficient to force rigidity (Theorem 6.2). Otherwise, additional protocol-level *structure* (such as independent MCA folds, DEEP/STIR out-of-domain sampling, or a global error locator) is required to restore soundness.

- **Strict gap ($\Delta \geq 2$; $k \leq m-2$).** A *vanishing rth differences* identity (Theorem 7.1) combined with a double-counting bound (Theorem 7.2) establishes *unconditional* rigidity under the simple quantitative condition
$$(r+1)k < m+1\,.$$
In parallel, a *local* combinatorial witness (such as a repeat support or a small union of supports) also suffices to trap the entire line in a single span (Theorem 6.2). Together, these results provide a complete picture: $\Delta = 0$ is vacuous; $\Delta = 1$ is a conditional knife-edge; and $\Delta \geq 2$ is rigid by geometry alone.

- **Restoring rigidity at $\Delta = 0$ (structure).** We classify the geometric remedies that suffice to restore soundness at the equality boundary: an *algebraic geometric remedy* via bundle constancy (Lemma 5.4), a *combinatorial geometric remedy* via a sunflower trap (Lemma 5.5), and protocol-level enforcement (e.g., committing to a single error set). For completeness, classical variants are recorded in Appendices E.1 and E.2.
- **Implications for MCA and STARKs (rank-exponent budgeting).** For $\Delta > 0$ ($t > d$), MCA follows under the *same* conditions as CA. With *independent* folds per round (achieved via domain-separated Fiat-Shamir over fully bound transcripts), rank margins *add* (Theorem 7.3). If the per-fold margins are Δ_1, Δ_2, then the total per-round exponent is $\Delta_{\text{tot}} = \Delta_1 + \Delta_2$, and the model-scoped soundness budget is $q^{-\Delta_{\text{tot}} s}$. This provides a clear, auditable rulebook for sizing safe near-capacity parameters in FRI/STARK systems.

3 Mathematical Primer

Equality-case notation. We call the setting $k = m$ (equivalently, $t = d$) the *equality case*: the error budget equals the number of parity checks, and the agreement threshold equals the code dimension. For a Reed-Solomon code $[n, d, d_{\min}]_q$, we use d for the code dimension and $d_{\min} = n - d + 1$ for the minimum distance. To avoid ambiguity, we will always write $d_{\min}$ when referring to distance. Table 2 collects the symbols used throughout the paper.
We use the shorthand $[n] := \{1, \ldots, n\}$ for the index set of coordinates.

Field and challenges. We assume only that $q \geq n$ and that whenever an r-fold step is used, the verifier works with $r+1$ *pairwise-distinct* challenges. If challenges are sampled with replacement, we simply resample collisions; among M draws the collision probability is at most $\binom{M}{2}/q$, which is negligible at cryptographic sizes. No assumption on $\mathrm{char}(\mathbb{F}_q)$ is needed in our arguments. For a subset $T \subseteq [n]$, we write H_T for the $m \times |T|$ submatrix of H with columns indexed by T.

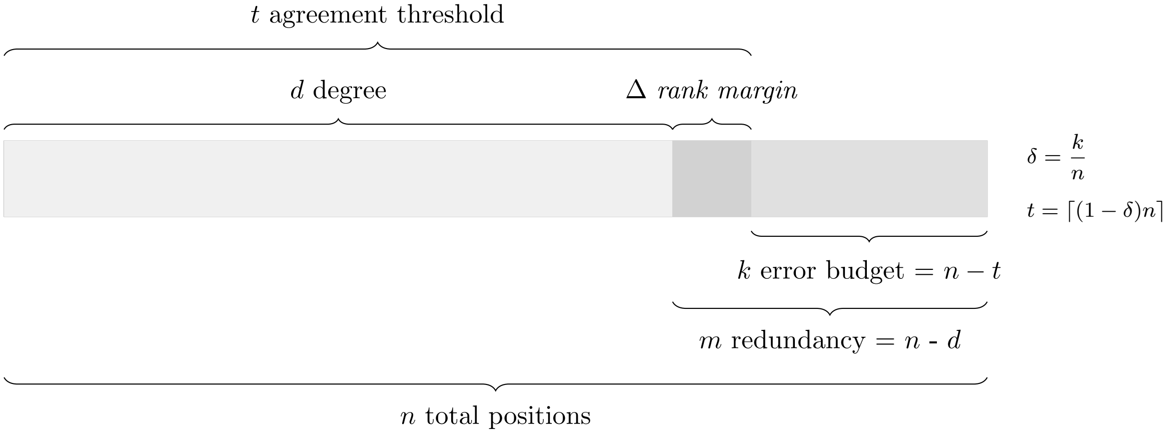

Figure 2: Single–row schematic of agreement vs. errors (length n). The verifier requires at least t agreements and tolerates up to $k = n - t$ errors. The code dimension is d, so the redundancy is $m = n - d$, and the *rank margin* is $\Delta = t - d$. The darkest shaded band highlights Δ, i.e. the gap between the degree threshold d and the agreement threshold t; the lighter regions illustrate the remaining portions. The normalized error rate is $\delta = k/n$, hence $t = \lceil (1-\delta)n \rceil$. This makes explicit how the parameters (n, d, t, m, k, Δ) interrelate.

Table 2: Notation and Definitions

Symbol	Meaning
n	Code length.
q	Field size ($\mathbb{F}_q$ is the finite field with q elements).
$D \subseteq \mathbb{F}_q$	Evaluation domain (set of n points), $\lvert D\rvert = n$.
d	Degree bound / dimension of the RS code.
ρ	Code rate, $\rho = d/n$.
m	Redundancy (number of parity checks), $m = n - d$.
$d_{\min}$	Minimum distance of $\mathrm{RS}_{n,d}$; for RS, $d_{\min} = n - d + 1$.
$\mathrm{RS}_{n,d}(D)$	Reed-Solomon code of length n and dimension d over domain D.
H	An $m \times n$ parity-check matrix for the RS code; H_T is the submatrix on columns T.
U_T	Syndrome subspace for $T \subseteq [n]$: $U_T = \mathrm{Span}(H_T) \subseteq \mathbb{F}_q^m$.
f, g	Functions $D \to \mathbb{F}_q$ provided by the prover (viewed as length-n vectors).
h_z	Folded word for challenge z: $h_z := f + z\,g$.
A, B	Syndromes of f and g: $A = Hf^T$, $B = Hg^T$.
$s(z)$	Affine syndrome line: $s(z) = A + z\,B$.
t	Agreement threshold (number of positions h_z must agree with some codeword).
k	Error budget per challenge, $k = n - t$ (max. errors in h_z).
Δ	Rank margin from capacity, $\Delta := m - k = t - d$ (so $\Delta = 0$ at $t = d$).
Δ_i	Per-fold margins in MCA.
r	Order in the r-fold (vanishing-differences) argument (uses r+1 points).
s	Number of independent repetitions (rounds) in budgeting; Pr[false accept] $\le q^{-\Delta s}$.
δ	Proximity / error fraction used in CA/MCA: $\delta := k/n$.
$\delta_{\min}$	Relative distance: $\delta_{\min} := d_{\min}/n = 1 - \rho$.
$\mathcal{Z}^*$	Set of successful challenges (indices z where the CA premise holds).
e_z	Error vector for challenge z; $e_z := h_z - p_z$.
p_z	A codeword in $\mathrm{RS}_{n,d}(D)$ agreeing with h_z outside the error support T_z.
T_z	Error support set for challenge z: $T_z = \mathrm{supp}(e_z)$.

Sizing parameters. Throughout, the operational sizing is governed by the rank margin $\Delta := m - k = t - d$, the number of independent rounds s, and the field size via $\log_2 q$. Our *model-scoped* budgeting law (for a per–round *effective* margin $\Delta_{\text{eff}} \ge 1$ and *independent* rounds) is

$$\Pr[\mathrm{FA}] \;\le\; q^{-\Delta_{\text{eff}} s} \quad\Longleftrightarrow\quad \Delta_{\text{eff}} \cdot s \cdot \log_2 q \;\ge\; \lambda_{\text{sec}}.$$

(We retain $\delta := k/n$ in the notation table for completeness, but the analysis and sizing rules are expressed in terms of Δ. For operational parameter floors we include a conservative +2-bit slack as spelled out in Appendix B.)

3.1 Reed-Solomon Codes and the MDS Property

Fix a finite field $\mathbb{F}_q$ and an evaluation set (domain) $D = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{F}_q$ of size n. The Reed-Solomon code $\mathrm{RS}_{n,d}(D)$ consists of all vectors $(p(\alpha_1), \ldots, p(\alpha_n)) \in \mathbb{F}_q^n$, where $p(x)$ is a univariate polynomial of degree $< d$. Key parameters include: **dimension** d and **minimum distance** $d_{\min} = n - d + 1$. RS codes are **Maximum Distance Separable (MDS)**: they achieve the Singleton bound $d_{\min} \le n - d + 1$ with equality.

Lemma 3.1. *Let $H \in \mathbb{F}_q^{m\times n}$ be a parity-check matrix for an MDS code (such as a Reed-Solomon code). For every $T \subseteq [n]$ with $|T| = m$, the submatrix H_T is invertible. In particular, the columns of H_T are linearly independent and $U_T = \mathrm{Span}(H_T) = \mathbb{F}_q^m$.*

Proof. This is a standard property of MDS codes. A code is MDS if and only if every $m \times m$ submatrix of its parity-check matrix H is invertible. For a concrete example with RS codes, one can construct H from a truncated Vandermonde matrix. Fix $D = \{\alpha_1, \dots, \alpha_n\} \subset \mathbb{F}_q$. Take $H \in \mathbb{F}_q^{m\times n}$ with $H_{j,i} = \alpha_i^{j-1}$ for $j = 1, \dots, m$ and $i = 1, \dots, n$. For any $T = \{i_1, \dots, i_m\} \subseteq [n]$ with corresponding evaluation points $\gamma_\ell := \alpha_{i_\ell}$, the submatrix H_T is a standard Vandermonde matrix whose determinant is $\prod_{1\le j<\ell\le m}(\gamma_\ell - \gamma_j) \neq 0$, since the evaluation points are distinct. Hence H_T is invertible. □

This linear-algebraic fact is central: for any set T of m coordinates, the corresponding columns of H form a basis of $\mathbb{F}_q^m$. Equivalently, every syndrome $S \in \mathbb{F}_q^m$ can be written as $S = He_T^T$ for some vector e_T supported on T.

3.2 Syndromes and the CA premise

Let $H \in \mathbb{F}_q^{m\times n}$ be a parity-check matrix for $\mathrm{RS}_{n,d}(D)$; by definition c is a codeword iff $Hc^T = 0$. For any vector $w \in \mathbb{F}_q^n$, its **syndrome** is $\mathrm{Syn}(w) := Hw^T \in \mathbb{F}_q^m$. For the prover's functions $f, g : D \to \mathbb{F}_q$, write $A := Hf^T$ and $B := Hg^T$. For a challenge $z \in \mathbb{F}_q$, the folded word is $h_z := f + z\,g$ with syndrome

$$s(z) = A + zB\,,$$

which traces an *affine syndrome line* parameterized by z. The CA premise states: for each challenge z in some large set $\mathcal{Z}^* \subseteq \mathbb{F}_q$, there exists an error vector e_z of weight at most $k = n - t$ such that $h_z - e_z \in \mathrm{RS}_{n,d}(D)$. This is equivalent to

$$H(h_z - e_z)^T = 0\,,$$

If we let $T_z := \mathrm{supp}(e_z)$ and $U_{T_z} := \mathrm{Span}(H_{T_z})$, this condition is precisely

$$s(z) \in U_{T_z} \qquad \text{for each } z \in \mathcal{Z}^*\,.$$

In words: the line $\{A + zB : z \in \mathcal{Z}^*\}$ in syndrome space is covered by a union of low-dimensional subspaces U_T, where each subspace is spanned by at most k columns of H. Figure 3 illustrates this core geometric concept. See Appendix G, for a miniature end-to-end Reed-Solomon example.

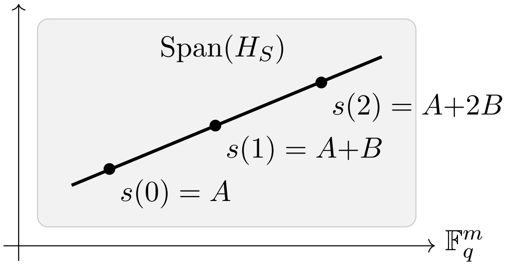

Figure 3: The affine syndrome line $s(z) = A + zB$ traversing syndrome space. The shaded region indicates a subspace $\mathrm{Span}(H_S)$ for some index set S. The CA premise implies that for successful challenges z, the point $s(z)$ lies in such a subspace for some small set S.

3.3 Auxiliary tools: Divided Differences, Hankel Lifts, and Johnson Radius

This subsection collects three tools used repeatedly: (i) divided differences and annihilation weights, (ii) Hankel matrices for capturing algebraic structure, and (iii) a reminder of the Johnson radius versus capacity.

Divided differences and annihilation weights. For any set of $r+1$ pairwise-distinct points $z_0, \dots, z_r \in \mathbb{F}_q$, there exist nonzero coefficients, or *annihilation weights*, $\lambda_0, \dots, \lambda_r$ such that for any polynomial $p(z)$ of degree less than r, we have

$$\sum_{i=0}^{r} \lambda_i p(z_i) = 0\,.$$

Lemma 3.2 (Characteristic-free annihilator)**.** *Let $z_0, \dots, z_r \in \mathbb{F}_q$ be pairwise distinct and let $V \in \mathbb{F}_q^{r\times(r+1)}$ be the Vandermonde matrix with entries $V_{j,i} = z_i^j$ for $j = 0, \dots, r-1$ and $i = 0, \dots, r$. Then* $\operatorname{rank}(V) = r$ *and* $\dim \operatorname{Ker}(V) = 1$. *For any nonzero vector $\lambda = (\lambda_0, \dots, \lambda_r)^T \in \operatorname{Ker}(V)$, we have*

$$\sum_{i=0}^{r} \lambda_i \, p(z_i) = 0 \qquad \textit{for every polynomial } p \textit{ with } \deg p < r\,.$$

Proof. The distinctness of the z_i implies the r rows of V are linearly independent, so $\operatorname{rank}(V) = r$ and the right kernel is one-dimensional. By definition, any $\lambda \in \operatorname{Ker}(V)$ satisfies $\sum_{i=0}^{r} \lambda_i z_i^j = 0$ for all $j = 0, \dots, r-1$. Since any polynomial p with $\deg p < r$ is a linear combination of the monomials $1, z, \dots, z^{r-1}$, the claim follows by linearity. □

Applying this coordinate-wise to the affine vector-valued map $s(z) = A + zB$ gives an immediate and crucial corollary.

Corollary 3.3 (Annihilation for affine lines)**.** *Let $z_0, \dots, z_r \in \mathbb{F}_q$ be pairwise distinct with $r \geq 2$. For any $A, B \in \mathbb{F}_q^m$, if $s(z) = A + zB$ and $\lambda \in \operatorname{Ker}(V) \setminus \{0\}$ as in Lemma 3.2, then $\sum_{i=0}^{r} \lambda_i s(z_i) = 0$.*

Proof. Apply Lemma 3.2 to the constant polynomial $p(z) = 1$ and the linear polynomial $p(z) = z$ (since $r \geq 2$). This gives $\sum_i \lambda_i = 0$ and $\sum_i \lambda_i z_i = 0$, respectively. Expanding $\sum_i \lambda_i (A + z_i B)$ yields $A(\sum_i \lambda_i) + B(\sum_i \lambda_i z_i) = A \cdot 0 + B \cdot 0 = 0$. □

Hankel lifting and matrix pencils. To analyze the algebraic structure of a syndrome vector $V = (v_0, \dots, v_{m-1})$, we use the *Hankel lifting* operator $\mathcal{C}_k$, which maps V to an $(m-k) \times (k+1)$ matrix whose anti-diagonals are constant: $(\mathcal{C}_k(V))_{i,j} = v_{i+j-2}$. A key fact from coding theory is that if V is the syndrome of an error vector of weight at most k, then $\operatorname{rank}(\mathcal{C}_k(V)) \leq k$. This is because the coefficients of the error- locator polynomial lie in the right kernel of this matrix. For the affine line of syndromes $s(z) = A + zB$, we can form the *Hankel matrix pencil* $M(z) := \mathcal{C}_k(A) + z\mathcal{C}_k(B)$. The rank of this pencil at different values of z reveals information about the shared structure of the underlying error supports.

Johnson radius vs. capacity. For an RS code of rate $\rho = d/n$, the classical Johnson bound guarantees a polynomially bounded list of candidate codewords whenever the agreement fraction exceeds $\sqrt{\rho}$ [7, 3]. In terms of agreement count, this is $t_J = n\sqrt{\rho} = \sqrt{nd}$. The soundness of FRI was proven up to this bound [3]. By contrast, the information-theoretic capacity limit for decoding corresponds to an agreement of $t = d$. The region between $t = d$ and $t = t_J$ was a long- standing zone of uncertainty. Our analysis resolves the behavior across this entire gap.

4 The Δ-Trichotomy: A Geometric Overview

The behavior of the Correlated Agreement problem undergoes a sharp phase transition, which becomes visible when organized by the rank margin

$$\Delta := t - d\,,$$

This parameter measures how far the agreement threshold t is from the code dimension d, or equivalently, how much smaller the error budget k is than the number of parity checks m. There are three distinct regimes: *capacity* ($\Delta = 0$), *knife-edge* ($\Delta = 1$), and *strict gap* ($\Delta \geq 2$). The geometric intuition in the syndrome space for each regime is as follows.

Capacity ($\Delta = 0$): Vacuity. At the capacity boundary, the error budget equals the number of parity checks ($k = m$). By the MDS property, any set of m columns of the parity-check matrix H spans the entire syndrome space $\mathbb{F}_q^m$. This means that for any syndrome vector $s \in \mathbb{F}_q^m$, we can always find an error vector of weight m that produces it. Consequently, for any folded word h_z, its syndrome $s(z)$ can always be explained by a weight-m error. The CA premise is always satisfied, imposing no restriction on the prover. The syndrome line $s(z) = A + zB$ is *free to roam* the ambient space, and no structural conclusion can be drawn.

Knife-edge ($\Delta = 1$): Conditional Rigidity. With an error budget of $k = m - 1$, each successful challenge confines the syndrome $s(z)$ to a subspace $U_T = \mathrm{Span}(H_T)$ of dimension at most $m - 1$. Such a subspace is a hyperplane (or is contained in one). A line can intersect a fixed hyperplane in at most one point unless the entire line lies within it. This creates a dichotomy: either the line is globally trapped in a single hyperplane (which implies rigidity), or each successful challenge must correspond to a different hyperplane (implying no shared global structure). Rigidity is not guaranteed by the premise alone; it requires a "witness"—a geometric coincidence like two error supports being identical or having a small union—or additional protocol structure to force the line into a single trap.

Strict gap ($\Delta \geq 2$): Unconditional Rigidity. When the error budget is $k \leq m - 2$, each success restricts $s(z)$ to a subspace $U_T = \mathrm{Span}(H_T)$ of dimension at most k, which has a codimension of at least $\Delta \geq 2$. A line cannot generically intersect many distinct subspaces of codimension 2 or more. A small number of successful challenges will force the line into the intersection of these subspaces, which is itself a low- dimensional subspace. This geometric constraint is so strong that it forces the entire line into a single, fixed low-dimensional span, yielding unconditional rigidity. This is formalized by the vanishing differences argument, which shows that a handful of successes create a linear dependency among the error vectors that is impossible unless they share a small, common support.

In short: *capacity is vacuous*; the *knife-edge is rigid only with a witness or structure*; and a *strict gap is rigid by geometry alone.*

5 Capacity: $\Delta = 0$ (Premise is Vacuous)

5.1 Vacuity of the CA Premise at Equality

At the capacity threshold, the CA premise is not just weak; it is entirely vacuous. It holds true for any pair of functions, regardless of their structure, and therefore provides no soundness guarantee.

Theorem 5.1 (Vacuity of the CA Premise at Equality)**.** *If $t = d$ (so $k = m$), then for **any** two functions $f, g \in \mathbb{F}_q^n$ and **any** challenge $z \in \mathbb{F}_q$, there exists a codeword $p_z \in \mathrm{RS}_{n,d}(D)$ that agrees with $h_z = f + zg$ on at least d positions.*

Proof. Let $s(z) = A + zB$ be the syndrome of h_z. We must show that for an arbitrary syndrome $s(z) \in \mathbb{F}_q^m$, one can find an error vector e_z of weight at most m such that $He_z^T = s(z)$. Choose any set of columns $T \subseteq [n]$ of size m. By the MDS property of Reed-Solomon codes (Lemma 3.1), the submatrix H_T is invertible, and thus its columns span the entire syndrome space $\mathbb{F}_q^m$. Therefore, for any $s(z)$, we have $s(z) \in \mathrm{Span}(H_T)$. This implies the existence of a vector $e_z \in \mathbb{F}_q^n$ supported only on the coordinates in T such that $He_z^T = s(z)$. The weight of this error vector is at most $|T| = m$. Now, define $p_z := h_z - e_z$. The syndrome of p_z is $Hp_z^T = H(h_z - e_z)^T = s(z) - s(z) = 0$, so p_z is a valid codeword in $\mathrm{RS}_{n,d}(D)$. By construction, p_z agrees with h_z on every coordinate outside of T. Since $|T| = m$, the number of agreement positions is $n - m = d$. Thus, the agreement threshold of $t = d$ is always met. □

5.2 Impossibility of Assumption-Free Rigidity at $t = d$

Not only is the $t = d$ premise always satisfied, but without extra structure, the desired rigidity conclusions can simultaneously fail. A malicious prover can pass every check without possessing any global low-degree structure.

Theorem 5.2. *At $t = d$ (so $k = m$), the CA premise holds for all challenges z and for every pair of functions (f, g). If $q^2 > \sum_{i=0}^{m-1} \binom{n}{i}$, then there exists a pair (f, g) for which neither of the standard rigidity conclusions holds:*

1. *There is no single set $L \subseteq [n]$ with $|L| \le m - 1$ such that $A, B \in \mathrm{Span}(H_L)$.*
2. *There do not exist codewords $P, Q \in \mathrm{RS}_{n,d}(D)$ and an index set L with $|L| \le m - 1$ such that f agrees with P outside L and g agrees with Q outside L.*

Counting proof (finite field). By Theorem 5.1, the CA premise holds for all (f, g) at $t = d$. The rigidity conclusion requires the syndrome pair (A, B) to lie in a small- dimensional subspace. Let $\mathcal{U}$ be the union of all "bad" pairs:

$$\mathcal{U} \;:=\; \bigcup_{\substack{L \subseteq [n] \\ |L| \le m-1}} \big(\mathrm{Span}(H_L) \times \mathrm{Span}(H_L)\big) \;\subset\; \mathbb{F}_q^m \times \mathbb{F}_q^m.$$

For each set L with $|L| \le m - 1$, $\mathrm{Span}(H_L)$ is a proper subspace of $\mathbb{F}_q^m$ of dimension at most $m - 1$. Thus, $|\mathrm{Span}(H_L) \times \mathrm{Span}(H_L)| \le q^{2(m-1)}$. By a union bound, the total number of such pairs is:

$$|\mathcal{U}| \;\le\; \left(\sum_{i=0}^{m-1} \binom{n}{i}\right) q^{2(m-1)}.$$

The total number of possible syndrome pairs (A, B) is q^{2m}. If $q^2 > \sum_{i=0}^{m-1} \binom{n}{i}$, then $|\mathcal{U}| < q^{2m}$. This guarantees the existence of a pair $(A, B) \notin \mathcal{U}$. Since the parity-check matrix H has full rank m, we can always find functions (f, g) such that $Hf^T = A$ and $Hg^T = B$. This pair (f, g) violates conclusion (1) by construction. If, toward a contradiction, conclusion (2) held, there would exist codewords $P, Q \in \mathrm{RS}_{n,d}(D)$ and a set L with $|L| \le m - 1$ such that $u := f - P$ and $v := g - Q$ are supported on L. Since $HP^T = HQ^T = 0$, the syndromes would be $A = H(P + u)^T = Hu^T \in \mathrm{Span}(H_L)$ and $B = H(Q + v)^T = Hv^T \in \mathrm{Span}(H_L)$. This would imply $(A, B) \in \mathcal{U}$, contradicting our choice. Hence, conclusion (2) also fails. □

Corollary 5.3 (Probabilistic impossibility at $t = d$ over small fields)**.** *Let $m = n - d$ and let $(A, B) \in \mathbb{F}_q^m \times \mathbb{F}_q^m$ be sampled uniformly at random. Then*

$$\Pr\{(A, B) \in \mathcal{U}\} \;\le\; \sum_{i=0}^{m-1} \binom{n}{i} q^{2(i-m)}\,,$$

where $\mathcal{U} := \bigcup_{|L| \le m-1} \big(\mathrm{Span}(H_L) \times \mathrm{Span}(H_L)\big)$. In particular, if this probability is less than 1, a counterexample to rigidity exists and can be found by random sampling.

5.3 Structural Rigidity at the Equality Boundary (Geometric Remedies)

While the basic CA premise at $t = d$ is vacuous, rigidity can be restored if additional structure is present or enforced. These remedies fall into three categories: algebraic, combinatorial, and protocol-level.

Algebraic remedy (bundle constancy / geometric KSP*). If the error supports, while changing with z, all share a common algebraic annihilator (an error-locator polynomial), this shared structure can be detected and used to force rigidity.

Lemma 5.4. *Fix $\ell \in \{1, \dots, m-1\}$ and define the Hankel pencil $M_\ell(z) := \mathcal{C}_\ell(A) + z\,\mathcal{C}_\ell(B)$. Assume there are $\ell + 2$ distinct challenges $z_0, \dots, z_{\ell+1}$ such that for each z_j, $\mathrm{rank}\, M_\ell(z_j) \le \ell$ and $\dim \mathrm{Ker}\, M_\ell(z_j) = 1$. Then there is a nonzero vector $q \in \mathbb{F}_q^{\ell+1}$ such that*

$$\mathcal{C}_\ell(A)\, q \;=\; -\, z\, \mathcal{C}_\ell(B)\, q\,,$$

implying that A and B are both explained by errors on a common set of at most ℓ locations.

Proof. Each $(\ell+1) \times (\ell+1)$ minor of $M_\ell(z)$ is a polynomial in z of degree at most $\ell + 1$. Vanishing at $\ell + 2$ distinct points forces it to be the zero polynomial, so $\mathrm{rank}\, M_\ell(z) \le \ell$ for all z. The one-dimensional kernel assumption implies that the kernel forms a line in projective space over the field of rational functions $\mathbb{F}_q(z)$. The identity $\mathcal{C}_\ell(A)\, q(z) \;=\; -\, z\, \mathcal{C}_\ell(B)\, q(z)$ shows that this line must be defined over $\mathbb{F}_q$, meaning the kernel is constant up to scaling. Taking a fixed representative q gives the result. □

Combinatorial remedy (sunflower trap). If the error supports exhibit significant overlap, this combinatorial structure can also trap the syndrome line.

Lemma 5.5 (Sunflower trap: affine-test form)**.** *Let the supports $\{T_{z_i}\}_{i=0}^m$ form a sunflower with core C of size $|C| = m - 1$ and distinct petals. Let $U := \mathrm{Span}(H_C)$ and pick a nonzero $\phi^\top$ with $\phi^\top U = 0$. Then exactly one of the following holds:*

1. *$A, B \in U$ and hence $s(z) \subseteq U$;*
2. *the petal scalars $\{\alpha_i\, \phi^\top h_{petal(i)}\}_{i=0}^m$ lie on a single affine function of z (equivalently, for every three indices i, j, k we have $\det\begin{pmatrix} 1 & z_i & y_i \\ 1 & z_j & y_j \\ 1 & z_k & y_k \end{pmatrix} = 0$ with $y_\ell := \phi^\top s(z_\ell)$).*

Proof. Write $s(z_i) = H_C u_i + \alpha_i h_{\mathrm{petal}(i)}$ and apply $\phi^\top$ to get $y_i = \phi^\top s(z_i) = \alpha_i\, \phi^\top h_{\mathrm{petal}(i)}$. Since $\phi^\top s(z) = \phi^\top A + z\, \phi^\top B$ is affine in z, the values y_i must satisfy $y_i = a + b z_i$ for some $a, b \in \mathbb{F}_q$. If

$a = b = 0$ we are in (1). Otherwise the $m+1$ values obey the affine relation in (2), equivalently: for every three indices i, j, k we have

$$\det \begin{pmatrix} 1 & z_i & y_i \\ 1 & z_j & y_j \\ 1 & z_k & y_k \end{pmatrix} = 0 \qquad \text{with } y_\ell := \phi^\top s(z_\ell).$$

□

Remark 5.6. A classical sunflower/Helly-style derivation of the same trap appears in Appendix 11.

Protocol-level remedy (enforced rigidity). A cryptographic protocol can enforce rigidity at $t = d$ by design. Examples include:

- **Explicit Commitments:** The prover commits to a single error locator polynomial or a small error support set L upfront, and all subsequent checks are verified against this commitment. This effectively reduces the problem to a case with a small global error set, which is rigid.
- **Independent Folds (MCA):** Using two independent folding challenges per round, as in Mutual Correlated Agreement, creates a stronger constraint. As shown in Theorem 7.3, this causes the security margins to add, effectively creating a strict gap ($\Delta_{\text{total}} \geq 2$) from two knife-edge instances ($\Delta_1 = 1, \Delta_2 = 1$).
- **Out-of-Domain Sampling (DEEP/STIR):** Techniques like DEEP-FRI ask the prover for evaluations outside the original domain. This provides a global constraint that links the error structures across different challenges, preventing the prover from exploiting the freedom of the $\Delta = 0$ or $\Delta = 1$ regimes.

6 Knife-Edge: $\Delta = 1$ (Witnesses or Structure Required)

At $k = m - 1$ ($\Delta = 1$), the problem enters a delicate "knife-edge" regime. The general-purpose rigidity argument from the strict-gap case (based on vanishing differences) fails, as the condition $(r+1)k < m+1$ cannot be satisfied for $r \geq 2$ when $m \geq 3$. However, the geometry of syndrome space provides a sharp alternative path to understanding this case.

Theorem 6.1. *Let $k = m - 1$ and consider the affine syndrome line $s(z) = A + zB$. For any support set T with $|T| = m - 1$, let $U_T := \mathrm{Span}(H_T)$ (a hyperplane in $\mathbb{F}_q^m$). If there exist distinct $z_1, z_2 \in \mathbb{F}_q$ with $s(z_1), s(z_2) \in U_T$, then necessarily $A, B \in U_T$, and hence $s(z) \in U_T$ for all $z \in \mathbb{F}_q$. Otherwise, the hyperplane U_T contains at most one point of the line $\{\, s(z) : z \in \mathbb{F}_q \,\}$. In particular, for each such T, either the entire line lies in U_T or $|\{z \in \mathbb{F}_q : \ s(z) \in U_T\}| \leq 1$.*

Proof. Fix a set T with $|T| = m - 1$. The subspace $U_T = \mathrm{Span}(H_T)$ has dimension at most $m - 1$, so it is a hyperplane (or is contained in one). If two distinct challenges $z_1 \neq z_2$ result in syndromes $s(z_1)$ and $s(z_2)$ that both lie in U_T, then their difference must also lie in U_T.

$$(z_2 - z_1)B \in U_T\,.$$

Since $z_1 \neq z_2$, this implies $B \in U_T$. Furthermore, since $s(z_1) = A + z_1 B \in U_T$ and $B \in U_T$, it must be that $A \in U_T$. If both A and B are in U_T, then the entire line $s(z) = A + zB$ is contained in U_T. Otherwise, the line can intersect the hyperplane at most once. □

This dichotomy is critical: either the line is globally trapped in a single hyperplane (which implies rigidity), or at most one successful challenge can correspond to any given error support of size $m-1$. This means a prover can pass many checks by using a different error structure each time, without possessing any global structure. Soundness, therefore, is not unconditional. It requires a "witness" that forces multiple points onto the same hyperplane.

Theorem 6.2 (Small-Union Rigidity for $\Delta \geq 1$)**.** *Assume the regime, where $k \leq m-1$ (equivalently, $\Delta \geq 1$). Suppose there exist two distinct successful challenges $z_i \neq z_j$ with the small-union property:*

$$\left| T_{z_i} \cup T_{z_j} \right| \;\leq\; m-1.$$

Then the entire syndrome line $s(z) = A + zB$ lies in the proper subspace $\mathrm{Span}\big(H_{T_{z_i} \cup T_{z_j}}\big)$*.*

Proof. Let $U := T_{z_i} \cup T_{z_j}$. By the CA premise, the syndromes $s(z_i)$ and $s(z_j)$ must lie in $\mathrm{Span}(H_{T_{z_i}})$ and $\mathrm{Span}(H_{T_{z_j}})$ respectively, and therefore both lie in the sum of these subspaces, $\mathrm{Span}(H_U)$.

Since two distinct points on the affine line $s(z)$ are in $\mathrm{Span}(H_U)$, the entire line must be contained in this subspace. Specifically, the difference $(z_j - z_i)B = s(z_j) - s(z_i)$ lies in $\mathrm{Span}(H_U)$, which implies the direction vector B is in $\mathrm{Span}(H_U)$ (since $z_i \neq z_j$). It follows that the starting point $A = s(z_i) - z_i B$ must also be in $\mathrm{Span}(H_U)$.

By hypothesis, the union of the supports has size $|U| \leq m-1$. Because the code is MDS, any set of fewer than m columns of H cannot span the full m-dimensional syndrome space. Therefore, $\mathrm{Span}(H_U)$ is a proper subspace of $\mathbb{F}_q^m$, and the line is trapped. □

Corollary 6.3 (Repeat Support Implies Rigidity)**.** *For $\Delta \geq 1$, if two distinct successful challenges $z_i \neq z_j$ have the same error support ($T_{z_i} = T_{z_j}$), then the syndrome line is trapped in the proper subspace* $\mathrm{Span}(H_{T_{z_i}})$*.*

Proof. This is a direct consequence of Theorem 6.2. If $T_{z_i} = T_{z_j}$, their union is simply T_{z_i}. The size of this union is $|T_{z_i}| \leq k = m - \Delta$. Since $\Delta \geq 1$, we have $|T_{z_i}| \leq m-1$. The hypothesis of the theorem is met, and the conclusion follows. □

7 Strict Gap: $\Delta \geq 2$ (Unconditional Rigidity)

7.1 Geometric Proof of Unconditional Rigidity

When the rank margin is strict ($\Delta \geq 2$, or $k \leq m-2$), the geometric constraints on the syndrome line become powerful enough to guarantee rigidity without any additional assumptions. The core of the argument is a "vanishing differences" identity derived from the annihilation property of polynomials.

Theorem 7.1. *Assume the CA premise holds on a set of challenges $\mathcal{Z}^*$. Let $z_0, \ldots, z_r$ be $r+1$ distinct challenges from $\mathcal{Z}^*$. If $(r+1)k < m+1$ for some integer $r \geq 2$, then the corresponding error vectors are linearly dependent:*

$$\sum_{i=0}^{r} \lambda_i \, e_{z_i} \;=\; 0\,,$$

where the λ_i are the annihilation weights from Corollary 3.3.

Proof. From the CA premise, we have $s(z_i) = He_{z_i}^T$ for each $i = 0, \ldots, r$. By Corollary 3.3, the affine nature of the syndrome line implies $\sum_{i=0}^{r} \lambda_i s(z_i) = 0$. Since $s(z_i) = He_{z_i}^T$, we obtain

$$\sum_{i=0}^{r} \lambda_i s(z_i) \;=\; \sum_{i=0}^{r} \lambda_i (He_{z_i}^T) \;=\; H\left(\sum_{i=0}^{r} \lambda_i e_{z_i}^T\right) \;=\; 0\,.$$

Thus the vector $w := \sum_{i=0}^{r} \lambda_i e_{z_i}$ is a codeword in $\mathrm{RS}_{n,d}(D)$. *Equivalently,*

$$Hw^T \;=\; H\left(\sum_{i=0}^{r} \lambda_i e_{z_i}^T\right) \;=\; \sum_{i=0}^{r} \lambda_i s(z_i) \;=\; 0\,,$$

so w has zero syndrome.

The support of w is contained within the union of the supports of the individual error vectors: $\mathrm{supp}(w) \subseteq \bigcup_{i=0}^{r} T_{z_i}$. Therefore,

$$\mathrm{wt}(w) \;\leq\; \sum_{i=0}^{r} |T_{z_i}| \;\leq\; (r+1)k\,.$$

By hypothesis, $(r+1)k < m+1$. For a Reed–Solomon code, the minimum distance is $d_{\min} = n - d + 1 = m + 1$, hence $\mathrm{wt}(w) < d_{\min}$. The only codeword with weight less than the minimum distance is the zero vector, so $w = 0$. □

This linear dependency is the only structural fact needed. It implies that the error values across different challenges are not independent but are linked by a low-degree polynomial structure. This structure can be used to trap all errors within a small global set of coordinates.

7.2 Unconditional CA and MCA Rigidity

The vanishing differences property has two powerful consequences: a global bound on the total number of error locations, and a soundness guarantee for protocols with independent folds like MCA.

Theorem 7.2. *Let $M = |\mathcal{Z}^*|$ be the number of challenges for which the CA premise holds. Suppose $(r+1)k < m+1$ for some $r \geq 2$ and $M > r-1$. Define the global error set $S := \bigcup_{z \in \mathcal{Z}^*} T_z$. Then the size of this set is bounded:*

$$|S| \;\leq\; \frac{M\,k}{M-(r-1)}\,.$$

In particular, if M is large enough such that this bound is at most $m-1$, then $|S| \leq m-1$ and the entire syndrome line lies in the single proper subspace $\mathrm{Span}(H_S)$*, implying $A, B \in$* $\mathrm{Span}(H_S)$.

Proof. Fix a coordinate $j \in S$. Consider the function $g_j(z) := e_z(j)$, which gives the error value at coordinate j for challenge z. Theorem 7.1 implies that for any $r+1$ distinct points in $\mathcal{Z}^*$, the values of $g_j(z)$ satisfy a linear recurrence. This means that $g_j(z)$ must agree with a polynomial in z of degree at most $r-1$ on the set $\mathcal{Z}^*$. A non-zero polynomial of degree at most $r-1$ can have at most $r-1$ roots. Therefore, if $g_j(z)$ is not identically zero, coordinate j must be in error for at least $M-(r-1)$ challenges in $\mathcal{Z}^*$. We can now count the total number of error incidences (pairs (z, j) where $j \in T_z$) in two ways:

$$\sum_{z \in \mathcal{Z}^*} |T_z| \;\leq\; M \cdot k, \qquad\qquad \sum_{j \in S} |\{\, z : j \in T_z \,\}| \;\geq\; |S| \cdot (M-(r-1))\,.$$

The left side is at most $M \cdot k$. The right side is at least $|S| \cdot (M - (r-1))$. Combining these gives:

$$|S| \cdot (M - (r-1)) \;\le\; Mk\,.$$

Rearranging yields the desired bound. If $|S| \le m-1$, then for every $z \in \mathcal{Z}^*$, the error vector e_z is supported on S. This means $s(z) = He_z^T \in \mathrm{Span}(H_S)$. Since this holds for at least two distinct z, it forces $A, B \in \mathrm{Span}(H_S)$. □

Theorem 7.3 (MCA Exponent Addition)**.** *Assume a protocol uses two independent folds per round (e.g., via domain-separated Fiat-Shamir), with per-fold rank margins $\Delta_1, \Delta_2 \ge 0$. If the soundness bounds for fooling each fold individually are $\Pr[\text{FA on fold } i] \le q^{-\Delta_i}$, then the soundness of fooling the round is multiplicative. Over s independent rounds, the total soundness error is:*

$$\Pr[\text{FA in a round}] \;\le\; q^{-(\Delta_1+\Delta_2)}\,, \qquad \Pr[\text{FA over } s \text{ rounds}] \;\le\; q^{-(\Delta_1+\Delta_2)s}\,.$$

Proof. Let $E_{i,t}$ be the event that the prover successfully fools the verifier on fold $i \in \{1,2\}$ in round t. The per-fold soundness bound is $\Pr(E_{i,t}) \le q^{-\Delta_i}$. The use of independent randomness for each fold (e.g., domain separation in the Fiat-Shamir transform) ensures that the events $E_{1,t}$ and $E_{2,t}$ are independent. A prover fools the round only if they fool both folds, so the probability is:

$$\Pr(E_{1,t} \wedge E_{2,t}) = \Pr(E_{1,t}) \cdot \Pr(E_{2,t}) \le q^{-\Delta_1} \cdot q^{-\Delta_2} = q^{-(\Delta_1+\Delta_2)}\,.$$

If the rounds also use independent randomness, the probabilities multiply across rounds, giving the final bound. □

Remark 7.4 (Alternative Rigidity via a Local Witness). In addition to the unconditional proof from vanishing differences, rigidity in the $\Delta \ge 2$ regime can also be established via a local 'small-union' witness. The geometric argument is identical to the one presented for the knife-edge case in Theorem 6.2. If a prover provides two successful claims whose error supports have a union of size at most $m-1$, the syndrome line is immediately trapped in the corresponding subspace.

8 Mutual Correlated Agreement (MCA) under the Δ-Spine

The Mutual Correlated Agreement (MCA) problem strengthens CA by requiring consistency across multiple folds simultaneously. The standard reduction from MCA to CA (as in BKS [3]) shows that if CA holds, so does MCA. Our results therefore directly apply. Below capacity ($\delta < 1-\rho$), MCA holds under the same conditions as CA: either via the algebraic r-fold route when $(r+1)k < m+1$, or via a local combinatorial witness.

Theorem 8.1. *Let $\rho = d/n$ and consider a proximity parameter $\delta < 1-\rho$ (so $t > d$). Assume either $(r+1)k < m+1$ for some $r \ge 2$, or a combinatorial witness occurs (repeat support or $|T_{z_i} \cup T_{z_j}| \le m-1$). If for a sufficiently large number of random challenges z, the folded word $h_z = f + zg$ is δ-close to $\mathrm{RS}_{n,d}(D)$, then there exist polynomials P, Q of degree $< d$ and a set $S' \subseteq D$ of size $|S'| \ge (1-\delta)n$ such that f and g agree with P and Q on S', respectively.*

Proof. The condition $\delta < 1-\rho$ is equivalent to $t > d$. Let $k = n-t$ and $m = n-d$. The rank margin is $\Delta := m - k = t - d \ge 1$. We are therefore in either the knife-edge or strict-gap regime. If the r-fold hypothesis holds (which requires $\Delta \ge 2$), Theorem 7.2 guarantees the existence of a single global error set S with $|S| \le m-1$ such that the syndromes A and B are in $\mathrm{Span}(H_S)$. This implies the existence of error vectors e_f and e_g supported on S such that $f - e_f$ and $g - e_g$ are codewords, which we call P and Q. Let $S' := D \setminus S$. Then on S', f agrees with P and g agrees with Q. If instead a combinatorial witness occurs (e.g., $|T_{z_i} \cup T_{z_j}| \leq m-1$), then Theorem 6.2 implies $A, B \in \mathrm{Span}(H_U)$ for the witnessed union U. Taking $S := U$ yields the same conclusion. The size of the agreement set is $|S'| = n - |S| \geq n - (m-1) = d+1$. The premise that each h_z is δ-close implies a much stronger agreement bound, consistent with $(1-\delta)n$. □

9 Discussion and Implications for Protocol Design

The syndrome-space lens provides a complete characterization of CA/MCA for Reed- Solomon codes, translating directly into concrete, actionable guidelines for cryptographic protocol design.

Regime	**Safety**	**What is required for soundness**	**Bits/round**	**Action for Designers**
$\Delta = 0$ $(t = d)$	Unsafe (vacuous)	—	0	**Do not operate here**; add a structural remedy (Sec. 5.3).
$\Delta = 1$, CA	Knife-edge	Witness/structure (*e.g.*, repeat support; small union $\lvert T_{z_i} \cup T_{z_j}\rvert \leq m-1$; DEEP/STIR; global locator).	$\log_2 q$ *only if* structure enforces margin	Add MCA (indep. folds) and/or DEEP/STIR.
$\Delta = 1$, MCA (indep. folds)	Safe (model-scoped)	Two *independent* folds (domain-separated Fiat-Shamir; transcript binding).	$2\log_2 q$	Budget soundness with q^{-2s}.
$\Delta \geq 2$	Safe (provable)	$(r+1)k < m+1$ for some $r \geq 2$.	$\Delta \log_2 q$	Budget soundness with $q^{-\Delta s}$.

Table 3: Operator guidance by regime. Bits/round and exponent additivity are *model-scoped*, requiring independent challenges/folds (Fiat-Shamir/ROM) and no cross-round adaptivity.

Principled Soundness Budgeting. *The most critical takeaway is a clear, auditable rule.*

> **Key budgeting rule.** For a per–round *effective* rank margin $\Delta_{\text{eff}} \geq 1$ and s *independent* query rounds (e.g., domain–separated Fiat-Shamir; no cross-round adaptivity),
>
> $$\Pr[\text{FA}] \leq q^{-\Delta_{\text{eff}} s} \quad \iff \quad \Delta_{\text{eff}}\, s \log_2 q \geq \lambda_{\text{sec}} + 2\,.$$
>
> *At $\Delta = 0$ the premise is vacuous; at $\Delta = 1$ a margin counts only with a witness/structure or via MCA; for $\Delta \geq 2$ the bound is unconditional (under $(r+1)k < m+1$).*

This provides a firm, auditable basis for parameter selection, replacing prior heuristics. In an MCA setting with two *independent* folds per round with margins Δ_1 and Δ_2, the effective margin adds, $\Delta_{\text{eff}} = \Delta_1 + \Delta_2$ (Theorem 7.3), yielding the (model–scoped) bound $\Pr[\text{FA}] \leq q^{-(\Delta_1+\Delta_2)s}$. For example, two independent folds in the knife–edge regime ($\Delta_1 = \Delta_2 = 1$) give $\Delta_{\text{eff}} = 2$ and thus $\Pr[\text{FA}] \leq q^{-2s}$. Appendix B lists concrete parameter floors. At capacity ($\Delta = 0$) the premise is vacuous (no soundness), at knife–edge ($\Delta = 1$) a margin counts only with a witness/structure or via MCA, and in the strict–gap ($\Delta \geq 2$) the bound holds unconditionally (under $(r+1)k < m+1$).

Parameter Transparency and Avoiding Silent Failures. Implementations must make *explicit* the proximity parameters and assumptions under which soundness is claimed. At minimum, publish (n, d, t, q), the number of rounds s, and the per–round effective margin Δ_{eff} (or, for MCA, the per–fold margins (Δ_1, Δ_2) with $\Delta_{\text{eff}} = \Delta_1 + \Delta_2$), together with the independence mechanism (e.g., domain–separated Fiat-Shamir / ROM) used to justify exponent additivity. Any protocol operating at $t = d$ ($\Delta = 0$) is *vacuous* unless accompanied by a structural remedy (Section 5.3); claiming security there without such safeguards yields a proof system with *zero soundness*. Likewise, at knife–edge ($\Delta = 1$), a margin counts only with a witness/structure (repeat support, small union, DEEP/STIR, global locator) or via MCA; otherwise the budgeting rule does not apply. This transparency prevents "silent failures," where a system appears to work yet operates in a regime (e.g., $\Delta = 0$ or $\Delta = 1$ without structure/independence) that cannot support the advertised bound. *Checklist:* disclose (n, d, t, q, s); the per–round Δ_{eff} (or (Δ_1, Δ_2) for MCA); and how round/fold *independence* is enforced. Do *not* claim budgeting at $\Delta = 0$, and only claim it at $\Delta = 1$ with structure or MCA.

9.1 Conjecture 8.4 (BCI$^+$20) — Revised Status

The results in this paper provide a complete resolution to Conjecture 8.4 of BKS20 [3]. The conjecture posits that for a proximity parameter $\delta \le 1 - \rho - \eta$, the soundness error ε of the affine CA test satisfies

$$\varepsilon \;\le\; \frac{1}{(\eta\rho)^{c_1}} \cdot \frac{n^{c_2}}{q}$$

Mapping parameters from their notation to ours reveals that their gap parameter η corresponds to our normalized rank margin:

$$\eta \;=\; (1-\rho) - \delta \;=\; \frac{m}{n} - \frac{k}{n} \;=\; \frac{m-k}{n} \;=\; \frac{\Delta}{n}$$

The conjecture therefore predicts a soundness error that scales as $1/q$, with polynomial factors in n, ρ, and Δ^{-1}. Our syndrome-space analysis reveals a more nuanced reality and provides a much stronger bound in the secure regime. Our bound is "model-scoped" in that exponent addition requires independent folds/rounds (e.g., domain-separated Fiat-Shamir) and the per-round margin is the *effective* Δ_{eff} as discussed earlier in Section 9.

Table 4: Parameter mapping between BKS20 Conjecture 8.4 and this paper.

BKS20 Symbol	**Meaning**	**This Paper**	**Relation**
n	code length	n	same
q	field size	q	same
ρ	rate $= d/n$	$\rho = d/n$	same
δ	error / proximity fraction	$\delta = k/n$	same
η	gap below capacity	$\eta = \Delta/n$	$\Delta := t - d = m - k$
ε	soundness error	false-accept probability	same quantity

Summary of Status. Conjecture 8.4 is false in its unconditional form at and near capacity, but it is superseded by a much stronger result in the strict-gap regime.

- For $\Delta = 0$ ($t = d$, $\eta = 0$), the premise is vacuous and there is no soundness.

- For $\Delta = 1$ ($t = d + 1$), unconditional rigidity does not hold. A prover can succeed with probability 1 by choosing different error supports for each challenge, violating the spirit of the conjecture. Soundness requires additional structure.
- For $\Delta \geq 2$ (and when the r-fold condition holds), we prove unconditional rigidity. Our soundness bound of $q^{-\Delta s}$ is exponentially stronger than the $1/q$ bound conjectured. It eliminates the polynomial dependency on n and replaces it with a clean exponential decay in the rank margin.

10 Open Problems and Extensions

Our syndrome-space lens opens several avenues for further research:

1. **Beyond Reed-Solomon:** Can a similar affine/linear-algebraic geometry approach be applied to proximity testing for other algebraic codes, such as *algebraic-geometric (AG) codes*? For AG codes, syndromes can be defined in a similar manner, and it is plausible that a sharp phase transition also exists, though the geometry may be more complex.
2. **Explicit Small-Field Adversaries:** Our impossibility result for $\Delta = 0$ relies on a counting argument for large fields. Constructing deterministic, worst-case adversarial functions (f, g) over small fields (e.g., $q = O(n)$) that pass the $t = d$ test yet defeat all rigidity conclusions would be a valuable contribution.
3. **Multidimensional Generalizations:** Recent work has begun to explore proximity gaps for interleaved and tensor-product codes [6]. The SSL might extend to these settings by considering higher-dimensional syndrome spaces or tensor products of parity-check matrices, potentially revealing analogous rigidity thresholds.
4. **Analytic Decoding Connections:** Recent work has explored analytic approaches to list decoding [8]. Is there a way to translate our syndrome-space criteria (like the rank conditions on Hankel pencils) into analytic constraints on the Fourier or other spectral transforms of the functions? Bridging these perspectives could yield new algorithms or deeper insights into the fundamental barriers to decoding.

11 Conclusion

By adopting a new perspective—the syndrome-space lens—this work has resolved the long-standing problem of Correlated Agreement for Reed-Solomon codes up to the information-theoretic capacity limit. This change of basis converts RS proximity testing from a complex list-decoding analysis into a transparent linear-algebraic geometry problem, revealing a sharp phase transition in the process. We established a precise trichotomy based on the rank margin $\Delta = t - d$. We showed that *unconditional* rigidity holds in the strict-gap regime ($\Delta \geq 2$), while the capacity boundary ($\Delta = 0$) is fundamentally vacuous unless additional structure is imposed, and the $\Delta = 1$ regime is a knife-edge requiring a witness or protocol-level enforcement. Our classification of the necessary structural remedies provides a complete map for achieving soundness in all cases. These results close a significant theoretical gap while simultaneously delivering immediate practical benefits for STARKs and other proof systems. The derived soundness law, $\Pr[\text{FA}] \leq q^{-(\sum \Delta_i)s}$, provides a clear, auditable rulebook for parameter selection, rigorously justifying efficiency improvements that were previously based on heuristics. Methodologically, the syndrome-space approach may prove broadly useful beyond RS codes, inviting exploration into other code families and continuing the fruitful interplay between classical coding theory and modern cryptographic protocol design.

The Lying Finger wrongs; and, having throngs,
Lies on: nor all thy Polity nor Writ
Shall remand it back to recant half a Lie,
Nor all thy Peers right out a Sentence for it.

Proof propels us from our Age of Impunity to a new Age of Accountability,
Where words don't merely matter, but words are matter,
An Age where words count...

The Arithmocene.

Appendix A: Additional Details for Barriers and Sharpness

A.1 Linear-Algebraic Normal Form for the Strict-Gap Barrier

Fix distinct challenges $z_0, \dots, z_r \in \mathbb{F}_q$. For each i, choose a support $T_i \subseteq [n]$ of size k. The CA constraints are $H_{T_i} u_i = A + z_i B$ for some unknown error values u_i and syndromes (A, B). The divided-difference barrier constructs a minimum-weight codeword v as a combination $\sum \lambda_i R_i u_i$, where R_i is a selector matrix picking out coordinates in T_i. This yields a linear system

$$\mathcal{M}\mathbf{x} = \mathbf{b}$$

with $(r{+}1)k + 2m$ unknowns and $m(r{+}1) + (m{+}1)$ equations. If $(r{+}1)k \geq r\,m + 1$, the system is underdetermined. One can show (using the MDS property) that the supports T_i can be chosen to make $\mathcal{M}$ full-row-rank, guaranteeing a solution exists (hence vacuity). The details involve certain polynomial constraints in the z_i which can be avoided by Schwartz-Zippel for random z_i over large q.

A.2 Distinct Challenges and Schwartz-Zippel Considerations

The weights $\lambda_i = \prod_{j \neq i} (z_i - z_j)^{-1}$ used in divided differences require the z_i to be distinct. The full-rank condition on $\mathcal{M}$ in A.1 depends on certain minors of a matrix (involving the z_i) being nonzero. These conditions define some algebraic variety (polynomial constraints on the tuple $(z_0, \dots, z_r)$). If the field size q is larger than the sum of their degrees, the Schwartz-Zippel lemma guarantees that a random choice of $(z_0, \dots, z_r)$ will avoid all these bad dependencies with high probability. In practical terms, this says that for sufficiently large q, one can safely pick distinct challenges without worrying about degenerate linear dependencies invalidating the barrier argument.

Appendix B: Parameter Floors (Reference)

Budgeting rule (conservative). For a single fold with per-round rank margin Δ repeated s times,

$$\Pr[\text{false accept}] \le q^{-\Delta s} \quad \Rightarrow \quad \Delta s \ \ge \ \left\lceil \frac{\lambda_{\text{sec}} + 2}{\log_2 q} \right\rceil \tag{B.1}$$

The "+2" is a conservative slack covering small statistical losses outside the core analysis (e.g., hash-to-field, Fiat-Shamir instantiation, and constant-factor effects).

For $\ell \ge 2$ independent folds per round with margins $\Delta_1, \dots, \Delta_\ell$, the exponents add:

$$\left(\sum_{i=1}^{\ell} \Delta_i \right) s \ \ge \ \left\lceil \frac{\lambda_{\text{sec}} + 2}{\log_2 q} \right\rceil \tag{B.2}$$

assuming independence (or a sound surrogate under Fiat-Shamir).

Table 5: Conservative parameter floors for $\lambda_{\text{sec}} = 128$ (minimal examples; CA and MCA)

Field	$\log_2 q$	**Required total exponent**	**Minimal example(s) meeting target**
Goldilocks-like	≈ 32	$\lceil 130/32 \rceil = 5$	CA: $\Delta = 5$, $s = 1$ (or $\Delta = 1$, $s = 5$); MCA: $(\Delta_1, \Delta_2) = (1, 1)$, $s = 3$
64-bit prime†	≈ 64	$\lceil 130/63 \rceil = 3$	CA: $\Delta = 3$, $s = 1$ (or $\Delta = 1$, $s = 3$); MCA: $(1, 1)$, $s = 2$
BLS12-381 scalar	≈ 252	$\lceil 130/252 \rceil = 1$	CA: $\Delta = 1$, $s = 1$

† We conservatively budget with 63 bits for "64-bit primes" to reflect that $\log_2 q$ can be < 64; using 63 prevents overstating the exponent.

Note. Examples are minimal; any configuration with (B.1) or (B.2) satisfied is acceptable. In particular, for MCA (two independent folds per round) the round exponent is $(\Delta_1 + \Delta_2) \log_2 q$; the examples shown take $\Delta_1 = \Delta_2 = 1$ where applicable/minimal. Optional per-line sampling or composition across folds is compatible with this budgeting without changing the exponent arithmetic.

Appendix C: Empirical Validation of the Δ-Trichotomy

To provide a direct, computational proof of concept for the paper's central Δ-trichotomy, we include a minimal, reviewer-friendly Python simulator that works over prime fields $\mathbb{F}_q$. The simulator constructs random linear systems whose *nullity* indicates a cheating degree of freedom (a non-trivial kernel). A "success" is defined as nullity > 0. With NumPy installed, the rank computation runs in a vectorized "turbo" mode (GF(q) elimination via broadcasting); otherwise a pure-Python modular elimination is used. The empirical estimates match the theoretical laws to within sampling error.

From linear constraints to nullity. The linear-algebraic heart of the analysis is that the CA/MCA premises induce linear constraints on error vectors. In our toy model, we proxy this by sampling random full-rank candidates (subject to the appropriate over-constraint margin) and measuring the frequency of rank deficiency over $\mathbb{F}_q$. This frequency approximates the soundness failure probability: once Δ over-constraints are in play, the chance of a (random) degenerate configuration is $\approx q^{-\Delta}$.

Regimes and predictions (paper terminology). The simulator supports the three regimes of the Δ-trichotomy (plus MCA), matching the theoretical predictions derived in the paper:

- **Underconstrained / Capacity ($\Delta = 0$):** Here $k = m$. Theory predicts a non-trivial kernel always exists (vacuity); the simulator reports success ≈ 1. This corresponds to Theorem 5.1.
- **Critically constrained / Knife-edge ($\Delta = 1$):** Here $k = m - 1$. For a random instance over $\mathbb{F}_q$, the probability of rank deficiency is $\approx q^{-1}$. This aligns with the geometric knife-edge dichotomy (Theorem 6.1).
- **Overconstrained / Strict-gap ($\Delta \geq 2$):** Here $k \leq m - 2$. Random instances are rank-deficient with probability $\approx q^{-\Delta}$, matching the strict-gap soundness. This echoes the unconditional rigidity implications (e.g., Theorems 7.1 and 7.2).
- **Mutual Correlated Agreement (MCA, $\Delta_{\text{tot}} = \Delta_1 + \Delta_2$):** Two independent folds per round behave multiplicatively; the success probability is $\approx q^{-(\Delta_1+\Delta_2)}$, demonstrating exponent addition (Theorem 7.3).

Usage and artifacts. The complete source code, a test-matrix driver with auto-sizing of trial counts, and deterministic examples are included in the supplementary materials (see `README.md` and `EXPECTED_OUTPUTS.md`). The table below gives copy-paste commands for representative runs.

Table 6: Experimental regimes and theoretical predictions (paper-aligned CLI).

Regime	**Parameters**	**Example command (CLI)**	**Theory**

Underconstrained	$\Delta = 0$	`--regime underconstrained`
		`--Delta 0`
		`--q 29`
		`--trials 2000`
			1

Knife-edge	$\Delta = 1$	`--regime knife`
		`--Delta 1`
		`--q 29`
		`--trials 4000`
			$\approx q^{-1}$

Overconstrained (strict-gap)	$\Delta = 2$ (example)	`--regime overconstrained`
		`--Delta 2`
		`--q 19`
		`--trials 10000`
			$\approx q^{-\Delta}$

MCA	$\Delta_1 = 1,\ \Delta_2 = 1$	`--regime mca`
		`--Delta1 1`
		`--Delta2 1`
		`--q 29`
		`--trials 6000`
			$\approx q^{-(\Delta_1+\Delta_2)}$

As documented in `EXPECTED_OUTPUTS.md`, the empirical rates closely track the theoretical values (e.g., $\approx 1/q$ at $\Delta = 1$, $\approx 1/q^2$ at $\Delta = 2$, and $\approx 1/q^{\Delta_1+\Delta_2}$ for MCA), with small Monte Carlo deviations. The NumPy vectorized path (when present) significantly accelerates GF(q) elimination without changing results.

Code and artifacts availability. Code, documentation, and pre-generated outputs (CSV) are available at: `https://github.com/arithmocene/proximity`.

Appendix D: The Syndrome–Space Lens Methodology

The *syndrome–space lens* is a methodological shift: applying the parity–check map turns each folded word $h_z = f + zg$ into the affine line of syndromes $s(z) = A + zB$ and the CA premise into the geometric statement "$s(z)$ lies in some small column span $U_T = \mathrm{Span}(H_T)$". In this basis:

- At equality $t = d$ ($\Delta = 0$) the premise is vacuous (every point lies in some $U_T = \mathbb{F}_q^m$), and we prove impossibility of assumption–free rigidity.
- At the one-symbol margin $\Delta = 1$, a line meets each hyperplane U_T in at most one point unless contained; this gives a new impossibility: no witness-free unconditional CA/MCA theorem exists here.
- For $\Delta \geq 2$, vanishing differences yield unconditional rigidity with explicit quantitative bounds.

This lens bypasses list decoding and makes the capacity phase transition *visible*: vacuous ($\Delta = 0$), knife-edge ($\Delta = 1$), rigid ($\Delta \geq 2$). It also extends to higher dimensions (MCA planes and multi-fold k-flats), explaining why independence and structure restore soundness near capacity.

Spectral Interpretation

Reed-Solomon codewords are bandlimited "low-frequency" signals (degree $\leq d$). The syndrome $s = Hw^T$ is exactly the *high-frequency tail* beyond degree d (dual/GRS side). Our phase transition has a spectral form:

- $\Delta = 0$: no constraint on the tail (vacuity);
- $\Delta = 1$: only pointwise annihilations of tail coefficients (no global suppression without witnesses/structure);
- $\Delta \geq 2$: unconditional spectral collapse (bandlimitation enforced) by vanishing differences.

This dual viewpoint cross-validates the geometric proofs.

Langlands Perspective (Speculative)

Syndrome space is the dual code, hence a spectral object; evaluation space is the arithmetic/low-degree side. This mirrors the arithmetic–spectral split in the Langlands program (in a finite-field avatar): codewords (arithmetic side) vs. syndromes (spectral side). Our safe/knife-edge/unsafe trichotomy looks like a bandlimiting threshold, hinting at deeper bridges between harmonic analysis, incidence geometry, and rigidity.

Appendix E: Algebraic Route: Hankel-Pencil Kernel Stability (KSP*)

We first give an algebraic criterion for equality-case rigidity, based on analyzing a polynomial matrix formed by the syndrome components.

Consider the $(m-k) \times (k+1)$ Hankel matrices $\mathcal{C}_k(A)$ and $\mathcal{C}_k(B)$ of the syndrome vectors A and B (see Section 3.3 for definitions). Define the linear matrix pencil

$$M(z) \;=\; \mathcal{C}_k(A) + z\,\mathcal{C}_k(B).$$

This is a matrix whose entries are linear functions in z. Note that for each challenge z, the condition that $s(z) = A + zB$ can be explained by at most k errors translates to saying $\dim \operatorname{Ker} M(z) \geq 1$ (there is an error-locator polynomial of degree $\leq k$ for T_z) or equivalently $\operatorname{rank} M(z) \leq k$. At $t = d$ (i.e. $k = m$), this condition is always trivially satisfied for any z, so it's not helpful by itself. However, suppose *further* that the pencil $M(z)$ has rank $\leq k$ for $k + 2$ distinct values of z. Then all its maximal minors (which are polynomials in z of degree $\leq k + 1$) have $k + 2$ roots, so by Schwartz-Zippel they must be the zero polynomial. Hence $\operatorname{rank} M(z) \leq k$ for *all* z. In particular, the kernel dimension is at least 1 for all z and cannot "jump" above 1 at any point. This is a strong algebraic constraint that we formalize below. Roughly, it implies the existence of a single polynomial vector $q(z)$ that spans the kernel for all z (a "globally stable" kernel), and evaluating $q(z)$ at different z shows that $\mathcal{C}_k(A), \mathcal{C}_k(B)$ share a common kernel vector (hence A, B share an error-locator polynomial).

The following theorem captures that condition. We denote it KSP* because it's related to the "Kernel Stability of Pencils" concept used in decoders like the Berlekamp-Massey algorithm (the $\star$ is to distinguish it from a version that allows kernel jumps of bounded size, not needed here).

Throughout this argument we view z as a formal indeterminate and work over $\mathbb{F}_q[z]$ (equivalently $\mathbb{F}_q(z)$), so "k+2 roots" implies an identically–zero polynomial identity without a large–q assumption; we then specialize back to $\mathbb{F}_q$.

Theorem .1 (KSP* Algebraic Rigidity). *At equality $k = m$ the size-$m{-}k$ Hankel window is row-degenerate; we therefore state the condition in a* windowed key–equation *form. Fix any window length $\ell \in \{1, \dots, m-1\}$ and let $\mathcal{C}_\ell(\cdot)$ be the $(m-\ell) \times (\ell+1)$ Hankel window operator. Consider the pencil*

$$M_\ell(z) \;=\; \mathcal{C}_\ell(A) \;+\; z\,\mathcal{C}_\ell(B).$$

Suppose there exist $\ell + 2$ distinct challenges $z_0, \dots, z_{\ell+1}$ at which all $(\ell{+}1) \times (\ell{+}1)$ minors of $M_\ell(z)$ vanish. Then either

1. *(*kernel jump*) there is z_j with $\dim \operatorname{Ker} M_\ell(z_j) \geq 2$, or*

2. *(*constant kernel*) there exists a nonzero constant vector $q \in \mathbb{F}_q^{\ell+1}$ with $\mathcal{C}_\ell(A)q = \mathcal{C}_\ell(B)q = 0$.*

In case (ii)*, q is a common error-locator of degree $\leq \ell$, hence $A, B \in \operatorname{Span}(H_L)$ for some $|L| \leq \ell$, giving equality-case rigidity.*

Proof. If case (i) occurs, we already have a kind of rigidity: one particular h_{z_0} itself had two linearly independent degree-ℓ error-locator polynomials (for the chosen window), which is very restrictive. We focus on case (ii), so assume no such jump exists and that $\dim \operatorname{Ker} M_\ell(z) = 1$ for the $\ell{+}2$ specified challenges (and hence, by the minors argument below, for all $z \in \mathbb{F}_q$). We show this forces a common kernel vector.

Step 1 (Polynomial kernel section). All $(\ell{+}1)$-minors of $M_\ell(z)$ are polynomials in z of degree at most $\ell{+}1$. Since they vanish at the $\ell{+}2$ distinct points $z_0, \dots, z_{\ell+1}$, each such minor is the zero polynomial. Hence $\operatorname{rank} M_\ell(z) \leq \ell$ for every z, and *generically* $\operatorname{rank} M_\ell(z) = \ell$ (one-dimensional kernel). Choose any ℓ rows that span the row space (for generic z), forming an $\ell \times (\ell{+}1)$ submatrix $\widetilde{M_\ell}(z)$. By the cofactor construction, there is a nonzero polynomial vector $\mathbf{q}(z) \in \mathbb{F}_q[z]^{\ell+1}$ with $\deg q_i(z) \leq \ell$ such that

$$\widetilde{M_\ell}(z)\,\mathbf{q}(z) \;\equiv\; 0, \qquad \text{hence} \qquad M_\ell(z)\,\mathbf{q}(z) \;\equiv\; 0.$$

Step 2 (Degree vs. roots). Differentiate the polynomial identity $M_\ell(z)\,\mathbf{q}(z) \equiv 0$:

$$\mathcal{C}_\ell(B)\,\mathbf{q}(z) \;+\; M_\ell(z)\,\mathbf{q}'(z) \;\equiv\; 0.$$

At each of the ℓ+2 challenges z_j, we have $M_\ell(z_j)\,\mathbf{q}(z_j) = 0$, so plugging $z = z_j$ gives

$$\mathcal{C}_\ell(B)\,\mathbf{q}(z_j) \;=\; -\,M_\ell(z_j)\,\mathbf{q}'(z_j) \in (\text{column space of } M_\ell(z_j)).$$

If $\mathcal{C}_\ell(B)\,\mathbf{q}(z_j) \neq 0$ for some j, pick any nonzero $\lambda(z_j)^T \in \ker M_\ell(z_j)^T$; multiplying the differentiated identity by $\lambda(z_j)^T$ shows $\lambda(z_j)^T \mathcal{C}_\ell(B)\,\mathbf{q}(z_j) = 0$, which would produce a second independent kernel vector at z_j, contradicting "no jump". Therefore $\mathcal{C}_\ell(B)\,\mathbf{q}(z_j) = 0$ for all j. Since each coordinate of $\mathcal{C}_\ell(B)\,\mathbf{q}(z)$ is a polynomial of degree $\leq \ell$, vanishing at ℓ+2 distinct points implies

$$\mathcal{C}_\ell(B)\,\mathbf{q}(z) \;\equiv\; 0 \quad \text{as a polynomial vector,}$$

and substituting back into $M_\ell(z)\,\mathbf{q}(z) \equiv \mathcal{C}_\ell(A)\,\mathbf{q}(z) + z\,\mathcal{C}_\ell(B)\,\mathbf{q}(z) \equiv 0$ yields

$$\mathcal{C}_\ell(A)\,\mathbf{q}(z) \;\equiv\; 0.$$

Step 3 (Common annihilator). Thus for every z, $\mathbf{q}(z) \in \operatorname{Ker}\mathcal{C}_\ell(A) \cap \operatorname{Ker}\mathcal{C}_\ell(B)$. Pick any z_1 with $\mathbf{q}(z_1) \neq 0$ and set $\mathbf{q} := \mathbf{q}(z_1) \in \mathbb{F}_q^{\ell+1} \setminus \{0\}$. Then

$$\mathcal{C}_\ell(A)\,\mathbf{q} \;=\; \mathcal{C}_\ell(B)\,\mathbf{q} \;=\; 0.$$

Let $L(x) = q_0 + q_1 x + \cdots + q_\ell x^\ell$ be the polynomial with coefficient vector $\mathbf{q}$, and define its root index set

$$\Lambda \;:=\; \big\{\, i \in [n] : L(\alpha_i) = 0 \,\big\},$$

where columns of H are indexed by the evaluation points $\alpha_i \in D$. By the key–equation interpretation of $\operatorname{Ker}\mathcal{C}_\ell(\cdot)$, L is a common error–locator for the syndrome sequences A and B; hence the error supports associated to A and to B are contained in Λ, so $A, B \in \operatorname{Span}(H_\Lambda)$. Since $\deg L \leq \ell$, we have $|\Lambda| \leq \ell$, and because $\ell \in \{1, \ldots, m-1\}$, it follows that $|\Lambda| \leq m-1$. Therefore the entire line $s(z) = A + zB$ lies in the single proper subspace $\operatorname{Span}(H_\Lambda)$. This is equality–case rigidity (case (ii)). □

The KSP* theorem provides a precise algebraic meaning to "common structured core" at $t = d$. In practice, a protocol designer can try to enforce this by adding a check that $M(z)$ has no kernel jumps across challenges (e.g. by some rank consistency check), or simpler, by requiring the prover to commit to a single error-locator polynomial for all rounds.

Appendix F: Combinatorial Route: Sunflower and Helly-Type Rigidity

We now turn to the combinatorial classification of how the family of error support sets $\{T_z : z \in \mathcal{Z}^*\}$ might force a global structure. At $t = d$, each T_z can be of size m. The question is: if you have "too many" such m-subsets of $[n]$, must there be either a repeat or a large common intersection among them? This is reminiscent of classical results in extremal set theory.

Primer on the Sunflower Lemma. A **sunflower** (or Δ-system) is a family of sets $S_1, \ldots, S_r$ such that every pair of sets has exactly the same intersection C. That is, $S_i \cap S_j = C$ for all $i \neq j$. The common intersection C is called the *core* of the sunflower, and the sets $S_i \setminus C$ are called the petals. The Sunflower Lemma of Erdős and Rado (1960) states that for any fixed r, if you have a sufficiently large family of sets of size K, then it must contain a sunflower of r petals [4, 1].

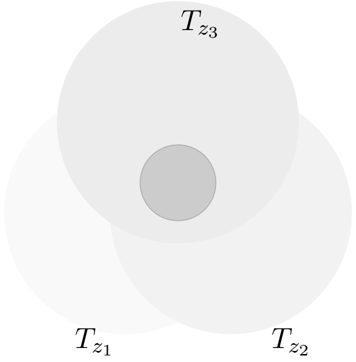

Figure 4: Sunflower-like overlaps among error supports: a large common core (dark gray) forces A, B into the span of the core's columns H_C, restoring rigidity.

Intuitively, a large collection of same-size sets cannot all avoid having a large common core - some r of them will fall into a sunflower structure.

We apply this principle to the family $\{T_z\}$ when $|\mathcal{Z}^*|$ is very large. The idea is: if no support repeats (so all T_z are distinct) and no r of them share a big core, then they behave somewhat like constant-weight codewords of length n and weight m with pairwise intersections bounded, which is reminiscent of a constant-weight code with minimum distance at least 4 for $r = 3$, etc. Known bounds (like the Plotkin bound in coding theory [11]) say you can't have too many such sets.

Combining these ideas yields the following Helly-type theorem for small-union supports (which we dub SU-Helly):

Theorem .2 (SU–Helly structure at equality)**.** *Assume the equality case $k = m$ (equivalently $t = d$) for an $[n, n-m, m+1]_q$ MDS (RS) code, so each error support has size m. Suppose the CA premise holds for a set of challenges $\mathcal{Z}^*$, yielding supports $\{T_z\}_{z\in\mathcal{Z}^*}$ with $|T_z| = m$. If $|\mathcal{Z}^*|$ is sufficiently large (for example, exceeding $A(n, 4, m) + m\binom{n}{m-1}$; Plotkin's bound gives an explicit exponential upper bound on $A(n, 4, m)$), then the family must contain either:*

1. ***Repeat support:*** *$T_{z_i} = T_{z_j}$ for some $i \neq j$; or*
2. ***Sunflower:*** *a subfamily of $m+1$ supports $T_{z_\ell} = C \cup \{p_\ell\}$ with a common core $|C| = m-1$ and pairwise distinct petals $p_\ell \notin C$.*

In case (1) (repeat), we obtain $A, B \in \mathrm{Span}(H_{T_{z_i}})$; if $|T_{z_i}| \leq m - 1$ (strict gap) this gives rigidity, while at equality $|T_{z_i}| = m$ it need not be a proper subspace. In case (2) (sunflower), a linear-algebraic trapping lemma (Lemma .3) yields equality-case rigidity under a mild non-degeneracy hypothesis.

Proof. If a repeat occurs we are in case (1) and are done by Theorem 6.2.

Otherwise, partition the supports by $(m-1)$-cores: for each $(m-1)$-subset $C \subseteq [n]$, let

$$\mathcal{F}_C := \{\, T \in \{T_z\} : C \subset T \,\}.$$

If some $\mathcal{F}_C$ has size $\geq m+1$, it yields a sunflower of size $m+1$ with core C (case (2)).

Otherwise each $\mathcal{F}_C$ has size $\leq m$. Let

$$\mathcal{G} \;:=\; \big\{\, T \in \{T_z\} : \ T \text{ shares no } (m-1)\text{-core with any other support} \,\big\}.$$

For distinct $T, T' \in \mathcal{G}$ we have $|T \cap T'| \leq m-2$, so their characteristic vectors have Hamming distance ≥ 4. Hence $|\mathcal{G}| \leq A(n, 4, m)$. The supports that do share some $(m-1)$-core are at most $\sum_C |\mathcal{F}_C| \leq m\binom{n}{m-1}$ (this upper bound may double count but is sufficient). Altogether,

$$|\mathcal{Z}^*| \;\leq\; |\mathcal{G}| + \sum_C |\mathcal{F}_C| \;\leq\; A(n,4,m) + m\binom{n}{m-1}.$$

Thus, if $|\mathcal{Z}^*|$ exceeds this bound, either a repeat (case (1)) or a sunflower (case (2)) must occur. In the sunflower branch, apply Lemma .3 to conclude equality-case rigidity under the non-degeneracy hypothesis. □

Lemma .3 (Sunflower core ⇒ common-core trapping under non-degeneracy)**.** *Let $H \in \mathbb{F}_q^{m\times n}$ be an MDS RS parity-check matrix, and assume $k = m$ (equivalently $t = d$). Suppose there exist $m+1$ distinct successes $z_0, \dots, z_m$ with supports*

$$T_{z_i} = C \cup \{p_i\}, \qquad |C| = m-1, \quad p_i \notin C \textit{ pairwise distinct.}$$

Write the syndromes as $s(z_i) = H_C u_i + \alpha_i h_{p_i}$ with $\alpha_i \in \mathbb{F}_q^\times$, and let ϕ^T span the left-nullspace of H_C (so $\phi^T H_C = 0$). Define $\beta_i := \alpha_i\, \phi^T h_{p_i} \in \mathbb{F}_q^\times$. If the three row vectors $(1, \dots, 1)$, $(z_0, \dots, z_m)$, and $(\beta_0, \dots, \beta_m)$ are linearly independent over $\mathbb{F}_q$ (equivalently, (β_i) is not an affine function of (z_i)), then

$$A, B \;\in\; \mathrm{Span}(H_C),$$

hence $s(z) = A + zB \in \mathrm{Span}(H_C)$ for all z (equality-case rigidity).

Remark .4 (Genericity heuristic for sunflower trapping)*.* The non-degeneracy hypothesis in Lemma .3 ensures the moment systems for μ, ν are consistent. In many practical settings (e.g., independently sampled challenges), the triples (z_i, β_i) are in general position with high probability, so the same trapping of A, B is expected to hold without explicitly checking $\mathrm{rank}(M) = 3$. We keep the lemma's rank-3 condition to avoid corner cases and for complete rigor.

Proof. Let $M = \begin{bmatrix} 1 & \cdots & 1 \\ z_0 & \cdots & z_m \\ \beta_0 & \cdots & \beta_m \end{bmatrix}$. The rank hypothesis implies $\mathrm{rank}(M) = 3$, so there exist $\mu, \nu \in \mathbb{F}_q^{m+1}$ with $M\mu = (0,1,0)^T$, $\qquad M\nu = (1,0,0)^T$. Then $\sum_i \mu_i s(z_i) = B$ and $\sum_i \nu_i s(z_i) = A$. Using $s(z_i) = H_C u_i + \alpha_i h_{p_i}$ and left-multiplying by ϕ^T gives $\sum_i \mu_i \beta_i = \sum_i \nu_i \beta_i = 0$, so both sums lie in $\mathrm{Span}(H_C)$, implying $B, A \in \mathrm{Span}(H_C)$. □

Appendix G: Miniature End-to-End Reed-Solomon Example

A concrete, step-by-step example of a Reed-Solomon proximity proof intended to make the abstract concepts of the paper—arithmetization, codeword commitment, and verification via the Syndrome-Space Lens—more tangible.

The Setup

First, the Prover and Verifier agree on a public set of parameters. We choose parameters small enough for manual calculation.

- **Finite Field:** $\mathbb{F}_7$. All arithmetic is performed modulo 7.

- **Evaluation Domain (D):** $D = \{1, 2, 3, 4, 5, 6\}$.
- **Code Length (n):** The size of the domain D, so $n = 6$.
- **Message Size (d):** The message will be a polynomial of degree less than 2, so $d = 2$.
- **Redundancy (m):** The number of parity checks is $m = n - d = 6 - 2 = 4$.

The Prover's Claim (Arithmetization)

The Prover takes their message, which consists of $d = 2$ numbers, and turns it into a polynomial.

- **The Message:** Let the Prover's message be the coefficients '[2, 1]'.
- **Arithmetization:** This corresponds to the polynomial $P(x) = 1x + 2$.
- **Evaluation:** The Prover evaluates $P(x)$ on the entire domain D to create their codeword commitment, the vector $\mathbf{c}$.

$$\begin{aligned} P(1) &= 1 + 2 = 3 \\ P(2) &= 2 + 2 = 4 \\ P(3) &= 3 + 2 = 5 \\ P(4) &= 4 + 2 = 6 \\ P(5) &= 5 + 2 = 7 \equiv 0 \pmod 7 \\ P(6) &= 6 + 2 = 8 \equiv 1 \pmod 7 \end{aligned}$$

The final codeword vector that the Prover commits to is $\mathbf{c} = (\mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{6}, \mathbf{0}, \mathbf{1})$.

The 'Rulebook' (Parity-Check Matrix H)

The public parameters also define the "rulebook"—the Parity-Check Matrix H that all valid codewords must obey. A vector is a valid codeword if and only if $Hc^T = 0$. For a Reed-Solomon code with evaluation domain $D = \{\alpha_1, \dots, \alpha_n\}$ and m parity checks, the $m \times n$ matrix H is constructed as follows:

$$H_{j,i} = \alpha_i^j \quad \text{for } j \in \{0, \dots, m-1\}, i \in \{1, \dots, n\}$$

For our example, this yields the 4×6 matrix:

$$H = \begin{pmatrix} 1^0 & 2^0 & 3^0 & 4^0 & 5^0 & 6^0 \\ 1^1 & 2^1 & 3^1 & 4^1 & 5^1 & 6^1 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \\ 1^3 & 2^3 & 3^3 & 4^3 & 5^3 & 6^3 \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 2 & 4 & 1 \\ 1 & 1 & 6 & 1 & 6 & 6 \end{pmatrix} \pmod 7$$

The Verifier's Check

Imagine the Prover makes a single mistake, reporting the value at position 3 as '0' instead of '5'. The Verifier receives the corrupted vector $\mathbf{c_{corr}} = (\mathbf{3}, \mathbf{4}, \mathbf{0}, \mathbf{6}, \mathbf{0}, \mathbf{1})$. The Syndrome-Space Lens provides a direct method to detect this error. The Verifier computes the syndrome, $\mathbf{s} = Hc_{corr}^T$:

$$\mathbf{s} = Hc_{corr}^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 2 & 4 & 1 \\ 1 & 1 & 6 & 1 & 6 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ 0 \\ 6 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 3+4+0+6+0+1 \\ 3+8+0+24+0+6 \\ 3+16+0+12+0+1 \\ 3+4+0+6+0+6 \end{pmatrix} = \begin{pmatrix} 14 \\ 41 \\ 32 \\ 19 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 6 \\ 4 \\ 5 \end{pmatrix} \pmod 7$$

The result is **not the zero vector**. The non-zero entries are the syndrome—a unique fingerprint of the error. The Verifier does not need to know the original polynomial. They only need to see a non-zero syndrome to know the received vector is invalid, and they therefore **reject** the proof. This demonstrates the power of the global, linear-algebraic check provided by the Syndrome-Space Lens.

References

[1] R. Alweiss, S. Lovett, K. Wu, and J. Zhang. Improved bounds for the sunflower lemma. *Ann. of Math.*, 194(3):795–815, 2021.

[2] E. Ben-Sasson, I. Bentov, Y. Horesh, and M. Riabzev. Scalable, Transparent, and Post-Quantum Secure Computational Integrity. *IACR ePrint*, 2018/046, 2018.

[3] E. Ben-Sasson, S. Kopparty, and S. Saraf. Proximity Gaps for Reed-Solomon Codes. In *Proc. 61st IEEE Symp. Foundations of Computer Science (FOCS)*, 2020.

[4] P. Erdős and R. Rado. Intersection theorems for systems of sets. *J. London Math. Soc.*, 35(1):85–90, 1960.

[5] V. Guruswami and M. Sudan. Improved Decoding of Reed-Solomon and Algebraic-Geometric Codes. *IEEE Trans. Inform. Theory*, 45(6):1757–1767, 1999.

[6] Anonymous Author(s). "Proximity Gaps in Interleaved Codes," *IACR ePrint* 2024/1351, 2024. Available at https://eprint.iacr.org/2024/1351.

[7] S. Johnson. A New Upper Bound for Error-Correcting Codes. *IRE Trans. Inform. Theory*, 8(3):203–207, 1962.

[8] S. Kopparty, Z. Li, and A. Restivo. "From Linear-Algebraic to Analytic Decoding: What Resists List Decoding at Capacity?" *IACR ePrint* 2024/1810, 2024. Available at https://eprint.iacr.org/2024/1810.

[9] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes.* North-Holland, 1977.

[10] M. Tsfasman and S. Vladut. *Algebraic-Geometric Codes.* Kluwer Academic Publishers, 1991.

[11] M. Plotkin. Binary Codes with Specified Minimum Distance. *IRE Trans. Inform. Theory*, 6(4):445–450, 1960.