Sequential Detection of Algebraic Coherence in Curve-Decoding

A statistical bridge from flight-path mixtures to proximity gaps

Jonathan R. Landers

July 2026

Abstract

This note develops a statistical counterpart to curve-decodability by carrying over the full logic of a recent flight-path mixture analysis. In that analysis, the product model $D = AQ$ first predicts that broad lognormal scale variation should suppress the non-Gaussian cumulants of bridge shape; the observed skewness and kurtosis sharply violate that one-regime prediction, forcing a heterogeneous or boundary mechanism. The argument then reverses direction: persistent small residuals identify which individual trajectories belong to the boundary population. Curve-decoding has the complementary structural statement that many local codeword matches force one common low-degree codeword curve. We make the analogy quantitative. For decoder outputs at parameters $\alpha_1, \ldots, \alpha_m$, an algebraic residual is the syndrome rank of their row span relative to a Reed–Solomon evaluation space. It vanishes exactly on a degree-ℓ codeword curve. If the ith decoder output has conditional point mass at most q^{-h_i}, even after adaptive dependence and local-proximity conditioning, then accidental coherence beyond the interpolation barrier has probability at most $q^{-\sum_{i=\ell+2}^{m} h_i}$. A scan version controls false coherent subsets. The iid-uniform model yields the exact random-matrix rank law, a Wasserstein detectability curve, and a posterior waiting-time rule. The resulting picture joins three steps: reject a false single mechanism, identify a latent mechanism from persistent local evidence, and measure when algebraic coherence becomes statistically compelling.

1 Narrative arc: from one regime to two, then back to one

Key conceptual points

- The flight-path note first derives what one mechanism must look like, then rejects it from the observed distribution.
- After heterogeneity is forced, persistent local evidence can identify which mechanism generated an individual trajectory.
- Curve-decodability begins at the third step: repeated local matches force one global algebraic explanation. This note adds the statistical evidence scale.

The full connection is easiest to see as a three-stage argument rather than a single analogy.

Stage I: a one-regime model makes a sharp distributional prediction

In the flight-path note [2], each RMS departure from an endpoint geodesic is factored as

$$D = AQ,$$

where A is a route-scale variable and Q is the dimensionless RMS shape of an endpoint-pinned bridge. In log space,

$$\log(D/d_0) = U + V, \qquad U = \log(A/d_0), \quad V = \log Q.$$

If $U \sim N(\mu, \sigma_A^2)$ is independent of V, then adding the broad Gaussian scale term smooths the shape term. Writing $\tau_Q^2 = \mathrm{Var}(V)$ and

$$\rho = \frac{\tau_Q^2}{\sigma_A^2 + \tau_Q^2},$$

the standardized higher cumulants satisfy the exact attenuation identities

$$\mathrm{Skew}(\log D) = \mathrm{Skew}(\log Q)\rho^{3/2}, \qquad \mathrm{ExKurt}(\log D) = \mathrm{ExKurt}(\log Q)\rho^2. \tag{4}$$

More generally, if G is the Gaussian moment match to V and W is the centered, variance-one version of $U + V$, then

$$d_{W_2}(W, N(0,1)) \le \frac{d_{W_2}(V, G)}{\sqrt{\sigma_A^2 + \tau_Q^2}}. \tag{5}$$

Thus a broad independent scale does not merely make the histogram look smoother; it gives a quantitative single-population prediction.

For the Brownian-bridge benchmark,

$$Q^2 = \sum_{j\geq 1} \frac{\xi_j^2}{\pi^2 j^2}, \qquad \xi_j \overset{\text{iid}}{\sim} N(0,1),$$

the estimated shape cumulants are modest. Matching the observed log-variance gives the one-regime predictions

$$\mathrm{Skew}(\log D) \approx 0.003, \qquad \mathrm{ExKurt}(\log D) \approx -0.002.$$

The observed cohort instead has

$$\mathrm{Skew}(\log D) = -2.40, \qquad \mathrm{ExKurt}(\log D) = 8.11. \tag{6}$$

The gap is far too large to be ordinary bridge-shape nonnormality surviving the smoothing in (4). Therefore at least one single-regime assumption fails: homogeneity, independence of A and Q, Gaussianity of $\log A$, or the Brownian law for Q. The visible cluster of 53 tracks at or below 0.1 km RMS supports a distinct near-geodesic boundary component, leading to

$$f_D(d) = \pi_0 f_0(d) + \sum_{c=1}^{C} \pi_c f_D(d \mid c). \tag{7}$$

This is the first direction of the story:

$$\boxed{\text{a proposed global mechanism predicts the wrong distribution} \implies \text{more than one mechanism is needed.}}$$

Stage II: persistent local evidence identifies one of the mechanisms

Once the mixture is forced at the cohort level, the question becomes individual and sequential. For the first n observations of one trajectory, define

$$R_n^2 = \frac{1}{n}\sum_{j=1}^{n} \eta(t_j)^2.$$

Let $C = 0$ denote the near-geodesic boundary component and $C = 1$ the ordinary bridge component. If, for some tolerance τ,

$$\Pr_1(R_n \leq \tau) \leq e^{-nI(\tau)}, \qquad \Pr_0(R_n \leq \tau) \geq 1 - \beta_n,$$

then Bayes' rule gives

$$\Pr(C = 0 \mid R_n \leq \tau) \geq \frac{\pi_0(1-\beta_n)}{\pi_0(1-\beta_n) + (1-\pi_0)e^{-nI(\tau)}}. \tag{8}$$

The point is not that every straight-looking prefix comes from the boundary regime. It is that continued near-zero residual energy spends exponentially decreasing probability under the ordinary regime. Sufficient persistence therefore turns local observations into a global mechanism assignment.

Stage III: curve-decoding supplies the algebraic local-to-global analogue

Goyal, Guruswami, Sun, and Wootters study an ambient degree-ℓ curve

$$u(x) = \sum_{j=0}^{\ell} u_j x^j \in \mathbb{F}_q^n$$

and codewords $f(\alpha) \in C$ close to $u(\alpha)$ for many parameters α [1]. Curve-decodability says that enough local decoding successes force many of those codewords to share one global explanation,

$$f(\alpha) = c(\alpha), \qquad c(x) = \sum_{j=0}^{\ell} c_j x^j, \quad c_j \in C. \tag{9}$$

Their row-span constrained local-property framework proves that this coherent subset must exist. The present note asks the complementary statistical question: once coherence appears, how unlikely would it have been under an accidental decoder mechanism?

The two rarity rates line up directly:

$$\boxed{\Pr_1(R_n \le \tau) \lesssim e^{-nI(\tau)} \qquad \longleftrightarrow \qquad \Pr_{\text{acc}}(r_m = 0) \lesssim q^{-\sum_{i=\ell+2}^{m} h_i}.} \tag{10}$$

Here $I(\tau)$ is the ordinary-flight small-ball rate, while h_i is the conditional min-entropy of the next decoder output after previous outputs and local-proximity information are known. In both problems, repeated local agreement becomes evidence for one latent mechanism only after the null mechanism has spent enough probability mass.

2 Algebraic residual and interpolation barrier

Key conceptual points

- $r_m = 0$ means one degree-ℓ codeword curve explains all observed decoder outputs.
- Positive rank counts independent algebraic directions that the proposed curve cannot explain.
- Any $\ell + 1$ observations can be interpolated, so evidence begins only at observation $\ell + 2$.

Let $C \subseteq \mathbb{F}_q^n$ be an $[n, k]$ linear code, let $\alpha_1, \ldots, \alpha_m$ be distinct, and let $f_i \in C$. Form

$$M_m = [f_1 \ \cdots \ f_m], \qquad U_m = \mathrm{RowSpan}(M_m) \subseteq \mathbb{F}_q^m,$$

and the degree-ℓ evaluation space

$$\mathrm{RS}_\ell(A_m) = \{(p(\alpha_1), \ldots, p(\alpha_m)) : \deg p \le \ell\}.$$

Let $H_{\ell,m}$ be a full-row-rank parity-check matrix for this evaluation code, with $s = (m - \ell - 1)_+$ rows.

Definition 1 (Algebraic coherence residual)**.**

$$r_m := \mathrm{rank}(M_m H_{\ell,m}^{\top}) = \dim U_m - \dim(U_m \cap \mathrm{RS}_\ell(A_m)).$$

Proposition 2 (Exact coherence test)**.** *For distinct α_i,*

$$r_m = 0 \Longleftrightarrow \exists\, c_0, \ldots, c_\ell \in C : \quad f_i = \sum_{j=0}^{\ell} c_j \alpha_i^j \quad (1 \le i \le m).$$

In particular, $r_m = 0$ for every choice of $f_1, \ldots, f_m$ when $m \le \ell + 1$.

Proof. The syndrome vanishes exactly when every row of M_m lies in $\mathrm{RS}_\ell(A_m)$, so every coordinate sequence is generated by a scalar polynomial of degree at most ℓ. Collecting the coordinatewise coefficients gives vectors $c_0, \ldots, c_\ell$. Vector-valued interpolation expresses each c_j as a linear combination of observed columns, hence $c_j \in C$ because C is linear. Conversely, such coefficient vectors make every row a degree-ℓ evaluation vector and therefore force zero syndrome. When $m \le \ell + 1$, interpolation always exists. □

The proposition creates a sharp diagnostic threshold. Before $\ell + 2$ observations, coherent and accidental data are algebraically indistinguishable. At $\ell + 2$, the fitted curve makes its first nontrivial prediction.

3 Decoder-conditioned accidental coherence

Key conceptual points

- Independence is not required: the decoder may be adaptive, biased, and conditioned on local proximity.
- Once $\ell + 1$ observations fix a curve, every later coherent output must equal one specific codeword.
- The false-coherence probability is therefore the product of the decoder's largest allowable conditional point masses.

Let $F_1, \dots, F_a$ be C-valued random decoder outputs with filtration $\mathcal{F}_i = \sigma(F_1, \dots, F_i)$. Their joint law may already be conditioned on local successes such as $\Delta(u(\alpha_i), F_i) \le \delta$. Define

$$\lambda_i := \operatorname{ess\,sup} \max_{c \in C} \Pr(F_i = c \mid \mathcal{F}_{i-1}), \qquad h_i := -\log_q \lambda_i.$$

Thus h_i is a conditional min-entropy in q-ary units. For $B = \{i_1 < \cdots < i_b\} \subseteq [a]$, let r_B be the residual formed from the selected columns and parameters, and define

$$S_{a,b} := \min_{B \subseteq [a],\ |B| = b} r_B.$$

Theorem 3 (Adaptive decoder-conditioned coherence bound)**.** *For $m > \ell + 1$,*

$$\boxed{\Pr(r_m = 0) \le \prod_{i=\ell+2}^{m} \lambda_i = q^{-\sum_{i=\ell+2}^{m} h_i}.} \tag{1}$$

More generally, for $b > \ell + 1$,

$$\Pr(S_{a,b} = 0) \le \sum_{B = \{i_1 < \cdots < i_b\} \subseteq [a]} \prod_{j=\ell+2}^{b} \lambda_{i_j}. \tag{2}$$

If $h_i \ge h$ for all i, then

$$\boxed{\Pr(S_{a,b} = 0) \le \binom{a}{b} q^{-h(b-\ell-1)}.} \tag{3}$$

Proof. The first $\ell + 1$ pairs (α_i, F_i) determine a unique degree-ℓ codeword curve $c(x)$. On the event $r_m = 0$, every later output must equal the single prescribed value $c(\alpha_i)$. Conditional on $\mathcal{F}_{i-1}$ and the previous coherence equalities, this value is measurable, so its conditional probability is at most λ_i. Iterating the chain rule gives (1).

For a fixed subset $B = \{i_1 < \cdots < i_b\}$, the first $\ell + 1$ selected outputs determine the candidate curve. Applying the same conditional argument to the remaining selected indices gives

$$\Pr(r_B = 0) \le \prod_{j=\ell+2}^{b} \lambda_{i_j}.$$

A union bound over all B proves (2), and $\lambda_i \le q^{-h}$ gives (3). □

Remark 4 (Where the code and decoder enter)**.** *If the conditional decoder law is at most q^{γ_i} times uniform on C, then $h_i \ge k - \gamma_i$ and*

$$\Pr(r_m = 0) \le q^{-\sum_{i=\ell+2}^{m}(k-\gamma_i)}.$$

The entropy loss γ_i records how strongly local proximity, code geometry, and decoder bias concentrate the next output. Estimating this term is the code-family-specific problem left by the general theorem.

4 Exact calibration, transport, and waiting time

Key conceptual points

- Independent uniform codewords are an exact baseline, not the final decoder model.
- In that baseline, the entire residual distribution is a classical finite-field random-matrix rank law.
- The resulting detectability curve is flat before the interpolation barrier and then separates exponentially fast.

Under the calibration null $F_i \overset{\text{iid}}{\sim} \mathrm{Unif}(C)$, one has $h_i = k$. This case isolates the algebraic cost of accidental interpolation before any decoder-specific concentration is introduced.

Proposition 5 (Exact accidental-coherence rank law)**.** *Let $s = (m - \ell - 1)_+$. Under the iid-uniform null, r_m has the same law as the rank of a uniformly random matrix in $\mathbb{F}_q^{k\times s}$. Consequently, for $0 \le r \le \min(k, s)$,*

$$\Pr(r_m = r) = \frac{\prod_{i=0}^{r-1}(q^k - q^i)(q^s - q^i)}{q^{ks}\prod_{i=0}^{r-1}(q^r - q^i)},$$

and, for $m > \ell + 1$,

$$\boxed{\Pr(r_m = 0) = q^{-k(m-\ell-1)}.}$$

Proof. Write $M_m = GZ$, where $G \in \mathbb{F}_q^{n\times k}$ is a full-column-rank generator matrix and $Z \in \mathbb{F}_q^{k\times m}$ has independent uniform columns. Since multiplication by G is injective,

$$\operatorname{rank}(M_m H_{\ell,m}^{\mathsf{T}}) = \operatorname{rank}(Z H_{\ell,m}^{\mathsf{T}}).$$

Each row of Z is uniform in $\mathbb{F}_q^m$, and multiplication by the surjective linear map $H_{\ell,m}^{\mathsf{T}} : \mathbb{F}_q^m \to \mathbb{F}_q^s$ sends it to a uniform vector in $\mathbb{F}_q^s$. The rows remain independent, so $Z H_{\ell,m}^{\mathsf{T}}$ is a uniform $k \times s$ matrix. Counting matrices of rank r gives the displayed formula. Rank zero means the entire $k \times s$ matrix vanishes, which has probability q^{-ks}. □

The exponent admits an elementary reading. The first $\ell + 1$ observations determine a unique degree-ℓ codeword curve. Every additional independent uniform codeword must hit one prescribed value in a set of size $|C| = q^k$, costing a factor q^{-k}. Thus the exact accidental-coherence rate is $k \log q$ per observation after the interpolation barrier.

Corollary 6 (Transport separation)**.** *Let $B_m = \mathbf{1}\{r_m > 0\}$. Under a coherent mechanism $B_m = 0$ almost surely, while under the iid-uniform accidental mechanism*

$$W_1(\mathcal{L}(B_m \mid H_{\mathrm{coh}}), \mathcal{L}(B_m \mid H_{\mathrm{acc}})) = \begin{cases} 0, & m \le \ell + 1, \\ 1 - q^{-k(m-\ell-1)}, & m > \ell + 1. \end{cases}$$

More generally, with $X_m = r_m / \min(k, s)$,

$$W_p(\delta_0, \mathcal{L}(X_m \mid H_{\mathrm{acc}})) = \left(\mathbb{E}_{\mathrm{acc}} X_m^p\right)^{1/p},$$

where the expectation is explicit from the rank law.

The binary Wasserstein curve is deliberately simple: it is exactly zero while interpolation is automatic, then rises exponentially toward one. The full-rank residual retains more information and can compare mechanisms whose failure is gradual rather than merely zero versus nonzero.

Corollary 7 (Posterior waiting time)**.** *With prior $\pi = \Pr(H_{\mathrm{coh}})$, observing $r_m = 0$ gives*

$$\Pr(H_{\mathrm{coh}} \mid r_m = 0) \ge \frac{\pi}{\pi + (1-\pi)q^{-\sum_{i=\ell+2}^{m} h_i}}.$$

If $h_i \ge h$, posterior confidence at least $1 - \alpha$ is guaranteed once

$$\boxed{m \ge \ell + 1 + \frac{\log\frac{1-\pi}{\pi} + \log\frac{1-\alpha}{\alpha}}{h \log q}.}$$

This is the coding analogue of the flight-path waiting-time theorem. Short prefixes are nondiagnostic because interpolation is free. After that point, every further zero-syndrome observation spends conditional probability under the accidental model, and the accumulated expenditure becomes posterior evidence for one coherent mechanism.

5 Relation to curve-decodability

Key conceptual points

- Curve-decodability asks whether at least one large subset of local matches has a shared polynomial explanation; $S_{a,b} = 0$ is exactly that event.
- The proximity theorem supplies existence of coherence. The entropy theorem supplies a finite-sample significance scale for that coherence.
- The two results are not competitors: one is structural, the other inferential.

The row-span constrained framework of [1] proves that sufficiently many local proximity events force a coherent subset for random linear, random Reed–Solomon, and random LDPC ensembles. In the present notation, its global conclusion is $S_{a,b} = 0$. The roles of the two statements are therefore

Curve-decodability:	local proximity forces a coherent subset;
entropy-rate theorem:	that subset is costly under an accidental decoder law.

The first statement answers an existence question. The second asks how much evidence the observed coherence carries against a competing stochastic explanation. In particular, deterministic curve-decodability can hold while the statistical distinction remains weak if decoder conditioning makes h_i small; conversely, a large entropy rate can make a modest amount of coherence highly diagnostic.

This is the precise connection to [2]. There, the ordinary flight population can mimic the near-geodesic regime for a short time, but its small-ball mass decays. Here, accidental decoder outputs can mimic a codeword curve through $\ell + 1$ observations, but their conditional coherence mass decays after the interpolation barrier. Both analyses separate three objects that are often conflated: a local residual, a null rarity rate, and a global mechanism assignment.

A useful conceptual reading is that the row span plays the role of a sufficient geometric state. It stores the cross-coordinate relationships needed to decide whether the observed columns admit a common polynomial source, just as the accumulated prefix residual stores the persistence needed to distinguish a boundary flight from an ordinary bridge.

6 Directions

Key conceptual points

- The main coding-theory task is to estimate h_i for actual ensembles and actual decoders.
- Approximate coherence should replace exact zero syndrome by low-rank or low-weight syndrome events.
- A genuine sequential version should control false discoveries under optional stopping, not only at a fixed sample size.

Decoder-specific entropy. Bound h_i after conditioning on $\Delta(u(\alpha_i), F_i) \le \delta$ for random linear, random Reed–Solomon, and LDPC ensembles. Such a result would turn (1) into a code-, radius-, and decoder-dependent exponent and reveal when proximity conditioning destroys most of the nominal k units of entropy.

Approximate and partial coherence. Replace $r_B = 0$ by $r_B \le \tau$, or by a low-weight syndrome event, to model imperfect algebraic agreement. For subset scans, exploit the overlap structure among interpolating sets to improve the union bound in (3); this is the natural place for second-moment, dependency-graph, or matroid methods.

Sequential validity. The product in (1) suggests a likelihood ratio or e-process whose stopping rule remains valid under optional stopping. One would then seek finite false-alarm control and expected detection-delay bounds, completing the analogy with sequential small-ball diagnosis rather than merely borrowing its posterior form.

Ensemble detectability. Code families with similar asymptotic proximity gaps may have different finite-m residual laws. Comparing

$$W_p(\mathcal{L}(r_m \mid C_1), \mathcal{L}(r_m \mid C_2))$$

or the first m at which coherent and accidental laws separate beyond sampling uncertainty could expose finite-length behavior hidden by asymptotic thresholds.

The conceptual endpoint is simple: a proximity theorem says that local agreements cannot remain arbitrary; the entropy theorem measures the statistical cost of pretending that the resulting global structure was accidental.

References

[1] R. Goyal, V. Guruswami, Y. Sun, and M. Wootters, "Locality of Curve-Decoding and Improved Proximity Gaps," *arXiv preprint arXiv:2607.08516*, 2026. https://arxiv.org/abs/2607.08516.

[2] J. R. Landers, "When Does a Random-Scale Flight Path Look Lognormal? A Spherical Bridge Note from OpenSky Trajectories," GitHub repository, 2026. https://github.com/jonland82/aerostatsome.

[3] C. Villani, *Optimal Transport: Old and New*. Springer, 2009.