# RECOVERY THRESHOLDS FOR IDEALIZED BEST-OF-$N$ RETRAINING

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Repeated best-of-$N$ distillation changes the model distribution over time, so rare errors may persist and the proportions of correct answers may drift. We study when mixing a fixed reference distribution into each retraining round guarantees recovery of the desired distribution. When all correct answers share the same score distribution, we give an exact necessary-and-sufficient condition for removing incorrect answers, based on a rare-error multiplier that measures whether selection amplifies a rare error. This shows that pairwise verifier accuracy alone does not determine long-run recovery. When correct answers have different score distributions, recovery further requires selection to preserve the reference distribution. Under this condition, we prove matching upper and lower bounds showing that, for fixed $N$, the worst-case reference weight needed for global recovery scales as $\Theta(\delta^2)$, where $\delta$ measures score-distribution differences. We also give sufficient conditions for the general heterogeneous setting, with exact or approximate recovery depending on preservation error.

## 1 INTRODUCTION

Verifiers and reward models can improve language-model output by selecting among candidate responses. Best-of-$N$ sampling draws $N$ responses and returns the one with the highest score, but it requires generating and scoring the entire candidate set. Best-of-$N$ distillation trains the model on selected responses to reduce inference cost (Sessa et al., 2025). In iterative methods, the updated model generates candidates for the next round (Sessa et al., 2025; Yang et al., 2025).

In repeated distillation, the responses selected in one round change the candidate distribution in later rounds. Incorrect answers that occasionally receive high scores may therefore persist. Even when all answers are correct, selection can favor some valid solutions over others without improving correctness. A curated collection of correct responses can define the desired proportions among these solutions, and we represent those proportions by a fixed reference distribution $q$. We can mix this reference into each retraining round to counteract these changes, as in models of curated recursive training (Ferbach et al., 2024; Zhao et al., 2025). To recover the reference, the process must remove incorrect answers and restore the desired proportions.

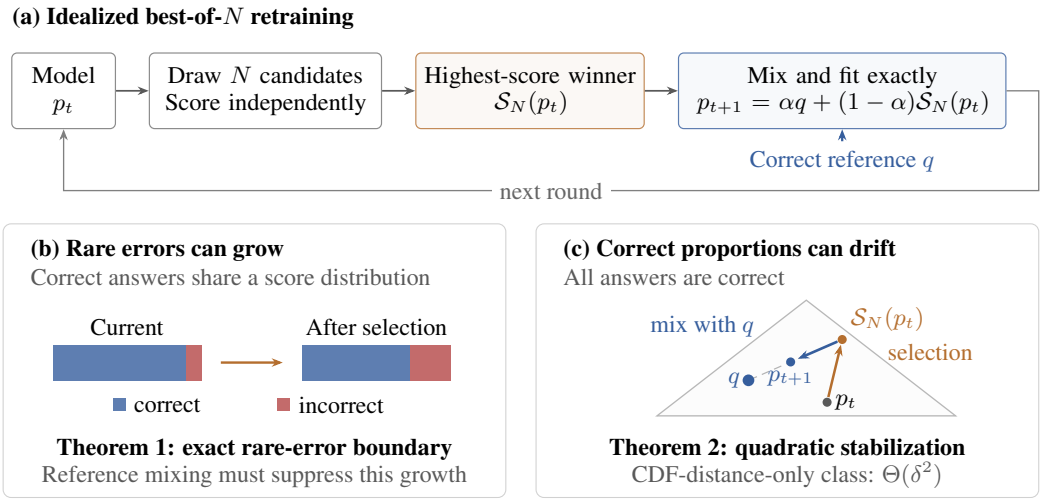
We study a population model on a finite answer space. Each sampled candidate receives an independent score from a fixed continuous distribution associated with its answer. At each round, the model exactly fits a mixture of the winner distribution and the reference. This setting isolates failures of recovery caused by repeated selection, even with fixed score distributions and exact fitting. We ask: *when does reference mixing guarantee recovery of the reference from every initialization, and how does the required reference weight depend on the score distributions?*

Existing finite-pool contraction results guarantee convergence when reference mixing is sufficiently strong (Zhao et al., 2025). They do not identify how score distributions affect the recovery boundary. We characterize this boundary under common correct scores and study how heterogeneous scores affect recovery of correct-answer proportions. In particular, high pairwise accuracy can coexist with amplification of rare errors.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

**Our main theoretical contributions are:**

- **Exact error-removal boundary.** When correct answers share a score distribution, we characterize recovery from a specified initialization and derive the corresponding condition for global recovery. A rare-error multiplier measures how strongly selection amplifies an incorrect answer when its probability is small. The exact condition includes the critical boundary and shows that pairwise accuracy alone does not determine recovery for three or more candidates.
- **Stabilization depends on the score class.** When all answers are correct and the reference assigns positive probability to each answer, selection must preserve the reference for recovery. The worst-case mixing scale is $\Theta(\delta^2)$ in the cumulative distribution function (CDF) distance class at fixed $N$. The lower bound uses score distributions with increasing oscillation frequency. On a fixed bounded score interval with uniformly bounded density derivatives, the worst-case threshold is cubic at $N = 2$. For fixed $N \geq 3$ and sufficiently small $\delta$, every positive reference weight suffices.
- **Joint recovery under heterogeneous scores.** When both correct and incorrect answers have heterogeneous score distributions, we give a sufficient condition for rare-error suppression and stable correct-answer proportions. Under this condition, exact preservation gives recovery of the reference, while small preservation errors give convergence to a nearby fixed point.




**Theorem 5: a joint sufficient condition**

Figure 1: Idealized best-of-$N$ retraining and two obstacles to recovery (schematic). (a) With reference weight $0 < \alpha < 1$, the current distribution generates candidates, and independent scores from fixed distributions select a winner. The update exactly fits the mixture of the winner distribution and the correct reference. (b) Selection can amplify rare incorrect answers; Theorem 1 gives the exact recovery boundary when correct answers share a score distribution. (c) Correct-answer proportions can drift even when all answers are correct. Theorem 2 gives the worst-case quadratic scale for preserved full-support targets in the CDF-distance-only class. This scale holds for fixed $N$ and vanishing $\delta$. Theorem 5 gives a sufficient condition that controls both effects.

**Related work.** For finite-pool curation with an arithmetic reference mixture, Ferbach et al. (2024) establish preference improvement and deviation bounds. Zhao et al. (2025) analyze several curation regimes. Under stationary preference noise, their finite-pool mixed result gives contraction in total variation (TV) when $N(1-\alpha) < 1$. They also analyze robustness to reward perturbations. We identify recovery thresholds that depend on the score distributions and show that a fixed update can fail to converge. Appendix G compares the assumptions and conclusions.

Table 1: Comparison with closely related finite-pool models. Recursive rows use arithmetic reference mixing. The Zhao row concerns the finite-pool mixed regime.

| Work | Feedback regime | Selection model | Main guarantee |
|---|---|---|---|
| Ferbach et al. (2024) | Recursive | Finite Plackett–Luce curation | Preference improvement; deviation bounds |
| Zhao et al. (2025) | Recursive | Stationary preference noise | TV contraction if $N(1-\alpha) < 1$; reward-perturbation robustness |
| Dorner et al. (2026) | Fixed generator | Imperfect verifier scores | Best-of-$N$ correctness |
| **This work** | Recursive | Independent scalar scores | Exact common-score boundary; CDF-class $\Theta(\delta^2)$ stabilization for fixed $N$ under preservation |

Recent inference-time theory studies verifier receiver operating characteristic (ROC) curves (Dorner et al., 2026) and Kullback–Leibler (KL) divergence and regret for soft best-of-$N$ (Aminian et al., 2026). Other work studies win-rate optimality in the presence of reward hacking (Sriraman & Block, 2026). Somerstep et al. (2026) compare best-of-$N$ with supervised fine-tuning for bit-string generation. These works address selection quality or adaptation costs. We study how selection changes the generator used in later rounds. Related analyses of recursive training study losses in quality and diversity (Alemohammad et al., 2024) and verification under confidence calibration (Fu et al., 2025). Other recursive-training work examines linear regression (Yi et al., 2026). Large-pool curation with a shared sampled rater (Falahati et al., 2026) uses a different sampling model from the finite pools and independent candidate scores considered here.

## 2 Problem Setup and Main Results

### 2.1 Setting and recovery

**Answers and reference.** Let $\mathcal{Y}$ be a finite answer space, with correct-answer set $\mathcal{C}$ and incorrect-answer set $\mathcal{B} = \mathcal{Y} \setminus \mathcal{C}$. Write $\Delta(\mathcal{Y})$ for the probability distributions on $\mathcal{Y}$. At round $t$, the model has distribution $p_t \in \Delta(\mathcal{Y})$; $p_{t,y}$ is the probability of answer $y$. A fixed reference $q \in \Delta(\mathcal{Y})$, supported on $\mathcal{C}$, specifies the desired correct-answer proportions.

**Scores and selection.** Answer $y$ has a fixed continuous score cumulative distribution function (CDF) $F_y$, where $F_y(s)$ is the probability of a score at most $s$. Draw $N \geq 2$ answers independently from $p$, then draw their scores independently conditional on their labels and return the highest-scoring answer. Repeated answers receive independent scores; continuity excludes ties. A random candidate has score CDF $F_p(s) = \sum_{y\in\mathcal{Y}} p_y F_y(s)$. The winner distribution $\mathcal{S}_N(p)$ satisfies

$$\mathcal{S}_N(p)_y = p_y W_y(p), \qquad W_y(p) = N \int_{\mathbb{R}} F_p(s)^{N-1} \, \mathrm{d}F_y(s). \tag{1}$$

The weight $W_y(p)$ multiplies $p_y$; the integral is an expectation under the score law $F_y$.

**Retraining and recovery.** Fix a reference weight $0 < \alpha < 1$ and an initial distribution $p_0$. Mixing the winner distribution with $q$ and fitting this mixture exactly gives the iteration in Figure 1(a):

$$p_{t+1} = T_\alpha(p_t) := \alpha q + (1-\alpha)\mathcal{S}_N(p_t), \qquad t = 0, 1, \ldots. \tag{2}$$

This is exact population cross-entropy fitting with freely parameterized answer probabilities. Appendix A derives both identities; Appendix H details the assumptions. At $\alpha = 1$, $p_1 = q$.

For distributions $p, r$, define their total variation distance by $\mathrm{TV}(p, r) = \frac{1}{2}\sum_{y\in\mathcal{Y}} |p_y - r_y|$. We say that the iteration *recovers $q$ from $p_0$* if

$$\lim_{t\to\infty} \mathrm{TV}(p_t, q) = 0.$$

Recovery is *global* if this limit holds for every $p_0 \in \Delta(\mathcal{Y})$. The incorrect-answer probability is $z_t := \sum_{j\in\mathcal{B}} p_{t,j} = p_t(\mathcal{B})$. Recovery implies $\lim_{t\to\infty} z_t = 0$, but this alone need not restore the correct-answer proportions.

### 2.2 RECOVERY UNDER COMMON CORRECT SCORES

**Assumption (common correct-score law).** There is one continuous CDF $G$ such that $F_i = G$ for every correct answer $i \in \mathcal{C}$. For each incorrect answer $j \in \mathcal{B}$, denote its score CDF $F_j$ by $H_j$. The incorrect-score laws $H_j$ are unrestricted and need not coincide.

For each incorrect answer, define its *rare-error multiplier* and the maximum over initially present errors by

$$m_j = N \int_{\mathbb{R}} G(s)^{N-1} \, \mathrm{d}H_j(s), \qquad m_* := \max_{\substack{j \in \mathcal{B} \\ p_{0,j} > 0}} m_j, \tag{3}$$

with $m_* = 0$ if $z_0 = 0$. The integral is the probability that one candidate of type $j$ beats $N-1$ correct candidates; $N$ counts its possible positions. Thus its multiplier after reference mixing is $(1-\alpha)m_j$. Only initially present errors matter: selection introduces no new answer, and the reference contains no errors (Appendix A).

**Theorem 1** (Exact recovery under common correct scores)**.** *Suppose all correct answers have the same score CDF $G$. For any initial distribution $p_0$, let $m_*$ be as in* (3)*. Then*

$$\lim_{t\to\infty} \mathrm{TV}(p_t, q) = 0 \quad \Longleftrightarrow \quad (1-\alpha)m_* \le 1,$$
$$\limsup_{t\to\infty} z_t \ge 1 - ((1-\alpha)m_j)^{-1/(N-1)} > 0 \quad \textit{if } (1-\alpha)m_j > 1 \textit{ and } p_{0,j} > 0.$$

*If $(1-\alpha)m_* < 1$, then $z_t \le c^t z_0$ with $c = (1-\alpha)\max\{1, m_*\} < 1$. If $(1-\alpha)m_* = 1$, then $z_t \le z_0/(1+\alpha z_0 t)$.*

Consequently, recovery holds for every initialization if and only if $(1-\alpha)\max_{j\in\mathcal{B}} m_j \le 1$, with the same empty-set convention.

### 2.3 STABILIZATION UNDER HETEROGENEOUS CORRECT SCORES

Now take $\mathcal{B} = \varnothing$ and index the $d$ correct answers by $i = 1, \dots, d$. Their score CDFs $G_i := F_i$ may differ; assume $q_i > 0$ for every $i$. Write $\Delta_d$ for distributions on these $d$ answers. Selection leaves the target unchanged when $\mathcal{S}_N(q) = q$, equivalently $W_i(q) = 1$ for all $i$. We call this *target preservation*. Continuity of (2) makes it necessary for recovery: any limit must be a fixed point. Distinct score distributions and cyclic preferences can satisfy this balance at $q$.

Measure heterogeneity by the CDF diameter $\delta = \max_{i,j} \|G_i - G_j\|_\infty$. Let $\alpha_{\mathrm{crit}}(N, \delta)$ be the infimum of $a \in [0,1]$ such that global recovery holds at every $\alpha \in [a, 1]$, uniformly over finite continuous-score families of diameter at most $\delta$ and their preserved full-support targets. The guarantee thus covers all weights above $a$.

**Theorem 2** (Quadratic stabilization of preserved targets)**.** *For every fixed integer $N \ge 2$, there are constants $0 < c_N \le C_N < \infty$ and $\delta_N > 0$ such that*

$$c_N \delta^2 \le \alpha_{\mathrm{crit}}(N, \delta) \le C_N \delta^2, \qquad 0 < \delta < \delta_N. \tag{4}$$

*The upper bound holds uniformly in $d$ and in the positive target probabilities. For the lower bound, continuous score distributions with positive densities and a uniform target preserved by selection have, at a reference weight of order $\delta^2$, a unique fixed point whose convergence basin has relative Lebesgue measure zero.*

The convergence basin consists of initial distributions whose iterates approach $q$. Thus almost every initialization in the construction fails to recover, with measure taken within the simplex. Section 4.2 combines error removal with stabilization of correct proportions.

The quadratic lower bound uses score laws whose oscillation frequency increases as their CDF diameter decreases. On a fixed score interval, a uniform density-derivative bound changes the recovery threshold: Theorem 7 gives the cubic $N = 2$ scale and recovery at every positive weight for fixed $N \ge 3$ and sufficiently small $\delta$.

## 3 Error Reproduction and Rotational Instability

For brevity, write $\beta = 1 - \alpha$ throughout the analysis. Linearizing at $q$ generally gives only local stability information. Under common correct scores, convexity turns the local growth condition into a global bound. At any state with incorrect mass $z > 0$, let $H$ be the conditional mixture of its incorrect score distributions. Summing the winner probabilities of incorrect answers gives the selected incorrect mass

$$R_H(z) = Nz \int [(1-z)G + zH]^{N-1} \, \mathrm{d}H. \tag{5}$$

The function $w_H(z) = R_H(z)/z$, extended to $z = 0$, is convex, with $w_H(0) = m_H := N \int G^{N-1} \, \mathrm{d}H$ and $w_H(1) = 1$. Convexity bounds this function by the line joining its endpoints:

$$R_H(z) \le m_H z - (m_H - 1)z^2 \le m_* z - (m_* - 1)z^2. \tag{6}$$

The last inequality uses the fact that $H$ mixes the score distributions of initially present incorrect answers. The bound therefore applies at every round, even though the conditional mixture changes. If $\beta m_* < 1$, (6) gives the geometric bound in Theorem 1. At the critical boundary it gives $z_{t+1} \le z_t - \alpha z_t^2$; summing the reciprocal increments yields the critical upper bound. Once $z_t \to 0$, the conditional distribution on correct answers converges to $q$: common scoring preserves their relative weights during selection, while each reference mixture contracts their discrepancy from the reference by a factor at most $\beta$. To prove necessity, we use the following lower bound for an initially present incorrect answer:

$$p_{t+1,j} \ge \beta m_j (1 - z_t)^{N-1} p_{t,j}. \tag{7}$$

If $\limsup_t z_t$ were strictly below the stated threshold, this coordinate would eventually grow by a fixed factor greater than one, contradicting $p_{t,j} \le 1$. Appendix A supplies the conditional-distribution and endpoint details.

The multiplier in (3) also shows which verifier statistic matters. For an incorrect score $S_{\mathrm{bad}}$, put $U = G(S_{\mathrm{bad}})$ and $\mu = \mathbb{E}U = \mathbb{P}(S_{\mathrm{bad}} > S_{\mathrm{correct}})$. Then

$$m_N = N\mathbb{E}U^{N-1}, \qquad N\mu^{N-1} \le m_N \le N\mu. \tag{8}$$

Both bounds are attained by continuous score distributions (Appendix A.4). When $N \ge 3$, this interval can contain values on both sides of $1/\beta$, so pairwise accuracy $1 - \mu$ alone does not determine whether recovery occurs. At $N = 2$ the bounds coincide.

For example, take $G = \mathrm{Unif}[0, 1]$ and an incorrect score that is uniform on $[2, 3]$ with probability $\epsilon$ and on $[-2, -1]$ otherwise. Its pairwise accuracy is $1 - \epsilon$, but

$$m_N = N\epsilon, \qquad R_H(z) = 1 - (1 - \epsilon z)^N + [(1 - \epsilon)z]^N. \tag{9}$$

With $99\%$ pairwise accuracy and $\alpha = 0.1$, recovery holds for $N = 100$; it fails for $N = 128$ whenever incorrect mass is present initially. Appendix E.2 translates the rate bounds into finite-round guarantees.

More generally, if $\mathbb{P}(U > 1 - u) \sim cu^r$ as $u \downarrow 0$, with $c > 0$ and $r \ge 0$, then $m_N \sim c\Gamma(r+1)N^{1-r}$. The upper tail of verifier scores can therefore make the rare-error multiplier grow or vanish as the candidate pool increases. ROC-based inference analyses and reward-hacking results already identify the importance of this tail (Dorner et al., 2026; Khalaf et al., 2025); here its moment supplies an exact recursive recovery criterion.

When all incorrect answers share a CDF $H$ and $\beta m_H > 1$, the scalar map $F(z) = \beta R_H(z)$ has a unique positive equilibrium $z_* \in (0, \beta)$. Every $z_0 > 0$ converges monotonically to it, and the limiting full distribution is $(1 - z_*)q + z_* p_0(\cdot \mid \mathcal{B})$. This limit differs from the reference and preserves the initial conditional proportions among incorrect answers. Appendix A.3 proves this using monotonicity of $R_H$ and convexity of $w_H$; Figure 3 illustrates the positive equilibrium.

When correct answers have different score distributions, recovery also requires control of their relative proportions (Figure 1(c)). Even when $\mathcal{S}_N(q) = q$, selection can move probability cyclically around $q$. A second-order effect of selection can determine the required reference weight. For a skew-symmetric map $\Omega$, the step $v \mapsto v + \varepsilon \Omega v$ increases squared distance at second order because $\Omega v$ is orthogonal to $v$. The next section proves a global upper bound and constructs continuous scalar score distributions exhibiting this instability.

# 4 STABILIZATION WITH HETEROGENEOUS SCORES

We first take all $d$ answers to be correct and $q_i > 0$; Section 4.2 then adds incorrect answers. Exact preservation is restrictive for ordered score families: if scores are $\mu_i + E$ with common, everywhere-positive continuous noise density, it forces all $\mu_i$ to coincide (Lemma B.3). This includes common-variance Gaussian and common-scale Gumbel scores; Appendix I.4 also classifies preserved symmetric scale families at a common location. Heterogeneous scores can therefore preserve the target through differences in distribution shape, rather than location shifts alone. The selection weights $W_i$ in (1) form a vector $W$ satisfying the following inequality.

**Proposition 3** (Monotonicity and the fixed point)**.** *For all $p, r \in \Delta_d$,*

$$(p - r) \cdot (W(p) - W(r)) \leq 0. \tag{10}$$

*For every $0 < \alpha < 1$, $T_\alpha$ has exactly one fixed point. This point equals $q$ if and only if $\mathcal{S}_N(q) = q$.*

To derive (10), let $\sum_i v_i = 0$ and $G_v = \sum_i v_i G_i$. Integration by parts gives, for $N \geq 3$,

$$v^\top DW(p)v = -\frac{N(N-1)(N-2)}{2} \int F_p^{N-3} G_v^2 \, \mathrm{d}F_p \leq 0.$$

For $N = 2$, this quadratic form is zero. Integrating along the segment from $r$ to $p$ proves the proposition's inequality. Appendix B proves fixed-point uniqueness from the same inequality and handles nonsmooth CDFs.

Failure of target preservation also gives a lower bound on the asymptotic recovery error. Coupling the $N$ candidate draws under the two distributions gives $\mathrm{TV}(\mathcal{S}_N(p), \mathcal{S}_N(r)) \leq N \, \mathrm{TV}(p, r)$. The triangle inequality along an orbit then implies

$$\limsup_t \mathrm{TV}(p_t, q) \geq \frac{\beta \ \mathrm{TV}(\mathcal{S}_N(q), q)}{1 + \beta N}. \tag{11}$$

Appendix B derives (11).

Viewed as payoffs, the selection weights satisfy the monotonicity condition for a stable population game (Hofbauer & Sandholm, 2009). Its consequences depend on the update rule. The discrete map $p_i \mapsto p_i W_i(p)$ can contain a rotational component, so (10) alone does not ensure that distance to the fixed point decreases. Cyclic nonconvergence of quadratic simplex maps is classical (Zakharevich, 1978; Ganikhodzhaev et al., 2011). Our lower bound realizes such dynamics with independent scalar scores for every $N$ and quantifies the reference weight required as CDF diameter vanishes.

## 4.1 A GLOBAL KL-DIVERGENCE BOUND

Under target preservation, (10) controls the first-order change in KL divergence (relative entropy). Since $W(q) = \mathbf{1}$ and $p \cdot W(p) = 1$, it implies $q \cdot W(p) \geq 1$. This removes the potentially first-order increase in $D_{\mathrm{KL}}(q\|p) = \sum_i q_i \log(q_i/p_i)$.

**Theorem 4** (Global recovery from bounded variation in selection weights)**.** *Suppose $\mathcal{S}_N(q) = q$ and, for some $0 \leq h < 1$,*

$$|W_i(p) - 1| \leq h \ \mathrm{TV}(p, q) \qquad \textit{for all } p, i. \tag{12}$$

*Set $C_h = h^2/[4(1-h)^2]$ and $\rho = \beta(1 + C_h)$. If $\rho < 1$, then for every initialization and $t \geq 1$,*

$$D_{\mathrm{KL}}(q\|p_t) \leq \rho^{t-1} \log(1/\alpha). \tag{13}$$

*For CDF diameter $\delta$, condition* (12) *holds with*

$$h = h_N(\delta) := N(N-1)\delta(a_N + \delta)^{N-2}, \qquad a_N = N^{-1/(N-1)}. \tag{14}$$

*Proof.* For interior $p$, write $\xi_i = W_i(p) - 1$. Taylor's inequality on $[-h, h]$, the cancellation $\sum_i q_i \xi_i \geq 0$, and Pinsker's inequality yield

$$\begin{aligned} D_{\mathrm{KL}}(q\|\mathcal{S}_N(p)) - D_{\mathrm{KL}}(q\|p) &= -\sum_i q_i \log(1 + \xi_i) \\ &\leq \frac{\sum_i q_i \xi_i^2}{2(1-h)^2} \leq C_h D_{\mathrm{KL}}(q\|p). \end{aligned} \tag{15}$$

Convexity of KL divergence in its second argument gives $D_{\mathrm{KL}}(q\|T_\alpha(p)) \le \beta D_{\mathrm{KL}}(q\|\mathcal{S}_N(p))$. The bound $p_1 \ge \alpha q$ handles boundary initializations and bounds the KL divergence after one iteration. Iterating the preceding contraction then proves (13). The CDF estimate follows by differentiating $W_i$ along the segment from $q$ to $p$ and using preservation to control the required moment; Appendix B.2 supplies the calculation. □

It follows that $\alpha > h_N(\delta)^2/[4(1-h_N(\delta))^2 + h_N(\delta)^2]$ suffices whenever $h_N(\delta) < 1$. Since $h_N(\delta) = O_N(\delta)$, this proves the upper half of Theorem 2. Neither step depends on the number of answers or on the smallest target probability. The dependence on $N$ can also be made polynomial. Lemma B.1 gives the admissible coefficient $L_\delta = N\delta[1 + (N-1)\delta]$. Substitution in Theorem 4 yields

$$\alpha_{\mathrm{crit}}(N,\delta) \le N^2\delta^2 \qquad \text{whenever } 0 < N\delta \le \tfrac{1}{4}. \tag{16}$$

The lower construction permits $c_N = N/(3456\pi^2)$ for $N \ge 3$ in an explicit $N$-dependent small-$\delta$ range (Appendix C). Matching dependence on growing $N$ remains open.

### 4.2 Joint error removal and convergence near the reference

Let $q$ be positive on $\mathcal{C}$ and zero on $\mathcal{B}$, with score CDFs $G_i$ and $H_j$, respectively. Write $\mathcal{S}_N^{\mathcal{C}}$ and $T_\alpha^{\mathcal{C}}$ for selection and retraining restricted to correct candidates. Define their CDF diameter and the largest relative change in a reference probability under selection by $\delta = \max_{i,i' \in \mathcal{C}} \|G_i - G_{i'}\|_\infty$ and $\varepsilon_{\mathrm{rel}} = \max_{i \in \mathcal{C}} \left| \frac{\mathcal{S}_N^{\mathcal{C}}(q)_i}{q_i} - 1 \right|$. To control errors independently of the current correct proportions, set

$$\overline{m} = \max_{i \in \mathcal{C},\, j \in \mathcal{B}} N \int G_i^{N-1}\, \mathrm{d}H_j, \tag{17}$$

Set $\overline{m} = 0$ if $\mathcal{B}$ is empty. This maximum against a single correct answer bounds the rare-error multiplier against every correct mixture by convexity. Put

$$g_\delta = N\delta, \qquad L_\delta = N\delta[1 + (N-1)\delta], \qquad A_\delta = \frac{(1+g_\delta)L_\delta^2}{4(1-g_\delta)^2},$$
$$\kappa_\alpha = \alpha - \beta\varepsilon_{\mathrm{rel}}\left(1 + \frac{\beta L_\delta}{2\alpha}\right) - \beta A_\delta.$$

**Theorem 5** (Joint extinction and convergence)**.** *Suppose $g_\delta < 1$, $\kappa_\alpha > 0$ and $\beta\overline{m} \le 1$. Every trajectory converges to the unique fixed point $p_\alpha$ of the update restricted to $\mathcal{C}$, and*

$$\mathrm{TV}(p_\alpha, q) \le \frac{\beta\varepsilon_{\mathrm{rel}}}{2\alpha}.$$

*If $\beta\overline{m} < 1$, then $z_t \le [\beta \max\{1, \overline{m}\}]^t z_0$; at equality, $z_t \le z_0/(1 + \alpha z_0 t)$. Under exact preservation on $\mathcal{C}$, $p_\alpha = q$; for $0 < N\delta \le 1/4$, the sufficient conditions simplify to $\alpha \ge N^2\delta^2$ and $\beta\overline{m} \le 1$.*

Here $\beta\overline{m} \le 1$ removes incorrect mass, while $\kappa_\alpha > 0$ stabilizes correct proportions; neither requires identical score distributions within either group. On the correct support, $0 < g_\delta \le 1/4$ and $\varepsilon_{\mathrm{rel}} \le g_\delta^3$ permit convergence at $\alpha = 2g_\delta^2$.

The proof measures KL divergence from the actual fixed point. The reference weights for individual coordinates $a_i = \alpha q_i / p_{\alpha,i}$ need not be equal. Applying Jensen's inequality to $\phi(x) = x - 1 - \log x \ge 0$ and then monotonicity yields

$$D_{\mathrm{KL}}(p_\alpha \| T_\alpha^{\mathcal{C}}(p)) \le (1 - \kappa_\alpha) D_{\mathrm{KL}}(p_\alpha \| p).$$

Coupling the candidate draws shows that incorrect mass perturbs the next conditional correct distribution by at most $\beta N z_t$ in TV. The resulting KL-divergence recursion has a vanishing additive term, so convergence follows even at the critical $1/t$ extinction rate. Appendices B.5–B.6 prove these steps. Because (17) uses the worst correct type, the sufficient criterion may be conservative; Figure 6(c) and Appendix E.5 quantify its feasibility and conservatism in an explicit parameter family. Proposition B.2 gives neighborhood bounds for larger preservation errors and bounds actual fitting error on the correct support under stronger TV contraction.

### 4.3 SCORE CONSTRUCTIONS AND THE ROLE OF REGULARITY

For three correct answers, put $\theta_i = 2\pi i/3$ and $0 < \varepsilon_{\text{amp}} \le 1/4$. A fixed low-degree family has densities on $[0,1]$

$$h_i(u) = 1 + \varepsilon_{\text{amp}}\{\cos\theta_i(6u^2 - 6u + 1) + \sin\theta_i(20u^3 - 30u^2 + 12u - 1)\},$$

These are mixtures of four fixed Beta distributions, with uniformly bounded density derivatives on $[0,1]$. At $N = 2$, the uniform target is preserved and the tangent eigenvalues are $\beta(1 \pm i\varepsilon_{\text{amp}}^2/70)$. If $0 < \alpha < 1 - (1 + \varepsilon_{\text{amp}}^4/4900)^{-1/2}$, every nonfixed trajectory fails to converge. Appendix B.7 proves this claim and transfers it to smooth positive densities on $\mathbb{R}$. Since $\delta$ is proportional to $\varepsilon_{\text{amp}}$, the obstruction has order $\delta^4$. Appendix B.9 gives a nontransitive-dice example and proves recovery for two answers under preservation.

For every $N$, Appendix C builds oscillatory score distributions $G_{i,K}$ with frequency $K$, CDF diameter $\delta_K$, and $W_i(q) = 1$ at the uniform target. Figure 4 illustrates the construction and the quadratic instability mechanism.

**Theorem 6** (Continuous-score obstruction at every candidate count)**.** *For each fixed $N \ge 2$, the explicit distributions in Appendix C admit constants $\gamma_N > 0$ and $K_*(N) < \infty$ such that, for all integers $K \ge K_*(N)$,*

$$G'_{i,K} \ge \frac{1}{32N} \text{ a.e. on } (0,1), \qquad \frac{1}{8K} \le \delta_K \le \frac{2}{K}.$$

*At $\alpha_K = \gamma_N/K^2$, the uniform target preserved by selection is the unique fixed point and has a basin of relative Lebesgue measure zero. For almost every initialization, $\limsup_t \|p_t - q\|_2 \ge \varepsilon_{N,K} > 0$, and the trajectory converges to no distribution.*

**Proof roadmap for Theorem 6.** Appendix C constructs positive densities with exact preservation. The selection derivative has an $O_N(K^{-1})$ rotational term and an $O_N(K^{-2})$ dissipative term; a Fourier mode retains positive radial expansion after mixing at $\alpha_K = \gamma_N/K^2$. A global Lipschitz inverse and a local expanding-cone argument then place the convergence basin in a countable union of null sets. Uniqueness excludes other fixed-point limits. Rounding $K = \lceil 2/\delta \rceil$ gives $\delta_K \le \delta$ and $\alpha_K \ge \gamma_N \delta^2/9$.

**Theorem 7** (Recovery under normalized density regularity)**.** *Suppose all answers are correct, $q_i > 0$ and $\mathcal{S}_N(q) = q$. Let the score laws have globally $C^1$ densities $f_i$, supported in a common interval of length $B$, with $\|f_i'\|_\infty \le L$. At $N = 2$, a weight $O(B^2 L\delta^3)$ suffices; the worst-case order is $\Theta(\delta^3)$ for $B = 1, L = 23$. For each fixed $N \ge 3, B, L$, every positive reference weight gives global recovery once $\delta$ is sufficiently small, with $D_{\text{KL}}(q\|p_t) \le \beta^{t-1}\log(1/\alpha)$ for $0 < \alpha < 1$, $t \ge 1$.*

CDF closeness and the derivative bound imply density closeness of order $\sqrt{L\delta}$. At $N = 2$, preservation cancels the common-density term, yielding the cubic bound; at $N \ge 3$, curvature of the winner formula supplies enough dissipation to dominate the KL remainder. Appendix I gives an explicit small-heterogeneity condition and both proofs. A zero threshold infimum does not imply recovery at $\alpha = 0$. The fixed score scale matters: common dilation preserves comparisons while reducing density derivatives, so an absolute derivative bound alone does not exclude a quadratic obstruction.

## 5 NUMERICAL EVIDENCE

We examine the recovery criteria through exact population recursions and diagnostics computed from public verifier scores. Figures 2–3 illustrate Theorem 1: increasing $N$ can cross the recovery boundary, and equal pairwise accuracy can coexist with different long-run limits.

**Pooling and error groups.** Two GSM8K questions, two generators and five verifiers (Dorner et al., 2026) yield 20 pooled estimates below one at $N = 128$. Pooling averages the subtype moments, $m_{\text{pool}} = \sum_j \pi_j m_j$, whereas Theorem 1 uses their maximum. Grouping errors by final numeric-answer tokens gives three calibration maxima above one (Figure 6(a)). All selected groups' held-out estimates are below one, and the three corresponding intervals include one; amplification is not replicated. Freezing these group-score histograms produces the predicted transition in Figure 6(b). The estimator check and all cases appear in Appendix F.

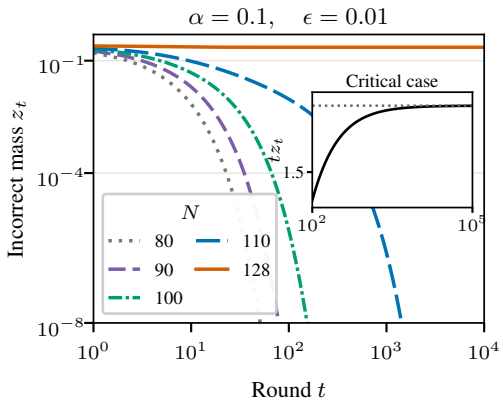



Figure 2: Theorem 1's recovery boundary in the band-mixture example, $z_0 = 0.25$. The inset uses exactly critical $N = 128$, $\alpha = 7/32$, with $tz_t \to 2/(127\epsilon)$ (dotted).

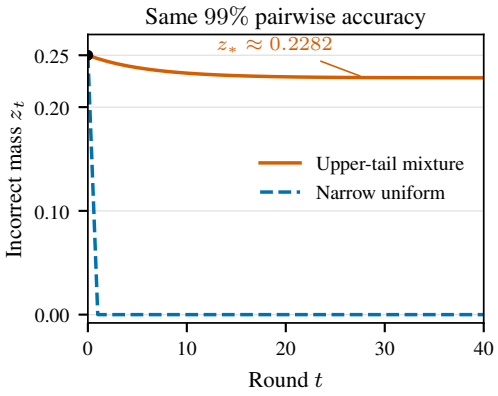



Figure 3: Equal 99% pairwise accuracy, different recovery outcomes, as predicted by (8) and Theorem 1. Here $N = 128$, $\alpha = 0.1$, $z_0 = 0.25$; only the incorrect-score law changes.

**Joint criterion.** For $N = 8$, correct-score diameters $\delta = 0.005, 0.015, 0.03$ and preservation errors $14\delta/9$ give sufficient-weight infima $0.0172, 0.0538, 0.1230$ when $\overline{m} = 0.8$. Four initializations also converge numerically at $\alpha = 0.001$, with TV bias 0.469–0.494. At $\overline{m} = 1.12$, the exact error boundary is $0.10714$ and the sufficient infima are $0.10714, 0.10714, 0.1230$. Runs above the certificates obey the bias bounds. These calculations quantify feasibility and conservatism in this specified family (Appendix E.5).

## 6 Discussion

Our analysis highlights a broader feature of repeated best-of-$N$ retraining: the verifier shapes both the current winner and the candidate distribution in later rounds. This feedback means that long-run behavior can depend on features of the score distributions that pairwise verifier accuracy does not capture. A verifier can perform well in pairwise comparisons and still amplify a rare incorrect answer, while differences among correct-answer score laws can shift probability across answers that are all correct. We therefore separate two questions that are often treated together: whether selection removes incorrect answers and whether it preserves the desired proportions among correct answers.

We show that reference mixing can control both effects, but the required weight depends on the structure of the score distributions. When correct answers share the same score law, the rare-error multiplier gives the exact recovery boundary. When correct score laws differ, recovery also depends on whether selection preserves the reference distribution, and the amount of mixing needed depends on the regularity of those score laws. These results suggest that analyses of recursive distillation should look beyond pairwise verifier accuracy and track the score features that repeated selection amplifies or suppresses. More broadly, they show why conclusions from one-shot best-of-$N$ selection may not carry over once selected outputs are used for retraining.

## AI USE STATEMENT

GPT-5.6 was used as an auxiliary research and writing tool for language editing, literature search, exploring examples and proof strategies, algebraic and numerical checks, and assistance with code and figures. It also contributed to several local technical developments: in particular, it suggested the basic oscillatory ansatz underlying the $N = 2$ cubic lower-bound construction in Appendix I.2 and was used to explore alternative parameterizations and proof routes for auxiliary results. GPT-5.6 and GPT-6 were additionally used to stress-test and review proofs for missing cases, algebraic errors, and notation inconsistencies. All AI-assisted mathematical suggestions were independently derived or verified by the authors before inclusion. The authors determined the research questions, assumptions, theorem statements, final proofs, experiments, and interpretations, and take full responsibility for the contents of the paper.

### REPRODUCIBILITY STATEMENT

Section 2 specifies the model; Appendices A–D, H and I give complete proofs. Appendices E–F provide exact recursions, parameters, pinned data, evaluation protocols and numerical checks. The supplementary code archive contains executable programs and instructions; the LaTeX archive includes editable figure sources.

## REFERENCES

Sina Alemohammad, Josue Casco-Rodriguez, Lorenzo Luzi, Ahmed Imtiaz Humayun, Hossein Babaei, Daniel LeJeune, Ali Siahkoohi, and Richard G. Baraniuk. Self-consuming generative models go MAD. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/ebc042e767de551803ccfcc45e2454f5-Paper-Conference.pdf.

Gholamali Aminian, Idan Shenfeld, Amir R. Asadi, Ahmad Beirami, and Youssef Mroueh. Best-of-N through the smoothing lens: KL divergence and regret analysis. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/abb7833fc3f38e785e3f9777c74610e5-Paper-Conference.pdf.

Florian E. Dorner, Yatong Chen, André F. Cruz, and Fanny Yang. ROC-n-reroll: How verifier imperfection affects test-time scaling. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=3Gy5mmyuxn.

Ali Falahati, Mohammad Mohammadi Amiri, Kate Larson, and Lukasz Golab. Curated synthetic data doesn't have to collapse: A theoretical study of generative retraining with pluralistic preferences. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, 2026. URL https://icml.cc/virtual/2026/poster/61989.

Damien Ferbach, Quentin Bertrand, Avishek Joey Bose, and Gauthier Gidel. Self-consuming generative models with curated data provably optimize human preferences. In *Advances in Neural Information Processing Systems*, volume 37, pp. 102531–102567, 2024. doi: 10.52202/079017-3256. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/b9e88ae0308cf82d0b0f634ddbdf809a-Paper-Conference.pdf.

Shi Fu, Yingjie Wang, Yuzhu Chen, Li Shen, and Dacheng Tao. Self-verification provably prevents model collapse in recursive synthetic training. In *Advances in Neural Information Processing Systems*, volume 38, pp. 40550–40603, 2025. doi: 10.52202/085713-1213. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/3380e8116452e0efbf36f35d95e88c94-Paper-Conference.pdf.

R. Ganikhodzhaev, F. Mukhamedov, and U. A. Rozikov. Quadratic stochastic operators and processes: Results and open problems. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 14(2):279–335, 2011. doi: 10.1142/S0219025711004365. URL https://doi.org/10.1142/S0219025711004365.

Hengzhi He, Shirong Xu, and Guang Cheng. Recursive learning without collapse: A weighting-based stabilization framework. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2026. doi: 10.1093/jrsssb/qkag099. URL https://doi.org/10.1093/jrsssb/qkag099.

Josef Hofbauer and William H. Sandholm. Stable games and their dynamics. *Journal of Economic Theory*, 144(4):1665–1693.e4, 2009. doi: 10.1016/j.jet.2009.01.007. URL https://doi.org/10.1016/j.jet.2009.01.007.

Hadi Khalaf, Claudio Mayrink Verdun, Alex Oesterling, Himabindu Lakkaraju, and Flavio du Pin Calmon. Inference-time reward hacking in large language models. In *Advances in Neural Information Processing Systems*, volume 38, pp. 68646–68686, 2025. doi: 10.52202/085713-2064. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/590a0cc0306c1c63e2d66a51a407718f-Paper-Conference.pdf.

Matteo Marchi, João Pedro Silvestre, Bahman Gharesifard, and Paulo Tabuada. Preventing model collapse: A Fisher–Rao perspective on the dynamics of training with synthetic data, 2026. URL https://arxiv.org/abs/2609.18878. arXiv:2609.18878.

Pier Giuseppe Sessa, Robert Dadashi, Léonard Hussenot-Desenonges, Johan Ferret, Nino Vieillard, Alexandre Ramé, Bobak Shahriari, Sarah Perrin, Abram L. Friesen, Geoffrey Cideron, Sertan Girgin, Piotr Stanczyk, Andrea Michi, Danila Sinopalnikov, Sabela Ramos Garea, Amélie Héliou, Aliaksei Severyn, Matthew Hoffman, Nikola Momchev, and Olivier Bachem. BOND: Aligning LLMs with best-of-N distillation. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/947f37882a394140f7add476bb99d1d3-Paper-Conference.pdf.

Seamus Somerstep, Vinod Raman, Unique Subedi, and Yuekai Sun. Learning to choose or choosing to learn: Best-of-N vs. supervised fine-tuning for bit string generation. In *Proceedings of The 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 1900–1908, 2026. URL https://proceedings.mlr.press/v300/somerstep26a.html.

Ved Sriraman and Adam Block. Revisiting the (sub)optimality of best-of-N for inference-time alignment. In *Proceedings of Thirty Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 5980–6028, 2026. URL https://proceedings.mlr.press/v336/sriraman26a.html.

Tong Yang, Jincheng Mei, Hanjun Dai, Zixin Wen, Shicong Cen, Dale Schuurmans, Yuejie Chi, and Bo Dai. Faster WIND: Accelerating iterative best-of-N distillation for LLM alignment. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 4537–4545, 2025. URL https://proceedings.mlr.press/v258/yang25e.html.

Bingji Yi, Qiyuan Liu, Yuwei Cheng, and Haifeng Xu. Escaping model collapse via synthetic data verification: Near-term improvements and long-term convergence. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/29806b2e4dcb35e42abed21d43e2b7d8-Paper-Conference.pdf.

M. I. Zakharevich. On the behaviour of trajectories and the ergodic hypothesis for quadratic mappings of a simplex. *Russian Mathematical Surveys*, 33(6):265–266, 1978. doi: 10.1070/RM1978v033n06ABEH003890. URL https://www.mathnet.ru/eng/rm3594.

Hongru Zhao, Jinwen Fu, and Tuan Pham. Convergence and stability analysis of self-consuming generative models with heterogeneous human curation, 2025. URL https://arxiv.org/abs/2511.09002. arXiv:2511.09002, version 3, August 2026.

# A Winner Distribution and Common-Score Recovery

## A.1 Selection and exact retraining

For a distinguished candidate to win with label $y$, it must have that label (probability $p_y$), receive a score $s$ with law $F_y$, and beat the other $N-1$ scores (conditional probability $F_p(s)^{N-1}$). The winner events for the $N$ candidate positions are disjoint almost surely. Summing over candidate positions proves (1). This argument treats each sampled candidate independently, including repeated labels. The identity $\sum_y \mathcal{S}_N(p)_y = 1$ also follows from $N \int F_p^{N-1}\,\mathrm{d}F_p = 1$, because $F_p$ is continuous.

For any target mixture $r = \alpha q + \beta \mathcal{S}_N(p)$, the population cross-entropy objective is $-\sum_y r_y \log p'_y$. Its excess over the value at $p' = r$ is $D_{\mathrm{KL}}(r\|p') \geq 0$. Equality forces $p'_y = r_y$ on the support of $r$, which already has total mass one, so the minimizer over the simplex is exactly $r$. This proves (2). If the reference is an empirical distribution, the same calculation uses that empirical distribution as $q$.

Plackett–Luce curation is an instance of the score model: assign score $r(y) + E$ to each sampled answer $y$, with independent standard Gumbel noise $E$ having CDF $\exp(-e^{-s})$. Conditional on a candidate list with rewards $r_1, \ldots, r_N$, the winning probability of position $k$ is

$$\int_{\mathbb{R}} e^{r_k - s} \exp\left(-e^{-s} \sum_{\ell=1}^{N} e^{r_\ell}\right) \mathrm{d}s = \frac{e^{r_k}}{\sum_{\ell=1}^{N} e^{r_\ell}}.$$

The calculation also applies when candidate labels repeat, since their score noises remain independent.

Selection cannot create a label outside the candidate support. Conversely, $\mathcal{S}_N(p)_y \geq p_y^N$ from the event that every candidate has label $y$. With $\beta > 0$, the set of present incorrect labels is therefore invariant at every finite time. In particular, the maximum in (3) remains the maximum over the initially present errors throughout the orbit.

## A.2 Convex amplification and proof of Theorem 1

At a state of incorrect mass $z > 0$, write $H = \sum_{j \in \mathcal{B}} (p_j/z) H_j$. The mixture score CDF is $(1-z)G + zH$, and summing the winner formula over incorrect labels gives (5). For every score $s$, the function $[(1-z)G(s) + zH(s)]^{N-1}$ is convex in $z \in [0,1]$: its second derivative is nonnegative for $N \geq 3$, and it is affine for $N = 2$. Integration preserves convexity. Consequently,

$$w_H(z) := N \int [(1-z)G + zH]^{N-1}\,\mathrm{d}H \leq (1-z)m_H + z,$$

where $w_H(1) = N \int H^{N-1}\,\mathrm{d}H = 1$ uses continuity. Since $m_H$ is a convex combination of active $m_j$, multiplying by $z$ proves (6). The chord bounds selected incorrect mass by $\max\{1, m_*\}z$. After reference mixing, induction gives the strict subcritical bound. At $\beta m_* = 1$, it becomes $z_{t+1} \leq z_t - \alpha z_t^2$. If $z_{t+1} > 0$, then $z_{t+1}^{-1} \geq [z_t(1-\alpha z_t)]^{-1} \geq z_t^{-1} + \alpha$. Summing yields the stated critical bound. If any $z_t = 0$, it remains zero and the same bound holds.

It remains to recover the correct proportions. For $t \geq 1$, $1 - z_t \geq \alpha > 0$. Let $u_t = p_t(\cdot \mid \mathcal{C})$, considering $q$ as a distribution on $\mathcal{C}$. All correct labels have the same selection weight. Selection therefore allocates its total correct mass in proportions $u_t$, and

$$u_{t+1} = \frac{\alpha}{1 - z_{t+1}} q + \left(1 - \frac{\alpha}{1 - z_{t+1}}\right) u_t. \tag{A.1}$$

The coefficient of the current conditional distribution lies in $[0, \beta]$. Iterating this contraction gives $\mathrm{TV}(u_t, q) \leq \beta^{t-1}\,\mathrm{TV}(u_1, q)$. Combining this contraction with the decomposition $\mathrm{TV}(p_t, q) \leq z_t + (1 - z_t)\,\mathrm{TV}(u_t, q)$ proves recovery whenever $\beta m_* \leq 1$. If $z_0 = 0$, the common-score selection map is the identity on the correct face and the same conclusion is immediate, including the convention $m_* = 0$.

For the converse, the contribution in which an initially present incorrect candidate $j$ beats $N-1$ correct candidates gives (7). Put $\lambda = \beta m_j > 1$ and $r = 1 - \lambda^{-1/(N-1)}$. If $\limsup_t z_t < r$, choose $r' < r$ with $z_t \leq r'$ for all sufficiently large $t$. Then $p_{t+1,j} \geq \lambda(1-r')^{N-1} p_{t,j}$ has a fixed multiplier greater than one. Since the coordinate stays positive, iterating this lower bound contradicts $p_{t,j} \leq 1$. Thus $\limsup_t z_t \geq r$. If $\beta m_* > 1$, such an active $j$ exists.

### A.3 A COMMON INCORRECT SCORE LAW

Suppose all incorrect labels have a common CDF $H$. Then $z_{t+1} = F(z_t)$ with $F(z) = \beta R_H(z)$. The map $R_H$ is nondecreasing. We prove this by a coupling that requires no ordering on $G$ and $H$: for each candidate pre-draw an independent correct score and incorrect score, and use a uniform random number to decide its type. Increasing $z$ switches some candidates from correct to incorrect. If the former winner was incorrect, that candidate remains; the remaining correct scores form a subset of the former correct scores. The new winner must still be incorrect.

**Proposition A.1.** *If $\beta m_H > 1$, there is exactly one fixed point $z_* \in (0, \beta)$ of $F$ other than zero. Every orbit with $z_0 > 0$ converges monotonically to $z_*$. The full distribution converges to $(1 - z_*)q + z_* p_0(\cdot \mid \mathcal{B})$.*

*Proof.* The convex function $\phi(z) = \beta w_H(z) - 1$ has $\phi(0) > 0$ and $\phi(1) = \beta - 1 < 0$. It has a unique zero in $(0, 1)$. Indeed, if it had two zeros $x < y < 1$, monotonicity of convex secant slopes would give $[\phi(1) - \phi(y)]/(1 - y) \geq [\phi(y) - \phi(x)]/(y - x) = 0$, a contradiction. By continuity, it is positive before its unique zero and negative after. Moreover $F(\beta) < \beta$, since at input $\beta < 1$ all candidates are correct with positive probability, so the positive zero is below $\beta$. Monotonicity of $F$ now implies $z < F(z) \leq z_*$ below $z_*$ and $z_* \leq F(z) < z$ above it. Thus every positive orbit is monotone and converges to the fixed point. The correct conditional distribution converges to $q$ by (A.1). The incorrect conditional distribution is unchanged because all incorrect labels have a common selection weight and reference mixing adds none of them. □

### A.4 PERCENTILE MOMENTS AND SCORE TAILS

For independent correct and incorrect scores, put $U = G(S_{\text{bad}})$. Conditioning on the incorrect score gives $\mu = \mathbb{P}(S_{\text{bad}} > S_{\text{correct}}) = \mathbb{E}G(S_{\text{bad}})$. The same conditioning against $N - 1$ correct scores gives $m_N = N\mathbb{E}U^{N-1}$. Jensen's inequality and $u^{N-1} \leq u$ on $[0, 1]$ prove (8). Both endpoints are attainable with continuous CDFs. For the lower endpoint, let the correct score be uniform on $[-2, -1]$ with probability $\mu$ and uniform on $[1, 2]$ otherwise; let the incorrect score be uniform on $[-1/4, 1/4]$. Then $U = \mu$ almost surely. For the upper endpoint, take $G$ uniform on $[0, 1]$ and an incorrect score uniform on $[2, 3]$ with probability $\mu$ and uniform on $[-2, -1]$ otherwise. Then $U$ is Bernoulli with mean $\mu$. Thus for $N \geq 3$, whenever $N\mu^{N-1} \leq 1/\beta < N\mu$, the first system recovers and the second fails whenever incorrect mass is present initially, despite equal pairwise accuracy. The three-band polynomial (9) follows because an incorrect winner occurs either when some high-band incorrect candidate appears, or when all $N$ candidates lie in the low incorrect band.

To derive the tail asymptotic, set $k = N - 1$. Tail integration and the beta integral give

$$\mathbb{E}U^k = k \int_0^1 t^{k-1} \mathbb{P}(U > t)\, \mathrm{d}t,$$

$$k \int_0^1 t^{k-1}(1 - t)^r\, \mathrm{d}t = \frac{\Gamma(k + 1)\Gamma(r + 1)}{\Gamma(k + r + 1)} \sim \Gamma(r + 1)k^{-r}.$$

If $\mathbb{P}(U > 1 - u) \sim cu^r$, the contribution from $t \leq 1 - \varepsilon$ is exponentially small in $k$. On the remaining interval the tail is between $(1 - \eta)cu^r$ and $(1 + \eta)cu^r$ for sufficiently small fixed $\varepsilon$. Letting $k \to \infty$ and then $\eta \downarrow 0$ in the upper and lower bounds proves $m_N \sim c\Gamma(r + 1)N^{1-r}$, including $r = 0$, which allows $\mathbb{P}(U = 1) > 0$ even though the score CDFs are continuous.

## B HETEROGENEOUS SCORES: FIXED POINTS AND RECOVERY

Except in Appendix B.6, we use the all-correct setting of Section 4, with $q_i > 0$ for every answer and $0 < \alpha < 1$.

### B.1 MONOTONICITY, UNIQUENESS, AND PRESERVATION

The vector of selection weights $W(p)$ is polynomial in the coordinates of $p$, so it is continuously differentiable on a neighborhood of the simplex. For a tangent vector $v$ with $\sum_i v_i = 0$, define $G_v = \sum_i v_i G_i$. Both endpoint values of $G_v$ are zero. Differentiating the winner formula and then

applying the product rule and integration by parts for continuous functions of bounded variation gives, for $N \geq 3$,

$$v^\top DW(p)v = N(N-1)\int F_p^{N-2} G_v \,\mathrm{d}G_v,$$
$$\int F_p^{N-2} G_v \,\mathrm{d}G_v = -\frac{N-2}{2}\int F_p^{N-3} G_v^2 \,\mathrm{d}F_p.$$

For $N = 2$, the integral is $\frac{1}{2}[G_v^2]_{-\infty}^{+\infty} = 0$. No density assumption is needed. Integrating the resulting quadratic inequality along $r + \theta(p - r)$ establishes (10).

The continuous map $T_\alpha$ maps the compact simplex to itself, so Brouwer's theorem gives a fixed point. Every fixed point satisfies $p_i \geq \alpha q_i > 0$, and the fixed-point equation becomes $W_i(p) = \frac{1-\alpha q_i/p_i}{\beta}$. If two fixed points $p, r$ were distinct, substitution would give

$$(p - r) \cdot (W(p) - W(r)) = \frac{\alpha}{\beta}\sum_i \frac{q_i(p_i - r_i)^2}{p_i r_i} > 0,$$

contradicting monotonicity. Hence they coincide. Finally $T_\alpha(q) = q$ if and only if $\mathcal{S}_N(q) = q$, proving the full proposition.

Failure of target preservation also gives a lower bound on the asymptotic recovery error. Couple each of the $N$ candidate labels optimally under $p$ and $r$, and use identical conditional scores whenever the labels agree. The winner labels agree whenever all candidate labels agree. A union bound therefore gives

$$\mathrm{TV}(\mathcal{S}_N(p), \mathcal{S}_N(r)) \leq N\,\mathrm{TV}(p, r). \tag{B.1}$$

Writing $e_t = \mathrm{TV}(p_t, q)$, the triangle inequality gives $\beta\,\mathrm{TV}(\mathcal{S}_N(q), q) = \mathrm{TV}(T_\alpha(q), q) \leq \beta N e_t + e_{t+1}$. Taking the limit superior yields

$$\limsup_t e_t \geq \frac{\beta\ \mathrm{TV}(\mathcal{S}_N(q), q)}{1 + \beta N}.$$

If $\beta N < 1$, (B.1) also makes $T_\alpha$ a contraction toward its unique fixed point, which need not be $q$.

This is the same finite-pool contraction factor as in Zhao et al. (2025, Theorem 4.6). On a finite answer set, their stationary-noise Plackett–Luce rule is equivalent to maximizing scores $u_i + E + Z$, where $u_i$ is deterministic utility, $E$ has their common noise law, and $Z$ is standard Gumbel. Each sampled candidate draws fresh independent noises; conditioning on the utilities gives selection probabilities proportional to their exponentials. The common noise $E + Z$ produces a location family, whereas our score laws may differ in shape. Their general answer spaces and infinite-pool results address additional settings. Continuous-time reference stabilization is studied, under a different dynamical approximation, by Marchi et al. (2026).

### B.2 Details of the KL-divergence estimate

Under preservation, $W(q) = \mathbf{1}$ and (10) imply $q \cdot W(p) \geq p \cdot W(p) = 1$. For $|\xi| \leq h < 1$, Taylor's theorem with second derivative bounded by $(1-h)^{-2}$ gives $-\log(1+\xi) \leq -\xi + \frac{\xi^2}{2(1-h)^2}$. Substituting $\xi_i = W_i(p) - 1$ and applying the selection-weight bound and Pinsker's inequality gives

$$-\sum_i q_i \log W_i(p) \leq \frac{h^2\,\mathrm{TV}(p,q)^2}{2(1-h)^2} \leq \frac{h^2}{4(1-h)^2} D_{\mathrm{KL}}(q\|p).$$

Here Pinsker's inequality uses natural logarithms. The identity $\mathcal{S}_N(p)_i = p_i W_i(p)$ establishes (15). Convexity in the second argument and $T_\alpha(p) \geq \alpha q$ complete the iteration in Theorem 4.

For the CDF bound, decompose $p - q = v^+ - v^-$, where both nonnegative parts have mass $d_{\mathrm{TV}} = \mathrm{TV}(p, q)$. If $d_{\mathrm{TV}} > 0$, the functions $\sum_i (v_i^\pm / d_{\mathrm{TV}}) G_i$ are convex combinations of the score CDFs, whose pointwise diameter is at most $\delta$. Therefore $\|F_p - F_q\|_\infty \leq d_{\mathrm{TV}}\delta$. Let $F_\theta = F_q + \theta(F_p - F_q)$. For $N \geq 3$, Minkowski's inequality under the probability measure $\mathrm{d}G_i$ gives

$$\left(\int F_\theta^{N-1}\,\mathrm{d}G_i\right)^{1/(N-1)} \leq \left(\int F_q^{N-1}\,\mathrm{d}G_i\right)^{1/(N-1)} + \delta d_{\mathrm{TV}} \leq a_N + \delta.$$

Monotonicity of probability-space norms then gives $\int F_\theta^{N-2}\,\mathrm{d}G_i \le (a_N+\delta)^{N-2}$. Differentiating along the mixture path and integrating in $\theta$ yields

$$|W_i(p)-W_i(q)| \le N(N-1)\delta d_{\mathrm{TV}} \int_0^1 \int F_\theta^{N-2}\,\mathrm{d}G_i\,\mathrm{d}\theta,$$

which proves (14). For $N=2$, the inner integral equals one.

### B.3 Score variation and an explicit candidate-count bound

The next estimate uses only the diameter of the continuous score CDFs; it does not require target preservation.

**Lemma B.1** (Score-dependent variation)**.** *Put $g_\delta = N\delta$ and $L_\delta = N\delta[1+(N-1)\delta]$. For every pair of distributions $p, r$ and every answer $i$,*

$$|W_i(p)-1| \le g_\delta, \qquad |W_i(p)-W_i(r)| \le L_\delta\,\mathrm{TV}(p,r).$$

*Consequently $\mathcal{S}_N$ is $(1+g_\delta+L_\delta/2)$-Lipschitz in total variation.*

*Proof.* Every mixture CDF $F_p$ is within $\delta$ of $G_i$. Integration by parts, with both endpoint differences zero, gives

$$W_i(p)-1 = N\int F_p^{N-1}\,\mathrm{d}(G_i-F_p) = N(N-1)\int (F_p-G_i)F_p^{N-2}\,\mathrm{d}F_p.$$

Since $\int F_p^{N-2}\,\mathrm{d}F_p = 1/(N-1)$, the first bound follows. These identities hold for continuous CDFs by Stieltjes integration; no density is needed.

Write $d_{\mathrm{TV}} = \mathrm{TV}(p,r)$ and $F_\theta = F_r+\theta(F_p-F_r)$. As in the preceding CDF estimate, $\|F_p-F_r\|_\infty \le \delta d_{\mathrm{TV}}$. For $N \ge 3$, a second integration by parts yields

$$\int F_\theta^{N-2}\,\mathrm{d}G_i = \frac{1}{N-1} + (N-2)\int (F_\theta-G_i)F_\theta^{N-3}\,\mathrm{d}F_\theta \le \frac{1}{N-1}+\delta.$$

Differentiating the winner weight along $\theta$ and integrating gives

$$|W_i(p)-W_i(r)| \le N(N-1)\delta d_{\mathrm{TV}}\left(\frac{1}{N-1}+\delta\right) = L_\delta d_{\mathrm{TV}}.$$

For $N=2$, the differentiated integral equals one, giving the stronger coefficient $2\delta \le L_\delta$. Finally, $\mathcal{S}_N(p)_i - \mathcal{S}_N(r)_i = (p_i-r_i)W_i(p) + r_i\big(W_i(p)-W_i(r)\big)$. The first term has half its $\ell^1$ norm at most $(1+g_\delta)d_{\mathrm{TV}}$, and the second at most $L_\delta d_{\mathrm{TV}}/2$. This proves the total-variation bound. □

Under preservation, $h = L_\delta$ is admissible in Theorem 4. If $0 < g_\delta \le 1/4$, then $L_\delta \le 5g_\delta/4 \le 5/16$, so $C_{L_\delta} = \frac{L_\delta^2}{4(1-L_\delta)^2} \le \frac{100}{121}g_\delta^2 < g_\delta^2$. For every $\alpha \ge g_\delta^2$ with $\alpha < 1$, this implies $\beta(1+C_{L_\delta}) < 1$; $\alpha = 1$ recovers in one step. Taking the infimum in the definition of $\alpha_{\mathrm{crit}}$ proves (16) uniformly in $N$. At $\delta = 0$, all score laws coincide, $\mathcal{S}_N(p) = p$, and every positive reference weight recovers $q$. For example, at $\delta = 10^{-3}$ this gives sufficient weights $6.4\times10^{-5}$, $1.024\times10^{-3}$ and $1.6384\times10^{-2}$ for $N = 8, 32, 128$, respectively. These sufficient weights may be conservative relative to the sharp threshold.

### B.4 Approximate preservation and bounded update errors

In this subsection, $\mathcal{B} = \varnothing$, $q_i > 0$, and $\delta$ is the score CDF diameter. Define $\varepsilon = \mathrm{TV}(\mathcal{S}_N(q), q)$ and $\varepsilon_{\mathrm{rel}} = \max_i |W_i(q)-1|$. The following guarantees apply beyond the regime of small preservation errors of Theorem 5.

**Proposition B.2** (Robustness to imperfect preservation)**.** *Set $g_\delta = N\delta$ and $L_\delta = N\delta[1+(N-1)\delta]$, without assuming $\mathcal{S}_N(q) = q$. If $g_\delta \le 1/2$ and $s_\alpha := \alpha - 2\beta L_\delta^2 > 0$, then every exact trajectory satisfies*

$$\limsup_t \mathrm{TV}(p_t, q) \le \frac{2\varepsilon_{\mathrm{rel}}}{s_\alpha}. \tag{B.2}$$

*If instead* $\chi := \beta(1 + g_\delta + L_\delta/2) < 1$*, every exact trajectory converges to the unique fixed point* $p_\alpha$*, with*

$$\mathrm{TV}(p_\alpha, q) \le \frac{\beta\varepsilon}{1-\chi}. \tag{B.3}$$

*More generally, any simplex-valued sequence satisfying* $\mathrm{TV}(\widehat{p}_{t+1}, T_\alpha(\widehat{p}_t)) \le \varepsilon_{\mathrm{upd}}$ *at every round obeys* $\limsup_t \mathrm{TV}(\widehat{p}_t, q) \le (\beta\varepsilon + \varepsilon_{\mathrm{upd}})/(1-\chi)$.

We prove Proposition B.2 in the same all-correct, full-support setting, without requiring $\mathcal{S}_N(q) = q$. Let $e_i = W_i(q) - 1$ and $\|e\|_\infty = \varepsilon_{\mathrm{rel}}$. Since $q \cdot W(q) = p \cdot W(p) = 1$, monotonicity gives $q \cdot W(p) \ge 2 - p \cdot W(q) = 1 - (p-q)\cdot e$. In particular, $1 - q \cdot W(p) \le (p-q)\cdot e \le 2\varepsilon_{\mathrm{rel}}\,\mathrm{TV}(p,q)$. For interior $p$, put $D = D_{\mathrm{KL}}(q\|p)$, $d_{\mathrm{TV}} = \mathrm{TV}(p,q)$, and $\xi_i = W_i(p) - 1$. If $g_\delta \le 1/2$, Lemma B.1 implies $|\xi_i| \le 1/2$ and $|\xi_i| \le \varepsilon_{\mathrm{rel}} + L_\delta d_{\mathrm{TV}}$. Taylor's inequality $-\log(1+x) \le -x + 2x^2$ on $[-1/2, 1/2]$ therefore gives

$$\begin{aligned} D_{\mathrm{KL}}(q\|\mathcal{S}_N(p)) - D &\le 2\varepsilon_{\mathrm{rel}} d_{\mathrm{TV}} + 2(\varepsilon_{\mathrm{rel}} + L_\delta d_{\mathrm{TV}})^2 \\ &\le 2\varepsilon_{\mathrm{rel}} d_{\mathrm{TV}} + 4\varepsilon_{\mathrm{rel}}^2 + 4L_\delta^2 d_{\mathrm{TV}}^2 \\ &\le \sqrt{2}\varepsilon_{\mathrm{rel}}\sqrt{D} + 4\varepsilon_{\mathrm{rel}}^2 + 2L_\delta^2 D, \end{aligned}$$

where the last step uses Pinsker's inequality. Convexity under reference mixing yields

$$D_{t+1} \le (1 - s_\alpha) D_t + \beta\sqrt{2}\varepsilon_{\mathrm{rel}}\sqrt{D_t} + 4\beta\varepsilon_{\mathrm{rel}}^2, \qquad D_t = D_{\mathrm{KL}}(q\|p_t).$$

For $s_\alpha > 0$, Young's inequality bounds the middle term by $s_\alpha D_t/2 + \beta^2\varepsilon_{\mathrm{rel}}^2/s_\alpha$. Since $p_1 \ge \alpha q$, iteration proves the finite-time estimate

$$D_t \le (1 - s_\alpha/2)^{t-1}\log(1/\alpha) + \frac{2\varepsilon_{\mathrm{rel}}^2}{s_\alpha}\left(\frac{\beta^2}{s_\alpha} + 4\beta\right), \qquad t \ge 1.$$

Pinsker's inequality now gives $\limsup_t \mathrm{TV}(p_t, q) \le \frac{\varepsilon_{\mathrm{rel}}}{s_\alpha}\sqrt{\beta^2 + 4\beta s_\alpha} \le \frac{2\varepsilon_{\mathrm{rel}}}{s_\alpha}$, because $0 < s_\alpha \le \alpha = 1 - \beta$ and $\beta^2 + 4\beta(1-\beta) \le 4/3 < 4$. This proves (B.2), including boundary initializations. It bounds the asymptotic neighborhood without asserting convergence within that neighborhood.

For the total-variation conclusion, Lemma B.1 makes $T_\alpha$ a contraction with coefficient $\chi < 1$ on the compact simplex. Its iterates converge to its unique fixed point $p_\alpha$. Since $\mathrm{TV}(T_\alpha(q), q) = \beta\varepsilon$, the fixed-point equation and the triangle inequality give $\mathrm{TV}(p_\alpha, q) \le \chi\,\mathrm{TV}(p_\alpha, q) + \beta\varepsilon$, proving (B.3). For a sequence of approximate updates, put $e_t = \mathrm{TV}(\widehat{p}_t, q)$. The assumed per-round error gives

$$e_{t+1} \le \varepsilon_{\mathrm{upd}} + \chi e_t + \beta\varepsilon, \qquad e_t \le \chi^t e_0 + \frac{\beta\varepsilon + \varepsilon_{\mathrm{upd}}}{1-\chi}(1 - \chi^t).$$

The result assumes a uniform bound on the realized per-round update error; a specific fitting or sampling scheme must be checked against this assumption separately. Finally, if only $\varepsilon$ is known, every coordinate satisfies $|\mathcal{S}_N(q)_i - q_i| \le \varepsilon$, so $\varepsilon_{\mathrm{rel}} \le \varepsilon/\min_i q_i$. The dimension-free quadratic bound uses relative preservation error; the stronger contraction bound uses TV preservation error.

### B.5 Convergence near the reference

We first work on the correct support alone. Put $g_\delta = N\delta < 1$, $L_\delta = N\delta[1 + (N-1)\delta]$, and

$$A_\delta = \frac{(1+g_\delta)L_\delta^2}{4(1-g_\delta)^2}, \qquad \varepsilon_{\mathrm{rel}} = \max_i |W_i(q) - 1|,$$

$$\underline{a} = \alpha - \beta\varepsilon_{\mathrm{rel}}\left(1 + \frac{\beta L_\delta}{2\alpha}\right), \qquad \kappa_\alpha = \underline{a} - \beta A_\delta.$$

Let $p_\alpha$ be the unique fixed point from Proposition 3. The following argument proves the convergence and bias bounds used in Theorem 5.

Write $e = W(q) - \mathbf{1}$. The fixed-point equation gives $W_i(p_\alpha) - 1 = (\alpha/\beta)(1 - q_i/p_{\alpha,i})$. Monotonicity applied to $p_\alpha, q$ therefore yields

$$\frac{\alpha}{\beta}\sum_i \frac{(p_{\alpha,i} - q_i)^2}{p_{\alpha,i}} \le (p_\alpha - q)\cdot e \le 2\varepsilon_{\mathrm{rel}}\,\mathrm{TV}(p_\alpha, q).$$

Cauchy–Schwarz bounds the sum on the left below by $4\,\mathrm{TV}(p_\alpha, q)^2$, so

$$\mathrm{TV}(p_\alpha, q) \le \frac{\beta\varepsilon_{\mathrm{rel}}}{2\alpha}. \tag{B.4}$$

Lemma B.1 now implies $W_i(p_\alpha) - 1 \le \varepsilon_{\mathrm{rel}}(1 + \beta L_\delta/(2\alpha))$. Define the reference weights for individual coordinates $a_i = \frac{\alpha q_i}{p_{\alpha,i}} = 1 - \beta W_i(p_\alpha)$ and $a_{\min} = \min_i a_i \ge \underline{a}$. Since $g_\delta < 1$, all $W_i(p_\alpha)$ are positive and $0 < a_i < 1$.

To obtain a Lyapunov function centered at $p_\alpha$, use the nonnegative convex function $\phi(x) = x - 1 - \log x$. For interior $p$, the update identity and Jensen's inequality give

$$\begin{aligned}
\frac{T_\alpha(p)_i}{p_{\alpha,i}} &= a_i + (1 - a_i)\frac{p_i W_i(p)}{p_{\alpha,i} W_i(p_\alpha)}, \\
D_{\mathrm{KL}}(p_\alpha \| T_\alpha(p)) &= \sum_i p_{\alpha,i}\phi\left(\frac{T_\alpha(p)_i}{p_{\alpha,i}}\right) \\
&\le \sum_i \beta p_{\alpha,i} W_i(p_\alpha)\phi\left(\frac{p_i W_i(p)}{p_{\alpha,i} W_i(p_\alpha)}\right) \\
&= \sum_i \beta p_{\alpha,i} W_i(p_\alpha)\phi\left(\frac{p_i}{p_{\alpha,i}}\right) - \beta(p - p_\alpha)\cdot W(p_\alpha) \\
&\quad - \beta\sum_i p_{\alpha,i} W_i(p_\alpha)\log\frac{W_i(p)}{W_i(p_\alpha)}.
\end{aligned}$$

The last equality uses $p \cdot W(p) = p_\alpha \cdot W(p_\alpha) = 1$. The first term is at most $(1 - a_{\min})D_{\mathrm{KL}}(p_\alpha \| p)$ because $\phi \ge 0$. Let $\Delta_i = W_i(p) - W_i(p_\alpha)$. Both weights lie in $[1 - g_\delta, 1 + g_\delta]$, so Taylor's theorem gives $-W_i(p_\alpha)\log\frac{W_i(p)}{W_i(p_\alpha)} \le -\Delta_i + \frac{1+g_\delta}{2(1-g_\delta)^2}\Delta_i^2$. The combined linear term is

$$-\beta(p - p_\alpha)\cdot W(p_\alpha) - \beta p_\alpha \cdot (W(p) - W(p_\alpha)) = \beta(p - p_\alpha)\cdot(W(p) - W(p_\alpha)) \le 0.$$

Using $|\Delta_i| \le L_\delta\,\mathrm{TV}(p, p_\alpha)$ and Pinsker's inequality proves

$$D_{\mathrm{KL}}(p_\alpha \| T_\alpha(p)) \le (1 - a_{\min} + \beta A_\delta)D_{\mathrm{KL}}(p_\alpha \| p) \le (1 - \kappa_\alpha)D_{\mathrm{KL}}(p_\alpha \| p). \tag{B.5}$$

Thus $\kappa_\alpha > 0$ gives global convergence to $p_\alpha$. Boundary initializations enter the interior after one round because $p_1 \ge \alpha q$.

This condition permits quadratic mixing with a nonzero preservation error. If $0 < g_\delta \le 1/4$, then $L_\delta \le 5g_\delta/4$ and $A_\delta \le 125g_\delta^2/144$. With $\alpha = 2g_\delta^2$ and $\varepsilon_{\mathrm{rel}} \le g_\delta^3$,

$$\beta\varepsilon_{\mathrm{rel}}\left(1 + \frac{\beta L_\delta}{2\alpha}\right) \le \frac{9}{16}g_\delta^2, \qquad \kappa_\alpha \ge \left(2 - \frac{125}{144} - \frac{9}{16}\right)g_\delta^2 = \frac{41}{72}g_\delta^2 > 0.$$

Under exact preservation, $\varepsilon_{\mathrm{rel}} = 0$, and every $\alpha \ge g_\delta^2$ with $\alpha < 1$ satisfies $\kappa_\alpha > 0$ in the same range. The case $\alpha = 1$ recovers the reference immediately.

### B.6 Joint extinction and stabilization

We now allow both heterogeneous correct and incorrect answers. Assume $q_i > 0$ on $\mathcal{C}$ and $q_j = 0$ on $\mathcal{B}$. For Theorem 5, the quantities $g_\delta, L_\delta, \varepsilon_{\mathrm{rel}}, \underline{a}, \kappa_\alpha$ are computed using selection restricted to $\mathcal{C}$. Let $\overline{m}$ be as in (17), with $\overline{m} = 0$ if $\mathcal{B}$ is empty.

For $0 < z = p(\mathcal{B}) < 1$, let $x = p(\cdot \mid \mathcal{C})$ and let $H$ be the current conditional mixture of incorrect score CDFs. The conditional correct score CDF is $G_x = \sum_i x_i G_i$. Convexity of $u \mapsto u^{N-1}$ gives

$$N\int G_x^{N-1}\,\mathrm{d}H \le \sum_{i,j} x_i \frac{p_j}{z}\, N\int G_i^{N-1}\,\mathrm{d}H_j \le \overline{m}.$$

The scalar chord bound from Appendix A applies at this round with $G_x, H$, even though these mixtures change in later rounds. Consequently

$$z_{t+1} \le \beta\max\{1, \overline{m}\}z_t; \qquad z_{t+1} \le z_t - \alpha z_t^2 \quad \text{if } \beta\overline{m} = 1.$$

For the critical bound, use $R_H(z) \le z^2 + \overline{m}z(1-z)$. Both bounds extend to $z = 0, 1$. Hence $\beta\overline{m} \le 1$ implies $z_t \to 0$, with the rates stated in Theorem 5.

Let $T_\alpha^{\mathcal{C}}$ be the update restricted to the correct support. For $t \ge 1$, $x_t = p_t(\cdot \mid \mathcal{C})$ is defined. Couple a draw from $p_t$ with a draw from $x_t$ so that they agree whenever the former is correct. Coupling all $N$ candidates gives $\mathrm{TV}\big(T_\alpha(p_t), T_\alpha^{\mathcal{C}}(x_t)\big) \le \beta N z_t$. For any probability $v$ and any probability $u$ supported on $\mathcal{C}$, normalization of $v|_{\mathcal{C}}$ and the triangle inequality imply $\mathrm{TV}(v(\cdot \mid \mathcal{C}), u) \le \mathrm{TV}(v, u)$, provided $v(\mathcal{C}) > 0$. Indeed, the extra $\ell^1$ distance from normalization is $v(\mathcal{B})$, which is already part of $\|v - u\|_1$. Thus

$$\mathrm{TV}\big(x_{t+1}, T_\alpha^{\mathcal{C}}(x_t)\big) \le \beta N z_t. \tag{B.6}$$

Let $p_\alpha$ be the unique fixed point of $T_\alpha^{\mathcal{C}}$. Both distributions in (B.6) have each coordinate at least $\alpha q_i$, while $p_{\alpha,i}/(\alpha q_i) = 1/a_i \le 1/\underline{a}$. The mean-value bound for $\log$ and (B.5) give

$$\begin{aligned} \big|D_{\mathrm{KL}}(p_\alpha \| x_{t+1}) - D_{\mathrm{KL}}(p_\alpha \| T_\alpha^{\mathcal{C}}(x_t))\big| &\le \frac{2\beta N}{\underline{a}} z_t, \\ D_{t+1} \le (1 - \kappa_\alpha) D_t + \frac{2\beta N}{\underline{a}} z_t, \qquad D_t &= D_{\mathrm{KL}}(p_\alpha \| x_t), \quad t \ge 1. \end{aligned} \tag{B.7}$$

Since $0 < \kappa_\alpha < 1$ and $z_t \to 0$, iterating this recursion gives $D_t \to 0$: the contribution of any fixed initial segment decays geometrically, and the tail is bounded by a constant times $\sup_{s \ge t_0} z_s$, which tends to zero as $t_0 \to \infty$. Pinsker's inequality implies $x_t \to p_\alpha$, and $\mathrm{TV}(p_t, p_\alpha) \le z_t + \mathrm{TV}(x_t, p_\alpha)$ proves convergence of the full distribution. Equation (B.4) gives its bias relative to $q$. If selection preserves $q$ on $\mathcal{C}$, then $p_\alpha = q$.

### B.7 A FIXED POLYNOMIAL FAMILY WITH UNSTABLE ROTATION

For $0 < \varepsilon_{\mathrm{amp}} \le 1/4$ and $i = 0, 1, 2$, put $\theta_i = 2\pi i/3$ and define densities on $[0, 1]$ by

$$\begin{aligned} h_i(u) &= 1 + \varepsilon_{\mathrm{amp}}\{\cos\theta_i P_2(u) + \sin\theta_i P_3(u)\}, \\ P_2(u) = 6u^2 - 6u + 1, \qquad P_3(u) &= 20u^3 - 30u^2 + 12u - 1. \end{aligned} \tag{B.8}$$

These are the shifted Legendre polynomials of degrees two and three. Their absolute values are at most one, their integrals vanish, and $h_i \ge 1 - \sqrt{2}/4 > 0$. The densities and their derivatives on $[0, 1]$ have bounds independent of $\varepsilon_{\mathrm{amp}}$. Extended by zero, their total variations are uniformly bounded as well. They also form a fixed four-component Beta-mixture family. In the degree-three Bernstein basis, the coefficient vectors of $P_2$ and $P_3$ are $(1, -1, -1, 1)$ and $(-1, 3, -3, 1)$. Hence the coefficients $b_{ik}$ of $h_i$ satisfy $b_{ik} \ge 1 - \sqrt{10}/4 > 0$ and $\sum_{k=0}^{3} b_{ik} = 4$. Taking weights $b_{ik}/4$ on $\mathrm{Beta}(k+1, 4-k)$ gives exactly (B.8).

Let

$$f(u) = 2u^3 - 3u^2 + u, \qquad h(u) = 5u^4 - 10u^3 + 6u^2 - u, \qquad f_i = \cos\theta_i f + \sin\theta_i h.$$

Then $G_i(u) = u + \varepsilon_{\mathrm{amp}} f_i(u)$, $\sum_i f_i = 0$, and $\int_0^1 f_i = 0$. For $N = 2$, these identities give $F_q(u) = u$ and $W_i(q) = 1$ at the uniform target. Direct polynomial integration gives

$$\int_0^1 f h' \,\mathrm{d}u = \frac{1}{70}, \qquad \int_0^1 f_i' f_j \,\mathrm{d}u = \frac{1}{70}\sin(\theta_i - \theta_j).$$

Consequently

$$W_i(p) = 1 + \gamma(p_{i-1} - p_{i+1}), \qquad \gamma = \frac{\sqrt{3}\,\varepsilon_{\mathrm{amp}}^2}{70}, \tag{B.9}$$

with cyclic indices. The two tangent eigenvalues of $DT_\alpha(q)$ are $\beta(1 \pm i\varepsilon_{\mathrm{amp}}^2/70)$. Thus $q$ is a strict local repeller whenever

$$0 < \alpha < 1 - \left(1 + \frac{\varepsilon_{\mathrm{amp}}^4}{4900}\right)^{-1/2}. \tag{B.10}$$

No point other than $q$ maps to $q$. Equation (B.9) gives $\sum_i W_i(p) = 3$ and $W_i(p) > 0$, hence

$$\prod_i \mathcal{S}_N(p)_i = \prod_i p_i \prod_i W_i(p) \le \prod_i p_i \le \frac{1}{27}.$$

If $T_\alpha(p) = q$, then $\mathcal{S}_N(p) = q$, so equality holds throughout and $p = q$. Near $q$ the tangent derivative is a rotation followed by expansion, and continuity gives $\|T_\alpha(p) - q\|_2 \geq \lambda\|p - q\|_2$ for some $\lambda > 1$ in a sufficiently small neighborhood. Any orbit converging to $q$ must therefore eventually equal $q$, and the preimage argument makes its initial point $q$. Proposition 3 excludes any other fixed-point limit, so every other trajectory fails to converge.

The CDF diameter is $\delta = \varepsilon_{\text{amp}} D_{\text{shape}}$, where $D_{\text{shape}} = \max_{i,j} \|f_i - f_j\|_\infty > 0$ is fixed. The threshold in (B.10) is asymptotic to $\varepsilon_{\text{amp}}^4/9800$, giving an obstruction of order $\delta^4$ with no increasing frequency. Appendix I proves cubic sharpness at $N = 2$ on a fixed score interval under uniformly bounded density derivatives, using a different, variable-frequency family; at each fixed $N \geq 3$, sufficiently small heterogeneity needs only arbitrarily small positive mixing. The sharp threshold for this fixed low-degree family is not determined by the displayed local obstruction. If densities smooth and positive on all of $\mathbb{R}$ are desired, apply the common transformation $u \mapsto \log(u/(1-u))$. The transformed density is $h_i(\sigma(x))\sigma'(x)$, with $\sigma(x) = (1 + e^{-x})^{-1}$; it is positive and smooth with uniformly bounded derivatives of every fixed order. This transformation preserves score comparisons and CDF uniform distances, so all dynamical conclusions remain unchanged.

### B.8 Target preservation in common location families

An ordered parametric family gives a useful boundary to the worst-case construction.

**Lemma B.3** (Common location noise)**.** *Suppose scores have the form $\mu_i + E$, where every occurrence draws $E$ independently from a common everywhere positive continuous density on $\mathbb{R}$. For $N \geq 2$ and full-support $q$, $\mathcal{S}_N(q) = q$ if and only if all $\mu_i$ coincide.*

*Proof.* The mixture CDF $F_q$ is strictly increasing on $\mathbb{R}$. If $\mu_i > \mu_j$, then $W_i(q) = N\mathbb{E}F_q(\mu_i + E)^{N-1} > N\mathbb{E}F_q(\mu_j + E)^{N-1} = W_j(q)$. Preservation requires every $W_i(q) = 1$, so no two locations can differ. Conversely, equal locations give identical score laws, for which $\mathcal{S}_N(p) = p$ for every $p$. The lemma applies to Gaussian scores with a common variance and to Gumbel scores with a common scale. The lemma is specific to location changes with a fixed noise law; arbitrary shape or scale changes fall outside this setting. □

### B.9 Three continuous dice and the two-answer case

Recall the face sets $D_0 = \{2, 4, 9\}$, $D_1 = \{1, 6, 8\}$ and $D_2 = \{3, 5, 7\}$, with independent uniform jitter on $(-1/4, 1/4)$. The nine comparisons between the faces of $D_0$ and $D_1$ have five wins for $D_0$; the same count holds for $D_1$ against $D_2$ and $D_2$ against $D_0$. The jitter leaves these comparisons unchanged because distinct faces are separated by at least one and the difference of jitters lies in $(-1/2, 1/2)$. Scores are continuous, and equal-label competitions are exchangeable. With cyclic indices $i = 0, 1, 2$,

$$\mathcal{S}_N(p)_i = p_i^2 + 2p_i\left(\frac{5}{9}p_{i+1} + \frac{4}{9}p_{i-1}\right) = p_i[1 + \gamma(p_{i+1} - p_{i-1})], \qquad \gamma = \frac{1}{9}.$$

Thus $q = (1/3, 1/3, 1/3)$ is preserved. On the two-dimensional tangent plane, the skew cyclic difference matrix $\Omega v = (v_{i+1} - v_{i-1})_i$ has eigenvalues $\pm i\sqrt{3}$. Hence $DT_\alpha(q) = \beta(I + \gamma\Omega/3)$ is a Euclidean dilation followed by a rotation, with dilation factor $\lambda = \beta\sqrt{1 + \gamma^2/3}$. Suppose $0 < \alpha < 1 - \sqrt{243/244}$, so $\lambda > 1$. Choose a small closed ball about $q$ and $\lambda_0 > 1$ such that $\|T_\alpha(p) - q\|_2 \geq \lambda_0\|p - q\|_2$ throughout that ball. There are no other preimages of $q$. Indeed, put $d_i = p_{i+1} - p_{i-1}$; the three positive factors $1 + \gamma d_i$ have sum three, so

$$\prod_i \mathcal{S}_N(p)_i = \left(\prod_i p_i\right)\prod_i(1 + \gamma d_i) \leq \prod_i p_i \leq \frac{1}{27}.$$

If $T_\alpha(p) = q$, then $\mathcal{S}_N(p) = q$, and equality forces $p = q$. On the compact set outside a ball of radius $r$ on which local expansion holds, $m = \min \|T_\alpha(p) - q\|_2 > 0$ since no point in this set maps to $q$. Every nonfixed trajectory starting inside the ball must eventually exit, since its positive distance expands while it remains inside. After its first exit, its distance is always at least $\min\{r, m\}$: outside the ball it is at least $r$, an entry from outside has distance at least $m$, and subsequent steps inside

expand. Thus every non-$q$ orbit stays a positive distance from $q$ eventually. A convergent orbit of a continuous map must converge to a fixed point, and Proposition 3 leaves only $q$.

With exactly two answers, write $p = (x, 1-x)$. The probability that answer one wins is nondecreasing in $x$. Couple label assignments by uniform thresholds, pre-drawing one independent score for each possible label at every position. Switching positions from label two to label one cannot turn a winner with label one into a winner with label two: the former winning candidate remains, and the competing label-two scores form a subset of their former set. The scalar update is therefore continuous and nondecreasing. Preservation and Proposition 3 give a unique fixed point $q_1$. For $x < q_1$, one has $x < T(x) \le q_1$: otherwise monotone iteration from $x$ would yield a fixed point below $q_1$. Reversing the inequalities gives the corresponding argument above $q_1$, proving global recovery.

## C An Explicit Lower Bound for Every Candidate Count

This appendix proves Theorem 6 by verifying the score laws at finite frequencies and then bounding the full convergence basin. All constants below depend only on the fixed integer $N \ge 2$.

For every $N$ and arbitrarily small CDF diameter, the construction must preserve balance while retaining rotation. Let $g_i$ be cyclic translates of a mean-zero triangle wave and let $A$ vanish at the endpoints of $[0, 1]$. With integer frequency $K$ and fixed amplitude $c \in (0, 1)$, set

$$G_{i,K}(u) = u + A(u)\left[\frac{c}{K} g_i(Ku) - \zeta_{i,K}\right].$$

The corrections $\zeta_{i,K} = O_N(K^{-3})$ make the perturbations sum to zero and have zero $u^{N-2}$-weighted integral. Thus $F_q(u) = u$ and integration by parts gives $W_i(q) = 1$ exactly for uniform $q$. The definitions below specify every coefficient and verify positive densities. Write $\delta_K$ for their CDF diameter. The program `figures/lower_bound_geometry.py` generates the editable figure and checks the calibration integrals and a positive density bound. Panels (b)–(c) are schematic.

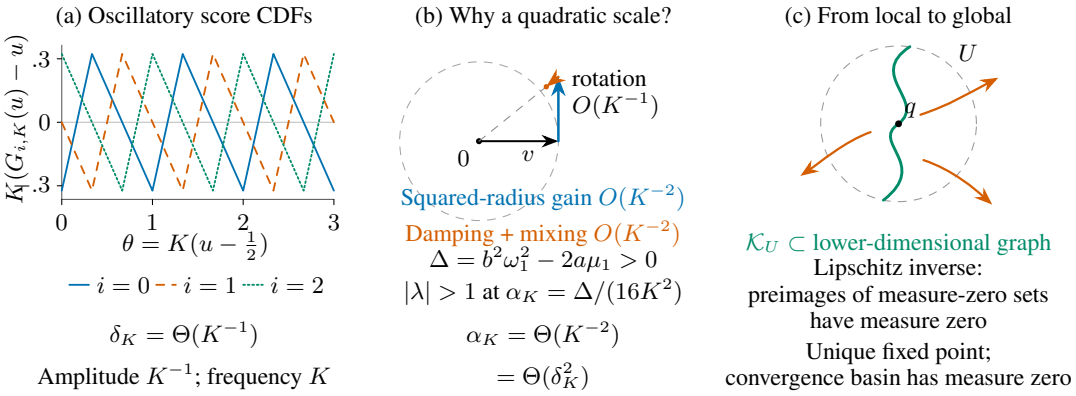



Figure 4: Construction and geometry underlying Theorem 6. (a) Centered score CDFs for $N = 2$, $K = 8192$, including the calibration corrections, on three central periods with rescaled axes. (b) Schematic transverse displacement and radial gain; damping and reference mixing act at order $K^{-2}$. (c) A schematic slice through the lower-dimensional graph containing points that stay near the fixed point. The global inverse bound makes its countably many preimages have measure zero. Panel (a) illustrates valid score CDFs; the instability theorem requires $K \ge K_*(N)$.

### C.1 Score distributions and exact target preservation

Set

$$d = \begin{cases} 3, & N = 2, \\ \lceil 2\pi\sqrt{N} \rceil, & N \ge 3, \end{cases} \qquad s = N - 2, \qquad J = 256N^2, \qquad c = 1 - \frac{1}{16N},$$

$$A(u) = (1 - u^J)^4[1 - (1 - u)^J]^4.$$

The envelope is defined on $[0, 1]$. Define a continuous one-periodic triangle wave by

$$g(v) = \begin{cases} (d-1)v - \dfrac{d-1}{2d}, & 0 \le v \le 1/d, \\ 1 - v - \dfrac{d-1}{2d}, & 1/d \le v \le 1, \end{cases} \qquad g_i(v) = g(v - i/d), \quad 0 \le i < d.$$

Each $g_i$ has mean zero and absolute value at most $1/2$. Almost everywhere, $g_i' = d\mathbf{1}_{[i/d,(i+1)/d)} - 1$, with the indicator understood modulo one. Hence $\sum_i g_i' = 0$ and, using the zero means, $\sum_i g_i = 0$. For an integer $K \ge 1$, the following definitions give perturbations with zero sum and zero weighted integral:

$$J_A = \int_0^1 u^s A(u)\,\mathrm{d}u, \qquad \zeta_{i,K} = \frac{c}{KJ_A}\int_0^1 u^s A(u) g_i(Ku)\,\mathrm{d}u,$$

$$f_{i,K}(u) = A(u)\left[\frac{c}{K} g_i(Ku) - \zeta_{i,K}\right], \qquad G_{i,K}(u) = u + f_{i,K}(u), \tag{C.1}$$

$$\sum_i f_{i,K} = 0, \qquad \int_0^1 u^s f_{i,K}(u)\,\mathrm{d}u = 0. \tag{C.2}$$

Extend $G_{i,K}$ by zero below zero and by one above one. The sum of the $\zeta_{i,K}$ is zero, giving the zero-sum identity; their definition gives the zero weighted integrals.

Once monotonicity of each $G_{i,K}$ is established, these are genuine score CDFs. For uniform $q_i = 1/d$, their mixture CDF is exactly $F_q(u) = u$. Since $f_{i,K}$ vanishes at both endpoints, integration by parts gives

$$W_i(q) = N\int_0^1 u^{N-1}\,\mathrm{d}G_{i,K}(u) = 1 - N(N-1)\int_0^1 u^{N-2} f_{i,K}(u)\,\mathrm{d}u = 1.$$

Thus preservation holds exactly at every admissible finite $K$.

To control the oscillatory corrections, we use the following averaging estimate. If $r$ is bounded, one-periodic, and has mean zero, it has a one-periodic mean-zero antiderivative with sup norm at most $\|r\|_\infty$. Indeed, the antiderivative's oscillation is at most $\int_0^1 |r| \le \|r\|_\infty$, and subtracting its mean preserves this bound. Repeating this step gives antiderivatives with the same bound. Integration by parts therefore yields

$$\left|\int_0^1 w(u) r(Ku)\,\mathrm{d}u\right| \le \frac{\|r\|_\infty \|w^{(k)}\|_1}{K^k}, \qquad k = 1, 2, \tag{C.3}$$

provided $w$ and its derivatives up to order $k-1$ vanish at both endpoints. For a nonzero-mean $r$, apply the estimate to $r - \int_0^1 r$, whose norm is at most $2\|r\|_\infty$. Piecewise continuous $r$, including the derivatives of $g_i$, satisfy the same estimate using absolutely continuous antiderivatives.

### C.2 Positive densities and CDF diameter

The constants $D$ and $B$ below control the envelope derivatives and the correction constants, respectively. The threshold $K_0$ keeps these corrections within the density margin $1 - c$ and below the CDF separation of order $1/K$. Let

$$D = 8J, \qquad B = 2(N-1)(s+D)^2, \qquad K_0 = \left\lceil \max\left\{6, 32ND, (64NBD)^{1/3}, \sqrt{12B/c}\right\}\right\rceil.$$

The product representation of $A$ gives $0 \le A \le 1$, $|A'| \le D$, and $|A''| \le D^2$. The functions $u^s A$ and their first derivatives vanish at both endpoints. Also $A(u)^2 \ge 1 - 8u^J - 8(1-u)^J$. It follows that, with $I_1 = \int_0^1 u^{N-2} A^2\,\mathrm{d}u$,

$$I_1 \ge \frac{1}{N-1} - \frac{16}{J} \ge \frac{c}{N-1}, \qquad J_A \ge I_1 \ge \frac{1}{2(N-1)}. \tag{C.4}$$

For the first bound, each of the two error integrals is at most $1/J$.

Applying (C.3) twice with $r = g_i$ and $\|(u^s A)''\|_1 \le (s+D)^2$ proves $|\zeta_{i,K}| \le B/K^3$. At points of differentiability, (C.1) gives

$$G_{i,K}' = 1 + cAg_i'(Ku) + A'\left[\frac{c}{K} g_i(Ku) - \zeta_{i,K}\right] \ge 1 - c - \frac{D}{2K} - \frac{BD}{K^3}.$$

For $K \geq K_0$, the last two terms are each at most $1/(64N)$, leaving $G'_{i,K} \geq 1/(32N)$. These CDFs are absolutely continuous, start at zero, and end at one.

The same bounds give $|f_{i,K}| \leq c/(2K) + B/K^3 \leq 1/K$ and $\|G_{i,K} - G_{j,K}\|_\infty \leq c/K + 2B/K^3 \leq 2/K$. To bound the diameter from below, choose an integer $m$ with $u = m/K \in [1/3, 2/3]$; such a choice exists for $K \geq 6$. At this point $g_0(Ku)$ is the minimum of $g$, while $g_{d-1}(Ku)$ is its maximum. Their difference is $(d-1)/d \geq 2/3$. Moreover $A(u) \geq [1 - (2/3)^J]^8 \geq 1/2$. Using $K^2 \geq 12B/c$,

$$G_{d-1,K}(u) - G_{0,K}(u) \geq \frac{1}{2}\left(\frac{2c}{3K} - \frac{2B}{K^3}\right) \geq \frac{c}{4K} \geq \frac{1}{8K}.$$

This proves the density and diameter bounds in Theorem 6.

### C.3 The tangent Jacobian and its remainder

Suppress the subscript $K$ on $f_i$. On the tangent space $E = \{v \in \mathbb{R}^d : \sum_i v_i = 0\}$, preservation and $q_i = 1/d$ give

$$D\mathcal{S}_N(q)v = v + \frac{N(N-1)}{d}\left(\int_0^1 u^s \sum_j v_j f_j(u)\,\mathrm{d}G_i(u)\right)_{i=0}^{d-1}.$$

The term with $\mathrm{d}u$ vanishes by (C.2). Thus, on $E$, the derivative minus the identity is represented by the matrix with entries $\frac{N(N-1)}{d}\int_0^1 u^s f'_i f_j\,\mathrm{d}u$. Define

$$\Omega_{ij} = \int_0^1 g'_i g_j\,\mathrm{d}v, \qquad Q_{ij} = \int_0^1 g_i g_j\,\mathrm{d}v.$$

Periodicity gives $\Omega^\top = -\Omega$, and $Q$ is a Gram matrix. Both are circulant and annihilate the constant vector.

In addition to $I_1$, for $N \geq 3$ put $I_0 = \int_0^1 u^{N-3}A^2\,\mathrm{d}u$, and set

$$b = \frac{N(N-1)c^2 I_1}{d}, \qquad a = \begin{cases} 0, & N = 2, \\ \dfrac{N(N-1)(N-2)c^2 I_0}{2d}, & N \geq 3. \end{cases} \tag{C.5}$$

To resolve the order-$K^{-2}$ stability margin, we bound the Jacobian remainder at order $K^{-3}$. We will show

$$D\mathcal{S}_N(q)|_E = \left(I + \frac{b}{K}\Omega - \frac{a}{K^2}Q\right)\Big|_E + R_K, \qquad \|R_K\|_2 \leq \frac{C_{\mathrm{rem}}}{K^3}, \tag{C.6}$$

where an admissible explicit constant is

$$E_0 = 4d(s+2D)^2 + sD + 2D^2 + B(d+D) + B^2 D, \qquad C_{\mathrm{rem}} = N(N-1)E_0.$$

Expand the integrand using (C.1):

$$f'_i f_j = \frac{c^2}{K}A^2 g'_i(Ku)g_j(Ku) + \frac{c^2}{K^2}AA' g_i(Ku)g_j(Ku) + \mathcal{R}_{ij}(u).$$

The terms containing the correction constants $\zeta_{i,K}$ obey $|\mathcal{R}_{ij}| \leq \frac{Bd}{K^3} + \frac{BD}{K^4} + \frac{B^2 D}{K^6}$. For the first main term, we apply the averaging estimate with two integrations by parts to $w = u^s A^2$ and the periodic function $g'_i g_j$ minus its mean $\Omega_{ij}$. Its centered sup norm is at most $d$, and $\|w''\|_1 \leq (s+2D)^2$. The resulting error, including the factor $c^2/K$, is at most $d(s+2D)^2/K^3$, which is bounded by the first term of $E_0/K^3$. For the second main term, apply one integration by parts to $w = u^s AA'$ and $g_i g_j - Q_{ij}$. The corresponding norm bounds give an error of at most $(sD + 2D^2)/(2K^3)$. Finally,

$$\int_0^1 u^s AA'\,\mathrm{d}u = \begin{cases} 0, & s = 0, \\ -\dfrac{s}{2}\int_0^1 u^{s-1}A^2\,\mathrm{d}u, & s > 0. \end{cases}$$

This identifies the damping coefficient in (C.5). For each matrix entry, the combined error is at most $E_0/K^3$. Multiplying by $N(N-1)/d$ and applying $\|X\|_2 \leq d\max_{ij}|X_{ij}|$ converts these entrywise bounds into (C.6). The tangent restriction is legitimate because both the exact derivative and its circulant approximation preserve $E$.

### C.4 Exact Fourier eigenvalues and the instability margin

For $1 \le \ell < d$, let $v_j^{(\ell)} = \exp(2\pi \mathrm{i}\ell j/d)$ and $x_\ell = \pi\ell/d$. The eigenvalues are

$$\Omega v^{(\ell)} = -i\omega_\ell v^{(\ell)}, \quad \omega_\ell = \frac{1}{2}\cot x_\ell, \qquad Qv^{(\ell)} = \mu_\ell v^{(\ell)}, \quad \mu_\ell = \frac{1+2\cos^2 x_\ell}{12d\sin^2 x_\ell}. \tag{C.7}$$

For $\Omega$, direct integration on the first cell gives $\Omega_{00} = 0$ and $\Omega_{0j} = j/d - 1/2$ for $1 \le j < d$. Summing this row against $\exp(2\pi i\ell j/d)$ yields $-\frac{i}{2}\cot x_\ell$. For $Q$, the Fourier coefficients of the mean-zero triangle wave are, for $n \neq 0$, $\widehat{g}_n = -\frac{d(1-e^{-2\pi i n/d})}{4\pi^2 n^2}$. The circulant Gram eigenvalue is $d\sum_{n\equiv\ell \pmod d} |\widehat{g}_n|^2$. Substitution gives (C.7), using

$$\sum_{k\in\mathbb{Z}} (k+x)^{-4} = \frac{\pi^4(2+\cos 2\pi x)}{3\sin^4 \pi x}, \qquad x \notin \mathbb{Z}.$$

This identity follows by twice differentiating $\sum_k (k+x)^{-2} = \pi^2 \csc^2(\pi x)$; the latter is the partial-fraction identity obtained from the logarithmic derivative of the sine product. All differentiated series converge locally uniformly away from the integers.

Set

$$\Delta = b^2\omega_1^2 - 2a\mu_1, \qquad \gamma_N = \Delta/16.$$

For $N = 2$, $d = 3$, $a = 0$, and $\omega_1^2 = 1/12$, so $\Delta = b^2/12 > 0$. For $N \ge 3$, note $I_0 \le 1/(N-2)$ and use (C.4) to obtain

$$\frac{b^2\omega_1^2}{2a\mu_1} = \frac{N(N-1)c^2 I_1^2}{(N-2)I_0}\,\frac{\omega_1^2}{d\mu_1} \ge \frac{N}{N-1}c^4 r_d, \qquad r_d = \frac{3\cos^2(\pi/d)}{1+2\cos^2(\pi/d)}.$$

Since $\pi/d \le 1/(2\sqrt{N})$, $r_d \ge 1 - 1/(8N)$ and $c^4 \ge 1 - 1/(4N)$. Hence $c^4 r_d \ge 1 - 3/(8N) > 1 - 1/N$, making the last ratio strictly greater than one. Thus $\Delta > 0$ for every $N$.

To make the frequency threshold explicit, put $\gamma_{\mathrm{sep}} = b\pi/(2d)$ and $C_{\mathrm{der}} = N(N-1)(1+\sqrt{d})$. These constants control spectral separation and the global derivative bound, respectively. Take

$$K_* = 1 + \left\lceil \max\left\{ K_0,\ 2b\omega_1,\ \sqrt{2a\mu_1},\ \sqrt{8C_{\mathrm{rem}}/\gamma_{\mathrm{sep}}},\ 8C_{\mathrm{rem}}/\Delta,\ \sqrt{\Delta},\ 4C_{\mathrm{der}} \right\} \right\rceil.$$

All terms are finite; those involving $a$ vanish when $N = 2$. To verify instability for $K \ge K_*$, let $\widetilde{D}_K = (I + b\Omega/K - aQ/K^2)|_E$. The matrices are simultaneously unitarily diagonalized by the Fourier basis, so $\widetilde{D}_K$ is normal. The distinct $\omega_\ell$ have consecutive gaps at least $\pi/(2d)$ by differentiating $\frac{1}{2}\cot x$. Thus distinct eigenvalues of $\widetilde{D}_K$ are separated by at least $\gamma_{\mathrm{sep}}/K$ through their imaginary parts. The perturbation bound $C_{\mathrm{rem}}/K^3 < \gamma_{\mathrm{sep}}/(8K)$ and continuity along $\widetilde{D}_K + tR_K$, $0 \le t \le 1$, imply that each separated eigenvalue has a corresponding perturbed eigenvalue within distance $C_{\mathrm{rem}}/K^3$. Indeed, normality gives $\|(zI - \widetilde{D}_K)^{-1}\|_2 = 1/\operatorname{dist}(z, \operatorname{spec}(\widetilde{D}_K))$. Outside these disks, the product of this norm and $\|tR_K\|_2$ is below one, so $zI - \widetilde{D}_K - tR_K$ is invertible. The spectrum stays in the disjoint disks for every $t \in [0,1]$, and continuity prevents an eigenvalue from moving between them. In particular, an eigenvalue remains near the eigenvalue associated with the first Fourier mode. The first eigenvalue of $\widetilde{D}_K$ has squared modulus $|1 - a\mu_1/K^2 - ib\omega_1/K|^2 = 1 + \Delta/K^2 + a^2\mu_1^2/K^4$. Since $K > \sqrt{\Delta}$, we can apply $\sqrt{1+x} \ge 1 + x/3$ for $0 \le x \le 1$. Using also $K > 8C_{\mathrm{rem}}/\Delta$ to bound the perturbation, we find that the corresponding exact eigenvalue of $D\mathcal{S}_N(q)$ has modulus at least $1 + \Delta/(8K^2)$. At $\alpha_K = \Delta/(16K^2)$, the reference update multiplies tangent eigenvalues by $1 - \alpha_K$. Since $(1 - x/16)(1 + x/8) > 1$ for $0 < x \le 1$, $DT_{\alpha_K}(q)$ has an eigenvalue of modulus greater than one.

### C.5 From instability to a convergence basin of measure zero

We first prove the required global inverse bound. Because $\|G_{i,K} - u\|_\infty \le 1/K$ and the mixture at $q$ is $u$, $|W_i(p) - 1| \le \frac{N(N-1)}{K}$. For $v \in E$, $\|\sum_i v_i G_{i,K}\|_\infty = \|\sum_i v_i f_{i,K}\|_\infty \le \sqrt{d}\,\|v\|_2/K$. Differentiate $\mathcal{S}_N(p)_i = p_i W_i(p)$ and use $\|p\|_2 \le 1$ to obtain

$$\sup_{p\in\Delta_a} \|D\mathcal{S}_N(p) - I\|_{E\to E} \le \frac{N(N-1)(1+\sqrt{d})}{K} = \frac{C_{\mathrm{der}}}{K}.$$

For $K \geq K_*$, $\alpha_K < 1/16$ and $C_{\mathrm{der}}/K < 1/4$, so $\sup_p \|DT_{\alpha_K}(p) - I\| < 1/2$. Integration along a segment in the simplex gives

$$\|T_{\alpha_K}(p) - T_{\alpha_K}(r)\|_2 \geq \frac{1}{2}\|p - r\|_2. \tag{C.8}$$

In particular, the map is injective and its inverse on its image is Lipschitz.

The remaining step is local: points whose orbits stay near an unstable fixed point form a set of measure zero.

**Lemma C.1** (Points remaining near an unstable fixed point)**.** *Let $T$ be a $C^1$ map near a fixed point in a finite-dimensional real affine space. If its derivative has an eigenvalue of modulus greater than one, there is a compact neighborhood $U$ of the fixed point such that the set $\mathcal{K}_U = \{x \in U : T^t(x) \in U \text{ for all } t \geq 0\}$ has Lebesgue measure zero.*

*Proof.* Translate the fixed point to zero and split the real invariant subspaces as $E^u \oplus E^{cs}$, placing all eigenvalues of modulus greater than one in $E^u$. This subspace has positive dimension. Choose equivalent norms and numbers $\mu > \sigma > 1$ such that the unstable linear part $A$ satisfies $\|Au\| \geq \mu\|u\|$ and the remaining part $B$ satisfies $\|Bv\| \leq \sigma\|v\|$. Such norms can be obtained by rescaling the finite Jordan blocks; equivalently, they follow from the spectral radii of $A^{-1}$ and $B$. Write $T(u, v) = (Au + f(u, v), Bv + g(u, v))$, where both nonlinear derivatives vanish at zero. On a sufficiently small convex compact product neighborhood, $f$ and $g$ each have Lipschitz constant at most $\epsilon$ in the norm $\|u\| + \|v\|$. Choose $\epsilon$ so that $\mu - 2\epsilon > \max\{1, \sigma + 2\epsilon\}$. Consider the difference of two points whose forward orbits stay in this neighborhood. If at some time $\|\Delta u\| \geq \|\Delta v\|$, the next difference satisfies

$$\|\Delta u'\| \geq (\mu - 2\epsilon)\|\Delta u\|, \qquad \|\Delta v'\| \leq (\sigma + 2\epsilon)\|\Delta u\| < \|\Delta u'\|.$$

This cone remains invariant, and a nonzero unstable difference grows geometrically, contradicting boundedness of the two orbits. Thus any two distinct points in $\mathcal{K}_U$ satisfy $\|\Delta u\| < \|\Delta v\|$.

Projection onto $E^{cs}$ is injective on $\mathcal{K}_U$, and its inverse represents $\mathcal{K}_U$ as a Lipschitz graph over a subset of $E^{cs}$. The set is compact, hence measurable. Since $\dim E^{cs}$ is less than the ambient dimension, this graph has ambient Lebesgue measure zero. If $E^{cs}$ is zero-dimensional, the same argument makes $\mathcal{K}_U$ contain at most one point. □

Apply the lemma in the simplex's affine hull, where the uniform target lies in the interior. The selection map is polynomial in $p$, so its $C^1$ extension is available in a neighborhood of $q$. By (C.8), each inverse iterate is Lipschitz on its image and therefore maps null sets to null sets. Every trajectory converging to $q$ eventually remains in $U$, so its initial point lies in the countable union $\bigcup_{t \geq 0} T_{\alpha_K}^{-t}(\mathcal{K}_U)$, which has relative measure zero.

Choose $\varepsilon_{N,K} > 0$ so that the Euclidean ball of that radius about $q$ lies in $U$. Outside this union of measure zero, an orbit leaves $U$ infinitely often, giving $\limsup_t \|p_t - q\|_2 \geq \varepsilon_{N,K}$. Proposition 3 makes $q$ the unique fixed point. Since every convergent orbit of a continuous map has a fixed-point limit, almost every orbit converges to no distribution. This proves Theorem 6.

### C.6 The uniform lower bound

For any sufficiently small $\delta > 0$, choose the integer $K = \lceil 2/\delta \rceil \geq K_*$. The constructed diameter obeys $\delta_K \leq 2/K \leq \delta$. For $\delta \leq 1$, one has $K \leq 3/\delta$, and hence $\alpha_K = \gamma_N/K^2 \geq \gamma_N \delta^2/9$. A uniform sufficient threshold $a \leq \alpha_K$ would require recovery at this reference weight, contradicting Theorem 6. Taking the infimum over valid $a$ proves $\alpha_{\mathrm{crit}}(N, \delta) \geq \gamma_N \delta^2/9$. Together with the explicit upper estimate in Appendix B, this establishes Theorem 2. The construction has positive densities for each $K$, while their oscillation frequency increases with $K$.

The lower prefactor can also be bounded explicitly in $N$. For $N \geq 3$, the same envelope estimate gives $I_0 \geq c/(N-2)$. The ratio estimate above yields

$$\frac{b^2\omega_1^2}{2a\mu_1} \geq 1 + \frac{5}{8(N-1)}, \qquad 2a\mu_1 \geq \frac{N(N-1)c^3}{12\pi^2},$$

where the second inequality uses $d^2 \sin^2(\pi/d) \le \pi^2$ and $1 + 2\cos^2(\pi/d) \ge 1$. Hence

$$\Delta \ge \frac{5Nc^3}{96\pi^2} \ge \frac{N}{24\pi^2}, \qquad \frac{\gamma_N}{9} \ge \frac{N}{3456\pi^2}.$$

The last step uses $c \ge 15/16$ and $c^3 \ge 4/5$. Thus $c_N = N/(3456\pi^2)$ is admissible for $N \ge 3$ whenever $0 < \delta \le \min\{1, 2/K_*(N)\}$. This prefactor does not decay with $N$, but the explicit frequency requirement restricts the admissible diameter as $N$ grows. It does not match the $N^2\delta^2$ upper bound over its full joint range.

# D Finite-Corpus Dynamics and Capped Reference Schedules

## D.1 Extinction and persistence with finitely many winners

For $M$ independent retained winners per round, let $\widehat{S}_{M,t}$ be their empirical distribution and set $p_{t+1} = \alpha q + \beta \widehat{S}_{M,t}$, with fixed $0 < \alpha < 1$. Define $\tau = \inf\{t : z_t = 0\}$. Let $\mathcal{F}_t$ contain the history through $p_t$. Each of the $M$ competitions at round $t$ is conditionally independent and has winner distribution $\mathcal{S}_N(p_t)$. Thus $\mathbb{E}[p_{t+1} \mid \mathcal{F}_t] = \alpha q + \beta \mathcal{S}_N(p_t)$. Under common correct scores, the chord bound applies to this conditional mean of incorrect mass. If $c = \beta \max\{1, m_*\} < 1$, induction gives $\mathbb{E}z_t \le c^t z_0$. For $t \ge 1$, $z_t$ is $\beta/M$ times the number of incorrect retained winners. Since zero incorrect mass is absorbing,

$$\mathbb{P}(\tau > t) = \mathbb{P}(z_t > 0) \le \frac{M}{\beta}\mathbb{E}z_t \le \frac{Mz_0}{\beta}c^t.$$

Combining this estimate with the trivial probability bound of one gives

$$\mathbb{P}(\tau > t) \le \min\{1, Mz_0c^t/\beta\}, \qquad t \ge 1.$$

At criticality, writing $x_t = \mathbb{E}z_t$, Jensen's inequality gives $x_{t+1} \le x_t - \alpha x_t^2$, and the reciprocal argument from Appendix A yields $x_t \le z_0/(1 + \alpha z_0 t)$. For any fixed $M$ and $\alpha > 0$, $p_t(\mathcal{C}) \ge \alpha$ for $t \ge 1$. The event that all $MN$ candidate labels at the next round are correct has conditional probability at least $\alpha^{MN}$ and implies extinction. Consequently,

$$\mathbb{P}(\tau > t) \le (1 - \alpha^{MN})^{t-1}, \quad t \ge 1, \qquad \mathbb{E}\tau \le 1 + \alpha^{-MN}.$$

These statements allow extinction in the first round and $\tau = 0$ when initialized without incorrect mass.

For persistence, suppose all incorrect scores share $H$ with $\beta m_H > 1$. The population map $F(z) = \beta R_H(z)$ has a unique positive fixed point $z_* \in (0, \beta)$. Choose $0 < \ell < z_* < u < \beta$, let $d_0 = \min\{F(\ell) - \ell, u - F(u)\} > 0$, and initialize $z_0 \in [\ell, u]$. We claim

$$\mathbb{P}\big(z_t \in [\ell, u] \text{ for every } 0 \le t \le T\big) \ge 1 - 2T\exp(-2Md_0^2/\beta^2). \tag{D.1}$$

Proposition A.1 gives a nondecreasing $F$ and $F([\ell, u]) \subseteq [F(\ell), F(u)] \subseteq [\ell + d_0, u - d_0]$. Conditional on $z_t$, the next number of incorrect winners is $\mathrm{Bin}(M, R_H(z_t))$. Hoeffding's inequality gives $\mathbb{P}\big(|z_{t+1} - F(z_t)| \ge d_0 \mid \mathcal{F}_t\big) \le 2\exp(-2Md_0^2/\beta^2)$. Apply this conditional bound before the first exit from $[\ell, u]$ and sum over $T$ possible exit steps. This proves (D.1). In particular, for any fixed $b < 2d_0^2/\beta^2$, the probability of staying for $\lfloor e^{bM} \rfloor$ rounds tends to one as $M \to \infty$.

For a common-$H$ system with $\beta m_H > 1$ and fixed initial $z_0 > 0$, the two orders of limits give

$$\lim_{M\to\infty}\lim_{t\to\infty} \mathbb{E}z_t^{(M)} = 0, \qquad \lim_{t\to\infty}\lim_{M\to\infty} \mathbb{E}z_t^{(M)} = z_* > 0.$$

The first identity follows from almost-sure extinction and bounded convergence. For the second, let $z_t$ be the population orbit. By conditional variance and (B.1),

$$\mathbb{E}|z_{t+1}^{(M)} - z_{t+1}| \le \frac{\beta}{2\sqrt{M}} + \beta N\,\mathbb{E}|z_t^{(M)} - z_t|.$$

Iterating this estimate shows that the error tends to zero for every fixed $t$. Applying Proposition A.1 to the population orbit then gives the second limit.

### D.2 THE STATIONARY DISTRIBUTION AFTER INCORRECT EXTINCTION

Assume that $q$ gives every correct answer positive probability and that their score laws are common. After extinction, selection satisfies $\mathcal{S}_N(p) = p$, so the recursion is

$$p_{t+1} = \alpha q + \frac{\beta}{M} X_t, \qquad X_t \mid p_t \sim \mathrm{Mult}(M, p_t).$$

After one step the state space is the finite set $\{\alpha q + \beta k/M : k_i \in \mathbb{Z}_{\geq 0},\ \sum_i k_i = M\}$. Every state has positive coordinates. Every count vector has positive conditional probability from every state. The transition matrix is therefore strictly positive, which gives a unique stationary law. In stationarity the expectation equation $\mathbb{E}p = \alpha q + \beta \mathbb{E}p$ yields $\mathbb{E}p = q$.

Let $V = \mathrm{Cov}(p)$ and $\Sigma_q = \mathrm{diag}(q) - qq^\top$. The conditional covariance formula and total covariance give

$$V = \frac{\beta^2}{M}\mathbb{E}[\mathrm{diag}(p) - pp^\top] + \beta^2 V = \frac{\beta^2}{M}(\Sigma_q - V) + \beta^2 V.$$

Solving this covariance identity gives

$$V = \frac{\beta^2}{M(1-\beta^2)+\beta^2}\big(\mathrm{diag}(q) - qq^\top\big).$$

With at least two correct answers, the covariance is nonzero at fixed $M$ and $\beta > 0$. These fluctuations arise from repeated winner sampling; weighted estimation from fresh real data concerns a different resource (He et al., 2026).

### D.3 OPTIMAL REFERENCE SCHEDULES UNDER A MIXING CAP

Return to the population model with a common correct score distribution. Let $m_* > 1$ be the maximum multiplier over initially present errors, and consider a deterministic reference schedule $0 \leq \alpha_t \leq a < 1$. Take $a > 0$, as the zero-cap case cannot recover. Define its cost as

$$\mathcal{A}(\alpha_\cdot) = \liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \alpha_t.$$

A schedule has exponential recovery rate at least $\kappa > 0$ if $\limsup_{t\to\infty} t^{-1}\log \mathrm{TV}(p_t, q) \leq -\kappa$. When recovery is feasible, its minimum average reference weight is

$$\inf \liminf_{T\to\infty} \frac{1}{T}\sum_{t<T} \alpha_t = \frac{a\log m_*}{-\log(1-a)}.$$

The theorem below also specifies the cost of a prescribed exponential rate. The cost averages the reference weights; it does not count independent human examples.

**Theorem D.1.** *Recovery is feasible if and only if $m_*(1-a) \leq 1$. For a prescribed $\kappa > 0$, recovery at rate at least $\kappa$ is feasible if and only if $0 < \kappa \leq -\log[m_*(1-a)]$. Whenever feasible, the respective minimum costs are*

$$\inf_{\alpha_\cdot:\, p_t \to q} \mathcal{A}(\alpha_\cdot) = \frac{a\log m_*}{-\log(1-a)},$$

$$\textit{minimum cost for rate at least } \kappa = \frac{a(\log m_* + \kappa)}{-\log(1-a)}.$$

*Both infima are attained by explicit schedules taking values only in $\{0, a\}$.*

*Proof.* We first bound the cost of any recovering schedule. Choose an initially present incorrect answer $j$ with multiplier $m_*$. Its coordinate remains positive at every finite time. Along any recovering schedule, $z_t \to 0$, so after some time $t_0$ one has $z_t < 1$. Taking logarithms in the coordinate bound and summing over rounds gives

$$\log p_{T,j} \geq \log p_{t_0,j} + (T - t_0)\log m_* + \sum_{t=t_0}^{T-1} \log(1-\alpha_t) + (N-1)\sum_{t=t_0}^{T-1} \log(1 - z_t).$$

The last sum is $o(T)$ by convergence of $z_t$ to zero. Since $\log p_{T,j} \le 0$, this implies

$$\liminf_{T\to\infty} \frac{1}{T} \sum_{t<T} -\log(1-\alpha_t) \ge \log m_*.$$

For recovery at rate $\kappa$, combining the inequality $p_{T,j} \le \mathrm{TV}(p_T, q)$ with the prescribed rate replaces the right side by $\log m_* + \kappa$. The line joining the endpoint values bounds the convex function from above, giving $-\log(1-u) \le [-\log(1-a)]u/a$ on $[0,a]$. Combining these bounds gives the stated cost lower bounds. Each logarithmic summand is also at most $-\log(1-a)$, giving both feasibility restrictions.

To attain these bounds, let $h_a = -\log(1-a)$, use $\kappa = 0$ for ordinary recovery, and set

$$f = \frac{\log m_* + \kappa}{h_a} \in (0,1], \qquad \alpha_t = a\big(\lfloor (t+1)f \rfloor - \lfloor tf \rfloor\big).$$

The number of nonzero reference rounds before time $T$ is exactly $\lfloor Tf \rfloor$, so its average cost converges to $af$. Define the cumulative rare-error multiplier $P_t = \prod_{r<t} m_*(1-\alpha_r) = m_*^t(1-a)^{\lfloor tf \rfloor}$. It satisfies $e^{-\kappa t} \le P_t \le e^{h_a} e^{-\kappa t}$.

Writing $k = (m_* - 1)/m_* > 0$, the global chord bound gives

$$z_{t+1} \le m_*(1-\alpha_t) z_t (1 - k z_t). \tag{D.2}$$

For $\kappa > 0$, dropping the final factor from the recurrence and iterating gives $z_t \le P_t z_0 \le e^{h_a} z_0 e^{-\kappa t}$. For $\kappa = 0$, let $x_t = z_t/P_t$ to remove the time-varying linear multiplier. Equation (D.2) yields $x_{t+1} \le x_t(1 - kP_t x_t)$. As $P_t \ge 1$, the reciprocal inequality gives $1/x_{t+1} \ge 1/x_t + k$, unless $z$ has already reached zero. Hence $z_t \le e^{h_a}/(z_0^{-1} + kt)$. After the first round using positive reference weight, the correct conditional distribution is defined. Its discrepancy from $q$ contracts by at most $\prod_{r<t}(1-\alpha_r) = P_t/m_*^t$, up to a fixed initial factor, by the time-dependent version of (A.1). This product decays at rate $\log m_* + \kappa$. Combining the resulting contraction with the incorrect-mass bounds proves ordinary recovery at $\kappa = 0$ and the requested total-variation exponent when $\kappa > 0$. □

# E Numerical Recursions and Reproducibility

Figures 2, 3, and 5 evaluate the stated population recursions deterministically. This appendix specifies the exact updates, parameters and numerical checks; the supplementary code archive contains the simulation programs.

## E.1 Scalar recovery and equal pairwise accuracy

For the band-mixture incorrect score law, (9) gives the exact update

$$z_{t+1} = (1-\alpha)\left\{1 - (1-\epsilon z_t)^N + [(1-\epsilon)z_t]^N\right\}.$$

We fix $\epsilon = 0.01$ and $z_0 = 0.25$. Figure 2 uses $\alpha = 0.1$ and $N \in \{80, 90, 100, 110, 128\}$ through $10^4$ rounds. The corresponding rare-error multipliers after reference mixing are $0.72$, $0.81$, $0.90$, $0.99$ and $1.152$. The curves are shown above $10^{-8}$; values below the plotted range are not reset to a positive floor. We evaluate $1 - (1-\epsilon z)^N$ as $-\mathrm{expm1}(N \log1p(-\epsilon z))$ to avoid cancellation near zero.

At these fixed $\epsilon, \alpha$, the formal critical candidate count is $1/[(1-\alpha)\epsilon] = 111.\overline{1}$, which is not an integer. The inset therefore uses $N = 128$ and $\alpha = 7/32$ so that $(1-\alpha)N\epsilon = 1$ exactly. Expanding this particular recurrence at zero gives

$$z_{t+1} = z_t - \frac{(N-1)\epsilon}{2} z_t^2 + O(z_t^3).$$

Since Theorem 1 implies $z_t \to 0$, the reciprocal increment tends to $(N-1)\epsilon/2$, and hence $tz_t \to 2/[(N-1)\epsilon]$. At $t = 10^5$, the computed $tz_t$ is $1.574652$, compared with the limit $1.574803$. The dotted line shows the instance-specific asymptote derived from this expansion.

For the second incorrect law, let $H = \mathrm{Unif}[a,b]$ with $a = 0.005$, $b = 0.015$, against $G = \mathrm{Unif}[0,1]$. Its pairwise error is $(a+b)/2 = 0.01$, exactly as for the band mixture. Integrating the winner formula

over $[a, b]$ yields

$$R_H(z) = z\,\frac{[(1-z)b+z]^N - [(1-z)a]^N}{(b-a)(1-z)+z}, \qquad m_H = \frac{b^N - a^N}{b-a}.$$

Figure 3 uses $N = 128$, $\alpha = 0.1$ and the same $z_0 = 0.25$. The first narrow-law update is about $2.112 \times 10^{-75}$; the next is about $6.585 \times 10^{-307}$, after which binary64 arithmetic underflows. The plotted zero is caused by binary64 underflow; the exact population orbit remains positive. The band-mixture orbit tends to the positive root $z_* \approx 0.2281791$ of $(1-\alpha)R_H(z) = z$.

### E.2 What the existing rates imply at finite horizons

Theorem 1 bounds incorrect mass at every integer time. For $0 < \varepsilon < z_0$, a sufficient number of rounds to reach $z_t \le \varepsilon$ is

$$t \ge \left\lceil \frac{\log(z_0/\varepsilon)}{-\log c} \right\rceil \quad (c = \beta \max\{1, m_*\} < 1), \qquad t \ge \left\lceil \frac{1/\varepsilon - 1/z_0}{\alpha} \right\rceil \quad (\beta m_* = 1).$$

These are sufficient round bounds and can overestimate the hitting time of the actual trajectory. Displayed bounds are rounded upward. For the band mixture with $\epsilon = 0.01$ and $z_0 = 0.25$, Table 2 gives the resulting upper bounds.

Table 2: Finite-round error bounds from Theorem 1 for the band-mixture example with $z_0 = 0.25$ and $\epsilon = 0.01$. Values are rounded upward.

| Upper bound on $z_t$ | $t = 10$ | $t = 50$ | $t = 100$ |
|---|---|---|---|
| $N = 100$, $\alpha = 0.1$: $0.25(0.9)^t$ | 0.08717 | 0.001289 | $6.65 \times 10^{-6}$ |
| $N = 128$, $\alpha = 7/32$: $0.25/(1 + 7t/128)$ | 0.1617 | 0.06695 | 0.03865 |

Incorrect mass and full distributional error are different quantities: the conditional correct proportions contract by at most $\beta$ per round under common scores (Appendix A). For a preserved all-correct target, Theorem 4 and Pinsker's inequality give the finite-round bound

$$\mathrm{TV}(p_t, q) \le \sqrt{\tfrac{1}{2}\rho^{t-1}\log(1/\alpha)}, \qquad t \ge 1, \quad \rho < 1.$$

The margins $1 - c$ and $1 - \rho$ determine the resulting round bounds.

### E.3 Regular-score rotation and numerical precision

For the fixed Beta-mixture family in Appendix B.7, the exact $N = 2$ recurrence is

$$p_{t+1,i} = \frac{\alpha}{3} + (1-\alpha)p_{t,i}\left[1 + \frac{\sqrt{3}\varepsilon_{\mathrm{amp}}^2}{70}(p_{t,i-1} - p_{t,i+1})\right],$$

with cyclic indices. We set $\varepsilon_{\mathrm{amp}} = 1/4$, $p_0 = (0.6, 0.25, 0.15)$ and run $4 \times 10^7$ rounds for $\alpha = 0.5\alpha_*$ and $1.5\alpha_*$, where $\alpha_* = 1 - (1 + \varepsilon_{\mathrm{amp}}^4/4900)^{-1/2}$. The threshold is evaluated using `log1p` and `expm1`. At each discrete update, the program stores two coordinates and obtains the third as $1 - p_1 - p_2$.

The simplex plots use Cartesian coordinates $(p_2 + p_3/2, \sqrt{3}p_3/2)$, with the vertices labeled by answer index. They show three separate 25,000-round windows beginning at 0, $2 \times 10^6$ and 39,975,000, sampled every 50 rounds. No lines connect omitted intervals. The distance plot is sampled every 2000 rounds, with all of the first 100 rounds retained. At the final round, the distances to $q$ are approximately $0.484616$ and $8.371 \times 10^{-5}$, respectively. The lower-weight nonconvergence is covered by the analytic result; the higher-weight trajectory illustrates convergence from this specified initial condition.

The threshold is approximately $3.986 \times 10^{-7}$, and this example tracks the rotational dynamics over tens of millions of iterations, far beyond a 10–100 round horizon.

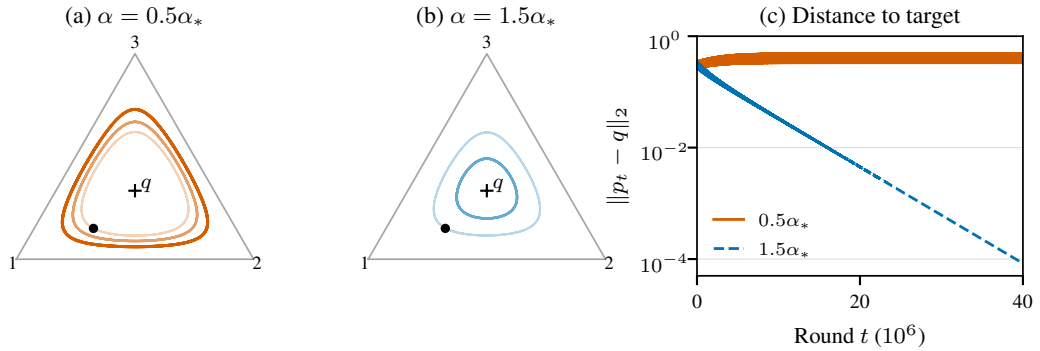



Figure 5: The quartic obstruction in Appendix B.7: fixed Beta-mixture dynamics at $N = 2$, $\varepsilon_{\text{amp}} = 1/4$, $q = (1/3, 1/3, 1/3)$, from $p_0 = (0.6, 0.25, 0.15)$ (dot). The local spectral threshold is $\alpha_* \approx 3.986 \times 10^{-7}$. Panels (a,b) show 25,000-round windows starting at 0, 2, and 39.975 million rounds (light to dark); (c) tracks distance over 40 million rounds. For the specified initialization, the above-threshold trajectory approaches the target over the plotted horizon. The proved regular-family obstruction has scale $\delta^4$.

All plotted iterates use binary64 arithmetic. Independent scalar quadrature agrees to relative error below $4 \times 10^{-13}$, and a 65-digit evaluation of the critical recursion agrees after 1000 rounds to absolute error below $2 \times 10^{-17}$. A rearranged three-coordinate rotation update, normalized at each step, changes the final coordinates by at most $6.7 \times 10^{-7}$ below the threshold and $1.9 \times 10^{-12}$ above it. These checks validate the numerical implementation and its rounding behavior; the universal convergence statements follow from the analytic arguments above.

### E.4 Code Availability

The JavaScript program `figures/simulate.js` in the supplementary code archive produces all plotted trajectories and declared sampling windows without external dependencies. The README gives running instructions and describes the output.

### E.5 Feasibility and Conservatism of the Joint Criterion

Take $N = 8$, two correct types with reference probabilities $(1/2, 1/2)$, and CDFs $G_{\pm}(u) = u \pm 2\delta u(1-u)$ on $[0, 1]$. Their CDF diameter is exactly $\delta$, their densities are positive for the values below, and their relative preservation defect is $\varepsilon_{\text{rel}} = 14\delta/9$. One incorrect type is uniform on $[2, 3]$ with probability $\epsilon$ and on $[-2, -1]$ otherwise. Thus $\overline{m} = 8\epsilon$ regardless of the correct proportions. The total incorrect mass follows (9), with exact extinction boundary $\alpha_{\text{err}} = \max\{0, 1 - (8\epsilon)^{-1}\}$.

For each $\delta \in \{0.005, 0.015, 0.03\}$ and $\epsilon \in \{0.10, 0.14\}$, bisection finds the zero of $\kappa_\alpha$. Let $\alpha_{\text{suff}}$ be the larger of that zero and $\alpha_{\text{err}}$. It is an infimum: a zero of $\kappa_\alpha$ does not satisfy the theorem's strict inequality. We evaluate at $\alpha = \alpha_{\text{suff}} + 0.01$. Table 3 compares fixed-point bias with the theorem's bound.

Table 3: Feasibility of Theorem 5 in the specified $N = 8$ family. Each tested weight exceeds the sufficient-weight infimum by 0.01; TV bias measures the fixed point's distance from $q$.

| $\delta$ | $\epsilon$ | $\varepsilon_{\text{rel}}$ | $\alpha_{\text{suff}}$ | Tested $\alpha$ | TV bias | Bound |
|---|---|---|---|---|---|---|
| 0.005 | 0.10 | 0.00778 | 0.01717 | 0.02717 | 0.12966 | 0.13922 |
| 0.015 | 0.10 | 0.02333 | 0.05382 | 0.06382 | 0.15403 | 0.17113 |
| 0.030 | 0.10 | 0.04667 | 0.12299 | 0.13299 | 0.13897 | 0.15211 |
| 0.005 | 0.14 | 0.00778 | 0.10714 | 0.11714 | 0.02920 | 0.02931 |
| 0.015 | 0.14 | 0.02333 | 0.10714 | 0.11714 | 0.08512 | 0.08793 |
| 0.030 | 0.14 | 0.04667 | 0.12299 | 0.13299 | 0.13897 | 0.15211 |

For each row, exact population iterations from $(0.45, 0.30, 0.25)$, $(0.01, 0.09, 0.90)$, $(1, 0, 0)$ and $(0, 1, 0)$ are within $1.1 \times 10^{-14}$ TV of the independently computed correct-face fixed point after 10,000 rounds. Correct winner probabilities are integrated by 16-point Gauss–Legendre quadrature; the integrands are polynomials of degree at most $2N - 1 = 15$, so the quadrature is exact in real arithmetic. The largest pre-normalization mass discrepancy is below $2.7 \times 10^{-15}$.

To test conservatism beyond the certified region, we also run the full three-type recursion for 100,000 rounds at $\alpha = 0.001$ when $\epsilon = 0.10$ and at $\alpha = 0.108$ when $\epsilon = 0.14$, computing the correct-face fixed point separately for each weight. All four initializations in every case end within $4.4 \times 10^{-15}$ TV of that point. For $\epsilon = 0.10$ all these weights fall below the certificates, but the fixed-point TV biases are $0.46869, 0.48920, 0.49448$. For $\epsilon = 0.14$, $\alpha = 0.108$ falls below the certificate only at $\delta = 0.03$. The below-certificate runs illustrate convergence to biased fixed points, with the reported biases measuring the remaining distance from $q$.

Separately, the scalar error recursion at $\epsilon = 0.14$ gives $z_{100000} = 3.2651 \times 10^{-4}$ at $\alpha = 0.107$ and $3.7644 \times 10^{-45}$ at $\alpha = 0.108$, bracketing the exact boundary $0.107142857$. Figure 6(c) plots this error-extinction boundary. Here $\varepsilon_{\text{rel}} = O(\delta)$, so the family explores first-order preservation error. The complete deterministic study is implemented in `figures/reviewer_checks.py`.

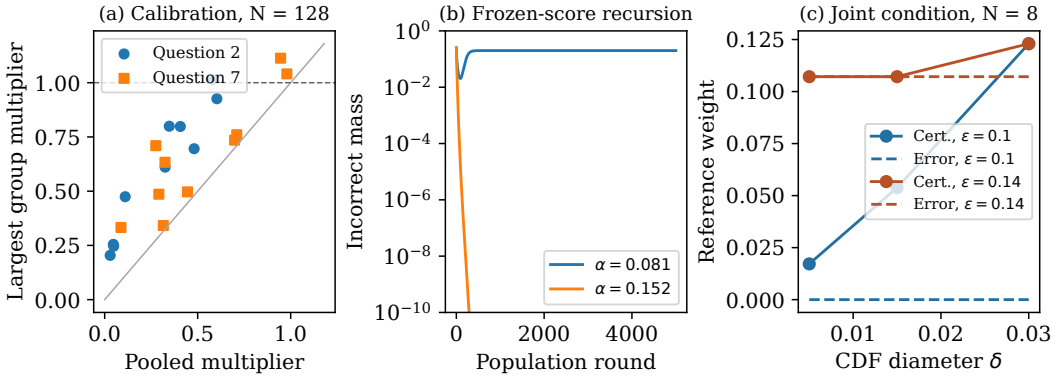



Figure 6: (a) Pooled versus largest answer-group moments in 20 calibration cases, illustrating the aggregation gap in Proposition H.1. (b) Theorem 1's transition at $\alpha_* = 0.10154$ for frozen group-score laws. (c) Theorem 5's sufficient boundary versus the exact error-extinction boundary in an $N = 8$ population family. Protocols and held-out results appear in Appendices F.5 and E.5.

# F A Diagnostic Using Public Verifier Scores

We evaluate the rare-error multiplier using public verifier scores from fixed generators. Within each question–generator setting, the conditional correct and incorrect score distributions are treated as fixed. The experiment evaluates the two-class score diagnostic used in Section 5.

## F.1 Data and separation of estimation from evaluation

We use all 20 comma-separated values (CSV) files containing scores and labels tracked in the public repository of Dorner et al. (2026), at commit

`aeb8eb114e3bc75fdf780dddb56b04f48222676f`.

The repository is https://github.com/socialfoundations/roc-n-reroll. Its examples cover GSM8K question indices 2 and 7, generators Qwen3-1.7B and Qwen3-4B, and five Qwen3 verifiers of sizes 1.7B, 4B, 8B, 14B and 32B. Each file contains 10,000 scores and supplied binary correctness labels. We use those labels without relabeling the responses. Generator metadata specify temperature 1, five examples in the prompt and seed 42; the stored scoring-verifier metadata also specify temperature 1 and seed 42. The scores include atoms, often at zero and one. We retain all scores, including the upstream parser's possible zero defaults.

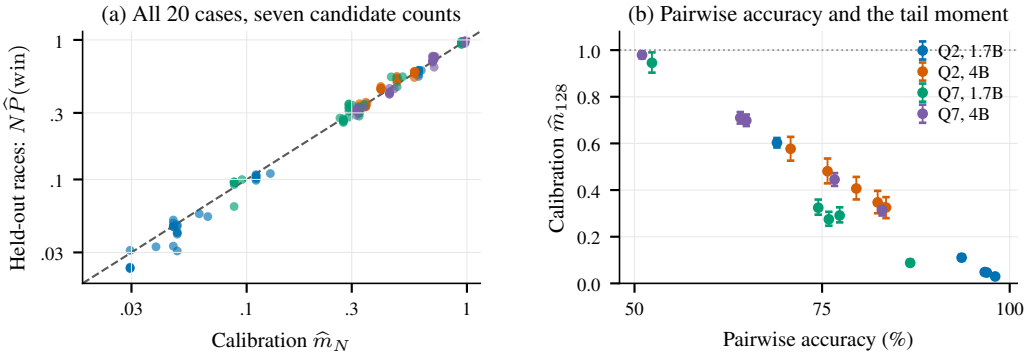



Figure 7: Checking the multiplier estimator (F.1) on two GSM8K questions, two generators and five verifiers. (a) Predictions versus 50,000 held-out competitions per point, $N = 2, 4, \ldots, 128$. (b) Pairwise accuracy and the $N = 128$ moment, with pointwise 95% bootstrap intervals. Colors identify question and generator size. Appendix F reports all cases and the protocol.

For each question–generator pair, a seeded random permutation assigns 5,000 response records to calibration and the remaining 5,000 to evaluation. All five verifiers share that split. Hashes of the complete numeric files are checked by the included program. Their binary label sequences agree across verifiers; for each question–generator pair, the 14B verifier CSV also matches the score and label arrays in the corresponding JavaScript Object Notation (JSON) response files. The numbers of correct records in the full pools are 4,531, 9,213, 586 and 7,240 for question 2/1.7B, question 2/4B, question 7/1.7B and question 7/4B, respectively. The small correct pool in the third case is reflected in its uncertainty intervals. Repeated generated texts remain distinct sampled records. We never pool score distributions across questions or generators.

Within one case, let $G$ and $H$ be the empirical conditional score distributions for correct and incorrect responses. This construction yields a pooled two-class multiplier for the fixed generator. The $m_*$ in Theorem 1 instead maximizes over initially present individual incorrect answer types; retraining may also change the within-class score mixture.

### F.2 Ties, the estimated moment and held-out competitions

Give every candidate an independent uniform priority to break score ties. For an incorrect score $s$, write $a_s = \mathbb{P}(S_{\text{correct}} < s)$ and $b_s = \mathbb{P}(S_{\text{correct}} = s)$. Conditioning on its priority $u$ gives winning probability $(a_s + ub_s)^{N-1}$ against $N - 1$ correct candidates. Thus the tie-aware multiplier is

$$m_N = \mathbb{E}_{s\sim H} r_N(a_s, b_s), \qquad r_N(a, b) = \begin{cases} \left[(a+b)^N - a^N\right]/b, & b > 0, \\ Na^{N-1}, & b = 0. \end{cases} \tag{F.1}$$

Replacing each score value by a disjoint uniform interval realizes these priorities with continuous score distributions and preserves the ordering of unequal scores. Consequently the moment bounds in (8) apply with tie-aware pairwise error $\mu = \mathbb{P}(S_{\text{bad}} > S_{\text{correct}}) + \frac{1}{2}\mathbb{P}(S_{\text{bad}} = S_{\text{correct}})$. In particular, $m_2 = 2\mu$. The formula is evaluated from calibration histograms to obtain $\widehat{m}_N$, for $N \in \{2, 4, 8, 16, 32, 64, 128\}$. The implementation uses `log1p` and `expm1` to avoid cancellation when a score atom is small.

Independently, for each case and $N$ we run 50,000 competitions. Each draws one incorrect and $N - 1$ correct records with replacement from the held-out pools and chooses uniformly among tied highest scores. The measured multiplier is $N$ times the incorrect winning frequency. These competitions contain no calibration records. Their conditional Monte Carlo uncertainty is computed by a Wilson interval for the winning probability, multiplied by $N$.

To measure response-sampling uncertainty, we also perform 2,000 multinomial bootstrap resamples of the correct and incorrect score histograms, separately in each split and conditional on its observed class counts. Figure 7(b) uses pointwise 95% percentile intervals from the calibration bootstrap.

The program additionally reports held-out bootstrap intervals and the exact moment of the held-out empirical score distributions, which removes Monte Carlo variation from the comparison of the two splits. The intervals quantify within-case sampling variability. Because the analysis covers two questions and reuses response pools across candidate counts and verifiers, we report them as case-level rather than benchmark-level uncertainty.

### F.3 RESULTS

Table 4 reports every case at $N = 128$. Figure 7(a) includes all 140 case–candidate-count combinations. The median absolute difference between the calibration and held-out empirical moments is $0.00903$, and the median relative difference is $3.95\%$; the largest absolute difference is $0.0893$. These split comparisons are descriptive for the four evaluated question–generator settings.

For a descriptive comparison, question 2 with a 4B generator and 4B verifier has calibration pairwise accuracy $75.72\%$ and $\widehat{m}_{128} = 0.4809$; question 7 with a 1.7B generator and 8B verifier has accuracy $75.90\%$ and $\widehat{m}_{128} = 0.2748$. Their held-out empirical moments are reported in the table. This pair illustrates the distinction between the two statistics; it was selected from the fully reported cases after analysis.

All 140 calibration point estimates are below one. At $\alpha = 0.1$, all their pointwise bootstrap intervals also lie below $1/(1-\alpha)$. By contrast, the plug-in pairwise bounds $N\widehat{\mu}^{N-1} \le m_N \le N\widehat{\mu}$ straddle that threshold in 94 of 140 cells, including all 20 cells at $N = 128$. The pairwise bounds quantify what remains unidentified from pairwise accuracy alone, whereas the bootstrap intervals quantify sampling uncertainty in the full-score estimate.

Table 4: All public-score cases at $N = 128$. Sizes are in billions of parameters. Accuracy and prediction use calibration records; brackets give pointwise 95% bootstrap intervals. The last columns use held-out records: the exact empirical moment and the multiplier measured in 50,000 competitions.

| Question | Generator | Verifier | Accuracy | $\widehat{m}_{128}$ [95%] | Held-out | Races |
|---|---|---|---|---|---|---|
| 2 | 1.7 | 1.7 | 68.98% | 0.603 [0.582, 0.623] | 0.591 | 0.548 |
| 2 | 1.7 | 4 | 93.59% | 0.110 [0.098, 0.123] | 0.103 | 0.105 |
| 2 | 1.7 | 8 | 96.67% | 0.048 [0.040, 0.057] | 0.042 | 0.031 |
| 2 | 1.7 | 14 | 96.95% | 0.046 [0.038, 0.054] | 0.047 | 0.051 |
| 2 | 1.7 | 32 | 98.06% | 0.030 [0.023, 0.037] | 0.023 | 0.023 |
| 2 | 4 | 1.7 | 70.80% | 0.577 [0.526, 0.628] | 0.586 | 0.576 |
| 2 | 4 | 4 | 75.72% | 0.481 [0.429, 0.535] | 0.525 | 0.497 |
| 2 | 4 | 8 | 79.56% | 0.407 [0.360, 0.456] | 0.454 | 0.440 |
| 2 | 4 | 14 | 83.52% | 0.325 [0.280, 0.370] | 0.323 | 0.300 |
| 2 | 4 | 32 | 82.40% | 0.347 [0.301, 0.397] | 0.345 | 0.340 |
| 7 | 1.7 | 1.7 | 52.32% | 0.946 [0.903, 0.990] | 0.953 | 0.973 |
| 7 | 1.7 | 4 | 74.49% | 0.324 [0.295, 0.359] | 0.319 | 0.284 |
| 7 | 1.7 | 8 | 75.90% | 0.275 [0.247, 0.307] | 0.262 | 0.256 |
| 7 | 1.7 | 14 | 77.34% | 0.291 [0.262, 0.326] | 0.320 | 0.346 |
| 7 | 1.7 | 32 | 86.74% | 0.088 [0.076, 0.101] | 0.091 | 0.064 |
| 7 | 4 | 1.7 | 51.00% | 0.979 [0.965, 0.994] | 0.970 | 0.957 |
| 7 | 4 | 4 | 64.90% | 0.698 [0.675, 0.724] | 0.709 | 0.765 |
| 7 | 4 | 8 | 64.09% | 0.710 [0.685, 0.735] | 0.751 | 0.637 |
| 7 | 4 | 14 | 76.66% | 0.445 [0.418, 0.473] | 0.421 | 0.451 |
| 7 | 4 | 32 | 82.99% | 0.315 [0.291, 0.342] | 0.308 | 0.310 |

### F.4 CODE AVAILABILITY AND NUMERICAL CHECKS

The supplementary code archive contains `figures/real_verifier.py`, which requires Python 3 and NumPy, together with running instructions. The program downloads the pinned numeric CSVs if absent, verifies their hashes and produces every estimate, interval, class count and held-out competition count.

An independent enumeration of all competitor score configurations agrees with (F.1) in 60 small-support cases to absolute error below $7 \times 10^{-15}$. Nine further checks, including a $10^{-9}$ score atom and $N = 128$, agree with 70-digit arithmetic to absolute error below $3 \times 10^{-13}$. These checks validate the estimator implementation.

### F.5 Answer-group diagnostics and frozen empirical recursion

The four pinned 14B-verifier JSON files retain the 10,000 response strings for each question–generator pair. Their scores and labels match the corresponding CSVs; all five verifier CSVs have the same label sequence. The pinned upstream exporter constructs these CSVs in response order. The supplementary program checks the four additional JSON hashes and every CSV hash.

We define a syntactic answer group as the first signed decimal token following the last `####` marker, with commas removed and decimal representations normalized. Responses lacking such a token receive an `unparsed` key. We retain the supplied correctness labels, even if token extraction and those labels appear inconsistent; grouping is not semantic relabeling. Within each 5,000-record calibration split, incorrect keys with at least 20 occurrences define separate groups; all remaining keys share an `other` group. This rule yields 23, 7, 18 and 6 error groups for question 2/1.7B, 2/4B, 7/1.7B and 7/4B. The groups are defined without verifier scores and without using evaluation records. All held-out incorrect records in these data map to a calibration-defined group.

For each group, compute (F.1) against the same pooled correct distribution, at all seven candidate counts. Its calibration-frequency-weighted mean equals the pooled multiplier to numerical precision. For each case and $N$, select the group with largest calibration moment, then estimate that fixed group's moment on the disjoint evaluation split. Pointwise 95% percentile intervals use 2,000 independent multinomial resamples of the held-out correct and selected-group score histograms. They condition on the observed group counts, and are neither simultaneous intervals across the 140 comparisons nor population-wide certificates.

At $N = 128$, the largest calibration group moment exceeds the pooled estimate by a factor as large as 6.90. Table 5 reports the three calibration maxima above one; all 20 selected held-out point estimates are below one, and all three corresponding intervals contain one. We therefore observe no replicated evidence of subtype amplification in these data.

Table 5: Held-out evaluation of the three groups selected by calibration maxima above one at $N = 128$. Intervals are pointwise 95% bootstrap intervals for the fixed selected groups.

| Question / generator / verifier | Group token | Pooled | Group | Held out | 95% interval |
|---|---|---|---|---|---|
| 2 / 4B / 1.7B | 195000 | 0.577 | 1.015 | 0.978 | $[0.915, 1.029]$ |
| 7 / 1.7B / 1.7B | 1.666 | 0.946 | 1.113 | 0.846 | $[0.657, 1.012]$ |
| 7 / 4B / 1.7B | 80 | 0.979 | 1.041 | 0.985 | $[0.884, 1.039]$ |

Supplementary code reports every group count and all 140 groupwise comparisons. Pooling correct responses is still a modeling choice, and finer error subtypes can remain hidden inside each group. These estimates therefore do not certify the original answer-level $m_*$.

For the recursion experiment, regard the pooled correct histogram as one answer type and each calibration error-group histogram as a separate type. Keep these laws fixed. Replace ordered score atoms by disjoint unit intervals, with uniform scores inside each interval; this realizes the tie-priority rule with continuous CDFs. If $h_{ik}$ is the probability of score bin $k$ for type $i$, write $f_k = \sum_i p_i h_{ik}$, $A_k = \sum_{\ell<k} f_\ell$ and $B_k = A_k + f_k$. The exact winner weight is

$$W_i(p) = \sum_k h_{ik} \begin{cases} (B_k^N - A_k^N)/f_k, & f_k > 0, \\ N A_k^{N-1}, & f_k = 0. \end{cases}$$

The implementation evaluates the difference stably with `log1p`/`expm1`. It starts with correct mass $0.75$ and distributes the remaining $0.25$ according to calibration error-group frequencies. The reference is concentrated on the correct type. Every initially represented group retains its own fixed score law, so Theorem 1 applies exactly to this constructed finite-support population model.

We run all 20 cases for 5,000 rounds. If $m_{\max} > 1$, let $\alpha_* = 1 - 1/m_{\max}$ and use $\max\{0.001, 0.8\alpha_*\}$ and $\alpha_* + 0.05$; otherwise use $0.01$ and $0.1$. Figure 6(b) displays the case with the largest calibration

maximum, question 7 / generator 1.7B / verifier 1.7B, selected after analysis. Its $\alpha_* = 0.101539$ separates the two runs: at $\alpha = 0.081231$, the final incorrect mass is $0.2000$; at $\alpha = 0.151539$, it is $7.40 \times 10^{-128}$. The maximum probability-sum correction before numerical renormalization is below $5.7 \times 10^{-14}$. No probability floor is introduced; trajectories leave the plot below $10^{-10}$. The transition follows from treating each group histogram as a fixed score law throughout the recursion.

# G Comparison with Heterogeneous Curation Theory

Table 6 compares the closest finite-pool mixed result in version 3 of Zhao et al. (2025), identifying their candidate count $K$ with our $N$. Their paper also treats pure-synthetic and infinite-pool regimes, including convergence beyond TV contraction.

Table 6: Theorem-level comparison for finite-pool retraining with reference mixing.

| Object | Zhao et al. (2025), Regime (iii) | This paper |
|---|---|---|
| State space | Densities on a measure space; finite entropy in Zhao et al. (2025, Theorem 4.6) | A finite probability simplex; categorical probabilities are free parameters |
| Selection | Plackett–Luce choice from utilities with independent preference noise | Maximum of independent continuous scalar scores, allowing answer-dependent shapes |
| Noise in the convergence result | Setting S2: the marginal noise law is independent of the candidate outcome | Common correct-score laws in Theorem 1; heterogeneous laws in Theorems 2 and 5 |
| Limit | Unique fixed point $p_*$ under Zhao et al. (2025, Theorem 4.6); not generally the reference | Exact recovery criterion under common scores; preservation identifies $q$ as the fixed point; small preservation errors allow a limit different from the reference |
| Sufficient mixing | $N(1-\alpha) < 1$ gives TV contraction and geometric convergence | Score-sensitive exact rare-error boundary; fixed-$N$ quadratic mixing for preserved targets |
| Meaning of stability | Uniform-in-time sensitivity to bounded reward perturbations (Zhao et al., 2025, Theorem 5.3) | Convergence of a fixed update; almost every initialization can fail to converge to the unique fixed point |

The selection models overlap in one important case. Conditional on noisy utilities $r(y)+\varepsilon_y$, Plackett–Luce selection equals maximization of $r(y) + \varepsilon_y + Z_y$, where the $Z_y$ are independent standard Gumbel variables. Under S2 and independent noise for each sampled candidate, $\varepsilon_y + Z_y$ has a common law, with an everywhere-positive continuous density obtained by convolution. On a finite all-correct support, Lemma B.3 therefore implies that preservation of a full-support target forces all $r(y)$ to coincide. Our preserved heterogeneous examples use differences in score shape beyond this common-location case. The common-correct-score regime allows distinct incorrect-score laws and addresses the separate removal of incorrect mass.

Zhao et al. (2025, Remark 4.9) leaves finite-pool mixed convergence below their TV-contraction threshold open. Our results give score-dependent guarantees in the regimes analyzed here. The two frameworks differ both in their selection models and in the stability questions addressed.

# H Assumptions, Pooling and Perturbed Retraining

## H.1 Which assumptions support which conclusions

The object analyzed here is the recursive selection operator in (2). The assumptions have distinct roles.

- *Finite fixed answer support.* Coordinates, support invariance and the maximum over initially present error types in Theorem 1 refer to this fixed set. Compactness and finite-dimensional arguments also enter fixed-point existence and the measure-zero basin construction.
- *Free simplex parameters and exact population fitting.* Together these identify cross-entropy minimization with the arithmetic mixture in (2). They are not needed to define selection, but all exact-recursion conclusions use that update. Proposition H.2 propagates a given bound on the fitting discrepancy through the recursion.
- *Fixed answer-conditioned score laws.* These make the main map autonomous and its fixed-point and nonconvergence claims meaningful. Proposition H.2 removes time homogeneity in the common-correct-score case under a uniform margin.

- *Candidate-level independence.* Independent label draws and conditionally independent score draws yield the power $F_p^{N-1}$ in (1). Correlated scoring, including a shared random rater, requires a corresponding joint winner law.
- *Continuous scores.* Continuity excludes ties. The ordered-bin construction in Appendix F.2 implements independent uniform tie priorities by continuous scores. Other tie-breaking mechanisms require their own winner law.

The preservation and common-correct-law assumptions are substantive as well. A multiplier condition controls errors; it does not by itself preserve the desired proportions among heterogeneous correct answers. Fixed finite corpora are a different approximation: Appendix D proves almost-sure extinction but potentially exponentially long persistence, rather than identifying a finite-sample path with the population recursion.

### H.2 What a pooled multiplier identifies

Let $G$ be the common correct-score CDF, and let $H_j$ be a finite collection of incorrect subtype CDFs with positive mixture weights $\pi_j$. Write $H_{\mathrm{pool}} = \sum_j \pi_j H_j$, $m_{\mathrm{pool}} = N \int G^{N-1}\,\mathrm{d}H_{\mathrm{pool}}$, and $m_{\max} = \max_j N \int G^{N-1}\,\mathrm{d}H_j$. This collection must cover the error types to which a proposed certificate applies.

**Proposition H.1** (Pooling and subtype control)**.** *Put* $\pi_{\min} = \min_j \pi_j$ *and* $r_H = \max_j \|H_j - H_{\mathrm{pool}}\|_\infty$. *Then*

$$m_{\mathrm{pool}} = \sum_j \pi_j m_j, \qquad m_{\mathrm{pool}} \le m_{\max} \le \min\{N, m_{\mathrm{pool}}/\pi_{\min}, m_{\mathrm{pool}} + N r_H\}. \tag{H.1}$$

*If the subtype CDF diameter is* $D_H$, *then* $r_H \le (1 - \pi_{\min}) D_H$.

*Proof.* Linearity gives the identity. Nonnegativity gives $m_{\mathrm{pool}} \ge \pi_j m_j$, and every $m_j \le N$. The continuous increasing function $G^{N-1}$ has total variation one. Integration by parts, whose boundary term vanishes, therefore gives

$$|m_j - m_{\mathrm{pool}}| = N \left| \int (H_j - H_{\mathrm{pool}})\,\mathrm{d}(G^{N-1}) \right| \le N \|H_j - H_{\mathrm{pool}}\|_\infty.$$

Finally $H_j - H_{\mathrm{pool}} = \sum_{k \ne j} \pi_k (H_j - H_k)$, proving the diameter bound. □

A small pooled moment alone cannot exclude a dangerous rare subtype. Take $G = \mathrm{Unif}[0,1]$, $H_0 = \mathrm{Unif}[-2,-1]$ and $H_1 = \mathrm{Unif}[2,3]$. If the latter has mixture weight $\pi$, then $m_0 = 0$, $m_1 = N$ and $m_{\mathrm{pool}} = \pi N$. Whenever $\beta N > 1$, choosing $0 < \pi < 1/(\beta N)$ makes the pooled condition subcritical while every initialization containing subtype 1 fails to recover by Theorem 1. For instance, $N = 4$, $\alpha = 1/4$ and $\pi = 1/10$ give $m_{\mathrm{pool}} = 0.4$ but $m_{\max} = 4$.

Thus a pooling-based certificate needs a bound on every relevant subtype's mass or score-law deviation. Empirical cluster masses do not bound the masses of finer subtypes, and finite-sample point estimates do not supply deterministic envelopes. The answer groups in Appendix F.5 are a diagnostic of this aggregation issue; they do not exhaust individual answer types or justify common correct-score laws.

### H.3 Changing score laws and inexact fitting

At each round $t$, let all correct answers share a continuous CDF $G_t$, and let error $j$ have CDF $H_{j,t}$. The laws can depend on the current state and the past, but are chosen before the independent scores in that round are drawn. Keep the finite label space, $q$, $N$ and $\alpha$ fixed. Define $m_{j,t} = N \int G_t^{N-1}\,\mathrm{d}H_{j,t}$ and let $T_t$ be the corresponding population update. Slow drift is not needed for the following uniform-margin result.

**Proposition H.2** (Uniform-margin recovery with drift and fitting error)**.** *Let* $J$ *contain every incorrect label that can occur along the trajectory, with the maximum over an empty* $J$ *defined as zero. Suppose,*

*for some $\eta > 0$, that $\beta \max_{j \in J} m_{j,t} \le 1 - \eta$ for every $t$, and put $c = \max\{\beta, 1 - \eta\} < 1$. If a simplex-valued sequence satisfies $\mathrm{TV}(p_{t+1}, T_t(p_t)) \le e_t$, then*

$$\mathrm{TV}(p_t, q) \le c^t \,\mathrm{TV}(p_0, q) + \sum_{s=0}^{t-1} c^{t-1-s} e_s. \tag{H.2}$$

*In particular, $e_t \to 0$ implies recovery; $\limsup_t e_t \le \bar{e}$ implies $\limsup_t \mathrm{TV}(p_t, q) \le \min\{1, \bar{e}/(1-c)\}$.*

*Proof.* We first prove a one-round bound toward $q$. For $0 < z = p(\mathcal{B}) < 1$, write $u = p(\cdot \mid \mathcal{C})$ and

$$D(z, u) = \mathrm{TV}(p, q) = \sum_{i \in \mathcal{C}} [q_i - (1 - z)u_i]_+,$$

where $[a]_+ = \max\{a, 0\}$. As a function of $z$, $D$ is nondecreasing, 1-Lipschitz and at least $z$. Common correct scoring preserves $u$ during selection. If $R$ is the selected incorrect mass, the chord bound gives $R/z \le \max\{1, \max_j m_{j,t}\}$. Hence

$$\mathrm{TV}(\mathcal{S}_{N,t}(p), q) = D(R, u) \le D(z, u) + (R - z)_+ \le \max\{1, R/z\} D(z, u).$$

Mixing with $q$ multiplies this distance by $\beta$, proving $\mathrm{TV}(T_t(p), q) \le c\,\mathrm{TV}(p, q)$. For $z = 0$, selection is the identity; for $z = 1$, both preselection and selected distributions have distance one from $q$. Thus the bound includes the endpoints. It is contraction toward this target, not a pairwise contraction between arbitrary states.

The triangle inequality now gives $D_{t+1} \le cD_t + e_t$, whose iteration is (H.2). Splitting the convolution into a finite initial segment and a uniformly bounded tail proves both limiting statements. □

For exact updates, $J$ can be the initially present errors. Fitting error can introduce new labels, so generally $J = \mathcal{B}$ is needed. Without a strict margin, exact updates still recover if $\beta \max_j m_{j,t} \le 1$ uniformly: the roundwise chord bound gives $z_{t+1} \le z_t - \alpha z_t^2$, and the common-score identity contracts conditional correct proportions by at most $\beta$. This gives a sufficient recovery condition for time-varying laws.

A uniform CDF perturbation bound can certify the margin. Relative to baseline laws $G, H_j$, let $a_t = \|G_t - G\|_\infty$ and $b_t = \max_j \|H_{j,t} - H_j\|_\infty$. The power-function Lipschitz bound and the same integration by parts as in Proposition H.1 imply

$$\max_j |m_{j,t} - m_j| \le N\{(N-1)a_t + b_t\}. \tag{H.3}$$

Consequently a baseline margin $\eta_0 = 1 - \beta \max_j m_j > 0$ survives whenever the right-hand side, multiplied by $\beta$, is uniformly smaller than $\eta_0$. Small successive increments alone do not control accumulated drift.

The fitting-error floor cannot generally be removed. With one correct and one incorrect answer having identical score laws, $\mathcal{S}_N(p) = p$. For $q = (1, 0)$ and $0 < e < \alpha$, the valid update $p_{t+1} = \alpha q + \beta p_t + e(-1, 1)$ has fitting discrepancy $e$ and limiting incorrect mass $e/\alpha$. Thus bounded fitting error does not imply exact recovery, even in this simplest common-score setting.

# I Recovery on a Fixed Score Scale

The CDF-distance-only class in Theorem 2 allows increasingly rapid changes in densities as $\delta$ decreases. We now impose a fixed score interval and a uniform density-derivative bound. These restrictions change the worst-case scale.

**Proposition I.1** (A cubic sufficient weight for uniformly regular densities)**.** *Suppose all answers are correct, $q_i > 0$, and $\mathcal{S}_N(q) = q$. Let the score laws have globally $C^1$ densities $f_i$, supported in a common interval of length $B$, with $\max_i \|f_i'\|_\infty \le L$. If their CDF diameter is at most $\delta$, then* (12) *holds with*

$$h_{\mathrm{reg}} = 2N(N-1)B\sqrt{L}\,\delta^{3/2} + \frac{N(N-1)(N-2)}{2}\delta^2. \tag{I.1}$$

*If $0 < h_{\mathrm{reg}} \le 1/2$, every $\alpha \ge h_{\mathrm{reg}}^2$ gives global recovery. Thus a weight $O_N(B^2 L\delta^3 + \delta^4)$ suffices, uniformly in the number of answers and in the positive target probabilities. At $\delta = 0$, every $\alpha > 0$ gives recovery.*

*Proof.* Let $f = \sum_i q_i f_i$ be the density of the reference score CDF $F_q$. Write $h_i = G_i - F_q$ for each answer's CDF deviation and $\eta_i = f_i - f$ for its density deviation. Then $\|h_i\|_\infty \le \delta$ and $\|h_i''\|_\infty \le 2L$. Taylor's theorem at $x+t$ and $x-t$ gives

$$2t|\eta_i(x)| \le |h_i(x+t) - h_i(x-t)| + 2Lt^2 \le 2\delta + 2Lt^2.$$

Choosing $t = \sqrt{\delta/L}$ yields $\|\eta_i\|_\infty \le 2\sqrt{L\delta}$ and $\|\eta_i\|_1 \le 2B\sqrt{L\delta}$. The nonempty compactly supported density class has $L > 0$; $\delta = 0$ is immediate. Global $C^1$ regularity ensures that this calculation also applies across support endpoints.

Preservation and integration by parts imply

$$0 = \int F_q^{N-1}\,\mathrm{d}h_i = -(N-1)\int F_q^{N-2} h_i\,\mathrm{d}F_q.$$

For $d_{\mathrm{TV}} = \mathrm{TV}(p,q)$ and $D = F_p - F_q$, the mixture-diameter bound gives $\|D\|_\infty \le \delta d_{\mathrm{TV}}$. Linearity also gives $\int F_q^{N-2} D\,\mathrm{d}F_q = 0$. Split the weight difference as

$$\begin{aligned} W_i(p) - 1 &= N\int (F_p^{N-1} - F_q^{N-1})\,\mathrm{d}F_q \\ &\quad + N\int (F_p^{N-1} - F_q^{N-1})\eta_i(x)\,\mathrm{d}x. \end{aligned}$$

In the first integral the linear Taylor term vanishes. The remainder is at most $N(N-1)(N-2)\delta^2 d_{\mathrm{TV}}^2/2$, with zero remainder at $N = 2$. The second integral has absolute value at most

$$N(N-1)\|D\|_\infty\|\eta_i\|_1 \le 2N(N-1)B\sqrt{L}\,\delta^{3/2} d_{\mathrm{TV}}.$$

Since $d_{\mathrm{TV}}^2 \le d_{\mathrm{TV}}$, this proves (12) with (I.1). For $0 < h_{\mathrm{reg}} \le 1/2$, the KL coefficient satisfies $C_{h_{\mathrm{reg}}} \le h_{\mathrm{reg}}^2$, so $\beta(1 + C_{h_{\mathrm{reg}}}) \le (1 - h_{\mathrm{reg}}^2)(1 + h_{\mathrm{reg}}^2) < 1$ whenever $\alpha \ge h_{\mathrm{reg}}^2$ and $\alpha < 1$. Apply Theorem 4; $\alpha = 1$ is immediate. □

### I.1 Three or more candidates: arbitrarily small positive mixing

For $N \ge 3$, the cubic upper bound can be strengthened. The curvature of the power in the winner formula supplies a nonnegative term that controls the second-order KL remainder when the densities are close.

**Proposition I.2** (Selection dissipation for $N \ge 3$)**.** *Under the assumptions of Proposition I.1, put $f = \sum_i q_i f_i$ and*

$$\chi_q = \int \frac{\sum_i q_i (f_i - f)^2}{f}\,\mathrm{d}x, \qquad C_N = \frac{N(N-1)(N-2)}{2},$$

*with integrand zero where $f = 0$. If $N \ge 3$, $N\delta < 1$ and*

$$C_N\delta^2 + N(N-1)^2\chi_q \le 2(1 - N\delta)^2, \tag{I.2}$$

*then selection does not increase $D_{\mathrm{KL}}(q\|p)$ on the interior. Every $0 < \alpha < 1$ gives global recovery, with*

$$D_{\mathrm{KL}}(q\|p_t) \le \beta^{t-1}\log(1/\alpha), \qquad t \ge 1.$$

*Moreover $\chi_q \le 2B\sqrt{L\delta}$. Therefore, for every fixed $N \ge 3, B, L$, every positive reference weight suffices uniformly over this class once $\delta$ is sufficiently small. The threshold infimum is zero; recovery at $\alpha = 0$ is not asserted.*

*Proof.* Full support of $q$ implies that all $f_i$ vanish where $f = 0$. The interpolation argument above, applied to each $G_i - G_j$, gives $\max_{i,j}\|f_i - f_j\|_\infty \le 2\sqrt{L\delta}$. At each score,

$$\sum_i q_i (f_i - f)^2 = \sum_i q_i f_i^2 - f^2 \le f(\max_i f_i - \min_i f_i).$$

Divide by $f$ where it is positive and integrate over the interval of length $B$ to obtain the bound on $\chi_q$.

Write $k = N - 1 \geq 2$. Let $D = F_p - F_q$ be the difference between the current and reference score CDFs, and let $M = \max\{F_p, F_q\}$ be their pointwise maximum. Preservation cancels the linear CDF term, giving $\int F_q^{k-1} D \, \mathrm{d}F_q = 0$. The average selection-weight excess is

$$T := q \cdot W(p) - 1 = N \int \{F_p^k - F_q^k - k F_q^{k-1} D\} \, \mathrm{d}F_q.$$

For $a, b \geq 0$, the exact identity

$$a^k - b^k - k b^{k-1}(a - b) = (a - b)^2 \sum_{j=0}^{k-2} (j + 1) a^{k-2-j} b^j$$

shows that the sum is at least $\max\{a, b\}^{k-2}$: use its first term when $a \geq b$ and its last term otherwise. The weighted squared CDF difference $E = \int D^2 M^{k-2} \, \mathrm{d}F_q$ therefore satisfies $T \geq NE \geq 0$. Taylor's upper remainder bound also gives $T \leq C_N \delta^2$.

To bound fluctuations of the selection weights, write $\eta_i = f_i - f$, $\xi_i = W_i(p) - 1$ and $V = F_p^k - F_q^k$. Preservation gives the decomposition $\xi_i = T + N \int V \eta_i \, \mathrm{d}x$, where the second term has zero $q$-weighted mean. Since $|V| \leq k M^{k-1} |D|$, Cauchy–Schwarz yields

$$\left| \int V \eta_i \, \mathrm{d}x \right|^2 \leq k^2 \left( \int D^2 M^{k-2} f \, \mathrm{d}x \right) \left( \int M^k \frac{\eta_i^2}{f} \, \mathrm{d}x \right) \leq k^2 E \int \frac{\eta_i^2}{f} \, \mathrm{d}x.$$

The zero-density convention is legitimate because $\eta_i = 0$ there; all other factors are nonnegative. Averaging over $i$ therefore gives

$$\sum_i q_i \xi_i^2 \leq T^2 + N^2 k^2 \chi_q E \leq (C_N \delta^2 + N k^2 \chi_q) T.$$

Lemma B.1 gives $|\xi_i| \leq N\delta < 1$. The logarithmic Taylor bound now proves

$$\begin{aligned} D_{\mathrm{KL}}(q \| \mathcal{S}_N(p)) - D_{\mathrm{KL}}(q \| p) &\leq -T + \frac{\sum_i q_i \xi_i^2}{2(1 - N\delta)^2} \\ &\leq -T \left[ 1 - \frac{C_N \delta^2 + N(N-1)^2 \chi_q}{2(1 - N\delta)^2} \right] \leq 0. \end{aligned}$$

Mixing with $q$ contracts this KL divergence by at most $\beta$. Boundary initializations satisfy $p_1 \geq \alpha q$, giving the stated rate. The sufficient condition follows uniformly for small $\delta$ from $\chi_q \leq 2B\sqrt{L\delta}$, since its left side tends to zero and its right side to two. The case $\alpha = 1$ recovers in one step. □

### I.2 A matching cubic obstruction at two candidates

The following construction shows that the cubic upper bound is sharp on a fixed interval under a fixed derivative bound. Its oscillation frequency still varies, unlike the fixed finite-dimensional Beta-mixture family in Appendix B.7.

**Proposition I.3** (Cubic sharpness at $N = 2$)**.** *Restrict the definition of $\alpha_{\mathrm{crit}}(2, \delta)$ to globally $C^1$ densities supported in $[0, 1]$ with $\|f_i'\|_\infty \leq 23$. The resulting worst-case threshold is $\Theta(\delta^3)$. The lower bound uses three answers and a uniform preserved target, with every other initialization failing to converge at a reference weight of order $\delta^3$.*

*Proof.* On $[0, 1]$ define $P(x) = 10x^3 - 15x^4 + 6x^5$ and $b(x) = P'(x) = 30x^2(1 - x)^2$. Extend $P$ by 0 and 1 outside the interval, and extend $b$ by zero. The density $b$ is globally $C^1$ and satisfies $\|b\|_\infty = 15/8$, $\|b'\|_\infty = 10\sqrt{3}/3$, and $b = b' = 0$ at both endpoints. For $a = 1/(16\pi)$, integer $K \geq 1$ and $\theta_i = 2\pi i/3$, put

$$\begin{aligned} v_K(u) = \sin(2\pi K u), \qquad w_K(u) = \cos(2\pi K u) - \cos(4\pi K u), \\ \phi_{i,K}(u) = \cos\theta_i \, v_K(u) + \sin\theta_i \, w_K(u), \qquad G_{i,K}(x) = P(x) + \frac{a}{K^2} \phi_{i,K}(P(x)). \end{aligned} \tag{I.3}$$

The perturbations have zero endpoint values, zero integrals, and sum to zero over $i$. Their derivatives obey

$$\|\phi_{i,K}'\|_\infty \leq 2\pi K \sqrt{10}, \qquad \|\phi_{i,K}''\|_\infty \leq 4\pi^2 K^2 \sqrt{26}.$$

Hence the derivative of $u + a\phi_{i,K}(u)/K^2$ is at least $1 - \sqrt{10}/(8K) > 0$. The actual score density is $f_{i,K}(x) = b(x)[1 + a\phi'_{i,K}(P(x))/K^2]$, with

$$\|f'_{i,K}\|_\infty \le \frac{10\sqrt{3}}{3}\left(1 + \frac{\sqrt{10}}{8}\right) + \frac{225\pi\sqrt{26}}{256} < 23. \tag{I.4}$$

The density and its derivative vanish at the endpoints, so the stated global regularity holds.

Work in the common coordinate $u = P(x)$, which preserves comparisons. The mixture CDF at $q = (1/3, 1/3, 1/3)$ is $u$. The zero integrals give $W_i(q) = 2\int_0^1 u\,\mathrm{d}(u + a\phi_{i,K}/K^2) = 1$. Orthogonality gives $\int_0^1 v_K w'_K\,\mathrm{d}u = -\pi K$, and thus

$$\int_0^1 \phi'_{i,K}\phi_{j,K}\,\mathrm{d}u = \pi K \sin(\theta_j - \theta_i).$$

Expanding the winner weights yields the exact recurrence weights

$$W_i(p) = 1 + \gamma_K(p_{i+1} - p_{i-1}), \qquad \gamma_K = \frac{\sqrt{3}\pi a^2}{K^3}, \tag{I.5}$$

with cyclic indices. The tangent eigenvalues at $q$ are $\beta(1 \pm \mathrm{i}\pi a^2/K^3)$, so $q$ is a strict local repeller when

$$0 < \alpha < \alpha_{*,K} := 1 - (1 + \pi^2 a^4/K^6)^{-1/2}. \tag{I.6}$$

The CDF diameter is exactly $\delta_K = 2\sqrt{3}a/K^2$. To verify this, write $c = \cos(2\pi K u)$. Then

$$v_K^2 + w_K^2 - 4 = 2(c+1)\{2c(c-1)^2 - 1\} \le 0.$$

For $c \le 0$ the bracket is negative; for $c \ge 0$, $2c(1-c)^2 \le 8/27$. Differences of phase vectors have norm $\sqrt{3}$, proving the upper diameter bound. The pair $i = 1, 2$ attains it at $u = 1/(2K)$. Consequently

$$\alpha_{*,K} \sim \frac{\pi^2 a^4}{2K^6} = \frac{\pi}{768\sqrt{3}}\delta_K^3.$$

This local instability implies failure of convergence from every $p_0 \ne q$. Equation (I.5) gives positive weights with $\sum_i W_i(p) = 3$. Therefore $\prod_i \mathcal{S}_N(p)_i \le \prod_i p_i \le 1/27$. If $T_\alpha(p) = q$, then $\mathcal{S}_N(p) = q$, so equality forces $p = q$. The tangent derivative at $q$ is a rotation followed by strict expansion; continuity gives a neighborhood on which distance from $q$ expands. Any orbit approaching $q$ must eventually equal it, and the preimage argument forces $p_0 = q$. Proposition 3 excludes convergence to any other point.

For an arbitrary sufficiently small allowed diameter $\delta$, take $K = \lceil\sqrt{2\sqrt{3}a/\delta}\rceil$. Then $\delta_K \le \delta$ and $K \le 2\sqrt{2\sqrt{3}a/\delta}$. The choice $\alpha = \alpha_{*,K}/2$ is bounded below by a positive constant times $\delta^3$. Proposition I.1 with $B = 1, L = 23$ supplies the matching upper bound. □

The distinction from Proposition I.2 is specific to $N = 2$: the dissipative Taylor remainder used there vanishes when the winner weight is linear in the mixture CDF. The variable-frequency construction attains the cubic class threshold; the fixed low-degree Beta-mixture construction in Appendix B.7 gives a quartic obstruction.

### I.3 Why score normalization matters

The fixed interval in Proposition I.1 is substantive. Replacing every score $X_i$ by $sX_i$ leaves comparisons, CDF uniform distances and all dynamics unchanged, while its density derivative becomes $s^{-2}f'_i(x/s)$. A bound on the absolute density derivative alone therefore does not fix the class's intrinsic score scale; the product $B^2L$ is invariant under this dilation.

For an explicit illustration at $N = 2$, replace $a/K^2$ in (I.3) by $a/K$. Positivity and preservation still hold, while the diameter is $2\sqrt{3}a/K$ and the rotational tangent component is $\pi a^2/K$. This gives a quadratic obstruction. Its density derivative is $O(K)$; a common dilation of order $\sqrt{K}$ makes that derivative uniformly bounded, on intervals whose length increases. Thus the literal derivative bound on $\mathbb{R}$ permits this quadratic example, whereas a fixed interval yields the cubic result above.

### I.4 SYMMETRIC LOCATION-SCALE FAMILIES

Let scores have the form $X_i = \mu_i + \sigma_i E_i$, with $\sigma_i > 0$ and independent noises sharing an everywhere-positive continuous density symmetric about zero. Gaussian and logistic scores are examples. At $N = 2$, target preservation holds if and only if all $\mu_i$ coincide; the scales may differ. Indeed $\sigma_i E_i - \sigma_j E_j$ is continuous and symmetric, so $\mathbb{P}(X_i > X_j) > 1/2$ exactly when $\mu_i > \mu_j$. If some locations differ, any maximal-location type has $W_i(q) > 1$, contradicting preservation at positive $q$. Conversely, common locations give every pairwise win probability $1/2$, so $W_i(p) = 1$ and $\mathcal{S}_N(p) = p$ for every $p$.

For $N \geq 3$, if the locations already coincide at $\mu$, preservation forces all scales to coincide too. The mixture $F_q$ is symmetric and strictly increasing. With $R = |E| > 0$ and $t_i = F_q(\mu + \sigma_i R) > 1/2$, symmetry gives

$$W_i(q) = \frac{N}{2}\mathbb{E}\{t_i^{N-1} + (1 - t_i)^{N-1}\}.$$

The function inside braces is strictly increasing on $(1/2, 1)$ for $N \geq 3$, hence $W_i(q) > W_j(q)$ whenever $\sigma_i > \sigma_j$. They cannot all equal one unless the scales agree. Thus the location-scale conclusions cover arbitrary locations at $N = 2$ and common locations at $N \geq 3$.