When Familiarity Is Not Enough: Geometry and Regret in Dual-Process Memory Routing

Anonymous authors
Paper under double-blind review

Abstract

Searching again can recover a missing memory, but a router must decide whether to incur the cost before seeing the results. RF-Mem makes this decision from initial similarity summaries, even though Recollection changes the query direction. We show that this mismatch creates an information barrier: corpora with identical complete indexed scores can require opposite optimal actions, with exact Bayes and minimax regret gaps. The separation persists under published Bag semantics across explicit single-beam ranges; under a stronger budget condition, it survives perturbations that remove duplicate embeddings. We trace the missing information to memory–memory geometry and ask which geometric observations can improve the decision before Recollection begins. In a constructed family, even the initial Top-K Gram block is uninformative, yet seed–candidate inner products yield an exact acquisition-cost versus regret tradeoff. Small observation budgets can leave the optimal action unchanged; a cached centroid resolves it through one scalar probe whose precision requirement tightens with candidate count. Charging for pool construction and cache maintenance gives explicit reuse thresholds. A stable Bag family with logistic relevance supplies a concrete learnability example, including i.i.d. labeled samples under explicit coverage conditions. Together, these results link RF-Mem's routing barrier to the observations, costs, and precision needed to overcome it.

1 Introduction

An adaptive memory system must decide whether to search again before seeing what that search would retrieve. Initial similarities describe the current query direction, but an updated query may expose evidence elsewhere. A more accurate classifier cannot resolve a decision if its inputs omit the information determining the next search's value.

RF-Mem (Zhang et al., 2026) routes between one-shot Familiarity and iterative Recollection using the mean M of initial Top-K similarities and entropy H of their softmax weights. Recollection clusters memories and updates the query, using memory–memory geometry absent from these summaries. Complementing the original empirical study and Appendix F analysis, we fix the mechanism and ask when its value is identifiable before execution and when geometric observations are worth acquiring. Our results are constructions and conditional guarantees, not estimates of deployment failure frequency.

Contributions. First, complete indexed initial scores can coincide while cost-optimal actions differ, yielding exact Bayes and minimax regret gaps. A designated off-query Gram entry can change expanded retrieval while every other entry remains fixed. Under published Bag semantics, action reversal extends to arbitrary retrieval multiplicities in a two-round single-beam regime and every finite round limit under explicit budget conditions. A positive perturbation radius removes duplicate embeddings at the two-round budget and a larger multiround budget. Degenerate regimes and competing beams delimit the scope.

Second, geometric insufficiency becomes a priced observation problem. In an explicit family, even the initial Top-K Gram block is uninformative, but seed–candidate inner products yield exact budget–regret and acquisition-cost optima. A cached centroid permits one aggregate probe at finer precision. A reusable-pool construction accounts for acquisition and maintenance through a

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.

break-even reuse count. A stable Bag family with logistic relevance further admits learned scalar gates from designed or i.i.d. labeled samples under explicit coverage.

Third, retrieval-success constructions make geometry, Top-K truncation-and-identity, and summary losses separately positive along RF-Mem's observation chain. Classical Gram factorization and value-of-information identities supply the background. Adaptive Gram blocks certify execution without specifying cheap routing observations. The mean–entropy, calibration, and upper-tail results provide conditional diagnostics: in particular, fixed-entropy regret bounds hold with the mean axes fixed, not for jointly tuned three-threshold gates.

2 Related Literature

Memory-augmented language models extend retrieval-augmented generation and dense retrieval (Lewis et al., 2020; Karpukhin et al., 2020) with persistent memory (Xu et al., 2022; Park et al., 2023; Packer et al., 2023; Wang et al., 2023a; Liu et al., 2023; Zhong et al., 2024; Pan et al., 2025; Xu et al., 2026b) and structured retrieval (Gutiérrez et al., 2024; Li et al., 2026; Xu et al., 2026a). Benchmarks evaluate memory and personalization (Salemi et al., 2024; Maharana et al., 2024; Wu et al., 2024; Jiang et al., 2025). RF-Mem (Zhang et al., 2026) operationalizes the familiarity–recollection distinction (Mandler, 1980; Jacoby, 1991; Yonelinas, 2002); Mnemis (Tang et al., 2026) uses a related System-1/System-2 architecture. We fix RF-Mem and ask what its routing observations identify. Adaptive retrieval (Asai et al., 2024; Jeong et al., 2024; Wang et al., 2023b; Cheng et al., 2024; Yan et al., 2024), interleaved retrieval and generation (Jiang et al., 2023; Trivedi et al., 2023; Shao et al., 2023; Su et al., 2024), relevance feedback and query expansion (Salton et al., 1975; Salton & Buckley, 1990; Harman, 1992; Lavrenko & Croft, 2001; Carpineto & Romano, 2012; Yu et al., 2021; Gao et al., 2023), and cost-aware model routing (Chen et al., 2023; Ding et al., 2024; Ong et al., 2025; Hu et al., 2024; Somerstep et al., 2025; Jitkrittum et al., 2026) address related decisions. Our focus is the value of RF-Mem's fixed expansion before execution.

Our acquisition objective instantiates cost-sensitive active classification (Greiner et al., 2002), combining observation costs with misclassification penalties. Our contribution is a feasible Gram family with identical initial Top-K geometry, a marked-entry correspondence to published Bag success, and its exact individual-versus-aggregate cost and precision consequences. Appendix A.2 gives the formal correspondence and the assumptions behind the closed form. Likewise, Gram factorization (Villar et al., 2021), comparisons by attainable Bayes risk (Blackwell, 1953; DeGroot, 1962), and plug-in margin bounds (Tsybakov & Audibert, 2007) are classical. Our constructions identify which omitted entries change retrieval value and make RF-Mem's representation losses separately strict; these losses are task-specific Bayes-risk gaps, not Le Cam deficiencies (Le, 1964).

Entropy bounds (Hellman & Raviv, 1970; Feder & Merhav, 1994; Sakai & Iwata, 2016) and calibration (Guo et al., 2017; Hébert-Johnson et al., 2018; Kumar et al., 2019; Gupta & Ramdas, 2021) distinguish concentration from correctness; selective prediction (Geifman & El-Yaniv, 2017; 2019) connects confidence to abstention. The Top-K mean connects to trimmed means (Stigler, 1973), conditional value-at-risk (Rockafellar et al., 2000; Rockafellar & Uryasev, 2002; Acerbi & Tasche, 2002), empirical-distribution bounds (Dvoretzky et al., 1956; Massart, 1990), and concentration for self-bounding functions and order statistics (Boucheron et al., 2009; Boucheron & Thomas, 2012). We use these tools to revisit RF-Mem's Appendix F: the action-risk difference determines the frontier, the lower entropy envelope gives the valid low-entropy certificate, and the selected mean is a dependent upper-tail functional. Appendix A.1 provides the claim-by-claim comparison.

3 RF-Mem and Decision-Theoretic Setup

We study RF-Mem (Zhang et al., 2026) as a canonical dual-process retriever with actions $\mathcal{A} = \{\mathrm{F}, \mathrm{R}\}$: one-shot Familiarity and iterative Recollection. Let $\mathbb{S}^{d-1} := \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$, let $q \in \mathbb{S}^{d-1}$ be the query, and let $z_1, \dots, z_n \in \mathbb{S}^{d-1}$ be the embeddings of a memory corpus $\mathcal{M}$. Define the initial similarities $V_i := \langle q, z_i \rangle$. With deterministic tie-breaking, let $S_1 \geq \cdots \geq S_K$ be the largest K values among $V_1, \dots, V_n$. RF-Mem routes from $T = (M, H)$, $M := \frac{1}{K}\sum_{j=1}^{K} S_j$, $H := -\sum_{j=1}^{K} P_j \log P_j$, $P_j := \frac{e^{\lambda S_j}}{\sum_{\ell=1}^{K} e^{\lambda S_\ell}}$, where $\lambda > 0$ is fixed. Recollection repeatedly retrieves by inner product, clusters the retrieved memories, mixes a cluster centroid with the current and

original queries, and beam-prunes the resulting queries while deduplicating retrieved memories. Let $V = (V_1, \ldots, V_n)$, $S = (S_1, \ldots, S_K)$, and $\mathcal{G} \coloneqq \sigma(q, z_1, \ldots, z_n)$, $\mathcal{V} \coloneqq \sigma(V)$, $\mathcal{S} \coloneqq \sigma(S)$, $\mathcal{T} \coloneqq \sigma(T)$. Any additional pre-routing randomness is included in $\mathcal{G}$. These representations form the information ladder $\mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{V} \subseteq \mathcal{G}$. The reductions $\mathcal{G} \to \mathcal{V}$, $\mathcal{V} \to \mathcal{S}$, and $\mathcal{S} \to \mathcal{T}$ discard, respectively, directions orthogonal to the initial query; all item identities together with scores below rank K; and the remaining configuration of the Top-K score values.

For $a \in \{\mathrm{F}, \mathrm{R}\}$, let L_a be an integrable cost-adjusted potential loss; for example, $L_a = \ell_a + \beta C_a$, where C_a measures computation, latency, or token cost. For an information set $\mathcal{H}$, an $\mathcal{H}$-measurable router $\pi : \Omega \to [0,1]$ chooses Recollection with conditional probability π. Its risk is $\mathcal{R}(\pi) \coloneqq \mathbb{E}[(1-\pi)L_{\mathrm{F}} + \pi L_{\mathrm{R}}]$. Let $\Pi(\mathcal{H}) \coloneqq \{\pi \colon \Omega \to [0,1] : \pi \text{ is } \mathcal{H}\text{-measurable}\}$ and define $\mathcal{R}^\star_{\mathcal{H}} \coloneqq \inf_{\pi \in \Pi(\mathcal{H})} \mathcal{R}(\pi)$, $D_{\mathcal{H}} \coloneqq \mathbb{E}[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \mathcal{H}]$. Thus $D_{\mathcal{H}} > 0$ favors Recollection and $D_{\mathcal{H}} < 0$ favors Familiarity.

4 Geometry of Recollection

We separate the initial query projections from the off-query geometry used by Recollection, and characterize exact and approximate Gram representations for routing.

At the first round $x^{(0)} = q$. For a cluster centroid $g \in \mathbb{R}^d$ with $\|g\|_2 \le 1$, define

$$\mathsf{U}_\alpha(q, g) \coloneqq \frac{(1+\alpha)q + (1-\alpha)g}{\|(1+\alpha)q + (1-\alpha)g\|_2}, \qquad \alpha \in [0,1],$$

whenever the numerator is nonzero. Decompose $c \coloneqq \langle q, g\rangle$, $g_\perp \coloneqq g - cq$, and $b \coloneqq \|g_\perp\|_2$. If $b > 0$, define $u \coloneqq \frac{g_\perp}{b}$. Then $g = cq + bu$, $\langle q, u\rangle = 0$, $\|u\|_2 = 1$.

Proposition 1 (Exact score change under one recollection update)**.** *Let $q, y \in \mathbb{S}^{d-1}$, $g \in \mathbb{R}^d$ and $\|g\|_2 \le 1$, and suppose that $\mathsf{U}_\alpha(q, g)$ is well-defined. Set $a \coloneqq \langle q, y\rangle \in (-1, 1)$. Then there exists a unit vector $v \in q^\perp$ such that $y = aq + \sqrt{1-a^2}\, v$. Define $A \coloneqq 1 + \alpha + (1-\alpha)c$, $B \coloneqq (1-\alpha)b$, and $D \coloneqq \sqrt{A^2 + B^2}$. If $b > 0$, let $\rho \coloneqq \langle u, v\rangle$. In that case,*

$$\langle \mathsf{U}_\alpha(q, g), y\rangle = \frac{Aa + B\sqrt{1-a^2}\,\rho}{D}, \tag{1}$$

and $\langle \mathsf{U}_\alpha(q, g), y\rangle - \langle q, y\rangle = \frac{B}{D}\left[\sqrt{1-a^2}\,\rho - \frac{aB}{D+A}\right]$. If, additionally, $B > 0$, the target score strictly improves if and only if $\rho > \frac{aB}{(D+A)\sqrt{1-a^2}}$. If $b > 0$ but $B = 0$, then $\mathsf{U}_\alpha(q, g) = q$ and the score change is zero. If $b = 0$, then $B = 0$, $\mathsf{U}_\alpha(q, g) = q$, and the score change is zero; in this case neither u nor ρ is required.

When $B > 0$, the score gain depends on $\rho = \langle u, v\rangle$, the alignment between the tangential components of the centroid and target. This quantity is not determined by their initial similarities to q. Hence even the complete initial score profile can fail to determine whether Recollection improves the target score.

Consider a one-expansion specialization with equal return budgets: $\mathcal{R}_{\mathrm{F}}(q; \mathcal{M}) \coloneqq \mathrm{Top}_3(q; \mathcal{M})$, $z_0 \coloneqq \mathrm{Top}_1(q; \mathcal{M})$, $q' \coloneqq \mathsf{U}_\alpha(q, z_0)$, $\mathcal{R}_{\mathrm{R}}(q; \mathcal{M}) \coloneqq \{z_0\} \cup \mathrm{Top}_2(q'; \mathcal{M} \setminus \{z_0\})$. This variant excludes the seed before expanded retrieval, unlike the published Bag semantics. Lemma 3 transfers the separation to the published Algorithm 3 with $B = F = 1$, $R = 2$, and $K = 3$. The impossibility is uniform over the implementation class, not a claim for every fixed budget or round limit.

Theorem 1 (Complete-score indistinguishability of recollection value)**.** *Let $d \ge 3$, $\alpha \in [0, 1)$ and $0 < a < s < 1$. Let $C_{\mathrm{R}} > C_{\mathrm{F}} \ge 0$, $\beta > 0$, and define $\Delta_c \coloneqq \beta\,(C_{\mathrm{R}} - C_{\mathrm{F}})$. Assume $0 < \Delta_c < 1$. Use the retrieval-success loss $L_a = \mathbf{1}\{y \notin \mathcal{R}_a(q; \mathcal{M})\} + \beta C_a$ for $a \in \{\mathrm{F}, \mathrm{R}\}$. Then there exist two memory corpora $\mathcal{M}_+$ and $\mathcal{M}_-$, a common query q, and a common unique relevant memory y such that: i) their complete indexed similarity profiles are identical; ii) Familiarity fails in both corpora; iii) Recollection succeeds in $\mathcal{M}_+$ and fails in $\mathcal{M}_-$; iv) Recollection is uniquely optimal in $\mathcal{M}_+$, whereas Familiarity is uniquely optimal in $\mathcal{M}_-$. Consequently, no router measurable only with respect to $\mathcal{V}$ can be Bayes-optimal in both environments. Under an equal prior over the two environments, $\mathcal{R}^\star_{\mathcal{V}} - \mathcal{R}^\star_{\mathcal{G}} = \frac{1}{2}\min\{\Delta_c, 1 - \Delta_c\}$. Moreover, the minimax regret over all score-only randomized routers equals $\inf_{\pi \in \Pi(\mathcal{V})} \max\left\{\mathrm{Reg}_+(\pi), \mathrm{Reg}_-(\pi)\right\} = \Delta_c(1 - \Delta_c)$.*

Equality of V implies equality of S, (M, H), and every deterministic score summary, so the obstruction precedes truncation and compression. Universal score-only optimality therefore requires assumptions linking visible scores to hidden geometry; stability is established in Corollary 11.

Theorem 2 (Single-beam separation under published Bag semantics)**.** *Fix $\alpha \in [0, 1)$, $d \geq 3$, and fixed costs with $\Delta_c = \beta(C_{\mathrm{R}} - C_{\mathrm{F}}) \in (0, 1)$. Use published Algorithm 3 with $B = 1$, an integer $F \geq 1$, exact full-corpus search, normalized centroids, fixed index ties, and no score floor. Either of the following parameter ranges admits complete-score-identical corpora with a common unique relevant index and opposite optimal actions: $R = 2, \quad K \geq 2F + 1$ or $R \geq 2, \quad K \geq (R + 1)F + 1$. The corpora have $n = K + 1$ indexed memories, allowing repeated embeddings, and all normalizations are nonzero. With $\mathcal{R}_{\mathrm{F}} = \mathrm{Top}_K(q; \mathcal{M})$ and $L_a = \mathbf{1}\{y \notin \mathcal{R}_a\} + \beta C_a$, their failure losses are $(\ell_{\mathrm{F},+}, \ell_{\mathrm{R},+}) = (1, 0)$ and $(\ell_{\mathrm{F},-}, \ell_{\mathrm{R},-}) = (1, 1)$. Their equal-prior Bayes gap is $\frac{1}{2}\min\{\Delta_c, 1 - \Delta_c\}$ and their score-only minimax regret is $\Delta_c(1 - \Delta_c)$. Every executed width $N_r = (1 + r)F$ satisfies the published restriction $N_r < K$. In the second range both runs execute all R rounds.*

These are sufficient ranges. Corollary 5 shows that small score-matched perturbations can remove all duplicate embeddings while preserving the losses when $K \geq F[R(R + 1)/2 - 1] + 1$. This is the same budget for $R = 2$ and a larger sufficient budget for $R > 2$; its proof controls Bag growth without preserving tied retrieval orders. Proposition 3 records regimes where Familiarity dominates. Competing beams remain unresolved (Appendix A.4).

Set $w_0 \coloneqq q$ and $w_i \coloneqq z_i$ for $i = 1, \ldots, n$, and define the indexed augmented Gram matrix $\Gamma(q, \mathcal{M}) \coloneqq (\langle w_i, w_j \rangle)_{i,j=0}^{n}$. The initial score profile is only its zeroth row: $V_i = \Gamma_{0i}$ for $i = 1, \ldots, n$. Let $e_0, \ldots, e_n$ be the standard basis of $\mathbb{R}^{n+1}$ and define $Wa \coloneqq a_0 q + \sum_{i=1}^{n} a_i z_i$, with $a \in \mathbb{R}^{n+1}$. Fix a finite round limit and all implementation choices of RF-Mem, including retrieval widths, beam width, deterministic tie-breaking, deduplication, stopping, and final ranking. Assume that K-means initialization, assignment, empty-cluster handling, and termination are deterministic functions of indexed Euclidean distances; a randomized implementation is covered after conditioning on its seed. Cluster representatives are arithmetic centroids, optionally unit-normalized by a fixed rule. The query update and cluster score are, respectively,

$$x' = \frac{\alpha x + (1 - \alpha) g + q}{\|\alpha x + (1 - \alpha) g + q\|_2}, \tag{2}$$

$$\mathrm{score}(C, x') = \sum_{i \in C} \langle x', z_i \rangle. \tag{3}$$

All other retrieval, beam, deduplication, stopping, and final-output decisions use only these inner-product scores, indexed set membership and cardinality, and the fixed tie-breaking rules. Assume that every normalization encountered by the run has a nonzero numerator. Appendix A.3 matches these conditions to the published procedure, including first-insertion Bag scores and final output. Lemmas 1 and 2 establish generic Gram factorization and verify it for the full RF-Mem trace.

Full-Gram sufficiency does not imply minimality. To state a sharper result, for index sets $A, B \subseteq \{0, 1, \ldots, n\}$, write $\Gamma_{A,B} \coloneqq (\Gamma_{ij})_{i \in A, j \in B}$, and for a coefficient vector $a \in \mathbb{R}^{n+1}$ write $J(a) \coloneqq \{j : a_j \neq 0\}$. Consider one active beam with query $x = Wa$, eligible memory set E, retrieved set $I \subseteq E$, and a K-means cluster $C \subseteq I$. Let $h_C \coloneqq \alpha a + (1 - \alpha)\widetilde{c}_C + e_0$, with $J_C \coloneqq J(h_C)$, where $\widetilde{c}_C$ is the coefficient vector of the raw or unit-normalized centroid. Let Q_C contain every memory index scored under the child query, including indices used for cluster scoring and final ranking.

Theorem 3 (Adaptive partial-Gram sufficiency and worst-case necessity)**.** *For the implementation class of Lemma 2, an exact simulation of the active beam step can be performed using only the following indexed blocks:*

$$\Gamma_{E,J(a)}, \qquad \Gamma_{I,I}, \qquad \Gamma_{J_C,J_C}, \qquad \Gamma_{Q_C,J_C}. \tag{4}$$

Their union, taken adaptively over all active beams and realized rounds, is sufficient for the complete execution trace. In the initial round, $a = e_0$, so the first block reduces to the complete initial score profile. Hence the full augmented Gram matrix is not generally minimal for a realized execution: any valid Gram matrix agreeing on all adaptively accessed entries produces the same trace. This reduction cannot be replaced by a fixed universal deletion of off-query entries. For every $\alpha \in [0, 1)$, already in a three-memory corpus each designated memory–memory position can be made

decisive: there are unit-vector corpora for which changing only that symmetric off-query pair, while preserving the entire initial score profile, changes the first expanded retrieval. Thus every off-query pair can be necessary in the worst case, although most need not be accessed on a particular run. The block collection in (4) *is an entrywise execution certificate; it is not claimed to be the unique minimal statistic, since RF-Mem may use only linear combinations of some listed entries.*

In particular, the relevant amount of geometry is controlled by active query supports and the clusters scored before pruning, not only by the corpus size. Ignoring overlap and symmetry, if $\mathcal{C}$ denotes the clusters spawned by one active beam, the number of entries requested at that beam is bounded by $|E|\,|J(a)| + |I|^2 + \sum_{C\in\mathcal{C}} \left(|J_C|^2 + |Q_C|\,|J_C|\right)$, with repeated entries cached across blocks, beams, and rounds. The blocks include every scored child, even if later pruned. In published Algorithm 3, $Q_C = C$ suffices for the same-round cluster and Bag scores; the next round's full-corpus search uses its own incoming block. Hence a pairwise inner product between two never-selected memories is unnecessary until the two enter a common clustering block or one enters an active query support.

Let $\mathfrak{C}_{n+1} := \left\{G \in \mathbb{R}^{(n+1)\times(n+1)} : G = G^\top,\ G \succeq 0,\ G_{ii} = 1 \text{ for } i = 0,\dots,n\right\}$ be the domain of unit-vector Gram matrices, and write $\|G\|_{\max} := \max_{i,j} |G_{ij}|$. Approximation error alone cannot control a discontinuous retrieval decision.

Proposition 2 (No margin-free approximate-Gram guarantee)**.** *Fix $\alpha \in [0,1)$ and a cost gap $\Delta_c \in (0,1)$. For every $\varepsilon > 0$, there exist two valid Gram matrices Γ_+, Γ_- and a common valid approximation $\widehat{\Gamma}$ such that*

$$\max\left\{\|\widehat{\Gamma} - \Gamma_+\|_{\max}, \|\widehat{\Gamma} - \Gamma_-\|_{\max}\right\} < \varepsilon, \tag{5}$$

but the corresponding one-expansion environments require opposite optimal actions. Any randomized router receiving only the common approximation must therefore satisfy

$$\inf_{p\in[0,1]} \max\left\{\mathrm{Reg}_+(p), \mathrm{Reg}_-(p)\right\} = \Delta_c(1-\Delta_c). \tag{6}$$

Consequently, no universal regret upper bound depending only on $\|\widehat{\Gamma} - \Gamma\|_{\max}$ can converge to zero for hard retrieval-success losses.

Let $\mathcal{K}_\Gamma = \sigma(\Gamma)$ and fix a Borel version $d_\Gamma(G) = \mathbb{E}[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \Gamma = G]$. On the domain $\mathfrak{C}$ of Assumption 2 in Appendix B, let $\Gamma, \widehat{\Gamma} \in \mathfrak{C}$ almost surely, with the approximation constructed from Γ and randomness conditionally independent of the potential losses given Γ. The oracle router $\widehat{\pi} = \mathbf{1}\{d_\Gamma(\widehat{\Gamma}) > 0\}$ has excess risk at most $L_\Gamma \mathbb{E}\|\widehat{\Gamma} - \Gamma\|_{\max}$; if the error exceeds ε with probability at most δ, the margin condition sharpens the bound to $C_{\mathrm{mar}}(L_\Gamma\varepsilon)^{\kappa+1} + B\delta$. Lemma 7 gives the complete statements and standard plug-in proof. These are representation stability bounds using the true advantage, not learning guarantees. The exponent is the classical margin conversion. Proposition 4 separately derives regularity from a concrete stable Bag model and adds an estimation term under an explicit labeled-data design.

5 Exact Information-Ladder Decomposition

Lemma 5 in the Appendix states the generic binary-action value-of-information identity. We apply it to RF-Mem's observation chain to audit the loss caused by geometry, truncation and identity removal, summary compression, and the implemented gate. For $x \in \mathbb{R}$, write $x_+ := \max\{x, 0\}$ and $x_- := \max\{-x, 0\}$. Then $x = x_+ - x_-$ and $|x| = x_+ + x_-$.

Corollary 1 (RF-Mem observation-loss audit)**.** *Let $\mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{V} \subseteq \mathcal{G}$ be the RF-Mem information ladder defined in Section 3. For every $\mathcal{T}$-measurable randomized router π,*

$$\begin{aligned}\mathcal{R}(\pi) - \mathcal{R}^\star_{\mathcal{G}} = \underbrace{\mathbb{E}\left[(1-\pi)(D_{\mathcal{T}})_+ + \pi(D_{\mathcal{T}})_-\right]}_{\mathfrak{E}_{\mathrm{gate}}(\pi)} + \underbrace{\frac{1}{2}\mathbb{E}\left[|D_{\mathcal{S}}| - |D_{\mathcal{T}}|\right]}_{\mathfrak{E}_{\mathrm{summary}}} \\ + \underbrace{\frac{1}{2}\mathbb{E}\left[|D_{\mathcal{V}}| - |D_{\mathcal{S}}|\right]}_{\mathfrak{E}_{\mathrm{trunc+id}}} + \underbrace{\frac{1}{2}\mathbb{E}\left[|D_{\mathcal{G}}| - |D_{\mathcal{V}}|\right]}_{\mathfrak{E}_{\mathrm{geometry}}}.\end{aligned}$$

Every term on the right-hand side is nonnegative.

The Gram results refine the first reduction in this ladder. Since V is the query row of Γ, $\mathcal{V} \subseteq \mathcal{K}_\Gamma \subseteq \mathcal{G}$.

Corollary 2 (Gram refinement of geometry loss)**.** *Let $D_{\mathcal{K}_\Gamma} := \mathbb{E}[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \mathcal{K}_\Gamma]$. Then*

$$\mathfrak{E}_{\text{geometry}} = \underbrace{\frac{1}{2}\mathbb{E}\left[|D_{\mathcal{K}_\Gamma}| - |D_{\mathcal{V}}|\right]}_{\mathfrak{E}_{\text{off-query}}} + \underbrace{\frac{1}{2}\mathbb{E}\left[|D_{\mathcal{G}}| - |D_{\mathcal{K}_\Gamma}|\right]}_{\mathfrak{E}_{\text{beyond-Gram}}}. \tag{7}$$

Both terms are nonnegative. If $D_{\mathcal{G}}$ is $\mathcal{K}_\Gamma$-measurable, then $\mathfrak{E}_{\text{beyond-Gram}} = 0$; in that case all geometry-induced routing loss is exactly the value lost by reducing the relevant Gram information to its initial-score row.

The generic identity does not imply that any particular RF-Mem reduction is strict. The next result supplies this mechanism-specific conclusion using retrieval-success losses rather than freely assigned potential outcomes.

Theorem 4 (Stagewise strictness of the RF-Mem information ladder)**.** *Fix $d \geq 3$, $\alpha \in [0, 1)$, and $\lambda > 0$, and consider the $K = 3$ one-expansion specialization preceding Theorem 1. In each environment, let y denote its unique relevant memory and set $\ell_a := \mathbf{1}\{y \notin \mathcal{R}_a(q; \mathcal{M})\}$ and $L_a := \ell_a + \beta C_a$. Suppose that $C_{\mathrm{R}} > C_{\mathrm{F}} \geq 0$ and $\Delta_c := \beta(C_{\mathrm{R}} - C_{\mathrm{F}}) \in (0, 1)$, and write*

$$g_c := \frac{1}{2}\min\{\Delta_c, 1 - \Delta_c\}.$$

For each $j \in \{\text{geometry}, \text{trunc} + \text{id}, \text{summary}\}$, there exists an equal-prior pair of admissible RF-Mem environments that are indistinguishable at the coarser endpoint of reduction j, distinguishable at its finer endpoint, and require opposite unique optimal actions, such that

$$\mathfrak{E}_j = g_c, \qquad \mathfrak{E}_{j'} = 0 \quad \textit{for every } j' \in \{\text{geometry}, \text{trunc} + \text{id}, \text{summary}\} \setminus \{j\}.$$

For the geometry pair, moreover,

$$\mathfrak{E}_{\text{off-query}} = g_c, \qquad \mathfrak{E}_{\text{beyond-Gram}} = 0.$$

The truncation pair's complete initial-score profiles can differ only in one coordinate below the initial Top-3 set, while the summary pair has distinct ordered Top-3 score vectors but exactly the same mean and softmax entropy. In each case the minimax regret of a randomized router observing only the coarser representation is $\Delta_c(1 - \Delta_c)$. Hence none of the three RF-Mem representation reductions is universally lossless.

Lemma 3 preserves these losses and stagewise gaps under the published two-round Bag procedure. The four terms represent gate, summary, truncation-and-identity, and geometry losses. Only the gate term depends on the implemented router; Corollary 12 characterizes exact losslessness of the remaining reductions.

6 Additional Observations Before Recollection

An execution certificate need not be cheap to acquire. For a singleton initial centroid z_{i_*}, however, the score of a chosen candidate z_j after the first update is

$$\langle q', z_j \rangle = \frac{(1+\alpha)V_j + (1-\alpha)\Gamma_{i_* j}}{\sqrt{(1+\alpha)^2 + (1-\alpha)^2 + 2(1-\alpha^2)V_{i_*}}}. \tag{8}$$

Once the initial scores and candidate identities are available, one cross-inner-product gives this candidate score without clustering or an expanded full-corpus search. For $B = 1$ and $F > 1$, forming the initial Top-F centroid costs $O(Fd)$; when nonzero, its unit normalization g replaces the seed in the same formula, using $\langle q, g \rangle$ and $\langle g, z_j \rangle$. For $B > 1$, identifying the actual child centroids additionally requires clustering. The next result quantifies both the value of such observations and the cost of not knowing which candidate matters.

Theorem 5 (Exact cost–regret tradeoff for individual Gram probes)**.** *Fix $\alpha \in [0, 1)$, $K \geq 3$, an integer $m \geq 1$, and $\Delta_c \in (0, 1)$. For published Algorithm 3 with $B = F = 1$ and $R = 2$, there is an explicit prior on $n = K + m$ indexed unit-vector corpora in dimension $m + 3$, with unique relevant indices and the loss of Theorem 2, having the following properties. The complete initial*

profile V, the initial Top-K index set I, and its entire augmented Gram block $\Gamma_{\{0\}\cup I,\{0\}\cup I}$ are constant. Only the m seed–candidate entries $\Gamma_{i_* j}$, $j \notin I$, can vary. With probability $1/2$ none is marked and Familiarity is optimal; otherwise a uniformly located single marked entry identifies a candidate recovered only by Recollection. All other Gram entries are constant.

Consider a policy supplied with the common observations and this family, which adaptively requests individual Gram entries and then routes. Let Π_t^{probe} allow at most t requests, $t \in \{0, \dots, m\}$, and let Q count requests. Its optimal routing excess risk, excluding acquisition cost, is

$$r_t := \inf_{\pi \in \Pi_t^{\mathrm{probe}}} \{\mathcal{R}(\pi) - \mathcal{R}_{\mathcal{G}}^\star\} = \frac{1}{2} \min\left\{\Delta_c, \left(1 - \frac{t}{m}\right)(1 - \Delta_c)\right\}. \tag{9}$$

A policy attaining r_t tests distinct candidates until a mark is found or t have been tested, with $\mathbb{E}Q = t - t(t-1)/(4m)$. If each request costs $c_p \geq 0$ in loss units, put $g_c = \frac{1}{2}\min\{\Delta_c, 1 - \Delta_c\}$. The optimal total excess risk over all such acquisition policies is

$$\inf_\pi \{\mathcal{R}(\pi) + c_p \mathbb{E}Q - \mathcal{R}_{\mathcal{G}}^\star\} = \min\left\{g_c,\ c_p \frac{3m+1}{4}\right\}. \tag{10}$$

Thus acquisition strictly improves the best unprobed decision exactly when $c_p < 4g_c/(3m+1)$. For $m = 1$, one scalar is sufficient and zero are insufficient. The attaining policies also work with uniformly bounded scalar error smaller than half the construction's marked–unmarked gap.

An optimal policy therefore either acquires nothing or tests distinct candidates until a mark is found or all m candidates have been checked; the latter is preferable below the displayed cost threshold, and either is optimal at equality. The query model charges for individual inner products and supplies candidate identities; it does not supply free coordinate-based tests of the hidden configuration. With available embeddings, each requested inner product costs $O(d)$ arithmetic; candidate acquisition and memory access may add cost. The theorem gives a decision rule without evaluating an unknown d_Γ, but only for its specified family and prior. It demonstrates why a small informative candidate pool can make pre-routing probes worthwhile, and why a trace certificate alone supplies no corpus-independent cheap policy in that observation model. For $\Delta_c < 1/2$, (9) remains $\Delta_c/2$ whenever $t/m \leq (1-2\Delta_c)/(1-\Delta_c)$: actually acquiring positive-cost observations in this range reduces uncertainty without reducing the best routing risk. Before the endpoint, Recollection remains uniquely optimal after either a mark or an unmarked history. The theorem solves a known-prior problem with at most one mark in dimension $m+3$; it gives neither a fixed-dimension lower bound as m grows nor a probe-selection guarantee for arbitrary corpora. Appendix A.2 separates these restrictions from the generic active-classification objective.

Aggregate observations can be more economical. In the same family, if the raw candidate centroid $\bar{y}$ is cached, the single scalar $\langle z_{i_*}, \bar{y}\rangle$ determines the optimal action for any m, while leaving the marked candidate unidentified when $m > 1$. Its marked–unmarked gap is only ζ/m, where ζ is the individual-entry gap. Corollary 8 gives the corresponding cost and precision requirements. Forming an uncached centroid costs $O(md)$ arithmetic; the cache is an additional observation, so this does not contradict the individual-entry lower bound. Corollary 9 realizes a reusable pool with fresh uncertainty at each decision. With setup cost A_0, online cost c_{on}, and average maintenance cost at most $\bar{c}_{\mathrm{maint}}$, probing pays when $N_{\mathrm{dec}} > A_0/(g_c - c_{\mathrm{on}} - \bar{c}_{\mathrm{maint}})$ and the denominator is positive, provided cache and scalar errors remain below the half-gap. This includes candidate acquisition, refresh, and memory access; it supplies no general pool-discovery algorithm or measured latency claim.

Proposition 4 supplies an existence example of learnable geometric advantage: a continuous two-round Bag family with positive retrieval gaps, bounded normalization, and logistic relevance conditional on one seed–candidate inner product. Its advantage is Lipschitz with margin exponent $\kappa = 1$. Designed labels at grid points give the bound $O(\varepsilon_t^2 + (\log N/N)^{2/3})$ with total sample size $N = J_g N_g$ (Corollary 6). Corollary 7 obtains the same order from ordinary i.i.d. labeled instances when the scalar density is bounded above and away from zero; it explicitly controls random bin counts and the coverage-dependent constants. Both results concern this relevance model with observed labels and scalar features, not arbitrary routing logs or clustering and ranking boundaries.

7 Optimal Routing on the Mean–Entropy Plane

We study the Bayes rule based on $T = (M, H)$. Let $\mathcal{Z} = [m_-, m_+] \times [h_-, h_+]$, with $m_- < m_+$ and $h_- < h_+$, be a compact ambient domain containing its support; the realized support may be nonrectangular. Fix Borel versions $r_a(m, h) = \mathbb{E}[L_a \mid M = m, H = h]$ and define $d = r_{\mathrm{F}} - r_{\mathrm{R}}$, so $D_{\mathcal{T}} = d(M, H)$ almost surely and the Bayes router is $\pi^\star(m, h) = \mathbf{1}\{d(m, h) > 0\}$, choosing Familiarity on ties. The relevant monotonicity is that of the risk difference, not either action risk alone; Appendix A.8 gives a counterexample.

Under the derivative and corner conditions of Assumption 3 in Appendix B, Lemma 25 supplies unique Bayes-axis thresholds $\theta_{\mathrm{low}} < \theta_{\mathrm{high}}$ and an increasing interior boundary $b(m)$ satisfying $d(m, b(m)) = 0$. Recollection is optimal below the lower mean threshold, Familiarity above the upper one, and the interior decision is $\mathbf{1}\{h > b(m)\}$. This is a conditional characterization using the implicit function theorem, not a monotonicity property derived from the RF-Mem mechanism.

To isolate the effect of a constant entropy cutoff, fix the mean thresholds at these values and define

$$\pi_\tau(m, h) = \begin{cases} 1, & m \le \theta_{\mathrm{low}}, \\ \mathbf{1}\{h > \tau\}, & \theta_{\mathrm{low}} < m < \theta_{\mathrm{high}}, \\ 0, & m \ge \theta_{\mathrm{high}}, \end{cases} \qquad \tau \in [h_-, h_+].$$

Write $I = \{\theta_{\mathrm{low}} < M < \theta_{\mathrm{high}}\}$, $\mathcal{J}_\tau(m) = [\min\{b(m), \tau\}, \max\{b(m), \tau\}]$, and, using the density from Assumption 4, define

$$\mathfrak{Q}(\tau) = \mathbb{E}\left[\mathbf{1}_I \int_{\mathcal{J}_\tau(M)} |h - b(M)| f_{H|M}(h \mid M)\, dh\right].$$

Lemma 26 gives $\mu_h \mathfrak{Q}(\tau) \le \mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}} \le L_h \mathfrak{Q}(\tau)$. Density bounds only on the disagreement segments yield its two-sided quadratic form. Corollary 10 characterizes positive regret after optimizing τ within this fixed-axis family. These bounds do not lower bound jointly tuned three-threshold gates: moving the mean thresholds can partly compensate for a fixed entropy threshold. Appendix A.6 gives an explicit example reducing excess risk from $1/24$ to $1/96$ in this way.

8 What Entropy Can and Cannot Certify

We derive the sharp low-entropy concentration certificate and identify the additional calibration needed to translate concentration into correctness. For $K \ge 2$ and the weights in the RF-Mem routing summary, let $\mathbf{P} = (P_1, \ldots, P_K) \in \Delta_K$, where

$$\Delta_K := \left\{ p = (p_1, \ldots, p_K) \in [0, 1]^K : \sum_{j=1}^K p_j = 1 \right\}.$$

For $p \in \Delta_K$, write $H(p) := -\sum_{j=1}^K p_j \log p_j$ with $0 \log 0 := 0$, and $p_{\max} := \max_j p_j$. For $x \in [1/K, 1]$, set $r_x := \lfloor 1/x \rfloor$, $u_x := 1 - r_x x$, and $H_K^{\min}(x) := -r_x x \log x - u_x \log u_x$, where $u_x \log u_x := 0$ when $u_x = 0$. This is the minimum entropy among vectors with maximum mass x; the full sharp lower and upper envelopes, their extremizers, and the majorization argument are stated in Appendix A.9. By Lemma 11, $H_K^{\min}$ is continuous and strictly decreasing from $\log K$ to 0. Define $\psi_K(\tau) := \left(H_K^{\min}\right)^{-1}(\tau)$ so that $H_K^{\min}(\psi_K(\tau)) = \tau$. Here and below, sharpness refers to the full simplex Δ_K. For $K \ge 2$, with cosine scores in $[-1, 1]$ and finite λ, RF-Mem softmax vectors form a strict subset of Δ_K. The certificate remains valid on that subset but can be conservative because a simplex extremizer need not be attainable. Proposition 5 characterizes the feasible set exactly.

Lemma 12 gives the sharp simplex certificate $H(p) \le \tau \implies p_{\max} \ge \psi_K(\tau)$ for $\tau \in [0, \log K]$. No stronger universal lower bound on $p_{\max}$ follows from $H(p) \le \tau$. The corresponding threshold-design corollary, the high-entropy upper certificate, and the direction of the two envelope inversions are collected in Appendix A.9. Let $C_{\mathrm{F}} \in \{0, 1\}$ indicate success of Familiarity under a fixed evaluation criterion, and define $A_\tau := \{H(\mathbf{P}) \le \tau\}$. Lemma 13 constructs identical laws of $\mathbf{P}$ with opposite deterministic correctness outcomes, even under low entropy. Thus even the full

marginal law of $\mathbf{P}$ does not identify correctness; a conditional link to the outcome is necessary. Define the conditional Familiarity success probability $\eta_{\mathrm{F}}(\mathbf{P}) \coloneqq \mathbb{E}\left[C_{\mathrm{F}} \mid \sigma(\mathbf{P})\right]$. The conditioning is on the complete realized vector rather than only on $P_{\max}$, so the bound below is stable under events in $\sigma(\mathbf{P})$. Under Assumption 5 in Appendix B, the tower-property argument in Lemma 14 yields $\mathbb{P}(C_{\mathrm{F}} = 1 \mid A_\tau) \geq [g(\psi_K(\tau)) - \varepsilon_{\mathrm{cal}}]_+$ whenever $\tau \in [0, \log K]$ and $\mathbb{P}(A_\tau) > 0$. Within-list and marginal $P_{\max}$ calibration can fail after entropy selection; Appendix A.10 gives explicit counterexamples. Thus entropy alone certifies concentration, not correctness.

9 The Probe Mean Is an Upper-Tail Statistic

Under the bounded i.i.d. baseline $X_1, \ldots, X_n \sim F$, which isolates Top-K selection rather than modeling dependence in real corpora, the probe mean is $\widehat{M}_{n,K} = K^{-1} \sum_{j=n-K+1}^{n} X_{(j)}$. Write $\ell \leq X_i \leq u$, with $\ell < u$, and $\Delta_s \coloneqq u - \ell$.

For $\gamma \coloneqq K/n$ and $X \sim F$, define

$$U_\gamma(F) \coloneqq \min_{t \in [\ell, u]} \left\{ t + \frac{1}{\gamma} \mathbb{E}[(X - t)_+] \right\}. \tag{11}$$

Define the lower quantile function by $Q(v) \coloneqq F^{-1}(v) \coloneqq \inf\{x \in \mathbb{R} : F(x) \geq v\}$ with $v \in (0, 1)$. Its quantile form is $U_\gamma(F) = \gamma^{-1} \int_{1-\gamma}^{1} Q(v)\, dv$. Consequently, $U_1(F) = \mathbb{E}[X]$, whereas for nondegenerate X and $\gamma < 1$ the target is strictly larger and increases as the retained fraction decreases. Finite-sample concentration around both the expectation and the population target is developed in Appendix A.11.

Fixed K targets the essential upper endpoint, whereas $K_n/n \to \gamma > 0$ targets the stable upper-tail functional $U_\gamma(F)$. The individual convergence theorems and the pathwise monotonicity result are stated in Appendix A.11; their consequences are consolidated in Lemma 24 of Appendix C. In particular, Lemma 24 gives $\mathbb{P}\Big(|\widehat{M}_{n,K} - U_{K/n}(F)| > \varepsilon\Big) \leq 2 \exp\Big\{-\frac{2K^2\varepsilon^2}{n\Delta_s^2}\Big\}$ for $\varepsilon > 0$. Let $x_F^\star \coloneqq \operatorname{ess\,sup} X_1$. For a deployed router whose threshold θ_{high} is held fixed as the corpus grows, item iv of Lemma 24 and the high-mean regime of Lemma 25 yield the following consequence.

Corollary 3. *Suppose the deployed threshold θ_{high} does not depend on n and satisfies $\theta_{\mathrm{high}} < x_F^\star$. Then $\mathbb{P}\left(\widehat{M}_{n,K} > \theta_{\mathrm{high}} \text{ for all sufficiently large } n\right) = 1$.*

Thus a high probe mean can reflect stronger query–memory similarity or merely more extreme selection from a larger corpus. The variables K, n, and K/n are therefore part of the statistic's meaning, not implementation details.

10 Discussion

RF-Mem can benefit from Recollection when even its complete initial scores cannot identify that benefit. The score-identical constructions and stagewise losses distinguish information missing from the observations from errors a better classifier can remove. The Bag extension controls later insertion and truncation across explicit single-beam ranges. Under a stronger sufficient budget, a perturbation argument removes exact duplicates; action reversal with competing beams remains open.

The probe result links missing geometry to acquisition cost. Individual-entry acquisition depends on candidate uncertainty; a cached centroid offers one aggregate probe at finer precision. A reusable-pool construction gives break-even conditions including acquisition, maintenance, and cache error. Finding such pools cheaply in arbitrary corpora remains open. The continuous logistic Bag model permits i.i.d. labeled training under density coverage, providing an existence example of learnable advantage within a stable retrieval region.

The remaining frontier, calibration, and corpus-growth results offer conditional diagnostics. Their scope is recorded in Appendix A.7; estimating population risk gaps requires action coverage or other identifying assumptions. The design objective is to expose enough information to value counterfactual retrieval while charging for the observations that make that decision possible.

AI use statement In this work, we used generative AI tools for spell and grammar check of English writing. We have not used generative AI tools for help develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, propose or refine hypotheses, and the rest of the required disclosure tasks are not applicable to this work. Additionally, we used generative AI tools for identify relevant literature. We have reviewed all AI-assisted work. We checked all LLM-assisted survey results. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

Reproducibility statement This paper presents a theoretical analysis and reports no new empirical experiments. Section 3 defines the model and cost-adjusted risks, while Appendix A.3 specifies the retrieval, tie-breaking, normalization, and Bag conventions. The appendices provide complete proofs, explicit geometric constructions, admissible parameter ranges, and perturbation bounds. The probe results specify their observation interfaces, acquisition and maintenance costs, and precision requirements. The relevance model, sampling schemes, and estimators underlying the learning guarantees are detailed in Appendix A.5 and Corollary 7. Appendix A.7 summarizes the assumptions and scope of the results to support independent verification.

REFERENCES

Carlo Acerbi and Dirk Tasche. On the coherence of expected shortfall. *Journal of banking & finance*, 26(7):1487–1503, 2002.

Akari Asai, Zeqiu Wu, Yizhong Wang, Avi Sil, and Hannaneh Hajishirzi. Self-rag: Learning to retrieve, generate, and critique through self-reflection. In *International conference on learning representations*, volume 2024, pp. 9112–9141, 2024.

David Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, pp. 265–272, 1953.

Stéphane Boucheron and Maud Thomas. Concentration inequalities for order statistics. 2012.

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. On concentration of self-bounding functions. 2009.

Claudio Carpineto and Giovanni Romano. A survey of automatic query expansion in information retrieval. *Acm Computing Surveys (CSUR)*, 44(1):1–50, 2012.

Lingjiao Chen, Matei Zaharia, and James Zou. Frugalgpt: How to use large language models while reducing cost and improving performance. *arXiv preprint arXiv:2305.05176*, 2023.

Qinyuan Cheng, Xiaonan Li, Shimin Li, Qin Zhu, Zhangyue Yin, Yunfan Shao, Linyang Li, Tianxiang Sun, Hang Yan, and Xipeng Qiu. Unified active retrieval for retrieval augmented generation. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pp. 17153–17166, 2024.

Morris H DeGroot. Uncertainty, information, and sequential experiments. *The Annals of Mathematical Statistics*, 33(2):404–419, 1962.

Dujian Ding, Ankur Mallick, Chi Wang, Robert Sim, Subhabrata Mukherjee, Victor Rühle, Laks Lakshmanan, and Ahmed H Awadallah. Hybrid llm: Cost-efficient and quality-aware query routing. In *International Conference on Learning Representations*, volume 2024, pp. 41348–41366, 2024.

Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, pp. 642–669, 1956.

Meir Feder and Neri Merhav. Relations between entropy and error probability. *IEEE Transactions on Information theory*, 40(1):259–266, 1994.

Luyu Gao, Xueguang Ma, Jimmy Lin, and Jamie Callan. Precise zero-shot dense retrieval without relevance labels. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1762–1777, 2023.

Yonatan Geifman and Ran El-Yaniv. Selective classification for deep neural networks. *Advances in neural information processing systems*, 30, 2017.

Yonatan Geifman and Ran El-Yaniv. Selectivenet: A deep neural network with an integrated reject option. In *International conference on machine learning*, pp. 2151–2159. PMLR, 2019.

Russell Greiner, Adam J Grove, and Dan Roth. Learning cost-sensitive active classifiers. *Artificial Intelligence*, 139(2):137–174, 2002.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pp. 1321–1330. PMLR, 2017.

Chirag Gupta and Aaditya Ramdas. Top-label calibration and multiclass-to-binary reductions. *arXiv preprint arXiv:2107.08353*, 2021.

Bernal J Gutiérrez, Yiheng Shu, Yu Gu, Michihiro Yasunaga, and Yu Su. Hipporag: Neurobiologically inspired long-term memory for large language models. *Advances in neural information processing systems*, 37:59532–59569, 2024.

Donna Harman. Relevance feedback revisited. In *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 1–10, 1992.

Ursula Hébert-Johnson, Michael Kim, Omer Reingold, and Guy Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In *International conference on machine learning*, pp. 1939–1948. PMLR, 2018.

Martin Hellman and Josef Raviv. Probability of error, equivocation, and the chernoff bound. *IEEE Transactions on Information Theory*, 16(4):368–372, 1970.

Qitian Jason Hu, Jacob Bieker, Xiuyu Li, Nan Jiang, Benjamin Keigwin, Gaurav Ranganath, Kurt Keutzer, and Shriyash Kaustubh Upadhyay. Routerbench: A benchmark for multi-llm routing system. *arXiv preprint arXiv:2403.12031*, 2024.

Larry L Jacoby. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of memory and language*, 30(5):513–541, 1991.

Soyeong Jeong, Jinheon Baek, Sukmin Cho, Sung Ju Hwang, and Jong C Park. Adaptive-rag: Learning to adapt retrieval-augmented large language models through question complexity. In *Proceedings of the 2024 conference of the north american chapter of the association for computational linguistics: Human language technologies (volume 1: Long papers)*, pp. 7036–7050, 2024.

Bowen Jiang, Zhuoqun Hao, Young-Min Cho, Bryan Li, Yuan Yuan, Sihao Chen, Lyle Ungar, Camillo J Taylor, and Dan Roth. Know me, respond to me: Benchmarking llms for dynamic user profiling and personalized responses at scale. *arXiv preprint arXiv:2504.14225*, 2025.

Zhengbao Jiang, Frank F Xu, Luyu Gao, Zhiqing Sun, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan, and Graham Neubig. Active retrieval augmented generation. In *Proceedings of the 2023 conference on empirical methods in natural language processing*, pp. 7969–7992, 2023.

Wittawat Jitkrittum, Harikrishna Narasimhan, Ankit Singh Rawat, Jeevesh Juneja, Congchao Wang, Zifeng Wang, Alec Go, Chen-Yu Lee, Pradeep Shenoy, Rina Panigrahy, et al. Universal model routing for efficient llm inference. In *International Conference on Learning Representations*, volume 2026, pp. 10169–10218, 2026.

Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. Dense passage retrieval for open-domain question answering. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, pp. 6769–6781, 2020.

Ananya Kumar, Percy S Liang, and Tengyu Ma. Verified uncertainty calibration. *Advances in neural information processing systems*, 32, 2019.

Victor Lavrenko and W Bruce Croft. Relevance-based language models: Estimation and analysis. *Croft and Lafferty [2]*, pp. 1–5, 2001.

L Le. Sufficiency and approximate sufficiency. *The Annals of Mathematical Statistics*, pp. 1419–1455, 1964.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33: 9459–9474, 2020.

Rui Li, Zeyu Zhang, Xiaohe Bo, Zihang Tian, Xu Chen, Quanyu Dai, Zhenhua Dong, and Ruiming Tang. Cam: A constructivist view of agentic memory for llm-based reading comprehension. *Advances in Neural Information Processing Systems*, 38:113381–113406, 2026.

Lei Liu, Xiaoyan Yang, Yue Shen, Binbin Hu, Zhiqiang Zhang, Jinjie Gu, and Guannan Zhang. Think-in-memory: Recalling and post-thinking enable llms with long-term memory. *arXiv preprint arXiv:2311.08719*, 2023.

Adyasha Maharana, Dong-Ho Lee, Sergey Tulyakov, Mohit Bansal, Francesco Barbieri, and Yuwei Fang. Evaluating very long-term conversational memory of llm agents. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 13851–13870, 2024.

George Mandler. Recognizing: The judgment of previous occurrence. *Psychological review*, 87(3): 252, 1980.

Pascal Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The annals of Probability*, pp. 1269–1283, 1990.

Isaac Ong, Amjad Almahairi, Vincent Wu, Wei-Lin Chiang, Tianhao Wu, Joseph E Gonzalez, Mohammed Kadous, and Ion Stoica. Routellm: Learning to route llms from preference data. In *International Conference on Learning Representations*, volume 2025, pp. 34433–34448, 2025.

Charles Packer, Sarah Wooders, Kevin Lin, Vivian Fang, Shishir G Patil, Ion Stoica, and Joseph E Gonzalez. Memgpt: Towards llms as operating systems. *arXiv preprint arXiv:2310.08560*, 2023.

Zhuoshi Pan, Qianhui Wu, Huiqiang Jiang, Xufang Luo, Hao Cheng, Dongsheng Li, Yuqing Yang, Chin-Yew Lin, H Vicky Zhao, Lili Qiu, et al. Secom: On memory construction and retrieval for personalized conversational agents. In *International Conference on Learning Representations*, volume 2025, pp. 91851–91885, 2025.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pp. 1–22, 2023.

R Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.

R Tyrrell Rockafellar, Stanislav Uryasev, et al. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

Yuta Sakai and Ken-ichi Iwata. Sharp bounds between two r\'enyi entropies of distinct positive orders. *arXiv preprint arXiv:1605.00019*, 2016.

Alireza Salemi, Sheshera Mysore, Michael Bendersky, and Hamed Zamani. Lamp: When large language models meet personalization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 7370–7392, 2024.

Gerard Salton and Chris Buckley. Improving retrieval performance by relevance feedback. *Journal of the American society for information science*, 41(4):288–297, 1990.

Gerard Salton, Anita Wong, and Chung-Shu Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.

Zhihong Shao, Yeyun Gong, Yelong Shen, Minlie Huang, Nan Duan, and Weizhu Chen. Enhancing retrieval-augmented large language models with iterative retrieval-generation synergy. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 9248–9274, 2023.

Seamus Somerstep, Felipe Maia Polo, Allysson Flavio Melo de Oliveira, Prattyush Mangal, Mírian Silva, Onkar Bhardwaj, Mikhail Yurochkin, and Subha Maity. Carrot: A cost aware rate optimal router. *arXiv preprint arXiv:2502.03261*, 2025.

Stephen M Stigler. The asymptotic distribution of the trimmed mean. *The Annals of Statistics*, pp. 472–477, 1973.

Weihang Su, Yichen Tang, Qingyao Ai, Zhijing Wu, and Yiqun Liu. Dragin: Dynamic retrieval augmented generation based on the real-time information needs of large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 12991–13013, 2024.

Zihao Tang, Xin Yu, Ziyu Xiao, Zengxuan Wen, Zelin Li, Jiaxi Zhou, Hualei Wang, Haohua Wang, Haizhen Huang, Weiwei Deng, et al. Mnemis: Dual-route retrieval on hierarchical graphs for long-term llm memory. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 23914–23928, 2026.

Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, and Ashish Sabharwal. Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions. In *Proceedings of the 61st annual meeting of the association for computational linguistics (volume 1: long papers)*, pp. 10014–10037, 2023.

AB Tsybakov and J-Y Audibert. Fast learning rates for plug-in classifiers. *Annals of Statistics*, 35 (2):608–633, 2007.

Soledad Villar, David W Hogg, Kate Storey-Fisher, Weichi Yao, and Ben Blum-Smith. Scalars are universal: Equivariant machine learning, structured like classical physics. *Advances in neural information processing systems*, 34:28848–28863, 2021.

Weizhi Wang, Li Dong, Hao Cheng, Xiaodong Liu, Xifeng Yan, Jianfeng Gao, and Furu Wei. Augmenting language models with long-term memory. *Advances in neural information processing systems*, 36:74530–74543, 2023a.

Yile Wang, Peng Li, Maosong Sun, and Yang Liu. Self-knowledge guided retrieval augmentation for large language models. In *Findings of the association for computational linguistics: EMNLP 2023*, pp. 10303–10315, 2023b.

Di Wu, Hongwei Wang, Wenhao Yu, Yuwei Zhang, Kai-Wei Chang, and Dong Yu. Longmemeval: Benchmarking chat assistants on long-term interactive memory. *arXiv preprint arXiv:2410.10813*, 2024.

Derong Xu, Yi Wen, Pengyue Jia, Yingyi Zhang, Wenlin Zhang, Yichao Wang, Huifeng Guo, Ruiming Tang, Xiangyu Zhao, Enhong Chen, et al. From single to multi-granularity: Toward long-term memory association and selection of conversational agents. In *International Conference on Learning Representations*, volume 2026, pp. 91581–91617, 2026a.

Jing Xu, Arthur Szlam, and Jason Weston. Beyond goldfish memory: Long-term open-domain conversation. In *Proceedings of the 60th annual meeting of the association for computational linguistics (volume 1: long papers)*, pp. 5180–5197, 2022.

Wujiang Xu, Zujie Liang, Kai Mei, Hang Gao, Juntao Tan, and Yongfeng Zhang. A-mem: Agentic memory for llm agents. *Advances in Neural Information Processing Systems*, 38:17577–17604, 2026b.

Shi-Qi Yan, Jia-Chen Gu, Yun Zhu, and Zhen-Hua Ling. Corrective retrieval augmented generation. *arXiv preprint arXiv:2401.15884*, 2024.

Andrew P Yonelinas. The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language*, 46(3):441–517, 2002.

HongChien Yu, Chenyan Xiong, and Jamie Callan. Improving query representations for dense retrieval with pseudo relevance feedback. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pp. 3592–3596, 2021.

Yingyi Zhang, Junyi Li, Wenlin Zhang, Pengyue Jia, Xianneng Li, Yichao Wang, Derong Xu, Yi Wen, Huifeng Guo, Yong Liu, and Xiangyu Zhao. Evoking user memory: Personalizing LLM via recollection-familiarity adaptive retrieval. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=f7p0F2X6XN.

Wanjun Zhong, Lianghong Guo, Qiqi Gao, He Ye, and Yanlin Wang. Memorybank: Enhancing large language models with long-term memory. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 19724–19731, 2024.

A SUPPLEMENTARY STATEMENTS

This section collects formal statements and clarifications moved from the main text. Their proofs and all supporting lemmas are retained below.

A.1 RELATION TO THE ORIGINAL RF-MEM ANALYSIS

Appendix F of RF-Mem (Zhang et al., 2026) provides a concise theoretical rationale for its routing rule. The present paper builds on that starting point but asks a different question: which routing conclusions are identified by the information actually observed before Recollection is run? We record the relationship claim by claim to distinguish genuinely new results from classical ingredients and from corrections of the original argument.

Risk-minimizing selection in Appendix F.1. The original monotone-likelihood-ratio argument motivates Familiarity by showing that its miss probability improves under a coordinatewise increase of the probe scores. Coordinatewise order, however, is stronger than order by the mean alone, and entropy does not totally order score configurations. Moreover, the optimal action depends on the difference between Familiarity and Recollection risks, not on Familiarity risk in isolation. Finally, a monotone two-dimensional decision region is generally a down-set with a varying boundary, rather than an axis-aligned region described by one fixed entropy threshold. Remark 1 gives an explicit risk-level counterexample. Our frontier results therefore impose monotonicity directly on the conditional action-risk difference, characterize the resulting boundary, and quantify the regret of replacing it by a fixed threshold. This sharpens the logical conditions supporting the gate; it does not claim that the original heuristic is empirically ineffective.

Entropy certification in Appendix F.2. The original analysis invokes the maximum entropy attainable at a fixed maximum mass and attempts to invert that upper envelope in the low-entropy direction. That implication does not follow from an upper bound on entropy. In fact, the maximum-based inverse used there is identically one because $H_K^{\max}(1) = 0$; see Remark 2. A valid low-entropy certificate instead follows by inverting the minimum entropy attainable at fixed maximum mass. The underlying extremal entropy and majorization facts are classical. Our contribution is to put the correct envelope into the RF-Mem observation model and to show that concentration of softmax mass still does not identify retrieval correctness without a conditional calibration link.

Mean-score concentration in Appendix F.3. The sub-Gaussian sample-mean bound in the original analysis is valid under its stated assumption that the selected probe scores are independent. Under the natural model in which independent raw corpus scores are ranked and the largest K are retained, however, the selected order statistics are dependent and do not form an i.i.d. sample from the raw-score distribution. We therefore place independence on the raw scores and analyze their Top-K mean as a single dependent functional. Self-bounding concentration, the variational representation of an upper-tail average, and the DKW inequality are classical tools. The RF-Mem-specific conclusions are that the relevant population target is $U_{K/n}(F)$ rather than the ordinary mean and that a fixed-K gate can drift solely as the corpus grows.

Complexity analysis in Appendix F.4. Appendix F.4 accounts for the computational overhead of the retrieve–cluster–mix procedure. The present paper studies the information available to the initial routing decision. Theorem 5 adds an exact acquisition-cost tradeoff in an individual-inner-product observation model, not a wall-clock complexity or coverage bound for a deployed retriever.

Novelty accounting. The new constructions are complete-score action separation, worst-case necessity of a designated off-query entry, stagewise strictness of RF-Mem's observation reductions, the explicit single-beam Bag parameter extension, and the priced pre-routing observation family. The acquisition objective itself belongs to cost-sensitive active classification (Greiner et al., 2002); the new content is its Gram-realizable RF-Mem instance, the Bag success correspondence, and the exact individual-versus-aggregate observation consequences. Their exact risk gaps and query-cost tradeoff use the RF-Mem retrieval and output mechanism. The adaptive blocks identify entries accessed along a realized run; they are execution certificates rather than a general minimal routing statistic. Proposition 4 supplies a concrete conditional model and a learned-gate example. Corollary 5

adds an explicit perturbation radius and Bag-cardinality control without duplicate embeddings. The cache accounting and i.i.d. histogram guarantee are applications of classical cost comparison and concentration arguments to these specified families.

Gram factorization and its RF-Mem verification, binary value-of-information identities, the implicit-function frontier, fixed-axis regret conversion, and plug-in margin bounds are background lemmas with complete proofs in Appendix C. The first two connect to invariant theory and comparison of experiments; the fast exponent follows the classical plug-in argument (Villar et al., 2021; Tsybakov & Audibert, 2007). Entropy envelopes, calibration implications, CVaR, DKW, self-bounding concentration, and order-statistic results support the remaining conditional diagnostics. These tools are not claimed as new general principles.

A.2 Cost-sensitive active classification and the probe model

The acquisition problem is a specialization of cost-sensitive active classification (Greiner et al., 2002). To make the correspondence exact, use the variables in the construction of Theorem 5 and define binary attributes

$$X_j = \frac{\Gamma_{i_* j} - sa}{\zeta}, \qquad W = \max_{j \le m} X_j.$$

The common initial observations reveal none of the X_j. Requesting X_j is precisely one paid seed–candidate Gram query. Set the classification penalties to

$$e(\mathrm{F}, 1) = 1 - \Delta_c, \quad e(\mathrm{R}, 0) = \Delta_c, \quad e(\mathrm{F}, 0) = e(\mathrm{R}, 1) = 0.$$

Because the published Bag procedure gives failure losses $(1, 0)$ when $W = 1$ and $(1, 1)$ when $W = 0$, every acquisition-and-routing policy satisfies

$$\mathcal{R}(\pi) + c_p \mathbb{E}Q - \mathcal{R}^\star_{\mathcal{G}} = \mathbb{E}\big[c_p Q + e(\widehat{a}, W)\big],$$

where $\widehat{a}$ is its final action. This follows pointwise by subtracting the better action's loss in each environment; the subtracted expectation is independent of the acquisition policy. Thus paying for attributes before classification is not a new formulation here.

The RF-Mem content is the joint realization: explicit unit vectors make all initial scores and the full initial Top-K Gram block identical, varying only the marked off-query pair; the mark determines success under the published Bag and final-output rules; and the resulting experiment admits the exact risk and cost expressions in the theorem. Its stopping calculation is an elementary solution of this special active-classification problem, not a new general stopping or learning principle. Corollary 8 further separates the decision target W from the retrieved identity J: a cached aggregate can identify W without identifying J, but its gap is ζ/m rather than the individual-entry gap ζ and forming an uncached centroid incurs additional cost.

Which assumptions determine the exact formula? The following distinctions concern (10), not generic RF-Mem routing.

Condition	Role and scope
Known candidate pool and common initial observations	Fix the m possible informative entries. The formula does not charge for finding that pool or supplying its identities.
Known fair mark prior and uniform marked location	Give unmarked-history probabilities $(m - t)/(2m)$ and $1/2$, and expected request count $t - t(t-1)/(4m)$. Unknown priors require a different learning or robust-decision problem.
At most one mark, with exact individual-entry responses	Make the first mark decisive and leave a single all-unmarked continuation path. Multiple marks or noisy responses can change the posterior and stopping calculation.
Common nonnegative per-entry cost and fixed action gap	Give the objective $c_p \mathbb{E}Q$ and fixed error penalties; their concave cutoff objective yields an endpoint optimum. Unequal costs need a different ordering and calculation.
Individual Gram-entry interface	Excludes free aggregate or coordinate-based tests. The cached centroid in Corollary 8 is a different observation channel.
Dimension $m + 3$	Certifies the stated geometric realization, not the cost calculation itself. No fixed-dimension family for unbounded m or corresponding fixed-dimension lower bound is established.

The fair prior fixes constants rather than being necessary for an endpoint solution. Indeed, keep the location uniform but let $\mathbb{P}(W = 1) = \vartheta \in (0, 1)$ be known. After t unmarked tests, the unconditional masses of the two unresolved cases are $\vartheta(1 - t/m)$ and $1 - \vartheta$. Therefore

$$r_t(\vartheta) = \min\{(1 - \vartheta)\Delta_c, \vartheta(1 - t/m)(1 - \Delta_c)\}, \qquad \mathbb{E}Q_t = t - \frac{\vartheta t(t - 1)}{2m}.$$

The first expression compares the two possible terminal errors. For the second, average t requests when $W = 0$ with $t - t(t - 1)/(2m)$ when $W = 1$. Their cost-weighted sum is again concave in $t \in [0, m]$, so its minimum is

$$\min\left\{r_0(\vartheta),\ c_p\left(m - \frac{\vartheta(m - 1)}{2}\right)\right\}.$$

This proves directly how the fair-prior formula changes; setting $\vartheta = 1/2$ recovers the theorem.

Uniform location is more consequential for the endpoint policy. For a counterexample to extending that policy unchanged, retain $m = 2$, a fair W, and the same Gram construction, but take $\mathbb{P}(J = 1 \mid W = 1) = 9/10$, $\mathbb{P}(J = 2 \mid W = 1) = 1/10$, $\Delta_c = 1/2$, and $c_p = 1/10$. No acquisition has excess risk $1/4$. Testing only candidate 1, then choosing Recollection on a mark and Familiarity otherwise, has total excess risk

$$\frac{1}{10} + \frac{1}{2} \cdot \frac{1}{10} \cdot \frac{1}{2} = \frac{1}{8}.$$

Testing until a mark or exhaustion has zero routing regret, but even its best order, candidate 1 first, costs

$$\frac{1}{10}\left[\frac{1}{2} \cdot 2 + \frac{1}{2}\left(\frac{9}{10} \cdot 1 + \frac{1}{10} \cdot 2\right)\right] = \frac{31}{200} > \frac{1}{8}.$$

Putting candidate 2 first only increases the expected count. Thus a proper partial search beats both endpoint policies in this nonuniform experiment. The main theorem's numerical formula and its optimal endpoint description are restricted to its specified prior and interface.

A.3 Published pseudocode and the implementation class

We use Algorithms 1–3 of RF-Mem (Zhang et al., 2026), with embeddings fixed, $x_t = q$ unit-normalized, exact inner-product search, and fixed index and tie orders. K-means is realized by distance-based initialization, assignment, empty-cluster handling, and termination, conditionally on a fixed seed if randomized. Every normalization must be nonzero. These are the hypotheses of Lemma 2 and Theorem 3, not guarantees for arbitrary clustering or retrieval backends. Write $[n] = \{1, \ldots, n\}$ and use the coefficient notation of Section 4.

Published operation	Gram implementation and required information
Algorithms 1–2: probe, mean, softmax entropy, and Familiarity	Initial scores are Γ_{i0}; Top-K, a fixed score floor, and the gate are deterministic functions of these indexed scores.
Algorithm 3, lines 1–3: initialize and set $N = (B + r)F$	$a = e_0$; Beam, Seen, Bag, round counters, and widths are discrete state. No off-query entries are required.
Lines 5–8: full-corpus Top-N, K-means, normalized centroids	$E = [n]$ and $I = \mathrm{Top}_N(Wa; \mathcal{M})$. $\Gamma_{E,J(a)}$ determines retrieval; $\Gamma_{I,I}$ determines clustering and normalized centroid coefficients.
Lines 9–10: construct every candidate child	$h_C = \alpha a + (1 - \alpha)\widetilde{c}_C + e_0$ and $a'_C = h_C/\sqrt{h_C^\top \Gamma h_C}$, using Γ_{J_C,J_C}.
Lines 13–15: empty-next stopping, cluster scores, and top-B beam	Score every child by $\sum_{i \in C} e_i^\top \Gamma a'_C$. Γ_{C,J_C} suffices, including for children later pruned.
Lines 16–21: insert retained clusters into Bag	If $i \notin$ Seen, store $(i, e_i^\top \Gamma a'_C)$ and mark i seen. Repeated hits do not replace the stored score.
Lines 22–26: budget stopping and final Top-K from Bag	Use Bag cardinality and a fixed ordering of stored scores and indices. No new Gram entries are required; fewer than K entries are all returned.

For completeness, the raw centroid coefficients are $c_C = |C|^{-1} \sum_{i \in C} e_i$, and $\widetilde{c}_C = c_C/\sqrt{c_C^\top \Gamma c_C}$. All point-to-centroid distances used by K-means have the form $(e_i - c_C)^\top \Gamma (e_i - c_C)$ and use only $\Gamma_{I,I}$. The displayed formula for a'_C implements the published query update exactly, and its

individual scores give both the published cluster sum and each Bag record. Given the previous state, every row of the table therefore determines the next state from the listed blocks. Induction over the finite rounds proves membership in the implementation class and gives the partial-Gram certificate. The same-round choice $Q_C = C$ suffices; retrieving under a surviving child at the next round requires $\Gamma_{[n],J(a'_C)}$ as that round's incoming block. This is an exact-arithmetic trace statement, not a claim of bitwise invariance under floating-point changes of coordinates.

Bag and output semantics. The published Top-N call searches the full corpus: Seen is tested at insertion, not before retrieval. Only clusters surviving beam selection enter Bag. An item's first stored score is computed with the child query, not the parent retrieval query; it is not refreshed on later hits. Final Top-K is taken from these records, so $|\mathcal{R}_{\mathrm{R}}| \le K$ and need not equal K. Any ambiguity in the last-line ordering or in tied insertion order is resolved by the fixed implementation choices above. Consequently, the seed-excluding one-expansion definition in the main text is not literally Algorithm 3. Lemma 3 identifies a published parameter setting that preserves its separation losses without asserting equality of the full returned sets.

Gate and backend scope. Algorithm 1 calls its probe with K despite listing k_p among its inputs; our S follows that call (equivalently, $k_p = K$). Its first Familiarity test uses an OR, giving Familiarity at simultaneously low mean and low entropy, whereas the low-mean branch of the paper's Equation (3) gives Recollection. Both are T-measurable gates; we do not identify them with one another or with the Bayes comparator analyzed here. The operation table certifies the published pseudocode, not every released code path. For example, the released PersonaBench `mean_group` backend uses MMR-based cumulative beam scores, epsilon-regularized normalization, and an insertion-prefix return. These are not the literal score, normalization, and output rules fixed above; a claim about that backend requires a separate implementation contract. Our loss-transfer lemma and new Bag and probe constructions concern Algorithm 3 only.

A.4 Parameter scope of published Bag separation

Here R is a round limit, B, F are positive integers, and retrieval widths follow $N_r = (B + r)F$. We retain full-corpus search, no score floor, insertion-time Seen checks, and final Top-K by first-insertion scores. Theorem 2 uses indexed copies to obtain its linear bound on cumulative Bag size. Corollary 5 proves stability under small score-matched perturbations with distinct embeddings, using $K \ge F[R(R+1)/2-1]+1$. This retains the full two-round range and strengthens the budget for $R > 2$; it does not prove robustness throughout the original linear multiround range or a sharp converse.

Parameters	Established conclusion
$B = 1$, $R = 2$, $K \ge 2F + 1$	Opposite optimal actions, any integer $F \ge 1$, with $n = K + 1$; stable under small score-matched perturbations that remove duplicate embeddings.
$B = 1$, $R \ge 2$, $K \ge (R + 1)F + 1$	The same separation, with later Bag growth and truncation explicitly controlled.
$B = 1$, $R > 2$, $K \ge F[R(R + 1)/2 - 1] + 1$	Perturbation-stable separation with distinct embeddings and all R rounds executed.
$R = 1$, $BF \le K$	Recollection's output is contained in Familiarity's; a positive cost gap precludes an action reversal.
$\alpha = 1$, all executed $N_r \le K$	All queries remain q; Recollection's output is contained in Familiarity's.
$K \ge n$, or $\Delta_c \ge 1$	Recollection cannot be uniquely optimal for the stipulated binary retrieval-success loss.
Other $B = 1$ ranges; $B > 1$ with competing children	Not settled by these constructions. The general Gram trace result does not establish action reversal in these regimes.

Proposition 3 (Degenerate regimes for Bag action reversal)**.** *Let y be a unique relevant index and $L_a = \mathbf{1}\{y \notin \mathcal{R}_a\} + \beta C_a$, with fixed $C_{\mathrm{R}} > C_{\mathrm{F}}$ and $\beta > 0$. Under the published contract above, Familiarity strictly dominates Recollection whenever $R = 1$ and $BF \le K$, or whenever $\alpha = 1$ and every executed width is at most K, or whenever $K \ge n$. If $\Delta_c \ge 1$, Familiarity weakly dominates Recollection for any parameters for which execution is defined.*

Proof. With $R = 1$, the only parent is q, and all inserted indices lie in $\mathrm{Top}_{BF}(q)$, which is contained in $\mathrm{Top}_K(q)$. Clustering, pruning, and final selection cannot add an index outside that set. With $\alpha = 1$, induction gives $x = q$ at every round, since the update is $\mathrm{norm}(x+q)$. Every retrieved index then belongs to the fixed initial Top-K whenever $N_r \le K$. Bag and its final output remain subsets of that Top-K. In either case, $\ell_{\mathrm{F}} \le \ell_{\mathrm{R}}$ pointwise, and the positive cost gap makes the dominance strict. If $K \ge n$, Familiarity contains y, so the same conclusion holds. Finally, $L_{\mathrm{F}} - L_{\mathrm{R}} = \ell_{\mathrm{F}} - \ell_{\mathrm{R}} - \Delta_c \le 1 - \Delta_c$, which proves the last assertion, including possible ties when $\Delta_c = 1$. □

If one relaxes the published annotation $N < K$ to $N \le K$, the same construction also covers $B = 1$, $K = 2F$, and any $R \ge 2$: round one fills Bag to K and stops, retaining y only in the plus corpus. This boundary case is outside the strict $N < K$ contract used in the main theorem. Changing K in these new families does not transfer the summary-matching construction of Theorem 4; its published Bag transfer remains the separately proved $K = 3$ case.

Why cloning does not prove a multibeam extension. There is a concrete failure of the naive padding argument. Take $\alpha = 0$, $B = 2$, $F = 1$, $R = 2$, $K = 4$, $q = e_1$, and two indexed copies of $z = 0.6e_1 + 0.8e_2$. Add $w_1 = 0.4e_1 + \sqrt{0.84}e_3$, $w_2 = 0.3e_1 + \sqrt{0.91}e_3$, and $y = 0.2e_1 + \sqrt{0.96}e_2$, the unique relevant index. Use distance-based global minimum-SSE K-means with nonempty clusters and fixed ties, a member of the stated implementation class. At round zero the two seed copies form singleton clusters and generate two identical children. At their query $q' = \mathrm{norm}(q + z)$, Top-3 consists of the two seeds and y. The unique minimum-SSE partition into two nonempty clusters, up to cluster order, is the duplicate seed cluster and the singleton target cluster: it has zero SSE, whereas every other partition mixes distinct vectors and has positive SSE. The two seed-cluster proposals, one from each parent, each score $2\sqrt{(1 + 0.6)/2}$; the two target-cluster proposals each score $\sqrt{(1 + 0.2)/2}$. Top-B therefore keeps only seed proposals. The target is retrieved but never inserted into Bag, and both actions miss it. Thus even favorable expanded retrieval does not prove a Bag success when multiple proposals compete. This is a counterexample to cloning as a proof method, not an impossibility theorem for $B > 1$.

A.5 A STABLE BAG MODEL WITH LEARNED PRE-ROUTING ADVANTAGE

Fixed retrieval membership and a smooth relevance model supply a concrete route to regular advantage. In general, if a unique relevant index has conditional probabilities $p_i(G)$ and both returned sets $A_{\mathrm{F}}, A_{\mathrm{R}}$ are constant on a domain, then

$$d_\Gamma(G) = \sum_{i \in A_{\mathrm{R}}} p_i(G) - \sum_{i \in A_{\mathrm{F}}} p_i(G) - \Delta_c.$$

The following example verifies the required membership stability for RF-Mem itself and permits both optimal actions within the same stable domain.

Proposition 4 (Stable Bag retrieval with logistic relevance)**.** *For every $\alpha \in [0, 1)$ and $\Delta_c \in (0, 1)$ there is a compact continuous family $\mathfrak{C} = \{G(t) : t \in [u_-, u_+]\}$ of four-memory unit-vector Gram matrices, with $u_- < u_+$, on which published Algorithm 3 with $B = F = 1$, $R = 2$, and $K = 3$ returns the fixed set $\{z, y\}$ while Familiarity returns $\{z, w_1, w_2\}$. The parameter is the single entry $t = \langle z, y\rangle$. Initial and expanded retrieval-membership gaps have positive uniform lower bounds, raw centroid norms have a positive uniform lower bound, and query-update numerator norms are at least one. There is one cluster per round, so clustering and beam selection introduce no additional branch instability.*

Let $\lambda_{\mathrm{rel}} > 0$ and specify the unique relevant index Y by

$$\mathbb{P}(Y = y \mid G(t)) = p(t) := \frac{1}{1 + e^{-\lambda_{\mathrm{rel}}(t - b_{\mathrm{rel}})}}, \qquad \mathbb{P}(Y = w_1 \mid G(t)) = 1 - p(t).$$

For $L_a = \mathbf{1}\{Y \notin \mathcal{R}_a\} + \beta C_a$,

$$d_\Gamma(G(t)) = 2p(t) - 1 - \Delta_c, \qquad L_\Gamma = \lambda_{\mathrm{rel}}/2 \tag{12}$$

is a valid Lipschitz constant in $\|\cdot\|_{\max}$. The intercept b_{rel} can be chosen so that the advantage has a unique zero $t_* \in (u_-, u_+)$, with both actions optimal on sets of positive probability when t is uniform. More generally, if t has density bounded by $f_{\max}$, then Assumption 2 holds with

$$\kappa = 1, \quad B = 1 + \Delta_c, \quad C_{\mathrm{mar}} = 2f_{\max}/\mu_*, \qquad \mu_* := \min_{t \in [u_-, u_+]} 2\lambda_{\mathrm{rel}} p(t)(1 - p(t)) > 0.$$

A learned gate is available under the following explicit data design. At each of $J_g \geq 2$ equally spaced points t_j of the interval, observe N_g independent relevance labels from this model, independently of the test instance, and estimate $p(t_j)$ by the fraction of labels equal to y. Let $N = J_g N_g$ be the total number of labels and let $h_g = (u_+ - u_-)/(J_g - 1)$. At test time obtain a scalar probe $\widehat{t}$ with almost-sure error at most ε_t, using randomness conditionally independent of Y given G, clip it to the interval, and use the nearest grid estimate $\widehat{p}$ to route by $\mathbf{1}\{2\widehat{p} - 1 - \Delta_c > 0\}$. For every $\delta \in (0,1)$ its risk, averaged over training and test randomness, satisfies

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^\star_{\mathcal{K}_\Gamma} \leq C_{\mathrm{mar}} \left(\frac{\lambda_{\mathrm{rel}}}{2}\varepsilon_t + \frac{\lambda_{\mathrm{rel}}}{4} h_g + \sqrt{\frac{2J_g \log(2J_g/\delta)}{N}} \right)^2 + (1 + \Delta_c)\delta. \tag{13}$$

This is a labeled-data guarantee for the specified family, not identification from arbitrary deterministic routing logs.

Proof. Choose $0 < a < s < 1$ and $0 < \rho_- < \rho_+ < 1$. Put $A = 1 + \alpha + (1-\alpha)s$, $b = (1-\alpha)\sqrt{1-s^2}$, and $D = \sqrt{A^2 + b^2}$. Choose

$$a < \eta_2 < \eta_1 < \min\{s, a + (b/A)\sqrt{1-a^2}\rho_-\}.$$

In dimension four let

$$q = e_1, \quad z = se_1 + \sqrt{1-s^2}e_2, \quad w_j = \eta_j e_1 + \sqrt{1-\eta_j^2}e_3,$$

and let

$$y(\rho) = ae_1 + \sqrt{1-a^2}\big(\rho e_2 + \sqrt{1-\rho^2}e_4\big), \qquad \rho \in [\rho_-, \rho_+].$$

The initial scores $s > \eta_1 > \eta_2 > a$ are fixed. Familiarity returns $\{z, w_1, w_2\}$ and the initial Recollection seed is z. Under $q' = (Ae_1 + be_2)/D$, the seed–distractor score gap is at least $[A(s - \eta_1) + b\sqrt{1-s^2}]/D > 0$ and the target–distractor gap is at least $[A(a - \eta_1) + b\sqrt{1-a^2}\rho_-]/D > 0$. Thus Top-2 is $\{z, y\}$ throughout the interval. Bag returns that same set, with no truncation, irrespective of the order of the two scores within Top-2. Every memory has projection at least a onto q. Every raw centroid therefore has norm at least a and positive q projection. Induction gives positive q projection for every query, so an update numerator has q projection at least one and consequently norm at least one. These verify the asserted margins and normalization bounds.

The parameter $t = sa + \sqrt{1-s^2}\sqrt{1-a^2}\rho$ maps the compact ρ interval affinely onto $[u_-, u_+]$. Only the symmetric seed–target pair varies in its Gram matrix: target–distractor inner products are $a\eta_j$, and all other pairs are fixed. Thus $\|G(t) - G(t')\|_{\max} = |t - t'|$.

If $Y = y$, only Recollection succeeds; if $Y = w_1$, only Familiarity succeeds. This gives $d_\Gamma(G(t)) = p(t) - (1 - p(t)) - \Delta_c$. Its derivative is $2\lambda_{\mathrm{rel}} p(t)(1 - p(t))$, which lies in $[\mu_*, \lambda_{\mathrm{rel}}/2]$ on the compact interval. Integration proves the Lipschitz bound. For any prescribed $t_* \in (u_-, u_+)$, choose

$$b_{\mathrm{rel}} = t_* - \lambda_{\mathrm{rel}}^{-1} \log \frac{1 + \Delta_c}{1 - \Delta_c}.$$

Then $p(t_*) = (1 + \Delta_c)/2$ and strict monotonicity gives the stated crossing. For any $v > 0$, the preimage of $[-v, v]$ under $d_\Gamma(G(t))$ is an interval of length at most $2v/\mu_*$, by the lower derivative bound. A density bound therefore yields $\mathbb{P}(0 < |d_\Gamma(G(t))| \leq v) \leq 2f_{\max} v/\mu_*$. The zero set has probability zero under a density, so the bound also holds at $v = 0$. Since $|2p - 1 - \Delta_c| \leq 1 + \Delta_c$, all asserted regularity constants follow.

For the learned rule, Hoeffding's inequality and a union bound give, with training probability at least $1 - \delta$,

$$\max_{j \leq J_g} |\widehat{p}_j - p(t_j)| \leq \sqrt{\frac{\log(2J_g/\delta)}{2N_g}}.$$

Clipping cannot increase the scalar error, and the nearest grid point is at distance at most $h_g/2$ from the clipped probe. Since p has Lipschitz constant $\lambda_{\mathrm{rel}}/4$, on this event the estimated advantage differs from the true advantage by at most

$$e_* := \frac{\lambda_{\mathrm{rel}}}{2}\varepsilon_t + \frac{\lambda_{\mathrm{rel}}}{4} h_g + \sqrt{\frac{2\log(2J_g/\delta)}{N_g}}.$$

A wrong sign at a point with nonzero advantage implies $0 < |d_\Gamma| \le e_*$. Conditional independence of training and probe randomness from the test label given G makes the conditional excess risk equal to $|d_\Gamma|$ on disagreement. On the good training event its expectation is bounded by $e_*\mathbb{P}(0 < |d_\Gamma| \le e_*) \le C_{\mathrm{mar}} e_*^2$. On the complement it is at most $1 + \Delta_c$. Averaging proves (13). Here $N = J_g N_g$ rewrites the estimation term without changing it. □

Corollary 7 replaces the designed grid labels by i.i.d. labeled instances, under a positive lower density bound for scalar-feature coverage. Its proof treats random bin counts explicitly. This extension changes the sampling design, while retaining the stipulated logistic relevance model and stable retrieval membership.

A.6 Jointly moving the mean thresholds can compensate

The bounds in Lemma 26 fix the mean thresholds at the Bayes-axis values. They do not give a lower bound over all jointly tuned triples. The following conditional-risk example makes the distinction quantitative; it does not assert realization by an RF-Mem score distribution.

Let M and H be independent, with $M \sim \mathrm{Unif}[0, 1/2]$ and $H \sim \mathrm{Unif}[1/2, 1]$, and set $d(m, h) = 2h - 2m - 1$ on $\mathcal{Z} = [-1/8, 5/8] \times [1/2, 1]$. For example, conditional risks $r_{\mathrm{F}} = (2 + d)/2$ and $r_{\mathrm{R}} = (2 - d)/2$ are nonnegative and realize this difference. Both derivative bounds hold with constants two, and the two required corner values are $1/4$ and $-1/4$. The Bayes-axis thresholds are 0 and $1/2$, and $b(m) = m + 1/2$.

Write $h = H - 1/2$ and $u = \tau - 1/2$, so (M, h) has joint density four on $[0, 1/2]^2$. A gate with fixed mean axes $0, 1/2$ disagrees with the Bayes rule between $h = M$ and $h = u$. Its excess risk is

$$4 \int_0^{1/2} \int_{\min(m,u)}^{\max(m,u)} 2|h - m| \, dh \, dm = 4 \int_0^{1/2} (m - u)^2 \, dm.$$

The minimum occurs at $u = 1/4$, or $\tau = 3/4$, and equals $1/24$. Now keep this entropy threshold but use mean cutoffs $\ell = 1/8$ and $v = 3/8$, with Recollection below ℓ, Familiarity above v, and the entropy test in between. Its excess risk is

$$4 \left[\int_0^{1/8} m^2 \, dm + \int_{1/8}^{3/8} (m - 1/4)^2 \, dm + \int_{3/8}^{1/2} (1/2 - m)^2 \, dm \right] = \frac{1}{96}.$$

The outer integrals are the Bayes-weighted losses from choosing a constant action there; the middle integral is the remaining entropy-threshold loss. Thus changing the mean axes reduces regret, despite moving away from their Bayes-axis values. This does not contradict the fixed-axis bounds, and precludes treating all mean-threshold displacement as an additional nonnegative term relative to their oracle-axis comparator.

A.7 Claim scope, assumptions, and empirical estimability

The results have different logical scopes. The assumptions used for one characterization should not be read as assumptions for the entire paper. We record the separation explicitly.

Distribution-free structural results. Proposition 1 is an algebraic identity for a well-defined one-step update. Theorem 1 is a finite explicit construction and requires no population model for memory corpora. Lemma 1 records the general inner-product factorization principle, while Lemma 2 verifies it for an exact finite RF-Mem run under the stated implementation and nonzero-normalization conditions. Its conclusion concerns the execution trace, not arbitrary relevance labels or semantic losses. The adaptive-block and worst-case claims of Theorem 3 are likewise finite algebraic statements. The listed blocks are sufficient entrywise certificates along a realized path, not a claim of unique minimality among arbitrary nonlinear statistics. Proposition 2 is a finite construction and requires no distributional regularity. Lemma 5 and Corollaries 1 and 2 require only integrable potential losses and nested information sets. Theorem 4 is a finite construction using the stated one-expansion retrieval-success loss. Lemma 3 establishes the same losses for published Algorithm 3 at $B = F = 1$, $R = 2$, and $K = 3$, with a common output cap rather than equal returned cardinalities across actions. Theorem 2 separately extends complete-score separation to its stated single-beam

ranges. Corollary 5 preserves the losses under small score-matched perturbations at its stated budget, with an explicit radius and no need for duplicate embeddings. These results do not generalize all stagewise witnesses or settle competing beams. Theorem 5 fixes a prior and a priced entry-query interface. Its lower bounds are information acquisition bounds in that interface, not lower bounds for all coordinate-based computations or aggregate observations. Corollary 8 analyzes a separately supplied candidate-centroid channel, including its precision requirement and the cost of forming an uncached centroid. Corollary 9 realizes a reusable candidate pool and adds setup, refresh, update, online, and cache error accounting. Its break-even condition compares specified policies under fresh conditional ambiguity; neither useful-pool discovery nor a latency benefit in deployment is established. None of these results asserts separation at every budget or backend. Likewise, the entropy envelopes and their low-entropy certificate use only the simplex constraint. These statements are unaffected by the monotonicity, density, calibration, and random-score assumptions introduced elsewhere. For bounded cosine logits, Proposition 5 identifies the smaller softmax-feasible subset: the simplex certificate remains valid on that subset, while simplex-level sharpness need not be attainable.

Conditional characterizations. The positive approximate-Gram rate is conditional on Assumption 2. Lipschitz continuity links matrix error to conditional-advantage error, the margin-mass bound converts sign instability into the fast exponent $\kappa+1$, and conditional independence prevents the Gram estimator's auxiliary randomness from carrying additional outcome information. The margin-free construction shows that hard retrieval losses need not satisfy this regularity near ranking ties. The gate in Lemma 7 uses the true advantage and provides stability, not estimation. Proposition 4 verifies positive retrieval gaps and normalization bounds in a concrete family, derives advantage regularity from its relevance probabilities, and adds estimation error under independent labeled sampling. Corollary 7 permits ordinary i.i.d. labels with measured scalar features and a density lower bound for bin coverage. This is an existence example within a specified relevance model; it neither learns that model from censored action logs nor certifies arbitrary K-means or beam boundaries.

The mean–entropy frontier is a conditional description of the Bayes rule. Weak monotonicity of the risk difference makes each Familiarity-optimal vertical section a lower interval. The strict derivative bounds in Assumption 3 add an interior crossing, uniqueness, differentiability, and quantitative slope control. The conditional density assumption is used only to express fixed-threshold regret through the support-weighted discrepancy in Lemma 26. Density bounds are needed only on a threshold-specific disagreement segment to obtain the two-sided quadratic specialization; no common rectangular conditional support is assumed. The mean cutoffs remain fixed throughout these bounds; they do not apply as lower bounds over jointly tuned triples. Appendix A.6 demonstrates compensation by moving the mean cutoffs. None of these conditions is asserted to hold in every deployment, and the logistic Bag example does not establish mean–entropy monotonicity.

The entropy certificate concerns concentration of the softmax weights without a correctness assumption. Assumption 5 is introduced only when translating that concentration into a Familiarity success guarantee. Lemma 13 shows why some such link is necessary. For the probe mean, bounded independent scores are sufficient for concentration around its finite-sample expectation, whereas the population upper-tail target and its asymptotic interpretations use the stated bounded i.i.d. baseline. Dependence or heterogeneity may change the population conclusions, but not the deterministic variational identity.

Population quantities versus empirical estimation. The information-ladder terms are population Bayes-risk gaps. Their definition does not imply identification from logs produced by a deterministic deployed router: such logs may never reveal the loss of the unchosen action. Estimating the terms therefore requires action coverage, such as randomized routing with known propensities, or another design that identifies both conditional action risks. This observability requirement concerns estimation of the magnitude of each term; it does not alter the decomposition identity.

The theory nevertheless specifies deployment-level diagnostics. One may examine whether score-matched cases have different cluster geometry or retrieval outcomes, estimate whether the conditional decision boundary bends across the mean–entropy plane, evaluate correctness calibration after entropy-based selection, and track the action frequency of a fixed gate as the corpus grows. Such measurements determine how prevalent the identified mechanisms are in a particular environment.

They are complementary to, rather than substitutes for, the structural identifiability and certification results.

A.8 Weak monotonicity and the extended routing frontier

The shape of the Bayes routing region is determined by the risk difference $d = r_{\mathrm{F}} - r_{\mathrm{R}}$, not by r_{F} alone.

Remark 1 (Monotonicity must concern the risk difference)**.** *For example, on $[0,1]^2$, let $r_{\mathrm{F}}(m,h) = 2 - m + h$ and $r_{\mathrm{R}}(m,h) = 2 - 2m + 2h$. Then $\partial_m r_{\mathrm{F}} = -1 < 0$ and $\partial_h r_{\mathrm{F}} = 1 > 0$, so Familiarity risk decreases with mean similarity and increases with entropy. Nevertheless,*

$$d(m,h) = (2 - m + h) - (2 - 2m + 2h) = m - h,$$

and therefore $\partial_m d = 1 > 0$ and $\partial_h d = -1 < 0$. Thus monotonicity of Familiarity risk alone does not imply the required monotonicity of the routing decision.

Even if Familiarity improves as mean similarity increases, Recollection may improve more quickly. Shape assumptions for the Bayes router must therefore be imposed directly on d.

Assumption 1 (Monotone conditional routing advantage)**.** *The conditional loss difference $d : \mathcal{Z} \to \mathbb{R}$ is continuous and satisfies: i) for every $h \in [h_-, h_+]$ and $m_1 \le m_2$, $d(m_2,h) \le d(m_1,h)$; ii) for every $m \in [m_-, m_+]$ and $h_1 \le h_2$, $d(m,h_1) \le d(m,h_2)$.*

The first condition formalizes that increasing mean similarity cannot increase the relative advantage of Recollection. The second formalizes that increasing entropy cannot make Familiarity more favorable relative to Recollection.

For every fixed mean score m, define the Familiarity-optimal vertical section $\mathcal{A}(m) := \{h \in [h_-, h_+] : d(m,h) \le 0\}$. Define the extended boundary

$$b^\star(m) := \begin{cases} -\infty, & \mathcal{A}(m) = \varnothing, \\ \max \mathcal{A}(m), & \mathcal{A}(m) \neq \varnothing. \end{cases}$$

If $\mathcal{A}(m) = [h_-, h_+]$, then $b^\star(m) = h_+$; the value $-\infty$ means that no admissible entropy makes Familiarity optimal.

Under Assumption 1, Lemma 8 shows that every nonempty $\mathcal{A}(m)$ is a lower interval and that $b^\star$ is nondecreasing. The stronger assumptions in the main text add an interior crossing, uniqueness, differentiability, and quantitative slope control.

A.9 Full entropy envelopes

For a vector $a \in \mathbb{R}^K$, let $a_1^\downarrow \ge \cdots \ge a_K^\downarrow$ denote its coordinates rearranged in nonincreasing order. For two vectors $a, b \in \mathbb{R}^K$ with the same total sum, we say that a majorizes b, and write $a \succeq b$, if $\sum_{j=1}^m a_j^\downarrow \ge \sum_{j=1}^m b_j^\downarrow$ for every $m = 1, \dots, K-1$. Shannon entropy is Schur concave by Lemma 10.

Recall from the main text that, for $x \in [1/K, 1]$, $r_x = \lfloor 1/x \rfloor$ and $u_x = 1 - r_x x$. Since $r_x \le 1/x < r_x + 1$, we have $u_x \ge 0$; if $u_x > 0$, then $u_x < x$.

Define the maximally concentrated vector $p^{\min}(x) \in \Delta_K$ as follows. If $u_x = 0$, let

$$p^{\min}(x) = \left(\underbrace{x, \dots, x}_{r_x \text{ entries}}, \underbrace{0, \dots, 0}_{K - r_x \text{ entries}} \right).$$

If $u_x > 0$, let

$$p^{\min}(x) = \left(\underbrace{x, \dots, x}_{r_x \text{ entries}}, u_x, \underbrace{0, \dots, 0}_{K - r_x - 1 \text{ entries}} \right).$$

In either case, $r_x x + u_x = 1$ and $\max_j p_j^{\min}(x) = x$. Its entropy is $H_K^{\min}(x)$.

At the opposite extreme, define

$$p^{\max}(x) := \left(x, \frac{1-x}{K-1}, \ldots, \frac{1-x}{K-1}\right).$$

Because $x \geq 1/K$, $(1-x)/(K-1) \leq x$, and hence $\max_j p_j^{\max}(x) = x$. Define

$$H_K^{\max}(x) := h_2(x) + (1-x)\log(K-1),$$

where

$$h_2(x) := -x\log x - (1-x)\log(1-x)$$

is the binary entropy.

Theorem 6 (Sharp entropy envelopes at fixed maximum mass)**.** *Let $p \in \Delta_K$ and set $x := p_{\max}$. Then $H_K^{\min}(x) \leq H(p) \leq H_K^{\max}(x)$. Both bounds are sharp. The lower bound is attained by $p^{\min}(x)$ and the upper bound by $p^{\max}(x)$, up to permutations of their coordinates.*

The result depends only on the simplex constraint. The next proposition records the smaller feasible set induced by bounded cosine scores and finite softmax temperature.

Proposition 5 (Feasible softmax set under bounded cosine scores)**.** *Fix $\lambda > 0$ and define*

$$\Delta_{K,\lambda}^{\cos} := \left\{ p \in (0,1)^K : \sum_{j=1}^K p_j = 1,\ \max_j \log p_j - \min_j \log p_j \leq 2\lambda \right\}. \tag{14}$$

Then $p \in \Delta_{K,\lambda}^{\cos}$ if and only if there exists $s = (s_1, \ldots, s_K) \in [-1,1]^K$ such that

$$p_j = \frac{\exp(\lambda s_j)}{\sum_{\ell=1}^K \exp(\lambda s_\ell)}, \qquad j = 1, \ldots, K. \tag{15}$$

Every feasible coordinate consequently satisfies

$$\frac{1}{1+(K-1)e^{2\lambda}} \leq p_j \leq \frac{1}{1+(K-1)e^{-2\lambda}}. \tag{16}$$

Hence the simplex envelopes remain valid for RF-Mem, but their simplex extremizers are attainable only when they satisfy (14)*; in particular, an extremizer with a zero coordinate is not attainable at finite λ.*

Thus a single entropy value generally does not identify $p_{\max}$ because the remaining mass can be arranged in different ways. On $\Delta_{K,\lambda}^{\cos}$, the simplex certificate is universally valid but may be conservative; deriving a restricted envelope is a distinct optimization problem.

Lemma 12 immediately yields the following threshold-design corollary.

Corollary 4 (Largest entropy threshold guaranteeing a desired maximum mass)**.** *Fix $\rho \in [1/K, 1]$. Then $H(p) \leq H_K^{\min}(\rho)$ implies $p_{\max} \geq \rho$. Moreover, $H_K^{\min}(\rho)$ is the largest universal threshold with this property.*

Thus an observed threshold τ yields the mass guarantee $\psi_K(\tau)$, while a desired mass ρ yields the largest valid threshold $H_K^{\min}(\rho)$. The simpler bound in Corollary 14 replaces $\psi_K(\tau)$ by $e^{-\tau}$ but is generally weaker.

The maximal envelope has the opposite logical direction.

Proposition 6 (The maximal envelope certifies high entropy, not low entropy)**.** *The function $H_K^{\max} : [1/K, 1] \to [0, \log K]$ is continuous and strictly decreasing. Let $\chi_K(\tau) := (H_K^{\max})^{-1}(\tau)$. Then $H(p) \geq \tau$ implies $p_{\max} \leq \chi_K(\tau)$.*

Remark 2 (Why the maximal envelope cannot be inverted in the low-entropy direction)**.** *Consider*

$$\phi_K(\tau) := \max\left\{x \in [1/K, 1] : H_K^{\max}(x) \leq \tau\right\}.$$

For every $\tau \geq 0$, $H_K^{\max}(1) = 0 \leq \tau$. Hence 1 belongs to the set, and $\phi_K(\tau) = 1$. This inverse therefore cannot yield a nontrivial low-entropy lower bound on $p_{\max}$. More generally, $H(p) \leq H_K^{\max}(p_{\max})$ is an upper bound on $H(p)$, so knowing $H(p) \leq \tau$ does not permit its inversion. The valid low-entropy certificate instead follows from $H_K^{\min}(p_{\max}) \leq H(p)$.

The lower envelope controls what low entropy forces, whereas the upper envelope controls what high entropy rules out. Reversing these implications produces a vacuous bound.

A.10 Additional calibration limitations

Full-vector calibration must be distinguished from calibration imposed only after the relevant item is known to be present in the probe list. Let $Y \in \{0, 1, \ldots, K\}$, where $Y = 0$ means that the required evidence is absent from the probe list and $Y = j$ means that candidate j is the designated relevant item. Define the probe-list coverage probability $c(\mathbf{P}) := \mathbb{P}(Y \neq 0 \mid \mathbf{P})$. Suppose that the softmax weights are calibrated only conditionally on the relevant item being present: $\mathbb{P}(Y = j \mid \mathbf{P}, Y \neq 0) = P_j$, for $j = 1, \ldots, K$. Lemma 15 shows that under this condition, $\mathbb{P}(Y = j \mid \mathbf{P}) = c(\mathbf{P})P_j$. In particular, the conditional probability that the highest-weight candidate is relevant equals $c(\mathbf{P})P_{\max}$. Entropy can control $P_{\max}$, but it does not control the coverage factor $c(\mathbf{P})$. Thus perfect calibration among candidates already present in the list does not certify that the probe contains the required evidence. A further distinction concerns marginal top-score calibration. One might require $\mathbb{E}[C_{\mathrm{F}} \mid P_{\max}] = P_{\max}$ and expect this condition to remain valid after selecting low-entropy cases. This is generally false because entropy depends on the complete residual-mass configuration. For $K \geq 3$, two vectors can have the same largest coordinate but different entropies.

Theorem 7 (Marginal top-score calibration does not survive entropy selection)**.** *Let $K \geq 3$ and choose $x \in (1/2, 1)$. Let $0 < \epsilon < (1 - x)/(K - 1)$. Define*

$$p^{\mathrm{L}} := \left(x,\, 1 - x - (K-2)\epsilon,\, \underbrace{\epsilon, \ldots, \epsilon}_{K-2 \text{ entries}} \right),$$

$$p^{\mathrm{H}} := \left(x,\, \frac{1-x}{K-1}, \ldots, \frac{1-x}{K-1} \right).$$

Then $p^{\mathrm{L}}_{\max} = p^{\mathrm{H}}_{\max} = x$ but $H(p^{\mathrm{L}}) < H(p^{\mathrm{H}})$. There exists a joint distribution of $(\mathbf{P}, C_{\mathrm{F}})$ satisfying $\mathbb{E}[C_{\mathrm{F}} \mid P_{\max}] = P_{\max}$ almost surely, while, for an entropy threshold τ, $\mathbb{P}(C_{\mathrm{F}} = 1 \mid H(\mathbf{P}) \leq \tau) = 0$.

The two vectors have the same maximum mass but distribute the residual mass differently. The low-entropy event can distinguish them even though conditioning on $P_{\max}$ cannot. Correct and incorrect cases can therefore be mixed so that marginal calibration holds at the common value $P_{\max} = x$, while entropy selection isolates only the incorrect component.

This is a selection effect: calibration with respect to one scalar statistic need not survive conditioning on another that contains additional information. A lower calibration condition given the complete vector remains valid on events defined from $\mathbf{P}$, including A_τ; marginal calibration given $P_{\max}$ alone does not.

A.11 Supporting results for the Top-K mean

Although the raw scores are independent in the baseline model, the selected Top-K scores are neither independent nor identically distributed. Lemma 16 makes this explicit. The Top-K mean must therefore be treated as a single functional of the raw scores.

Dependence among the selected order statistics does not prevent concentration. Changing one raw score can affect the Top-K sum only through its contribution to the selected set, permitting a self-bounding argument.

Theorem 8 (Concentration of the dependent Top-K mean)**.** *Suppose that $X_1, \ldots, X_n$ are independent, but not necessarily identically distributed, and satisfy $\ell \leq X_i \leq u$ almost surely. Then, for every $\varepsilon \geq 0$,*

$$\mathbb{P}\left(\widehat{M}_{n,K} - \mathbb{E}[\widehat{M}_{n,K}] \geq \varepsilon \right) \leq \exp\left\{ -\frac{K\varepsilon^2}{2\Delta_s^2 + 2\Delta_s\varepsilon/3} \right\},$$

$$\mathbb{P}\left(\mathbb{E}[\widehat{M}_{n,K}] - \widehat{M}_{n,K} \geq \varepsilon \right) \leq \exp\left\{ -\frac{K\varepsilon^2}{2\Delta_s^2} \right\}.$$

Consequently,

$$\mathbb{P}\left(\left| \widehat{M}_{n,K} - \mathbb{E}[\widehat{M}_{n,K}] \right| \geq \varepsilon \right) \leq 2\exp\left\{ -\frac{K\varepsilon^2}{2\Delta_s^2 + 2\Delta_s\varepsilon/3} \right\}.$$

Moreover, $\mathrm{Var}(\widehat{M}_{n,K}) \le \Delta_s^2/K$.

This result centers at the finite-sample expectation and requires neither identical score distributions nor a density. It controls the Top-K mean directly rather than approximating the selected order statistics by independent variables.

Under the i.i.d. baseline, the variational representation and a uniform empirical-cdf bound yield concentration around the population target.

Theorem 9 (Distribution-free concentration around the population upper-tail target)**.** *Suppose* $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} F$*, where* F *is supported on* $[\ell, u]$*. Let* $\gamma := K/n$*. Then, for every* $\varepsilon > 0$*,*

$$\mathbb{P}\left(\left|\widehat{M}_{n,K} - U_\gamma(F)\right| > \varepsilon\right) \le 2\exp\left\{-\frac{2n\gamma^2\varepsilon^2}{\Delta_s^2}\right\}.$$

Equivalently,

$$\mathbb{P}\left(\left|\widehat{M}_{n,K} - U_{K/n}(F)\right| > \varepsilon\right) \le 2\exp\left\{-\frac{2K^2\varepsilon^2}{n\Delta_s^2}\right\}.$$

When K/n is bounded away from zero, the exponent is of order n. The associated confidence radius and expectation approximation are given in Corollaries 15 and 16. With fixed K, however, the target itself moves as $K/n \to 0$, so the appropriate limit is the upper endpoint.

For a fixed probe size, adding a score cannot decrease the average of the largest K scores. The pathwise monotonicity in Lemma 23 leads to the following limit.

Theorem 10 (Fixed-K convergence to the upper endpoint)**.** *Let* $X_1, X_2, \dots$ *be i.i.d. from a distribution* F *supported on* $[\ell, u]$*. Define* $x_F^\star := \operatorname{ess\,sup} X_1$*. Fix an integer* $K \ge 1$*. Then* $\widehat{M}_{n,K} \to x_F^\star$ *almost surely as* $n \to \infty$*. Because* $\widehat{M}_{n,K}$ *is uniformly bounded,* $\mathbb{E}[\widehat{M}_{n,K}] \to x_F^\star$*.*

Thus a fixed-K probe becomes increasingly extreme even when every raw score follows the same stationary law. A different regime results when the probe grows with the corpus.

Theorem 11 (Consistency under proportional Top-K growth)**.** *Let* $X_1, X_2, \dots$ *be i.i.d. from a distribution* F *supported on* $[\ell, u]$*. Let* K_n *satisfy* $1 \le K_n \le n$ *and* $K_n/n \to \gamma \in (0, 1]$*. Then* $\widehat{M}_{n,K_n} \to U_\gamma(F)$ *almost surely.*

For $\gamma < 1$, the limit is the average of the upper γ-fraction of the distribution; for $\gamma = 1$, it is the ordinary mean. Hence retaining a fixed number of memories targets the endpoint, whereas retaining an asymptotically fixed fraction targets a stable upper-tail functional.

B Assumptions

Assumption 2 (Lipschitz Gram advantage and routing margin)**.** *On a measurable domain* $\mathfrak{C} \subseteq \mathfrak{C}_{n+1}$*, there are constants* $L_\Gamma < \infty$*,* $B < \infty$*,* $C_{\mathrm{mar}} < \infty$*, and* $\kappa \ge 0$ *such that*

$$|d_\Gamma(G) - d_\Gamma(G')| \le L_\Gamma \|G - G'\|_{\max}, \qquad G, G' \in \mathfrak{C}, \qquad |d_\Gamma(\Gamma)| \le B \quad \textit{almost surely}, \tag{17}$$

$$\mathbb{P}\left(0 < |d_\Gamma(\Gamma)| \le t\right) \le C_{\mathrm{mar}} t^\kappa, \qquad t \ge 0. \tag{18}$$

Assumption 3 (Regular and nondegenerate crossing)**.** *The conditional loss difference satisfies* $d \in C^1(\mathcal{Z})$*. There exist constants* $0 < \mu_m \le L_m < \infty$ *and* $0 < \mu_h \le L_h < \infty$ *such that, for every* $(m, h) \in \mathcal{Z}$*,* $\mu_m \le -\partial_m d(m, h) \le L_m$ *and* $\mu_h \le \partial_h d(m, h) \le L_h$*. We additionally assume* $d(m_-, h_-) > 0$ *and* $d(m_+, h_+) < 0$*.*

Assumption 4 (Conditional entropy densities)**.** *There exists a jointly Borel function* $(m, h) \mapsto f_{H|M}(h \mid m)$ *such that, for* P_M*-almost every* $m \in (\theta_{\mathrm{low}}, \theta_{\mathrm{high}})$*,* $f_{H|M}(\cdot \mid m)$ *is a density of the conditional law of* H *given* $M = m$ *with respect to Lebesgue measure. No positive density lower bound is imposed on entropy values outside the realized conditional support.*

Assumption 5 (Full-vector lower calibration)**.** *There exist a nondecreasing function* $g : [1/K, 1] \to [0, 1]$ *and a constant* $\varepsilon_{\mathrm{cal}} \ge 0$ *such that* $\eta_{\mathrm{F}}(\mathbf{P}) \ge g(P_{\max}) - \varepsilon_{\mathrm{cal}}$ *almost surely.*

C LEMMAS

Lemma 1 (Generic Gram factorization of a finite inner-product algorithm)**.** *Let $w_0, \ldots, w_n$ be vectors in a real inner-product space, let $Wa := \sum_{i=0}^{n} a_i w_i$, and let $\Gamma := (\langle w_i, w_j \rangle)_{i,j=0}^{n}$. Consider a finite deterministic algorithm satisfying the following conditions:*

1. *every vector in its initial state is Wa for a fixed known coefficient vector a;*
2. *every subsequently generated vector is Wa' for a coefficient vector a' that is a deterministic function of the previous coefficient vectors, discrete history, and previously computed inner products and norms; and*
3. *every scalar comparison and discrete decision is a deterministic function of the same information.*

Then its complete execution trace is a deterministic function of Γ. In particular, two indexed configurations with the same Gram matrix have identical coefficient traces, scalar comparisons, and discrete decisions. The same conclusion holds for a randomized implementation conditionally on its random seed.

Proof. For arbitrary $a, b \in \mathbb{R}^{n+1}$, bilinearity gives

$$\langle Wa, Wb \rangle = \left\langle \sum_{i=0}^{n} a_i w_i, \sum_{j=0}^{n} b_j w_j \right\rangle = \sum_{i=0}^{n} \sum_{j=0}^{n} a_i b_j \langle w_i, w_j \rangle = \sum_{i=0}^{n} \sum_{j=0}^{n} a_i \Gamma_{ij} b_j = a^\top \Gamma b. \quad (19)$$

Taking $b = a$ yields

$$\|Wa\|^2 = a^\top \Gamma a. \quad (20)$$

Thus every inner product and norm of vectors whose coefficient vectors are already known is determined by Γ and those coefficients.

We prove the trace claim by induction over the finite sequence of algorithmic operations. At the initial operation, all vector coefficients are fixed by condition 1, and the initial discrete state is fixed by the algorithm. Hence the initial state is a function of Γ.

Suppose that the coefficient vectors, computed scalars, and discrete history through some operation are determined by Γ. Equations (19) and (20) determine every inner product and norm requested at the next operation. Conditions 2 and 3 then determine every new coefficient vector, scalar comparison, and discrete decision. Therefore the state after the next operation is also a function of Γ. Induction proves this assertion through the final operation, so the complete finite trace is a deterministic function of Γ.

If two indexed configurations have the same Γ, the same induction gives identical initial states and identical updates at every operation, proving equality of their traces. For a randomized implementation, conditioning on the seed makes all random choices fixed and reduces the algorithm to the deterministic case just proved. □

Lemma 2 (RF-Mem trace factorization and score-row insufficiency)**.** *For every fixed implementation satisfying the conditions in Section 4 and every finite round limit R, RF-Mem satisfies the conditions of Lemma 1. Hence there exists a deterministic map Φ_R such that the complete indexed RF-Mem execution trace satisfies*

$$\mathsf{Tr}_R(q, \mathcal{M}) = \Phi_R(\Gamma(q, \mathcal{M})). \quad (21)$$

The trace includes every retrieved index set and K-means partition, all centroid and query coefficient vectors, all retrieval and cluster scores, every beam selection, deduplication and stopping decision, and the final returned indices. Consequently, two indexed query–corpus configurations with the same augmented Gram matrix have identical scalar scores, discrete decisions, and final returned indices at every round. If their relevance labels are also matched by index, their retrieval-success outcomes agree. By contrast, the complete initial score profile V is not sufficient for this trace, even for the one-beam, one-expansion special case: the two corpora of Theorem 1 have the same V but different expanded retrievals and opposite optimal initial actions. Thus the missing off-query entries

$\{\Gamma_{ij} : 1 \le i < j \le n\}$ *can be decision-relevant. This is an algorithmic sufficiency statement for the RF-Mem execution, not a claim that* Γ *is minimal or statistically sufficient for arbitrary semantic losses.*

Proof. By Lemma 1, it suffices to verify that RF-Mem maintains coefficient representations over the original query and memories and that each of its scalar and discrete operations uses only inner products, norms, and the preceding trace. Let $e_0, e_1, \ldots, e_n$ denote the standard basis of $\mathbb{R}^{n+1}$. Define the linear map $W : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^d$ and $Wa := a_0 q + \sum_{i=1}^n a_i z_i$. By the definition of $w_0, w_1, \ldots, w_n$, $We_i = w_i$ for every $i \in \{0, 1, \ldots, n\}$. In particular, $We_0 = q$ and $We_i = z_i$ for $i \ge 1$. Equations (19) and (20) give

$$\|Wa\|_2^2 = a^\top \Gamma a. \tag{22}$$

Taking $b = e_i$ in (19) and using symmetry of Γ yields

$$\langle Wa, z_i \rangle = \langle Wa, We_i \rangle = a^\top \Gamma e_i = e_i^\top \Gamma a, \qquad i = 1, \ldots, n. \tag{23}$$

Equations (22) and (23) are the RF-Mem specializations used below.

We first verify that every operation inside a single round is determined by Γ once the incoming query coefficient vector and the discrete history are determined. Let an active beam record have unit query $x = Wa$, where $a \in \mathbb{R}^{n+1}$ is its stored coefficient vector. For every memory index i,

$$\langle x, z_i \rangle = e_i^\top \Gamma a \tag{24}$$

by (23). Therefore all retrieval scores are functions of (Γ, a). The set of currently eligible indices is determined by the discrete history. Sorting the eligible scores with the fixed tie-breaking rule determines the retrieved index set.

We next consider K-means on any retrieved, nonempty candidate set. For two candidate indices i and j,

$$\begin{aligned} \|z_i - z_j\|_2^2 &= \langle z_i - z_j, z_i - z_j \rangle \\ &= \langle z_i, z_i \rangle + \langle z_j, z_j \rangle - 2\langle z_i, z_j \rangle = \Gamma_{ii} + \Gamma_{jj} - 2\Gamma_{ij}. \end{aligned} \tag{25}$$

Because the memory embeddings are unit vectors, $\Gamma_{ii} = \Gamma_{jj} = 1$, so the last expression also equals $2 - 2\Gamma_{ij}$. Hence every indexed pairwise squared distance used for initialization is determined by Γ.

Let $C \subseteq \{1, \ldots, n\}$ be any nonempty cluster produced during a K-means iteration. Define $c_C := \frac{1}{|C|} \sum_{j \in C} e_j$. Its arithmetic centroid is

$$g_C^{\mathrm{raw}} = \frac{1}{|C|} \sum_{j \in C} z_j = \frac{1}{|C|} \sum_{j \in C} We_j = W \left(\frac{1}{|C|} \sum_{j \in C} e_j \right) = Wc_C.$$

For a candidate z_i, its squared distance to this centroid is

$$\begin{aligned} \|z_i - g_C^{\mathrm{raw}}\|_2^2 &= \|We_i - Wc_C\|_2^2 = (e_i - c_C)^\top \Gamma (e_i - c_C) \\ &= e_i^\top \Gamma e_i - 2e_i^\top \Gamma c_C + c_C^\top \Gamma c_C = \Gamma_{ii} - \frac{2}{|C|} \sum_{j \in C} \Gamma_{ij} + \frac{1}{|C|^2} \sum_{j \in C} \sum_{k \in C} \Gamma_{jk}. \end{aligned} \tag{26}$$

Thus every assignment distance in every K-means iteration is a function of Γ and the current indexed partition. Likewise, if C and D are nonempty clusters, then the squared displacement between their arithmetic centroids is

$$\|g_C^{\mathrm{raw}} - g_D^{\mathrm{raw}}\|_2^2 = \|Wc_C - Wc_D\|_2^2 = (c_C - c_D)^\top \Gamma (c_C - c_D), \tag{27}$$

and the within-cluster sum of squares is obtained by summing (26). Consequently, initialization, assignment, objective evaluation, convergence checking, and termination are all determined by Γ under the stated distance-based implementation. Fixed indexed tie-breaking and the stated empty-cluster rule make the next partition unique whenever an equality or empty cluster occurs. It follows inductively over the internal K-means iterations that the final indexed partition and every coefficient c_C are determined by Γ.

If the implementation retains the arithmetic centroid, define $\widetilde{c}_C \coloneqq c_C$. If it unit-normalizes the centroid, then the nonzero-normalization assumption and (22) give

$$\|g_C^{\mathrm{raw}}\|_2 = \|Wc_C\|_2 = \sqrt{c_C^\top \Gamma c_C} > 0, \tag{28}$$

and we define $\widetilde{c}_C \coloneqq \frac{c_C}{\sqrt{c_C^\top \Gamma c_C}}$. In the latter case,

$$W\widetilde{c}_C = W\left(\frac{c_C}{\sqrt{c_C^\top \Gamma c_C}}\right) = \frac{Wc_C}{\|Wc_C\|_2} = \frac{g_C^{\mathrm{raw}}}{\|g_C^{\mathrm{raw}}\|_2}.$$

Therefore, under either fixed rule, the centroid representative used by the update has the form

$$g_C = W\widetilde{c}_C,$$

where $\widetilde{c}_C$ is determined by Γ and the indexed cluster C.

The unnormalized query-update vector in (2) is

$$\alpha x + (1-\alpha)g_C + q = \alpha Wa + (1-\alpha)W\widetilde{c}_C + We_0 = W\left[\alpha a + (1-\alpha)\widetilde{c}_C + e_0\right].$$

Define its coefficient vector by

$$h_C \coloneqq \alpha a + (1-\alpha)\widetilde{c}_C + e_0. \tag{29}$$

Its squared norm is $\|Wh_C\|_2^2 = h_C^\top \Gamma h_C$ by (22). The nonzero-numerator assumption makes this quantity strictly positive. Hence the updated query has the coefficient vector

$$a'_C \coloneqq \frac{h_C}{\sqrt{h_C^\top \Gamma h_C}}, \tag{30}$$

because

$$\begin{aligned} Wa'_C &= W\left(\frac{h_C}{\sqrt{h_C^\top \Gamma h_C}}\right) = \frac{Wh_C}{\|Wh_C\|_2} \\ &= \frac{\alpha x + (1-\alpha)g_C + q}{\|\alpha x + (1-\alpha)g_C + q\|_2} = x'. \end{aligned}$$

Moreover, $(a'_C)^\top \Gamma a'_C = \frac{h_C^\top \Gamma h_C}{h_C^\top \Gamma h_C} = 1$, so Wa'_C is indeed a unit vector. Equations (29) and (30) show explicitly that the new query coefficient vector is determined by (Γ, a, C).

For this child query, each individual memory score is

$$\langle x', z_i \rangle = e_i^\top \Gamma a'_C \tag{31}$$

by (23). The cluster score in (3) is therefore

$$\mathrm{score}(C, x') = \sum_{i \in C} e_i^\top \Gamma a'_C = \left(\sum_{i \in C} e_i\right)^\top \Gamma a'_C. \tag{32}$$

Thus all individual and aggregate scores used in beam pruning are determined by Γ, the incoming coefficient vector, and the indexed cluster. The fixed tie-breaking rule then determines the surviving beams. Since the retrieved and surviving index sets are already determined, all deduplication operations, bag updates, cardinality-based stopping decisions, and final score-based rankings are determined as well.

At round zero the only query is $x^{(0)} = q = We_0$, and all initial seen sets, output bags, beam labels, and counters are fixed by the implementation. The preceding calculations show that every later query remains of the form Wa, with its coefficient determined by the preceding trace and Gram-evaluable scalars, and that every retrieval, clustering, beam, deduplication, stopping, and ranking decision has the same property. RF-Mem therefore satisfies all conditions of Lemma 1. Since the round limit is finite, that lemma supplies a deterministic map Φ_R for the complete trace and proves (21). If two

indexed query–corpus configurations have the same Γ, every scalar expression displayed above has the same value in both configurations. The coefficient vectors generated by the deterministic recursions, all indexed partitions and sets, and all discrete decisions are therefore identical. The ambient query and centroid vectors need not have identical coordinates, but their coefficient representations and every inner product observed by RF-Mem agree. If relevance labels are additionally matched by index, identical returned indices imply identical retrieval-success outcomes.

It remains to prove initial-score insufficiency. Index the four memories in the construction of Theorem 1 as (z_+, w_1, w_2, y) in $\mathcal{M}_+$ and (z_-, w_1, w_2, y) in $\mathcal{M}_-$. Equation (91) gives

$$V_+ = V_- = (s, \eta_1, \eta_2, a). \tag{33}$$

Nevertheless, after the common top-index position is selected and the query is updated, the next Top-2 set contains the relevant index y in $\mathcal{M}_+$ and excludes it in $\mathcal{M}_-$, as shown in (92)– (93). Hence the two indexed execution traces are different. If a deterministic map from V to the execution trace existed uniformly over all admissible corpora, it would have to assign the same trace to the identical inputs in (33), which is a contradiction. The same theorem also proves that the two environments require opposite optimal initial actions under the stated cost-adjusted loss.

Finally, because $V_i = \Gamma_{0i}$, the equality of V_+ and V_- fixes only the query–memory entries of the two augmented Gram matrices. Their off-query entries differ; for example,

$$\begin{aligned} \langle z_+, y\rangle &= sa + \sqrt{1-s^2}\sqrt{1-a^2}, \\ \langle z_-, y\rangle &= sa - \sqrt{1-s^2}\sqrt{1-a^2}. \end{aligned} \tag{34}$$

Since $0 < a < s < 1$, both square roots in (34) are strictly positive, so these two inner products are unequal. This explicitly locates the information absent from V and completes the proof. □

Lemma 3 (Loss transfer to the published Bag procedure)**.** *Fix $\alpha \in [0, 1)$ and use the exact realization of published Algorithm 3 specified in Appendix A.3, with $B = F = 1$, $R = 2$, $K = 3$, and no score floor or search-time exclusion of Seen. Then the following statements hold.*

1. *If $n \geq 3$, all $V_i > 0$, and the initial Top-1 memory z_0 is unique, every normalization is well-defined and the returned index set is*

$$\mathcal{R}_{\mathrm{R}}^{\mathrm{Bag}}(q; \mathcal{M}) = \{z_0\} \cup \mathrm{Top}_2(q'; \mathcal{M}), \qquad q' = \mathsf{U}_\alpha(q, z_0). \tag{35}$$

   Here the notation identifies memories with their indices, and the Top-2 search includes z_0.

2. *The witness pairs constructed for Theorems 1 and 4 retain their Familiarity and Recollection failure losses under $\mathcal{R}_{\mathrm{F}} = \mathrm{Top}_3(q; \mathcal{M})$ and $\mathcal{R}_{\mathrm{R}} = \mathcal{R}_{\mathrm{R}}^{\mathrm{Bag}}$:*

$$(\ell_{\mathrm{F},+}, \ell_{\mathrm{R},+}) = (1, 0), \qquad (\ell_{\mathrm{F},-}, \ell_{\mathrm{R},-}) = (1, 1).$$

   With the same fixed costs and $\Delta_c \in (0, 1)$, all stated Bayes gaps, stagewise localizations, and minimax regrets are unchanged. This uses a common output cap $K = 3$, not equal output cardinalities.

3. *The three-memory witness in the proof of Theorem 3 also changes the published returned index set by changing only its designated off-query pair. This is an output-necessity statement, not an opposite-action claim for that three-memory corpus at $K = 3$.*

Proof. At $r = 0$, the retrieval width is $N = (1 + 0)1 = 1$, so the only retrieved cluster is $\{z_0\}$. Its normalized centroid is z_0, its child query is q', and the sole beam survives. Bag therefore contains exactly z_0. Its size is below $K = 3$, so execution continues. At $r = 1$, $N = (1 + 1)1 = 2$, and the sole cluster is $I_1 = \mathrm{Top}_2(q'; \mathcal{M})$, since $k = \min(B, |I_1|) = 1$. No child is pruned. Insertion with Seen therefore leaves precisely $\{z_0\} \cup I_1$ in Bag. This union has at most three indices, so final Top-3 returns all of them irrespective of their stored scores or order. There is no third retrieval because $R = 2$. This proves (35).

To check the normalizations, any nonempty centroid has strictly positive projection onto q, because every $V_i > 0$. Hence it is nonzero, and normalization preserves that positive projection. Starting from $\langle q, q\rangle = 1$, an update numerator has projection $\alpha\langle x, q\rangle + (1 - \alpha)\langle g, q\rangle + 1 > 0$ for $\alpha \in [0, 1]$.

Thus both child queries, including the last child whose score is stored without a further search, are well-defined.

For the complete-score construction, use its parameters $0 < a < \eta_2 < \eta_1 < s < 1$, with $\eta_1 < a + (b/A)\sqrt{1-a^2}$, where $A = 1 + \alpha + (1-\alpha)s$, $b = (1-\alpha)\sqrt{1-s^2} > 0$ and $D = \sqrt{A^2 + b^2}$. The construction has $q = e_1$, $z_\pm = se_1 \pm \sqrt{1-s^2}e_2$, $y = ae_1 + \sqrt{1-a^2}e_2$, and $w_j = \eta_j e_1 + \sqrt{1-\eta_j^2}e_3$. Consequently $q'_\pm = (Ae_1 \pm be_2)/D$ and $\langle q'_\pm, z_\pm\rangle = (As + b\sqrt{1-s^2})/D > A\eta_1/D$, $\langle q'_\pm, w_j\rangle = A\eta_j/D$ and $\langle q'_\pm, y\rangle = (Aa \pm b\sqrt{1-a^2})/D$. The plus target score strictly exceeds $A\eta_1/D$, whereas the minus target score is below $A\eta_2/D$. The full-corpus Top-2 sets are therefore $\{z_+, y\}$ and $\{z_-, w_1\}$, respectively. Equation (35) shows that these are also the final Bag outputs; in this pair each contains only two items. Familiarity still returns $\{z_\pm, w_1, w_2\}$ and misses y. This proves the claimed failure losses for the complete-score pair.

The same reasoning applies to all three stagewise witnesses. Each has four memories with positive initial scores, a unique seed, two distractors initially above the target, and hence a Familiarity miss. In its plus environment the target outranks both nonseed distractors after the first update; at most the seed can outrank it, so it belongs to the full-corpus Top-2. In its minus environment both distractors outrank the target, so it does not belong to that Top-2. For the geometry pair these are the inequalities just proved. For the truncation pair they follow from (122), (125), (126), and (127). For the summary pair they are precisely the positive margin in (135) and the minus-score comparison proved in the summary construction. Thus (35) preserves the failure losses in every pair. The configurations and their indexed relevance labels have not changed, so neither have the observations Γ, V, S, T or which endpoints of the ladder distinguish the two environments. Preserving both potential losses therefore preserves each of their information-conditional Bayes risks and every stagewise localization.

Explicitly, the two cost-adjusted advantages are $1 - \Delta_c$ and $-\Delta_c$. A randomized router that cannot distinguish the pair chooses Recollection with one common probability p, giving regrets $(1-p)(1-\Delta_c)$ and $p\Delta_c$. Their equal-prior average is minimized at an endpoint and equals $\frac{1}{2}\min\{\Delta_c, 1-\Delta_c\}$. Their maximum is minimized where they agree, at $p = 1 - \Delta_c$, and equals $\Delta_c(1-\Delta_c)$. These cost conclusions use the stipulated fixed action costs, not a newly inferred latency model.

Finally consider the three-memory construction for pairwise necessity. Its scores satisfy $s > \eta > a > 0$ and its common child query is $q' = (Ae_1 + be_2)/D$. The construction changes only the symmetric pair linking z_0 to $y_\pm$, while

$$\langle q', z_0\rangle = \frac{As + b\sqrt{1-s^2}}{D} > \frac{A\eta}{D} = \langle q', w\rangle, \qquad \langle q', y_\pm\rangle = \frac{A\eta \pm b\sqrt{1-a^2}h}{D},$$

where $h > 0$. Thus the full-corpus Top-2, and by (35) the final Bag output, is $\{z_0, y_+\}$ in the plus environment and $\{z_0, w\}$ in the minus environment. All normalizations are covered by the positive projection argument. Relabeling the three memory indices makes any designated off-query position decisive, as in the original construction. Familiarity Top-3 includes all three memories here, so this last assertion concerns output necessity only. □

Lemma 4 (Bayes-optimal router under a fixed information set)**.** *For every sub-σ-field $\mathcal{H} \subseteq \mathcal{F}$, an optimal router is*

$$\pi^\star_{\mathcal{H}} = \begin{cases} 1, & D_{\mathcal{H}} > 0, \\ 0, & D_{\mathcal{H}} < 0, \\ \textit{any value in } [0,1], & D_{\mathcal{H}} = 0. \end{cases}$$

Moreover,

$$\begin{aligned} \mathcal{R}^\star_{\mathcal{H}} &= \mathbb{E}\left[\min\left\{\mathbb{E}[L_{\mathrm{F}} \mid \mathcal{H}], \mathbb{E}[L_{\mathrm{R}} \mid \mathcal{H}]\right\}\right] \\ &= \frac{1}{2}\mathbb{E}\left[L_{\mathrm{F}} + L_{\mathrm{R}}\right] - \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{H}}|\right]. \end{aligned} \tag{36}$$

Proof. Let $m_{\mathrm{F},\mathcal{H}} \coloneqq \mathbb{E}[L_{\mathrm{F}} \mid \mathcal{H}]$ and$m_{\mathrm{R},\mathcal{H}} \coloneqq \mathbb{E}[L_{\mathrm{R}} \mid \mathcal{H}]$. For an $\mathcal{H}$-measurable router π,

$$\mathcal{R}(\pi) = \mathbb{E}\left[\mathbb{E}\left[(1-\pi)L_{\mathrm{F}} + \pi L_{\mathrm{R}} \mid \mathcal{H}\right]\right] = \mathbb{E}\left[(1-\pi)m_{\mathrm{F},\mathcal{H}} + \pi m_{\mathrm{R},\mathcal{H}}\right].$$

Since $D_{\mathcal{H}} = m_{\mathrm{F},\mathcal{H}} - m_{\mathrm{R},\mathcal{H}}$, we have $m_{\mathrm{R},\mathcal{H}} = m_{\mathrm{F},\mathcal{H}} - D_{\mathcal{H}}$. Thus

$$(1-\pi)m_{\mathrm{F},\mathcal{H}} + \pi m_{\mathrm{R},\mathcal{H}} = (1-\pi)m_{\mathrm{F},\mathcal{H}} + \pi\left(m_{\mathrm{F},\mathcal{H}} - D_{\mathcal{H}}\right) = m_{\mathrm{F},\mathcal{H}} - \pi D_{\mathcal{H}}.$$

Pointwise in ω, this is affine in $\pi(\omega) \in [0,1]$.

If $D_{\mathcal{H}}(\omega) > 0$, it is strictly decreasing in $\pi(\omega)$, so its minimum is attained at $\pi(\omega) = 1$. If $D_{\mathcal{H}}(\omega) < 0$, it is strictly increasing in $\pi(\omega)$, so its minimum is attained at $\pi(\omega) = 0$. If $D_{\mathcal{H}}(\omega) = 0$, it is constant in $\pi(\omega)$. This proves the optimal-router statement and (36).

For any real x, y, $\min\{x,y\} = \frac{x+y-|x-y|}{2}$. Applying this identity,

$$\min\left\{m_{\mathrm{F},\mathcal{H}}, m_{\mathrm{R},\mathcal{H}}\right\} = \frac{1}{2}\left[m_{\mathrm{F},\mathcal{H}} + m_{\mathrm{R},\mathcal{H}} - |m_{\mathrm{F},\mathcal{H}} - m_{\mathrm{R},\mathcal{H}}|\right] = \frac{1}{2}\left[m_{\mathrm{F},\mathcal{H}} + m_{\mathrm{R},\mathcal{H}} - |D_{\mathcal{H}}|\right].$$

Taking expectations and using the tower property,

$$\mathcal{R}^{\star}_{\mathcal{H}} = \frac{1}{2}\mathbb{E}\left[m_{\mathrm{F},\mathcal{H}} + m_{\mathrm{R},\mathcal{H}}\right] - \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{H}}|\right] = \frac{1}{2}\mathbb{E}\left[L_{\mathrm{F}} + L_{\mathrm{R}}\right] - \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{H}}|\right].$$

□

Lemma 5 (Binary-action value of nested information)**.** *Let L_{F} and L_{R} be integrable, let $\mathcal{H}_0 \subseteq \mathcal{H}_1 \subseteq \cdots \subseteq \mathcal{H}_m$ be sub-σ-fields, let $D_{\mathcal{H}_j} := \mathbb{E}[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \mathcal{H}_j]$, and let π be an $\mathcal{H}_0$-measurable randomized router. Then*

$$\mathcal{R}(\pi) - \mathcal{R}^{\star}_{\mathcal{H}_m} = \mathbb{E}\left[(1-\pi)(D_{\mathcal{H}_0})_+ + \pi(D_{\mathcal{H}_0})_-\right] + \frac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[|D_{\mathcal{H}_j}| - |D_{\mathcal{H}_{j-1}}|\right]. \tag{37}$$

Every term on the right-hand side is nonnegative. In particular, for any $\mathcal{H}_0 \subseteq \mathcal{H}_1$,

$$\mathcal{R}^{\star}_{\mathcal{H}_0} - \mathcal{R}^{\star}_{\mathcal{H}_1} = \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{H}_1}| - |D_{\mathcal{H}_0}|\right] \geq 0. \tag{38}$$

Proof. Let $m_{\mathrm{F},0} := \mathbb{E}[L_{\mathrm{F}} \mid \mathcal{H}_0]$ and $m_{\mathrm{R},0} := \mathbb{E}[L_{\mathrm{R}} \mid \mathcal{H}_0]$. Since π is $\mathcal{H}_0$-measurable, the tower property gives $\mathcal{R}(\pi) = \mathbb{E}\left[\mathbb{E}\left[(1-\pi)L_{\mathrm{F}} + \pi L_{\mathrm{R}} \mid \mathcal{H}_0\right]\right] = \mathbb{E}\left[(1-\pi)m_{\mathrm{F},0} + \pi m_{\mathrm{R},0}\right]$. Because $D_{\mathcal{H}_0} = m_{\mathrm{F},0} - m_{\mathrm{R},0}$, $(1-\pi)m_{\mathrm{F},0} + \pi m_{\mathrm{R},0} = m_{\mathrm{F},0} - \pi D_{\mathcal{H}_0}$. Lemma 4 gives the conditional Bayes loss $m_{\mathrm{F},0} - (D_{\mathcal{H}_0})_+$. Hence

$$\mathcal{R}(\pi) - \mathcal{R}^{\star}_{\mathcal{H}_0} = \mathbb{E}\left[(D_{\mathcal{H}_0})_+ - \pi D_{\mathcal{H}_0}\right] = \mathbb{E}\left[(1-\pi)(D_{\mathcal{H}_0})_+ + \pi(D_{\mathcal{H}_0})_-\right], \tag{39}$$

where the second equality uses $D_{\mathcal{H}_0} = (D_{\mathcal{H}_0})_+ - (D_{\mathcal{H}_0})_-$.

For each $j = 1, \ldots, m$, Lemma 4 also gives

$$\begin{aligned}\mathcal{R}^{\star}_{\mathcal{H}_{j-1}} - \mathcal{R}^{\star}_{\mathcal{H}_j} &= \left\{\frac{1}{2}\mathbb{E}[L_{\mathrm{F}} + L_{\mathrm{R}}] - \frac{1}{2}\mathbb{E}|D_{\mathcal{H}_{j-1}}|\right\} - \left\{\frac{1}{2}\mathbb{E}[L_{\mathrm{F}} + L_{\mathrm{R}}] - \frac{1}{2}\mathbb{E}|D_{\mathcal{H}_j}|\right\} \\ &= \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{H}_j}| - |D_{\mathcal{H}_{j-1}}|\right]. \end{aligned} \tag{40}$$

This proves (38) when $m = 1$.

Adding and subtracting the intermediate optimal risks yields

$$\mathcal{R}(\pi) - \mathcal{R}^{\star}_{\mathcal{H}_m} = \left[\mathcal{R}(\pi) - \mathcal{R}^{\star}_{\mathcal{H}_0}\right] + \sum_{j=1}^{m}\left[\mathcal{R}^{\star}_{\mathcal{H}_{j-1}} - \mathcal{R}^{\star}_{\mathcal{H}_j}\right].$$

Substitution of (39) and (40) proves (37).

The gate term is nonnegative because $0 \leq \pi \leq 1$ and both positive and negative parts are nonnegative. For each j, $\Pi(\mathcal{H}_{j-1}) \subseteq \Pi(\mathcal{H}_j)$, so $\mathcal{R}^{\star}_{\mathcal{H}_j} \leq \mathcal{R}^{\star}_{\mathcal{H}_{j-1}}$. Equation (40) therefore shows that every adjacent information term is nonnegative. □

Lemma 6 (Exact action-disagreement identity)**.** *For every deterministic $\mathcal{T}$-measurable router π,*

$$\mathcal{R}(\pi) - \mathcal{R}^{\star}_{\mathcal{T}} = \mathbb{E}\left[|d(M,H)|\mathbf{1}\left\{\pi(M,H) \neq \pi^{\star}(M,H)\right\}\right].$$

Proof. Condition on $M = m$ and $H = h$. If $d(m,h) > 0$, then Recollection is uniquely optimal. Choosing Familiarity incurs conditional excess risk $r_{\mathrm{F}}(m,h) - r_{\mathrm{R}}(m,h) = d(m,h) = |d(m,h)|$. Choosing Recollection incurs zero excess risk. If $d(m,h) < 0$, then Familiarity is uniquely optimal. Choosing Recollection incurs $r_{\mathrm{R}}(m,h) - r_{\mathrm{F}}(m,h) = -d(m,h) = |d(m,h)|$. Choosing Familiarity incurs zero excess risk. If $d(m,h) = 0$, both actions have equal conditional risk, and $|d(m,h)| = 0$. Thus, pointwise, the conditional excess risk equals $|d(m,h)|\mathbf{1}\{\pi(m,h) \neq \pi^{\star}(m,h)\}$. Taking expectation proves the identity. □

Lemma 7 (Approximate-Gram routing-regret bounds)**.** *Use the Gram advantage and plug-in router defined in Section 4, with $\Gamma, \widehat{\Gamma} \in \mathfrak{C}$ almost surely and the stated conditional independence of the approximation randomness. Under the Lipschitz condition* (17) *in Assumption 2,*

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^{\star}_{\mathcal{K}_\Gamma} \le L_\Gamma \mathbb{E}\left[\|\widehat{\Gamma} - \Gamma\|_{\max}\right]. \tag{41}$$

If Assumption 2 holds in full and there are $\varepsilon \ge 0$ and $\delta \in [0,1]$ such that

$$\mathbb{P}\left(\|\widehat{\Gamma} - \Gamma\|_{\max} > \varepsilon\right) \le \delta, \tag{42}$$

then

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^{\star}_{\mathcal{K}_\Gamma} \le C_{\mathrm{mar}}(L_\Gamma \varepsilon)^{\kappa+1} + B\delta. \tag{43}$$

In particular, with an almost-sure error bound and a hard routing margin $|d_\Gamma(\Gamma)| \ge \gamma > 0$ whenever it is nonzero, $L_\Gamma \varepsilon < \gamma$ gives zero excess routing risk. The same conclusions hold for a partial-Gram representation $P(\Gamma)$, with an approximation $\widehat{P}$, whenever $d_\Gamma(G) = \overline{d}(P(G))$ for a function $\overline{d}$ that is Lipschitz in the corresponding representation norm; the error terms are then $\|\widehat{P} - P(\Gamma)\|$, with $\widehat{P}$ constructed using randomness conditionally independent of the potential losses given Γ.

Proof. Write $d \coloneqq d_\Gamma(\Gamma)$ and $\widehat{d} \coloneqq d_\Gamma(\widehat{\Gamma})$. The Bayes router based on the exact Gram information, with Familiarity on ties, is $\pi^{\star}_{\Gamma} \coloneqq \mathbf{1}\{d > 0\}$. By definition of d_Γ,

$$d = \mathbb{E}\left[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \Gamma\right] = \mathbb{E}[L_{\mathrm{F}} \mid \Gamma] - \mathbb{E}[L_{\mathrm{R}} \mid \Gamma]. \tag{44}$$

The auxiliary randomness used to form $\widehat{\Gamma}$ is conditionally independent of $(L_{\mathrm{F}}, L_{\mathrm{R}})$ given Γ. Therefore, conditional on $(\Gamma, \widehat{\Gamma})$, the action risks remain the two conditional risks in (44). If $d > 0$, Recollection is optimal and choosing Familiarity incurs conditional excess risk $\mathbb{E}[L_{\mathrm{F}} \mid \Gamma] - \mathbb{E}[L_{\mathrm{R}} \mid \Gamma] = d = |d|$. If $d < 0$, Familiarity is optimal and choosing Recollection incurs conditional excess risk $\mathbb{E}[L_{\mathrm{R}} \mid \Gamma] - \mathbb{E}[L_{\mathrm{F}} \mid \Gamma] = -d = |d|$. If $d = 0$, both actions have the same conditional risk and the excess risk is zero. Taking expectations gives the exact identity

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^{\star}_{\mathcal{K}_\Gamma} = \mathbb{E}\left[|d|\mathbf{1}\left\{\widehat{\pi} \ne \pi^{\star}_{\Gamma}\right\}\right]. \tag{45}$$

We next establish the fundamental plug-in inequality. On the event $\{d > 0, \widehat{d} \le 0\}$, $|d - \widehat{d}| = d - \widehat{d} \ge d = |d|$. On the event $\{d < 0, \widehat{d} > 0\}$,

$$|d - \widehat{d}| = \widehat{d} - d \ge -d = |d|.$$

When $d = 0$, the product of $|d|$ and any indicator is zero. Consequently, pointwise,

$$|d|\mathbf{1}\left\{\widehat{\pi} \ne \pi^{\star}_{\Gamma}\right\} \le |d - \widehat{d}|. \tag{46}$$

The Lipschitz condition (17) yields

$$|d - \widehat{d}| = |d_\Gamma(\Gamma) - d_\Gamma(\widehat{\Gamma})| \le L_\Gamma \|\widehat{\Gamma} - \Gamma\|_{\max}. \tag{47}$$

Combining (45), (46), and (47), and then taking expectations, gives

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^{\star}_{\mathcal{K}_\Gamma} \le \mathbb{E}[|d - \widehat{d}|] \le L_\Gamma \mathbb{E}\left[\|\widehat{\Gamma} - \Gamma\|_{\max}\right].$$

This proves (41).

For the fast bound, define the good approximation event $\mathcal{E}_\varepsilon \coloneqq \left\{\|\widehat{\Gamma} - \Gamma\|_{\max} \le \varepsilon\right\}$. On $\mathcal{E}_\varepsilon \cap \{\widehat{\pi} \ne \pi^{\star}_{\Gamma}\}$, equations (46) and (47) imply

$$0 < |d| \le |d - \widehat{d}| \le L_\Gamma \varepsilon, \tag{48}$$

except when $d = 0$, whose regret contribution is zero. Split the exact excess risk according to $\mathcal{E}_\varepsilon$:

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^{\star}_{\mathcal{K}_\Gamma} = \mathbb{E}\left[|d|\mathbf{1}\{\widehat{\pi} \ne \pi^{\star}_{\Gamma}\}\mathbf{1}_{\mathcal{E}_\varepsilon}\right] + \mathbb{E}\left[|d|\mathbf{1}\{\widehat{\pi} \ne \pi^{\star}_{\Gamma}\}\mathbf{1}_{\mathcal{E}_\varepsilon^c}\right]. \tag{49}$$

For the first term, use (48) and then the margin condition (18):

$$\mathbb{E}\left[|d|\mathbf{1}\{\widehat{\pi} \neq \pi^\star_\Gamma\}\mathbf{1}_{\mathcal{E}_\varepsilon}\right] \leq L_\Gamma \varepsilon \mathbb{P}\left(0 < |d| \leq L_\Gamma \varepsilon\right) \leq L_\Gamma \varepsilon C_{\mathrm{mar}} (L_\Gamma \varepsilon)^\kappa = C_{\mathrm{mar}} (L_\Gamma \varepsilon)^{\kappa+1}. \tag{50}$$

For the second term, use $|d| \leq B$ and (42):

$$\mathbb{E}\left[|d|\mathbf{1}\{\widehat{\pi} \neq \pi^\star_\Gamma\}\mathbf{1}_{\mathcal{E}^c_\varepsilon}\right] \leq B\mathbb{P}(\mathcal{E}^c_\varepsilon) \leq B\delta. \tag{51}$$

Substituting (50) and (51) into (49) proves (43).

Suppose next that the approximation error is at most ε almost surely and every nonzero $|d|$ is at least γ. If $L_\Gamma \varepsilon < \gamma$, then (48) would require a nonzero number $|d|$ to satisfy simultaneously $|d| \geq \gamma > L_\Gamma \varepsilon \geq |d|$, which is impossible. Thus a mismatch can occur only when $d = 0$, where it incurs zero regret. The excess risk is therefore zero.

Finally, suppose $d_\Gamma(G) = \overline{d}(P(G))$ for a partial-Gram representation P, and suppose $\overline{d}$ is Lipschitz in a representation norm. Conditional independence and the tower property give

$$\mathbb{E}[L_{\mathrm{F}} - L_{\mathrm{R}} \mid P(\Gamma), \widehat{P}] = \mathbb{E}[d_\Gamma(\Gamma) \mid P(\Gamma), \widehat{P}] = \overline{d}(P(\Gamma)).$$

Thus $P(\Gamma)$ and Γ have the same Bayes action and risk. Use $d = \overline{d}(P(\Gamma))$ and $\widehat{d} = \overline{d}(\widehat{P})$ in the preceding argument. This gives the same disagreement identity and all the bounds in the representation norm, proving the partial-representation claim. □

Lemma 8 (Monotone Bayes-frontier characterization)**.** *Under Assumption 1, the following statements hold.*

i) For every $m \in [m_-, m_+]$, the set $\mathcal{A}(m)$ is one of $\varnothing$, $[h_-, h_+]$, $[h_-, b^\star(m)]$ for some $b^\star(m) \in [h_-, h_+]$.

ii) The Bayes-optimal routing rule is

$$a^\star(m, h) = \begin{cases} \mathrm{F}, & h \leq b^\star(m), \\ \mathrm{R}, & h > b^\star(m), \end{cases} \tag{52}$$

where no $h \in [h_-, h_+]$ is regarded as satisfying $h \leq -\infty$.

iii) The boundary $b^\star$ is nondecreasing: $m_1 \leq m_2 \implies b^\star(m_1) \leq b^\star(m_2)$.

iv) As an extended-real-valued monotone function, $b^\star$ is Borel measurable.

Proof. Fix $m \in [m_-, m_+]$. We first show that $\mathcal{A}(m)$ is a lower set in the entropy coordinate. Let $h_2 \in \mathcal{A}(m)$ and let $h_- \leq h_1 \leq h_2$. Since $h_2 \in \mathcal{A}(m)$, $d(m, h_2) \leq 0$. By monotonicity in entropy, $d(m, h_1) \leq d(m, h_2)$. Combining the two inequalities, $d(m, h_1) \leq d(m, h_2) \leq 0$. Hence $h_1 \in \mathcal{A}(m)$. Thus, whenever $\mathcal{A}(m)$ contains an entropy value, it contains every smaller admissible entropy value. Because d is continuous, the function $h \longmapsto d(m, h)$ is continuous on the compact interval $[h_-, h_+]$. Consequently, $\mathcal{A}(m) = \{h : d(m, h) \leq 0\}$ is closed. A closed lower subset of an interval must be either empty, the entire interval, or an interval beginning at its lower endpoint. We verify this explicitly. Suppose $\mathcal{A}(m)$ is nonempty and is not the entire interval. Since it is a nonempty closed subset of the compact set $[h_-, h_+]$, it possesses a maximum: $b \coloneqq \max \mathcal{A}(m)$. By the lower-set property, $[h_-, b] \subseteq \mathcal{A}(m)$. By maximality of b, no $h > b$ belongs to $\mathcal{A}(m)$. Therefore, $\mathcal{A}(m) = [h_-, b]$. By definition, $b = b^\star(m)$. This proves statement i).

The Bayes rule selects Familiarity precisely on the event $d(m, h) \leq 0$, which is equivalent to $h \in \mathcal{A}(m)$. The representation established in statement i) therefore gives (52). This proves statement ii).

We next prove monotonicity of $b^\star$. Let $m_1 \leq m_2$. Take any $h \in \mathcal{A}(m_1)$. Then $d(m_1, h) \leq 0$. By monotonicity in the mean score, $d(m_2, h) \leq d(m_1, h) \leq 0$. Hence $h \in \mathcal{A}(m_2)$. We have shown

$$\mathcal{A}(m_1) \subseteq \mathcal{A}(m_2). \tag{53}$$

If $\mathcal{A}(m_1) = \varnothing$, then $b^\star(m_1) = -\infty \leq b^\star(m_2)$. If $\mathcal{A}(m_1) \neq \varnothing$, the inclusion (53) implies that the maximum of $\mathcal{A}(m_1)$ cannot exceed the maximum of $\mathcal{A}(m_2)$. Therefore, $b^\star(m_1) \leq b^\star(m_2)$. This proves statement iii).

Finally, every monotone extended-real-valued function on an interval is Borel measurable. For completeness, for every $c \in \mathbb{R}$, monotonicity implies that $\{m : b^\star(m) > c\}$ is an interval, possibly empty or the full domain, and is therefore Borel. Thus $b^\star$ is measurable. □

Lemma 9 (Exact regret of a fixed entropy threshold)**.** *Under Assumption 3,*

$$\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}} = \mathbb{E}\left[|d(M,H)|\mathbf{1}_I \mathbf{1}\left\{\mathbf{1}\{H \le \tau\} \neq \mathbf{1}\{H \le b(M)\}\right\}\right].$$

If, conditionally on $M = m$, H admits a density $f_{H|M}(h \mid m)$, then

$$\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}} = \mathbb{E}\left[\mathbf{1}_I \int_{\min\{b(M),\tau\}}^{\max\{b(M),\tau\}} |d(M,h)| f_{H|M}(h \mid M)\, dh\right].$$

Proof. For $m < \theta_{\text{low}}$, the Bayes rule and π_τ both choose Recollection for every h. For $m > \theta_{\text{high}}$, both choose Familiarity for every h. At $m = \theta_{\text{low}}$, the only possible action disagreement occurs at $h = h_-$, where $d(\theta_{\text{low}}, h_-) = 0$. Its contribution to regret is therefore zero.

At $m = \theta_{\text{high}}$, Familiarity is Bayes-optimal for every $h \le h_+$, with a tie at $h = h_+$, and π_τ also chooses Familiarity.

It remains to consider the middle region I. There, the Bayes rule chooses Familiarity if and only if $H \le b(M)$, whereas the rectangular rule chooses Familiarity if and only if $H \le \tau$. Hence the two rules disagree exactly when $\mathbf{1}\{H \le \tau\} \neq \mathbf{1}\{H \le b(M)\}$. Lemma 6 gives the first identity.

If a conditional density exists, then conditional on $M = m$ the disagreement set is the interval between $b(m)$ and τ, up to endpoints. Endpoints have zero conditional Lebesgue measure. Therefore the conditional excess risk is $\int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |d(m,h)| f_{H|M}(h \mid m)\, dh$. Taking expectation over M proves the second identity. □

Lemma 10 (Schur concavity of Shannon entropy)**.** *Let $a, b \in \Delta_K$. If $a \succeq b$, then $H(a) \le H(b)$.*

Proof. By the Hardy–Littlewood–Pólya characterization of majorization, $a \succeq b$ implies that there exists a doubly stochastic matrix $D = (D_{ij})_{i,j=1}^K$ such that

$$b = Da. \tag{54}$$

Thus

$$D_{ij} \ge 0, \qquad \sum_{j=1}^K D_{ij} = 1, \qquad \sum_{i=1}^K D_{ij} = 1. \tag{55}$$

Define $\varphi(x) := -x \log x$ for $x \in [0,1]$, with $\varphi(0) = 0$. For $x \in (0,1]$, $\varphi'(x) = -\log x - 1$ and $\varphi''(x) = -\frac{1}{x} < 0$. Hence φ is concave. By (54), $b_i = \sum_{j=1}^K D_{ij} a_j$. Since the entries in row i sum to one, Jensen's inequality gives $\varphi(b_i) = \varphi\left(\sum_{j=1}^K D_{ij} a_j\right) \ge \sum_{j=1}^K D_{ij} \varphi(a_j)$. Summing over i,

$$H(b) = \sum_{i=1}^K \varphi(b_i) \ge \sum_{i=1}^K \sum_{j=1}^K D_{ij} \varphi(a_j) = \sum_{j=1}^K \varphi(a_j) \sum_{i=1}^K D_{ij} = \sum_{j=1}^K \varphi(a_j) = H(a),$$

where the penultimate equality uses the column-sum condition in (55). □

Lemma 11 (Continuity and strict monotonicity of the minimal-entropy envelope)**.** *The function $H_K^{\min} : [1/K, 1] \to [0, \log K]$ is continuous and strictly decreasing. Moreover, $H_K^{\min}(1/r) = \log r$ for $r = 1, \dots, K$.*

Proof. Fix $r \in \{1, \dots, K-1\}$ and consider $x \in \left(\frac{1}{r+1}, \frac{1}{r}\right)$. On this interval, $r_x = \left\lfloor \frac{1}{x} \right\rfloor = r$ and $u_x = 1 - rx$. Thus $H_K^{\min}(x) = -rx \log x - (1 - rx) \log(1 - rx)$. Differentiate the first term: $\frac{d}{dx}(-rx \log x) = -r(\log x + 1)$. Let $u(x) := 1 - rx$. Then $u'(x) = -r$. Therefore,

$$\frac{d}{dx}\left(-u(x) \log u(x)\right) = -u'(x)\left(\log u(x) + 1\right) = r\left(\log u(x) + 1\right).$$

Adding the derivatives,

$$\frac{d}{dx}H_K^{\min}(x) = -r(\log x + 1) + r(\log u(x) + 1) = r\left[\log u(x) - \log x\right] = r\log\frac{u(x)}{x}.$$

On the current interval, $0 < u(x) < x$. Hence $0 < \frac{u(x)}{x} < 1$, and therefore $\log\frac{u(x)}{x} < 0$. It follows that $\frac{d}{dx}H_K^{\min}(x) < 0$. Thus $H_K^{\min}$ is strictly decreasing between every pair of consecutive breakpoints. At $x = 1/r$, $r_x = r$ and $u_x = 0$, and therefore $H_K^{\min}(1/r) = -r\frac{1}{r}\log\frac{1}{r} = -\log\frac{1}{r} = \log r$. We now verify continuity at $x = 1/r$. Approaching from below, $r_x = r$ and $u_x = 1 - rx \longrightarrow 0$. Since $\lim_{t\downarrow 0} t\log t = 0$, we obtain $\lim_{x\uparrow 1/r} H_K^{\min}(x) = -r\frac{1}{r}\log\frac{1}{r} = \log r$.

Approaching from above, $r_x = r - 1$ and $u_x = 1 - (r-1)x \longrightarrow 1 - \frac{r-1}{r} = \frac{1}{r}$. Hence

$$\lim_{x\downarrow 1/r} H_K^{\min}(x) = -\frac{r-1}{r}\log\frac{1}{r} - \frac{1}{r}\log\frac{1}{r} = -\log\frac{1}{r} = \log r.$$

Thus the function is continuous at every breakpoint. Since it is strictly decreasing on each interval and its breakpoint values satisfy $\log(r+1) > \log r$, it is strictly decreasing on the full interval $[1/K, 1]$. Finally, $H_K^{\min}(1/K) = \log K$ and $H_K^{\min}(1) = 0$. □

Lemma 12 (Sharp low-entropy certificate)**.** *Let $K \geq 2$ and use ψ_K defined in Section 8. For every $p \in \Delta_K$ and every $\tau \in [0, \log K]$, $H(p) \leq \tau \Longrightarrow p_{\max} \geq \psi_K(\tau)$. The certificate is sharp: for every τ, there exists $p \in \Delta_K$ such that $H(p) = \tau$ and $p_{\max} = \psi_K(\tau)$.*

Proof. Set $x \coloneqq p_{\max}$. By Theorem 6, $H_K^{\min}(x) \leq H(p)$. If $H(p) \leq \tau$, then $H_K^{\min}(x) \leq \tau$. By Lemma 11, $H_K^{\min}$ is strictly decreasing. Since $H_K^{\min}(\psi_K(\tau)) = \tau$, the inequality $H_K^{\min}(x) \leq H_K^{\min}(\psi_K(\tau))$ implies $x \geq \psi_K(\tau)$. Thus $p_{\max} \geq \psi_K(\tau)$.

For sharpness, set $x \coloneqq \psi_K(\tau)$ and choose $p = p^{\min}(x)$. Then $p_{\max} = x = \psi_K(\tau)$ and $H(p) = H_K^{\min}(x) = H_K^{\min}(\psi_K(\tau)) = \tau$. □

Lemma 13 (Correctness is not identifiable from entropy alone)**.** *Fix $K \geq 2$ and $\tau > 0$. There exist two joint distributions of $(\mathbf{P}, C_{\mathrm{F}})$ such that: i) the marginal distribution of $\mathbf{P}$ is identical under the two models; ii) $H(\mathbf{P}) \leq \tau$ almost surely under both models; iii) under the first model, $C_{\mathrm{F}} = 0$ almost surely; iv) under the second model, $C_{\mathrm{F}} = 1$ almost surely. Consequently, entropy alone implies no nontrivial universal lower or upper bound on $\mathbb{P}(C_{\mathrm{F}} = 1 \mid A_\tau)$.*

Proof. For $\epsilon \in (0, 1/K)$, define $p^{(\epsilon)} \coloneqq \Big(1 - (K-1)\epsilon,\ \underbrace{\epsilon, \ldots, \epsilon}_{K-1 \text{ entries}}\Big)$. Its coordinates are positive and sum to one. Its entropy is $H(p^{(\epsilon)}) = -\left[1 - (K-1)\epsilon\right]\log\left[1 - (K-1)\epsilon\right] - (K-1)\epsilon\log\epsilon$. As $\epsilon \downarrow 0$, $1 - (K-1)\epsilon \longrightarrow 1$, and $-t\log t \longrightarrow 0$ both as $t \downarrow 0$ and as $t \uparrow 1$. Therefore, $H(p^{(\epsilon)}) \longrightarrow 0$. Because $\tau > 0$, choose ϵ sufficiently small that $H(p^{(\epsilon)}) \leq \tau$. In both models, define $\mathbf{P} = p^{(\epsilon)}$ deterministically. In the first model, define $C_{\mathrm{F}} = 0$ almost surely. In the second model, define $C_{\mathrm{F}} = 1$ almost surely. The marginal law of $\mathbf{P}$ is therefore identical, and the low-entropy condition holds in both models, while the success probabilities are zero and one, respectively. □

Lemma 14 (Correctness guarantee under calibrated low entropy)**.** *Suppose Assumption 5, stated in Appendix B, holds. For every $\tau \in [0, \log K]$ such that $\mathbb{P}(A_\tau) > 0$, we have $\mathbb{P}\left(C_{\mathrm{F}} = 1 \mid A_\tau\right) \geq \left[g(\psi_K(\tau)) - \varepsilon_{\mathrm{cal}}\right]_+$, where $[x]_+ \coloneqq \max\{x, 0\}$.*

Proof. On the event $A_\tau = \{H(\mathbf{P}) \leq \tau\}$, Lemma 12 gives $P_{\max} \geq \psi_K(\tau)$. Since g is nondecreasing, $g(P_{\max}) \geq g(\psi_K(\tau))$ on A_τ. By the calibration assumption, $\eta_{\mathrm{F}}(\mathbf{P}) \geq g(P_{\max}) - \varepsilon_{\mathrm{cal}}$. Therefore, on A_τ, $\eta_{\mathrm{F}}(\mathbf{P}) \geq g(\psi_K(\tau)) - \varepsilon_{\mathrm{cal}}$. Because $A_\tau \in \sigma(\mathbf{P})$, the tower property gives

$$\mathbb{P}\left(C_{\mathrm{F}} = 1 \mid A_\tau\right) = \mathbb{E}\left[C_{\mathrm{F}} \mid A_\tau\right] = \mathbb{E}\left[\mathbb{E}\left[C_{\mathrm{F}} \mid \mathbf{P}\right] | A_\tau\right] = \mathbb{E}\left[\eta_{\mathrm{F}}(\mathbf{P}) \mid A_\tau\right] \geq g(\psi_K(\tau)) - \varepsilon_{\mathrm{cal}}.$$

Since a probability is nonnegative, taking the positive part gives the stated bound. □

Lemma 15 (Within-list calibration does not certify probe coverage)**.** *Under the above assumption,* $\mathbb{P}(Y = j \mid \mathbf{P}) = c(\mathbf{P})P_j$. *Let* $\widehat{j}(\mathbf{P}) \in \arg\max_{1\le j\le K} P_j$ *be chosen using a fixed tie-breaking rule. Then* $\mathbb{P}\left(Y = \widehat{j}(\mathbf{P}) \mid \mathbf{P}\right) = c(\mathbf{P})P_{\max}$. *Moreover,* $\mathbb{P}(Y \neq 0 \mid \mathbf{P}) = c(\mathbf{P})$. *Thus entropy controls* $P_{\max}$*, but does not control* $c(\mathbf{P})$.

Proof. By the conditional multiplication rule,

$$\mathbb{P}(Y = j \mid \mathbf{P}) = \mathbb{P}(Y \neq 0 \mid \mathbf{P}) \times \mathbb{P}\left(Y = j \mid \mathbf{P}, Y \neq 0\right) = c(\mathbf{P})P_j.$$

Setting $j = \widehat{j}(\mathbf{P})$ gives $\mathbb{P}\left(Y = \widehat{j}(\mathbf{P}) \mid \mathbf{P}\right) = c(\mathbf{P})P_{\widehat{j}(\mathbf{P})} = c(\mathbf{P})P_{\max}$. The final identity is the definition of $c(\mathbf{P})$. □

Lemma 16 (Dependence induced by Top-K selection)**.** *Suppose that* $X_1, \dots, X_n$ *are i.i.d. from a continuous, nondegenerate distribution satisfying* $\mathbb{E}[|X_1|] < \infty$*, with* $n \ge 2$*. Then* $X_{(n-1)}$ *and* $X_{(n)}$ *are not independent. Moreover,* $\mathbb{E}[X_{(n-1)}] < \mathbb{E}[X_{(n)}]$*. Consequently, whenever* $K \ge 2$*, the selected Top-*K *scores cannot in general be treated as independent random variables having a common mean.*

Proof. Because the common distribution is nondegenerate and continuous, there exists $c \in \mathbb{R}$ such that $0 < F(c) < 1$. Set $p := \mathbb{P}(X_1 > c) = 1 - F(c)$. Then $0 < p < 1$. Define $A := \{X_{(n-1)} > c\}$ and $B := \{X_{(n)} \le c\}$. The event A occurs if and only if at least two of the n observations exceed c. Hence

$$\begin{aligned}\mathbb{P}(A) &= 1 - \mathbb{P}(\text{no observation exceeds } c) - \mathbb{P}(\text{exactly one observation exceeds } c)\\ &= 1 - (1-p)^n - np(1-p)^{n-1}.\end{aligned}$$

Since $n \ge 2$ and $p \in (0, 1)$,

$$\mathbb{P}(A) > 0.$$

The event B occurs if and only if every observation is at most c. Thus $\mathbb{P}(B) = F(c)^n = (1-p)^n > 0$. However, A and B are disjoint. Indeed, on B, $X_{(n)} \le c$. Since $X_{(n-1)} \le X_{(n)}$, we also have $X_{(n-1)} \le c$, which contradicts A. Therefore, $A \cap B = \varnothing$ and $\mathbb{P}(A \cap B) = 0$. If $X_{(n-1)}$ and $X_{(n)}$ were independent, then the events A and B, which are respectively measurable with respect to $X_{(n-1)}$ and $X_{(n)}$, would be independent. This would imply $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) > 0$, contradicting $\mathbb{P}(A \cap B) = 0$. Thus the two order statistics are dependent. Because the common distribution is continuous, ties occur with probability zero. Hence $X_{(n-1)} < X_{(n)}$ almost surely. For every $j \in \{1, \dots, n\}$, $|X_{(j)}| \le \max_{1\le i\le n} |X_i| \le \sum_{i=1}^n |X_i|$. Therefore, $\mathbb{E}[|X_{(j)}|] \le \sum_{i=1}^n \mathbb{E}[|X_i|] = n\mathbb{E}[|X_1|] < \infty$, so in particular $X_{(n-1)}$ and $X_{(n)}$ are integrable. Thus $X_{(n)} - X_{(n-1)} > 0$ almost surely, and consequently $\mathbb{E}\left[X_{(n)} - X_{(n-1)}\right] > 0$. Equivalently, $\mathbb{E}[X_{(n-1)}] < \mathbb{E}[X_{(n)}]$. □

Lemma 17 (Variational representation of the Top-K mean)**.** *For every deterministic vector* $x = (x_1, \dots, x_n) \in [\ell, u]^n$*, let* $x_{[1]} \ge x_{[2]} \ge \cdots \ge x_{[n]}$ *be its decreasing rearrangement. Then*

$$\frac{1}{K}\sum_{j=1}^K x_{[j]} = \min_{t\in[\ell,u]} \left\{ t + \frac{1}{K}\sum_{i=1}^n (x_i - t)_+ \right\}.$$

Proof. Define $\Phi_x(t) := t + \frac{1}{K}\sum_{i=1}^n (x_i - t)_+$. For every real number z, $z_+ \ge z$. Therefore, for every $t \in [\ell, u]$,

$$\begin{aligned}\Phi_x(t) &= t + \frac{1}{K}\sum_{i=1}^n (x_i - t)_+ \ge t + \frac{1}{K}\sum_{j=1}^K (x_{[j]} - t)_+\\ &\ge t + \frac{1}{K}\sum_{j=1}^K (x_{[j]} - t) = t + \frac{1}{K}\sum_{j=1}^K x_{[j]} - \frac{1}{K}\sum_{j=1}^K t = t + \frac{1}{K}\sum_{j=1}^K x_{[j]} - t = \frac{1}{K}\sum_{j=1}^K x_{[j]}.\end{aligned}$$

Hence

$$\min_{t\in[\ell,u]} \Phi_x(t) \geq \frac{1}{K}\sum_{j=1}^{K} x_{[j]}. \tag{56}$$

We next exhibit a threshold attaining equality. Suppose first that $K < n$. Choose any $t \in [x_{[K+1]}, x_{[K]}]$. Then, for $j \leq K$, $x_{[j]} \geq x_{[K]} \geq t$, whereas for $j \geq K+1$, $x_{[j]} \leq x_{[K+1]} \leq t$. Consequently,

$$(x_{[j]} - t)_+ = \begin{cases} x_{[j]} - t, & j \leq K, \\ 0, & j \geq K+1. \end{cases}$$

Therefore,

$$\Phi_x(t) = t + \frac{1}{K}\sum_{j=1}^{K}(x_{[j]} - t) = t + \frac{1}{K}\sum_{j=1}^{K} x_{[j]} - t = \frac{1}{K}\sum_{j=1}^{K} x_{[j]}.$$

Suppose next that $K = n$. Choose $t = \ell$. Since $x_i \geq \ell$, $(x_i - \ell)_+ = x_i - \ell$. Hence

$$\Phi_x(\ell) = \ell + \frac{1}{n}\sum_{i=1}^{n}(x_i - \ell) = \ell + \frac{1}{n}\sum_{i=1}^{n} x_i - \ell = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Thus equality is attained in all cases. Combining this with (56) proves the result. □

Lemma 18 (Quantile representation of the population upper-tail functional)**.** *For $\gamma \in (0,1)$, define $q_\gamma := Q(1-\gamma)$. Then $\mathbb{P}(X > q_\gamma) \leq \gamma \leq \mathbb{P}(X \geq q_\gamma)$. The value q_γ minimizes* (11)*, and*

$$U_\gamma(F) = q_\gamma + \frac{1}{\gamma}\mathbb{E}[(X - q_\gamma)_+] = \frac{1}{\gamma}\Big\{\mathbb{E}[X\mathbf{1}\{X > q_\gamma\}] + q_\gamma[\gamma - \mathbb{P}(X > q_\gamma)]\Big\} \tag{57}$$

$$= \frac{1}{\gamma}\int_{1-\gamma}^{1} Q(v)\,dv. \tag{58}$$

For $\gamma = 1$, $U_1(F) = \mathbb{E}[X]$. If $\mathbb{P}(X = q_\gamma) = 0$, $\mathbb{P}(X > q_\gamma) = \gamma$ and $U_\gamma(F) = \mathbb{E}[X \mid X > q_\gamma]$.

Proof. By the definition $q_\gamma = \inf\{x : F(x) \geq 1-\gamma\}$, right continuity of F gives $F(q_\gamma) \geq 1-\gamma$. Therefore,

$$\begin{aligned} \mathbb{P}(X > q_\gamma) &= 1 - F(q_\gamma) \\ &\leq \gamma. \end{aligned} \tag{59}$$

For every $x < q_\gamma$, $F(x) < 1-\gamma$. Taking $x \uparrow q_\gamma$, $F(q_\gamma-) := \lim_{x\uparrow q_\gamma} F(x) \leq 1-\gamma$. Hence

$$\mathbb{P}(X \geq q_\gamma) = 1 - F(q_\gamma-) \geq \gamma. \tag{60}$$

Define $\Phi(t) := t + \frac{1}{\gamma}\mathbb{E}[(X-t)_+]$. We first compare q_γ with a larger threshold. Let $t = q_\gamma + h$ and $h > 0$, with $t \leq u$. For every real x, $(x - q_\gamma - h)_+ - (x - q_\gamma)_+ \geq -h\mathbf{1}\{x > q_\gamma\}$. Indeed:

- If $x \leq q_\gamma$, both positive parts are zero, so both sides equal zero.
- If $q_\gamma < x \leq q_\gamma + h$, the left-hand side is $0 - (x - q_\gamma) = -(x - q_\gamma) \geq -h$.
- If $x > q_\gamma + h$, the left-hand side is $(x - q_\gamma - h) - (x - q_\gamma) = -h$.

Therefore,

$$\begin{aligned} \Phi(q_\gamma + h) - \Phi(q_\gamma) &= h + \frac{1}{\gamma}\mathbb{E}\left[(X - q_\gamma - h)_+ - (X - q_\gamma)_+\right] \\ &\geq h - \frac{h}{\gamma}\mathbb{P}(X > q_\gamma) = h\left[1 - \frac{\mathbb{P}(X > q_\gamma)}{\gamma}\right] \geq 0, \end{aligned}$$

where the final inequality uses (59). We next compare q_γ with a smaller threshold. Let $t = q_\gamma - h$ and $h > 0$, with $t \geq \ell$. For every real x, $(x - q_\gamma + h)_+ - (x - q_\gamma)_+ \geq h\mathbf{1}\{x \geq q_\gamma\}$. Indeed:

- If $x < q_\gamma$, the right-hand side is zero and the left-hand side is nonnegative.
- If $x \geq q_\gamma$, both positive parts are active and their difference equals h.

Thus

$$\begin{aligned}\Phi(q_\gamma - h) - \Phi(q_\gamma) &= -h + \frac{1}{\gamma}\mathbb{E}\left[(X - q_\gamma + h)_+ - (X - q_\gamma)_+\right] \\ &\geq -h + \frac{h}{\gamma}\mathbb{P}(X \geq q_\gamma) = h\left[\frac{\mathbb{P}(X \geq q_\gamma)}{\gamma} - 1\right] \geq 0,\end{aligned}$$

where the final inequality uses (60). Every $t \in [\ell, u]$ can be written as either $q_\gamma + h$, $q_\gamma - h$, or q_γ. Hence $\Phi(t) \geq \Phi(q_\gamma)$ for every t, and q_γ is a minimizer. We now derive the atom-adjusted representation. Since $(X - q_\gamma)_+ = (X - q_\gamma)\mathbf{1}\{X > q_\gamma\}$, we have

$$\begin{aligned}U_\gamma(F) &= q_\gamma + \frac{1}{\gamma}\mathbb{E}\left[(X - q_\gamma)\mathbf{1}\{X > q_\gamma\}\right] \\ &= q_\gamma + \frac{1}{\gamma}\Big\{\mathbb{E}[X\mathbf{1}\{X > q_\gamma\}] - q_\gamma\mathbb{P}(X > q_\gamma)\Big\} \\ &= \frac{1}{\gamma}\Big\{\gamma q_\gamma + \mathbb{E}[X\mathbf{1}\{X > q_\gamma\}] - q_\gamma\mathbb{P}(X > q_\gamma)\Big\} \\ &= \frac{1}{\gamma}\Big\{\mathbb{E}[X\mathbf{1}\{X > q_\gamma\}] + q_\gamma[\gamma - \mathbb{P}(X > q_\gamma)]\Big\}.\end{aligned}$$

It remains to prove the quantile-integral form. Let $U \sim \mathrm{Uniform}(0, 1)$. Then the generalized inverse transform implies $Q(U) \sim F$. Moreover, $\{v : Q(v) > q_\gamma\} = (F(q_\gamma), 1)$ up to endpoints of Lebesgue measure zero. Indeed, if $v \leq F(q_\gamma)$, then q_γ belongs to the set $\{x : F(x) \geq v\}$, and therefore $Q(v) \leq q_\gamma$. If $v > F(q_\gamma)$, then for every $x \leq q_\gamma$, $F(x) \leq F(q_\gamma) < v$, so no $x \leq q_\gamma$ can satisfy $F(x) \geq v$, and hence $Q(v) > q_\gamma$. It follows that

$$\int_{F(q_\gamma)}^1 Q(v)\,dv = \mathbb{E}\left[Q(U)\mathbf{1}\{Q(U) > q_\gamma\}\right] = \mathbb{E}\left[X\mathbf{1}\{X > q_\gamma\}\right].$$

Since $F(q_\gamma-) \leq 1 - \gamma \leq F(q_\gamma)$, the quantile function equals q_γ almost everywhere on $(1 - \gamma, F(q_\gamma)]$. Therefore,

$$\begin{aligned}\int_{1-\gamma}^1 Q(v)\,dv &= q_\gamma\left[F(q_\gamma) - (1 - \gamma)\right] + \int_{F(q_\gamma)}^1 Q(v)\,dv \\ &= q_\gamma\left[\gamma - \{1 - F(q_\gamma)\}\right] + \mathbb{E}\left[X\mathbf{1}\{X > q_\gamma\}\right] \\ &= q_\gamma\left[\gamma - \mathbb{P}(X > q_\gamma)\right] + \mathbb{E}\left[X\mathbf{1}\{X > q_\gamma\}\right].\end{aligned}$$

Division by γ gives (58). For $\gamma = 1$, choose $t = \ell$ in the variational definition. Since $X \geq \ell$, $(X-\ell)_+ = X-\ell$. Hence $U_1(F) \leq \ell+\mathbb{E}[X-\ell] = \mathbb{E}[X]$. Conversely, for every t, $(X-t)_+ \geq X-t$, so $t+\mathbb{E}[(X-t)_+] \geq t+\mathbb{E}[X-t] = \mathbb{E}[X]$. Therefore, $U_1(F) = \mathbb{E}[X]$. Finally, if $\mathbb{P}(X = q_\gamma) = 0$, then $\mathbb{P}(X > q_\gamma) = \mathbb{P}(X \geq q_\gamma)$. The sandwich $\mathbb{P}(X > q_\gamma) \leq \gamma \leq \mathbb{P}(X \geq q_\gamma)$ therefore forces $\mathbb{P}(X > q_\gamma) = \gamma$. Substituting into (57), $U_\gamma(F) = \frac{1}{\gamma}\mathbb{E}\left[X\mathbf{1}\{X > q_\gamma\}\right] = \mathbb{E}[X \mid X > q_\gamma]$. □

Lemma 19 (Upper-tail selection bias)**.** *For every integrable score distribution F and every $\gamma \in (0, 1]$, $U_\gamma(F) \geq \mathbb{E}[X]$. If $\gamma < 1$ and X is nondegenerate, then $U_\gamma(F) > \mathbb{E}[X]$.*

Proof. By Lemma 18, $U_\beta(F) = \frac{1}{\beta}\int_{1-\beta}^1 Q(v)\,dv$ for $\beta \in (0, 1]$, where the case $\beta = 1$ equals $\mathbb{E}[X]$. The case $\gamma = 1$ follows from $U_1(F) = \mathbb{E}[X]$. Assume therefore that $0 < \gamma < 1$. Define the lower-part average $L_\gamma := \frac{1}{1-\gamma}\int_0^{1-\gamma} Q(v)\,dv$ and the upper-tail average $B_\gamma := \frac{1}{\gamma}\int_{1-\gamma}^1 Q(v)\,dv = U_\gamma(F)$. Because Q is nondecreasing, for every $v \in (0, 1 - \gamma)$ and $w \in (1 - \gamma, 1)$, we have $Q(v) \leq Q(w)$. Therefore,

$$B_\gamma - L_\gamma = \frac{1}{\gamma(1 - \gamma)}\int_0^{1-\gamma}\int_{1-\gamma}^1 \left[Q(w) - Q(v)\right]\,dw\,dv \geq 0. \tag{61}$$

Hence $L_\gamma \le B_\gamma$. The ordinary mean decomposes as

$$\begin{aligned}\mathbb{E}[X] &= \int_0^1 Q(v)\,dv = \int_0^{1-\gamma} Q(v)\,dv + \int_{1-\gamma}^1 Q(v)\,dv \\ &= (1-\gamma)L_\gamma + \gamma B_\gamma \le (1-\gamma)B_\gamma + \gamma B_\gamma = B_\gamma = U_\gamma(F).\end{aligned}$$

It remains to establish strictness for a nondegenerate distribution. Suppose equality holds: $U_\gamma(F) = \mathbb{E}[X]$. Then the preceding convex-combination identity and $L_\gamma \le B_\gamma$ imply $L_\gamma = B_\gamma$. Consequently, the nonnegative double integral in (61) is zero: $\int_0^{1-\gamma}\int_{1-\gamma}^1 [Q(w) - Q(v)]\,dw\,dv = 0$. Thus $Q(w) = Q(v)$ for Lebesgue-almost every pair $(v, w) \in (0, 1-\gamma)\times(1-\gamma, 1)$. By Fubini's theorem, there exists $v_0 \in (0, 1-\gamma)$ such that $Q(w) = Q(v_0)$ for almost every $w \in (1-\gamma, 1)$. Likewise, choosing such a w_0, we have $Q(v) = Q(w_0)$ for almost every $v \in (0, 1-\gamma)$. Therefore, Q is equal to the same constant almost everywhere on $(0, 1)$. This implies that X is degenerate. Hence, if X is nondegenerate and $\gamma < 1$, equality is impossible and the inequality is strict. □

Lemma 20 (The upper-tail target is monotone in the tail fraction)**.** *Let $0 < \gamma_1 < \gamma_2 \le 1$. Then $U_{\gamma_1}(F) \ge U_{\gamma_2}(F)$.*

Proof. By Lemma 18, the upper-tail functional has the quantile-integral representation used below. Define $B := \frac{1}{\gamma_1}\int_{1-\gamma_1}^1 Q(v)\,dv$ and $C := \frac{1}{\gamma_2-\gamma_1}\int_{1-\gamma_2}^{1-\gamma_1} Q(v)\,dv$. Because Q is nondecreasing, every value of Q on $[1-\gamma_2, 1-\gamma_1]$ is no larger than every value on $[1-\gamma_1, 1]$. Therefore, $C \le B$. Now decompose the upper γ_2-tail integral:

$$\begin{aligned}U_{\gamma_2}(F) &= \frac{1}{\gamma_2}\int_{1-\gamma_2}^1 Q(v)\,dv = \frac{1}{\gamma_2}\left[\int_{1-\gamma_2}^{1-\gamma_1} Q(v)\,dv + \int_{1-\gamma_1}^1 Q(v)\,dv\right] \\ &= \frac{\gamma_2-\gamma_1}{\gamma_2}C + \frac{\gamma_1}{\gamma_2}B \le \frac{\gamma_2-\gamma_1}{\gamma_2}B + \frac{\gamma_1}{\gamma_2}B = B = U_{\gamma_1}(F).\end{aligned}$$

□

Lemma 21 (Standard self-bounding concentration inequality)**.** *Let $Y_1, \dots, Y_n$ be independent random variables and let $Z = f(Y_1, \dots, Y_n) \ge 0$. Suppose that, for every i, there exists a function f_i, independent of Y_i, such that $0 \le Z - f_i(Y_{-i}) \le 1$ and $\sum_{i=1}^n [Z - f_i(Y_{-i})] \le Z$. Then, for every $t \ge 0$, $\mathbb{P}(Z - \mathbb{E}Z \ge t) \le \exp\left\{-\frac{t^2}{2\mathbb{E}Z + 2t/3}\right\}$ and $\mathbb{P}(\mathbb{E}Z - Z \ge t) \le \exp\left\{-\frac{t^2}{2\mathbb{E}Z}\right\}$, and $\mathrm{Var}(Z) \le \mathbb{E}Z$.*

Lemma 22 (Dvoretzky–Kiefer–Wolfowitz inequality)**.** *Let $F_n(t) := \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{X_i \le t\}$ be the empirical distribution function. Then, for every $\eta > 0$, $\mathbb{P}\left(\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| > \eta\right) \le 2e^{-2n\eta^2}$.*

Lemma 23 (Pathwise monotonicity in corpus size)**.** *Fix $K \ge 1$. For an arbitrary deterministic sequence $x_1, x_2, \dots$, define $\widehat{M}_{n,K}$ from the first n values whenever $n \ge K$. Then $\widehat{M}_{n+1,K} \ge \widehat{M}_{n,K}$ for every $n \ge K$.*

Proof. The Top-K sum among the first n values is $T_{n,K} := \max_{\substack{I\subseteq\{1,\dots,n\}\\|I|=K}} \sum_{i\in I} x_i$. The collection of feasible subsets for $T_{n,K}$ is contained in the collection of feasible subsets for $T_{n+1,K}$, because every subset of $\{1, \dots, n\}$ is also a subset of $\{1, \dots, n+1\}$. Therefore, $T_{n+1,K} \ge T_{n,K}$. Dividing by the same positive value K, $\widehat{M}_{n+1,K} = \frac{T_{n+1,K}}{K} \ge \frac{T_{n,K}}{K} = \widehat{M}_{n,K}$. □

Lemma 24 (Standard upper-tail facts for the Top-K mean)**.** *Let $X_1, X_2, \dots$ be i.i.d. with law F supported on $[\ell, u]$, where $\ell < u$, and put $\Delta_s := u - \ell$. For integers $1 \le K \le n$, use $\widehat{M}_{n,K}$ and $U_{K/n}(F)$ defined in Section 9. i) The Top-K mean admits the exact representation $\widehat{M}_{n,K} = \min_{t\in[\ell,u]}\left\{t + \frac{1}{K}\sum_{i=1}^n (X_i - t)_+\right\}$; ii) Its population target is $U_{K/n}(F) = \frac{n}{K}\int_{1-K/n}^1 Q(v)\,dv$, which equals $\mathbb{E}[X]$ when $K = n$ and strictly exceeds it when $K < n$ and X is nondegenerate; iii) For every $\varepsilon > 0$, $\mathbb{P}\left(\left|\widehat{M}_{n,K} - U_{K/n}(F)\right| > \varepsilon\right) \le 2\exp\left\{-\frac{2K^2\varepsilon^2}{n\Delta_s^2}\right\}$; iv) If K is fixed, then $\widehat{M}_{n,K} \to \operatorname{ess\,sup} X$ almost surely; v) If $K_n/n \to \gamma \in (0, 1]$, then $\widehat{M}_{n,K_n} \to U_\gamma(F)$ almost surely.*

Proof. Item i follows directly from Lemma 17. Item ii follows from Lemma 18 after substituting $\gamma = K/n$: $U_{K/n}(F) = \frac{1}{K/n}\int_{1-K/n}^{1} Q(v)\,dv = \frac{n}{K}\int_{1-K/n}^{1} Q(v)\,dv$. The equality and strict inequality relative to $\mathbb{E}[X]$ then follow from Lemma 19, with $\gamma = K/n$. Item iii is the second form of Theorem 9. Item iv is Theorem 10, and item v is Theorem 11. The cited lemmas are proved in this section and the cited theorems in Appendix E, so all five claims follow under the bounded i.i.d. assumptions of the lemma. □

Lemma 25 (Two mean thresholds and a unique increasing frontier)**.** *Under Assumption 3, stated in Appendix B, there exist unique numbers $\theta_{\mathrm{low}}, \theta_{\mathrm{high}} \in (m_-, m_+)$ satisfying $d(\theta_{\mathrm{low}}, h_-) = 0$ and $d(\theta_{\mathrm{high}}, h_+) = 0$. They satisfy $\theta_{\mathrm{low}} < \theta_{\mathrm{high}}$. The Bayes-optimal rule has the following structure.*

i) If $m < \theta_{\mathrm{low}}$, then $d(m, h) > 0$ for every $h \in [h_-, h_+]$, so Recollection is uniquely optimal.

ii) If $m \in (\theta_{\mathrm{low}}, \theta_{\mathrm{high}})$, then there exists a unique $b(m) \in (h_-, h_+)$ such that $d(m, b(m)) = 0$. Moreover,

$$a^\star(m, h) = \begin{cases} \mathrm{F}, & h \le b(m), \\ \mathrm{R}, & h > b(m). \end{cases}$$

iii) If $m > \theta_{\mathrm{high}}$, then $d(m, h) < 0$ for every $h \in [h_-, h_+]$, so Familiarity is uniquely optimal.

The interior boundary $b : (\theta_{\mathrm{low}}, \theta_{\mathrm{high}}) \to (h_-, h_+)$ belongs to C^1 and satisfies

$$b'(m) = -\frac{\partial_m d(m, b(m))}{\partial_h d(m, b(m))}.$$

In particular, b is strictly increasing, and

$$\frac{\mu_m}{L_h} \le b'(m) \le \frac{L_m}{\mu_h}. \tag{62}$$

Finally, $\lim_{m \downarrow \theta_{\mathrm{low}}} b(m) = h_-$ and $\lim_{m \uparrow \theta_{\mathrm{high}}} b(m) = h_+$.

Proof. The derivative bounds in Assumption 3 imply Assumption 1; hence Lemma 8 applies. We now establish the stronger threshold and regularity conclusions. Define $g_-(m) := d(m, h_-)$. The function g_- is continuous. By assumption, $g_-(m_-) = d(m_-, h_-) > 0$. Since d is increasing in entropy, $d(m_+, h_-) \le d(m_+, h_+) < 0$. Hence $g_-(m_+) < 0$. The intermediate value theorem gives at least one $\theta_{\mathrm{low}} \in (m_-, m_+)$ such that $g_-(\theta_{\mathrm{low}}) = 0$. The root is unique because

$$g_-'(m) = \partial_m d(m, h_-) \le -\mu_m < 0.$$

Thus g_- is strictly decreasing.

Next, define $g_+(m) := d(m, h_+)$. Since d is increasing in entropy, $d(m_-, h_+) \ge d(m_-, h_-) > 0$. Thus $g_+(m_-) > 0$. By assumption, $g_+(m_+) = d(m_+, h_+) < 0$. The intermediate value theorem gives a root $\theta_{\mathrm{high}} \in (m_-, m_+)$. It is unique because

$$g_+'(m) = \partial_m d(m, h_+) \le -\mu_m < 0.$$

We prove $\theta_{\mathrm{low}} < \theta_{\mathrm{high}}$. At $m = \theta_{\mathrm{low}}$, $d(\theta_{\mathrm{low}}, h_-) = 0$. By the fundamental theorem of calculus,

$$\begin{aligned} d(\theta_{\mathrm{low}}, h_+) - d(\theta_{\mathrm{low}}, h_-) &= \int_{h_-}^{h_+} \partial_h d(\theta_{\mathrm{low}}, u)\,du \\ &\ge \int_{h_-}^{h_+} \mu_h\,du \\ &= \mu_h(h_+ - h_-) \\ &> 0. \end{aligned}$$

Therefore,

$$d(\theta_{\mathrm{low}}, h_+) > 0.$$

The strictly decreasing function $m \longmapsto d(m, h_+)$ has its unique zero at θ_{high}. Since it is still positive at θ_{low}, its zero must lie strictly to the right: $\theta_{\mathrm{low}} < \theta_{\mathrm{high}}$.

Suppose $m < \theta_{\text{low}}$. Because g_- is strictly decreasing and vanishes at θ_{low}, $d(m, h_-) > 0$. For every $h \geq h_-$, monotonicity in entropy gives $d(m, h) \geq d(m, h_-) > 0$. Hence Recollection is uniquely optimal.

Suppose $m > \theta_{\text{high}}$. Because g_+ is strictly decreasing and vanishes at θ_{high}, $d(m, h_+) < 0$. For every $h \leq h_+$, $d(m, h) \leq d(m, h_+) < 0$. Hence Familiarity is uniquely optimal.

Now suppose $m \in (\theta_{\text{low}}, \theta_{\text{high}})$. Since $m > \theta_{\text{low}}$ and $m \mapsto d(m, h_-)$ is strictly decreasing, $d(m, h_-) < d(\theta_{\text{low}}, h_-) = 0$. Likewise, since $m < \theta_{\text{high}}$, $d(m, h_+) > d(\theta_{\text{high}}, h_+) = 0$. Therefore, $d(m, h_-) < 0 < d(m, h_+)$.

By continuity in h, there exists at least one $b(m) \in (h_-, h_+)$ such that $d(m, b(m)) = 0$. The root is unique because $\partial_h d(m, h) \geq \mu_h > 0$, so $h \mapsto d(m, h)$ is strictly increasing.

It follows that

$$d(m, h) < 0 \quad \text{for } h < b(m),$$

and

$$d(m, h) > 0 \quad \text{for } h > b(m),$$

which proves the stated optimal rule.

At every point $(m, b(m))$,

$$\partial_h d(m, b(m)) \geq \mu_h > 0.$$

The implicit function theorem therefore yields a locally C^1 solution. Uniqueness of the root implies that these local solutions coincide on overlaps, so $b \in C^1$ on the whole middle interval.

Differentiating $d(m, b(m)) = 0$ with respect to m,

$$\begin{aligned} 0 &= \frac{d}{dm} d(m, b(m)) \\ &= \partial_m d(m, b(m)) + \partial_h d(m, b(m)) b'(m). \end{aligned}$$

Therefore,

$$b'(m) = -\frac{\partial_m d(m, b(m))}{\partial_h d(m, b(m))}.$$

Using $\mu_m \leq -\partial_m d \leq L_m$ and $\mu_h \leq \partial_h d \leq L_h$, we obtain

$$b'(m) \geq \frac{\mu_m}{L_h}$$

and

$$b'(m) \leq \frac{L_m}{\mu_h}.$$

It remains to prove the endpoint limits. Because b is increasing and bounded in (h_-, h_+), the one-sided limit $\ell_- := \lim_{m \downarrow \theta_{\text{low}}} b(m)$ exists in $[h_-, h_+]$. Choose any sequence $m_r \downarrow \theta_{\text{low}}$. Since $d(m_r, b(m_r)) = 0$, continuity gives $d(\theta_{\text{low}}, \ell_-) = 0$. At $m = \theta_{\text{low}}$, the function $h \mapsto d(\theta_{\text{low}}, h)$ is strictly increasing and satisfies $d(\theta_{\text{low}}, h_-) = 0$. Its zero is therefore unique, so $\ell_- = h_-$.

The upper endpoint is analogous. The limit $\ell_+ := \lim_{m \uparrow \theta_{\text{high}}} b(m)$ exists and continuity gives $d(\theta_{\text{high}}, \ell_+) = 0$. Since $d(\theta_{\text{high}}, h_+) = 0$ and the zero is unique, $\ell_+ = h_+$. □

Lemma 26 (Support-aware and local quadratic regret bounds for rectangular gating)**.** *Under Assumptions 3 and 4, stated in Appendix B, for every $\tau \in [h_-, h_+]$,*

$$\mu_h \mathfrak{Q}(\tau) \leq \mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}} \leq L_h \mathfrak{Q}(\tau). \tag{63}$$

Moreover, fix τ and suppose that there exist constants $0 < \underline{f}_\tau \leq \overline{f}_\tau < \infty$ such that, for P_M-almost every $m \in (\theta_{\text{low}}, \theta_{\text{high}})$,

$$\underline{f}_\tau \leq f_{H|M}(h \mid m) \leq \overline{f}_\tau \quad \textit{for Lebesgue-a.e. } h \in \mathcal{J}_\tau(m). \tag{64}$$

Then

$$\frac{\mu_h \underline{f}_\tau}{2} \mathbb{E}\left[(b(M) - \tau)^2 \mathbf{1}_I\right] \leq \mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}} \leq \frac{L_h \overline{f}_\tau}{2} \mathbb{E}\left[(b(M) - \tau)^2 \mathbf{1}_I\right]. \tag{65}$$

Proof. Fix $m \in (\theta_{\mathrm{low}}, \theta_{\mathrm{high}})$ for which the conditional density exists, and write $b := b(m)$. We have $d(m, b) = 0$. Suppose first that $h \geq b$. By the fundamental theorem of calculus,

$$d(m, h) - d(m, b) = \int_b^h \partial_h d(m, u)\, du.$$

Since $d(m, b) = 0$,

$$d(m, h) = \int_b^h \partial_h d(m, u)\, du.$$

Using $\mu_h \leq \partial_h d(m, u) \leq L_h$, we obtain

$$\mu_h(h - b) \leq d(m, h) \leq L_h(h - b).$$

Because $h \geq b$, $d(m, h) \geq 0$, and therefore

$$\mu_h|h - b| \leq |d(m, h)| \leq L_h|h - b|.$$

Suppose next that $h \leq b$. Then

$$d(m, b) - d(m, h) = \int_h^b \partial_h d(m, u)\, du.$$

Using $d(m, b) = 0$,

$$-d(m, h) = \int_h^b \partial_h d(m, u)\, du.$$

Hence

$$\mu_h(b - h) \leq -d(m, h) \leq L_h(b - h).$$

Because $d(m, h) \leq 0$, $|d(m, h)| = -d(m, h)$, and therefore

$$\mu_h|h - b| \leq |d(m, h)| \leq L_h|h - b|.$$

Combining the two cases,

$$\mu_h|h - b(m)| \leq |d(m, h)| \leq L_h|h - b(m)| \tag{66}$$

for every $h \in [h_-, h_+]$.

Define the conditional gate regret

$$\Delta_\tau(m) := \int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |d(m, h)| f_{H|M}(h \mid m)\, dh.$$

Since the conditional density is nonnegative, (66) gives

$$\Delta_\tau(m) \geq \mu_h \int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |h - b(m)| f_{H|M}(h \mid m)\, dh, \tag{67}$$

$$\Delta_\tau(m) \leq L_h \int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |h - b(m)| f_{H|M}(h \mid m)\, dh. \tag{68}$$

Lemma 9 gives

$$\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}} = \mathbb{E}\left[\mathbf{1}_I \Delta_\tau(M)\right]. \tag{69}$$

Taking expectations in (67) and (68), and using the definition of $\mathfrak{Q}(\tau)$, proves (63).

Now fix a threshold τ for which the local bounds (64) hold. Combining those bounds with (66) yields

$$\Delta_\tau(m) \geq \mu_h \underline{f}_\tau \int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |h - b(m)|\, dh, \tag{70}$$

$$\Delta_\tau(m) \leq L_h \overline{f}_\tau \int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |h - b(m)|\, dh. \tag{71}$$

We calculate the common integral exactly.

If $\tau \geq b(m)$, then

$$\begin{aligned}\int_{b(m)}^{\tau} |h - b(m)|\, dh &= \int_{b(m)}^{\tau} (h - b(m))\, dh \\ &= \left[\frac{(h - b(m))^2}{2}\right]_{h=b(m)}^{h=\tau} \\ &= \frac{(\tau - b(m))^2}{2}.\end{aligned}$$

If $\tau < b(m)$, then

$$\begin{aligned}\int_{\tau}^{b(m)} |h - b(m)|\, dh &= \int_{\tau}^{b(m)} (b(m) - h)\, dh \\ &= \left[b(m)h - \frac{h^2}{2}\right]_{h=\tau}^{h=b(m)} \\ &= \frac{b(m)^2}{2} - b(m)\tau + \frac{\tau^2}{2} \\ &= \frac{(b(m) - \tau)^2}{2}.\end{aligned}$$

Thus, in both cases,

$$\int_{\min\{b(m),\tau\}}^{\max\{b(m),\tau\}} |h - b(m)|\, dh = \frac{(b(m) - \tau)^2}{2}.$$

Substitution into (70) and (71) gives

$$\Delta_\tau(m) \geq \frac{\mu_h \underline{f}_\tau}{2}(b(m) - \tau)^2,$$

$$\Delta_\tau(m) \leq \frac{L_h \overline{f}_\tau}{2}(b(m) - \tau)^2.$$

Multiplying the conditional bounds by $\mathbf{1}_I$ and taking expectations in (69) proves the local quadratic inequalities. □

D COROLLARIES

Corollary 5 (Bag separation under perturbations of repeated embeddings)**.** *Fix $\alpha \in [0, 1)$, $d \geq 3$, $\Delta_c \in (0, 1)$, integers $F \geq 1$, $R \geq 2$, and*

$$K \geq H_{F,R} + 1, \qquad H_{F,R} := F\left(\frac{R(R+1)}{2} - 1\right). \tag{72}$$

Under the contract of Theorem 2, there is a pair of its $n = K + 1$-memory constructions and an explicit radius $\varepsilon_0 > 0$ with the following property. Keep the query fixed and perturb each memory by less than ε_0 in Euclidean norm, retaining unit norms, a common relevant memory, and equality of the two complete indexed initial score profiles. Every such pair has failure losses $(1, 0)$ and $(1, 1)$, executes all R rounds, and returns its whole Bag. It therefore has the same equal-prior Bayes gap and score-only minimax regret as the theorem. Such perturbations can make all embeddings within each corpus distinct, even in dimension three.

For $R = 2$, (72) is exactly $K \geq 2F + 1$. For $R > 2$ it is a stronger sufficient condition than the theorem's linear budget bound; robustness throughout that smaller range is not asserted. The neighborhood is relative to the unit-norm and score-matching constraints, and need not preserve the trace.

Proof. Choose $0 < a < s < 1$ and set

$$A = 1 + \alpha + (1-\alpha)s, \quad b = (1-\alpha)\sqrt{1-s^2} > 0, \quad D = \sqrt{A^2 + b^2} \geq 1.$$

Choose $a < \eta < \min\{s, a + (b/A)\sqrt{1-a^2}\}$. Use $q = e_1$ and the original vectors

$$z_\pm = se_1 \pm \sqrt{1-s^2}e_2, \qquad w = \eta e_1 + \sqrt{1-\eta^2}e_3, \qquad y = ae_1 + \sqrt{1-a^2}e_2.$$

There are F seed indices, $K - F$ distractor indices, and the target index. Define the strictly positive gaps

$$\begin{aligned} m_0 &= \min\{s - \eta, \eta - a\}, \\ m_1 &= \frac{1}{D}\min\big\{A(s-\eta) + b\sqrt{1-s^2}, \\ &\qquad b\sqrt{1-a^2} - A(\eta - a),\ A(\eta - a) + b\sqrt{1-a^2}\big\}. \end{aligned}$$

The three terms in m_1 bound, respectively, seed over distractor in both first child queries, target over distractor in the plus query, and distractor over target in the minus query. One valid radius is

$$\varepsilon_0 = \min\left\{\frac{a}{2}, \frac{m_0}{4}, \frac{m_1}{20}, \frac{\eta - a}{8 + 16/\eta}\right\} > 0. \tag{73}$$

Let the maximum perturbation norm be $\varepsilon < \varepsilon_0$. Every memory still has positive first coordinate, at least $a - \varepsilon > a/2$. Initial comparisons change by at most $2\varepsilon < m_0/2$. Thus Familiarity still returns all K nonrelevant indices, and round zero still retrieves all F seeds.

The raw initial seed centroid differs from $z_\pm$ by at most ε. For nonzero v, v_0, the triangle inequality gives

$$\left\| \frac{v}{\|v\|} - \frac{v_0}{\|v_0\|} \right\| \leq \frac{2\|v - v_0\|}{\|v_0\|}.$$

Indeed, insert $v/\|v_0\|$ between the two terms and use $|\|v\| - \|v_0\|| \leq \|v - v_0\|$. The normalized centroid therefore differs by at most 2ε, and the first child query differs from $q'_\pm = (Ae_1 \pm be_2)/D$ by at most $4(1-\alpha)\varepsilon/D \leq 4\varepsilon$. For any memory its score changes by at most 5ε; a pairwise score difference changes by at most 10ε. All comparisons defining m_1 consequently remain strict. Writing I_r for the indices retrieved at round r, both runs have $I_0 \subset I_1$. The plus run's I_1 contains the target and all F seeds, and the minus run's I_1 contains all seeds and excludes the target. This remains true for $F = 1$, since $F + 1 \leq 2F$.

We next exclude the target from every later minus retrieval without requiring the tied distractor order to remain fixed. Put $h = \varepsilon/(\eta - \varepsilon) \leq 2\varepsilon/\eta$. Every perturbed nonrelevant vector in the minus corpus has first coordinate at least $\eta - \varepsilon$, second coordinate at most ε, and third coordinate at least $-\varepsilon$. Hence its normalized nonempty centroid g satisfies $g_1 > 0$, $g_2 \leq h$, and $g_3 \geq -h$: the raw centroid norm is at least $\eta - \varepsilon$. For any unit parent with $x_1 > 0$, $x_2 \leq h$, and $x_3 \geq -h$, the update numerator $v = \alpha x + (1-\alpha)g + q$ has

$$v_1 \geq 1, \qquad \|v\| \leq 2, \qquad v_2 \leq h, \qquad v_3 \geq -h.$$

Since $\|v\| \geq 1$, its normalization has first coordinate at least $1/2$, second coordinate at most h, and third coordinate at least $-h$. These bounds also hold at the initial query q. For any such parent, comparisons with the original vectors give

$$\begin{aligned} \langle x, z_- - y\rangle &= x_1(s-a) - x_2\big(\sqrt{1-s^2} + \sqrt{1-a^2}\big) \geq \frac{\eta - a}{2} - 2h, \\ \langle x, w - y\rangle &= x_1(\eta - a) - x_2\sqrt{1-a^2} + x_3\sqrt{1-\eta^2} \geq \frac{\eta - a}{2} - 2h. \end{aligned}$$

Replacing both compared vectors by their perturbations costs at most 2ε. Therefore every nonrelevant index still exceeds the target by at least

$$\frac{\eta - a}{2} - 2h - 2\varepsilon \geq \frac{\eta - a}{2} - \left(\frac{4}{\eta} + 2\right)\varepsilon > \frac{\eta - a}{4} > 0.$$

There are K nonrelevant indices and every width is at most $RF \leq H_{F,R} < K$. Induction now shows that the minus run never retrieves the target, and every centroid used in this induction indeed consists only of nonrelevant memories.

Because $B = 1$, each retrieved set is the sole retained cluster and every new index in it enters Bag. The target enters the plus Bag at round one. In either run $I_0 \subset I_1$, so for any executed prefix through round $r \geq 1$ the number of inserted indices is bounded by

$$\left|\bigcup_{\ell=0}^{r} I_\ell\right| \leq \sum_{\ell=1}^{r} |I_\ell| = F \sum_{j=2}^{r+1} j \leq H_{F,R} < K.$$

At round zero Bag has only $F < K$ indices, so no budget stop occurs. Retrieval is nonempty, all memories have positive first coordinate, and all centroid and update normalizations are nonzero. Thus both runs execute R rounds, and final Top-K returns the entire Bag. Neither changing insertion scores nor changing the identities of selected distractors can remove the target from the plus output. The two failure-loss pairs follow. Score matching forces a common Recollection probability p in the two environments, with regrets $(1-p)(1-\Delta_c)$ and $p\Delta_c$. Their optimized average and maximum are the claimed $\frac{1}{2}\min\{\Delta_c, 1-\Delta_c\}$ and $\Delta_c(1-\Delta_c)$.

Finally, the exact duplicates can be removed inside every such neighborhood. In the three-dimensional subspace already used, replace seed copy j by

$$se_1 + \sqrt{1-s^2}\big(\pm\cos\theta_j e_2 + \sin\theta_j e_3\big),$$

with distinct sufficiently small positive θ_j. Replace distractor copy j, in both corpora, by

$$\eta e_1 + \sqrt{1-\eta^2}\big(\sin\phi_j e_2 + \cos\phi_j e_3\big),$$

with distinct sufficiently small positive ϕ_j, and keep y fixed. These unit vectors retain the exact common initial profile $(s^{[F]}, \eta^{[K-F]}, a)$ and converge to the original pair as the angles vanish. Within-group embeddings are distinct; different groups have different first coordinates. All conclusions therefore hold with no repeated embeddings. The proof controls losses and Bag cardinality, not equality of perturbed traces. □

Corollary 6 (Total label budget for the grid-estimated gate)**.** *Under the model and sampling design of Proposition 4, write $w_g = u_+ - u_- > 0$. For $N = J_g N_g$ labels its bound is*

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^\star_{\mathcal{K}_\Gamma} \leq C_{\mathrm{mar}} \left[\frac{\lambda_{\mathrm{rel}}}{2}\varepsilon_t + \frac{\lambda_{\mathrm{rel}} w_g}{4(J_g-1)} + \sqrt{\frac{2J_g \log(2J_g/\delta)}{N}}\right]^2 + (1+\Delta_c)\delta. \tag{74}$$

Here a finer grid decreases the deterministic approximation term but leaves fewer independent labels at each grid point.

For an integer available budget $N_{\mathrm{lab}} \geq 2$, let $\ell_{\mathrm{lab}} = \log(2N_{\mathrm{lab}}/\delta)$, with $\delta \in (0,1)$, and assume $N_{\mathrm{lab}} \geq 8\ell_{\mathrm{lab}}$. Choose

$$J_g = \left\lceil (N_{\mathrm{lab}}/\ell_{\mathrm{lab}})^{1/3} \right\rceil, \qquad N_g = \lfloor N_{\mathrm{lab}}/J_g \rfloor.$$

Then the number used is $N = J_g N_g \in [N_{\mathrm{lab}}/2, N_{\mathrm{lab}}]$, and

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^\star_{\mathcal{K}_\Gamma} \leq C_{\mathrm{mar}} \left[\frac{\lambda_{\mathrm{rel}}}{2}\varepsilon_t + \left(\frac{\lambda_{\mathrm{rel}} w_g}{2} + \sqrt{6}\right)\left(\frac{\ell_{\mathrm{lab}}}{N_{\mathrm{lab}}}\right)^{1/3}\right]^2 + (1+\Delta_c)\delta. \tag{75}$$

In particular, with fixed model constants and $\delta = N_{\mathrm{lab}}^{-1}$, the bound is $O(\varepsilon_t^2 + (\log N/N)^{2/3})$ as $N \to \infty$. This is a guarantee for the stated grid estimator under designed independent labeling, not a minimax or parametric optimality claim and not a guarantee from uncontrolled routing logs.

Proof. Substitute $N_g = N/J_g$ and $h_g = w_g/(J_g - 1)$ in (13) to obtain (74). Apart from fixed constants, balancing its two label-dependent terms gives $J_g^{-1} \asymp \sqrt{J_g \log(2J_g/\delta)/N}$, or $J_g^3 \log(2J_g/\delta) \asymp N$.

For the explicit rounded allocation, set $v = (N_{\mathrm{lab}}/\ell_{\mathrm{lab}})^{1/3} \geq 2$. Then

$$2 \leq J_g = \lceil v \rceil \leq v + 1 \leq \tfrac{3}{2}v, \qquad J_g - 1 \geq v - 1 \geq v/2.$$

Because $\ell_{\mathrm{lab}} > 1$ and $N_{\mathrm{lab}} > 8$, $J_g \le \frac{3}{2} N_{\mathrm{lab}}^{1/3} \le N_{\mathrm{lab}}/2$. Consequently $N_{\mathrm{lab}}/J_g \ge 2$ and

$$N_g = \lfloor N_{\mathrm{lab}}/J_g \rfloor \ge \frac{N_{\mathrm{lab}}}{2J_g}, \qquad \frac{N_{\mathrm{lab}}}{2} \le J_g N_g \le N_{\mathrm{lab}}.$$

The deterministic grid term is bounded by

$$\frac{\lambda_{\mathrm{rel}} w_g}{4(J_g - 1)} \le \frac{\lambda_{\mathrm{rel}} w_g}{2v} = \frac{\lambda_{\mathrm{rel}} w_g}{2} \left(\frac{\ell_{\mathrm{lab}}}{N_{\mathrm{lab}}} \right)^{1/3}.$$

Since $J_g \le N_{\mathrm{lab}}$, $\log(2J_g/\delta) \le \ell_{\mathrm{lab}}$. Therefore

$$\begin{aligned} \sqrt{\frac{2\log(2J_g/\delta)}{N_g}} &\le \sqrt{\frac{4J_g \ell_{\mathrm{lab}}}{N_{\mathrm{lab}}}} \\ &\le \sqrt{\frac{6v\ell_{\mathrm{lab}}}{N_{\mathrm{lab}}}} = \sqrt{6} \left(\frac{\ell_{\mathrm{lab}}}{N_{\mathrm{lab}}} \right)^{1/3}. \end{aligned}$$

These inequalities prove (75). For $\delta = N_{\mathrm{lab}}^{-1}$, $\ell_{\mathrm{lab}} = \log(2N_{\mathrm{lab}}^2) = O(\log N_{\mathrm{lab}})$; the size condition holds for all sufficiently large budgets. Apply $(a+b)^2 \le 2a^2 + 2b^2$, absorb the N_{lab}^{-1} term, and use $N_{\mathrm{lab}}/2 \le N \le N_{\mathrm{lab}}$ to obtain the stated rate in N. Probe error must be at most of order $(\log N/N)^{1/3}$ for it not to dominate this bound. □

Corollary 7 (Learning a scalar Bag gate from i.i.d. labeled instances)**.** *Use the family and relevance model of Proposition 4. Instead of its designed grid labels, observe N i.i.d. pairs (t_i, Z_i), where $Z_i = \mathbf{1}\{Y_i = y\}$ and $\mathbb{P}(Z_i = 1 \mid t_i) = p(t_i)$. The test pair is an independent draw from the same distribution. Suppose the scalar density on $[u_-, u_+]$ satisfies $0 < f_{\min} \le f(t) \le f_{\max} < \infty$. Set $w_g = u_+ - u_-$ and $\nu = f_{\min} w_g \in (0, 1]$. Partition the interval into $J \ge 1$ bins of equal width w_g/J, with a fixed convention at endpoints. In each bin use the empirical fraction of $Z_i = 1$, assigning $1/2$ to an empty bin. At test time clip a probe $\widehat{t}$ with almost-sure error at most ε_t to the interval and use the estimate in its bin to route by $\mathbf{1}\{2\widehat{p} - 1 - \Delta_c > 0\}$. The probe is conditionally independent of the test label given $G(t)$ and the training data.*

For $\delta \in (0, 1)$, if

$$N\nu \ge 8J \log(2J/\delta), \tag{76}$$

then the risk averaged over training and test randomness obeys

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^\star_{\mathcal{K}_\Gamma} \le C_{\mathrm{mar}} \left[\frac{\lambda_{\mathrm{rel}}}{2} \left(\varepsilon_t + \frac{w_g}{J} \right) + 2\sqrt{\frac{J \log(4J/\delta)}{N\nu}} \right]^2 + (1 + \Delta_c)\delta, \tag{77}$$

where $C_{\mathrm{mar}} = 2f_{\max}/\mu_*$ as in the proposition. In particular, let $\ell_N = \log(4N/\delta)$. If $N\nu \ge 27\ell_N$, the integer choice $J = \lfloor (N\nu/\ell_N)^{1/3} \rfloor$ satisfies (76) *and gives*

$$\mathcal{R}(\widehat{\pi}) - \mathcal{R}^\star_{\mathcal{K}_\Gamma} \le C_{\mathrm{mar}} \left[\frac{\lambda_{\mathrm{rel}}}{2} \varepsilon_t + \left(\frac{3\lambda_{\mathrm{rel}} w_g}{4} + 2 \right) \left(\frac{\ell_N}{N\nu} \right)^{1/3} \right]^2 + (1 + \Delta_c)\delta. \tag{78}$$

For fixed model and density constants and $\delta = N^{-1}$, this is $O(\varepsilon_t^2 + (\log N/N)^{2/3})$. The lower density bound is a sufficient coverage condition for this histogram estimator; the rate is not claimed to be optimal for the parametric logistic model.

Proof. Write B_j for bin j, n_j for its sample count, and $\mu_j = N\mathbb{P}(t \in B_j) \ge N\nu/J$ for its expected count. The count is binomial. For completeness, exponential Markov at $\log 2$ and $1 - x \le e^{-x}$ give

$$\begin{aligned} \mathbb{P}(n_j \le \mu_j/2) &\le 2^{\mu_j/2} \mathbb{E}[2^{-n_j}] \\ &= 2^{\mu_j/2} \left(1 - \frac{\mathbb{P}(t \in B_j)}{2} \right)^N \\ &\le \exp\left(-\frac{1 - \log 2}{2} \mu_j \right) \le e^{-\mu_j/8}. \end{aligned}$$

Since $N\nu/(2J) \le \mu_j/2$, a union bound and (76) imply

$$\mathbb{P}\left(\min_{j\le J} n_j < \frac{N\nu}{2J}\right) \le Je^{-N\nu/(8J)} \le \delta/2.$$

Call the complementary event E_{cnt}. On this event all bins are nonempty.

Conditional on the scalar training features, the labels within a bin are independent Bernoulli variables, although their means need not coincide. For a nonempty bin put $\bar{p}_j = n_j^{-1}\sum_{i:t_i\in B_j} p(t_i)$ and $\widehat{p}_j = n_j^{-1}\sum_{i:t_i\in B_j} Z_i$. Hoeffding's inequality yields

$$\mathbb{P}\left(|\widehat{p}_j - \bar{p}_j| > \sqrt{\frac{\log(4J/\delta)}{2n_j}} \,\middle|\, t_1,\dots,t_N\right) \le \frac{\delta}{2J}.$$

Conditioning first on any feature configuration in E_{cnt}, taking a union bound over its bins, and then integrating over the features gives a joint event E with $\mathbb{P}(E) \ge 1-\delta$ on which

$$\max_{j\le J} |\widehat{p}_j - \bar{p}_j| \le \sqrt{\frac{J\log(4J/\delta)}{N\nu}}.$$

The two failure probabilities are at most $\delta/2$ each; no independence of bin counts is needed.

Let $\widetilde{t}$ be the clipped test probe. Its distance from t is at most ε_t. If $\widetilde{t} \in B_j$, every training feature in that bin has distance at most $w_g/J + \varepsilon_t$ from t. The logistic derivative satisfies $|p'| \le \lambda_{\mathrm{rel}}/4$, so

$$|\bar{p}_j - p(t)| \le \frac{\lambda_{\mathrm{rel}}}{4}\left(\frac{w_g}{J} + \varepsilon_t\right).$$

Consequently, on E the estimated advantage differs from $d_\Gamma(G(t)) = 2p(t) - 1 - \Delta_c$ uniformly by at most

$$e_{N,J} = \frac{\lambda_{\mathrm{rel}}}{2}\left(\varepsilon_t + \frac{w_g}{J}\right) + 2\sqrt{\frac{J\log(4J/\delta)}{N\nu}}.$$

Disagreement with the Bayes action at a nonzero advantage implies $0 < |d_\Gamma| \le e_{N,J}$. Training independence and the stated conditional independence of the probe make conditional excess risk equal to $|d_\Gamma|$ on this disagreement event. The margin bound from Proposition 4 gives expected excess risk at most $C_{\mathrm{mar}}e_{N,J}^2$ on the good training event. On its complement it is at most $1 + \Delta_c$. Averaging proves (77).

To check the integer allocation, put $v = (N\nu/\ell_N)^{1/3} \ge 3$ and $J = \lfloor v \rfloor$. Then $2v/3 \le v - 1 \le J \le v \le N$, and

$$\frac{N\nu}{J} \ge v^2\ell_N \ge 9\ell_N \ge 8\log(2J/\delta).$$

Thus the coverage condition holds. Moreover,

$$\frac{\lambda_{\mathrm{rel}}w_g}{2J} \le \frac{3\lambda_{\mathrm{rel}}w_g}{4v},$$
$$2\sqrt{\frac{J\log(4J/\delta)}{N\nu}} \le 2\sqrt{\frac{v\ell_N}{N\nu}} = \frac{2}{v}.$$

Substitution proves (78). For $\delta = N^{-1}$, $\ell_N = \log(4N^2)$, and the size condition holds for all sufficiently large N at fixed $\nu > 0$. Using $(a+b)^2 \le 2a^2 + 2b^2$ and absorbing N^{-1} proves the stated rate. These are histogram concentration and the classical plug-in margin step (Tsybakov & Audibert, 2007), applied to the specified RF-Mem family. The observations are labeled instances with measured scalar features, not arbitrary logs containing only the selected action's outcome. □

Corollary 8 (A cached aggregate can suffice for routing)**.** *Use the family and equal prior of Theorem 5, and let $\bar{y} = m^{-1}\sum_{j=1}^m y_j$ be the raw centroid of its candidate pool. With the notation of its construction,*

$$\langle z, \bar{y}\rangle = sa + \frac{\zeta}{m}W, \qquad \zeta = \sqrt{1-s^2}\sqrt{1-a^2}\rho > 0.$$

Thus observing this scalar yields zero routing excess risk over $\mathcal{R}^\star_{\mathcal{G}}$. Uniform scalar error $\varepsilon_{\mathrm{agg}} < \zeta/(2m)$ preserves this guarantee. If a policy may either acquire this scalar at total cost $c_{\mathrm{agg}} \ge 0$ or

use only the common initial observations, its optimal total excess risk is $\min\{g_c, c_{\mathrm{agg}}\}$. *Conditional on* $W = 1$, *however, the scalar and the common observations leave the marked location* J *uniform; for* $m > 1$ *they do not determine the returned index. A precomputed centroid permits* $O(d)$ *online arithmetic, while forming it from uncached candidate embeddings costs* $O(md)$.

Proof. Averaging (142) gives the displayed formula because at most one entry is marked. Thresholding at $sa + \zeta/(2m)$ identifies W exactly, including under the stated error bound. Routing by Recollection for $W = 1$ and Familiarity for $W = 0$ is optimal by the theorem's retrieval construction. Without acquiring the scalar the common observations yield excess risk g_c. Before acquisition those observations are constant, so randomizing whether to acquire gives a convex combination of g_c and c_{agg}, minimized at an endpoint. For $W = 1$ the observed scalar is the same for all J, whereas the returned candidate is J, proving the trace-insufficiency claim. The arithmetic bounds follow from a dense vector inner product and from summing m vectors of dimension d. Any materialization or cache-construction cost must be included in c_{agg} if charged to the decision. These statements concern the specified aggregate channel, not arbitrary additional information in approximate responses. □

Corollary 9 (Reusable candidate pools and total probe cost)**.** *The individual-probe family of Theorem 5 admits a realization with a fixed candidate pool and a fresh seed at each decision. Independent fair mark variables and independent uniform marked locations then give arbitrarily many decisions with the same cached candidate centroid and the same unprobed excess risk* g_c, *even conditional on all previous decisions and probe outcomes.*

Consider $N_{\mathrm{dec}} \geq 1$ *decisions of this kind, or a sequence of pool versions satisfying the same conditional experiment. Charge all incremental costs in loss units:* $A_0 = A_{\mathrm{pool}} + A_{\mathrm{build}}$ *for initially obtaining the prescribed candidate pool and forming its centroid,* A_{ref} *for later candidate identification or rematerialization,* Uc_u *for* U *applied vector updates, and* c_{on} *per decision for access and the scalar probe at its stipulated precision. These charges are additional to the action costs already in* L_a*; costs shared with the initial score probe cancel in the comparison. The displayed accounting uses deterministic cost totals; for random costs use their expectations.*

For each current pool, let the cached vector approximate a snapshot centroid with error at most $\varepsilon_{\mathrm{cache}}$. *Suppose scalar evaluation contributes error at most* $\varepsilon_{\mathrm{arith}}$ *and*

$$\varepsilon_{\mathrm{arith}} + \varepsilon_{\mathrm{cache}} + \frac{1}{m}\sum_{j=1}^{m} \|y_j^{\mathrm{cur}} - y_j^{\mathrm{snap}}\|_2 < \frac{\zeta}{2m} \tag{79}$$

at every decision, with the known baseline and gap of the current experiment. The policy that maintains the cache, probes at every decision, and thresholds as in Corollary 8 has zero routing excess risk. Its average excess risk including all these charges is

$$\frac{A_0 + A_{\mathrm{ref}} + Uc_u}{N_{\mathrm{dec}}} + c_{\mathrm{on}}. \tag{80}$$

It strictly improves on the best policy using only current and past initial scores, whose average excess risk is g_c, *exactly when* (80) *is less than* g_c. *This compares the two specified policies; it is not an optimality claim over all ways to discover pools or process coordinates.*

In particular, if $A_{\mathrm{ref}} + Uc_u \leq N_{\mathrm{dec}}\bar{c}_{\mathrm{maint}}$, *a sufficient break-even condition is*

$$c_{\mathrm{on}} + \bar{c}_{\mathrm{maint}} < g_c, \qquad N_{\mathrm{dec}} > \frac{A_0}{g_c - c_{\mathrm{on}} - \bar{c}_{\mathrm{maint}}}. \tag{81}$$

For a static pool $\bar{c}_{\mathrm{maint}} = 0$ *and the condition is exact, with the smallest admissible integer* $\lfloor A_0/(g_c - c_{\mathrm{on}})\rfloor + 1$ *when* $c_{\mathrm{on}} < g_c$.

Proof. We first justify reuse without assuming that repeated observations of a single hidden mark retain their original value. Retain a, s, ρ, η from the proof of Theorem 5. In $\mathbb{R}^{m+3}$ fix

$$q = e_1, \qquad w = \eta e_1 + \sqrt{1-\eta^2}e_2, \qquad y_j = ae_1 + \sqrt{1-a^2}e_{j+2}, \quad 1 \leq j \leq m.$$

The m candidates and the $K-1$ indexed distractor copies of w are shared across decisions. Put

$$z^0 = se_1 + \sqrt{1-s^2}e_{m+3},$$
$$z_j^1 = se_1 + \sqrt{1-s^2}\big(\rho e_{j+2} + \sqrt{1-\rho^2}e_{m+3}\big).$$

All displayed vectors have unit norm. At each decision draw $W \sim \mathrm{Bernoulli}(1/2)$ and J uniform on $[m]$, independently of one another and of the past. Use z^0 as the seed if $W = 0$, and z_J^1 otherwise; let candidate J be the unique relevant index in either case. Direct calculation gives

$$\langle q, z\rangle = s, \qquad \langle z, w\rangle = s\eta,$$
$$\langle z, y_j\rangle = sa + \zeta \mathbf{1}\{W = 1, J = j\}, \qquad \zeta = \sqrt{1-s^2}\sqrt{1-a^2}\rho.$$

Also $\langle q, y_j\rangle = a$, $\langle w, y_j\rangle = a\eta$, and $\langle y_i, y_j\rangle = a^2$ for $i \neq j$. Thus the indexed augmented Gram matrix at each decision is exactly that of the corresponding environment in Theorem 5. By Lemma 2, the published Bag trace and the matched retrieval-success losses agree with that construction. In particular, Recollection succeeds exactly when $W = 1$, while Familiarity always fails.

The raw candidate centroid

$$\bar{y} = ae_1 + \frac{\sqrt{1-a^2}}{m}\sum_{j=1}^{m} e_{j+2}$$

is constant across these decisions, whereas $\langle z, \bar{y}\rangle = sa + \zeta W/m$ uses the fresh seed. Past probe answers and labels therefore reveal no information about the new W or J. Conditional on the past, the current initial scores are still constant and the best unprobed regret is g_c. This supplies an explicit reuse model; it is a sequence of instances sharing a candidate pool, not a claim about arbitrary queries on one fixed complete corpus. If the same seed and mark were simply repeated, an earlier probe could settle later decisions, and charging g_c as their residual uncertainty would be incorrect. The fresh conditional experiment excludes that case.

For the cache guarantee, write $\bar{y}_{\mathrm{cur}} = m^{-1}\sum_j y_j^{\mathrm{cur}}$, $\bar{y}_{\mathrm{snap}} = m^{-1}\sum_j y_j^{\mathrm{snap}}$, and let $\widehat{\bar{y}}$ be the stored vector. The triangle inequality gives

$$\|\widehat{\bar{y}} - \bar{y}_{\mathrm{cur}}\|_2 \le \varepsilon_{\mathrm{cache}} + \frac{1}{m}\sum_j \|y_j^{\mathrm{cur}} - y_j^{\mathrm{snap}}\|_2.$$

Since the current seed is unit norm, Cauchy–Schwarz and the scalar arithmetic bound imply

$$|\widehat{s} - \langle z, \bar{y}_{\mathrm{cur}}\rangle| \le \varepsilon_{\mathrm{arith}} + \varepsilon_{\mathrm{cache}} + \frac{1}{m}\sum_j \|y_j^{\mathrm{cur}} - y_j^{\mathrm{snap}}\|_2.$$

Condition (79) therefore preserves the threshold at $sa + \zeta/(2m)$ and identifies W exactly. Only cache error is being allowed here: any actual pool change must still satisfy the stipulated current observation and retrieval-success model. Arbitrary changes of corpus geometry are not certified by this error calculation.

Subtracting each current full-information Bayes risk leaves zero routing regret for the maintained-cache policy. Adding the one-time, refresh, update, and online charges and dividing by N_{dec} gives (80). Conditional on every past history, a policy using initial scores alone incurs at least g_c, and the same constant action attains that value at each decision. Summing proves the exact comparison. Replacing maintenance charges by their stated upper bound and solving the strict inequality yields (81). The static-pool integer formula follows directly.

For a fixed pool size, forming the raw sum initially uses $O(md)$ arithmetic. A supplied vector replacement updates it by $y_j^{\mathrm{new}} - y_j^{\mathrm{old}}$ in $O(d)$ arithmetic, so U such replacements use $O(Ud)$; one online scalar uses $O(d)$ arithmetic. Locating, fetching, and verifying replacements is included separately in the charges, and numerical accumulation error belongs to $\varepsilon_{\mathrm{cache}}$. Unapplied changes contribute to the drift term in (79). Thus these arithmetic counts do not assign zero cost to candidate discovery, memory traffic, cache freshness, or achieving the required precision. □

As an illustration in loss units, take $\Delta_c = 1/2$, hence $g_c = 1/4$, with $A_0 = 20$, $c_{\mathrm{on}} = 0.01$, and average maintenance cost at most 0.005. If (79) holds, the sufficient threshold is $N_{\mathrm{dec}} > 20/0.235$,

namely $N_{\mathrm{dec}} \geq 86$. These are illustrative charges, not measured RF-Mem timings. For a query-specific pool rebuilt at every decision, its full acquisition and construction cost is recurring and belongs in the per-decision cost; it cannot be amortized by this argument.

Corollary 10 (Support-aware criterion for irreducible fixed-threshold regret)**.** *Assume the conditions of Lemma 26. Suppose* $p_I := \mathbb{P}(I) > 0$*. Define* $\mathfrak{Q}^\star := \inf_{\tau\in[h_-,h_+]} \mathfrak{Q}(\tau)$*. Then*

$$\mu_h \mathfrak{Q}^\star \leq \inf_{\tau\in[h_-,h_+]} \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \leq L_h \mathfrak{Q}^\star. \tag{82}$$

In particular, every entropy threshold in this fixed-mean-axis family has irreducible positive regret whenever $\mathfrak{Q}^\star > 0$*. As a useful sufficient condition, suppose additionally that there exist constants* $0 < \underline{f} \leq \overline{f} < \infty$ *for which the local bounds in* (64) *hold with* $\underline{f}_\tau = \underline{f}$ *and* $\overline{f}_\tau = \overline{f}$ *for every* $\tau \in [h_-, h_+]$*. Define* $\overline{b}_I := \mathbb{E}[b(M) \mid I]$ *and* $\sigma^2_{b,I} := \mathrm{Var}(b(M) \mid I)$*. Then*

$$\frac{\mu_h \underline{f}}{2} p_I \sigma^2_{b,I} \leq \inf_{\tau\in[h_-,h_+]} \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \leq \frac{L_h \overline{f}}{2} p_I \sigma^2_{b,I}.$$

Under this uniform local-coverage condition, if $\sigma^2_{b,I} > 0$*, then* $\inf_{\tau\in[h_-,h_+]} \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} > 0$*.*

Proof. Taking the infimum over $\tau \in [h_-, h_+]$ in (63) gives

$$\inf_\tau \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \geq \mu_h \inf_\tau \mathfrak{Q}(\tau) = \mu_h \mathfrak{Q}^\star,$$
$$\inf_\tau \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \leq L_h \inf_\tau \mathfrak{Q}(\tau) = L_h \mathfrak{Q}^\star.$$

This proves (82) and its strict-positivity conclusion.

We now impose the additional uniform local-coverage condition in the corollary and derive its variance-based specialization. For every constant τ,

$$\mathbb{E}\left[(b(M) - \tau)^2 \mathbf{1}_I\right] = p_I \mathbb{E}\left[(b(M) - \tau)^2 \mid I\right].$$

Write

$$b(M) - \tau = \left(b(M) - \overline{b}_I\right) + \left(\overline{b}_I - \tau\right).$$

Squaring,

$$(b(M) - \tau)^2 = (b(M) - \overline{b}_I)^2 + 2(b(M) - \overline{b}_I)(\overline{b}_I - \tau) + (\overline{b}_I - \tau)^2.$$

Taking conditional expectation given I,

$$\mathbb{E}\left[(b(M) - \tau)^2 \mid I\right] = \mathbb{E}\left[(b(M) - \overline{b}_I)^2 \mid I\right] + 2(\overline{b}_I - \tau)\mathbb{E}\left[b(M) - \overline{b}_I \mid I\right] + (\overline{b}_I - \tau)^2.$$

By definition of $\overline{b}_I$,

$$\mathbb{E}\left[b(M) - \overline{b}_I \mid I\right] = 0.$$

Therefore,

$$\mathbb{E}\left[(b(M) - \tau)^2 \mid I\right] = \sigma^2_{b,I} + (\overline{b}_I - \tau)^2.$$

The right-hand side is minimized at $\tau = \overline{b}_I$, with minimum $\sigma^2_{b,I}$. Since

$$b(M) \in [h_-, h_+] \quad \text{on } I,$$

its conditional mean also belongs to $[h_-, h_+]$, so the minimizer is admissible.

Consequently,

$$\inf_{\tau\in[h_-,h_+]} \mathbb{E}\left[(b(M)-\tau)^2\mathbf{1}_I\right] = p_I\sigma_{b,I}^2.$$

The lower bound follows by taking the infimum

$$\begin{aligned}
&\inf_\tau \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \\
&\geq \frac{\mu_h \underline{f}}{2} \inf_\tau \mathbb{E}\left[(b(M)-\tau)^2\mathbf{1}_I\right] \\
&= \frac{\mu_h \underline{f}}{2} p_I\sigma_{b,I}^2.
\end{aligned}$$

For the upper bound, evaluate (65) at

$$\tau = \bar{b}_I :$$

$$\begin{aligned}
&\inf_\tau \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \\
&\leq \mathcal{R}(\pi_{\bar{b}_I}) - \mathcal{R}^\star_{\mathcal{T}} \\
&\leq \frac{L_h \overline{f}}{2} \mathbb{E}\left[(b(M)-\bar{b}_I)^2\mathbf{1}_I\right] \\
&= \frac{L_h \overline{f}}{2} p_I\sigma_{b,I}^2.
\end{aligned}$$

If $\sigma_{b,I}^2 > 0$, the lower bound is strictly positive. □

Corollary 11 (Perturbation stability of the score-indistinguishable pair)**.** *This corollary strengthens Theorem 1 within the class of score-matched corpus pairs. Fix parameters satisfying that theorem and let q, y, $\mathcal{M}_+ = (z_+, w_1, w_2, y)$ and $\mathcal{M}_- = (z_-, w_1, w_2, y)$ be the indexed construction in its proof. There exists $\varepsilon > 0$ such that the following holds. Let $\widetilde{q}$ and $\widetilde{y}$ be unit vectors, and let $\widetilde{\mathcal{M}}_+ = (\widetilde{z}_+, \widetilde{w}_{+,1}, \widetilde{w}_{+,2}, \widetilde{y})$ and $\widetilde{\mathcal{M}}_- = (\widetilde{z}_-, \widetilde{w}_{-,1}, \widetilde{w}_{-,2}, \widetilde{y})$ be unit-vector corpora satisfying*

$$\max\Big\{\|\widetilde{q}-q\|_2, \|\widetilde{y}-y\|_2, \|\widetilde{z}_+ - z_+\|_2, \|\widetilde{z}_- - z_-\|_2, \max_{\sigma\in\{+,-\},\, j\in\{1,2\}} \|\widetilde{w}_{\sigma,j} - w_j\|_2\Big\} < \varepsilon,$$

and suppose their indexed initial score profiles agree:

$$\langle \widetilde{q}, \widetilde{z}_+\rangle = \langle \widetilde{q}, \widetilde{z}_-\rangle, \tag{83}$$

$$\langle \widetilde{q}, \widetilde{w}_{+,j}\rangle = \langle \widetilde{q}, \widetilde{w}_{-,j}\rangle, \qquad j = 1, 2. \tag{84}$$

Then all conclusions of Theorem 1, including its Bayes-gap and minimax-regret formulas, remain valid for the perturbed pair. Hence the construction is stable on a relative neighborhood of the set of score-matched corpus pairs.

Proof. Let q'_+ and q'_- be the two updated queries in the proof of Theorem 1. Define the initial-ranking margin

$$\delta_{\text{init}} := \min\{s-\eta_1, \eta_1-\eta_2, \eta_2-a\} > 0. \tag{85}$$

The strict inequalities established in that proof also give

$$\begin{aligned}
\delta_+ &:= \min_{j\in\{1,2\}} \left\{\langle q'_+, y\rangle - \langle q'_+, w_j\rangle\right\} > 0, \\
\delta_- &:= \min_{j\in\{1,2\}} \left\{\langle q'_-, w_j\rangle - \langle q'_-, y\rangle\right\} > 0.
\end{aligned} \tag{86}$$

Set $\delta := \min\{\delta_{\text{init}}, \delta_+, \delta_-\} > 0$.

At (q, z_+) and (q, z_-), the unnormalized update vectors have norm $D = \sqrt{A^2+B^2} > 0$. Consequently, $(x, z) \mapsto \mathsf{U}_\alpha(x, z)$ is continuous in neighborhoods of both pairs. Inner products are continuous as well. The finitely many initial-score differences and updated-score differences appearing in (85)– (86) are therefore continuous functions of the query and corpus vectors.

Choose $\varepsilon > 0$ small enough that both perturbed updates are well-defined and that every one of these comparison functions changes from its value at the constructed pair by less than $\delta/2$. All comparison signs are then preserved. In each perturbed corpus, the first listed memory remains the unique Top-1 item, and the common relevant memory remains below both distractors in the initial ranking. Familiarity therefore fails in both corpora. After the update, $\widetilde{y}$ outranks both remaining distractors in the $+$ corpus and is outranked by both in the $-$ corpus. Recollection consequently succeeds in the $+$ corpus and fails in the $-$ corpus.

Equations (83) and (84), together with the common relevant memory $\widetilde{y}$, imply equality of the two complete indexed initial score profiles. Finally, the costs are unchanged and $0 < \Delta_c < 1$, so the strict loss comparisons, equal-prior Bayes gap, and minimax-regret calculation from the original proof apply without modification. □

Corollary 12 (Exact losslessness condition)**.** *This corollary follows from Lemma 5, specialized to the nested information sets $\mathcal{H}_0 \subseteq \mathcal{H}_1$. Let $\mathcal{H}_0 \subseteq \mathcal{H}_1$. Define $P_0 := \mathbb{E}\left[(D_{\mathcal{H}_1})_+ \mid \mathcal{H}_0\right]$ and $N_0 := \mathbb{E}\left[(D_{\mathcal{H}_1})_- \mid \mathcal{H}_0\right]$. Then*

$$\mathcal{R}^\star_{\mathcal{H}_0} - \mathcal{R}^\star_{\mathcal{H}_1} = \mathbb{E}\left[\min\{P_0, N_0\}\right]. \tag{87}$$

Consequently, $\mathcal{R}^\star_{\mathcal{H}_0} = \mathcal{R}^\star_{\mathcal{H}_1}$ if and only if $P_0 N_0 = 0$ almost surely.

Proof. By Lemma 4, $\mathcal{R}^\star_{\mathcal{H}_0} - \mathcal{R}^\star_{\mathcal{H}_1} = \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{H}_1}| - |D_{\mathcal{H}_0}|\right]$. By the tower property,

$$D_{\mathcal{H}_0} = \mathbb{E}\left[D_{\mathcal{H}_1} \mid \mathcal{H}_0\right] = \mathbb{E}\left[(D_{\mathcal{H}_1})_+ - (D_{\mathcal{H}_1})_- \mid \mathcal{H}_0\right] = P_0 - N_0.$$

Also, $\mathbb{E}\left[|D_{\mathcal{H}_1}| \mid \mathcal{H}_0\right] = \mathbb{E}\left[(D_{\mathcal{H}_1})_+ + (D_{\mathcal{H}_1})_- \mid \mathcal{H}_0\right] = P_0 + N_0$. Therefore,

$$\frac{1}{2}\left[\mathbb{E}\left[|D_{\mathcal{H}_1}| \mid \mathcal{H}_0\right] - |D_{\mathcal{H}_0}|\right] = \frac{1}{2}\left[P_0 + N_0 - |P_0 - N_0|\right] = \min\{P_0, N_0\}.$$

Taking expectations proves (87).

The random variable $\min\{P_0, N_0\}$ is nonnegative. Therefore its expectation is zero if and only if it is zero almost surely. Since $P_0, N_0 \geq 0$, $\min\{P_0, N_0\} = 0$ if and only if $P_0 N_0 = 0$. □

Corollary 13 (Explicit regret lower bound from mean-score variability)**.** *This corollary follows from Lemma 26 together with the derivative lower bound in Lemma 25, and refines Corollary 10. Assume the conditions of Corollary 10, including the uniform local-coverage condition in its variance-based specialization. Then $\mathrm{Var}(b(M) \mid I) \geq \left(\frac{\mu_m}{L_h}\right)^2 \mathrm{Var}(M \mid I)$. Consequently,*

$$\inf_{\tau \in [h_-, h_+]} \{\mathcal{R}(\pi_\tau) - \mathcal{R}^\star_{\mathcal{T}}\} \geq \frac{\mu_h \underline{f}}{2} p_I \left(\frac{\mu_m}{L_h}\right)^2 \mathrm{Var}(M \mid I).$$

Proof. Let M' be an independent copy of M under the conditional distribution given I.

From (62), $b'(m) \geq c_b := \frac{\mu_m}{L_h} > 0$. For arbitrary $m_1 < m_2$ in the middle interval,

$$b(m_2) - b(m_1) = \int_{m_1}^{m_2} b'(u)\, du \geq \int_{m_1}^{m_2} c_b\, du = c_b(m_2 - m_1).$$

Thus, for arbitrary m_1, m_2, $|b(m_2) - b(m_1)| \geq c_b |m_2 - m_1|$.

For any square-integrable random variable X and an independent copy X', $\mathrm{Var}(X) = \frac{1}{2}\mathbb{E}[(X - X')^2]$. Indeed,

$$\mathbb{E}[(X - X')^2] = \mathbb{E}[X^2] + \mathbb{E}[(X')^2] - 2\mathbb{E}[XX'] = 2\mathbb{E}[X^2] - 2(\mathbb{E}[X])^2 = 2\,\mathrm{Var}(X).$$

Applying this identity conditionally on I,

$$\begin{aligned}\mathrm{Var}(b(M) \mid I) &= \frac{1}{2}\mathbb{E}\left[(b(M) - b(M'))^2 \mid I\right] \\ &\geq \frac{1}{2} c_b^2 \mathbb{E}\left[(M - M')^2 \mid I\right] = c_b^2\, \mathrm{Var}(M \mid I) = \left(\frac{\mu_m}{L_h}\right)^2 \mathrm{Var}(M \mid I).\end{aligned}$$

Substitution into the lower bound in Corollary 10 proves the result. □

Corollary 14 (A simple min-entropy certificate)**.** *This corollary is an explicit relaxation of Lemma 12, obtained by lower-bounding its sharp threshold by $e^{-\tau}$. For every $p \in \Delta_K$, $H(p) \geq -\log p_{\max}$. Consequently, $H(p) \leq \tau \Longrightarrow p_{\max} \geq e^{-\tau}$. Moreover, $\psi_K(\tau) \geq e^{-\tau}$.*

Proof. For every j with $p_j > 0$, $p_j \leq p_{\max}$. Therefore, $\log \frac{1}{p_j} \geq \log \frac{1}{p_{\max}}$. Multiplying by $p_j \geq 0$, $p_j \log \frac{1}{p_j} \geq p_j \log \frac{1}{p_{\max}}$. Summing, $H(p) = \sum_{j:p_j>0} p_j \log \frac{1}{p_j} \geq \log \frac{1}{p_{\max}} \sum_{j=1}^{K} p_j = -\log p_{\max}$. If $H(p) \leq \tau$, then $-\log p_{\max} \leq \tau$. Multiplying by -1 reverses the inequality: $\log p_{\max} \geq -\tau$. Exponentiation gives $p_{\max} \geq e^{-\tau}$. Applying this inequality to the sharp extremizer $p^{\min}(\psi_K(\tau))$, whose entropy is τ, yields $\psi_K(\tau) \geq e^{-\tau}$. □

Corollary 15 (Distribution-free upper-tail confidence radius)**.** *This corollary follows directly from Theorem 9. Let $\delta \in (0,1)$. Then, with probability at least $1-\delta$, $\left|\widehat{M}_{n,K} - U_{K/n}(F)\right| \leq r^{\mathrm{tail}}_{n,K,\delta}$, where*

$$r^{\mathrm{tail}}_{n,K,\delta} := \min\left\{\Delta_s,\ \frac{\Delta_s}{K/n}\sqrt{\frac{\log(2/\delta)}{2n}}\right\}.$$

Equivalently,

$$r^{\mathrm{tail}}_{n,K,\delta} = \min\left\{\Delta_s,\ \frac{\Delta_s\sqrt{n}}{K}\sqrt{\frac{\log(2/\delta)}{2}}\right\}.$$

Corollary 16 (Finite-sample expectation approximation)**.** *As a consequence of Theorem 9, under its assumptions,*

$$\left|\mathbb{E}[\widehat{M}_{n,K}] - U_{K/n}(F)\right| \leq \min\left\{\Delta_s,\ \frac{\Delta_s}{K/n}\sqrt{\frac{\pi}{2n}}\right\}.$$

Proof. Let $D_n := \sup_{s\in\mathbb{R}} |F_n(s) - F(s)|$. From (149),

$$\left|\widehat{M}_{n,K} - U_{K/n}(F)\right| \leq \frac{\Delta_s}{K/n} D_n.$$

Therefore,

$$\left|\mathbb{E}[\widehat{M}_{n,K}] - U_{K/n}(F)\right| \leq \mathbb{E}\left[\left|\widehat{M}_{n,K} - U_{K/n}(F)\right|\right] \leq \frac{\Delta_s}{K/n}\mathbb{E}[D_n].$$

Using the tail-integral formula, $\mathbb{E}[D_n] = \int_0^\infty \mathbb{P}(D_n > t)\,dt \leq \int_0^\infty 2e^{-2nt^2}\,dt$. The Gaussian integral satisfies $\int_0^\infty e^{-at^2}\,dt = \frac{\sqrt{\pi}}{2\sqrt{a}}$. Setting $a = 2n$, $\int_0^\infty 2e^{-2nt^2}\,dt = 2\frac{\sqrt{\pi}}{2\sqrt{2n}} = \sqrt{\frac{\pi}{2n}}$. Thus $\left|\mathbb{E}[\widehat{M}_{n,K}] - U_{K/n}(F)\right| \leq \frac{\Delta_s}{K/n}\sqrt{\frac{\pi}{2n}}$. Since both quantities lie in $[\ell, u]$, their difference is also deterministically at most Δ_s. Taking the minimum gives the result. □

E PROOFS

E.1 PROOF FOR PROPOSITION 1

Proof. Because $c = \langle q, g\rangle$, the vector $g_\perp = g - cq$ satisfies $\langle q, g_\perp\rangle = \langle q, g - cq\rangle = \langle q, g\rangle - c\langle q, q\rangle = c - c = 0$. Thus $g_\perp \in q^\perp$.

We first consider the case $b = 0$. Then $g_\perp = 0$, so $g = cq$ and $B = 0$. The unnormalized update is therefore $(1+\alpha)q + (1-\alpha)g = [1 + \alpha + (1-\alpha)c]\,q = Aq$. Since $\|q\|_2 = 1$ and $\|g\|_2 \leq 1$, Cauchy–Schwarz gives $c = \langle q, g\rangle \geq -1$, and hence

$$A = 1 + \alpha + (1-\alpha)c \geq 1 + \alpha - (1-\alpha) = 2\alpha \geq 0.$$

Because the update is well-defined, $Aq \neq 0$; thus $A > 0$. Consequently, $\mathsf{U}_\alpha(q, g) = \frac{Aq}{\|Aq\|_2} = \frac{Aq}{A} = q$, and the score change is zero.

For the remainder of the proof, suppose $b > 0$. Then $u = \frac{g_\perp}{b}$ is a unit vector orthogonal to q, $g = cq + bu$, and $\rho = \langle u, v\rangle$ is well-defined. Substituting this decomposition into the unnormalized update,

$$\begin{aligned}(1+\alpha)q + (1-\alpha)g &= (1+\alpha)q + (1-\alpha)(cq + bu)\\ &= [1 + \alpha + (1-\alpha)c]\, q + (1-\alpha)b\, u = Aq + Bu.\end{aligned}$$

Since q and u are orthonormal,

$$\begin{aligned}\|Aq + Bu\|_2^2 &= \langle Aq + Bu, Aq + Bu\rangle\\ &= A^2\langle q, q\rangle + 2AB\langle q, u\rangle + B^2\langle u, u\rangle = A^2 + B^2.\end{aligned}$$

Therefore,

$$\mathsf{U}_\alpha(q, g) = \frac{Aq + Bu}{D}. \tag{88}$$

We next verify that $D + A > 0$. Since $\|q\|_2 = 1$ and $\|g\|_2 \le 1$, Cauchy–Schwarz gives $c = \langle q, g\rangle \ge -1$. Hence $A = 1 + \alpha + (1-\alpha)c \ge 1 + \alpha - (1-\alpha) = 2\alpha \ge 0$. The update is assumed to be well-defined, so $D = \sqrt{A^2 + B^2} > 0$. Thus $D + A > 0$.

Now consider the target memory y. Since $a = \langle q, y\rangle \in (-1, 1)$, the vector $y - aq$ is nonzero. Its squared norm is $\|y - aq\|_2^2 = \|y\|_2^2 - 2a\langle q, y\rangle + a^2\|q\|_2^2 = 1 - 2a^2 + a^2 = 1 - a^2$. Therefore, $v \coloneqq \frac{y - aq}{\sqrt{1-a^2}}$ is unit-normalized. Moreover, $\langle q, v\rangle = \frac{\langle q,y\rangle - a\langle q,q\rangle}{\sqrt{1-a^2}} = \frac{a-a}{\sqrt{1-a^2}} = 0$. Hence $y = aq + \sqrt{1-a^2}\, v$.

Using (88),

$$\begin{aligned}\langle \mathsf{U}_\alpha(q, g), y\rangle &= \left\langle \frac{Aq + Bu}{D}, aq + \sqrt{1-a^2}\, v\right\rangle\\ &= \frac{1}{D}\Big[Aa\langle q, q\rangle + A\sqrt{1-a^2}\langle q, v\rangle + Ba\langle u, q\rangle + B\sqrt{1-a^2}\langle u, v\rangle\Big]\\ &= \frac{Aa + B\sqrt{1-a^2}\rho}{D}.\end{aligned}$$

This proves (1).

Subtracting the initial score a,

$$\langle \mathsf{U}_\alpha(q, g), y\rangle - a = \frac{Aa + B\sqrt{1-a^2}\rho - aD}{D} = \frac{B\sqrt{1-a^2}\rho - a(D - A)}{D}. \tag{89}$$

Because $D + A > 0$,

$$D - A = \frac{(D-A)(D+A)}{D+A} = \frac{D^2 - A^2}{D+A} = \frac{A^2 + B^2 - A^2}{D+A} = \frac{B^2}{D+A}.$$

Substituting this into (89),

$$\langle \mathsf{U}_\alpha(q, g), y\rangle - a = \frac{B\sqrt{1-a^2}\rho - aB^2/(D+A)}{D} = \frac{B}{D}\left[\sqrt{1-a^2}\rho - \frac{aB}{D+A}\right].$$

If $B > 0$, then $B/D > 0$, and therefore the score gain is strictly positive if and only if $\sqrt{1-a^2}\rho > \frac{aB}{D+A}$. Since $\sqrt{1-a^2} > 0$, division by this positive number gives $\rho > \frac{aB}{(D+A)\sqrt{1-a^2}}$.

Finally, suppose $B = 0$. Then the unnormalized update is Aq. Since the update is well-defined and $A \ge 0$, we must have $A > 0$. Hence $\mathsf{U}_\alpha(q, g) = \frac{Aq}{\|Aq\|_2} = \frac{Aq}{A} = q$. The score change is therefore zero. □

E.2 Proof for Theorem 1

Proof. Let e_1, e_2, e_3 be orthonormal vectors in $\mathbb{R}^d$, and set $q \coloneqq e_1$.

Define $A \coloneqq 1 + \alpha + (1-\alpha)s$ and $B \coloneqq (1-\alpha)\sqrt{1-s^2}$. Because $\alpha < 1$ and $s < 1$, we have $B > 0$. Also, $A > 0$. Since $s > a$ and $a + \frac{B}{A}\sqrt{1-a^2} > a$, we have $\min\left\{s,\, a + \frac{B}{A}\sqrt{1-a^2}\right\} > a$. Therefore, we may choose

$$a < \eta_2 < \eta_1 < \min\left\{s,\, a + \frac{B}{A}\sqrt{1-a^2}\right\}. \tag{90}$$

Define the relevant memory $y := ae_1 + \sqrt{1-a^2}\, e_2$, two alternative first-retrieved memories

$$z_+ := se_1 + \sqrt{1-s^2}\, e_2,$$
$$z_- := se_1 - \sqrt{1-s^2}\, e_2,$$

and two distractors $w_j := \eta_j e_1 + \sqrt{1-\eta_j^2}\, e_3$ for $j = 1, 2$. All five vectors are unit-normalized. For example, $\|y\|_2^2 = a^2 + (1-a^2) = 1$, and the same calculation applies to $z_\pm$ and w_j.

Set $\mathcal{M}_+ := \{z_+, w_1, w_2, y\}$ and$\mathcal{M}_- := \{z_-, w_1, w_2, y\}$. We first calculate the complete initial similarity profile. Since $q = e_1$, $\langle q, z_\pm\rangle = s$, $\langle q, w_1\rangle = \eta_1$, $\langle q, w_2\rangle = \eta_2$ and $\langle q, y\rangle = a$. Hence both indexed score profiles equal

$$V_+ = V_- = (s, \eta_1, \eta_2, a). \tag{91}$$

In particular, the complete score profile, Top-3 score vector, mean score, and entropy are identical in the two corpora.

By (90), $s > \eta_1 > \eta_2 > a$. Therefore, Familiarity returns

$$\mathcal{R}_{\mathrm{F}}(q; \mathcal{M}_+) = \{z_+, w_1, w_2\},$$
$$\mathcal{R}_{\mathrm{F}}(q; \mathcal{M}_-) = \{z_-, w_1, w_2\}.$$

The relevant memory y is absent in both cases. Thus $\ell_{\mathrm{F},+} = \ell_{\mathrm{F},-} = 1$.

We next analyze Recollection. Since $z_\pm$ has the unique largest initial score, the first retrieved memory is z_+ in $\mathcal{M}_+$ and z_- in $\mathcal{M}_-$. The two updated queries are $q'_+ = \mathsf{U}_\alpha(q, z_+)$ and $q'_- = \mathsf{U}_\alpha(q, z_-)$. Using the definitions of A and B,

$$(1+\alpha)q + (1-\alpha)z_+ = Ae_1 + Be_2,$$
$$(1+\alpha)q + (1-\alpha)z_- = Ae_1 - Be_2.$$

Let $D \coloneqq \sqrt{A^2 + B^2}$. Then $q'_+ = \frac{Ae_1 + Be_2}{D}$ and $q'_- = \frac{Ae_1 - Be_2}{D}$. Under q'_+, the relevant-memory score is

$$\langle q'_+, y\rangle = \left\langle \frac{Ae_1 + Be_2}{D},\, ae_1 + \sqrt{1-a^2}\, e_2 \right\rangle = \frac{Aa + B\sqrt{1-a^2}}{D}. \tag{92}$$

For $j = 1, 2$,

$$\langle q'_+, w_j\rangle = \left\langle \frac{Ae_1 + Be_2}{D},\, \eta_j e_1 + \sqrt{1-\eta_j^2}\, e_3 \right\rangle = \frac{A\eta_j}{D}.$$

By (90), $\eta_j < a + \frac{B}{A}\sqrt{1-a^2}$. Multiplying by $A > 0$, $A\eta_j < Aa + B\sqrt{1-a^2}$. Dividing by $D > 0$, $\langle q'_+, w_j\rangle < \langle q'_+, y\rangle$. Thus y outranks both remaining distractors and belongs to the next-round Top-2 set. Therefore, $\ell_{\mathrm{R},+} = 0$.

Under q'_-, $\langle q'_-, y\rangle = \frac{Aa - B\sqrt{1-a^2}}{D}$ and$\langle q'_-, w_j\rangle = \frac{A\eta_j}{D}$. Since $\eta_j > a$ and $A > 0$, $A\eta_j > Aa$. Since $B > 0$ and $a < 1$, $Aa > Aa - B\sqrt{1-a^2}$. Therefore, $A\eta_j > Aa - B\sqrt{1-a^2}$, and hence $\langle q'_-, w_j\rangle > \langle q'_-, y\rangle$ for $j = 1, 2$. Both distractors outrank y, so y is absent from the next-round Top-2 set. Thus

$$\ell_{\mathrm{R},-} = 1. \tag{93}$$

Now use the cost-adjusted loss $L_a = \ell_a + \beta C_a$. In the $+$ environment, $L_{\mathrm{F},+} = 1 + \beta C_{\mathrm{F}}$ and $L_{\mathrm{R},+} = \beta C_{\mathrm{R}}$. Therefore,

$$L_{\mathrm{F},+} - L_{\mathrm{R},+} = 1 - \beta(C_{\mathrm{R}} - C_{\mathrm{F}}) = 1 - \Delta_c > 0.$$

Thus Recollection is uniquely optimal.

In the $-$ environment, $L_{\mathrm{F},-} = 1 + \beta C_{\mathrm{F}}$ and $L_{\mathrm{R},-} = 1 + \beta C_{\mathrm{R}}$. Therefore,

$$L_{\mathrm{F},-} - L_{\mathrm{R},-} = -\beta(C_{\mathrm{R}} - C_{\mathrm{F}}) = -\Delta_c < 0.$$

Thus Familiarity is uniquely optimal.

Because $V_+ = V_-$, a $\mathcal{V}$-measurable router must use the same conditional action distribution in both environments. Let $p \coloneqq \Pr(\text{choose Recollection} \mid V)$. Its regret in the $+$ environment is $\mathrm{Reg}_+(p) = (1-p)(1-\Delta_c)$, because only choosing Familiarity is suboptimal there. Its regret in the $-$ environment is $\mathrm{Reg}_-(p) = p\Delta_c$, because only choosing Recollection is suboptimal there.

Under an equal prior, the expected regret is

$$\frac{1}{2}\left[\mathrm{Reg}_+(p) + \mathrm{Reg}_-(p)\right] = \frac{1}{2}\left[(1-p)(1-\Delta_c) + p\Delta_c\right] = \frac{1}{2}\left[1 - \Delta_c + p(2\Delta_c - 1)\right].$$

This is affine in p. If $\Delta_c < 1/2$, it is minimized at $p = 1$, giving $\Delta_c/2$. If $\Delta_c > 1/2$, it is minimized at $p = 0$, giving $(1-\Delta_c)/2$. If $\Delta_c = 1/2$, it is constant and equals $1/4$. Therefore,

$$\inf_p \frac{1}{2}\left[\mathrm{Reg}_+(p) + \mathrm{Reg}_-(p)\right] = \frac{1}{2}\min\left\{\Delta_c, 1-\Delta_c\right\}.$$

The full geometry $\mathcal{G}$ distinguishes the two environments and chooses the optimal action in each. Thus the equal-prior score deficiency equals the preceding quantity.

For minimax regret, minimize $\max\{(1-p)(1-\Delta_c), p\Delta_c\}$. The first term decreases in p, whereas the second increases. The minimum is attained when they are equal: $(1-p)(1-\Delta_c) = p\Delta_c$ and $1 - \Delta_c = p$, so $p^\star = 1 - \Delta_c$. Substitution gives $\mathrm{Reg}_+(p^\star) = \mathrm{Reg}_-(p^\star) = \Delta_c(1-\Delta_c)$. □

E.3 Proof for Theorem 2

Proof. Choose $0 < a < s < 1$, and put

$$A = 1 + \alpha + (1-\alpha)s, \quad b = (1-\alpha)\sqrt{1-s^2} > 0, \quad D = \sqrt{A^2 + b^2}.$$

Choose $a < \eta < \min\{s, a + (b/A)\sqrt{1-a^2}\}$, and set

$$q = e_1, \quad z_\pm = se_1 \pm \sqrt{1-s^2}e_2, \quad w = \eta e_1 + \sqrt{1-\eta^2}e_3, \quad y = ae_1 + \sqrt{1-a^2}e_2.$$

Each corpus contains F indexed copies of $z_\pm$, $K - F$ indexed copies of w, and the single relevant memory y. Only embeddings, not indices, are repeated. Fix the same tie order on corresponding indices. The common initial profile is $(s^{[F]}, \eta^{[K-F]}, a)$, so Familiarity misses y.

At round zero, Top-F consists of the seed copies. There is one cluster, its normalized centroid is $z_\pm$, its child query is $q'_\pm = (Ae_1 \pm be_2)/D$, and Bag contains the F seed indices. At this child query,

$$\langle q'_\pm, z_\pm\rangle = \frac{As + b\sqrt{1-s^2}}{D} > \frac{A\eta}{D} = \langle q'_\pm, w\rangle, \qquad \langle q'_\pm, y\rangle = \frac{Aa \pm b\sqrt{1-a^2}}{D}.$$

The target score is above the distractor score in the plus corpus and below it in the minus corpus. Thus round one's full-corpus Top-$2F$ contains all seed indices and, respectively, y and $F-1$ distractor indices, or F distractor indices. There is again one cluster and no pruning. Both Bags have $2F$ indices. When $R = 2$ and $K \geq 2F + 1$, final Top-K returns the whole Bag, so the required failure losses follow.

Now let $R \geq 2$ and $K \geq (R+1)F + 1$. In the minus corpus every parent query belongs to the cone

$$\{x \in \mathrm{span}(e_1, e_2, e_3) : x_1 > 0,\ x_2 \leq 0,\ x_3 \geq 0\}.$$

For such a query,

$$\langle x, z_- - y\rangle = x_1(s-a) - x_2\big(\sqrt{1-s^2} + \sqrt{1-a^2}\big) > 0,$$
$$\langle x, w - y\rangle = x_1(\eta - a) - x_2\sqrt{1-a^2} + x_3\sqrt{1-\eta^2} > 0.$$

There are K nonrelevant indices, and every retrieval width is at most $RF < K$. Hence y is never retrieved. Centroid formation from the seed and distractor vectors, normalization, and the query update preserve this cone, proving the induction even if execution stops early.

In the plus corpus, y entered Bag at round one. Because all distractor embeddings coincide and ties are fixed by index, every retrieval selects a prefix of the same distractor order. The union of distractor indices ever selected therefore contains at most RF indices. Including the F seeds and y gives the pathwise bound

$$|\mathrm{Bag}| \le F + 1 + RF \le K.$$

Thus no final truncation can remove y, regardless of later query directions or first-insertion scores. Budget stopping also preserves its inclusion. More precisely, after any round $r \le R-2$, either Bag has at most $F + 1 + (r+1)F \le RF + 1 < K$ indices. No earlier budget stop occurs, and every retrieval is nonempty. Both runs therefore execute all R rounds.

Every memory has positive projection onto q. Every nonempty centroid is therefore nonzero, and every update numerator has strictly positive projection onto q, by induction. All normalizations, including those used only for final-round Bag scores, are well-defined. The maximum widths are $2F < K$ in the first range and $RF < K$ in the second.

The two advantages are $1 - \Delta_c$ and $-\Delta_c$. A score-only router uses a common probability p of Recollection and has regrets $(1-p)(1-\Delta_c)$ and $p\Delta_c$. Minimizing their average gives $\frac{1}{2}\min\{\Delta_c, 1-\Delta_c\}$; minimizing their maximum at $p = 1-\Delta_c$ gives $\Delta_c(1-\Delta_c)$. This proves all claims for the stipulated fixed action costs. □

E.4 Proof for Theorem 3

Proof. We first prove the adaptive sufficiency statement. Retain the linear map W and standard basis vectors $e_0, e_1, \ldots, e_n$ from the proof of Lemma 2. Let the incoming query of an active beam be $x = Wa$, and let $J(a) = \{j : a_j \neq 0\}$.

For every eligible index $i \in E$,

$$\langle x, z_i\rangle = e_i^\top \Gamma a = \sum_{j=0}^{n} \Gamma_{ij} a_j = \sum_{j \in J(a)} \Gamma_{ij} a_j. \tag{94}$$

Every entry on the right-hand side belongs to $\Gamma_{E,J(a)}$. Hence this cross-block determines all scores used to retrieve the indexed set $I \subseteq E$, including their ties and their ordering under the fixed tie-breaking rule.

All points passed to K-means have indices in I. For $i, j \in I$, their squared distance is

$$\|z_i - z_j\|_2^2 = \Gamma_{ii} + \Gamma_{jj} - 2\Gamma_{ij},$$

which is determined by the principal block $\Gamma_{I,I}$. For any nonempty intermediate K-means cluster $A \subseteq I$, define $c_A \coloneqq \frac{1}{|A|}\sum_{j\in A} e_j$. Then its arithmetic centroid is Wc_A, and for every $i \in I$,

$$\|z_i - Wc_A\|_2^2 = (e_i - c_A)^\top \Gamma (e_i - c_A) = \Gamma_{ii} - \frac{2}{|A|}\sum_{j\in A}\Gamma_{ij} + \frac{1}{|A|^2}\sum_{j\in A}\sum_{k\in A}\Gamma_{jk}. \tag{95}$$

Every index in the sums in (95) lies in I. Thus $\Gamma_{I,I}$ determines every assignment distance, within-cluster objective, and centroid displacement used by the stated K-means implementation. It consequently determines every internal partition and the final clusters $C \subseteq I$.

For a final cluster C, its raw centroid coefficient is c_C. If the raw centroid is retained, then $\widetilde{c}_C = c_C$. If it is unit-normalized, then

$$\widetilde{c}_C = \frac{c_C}{\sqrt{c_C^\top \Gamma c_C}}. \tag{96}$$

Because $J(c_C) = C$, the quadratic form in (96) expands as

$$c_C^\top \Gamma c_C = \sum_{j\in C}\sum_{k\in C}(c_C)_j \Gamma_{jk}(c_C)_k = \frac{1}{|C|^2}\sum_{j\in C}\sum_{k\in C}\Gamma_{jk}.$$

This quantity is determined by $\Gamma_{C,C}$, which is contained in $\Gamma_{I,I}$. Hence $\widetilde{c}_C$ is determined by the second block in (4).

The unnormalized child-query coefficient is $h_C = \alpha a + (1-\alpha)\widetilde{c}_C + e_0$. Once a and $\widetilde{c}_C$ are known, this vector and its support $J_C = J(h_C)$ are known without any additional Gram entry. Its squared vector norm is

$$\|Wh_C\|_2^2 = h_C^\top \Gamma h_C = \sum_{j\in J_C}\sum_{k\in J_C}(h_C)_j \Gamma_{jk}(h_C)_k.$$

Therefore the principal block Γ_{J_C,J_C} determines the normalizing denominator and the child-query coefficient $a'_C = \frac{h_C}{\sqrt{h_C^\top \Gamma h_C}}$.

For every index $i \in Q_C$ scored under this child query, $\langle Wa'_C, z_i\rangle = e_i^\top \Gamma a'_C = \sum_{j\in J_C}\Gamma_{ij}(a'_C)_j$. The right-hand side uses only Γ_{Q_C,J_C}. In particular, for every cluster A scored under the child query,

$$\mathrm{score}(A, Wa'_C) = \sum_{i\in A}\langle Wa'_C, z_i\rangle = \sum_{i\in A}\sum_{j\in J_C}\Gamma_{ij}(a'_C)_j, \tag{97}$$

and $A \subseteq Q_C$ by the definition of Q_C. Thus the fourth block determines all child retrieval, beam, cluster, and final-ranking scores. Fixed indexed rules then determine beam selection, deduplication, bag updates, stopping, and output.

Equations (94)– (97) prove that the four blocks in (4) determine one active beam step. At round zero, $a = e_0$, so

$$J(a) = \{0\}, \qquad \Gamma_{E,J(a)} = (\Gamma_{i0})_{i\in E}, \tag{98}$$

which is exactly the indexed initial score vector on E because $\Gamma_{i0} = \langle z_i, q\rangle = V_i$.

We now induct over active beams and rounds. Suppose all preceding discrete states, coefficient vectors, and adaptively requested Gram entries have been determined. The first block determines the next retrieved set; that set determines which second block is requested; the resulting clusters determine the supports J_C; and those supports determine the third and fourth requested blocks. The calculations above then determine the next complete state. Since the round limit and beam width are finite, this adaptive induction terminates after finitely many requests and determines the full trace.

Let Γ' be any other valid Gram matrix that agrees with Γ on every entry requested along this execution. The initial scores in (98) agree. Assume inductively that the executions under Γ and Γ' have agreed up to an active beam step. They consequently request the same four indexed blocks. Those block values agree by hypothesis, so all quantities in (94)– (97) agree and the next states agree. Induction gives identical traces. Hence no entry outside the adaptive union is required for that realized execution. Whenever this union is a strict subset of the full matrix and there exists another valid completion of its entries, full Gram is not an instance-minimal entrywise representation.

It remains to prove the worst-case necessity claim. Fix $\alpha \in [0, 1)$, choose $0 < a < s < 1$, and define

$$A \coloneqq 1 + \alpha + (1-\alpha)s, \qquad B \coloneqq (1-\alpha)\sqrt{1-s^2}. \tag{99}$$

Then $A > 0$ and $B > 0$. Because both s and $a + \frac{B}{A}\sqrt{1-a^2}$ strictly exceed a, choose

$$a < \eta < \min\left\{s, a + \frac{B}{A}\sqrt{1-a^2}\right\}. \tag{100}$$

Define

$$c_0 \coloneqq \frac{A(\eta - a)}{B\sqrt{1-a^2}}. \tag{101}$$

The strict inequalities in (100) imply $0 < c_0 < 1$. Choose

$$0 < h < \min\{c_0, 1 - c_0\}. \tag{102}$$

Then $0 < c_0 - h < c_0 + h < 1$.

Let e_1, e_2, e_3, e_4 be orthonormal vectors in $\mathbb{R}^d$ for $d \geq 4$, and define

$$\begin{aligned} q &\coloneqq e_1, \\ z_0 &\coloneqq s e_1 + \sqrt{1-s^2}\, e_2, \\ w &\coloneqq \eta e_1 + \sqrt{1-\eta^2}\, e_4, \\ y_\pm &\coloneqq a e_1 + \sqrt{1-a^2}\left[(c_0 \pm h) e_2 + \sqrt{1-(c_0 \pm h)^2}\, e_3\right]. \end{aligned}$$

Each vector is unit-normalized. For $y_\pm$, this follows from

$$\|y_\pm\|_2^2 = a^2 + (1-a^2)\left[(c_0 \pm h)^2 + 1 - (c_0 \pm h)^2\right] = a^2 + (1-a^2) = 1.$$

The other norm calculations are immediate from the orthonormality of e_1, e_2, e_3, e_4.

Consider the two indexed corpora $\mathcal{M}_+ = (z_0, w, y_+)$ and $\mathcal{M}_- = (z_0, w, y_-)$. Their complete initial score profiles agree because $\langle q, z_0\rangle = s$, $\langle q, w\rangle = \eta$ and $\langle q, y_\pm\rangle = a$. Moreover, $s > \eta > a$, so z_0 is the unique initial Top-1 memory and Familiarity Top-2 returns (z_0, w) in both corpora.

The one-cluster update based on z_0 is common to the two corpora. By (99), its unnormalized numerator and normalized query are $(1+\alpha)q + (1-\alpha)z_0 = Ae_1 + Be_2$ and $x' = \frac{Ae_1 + Be_2}{D}$ with $D \coloneqq \sqrt{A^2 + B^2}$. The remaining-memory scores are

$$\begin{aligned} \langle x', w\rangle &= \frac{A\eta}{D}, \\ \langle x', y_\pm\rangle &= \frac{Aa + B\sqrt{1-a^2}(c_0 \pm h)}{D}. \end{aligned} \tag{103}$$

From (101), $B\sqrt{1-a^2}c_0 = A(\eta - a)$. Substituting this equality into (103) gives

$$\langle x', y_\pm\rangle = \frac{Aa + A(\eta - a) \pm B\sqrt{1-a^2}h}{D} = \frac{A\eta \pm B\sqrt{1-a^2}h}{D}.$$

Since $B > 0$, $h > 0$, and $a < 1$,

$$\langle x', y_+\rangle > \langle x', w\rangle > \langle x', y_-\rangle. \tag{104}$$

Therefore, after excluding z_0, the expanded Top-1 retrieval returns y_+ in $\mathcal{M}_+$ and w in $\mathcal{M}_-$. The indexed outputs differ.

We finally compare the two augmented Gram matrices. Inner products among q, z_0, w are common. In addition,

$$\begin{aligned} \langle q, y_\pm\rangle &= a, \\ \langle w, y_\pm\rangle &= \eta a, \\ \langle y_\pm, y_\pm\rangle &= 1. \end{aligned} \tag{105}$$

The only differing off-diagonal pair is

$$\langle z_0, y_\pm\rangle = sa + \sqrt{1-s^2}\sqrt{1-a^2}(c_0 \pm h). \tag{106}$$

Thus changing only the symmetric Gram positions associated with (z_0, y) changes the expanded output while all initial scores remain fixed. Any designated pair of memory indices can play the roles of (z_0, y) after relabeling; the third index plays the role of w. Hence no fixed memory–memory position can be deleted from a universally sufficient entry set. This proves the worst-case claim and completes the proof. □

E.5 Proof for Proposition 2

Proof. Use the two corpora constructed in the proof of Theorem 3. Define $c_\star \coloneqq \sqrt{1-s^2}\sqrt{1-a^2}$. Because $0 < a < s < 1$, we have $c_\star > 0$. In addition to (102), choose h small enough that $c_\star h < \varepsilon$. Such an $h > 0$ exists because both $\min\{c_0, 1 - c_0\}$ and $\varepsilon / c_\star$ are strictly positive.

Let Γ_+ and Γ_- be the augmented Gram matrices of (q, z_0, w, y_+) and (q, z_0, w, y_-), respectively. By (105)– (106), the matrices agree in every entry except the symmetric (z_0, y) positions, and

$$|(\Gamma_+)_{z_0,y} - (\Gamma_-)_{z_0,y}| = |c_\star(c_0 + h) - c_\star(c_0 - h)| = 2c_\star h. \tag{107}$$

Define the common approximation

$$\widehat{\Gamma} \coloneqq \frac{1}{2}(\Gamma_+ + \Gamma_-). \tag{108}$$

Both Γ_+ and Γ_- are Gram matrices of unit vectors, so they are symmetric, positive semidefinite, and have unit diagonal. Symmetry and the unit diagonal are preserved by their average. For every $v \in \mathbb{R}^4$,

$$v^\top \widehat{\Gamma} v = \frac{1}{2} v^\top \Gamma_+ v + \frac{1}{2} v^\top \Gamma_- v \geq 0,$$

so $\widehat{\Gamma} \succeq 0$. Hence $\widehat{\Gamma} \in \mathfrak{C}_4$. Explicitly, the spectral theorem gives a matrix H such that $\widehat{\Gamma} = H^\top H$; its ith column has squared norm $(\widehat{\Gamma})_{ii} = 1$. Thus these columns realize $\widehat{\Gamma}$ as a Gram matrix of unit vectors.

Since only one symmetric pair differs, equations (107) and (108) give

$$\|\widehat{\Gamma} - \Gamma_+\|_{\max} = \frac{1}{2}\|\Gamma_- - \Gamma_+\|_{\max} = c_\star h < \varepsilon,$$
$$\|\widehat{\Gamma} - \Gamma_-\|_{\max} = \frac{1}{2}\|\Gamma_+ - \Gamma_-\|_{\max} = c_\star h < \varepsilon.$$

This proves (5).

Designate the third indexed memory as the unique relevant memory. Familiarity returns (z_0, w) and therefore fails in both environments. By (104), Recollection returns (z_0, y_+) and succeeds in the $+$ environment, whereas it returns (z_0, w) and fails in the $-$ environment.

Set $C_{\mathrm{F}} = 0$, $C_{\mathrm{R}} = 1$, and $\beta = \Delta_c$. For $L_a = \ell_a + \beta C_a$, the two action losses in the $+$ environment are $L_{\mathrm{F},+} = 1$ and $L_{\mathrm{R},+} = \Delta_c$. Because $0 < \Delta_c < 1$, Recollection is uniquely optimal and the regret of choosing Familiarity is $1 - \Delta_c$. In the $-$ environment, $L_{\mathrm{F},-} = 1$ and $L_{\mathrm{R},-} = 1 + \Delta_c$. Thus Familiarity is uniquely optimal and the regret of choosing Recollection is Δ_c.

A router observing only the common matrix $\widehat{\Gamma}$ must use the same action distribution in both environments. Let $p \in [0, 1]$ be its probability of choosing Recollection. Its two regrets are $\mathrm{Reg}_+(p) = (1 - p)(1 - \Delta_c)$ and $\mathrm{Reg}_-(p) = p\Delta_c$. The first expression is decreasing in p and the second is increasing. Their maximum is minimized at their unique intersection:

$$\begin{aligned}(1 - p)(1 - \Delta_c) &= p\Delta_c,\\ 1 - \Delta_c - p(1 - \Delta_c) &= p\Delta_c,\\ 1 - \Delta_c &= p[(1 - \Delta_c) + \Delta_c] = p.\end{aligned}$$

Hence $p^\star = 1 - \Delta_c$, and substitution yields

$$\mathrm{Reg}_+(p^\star) = \Delta_c(1 - \Delta_c),$$
$$\mathrm{Reg}_-(p^\star) = \Delta_c(1 - \Delta_c).$$

This proves (6). The value is strictly positive and independent of ε. Because the construction is available for every $\varepsilon > 0$, no error-only upper bound tending to zero can hold uniformly over these hard retrieval losses. □

E.6 Proof for Corollary 1

Proof. Apply Lemma 5 with $\mathcal{H}_0 = \mathcal{T}$, $\mathcal{H}_1 = \mathcal{S}$, $\mathcal{H}_2 = \mathcal{V}$ and $\mathcal{H}_3 = \mathcal{G}$. The gate term in (37) is exactly $\mathfrak{E}_{\mathrm{gate}}(\pi)$. Its three adjacent information terms are, in order, $\mathfrak{E}_{\mathrm{summary}}$, $\mathfrak{E}_{\mathrm{trunc+id}}$ and$\mathfrak{E}_{\mathrm{geometry}}$, by their definitions in the corollary. Substitution into (37) proves the displayed RF-Mem decomposition. The nonnegativity of every term follows from the final assertion of the same lemma. □

E.7 Proof for Corollary 2

Proof. Because the initial score vector is the query row of the augmented Gram matrix and the Gram matrix is a measurable function of the full embedding configuration,

$$\mathcal{V} \subseteq \mathcal{K}_\Gamma \subseteq \mathcal{G}. \tag{109}$$

Add and subtract $\mathcal{R}^\star_{\mathcal{K}_\Gamma}$ in the geometry-risk difference:

$$\mathfrak{E}_{\text{geometry}} = \mathcal{R}^\star_{\mathcal{V}} - \mathcal{R}^\star_{\mathcal{G}} = \left(\mathcal{R}^\star_{\mathcal{V}} - \mathcal{R}^\star_{\mathcal{K}_\Gamma}\right) + \left(\mathcal{R}^\star_{\mathcal{K}_\Gamma} - \mathcal{R}^\star_{\mathcal{G}}\right). \tag{110}$$

Apply (38) from Lemma 5 first to $\mathcal{V} \subseteq \mathcal{K}_\Gamma$ and then to $\mathcal{K}_\Gamma \subseteq \mathcal{G}$. This gives

$$\mathcal{R}^\star_{\mathcal{V}} - \mathcal{R}^\star_{\mathcal{K}_\Gamma} = \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{K}_\Gamma}| - |D_{\mathcal{V}}|\right], \tag{111}$$

$$\mathcal{R}^\star_{\mathcal{K}_\Gamma} - \mathcal{R}^\star_{\mathcal{G}} = \frac{1}{2}\mathbb{E}\left[|D_{\mathcal{G}}| - |D_{\mathcal{K}_\Gamma}|\right]. \tag{112}$$

Substitution of (111) and (112) into (110) proves (7).

The nesting in (109) implies $\Pi(\mathcal{V}) \subseteq \Pi(\mathcal{K}_\Gamma) \subseteq \Pi(\mathcal{G})$. Taking infima of the same risk over these successively larger router classes yields $\mathcal{R}^\star_{\mathcal{G}} \le \mathcal{R}^\star_{\mathcal{K}_\Gamma} \le \mathcal{R}^\star_{\mathcal{V}}$. Therefore both terms in (110), and equivalently both terms in (7), are nonnegative.

Finally, if $D_{\mathcal{G}}$ is $\mathcal{K}_\Gamma$-measurable, the tower property gives

$$\begin{aligned} D_{\mathcal{K}_\Gamma} &= \mathbb{E}\left[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \mathcal{K}_\Gamma\right] \\ &= \mathbb{E}\left[\mathbb{E}[L_{\mathrm{F}} - L_{\mathrm{R}} \mid \mathcal{G}] \mid \mathcal{K}_\Gamma\right] = \mathbb{E}\left[D_{\mathcal{G}} \mid \mathcal{K}_\Gamma\right] = D_{\mathcal{G}} \qquad \text{almost surely.} \end{aligned}$$

It follows that $|D_{\mathcal{G}}| - |D_{\mathcal{K}_\Gamma}| = 0$ almost surely and $\mathfrak{E}_{\text{beyond-Gram}} = 0$. □

E.8 Proof for Theorem 4

Proof. We first record a calculation common to all three constructions. Let the environment variable take values in $\{+, -\}$, each with probability $1/2$. We construct each pair so that

$$\ell_{\mathrm{F},+} = 1, \qquad \ell_{\mathrm{R},+} = 0, \qquad \ell_{\mathrm{F},-} = 1, \qquad \ell_{\mathrm{R},-} = 1. \tag{113}$$

Consequently, in the $+$ environment,

$$L_{\mathrm{F},+} - L_{\mathrm{R},+} = 1 + \beta C_{\mathrm{F}} - \beta C_{\mathrm{R}} = 1 - \beta(C_{\mathrm{R}} - C_{\mathrm{F}}) = 1 - \Delta_c > 0, \tag{114}$$

so Recollection is uniquely optimal. In the $-$ environment,

$$L_{\mathrm{F},-} - L_{\mathrm{R},-} = 1 + \beta C_{\mathrm{F}} - 1 - \beta C_{\mathrm{R}} = -\beta(C_{\mathrm{R}} - C_{\mathrm{F}}) = -\Delta_c < 0, \tag{115}$$

so Familiarity is uniquely optimal.

Suppose that a coarser representation is identical in the two environments, while a finer representation distinguishes them. A router based on the coarser representation must choose Recollection with the same probability $p \in [0, 1]$ in both environments. Its regrets are

$$\mathrm{Reg}_+(p) = (1-p)(1-\Delta_c), \tag{116}$$

$$\mathrm{Reg}_-(p) = p\Delta_c. \tag{117}$$

Because the finer representation distinguishes the two atoms, it can select the uniquely optimal action in each. Its Bayes-risk improvement over the coarser representation is therefore

$$\begin{aligned} \inf_{p\in[0,1]} \frac{1}{2}\left[(1-p)(1-\Delta_c) + p\Delta_c\right] &= \inf_{p\in[0,1]} \frac{1}{2}\left[1 - \Delta_c + p(2\Delta_c - 1)\right] \\ &= \frac{1}{2}\min\{\Delta_c, 1-\Delta_c\} = g_c. \end{aligned} \tag{118}$$

Indeed, the affine expression is minimized at $p = 1$ when $\Delta_c < 1/2$, at $p = 0$ when $\Delta_c > 1/2$, and is constant when $\Delta_c = 1/2$.

For minimax regret, the first expression in (116) decreases in p, whereas the second expression in (117) increases in p. Their maximum is minimized at their intersection. Solving

$$\begin{aligned} (1-p)(1-\Delta_c) &= p\Delta_c, \\ 1 - \Delta_c - p(1-\Delta_c) &= p\Delta_c, \\ 1 - \Delta_c &= p[(1-\Delta_c) + \Delta_c] \end{aligned}$$

gives $p = 1 - \Delta_c$. Substitution yields

$$\inf_{p\in[0,1]} \max\{\mathrm{Reg}_+(p), \mathrm{Reg}_-(p)\} = \Delta_c(1-\Delta_c). \tag{119}$$

Geometry reduction. Use the two corpora from Theorem 1. That theorem gives the retrieval losses in (113) and $V_+ = V_-$. Since S and T are deterministic functions of V, $S_+ = S_-$ and $T_+ = T_-$. Thus, on the equal-prior two-point space, $\mathcal{R}^\star_{\mathcal{T}} = \mathcal{R}^\star_{\mathcal{S}} = \mathcal{R}^\star_{\mathcal{V}}$. The full configurations are distinct and hence generate the full two-point sigma-field. Equation (118) therefore gives $\mathcal{R}^\star_{\mathcal{V}} - \mathcal{R}^\star_{\mathcal{G}} = g_c$. It follows that

$$\mathfrak{E}_{\text{geometry}} = g_c, \qquad \mathfrak{E}_{\text{trunc+id}} = 0, \qquad \mathfrak{E}_{\text{summary}} = 0. \tag{120}$$

We further localize this gap within the Gram refinement. In the construction used by Theorem 1, $y = ae_1 + \sqrt{1-a^2}\, e_2$ and $z_\pm = se_1 \pm \sqrt{1-s^2}\, e_2$. Hence

$$\langle z_+, y\rangle = sa + \sqrt{1-s^2}\sqrt{1-a^2},$$
$$\langle z_-, y\rangle = sa - \sqrt{1-s^2}\sqrt{1-a^2},$$

and therefore $\langle z_+, y\rangle - \langle z_-, y\rangle = 2\sqrt{1-s^2}\sqrt{1-a^2} > 0$. The augmented Gram matrices are distinct. Thus $\mathcal{K}_\Gamma = \sigma(\Gamma)$ also distinguishes the two atoms, so $\mathcal{R}^\star_{\mathcal{K}_\Gamma} = \mathcal{R}^\star_{\mathcal{G}}$. Together with (120), this yields

$$\mathfrak{E}_{\text{off-query}} = g_c, \qquad \mathfrak{E}_{\text{beyond-Gram}} = 0. \tag{121}$$

Top-3 truncation reduction. Let e_1, e_2, e_3 be orthonormal and set $q := e_1$. Choose $0 < a < s < 1$ and define

$$A := 1 + \alpha + (1-\alpha)s, \qquad B := (1-\alpha)\sqrt{1-s^2}.$$

Because $\alpha < 1$ and $s < 1$, we have $A > 0$ and $B > 0$. Furthermore, $\min\left\{s, a + \frac{B}{A}\sqrt{1-a^2}\right\} > a$. Choose

$$a < \eta_2 < \eta_1 < \min\left\{s, a + \frac{B}{A}\sqrt{1-a^2}\right\}, \tag{122}$$

and choose any $a_- \in (0, a)$.

Define

$$z := se_1 + \sqrt{1-s^2}\, e_2,$$
$$w_j := \eta_j e_1 + \sqrt{1-\eta_j^2}\, e_3, \qquad j = 1, 2,$$
$$y_+ := ae_1 + \sqrt{1-a^2}\, e_2,$$
$$y_- := a_- e_1 - \sqrt{1-a_-^2}\, e_2.$$

All these vectors have unit norm. Consider the indexed corpora $\mathcal{M}_+ = (z, w_1, w_2, y_+)$ and $\mathcal{M}_- = (z, w_1, w_2, y_-)$, where the fourth item is uniquely relevant in each environment. Their complete initial score vectors are

$$V_+ = (s, \eta_1, \eta_2, a), \qquad V_- = (s, \eta_1, \eta_2, a_-). \tag{123}$$

Thus $V_+ \neq V_-$. However, $s > \eta_1 > \eta_2 > a > a_-$, so their ordered Top-3 score vectors agree:

$$S_+ = S_- = (s, \eta_1, \eta_2). \tag{124}$$

In particular, $T_+ = T_-$.

Familiarity returns $\{z, w_1, w_2\}$ in both environments and therefore fails in both. Since z is the unique initial Top-1 memory, both environments use the same updated query $q' = \frac{Ae_1 + Be_2}{D}$ and $D := \sqrt{A^2 + B^2}$. In the $+$ environment,

$$\langle q', y_+\rangle = \frac{Aa + B\sqrt{1-a^2}}{D}, \tag{125}$$

$$\langle q', w_j\rangle = \frac{A\eta_j}{D}, \qquad j = 1, 2. \tag{126}$$

By (122), $A\eta_j < Aa + B\sqrt{1-a^2}$. Hence y_+ strictly outranks both distractors, Recollection returns it, and $\ell_{\mathrm{R},+} = 0$.

In the $-$ environment,

$$\langle q', y_- \rangle = \frac{Aa_- - B\sqrt{1-a_-^2}}{D}. \tag{127}$$

Since $\eta_j > a_- > 0$ and $A, B > 0$, $A\eta_j > Aa_- > Aa_- - B\sqrt{1-a_-^2}$. Equations (126) and (127) show that both distractors strictly outrank y_-. Thus Recollection fails and $\ell_{\mathrm{R},-} = 1$.

The losses again satisfy (113). Equation (123) shows that $\mathcal{V}$ distinguishes the two atoms, whereas (124) shows that $\mathcal{S}$ and $\mathcal{T}$ do not. Therefore $\mathcal{R}^\star_{\mathcal{T}} = \mathcal{R}^\star_{\mathcal{S}}$ and $\mathcal{R}^\star_{\mathcal{V}} = \mathcal{R}^\star_{\mathcal{G}}$, and (118) gives

$$\mathfrak{E}_{\mathrm{trunc+id}} = g_c, \qquad \mathfrak{E}_{\mathrm{summary}} = 0, \qquad \mathfrak{E}_{\mathrm{geometry}} = 0. \tag{128}$$

The two score vectors differ only in their fourth, below-Top-3 coordinate, so this construction needs truncation alone rather than loss of Top-3 identity.

Mean–entropy summary reduction. Set $m := 1/2$ and define $A_0 := 1 + \alpha + (1-\alpha)m$ and $B_0 := (1-\alpha)\sqrt{1-m^2}$. Both values are strictly positive. The function $\psi(a) := a + \frac{B_0}{A_0}\sqrt{1-a^2}$ is continuous on $(0,1)$ and satisfies $\psi(m) = m + \frac{B_0}{A_0}\sqrt{1-m^2} > m$. There therefore exists $a \in (0,m)$, sufficiently close to m, such that

$$m < a + \frac{B_0}{A_0}\sqrt{1-a^2}. \tag{129}$$

We next construct distinct Top-3 score vectors having the same mean and softmax entropy. Let $d^+ := (2,-1,-1)$ and $d^- := (2,0,-2)$. For a nonconstant $d = (d_1,d_2,d_3)$ and $t \geq 0$, define

$$p_i^d(t) := \frac{\exp(\lambda t d_i)}{\sum_{k=1}^3 \exp(\lambda t d_k)}, \qquad \mathsf{H}_d(t) := -\sum_{i=1}^3 p_i^d(t)\log p_i^d(t).$$

Writing $Z_d(t) := \sum_{i=1}^3 \exp(\lambda t d_i)$ and $\mu_d(t) := \sum_{i=1}^3 p_i^d(t) d_i$, we have

$$\mathsf{H}_d(t) = \log Z_d(t) - \lambda t \mu_d(t), \tag{130}$$

$$\frac{d}{dt}\log Z_d(t) = \lambda \mu_d(t), \tag{131}$$

$$\mu_d'(t) = \lambda\left[\sum_{i=1}^3 p_i^d(t) d_i^2 - \mu_d(t)^2\right] = \lambda \operatorname{Var}_{p^d(t)}(d). \tag{132}$$

Differentiating (130) and using (131)–(132) gives

$$\mathsf{H}_d'(t) = -\lambda^2 t \operatorname{Var}_{p^d(t)}(d) < 0, \qquad t > 0. \tag{133}$$

The strict inequality holds because all softmax probabilities are positive and d is nonconstant.

At $t = 0$, the softmax distribution is uniform, so $\mathsf{H}_d(0) = \log 3$. Both d^+ and d^- have a unique largest coordinate. Their softmax distributions therefore converge to $(1,0,0)$ as $t \to \infty$, and $\lim_{t\to\infty} \mathsf{H}_{d^+}(t) = \lim_{t\to\infty} \mathsf{H}_{d^-}(t) = 0$. It follows from continuity and strict monotonicity that, for every $h \in (0, \log 3)$, there are unique $u(h), v(h) > 0$ satisfying

$$\mathsf{H}_{d^+}(u(h)) = h = \mathsf{H}_{d^-}(v(h)). \tag{134}$$

Moreover, $u(h) \to 0$, $v(h) \to 0$ and $h \uparrow \log 3$. Indeed, any sequence of inverse values bounded away from zero would, by strict decrease of the corresponding entropy curve, have entropy bounded strictly below $\log 3$.

For u near zero, define

$$A_+(u) := 1 + \alpha + (1-\alpha)(m + 2u),$$
$$B_+(u) := (1-\alpha)\sqrt{1-(m+2u)^2}.$$

At $u = 0$, these equal A_0 and B_0. By (129) and continuity, for all sufficiently small $u > 0$,

$$m - u < a + \frac{B_+(u)}{A_+(u)}\sqrt{1-a^2}. \tag{135}$$

Choose $h < \log 3$ sufficiently close to $\log 3$ that, with $u := u(h)$ and $v := v(h)$,

$$m + 2u < 1, \qquad m + 2v < 1, \tag{136}$$

$$m - u > a, \qquad m - 2v > a, \tag{137}$$

and (135) holds. Such an h exists because $u(h), v(h) \to 0$, $a < m$, and the inequality in (129) is strict.

Define the common query and relevant memory by

$$q := e_1, \qquad y := ae_1 + \sqrt{1 - a^2}\, e_2.$$

In the $+$ environment, define

$$\begin{aligned} z_+ &:= (m + 2u)e_1 + \sqrt{1 - (m + 2u)^2}\, e_2, \\ w_{+,1} &:= (m - u)e_1 + \sqrt{1 - (m - u)^2}\, e_3, \\ w_{+,2} &:= (m - u)e_1 - \sqrt{1 - (m - u)^2}\, e_3. \end{aligned}$$

In the $-$ environment, define

$$\begin{aligned} z_- &:= (m + 2v)e_1 - \sqrt{1 - (m + 2v)^2}\, e_2, \\ w_{-,1} &:= me_1 + \sqrt{1 - m^2}\, e_3, \\ w_{-,2} &:= (m - 2v)e_1 - \sqrt{1 - (m - 2v)^2}\, e_3. \end{aligned}$$

All vectors have unit norm. Set

$$\mathcal{M}_+ := \{z_+, w_{+,1}, w_{+,2}, y\}, \qquad \mathcal{M}_- := \{z_-, w_{-,1}, w_{-,2}, y\}.$$

Their ordered Top-3 score vectors are

$$S_+ = (m + 2u, m - u, m - u) = m\mathbf{1}_3 + ud^+, \tag{138}$$

$$S_- = (m + 2v, m, m - 2v) = m\mathbf{1}_3 + vd^-. \tag{139}$$

They are distinct: equality of their second coordinates would imply $m - u = m$, contradicting $u > 0$. Their means are both exactly m because the coordinates of both d^+ and d^- sum to zero. Adding the common constant m to all logits leaves their softmax probabilities unchanged. Equation (134) therefore gives

$$H(S_+) = \mathsf{H}_{d^+}(u) = h = \mathsf{H}_{d^-}(v) = H(S_-).$$

Consequently,

$$T_+ = (m, h) = T_-, \qquad S_+ \neq S_-. \tag{140}$$

By (137), every distractor has a larger initial score than y. Thus Familiarity fails in both environments. In the $+$ environment, z_+ is the unique initial Top-1 memory, and the updated query is

$$q'_+ = \frac{A_+(u)e_1 + B_+(u)e_2}{D_+}, \qquad D_+ := \sqrt{A_+(u)^2 + B_+(u)^2}.$$

Its relevant-memory and distractor scores are

$$\begin{aligned} \langle q'_+, y \rangle &= \frac{A_+(u)a + B_+(u)\sqrt{1 - a^2}}{D_+}, \\ \langle q'_+, w_{+,j} \rangle &= \frac{A_+(u)(m - u)}{D_+}, \qquad j = 1, 2. \end{aligned}$$

By (135), y strictly outranks both distractors. Recollection therefore succeeds and $\ell_{\mathrm{R},+} = 0$.

In the $-$ environment, define

$$A_- := 1 + \alpha + (1 - \alpha)(m + 2v), \qquad B_- := (1 - \alpha)\sqrt{1 - (m + 2v)^2},$$

$$D_- := \sqrt{A_-^2 + B_-^2}.$$

Since z_- is the unique initial Top-1 memory,

$$q'_- = \frac{A_- e_1 - B_- e_2}{D_-}.$$

The relevant-memory score is

$$\langle q'_-, y \rangle = \frac{A_- a - B_- \sqrt{1-a^2}}{D_-}.$$

The two distractor scores are

$$\langle q'_-, w_{-,1} \rangle = \frac{A_- m}{D_-}, \qquad \langle q'_-, w_{-,2} \rangle = \frac{A_-(m-2v)}{D_-}.$$

Both m and $m-2v$ exceed a. Hence, for $\eta \in \{m, m-2v\}$,

$$A_- \eta > A_- a > A_- a - B_- \sqrt{1-a^2}.$$

Both distractors strictly outrank y, so Recollection fails and $\ell_{\mathrm{R},-} = 1$.

The losses again satisfy (113). By (140), $\mathcal{T}$ does not distinguish the environments, whereas $\mathcal{S}$ does. Thus

$$\mathcal{R}^\star_{\mathcal{S}} = \mathcal{R}^\star_{\mathcal{V}} = \mathcal{R}^\star_{\mathcal{G}},$$

and (118) gives

$$\mathfrak{E}_{\mathrm{summary}} = g_c, \qquad \mathfrak{E}_{\mathrm{trunc+id}} = 0, \qquad \mathfrak{E}_{\mathrm{geometry}} = 0. \tag{141}$$

Equation (119) applies to the coarser representation in all three constructions. This proves the minimax claim and completes the proof. □

E.9 Proof for Theorem 5

Proof. Choose $0 < a < s < 1$, $0 < \rho < 1$, and put

$$A = 1 + \alpha + (1-\alpha)s, \quad b = (1-\alpha)\sqrt{1-s^2}, \quad D = \sqrt{A^2 + b^2}.$$

Choose $a < \eta < \min\{s, a + (b/A)\sqrt{1-a^2}\rho\}$. In $\mathbb{R}^{m+3}$, let

$$q = e_1, \qquad z = s e_1 + \sqrt{1-s^2} e_2, \qquad w = \eta e_1 + \sqrt{1-\eta^2} e_3.$$

The initial Top-K consists of z and $K-1$ indexed copies of w. For candidate indices $j = 1, \ldots, m$ outside this set, define

$$y^0_j = a e_1 + \sqrt{1-a^2} e_{j+3}, \qquad y^1_j = a e_1 + \sqrt{1-a^2}\big(\rho e_2 + \sqrt{1-\rho^2} e_{j+3}\big).$$

Let $W \in \{0,1\}$ be fair and let J be uniform on $[m]$, independent of W. If $W = 0$, include all y^0_j; if $W = 1$, replace only y^0_J by y^1_J. In both cases the unique relevant index is candidate J. These are $2m$ equally likely environments; the $W = 0$ environments share the same embeddings and differ only in relevance labels. All have the same initial profile $(s, \eta^{[K-1]}, a^{[m]})$ and initial Top-K indices. Familiarity therefore always fails.

Only the seed–candidate entries depend on W, J:

$$\langle z, y_j \rangle = sa + \zeta \mathbf{1}\{W = 1, J = j\}, \qquad \zeta = \sqrt{1-s^2}\sqrt{1-a^2}\rho > 0. \tag{142}$$

Indeed, each candidate has projection a onto q and inner product $a\eta$ with every distractor. Distinct candidates have inner product a^2, since at most one has a nonzero e_2 component and their remaining tangential coordinates are distinct. Self-inner-products are one. The seed, distractors, and all their mutual entries are fixed. This verifies both Gram validity by explicit vectors and constancy of every unlisted entry.

At round zero, Algorithm 3 retrieves the unique seed and forms $q' = (Ae_1 + be_2)/D$. Its seed score is strictly greater than $A\eta/D$, its distractor score. Unmarked candidates score $Aa/D < A\eta/D$, while the marked candidate, if present, scores $(Aa + b\sqrt{1-a^2}\rho)/D > A\eta/D$. Thus round one's

full-corpus Top-2 contains the seed and the marked candidate when $W = 1$, and the seed and the first distractor when $W = 0$. There is one cluster and no pruning; Bag contains exactly these two indices, which final Top-K returns since $K \geq 3$. All projections onto q are positive, so all centroid and query normalizations are nonzero. Hence the failure losses are $(1, 0)$ when $W = 1$ and $(1, 1)$ when $W = 0$. The full Gram matrix identifies W, so its Bayes risk equals $\mathcal{R}^\star_{\mathcal{G}}$. Conditional on $W = 0$, the unobserved label J does not affect either potential loss.

A request outside the listed m entries has a deterministic answer and cannot improve routing. Repeated requests likewise add no information. If a marked entry is found, $W = 1$ is known and Recollection is uniquely optimal; further requests cannot improve the decision. Conditional on any sequence of u distinct unmarked requests, the remaining location J, given $W = 1$, is uniform on the $m - u$ untested indices. Consequently their order does not matter. For a fixed request limit t, testing until the mark or that limit gives

$$\mathbb{P}(W = 1, \text{no mark}) = \frac{m - t}{2m}, \qquad \mathbb{P}(W = 0, \text{no mark}) = \frac{1}{2}.$$

On a no-mark history, choosing Familiarity therefore contributes unconditional excess risk $(1 - t/m)(1-\Delta_c)/2$, whereas choosing Recollection contributes $\Delta_c/2$. Taking the smaller gives (9). No other policy with at most t requests can do better: it can be augmented to test t distinct candidates on its no-mark path, and the additional information cannot increase its optimized risk. This augmented policy has exactly the experiment just analyzed. Randomization averages such policies and does not improve their minimum.

For the attaining policy, conditional on $W = 0$ there are t requests. Conditional on $W = 1$, the marked position in the test order is uniform, so

$$\mathbb{E}[Q \mid W = 1] = \frac{1}{m}\left(\sum_{j=1}^{t} j + (m - t)t\right) = t - \frac{t(t-1)}{2m}.$$

Averaging yields $\mathbb{E}Q = t - t(t-1)/(4m)$.

To optimize total cost, first remove all uninformative or repeated requests and stop immediately after a mark. Conditional on its private random seed, any remaining policy has a single all-unmarked path and thus a deterministic cutoff $t \in \{0, \ldots, m\}$. Its best final action yields total excess risk

$$f(t) = c_p\left(t - \frac{t(t-1)}{4m}\right) + \frac{1}{2}\min\left\{\Delta_c, \left(1 - \frac{t}{m}\right)(1 - \Delta_c)\right\}.$$

As a function on the real interval $[0, m]$, f is concave: the first term is a concave quadratic and the second is the minimum of two affine functions. A concave function is bounded below by the smaller of its endpoint values. Here $f(0) = g_c$ and $f(m) = c_p(3m + 1)/4$. Both endpoints are attainable, by not probing or by testing until a mark or exhaustion. Mixtures cannot improve the smaller value. This proves (10) and its strict improvement criterion, including $m = 1$.

Finally, an approximate response $\widehat{t}$ with error at most $\varepsilon < \zeta/2$ is classified correctly by the known threshold $sa + \zeta/2$. Every branch of the attaining policies is therefore unchanged. This robustness assertion is an achievability guarantee; it does not permit an otherwise unspecified approximation mechanism to encode extra information in its errors. The exact lower bounds above concern the stated individual-entry observation model. □

E.10 Proof for Theorem 6

Proof. Let

$$p^{\downarrow}_1 \geq p^{\downarrow}_2 \geq \cdots \geq p^{\downarrow}_K$$

be the coordinates of p in nonincreasing order. Because $p_{\max} = x$, we have $p^{\downarrow}_j \leq x$ for every j.

We first prove

$$p^{\min}(x) \succeq p.$$

For every $m \le r_x$,

$$\begin{aligned}\sum_{j=1}^{m} p_j^{\downarrow} &\le \sum_{j=1}^{m} x \\ &= mx \\ &= \sum_{j=1}^{m} p_j^{\min}(x).\end{aligned}$$

Suppose first that $u_x > 0$. For every $m \ge r_x + 1$,

$$\sum_{j=1}^{m} p_j^{\min}(x) = r_x x + u_x = 1.$$

Since p is a probability vector,

$$\sum_{j=1}^{m} p_j^{\downarrow} \le 1.$$

Hence

$$\sum_{j=1}^{m} p_j^{\min}(x) \ge \sum_{j=1}^{m} p_j^{\downarrow}.$$

If $u_x = 0$, then $r_x x = 1$. Therefore, for every $m \ge r_x$, $\sum_{j=1}^{m} p_j^{\min}(x) = 1$ and the same inequality follows. Finally,

$$\sum_{j=1}^{K} p_j^{\min}(x) = 1 = \sum_{j=1}^{K} p_j.$$

Thus

$$p^{\min}(x) \succeq p.$$

By Lemma 10,

$$H(p^{\min}(x)) \le H(p).$$

Direct calculation yields

$$\begin{aligned}H(p^{\min}(x)) &= -\sum_{j=1}^{r_x} x \log x - u_x \log u_x \\ &= -r_x x \log x - u_x \log u_x \\ &= H_K^{\min}(x).\end{aligned}$$

Therefore, $H_K^{\min}(x) \le H(p)$. Since $p^{\min}(x)$ itself satisfies $p_{\max}^{\min}(x) = x$, the lower bound is attainable and is therefore sharp. If $x = 1$, then p is a permutation of $(1, 0, \dots, 0)$ and $H(p) = 0 = H_K^{\max}(1)$. Suppose henceforth that $x < 1$. Choose an index $j_\star$ such that $p_{j_\star} = x$. For every $j \ne j_\star$, define

$$v_j := \frac{p_j}{1 - x}.$$

Then

$$v_j \ge 0$$

and

$$\begin{aligned}\sum_{j \ne j_\star} v_j &= \frac{\sum_{j \ne j_\star} p_j}{1 - x} \\ &= \frac{1 - x}{1 - x} \\ &= 1.\end{aligned}$$

Thus v is a probability vector on $K - 1$ outcomes.

We decompose the entropy:

$$\begin{aligned} H(p) &= -x\log x - \sum_{j\neq j_\star} p_j \log p_j \\ &= -x\log x - \sum_{j\neq j_\star} (1-x)v_j \log\big((1-x)v_j\big) \\ &= -x\log x - (1-x)\sum_{j\neq j_\star} v_j\,[\log(1-x) + \log v_j] \\ &= -x\log x - (1-x)\log(1-x)\sum_{j\neq j_\star} v_j \\ &\quad - (1-x)\sum_{j\neq j_\star} v_j \log v_j \\ &= h_2(x) + (1-x)H(v). \end{aligned} \tag{143}$$

The entropy of a distribution on $K-1$ outcomes is at most

$$\log(K-1).$$

To verify this, let

$$u_j := \frac{1}{K-1}.$$

Nonnegativity of Kullback–Leibler divergence gives

$$\begin{aligned} 0 &\le D_{\mathrm{KL}}(v\|u) \\ &= \sum_{j\neq j_\star} v_j \log \frac{v_j}{1/(K-1)} \\ &= \sum_{j\neq j_\star} v_j \log v_j + \log(K-1)\sum_{j\neq j_\star} v_j \\ &= -H(v) + \log(K-1). \end{aligned}$$

Hence

$$H(v) \le \log(K-1).$$

Substitution into (143) yields

$$\begin{aligned} H(p) &\le h_2(x) + (1-x)\log(K-1) \\ &= H_K^{\max}(x). \end{aligned}$$

Equality holds when v is uniform, which gives $p^{\max}(x)$. Thus the upper bound is also sharp. □

E.11 Proof for Proposition 5

Proof. Suppose first that $s \in [-1,1]^K$ and that p is given by (15). Writing

$$Z := \sum_{\ell=1}^{K} e^{\lambda s_\ell},$$

we have

$$\log p_j = \lambda s_j - \log Z.$$

Therefore,

$$\begin{aligned} \max_j \log p_j - \min_j \log p_j &= \lambda\left(\max_j s_j - \min_j s_j\right) \\ &\le 2\lambda. \end{aligned}$$

All coordinates are strictly positive and sum to one, so $p \in \Delta_{K,\lambda}^{\cos}$.

Conversely, let $p \in \Delta^{\cos}_{K,\lambda}$ and set

$$r_j := \frac{1}{\lambda} \log p_j.$$

The defining range bound gives

$$\max_j r_j - \min_j r_j \le 2. \tag{144}$$

Let

$$c := \frac{1}{2}\left(\max_j r_j + \min_j r_j\right), \qquad s_j := r_j - c.$$

By (144), $s_j \in [-1, 1]$ for every j. Moreover,

$$\begin{aligned}\frac{e^{\lambda s_j}}{\sum_{\ell=1}^K e^{\lambda s_\ell}} &= \frac{e^{\log p_j - \lambda c}}{\sum_{\ell=1}^K e^{\log p_\ell - \lambda c}} \\ &= \frac{p_j}{\sum_{\ell=1}^K p_\ell} \\ &= p_j.\end{aligned}$$

This proves the equivalence.

Finally, for any feasible score vector and any j,

$$p_j = \frac{1}{1 + \sum_{\ell \ne j} e^{\lambda(s_\ell - s_j)}}. \tag{145}$$

Because $-2 \le s_\ell - s_j \le 2$,

$$(K-1)e^{-2\lambda} \le \sum_{\ell \ne j} e^{\lambda(s_\ell - s_j)} \le (K-1)e^{2\lambda}.$$

Substitution into (145) yields (16). Strict positivity excludes every zero-coordinate simplex extremizer. □

E.12 Proof for Corollary 4

Proof. Set $\tau := H_K^{\min}(\rho)$. Then $\psi_K(\tau) = \rho$, so the first conclusion follows from Lemma 12. To prove maximality, suppose $\tau' > H_K^{\min}(\rho)$. If $\rho > 1/K$, continuity and strict monotonicity from Lemma 11 imply that there exists $x < \rho$, sufficiently close to ρ, such that $H_K^{\min}(x) < \tau'$. Taking $p = p^{\min}(x)$, we obtain $H(p) < \tau'$ but $p_{\max} = x < \rho$. Therefore, τ' does not provide the desired guarantee. When $\rho = 1/K$, the guarantee $p_{\max} \ge 1/K$ holds for every probability vector, and $H_K^{\min}(1/K) = \log K$ is already the largest possible entropy. □

E.13 Proof for Proposition 6

Proof. Differentiate

$$H_K^{\max}(x) = h_2(x) + (1-x)\log(K-1).$$

Since

$$h_2'(x) = \log \frac{1-x}{x},$$

we have

$$\begin{aligned}\frac{d}{dx} H_K^{\max}(x) &= \log \frac{1-x}{x} - \log(K-1) \\ &= \log \frac{1-x}{x(K-1)}.\end{aligned}$$

For $x \ge 1/K$, $1 - x \le x(K-1)$, so $\frac{1-x}{x(K-1)} \le 1$. The inequality is strict for $x > 1/K$. Hence $H_K^{\max}$ is strictly decreasing. From the upper entropy envelope, $H(p) \le H_K^{\max}(p_{\max})$. If $H(p) \ge \tau$, then $H_K^{\max}(p_{\max}) \ge \tau$. Since $H_K^{\max}$ is decreasing, $p_{\max} \le \chi_K(\tau)$. □

E.14 Proof for Theorem 7

Proof. Lemma 15 shows that within-list calibration does not control probe coverage. The present construction establishes the distinct failure of marginal top-score calibration after entropy selection. The coordinates of p^{L} sum to

$$x + [1 - x - (K-2)\epsilon] + (K-2)\epsilon = 1.$$

Here,

$$\begin{aligned} 1 - x - (K-2)\epsilon &> 1 - x - (K-2)\frac{1-x}{K-1} \\ &= \frac{1-x}{K-1} \\ &> 0. \end{aligned}$$

Thus all coordinates are positive. Since $x > 1/2$, $1 - x < x$. Therefore,

$$1 - x - (K-2)\epsilon < 1 - x < x.$$

Also, $\epsilon < x$. Thus the maximum coordinate of p^{L} is x. For p^{H},

$$\frac{1-x}{K-1} < 1 - x < x,$$

so its maximum coordinate is also x. We next compare the entropies. Conditional on the first coordinate having mass x, the normalized residual distribution of p^{H} is uniform:

$$\left(\frac{1}{K-1}, \ldots, \frac{1}{K-1}\right).$$

The normalized residual distribution of p^{L} is

$$\left(\frac{1-x-(K-2)\epsilon}{1-x}, \underbrace{\frac{\epsilon}{1-x}, \ldots, \frac{\epsilon}{1-x}}_{K-2 \text{ entries}}\right). \tag{146}$$

The first coordinate in (146) is strictly larger than the others because

$$1 - x - (K-2)\epsilon > \epsilon \iff 1 - x > (K-1)\epsilon.$$

Hence the residual distribution of p^{L} is nonuniform. Since Shannon entropy is strictly maximized by the uniform distribution, $H(p^{\mathrm{L}}) < H(p^{\mathrm{H}})$. Choose

$$\tau := \frac{H(p^{\mathrm{L}}) + H(p^{\mathrm{H}})}{2}.$$

Then

$$H(p^{\mathrm{L}}) < \tau < H(p^{\mathrm{H}}).$$

Define the random vector $\mathbf{P}$ by

$$\begin{aligned} \mathbb{P}(\mathbf{P} = p^{\mathrm{H}}) &= x, \\ \mathbb{P}(\mathbf{P} = p^{\mathrm{L}}) &= 1 - x. \end{aligned}$$

Define

$$C_{\mathrm{F}} = \begin{cases} 1, & \mathbf{P} = p^{\mathrm{H}}, \\ 0, & \mathbf{P} = p^{\mathrm{L}}. \end{cases}$$

Under both possible values of $\mathbf{P}$,

$$P_{\max} = x.$$

Hence $P_{\max}$ is constant and

$$\begin{aligned} \mathbb{E}[C_{\mathrm{F}} \mid P_{\max}] &= \mathbb{E}[C_{\mathrm{F}}] \\ &= 1 \cdot \mathbb{P}(\mathbf{P} = p^{\mathrm{H}}) + 0 \cdot \mathbb{P}(\mathbf{P} = p^{\mathrm{L}}) \\ &= x \\ &= P_{\max}. \end{aligned}$$

Thus perfect scalar calibration holds. However, by the choice of τ,

$$\{H(\mathbf{P}) \le \tau\} = \{\mathbf{P} = p^{\mathrm{L}}\}.$$

On this event,

$$C_{\mathrm{F}} = 0.$$

Therefore,

$$\mathbb{P}\left(C_{\mathrm{F}} = 1 \mid H(\mathbf{P}) \le \tau\right) = 0.$$

□

E.15 Proof for Theorem 8

Proof. Although the selected order statistics are dependent as shown in Lemma 16, the raw scores remain independent; we therefore apply the self-bounding argument to the raw inputs. Normalize the raw scores: $Y_i := \frac{X_i - \ell}{\Delta_s}$. Then $0 \le Y_i \le 1$ almost surely, and the Y_i remain independent. Let $Y_{[1]} \ge \cdots \ge Y_{[n]}$ be the decreasing order statistics and define

$$Z := \sum_{j=1}^{K} Y_{[j]}.$$

Because each selected value lies in $[0,1]$, $0 \le Z \le K$. The positive affine transformation preserves ordering, so

$$\begin{aligned}
\widehat{M}_{n,K} &= \frac{1}{K} \sum_{j=1}^{K} \left(\ell + \Delta_s Y_{[j]}\right) \\
&= \ell + \frac{\Delta_s}{K} \sum_{j=1}^{K} Y_{[j]} \\
&= \ell + \frac{\Delta_s}{K} Z. \qquad (147)
\end{aligned}$$

For a deterministic vector $y = (y_1, \ldots, y_n) \in [0,1]^n$, write

$$Z(y) = \max_{\substack{I \subseteq \{1,\ldots,n\} \\ |I| = K}} \sum_{j \in I} y_j.$$

For every coordinate i, define

$$Z_i(y_{-i}) := Z(y_1, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_n).$$

This function does not depend on y_i. Because $y_i \ge 0$, replacing y_i by zero cannot increase the sum of the largest K coordinates. Therefore,

$$Z(y) - Z_i(y_{-i}) \ge 0.$$

Choose a deterministic maximizing set $I^\star(y)$, for example the lexicographically smallest maximizer, satisfying

$$Z(y) = \sum_{j \in I^\star(y)} y_j.$$

Suppose first that $i \notin I^\star(y)$. After setting y_i to zero, the same set $I^\star(y)$ remains feasible and its sum is unchanged. Thus $Z_i(y_{-i}) \ge Z(y)$. Together with $Z_i(y_{-i}) \le Z(y)$, we obtain

$$Z(y) - Z_i(y_{-i}) = 0 \qquad \text{if } i \notin I^\star(y).$$

Suppose next that $i \in I^\star(y)$. After replacing y_i by zero, the same selected set has value $Z(y) - y_i$. Since Z_i is the maximum over all size-K sets after the replacement, $Z_i(y_{-i}) \ge Z(y) - y_i$. Therefore,

$$Z(y) - Z_i(y_{-i}) \le y_i \le 1.$$

Combining the two cases, $0 \le Z - Z_i \le 1$. Furthermore,

$$\begin{aligned}\sum_{i=1}^{n}[Z(y) - Z_i(y_{-i})] &= \sum_{i \in I^\star(y)} [Z(y) - Z_i(y_{-i})] \\ &\le \sum_{i \in I^\star(y)} y_i \\ &= Z(y).\end{aligned}$$

Thus Z is self-bounding. Applying Lemma 21, for every $t \ge 0$,

$$\mathbb{P}(Z - \mathbb{E}Z \ge t) \le \exp\left\{-\frac{t^2}{2\mathbb{E}Z + 2t/3}\right\},$$
$$\mathbb{P}(\mathbb{E}Z - Z \ge t) \le \exp\left\{-\frac{t^2}{2\mathbb{E}Z}\right\}.$$

Since $Z \le K$, $\mathbb{E}Z \le K$. Set

$$t := \frac{K\varepsilon}{\Delta_s}.$$

From (147), $\widehat{M}_{n,K} - \mathbb{E}[\widehat{M}_{n,K}] = \frac{\Delta_s}{K}(Z - \mathbb{E}Z)$. Therefore,

$$\begin{aligned}\mathbb{P}\left(\widehat{M}_{n,K} - \mathbb{E}[\widehat{M}_{n,K}] \ge \varepsilon\right) &= \mathbb{P}\left(Z - \mathbb{E}Z \ge \frac{K\varepsilon}{\Delta_s}\right) \\ &\le \exp\left\{-\frac{K^2\varepsilon^2/\Delta_s^2}{2\mathbb{E}Z + 2K\varepsilon/(3\Delta_s)}\right\} \\ &\le \exp\left\{-\frac{K^2\varepsilon^2/\Delta_s^2}{2K + 2K\varepsilon/(3\Delta_s)}\right\} \\ &= \exp\left\{-\frac{K\varepsilon^2}{2\Delta_s^2 + 2\Delta_s\varepsilon/3}\right\}.\end{aligned}$$

This proves the upper-tail inequality. Similarly,

$$\begin{aligned}\mathbb{P}\left(\mathbb{E}[\widehat{M}_{n,K}] - \widehat{M}_{n,K} \ge \varepsilon\right) &= \mathbb{P}\left(\mathbb{E}Z - Z \ge \frac{K\varepsilon}{\Delta_s}\right) \\ &\le \exp\left\{-\frac{K^2\varepsilon^2/\Delta_s^2}{2\mathbb{E}Z}\right\} \\ &\le \exp\left\{-\frac{K^2\varepsilon^2/\Delta_s^2}{2K}\right\} = \exp\left\{-\frac{K\varepsilon^2}{2\Delta_s^2}\right\}.\end{aligned}$$

The union bound gives

$$\begin{aligned}\mathbb{P}\left(\left|\widehat{M}_{n,K} - \mathbb{E}[\widehat{M}_{n,K}]\right| \ge \varepsilon\right) &\le \exp\left\{-\frac{K\varepsilon^2}{2\Delta_s^2}\right\} + \exp\left\{-\frac{K\varepsilon^2}{2\Delta_s^2 + 2\Delta_s\varepsilon/3}\right\} \\ &\le 2\exp\left\{-\frac{K\varepsilon^2}{2\Delta_s^2 + 2\Delta_s\varepsilon/3}\right\}.\end{aligned}$$

Finally, the variance part of the self-bounding lemma gives $\mathrm{Var}(Z) \le \mathbb{E}Z \le K$. Using (147),

$$\mathrm{Var}(\widehat{M}_{n,K}) = \frac{\Delta_s^2}{K^2}\,\mathrm{Var}(Z) \le \frac{\Delta_s^2}{K^2}K = \frac{\Delta_s^2}{K}.$$

□

E.16 Proof for Theorem 9

Proof. Define the empirical objective

$$\widehat{\Phi}_n(t) := t + \frac{1}{\gamma n}\sum_{i=1}^{n}(X_i - t)_+$$

and the population objective

$$\Phi(t) := t + \frac{1}{\gamma}\mathbb{E}[(X-t)_+].$$

Because $\gamma n = K$, Lemma 17 gives $\widehat{M}_{n,K} = \inf_{t\in[\ell,u]} \widehat{\Phi}_n(t)$. By definition, $U_\gamma(F) = \inf_{t\in[\ell,u]} \Phi(t)$. For arbitrary bounded functions f, g on a common domain,

$$\left|\inf_t f(t) - \inf_t g(t)\right| \le \sup_t |f(t) - g(t)|.$$

Indeed, if $\eta := \sup_t |f(t) - g(t)|$, then $g(t) - \eta \le f(t) \le g(t) + \eta$ for every t. Taking infima gives $\inf_t g(t) - \eta \le \inf_t f(t) \le \inf_t g(t) + \eta$. Therefore,

$$\begin{aligned}\left|\widehat{M}_{n,K} - U_\gamma(F)\right| &\le \sup_{t\in[\ell,u]} \left|\widehat{\Phi}_n(t) - \Phi(t)\right| \\ &= \frac{1}{\gamma}\sup_{t\in[\ell,u]} \left|\frac{1}{n}\sum_{i=1}^n (X_i - t)_+ - \mathbb{E}[(X-t)_+]\right|. \end{aligned} \tag{148}$$

For every $x, t \in [\ell, u]$,

$$(x-t)_+ = \int_t^u \mathbf{1}\{x > s\}\, ds.$$

Indeed, if $x \le t$, both sides are zero. If $x > t$, the integral is over $s \in [t, x)$ and equals $x - t$. Consequently,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n (X_i - t)_+ - \mathbb{E}[(X-t)_+] &= \int_t^u \left[\frac{1}{n}\sum_{i=1}^n \mathbf{1}\{X_i > s\} - \mathbb{P}(X > s)\right] ds \\ &= \int_t^u [\{1 - F_n(s)\} - \{1 - F(s)\}]\, ds \\ &= \int_t^u [F(s) - F_n(s)]\, ds.\end{aligned}$$

Taking absolute values,

$$\begin{aligned}\left|\frac{1}{n}\sum_{i=1}^n (X_i - t)_+ - \mathbb{E}[(X-t)_+]\right| &\le \int_t^u |F_n(s) - F(s)|\, ds \\ &\le (u-t)\sup_{s\in\mathbb{R}} |F_n(s) - F(s)| \\ &\le \Delta_s \sup_{s\in\mathbb{R}} |F_n(s) - F(s)|.\end{aligned}$$

Substitution into (148) gives the deterministic inequality

$$\left|\widehat{M}_{n,K} - U_\gamma(F)\right| \le \frac{\Delta_s}{\gamma}\sup_{s\in\mathbb{R}} |F_n(s) - F(s)|. \tag{149}$$

Therefore,

$$\begin{aligned}\mathbb{P}\left(\left|\widehat{M}_{n,K} - U_\gamma(F)\right| > \varepsilon\right) &\le \mathbb{P}\left(\sup_s |F_n(s) - F(s)| > \frac{\gamma\varepsilon}{\Delta_s}\right) \\ &\le 2\exp\left\{-2n\left(\frac{\gamma\varepsilon}{\Delta_s}\right)^2\right\} \\ &= 2\exp\left\{-\frac{2n\gamma^2\varepsilon^2}{\Delta_s^2}\right\},\end{aligned}$$

where the second inequality follows from Lemma 22. Finally, $n\gamma^2 = n\left(\frac{K}{n}\right)^2 = \frac{K^2}{n}$, which gives the equivalent form. □

E.17 Proof for Theorem 10

Proof. By Lemma 23, the sequence $\widehat{M}_{n,K}$ is nondecreasing on every sample path. We now identify its limit. By the definition of essential supremum, $X_i \le x_F^\star$ almost surely for every i. Since there are countably many i, there is a probability-one event on which $X_i \le x_F^\star$ simultaneously for all i. On this event, $\widehat{M}_{n,K} \le x_F^\star$ for every n. For every integer $r \ge 1$, define $\varepsilon_r := \frac{1}{r}$. By the definition of essential supremum,

$$p_r := \mathbb{P}\left(X_1 > x_F^\star - \varepsilon_r\right) > 0.$$

Define the count

$$N_n^{(r)} := \sum_{i=1}^{n} \mathbf{1}\left\{X_i > x_F^\star - \varepsilon_r\right\}.$$

By the strong law of large numbers,

$$\frac{N_n^{(r)}}{n} \longrightarrow p_r \qquad \text{almost surely.} \tag{150}$$

Since $p_r > 0$, it follows that, almost surely, $N_n^{(r)} \ge K$ for all sufficiently large n. Whenever $N_n^{(r)} \ge K$, at least K observations exceed $x_F^\star - \varepsilon_r$. Therefore every member of the Top-K set exceeds this threshold, and hence $\widehat{M}_{n,K} > x_F^\star - \varepsilon_r$. For each r, let E_r be the probability-one event on which (150) holds. The countable intersection $E := \bigcap_{r=1}^{\infty} E_r$ also has probability one. On E, for every r, there exists n_r such that for all $n \ge n_r$, $x_F^\star - \frac{1}{r} < \widehat{M}_{n,K} \le x_F^\star$. Therefore, $\liminf_{n\to\infty} \widehat{M}_{n,K} \ge x_F^\star - \frac{1}{r}$ for every r. Letting $r \to \infty$, $\liminf_{n\to\infty} \widehat{M}_{n,K} \ge x_F^\star$. Together with the upper bound, $\limsup_{n\to\infty} \widehat{M}_{n,K} \le x_F^\star$, we obtain $\widehat{M}_{n,K} \to x_F^\star$ almost surely. Finally, $\ell \le \widehat{M}_{n,K} \le u$ for every n. The bounded convergence theorem therefore gives $\mathbb{E}[\widehat{M}_{n,K}] \to x_F^\star$. □

E.18 Proof for Corollary 3

Proof. Set $\epsilon := \frac{x_F^\star - \theta_{\text{high}}}{2} > 0$. By Theorem 10, almost surely there exists N such that for all $n \ge N$, $\left|\widehat{M}_{n,K} - x_F^\star\right| < \epsilon$. Hence

$$\begin{aligned}
\widehat{M}_{n,K} &> x_F^\star - \epsilon \\
&= x_F^\star - \frac{x_F^\star - \theta_{\text{high}}}{2} \\
&= \frac{x_F^\star + \theta_{\text{high}}}{2} > \theta_{\text{high}}.
\end{aligned}$$

□

E.19 Proof for Theorem 11

Proof. Define $\gamma_n := \frac{K_n}{n}$. The deterministic DKW reduction (149) gives

$$\left|\widehat{M}_{n,K_n} - U_{\gamma_n}(F)\right| \le \frac{\Delta_s}{\gamma_n} \sup_x |F_n(x) - F(x)|.$$

By the Glivenko–Cantelli theorem,

$$\sup_x |F_n(x) - F(x)| \longrightarrow 0 \qquad \text{almost surely.} \tag{151}$$

For completeness, this also follows from Lemma 22 and the Borel–Cantelli lemma. For every fixed $\epsilon > 0$,

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}\left(\sup_x |F_n(x) - F(x)| > \epsilon\right) &\le \sum_{n=1}^{\infty} 2e^{-2n\epsilon^2} \\
&< \infty.
\end{aligned}$$

Therefore, for every rational $\epsilon > 0$, the event occurs only finitely often almost surely, which implies (151). Since $\gamma_n \to \gamma > 0$, there exists n_0 such that $\gamma_n \geq \frac{\gamma}{2}$ for all $n \geq n_0$. Hence $\frac{1}{\gamma_n} \leq \frac{2}{\gamma}$ eventually. Combining this with (151),

$$\widehat{M}_{n,K_n} - U_{\gamma_n}(F) \longrightarrow 0 \qquad \text{almost surely.}$$

It remains to prove $U_{\gamma_n}(F) \longrightarrow U_\gamma(F)$. Using the quantile representation in Lemma 18,

$$U_\beta(F) = \frac{1}{\beta} \int_{1-\beta}^{1} Q(v)\, dv, \qquad \beta \in (0, 1].$$

Because $Q(v) \in [\ell, u]$, it is bounded. Therefore,

$$\left| \int_{1-\gamma_n}^{1} Q(v)\, dv - \int_{1-\gamma}^{1} Q(v)\, dv \right| \leq \max\{|\ell|, |u|\} |\gamma_n - \gamma| \longrightarrow 0.$$

Also, $\frac{1}{\gamma_n} \longrightarrow \frac{1}{\gamma}$. Hence $U_{\gamma_n}(F) \longrightarrow U_\gamma(F)$. Finally,

$$\left| \widehat{M}_{n,K_n} - U_\gamma(F) \right| \leq \left| \widehat{M}_{n,K_n} - U_{\gamma_n}(F) \right| + \left| U_{\gamma_n}(F) - U_\gamma(F) \right|.$$

Both terms converge to zero almost surely, proving the result. □