BEYOND GROUP SIZE: THE SPECTRAL GEOMETRY OF PLACKETT–LUCE SUPERVISION

Anonymous authors
Paper under double-blind review

ABSTRACT

LLM post-training increasingly uses ranked groups of candidate responses rather than isolated pairwise choices. But a complete ranking is not simply a collection of independent pairwise comparisons: under the Plackett–Luce model, candidates are selected sequentially, so a pair contributes information only while both members remain in the ranking.

We show that the Fisher information in a complete ranking has an exact graph representation with two distinct ingredients. The first is classical pairwise discriminability: candidates with similar scores reveal more about their score difference than candidates whose ordering is nearly certain. The second is listwise exposure: even when a pair's own scores stay fixed, its contribution changes with the other candidates because they determine how long the pair remains jointly active. We derive universal bounds on this exposure and use them to relate the exact Plackett–Luce Fisher operator to a much simpler reward-gap graph. The graph's connectivity then controls the relative-score direction that the ranking identifies least well, within a factor determined by group size.

At complete symmetry, this weakest-direction information increases with group size but quickly saturates, even while ranking entropy and total Fisher information continue to grow. More generally, group size and reward range do not determine the weakest direction. Across 9,696 exact instances, our spectral bounds hold without genuine violation. In a controlled example with the same group size and reward range, weakest-direction curvature changes by a factor of 8384 while total Fisher information moves in the opposite direction; nearly trace-matched configurations can still differ by hundreds-fold. These are local information-geometric results, not finite-sample or downstream-training guarantees.

1 INTRODUCTION

Modern preference-learning systems increasingly supervise models with ranked groups rather than isolated pairwise choices. Pairwise methods learn from a single binary preference (Rafailov et al., 2023; Azar et al., 2024; Ethayarajh et al., 2024); listwise methods consume an ordering over K candidates (Yuan et al., 2023; Zhao et al., 2023; Liu et al., 2024; Song et al., 2024); and grouped rollout methods score a batch of samples at once (Shao et al., 2024; Guo et al., 2025). We study the complete-ranking case that arises naturally in LLM post-training: a policy generates K candidate responses to the same prompt, an evaluator (for example a reward model, verifier, or human) scores or ranks them, and the resulting grouped feedback is used to update the policy. Our question is how much information one such ranked group carries about the underlying relative scores.

From scores to ranking geometry. Let $r = (r_1, \ldots, r_K)$ be the candidate scores; they determine a distribution over complete rankings. A score direction is identifiable to the extent that perturbing r along it changes that distribution, and Fisher information measures this local sensitivity. A ranking can therefore be informative in aggregate while leaving one relative-score contrast weakly constrained, and estimating that contrast then needs many more ranking observations.

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.

This matters because a complete ranking is not equivalent to observing all $\binom{K}{2}$ pairwise comparisons independently. Under the Plackett–Luce (PL) model (Luce, 1959; Plackett, 1975; Yellott, 1977), the ranking is generated sequentially: once a candidate is selected, it disappears from all later stages. Consequently, the information contributed by a pair has two conceptually different sources.

The first is classical. A pair is most informative when its two scores are close, since the comparison probability is then most sensitive to the gap. The unit here is the evaluator's resolution scale $\sigma > 0$: ranking probabilities depend on scores only through r/σ, so σ fixes how large a difference must be before the evaluator reliably resolves it. Once the gap is large relative to σ the comparison is nearly deterministic and says little about the gap's size.

The second is listwise. Fix a and b, so their gap $r_a - r_b$ and their direct pairwise Fisher information are fixed, and move only a third candidate c. This changes how likely c is to be selected before either of them, and hence how long a and b remain jointly active. Their contribution to the Fisher information of the *complete ranking* changes even though their own gap does not. This paper asks how the two effects combine across the full ranking.

Adding a constant to every score changes no ranking probability, so only relative scores are identifiable and the Fisher operator acts on $V_\perp = \{v \in \mathbb{R}^K : v^\top \mathbf{1} = 0\}$. Its smallest eigenvalue γ_{PL} is the relative-score contrast the supervision constrains least. Ranking entropy and Fisher trace do not determine it: entropy measures uncertainty in the observed ranking and trace measures total local sensitivity, and neither prevents the information from concentrating in some directions while one stays weak.

Main result. That Fisher operator has an exact representation as a weighted graph on the K candidates, in which the weight of edge (a, b) is a product of two positive numbers:

$$w_{ab}^{\mathrm{PL}} = \underbrace{\kappa_{\log}(r_a - r_b; \sigma)}_{\text{pairwise discriminability}} \times \underbrace{\alpha_{ab}(r, \sigma)}_{\text{listwise exposure}}$$

They differ in what each may depend on. The first is the classical Bradley–Terry Fisher kernel, a function of this pair's own gap $r_a - r_b$ alone. The second we introduce: the exposure coefficient α_{ab} is a function of the entire score vector r, and measures how much opportunity the sequential ranking process leaves the pair to contribute.

We prove the universal bounds $1/(K-1) \le \alpha_{ab}(r, \sigma) \le 1$, and hence the operator comparison

$$\frac{1}{K-1} L_{\mathrm{BT}}(r, \sigma) \preceq H_\sigma(r; r) \preceq L_{\mathrm{BT}}(r, \sigma)$$

where $H_\sigma(r; r)$ is the exact full-ranking PL Fisher operator and L_{BT} is the reward-gap Laplacian whose edge weights are $\kappa_{\log}(r_a - r_b; \sigma)$. Taking the smallest eigenvalue on $V_\perp$ turns this into the same two-sided bound on γ_{PL} in terms of $\lambda_2(L_{\mathrm{BT}})$. Thus the weakest identifiable direction of the full ranking is controlled, within a factor $K - 1$, by the algebraic connectivity of a much simpler reward-gap graph. A weak cut in this graph corresponds to a coordinated relative-score perturbation that changes the ranking distribution only slightly.

Symmetry as a calibration. Complete symmetry, $r_1 = \cdots = r_K$, removes score geometry and isolates the effect of group size K. The Fisher operator is then exactly computable (Theorem 3.1), and increasing K improves the weakest identifiable curvature only up to a finite limit, even while ranking entropy and total Fisher trace continue to grow. Symmetry is therefore a calibration rather than the general explanation: once the scores are nonsymmetric, the geometry of the group matters. Moreover, positive curvature from individual PL stages cannot simply be added, because every proper-subset stage is blind to some directions in the full relative-score space. This obstruction is why the operators must be aggregated before their weakest direction is assessed.

Why the weak direction matters. For LLM post-training, two groups with the same number of candidate rollouts can therefore provide very different supervision geometry even before the policy optimizer or model parameterization enters. Group size and reward range do not determine the least-constrained score direction, and our experiments show that even nearly matched total Fisher mass can hide large differences. The gap is paid in labels: once the weak direction is adequately sampled, estimation error along it tracks $1/(N\gamma_{\mathrm{PL}})$, so a curvature deficit becomes a proportional error deficit

at equal labelling budget (Section 6). We claim no nonasymptotic sample-complexity theorem or downstream language-model improvement, and the severely undersampled regime fails rather than degrading gracefully.

How we establish and validate the result. We first derive the exact Fisher contribution of each PL stage and show that every pair separates into two factors: its classical Bradley–Terry discriminability $\kappa_{\log}$ and its listwise exposure α_{ab}. A first-hit argument bounds the cumulative exposure of every pair, and these pairwise bounds combine to give the spectral sandwich. We then verify the theory computationally without reusing the same central calculation. For 9,696 score configurations, one implementation computes the PL Fisher operator directly by enumerating all rankings, while a second independently reconstructs it from the discriminability–exposure factorization. The two agree to numerical precision, and every instance satisfies the predicted operator bounds.

The novelty is not the Bradley–Terry Fisher kernel, stagewise covariance decomposition, or the general use of graph Laplacians in ranking; it is the exact discriminability–exposure representation of the full-ranking PL Fisher operator, the resulting uniform operator comparison, and its characterization of the weakest identifiable score direction. The theoretical claims remain local score-space Fisher results: we establish no general finite-sample rate and no downstream language-model guarantee (Section 8 states the limits in full).

2 Setup: Ranking Laws and Fisher Geometry

For one prompt with K candidates, let $r, u \in \mathbb{R}^K$ be target and model scores, let $\omega \in S_K$ be a complete ranking, and let $S_i(\omega) = \{\omega(i), \dots, \omega(K)\}$ be the remaining set at stage i. For evaluator resolution $\sigma > 0$ the Plackett–Luce law is

$$P_u^{(\sigma)}(\omega) = \prod_{i=1}^{K-1} \frac{\exp(u_{\omega(i)}/\sigma)}{\sum_{j \in S_i(\omega)} \exp(u_j/\sigma)}.$$

Adding a constant to all scores changes no ranking probability, so the identifiable relative-score space and its projector are

$$V_\perp = \{v \in \mathbb{R}^K : v^\top \mathbf{1} = 0\}, \qquad P_\perp = I_K - \tfrac{1}{K}\mathbf{1}\mathbf{1}^\top.$$

Writing $\mathcal{L}_\sigma(u; r) = \mathbb{E}_{\omega \sim P_r^{(\sigma)}}[-\log P_u^{(\sigma)}(\omega)]$ for the population loss, we have the exact identity $\mathcal{L}_\sigma(u; r) - \mathcal{L}_\sigma(r; r) = \mathrm{KL}(P_r^{(\sigma)} \| P_u^{(\sigma)})$, so at the truth

$$H_\sigma(r; r) = \nabla_u^2 \mathcal{L}_\sigma(u; r)\big|_{u=r}$$

is the Fisher information matrix, and the quantity this paper studies is its weakest identifiable eigenvalue

$$\gamma_{\mathrm{PL}}(r, \sigma) = \min_{\substack{v \in V_\perp \\ \|v\|_2 = 1}} v^\top H_\sigma(r; r) v.$$

We also write $\gamma_{\mathrm{PL}} = \lambda_2(H_\sigma)$, its smallest nonzero eigenvalue.

Throughout, "information" means Fisher information about relative-score parameters, especially this weakest identifiable eigenvalue — not Shannon entropy, mutual information under a prior, or Fisher trace, none of which we claim saturates with K.

Two graphs that must not be identified. Each PL stage contributes a categorical covariance $\Sigma(p) = \mathrm{diag}(p) - pp^\top$, which decomposes exactly into pairwise contrasts $\sum_{a<b \in S} p_a p_b B_{ab}$ with $B_{ab} := (e_a - e_b)(e_a - e_b)^\top$ for standard basis vectors e_a. Embedding stages into $\mathbb{R}^K$, summing, and averaging over rankings gives the *exact* PL Fisher Laplacian $L_{\mathrm{PL}}(u; r, \sigma)$ with

$$H_\sigma(u; r) = L_{\mathrm{PL}}(u; r, \sigma),$$

an identity, not an approximation (Appendix A). Separately, a single direct two-way logistic comparison with gap Δ has Fisher information

$$\kappa_{\log}(\Delta; \sigma) = \frac{1}{4\sigma^2} \operatorname{sech}^2\left(\frac{\Delta}{2\sigma}\right), \qquad L_{\mathrm{BT}}(r, \sigma) = \sum_{a<b} \kappa_{\log}(r_a - r_b; \sigma)\, B_{ab}.$$

These are different objects. L_{PL} is the exact PL Fisher/Hessian operator; L_{BT} is a simpler comparison graph whose weights depend only on target-score gaps relative to σ; those weights peak at $\Delta = 0$ and decay like $e^{-|\Delta|/\sigma}$, so nearly tied candidates form the most informative edges. The exact PL edge coefficient at a stage is $\sigma^{-2} p_a^{(i)} p_b^{(i)}$, *not* $\kappa_{\log}(r_a - r_b; \sigma)$. Their relation is the theorem of Section 5, not a definition.

3 Symmetry: Group Size Helps, but Only Boundedly

Theorem 3.1 (Exact symmetric PL curvature)**.** *For $K \geq 2$, $\sigma > 0$, and $r_1 = \cdots = r_K$,*

$$H_{\mathrm{sym},\sigma} = \frac{1}{\sigma^2} \frac{K - H_K}{K - 1} P_\perp, \qquad H_K = \sum_{m=1}^{K} \frac{1}{m}.$$

Consequently

$$\gamma_{\mathrm{sym}}(K, \sigma) = \frac{1}{\sigma^2} \frac{K - H_K}{K - 1}.$$

Hence $\gamma_{\mathrm{sym}}(2, \sigma) = 1/(2\sigma^2)$, and $\gamma_{\mathrm{sym}}(K, \sigma)$ is strictly increasing in K with

$$\gamma_{\mathrm{sym}}(K, \sigma) \uparrow \frac{1}{\sigma^2} \qquad \text{as } K \to \infty.$$

Larger symmetric lists therefore do increase the weakest identifiable Fisher curvature; the gain is bounded. Relative to pairwise supervision,

$$\frac{\gamma_{\mathrm{sym}}(K, \sigma)}{\gamma_{\mathrm{sym}}(2, \sigma)} = 2 \frac{K - H_K}{K - 1} < 2 \qquad \text{for every finite } K,$$

with the ratio tending to 2 only in the limit: per complete-ranking observation, the weakest-direction gain over one pairwise observation is less than a factor of two.

The trace behaves differently:

$$\mathrm{tr}(H_{\mathrm{sym},\sigma}) = \frac{K - H_K}{\sigma^2} \sim \frac{K}{\sigma^2}.$$

So total Fisher mass grows roughly linearly while the *weakest identifiable Fisher curvature* saturates. Ranking entropy is a third quantity again: at symmetry it equals $\log_2(K!)$ bits and grows without bound. There is no contradiction: entropy measures uncertainty of the observed ranking, γ_{sym} local distinguishability of nearby score parameters.

The proof averages an embedded subset projector over uniformly random remaining sets and sums a harmonic series; Appendix B gives it in full. Symmetry calibrates the phenomenon but does not explain nonsymmetric groups; the obvious extension, lower-bounding each stage separately, fails for the reason Section 4 gives.

4 Why Positive Stagewise Curvature Does Not Naively Add

At a PL stage with active set $S \subseteq [K]$, $|S| = n$, the categorical Fisher is strictly positive on the local tangent space

$$T_S = \{v \in \mathbb{R}^K : \mathrm{supp}(v) \subseteq S,\ \mathbf{1}_S^\top v = 0\}.$$

Yet proper-subset stages are degenerate on the global identifiable space. Write $\overline{\Sigma}_S(p)$ for the stage covariance $\Sigma(p)$ of Section 2 zero-padded outside S.

Theorem 4.1 (Proper-subset stage obstruction)**.** *Let $\varnothing \neq S \subsetneq [K]$, and let $p_j > 0$ on S. Define*

$$A_{S,\sigma}(p) = \frac{1}{\sigma^2} \overline{\Sigma}_S(p).$$

Then

$$\lambda_{\min}\Big(A_{S,\sigma}(p)\big|_{V_\perp} \Big) = 0,$$

and

$$\dim\left(\ker A_{S,\sigma}(p) \cap V_\perp\right) = K - |S|.$$

An explicit witness, with $\mathbf{1}_S$ the indicator vector of S, is

$$v_S^\star = \mathbf{1}_S - \frac{|S|}{|S^c|}\mathbf{1}_{S^c}.$$

It is globally mean-zero but constant on S, so

$$A_{S,\sigma}(p)v_S^\star = 0.$$

Each stage operator is strictly positive definite on its own tangent space T_{S_i}, which invites lower-bounding $\lambda_{\min}(A_{S_i,\sigma})$ stage by stage and adding the scalars via superadditivity. That step is valid only when every minimum is taken over the *same* space: the positive local minima live on the spaces T_{S_i}, whereas the quantity of interest is a minimum over $V_\perp$, where Theorem 4.1 gives exactly 0. This does not make later stages uninformative. Each nullspace moves with the realized remaining set, so differently oriented degenerate operators can sum to a well-conditioned global operator. The lesson is to aggregate operators first and take a minimum eigenvalue last, which is what Section 5 does. Appendix C expands both points.

5 Beyond Symmetry: Discriminability × Exposure

We now seek the aggregate representation the obstruction requires. Write $\Delta_{ab} = r_a - r_b$, and recall from Section 2 the direct pairwise kernel $\kappa_{\log}(\Delta_{ab};\sigma)$, which is fixed for the pair across all stages, together with the reward-gap Laplacian $L_{\mathrm{BT}}(r,\sigma) = \sum_{a<b}\kappa_{\log}(\Delta_{ab};\sigma)B_{ab}$. The results below establish the relation between L_{BT} and the exact operator $H_\sigma(r;r) = L_{\mathrm{PL}}(r;r,\sigma)$.

At a stage i where $a, b \in S_i$, let

$$p_a^{(i)} = \frac{e^{r_a/\sigma}}{\sum_{j\in S_i} e^{r_j/\sigma}}, \qquad p_b^{(i)} = \frac{e^{r_b/\sigma}}{\sum_{j\in S_i} e^{r_j/\sigma}},$$

and

$$q_{ab}^{(i)} = p_a^{(i)} + p_b^{(i)}.$$

Lemma 5.1 (Stagewise discriminability–exposure factorization)**.** *Whenever $a, b \in S_i$,*

$$\frac{1}{\sigma^2}p_a^{(i)}p_b^{(i)} = \left(q_{ab}^{(i)}\right)^2 \kappa_{\log}(\Delta_{ab};\sigma).$$

A pair contributes only while both of its members remain active. Define the stage hazard by the piecewise rule

$$\widetilde{q}_i^{ab} := \begin{cases} p_a^{(i)} + p_b^{(i)}, & a, b \in S_i, \\ 0, & \text{otherwise}, \end{cases} \qquad 0 \le \widetilde{q}_i^{ab} \le 1.$$

Setting it to zero after either member leaves means it never refers to an inactive candidate's probability. Conditional on the ranking history, $\widetilde{q}_i^{ab}$ is exactly the hazard that the first member of the pair is selected at stage i; Appendix D gives the filtration-level statement.

Definition 5.2 (Listwise exposure coefficient)**.**

$$\alpha_{ab}(r,\sigma) = \mathbb{E}\left[\sum_{i=1}^{K-1}(\widetilde{q}_i^{ab})^2\right].$$

The square is inherited from Lemma 5.1, where a stage contributes the pair's joint mass squared.

Theorem 5.3 (Exact exposure-weighted Fisher representation)**.** *For every finite $r \in \mathbb{R}^K$ and $\sigma > 0$,*

$$H_\sigma(r;r) = \sum_{a<b}\alpha_{ab}(r,\sigma)\kappa_{\log}(r_a - r_b;\sigma)(e_a - e_b)(e_a - e_b)^\top.$$

Thus

$$\boxed{\text{effective edge information} = \text{pair discriminability} \times \text{listwise exposure}.}$$

Let $\tau_{ab} = \min\{i : \omega(i) \in \{a, b\}\}$ be the stage at which the first member of the pair is selected. The events $\{\tau_{ab} = i\}$, $i = 1, \dots, K-1$, are disjoint and exhaustive, and summing the stagewise hazards gives

$$\mathbb{E}\left[\sum_{i=1}^{K-1} \widetilde{q}_i^{ab}\right] = 1.$$

The identity holds in expectation, not pathwise: $\sum_i \widetilde{q}_i^{ab}$ need *not* equal 1 on a realized ranking, and the lower bound below is proved so as not to require it.

Theorem 5.4 (Universal exposure bounds)**.** *For every pair, finite r, and $\sigma > 0$,*

$$\frac{1}{K-1} \le \alpha_{ab}(r, \sigma) \le 1.$$

Because each B_{ab} is positive semidefinite, these edgewise bounds lift to the reward-gap Laplacian $L_{\mathrm{BT}}(r, \sigma)$ of Section 2.

Theorem 5.5 (Spectral sandwich)**.** *For every finite $r \in \mathbb{R}^K$ and $\sigma > 0$,*

$$\frac{1}{K-1} L_{\mathrm{BT}}(r, \sigma) \preceq H_\sigma(r; r) \preceq L_{\mathrm{BT}}(r, \sigma).$$

Consequently,

$$\frac{1}{K-1} \lambda_2(L_{\mathrm{BT}}(r, \sigma)) \le \gamma_{\mathrm{PL}}(r, \sigma) \le \lambda_2(L_{\mathrm{BT}}(r, \sigma)).$$

Here $\preceq$ is the Loewner order ($B - A$ positive semidefinite), so the comparison holds in every direction at once, and λ_2 is the graph's algebraic connectivity. This is a factor-$(K-1)$ characterization of exact PL Fisher curvature by a graph whose weights depend only on reward gaps relative to evaluator resolution. Geometrically L_{BT} is heavy between candidates with close scores, so scores clumped into well-separated groups leave exponentially light edges between the groups and nearly cut the graph; the weakest direction is that between-group contrast.

Boundary checks. Three cases pin the normalization down. For $K = 2$ there is a single stage with $\widetilde{q}_1^{12} = 1$, so $\alpha_{12} = 1$ and

$$H_\sigma(r; r) = \kappa_{\log}(r_1 - r_2; \sigma) B_{12}$$

is exactly the Bradley–Terry Fisher matrix; the sandwich collapses to equality because $K - 1 = 1$. For $K = 3$ with first-stage pair mass $q = p_a^{(1)} + p_b^{(1)}$, either the pair is hit immediately or the third candidate goes first and the pair mass becomes 1, so

$$\alpha_{ab} = q^2 + (1 - q) = 1 - q + q^2 = \left(q - \tfrac{1}{2}\right)^2 + \tfrac{3}{4} \ge \tfrac{3}{4},$$

which is strictly stronger than the $1/2$ guaranteed by Theorem 5.4. Because q changes when the third candidate moves even though $r_a - r_b$ is fixed, this is the smallest explicit instance of pure listwise exposure; the universal constant is deliberately conservative, and whether $1/(K-1)$ can be improved is open. At symmetry, summing $\Pr(a, b \in S_n)\,(2/n)^2 = 4(n-1)/(nK(K-1))$ over $n = 2, \dots, K$ gives

$$\alpha_{ab}^{\mathrm{sym}} = \frac{4(K - H_K)}{K(K-1)},$$

so every effective edge has weight $(K - H_K)/(\sigma^2 K(K-1))$, and the complete-graph Laplacian's nonzero eigenvalue is K times that, recovering Theorem 3.1.

Corollary 5.6 (Group size and reward range do not determine curvature)**.** *Fix $K = 4$ and $R > 0$. Let*

$$r^{\mathrm{cluster}} = (-R, -R, R, R), \qquad r^{\mathrm{spread}} = \left(-R, -\frac{R}{3}, \frac{R}{3}, R\right).$$

Both have range $2R$, but

$$\frac{\gamma_{\mathrm{PL}}(r^{\mathrm{spread}}, \sigma)}{\gamma_{\mathrm{PL}}(r^{\mathrm{cluster}}, \sigma)} \ge \frac{2 - \sqrt{2}}{12} \frac{\operatorname{sech}^2(R/(3\sigma))}{\operatorname{sech}^2(R/\sigma)}.$$

Hence the ratio grows as $\Omega(\exp(4R/(3\sigma)))$ as $R/\sigma \to \infty$.

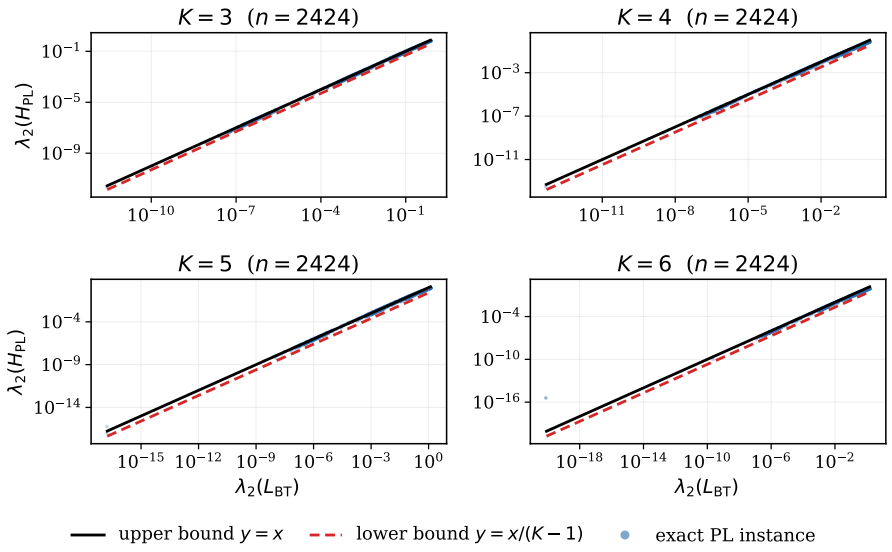

Figure 1: Exact PL weakest identifiable Fisher curvature against reward-gap graph connectivity, all 9,696 instances, one panel per K. Solid: upper bound $\lambda_2(L_{\mathrm{BT}})$; dashed: lower bound $\lambda_2(L_{\mathrm{BT}})/(K-1)$. Every instance lies inside the band of Theorem 5.5.

Corollary 5.7 (Greedy group selection)**.** *Fix K and σ and let $\mathcal{S}$ be a finite collection of candidate groups of size K. If S_E maximizes $\lambda_2(L_{\mathrm{BT}}(S,\sigma))$ over $\mathcal{S}$ and $S^\star$ maximizes $\gamma_{\mathrm{PL}}(S,\sigma)$ over $\mathcal{S}$, then*

$$\gamma_{\mathrm{PL}}(S_E,\sigma) \geq \tfrac{1}{K-1}\,\gamma_{\mathrm{PL}}(S^\star,\sigma).$$

This is immediate from Theorem 5.5, chaining its lower bound at S_E, the maximality of λ_2, and its upper bound at $S^\star$. It says that selecting groups on the computable $\lambda_2(L_{\mathrm{BT}})$ costs at most a factor $K-1$ against selecting on the exact curvature; designing and evaluating acquisition policies is outside the scope of this paper. The factor-$(K-1)$ guarantee is worst-case; we do not characterize its tightness over realistic candidate pools here.

6 What Simple Summaries Miss

The theory shows that group size and reward range do not determine the weakest identifiable Fisher curvature. Our controlled deformation makes the distinction quantitative: with both held fixed, γ_{PL} changes by $8384\times$ while total Fisher trace changes by about $35\times$ in the opposite direction. Practically, selection or sample-allocation rules based only on these summaries can therefore treat preference groups with substantially different information geometry as equivalent. We then ask whether the trace itself can nevertheless serve as a sufficient summary. We pre-specified 9,696 exact PL instances with $K \in \{3,4,5,6\}$, $\sigma = 1$, six score scales, and eight families, exhaustively enumerating all $K!$ rankings. Because $\kappa_{\log}$ depends on scores only through Δ/σ, varying score scale at fixed σ sweeps the gap-to-resolution regime rather than a single evaluator resolution. We independently validate the Fisher matrices by two code paths: embedded stage covariances and the survival-aware α_{ab} construction of Theorem 5.3. Their maximum relative Frobenius discrepancy is 1.03×10^{-14}, below the pre-specified 10^{-10} gate, with all ten blocking checks passing. Since $\rho = \lambda_2(H_\sigma)/\lambda_2(L_{\mathrm{BT}})$ is ill-conditioned near the noise floor, PSD comparisons are primary; Appendix F gives the full protocol and conditioning analysis.

6.1 Independent Verification of the Factorization and Spectral Sandwich

Across all instances both PSD inequalities held to roundoff (minimum eigenvalues -6.1×10^{-15} and -2.3×10^{-15}) (Figure 1), and the exposure bounds of Theorem 5.4 held for every pair. Thirteen unstable eigenvalue-ratio cases occurred only when both eigenvalues sat at the numerical noise floor; all satisfied the bounds at 50-digit precision (Appendix F).

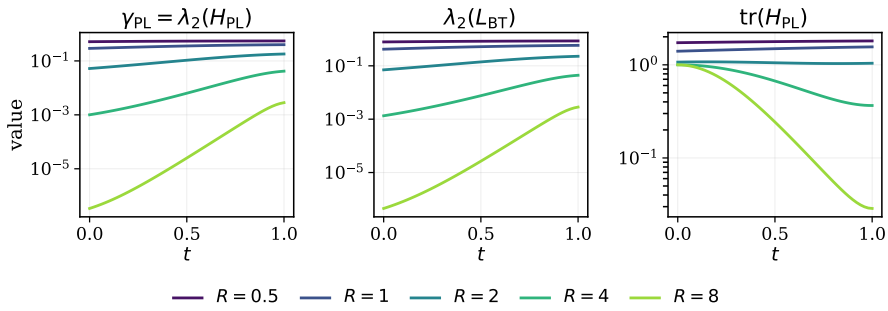

Figure 2: Controlled $K = 4$ deformation with range held at exactly $2R$, from clustered $(-R, -R, R, R)$ at $t = 0$ to spread $(-R, -R/3, R/3, R)$ at $t = 1$. Left: γ_{PL}; centre: $\lambda_2(L_{\mathrm{BT}})$; right: $\mathrm{tr}(H_\sigma)$; curves are the five values of R.

6.2 Fixed group size and range do not determine weakest identifiable Fisher curvature

To test Corollary 5.6 without confounds we interpolated between its two configurations at $K = 4$, verifying programmatically that the range equals $2R$ at every point to machine precision. The effect is large and grows with R/σ, as the theory predicts (full sweep in Appendix F, Table 2). At $R = 8$, weakest-direction curvature rises from $\gamma_{\mathrm{PL}} = 3.38 \times 10^{-7}$ at the clustered endpoint to 2.83×10^{-3} at the spread endpoint, a factor of 8384, while the Fisher trace *falls* by about $35\times$: total Fisher mass and weakest-direction curvature move in opposite directions along this controlled deformation, and neither K nor the range changes anywhere along it. That conditioning has a measurable label cost. In the adequately sampled regime $N\gamma_{\mathrm{PL}} \geq 10$, identified post hoc, weak-direction estimation error follows the expected $1/(N\gamma_{\mathrm{PL}})$ behaviour, with log-log slope -1.0045 ($R^2 = 0.996$, no free parameters). At $R = 2$, well inside this regime, the two configurations differ by $3.46\times$ in weakest-direction curvature and show a comparably sized difference in weak-direction error at fixed N. Below roughly ten effective observations in that direction the estimator does not degrade gracefully but fails to recover the contrast at all, so the frozen form of this hypothesis, which set no such regime restriction, is *not* supported as stated and the restriction is post hoc (Appendix K).

6.3 Nearly matched Fisher trace can hide large curvature differences

Section 6.2 holds K and the range fixed but lets the trace move freely — indeed opposite to γ_{PL}. This experiment is its adversarial extension, asking whether the separation survives once the trace is *also* approximately matched, which closes the escape route that total Fisher mass explains the effect. Theorem 5.5 says nothing about the trace, so this is empirical. We ran a pre-specified search over $7{,}424$ candidates per (K, range) cell, scanning exhaustively within each trace-tolerance band so the reported optimum is the bank optimum, not a hand-picked example.

At matched K and matched range, we find realizable score configurations whose Fisher traces agree to within 1% but whose weakest identifiable Fisher curvatures differ by hundreds-fold. At range 16 and 1% trace tolerance the best curvature ratios were $127.5\times$ at $K = 4$ (traces agreeing to 0.71%), $249.5\times$ at $K = 5$ (0.37%), and $478.3\times$ at $K = 6$ (0.68%). The tolerance ladder never had to be relaxed; the 1% band already yields these separations. Score vectors for all three pairs are in Appendix F, Table 1, together with the trace/curvature scatter (Figure 3); both members of every pair were reverified at 50 decimal digits.

This is an *empirical separation over the tested candidate families, not a theorem*: it shows such configurations exist and are easy to find, which rules out total Fisher mass as a sufficient summary, but establishes no general statement about the trace.

Correlation between $\lambda_2(H_\sigma)$ and $\lambda_2(L_{\mathrm{BT}})$ is not independent evidence here: Theorem 5.5 already confines the ratio to $[1/(K-1), 1]$, so a high correlation is forced and could not have come out otherwise. Measuring it checks the implementation and how tight the envelope is in practice, not whether $\lambda_2(L_{\mathrm{BT}})$ is predictive. Appendix F gives the descriptive comparison, the within-K breakdown, and the realized ratios.

7 Related Work

PL inference, comparison topology, and spectral structure. Bradley–Terry and Plackett–Luce are classical (Bradley & Terry, 1952; Luce, 1959; Plackett, 1975; Hunter, 2004; Yellott, 1977), and a literature ties ranking estimation to graph structure: rates governed by algebraic connectivity of item-pair co-occurrence (Vojnović et al., 2023), structure-dependent risk bounds for repeated selection (Seshadri et al., 2020), minimax rates whose Cramér–Rao bound depends inversely on the spectral gap of a comparison-graph Laplacian (Hajek et al., 2014), and unified asymptotics under deterministic comparison-graph conditions (Han & Xu, 2025). Our design is instead fixed to one complete ranking, whose internally generated Fisher graph we characterize.

Three Laplacians that must be kept distinct. *Comparison-design Laplacians* encode which items co-occur across observations, a property of the collection protocol. Our $L_{\mathrm{BT}}(r, \sigma)$ is not one: it is score-dependent, its weights the Fisher sensitivities of hypothetical direct Bradley–Terry comparisons. The third, $H_\sigma(r; r) = L_{\mathrm{PL}}(r; r, \sigma)$, is the exact Fisher operator of a complete PL ranking. Theorem 5.3 connects the latter two edge by edge, and Theorem 5.5 turns the exposure bounds into a two-sided operator comparison. We are not aware of prior work giving this exact discriminability–exposure factorization for the full-ranking score-space Fisher operator, or the resulting comparison with the reward-gap Laplacian.

Subset-wise feedback, design, and listwise objectives. That larger comparison sets need not give proportional gains is known: winner-only subset feedback barely improves on pairwise as the subset grows, while top-m feedback does (Saha & Gopalan, 2019a), and full ranking feedback can improve regret in online subset selection (Saha & Gopalan, 2019b). Those study repeated adaptive interaction; we ask how much local information about the relative-score vector is in *one* complete ranking. A second line asks which comparisons to acquire: D-optimal design for K-subsets (Thekumparampil et al., 2025), PL-specific Fisher objectives for negative selection (Surana et al., 2026), and classical optimal design (Pukelsheim, 2006). Our goal is prior to acquisition, so α_{ab} is not a sampling probability but the attenuation the sequential PL process applies to edge (a, b). Listwise and grouped alignment objectives (Yuan et al., 2023; Zhao et al., 2023; Liu et al., 2024; Song et al., 2024; Shao et al., 2024; Guo et al., 2025) alongside pairwise DPO, IPO and KTO (Rafailov et al., 2023; Azar et al., 2024; Ethayarajh et al., 2024) motivate the setting, and recent theory writes the Hessian of a log-linear PL listwise loss as stagewise conditional covariances (Wu et al., 2026). Beyond the priority disclaimers of Section 1, we do *not* claim the nullspace statement behind Theorem 4.1 is anything but elementary; the contribution is the exact factorization, the two-sided spectral comparison, and the structural role of the obstruction.

8 Discussion and Conclusion

A full ranking is governed not only by which pairs are discriminable, but by how the sequential process exposes those pairs and how their information connects across relative-score space. The exact PL Fisher edge weight factors as $w_{ab}^{\mathrm{PL}} = \alpha_{ab}\kappa_{\log}(\Delta_{ab}; \sigma)$, and universal exposure bounds place the operator between $L_{\mathrm{BT}}/(K-1)$ and L_{BT}: grouped preference information is geometric rather than merely combinatorial.

What this gives a practitioner. The reward-gap Laplacian is computable from target or estimated scores and the evaluator-resolution scale alone, so $\lambda_2(L_{\mathrm{BT}})$ is available *before* a ranking is collected. Selecting groups on it costs at most a factor $K-1$ against exact-curvature selection (Corollary 5.7); Section 6 shows that equal-size, equal-range groups can still differ sharply in weakest-direction curvature. That conditioning is paid in labels: well inside the adequately sampled regime, a $3.46\times$ curvature deficit shows up as a comparably sized error deficit at equal budget.

What this does not establish. All main theorems characterize the Fisher operator at the truth: no global quadratic growth, no nonasymptotic sample-complexity rate, no convergence rate for a nonlinear policy optimizer, and the experiments show no downstream language-model consequence and no evidence that $\lambda_2(L_{\mathrm{BT}})$ predicts curvature outside the sampled geometries; Appendix H's historical results validate mechanisms, not a state-of-the-art method. Conditioning may weaken under $J_\theta^\top H_\sigma J_\theta$; Corollary 5.7 motivates spectral group selection for LLM post-training.

Reproducibility Statement

Complete proofs are provided in the appendices. The supplementary material contains the frozen experimental specifications, exact-enumeration implementation, independent computational cross-checks, numerical validation gates, and scripts used to reproduce every table and figure generated by the spectral-geometry and label-cost experiments; the historical results in Appendix H are reused from an earlier submission and ship no artifacts. The experiment was frozen before result inspection, and numerically degenerate cases were additionally recomputed at high precision.

AI Use Statement

Generative AI tools were used as research assistants during the development of this work. They assisted with brainstorming and refining theoretical formulations, checking mathematical arguments and proofs, designing and implementing computational experiments, including code for generating synthetic score configurations and ranking observations, analyzing and stress-testing results, surveying and organizing related literature, and drafting and editing portions of the manuscript. All AI-assisted mathematical claims, code, and experimental results were independently checked against explicit derivations, validation gates, or frozen artifacts. The author takes responsibility for the final content of the work.

References

Mohammad Gheshlaghi Azar, Zhaohan Daniel Guo, Bilal Piot, Rémi Munos, Mark Rowland, Michal Valko, and Daniele Calandriello. A general theoretical paradigm to understand learning from human preferences. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238, pp. 4447–4455, 2024.

Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

Kawin Ethayarajh, Winnie Xu, Niklas Muennighoff, Dan Jurafsky, and Douwe Kiela. Model alignment as prospect theoretic optimization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 12634–12651, 2024.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Peiyi Wang, et al. DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. *Nature*, 645(8081):633–638, 2025. doi: 10.1038/s41586-025-09422-z.

Bruce Hajek, Sewoong Oh, and Jiaming Xu. Minimax-optimal inference from partial rankings. In *Advances in Neural Information Processing Systems*, volume 27, pp. 1475–1483, 2014. URL https://proceedings.neurips.cc/paper_files/paper/2014/hash/daadbd06d5082478b7677bea9812b575-Abstract.html.

Ruijian Han and Yiming Xu. A unified analysis of likelihood-based estimators in the Plackett–Luce model. *The Annals of Statistics*, 53(5):2077–2102, 2025. doi: 10.1214/25-AOS2530. URL https://doi.org/10.1214/25-AOS2530.

David R. Hunter. MM algorithms for generalized Bradley–Terry models. *The Annals of Statistics*, 32(1):384–406, 2004.

Tianqi Liu, Zhen Qin, Junru Wu, Jiaming Shen, Misha Khalman, Rishabh Joshi, Yao Zhao, Mohammad Saleh, Simon Baumgartner, Jialu Liu, Peter J. Liu, and Xuanhui Wang. LiPO: Listwise preference optimization through learning-to-rank. *arXiv preprint arXiv:2402.01878*, 2024.

R. Duncan Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, 1959.

Robin L. Plackett. The analysis of permutations. *Applied Statistics*, 24(2):193–202, 1975.

Friedrich Pukelsheim. *Optimal Design of Experiments*. SIAM Classics in Applied Mathematics, 2006.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D. Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, 2023.

Aadirupa Saha and Aditya Gopalan. Active ranking with subset-wise preferences. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 3312–3321. PMLR, 2019a. URL https://proceedings.mlr.press/v89/saha19a.html.

Aadirupa Saha and Aditya Gopalan. Combinatorial bandits with relative feedback. In *Advances in Neural Information Processing Systems*, volume 32, pp. 983–993, 2019b. URL https://proceedings.neurips.cc/paper/2019/hash/5e388103a391daabe3de1d76a6739ccd-Abstract.html.

Arjun Seshadri, Stephen Ragain, and Johan Ugander. Learning rich rankings. In *Advances in Neural Information Processing Systems*, volume 33, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/6affee954d76859baa2800e1c49e2c5d-Abstract.html.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. DeepSeekMath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Feifan Song, Bowen Yu, Minghao Li, Haiyang Yu, Fei Huang, Yongbin Li, and Houfeng Wang. Preference ranking optimization for human alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 18990–18998, 2024. doi: 10.1609/aaai.v38i17.29865.

Rohan Surana, Xintong Li, Sheldon Yu, Yiran Jenny Shen, Chuhan Wang, Tong Yu, Prithviraj Ammanabrolu, Jingbo Shang, Julian McAuley, and Junda Wu. MASS-DPO: Multi-negative active sample selection for direct policy optimization. *arXiv preprint arXiv:2605.10784*, 2026. doi: 10.48550/arXiv.2605.10784. URL https://arxiv.org/abs/2605.10784.

Kiran Koshy Thekumparampil, Gaurush Hiranandani, Kousha Kalantari, Shoham Sabach, and Branislav Kveton. Comparing few to rank many: Active human preference learning using randomized frank-Wolfe method. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 59355–59376. PMLR, 2025. URL https://proceedings.mlr.press/v267/thekumparampil25a.html.

Milan Vojnović, Se-Young Yun, and Kaifang Zhou. Accelerated MM algorithms for inference of ranking scores from comparison data. *Operations Research*, 71(4):1318–1342, 2023. doi: 10.1287/opre.2022.2264. URL https://doi.org/10.1287/opre.2022.2264.

Xudong Wu, Jian Qian, Pangpang Liu, Vaneet Aggarwal, and Jiayu Chen. Distributionally robust listwise preference optimization. *arXiv preprint arXiv:2607.01715*, 2026. doi: 10.48550/arXiv.2607.01715. URL https://arxiv.org/abs/2607.01715.

John I. Yellott. The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, 15(2): 109–144, 1977.

Hongyi Yuan, Zheng Yuan, Chuanqi Tan, Wei Wang, Songfang Huang, and Fei Huang. RRHF: Rank responses to align language models with human feedback. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, 2023.

Yao Zhao, Rishabh Joshi, Tianqi Liu, Misha Khalman, Mohammad Saleh, and Peter J. Liu. SLiC-HF: Sequence likelihood calibration with human feedback. *arXiv preprint arXiv:2305.10425*, 2023.

A Setup Derivations

This appendix records the derivations compressed out of Section 2. Nothing here is new; it is the explicit form of the identities quoted there.

Pairwise decomposition of a categorical Fisher. For a categorical probability vector p supported on an active set S,

$$\Sigma(p) = \mathrm{diag}(p) - pp^\top = \sum_{a<b\in S} p_a p_b (e_a - e_b)(e_a - e_b)^\top.$$

Thus a PL stage is already a weighted graph of pairwise contrast directions, with edge weight $p_a p_b$ on the pair (a, b).

The exact PL Fisher operator is a graph Laplacian. Embedding each stage covariance into the original K coordinates, summing stages, and averaging over rankings defines

$$L_{\mathrm{PL}}(u; r, \sigma) := \frac{1}{\sigma^2} \sum_{i=1}^{K-1} \mathbb{E}_{\omega\sim P_r^{(\sigma)}} \left[\sum_{a<b\in S_i(\omega)} p_a^{(i)}(u,\omega)\, p_b^{(i)}(u,\omega)\, B_{ab} \right], \qquad B_{ab} := (e_a - e_b)(e_a - e_b)^\top,$$

and by construction $H_\sigma(u; r) = L_{\mathrm{PL}}(u; r, \sigma)$. This is an identity rather than an approximation. Its edge weights depend on the stage, the realized active set, and the full PL probabilities, which is precisely why the exact operator is not the reward-gap graph L_{BT}.

Equivalent form of the pairwise kernel. The direct two-way logistic Fisher kernel admits the equivalent exponential form

$$\kappa_{\log}(\Delta; \sigma) = \frac{1}{4\sigma^2} \operatorname{sech}^2\left(\frac{\Delta}{2\sigma}\right) = \frac{1}{\sigma^2} \frac{e^{r_a/\sigma} e^{r_b/\sigma}}{\left(e^{r_a/\sigma} + e^{r_b/\sigma}\right)^2}, \qquad \Delta = r_a - r_b,$$

which is the form used in the numerically stable implementation of Section 6.

B Proof of the Symmetric Curvature Theorem

For $S \subseteq [K]$ with $|S| = n$, define

$$P_{S,\perp}^{\mathrm{emb}} = D_S - \frac{1}{n} \mathbf{1}_S \mathbf{1}_S^\top.$$

At symmetry the stage Fisher for remaining set S is

$$\frac{1}{\sigma^2 n} P_{S,\perp}^{\mathrm{emb}}.$$

Proof mechanics in brief. At a stage with n remaining candidates, the local symmetric Fisher is $\sigma^{-2} n^{-1}$ times the local mean-zero projector. Averaging that projector over all uniformly random size-n remaining sets yields

$$\mathbb{E}[P_{S,\perp}^{\mathrm{emb}}] = \frac{n-1}{K-1} P_\perp,$$

and summing over $n = K, \ldots, 2$ gives

$$\sum_{n=2}^{K} \frac{n-1}{n} = K - H_K.$$

Lemma B.1 (Uniform remaining set)**.** *At the stage with n remaining candidates, S is uniform over the $\binom{K}{n}$ size-n subsets.*

Proof. Fix S with $|S| = n$. The $K - n$ removed items can occupy the first $K - n$ positions in $(K - n)!$ orders and the items in S can occupy the last n positions in $n!$ orders. Thus the event has probability $(K - n)!n!/K! = \binom{K}{n}^{-1}$. □

Lemma B.2 (Average subset projector)**.**

$$\mathbb{E}_S[P_{S,\perp}^{\mathrm{emb}}] = \frac{n-1}{K-1} P_\perp.$$

Proof. Permutation invariance implies

$$M_n := \mathbb{E}[P^{\mathrm{emb}}_{S,\perp}] = aI + b\mathbf{1}\mathbf{1}^\top.$$

Every subset projector annihilates $\mathbf{1}$, so $M_n = aP_\perp$. Taking traces,

$$n - 1 = \mathrm{tr}(M_n) = a(K - 1),$$

which gives $a = (n-1)/(K-1)$. □

Proof of Theorem 3.1. Summing expected stage contributions over $n = K, \ldots, 2$,

$$H_{\mathrm{sym},\sigma} = \frac{1}{\sigma^2(K-1)} \sum_{n=2}^{K} \frac{n-1}{n} P_\perp.$$

Since

$$\sum_{n=2}^{K} \frac{n-1}{n} = K - H_K,$$

the result follows. □

C Proof of the Stagewise Obstruction

Proof of Theorem 4.1. Let $n = |S|$ and $m = |S^c| = K - n$. Define

$$v^\star_S = \mathbf{1}_S - \frac{n}{m}\mathbf{1}_{S^c}.$$

Then $(v^\star_S)^\top \mathbf{1} = 0$, so $v^\star_S \in V_\perp$. On S, the vector is constant. Because

$$\Sigma_S(p)\mathbf{1}_S = 0,$$

the embedded Fisher satisfies

$$A_{S,\sigma}(p)v^\star_S = 0.$$

Since $A_{S,\sigma}(p)$ is PSD, its restricted minimum Rayleigh quotient on $V_\perp$ is zero.

For the nullity statement, positivity of every p_j implies the local categorical Fisher has nullspace exactly the constant direction on S. After embedding, coordinates on S^c are arbitrary, so

$$\dim\ker A_{S,\sigma} = 1 + (K - n).$$

Intersecting with $v^\top \mathbf{1} = 0$ leaves dimension $K - n$. □

The invalid argument this rules out. Each stage operator is strictly positive definite on its own tangent space T_{S_i}, which invites the following reasoning: lower-bound $\lambda_{\min}(A_{S_i,\sigma})$ stage by stage, then add the scalars using superadditivity,

$$\lambda_{\min}\left(\sum_i A_{S_i,\sigma}\right) \geq \sum_i \lambda_{\min}(A_{S_i,\sigma}).$$

The superadditivity step is valid only when every minimum is taken over the same space. Here the positive local minima live on the spaces T_{S_i}, whereas the quantity of interest is a minimum over $V_\perp$, and Theorem 4.1 shows that each proper-subset stage contributes exactly 0 there. The scalars being added are curvatures of different spaces, so the conclusion does not follow.

What remains true. The theorem does not say that later stages are uninformative. Each proper-subset stage is positive on its own local tangent space and globally degenerate on $V_\perp$, and its nullspace moves with the realized remaining set. Sums and expectations of differently oriented degenerate operators can therefore recover positive global curvature. The lesson is to aggregate operators first and take a minimum eigenvalue last, which is what Section 5 does.

D Proofs for Exposure Factorization and Spectral Sandwich

Proof of Lemma 5.1. When $a, b \in S_i$,

$$\frac{p_a^{(i)}}{p_a^{(i)} + p_b^{(i)}} = \frac{e^{r_a/\sigma}}{e^{r_a/\sigma} + e^{r_b/\sigma}}.$$

Therefore

$$\frac{1}{\sigma^2} p_a^{(i)} p_b^{(i)} = \left(p_a^{(i)} + p_b^{(i)}\right)^2 \kappa_{\log}(r_a - r_b; \sigma).$$

□

Proof of Theorem 5.3. For one fixed pair,

$$w_{ab}^{\mathrm{PL}} = \mathbb{E}\left[\sum_i X_i^{ab} \frac{1}{\sigma^2} p_a^{(i)} p_b^{(i)}\right].$$

Applying Lemma 5.1,

$$w_{ab}^{\mathrm{PL}} = \mathbb{E}\left[\sum_i (\widetilde{q}_i^{ab})^2\right] \kappa_{\log}(\Delta_{ab}; \sigma) = \alpha_{ab} \kappa_{\log}(\Delta_{ab}; \sigma).$$

Summing all pair edges gives the theorem. □

Lemma D.1 (First-hit compensator identity)**.** *For every pair,*

$$\mathbb{E}\left[\sum_{i=1}^{K-1} \widetilde{q}_i^{ab}\right] = 1.$$

Proof. Let $\mathcal{F}_i = \sigma(\omega(1), \ldots, \omega(i-1))$ and let $\tau_{ab} = \min\{i : \omega(i) \in \{a, b\}\}$. The event that both a and b are still active at stage i is $\mathcal{F}_i$-measurable, and by the piecewise definition $\widetilde{q}_i^{ab}$ is $\mathcal{F}_i$-measurable and never evaluates a probability of an inactive candidate. Conditioned on $\mathcal{F}_i$, stage i selects the first member of the pair with probability exactly $\widetilde{q}_i^{ab}$, so

$$\Pr(\tau_{ab} = i \mid \mathcal{F}_i) = \widetilde{q}_i^{ab}.$$

Taking total expectation gives the pointwise-in-i identity $\mathbb{E}[\widetilde{q}_i^{ab}] = \Pr(\tau_{ab} = i)$. The events $\{\tau_{ab} = i\}$ for $i = 1, \ldots, K-1$ are disjoint and exhaustive: a complete ranking eventually selects one of a, b, and if the two are the final survivors this happens at stage $K-1$. Summing over i gives the claim. □

Remark D.2. The stronger pathwise statement $\sum_i \widetilde{q}_i^{ab} = 1$ is false in general and is not used. Only the expectation is pinned to 1, which is why the lower bound below applies Cauchy–Schwarz pathwise and passes to expectations afterwards.

Proof of Theorem 5.4. Upper bound. Since $0 \le \widetilde{q}_i^{ab} \le 1$ we have $(\widetilde{q}_i^{ab})^2 \le \widetilde{q}_i^{ab}$ pathwise, so by Lemma D.1

$$\alpha_{ab} = \mathbb{E}\sum_i (\widetilde{q}_i^{ab})^2 \le \mathbb{E}\sum_i \widetilde{q}_i^{ab} = 1.$$

Lower bound. There are at most $K-1$ nontrivial stages, so on every realized ranking path Cauchy–Schwarz gives

$$\sum_{i=1}^{K-1} (\widetilde{q}_i^{ab})^2 \ge \frac{1}{K-1}\left(\sum_{i=1}^{K-1} \widetilde{q}_i^{ab}\right)^2.$$

Taking expectations and then applying Jensen's inequality to $Z = \sum_i \widetilde{q}_i^{ab}$,

$$\alpha_{ab} \ge \frac{1}{K-1}\mathbb{E}[Z^2] \ge \frac{1}{K-1}(\mathbb{E}Z)^2 = \frac{1}{K-1},$$

the last equality by Lemma D.1. Note the order: the pathwise inequality comes first and the expectation identity is applied only at the end, so no pathwise normalization of $\sum_i \widetilde{q}_i^{ab}$ is assumed. □

Proof of Theorem 5.5. Every edge matrix $B_{ab} = (e_a - e_b)(e_a - e_b)^\top$ is positive semidefinite, so Theorem 5.4 gives the edgewise Loewner inequalities

$$\frac{1}{K-1}\kappa_{\log}(\Delta_{ab};\sigma)B_{ab} \preceq \alpha_{ab}\kappa_{\log}(\Delta_{ab};\sigma)B_{ab} \preceq \kappa_{\log}(\Delta_{ab};\sigma)B_{ab}.$$

Summing over pairs preserves the Loewner order, and Theorem 5.3 identifies the middle sum as $H_\sigma(r;r)$, which gives the operator sandwich. All three operators are graph Laplacians and therefore annihilate $\mathbf{1}$, so restricting the common Rayleigh quotient to

$$V_\perp = \{v : v^\top \mathbf{1} = 0\}$$

preserves the order of the minima. The minimum of the Rayleigh quotient of a Laplacian over unit vectors in $V_\perp$ is its second-smallest eigenvalue, which yields

$$\frac{1}{K-1}\lambda_2(L_{\mathrm{BT}}) \le \gamma_{\mathrm{PL}}(r,\sigma) \le \lambda_2(L_{\mathrm{BT}}).$$

No commuting or shared eigenvectors are required. □

E Proof of the Fixed-Range Separation

For

$$r^{\mathrm{cluster}} = (-R, -R, R, R),$$

writing $w(\Delta) := \kappa_{\log}(\Delta;\sigma)$ for the kernel of Section 2, the two within-cluster BT edges have weight $w(0)$ and all four cross-cluster edges have weight

$$w(2R).$$

The nonzero eigenvalues of $L_{\mathrm{BT}}(r^{\mathrm{cluster}},\sigma)$ are

$$4w(2R), \qquad 2\big(w(0)+w(2R)\big), \qquad 2\big(w(0)+w(2R)\big),$$

the first belonging to the cluster-contrast vector $(1,1,-1,-1)^\top$ and the other two to the within-cluster contrasts. Since $w(0) \ge w(2R)$, the smallest of these is the first. Hence

$$\lambda_2(L_{\mathrm{BT}}(r^{\mathrm{cluster}},\sigma)) = 4w(2R),$$

so

$$\gamma_{\mathrm{PL}}(r^{\mathrm{cluster}},\sigma) \le 4w(2R).$$

For

$$r^{\mathrm{spread}} = \left(-R, -\frac{R}{3}, \frac{R}{3}, R\right),$$

the three adjacent score gaps all equal $2R/3$, so the BT graph contains, as a subgraph, the four-vertex path $1-2-3-4$ with all three edge weights $w(2R/3)$. Deleting the remaining edges removes positive semidefinite terms and therefore only decreases the Laplacian in Loewner order. The algebraic connectivity of the unweighted four-vertex path is $2-\sqrt{2}$, so

$$\lambda_2(L_{\mathrm{BT}}(r^{\mathrm{spread}},\sigma)) \ge (2-\sqrt{2})w(2R/3).$$

Since $K-1=3$,

$$\gamma_{\mathrm{PL}}(r^{\mathrm{spread}},\sigma) \ge \frac{2-\sqrt{2}}{3}w(2R/3).$$

Combining the lower bound for the spread configuration with the upper bound for the clustered one—directionally valid because all curvatures here are positive—yields the ratio of Corollary 5.6. Using $\operatorname{sech}^2 x \sim 4e^{-2x}$,

$$\frac{\operatorname{sech}^2(R/(3\sigma))}{\operatorname{sech}^2(R/\sigma)} \sim \exp\left(\frac{4R}{3\sigma}\right),$$

which gives the stated exponential growth. The two configurations share K and reward range but not total Fisher trace; a trace-matched separation is not claimed here and remains open.

F SPECTRAL EXPERIMENT: PROTOCOL, VALIDATION, AND ADVERSARIAL SEARCH

Pre-specification. The protocol was frozen in a specification committed to version control before any result was generated. It was not lodged in an external registry, so we describe the experiment as pre-specified rather than preregistered; the commit history is included in the supplement and fixes the ordering. The specification fixed the hypotheses, score families, sample counts, master seed, tolerances, figures, and interpretation rules, and was not edited afterwards. Three deviations occurred and are recorded with the experiment: a dependency version file was corrected to match the versions actually captured at runtime; one companion figure was added showing the same data with no subsetting; and a precision-stratification flag was reported alongside, never instead of, the full statistics. No hypothesis, family, sample count, tolerance, or interpretation rule was changed.

Instances. Experiment A covers $K \in \{3, 4, 5, 6\}$, $\sigma = 1$, scales $s \in \{0.25, 0.5, 1, 2, 4, 8\}$, and eight families: Gaussian, uniform, evenly spaced, two-cluster, jittered two-cluster, one dominant item, one suppressed item, and a random clustered mixture, with 100 vectors per stochastic cell and master seed 20260917, giving 9,696 instances. Experiment B is the controlled $K = 4$ path at $R \in \{0.5, 1, 2, 4, 8\}$ with $t = 0, 0.02, \ldots, 1$ (255 points). Experiment C searches $K \in \{4, 5, 6\}$ at target ranges $2R \in \{1, 2, 4, 8, 16\}$.

Two independent implementations. The two code paths never call each other. The direct route computes ranking probabilities in log-sum form and forms embedded categorical covariances; the factorized route computes ranking probabilities as a forward product of stage selection probabilities, never forms a covariance matrix, and reaches the operator only through the exposure identity of Theorem 5.3. Their agreement is therefore evidence about the identity rather than a shared assumption.

Blocking validation. All ten gates passed before any scientific run: ranking probabilities sum to one (3.2×10^{-15}); H_σ symmetric (exactly); H_σ positive semidefinite (-5.2×10^{-15}); $H_\sigma \mathbf{1} = 0$ (7.4×10^{-15}); direct equals factorized (7.5×10^{-15}, gate 10^{-10}); lower sandwich (-5.1×10^{-15}); upper sandwich (-3.3×10^{-16}); $K = 2$ Bradley–Terry equality (1.2×10^{-16}); the symmetric closed form of Theorem 3.1 (7.6×10^{-15}); and the $K = 3$ exposure identity $\alpha_{ab} = 1 - q + q^2$ with $\alpha_{ab} \geq 3/4$ (6.7×10^{-16}, minimum 0.750141). Over the full sweep the direct-versus-factorized relative Frobenius error never exceeded 1.03×10^{-14}.

Numerical conditioning of ρ. Thirteen of the 9,696 instances produced ρ outside $[1/(K-1), 1]$ at double precision. Every one occurs at large score separation where both λ_2 values fall between 10^{-8} and 10^{-21}, i.e. at or below the double-precision eigenvalue noise floor, so ρ is a quotient of two unresolvable quantities. All thirteen were recomputed at 50 decimal digits and all lie inside the envelope. The operator form of the theorem, which is far better conditioned, held at double precision across every instance. We therefore report the positive semidefinite comparisons as the primary validation object and treat ρ as a diagnostic only.

Adversarial trace-matched search. For each (K, range) cell the candidate bank held 7,424 vectors: the 2,424 Experiment A vectors for that K renormalized to the target range, plus 5,000 fresh vectors drawn from Gaussian, uniform, two-cluster, three-cluster, and randomly spaced ordered families. Pairs were compared by relative trace difference and curvature ratio, scanning exhaustively within each tolerance band, so the reported value is the bank optimum rather than a selected example. Tolerances were applied in the frozen order 1%, 2%, 5%; the 1% band already produced the separations reported in the main text and the ladder was never relaxed further. Table 1 gives the strongest pair at each K.

Controlled path, full sweep. Table 2 gives every R value of the controlled deformation.

Realized spectral ratios for the trace-matched pairs. For the pairs of Table 1, the ratio of $\lambda_2(L_{\mathrm{BT}})$ between the two members tracks the ratio of exact curvature but is not equal to it, which is expected: Theorem 5.5 constrains each member only to within a factor $K - 1$, so the two ratios may differ by up to that factor squared in the worst case. Observed differences are far smaller.

		Score vector	$\mathrm{tr}(H_\sigma)$	γ_{PL}	trace gap	γ_{PL} ratio
$K = 4$	A	$(-0.15692, -7.9407, 0.038323, 8.0593)$	0.4972135	5.84×10^{-4}	0.71%	127.5×
	B	$(-6.2458, -7.9936, 6.2331, 8.0063)$	0.5007810	4.58×10^{-6}		
$K = 5$	A	$(0.083131, -0.124, -0.00023028, 8.0206, -7.9794)$	1.1639773	7.32×10^{-4}	0.37%	249.5×
	B	$(8.2913, -6.574, -6.5589, -4.5845, 9.426)$	1.1682513	2.93×10^{-6}		
$K = 6$	A	$(0.17856, 0.072203, -7.8465, -0.20526, 8.1535, -0.3525)$	1.8811721	9.83×10^{-4}	0.68%	478.3×
	B	$(-7.8589, -8.1204, -6.4532, 6.7232, 7.8796, 7.8297)$	1.8939865	2.06×10^{-6}		

Table 1: Strongest trace-matched pairs at 1% trace tolerance and range 16, from the frozen adversarial search. Both members of each pair were reverified at 50 decimal digits. This is an empirical separation over the tested candidate families, not a theorem.

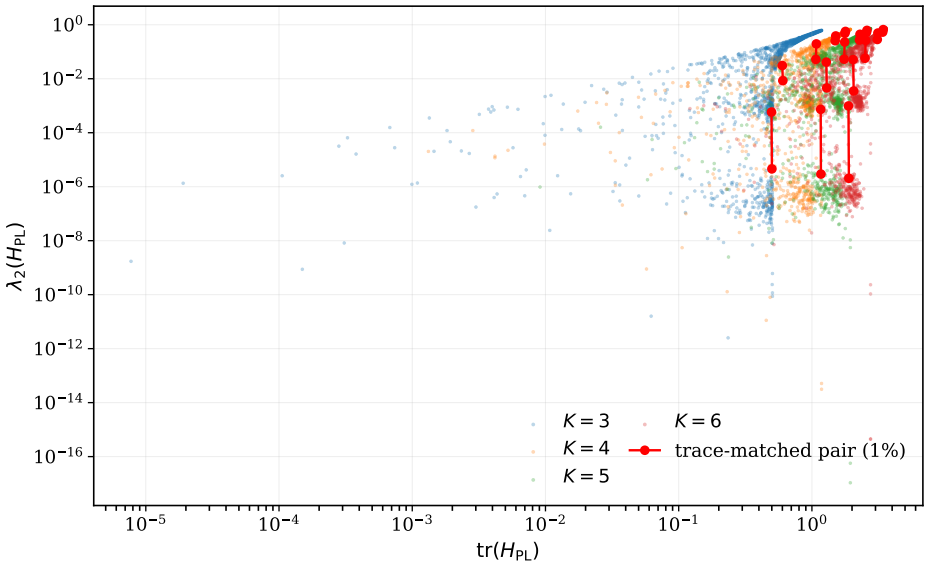

Figure 3: Total Fisher mass against weakest identifiable direction over all exact instances. At a fixed trace the attainable $\lambda_2(H_\sigma)$ spans many orders of magnitude. Red segments join the trace-matched pairs located by the frozen adversarial search at 1% trace tolerance.

Why the pooled correlation is a consistency check. Theorem 5.5 confines $\lambda_2(H_\sigma)$ to $[\lambda_2(L_{\mathrm{BT}})/(K-1), \lambda_2(L_{\mathrm{BT}})]$. Over instances whose curvature spans many orders of magnitude, this forces a high rank and log-space correlation between the two quantities regardless of the score geometry, so a large R^2 cannot discriminate between hypotheses and is not independent empirical support for $\lambda_2(L_{\mathrm{BT}})$ as a predictor. What it does establish is that the implementation respects the envelope and that the envelope is tight in practice. The comparisons that carry information are with the summaries the theorem leaves unconstrained — group size, reward range, ranking entropy, and total Fisher mass — and with the controlled and adversarial experiments of Section 6, where geometry is varied deliberately rather than sampled.

Predictor comparison. Table 4 gives the full descriptive comparison. Pooled correlations mix different K and are reported for completeness only; within-K values are the informative ones, and the controlled and adversarial experiments of Section 6 are the substantive tests.

G Realized Exposure and Spectral Ratio

Every pair in every instance satisfied the exposure bounds of Theorem 5.4, with no violation. The smallest numerically resolvable values of ρ were 0.750000 at $K = 3$, 0.610182 at $K = 4$, 0.517919 at $K = 5$, and 0.451452 at $K = 6$, against the theoretical floors 0.5, $1/3$, 0.25, and 0.2. The $K = 3$ minimum coincides with the exact bound $\alpha_{ab} \geq 3/4$ derived in Section 5.

R	γ_{PL} (clustered)	γ_{PL} (spread)	γ_{PL} ratio	trace ratio
0.5	5.16×10^{-1}	5.51×10^{-1}	$1.07\times$	$1.04\times$
1	2.91×10^{-1}	4.00×10^{-1}	$1.38\times$	$1.11\times$
2	5.22×10^{-2}	1.81×10^{-1}	$3.46\times$	$0.97\times$
4	1.01×10^{-3}	4.17×10^{-2}	$41.4\times$	$0.37\times$
8	3.38×10^{-7}	2.83×10^{-3}	$8384\times$	$0.029\times$

Table 2: Controlled $K = 4$, fixed-range path of Section 6.2, full R sweep. Group size and reward range are identical at both endpoints of every row; the range equals $2R$ at every point of the path to machine precision.

K	exact γ_{PL} ratio	$\lambda_2(L_{\mathrm{BT}})$ ratio	relative difference
4	$127.5\times$	$126.3\times$	1.0%
5	$249.5\times$	$280.2\times$	12.3%
6	$478.3\times$	$414.2\times$	13.4%

Table 3: Exact versus reward-gap-graph curvature ratios for the strongest trace-matched pair at each K (range 16, 1% trace tolerance).

The only conclusion we draw is that the universal lower bound is *conservative in the tested families*. These observations do not establish that a better universal constant exists, that $1/(K-1)$ is or is not optimal, or that symmetry is extremal in either direction; none of those questions is settled by this experiment, and Section 5 leaves the sharpness question open.

Figure 5 plots the same instances as Figure 1, none removed from the analysis, in a form where the width of the band is legible.

H Historical Supporting Experiments

The results in this appendix predate the spectral experiment of Section 6 and are retained as supporting evidence. They are not the primary empirical support for any claim in this paper, and none of them is used to support a nonsymmetric curvature statement.

H.1 Symmetric validation and the stagewise obstruction

An earlier synthetic PL teacher experiment computed the Hessian exactly for $K \leq 8$ and by Monte Carlo for larger K. The observed nonzero symmetric eigenvalue matched

$$\sigma^{-2} \frac{K - H_K}{K - 1}$$

to numerical precision in the exact regime and followed the predicted saturating curve under Monte Carlo estimation. This is superseded, for the symmetric case, by validation gate nine of the frozen experiment (Appendix F), which reproduces the same closed form to 7.6×10^{-15}.

The same environment directly visualized Theorem 4.1. For the explicit witness $v_S^\star$, visibility on the chosen proper subset was numerically zero, while random subsets of the same size generally retained nonzero projection energy.

H.2 Controlled evaluator-resolution sweep

An earlier synthetic preference-bandit experiment used a controlled noise variable δ with

$$\sigma = \sqrt{1 + \delta^2}.$$

This mapping is an experimental mechanism, not a theoretical assumption.

Predictor	Spearman	R^2 (log)	within-K R^2	Note
$\lambda_2(L_{\mathrm{BT}})$	0.9871	0.9970	0.9941–0.9996	theoretically justified comparison object
reward range	−0.9652	0.6370	0.6329–0.6544	monotone but not determining
ranking entropy (bits)	0.6361	0.2307	0.3902–0.4392	grows with $K!$, not with γ_{PL}
$\mathrm{tr}(H_\sigma)$	0.6322	0.2038	0.3228–0.3639	Spearman 0.94–0.96 within K
K	0.1008	0.0007	—	pooled only; K is the stratifier

Table 4: Descriptive predictor comparison for $\log \gamma_{\mathrm{PL}}$ over the 9,696 exact instances. Pooled correlations mix different K and are reported for completeness only; the controlled and adversarial experiments above are the substantive tests.

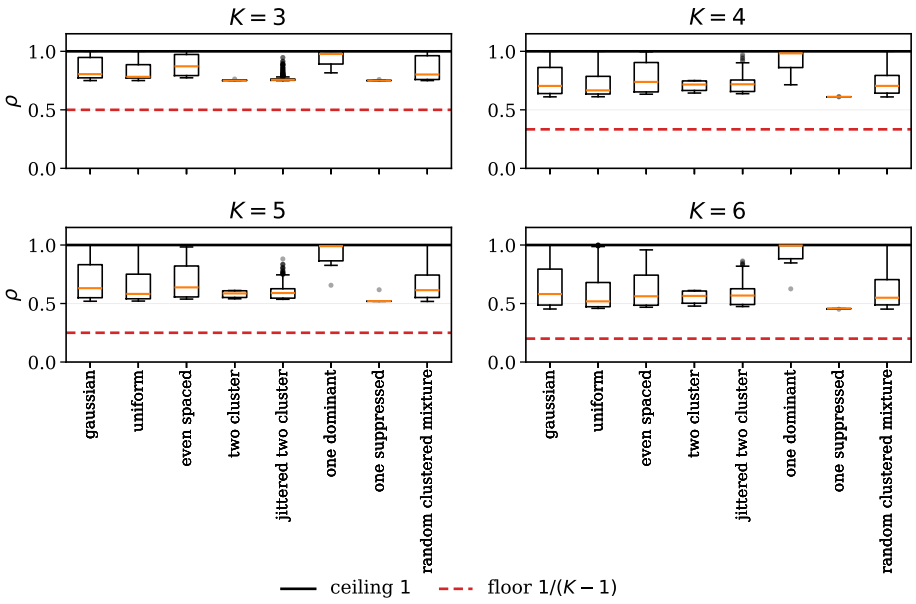

Figure 4: Realized $\rho = \lambda_2(H_\sigma)/\lambda_2(L_{\mathrm{BT}})$ by K and score family across the 9,696 exact instances, with the theoretical floor $1/(K-1)$ and ceiling 1. Thirteen instances for which ρ is a ratio of two eigenvalues at the float64 numerical noise floor fall outside the displayed range; all were recomputed at 50-digit precision and produced no genuine violation of the spectral bounds. They remain in the experiment and in every numerical summary.

H.3 Small real-model anchor

An earlier study ran Pythia-410M-deduped with $K = 4$, three seeds per cell, and contamination levels $\rho \in \{0, 0.5\}$. The utility-scaled information-aware variant was directionally stronger on mean test reward in both conditions, but paired-bootstrap confidence intervals crossed zero at three seeds. This anchor is directional only and carries no headline claim in this paper; in particular it is not evidence that spectral connectivity determines language-model post-training performance.

H.4 Grouped post-training results

For completeness, Table 7 reproduces the five-seed regex-lite grouped post-training results from the earlier submission. These results are secondary here because they probe an algorithmic consequence rather than the new spectral theorem.

I Reproducibility and Code Audit

The spectral experiment is reproducible with a single command from the frozen experiment package, which contains the specification, both implementations, the runners, the blocking validation script,

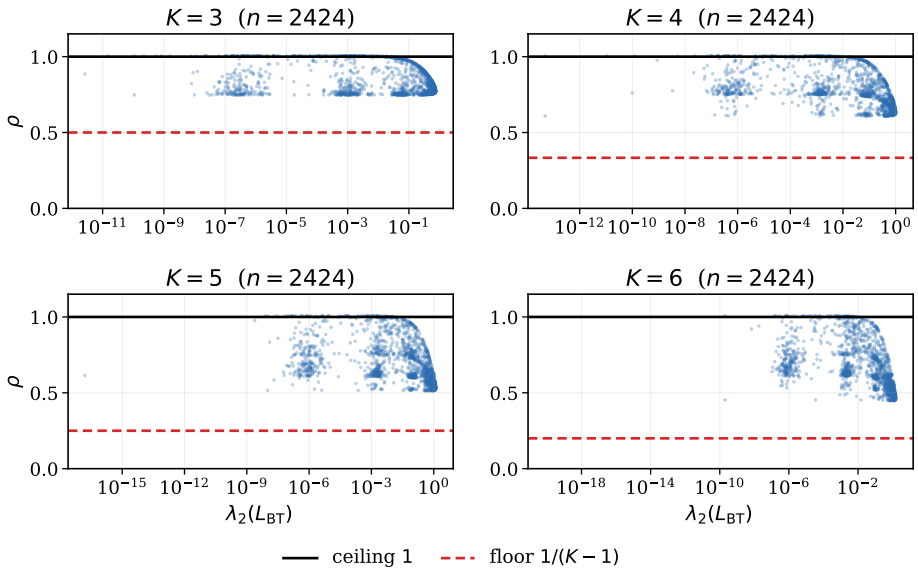

Figure 5: Companion view of the data in Figure 1: realized ρ against $\lambda_2(L_{\mathrm{BT}})$. Same instances as Figure 1, with none removed from the analysis; the display axis is clipped at 1.15 to avoid numerically unresolved near-zero ratios dominating the view.

δ	σ	$1/(1+\delta^2)$	NMSE@200	Kendall τ@200
0.0	1.00	1.000	0.559	0.791
0.3	1.04	0.917	0.560	0.810
0.6	1.17	0.735	0.597	0.759
1.0	1.41	0.500	0.622	0.739
2.0	2.24	0.200	0.718	0.662
4.0	4.12	0.059	0.869	0.417

Table 5: Controlled-noise synthetic experiment reused from the earlier submission.

the raw per-instance data, the processed summaries, the figures, and the logs. Every run records its exact command, git commit, Python and package versions, seed, timestamp, and configuration alongside its output, and no run overwrites a previous result directory.

Prior validation code. Earlier symmetric validation code used uniform ranking averaging, which is exact only at symmetry; all nonsymmetric results reported here use the independently implemented probability-weighted exact calculation. Concretely, the earlier routine averaged the $K!$ stage-Fisher matrices uniformly rather than weighting each ranking by $P_r^{(\sigma)}(\omega)$, and fixed $\sigma = 1$. At exact symmetry the PL law over rankings is uniform, so the two agree there to machine precision, and the earlier symmetric validation remains valid. Away from symmetry the two differ, by up to 63% in relative Frobenius norm on the cases we checked. No claim in this paper about nonsymmetric curvature derives from that earlier code.

J Extended Positioning

This appendix preserves the longer discussion compressed out of Section 7 for the main-text page budget. No citation or claim differs between the two; only the exposition is fuller here.

Comparison topology versus an internally generated Fisher graph. Vojnović, Yun, and Zhou show that optimization rates for ranking-score inference are governed by the algebraic connectivity of item-pair co-occurrence structure, with the framework covering Luce and Plackett–Luce ranking models (Vojnović et al., 2023). Seshadri, Ragain, and Ugander derive structure-dependent

Method	ρ	Val. Exact	Test Exact	Val. Reward	Test Reward	Avg. Utility
Blind GRPO	0.00	0.609 ± 0.019	0.560 ± 0.030	0.670 ± 0.023	0.616 ± 0.027	0.640 ± 0.119
IW-GRPO + step	0.00	0.551 ± 0.068	0.577 ± 0.051	0.634 ± 0.038	0.636 ± 0.035	0.644 ± 0.066
Blind GRPO	0.50	0.575 ± 0.025	0.577 ± 0.042	0.649 ± 0.018	0.633 ± 0.033	0.634 ± 0.155
IW-GRPO + step	0.50	0.623 ± 0.038	0.600 ± 0.020	0.665 ± 0.034	0.646 ± 0.019	0.748 ± 0.002

Table 6: Pythia-410M anchor reused from the earlier submission; mean $\pm$ sample standard deviation over seeds $\{0, 1, 2\}$.

Method	ρ	Val. Exact	Test Exact	Val. Reward	Test Reward	Avg. Utility
Blind GRPO	0.00	0.511 ± 0.045	0.524 ± 0.012	0.570 ± 0.037	0.580 ± 0.009	0.719 ± 0.003
IW-GRPO	0.00	0.489 ± 0.058	0.515 ± 0.029	0.554 ± 0.044	0.574 ± 0.017	0.740 ± 0.001
IW-GRPO + step	0.00	0.490 ± 0.044	0.536 ± 0.011	0.551 ± 0.039	0.588 ± 0.010	0.741 ± 0.001
Blind GRPO	0.25	0.492 ± 0.015	0.526 ± 0.018	0.555 ± 0.016	0.583 ± 0.012	0.721 ± 0.002
IW-GRPO	0.25	0.483 ± 0.031	0.527 ± 0.008	0.552 ± 0.031	0.585 ± 0.004	0.740 ± 0.001
IW-GRPO + step	0.25	0.490 ± 0.051	0.532 ± 0.006	0.553 ± 0.044	0.587 ± 0.003	0.740 ± 0.001
Blind GRPO	0.50	0.493 ± 0.044	0.513 ± 0.029	0.557 ± 0.035	0.575 ± 0.024	0.725 ± 0.001
IW-GRPO	0.50	0.498 ± 0.053	0.520 ± 0.006	0.556 ± 0.040	0.577 ± 0.005	0.741 ± 0.000
IW-GRPO + step	0.50	0.499 ± 0.053	0.532 ± 0.009	0.558 ± 0.042	0.587 ± 0.009	0.740 ± 0.001

Table 7: Five-seed regex-lite experiment reused from the earlier submission. Paired-bootstrap test-reward differences for IW-GRPO+step minus Blind were $+0.008$ with 95% CI $[+0.002, +0.015]$ at $\rho = 0$, $+0.004$ with CI $[-0.003, +0.012]$ at $\rho = 0.25$, and $+0.012$ with CI $[-0.008, +0.036]$ at $\rho = 0.5$.

risk guarantees for repeated-selection models, including tight risk bounds for multinomial-logit and Plackett–Luce estimation (Seshadri et al., 2020). Han and Xu give a unified asymptotic analysis of full, marginal, and quasi-likelihood estimators under deterministic comparison-graph and hypergraph conditions (Han & Xu, 2025). These works establish that spectral connectivity and comparison topology matter for PL estimation. Our object is different: the comparison design is fixed to one complete ranking, and we characterize the *score-dependent Fisher graph generated internally by that ranking*. In particular, $L_{\mathrm{BT}}(r, \sigma)$ in this paper is not the empirical pair-co-occurrence graph of a data-collection design; it is a reward-gap graph whose edge weights are the Fisher sensitivities of hypothetical direct Bradley–Terry comparisons.

Subset-wise feedback and the value of larger groups. Under winner-only subset feedback, Saha and Gopalan obtain PAC-ranking rates with essentially no improvement over the pairwise setting as subset size grows, whereas top-m ranking feedback yields an m-fold improvement in their ranking objective (Saha & Gopalan, 2019a). In an online subset-selection setting, full ranking feedback can yield regret improvements that scale with the played subset size (Saha & Gopalan, 2019b). These results are complementary rather than contradictory to our symmetric saturation theorem: they study repeated adaptive interaction over a universe of items, with different feedback protocols and different statistical objectives. We instead ask how much local information about the full relative-score vector is contained in *one complete PL ranking*. The distinction separates the combinatorics of the observation protocol from the weakest identifiable direction of the score-space Fisher operator.

Optimal design and active preference acquisition. Thekumparampil et al. use D-optimal design to select K-subsets for learning a PL model under a comparison budget (Thekumparampil et al., 2025). In language-model alignment, MASS-DPO derives a PL-specific Fisher/curvature objective and uses log-determinant optimal design to select compact, complementary negative subsets within a prompt (Surana et al., 2026). More broadly, Fisher- and D-optimal-design criteria are standard tools for choosing informative observations (Pukelsheim, 2006). Our goal is prior to such acquisition decisions: we characterize the exact score-space Fisher operator of a realized complete ranking. Consequently, our exposure coefficient α_{ab} is not a sampling probability or a learned acquisition weight; it is the exact amount by which the sequential PL ranking process attenuates the direct pairwise Fisher edge (a, b).

Position of the present result. Three Laplacian objects should be kept distinct. First, classical comparison-design Laplacians encode which items co-occur across collected observations. Second, our

$$L_{\mathrm{BT}}(r,\sigma)=\sum_{a<b}\kappa_{\log}(r_a-r_b;\sigma)(e_a-e_b)(e_a-e_b)^\top$$

is a score-dependent graph of *direct pairwise discriminability*. Third, $H_\sigma(r;r)=L_{\mathrm{PL}}(r;r,\sigma)$ is the exact Fisher operator induced by a complete PL ranking. Theorem 5.3 connects the latter two edge by edge, and Theorem 5.5 turns the exposure bounds into the uniform operator comparison of Section 5. Likewise, the nullspace statement behind Theorem 4.1 is elementary linear algebra; the contribution here is its structural role: it explains why positive local curvature on shrinking PL stages cannot be converted by naive scalar addition into a global bound on the weakest identifiable Fisher curvature. The exact harmonic formula of Theorem 3.1 and the fixed-range separation are consequences of this same score-space viewpoint.

K Label Cost of the Weakest Direction

The main text characterizes γ_{PL} but does not show that it costs anything. This appendix asks whether γ_{PL} predicts estimation error, using the same controlled configurations as Section 6.2. The protocol was frozen before any result was generated and is included with the supplement. Nothing here is a new theorem: asymptotic normality of the maximum-likelihood estimator is classical, and we are checking that the quantity characterized in this paper behaves as Fisher information should at finite N.

Design. For $K=4$, $\sigma=1$, $R\in\{0.5,1,2,4\}$ and $t\in\{0,\frac{1}{2},1\}$ along the path of Section 6.2, we draw $N\in\{100,300,1000,3000,10^4\}$ rankings exactly from the PL law by the Gumbel–max construction, fit $\hat{r}$ by maximum likelihood on $V_\perp$ (the PL log-likelihood is concave), and record the squared error along the eigenvector of $H_\sigma(r;r)$ for γ_{PL} and, as a control, along the eigenvector for $\lambda_{\max}$. Twelve configurations, 200 replicates per cell. $R=8$ was excluded *in advance*, since $\gamma_{\mathrm{PL}}\approx 3.4\times10^{-7}$ there would require $N\sim10^6$. Three blocking checks passed before the run: the sampler reproduces exact PL ranking probabilities (total variation 0.0015), the estimator recovers r on large samples, and γ_{PL} agrees with the exact operator.

The frozen hypothesis fails as stated. Over all 60 cells, regressing log weak-direction MSE on $\log(N\gamma_{\mathrm{PL}})$ gives slope -3.750 with $R^2=0.386$. The frozen hypothesis predicted -1 and is *not supported*. The failure is confined to the eight cells with $N\gamma_{\mathrm{PL}}\le 6.34$, where the estimator runs away along a direction the data barely constrains, reaching mean squared errors of order 10^{27} — not a different scaling exponent but a breakdown of estimation.

Restricted analysis (post hoc). Excluding cells with $N\gamma_{\mathrm{PL}}<10$, a criterion motivated by the effective sample size in the measured direction rather than fitted to the data, leaves 51 cells and gives slope -1.0045 with $R^2=0.9955$, and MSE/prediction of median 1.038 over the range $[0.72,1.28]$, with no free parameters. The control direction behaves the same way against $\lambda_{\max}$: slope -1.0101, $R^2=0.9975$. Within a single configuration the weak- and strong-direction errors differ by up to five orders of magnitude, so the dependence is direction-specific. This restriction was not part of the frozen specification and is reported as post hoc.

Cost in labels. At matched K, matched range and matched N, the ratio of weak-direction errors tracks the inverse ratio of curvatures: at $R=2$ the curvature ratio is 3.46 and the observed error ratios are 3.80, 2.89 and 3.18 at $N=10^3, 3\times10^3, 10^4$; at $R=4$, where both endpoints are adequately sampled only at $N=10^4$, the curvature ratio is 41.44 and the observed error ratio is 39.85. The more extreme $R=4$ comparison lies at the boundary of the included regime: the clustered endpoint has $N\gamma_{\mathrm{PL}}=10.05$ at $N=10^4$. Reading this operationally, the clustered configuration needs roughly $\gamma_{\mathrm{spread}}/\gamma_{\mathrm{cluster}}$ times as many rankings to pin the weakest contrast down as well — about $41\times$ at $R=4$ — with the same number of candidates and the same reward range.

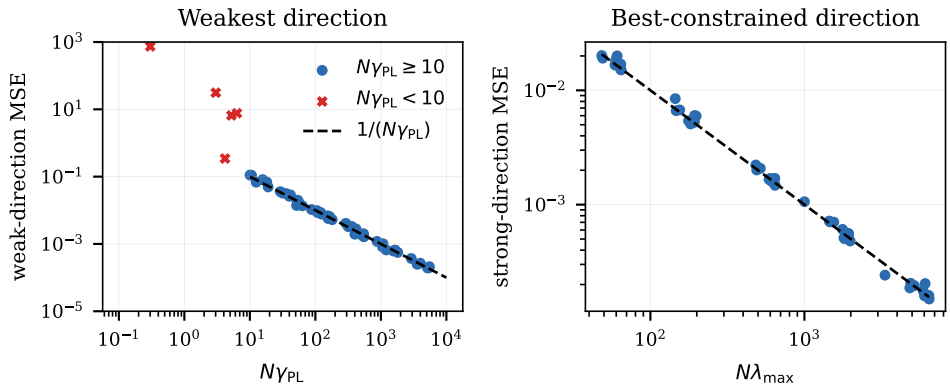

Figure 6: Estimation error against effective sample size in each direction. Left: weakest direction; crosses are the cells with $N\gamma_{\mathrm{PL}} < 10$, where the estimator fails rather than degrades. Right: best-constrained direction. Dashed lines are $1/(N\gamma_{\mathrm{PL}})$ and $1/(N\lambda_{\max})$, not fits.

Scope. This establishes no finite-sample rate, no minimax statement, no claim about how groups should be *selected*, and no downstream language-model consequence.