THE HIGHER-ORDER IDENTIFIABILITY PARADOX: LEARNING AND CERTIFYING HYPERGRAPH ORDER

Anonymous authors
Paper under double-blind review

ABSTRACT

When does higher-order information become usable under distribution shift, and when can a learner certify that a simpler view is sufficient? We study this *higher-order identifiability paradox* in hypergraphs. Projection can merge contexts whose changing mixture reverses the Bayes decision under a stable full-state mechanism. We derive the exact attainable score interval and minimax Brier regret under joint total-variation and likelihood-ratio constraints. Our main learning-theoretic result separates the environments needed for mechanism estimation from those needed for efficient certification. For a realizable quadratic family with d hidden orientations and fixed-budget Gaussian population-moment acquisition, a designed collection of $d - \Theta(\sqrt{d})$ environments already attains the optimal integrated estimation rate. Within the centered isotropic design class, matching full-span certification efficiency requires only $O(d^{1/3})$ directions to remain unspanned. Zero-diagonal matrix completion yields matching estimation and safety–power bounds, with certification power at exact sufficiency. Our designs attain both rates; estimation-optimal designs can otherwise incur a $d^{1/4}$ certification penalty. These results inform CHORD (*Certified Hypergraph Order Selection for Robust Deployment*). Controlled studies test its spectral and power predictions. Around frozen predictors on AskUbuntu and Math StackExchange, its separate temporal mode improves all four run-level metrics in the seven affected runs and exactly replays the remaining 41 of 48 arity–seed decisions.

1 INTRODUCTION

Why can preserving more of a hyperedge make an out-of-distribution (OOD) prediction harder to justify? Hypergraphs represent group interactions that none of their constituent pairs can express (Benson et al., 2018; Yoon et al., 2020; Zhang et al., 2020; Chien et al., 2022). A richer view gives an oracle more information and can only improve its best achievable prediction. A learner must also establish which effects remain supported outside its source environments (Peters et al., 2016; Arjovsky et al., 2019; Krueger et al., 2021; Chen et al., 2022). This creates a less familiar question: can environments that are already optimal for learning a mechanism still be inefficient for certifying which parts of its representation can safely be discarded?

The difficulty starts before any training objective is chosen. A coarse view merges distinct contexts; changing their relative frequency can reverse the predicted outcome even when every context-specific mechanism stays fixed. This is a projection-induced Simpson reversal (Simpson, 1951). Higher-order information separates the contexts, but also exposes effects that aggregate source observations may leave unidentified. Diverse environments can recover these effects. Our central finding is that recovering a mechanism and certifying a coarser representation require different kinds of evidence: their optimal designs depend differently on weakly observed directions of the source equations. We call this tension the *higher-order identifiability paradox*.

Figure 1 introduces the projection and identification boundaries. We realize the first with hypergraphs whose labelled views agree at every proper order but whose hidden mixture supports opposite decisions. Transport is invariant precisely when merged states have equal outcome means, the standard factorization property we call *causal lumpability*. Inside each projection cell, joint total-variation and likelihood-ratio constraints give an exact attainable score interval,

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.

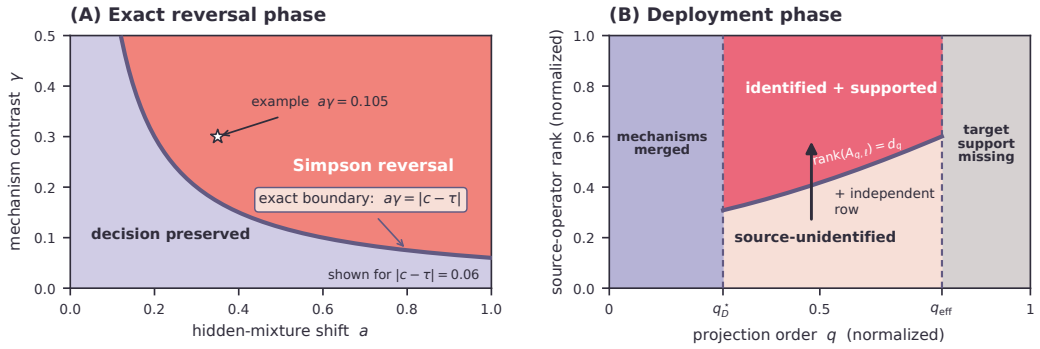

Figure 1: **Exact reversal and deployment phases.** (A) For mixture shift $a \in [0, 1]$, hidden-state probabilities $c \pm \gamma$, and decision threshold τ, reversal occurs exactly when $a\gamma > |c - \tau|$; the star marks the displayed example. (B) Order q lies between the first mechanism-preserving order $q^\star_{\mathrm{D}}$ and the largest supported order q_{eff}. The stronger full-mechanism condition requires source operator $A_{q,\ell}$, at moment resolution ℓ, to have rank equal to its parameter dimension d_q. One linearly independent row closes the depicted one-dimensional rank deficiency; (B) uses schematic boundary locations.

a sharp reversal boundary, and matching minimax Brier regret. A quantile formulation extends the same calculation to general fibers, with explicit sampling and approximation error.

The sharper learning-theoretic distinction appears under a fixed acquisition budget. For quadratic mechanisms on d hidden trade orientations, we study Gaussian releases of source population values and first moments. Centered environment designs first identify the mechanism at $d - \Theta(\sqrt{d})$ environments; an explicit design already attains the optimal integrated estimation rate there. Yet matching the certification efficiency of a full-span design requires the number of unspanned directions to fall to $O(d^{1/3})$. Between these regimes, additional environments can reduce the certifiable tolerance by a factor $\Theta(\sqrt{d})$ at the same budget, while leaving the optimal estimation-rate order unchanged. The obstruction comes from completing a zero-diagonal Hessian, and we give matching upper and lower bounds over the declared design class.

CHORD turns these distinctions into order-selection rules, using structural certificates when their inputs are available and a separate temporal rule around frozen Deep Sets and self-attention predictors on AskUbuntu and Math StackExchange. Controlled Gaussian experiments isolate the new spectral predictions; real temporal experiments test the frozen-predictor wrapper. Our contributions are threefold:

- **We characterize the ambiguity created by hypergraph projection.** Stable full-state mechanisms can induce opposite projected decisions. Exact score intervals determine reversal and minimax regret, with a general-fiber extension and explicit statistical error.
- **We prove distinct environment requirements for learning and certification.** A structured completion argument yields a fixed-budget minimax learning rate and a matching safe-coarsening power law. The resulting $\sqrt{d}$ and $d^{1/3}$ unspanned-dimension scales separate algebraic recovery from statistically efficient certification.
- **We construct matching designs and evaluate order selection.** A weighted environment construction attains both laws. Controlled studies test its spectrum and power; on AskUbuntu and Math StackExchange, all four run-level metrics improve in the seven runs affected by the wrapper's 7 of 48 arity–seed switches, with exact replay elsewhere.

2 Related Work

Higher-order learning connects three questions: what a representation retains, what source observations identify, and what transfers to a new environment. We study their interaction by asking when compression preserves a stable conditional and when source diversity determines its target value.

Higher-order representation, reconstruction, and time. Set encoders aggregate unordered members (Zaheer et al., 2017; Soelch et al., 2019; Lee et al., 2019); incidence-aware models retain

hyperedge membership (Zhang et al., 2020; Chien et al., 2022; Kim et al., 2022). Higher-order and temporal prediction expose signal beyond pairs (Benson et al., 2018; Yoon et al., 2020; Wan et al., 2024; Li et al., 2026; Behrouz et al., 2023). Theory studies generalization and order reduction (Wang et al., 2025; Yadati, 2025; Bian et al., 2025); reconstruction asks whether a graph shadow determines the full hypergraph (Young et al., 2021; Wang & Kleinberg, 2024; Gu & Polyanskiy, 2023; Bresler et al., 2025). We need only equal outcome means among merged states, yet combinatorial trades show that identical shadows need not supply them (Khosrovshahi & Tayfeh-Rezaie, 2009).

Invariance and graph OOD generalization. Invariant prediction and robust training seek stable relations (Peters et al., 2016; Arjovsky et al., 2019; Krueger et al., 2021; Sagawa et al., 2020). Graph methods optimize stable subgraphs or learnable walks (Chen et al., 2022; Sun et al., 2025); link-prediction theory studies larger test graphs (Zhou et al., 2022), and mixture geometry separates interpolation from extrapolation (Zhang et al., 2024). CHORD asks a prior question around a frozen predictor: does its projection preserve and identify the stable conditional?

Causal abstraction and heterogeneous environments. Causal abstraction asks when compression preserves a mechanism (Pearl, 2009; Rubenstein et al., 2017; Geiger et al., 2025). Auxiliary variables identify latent structure (Hyvärinen et al., 2019; Yin et al., 2024), and DAG-FM exploits mechanism diversity for directed acyclic graph discovery (Chen et al., 2026). Here projection-cell constancy gives preservation; the source row space gives environmental identification.

Interpolation, outer-product frames, and inverse testing. Value–gradient constraints belong to multivariate Hermite interpolation (Sauer & Xu, 1995; Alexander & Hirschowitz, 1995). Rank-one recovery and outer-product frames characterize spanning and stability (Cai & Zhang, 2015; Casazza et al., 2016; Rauhut & Terstiege, 2019); Gaussian inverse testing characterizes minimax separation (Schluttenhofer & Johannes, 2020). These works supply the tools. Our contribution is the zero-diagonal Hessian completion induced by environment means: learning and safe order certification then require different environment budgets. This is the learning–certification crossover; the frame basis supplies its construction.

Robust statistics and partial identification. Robust testing and distributionally robust optimization use ambiguity sets (Kassam, 1981; Kuhn et al., 2025); causal sensitivity analysis derives sharp bounds under outcome shift (Asiaee et al., 2026; Dorn & Guo, 2023). Placing the ambiguity set inside each projection cell, we obtain an exact score interval and minimax Brier rule under joint total-variation and likelihood-ratio constraints, tying robustness to identifiable order.

3 SETTING AND OBSERVATION MODEL

Stable mechanism and two orders. For a candidate hyperedge S queried at time t, let $K = |S|$ be its observed size, $H = (G_{<t}, S) \in \mathcal{H}$ its complete rooted history after S is withheld from the past hypergraph $G_{<t}$, and $Y \in \{0, 1\}$ its future label. Index source environments by $e \in [E] := \{1, \dots, E\}$ and the target by $\star$. Their laws of H may differ; the conditional mean $\eta : \mathcal{H} \to [0, 1]$ is stable:

$$Y = \mathbf{1}\{\varepsilon \le \eta(H)\}, \quad \varepsilon \sim \mathrm{Uniform}(0, 1), \quad \varepsilon \perp H \mid e, \tag{1}$$

$$\eta(H) = \chi(H)^\top \theta, \quad \chi : \mathcal{H} \to \mathbb{R}^p, \quad \theta \in \Theta \subseteq \{v \in \mathbb{R}^p : \|v\|_2 \le B\}, \quad 0 < B < \infty, \tag{2}$$

$$Z_q = \Pi_q(H) := T_q(G_{<t}, S), \quad q \in \{0, \dots, q_{\max}\}, \sigma(Z_q) \subseteq \sigma(Z_{q+1}), \quad T_{q_{\max}}(H) = H. \tag{3}$$

Here "causal" means stability under declared covariate interventions. One-hot χ is exact on finite $\mathcal{H}$; otherwise the model class also requires $\chi(h)^\top \theta \in [0, 1]$. The projection order q masks both history and candidate before encoding; $\ell \ge 0$ indexes source-only moment resolution.

Aggregate observation model. For an outcome-free source feature $\psi_{q,\ell}(H) \in \mathbb{R}^{s_{q,\ell}}$ whose products with χ and Y are integrable, independent cohorts or a dependence-valid joint release provide only

$$\widehat{A}^e_{q,\ell} \approx \mathbb{E}_e[\psi_{q,\ell}\chi^\top], \qquad \widehat{b}^e_{q,\ell} \approx \mathbb{E}_e[\psi_{q,\ell}Y]. \tag{4}$$

Environment blocks constrain the same mechanism without releasing record-level (χ, Y) pairs. Write $A_{q,\ell}$ and $b_{q,\ell}$, and their hatted versions, for the vertical stacks of these E population moments and

releases. Target adaptation stores $a_{\star,q}(z) := \mathbb{E}_\star[\chi(H) \mid Z_q = z]$ on a declared covered set $\mathcal{Z}_q^{\mathrm{cov}}$ using a known design or independent unlabeled full contexts. A target-agnostic alternative assumes $\eta(H) = \phi_q(Z_q)^\top \theta_q$, with $\phi_q : \mathcal{Z}_q \to \mathbb{R}^{d_q}$, $\theta_q \in \mathbb{R}^{d_q}$, and declared target support; $\mathcal{Z}_q$ is the range of Z_q. Approximation incurs a one-sided risk bound. Rank is always relative to the declared map and release.

Estimated inputs and conservative defaults. Simultaneous radii satisfy $\|\widehat{A}_{q,\ell} - A_{q,\ell}\|_{\mathrm{op}} \le \epsilon_{A,q,\ell}$, $\|\widehat{b}_{q,\ell} - b_{q,\ell}\|_2 \le \epsilon_{b,q,\ell}$, and $\sup_{z \in \mathcal{Z}_q^{\mathrm{cov}}} \|\widehat{a}_{\star,q}(z) - a_{\star,q}(z)\|_2 \le \epsilon_{a,q}$; the last radius adds $B\epsilon_{a,q}$ to the score radius. Support holes and dependence are charged separately; absent full contexts, overlap and source–target shift are sensitivity parameters, and missing radii force fallback or abstention. Appendix G gives estimators and separates certified excess risk, certified paired improvement, and temporal validation. The first abstains on structural failure, the second replays its baseline when the required source–target bound is unavailable, and the third makes only an empirical claim. Appendix B.3 adds a declared uniform mechanism-drift radius to the score and paired-risk bounds.

Frozen representations. Train the richest view once, freeze its checkpoint, heads, and calibration, and recompute both history and candidate from each T_q; a cached full incidence violates the order restriction. Conditional on the fitting data,

$$s_q = f_{\widehat{\omega}}(Z_q, K), \qquad \phi_{q,K}^{\mathrm{sc}}(Z_q) = (1, s_q)^\top. \tag{5}$$

Here $\widehat{\omega}$ denotes frozen fitted parameters; conditioning on K gives two-dimensional moments. Score sufficiency and equal merged-state means remain distinct requirements.

4 Projection Can Reverse OOD Decisions

For a finite projection $Z = \Pi(H) \in \mathcal{Z}$, enumerate the positive-reference-mass states in the fiber $\mathcal{F}_z := \Pi^{-1}(z)$ as $h_1, \dots, h_J$, and set $\mu_i := P_0(H = h_i \mid Z = z)$ and $\eta_i := \eta(h_i)$, with $\mu_z = (\mu_1, \dots, \mu_J)$. Writing P_Z for the marginal law of Z, the reference law factorizes as $P_0(H = h_i) = P_Z(z)\mu_i$ and has projected mean $m_0(z) := \sum_{i=1}^J \mu_i \eta_i$. A target preserves the declared marginal P_Z but changes the conditional mixture on $\mathcal{F}_z$ from μ_z to ν_z. With total variation $\mathrm{TV}(\nu, \mu) := \frac{1}{2}\sum_i |\nu_i - \mu_i|$, radius $\rho_{\mathrm{fib}} \in [0, 1]$, and likelihood-ratio bound $\kappa \in (1, \infty]$, define

$$m_\nu(z) = \sum_{h \in \mathcal{F}_z} \nu_z(h)\eta(h), \tag{6}$$

$$\mathcal{U}_z(\rho_{\mathrm{fib}}, \kappa) = \{\nu \ll \mu_z : \mathrm{TV}(\nu, \mu_z) \le \rho_{\mathrm{fib}},\ \kappa^{-1} \le \nu(h)/\mu_z(h) \le \kappa\}. \tag{7}$$

This intervention changes only the mixture over projection-hidden states. We use the usual extended-real convention at $\kappa = \infty$; the limiting case $\kappa = 1$ permits no reweighting.

Definition 4.1 (Projection-level mechanism invariance (causal lumpability))**.** The mechanism factors through Π when $\eta(h) = \bar{\eta}\{\Pi(h)\}$ on every relevant fiber: merged states have the same outcome probability. Equivalently, every within-fiber reweighting preserves the projected mean.

Relabel the J states so that $\eta_1 \le \cdots \le \eta_J$, set $c_i^- = \mu_i(1 - \kappa^{-1})$, $c_i^+ = \mu_i(\kappa - 1)$, and let $G_{\mathrm{hi}}(t; c, \eta) := \max\{s^\top \eta : 0 \le s \le c, \mathbf{1}^\top s = t\}$ and $G_{\mathrm{lo}}(t; c, \eta) := \min\{s^\top \eta : 0 \le s \le c, \mathbf{1}^\top s = t\}$. With $\bar{\rho}_{\mathrm{fib}} = \min\{\rho_{\mathrm{fib}}, \sum_i c_i^-, \sum_i c_i^+\}$,

$$\omega_z^\pm = \max_{\nu \in \mathcal{U}_z} \{\pm(m_\nu(z) - m_0(z))\}, \tag{8}$$

$$\omega_z^+ = \max_{0 \le t \le \bar{\rho}_{\mathrm{fib}}} \{G_{\mathrm{hi}}(t; c^+, \eta) - G_{\mathrm{lo}}(t; c^-, \eta)\},$$

$$\omega_z^- = \max_{0 \le t \le \bar{\rho}_{\mathrm{fib}}} \{G_{\mathrm{hi}}(t; c^-, \eta) - G_{\mathrm{lo}}(t; c^+, \eta)\}. \tag{9}$$

Set $D_z = \omega_z^+ + \omega_z^-$. The adversary fills likelihood-ratio capacities from low to high outcome means, so one sort computes the exact value in $O(J \log J)$ time.

Theorem 4.2 (Exact range, reversal, and minimax risk)**.** *On each fiber restricted to* $\mathrm{supp}(\mu_z)$*, the exact attainable score range is*

$$I_z = [m_0(z) - \omega_z^-,\ m_0(z) + \omega_z^+]. \tag{10}$$

Thus the same observed input can support every score in I_z as its hidden mixture changes. With the extrema taken over the active fiber, $\operatorname{osc}_z(\eta) := \max_{h\in\operatorname{supp}(\mu_z)} \eta(h) - \min_{h\in\operatorname{supp}(\mu_z)} \eta(h)$,

$$\omega_z^{\pm} \le \rho_{\mathrm{fib}} \operatorname{osc}_z(\eta), \qquad D_z \le 2\rho_{\mathrm{fib}} \operatorname{osc}_z(\eta). \tag{11}$$

The one-sided bounds are attained for small enough ρ_{fib}, and for every $\rho_{\mathrm{fib}} > 0, \kappa > 1$, $D_z = 0$ iff Π is causally lumpable there. Allowing the conditional law to vary independently within each fiber, if $m_0(z_1) > m_0(z_0)$, a strict ranking reversal exists iff

$$m_0(z_1) - m_0(z_0) < \omega_{z_1}^- + \omega_{z_0}^+. \tag{12}$$

Admissible mixture motion must exceed the reference gap. If $m_0(z) > \tau$, a strict downward crossing exists iff $m_0(z) - \omega_z^- < \tau$; for $\mathbf{1}\{m > \tau\}$, reversal occurs iff $m_0(z) - \omega_z^- \le \tau$. The exact minimax Brier regret relative to the target-specific Z-Bayes rule is

$$\inf_{g:\mathcal{Z}\to\mathbb{R}} \sup_{(\nu_z)_z \in \prod_z \mathcal{U}_z} \mathbb{E}_{P_Z}(g(Z) - m_\nu(Z))^2 = \mathbb{E}_{P_Z} \frac{D_Z^2}{4}. \tag{13}$$

The midpoint is optimal; half-width remains because Z hides the target mixture.

The same interval law holds for bounded real outcomes under squared loss. For multiclass Brier loss, the optimal prediction is the center of the smallest ball containing all attainable probability vectors (the Chebyshev center; Appendix B.2). The robust optimization (Kassam, 1981; Kuhn et al., 2025) becomes a finite-sample diagnostic with simultaneous bounds on full-state means and weights; without those inputs, it gives a sensitivity curve. Appendix B.7 gives attained quantile endpoints for general fibers and explicit sampling error.

Octahedral trades. For an integer $r \ge 2$, let $[r] = \{1, \dots, r\}$, $V_r = \{v_i^0, v_i^1 : i \in [r]\}$, and $E_x = \{v_i^{x_i} : i \in [r]\}$ for $x \in \{0,1\}^r$. The standard design trade (Khosrovshahi & Tayfeh-Rezaie, 2009)

$$\mathcal{H}_+ = \{E_x : \textstyle\sum_i x_i \text{ even}\}, \qquad \mathcal{H}_- = \{E_x : \textstyle\sum_i x_i \text{ odd}\} \tag{14}$$

is edge-disjoint. For an r-uniform hypergraph G and a q-set F, write $C_q(G)_F := \sum_{E\in G} \mathbf{1}\{F \subseteq E\}$ for its labelled q-shadow count.

Theorem 4.3 (All lower-order projections are identical)**.** *For every $q < r$, $C_q(\mathcal{H}_+) = C_q(\mathcal{H}_-)$. Each compatible q-set has 2^{r-q-1} extensions in either orientation; at $q = r$ the systems are disjoint. Hence every proper-order view agrees while full edges reveal opposite parity (Appendix B).*

Let the latent orientation $S \in \{+1, -1\}$ have environment-specific weight $w_e := P_e(S = +1)$ and the stable outcome mechanism

$$Y \mid S = s \sim \mathrm{Bernoulli}(c + s\gamma), \qquad 0 < \gamma \le \min\{c, 1-c\}. \tag{15}$$

Every $q < r$ score is $c + \gamma(2w_e - 1)$. For $0 < a < 1$, let $\mu_\pm$ be orientation laws with $w_\pm = (1 \pm a)/2$. Taking $c = 1/2$ gives

$$m_\pm = \tfrac{1}{2} \pm a\gamma, \quad \mathrm{TV}(\mu_+, \mu_-) = a, \quad \kappa = \frac{1+a}{1-a}, \quad |m_+ - m_-| = 2a\gamma. \tag{16}$$

The two displayed environments lie at fiber radius $\rho_{\mathrm{fib}} = a$ and likelihood-ratio bound $\kappa = (1 + a)/(1 - a)$; swapping them attains the oscillation bound and reverses the decision. Over these two environments, midpoint minimax Brier excess is $a^2\gamma^2$, versus $4a^2\gamma^2$ for source transfer. The full view separates S; Appendix B extends equality to overlaps, random embeddings, and thinning.

5 What Heterogeneous Environments Identify

Mixture diversity supplies equations. In a projection cell with J states, collect their stable outcome means in $k \in [0,1]^J$ and the E source mixture weights in $W \in [0,1]^{E\times J}$. Source label means satisfy $m = Wk$. For a coefficient class with nonempty interior, a target mixture mean $w_\star^\top k$ is identified exactly when its probability vector $w_\star$ belongs to $\mathrm{Row}(W)$ (Proposition E.1). An independent row identifies another direction; duplicating a row improves precision along directions already identified. The general source-moment model has the same form:

$$a_{\star,q}(z) = \mathbb{E}_\star[\chi(H) \mid Z_q = z], \qquad m_{\star,q}(z) = a_{\star,q}(z)^\top \theta, \tag{17}$$

$$A_{q,\ell} = \operatorname{stack}_e \mathbb{E}_e[\psi_{q,\ell}\chi^\top], \qquad b_{q,\ell} = \operatorname{stack}_e \mathbb{E}_e[\psi_{q,\ell} Y] = A_{q,\ell}\theta. \tag{18}$$

The target score is identified iff $a_{\star,q}(z) \in \mathrm{Row}(A_{q,\ell})$; full column rank identifies every coefficient. Order q determines the deployed view and ℓ the source features. Independent trade mixtures replace an orientation moment by constant moments (Proposition E.2).

Refinement changes the question asked of the sources. Let $\mathcal{R}_\star(g_q) = \mathbb{E}_\star(g_q(Z_q) - Y)^2$ and $\mathcal{R}^*_{\star,H} = \mathbb{E}_\star(\eta(H) - Y)^2$. Write $\mathcal{F}_q = \sigma(Z_q)$, $a_q = \mathbb{E}_\star[\chi(H) \mid \mathcal{F}_q]$, and $M_q = \mathbb{E}_\star[a_q a_q^\top]$.

Theorem 5.1 (Projection loss and fixed-source ambiguity)**.** *For every Z_q-measurable predictor g_q,*

$$\mathcal{R}_\star(g_q) - \mathcal{R}^*_{\star,H} = V_{\star,q} + \mathbb{E}_\star(g_q - m_{\star,q})^2, \qquad V_{\star,q} = \mathbb{E}_\star \mathrm{Var}(\eta(H) \mid Z_q). \tag{19}$$

Assume $\mathbb{E}_\star\|\chi(H)\|_2^2 < \infty$ and nested projections. Fix a source operator A with nullspace $\mathcal{N} = \ker A$, its orthogonal projector $P_\mathcal{N}$, and a radius $R \geq 0$. Then

$$M_{q+1} - M_q = \mathbb{E}_\star[(a_{q+1} - a_q)(a_{q+1} - a_q)^\top] \succeq 0, \tag{20}$$

$$\sup_{v \in \mathcal{N}: \|v\|_2 \leq R} \mathbb{E}_\star(a_q^\top v)^2 = R^2 \lambda_{\max}(P_\mathcal{N} M_q P_\mathcal{N}). \tag{21}$$

Projection loss is nonincreasing; exposure and $\mathrm{tr}(P_\mathcal{N} M_q)$ are nondecreasing.

Equation (19) separates information lost by projection from error in learning the projected answer. The second monotonicity says that finer target questions can involve additional unidentified coefficients. It does not imply increasing optimal total risk: with the same training experiment, parameter class, target, and loss, a learner observing Z_{q+1} can emulate every Z_q rule. The next section holds the acquisition budget fixed and asks how much evidence is needed to learn the mechanism and to certify that a coarse view loses little.

6 Learning and Certifying Under a Fixed Budget

Algebraic rank answers whether infinite-precision equations determine a mechanism. Learning and safe order selection ask different finite-budget questions. We make this distinction quantitative on a hypergraph family whose hidden interactions induce a structured quadratic inverse problem.

A realizable family and a fixed acquisition channel. Take $d \geq 10$ disjoint labelled r-way trades in the history of a disjoint query, and encode their orientations by $X \in \{-1, 1\}^d$. Every $q < r$ shadow hides X, while the complete edges reveal it. The target orientations are independent and uniform. Consider the multiaffine mechanism

$$\eta(X) = \tfrac{1}{2} + b^\top X + \tfrac{1}{2} X^\top Q X, \qquad Q = Q^\top, \quad \mathrm{diag}\, Q = 0, \quad V := \|b\|_2^2 + \tfrac{1}{2}\|Q\|_F^2 \leq R^2. \tag{22}$$

Here $b \in \mathbb{R}^d$, $Q \in \mathbb{R}^{d \times d}$, and $0 < R \leq [4\sqrt{d + \binom{d}{2}}]^{-1}$ ensures $\eta \in [0, 1]$. The coarse view loses exactly V units of Brier risk. Quadratic degree counts interactions between trade orientations; their hypergraph visibility threshold is r. In a source with independent orientations of mean x, the population mean is $f(x) = \frac{1}{2} + b^\top x + \frac{1}{2} x^\top Q x$; normalized first-order moments give $\nabla f(x)$ (Appendix C). For known means x_e and allocations n_e totaling N, observe only

$$(\widehat{f}_e, \widehat{g}_e) = (f(x_e), \nabla f(x_e)) + \frac{\sigma}{\sqrt{n_e}} \zeta_e, \qquad \zeta_e \overset{\text{iid}}{\sim} N(0, I_{d+1}), \quad \sigma > 0. \tag{23}$$

These bounds apply to the declared Gaussian channel. Empirical moments and raw records have different sampling distributions and therefore define separate statistical experiments.

Environment designs. Fix $0 < \alpha \leq 1/2$, let $m \leq m_d := \lfloor(\sqrt{8d+1} - 1)/2\rfloor$ be a nonnegative integer, and put $k = d - m$, $E = k + 1$. The class $\mathfrak{D}_{d,m}(\alpha)$ consists of centered regular-simplex designs in a k-dimensional subspace $\mathcal{V} \subseteq \mathbb{R}^d$, with equal allocation $n_e = N/E$ and N/E a positive integer. Equivalently, $\|x_e\|_2 = \alpha$, $\sum_e x_e = 0$, and $E^{-1}\sum_e x_e x_e^\top = (\alpha^2/k) P_\mathcal{V}$, where $P_\mathcal{V}$ is the orthogonal projector. Thus m counts directions outside the span of the environment means. Centered gradients give $QV_{\mathrm{env}} + s_N Z$, where the columns of V_{env} are an orthonormal basis of $\mathcal{V}$, Z has independent standard normal entries, and $s_N^2 = \sigma^2 k/(N\alpha^2)$.

Theorem 6.1 (Optimal learning at the first identifiable environment count)**.** *In the preceding model, write $\Theta_R = \{(b, Q) : V \leq R^2\}$, and let $\widehat{\eta}$ use the complete target view. Uniformly over the declared $d, m, N, \alpha, R, \sigma$,*

$$\inf_{D \in \mathfrak{D}_{d,m}(\alpha)} \inf_{\widehat{\eta}} \sup_{(b,Q) \in \Theta_R} \mathbb{E}\|\widehat{\eta} - \eta\|^2_{L_2(P_\star)} \asymp \min\left\{R^2, \frac{\sigma^2 d^3}{N\alpha^2}\right\}. \tag{24}$$

The comparison constants are universal. The lower bound holds even among arbitrary radius-α source designs with the same total budget. An explicit design and a linear completion estimator attain the upper bound. For centered designs with $\binom{m+1}{2} > d$, a nonzero ambiguity remains at every budget.

The first feasible count is $E = d + 1 - m_d = d - \Theta(\sqrt{d})$ (Theorem E.10). A suitable design attains the optimal learning rate at this threshold; efficient certification requires more environment diversity.

Theorem 6.2 (A different environment threshold for safe coarsening)**.** *Fix $\delta, \beta \in (0, 1/4)$. A rule $C \in \{0, 1\}$ certifies $V < \varepsilon$ by returning $C = 1$ and is required to satisfy*

$$\sup_{(b,Q)\in\Theta_R:\, V\geq\varepsilon} \Pr(C = 1) \leq \delta, \qquad \Pr_{b=0,Q=0}(C = 1) \geq 1 - \beta. \tag{25}$$

The inequalities require uniform safety and power at exact sufficiency. Define

$$r_{d,m,N} := \frac{\sigma^2(d-m)}{N\alpha^2}\sqrt{d^2 + m^6}. \tag{26}$$

If $r_{d,m,N} \leq \bar{c}_{\delta,\beta} R^2$ for a sufficiently small positive constant $\bar{c}_{\delta,\beta}$, the smallest achievable $\varepsilon \in (0, R^2]$, optimized over designs $D \in \mathfrak{D}_{d,m}(\alpha)$ and all rules, is $\asymp_{\delta,\beta} r_{d,m,N}$. A quadratic-energy certificate and a Gaussian sign-mixture lower bound attain the two sides of this rate.

At the algebraic threshold $m \asymp \sqrt{d}$, certification needs tolerance of order $\sigma^2 d^{5/2}/(N\alpha^2)$. Its full-span benchmark is $\sigma^2 d^2/(N\alpha^2)$, attained precisely in the scaling regime $m = O(d^{1/3})$. Thus $O(\sqrt{d})$ additional, suitably placed environments reduce the certifiable tolerance by a factor $\Theta(\sqrt{d})$ at fixed N. Equivalently, the same tolerance needs a $\Theta(\sqrt{d})$-fold smaller budget, while the optimal learning-rate order is unchanged. This statistical crossover concerns the declared centered design class.

Why the thresholds differ, and how to attain them. Let $U \in \mathbb{R}^{d\times m}$ have orthonormal columns spanning $\mathcal{V}^\perp$, with rows $u_i^\top$, and define $D_U(B) = \operatorname{diag}(UBU^\top)$ for symmetric $m \times m$ matrices B. Completing QV_{env} uses the constraint $\operatorname{diag} Q = 0$. If a_j are the $t = m(m+1)/2$ singular values of D_U, its rank-one rows satisfy $\sum_j a_j \leq m$. This enforces $\sum_j a_j^{-2} \gtrsim m^4$ and $\sum_j a_j^{-4} \gtrsim m^6$. The former controls estimation error; the latter controls quadratic-energy testing. Their different dependence on the weak directions produces the two environment scales.

For $m \geq 1$, an explicit design starts from rows $\sqrt{w}\, e_i$ and $e_i + e_j$ $(1 \leq i < j \leq m)$, pads with zeros to d rows, and orthonormalizes the columns to obtain U; e_i denotes the ith coordinate vector. This outer-product basis is classical (Casazza et al., 2016). Taking $w = m^{1/4}$ controls both inverse spectral sums and attains both theorems. At $d = t$, choosing $w = 1$ still attains the learning rate but increases the certification threshold by $\Theta(d^{1/4})$. Appendices C–D give the completion algorithm, exact covariance, and matching proofs.

7 Certified Order Selection

The structural mode of CHORD uses two complementary certificates. The interval in Theorem 4.2 bounds a projected target score under a declared shift envelope; Theorem 6.2 tests whether the coarse view loses at most a specified mechanism energy. Known support and source-moment radii remain necessary inputs. For a finite candidate set, allocating failure probability across candidates preserves validity after selection; the power lower bound concerns each declared coarse–full comparison.

Relative deployment guarantee. Fix an observed arity K, a frozen baseline $g_{Q_K,K}$, and frozen $[0, 1]$-valued candidates $g_{q,K}$ for a finite order set $\mathcal{Q}_K$. Let $\mathcal{R}_{\star,K}(g) := \mathbb{E}_\star[(g - Y)^2 \mid K]$. Source paired losses, simultaneous confidence radii, and a declared shift allowance provide upper bounds $\widehat{\Delta}^+_{qK}$ on $\mathcal{R}_{\star,K}(g_{q,K}) - \mathcal{R}_{\star,K}(g_{Q_K,K})$ (Theorem E.5). For a declared margin $\gamma_K \geq 0$, select the minimizing order $\widehat{q}_K$ only if its bound is below $-\gamma_K$; otherwise set $\widehat{q}_K = Q_K$. On the simultaneous coverage event,

$$\mathcal{R}_{\star,K}(g_{\widehat{q}_K,K}) - \mathcal{R}_{\star,K}(g_{Q_K,K}) \leq -\gamma_K \mathbf{1}\{\widehat{q}_K \neq Q_K\}.$$

An exact source mixture needs zero allowance; total-variation distance ρ_K from one gives $2\rho_K$. When no valid source–target bound is available, four-window temporal selection makes an empirical comparison and replays the baseline on rejection (Appendix G.8).

Algorithm 1 CHORD: order selection around one frozen baseline

Require: Recomputable backbone; predeclared projection orders $q \le Q_K$ for each observed arity K; source moment orders ℓ; aggregate moments, support information, and simultaneous radii; requested claim type.

1: Predeclare data access, failure probability, margin $\gamma_K \ge 0$, and either a source–target risk inequality or chronological validation windows; train the richest view once and freeze every candidate and baseline $g_{Q_K,K}$.

2: Recompute history and candidate from each joint map T_q without a full-incidence cache; check support and moment validity, then choose the eligible source moment order $\widehat{\ell}(q,K)$.

3: For excess risk, minimize the simultaneous upper bound $\widehat{U}_{q,\ell}$ on Brier excess or ABSTAIN if no pair passes. In the quadratic acquisition model, use the weighted completion design and the energy test of Theorem 6.2 to certify the coarse-loss tolerance before selecting that view.

4: For paired risk, switch to the minimizer of the separately released upper differences $\widehat{\Delta}^{+}_{qK}$ only when its value is below $-\gamma_K$; otherwise return $g_{Q_K,K}$ unchanged.

5: For temporal validation, apply the fixed earlier-window rule and make only an empirical claim; if it selects no candidate, return $g_{Q_K,K}$.

8 EXPERIMENTS

Theory and frozen-predictor tests. Controlled studies test the exact reversal boundary and the distinct estimation and certification costs. Temporal studies test order selection around the same frozen predictor on the official AskUbuntu and Math StackExchange temporal hypergraphs (Benson et al., 2018). On each real stream, Deep Sets (Zaheer et al., 2017) and self-attention (Lee et al., 2019) scorers are trained on the first 40% of events using the complete-event view and then frozen. Here q denotes the largest subset order visible to both history and candidate; a q-view recomputes those projected statistics and removes the baseline's complete-event feature block. Thus even $q = K$ is a genuine projection for a candidate of arity K. The 40–60% interval supplies four chronological source blocks for rank screening and probability calibration, four fixed windows in 60–80% select q, and 80–100% is opened once for held-out evaluation. We report mean reciprocal rank (MRR), area under the precision–recall curve (AUPRC), Brier score, and minimum-window MRR.

Table 1: **Frozen baseline + CHORD on real streams.** For candidate arity $K \in \{2,3,4,5\}$, the full-view baseline retains complete-event statistics; the candidate $q \le K$ suppresses that block even when $q = K$. Each comparison shares checkpoint, candidates, calibration, and split. Mean reciprocal rank (MRR), area under the precision–recall curve (AUPRC), and Brier are macro-averaged over arity. MRR columns show mean ± standard deviation over three seeds; Δ gives paired means of CHORD minus baseline. Brier gain reverses the sign. Worst MRR is the minimum of window-level macro MRR across four test windows.

Target stream	Frozen scorer	Baseline MRR	+CHORD MRR	ΔMRR	ΔAUPRC	Brier gain	Δworst MRR
AskUbuntu	Deep Sets	0.735±0.006	0.739±0.009	+0.0038	+0.0022	+0.0007	+0.0040
AskUbuntu	Self-attention	0.715±0.007	0.719±0.010	+0.0046	+0.0067	+0.0005	+0.0059
Math StackExchange	Deep Sets	0.787±0.005	0.788±0.004	+0.0008	+0.0002	+0.0004	+0.0012
Math StackExchange	Self-attention	0.766±0.006	0.768±0.007	+0.0019	+0.0014	+0.0005	+0.0019

Figures 2–3 expose the geometry behind the controlled study: independent source equations contract the compatible mechanism set, whereas duplication leaves its unidentified direction unchanged.

The new fixed-budget distinction. In the Gaussian gradient channel at $d = 36$, increasing the centered environment count from 29 to 33 reduces estimation variance by 3.57 times and the quadratic-energy testing scale by 6.86 times at the same total budget. Across critical designs with $m = 4$ through 32, weighting keeps both inverse spectral sums at their predicted optimal orders; the unweighted design retains an extra certification factor. A separate 20 000-repetition test gives null acceptance 0.9845 at required power 0.90, with no false acceptances on the two tested unsafe energy levels. Appendix I.1 specifies the channel, design, covariance, and analytic calibration for these fixed-budget comparisons.

Held-out results and the comparison that matters. Across 12 seed runs and their 48 arity–seed decisions in Table 1, CHORD changes 7 decisions, one in each of seven runs. Every changed run

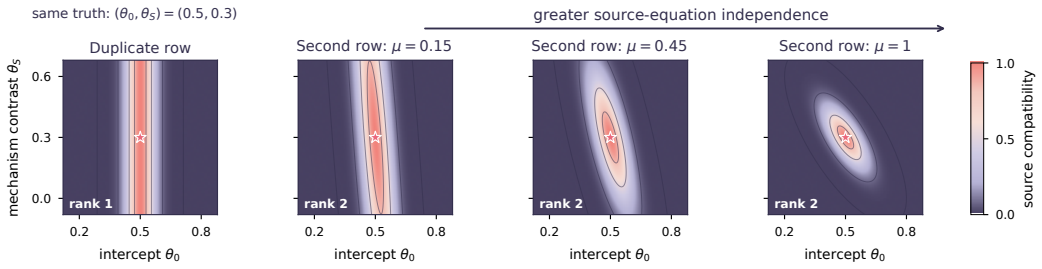

Figure 2: **Linearly independent environment equations contract the compatible mechanism set.** For binary orientation $S \in \{-1, 1\}$, the feature $\chi = (1, S)$ has intercept $\theta_0 = 0.5$ and contrast $\theta_S = 0.3$. Color shows the same exponential transform of squared moment residuals; the star marks the true pair. Duplicating a balanced row preserves the ridge; adding $(1, \mu)$, where $\mu = \mathbb{E}[S] > 0$, makes the two rows linearly independent, and larger μ improves conditioning of the source equations.

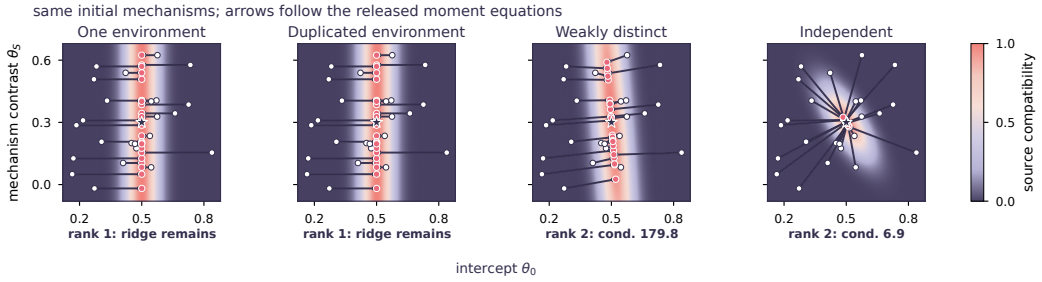

Figure 3: **Linearly independent equations identify the missing direction.** From the same 24 initial parameters (white), arrows follow 18 gradient steps on mean squared moment error to the coral endpoints; the star is the true intercept–contrast pair $(0.5, 0.3)$. Duplicate rows preserve a rank-one solution set. Adding a weakly distinct row identifies the mechanism with poor conditioning; adding a well-separated row concentrates updates at the truth.

improves all four reported held-out run-level metrics; the other 41 arity decisions replay the saved full-view predictions exactly. Every fixed $q \in \{1, 2, 3\}$ lowers mean MRR. Pooled temporal cross-validation produces 2 harmful seed-level changes among 12 runs; the four-window robust rule used by CHORD produces 0 and agrees with worst-window validation on this grid. The same wrapper gives positive mean MRR, Brier, and minimum-window MRR changes around empirical risk minimization, variance risk extrapolation, and group distributionally robust optimization (Group DRO) checkpoints on AskUbuntu, with exact fallback on Math StackExchange (Appendix I.1). Each pair shares its checkpoint, candidates, calibration, and split. A three-seed AskUbuntu/Deep Sets isotonic-calibration audit gives mean MRR gain 0.0020 and Brier-score reduction 0.00019 (Appendix I.1).

Theory-matched stress tests. For a reference score gap of 0.265 and total-variation radius $\rho_{\mathrm{fib}} = 0.20$, the worst target gap is 0.020 at likelihood-ratio cap $\kappa = 2.0$ and -0.015 at $\kappa = 4.0$, crossing the exact water-filling reversal boundary. Across 216 joint order settings, all 76 accepted actions satisfy the known mechanism and support conditions, with 76/76 excess-risk bound coverage. Appendix I gives the complete design, overlap sensitivity, coverage, calibration, and runtime results.

9 CONCLUSION

Representation order determines both what a predictor can distinguish and what its evidence must justify. We connect projection-induced reversal to a sharper statistical distinction: environments sufficient for optimal mechanism estimation can still be inefficient for certifying a coarse view. Matching learning and safety–power bounds identify two environment scales, and an explicit design attains both. Together with exact projection intervals and the CHORD wrapper, these results make environment collection part of representation design. Reliable higher-order learning depends on collecting the directions needed for the intended decision and its certificate.

REFERENCES

J. Alexander and A. Hirschowitz. Polynomial interpolation in several variables. *Journal of Algebraic Geometry*, 4(2):201–222, 1995.

Martin Arjovsky, Léon Bottou, Ishaan Gulrajani, and David Lopez-Paz. Invariant risk minimization. *arXiv preprint arXiv:1907.02893*, 2019. URL https://arxiv.org/abs/1907.02893.

Amir Asiaee, Samhita Pal, Cole Beck, and Jared Davis Huling. Sharp bounds for treatment effect generalization under outcome distribution shift. In *Proceedings of the Fifth Conference on Causal Learning and Reasoning*, volume 323 of *Proceedings of Machine Learning Research*, pp. 1572–1603. PMLR, 2026. URL https://proceedings.mlr.press/v323/asiaee26a.html.

Ali Behrouz, Farnoosh Hashemi, Sadaf Sadeghian, and Margo Seltzer. CAT-Walk: Inductive hypergraph learning via set walks. In *Advances in Neural Information Processing Systems*, volume 36, pp. 32636–32671, 2023. doi: 10.52202/075280-1416. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/6739d8df16b5bce3587ca5f18662a6aa-Abstract-Conference.html.

Austin R. Benson, Rediet Abebe, Michael T. Schaub, Ali Jadbabaie, and Jon Kleinberg. Simplicial closure and higher-order link prediction. *Proceedings of the National Academy of Sciences*, 115(48):E11221–E11230, 2018. doi: 10.1073/pnas.1800683115. URL https://doi.org/10.1073/pnas.1800683115.

Junhao Bian, Tao Zhou, and Yilin Bi. Unveiling the role of higher-order interactions via stepwise reduction. *Communications Physics*, 8:228, 2025. doi: 10.1038/s42005-025-02157-3.

Guy Bresler, Chenghao Guo, Yury Polyanskiy, and Andrew Yao. Partial and exact recovery of a random hypergraph from its graph projection. In *Proceedings of Thirty Eighth Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pp. 534–593. PMLR, 2025. URL https://proceedings.mlr.press/v291/bresler25a.html.

Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3, 1950.

T. Tony Cai and Anru Zhang. ROP: Matrix recovery via rank-one projections. *The Annals of Statistics*, 43(1):102–138, 2015. doi: 10.1214/14-AOS1267. URL https://doi.org/10.1214/14-AOS1267.

Peter G. Casazza, Eric Pinkham, and Brian Tuomanen. Riesz outer product hilbert space frames: Quantitative bounds, topological properties, and full geometric characterization. *Journal of Mathematical Analysis and Applications*, 441(1):475–498, 2016. doi: 10.1016/j.jmaa.2016.04.001. URL https://doi.org/10.1016/j.jmaa.2016.04.001.

Yikang Chen, Zhengkang Guan, Haoyuan Qian, Xingxuan Zhang, Peng Cui, Yi Yang, Fei Wu, and Kun Kuang. DAG-FM: A foundation model for causal discovery under heterogeneous causal mechanisms, 2026. URL https://arxiv.org/abs/2607.11510v2. Version 2.

Yongqiang Chen, Yonggang Zhang, Yatao Bian, Han Yang, Kaili Ma, Binghui Xie, Tongliang Liu, Bo Han, and James Cheng. Learning causally invariant representations for out-of-distribution generalization on graphs. In *Advances in Neural Information Processing Systems*, volume 35, pp. 22131–22148, 2022. doi: 10.52202/068431-1608. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/8b21a7ea42cbcd1c29a7a88c444cce45-Abstract-Conference.html.

Eli Chien, Chao Pan, Jianhao Peng, and Olgica Milenkovic. You are AllSet: A multiset function framework for hypergraph neural networks. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=hpBTIv2uy_E.

Jacob Dorn and Kevin Guo. Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. *Journal of the American Statistical Association*, 118 (544):2645–2657, 2023. doi: 10.1080/01621459.2022.2069572. URL https://doi.org/10.1080/01621459.2022.2069572.

Atticus Geiger, Duligur Ibeling, Amir Zur, Maheep Chaudhary, Sonakshi Chauhan, Jing Huang, Aryaman Arora, Zhengxuan Wu, Noah Goodman, Christopher Potts, and Thomas Icard. Causal abstraction: A theoretical foundation for mechanistic interpretability. *Journal of Machine Learning Research*, 26(83):1–64, 2025. URL https://www.jmlr.org/papers/v26/23-0058.html.

Yuzhou Gu and Yury Polyanskiy. Weak recovery threshold for the hypergraph stochastic block model. In *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pp. 885–920. PMLR, 2023. URL https://proceedings.mlr.press/v195/gu23b.html.

James A. Hanley and Barbara J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, 1982. doi: 10.1148/radiology.143.1.7063747.

Aapo Hyvärinen, Hiroaki Sasaki, and Richard Turner. Nonlinear ICA using auxiliary variables and generalized contrastive learning. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 859–868. PMLR, 2019. URL https://proceedings.mlr.press/v89/hyvarinen19a.html.

Saleem A. Kassam. Robust hypothesis testing for bounded classes of probability densities (corresp.). *IEEE Transactions on Information Theory*, 27(2):242–247, 1981. doi: 10.1109/TIT.1981.1056314.

Gholamreza B. Khosrovshahi and Behrouz Tayfeh-Rezaie. Trades and t-designs. In Sophie Huczynska, James D. Mitchell, and Colva M. Roney-Dougal (eds.), *Surveys in Combinatorics 2009*, volume 365 of *London Mathematical Society Lecture Note Series*, pp. 91–112. Cambridge University Press, 2009. doi: 10.1017/CBO9781107325975.005.

Jinwoo Kim, Saeyoon Oh, Sungjun Cho, and Seunghoon Hong. Equivariant hypergraph neural networks. In *Computer Vision – ECCV 2022*, volume 13681 of *Lecture Notes in Computer Science*, pp. 86–103. Springer, 2022. doi: 10.1007/978-3-031-19803-8_6. URL https://doi.org/10.1007/978-3-031-19803-8_6.

David Krueger, Ethan Caballero, Jörn-Henrik Jacobsen, Amy Zhang, Jonathan Binas, Dinghuai Zhang, Rémi Le Priol, and Aaron Courville. Out-of-distribution generalization via risk extrapolation (REx). In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 5815–5826. PMLR, 2021. URL https://proceedings.mlr.press/v139/krueger21a.html.

Daniel Kuhn, Soroosh Shafiee, and Wolfram Wiesemann. Distributionally robust optimization. *Acta Numerica*, 34:579–804, 2025. doi: 10.1017/S0962492924000084.

Juho Lee, Yoonho Lee, Jungtaek Kim, Adam Kosiorek, Seungjin Choi, and Yee Whye Teh. Set transformer: A framework for attention-based permutation-invariant neural networks. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 3744–3753. PMLR, 2019. URL https://proceedings.mlr.press/v97/lee19d.html.

Ming Li, Huiting Wang, Yuting Chen, Lu Bai, Lixin Cui, Feilong Cao, and Ke Lv. Self-supervised hypergraph learning with substructure awareness for hyperedge prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(27):23055–23062, 2026. doi: 10.1609/aaai.v40i27.39471. URL https://doi.org/10.1609/aaai.v40i27.39471.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=Bkg6RiCqY7.

Pascal Massart. The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *The Annals of Probability*, 18(3):1269–1283, 1990. doi: 10.1214/aop/1176990746. URL https://doi.org/10.1214/aop/1176990746.

Alexandru Niculescu-Mizil and Rich Caruana. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning*, 2005. URL https://www.cs.cornell.edu/~alexn/papers/calibration.icml05.crc.rev3.pdf.

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2 edition, 2009. doi: 10.1017/CBO9780511803161. URL https://doi.org/10.1017/CBO9780511803161.

Jonas Peters, Peter Bühlmann, and Nicolai Meinshausen. Causal inference by using invariant prediction: Identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(5):947–1012, 2016. doi: 10.1111/rssb.12167. URL https://doi.org/10.1111/rssb.12167.

Holger Rauhut and Ulrich Terstiege. Low-rank matrix recovery via rank one tight frame measurements. *Journal of Fourier Analysis and Applications*, 25(2):588–593, 2019. doi: 10.1007/s00041-017-9579-x. URL https://doi.org/10.1007/s00041-017-9579-x.

Paul K. Rubenstein, Sebastian Weichwald, Stephan Bongers, Joris M. Mooij, Dominik Janzing, Moritz Grosse-Wentrup, and Bernhard Schölkopf. Causal consistency of structural equation models. In *Proceedings of the 33rd Annual Conference on Uncertainty in Artificial Intelligence (UAI 2017)*. AUAI Press, 2017. URL https://auai.org/uai2017/proceedings/papers/11.pdf.

Shiori Sagawa, Pang Wei Koh, Tatsunori B. Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=ryxGuJrFvS.

Thomas Sauer and Yuan Xu. On multivariate hermite interpolation. *Advances in Computational Mathematics*, 4:207–259, 1995. doi: 10.1007/BF03177515.

Sandra Schluttenhofer and Jan Johannes. Adaptive minimax testing in inverse gaussian sequence space models, 2020. URL https://arxiv.org/abs/2002.07623.

Edward H. Simpson. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 13(2):238–241, 1951. doi: 10.1111/j.2517-6161.1951.tb00088.x. URL https://doi.org/10.1111/j.2517-6161.1951.tb00088.x.

Maximilian Soelch, Adnan Akhundov, Patrick van der Smagt, and Justin Bayer. On deep set learning and the choice of aggregations. In *Artificial Neural Networks and Machine Learning–ICANN 2019*, pp. 444–457, 2019. doi: 10.1007/978-3-030-30487-4_35. URL https://doi.org/10.1007/978-3-030-30487-4_35.

Henan Sun, Xunkai Li, Lei Zhu, Junyi Han, Guang Zeng, Ronghua Li, and Guoren Wang. Rethinking graph Out-of-Distribution generalization: A learnable random walk perspective, 2025. URL https://arxiv.org/abs/2505.05785.

Changlin Wan, Muhan Zhang, Pengtao Dang, Wei Hao, Sha Cao, Pan Li, and Chi Zhang. Ambiguities in neural-network-based hyperedge prediction. *Journal of Applied and Computational Topology*, 8(5):1333–1361, 2024. doi: 10.1007/s41468-024-00172-x. URL https://doi.org/10.1007/s41468-024-00172-x.

Yanbang Wang and Jon Kleinberg. From graphs to hypergraphs: Hypergraph projection and its reconstruction. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=qwYKE3VB2h.

Yifan Wang, Gonzalo R. Arce, and Guangmo Tong. Generalization performance of hypergraph neural networks. In *Proceedings of the ACM on Web Conference 2025*, pp. 1273–1291, 2025. doi: 10.1145/3696410.3714586. URL https://doi.org/10.1145/3696410.3714586.

Naganand Yadati. Oversquashing in hypergraph neural networks. In *Proceedings of the Third Learning on Graphs Conference*, volume 269 of *Proceedings of Machine Learning Research*, pp. 30:1–30:16. PMLR, 2025. URL https://proceedings.mlr.press/v269/yadati25a.html.

Mingzhang Yin, Yixin Wang, and David M. Blei. Optimization-based causal estimation from heterogeneous environments. *Journal of Machine Learning Research*, 25(168):1–44, 2024. URL https://www.jmlr.org/papers/v25/21-1028.html.

Se-eun Yoon, Hyungseok Song, Kijung Shin, and Yung Yi. How much and when do we need higher-order information in hypergraphs? a case study on hyperedge prediction. In *Proceedings of The Web Conference 2020*, pp. 2627–2633, 2020. doi: 10.1145/3366423.3380016. URL https://doi.org/10.1145/3366423.3380016.

Jean-Gabriel Young, Giovanni Petri, and Tiago P. Peixoto. Hypergraph reconstruction from network data. *Communications Physics*, 4(1):135, 2021. doi: 10.1038/s42005-021-00637-w. URL https://doi.org/10.1038/s42005-021-00637-w.

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabás Póczos, Ruslan Salakhutdinov, and Alexander J. Smola. Deep sets. In *Advances in Neural Information Processing Systems*, volume 30, pp. 3391–3401, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/f22e4747da1aa27e363d86d40ff442fe-Abstract.html.

Ruochi Zhang, Yuesong Zou, and Jian Ma. Hyper-SAGNN: A self-attention based graph neural network for hypergraphs. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=ryeHuJBtPH.

Songming Zhang, Yuxiao Luo, Qizhou Wang, Haoang Chi, Weikai Li, Bo Han, and Jinyan Li. Are all unseen data Out-of-Distribution?, 2024. URL https://arxiv.org/abs/2312.16243v2. Version 2.

Yangze Zhou, Gitta Kutyniok, and Bruno Ribeiro. OOD link prediction generalization capabilities of message-passing GNNs in larger test graphs. In *Advances in Neural Information Processing Systems*, volume 35, pp. 20257–20272, 2022. doi: 10.52202/068431-1473. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/7f88a8478c4ae97819ccffa1e80e7a7b-Abstract-Conference.html.

Appendix

Contents

A Mathematical notation **15**

B Projection reversal and octahedral trades **18**

B.2 Exact score range, reversal, and minimax risk 19

B.7 General fibers and quantile sampling error 22

C Fixed-budget learning and environment geometry **23**

D Safety, switching power, and matching lower bounds **29**

E Heterogeneous-environment identification and CHORD **33**

E.1 Mixture-weight identification and target transport 33

E.2 The aggregate source moment map 34

E.4 Projection loss and unidentified source directions 38

E.7 Generic quadratic rank deficiency 41

F Set-size pooling and pure high-order testing **46**

F.2 Set-size erasure and product pooling 46

F.3 Pure r-way dependence experiment 47

G Finite-sample guarantees and frozen-baseline CHORD **52**

G.7 Selecting projection and source-moment orders 59

G.8 Paired Brier comparison with a frozen baseline 59

G.9 Temporal validation on natural streams 61

H Experimental protocols and implementation **63**

I Additional quantitative evidence **66**

A MATHEMATICAL NOTATION

Table 2 collects the notation used in the main results and their certificates. A symbol that is reused locally is explicitly scoped to the corresponding construction.

Table 2: Notation, grouped by semantic role, with domains and theorem-specific meanings.

Symbol	Meaning
Indices, environments, and the two selected orders	
$e \in [E], \star$	Source-environment index and target-environment marker; P_e and $\mathbb{E}_e$ denote probability and expectation in environment e.
r	Uniformity of the rooted hypergraph context or theoretical construction.
q	Observed projection order: the target scorer observes $Z_q = \Pi_q(H)$.
ℓ	Source moment order: source sites may form moment features $\psi_{q,\ell}(H)$ using structure up to order ℓ.
K	Observable candidate size (arity).
p, p_q, d_q	Dimensions of the full-state feature map, direct-factorization feature map, and data-setting-specific operator columns, respectively.
$s_{q,\ell}$	Dimension of the source moment feature $\psi_{q,\ell}$.
$m_{q,\ell}$	Number of rows in the stacked source moment operator $A_{q,\ell} \in \mathbb{R}^{m_{q,\ell} \times d_q}$.
N, n, M	Context-dependent size parameters: ambient vertices, observations or outside nodes, and planted trade cells, respectively.
Hypergraph states, projections, and fibers	
H, Y	Complete rooted historical state after withholding the queried edge, and its future binary label.
$\Pi_q,\ Z_q = \Pi_q(H)$	Order-q projection and the corresponding deployed observation.
$\mathcal{Z}_q^{\mathrm{cov}}$	Target values covered by the declared conditional-feature-mean map and support certificates.
$\mathcal{F}_z = \Pi^{-1}(z)$	Projection fiber: full states that share observed value z.
$P_Z,\ \mu_z,\ \nu_z$	Projected marginal, reference conditional law on $\mathcal{F}_z$, and a shifted conditional law on that fiber.
$V_r,\ v_i^b,\ E_x$	Paired vertex set $V_r = \{v_i^0, v_i^1 : i \in [r]\}$ and transversal $E_x = \{v_i^{x_i} : i \in [r]\}$ in the octahedral trade.
$\mathcal{H}_+, \mathcal{H}_-$	Edge-disjoint even- and odd-parity orientations of the octahedral trade.
$C_q(G)_F$	Complete labelled q-shadow count of q-set F in r-graph G.
$S \in \{-1, +1\}$	Full-order orientation of the trade; no proper shadow reveals S.
$G^{(r)}(N, p_{\mathrm{bg}})$	Independent r-uniform background hypergraph in the overlap-stress construction.
n_r	Number of trade cells retained by coupled whole-cell thinning; ρ denotes the retention probability in this local overlap statement.
Stable mechanisms and environment mixing	
$\varepsilon,\ \eta(H)$	Independent $\mathrm{Uniform}(0,1)$ noise and the invariant conditional mean in $Y = \mathbf{1}\{\varepsilon \le \eta(H)\}$.
$\bar{\eta}$	Projected mechanism satisfying $\eta(h) = \bar{\eta}\{\Pi(h)\}$ when the outcome mechanism factors through the projection (causal lumpability).
$\chi(H),\ \theta,\ \Theta,\ B$	Full-state feature map, its coefficient, declared parameter set, and radius with $\eta(H) = \chi(H)^\top \theta$ and $\|\theta\|_2 \le B$.
$\phi_q(Z_q),\ \theta_q$	Direct-factorization feature map and coefficient when $\eta(H) = \phi_q(Z_q)^\top \theta_q$.
$\chi_q, \widetilde{\chi}_q, \widetilde{\theta}_q$	Direct feature map $\chi_q = \phi_q \circ \Pi_q$ and the data-setting-specific feature map/parameter used in aggregate moment equations.
$a_{\star,q}(z)$	Target conditional feature mean $\mathbb{E}_\star[\chi(H) \mid Z_q = z]$; in the direct-factorization setting it equals $\phi_q(z)$.
$m_{\star,q}(z)$	Target projected Bayes mean $a_{\star,q}(z)^\top \theta$.
$k_j,\ k$	Stable full-state outcome probabilities $k_j = P(Y = 1 \mid H = h_j)$ and their vector over one fiber.
$W_q,\ m_q$	Environment-by-state mixture matrix and projected label means, related by $m_q = W_q k$.
$w_\star$	Target mixture row; the corresponding target outcome probability is $w_\star^\top k$.
w_e, c, γ, a	Trade orientation weight, score center, mechanism contrast, and symmetric source–target mixture separation.
$\mathsf{D}_q, \mathsf{T}_q$	Target-agnostic factorized and target-adaptive data-access settings, respectively.
Within-projection shifts and attainable score ranges	
$\mathrm{TV}(\nu, \mu)$	Total-variation distance.

Continued on the next page

Table 2 continued

Symbol	Meaning
$\rho_{\mathrm{fib}}, \kappa$	Total-variation radius and two-sided likelihood-ratio bound for within-fiber shift.
$\mathcal{U}_z(\rho_{\mathrm{fib}}, \kappa)$	Conditional laws on $\mathcal{F}_z$ satisfying the declared total-variation and likelihood-ratio constraints.
$m_\nu(z), m_0(z)$	Projected Bayes scores under a shifted and the reference fiber law.
ω_z^+, ω_z^-	Exact maximum upward and downward changes of the projected score on fiber z.
$D_z = \omega_z^+ + \omega_z^-$	Diameter of the attainable projected-score interval.
ϵ_z	A declared or simultaneously certified upper bound on D_z used for quantitative within-fiber stability.
I_z	Attainable score interval $[m_0(z) - \omega_z^-, m_0(z) + \omega_z^+]$.
$\mathrm{osc}_z(\eta)$	Range of full-state outcome probabilities within fiber z.
c_i^-, c_i^+	Removable and addable probability-mass capacities in the capped water-filling program.
$G_{\mathrm{lo}}, G_{\mathrm{hi}}, \bar{\rho}_{\mathrm{fib}}$	Cumulative removal cost, cumulative addition gain, and the feasible transported mass in the sorted water-filling solution.
τ (reversal theorem)	Decision threshold whose Bayes classification can reverse under an admissible within-fiber shift.
δ_{mech}	Uniform bound on a declared additive change of the conditional outcome mean in Appendix B.3.
Aggregate moments, rank, and nullspace geometry	
$\psi_{q,\ell}(H)$	Outcome-free source moment feature used to form order-pair (q, ℓ) moment equations.
$A^e_{q,\ell}, b^e_{q,\ell}$	Environment-e population operator and response moment; under an exact data-access setting $b^e_{q,\ell} = A^e_{q,\ell}\widetilde{\theta}_q$.
$A_{q,\ell}, b_{q,\ell}$	Operators and response moments stacked across source environments.
$\widehat{A}_{q,\ell}, \widehat{b}_{q,\ell}$	Released or cross-fitted estimates of the stacked moments.
$\mathrm{rank}(A), \mathrm{Row}(A), \ker(A)$	Matrix rank, row space, and nullspace.
$A^\dagger$	Moore–Penrose pseudoinverse.
a_q, M_q	Target conditional feature mean $a_q = \mathbb{E}_\star[\chi(H) \mid \sigma(Z_q)]$ and its second moment $M_q = \mathbb{E}_\star[a_q a_q^\top]$.
$\mathcal{N}, P_{\mathcal{N}}$	Fixed source nullspace $\mathcal{N} = \ker(A)$ in the refinement theorem and its orthogonal projector.
R	Radius of the unidentified nullspace component in ambiguity and minimax statements.
$\lambda_\star, L_{q,\ell}$	Linear weights $\lambda_\star = (A^\dagger_{q,\ell})^\top a_{\star,q}$ and the associated inverse-design variance factor.
$d_{q,\ell}(a)$	Distance of target conditional feature mean a from $\mathrm{Row}(A_{q,\ell})$, equivalently its projected nullspace norm.
$a_\parallel, a_\perp$	Row-space and nullspace components of a target conditional feature mean in the Gaussian minimax law.
$\sigma, n^{\mathrm{src}}_{q,\ell}$	Gaussian release-noise scale and effective source sample size for pair (q, ℓ).
$\widehat{s}_{q,\ell}, \widehat{C}_{q,\ell}$	Certified smallest singular value and the resulting parameter-error radius $C_{q,\ell} = \eta_{q,\ell}/\widehat{s}_{q,\ell}$.
$\widehat{\mathfrak{E}}_{q,\ell}$	Confidence equivalence set of bounded coefficients compatible with the released moment equation.
$\underline{T}_{q,\ell}, \overline{T}_{q,\ell}$	Certified lower and upper target-score endpoints over the equivalence set.
Generic rank threshold and feasible order pairs	
d, E	Ambient mechanism dimension and number of source environments in the generic quadratic rank result.
$f(x), c, b, Q$	Quadratic response, intercept, linear coefficient, and zero-diagonal symmetric quadratic matrix in the generic-rank construction.
$\mathcal{J}_{d,E}$	Linear map taking the coefficients of a d-variable multiaffine quadratic to its values and gradients at E environment means.
$m = \max\{d - E + 1, 0\}$	Residual dimension appearing in the exact generic quadratic nullity formula.
$E^{\mathrm{gen}}_{\mathrm{min}}$	Sharp minimum number of generic environments for full quadratic rank.
$q^\star_{\mathrm{ID}}, q_{\mathrm{eff}}$	Smallest order through which the mechanism factors, and largest order passing the declared target-support check.
$\ell^{\mathrm{dir}}_{\mathrm{id}}(q; E)$	Least source moment order giving full column rank for the direct-factorization feature map at projection order q.
$\ell_{\max}(q)$	Largest source moment order estimable for projection order q.

Continued on the next page

Table 2 continued

Symbol	Meaning
$\mathcal{W}_E^{\mathrm{dir}}$	Feasible pairs of projection and source moment orders in the direct-factorization setting.
Risk, support, and finite-sample certificates	
$\mathcal{R}_\star(g_q), \mathcal{R}^*_{\star,H}$	Target Brier risk of an order-q predictor and the full-state Bayes Brier risk.
$V_{\star,q}, v_{\star,q}(z)$	Average and cellwise information lost by projection, $\mathbb{E}_\star \mathrm{Var}\{\eta(H) \mid Z_q\}$ and $\mathrm{Var}_\star\{\eta(H) \mid Z_q = z\}$.
$G_{\star,q}, \widehat{G}_q, \overline{G}_q$	Target conditional-covariance operator, its cross-fitted estimate, and its positive-semidefinite upper certificate.
$\overline{V}_{q,\ell}$	Certified upper bound on information lost by projection for order pair (q,ℓ).
$\epsilon_{A,q,\ell}, \epsilon_{b,q,\ell}, \epsilon_{a,q}$	Simultaneous operator, response-moment, and target conditional-feature-mean radii.
$\eta_{q,\ell}$	Source release radius under an exact data-access setting, $\epsilon_{b,q,\ell} + B\epsilon_{A,q,\ell}$.
$\beta^{\mathrm{mom}}_{q,\ell}, \eta^{\mathrm{ctr}}_{q,\ell}$	One-sided moment-defect bound and the data-setting-specific moment radius used by the approximate direct setting.
$\underline{n}^{\mathrm{src}}_{q,\ell}, \underline{n}^{\mathrm{tar}}_{q}$	Certified effective source-release count and target-cell support count.
$C_{\star,q}(F), \mathcal{F}_{\star,q}$	Number of observable extensions of a target q-cell F, and the declared collection of target q-cells.
$\overline{\pi}^{\mathrm{hole}}_q, \epsilon^{\mathrm{rare}}_q$	Upper bound on unsupported target mass and the declared rare-cell radius.
Δ^{sup}_q	Brier-risk allowance for unsupported and rare target cells.
$\vartheta^{\mathrm{rep}}_{q,\ell}, u_{q,\ell}$	Frozen approximate-direct coefficient and its representation residual $u_{q,\ell} = \eta - \widetilde{\chi}_q^\top \vartheta^{\mathrm{rep}}_{q,\ell}$.
$\overline{\mathcal{E}}^{\mathrm{src}}_{q,\ell}$	One-sided source upper certificate for the representation residual.
$\kappa^{\mathrm{rep}}_q, \tau^{\mathrm{rep}}_q$	Constants in the assumed upper bound relating target representation error to its source value.
$\mathrm{repUCB}_{q,\ell}$	Upper confidence bound on target representation error.
$\delta^{\mathrm{rep}}_{q,\ell}$	Predeclared tolerance for the approximate-direct representation-error check.
$D_{q,\ell}(z), \Gamma^{\mathrm{stab}}_{q,\ell}(z)$	Half-width of the certified target-score interval and any separately charged target-bias allowance.
$\overline{T}_{q,\ell}$	Squared score-error bound on supported target values.
$\mathcal{B}_{q,\ell}$	Data-setting-specific risk bound on supported target values.
$\widehat{U}_{q,\ell}$	Complete certified excess-risk upper bound, $\mathcal{B}_{q,\ell} + \Delta^{\mathrm{sup}}_q$.
$\mathsf{C}_q, \mathsf{S}^{\mathrm{tar}}_q, \mathsf{S}^{\mathrm{src}}_{q,\ell}$	Data-setting, target-support, and source-support checks.
$\mathsf{G}_{q,\ell}, \mathsf{M}^{\mathrm{val}}_{q,\ell}, \mathsf{I}_{q,\ell}$	Representation-error check, independent source-validation moment check used only for falsification, and the requested identification certificate.
$\mathcal{C}$	Set of order pairs passing every required check and the declared risk threshold.
$\delta, \eta^2_{\mathrm{risk}}$ (certificates)	Simultaneous failure probability and the predeclared admissible-risk threshold.
$\widehat{\ell}(q), \widehat{q}$	Selected source moment order at fixed q and the final selected projection order.
λ_ℓ, λ_q	Weights on source-moment and projection costs in the certified selection rule.
Fixed-budget mechanism learning and safe coarsening	
d, $X \in \{-1,1\}^d$	Number of labelled trade cells and their hidden orientations; the target law is uniform on the sign cube.
b, Q	Linear coefficients in $\mathbb{R}^d$ and a symmetric $d \times d$ interaction matrix with zero diagonal.
ϑ, V, Θ_R	Nonconstant coefficients $(b, \{Q_{ij}\}_{i<j})$, omitted energy $V = \|b\|_2^2 + \|Q\|_F^2/2$, and the ball $V \le R^2$.
x_e, α, σ	Known environment mean, common bound $\|x_e\|_2 = \alpha \le 1/2$, and Gaussian acquisition-noise scale.
n_e, N	Per-environment precision budget and total budget $N = \sum_e n_e$ in the Gaussian population-moment channel.
$\mathfrak{D}_{d,m}(\alpha)$	Centered radius-α regular-simplex designs with equal allocations and mean-span dimension $k = d - m$.
m_d, m, k, E	$m_d = \lfloor(\sqrt{8d+1}-1)/2\rfloor$; unspanned dimension $0 \le m \le m_d$; $k = d-m$; environment count $E = k+1$.
$\mathcal{V}$, V_{env}, U	Environment mean subspace, its orthonormal basis matrix, and an orthonormal basis of its complement.
$s^2_N = v_N$	Contrast-channel noise variance $\sigma^2(d-m)/(N\alpha^2)$.
D_U, a_j, t	Diagonal completion map $B \mapsto \mathrm{diag}(UBU^\top)$ on symmetric $m \times m$ matrices, its singular values, and its domain dimension $t = m(m+1)/2$.

Continued on the next page

Table 2 continued

Symbol	Meaning
G_w, w	Raw outer-product frame and its coordinate-row weight; the balanced construction uses rows $\sqrt{w}\, e_i$ with $w = m^{1/4}$.
Σ, F, M	Gaussian coefficient-estimation covariance, its Frobenius norm, and its operator norm in the certification lemma.
ε, δ, β, C	Coarse-loss tolerance, false certification probability, null rejection probability, and binary acceptance rule.
$r_{d,m,N}$	Optimal certification scale $\sigma^2(d-m)\sqrt{d^2+m^6}/(N\alpha^2)$ over the declared design class.
General-fiber quantile intervals	
F, Q_F, m_F	Reference distribution of the bounded full-state score, its quantile function, and its mean.
$L(F)$, $U(F)$	Smallest and largest score attainable under the fixed total-variation and finite likelihood-ratio envelope.
$W_1(F,G)$, h	First Wasserstein distance between score distributions; uniform score-approximation or quantization error.
Pure r-way dependence, sparse additive model, and set size	
$X_j = (X_{j1}, \ldots, X_{jr})$	Replacement-edge vector contributed by outside node j in the pure r-way construction.
P_y, P_0, τ, ρ	Label-conditional law, independent Bernoulli base law, interaction strength, and edge probability; here P_0 and ρ are scoped to the pure r-way model.
$Y, y \in \{-1, +1\}$	Balanced label coding used locally in the pure r-way and set-size constructions.
$R_{B,n}$, $R_{M,n}$	Equal-prior Bayes testing error and worst-class minimax testing error for the complete array $\{X_j\}_{j=1}^n$.
I_n	Pure r-way information scale $n\tau^2\{\rho(1-\rho)\}^r$.
W	Number of all-one rows, $\sum_{j=1}^n \prod_{i=1}^r X_{ji}$.
α	Sparsity exponent, with $\rho = n^{-\alpha}$ or $\rho_N = N^{-\alpha}$ in the two sparse constructions.
Z_i, A_C, ρ_N, β	Latent Rademacher sign, hyperedge indicator, sparse edge rate, and additive signal strength for candidate-face-incidence-restricted versus degree-based separation.
c_r	Exact set-size-dependent constant $\binom{2r}{r}/(2\,4^r)$ in the degree-based area under the receiver operating characteristic curve (AUROC) gap.
π	Positive-class prior in the set-size construction.
ε (set size)	Member error probability $P(X_i \neq Y \mid Y)$; this local use is distinct from the uniform noise in the main outcome equation.
X_i, M_K	Member observation; $M_K := K^{-1}\sum_i X_i$ is its mean-pooled representation.
b, λ	Prior log-odds and member log-likelihood ratio, $b = \log\{\pi/(1-\pi)\}$ and $\lambda = \log\{(1-\varepsilon)/\varepsilon\}$.
$\sigma(\cdot)$	Logistic sigmoid.
T_K	Product-pooled statistic $\prod_{i=1}^K X_i$.
$\eta_K(m)$	Set-size-conditioned Bayes posterior at pooled value m.

B Proofs for Score Reversal under Projection and the Octahedral-Trade Construction

B.1 Information restriction and causal interpretation

For any fixed target distribution and proper loss, the predictors measurable with respect to $Z = \Pi(H)$ form a subset of those measurable with respect to H. Hence no result in the paper says that projection beats the Bayes rule given the complete state H on one known target. The comparison is between source-trained or projection-restricted procedures across a declared intervention class.

Under (1), a soft intervention changes only the law of H. If $\eta = \bar{\eta} \circ \Pi$, marginalizing H within any fiber gives $P(Y = 1 \mid Z = z) = \bar{\eta}(z)$ under every such intervention. Conversely, if $\eta(h_1) \neq \eta(h_2)$ in one fiber, a sufficiently small positive-mass transfer between h_1 and h_2 changes the projected conditional. This proves the intervention statement in Definition 4.1.

B.2 Proof of Theorem 4.2

Fix a fiber and abbreviate $\mu = \mu_z$, $\eta_h = \eta(h)$, and $m(\nu) = \sum_h \nu_h \eta_h$. The feasible set $\mathcal{U}_z(\rho_{\mathrm{fib}}, \kappa)$ is a nonempty compact convex polytope: it is the intersection of the probability simplex with linear TV and likelihood-ratio constraints. A linear functional maps a compact convex set in $\mathbb{R}^{|\mathcal{F}_z|}$ to a compact interval. Its endpoints are exactly $m_0 - \omega_z^-$ and $m_0 + \omega_z^+$, proving (10).

To obtain (9), write $r_i = (\mu_i - \nu_i)_+$ and $s_i = (\nu_i - \mu_i)_+$. Then $\sum_i r_i = \sum_i s_i = t = \mathrm{TV}(\nu, \mu)$,

$$0 \le r_i \le \mu_i(1 - \kappa^{-1}) = c_i^-, \qquad 0 \le s_i \le \mu_i(\kappa - 1) = c_i^+, \tag{27}$$

and the score increase is $\sum_i s_i \eta_i - \sum_i r_i \eta_i$. For fixed t, an exchange argument is exact: if removal assigns positive mass to a larger η_j while a smaller η_i has unused capacity, moving the minimum of those two amounts from j to i weakly decreases removal cost. Iterating fills c_i^- in ascending η_i order. The symmetric exchange fills c_i^+ in descending order for additions. Conversely these greedy vectors satisfy all coordinate bounds and have equal total mass, so $\nu = \mu - r + s$ is a probability vector satisfying the likelihood-ratio band. If removal and addition overlap at one coordinate, cancelling their common part leaves ν and the objective unchanged; hence $\mathrm{TV}(\nu, \mu) \le t \le \rho_{\mathrm{fib}}$. The constructed law is feasible. Optimizing their piecewise-linear difference over $0 \le t \le \bar{\rho}_{\mathrm{fib}}$ proves the formula. Its slope changes only when a removal or addition capacity is exhausted, so a maximizer lies at a listed breakpoint or at $\bar{\rho}_{\mathrm{fib}}$. Replacing η by $-\eta$ gives the downward formula.

Let $a = \min_h \eta_h$, $b = \max_h \eta_h$, and write $\Delta_h = \nu_h - \mu_h$. Since $\sum_h \Delta_h = 0$, the positive and negative parts of Δ each have mass $\mathrm{TV}(\nu, \mu)$. Therefore

$$m(\nu) - m(\mu) = \sum_h \Delta_h(\eta_h - a) \le (b - a) \sum_{\Delta_h > 0} \Delta_h \le \rho_{\mathrm{fib}}(b - a), \tag{28}$$

Applying the argument to $-\eta$ proves the lower bound in (11).

If $a < b$, choose positive-reference-mass states h_- and h_+ attaining the minimum and maximum on the relevant fiber. For sufficiently small $t > 0$,

$$\nu^{\uparrow} = \mu + t(e_{h_+} - e_{h_-}) \tag{29}$$

obeys positivity, the TV constraint, and both likelihood-ratio inequalities. It has TV t and increases the score by $t(b - a)$. Reversing the transfer attains the downward bound. This proves local sharpness. It also shows that $D_z > 0$ whenever the mechanism is nonconstant. The reverse implication is immediate, completing the lumpability equivalence.

For two projected values, independent fiber shifts can simultaneously attain the lower endpoint at z_1 and upper endpoint at z_0. Their difference is

$$m_0(z_1) - m_0(z_0) - \omega_{z_1}^- - \omega_{z_0}^+. \tag{30}$$

It is negative exactly under (12); equality yields a tie. Starting from $m_0(z) > \tau$, the threshold statement follows by the same endpoint argument. Under $\mathbf{1}\{m > \tau\}$, attaining the threshold already changes the binary output, so that endpoint condition is weak.

Finally, conditional Brier regret relative to the target's Z-Bayes rule at z is $(g(z) - m_\nu(z))^2$. Over an interval $[L_z, U_z]$, its smallest worst-case value is $(U_z - L_z)^2/4$, uniquely attained at the midpoint. The uncertainty set is a product across fibers and P_Z is fixed, so the coordinatewise midpoints can be chosen simultaneously. Averaging proves (13).

If a declared tolerance $\epsilon_z \ge 0$ satisfies $D_z \le \epsilon_z$, the midpoint has worst-case conditional Brier regret at most $\epsilon_z^2/4$. A strict reference ranking is preserved throughout the shift class whenever

$$m_0(z_1) - m_0(z_0) > \omega_{z_1}^- + \omega_{z_0}^+. \tag{31}$$

The same argument covers a bounded real response under squared loss because the conditional excess risk of a scalar prediction g over the conditional mean m is exactly $(g - m)^2$. For a C-class response, encode the label by its one-hot vector and write the stable probability vector as $\eta(h) \in \Delta^{C-1}$. The attainable set on fiber z is the compact convex image

$$\mathcal{K}_z = \left\{ \sum_h \nu_z(h)\eta(h) : \nu_z \in \mathcal{U}_z \right\} \subseteq \Delta^{C-1}.$$

Multiclass Brier excess equals $\|g-u\|_2^2$ at target probability vector u. Consequently its conditional minimax value is $\min_{g\in\Delta^{C-1}}\max_{u\in\mathcal{K}_z}\|g-u\|_2^2$, the squared Chebyshev radius of $\mathcal{K}_z$. A smallest enclosing-ball center may be chosen in the convex hull of its contact points and hence in the simplex. This reduces to the scalar midpoint formula only when $\mathcal{K}_z$ lies on a line segment; no unspecified multiclass ordering is needed for the stated outcome-scope result.

B.3 A BOUNDED VIOLATION OF MECHANISM STABILITY

Let $\eta_0 : \mathcal{H} \to [0,1]$ be the reference conditional mean and fix a declared mechanism-drift radius $\delta_{\mathrm{mech}} \in [0,1]$. Consider exactly the drift class

$$\mathcal{D}_{\delta_{\mathrm{mech}}}(\eta_0) := \{d : \ |d(h)| \le \delta_{\mathrm{mech}}, \quad 0 \le \eta_0(h) + d(h) \le 1 \text{ for every } h\}.$$

For the reference attainable interval $I_z = [L_z, U_z]$ from Theorem 4.2, the exact endpoints after allowing both $\nu_z \in \mathcal{U}_z$ and $d \in \mathcal{D}_{\delta_{\mathrm{mech}}}(\eta_0)$ are

$$L_z^{\mathrm{drift}} = \min_{\nu_z\in\mathcal{U}_z} \sum_h \nu_z(h)\max\{0, \eta_0(h) - \delta_{\mathrm{mech}}\},$$
$$U_z^{\mathrm{drift}} = \max_{\nu_z\in\mathcal{U}_z} \sum_h \nu_z(h)\min\{1, \eta_0(h) + \delta_{\mathrm{mech}}\}. \tag{32}$$

For fixed ν_z, coordinatewise minimization and maximization over d give the two clipped integrands. On the finite fiber, the product of the compact convex sets $\mathcal{U}_z$ and the restriction of $\mathcal{D}_{\delta_{\mathrm{mech}}}(\eta_0)$ is connected, and the projected mean is continuous on it; its image in $\mathbb{R}$ is therefore the interval with these endpoints. Consequently the new attainable interval is contained in the clipped δ_{mech}-inflation

$$[L_z^{\mathrm{drift}}, U_z^{\mathrm{drift}}] \subseteq [\max\{0, L_z - \delta_{\mathrm{mech}}\}, \min\{1, U_z + \delta_{\mathrm{mech}}\}]. \tag{33}$$

Its exact minimax score radius is $(U_z^{\mathrm{drift}} - L_z^{\mathrm{drift}})/2$, which is at most $(U_z - L_z)/2 + \delta_{\mathrm{mech}}$. If every relevant reference mean lies in $[\delta_{\mathrm{mech}}, 1-\delta_{\mathrm{mech}}]$, the constant perturbations $d \equiv \pm\delta_{\mathrm{mech}}$ are feasible, and the untruncated inflation is attained at both endpoints. Equation (32), instead of the outer envelope, remains exact when probability boundaries bind.

For two frozen predictions $g, f \in [0,1]$ at the same state, their conditional Brier-risk difference at outcome mean η is $g^2 - f^2 - 2\eta(g-f)$. Replacing η_0 by $\eta_0 + d$ changes this difference by $-2d(g-f)$, whose absolute value is at most $2\delta_{\mathrm{mech}}$. Thus any valid paired source–target allowance Γ^{shift} under η_0 remains valid under the declared drift class after replacing it by $\Gamma^{\mathrm{shift}} + 2\delta_{\mathrm{mech}}$. These statements are sharp for the declared uniform class whenever the endpoint perturbation and $|g-f| = 1$ are feasible. An unrestricted change of the outcome mechanism has no nontrivial guarantee from the source release.

B.4 PROOFS FOR THE OCTAHEDRAL TRADE

Proof of Theorem 4.3. If a q-set F contains both v_i^0 and v_i^1, it belongs to no transversal, so both counts are zero. Otherwise F fixes q bits in distinct coordinates. Since $q < r$, at least one bit remains free. Exactly half of the 2^{r-q} completions have even parity and half have odd parity, giving 2^{r-q-1} extensions in each orientation. Hence

$$C_q(\mathcal{H}_+) = C_q(\mathcal{H}_-), \qquad C_q(\mathcal{H}_\pm)_F = \begin{cases} 2^{r-q-1}, & F \text{ uses at most one vertex from each pair,} \\ 0, & \text{otherwise,} \end{cases} \tag{34}$$

At $q = r$, every transversal has one parity and lies in exactly one edge system; together with the proper-order count above, this proves the count identity, disjointness, and full-order separation. □

For the symmetric environments, where $0 < a < 1$, the fiber distributions are

$$\mu_+ = \left(\frac{1+a}{2}, \frac{1-a}{2}\right), \qquad \mu_- = \left(\frac{1-a}{2}, \frac{1+a}{2}\right). \tag{35}$$

Their TV distance is a, and their two coordinate likelihood ratios are $(1-a)/(1+a)$ and $(1+a)/(1-a)$. The two mechanism values are $c \pm \gamma$, so their oscillation is 2γ and direct multiplication gives (16). Relative to μ_+, any target in the TV-and-overlap class has plus weight

$$u \ge \max\left\{\frac{1+a}{2} - a, \frac{1+a}{2\kappa}\right\} = \frac{1-a}{2}. \tag{36}$$

Thus μ_- attains the smallest feasible score and the downward modulus.

For the restricted two-environment class, write a projected prediction as $g = c+d$. Its endpoint regrets are $(d - a\gamma)^2$ and $(d + a\gamma)^2$. Their maximum is minimized at $d = 0$ with value $a^2\gamma^2$. Transferring one endpoint score to the other gives $4a^2\gamma^2$. More generally, let fiber j have mass λ_j, score center c_j, endpoint half-gap $a_j\gamma_j$, and prediction $g_j = c_j + d_j$. Averaging the two environment risks gives

$$\frac{\mathcal{R}_+(g) + \mathcal{R}_-(g)}{2} = \sum_j \lambda_j (d_j^2 + a_j^2\gamma_j^2), \tag{37}$$

which proves the aggregate minimax $\sum_j \lambda_j a_j^2\gamma_j^2$ and direct-transfer regret four times as large. These identities concern the two endpoints; for a larger uncertainty ball, Theorem 4.2 supplies its exact attainable interval and the corresponding minimax radius.

B.5 Identification from environment-specific mixture weights

For $p = (p_+, p_-)^\top$, projected environment means obey

$$m = Wp, \qquad W[e,:] = (w_e, 1 - w_e). \tag{38}$$

With one environment, $(1 - w, -w)^\top$ spans $\ker W$. Thus $p \pm t(1 - w, -w)^\top$ are observationally identical in the source for small t but have target means separated by $2|t(w_\star - w)|$. The midpoint argument above gives exact two-mechanism minimax regret $t^2(w_\star - w)^2$. With two environments,

$$\det \begin{pmatrix} w_1 & 1 - w_1 \\ w_2 & 1 - w_2 \end{pmatrix} = w_1 - w_2, \tag{39}$$

so distinct mixtures identify both mechanisms. Moreover $\sigma_{\min}(W) = |\det W|/\sigma_{\max}(W) \geq |w_1 - w_2|/\sqrt{2}$ because $\|W\|_F \leq \sqrt{2}$.

If each environment supplies M_e independent rooted gadgets and $M_{\min} := \min_{e\in[E]} M_e$, Hoeffding's inequality and a union bound yield

$$P\Big(\max_e \max\{|\widehat{w}_e - w_e|, |\widehat{m}_e - m_e|\} > t\Big) \leq 4E \exp(-2M_{\min}t^2). \tag{40}$$

For the symmetric pair with M gadgets in each environment, applying two-sided Hoeffding only to the two label-mean estimates at $t = a\gamma$ proves reversal recovery with probability at least $1 - 4e^{-2Ma^2\gamma^2}$. If $\Delta_w = |w_1 - w_2|$, setting $t = \Delta_w/4$ gives, with probability at least $1 - 4e^{-M\Delta_w^2/8}$, $|\widehat{w}_1 - \widehat{w}_2| \geq \Delta_w/2$ and $\sigma_{\min}(\widehat{W}) \geq \Delta_w/(2\sqrt{2})$.

B.6 Random embeddings, overlap, and thinning

For $j \in [M]$, let $\varphi_j : V_r \to [N]$ be an injective embedding and let $S_j \in \{+,-\}$ be its trade orientation. In a multihypergraph, complete shadow counts add:

$$C_q\left(\biguplus_{j=1}^{M} \varphi_j(\mathcal{H}_{S_j})\right) = \sum_{j=1}^{M} C_q(\varphi_j(\mathcal{H}_{S_j})). \tag{41}$$

Theorem 4.3 makes each summand independent of S_j for $q < r$, so arbitrary vertex overlaps preserve exact equality. A coupled, environment-invariant background adds the same vector to both sides.

For a simple union, exact equality holds whenever all potential planted r-edges are distinct. With M independent uniform embeddings and an independent Erdős–Rényi r-uniform background $G^{(r)}(N, p_{\mathrm{bg}})$, where $p_{\mathrm{bg}} \in [0,1]$ is its edge probability, a union bound gives

$$\Pr(\text{planted collision or background hit}) \leq \binom{M}{2} \frac{4^r}{\binom{N}{r}} + M2^r p_{\mathrm{bg}}. \tag{42}$$

Indeed, two cells have at most 4^r pairs of potential edges, independently uniform labelled r-sets agree with probability $1/\binom{N}{r}$, and the background has at most $M2^r$ planted edges to hit. Bounded

vertex overlap alone is insufficient; the relevant event is edge collision. If duplicate excesses are D_s, D_t, deleting one repeated r-edge changes $\binom{r}{q}$ shadow coordinates, hence

$$\|C_q(G_s^{\text{simple}}) - C_q(G_t^{\text{simple}})\|_1 \le \binom{r}{q}(D_s + D_t). \tag{43}$$

Under gadget-level thinning, retain each of the M planted cells independently with probability $\rho \in [0,1]$ and condition on the retention indicators. A retained cell has identical proper shadows in both orientations and a removed cell is empty, so equality remains exact. The number of retained trade cells is $n_r \sim \text{Bin}(M,\rho)$ and Chernoff gives

$$P\{n_r < (1-t)M\rho\} \le e^{-t^2 M\rho/2}. \tag{44}$$

Independent *edge* thinning is weaker. Each valid shadow coordinate has the same marginal law $\text{Bin}(2^{r-q-1},\rho)$ under both orientations, but dependencies among coordinates can differ. With known embeddings, any retained full edge reveals orientation, so the effective probability per cell is $1-(1-\rho)^{2^{r-1}} \sim 2^{r-1}\rho$. We make no claim of exact joint failure of the projection to reveal the orientation for arbitrary nonlinear learners after independent edge thinning.

Recovering full order from overlapping unions requires marked roots, known embeddings, or separable cells; decomposing an anonymous dense union into trades is a separate problem.

B.7 General fibers, quantile computation, and sampling error

The same interval can be computed without enumerating the full states in a fiber. Fix its reference probability law μ on a standard Borel space, a measurable score $\eta : \mathcal{H} \to [0,1]$, a total-variation radius $\rho \ge 0$, and a finite likelihood-ratio cap $1 < \kappa < \infty$. The admissible target densities satisfy $\kappa^{-1} \le r \le \kappa$, $\mathbb{E}_\mu r = 1$, and $\mathbb{E}_\mu|r-1|/2 \le \rho$. Let F be the distribution of $\eta(H)$ for $H \sim \mu$, let $Q_F(u) = \inf\{v : F(v) \ge u\}$ be its quantile for $0 < u < 1$, and put $m_F = \int_0^1 Q_F(u)\,du$. This is the quantile form of the robust expectation problem; threshold weights and quantile balancing have established roles in sensitivity analysis (Asiaee et al., 2026; Dorn & Guo, 2023).

Proposition B.1 (Attainable interval on a general fiber)**.** *Set $a = \kappa - 1$, $b = 1 - \kappa^{-1}$, and $t = \min\{\rho, (\kappa-1)/(\kappa+1)\}$. The largest and smallest admissible target scores are attained and equal*

$$U(F) = m_F + a\int_{1-t/a}^{1} Q_F(u)\,du - b\int_0^{t/b} Q_F(u)\,du, \tag{45}$$

$$L(F) = m_F - b\int_{1-t/b}^{1} Q_F(u)\,du + a\int_0^{t/a} Q_F(u)\,du. \tag{46}$$

For any other score distribution G on $[0,1]$, both endpoint errors are at most $\kappa W_1(F,G)$, where $W_1(F,G) = \int_0^1 |Q_F(u) - Q_G(u)|\,du$ is the first Wasserstein distance. Consequently, with n independent full-context reference observations and known η, the empirical endpoints satisfy, simultaneously with probability at least $1-\delta$,

$$\max\{|U(\widehat{F}_n) - U(F)|, |L(\widehat{F}_n) - L(F)|\} \le \kappa\sqrt{\frac{\log(2/\delta)}{2n}}, \qquad 0 < \delta < 1. \tag{47}$$

Replacing η by a uniformly accurate $\widehat{\eta}$ with $\|\widehat{\eta} - \eta\|_\infty \le h$ adds at most h to either endpoint. The same statement applies to deterministic score quantization of mesh h.

Proof. Write $r - 1 = s - v$ with $s = (r-1)_+$ and $v = (1-r)_+$. Their common integral is a transported mass $t' \le \rho$, and their pointwise capacities are a and b. To maximize the score, rearrangement puts s on the upper score quantiles and v on the lower quantiles. These portions have reference mass t'/a and t'/b. They are disjoint until $t'(a^{-1} + b^{-1}) = 1$, or $t' = (\kappa - 1)/(\kappa + 1)$; before that point the gain is nondecreasing. Thus the optimal mass is the displayed t. Interchanging upper and lower portions gives the minimum. At a quantile atom, use the fractional average of the threshold density; this preserves the expectation and can only reduce total variation. Hence atoms require neither an atomless state law nor a state-space discretization.

The displayed optimizers are fixed functions of quantile position with values between κ^{-1} and κ. Coupling Q_F and Q_G by the same uniform quantile therefore bounds the endpoint difference by $\kappa W_1(F, G)$. On $[0, 1]$, $W_1(F, \widehat{F}_n)$ is bounded by the uniform empirical cumulative-distribution error. The Dvoretzky–Kiefer–Wolfowitz inequality with its sharp constant gives (47) (Massart, 1990). Finally, every admissible density integrates to one, so a uniform score perturbation of size h changes every feasible expectation by at most h. Taking extrema proves the last assertion. □

Sorting the n reference scores costs $O(n \log n)$; each fixed-radius query then uses two quantile-tail sums. The guarantee concerns the declared reference fiber law and sensitivity envelope. Projected observations alone do not supply its hidden scores or certify its overlap radii. At $\kappa = 1$ the interval is the singleton m_F. Removing the finite density cap can destroy endpoint attainment; bounded-score infima and suprema still define the corresponding total-variation interval.

C Fixed-Budget Learning and Environment Geometry

This appendix separates two questions that coincide at population level but need not coincide statistically. The first asks whether a quadratic multiaffine mechanism can be estimated in integrated prediction loss under a fixed total source budget. The second asks whether its quadratic component can be detected uniformly. We prove the learning statement and the geometric spectral bounds here. Appendix D uses the same bounds to prove the certification statement. Throughout this appendix, the data are an explicitly declared Gaussian release of population moments; they are not asserted to have the exact finite-sample likelihood of empirical Bernoulli moments.

C.1 Experiment, target loss, and centered environment class

Let $d \geq 10$, let

$$\mathbb{S}_0^d := \{Q \in \mathbb{R}^{d\times d} : Q = Q^\top,\ \operatorname{diag}(Q) = 0\}, \qquad p := \dim(\mathbb{S}_0^d) = \binom{d}{2},$$

and equip every matrix space with the Frobenius inner product. The stable binary mechanism is

$$\eta_{b,Q}(X) = \frac{1}{2} + b^\top X + \frac{1}{2} X^\top Q X, \qquad (b, Q) \in \Theta_R := \left\{(b, Q) : \|b\|_2^2 + \frac{1}{2}\|Q\|_F^2 \leq R^2\right\}. \tag{48}$$

Here $X \in \{-1, 1\}^d$. Because the $d + p$ nonconstant monomials in (48) all have absolute value one, $R \leq [4\sqrt{d+p}]^{-1}$ implies $\eta_{b,Q}(X) \in [1/4, 3/4]$ for every $(b, Q) \in \Theta_R$ and every X. This radius condition only supplies a valid Bernoulli interpretation; the Gaussian experiment below is well defined for every $R > 0$.

An environment is indexed by a vector $x_e \in \mathbb{R}^d$ with $\|x_e\|_2 \leq \alpha \leq 1/2$. Given allocations $n_e \geq 1$ satisfying $\sum_{e=1}^E n_e = N$, the released value and gradient moments are independent and obey

$$\widehat{v}_e = \frac{1}{2} + b^\top x_e + \frac{1}{2} x_e^\top Q x_e + \frac{\sigma}{\sqrt{n_e}}\zeta_e, \qquad \widehat{g}_e = b + Q x_e + \frac{\sigma}{\sqrt{n_e}}\xi_e, \tag{49}$$

where $\zeta_e \sim N(0, 1)$, $\xi_e \sim N(0, I_d)$, and all noises are independent. Under a product-Rademacher environment with $\mathbb{E}_e X = x_e$, the two population means in (49) equal $\mathbb{E}_e Y$ and the normalized first moments $\operatorname{Cov}_e(X_i, Y)/(1 - x_{e,i}^2)$, respectively. Their empirical counterparts are generally correlated and heteroscedastic. Thus (49) is a Gaussian population-moment oracle, not a relabeling of the complete Bernoulli sample.

The target has independent Rademacher coordinates and is independent of the source releases; write $P_\star$ for its law. Orthogonality of its linear and quadratic multiaffine monomials gives the exact prediction loss

$$\mathcal{L}_\star((\widehat{b}, \widehat{Q}), (b, Q)) := \mathbb{E}_X\left[\left\{\eta_{\widehat{b},\widehat{Q}}(X) - \eta_{b,Q}(X)\right\}^2\right] = \|\widehat{b} - b\|_2^2 + \frac{1}{2}\|\widehat{Q} - Q\|_F^2. \tag{50}$$

We next define the centered environment class used for the sharp upper bound. Choose an integer

$$m_d := \left\lfloor \frac{\sqrt{8d+1} - 1}{2} \right\rfloor, \qquad 0 \leq m \leq m_d, \qquad k := d - m, \qquad E := k + 1,$$

and let $[U, V_{\text{env}}] \in \mathbb{R}^{d\times d}$ be orthogonal, with U having m columns and V_{env} having k columns. Let $\omega_1, \dots, \omega_E \in \mathbb{R}^k$ be the vertices of a unit regular simplex:

$$\sum_{e=1}^{E} \omega_e = 0, \qquad \frac{1}{E}\sum_{e=1}^{E} \omega_e \omega_e^\top = \frac{1}{k} I_k. \tag{51}$$

Set $x_e = \alpha V_{\text{env}} \omega_e$ and take the balanced allocation $n_e = N/E$, assumed integer for notational simplicity. Rounding the allocations changes only universal constants when $N \geq 2E$.

Stack the gradient releases in $\widehat{G} = [\widehat{g}_1, \dots, \widehat{g}_E]$. An orthonormal contrast of the columns, using (51), removes b and is exactly equivalent to

$$\widehat{M}_g = QV_{\text{env}} + \sqrt{v_N}\, Z \in \mathbb{R}^{d\times k}, \qquad v_N := \frac{\sigma^2 k}{N\alpha^2}, \qquad Z_{ij} \overset{\text{iid}}{\sim} N(0,1). \tag{52}$$

The discarded constant contrast gives independently $\widehat{b} = b + (\sigma/\sqrt{N}) z_b$ with $z_b \sim N(0, I_d)$. The value releases are redundant for the upper bound and can only improve an optimal procedure.

They also change the information in the centered experiment by only a constant factor. This fact is needed when a lower bound proved from the gradient contrast is transferred to the full release.

Lemma C.1 (Value moments preserve the centered-channel spectrum)**.** *Let $\sum_e x_e = 0$, let the allocation be balanced, and suppose $\|x_e\|_2 \leq \alpha \leq 1/2$. Denote by $\mathcal{I}_{\text{grad}}$ and $\mathcal{I}_{\text{full}}$ the Fisher information operators for (b, Q) from, respectively, the gradient releases and the joint value–gradient releases in (49). Then*

$$\mathcal{I}_{\text{grad}} \preceq \mathcal{I}_{\text{full}} \preceq (1 + 2\alpha^2)\mathcal{I}_{\text{grad}} \preceq \frac{3}{2}\mathcal{I}_{\text{grad}}. \tag{53}$$

In particular, both experiments have the same identifiable parameter range, and their generalized least-squares covariance spectra on that range agree within a factor $3/2$. Equivalently, a gradient experiment with its noise variance divided by $1 + 2\alpha^2$ Blackwell-dominates the full release. Thus a minimax lower bound for the gradient contrast remains a lower bound for the full release after this universal variance adjustment.

Proof. For a perturbation $(h, H) \in \mathbb{R}^d \times \mathbb{S}_0^d$, the gradient signals are $h + Hx_e$. Centering gives

$$\sum_{e=1}^{E} \|h + Hx_e\|_2^2 = E\|h\|_2^2 + \sum_{e=1}^{E} \|Hx_e\|_2^2. \tag{54}$$

The corresponding value signals are $h^\top x_e + x_e^\top H x_e/2$. Since $(a+b)^2 \leq 2a^2 + 2b^2$ and $\|x_e\|_2 \leq \alpha$,

$$\begin{aligned}\sum_{e=1}^{E} \left(h^\top x_e + \frac{1}{2} x_e^\top H x_e\right)^2 &\leq 2\alpha^2 E\|h\|_2^2 + \frac{\alpha^2}{2}\sum_{e=1}^{E}\|Hx_e\|_2^2 \\ &\leq 2\alpha^2 \left(E\|h\|_2^2 + \sum_{e=1}^{E}\|Hx_e\|_2^2\right).\end{aligned}$$

Equal variances give (53); inversion on the identifiable range proves both claims. □

Lemma C.2 (Rank threshold inside the centered class)**.** *For a centered design spanning $\mathcal{V} = \text{Col}(V_{\text{env}})$, the noiseless gradient channel $Q \mapsto QV_{\text{env}}$ is injective on $\mathbb{S}_0^d$ exactly when $\mathcal{D}_U$ is injective on Sym_m, where $\text{Col}(U) = \mathcal{V}^\perp$ and $\mathcal{D}_U(B) := \text{diag}(UBU^\top)$. Therefore no centered design with $\binom{m+1}{2} > d$ is identifying. When $m = 0$, the channel directly observes all columns of Q and is injective. For $m \geq 1$, the frame (61) below is identifying whenever $\binom{m+1}{2} \leq d$.*

Proof. If $QV_{\text{env}} = 0$ and Q is symmetric, then both the rows and columns of Q vanish on $\mathcal{V}$, so $Q = UBU^\top$ for a unique $B \in \text{Sym}_m$. This matrix belongs to $\mathbb{S}_0^d$ exactly when $\mathcal{D}_U(B) = 0$. Hence the channel kernel and $\ker(\mathcal{D}_U)$ are linearly isomorphic. Since $\dim(\text{Sym}_m) = \binom{m+1}{2}$ and $\mathcal{D}_U$ has only d rows, injectivity requires the displayed dimension inequality. Lemma C.4 proves injectivity for the explicit frame. The centered value signal of any channel-kernel element is also zero, because every x_e lies in $\mathcal{V}$; fixing the intercept at $1/2$ does not add another identifying equation. □

C.2 FIXED-BUDGET PREDICTION LAW

Proposition C.3 (Explicit bounds underlying Theorem 6.1)**.** *Consider the Gaussian release (49) and the loss (50). There are universal constants $c, C > 0$ with the following properties.*

First, for every number of environments, every placement satisfying $\|x_e\|_2 \le \alpha \le 1/2$, and every allocation with total budget N,

$$\inf_{(\widehat{b},\widehat{Q})} \sup_{(b,Q)\in\Theta_R} \mathbb{E}_{b,Q}\mathcal{L}_\star \ge c \min\left\{R^2, \frac{\sigma^2 d^3}{N\alpha^2}\right\}. \tag{55}$$

Second, for every $0 \le m \le m_d$, there is an explicit orthogonal decomposition $[U, V_{\mathrm{env}}]$ for which the centered simplex design above and a linear estimator satisfy

$$\begin{aligned}\sup_{(b,Q)\in\Theta_R} \mathbb{E}_{b,Q}\mathcal{L}_\star &\le C \min\left\{R^2, \frac{\sigma^2 d}{N} + v_N(d^2+m^4)\right\} \\ &\le C \min\left\{R^2, \frac{\sigma^2 d^3}{N\alpha^2}\right\}.\end{aligned} \tag{56}$$

At $m = m_d$, this design uses $E = d - m_d + 1 = d - \Theta(\sqrt{d})$ environments, the smallest count in the centered class for which the quadratic channel can be injective. Hence the first algebraically identifying environment count already attains the global fixed-budget minimax rate for integrated prediction. Its sample transition is $N \asymp \sigma^2 d^3/(\alpha^2 R^2)$; critical algebraic identification creates no additional average-risk factor.

The lower bound covers all radius designs; the environment minimum uses the centered channel (52), while Theorem E.10 permits a free affine-anchor intercept.

Proof of the lower bound in Proposition C.3. The restriction to coefficient-valued estimators loses nothing. Conditional on the source release, project any proposed function $\widehat{\eta} : \{-1,1\}^d \to \mathbb{R}$ orthogonally in $L_2(P_\star)$ onto the known affine space

$$\frac{1}{2} + \mathrm{span}\{X_i : i \in [d]\} + \mathrm{span}\{X_i X_j : 1 \le i < j \le d\}.$$

Every truth in (48) belongs to this space, so the Pythagorean identity shows that the projection cannot increase target loss. Its coefficients define an estimator $(\widehat{b}, \widehat{Q})$, and (50) is its exact error. Thus the lower bound below applies to the infimum over all measurable prediction rules in Theorem 6.1.

Give b to the learner. This makes the experiment more informative and leaves only Q. Use the Frobenius-orthonormal basis

$$F_{ij} := \frac{e_i e_j^\top + e_j e_i^\top}{\sqrt{2}}, \qquad 1 \le i < j \le d,$$

of $\mathbb{S}_0^d$. At a location x, the Hilbert–Schmidt norm of the gradient map $Q \mapsto Qx$ and the squared norm of the value functional $Q \mapsto x^\top Qx/2$ are

$$\sum_{i<j} \|F_{ij}x\|_2^2 = \frac{d-1}{2}\|x\|_2^2, \qquad \sum_{i<j}\left(\frac{x^\top F_{ij} x}{2}\right)^2 = \frac{1}{2}\sum_{i<j} x_i^2 x_j^2 \le \frac{1}{4}\|x\|_2^4. \tag{57}$$

Let $\mathcal{I}_Q$ be the Gaussian Fisher information operator for the p orthonormal quadratic coordinates and put $T := \mathrm{tr}(\mathcal{I}_Q)$. Then

$$T \le \frac{1}{\sigma^2}\sum_{e=1}^{E} n_e \left\{\frac{d-1}{2}\alpha^2 + \frac{1}{4}\alpha^4\right\} \le \frac{Nd\alpha^2}{2\sigma^2}. \tag{58}$$

We use a bounded smooth prior so that the trace calculation is not mistaken for an unconstrained Gaussian-parameter formula. In each orthonormal coordinate take independently

$$\pi_r(t) = r^{-1}\cos^2\left(\frac{\pi t}{2r}\right)\mathbf{1}\{|t| \le r\}, \qquad r = R/\sqrt{p}.$$

The product prior is supported on $\|Q\|_F \le R$, hence on Θ_R with $b = 0$, and has prior Fisher information $\pi^2 p/R^2$ per coordinate. The multivariate van Trees inequality followed by the harmonic-mean inequality gives

$$\inf_{\widehat{Q}} \mathbb{E}_\pi \|\widehat{Q} - Q\|_F^2 \ge \operatorname{tr}\left(\mathcal{I}_Q + \frac{\pi^2 p}{R^2} I_p\right)^{-1} \ge \frac{p^2}{T + \pi^2 p^2/R^2}$$
$$\ge \frac{1}{1+\pi^2} \min\left\{R^2, \frac{p^2}{T}\right\}. \tag{59}$$

For $d \ge 3$, the bound $p^2/T \ge 2\sigma^2 d^3/(9N\alpha^2)$, combined with (50), proves (55). □

C.3 An explicit outer-product frame and its inverse spectrum

Put $t := m(m+1)/2 \le d$. For $1 \le w \le m$, form $G_w \in \mathbb{R}^{d\times m}$ from the t nonzero rows

$$\sqrt{w}\, e_i^\top \quad (i \in [m]), \qquad (e_i + e_j)^\top \quad (1 \le i < j \le m), \tag{60}$$

followed by $d - t$ zero rows, and define

$$S_w := G_w^\top G_w = (w + m - 2) I_m + \mathbf{1}\mathbf{1}^\top, \qquad U_w := G_w S_w^{-1/2}. \tag{61}$$

Then $U_w^\top U_w = I_m$. Coordinate and pairwise-sum outer-product frames such as (60) are classical; our use of this frame is to obtain the fixed-budget completion spectrum below (Casazza et al., 2016).

For an arbitrary $d \times m$ matrix U with orthonormal columns, write its ath row as $u_a^\top$ and define

$$\mathcal{D}_U : \mathrm{Sym}_m \longrightarrow \mathbb{R}^d, \qquad [\mathcal{D}_U(B)]_a := u_a^\top B u_a. \tag{62}$$

The singular values below use the Frobenius norm on Sym_m and the Euclidean norm on $\mathbb{R}^d$.

Lemma C.4 (Completion spectrum of the weighted frame)**.** *For $U = U_w$ in (61), the map $\mathcal{D}_U$ has full column rank. If $a_1, \dots, a_t$ are its singular values, then, uniformly over $1 \le w \le m$,*

$$\sum_{j=1}^t a_j^{-2} \asymp m^4, \qquad \sum_{j=1}^t a_j^{-4} \asymp m^6 + \frac{m^7}{w^4}, \tag{63}$$

where the comparison constants are universal. Every w is rate-optimal for integrated estimation; $w \ge m^{1/4}$ also removes the unweighted frame's extra fourth-moment term.

Here $\|A\|_{S_q} := (\sum_j \sigma_j(A)^q)^{1/q}$ denotes the Schatten q-norm, with $\sigma_j(A)$ the singular values of A.

Proof. For $B \in \mathrm{Sym}_m$, put $C = S_w^{-1/2} B S_w^{-1/2}$ and vectorize C in the orthonormal symmetric coordinates $(C_{11}, \dots, C_{mm}, \sqrt{2}C_{12}, \dots)$. On the nonzero rows of G_w, the raw diagonal map is

$$T_w = \begin{pmatrix} wI_m & 0 \\ A & \sqrt{2} I_{\binom{m}{2}} \end{pmatrix}, \qquad T_w^{-1} = \begin{pmatrix} w^{-1} I_m & 0 \\ -(\sqrt{2}w)^{-1} A & 2^{-1/2} I_{\binom{m}{2}} \end{pmatrix}, \tag{64}$$

where $A \in \mathbb{R}^{\binom{m}{2}\times m}$ is the unsigned incidence matrix of the complete graph: its row for $\{i,j\}$ is $e_i^\top + e_j^\top$. In particular, T_w is invertible. Direct multiplication gives

$$A^\top A = (m-2) I_m + \mathbf{1}\mathbf{1}^\top, \tag{65}$$

whose eigenvalues are $m-2$ with multiplicity $m-1$ and $2m-2$ in the all-one direction. Hence

$$\|T_w^{-1}\|_{S_2}^2 \asymp m^2, \tag{66}$$

$$\|T_w^{-1}\|_{S_4}^4 = \operatorname{tr}\{(T_w^{-\top} T_w^{-1})^2\} \asymp m^2 + \frac{m^3}{w^4}. \tag{67}$$

For clarity, the second comparison follows by expanding $T_w^{-1} T_w^{-\top}$: its four contributions are bounded by constant multiples of m/w^4, $\|A\|_F^2/w^4$, $\|AA^\top\|_F^2/w^4$, and $\binom{m}{2} + \operatorname{tr}(AA^\top)/w^2$. Equation (65) makes their sum $\asymp m^2 + m^3/w^4$.

The eigenvalues of S_w are $w + m - 2$ on $\mathbf{1}^\perp$ and $w + 2m - 2$ on $\operatorname{span}(\mathbf{1})$, both between universal multiples of m for $1 \le w \le m$. The congruence $C \mapsto S_w^{1/2} C S_w^{1/2}$ therefore has every singular value between universal multiples of m. Since

$$\mathcal{D}_{U_w}^\dagger = \left[C \mapsto S_w^{1/2} C S_w^{1/2}\right] T_w^{-1}$$

on the t active coordinates, the singular-value product inequalities and (67) give (63). □

C.4 TALL COMPLETION AND COVARIANCE GEOMETRY

We now allow $t < d$, so $\mathcal{D}_U$ is tall. Let $\mathcal{D} : \mathrm{Sym}_d \to \mathbb{R}^d$ be diagonal extraction. It is a coisometry: $\mathcal{D}\mathcal{D}^* = I_d$. Relative to the orthogonal decomposition

$$\mathrm{Sym}_d = \{UBU^\top : B \in \mathrm{Sym}_m\} \oplus \mathcal{R}_{U,V_{\mathrm{env}}}, \tag{68}$$

where $\mathcal{R}_{U,V_{\mathrm{env}}}$ contains the UV_{env}, $V_{\mathrm{env}}U$, and $V_{\mathrm{env}}V_{\mathrm{env}}$ blocks, write $\mathcal{D} = (\mathcal{D}_U, \mathcal{D}_{\mathrm{rest}})$. Coisometry gives the exact identity

$$\mathcal{D}_{\mathrm{rest}}\mathcal{D}_{\mathrm{rest}}^* = I_d - \mathcal{D}_U\mathcal{D}_U^*. \tag{69}$$

Lemma C.5 (Tall completion identity)**.** *Suppose $\mathcal{D}_U$ has full column rank and has singular values $a_1, \dots, a_t$. Put*

$$M := \mathcal{D}_U^\dagger \mathcal{D}_{\mathrm{rest}}, \qquad L(r) := r - U\{M(r)\}U^\top,$$

where $r \in \mathcal{R}_{U,V_{\mathrm{env}}}$ and $M(r) \in \mathrm{Sym}_m$. Then

$$MM^* = (\mathcal{D}_U^*\mathcal{D}_U)^{-1} - I_t. \tag{70}$$

Hence, counting multiplicity, L has squared spectrum $\{a_j^{-2}\}_{j=1}^t$ plus ones.

Let $\widehat{Q}$ be the linear completion obtained from (52): retain $U^\top \widehat{M}_g$ and $\mathrm{sym}(V_{\mathrm{env}}^\top \widehat{M}_g)$, apply L, and project orthogonally onto $\mathbb{S}_0^d$. It is unbiased for every $Q \in \mathbb{S}_0^d$. If $\Gamma_{U,V_{\mathrm{env}}}$ is its noise covariance as an operator on $\mathbb{S}_0^d$, then

$$\mathrm{tr}(\Gamma_{U,V_{\mathrm{env}}}) \le 2v_N \left\{ \dim(\mathcal{R}_{U,V_{\mathrm{env}}}) - t + \sum_{j=1}^t a_j^{-2} \right\}, \tag{71}$$

$$\mathrm{tr}(\Gamma_{U,V_{\mathrm{env}}}^2) \le 4v_N^2 \left\{ \dim(\mathcal{R}_{U,V_{\mathrm{env}}}) - t + \sum_{j=1}^t a_j^{-4} \right\}. \tag{72}$$

Proof. Take a thin singular value decomposition $\mathcal{D}_U = P\,\mathrm{diag}(a_1, \dots, a_t)O^*$, where P and O have orthonormal columns. Using (69),

$$\begin{aligned} MM^* &= \mathcal{D}_U^\dagger(I_d - \mathcal{D}_U\mathcal{D}_U^*)(\mathcal{D}_U^\dagger)^* \\ &= O\,\mathrm{diag}(a_1^{-2} - 1, \dots, a_t^{-2} - 1)O^*, \end{aligned}$$

which proves (70). The two summands in $L(r)$ occupy orthogonal blocks, so $L^*L = I + M^*M$. Its t distinguished eigenvalues are a_j^{-2} and the rest are one.

To connect this graph map to (52), parameterize the rest block as

$$r = UAV_{\mathrm{env}}^\top + V_{\mathrm{env}}A^\top U^\top + V_{\mathrm{env}}CV_{\mathrm{env}}^\top, \qquad C = C^\top.$$

The sufficient statistics $U^\top \widehat{M}_g$ and $\mathrm{sym}(V_{\mathrm{env}}^\top \widehat{M}_g)$ observe A and C, where $\mathrm{sym}(M) := (M + M^\top)/2$. In Frobenius-orthonormal coordinates on $\mathcal{R}_{U,V_{\mathrm{env}}}$, their information weights are $1/2$ on the cross block and one on the symmetric VV block. Equivalently, the reconstructed noisy rest block is $r + \sqrt{v_N}W^{-1}\varepsilon$, where ε is standard Gaussian and

$$I \preceq W^{-2} \preceq 2I. \tag{73}$$

For a true zero-diagonal Q, its blocks satisfy $\mathcal{D}_U(B) + \mathcal{D}_{\mathrm{rest}}(r) = 0$. Full column rank makes $B = -M(r)$, hence $L(r) = Q$ before the final projection. The projection is therefore exact on the signal and contractive on the noise. Applying the Schatten norm inequalities to P_0LW^{-1}, where P_0 is projection onto $\mathbb{S}_0^d$, yields

$$\|P_0LW^{-1}\|_{S_2}^2 \le 2\|L\|_{S_2}^2, \qquad \|P_0LW^{-1}\|_{S_4}^4 \le 4\|L\|_{S_4}^4.$$

The singular values of L found above now give (71)–(72). □

Proof of the upper bound in Proposition C.3. When $m = 0$, take U to have no columns and $V_{\text{env}} = I_d$. The completion map is the identity on the observed quadratic matrix, and its covariance trace is $O(v_N d^2)$. Suppose henceforth $m \geq 1$ and use the weighted frame in Lemma C.4. Then $\dim(\mathcal{R}_{U,V_{\text{env}}}) = \binom{d+1}{2} - t = O(d^2)$. Lemmas C.4 and C.5 therefore give

$$\operatorname{tr}(\Gamma_{U,V_{\text{env}}}) \lesssim v_N(d^2 + m^4), \qquad \{\operatorname{tr}(\Gamma^2_{U,V_{\text{env}}})\}^{1/2} \lesssim v_N\sqrt{d^2 + m^6 + m^7/w^4}. \tag{74}$$

Adding $\mathbb{E}\|\widehat{b} - b\|_2^2 = \sigma^2 d/N$ gives the linear-estimator branch, and the zero estimator gives the R^2 branch. For $m \leq m_d$, $m^4 \leq 4d^2$ and $k = d - m \asymp d$ yield the second inequality. □

The covariance Frobenius norm in the second bound of (74) governs dense quadratic-signal testing. Choosing $w \geq m^{1/4}$ reduces it to

$$v_N\sqrt{d^2 + m^6} \asymp v_N \max\{d, m^3\}. \tag{75}$$

Appendix D proves necessity and constructs a matching certificate with nontrivial power.

Corollary C.6 (Exact penalty of the unweighted critical frame)**.** *Suppose the ambient dimension is the triangular number $d = t = m(m+1)/2$. For the centered design generated by U_w, let $\Sigma_{\text{grad}}(w)$ be the efficient covariance for estimating Q from the gradient releases after profiling b, and let $\Sigma_{\text{full}}(w)$ be its counterpart for the joint value–gradient release. Uniformly over $1 \leq w \leq m$,*

$$\operatorname{tr}\{\Sigma_{\text{grad}}(w)\} \asymp \operatorname{tr}\{\Sigma_{\text{full}}(w)\} \asymp v_N(d^2 + m^4), \tag{76}$$

$$\|\Sigma_{\text{grad}}(w)\|_F \asymp \|\Sigma_{\text{full}}(w)\|_F \asymp v_N\sqrt{d^2 + m^6 + m^7/w^4}. \tag{77}$$

Thus $w = 1$ and $w = m^{1/4}$ have the same integrated estimation rate, while their covariance Frobenius norms differ by $\Theta(m^{1/2}) = \Theta(d^{1/4})$. The testing result in Appendix D turns this exact spectral gap into the same factor in certification tolerance.

Proof. When $d = t$, $\mathcal{D}_{U_w}$ is square and invertible. Therefore $\mathcal{D}_{U_w}\mathcal{D}_{U_w}^{-1} = I_d$, the graph completion $L(r)$ has zero diagonal for every rest block r, and the projection P_0 in Lemma C.5 acts as the identity. The rest-block space and $\mathbb{S}_0^d$ both have dimension $\binom{d}{2}$, so L is a bijection between them. Equation (73) now gives the two-sided Loewner comparison

$$v_N LL^* \preceq \Sigma_{\text{grad}}(w) \preceq 2v_N LL^*. \tag{78}$$

The squared singular values of L are the a_j^{-2} together with unit values. Taking the trace and the squared Frobenius norm in (78), then applying Lemma C.4, proves both comparisons for $\Sigma_{\text{grad}}(w)$. Lemma C.1 transfers them to $\Sigma_{\text{full}}(w)$ within universal constants. At $d \asymp m^2$, the dominant fourth sums are m^7 and m^6, giving the factor $\Theta(d^{1/4})$ in the optimal certification tolerance. □

C.5 Why the testing scale cannot be improved by another frame

The following argument gives geometric lower bounds for every orthonormal U, independent of the explicit construction. The map $\mathcal{D}_U$ is the sum of the d rank-one row operators $B \mapsto \langle B, u_a u_a^\top\rangle e_a$. The triangle inequality for the nuclear norm and $\sum_a \|u_a\|_2^2 = m$ imply

$$\|\mathcal{D}_U\|_{S_1} \leq \sum_{a=1}^{d} \|u_a u_a^\top\|_F = m. \tag{79}$$

If $\mathcal{D}_U$ has full column rank with singular values $a_1, \ldots, a_t$, Jensen's inequality applied to $x \mapsto x^{-q}$ gives

$$\sum_{j=1}^{t} a_j^{-2} \geq \frac{t^3}{m^2} \gtrsim m^4, \qquad \sum_{j=1}^{t} a_j^{-4} \geq \frac{t^5}{m^4} \gtrsim m^6. \tag{80}$$

Thus $w \geq m^{1/4}$ matches both universal inverse-spectrum lower bounds.

For completeness, we exhibit the actual zero-diagonal directions behind the fourth-moment obstruction. At most $2m$ singular values can exceed $1/2$ by (79). For $m \geq 6$, there is therefore an index set I with

$$\ell := |I| \geq t - 2m \geq \frac{m^2}{4}, \qquad a_j \leq \frac{1}{2} \quad (j \in I). \tag{81}$$

Let B_j and y_j be corresponding orthonormal right and left singular vectors, so $\mathcal{D}_U(B_j) = a_j y_j$. Equation (69) shows that y_j is an eigenvector of $\mathcal{D}_{\mathrm{rest}}\mathcal{D}_{\mathrm{rest}}^*$ with eigenvalue $1 - a_j^2$. The minimum-norm rest block satisfying $\mathcal{D}_{\mathrm{rest}}(r_j) = -a_j y_j$ is

$$r_j = -\frac{a_j}{1-a_j^2}\mathcal{D}_{\mathrm{rest}}^* y_j, \qquad \|r_j\|_F^2 = \frac{a_j^2}{1-a_j^2}. \tag{82}$$

The zero-diagonal matrices

$$Q_j := UB_jU^\top + r_j \tag{83}$$

are mutually Frobenius-orthogonal and satisfy $\|Q_j\|_F^2 = (1-a_j^2)^{-1} \in [1, 4/3]$. Their observable rest components are also orthogonal. Moreover,

$$\sum_{j\in I} a_j^{-2} \geq \frac{\ell^3}{(\sum_{j\in I} a_j)^2} \geq \frac{\ell^3}{m^2} \gtrsim m^4. \tag{84}$$

These are not abstract worst singular vectors: they are an $\Omega(m^2)$-dimensional family of valid multiaffine mechanisms whose target energy is order one while the corresponding rest component has norm $a_j/\sqrt{1-a_j^2}$. The rest components are mutually orthogonal. The actual gradient contrast applies the cross/VV information weight from (73), so its signals are spectrally equivalent, though not necessarily pairwise orthogonal. A more informative direct-rest Gaussian channel makes them orthogonal and is sufficient for the lower-bound comparison. A dense Rademacher mixture over these directions gives the $v_N m^3$ component of the testing lower bound, while an ordinary p-dimensional mixture gives the $v_N d$ component. The likelihood calculation and the matching quadratic test are provided in Appendix D.

Combining (75) with that lower bound produces the two environment thresholds. Integrated estimation is already budget-optimal for the largest admissible hidden dimension $m = m_d = \Theta(\sqrt{d})$, hence at $E = d - \Theta(\sqrt{d})$. Certification matches the fully observed dense-testing scale $v_N d$ exactly when $m^3 \lesssim d$, or $m = O(d^{1/3})$, which requires $E = d - O(d^{1/3})$. Adding environments across this interval does not improve the minimax prediction-error order, but it reduces the source budget needed to certify quadratic sufficiency by as much as a factor $\sqrt{d}$.

D Proof of the Safe-Coarsening Power Law

This section proves Theorem 6.2 for the Gaussian value–gradient channel in (23); Appendix C supplies its fixed-budget geometry. Accepting the coarse view certifies that its integrated omitted-mechanism energy lies below a declared tolerance. This is an approximate projection-sufficiency claim; paired risk against an arbitrary frozen baseline is a separate target.

D.1 A Gaussian energy certificate

We first isolate the testing statement used by the structured environment calculation. Let

$$X \sim N(\vartheta, \Sigma), \qquad \vartheta \in \mathbb{R}^D, \qquad \|\vartheta\|_2^2 \leq R^2, \tag{85}$$

where $\Sigma \succ 0$ is known, and abbreviate

$$F := \|\Sigma\|_{\mathrm{F}}, \qquad M := \|\Sigma\|_{\mathrm{op}}.$$

An acceptance rule $\phi(X) \in \{0,1\}$ is (ε,δ)-safe if

$$\sup_{\vartheta:\, \varepsilon \leq \|\vartheta\|_2^2 \leq R^2} P_\vartheta\{\phi(X) = 1\} \leq \delta. \tag{86}$$

Its null power is $P_0\{\phi(X) = 1\}$.

Lemma D.1 (Matching Gaussian safety–power boundary). *Fix $\delta, \beta \in (0, 1/4)$ with $\delta + \beta < 1$, and let $0 < \varepsilon \leq R^2$. If a rule satisfies (86) and has null power at least $1 - \beta$, then*

$$\varepsilon \geq c_{\delta,\beta} F, \qquad c_{\delta,\beta} := \left[2\log\{1 + 4(1-\delta-\beta)^2\}\right]^{1/2}. \tag{87}$$

Conversely, if

$$\varepsilon \geq C_{\delta,\beta} F, \qquad C_{\delta,\beta} := \max\{\sqrt{8/\beta}, 32/\delta\}, \tag{88}$$

then the rule

$$\phi_\varepsilon(X) := \mathbf{1}\{\|X\|_2^2 - \mathrm{tr}(\Sigma) \leq \varepsilon/2\} \tag{89}$$

is (ε, δ)*-safe and has null power at least* $1 - \beta$*.*

Proof. Diagonalize Σ and denote its eigenvalues by $\lambda_1, \ldots, \lambda_D$. Put $F^2 = \sum_i \lambda_i^2$. For independent Rademacher signs ω_i, define the prior

$$\vartheta_i(\omega) = \omega_i \frac{\sqrt{\varepsilon}\lambda_i}{F}. \tag{90}$$

Every prior draw satisfies $\|\vartheta(\omega)\|_2^2 = \varepsilon$ and hence belongs to the alternative in (86). Let $\overline{P}_\varepsilon$ be the induced mixture and $P_0 = N(0, \Sigma)$. In the eigenbasis, the sign-mixture likelihood ratio in coordinate i is

$$L_i(X_i) = \exp\left(-\frac{\vartheta_i^2}{2\lambda_i}\right) \cosh\left(\frac{\vartheta_i X_i}{\lambda_i}\right).$$

Independence and a direct Gaussian moment calculation give

$$\mathbb{E}_0 L_i^2 = \cosh\left(\frac{\vartheta_i^2}{\lambda_i}\right).$$

Consequently, using $\log\cosh(x) \leq x^2/2$,

$$\begin{aligned}\chi^2(\overline{P}_\varepsilon, P_0) &= \prod_{i=1}^{D} \cosh\left(\frac{\varepsilon\lambda_i}{F^2}\right) - 1 \\ &\leq \exp\left\{\frac{\varepsilon^2}{2F^4}\sum_{i=1}^{D}\lambda_i^2\right\} - 1 = \exp\left(\frac{\varepsilon^2}{2F^2}\right) - 1.\end{aligned} \tag{91}$$

If $\mathcal{A} = \{x : \phi(x) = 1\}$, null power and safety imply

$$P_0(\mathcal{A}) - \overline{P}_\varepsilon(\mathcal{A}) \geq 1 - \beta - \delta.$$

On the other hand,

$$\mathrm{TV}(\overline{P}_\varepsilon, P_0) \leq \frac{1}{2}\sqrt{\chi^2(\overline{P}_\varepsilon, P_0)}.$$

Combining this inequality with (91) and rearranging proves (87).

For the upper bound, write

$$T = \|X\|_2^2 - \mathrm{tr}(\Sigma).$$

Under the null,

$$\mathbb{E}_0 T = 0, \qquad \mathrm{Var}_0(T) = 2F^2.$$

Chebyshev's inequality therefore gives

$$P_0\{\phi_\varepsilon(X) = 0\} = P_0(T > \varepsilon/2) \leq \frac{8F^2}{\varepsilon^2} \leq \beta$$

under (88). For a general ϑ, set $v = \|\vartheta\|_2^2$. Gaussian quadratic-form moments give

$$\mathbb{E}_\vartheta T = v, \qquad \mathrm{Var}_\vartheta(T) = 2F^2 + 4\vartheta^\top\Sigma\vartheta \leq 2F^2 + 4Mv. \tag{92}$$

If $v \geq \varepsilon$, then $v - \varepsilon/2 \geq v/2$, and hence

$$\begin{aligned}P_\vartheta\{\phi_\varepsilon(X) = 1\} &\leq \frac{2F^2 + 4Mv}{(v - \varepsilon/2)^2} \\ &\leq 8\frac{F^2}{\varepsilon^2} + 16\frac{M}{\varepsilon} \leq \delta.\end{aligned}$$

Here $M \leq F$ and $\varepsilon \geq 32F/\delta$, with $\delta < 1/4$, justify the last inequality and complete the proof. □

Corollary D.2 (Executable analytic calibration)**.** *In the experiment of Lemma D.1, define*

$$t_\beta := \sqrt{2/\beta}\, F \tag{93}$$

and

$$\varepsilon_{\delta,\beta} := t_\beta + \frac{2M + \sqrt{4M^2 + 4\delta M t_\beta + 2\delta F^2}}{\delta}. \tag{94}$$

If $\varepsilon_{\delta,\beta} \le R^2$, *the rule*

$$\phi_{\delta,\beta}(X) := \mathbf{1}\{\|X\|_2^2 - \mathrm{tr}(\Sigma) \le t_\beta\}$$

has null power at least $1-\beta$ *and satisfies*

$$\sup_{\vartheta:\, \varepsilon_{\delta,\beta} \le \|\vartheta\|_2^2 \le R^2} P_\vartheta\{\phi_{\delta,\beta}(X) = 1\} \le \delta.$$

Proof. The null claim follows from $\mathrm{Var}_0(T) = 2F^2$ and Chebyshev's inequality. For $v = \|\vartheta\|_2^2 > t_\beta$, the one-sided form of the same inequality and (92) give

$$P_\vartheta(T \le t_\beta) \le h(v) := \frac{2F^2 + 4Mv}{(v - t_\beta)^2}. \tag{95}$$

The function h is strictly decreasing on (t_β, ∞), because its derivative has numerator $-4Mv - 4Mt_\beta - 4F^2 < 0$. The value in (94) is the larger solution of

$$\delta(\varepsilon - t_\beta)^2 = 2F^2 + 4M\varepsilon.$$

Therefore $h(v) \le \delta$ for every $v \ge \varepsilon_{\delta,\beta}$. □

Why a separation margin is necessary. Lemma D.1 compares exact sufficiency $\|\vartheta\|_2^2 = 0$ with loss at least ε. If power is instead required uniformly over $\|\vartheta\|_2^2 \le \varepsilon_0$, honest adaptation requires a positive gap between ε_0 and the unsafe region. This is the standard statistical separation needed to distinguish two touching composite hypotheses; it is distinct from the algebraic identification condition.

D.2 Application to value–gradient environment geometry

We now apply the lemma to the coordinates used in Theorem 6.1. Write

$$\vartheta := (b_1, \dots, b_d, \{Q_{ij} : 1 \le i < j \le d\})\,. \tag{96}$$

For the balanced product-Rademacher target in that theorem,

$$V(\vartheta) = \|\vartheta\|_2^2 = \|b\|_2^2 + \frac{1}{2}\|Q\|_{\mathrm{F}}^2. \tag{97}$$

Thus the coarse view is sufficient exactly when all nonconstant coefficients vanish, which requires both $b = 0$ and $Q = 0$. The fixed intercept $1/2$ does not enter (97).

Fix the centered isotropic environment design class in Theorem 6.2. It has

$$E = d - m + 1, \qquad k = d - m, \qquad v_N := \frac{\sigma^2 k}{N\alpha^2},$$

where $m(m+1)/2 \le d$ and $\alpha \le 1/2$. If the resulting source operator is rank deficient, a nonzero parameter direction has exactly the same observation law as the null. Scaling that direction to energy $\varepsilon \le R^2$ makes (86) and null power $1-\beta$ incompatible because $1-\beta > \delta$.

It remains to consider a full-rank design. Let Σ_m^{eff} be the covariance of the generalized least-squares sufficient statistic for ϑ in the complete value–gradient channel. Let Γ_m be the covariance of the explicit completion estimator under the weighted design in Appendix C. We next show

$$\|\Sigma_m^{\mathrm{eff}}\|_{\mathrm{F}} \ge c_0 v_N \sqrt{d^2 + m^6}, \qquad \|\Gamma_m\|_{\mathrm{F}} \le C_0 v_N \sqrt{d^2 + m^6}. \tag{98}$$

For the lower inequality, first consider only the gradient experiment (52). The map $Q \mapsto QV_{\mathrm{env}}$ is a contraction in Frobenius norm. Hence its Fisher-information eigenvalues on the

$p = \binom{d}{2}$ Frobenius-orthonormal quadratic coordinates are at most $1/v_N$, and the inverse Fisher covariance contributes at least a constant multiple of $v_N^2 d^2$ to its squared Frobenius norm. Lemma C.1 and (53) transfer this term to the complete value–gradient experiment, changing every inverse-information eigenvalue by at most a factor $3/2$.

For $m \geq 6$, the sharper completion term follows from the valid multiaffine directions Q_j in (83). For the index set I in (81), these directions are mutually orthogonal and have norms bounded above and below by universal constants. Their observable rest blocks are also mutually orthogonal, with squared norms $a_j^2/(1-a_j^2)$. The actual contrast applies the weight in (73); its information is bounded above by the information of the direct rest-block channel. Combining this domination with (53), the full-channel Fisher form on the normalized Q_j span is bounded above by a universal constant times $v_N^{-1}\,\mathrm{diag}(a_j^2 : j \in I)$. Its inverse therefore has eigenvalues at least $c v_N a_j^{-2}$. By (84) and Cauchy–Schwarz,

$$\sum_{j\in I} a_j^{-4} \geq \frac{\left(\sum_{j\in I} a_j^{-2}\right)^2}{|I|} \gtrsim m^6. \tag{99}$$

The same value–gradient comparison preserves this bound up to a universal constant. When $m \leq 5$, its m^6 term is absorbed into a universal multiple of the d^2 bulk. Converting to the target-energy coordinates rescales quadratic covariances by $1/2$. Combining the ordinary d^2 bulk with (99) proves the first inequality in (98). Because the argument uses the efficient covariance of the full release, its lower bound applies to every rule.

For the upper inequality, use the full-span estimator when $m = 0$ and, for $m \geq 1$, the weighted frame with $w = m^{1/4}$. At $m = 0$, the Gaussian full-span estimator has $\|\Gamma_{Q,0}\|_F = O(v_N d)$ because it has $\binom{d}{2}$ orthonormal quadratic coordinates with variance $O(v_N)$ each. For $m \geq 1$, equations (74)–(75) give

$$\{\mathrm{tr}(\Gamma_{Q,m}^2)\}^{1/2} \lesssim v_N \sqrt{d^2 + m^6}$$

for the completed quadratic coordinates. The independent estimator of b has covariance $(\sigma^2/N) I_d$, whose squared Frobenius contribution is lower order. This proves the second inequality in (98).

For the upper bound in Theorem 6.2, use that corresponding estimator and apply Lemma D.1 with $\Sigma = \Gamma_m$. The theorem's small-radius condition ensures that the resulting constant multiple of $v_N\sqrt{d^2+m^6}$ lies in $(0, R^2]$. For the lower bound, suppose that an (ε, δ)-safe rule has null power at least $1-\beta$. The full-rank Gaussian linear model is equivalent to its generalized least-squares sufficient statistic, so Lemma D.1 applies with $\Sigma = \Sigma_m^{\mathrm{eff}}$. The lower half of (98) and (87) imply

$$\varepsilon \geq c_{\delta,\beta} c_0 v_N \sqrt{d^2 + m^6}.$$

Together these two directions prove the matching rate

$$\varepsilon_{\mathrm{cert}}(d, m, N, \alpha, \sigma) \asymp_{\delta,\beta} \frac{\sigma^2(d-m)}{N\alpha^2}\sqrt{d^2+m^6}, \tag{100}$$

which is (26).

At the algebraic threshold $m = m_d \asymp \sqrt{d}$, the right side of (100) is

$$\frac{\sigma^2 d^{5/2}}{N\alpha^2}.$$

For $m \leq d^{1/3}$ it is instead

$$\frac{\sigma^2 d^2}{N\alpha^2}.$$

In both regimes, Theorem 6.1 keeps the fixed-budget mechanism-estimation order at $\sigma^2 d^3/(N\alpha^2)$. Additional well-placed environments improve the certification boundary by a factor $\sqrt{d}$ while preserving this rate. This gain quantifies the evidence price of approximate projection sufficiency; finer observations retain their usual minimax information advantage.

E PROOFS AND STATISTICAL DETAILS FOR CHORD

This appendix proves the claims in Section 5. The two orders are kept separate throughout: q indexes the projection actually evaluated at deployment, whereas ℓ indexes outcome-free source moment features. At a fixed (q, ℓ), write p for the number of columns of the declared full-state feature map (in the direct prediction setting this is p_q). All identification statements are relative to the declared linear feature map and to the fiber signatures or aggregate moment release. They are not claims that a rank-deficient release exhausts the information in an unrestricted raw joint distribution. The mixture matrices and target conditional feature means used below are known by design or estimated from the explicitly allowed unlabelled full contexts; they are not recovered jointly with the mechanism from projected label means.

E.1 IDENTIFICATION FROM MIXTURE WEIGHTS AND TARGET TRANSPORT

Fix an order q, a projected value z, and J full states $h_1, \dots, h_J$ in its fiber. For every source $e \in [E]$ with $P_e(Z_q = z) > 0$, define the stable probabilities $k_j := P_e(Y = 1 \mid H = h_j) = \eta(h_j)$, the mixture weights $W_q[e, j] := P_e(H = h_j \mid Z_q = z)$, and the projected means $m_q[e] := P_e(Y = 1 \mid Z_q = z)$. In vector form,

$$W_q \in [0,1]^{E \times J}, \qquad k = (k_1, \dots, k_J)^\top, \qquad m_q = W_q k. \tag{101}$$

Rows use a known design or allowed unlabeled full contexts. Repeating a row reduces noise and leaves $\ker(W_q)$ unchanged; a distinct mixture can identify an additional outcome contrast.

Proposition E.1 (Identification and target transport)**.** *On any parameter set with nonempty interior in $[0,1]^J$, W_q globally identifies k iff $\operatorname{rank}(W_q) = J$. For a declared target mixture row $w_\star \in [0,1]^J$ with $\mathbf{1}^\top w_\star = 1$, the functional $w_\star^\top k$ is globally identified iff $w_\star \in \operatorname{Row}(W_q)$; at a boundary truth, additional inequality constraints can yield accidental point identification. For the two-orientation trade, two sources identify both mechanism values iff their orientation weights differ, despite their identical proper-order shadows.*

Suppress the fixed projection index below, writing $W = W_q$ and $m = m_q$.

Proof of Proposition E.1. Denote the declared parameter set in the proposition by $\mathcal{K} \subseteq [0,1]^J$. If $\operatorname{rank}(W) = J$, then W has a left inverse and $k = W^\dagger m$ is unique. Conversely, suppose the parameter set $\mathcal{K} \subseteq [0,1]^J$ has nonempty interior and $\operatorname{rank}(W) < J$. Choose an interior k_0 and a nonzero $v \in \ker(W)$. For sufficiently small $t > 0$, both $k_0 + tv$ and $k_0 - tv$ belong to $\mathcal{K}$, are distinct, and have the same release. Thus the map is not globally injective on $\mathcal{K}$. The same argument proves the pointwise statement at any k_0 in the Euclidean interior of the declared parameter set.

All algebraic solutions of $W\widetilde{k} = m$ differ by an element of $\ker(W)$. Hence a target mean $w_\star^\top k$ is constant on every such equivalence class exactly when $w_\star^\top v = 0$ for every $v \in \ker(W)$. By the fundamental theorem of linear algebra, this is equivalent to $w_\star \in \ker(W)^\perp = \operatorname{Row}(W)$. If the condition fails, an interior point and a sufficiently short null perturbation give two admissible mechanisms with identical source releases but different target means, proving the target-specific converse.

The neighbourhood qualification is necessary for a pointwise iff. A rank-deficient map can accidentally identify an isolated or constrained boundary point because its affine null line may leave the parameter set in both directions. For example, on $[0,1]^2$, the valid probability row $W = (1/2, 1/2)$ point-identifies $k = (0,0)$ from $Wk = 0$ despite $\operatorname{rank}(W) = 1$. This boundary coincidence neither restores global identification nor persists after moving into the parameter-set interior.

For the two orientations of the octahedral trade, an environment with mixture weight w_e contributes the row $(w_e, 1 - w_e)$. One such row has a one-dimensional null space spanned by $(1 - w_e, -w_e)^\top$, whose inner product with the target mixture vector is $w_\star - w_e$; it is target-relevant exactly when $w_\star \neq w_e$. Two environments with weights $w_1 \neq w_2$ give

$$\det \begin{pmatrix} w_1 & 1 - w_1 \\ w_2 & 1 - w_2 \end{pmatrix} = w_1 - w_2 \neq 0,$$

so both orientation-specific mechanisms are recovered even though the two orientations have identical proper shadows. This is a qualitative rank condition: uniform recovery over a class also requires a relevant singular-value lower bound or, for one target, bounded transport leverage. □

Why pooling environment equations destroys an identifying row. For the two-environment trade, pooling the source blocks with weights α and $1-\alpha$ retains only

$$\bar{r} = (\bar{w}, 1-\bar{w}), \qquad \bar{m} = \bar{r}^\top k, \qquad \bar{w} = \alpha w_1 + (1-\alpha) w_2.$$

Because the entries of $\bar{r}$ sum to one, this compressed operator has rank exactly one. Its null space contains $v_{\bar{w}} = (1-\bar{w}, -\bar{w})^\top$. For any interior k_0 and sufficiently small t, define the admissible pair

$$k_\pm = k_0 \pm t(1-\bar{w}, -\bar{w})^\top. \tag{102}$$

It is indistinguishable after pooling but separated under the target mixture:

$$\bar{r}^\top k_+ = \bar{r}^\top k_-, \qquad r_\star^\top (k_+ - k_-) = 2t(w_\star - \bar{w}), \quad r_\star = (w_\star, 1-w_\star).$$

For a Bernoulli target with success probability p, Brier excess of a score s is $(s-p)^2$. Two indistinguishable target probabilities separated by $2t|w_\star - \bar{w}|$ therefore give, for every estimator,

$$\max_{\pm} \mathbb{E}_\pm (\widehat{s} - r_\star^\top k_\pm)^2 \geq t^2 (w_\star - \bar{w})^2,$$

because the maximum squared distance to two points is at least the squared half-separation. This lower bound covers the declared aggregate-release experiment and every common noise channel applied to the same aggregate. A learner additionally given raw labelled full states lies outside this experiment. In contrast, before compression the two rows have determinant $w_1 - w_2$, so $w_1 \neq w_2$ identifies k exactly at population level. The lower bound is zero when $w_\star = \bar{w}$, as it should be: that one target outcome probability is identified even though the full mechanism is not.

E.2 The aggregate source moment map

Proposition E.2 (Environment diversity trades against source moment order)**.** *Encode the octahedral orientation by $S \in \{-1, 1\}$ and let $\eta(S) = c + \gamma S$, with both orientations supported. No proper shadow reveals S, and it admits an exact direct scorer iff $\gamma = 0$; full order reveals S. With $\chi = (1, S)^\top$ and $\mu_e = \mathbb{E}_e S$, the constant source moment gives rows $A^e_{r,0} = (1, \mu_e)$ and rank two iff two source means differ. The orientation moment gives*

$$A^e_{r,r} = \begin{pmatrix} 1 & \mu_e \\ \mu_e & 1 \end{pmatrix}, \tag{103}$$

which has full rank in one nondegenerate environment. Thus heterogeneity can replace source moment order, while an exact target scorer that omits S remains impossible when $\gamma \neq 0$.

Proof of Proposition E.2. The two trade orientations have identical complete labelled shadows at every order below r, so a proper-shadow direct scorer is constant across $S = \pm 1$. It equals $c + \gamma S$ on both supported orientations exactly when $\gamma = 0$; at order r, S itself is observed. For $\chi = (1, S)^\top$ and $\psi_0 = 1$, $\mathbb{E}_e[\psi_0 \chi^\top] = (1, \mu_e)$, and stacked rows have rank two exactly when two μ_e differ. For $\psi_r = (1, S)^\top$, using $S^2 = 1$ gives (103), whose determinant is $1 - \mu_e^2 > 0$ in a nondegenerate environment. Without full-order target measurements, exact direct prediction is invalid for $\gamma \neq 0$. □

E.3 Paired Brier-risk switching relative to a frozen baseline

Corollary E.3 (Environment diversity for a frozen score head)**.** *Let $\mathcal{F}_{\text{fit}}$ be the sigma-field generated by all data and randomness used to fit the score s_q. Conditional on $\mathcal{F}_{\text{fit}}$, suppose $\mathbb{E}_e[Y \mid s_q, K] = \alpha_{qK} + \beta_{qK} s_q$ in every environment. With a constant source moment, the affine head has rank two iff two environment score means differ; with a within-score moment in one environment, it has rank two iff $\mathrm{Var}_e(s_q \mid K) > 0$. Finite-sample validity additionally requires the strict margin in (184), together with an exact common model or its declared one-sided approximation allowance.*

Proof of Corollary E.3. Condition throughout on $\mathcal{F}_{\text{fit}}$, so the learned score is a fixed measurable function. The assumed shared model is

$$\mathbb{E}_e[Y \mid s_q, K] = \alpha_{qK} + \beta_{qK} s_q, \qquad \phi^{\text{sc}}_{q,K} = (1, s_q)^\top. \tag{104}$$

With the constant source moment feature, the stacked design is

$$A_{q,0} = \begin{pmatrix} 1 & \mu_{1,qK} \\ \vdots & \vdots \\ 1 & \mu_{E,qK} \end{pmatrix}.$$

Here $\mu_{e,qK} := \mathbb{E}_e[s_q \mid K]$ is the conditional score mean in environment e. Writing $\bar{\mu}_{qK} = E^{-1}\sum_e \mu_{e,qK}$ and expanding gives

$$\det(A_{q,0}^\top A_{q,0}) = E\sum_{e=1}^{E}(\mu_{e,qK} - \bar{\mu}_{qK})^2. \tag{105}$$

This is positive exactly when the environment means are not all equal, which is equivalent to column rank two. If the source moment feature is instead $\phi^{\mathrm{sc}}_{q,K} = (1, s_q)^\top$ within one environment, then

$$A^e_{q,q} = \begin{pmatrix} 1 & \mathbb{E}_e[s_q \mid K] \\ \mathbb{E}_e[s_q \mid K] & \mathbb{E}_e[s_q^2 \mid K] \end{pmatrix},$$

whose determinant is

$$\det(A^e_{q,q}) = \mathbb{E}_e[s_q^2 \mid K] - \mathbb{E}_e[s_q \mid K]^2 = \mathrm{Var}_e(s_q \mid K). \tag{106}$$

Finite-sample rank also requires the corollary's perturbation event and singular-value margin. □

Lemma E.4 (Conditional validity for a frozen learned representation)**.** *Let $\mathcal{F}_{\mathrm{fit}}$ contain exactly the data and randomness used to fit all attempted predictors. Freeze every searched backbone and order-restricted map before opening the source-moment, support, specification, or paired-loss releases. Conditional on $\mathcal{F}_{\mathrm{fit}}$, if the simultaneous assumptions of Proposition G.1, or the paired event and assumed source–target risk bound of Theorem E.5, hold with probability $1-\delta$, their conclusions hold conditionally and unconditionally at that level, provided every later search is covered by the same simultaneous event.*

Proof of Lemma E.4. Let $\mathcal{E}$ be the respective simultaneous certification event. Conditional on $\mathcal{F}_{\mathrm{fit}}$, every learned representation and every candidate score is fixed; hence all deterministic rank, interval, support, and risk arguments in Proposition G.1 apply on $\mathcal{E}$ without conditioning on a data-selected function. The same is true of the paired inequalities in Theorem E.5. By assumption, $P(\mathcal{E} \mid \mathcal{F}_{\mathrm{fit}}) \geq 1-\delta$ almost surely, and thus

$$P(\mathcal{E}) = \mathbb{E}\{P(\mathcal{E} \mid \mathcal{F}_{\mathrm{fit}})\} \geq 1-\delta.$$

This proves both the conditional and unconditional conclusions. If a backbone, order, or arity is selected after a certification release is opened, then it is no longer fixed under this conditioning; simultaneous coverage over that search is therefore necessary, exactly as claimed. □

Theorem E.5 (Paired Brier switching around a fixed baseline)**.** *For frozen $[0,1]$-valued candidates indexed by a finite set $\mathcal{Q}_K$, let Q_K be the baseline order and define $\mathcal{R}_{e,K}(g) := \mathbb{E}_e[(g-Y)^2 \mid K]$ and $\Delta_{e,qK} := \mathcal{R}_{e,K}(g_{q,K}) - \mathcal{R}_{e,K}(g_{Q_K,K})$, with target analogs $\mathcal{R}_{\star,K}$ and $\Delta_{\star,qK}$. Fix a nonempty weight set $\mathcal{W}_K \subseteq \{w \in \mathbb{R}_+^E : \sum_e w_e = 1\}$ and allowances $\Gamma^{\mathrm{shift}}_{qK} \geq 0$ satisfying*

$$\Delta_{\star,qK} \leq \sup_{w\in\mathcal{W}_K} \sum_e w_e \Delta_{e,qK} + \Gamma^{\mathrm{shift}}_{qK}. \tag{107}$$

On the simultaneous event $\Delta_{e,qK} \leq \widehat{\Delta}_{e,qK} + r_{e,qK}$, where $\widehat{\Delta}_{e,qK}$ is a paired estimate and $r_{e,qK} \geq 0$ its radius, set

$$\widehat{\Delta}^+_{qK} := \sup_{w\in\mathcal{W}_K} \sum_e w_e(\widehat{\Delta}_{e,qK} + r_{e,qK}) + \Gamma^{\mathrm{shift}}_{qK}. \tag{108}$$

For a declared margin $\gamma_K \geq 0$, choose its minimizing order $\widehat{q}_K$ only if the minimum is below $-\gamma_K$; otherwise set $\widehat{q}_K = Q_K$. Then

$$\mathcal{R}_{\star,K}(g_{\widehat{q}_K,K}) - \mathcal{R}_{\star,K}(g_{Q_K,K}) \leq -\gamma_K \mathbf{1}\{\widehat{q}_K \neq Q_K\}. \tag{109}$$

Proof of Theorem E.5. Work on the stated simultaneous event. For every admissible $q < Q_K$, nonnegative weights and the coordinatewise one-sided bounds imply

$$\begin{aligned}\Delta_{\star,qK} &\le \sup_{w\in\mathcal{W}_K} \sum_e w_e \Delta_{e,qK} + \Gamma_{qK}^{\text{shift}} \\ &\le \sup_{w\in\mathcal{W}_K} \sum_e w_e (\widehat{\Delta}_{e,qK} + r_{e,qK}) + \Gamma_{qK}^{\text{shift}} = \widehat{\Delta}_{qK}^{+}.\end{aligned}$$

If the rule switches, the upper bound at its selected order is below $-\gamma_K$, so the corresponding target Brier-risk difference is at most $-\gamma_K$. If there is insufficient evidence to switch, the rule returns exactly the same frozen function $g_{Q_K,K}$; its risk difference is zero. Combining the two cases proves (109). The proof invokes only the declared assumed source–target risk bound (107), so it yields no statement for other losses or shifts and no causal-validity claim for the baseline. □

Corollary E.6 (Two sufficient source–target risk bounds)**.** *Conditional on K, if $P_{\star,K} = \sum_e w_e P_{e,K}$ for some $w \in \mathcal{W}_K$, then (107) holds with $\Gamma_{qK}^{\text{shift}} = 0$. If its half-$L_1$ total-variation distance from such a mixture is at most ρ_K, the bound holds with $\Gamma_{qK}^{\text{shift}} = 2\rho_K$ for predictors in $[0,1]$.*

Proof of Corollary E.6. Fix (q, K) and define the measurable paired loss difference

$$D_{qK}(H,Y) = \{g_{q,K}(Z_q) - Y\}^2 - \{g_{Q_K,K}(Z_{Q_K}) - Y\}^2.$$

Both predictors are frozen, so this is one fixed function. Because the predictions and Y lie in $[0,1]$, each Brier loss lies in $[0,1]$ and hence $D_{qK} \in [-1,1]$. If $P_{\star,K} = \sum_e w_e P_{e,K}$, linearity of expectation gives

$$\Delta_{\star,qK} = \mathbb{E}_{P_{\star,K}} D_{qK} = \sum_e w_e \mathbb{E}_{P_{e,K}} D_{qK} = \sum_e w_e \Delta_{e,qK}.$$

Taking the supremum over $\mathcal{W}_K$ proves (107) with zero allowance.

For the second case, put $P_{w,K} = \sum_e w_e P_{e,K}$. Under the half-$L_1$ definition of total variation, every measurable $f \in [a,b]$ obeys $|\mathbb{E}_P f - \mathbb{E}_Q f| \le (b-a)\,\mathrm{TV}(P,Q)$. Applying this inequality to $D_{qK} \in [-1,1]$ yields

$$\Delta_{\star,qK} \le \mathbb{E}_{P_{w,K}} D_{qK} + 2\rho_K = \sum_e w_e \Delta_{e,qK} + 2\rho_K \le \sup_{v\in\mathcal{W}_K} \sum_e v_e \Delta_{e,qK} + 2\rho_K,$$

which proves the claimed risk bound with $\Gamma_{qK}^{\text{shift}} = 2\rho_K$. □

Proposition E.7 (A forward-shift restriction is necessary)**.** *Fix an arity K, an alternative $q < Q_K$, and two frozen predictors with values in $[0,1]$. Let every predeployment covariate law be supported on $\mathcal{H}_{\text{pre}}$, and suppose there is a state $h_0 \notin \mathcal{H}_{\text{pre}}$ such that*

$$a := g_{q,K}\{\Pi_q(h_0)\} \neq b := g_{Q_K,K}\{\Pi_{Q_K}(h_0)\}.$$

Suppose the model class contains two stable Bernoulli outcome mechanisms $\eta^{(0)}$ and $\eta^{(1)}$ that agree on $\mathcal{H}_{\text{pre}}$ and satisfy $\eta^{(j)}(h_0) = j$ for $j \in \{0,1\}$. If the target shift class permits the point mass $H = h_0$ under either mechanism, then the two worlds have the same predeployment observation law but opposite strict orderings of the target Brier risks of $g_{q,K}$ and $g_{Q_K,K}$. Consequently, no rule measurable with respect to the predeployment release that selects $g_{q,K}$ with positive probability can guarantee a nonpositive target Brier-risk difference uniformly over this model class.

Proof. Use the same predeployment covariate laws in the two worlds. Their support is contained in $\mathcal{H}_{\text{pre}}$, where the two outcome mechanisms agree, so the complete joint law of every predeployment observation, and hence of any release computed from those observations, is identical. Under target point mass at h_0, the paired Brier-risk difference in world j is

$$\Delta_\star^{(j)} = (a-j)^2 - (b-j)^2 = (a-b)(a+b-2j). \tag{110}$$

Because $a, b \in [0,1]$ and $a \neq b$, we have $a+b > 0$ and $a+b-2 < 0$; hence $\Delta_\star^{(0)}\Delta_\star^{(1)} < 0$. The preferred predictor therefore reverses between two worlds that no predeployment rule can distinguish. Such a rule selects $g_{q,K}$ with the same probability $p > 0$ in both worlds. In the world where $\Delta_\star^{(j)} > 0$, its expected risk difference relative to fallback that returns the same baseline is $p\Delta_\star^{(j)} > 0$. Uniform noninferiority therefore requires (107), target-support identification, or an equivalent forward restriction. □

Let $\psi_{q,\ell}(H)$ be vector valued, outcome-free, and integrable; suppress q when it is unambiguous. The stable full-state model (2) gives

$$\mathbb{E}_e[\psi_\ell Y] = \mathbb{E}_e[\psi_\ell \mathbb{E}_e(Y \mid H)] = \mathbb{E}_e[\psi_\ell \chi(H)^\top]\theta. \tag{111}$$

Stacking these identities yields

$$b_{q,\ell} := \operatorname{stack}_e \mathbb{E}_e[\psi_{q,\ell} Y] = \operatorname{stack}_e \mathbb{E}_e[\psi_{q,\ell}\chi(H)^\top]\theta = A_{q,\ell}\theta, \tag{112}$$

which proves (18). No coordinate independence and no record-level pairing between the samples used for A and b is needed. In particular, an unlabelled sample from $P_e(H)$ may estimate A^e, while an independent labelled sample from the same environment law estimates b^e. What is essential is that the two aggregates refer to the same population law (or use correct importance weights); arbitrary unmatched environment laws would not satisfy (112). A known nonspherical block covariance is handled by applying the same nonsingular whitening map to both $A_{q,\ell}$ and $b_{q,\ell}$.

Proposition E.8 (Full and target-specific moment identification)**.** *Fix (q,ℓ), let $A_{q,\ell} \in \mathbb{R}^{m_{q,\ell}\times p}$ and $b_{q,\ell} = A_{q,\ell}\theta$, and fix a requested target row $a_\star \in \mathbb{R}^p$. Suppose the admissible parameter set has nonempty interior in $\mathbb{R}^p$. Then the declared moments globally identify θ if and only if $\operatorname{rank}(A_{q,\ell}) = p$. They globally identify $a_\star^\top\theta$ if and only if $a_\star \in \operatorname{Row}(A_{q,\ell})$. In the latter case the minimum-norm transport vector and its leverage are*

$$\lambda_\star = (A_{q,\ell}^\dagger)^\top a_\star, \qquad a_\star^\top\theta = \lambda_\star^\top b_{q,\ell}, \qquad L_{q,\ell}^2(a_\star) := a_\star^\top (A_{q,\ell}^\top A_{q,\ell})^\dagger a_\star. \tag{113}$$

The same iff statements hold pointwise at every Euclidean-interior parameter. At a constrained boundary only the forward implications are automatic; for a genuinely lower-dimensional model, the converse uses A and its kernel restricted to the model's affine tangent span.

Proof. Write $A = A_{q,\ell}$. By (112), two parameters give the same released moments exactly when their difference belongs to $\ker(A)$. Thus every coordinate is unique exactly when $\ker(A) = \{0\}$, equivalently when A has full column rank. If the null space is nontrivial, the interior assumption permits a sufficiently short pair $\theta_0 \pm tv$ along any $v \in \ker(A)$, so the algebraic ambiguity is also realized inside the parameter model. A functional $a_\star^\top\theta$ is constant on each affine equivalence class $\theta + \ker(A)$ exactly when $a_\star \perp \ker(A)$, or equivalently $a_\star \in \operatorname{Row}(A)$. Under this condition, the solutions of $A^\top\lambda = a_\star$ transport the target from $b_{q,\ell}$. The minimum-Euclidean-norm solution is $\lambda_\star = (A^\dagger)^\top a_\star$. Therefore

$$\|\lambda_\star\|_2^2 = a_\star^\top A^\dagger (A^\dagger)^\top a_\star = a_\star^\top (A^\top A)^\dagger a_\star,$$

Singular value decomposition (SVD) gives the last identity. Interior points admit null perturbations; at a constrained boundary, a rank-deficient operator can still identify the target functional. □

The source moment system contains the two advertised special cases. First, condition on a fixed cell $Z_q = z$ and let $\chi_j(H) = \mathbf{1}\{H = h_j\}$ for its J full states. Taking the constant source moment feature in every environment gives

$$A[e,j] = P_e(H = h_j \mid Z_q = z) = W_q[e,j], \qquad b[e] = P_e(Y = 1 \mid Z_q = z) = m_q[e].$$

Equivalently, an unconditional release can use the normalized cell indicator $\mathbf{1}\{Z_q = z\}/P_e(Z_q = z)$ under positivity. Thus the one-hot full-state moment map is exactly identification from environment-specific mixture weights. Second, if the mechanism is causally lumpable and has the direct representation $\eta(H) = \phi_q(Z_q)^\top\theta_q$, choose $\chi(H) = \phi_q(Z_q)$. Then $a_{\star,q}(z) = \mathbb{E}_\star[\chi(H) \mid Z_q = z] = \phi_q(z)$, so the same moment system is target-agnostic and deployment requires no full target contexts. With a general full-state feature map, the moment equations instead transport the target conditional feature mean $\mathbb{E}_\star[\chi(H) \mid Z_q = z]$. The martingale comparison in Theorem 5.1 holds one fixed full-state feature map χ across q; replacing it separately by each setting-specific ϕ_q would compare different parameter spaces and is not that theorem.

Because nested source moment features and additional environments append row blocks, their population kernels can only shrink. Repeated or linearly redundant environments may leave the rank unchanged. When a family of direct feature dictionaries is nested in q, increasing q appends columns, so identifying the larger feature map implies identification of its lower-order submap. These algebraic monotonicities concern a fixed unnormalised (or consistently whitened) release. Splitting a fixed sample budget across more environments changes noise and need not improve risk.

E.4 Information lost by projection and directions in the source moment nullspace

Proof of the risk decomposition in Theorem 5.1. Under (1), $\eta(H) = \mathbb{E}_\star[Y \mid H]$. For every Z_q-measurable g_q,

$$\mathbb{E}_\star[(g_q - Y)^2 - (\eta - Y)^2] = \mathbb{E}_\star(g_q - \eta)^2,$$

because the cross term has conditional expectation zero given H. Put $m_{\star,q} = \mathbb{E}_\star[\eta(H) \mid Z_q]$. Conditional expectation is an orthogonal projection in $L_2(P_\star)$, and hence

$$\mathbb{E}_\star(g_q - \eta)^2 = \mathbb{E}_\star(\eta - m_{\star,q})^2 + \mathbb{E}_\star(g_q - m_{\star,q})^2.$$

The first term is $V_{\star,q} = \mathbb{E}_\star \mathrm{Var}_\star\{\eta(H) \mid Z_q\}$, proving (19).

Let $\mathcal{F}_q = \sigma(Z_q)$ and suppose $\mathcal{F}_q \subseteq \mathcal{F}_{q+1}$. The conditional-variance identity gives the sharper relation

$$V_{\star,q} - V_{\star,q+1} = \mathbb{E}_\star\{\mathbb{E}_\star(\eta \mid \mathcal{F}_{q+1}) - \mathbb{E}_\star(\eta \mid \mathcal{F}_q)\}^2 \geq 0. \tag{114}$$

Moreover $V_{\star,q} = 0$ exactly when η is $\mathcal{F}_q$-measurable $P_\star$-almost surely, or equivalently when it is constant on almost every target-supported fiber. Uniform zero loss over a declared full-support intervention class is equivalent to constancy on every relevant fiber, which is causal lumpability in Definition 4.1. The full-support qualification is essential: $V_{\star,q} = 0$ imposes no restriction on mechanism values at full states outside the declared target support.

For the octahedral trade at $q < r$, the projection does not reveal the orientation $S \in \{-1, +1\}$. Under target plus weight $w_\star$, $\eta = c + \gamma S$, and therefore

$$V_{\star,q} = \gamma^2 \mathrm{Var}_\star(S) = \gamma^2\{1 - (2w_\star - 1)^2\} = 4\gamma^2 w_\star(1 - w_\star).$$

At $q = r$, observing S makes the conditional variance and information loss zero. □

There is also a cellwise version used below. With $v_{\star,q}(z) = \mathrm{Var}_\star\{\eta(H) \mid Z_q = z\}$,

$$\mathbb{E}_\star[(g_q(z) - Y)^2 - (\eta - Y)^2 \mid Z_q = z] = v_{\star,q}(z) + \{g_q(z) - m_{\star,q}(z)\}^2. \tag{115}$$

Nested refinement guarantees monotonicity only for the target average $V_{\star,q} = \mathbb{E}v_{\star,q}(Z_q)$; pointwise variances belong to different cells across orders and need not be comparable.

Proof of the monotonicity statements in Theorem 5.1. Assume $\mathbb{E}_\star\|\chi(H)\|_2^2 < \infty$. Let $a_q = \mathbb{E}_\star[\chi(H) \mid \mathcal{F}_q]$ and $D_q = a_{q+1} - a_q$. The vector martingale property gives $\mathbb{E}[D_q \mid \mathcal{F}_q] = 0$. Since a_q is $\mathcal{F}_q$-measurable, both cross terms in the expansion of $\mathbb{E}[a_{q+1}a_{q+1}^\top]$ vanish, yielding

$$M_{q+1} - M_q = \mathbb{E}[D_q D_q^\top] \succeq 0,$$

which is (20). Because $\mathbb{E}(\eta \mid \mathcal{F}_q) = \theta^\top a_q$,

$$V_{\star,q} = \mathbb{E}_\star\eta(H)^2 - \theta^\top M_q \theta,$$

consistent with (114).

Now hold a source-release null space $\mathcal{N}$ fixed and let $P_\mathcal{N}$ be its orthogonal projector. Then

$$\mathrm{tr}(P_\mathcal{N} M_q) = \mathbb{E}\|P_\mathcal{N} a_q\|_2^2$$

is nondecreasing by the positive-semidefinite increment above. Likewise, the worst average squared contribution from a coefficient vector $v \in \mathcal{N}$ with $\|v\|_2 \leq R$ is

$$\sup_{v \in \mathcal{N}: \|v\|_2 \leq R} \mathbb{E}(a_q^\top v)^2 = R^2 \lambda_{\max}(P_\mathcal{N} M_q P_\mathcal{N}),$$

Compression is monotone for fixed $\mathcal{N}$; new environments may shrink this null space. □

Derivation of the two-dimensional witness. Take $\chi = (1, S)^\top$ and $P_\star(S = 1) = P_\star(S = -1) = 1/2$. If the coarse observation is constant, then its conditional feature mean is $a_{\text{coarse}} = (1, 0)^\top$ and therefore

$$M_{\text{coarse}} = a_{\text{coarse}} a_{\text{coarse}}^\top = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

If the fine observation is S, then $a_{\text{fine}} = \chi$ and

$$M_{\text{fine}} = \mathbb{E}_\star[\chi\chi^\top] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A balanced source with the constant moment contributes $A^{(1)} = (1, 0)$, whose nullspace is $\mathcal{N}_1 = \operatorname{span}(e_2)$ and whose nullspace projector is $P_{\mathcal{N}_1} = \operatorname{diag}(0, 1)$. Direct multiplication gives

$$\operatorname{tr}(P_{\mathcal{N}_1} M_{\text{coarse}}) = 0, \qquad \operatorname{tr}(P_{\mathcal{N}_1} M_{\text{fine}}) = 1. \tag{116}$$

A second constant-moment source with $\mathbb{E}_2 S = \mu \neq 0$ changes the stacked operator to

$$A^{(1:2)} = \begin{pmatrix} 1 & 0 \\ 1 & \mu \end{pmatrix}, \qquad \det A^{(1:2)} = \mu \neq 0.$$

Thus $\ker(A^{(1:2)}) = \{0\}$ and both nullspace traces vanish. The example separates the two operations exactly: refining the target observation enlarges the component exposed to a fixed source nullspace, whereas adding an independent environment row removes that nullspace.

E.5 Exact minimax risk from source moments

E.6 Canonical Gaussian release and algebraic rank statements

The next Gaussian result is a linear-inverse benchmark on an unbounded row-space class. It separates ambiguity from sampling noise. The bounded-ball learning results in Section 6 use a different parameter class and an explicitly budgeted population-moment acquisition channel.

Theorem E.9 (Exact minimax risk from source moments)**.** *Let the known stacked operator be $A_{q,\ell} \in \mathbb{R}^{m_{q,\ell} \times p}$, fix a target row $a \in \mathbb{R}^p$, and let $\sigma > 0$, $n_{q,\ell}^{\text{src}} > 0$, and $R \geq 0$. In the whitened Gaussian experiment*

$$\widehat{b} = A_{q,\ell}\theta + \xi, \qquad \xi \sim N(0, \sigma^2 I_{m_{q,\ell}}/n_{q,\ell}^{\text{src}}), \tag{117}$$

consider the parameter class $\Theta_R(A_{q,\ell}) := \{u + v : u \in \operatorname{Row}(A_{q,\ell}), v \in \ker(A_{q,\ell}), \|v\|_2 \leq R\}$. Decompose $a = a_\parallel + a_\perp$ along the same row-space–nullspace direct sum. Then

$$\inf_{\widehat{T}(\widehat{b})} \sup_{\theta \in \Theta_R(A_{q,\ell})} \mathbb{E}_\theta(\widehat{T} - a^\top\theta)^2 = R^2\|a_\perp\|_2^2 + \frac{\sigma^2}{n_{q,\ell}^{\text{src}}} a_\parallel^\top (A_{q,\ell}^\top A_{q,\ell})^\dagger a_\parallel. \tag{118}$$

The two terms are irreducible ambiguity and sampling variance, jointly attained by $a_\parallel^\top A_{q,\ell}^\dagger \widehat{b}$.

The following algebraic statement allows an unknown intercept c. The fixed-intercept budget model uses centered designs, whose kernel has the same zero-diagonal completion reduction.

Algebraic environment geometry. Let $d \geq 1$ and let $x_1, \ldots, x_E \in \mathbb{R}^d$ be environment means. For the multiaffine quadratic $f(x) = c + b^\top x + \frac{1}{2} x^\top Q x$, with $c \in \mathbb{R}$, $b \in \mathbb{R}^d$, and zero-diagonal $Q = Q^\top$, let $\mathcal{J}_{d,E}$ map (c, b, Q) to $\{f(x_e), \nabla f(x_e)\}_{e=1}^E$. They are multivariate Hermite data (Sauer & Xu, 1995; Alexander & Hirschowitz, 1995) and the normalized first-order moments derived in Appendix E.7.

Theorem E.10 (Exact generic quadratic rank deficiency)**.** *For $E \geq 1$ means outside a proper algebraic subset of any nonempty open feasible set, let $m = \max\{d - E + 1, 0\}$. The exact nullity is*

$$\dim\ker \mathcal{J}_{d,E} = \left[\binom{m+1}{2} - d\right]_+. \tag{119}$$

Full rank first occurs at

$$E_{\min}^{\text{gen}} = d + 1 - \left\lfloor \frac{\sqrt{8d+1} - 1}{2} \right\rfloor. \tag{120}$$

Thus $d = 100$ needs 88 generic environments; equation counting suggests 51 because it misses the shared-Hessian constraints.

The shared Hessian couples the equations. Even optimally placed means need this environment count; above it, design improves conditioning (Corollary E.11).

Proof of Theorem E.9. Abbreviate $A = A_{q,\ell}$, let $\mathcal{R} = \mathrm{Row}(A)$ and $\mathcal{K} = \ker(A)$, and decompose $a = a_\parallel + a_\perp$ and $\theta = u + v$ along $\mathcal{R} \oplus \mathcal{K}$. Since $A^\dagger A = P_{\mathcal{R}}$, the estimator in the theorem satisfies

$$\widehat{T}_0 - a^\top \theta = a_\parallel^\top A^\dagger \xi - a_\perp^\top v. \tag{121}$$

The first term is centered Gaussian with variance

$$\frac{\sigma^2}{n_{q,\ell}^{\mathrm{src}}} a_\parallel^\top A^\dagger (A^\dagger)^\top a_\parallel = \frac{\sigma^2}{n_{q,\ell}^{\mathrm{src}}} a_\parallel^\top (A^\top A)^\dagger a_\parallel.$$

The cross term in the squared error has expectation zero, and

$$\sup_{v \in \mathcal{K}:\, \|v\|_2 \le R} (a_\perp^\top v)^2 = R^2 \|a_\perp\|_2^2.$$

This proves the upper bound with equality for $\widehat{T}_0$.

For the matching lower bound, put independent priors on the two orthogonal components. When $a_\perp \neq 0$, let v take the two values $\pm R a_\perp / \|a_\perp\|_2$ with equal probability; take $v = 0$ otherwise. The distribution of the observation does not depend on v, so its posterior variance contribution to the target is exactly $R^2 \|a_\perp\|_2^2$.

For the row-space component, take $u \sim N(0, \tau^2 P_{\mathcal{R}})$. If $A = U_s D_s V_s^\top$ is a thin SVD, then the informative transformed observation is the normal means experiment

$$D_s^{-1} U_s^\top \widehat{b} = V_s^\top u + \zeta, \qquad \zeta \sim N\left(0, \frac{\sigma^2}{n_{q,\ell}^{\mathrm{src}}} D_s^{-2}\right).$$

Its posterior covariance on $\mathcal{R}$ is

$$V_s \left(\tau^{-2} I_s + \frac{n_{q,\ell}^{\mathrm{src}}}{\sigma^2} D_s^2\right)^{-1} V_s^\top.$$

Therefore the Bayes risk for $a_\parallel^\top u$ converges, as $\tau \to \infty$, to

$$\frac{\sigma^2}{n_{q,\ell}^{\mathrm{src}}} a_\parallel^\top V_s D_s^{-2} V_s^\top a_\parallel = \frac{\sigma^2}{n_{q,\ell}^{\mathrm{src}}} a_\parallel^\top (A^\top A)^\dagger a_\parallel.$$

The posterior variance of the sum is the sum of the two contributions because the priors are independent and the data contain no information about v. Every Bayes risk lower-bounds the minimax risk, and taking the diffuse-prior limit proves the reverse inequality and hence (118).

For known positive-definite Σ and $\xi \sim N(0, \sigma^2 \Sigma / n_{q,\ell}^{\mathrm{src}})$, premultiplication by $\Sigma^{-1/2}$ gives the same argument, with $A^\top \Sigma^{-1} A$ replacing $A^\top A$ throughout the identifiable variance term. □

The exact equality uses known A, a known target conditional feature mean, Gaussian released moments, squared loss, an *unrestricted* row-space component, and a radius constraint only on the unobserved kernel component. In particular, Theorem E.9 does not impose $\|\theta\| \le R$; doing so would change the row-space minimax problem. For estimated operators or rows, Bernoulli observations, dependent hyperedges, or other parameter classes, the theorem supplies the geometry. With known A, independent bounded moment noise gives an $O\{L_{q,\ell}^2 / n_{q,\ell}^{\mathrm{src}}\}$ upper rate; a matching lower rate additionally needs nondegenerate noise in the target direction and a separate minimax argument.

The mixture-weight system has the same geometry with $A = W$, $\theta = k$, and $a_\star = w_\star$. Thus its null-space term vanishes exactly under the target row-space condition, while the identifiable term is controlled by the conditioning of the source mixtures. Probability constraints can change a global minimax constant at the simplex boundary, while leaving the row-space identification geometry and its interior Gaussian limit unchanged.

Exact scope of the three-term display. Fix (q,ℓ) and a target cell z, abbreviate $A := A_{q,\ell}$ and $a := a_{\star,q}(z)$, and decompose $a = a_\parallel + a_\perp$ along $\mathrm{Row}(A) \oplus \ker(A)$. Define

$$d_{q,\ell}(a) := \|a_\perp\|_2, \qquad L^2_{q,\ell}(a) := a_\parallel^\top (A^\top A)^\dagger a_\parallel.$$

For a fixed coefficient θ, write $v_{\star,q,\theta}(z) := \mathrm{Var}_\star\{\chi(H)^\top\theta \mid Z_q = z\}$ and suppress the subscript θ when the mechanism is fixed. With nullspace radius R, total coefficient bound B, and conditional-feature-mean radius $\epsilon_{a,q}$, the pointwise design criterion is

$$\underbrace{v_{\star,q}(z)}_{\text{information lost by projection}} + \left(\sqrt{\underbrace{R^2 d^2_{q,\ell}(a)}_{\text{unidentified component}} + \underbrace{\frac{\sigma^2 L^2_{q,\ell}(a)}{n^{\mathrm{src}}_{q,\ell}}}_{\text{estimation variance}}} + \underbrace{B\epsilon_{a,q}}_{\text{conditional-mean estimation}} \right)^2 . \tag{122}$$

Let the Gaussian aggregate release be independent of a fresh target outcome. Combining (115) with the release experiment gives, for every parameter θ and every release-based predictor $\widehat{T}$,

$$\mathbb{E}_\theta[(\widehat{T} - Y)^2 - (\eta - Y)^2 \mid Z_q = z] = v_{\star,q,\theta}(z) + \mathbb{E}_\theta(\widehat{T} - a^\top\theta)^2. \tag{123}$$

For a fixed row $c \in \mathbb{R}^p$, let $c_\parallel$ denote its projection onto $\mathrm{Row}(A)$ and abbreviate the exact Gaussian transport minimax value by

$$\mathfrak{M}_{q,\ell}(c) := R^2 \|P_{\ker A} c\|_2^2 + \frac{\sigma^2}{n^{\mathrm{src}}_{q,\ell}} c_\parallel^\top (A^\top A)^\dagger c_\parallel.$$

If the true target conditional feature mean is known and $v_{\star,q,\theta}(z)$ is fixed across the Gaussian parameter class (or is a fixed orthogonal component in a declared product experiment), minimising the worst-case right side of (123) gives exactly $v_{\star,q}(z) + \mathfrak{M}_{q,\ell}(a)$. This exact sum belongs to that canonical product experiment; it does not assert that an unrestricted Gaussian row-space parameter is simultaneously a globally valid Bernoulli probability model; probability constraints may alter the global minimax constant outside this canonical product experiment.

Now suppose an unlabelled target sample provides $\widehat{a}$ with $\|\widehat{a} - a\|_2 \le \epsilon_{a,q}$ and impose the separate finite-sample bound $\|\theta\|_2 \le B$. Let $\widehat{T}_{\widehat{a}}$ be the Gaussian transport estimator for the nominal row $\widehat{a}$. Minkowski's inequality and Cauchy–Schwarz give, uniformly on the intersection of the Gaussian class with this norm ball,

$$\left\{\mathbb{E}_\theta(\widehat{T}_{\widehat{a}} - a^\top\theta)^2\right\}^{1/2} \le \sqrt{\mathfrak{M}_{q,\ell}(\widehat{a})} + B\epsilon_{a,q}. \tag{124}$$

Squaring (124) and adding the cellwise projection term yields the form in (122) (where the hat on the nominal certified row is suppressed). The cross term is generally real, so the conditional-feature-mean contribution cannot be added as an independent squared error. This is a conservative criterion when $\epsilon_{a,q} > 0$; the transport part is exact for the canonical class when $\epsilon_{a,q} = 0$. The total norm bound B is indispensable under row uncertainty: with the row-space component unrestricted, an error in that direction could otherwise have unbounded target effect.

The source size $n^{\mathrm{src}}_{q,\ell}$ is the effective size of the released source moment. Separately, if a target design gives each order-q face effective count $n^{\mathrm{tar}}_q \asymp N^{r-q}\rho$, then $\epsilon_{a,q} = O((n^{\mathrm{tar}}_q)^{-1/2})$. No equality between these source and target effective sizes is assumed. If the projection variance varies with θ, minimax and addition also do not generally commute; then (123), the Gaussian transport theorem, and the row-error bound are three exact statements with distinct scopes. Integrated-risk equalities require averaging (123); they are not obtained by comparing unmatched cellwise variances across orders.

E.7 Rank deficiency of quadratic value–gradient observations

To connect the value–gradient operator to the moment construction, specialize (135) to anchor zero, $q = 2$, and $\ell = 1$. For

$$f(x) = c + b^\top x + \tfrac{1}{2} x^\top Q x, \qquad Q = Q^\top, \qquad \mathrm{diag}(Q) = 0,$$

an environment with mean x_e releases $\beta^e_\varnothing = f(x_e)$ and $\beta^e_{\{i\}} = b_i + (Qx_e)_i = \partial_i f(x_e)$. Hence the structured operator is exactly $\mathcal{J}_{d,E}$, the stacked value–gradient map defined before Theorem E.10. The matrix Q is local to this construction and is distinct from the rooted history H.

Proof of Theorem E.10. Assume $E \geq 1$. A member of the kernel is a multiaffine quadratic

$$f(x) = c + b^\top x + \tfrac{1}{2} x^\top Q x, \qquad Q = Q^\top, \qquad \operatorname{diag}(Q) = 0,$$

whose value and gradient vanish at $x_1, \ldots, x_E$. The gradient equations are

$$b + Q x_e = 0, \qquad e = 1, \ldots, E. \tag{125}$$

Let $V = \operatorname{span}\{x_e - x_1 : e = 2, \ldots, E\}$. Subtracting the first equation in (125) from the others gives $QV = 0$, after which $b = -Q x_1$. The value equation at x_1 fixes $c = \frac{1}{2} x_1^\top Q x_1$. Every other value equation is automatic: if $v = x_e - x_1 \in V \subseteq \ker(Q)$, the exact Taylor identity gives

$$f(x_1 + v) = f(x_1) + \nabla f(x_1)^\top v + \tfrac{1}{2} v^\top Q v = 0.$$

Thus the value–gradient kernel is linearly isomorphic to

$$\{Q \in \mathbb{R}^{d \times d} : Q = Q^\top, \qquad \operatorname{diag}(Q) = 0,\ QV = 0\}. \tag{126}$$

For generic means in an open feasible set, $\dim(V) = \min\{E - 1, d\}$, so $m := \dim(V^\perp) = \max\{d - E + 1, 0\}$. Let $U \in \mathbb{R}^{d \times m}$ have orthonormal columns spanning $V^\perp$, and write its ith row as $u_i^\top$. Symmetry and $QV = 0$ are equivalent to $Q = UBU^\top$ for a unique $B \in \operatorname{Sym}_m$, the vector space of real symmetric $m \times m$ matrices. The zero-diagonal conditions become the d linear equations

$$\langle B, u_i u_i^\top \rangle = 0, \qquad i = 1, \ldots, d. \tag{127}$$

Generically, the matrices $u_i u_i^\top$ span a subspace of Sym_m of dimension $\min\{d, \binom{m+1}{2}\}$. To see attainability, start with a full-column rank $d \times m$ matrix whose rows include $e_1, \ldots, e_m$ and then as many $e_j + e_k$, $j < k$, as are needed. The corresponding outer products are linearly independent until all $\binom{m+1}{2}$ symmetric coordinates have been spanned. Multiplying every row on the right by the same invertible matrix to orthonormalize the columns applies an invertible congruence to all outer products and preserves their span dimension. Maximal rank is therefore attained. Its failure is the common zero set of maximal minors, hence a proper algebraic subset; the asserted rank holds generically.

Subtracting the number of independent constraints in (127) from $\dim(\operatorname{Sym}_m) = \binom{m+1}{2}$ yields

$$\dim \ker \mathcal{J}_{d,E} = \binom{m+1}{2} - \min\left\{d, \binom{m+1}{2}\right\} = \left[\binom{m+1}{2} - d\right]_+,$$

which is (119). Let

$$t_d = \left\lfloor \frac{\sqrt{8d+1} - 1}{2} \right\rfloor,$$

the largest integer satisfying $t_d(t_d + 1)/2 \leq d$. Full column rank is equivalent to $m \leq t_d$; for $E \leq d + 1$, this is precisely $\binom{d-E+2}{2} \leq d$, while $E \geq d + 1$ gives $m = 0$ and is automatically full rank. The smallest generic environment count is $E = d + 1 - t_d$, proving (120). In particular, $d = 100$ gives $E_{\min}^{\text{gen}} = 88$, whereas the raw equation-count bound is $\lceil (1 + d + \binom{d}{2})/(d + 1) \rceil = 51$. □

Corollary E.11 (Global minimum environment count)**.** *For the value–gradient operator $\mathcal{J}_{d,E}$ in Theorem E.10, every placement of fewer than $E_{\min}^{\text{gen}}$ environment means is rank deficient. Once this threshold is met, design improves conditioning while the minimum count stays fixed.*

Proof of Corollary E.11. Fix E and write the value–gradient operator as a matrix whose entries depend on the environment means $(x_1, \ldots, x_E)$. Its entries, and hence all its minors, are polynomials in these means. The maximum rank attained on an open feasible domain is therefore the generic rank: if a minor is nonzero at any feasible placement, it is a nonzero polynomial and remains nonzero on a nonempty open subset. Theorem E.10 shows that the generic rank is deficient whenever $E < E_{\min}^{\text{gen}}$. A full-rank special placement below that threshold would make a full-column minor nonzero and would therefore make the generic operator full rank, a contradiction. At $E = E_{\min}^{\text{gen}}$, the theorem gives full rank on a generic open subset. Rank alone does not determine the smallest singular value, so placements at or above the threshold may still be optimized for conditioning or target leverage. □

E.8 FINITE-SAMPLE RANK, PARAMETER, AND TARGET CERTIFICATES

Proposition E.12 (Simultaneous deterministic certificates)**.** *Fix (q, ℓ), a true target conditional feature mean a, and its estimate $\widehat{a}$. Suppose that, on one event,*

$$\|\widehat{A} - A\|_{\mathrm{op}} \le \epsilon_A, \qquad \|\widehat{b} - b\|_2 \le \epsilon_b, \qquad \|\widehat{a} - a\|_2 \le \epsilon_a, \qquad \|\theta\|_2 \le B, \tag{128}$$

where $b = A\theta$ and A has p columns. Let $\widehat{s} = \sigma_p(\widehat{A})$, set to zero if $\widehat{A}$ has fewer than p rows, and put $\rho_b = \epsilon_b + B\epsilon_A$. If $\widehat{s} > \epsilon_A$, then A has full column rank, $\sigma_p(A) \ge \widehat{s} - \epsilon_A$, and

$$\|\widehat{A}^{\dagger}\widehat{b} - \theta\|_2 \le \frac{\rho_b}{\widehat{s}}. \tag{129}$$

Consequently the plug-in target estimate obeys

$$|\widehat{a}^{\top}\widehat{A}^{\dagger}\widehat{b} - a^{\top}\theta| \le \frac{\|\widehat{a}\|_2\rho_b}{\widehat{s}} + B\epsilon_a. \tag{130}$$

Without full rank, define the confidence set

$$\widehat{\mathcal{E}} = \{\vartheta : \|\vartheta\|_2 \le B,\ \|\widehat{A}\vartheta - \widehat{b}\|_2 \le \rho_b\}. \tag{131}$$

For a truncated SVD retaining singular values above a threshold $\tau \ge \epsilon_A$, let $\widehat{A}_\tau^{\dagger}$ be its pseudoinverse and $\widehat{P}_\tau$ its row-space projector. Define

$$\widehat{d}_\tau(\widehat{a}) = \|(I - \widehat{P}_\tau)\widehat{a}\|_2, \qquad \widehat{L}_\tau(\widehat{a}) = \|(\widehat{A}_\tau^{\dagger})^{\top}\widehat{P}_\tau\widehat{a}\|_2.$$

Then $\theta \in \widehat{\mathcal{E}}$. Define the target interval inflated for conditional-feature-mean estimation by

$$\widehat{I}_\star = \left[\inf_{\vartheta\in\widehat{\mathcal{E}}} \widehat{a}^{\top}\vartheta - B\epsilon_a,\ \sup_{\vartheta\in\widehat{\mathcal{E}}} \widehat{a}^{\top}\vartheta + B\epsilon_a\right]. \tag{132}$$

It contains $a^{\top}\theta$, and its half-diameter is at most

$$\widehat{L}_\tau(\widehat{a})\rho_b + B\widehat{d}_\tau(\widehat{a}) + B\epsilon_a. \tag{133}$$

Moreover, for any population subspace S with projector P_S, $\|(I - P_S)a\|_2 \le \|(I - P_S)\widehat{a}\|_2 + \epsilon_a$; thus an estimated row cannot be declared exactly in a population row space without accounting for ϵ_a.

Proof. Weyl's inequality gives $\sigma_p(A) \ge \widehat{s} - \epsilon_A$, proving the population-rank certificate. Moreover,

$$\|\widehat{b} - \widehat{A}\theta\|_2 \le \|\widehat{b} - b\|_2 + \|\widehat{A} - A\|_{\mathrm{op}}\|\theta\|_2 \le \rho_b.$$

When $\widehat{A}$ has full column rank, multiplying this residual by $\widehat{A}^{\dagger}$ proves (129). The same display proves $\theta \in \widehat{\mathcal{E}}$. The triangle inequality then gives

$$|\widehat{a}^{\top}(\widehat{A}^{\dagger}\widehat{b} - \theta) + (\widehat{a} - a)^{\top}\theta| \le \|\widehat{a}\|_2\frac{\rho_b}{\widehat{s}} + B\epsilon_a,$$

which is (130).

For any $\vartheta_1, \vartheta_2 \in \widehat{\mathcal{E}}$, put $h = \vartheta_1 - \vartheta_2$. Then $\|h\|_2 \le 2B$ and $\|\widehat{A}h\|_2 \le 2\rho_b$. With $a_P = \widehat{P}_\tau\widehat{a}$ and $\lambda = (\widehat{A}_\tau^{\dagger})^{\top}a_P$, the SVD identities give $\lambda^{\top}\widehat{A} = a_P^{\top}$. Therefore

$$|\widehat{a}^{\top}h| \le \|\lambda\|_2\|\widehat{A}h\|_2 + \|(I - \widehat{P}_\tau)\widehat{a}\|_2\|h\|_2 \le 2\widehat{L}_\tau(\widehat{a})\rho_b + 2B\widehat{d}_\tau(\widehat{a}).$$

Since $|(a - \widehat{a})^{\top}\theta| \le B\epsilon_a$ for $\theta \in \widehat{\mathcal{E}}$, the added $2B\epsilon_a$ guarantees coverage in (133); the triangle inequality and orthogonal-projection contractivity give the stated subspace bound. □

The radii in (128) must be simultaneous over every pair and target conditional feature mean searched by CHORD and must respect sample dependence. Set $\epsilon_a = 0$ when that conditional feature mean is known by design. Cross-fitting, block concentration, or cluster-robust resampling can supply such radii under their respective assumptions. Empirical near-membership in a row space is not exact population membership, so in the rank-deficient target-specific setting the interval (132) is the certificate.

E.9 Multilinear moment identification by Möbius inversion

The following structured model gives a constructive, explicitly scoped sense in which environment heterogeneity can lower the source moment order. It is not needed for the general rank results above.

Let $\mathcal{S}_q = \{S \subseteq [d] : |S| \le q\}$ and consider the stable multilinear mean centered at a

$$m_\theta(x) = \sum_{S \in \mathcal{S}_q} \theta_S \prod_{i \in S} (x_i - a_i). \tag{134}$$

In environment e, suppose the coordinates are independent, with mean μ_i^e, variance $v_i^e > 0$, and $\nu_i^e = \mu_i^e - a_i$. For $|T| \le \ell$, use the normalized centered-product source moment feature

$$\psi_T^e(X) = \frac{\prod_{i \in T}(X_i - \mu_i^e)}{\prod_{i \in T} v_i^e},$$

where the empty product equals one. If $\mathbb{E}_e[Y \mid X] = m_\theta(X)$, then

$$\beta_T^e := \mathbb{E}_e[Y \psi_T^e(X)] = \sum_{S \supseteq T} \theta_S \prod_{j \in S \setminus T} \nu_j^e. \tag{135}$$

Indeed, expand each factor as $X_i - a_i = (X_i - \mu_i^e) + \nu_i^e$. Independence and centering make distinct centered multiaffine products orthogonal; division by the variances normalizes the surviving product. The conditional-mean residual is orthogonal to every source moment feature, yielding (135). These equations are precisely a structured instance of (112).

Proposition E.13 (Explicit tradeoff between environment diversity and source moment order)**.** *Fix $0 \le \ell \le q$ and a nonzero c. Suppose the source collection contains one product environment for every $C \subseteq [d]$ with $|C| \le q - \ell$, having mean offset $\nu^C = c\mathbf{1}_C$ and positive coordinate variances. Then moments $\{\beta_T^C : |T| \le \ell\}$ identify every coefficient in (134); equivalently, $A_{q,\ell}$ has full column rank. The construction uses*

$$E_{q,\ell}^{\mathrm{constr}} = \sum_{j=0}^{q-\ell} \binom{d}{j} \tag{136}$$

environments. This construction establishes existence at the stated environment count.

Proof. Fix $S \in \mathcal{S}_q$. Choose $T \subseteq S$ with $|T| = \min\{\ell, |S|\}$ and put $B = S \setminus T$, so $|B| \le q - \ell$. For each $C \subseteq B$, the transport identity reduces to

$$\beta_T^C = \sum_{D \subseteq C} c^{|D|} \theta_{T \cup D}, \tag{137}$$

because an offset outside C is zero and $C \cap T = \varnothing$. Boolean lattice Möbius inversion gives

$$\theta_S = c^{-|B|} \sum_{C \subseteq B} (-1)^{|B|-|C|} \beta_T^C. \tag{138}$$

Möbius inversion recovers every θ_S from the order-ℓ moments, proving injectivity. □

The raw row-count condition

$$E \sum_{j=0}^{\ell} \binom{d}{j} \ge \sum_{j=0}^{q} \binom{d}{j}$$

is necessary for dense full-column-rank recovery, but the quadratic theorem shows it is not sufficient. Feasibility of the Möbius construction is also explicit: for Bernoulli covariates, take an interior anchor a and choose $|c|$ small enough that every mean $a + c\mathbf{1}_C$ remains coordinatewise in $(0, 1)$; all variances are then positive. Since the entries of $A_{q,\ell}$ are polynomial in the means, the nonzero minor constructed above remains nonzero under sufficiently small perturbations. Thus the construction also establishes full rank on an open set of nearby designs.

Equation (138) lowers source moment order; deployment still uses the full-order vector $\phi_q(z_\star)$.

E.10 THE TWO-ORDER FEASIBLE REGION AND OBSERVATION SCOPE

For the direct setting, let p_q be the output dimension of ϕ_q, define $q^\star_{\mathrm{D}}$ as the smallest order admitting the exact factorization $\eta(H) = \phi_q(Z_q)^\top \theta_q$, and let q_{eff} be the largest order passing the declared target-support check. Let $\ell^{\mathrm{dir}}_{\mathrm{id}}(q; E) = \min\{\ell : \mathrm{rank}(A^{\mathrm{dir},(E)}_{q,\ell}) = p_q\}$ and let $\ell_{\max}(q)$ be the largest estimable source moment order. Assume exact direct feature spaces are nested in q, source moments are nested in ℓ, and the target-support check passes precisely for orders no larger than q_{eff}. Under these stated monotonicity assumptions, the exact feasible set is

$$\boxed{\mathcal{W}^{\mathrm{dir}}_E = \{(q,\ell) : q^\star_{\mathrm{D}} \le q \le q_{\mathrm{eff}},\ \ell^{\mathrm{dir}}_{\mathrm{id}}(q; E) \le \ell \le \ell_{\max}(q)\}.} \tag{139}$$

Indeed, full-parameter recovery in the target-agnostic factorized setting requires $q \ge q^\star_{\mathrm{D}}$, while the declared target support certificate requires $q \le q_{\mathrm{eff}}$. At a fixed such q, dense recovery from the chosen moments requires $\ell \ge \ell^{\mathrm{dir}}_{\mathrm{id}}(q; E)$ and moment estimability requires $\ell \le \ell_{\max}(q)$. Intersecting these independent requirements gives (139). Under correct specification, consistent moment estimation, and a nonvanishing singular-value margin, these conditions are also sufficient for the corresponding plug-in estimator. For one prescribed target, full rank is replaced by its row-space condition and finite-sample accuracy is governed by (133). The target-adaptive setting is not ordered by $q^\star_{\mathrm{D}}$: if the offline conditional-feature-mean map is available, (19) shows that its full-state approximation term is $V_{\star,q} \ge 0$, zero exactly on target-almost-every set of merged states with a common outcome mean. Consequently, environment diversity can lower $\ell^{\mathrm{dir}}_{\mathrm{id}}$ (or the source moment order needed to identify the target conditional feature mean), remove a target-relevant null direction, or reduce leverage at a fixed q; it cannot lower $q^\star_{\mathrm{D}}$, enlarge the target-supported order range, recover information erased at fixed q, or repair $q^\star_{\mathrm{D}} > q_{\mathrm{eff}}$ for the declared direct prediction setting.

Moment validity precedes rank. The stable full-state conditional mean $\mathbb{E}_e[Y \mid H] = \chi(H)^\top \theta$ implies (111) for every integrable, H-measurable source moment feature. In the direct prediction setting, the weaker statement $\mathbb{E}_e[Y \mid Z_q] = \phi_q(Z_q)^\top \theta_q$ suffices when $\psi_{q,\ell}$ is measurable with respect to Z_q. A source moment feature that uses additional source variables requires the stronger condition. Alternatively one may postulate the required moment orthogonality condition directly. Rank cannot validate an invalid moment condition or turn a misspecified feature map into the target conditional mean.

The observation model is declared. Rank deficiency proves nonidentification from the released $(A_{q,\ell}, b_{q,\ell})$, or for a learner restricted to those moments. It does not by itself produce two mechanisms with the same unrestricted raw-data law. Likewise, identification from mixture weights assumes W_q is known by design or estimable from unlabelled full contexts. If both W_q and k are unknown while only projected label means are released, the resulting system is bilinear and the displayed linear rank test identifies neither factor.

Dependence and misspecification. No covariate independence is needed for $b_{q,\ell} = A_{q,\ell}\theta$ once (111) is assumed; independence is used only for the closed-form moment identity (135). If instead $b_{q,\ell} = A_{q,\ell}\theta + r_{\mathrm{app}}$, deterministic parameter bounds acquire an additional $\|r_{\mathrm{app}}\|_2/\widehat{s}$ term. Dependence among candidates or hyperedges changes the valid radii and effective sample size while leaving the population row/kernel geometry unchanged.

Deployment and validation remain separate. A small ℓ never licenses hidden order-q test-time features, and source rank does not certify target support. In the direct, target-agnostic data-access setting, any data-driven selection claiming full-state causal sufficiency additionally needs a coverage certificate that rejects or charges every $q < q^\star_{\mathrm{D}}$ under the declared direct relation. The target-adaptive setting is not constrained by this direct-feature threshold: with a certified offline conditional feature mean it may estimate the projected Bayes mean at any declared q, but it must report $V_{\star,q}$ relative to the Bayes rule given H and may call the projection causally sufficient only when zero is certified. If no pair has valid moments, target support, and a rank or target-interval certificate, the sound output is an equivalence class, a target interval, or abstention.

F Proofs for Set-Size-Dependent Pooling and Pure High-Order Testing

F.1 Two consequences for higher-order aggregation

Pure r-way dependence. Fix $r \geq 2$ and a candidate r-set. For outside node $j \in [n] := \{1, \dots, n\}$, let $X_j \in \{0,1\}^r$ record its replacement-edge indicators. Draw Y uniformly from $\{-1, +1\}$ and, given $Y = y$, draw rows independently with

$$P_y(x) = P_0(x)\left[1 + y\tau \prod_{i=1}^{r}(x_i - \rho)\right], \quad P_0 = \text{Bernoulli}(\rho)^{\otimes r}, \quad 0 < \rho < 1, \quad 0 < \tau < 1, \tag{140}$$

$$I_n = n\tau^2\{\rho(1-\rho)\}^r. \tag{141}$$

Only joint r-way coincidences carry label information; every proper marginal is label-invariant.

Theorem F.1 (Pure r-way testing threshold)**.** *For fixed r and $0 < \tau \leq \tau_0 < 1$, both the equal-prior Bayes error and the smallest worst-class error tend to $1/2$ when $I_n \to 0$, to 0 when $I_n \to \infty$, and stay bounded away from both limits when $I_n = \Theta(1)$. If τ is fixed, $\rho = n^{-\alpha}$, and $1/r < \alpha < 1$, each replacement-face count $\sum_{j=1}^{n} X_{ji}$ diverges in expectation although the complete array becomes uninformative. Appendix F gives the exact laws, an asymptotically sufficient count, and a face-versus-degree separation.*

Discarding set size can reverse decisions. Let the observed positive-integer size K be independent of $Y \in \{-1, +1\}$, with $P(Y = 1) = \pi > 1/2$. Given (Y, K), let the $X_i \in \{-1, +1\}$ be independent with $P(X_i = Y \mid Y, K) = 1 - \varepsilon > 1/2$. For $M_K = K^{-1}\sum_{i=1}^{K} X_i$, $b = \log\frac{\pi}{1-\pi}$, and $\lambda = \log\frac{1-\varepsilon}{\varepsilon}$,

$$\eta_K(m) := P(Y = 1 \mid M_K = m, K) = \sigma(b + K\lambda m), \qquad \sigma(u) := (1 + e^{-u})^{-1}. \tag{142}$$

The factor K records how much evidence supports the same observed average.

Proposition F.2 (Set-size erasure reverses decisions)**.** *For $m < 0$ attainable at supported sizes $K_1 < K_2$, decisions reverse when $K_1\lambda|m| < b < K_2\lambda|m|$. On the shared cell $M_K = m$ restricted to $K \in \{K_1, K_2\}$, set $w := P(K = K_1 \mid M_K = m, K \in \{K_1, K_2\}) \in (0,1)$. Every real-valued predictor $g(M_K)$, scored by $(g(M_K) - \mathbf{1}\{Y = 1\})^2$, has conditional Brier regret at least*

$$w(1-w)\{\eta_{K_1}(m) - \eta_{K_2}(m)\}^2. \tag{143}$$

Appendix F proves the bound and extends the same size-erasure mechanism to product pooling.

This appendix makes explicit the scope of both separations. The set-size result assumes that the observed size is discarded by the shared head. The testing lower bound is for the correlated-edge construction in (140); the face–degree separation instead uses the conditionally independent model in (159). Neither statement claims that all hypergraph learners fail.

F.2 Bayes decisions after discarding set size and information loss under product pooling

We first record all assumptions hidden in the compact description in the main text. The set size K is observed and is independent of the label (equivalently, $\Pr(Y = 1 \mid K) = \pi$). Conditional on (Y, K), the variables $X_1, \dots, X_K \in \{-1, +1\}$ are independent and satisfy $\Pr(X_i = Y \mid Y, K) = 1 - \varepsilon$. Thus the member channel is shared across sizes; only the number of independent pieces of evidence changes.

Lemma F.3 (Set-size-dependent posterior)**.** *Under these assumptions,*

$$\log\frac{\Pr(Y = 1 \mid X_{1:K}, K)}{\Pr(Y = -1 \mid X_{1:K}, K)} = b + \lambda\sum_{i=1}^{K} X_i = b + K\lambda M_K, \tag{144}$$

where $b = \log\{\pi/(1-\pi)\}$ and $\lambda = \log\{(1-\varepsilon)/\varepsilon\}$. Hence (K, M_K) is sufficient for Y.

Proof. For $x \in \{-1, +1\}$, the one-member log likelihood ratio is

$$\log \frac{\Pr(X_i = x \mid Y = 1, K)}{\Pr(X_i = x \mid Y = -1, K)} = x\lambda.$$

Multiplying the likelihood ratios gives (144), whose logistic transform is (142). □

Proof of Proposition F.2. Write $\eta_K(m) = \sigma(b + K\lambda m)$. At a common attainable $m < 0$,

$$K_1\lambda|m| < b < K_2\lambda|m| \quad \Longleftrightarrow \quad b + K_1\lambda m > 0 > b + K_2\lambda m.$$

Thus the same mean representation induces opposite Bayes decisions. For the Brier claim, condition on the representation cell $M_K = m$ and on $K \in \{K_1, K_2\}$, whose conditional weights are w and $1 - w$. If a set-size-agnostic head outputs $c \in \mathbb{R}$ on this cell, its excess Brier risk over the set-size-conditioned posterior equals

$$\begin{aligned} &w\{c - \eta_{K_1}(m)\}^2 + (1 - w)\{c - \eta_{K_2}(m)\}^2 \\ &\quad = [c - w\eta_{K_1}(m) - (1 - w)\eta_{K_2}(m)]^2 + w(1 - w)\{\eta_{K_1}(m) - \eta_{K_2}(m)\}^2. \end{aligned} \tag{145}$$

Dropping the first nonnegative term proves (143). Retaining K recovers the posterior in (142), confirming that the regret is caused exactly by the imposed erasure of set size.

For product pooling, write $X_i = YE_i$. The E_i are independent and identically distributed (IID), independent of Y, and satisfy $\Pr(E_i = 1) = 1 - \varepsilon$ and $\mathbb{E}E_i = a := 1 - 2\varepsilon$. Hence

$$T_K = Y^K \prod_{i=1}^{K} E_i. \tag{146}$$

If K is even, $T_K = \prod_i E_i$ and is independent of Y. If K is odd, $T_K = Y \prod_i E_i$ and

$$\Pr(T_K = t \mid Y = y, K) = \frac{1 + tya^K}{2}.$$

Bayes' rule therefore gives

$$\Pr(Y = 1 \mid T_K = t, K) = \sigma\left(b + t \log \frac{1 + a^K}{1 - a^K}\right), \tag{147}$$

Product pooling erases label information for even K and leaves an a^K-attenuated contrast for odd K. This claim is specific to T_K; predictors observing K and $X_1, \ldots, X_K$ lie outside its scope. □

The shared-cell condition is non-vacuous. Let $d \geq 3$ be odd, set $m = -1/d$, and choose $K_1 < K_2$ to be multiples of d. Then $K_j m$ is an integer with the parity required of a sum of K_j signs, so m is attainable at both sizes. Any b strictly between $K_1\lambda/d$ and $K_2\lambda/d$ realizes the reversal.

F.3 The pure r-way dependence experiment

Put

$$h_\rho(x) := \prod_{i=1}^{r}(x_i - \rho), \qquad v_\rho := \mathbb{E}_0 h_\rho(X)^2 = \{\rho(1 - \rho)\}^r,$$

where $P_0 = \text{Bernoulli}(\rho)^{\otimes r}$. Throughout this subsection $r \geq 2$ is fixed and $0 < \tau \leq \tau_0 < 1$.

Lemma F.4 (Every proper marginal is label-invariant)**.** *The two laws $dP_y/dP_0 = 1 + y\tau h_\rho$, $y \in \{-1, +1\}$, are strictly positive probability distributions. For every $A \subsetneq [r]$,*

$$P_y(X_A = x_A) = \prod_{i \in A} \rho^{x_i}(1 - \rho)^{1 - x_i}, \tag{148}$$

independently of y.

Proof. Independence and centering give $\mathbb{E}_0 h_\rho = 0$, so each density integrates to one. Also $|h_\rho(x)| \le 1$, whence $1 + y\tau h_\rho(x) \ge 1 - \tau_0 > 0$. For proper A, at least one coordinate is integrated out, and

$$\begin{aligned} P_y(X_A = x_A) &= P_{0,A}(x_A) \left[1 + y\tau \prod_{i \in A} (x_i - \rho) \mathbb{E}_0 \prod_{j \notin A} (X_j - \rho) \right] \\ &= P_{0,A}(x_A). \end{aligned}$$

Consequently, with independent and identically distributed (IID) outside nodes, each individual face support is exactly $\mathrm{Bin}(n, \rho)$ in both classes even though their joint r-way dependence may differ. □

The construction is a genuine hypergraph distribution: for a fixed candidate r-set, coordinate i of X_j is the replacement edge obtained by swapping candidate member i for outside node j. Distinct outside nodes use disjoint edge sets. Independent examples can be placed on disjoint vertex gadgets. This establishes realizability, but does not assert an exchangeable independent-edge realization for many overlapping roots.

Lemma F.5 (One-sample divergence scale)**.** *Write $P_+ := P_{+1}$ and $P_- := P_{-1}$, and let D_{KL} denote Kullback–Leibler divergence. With squared Hellinger distance normalized as $H^2(P, Q) = 1 - \sum_x \sqrt{P(x)Q(x)}$,*

$$\frac{1}{2}\tau^2 v_\rho \le H^2(P_+, P_-) \le \frac{\tau^2 v_\rho}{1 + \sqrt{1 - \tau_0^2}}, \tag{149}$$

$$\tau^2 v_\rho \le D_{\mathrm{KL}}(P_+ \| P_-) \le \frac{4\tau^2 v_\rho}{1 - \tau_0}. \tag{150}$$

The Kullback–Leibler bounds also hold with P_+ and P_- interchanged.

Proof. With $u = \tau h_\rho$, direct substitution gives

$$H^2(P_+, P_-) = \mathbb{E}_0\{1 - \sqrt{1 - u^2}\}.$$

For $0 \le t \le \tau_0^2$,

$$\frac{t}{2} \le 1 - \sqrt{1 - t} = \frac{t}{1 + \sqrt{1 - t}} \le \frac{t}{1 + \sqrt{1 - \tau_0^2}}.$$

Taking expectations proves (149). Moreover, $D_{\mathrm{KL}}(P\|Q) \ge -2 \log \sum_x \sqrt{P(x)Q(x)} \ge 2H^2(P, Q)$, while

$$\begin{aligned} D_{\mathrm{KL}}(P_+ \| P_-) &\le \chi^2(P_+ \| P_-) \\ &= 4\tau^2 \mathbb{E}_0 \frac{h_\rho^2}{1 - \tau h_\rho} \le \frac{4\tau^2 v_\rho}{1 - \tau_0}. \end{aligned}$$

Replacing h_ρ by $-h_\rho$ proves the reverse bound. □

Bayes and genuine minimax risks. For a randomized test ϕ (interpreted as its probability of deciding $+$), define

$$R_{B,n} := \inf_\phi \frac{1}{2}\{P_+^n(\phi = -) + P_-^n(\phi = +)\} = \frac{1 - \mathrm{TV}(P_+^n, P_-^n)}{2}, \tag{151}$$

$$R_{M,n} := \inf_\phi \max\{P_+^n(\phi = -), P_-^n(\phi = +)\}. \tag{152}$$

The two finite-sample quantities differ: only $R_{M,n}$ is a worst-class minimax risk.

Proof of the testing claims in Theorem F.1. Let $I_n = n\tau^2 v_\rho$. If $I_n \to 0$, tensorization of (150) and Pinsker's inequality imply $\mathrm{TV}(P_+^n, P_-^n) \to 0$. Thus $R_{B,n} \to 1/2$; since $R_{B,n} \le R_{M,n} \le 1/2$, also $R_{M,n} \to 1/2$.

If $I_n \to \infty$, consider $T_n = \sum_{j=1}^n h_\rho(X_j)$. Under class y,

$$\mathbb{E}_y T_n = yn\tau v_\rho, \qquad \mathrm{Var}_y(T_n) \le n(1+\tau_0)v_\rho.$$

Chebyshev's inequality bounds each conditional error of $\mathrm{sign}(T_n)$ by $(1+\tau_0)/I_n$. Hence both risks tend to zero. Finally suppose $0 < c \le I_n \le C < \infty$. The Hellinger affinity $\mathcal{A}(P,Q) = \sum_x \sqrt{P(x)Q(x)}$ tensorizes, so

$$\mathcal{A}(P_+^n, P_-^n) = \{1 - H^2(P_+, P_-)\}^n.$$

Equation (149) keeps this quantity bounded away from both zero and one. The inequalities

$$H^2(P,Q) \le \mathrm{TV}(P,Q) \le \sqrt{1 - \mathcal{A}(P,Q)^2}$$

therefore keep $\mathrm{TV}(P_+^n, P_-^n)$ bounded away from both zero and one, proving the nondegeneracy of $R_{B,n}$. The lower bound for $R_{M,n}$ follows from $R_{M,n} \ge R_{B,n}$. For its upper bound, choose an event A with $P_+^n(A) - P_-^n(A) = \delta_n := \mathrm{TV}(P_+^n, P_-^n)$. The randomized test

$$\phi(x) = c_n + \frac{1}{2}\mathbf{1}_A(x), \qquad c_n = \frac{1 - \{P_+^n(A) + P_-^n(A)\}/2}{2},$$

takes values in $[0,1]$ and equalizes the two conditional errors at $1/2 - \delta_n/4$. The lower bound on δ_n therefore keeps $R_{M,n}$ uniformly and strictly below $1/2$ throughout the critical regime. □

Lemma F.6 (The all-one row count and sparse asymptotic sufficiency)**.** *For $W = \sum_{j=1}^n \prod_i X_{ji}$,*

$$W \mid Y = y \sim \mathrm{Bin}(n, p_y), \qquad p_y = \rho^r\{1 + y\tau(1-\rho)^r\}. \tag{153}$$

If $\rho \le 1/2$, its testing transition is governed, up to constants depending only on r and τ_0, by the same I_n as the full array. Moreover, if $\rho \to 0$ and $I_n = O(1)$, the binary experiment generated by W is asymptotically sufficient for the full array. In the particular window $n\rho^r \to \lambda \in (0,\infty)$ and fixed $\tau \in (0,1)$,

$$W \mid Y = y \Rightarrow \mathrm{Pois}\{\lambda(1 + y\tau)\}. \tag{154}$$

Proof. The event $\prod_i X_i = 1$ is the all-one cell, where $h_\rho(\mathbf{1}) = (1-\rho)^r$. Its probability under P_y is p_y above; IID sampling over j proves (153). The midpoint rule that decides $+$ when W exceeds $n\rho^r$ has each conditional error at most

$$\frac{1+\tau_0}{n\tau^2\rho^r(1-\rho)^{2r}}. \tag{155}$$

For fixed r and $\rho \le 1/2$, the denominator is within constant factors of $I_n = n\tau^2\rho^r(1-\rho)^r$. In the vanishing regime no statistic can outperform the full-array lower bound. In the critical regime, direct Hellinger bounds for the two Bernoulli laws in (153) make this explicit. With $\Delta = p_+ - p_- = 2\tau\rho^r(1-\rho)^r$, the identity

$$H^2\{\mathrm{Bernoulli}(p_+), \mathrm{Bernoulli}(p_-)\} = \frac{1}{2}\left\{(\sqrt{p_+} - \sqrt{p_-})^2 + (\sqrt{1-p_+} - \sqrt{1-p_-})^2\right\}$$

and $p_\pm \asymp \rho^r$ give

$$H^2\{\mathrm{Bernoulli}(p_+), \mathrm{Bernoulli}(p_-)\} \asymp \frac{\Delta^2}{\rho^r} \asymp \tau^2\rho^r(1-\rho)^{2r}. \tag{156}$$

The constants depend only on (r, τ_0). This is comparable to I_n/n when $\rho \le 1/2$, so tensorization gives the same three-regime transition. For sufficiency, set $B = \mathbf{1}\{X = \mathbf{1}\}$ and let $Q_0 = P_0(\cdot \mid B = 0)$. Conditional on $B = 0$, $|h_\rho(X)| \le \rho$, and

$$\mathbb{E}_0[h_\rho^2 \mathbf{1}\{B = 0\}] = \rho^r(1-\rho)^r\{1 - (1-\rho)^r\} = O(\rho^{r+1}). \tag{157}$$

Writing $\mu_0 = \mathbb{E}_{Q_0} h_\rho$, the two conditional densities relative to Q_0 are

$$q_y(x) = \frac{1 + y\tau h_\rho(x)}{1 + y\tau\mu_0}.$$

Here $|\mu_0| = O(\rho^r)$, and direct subtraction of q_+ and q_- followed by $D_{\mathrm{KL}} \le \chi^2$ gives

$$D_{\mathrm{KL}}(q_+Q_0 \| q_-Q_0) \le C_{\tau_0}\tau^2 \mathbb{E}_{Q_0}\{h_\rho - \mu_0\}^2.$$

Equation (157) therefore yields

$$D_{\mathrm{KL}}\{P_+(X \mid B=0) \| P_-(X \mid B=0)\} = O(\tau^2\rho^{r+1}), \tag{158}$$

and likewise in reverse. By the KL chain rule, mapping each observation to B discards at most $O(n\tau^2\rho^{r+1})$. If $\rho \to 0$ and $I_n = O(1)$, this is

$$O\left(I_n \frac{\rho}{(1-\rho)^r}\right) = o(1).$$

The binary-vector likelihood depends only on W, so the residual experiment is asymptotically negligible. Sufficiency follows, while $np_y \to \lambda(1+y\tau)$ and $p_y \to 0$ give the Poisson limit. □

The all-one count is the appropriate statistic on the sparse side used in the paper. If instead $\rho \to 1$, the factor $(1-\rho)^r$ in (156) is no longer constant relative to the full information scale; the dual all-zero count then becomes the appropriate row-count statistic.

Combining Lemmas F.4–F.6 completes Theorem F.1. In particular, for fixed nonzero τ and $\rho = n^{-\alpha}$ with $1/r < \alpha < 1$,

$$n\rho \longrightarrow \infty, \qquad I_n \asymp n\rho^r \longrightarrow 0.$$

Every individual replacement-face count $\sum_{j=1}^n X_{ji}$ therefore has mean $n\rho \to \infty$, while even the complete array has Bayes and worst-class minimax error tending to $1/2$.

F.4 Face support versus node degrees

Let $Z_1, \dots, Z_N$ be independent and identically distributed Rademacher variables and, conditional on them, generate independent r-edge indicators by

$$\Pr(A_C = 1 \mid Z_C) = \rho_N \left(1 + \frac{\beta}{r}\sum_{i \in C} Z_i\right), \qquad \rho_N = N^{-\alpha}, \quad 0 < \beta < 1. \tag{159}$$

Here $r \ge 3$ and $0 < \beta < 1$ are fixed, and $\rho_N(1+\beta) \le 1$ for all large N. The queried edge is always withheld, so no representation receives the candidate's own indicator as an input.

Definition F.7 (Predictors restricted to candidate-face incidences)**.** A predictor in this class is a shared, permutation-equivariant function of the rooted incidence object obtained by connecting each of a candidate's r abstract $(r-1)$-faces only to observed hyperedges containing that face. It receives constant initial features and no node identity, node degree, global statistic, positional feature, or lower-order skip connection. Its parameters may depend on an independent training sample. Consequently, every candidate whose rooted face incidences are all empty receives the same score.

Theorem F.8 (Candidate-face-incidence-restricted and degree-based predictors separate)**.** *If $1 < \alpha < r-1$, every predictor restricted as in Definition F.7 has balanced edge-versus-nonedge AUROC $1/2 + o(1)$. Thresholded node degrees recover a $1-o(1)$ fraction of the latent signs, and their score satisfies*

$$\mathrm{AUROC}_{\mathrm{degree}} = \frac{1}{2} + \beta c_r + o(1), \qquad c_r = \frac{\binom{2r}{r}}{2\,4^r} > 0. \tag{160}$$

Thus candidate-face incidence vanishes before node degrees lose their signal. The interval exists iff $r \ge 3$, so this construction has no graph analogue.

Lemma F.9 (Vanishing candidate-face support)**.** *For a fixed candidate C, after withholding A_C,*

$$\Pr(\textit{some } (r-1)\textit{-face of } C \textit{ has positive support} \mid Z, C, A_C) \le r(N-r)(1+\beta)\rho_N. \tag{161}$$

Thus if $N\rho_N \to 0$, all candidate faces are empty with probability $1-o(1)$ under the uniform-candidate, edge-case, and nonedge-control laws.

Proof. Each $(r-1)$-face has $N-r$ admissible extensions other than C, and every extension has conditional probability at most $(1+\beta)\rho_N$. Conditional edge independence makes these other edges independent of the withheld A_C. Markov's inequality followed by a union bound over the r faces gives (161); the resulting inequality is uniform in Z, the candidate C, and the withheld value A_C. □

Lemma F.10 (Degree recovery)**.** *Let*

$$D_i = \sum_{C \ni i} A_C, \qquad M_N = \binom{N-1}{r-1}, \qquad \Lambda_N = M_N \rho_N,$$

and set $\widehat{Z}_i = \operatorname{sign}(D_i - \Lambda_N)$*, with arbitrary tie breaking. If* $\Lambda_N \to \infty$*, then for constants depending only on* (r, β)*,*

$$\Pr(\widehat{Z}_i \neq Z_i) \le 2e^{-c_{r,\beta}\Lambda_N} + 2e^{-c_r N} \tag{162}$$

for a uniform node and all large N*. The conclusion remains valid after withholding one queried edge, and the fraction of incorrectly recovered signs converges to zero in probability, denoted* $o_p(1)$*.*

Proof. Using $\binom{N-2}{r-2}/\binom{N-1}{r-1} = (r-1)/(N-1)$ gives

$$\mathbb{E}(D_i \mid Z) = \Lambda_N \left[1 + \frac{\beta}{r} Z_i + \frac{\beta}{r} \frac{r-1}{N-1} \sum_{j \neq i} Z_j\right]. \tag{163}$$

Hoeffding's inequality shows that

$$\left| \frac{r-1}{N-1} \sum_{j \neq i} Z_j \right| \le \frac{1}{2}$$

except with probability at most $2e^{-c_r N}$. On this event the conditional mean is on the Z_i side of Λ_N by at least $(\beta/(2r))\Lambda_N$. Conditional on Z, the degree is a sum of independent Bernoulli variables with variance at most $(1+\beta)\Lambda_N$. Bernstein's inequality proves (162). Withholding the queried edge changes a degree by at most one, negligible relative to the diverging margin. Exchangeability and Markov's inequality yield convergence in probability of the fraction of misclassified nodes. □

Lemma F.11 (Exact oracle case–control AUROC)**.** *Let* $S = \sum_{i \in C} Z_i$ *for a uniform candidate, and score candidates by* S*. If the positive is drawn conditional on* $A_C = 1$ *and the control conditional on* $A_C = 0$*, independently from their population laws, then with half credit for ties,*

$$\mathrm{AUROC}_\rho(S) = \frac{1}{2} + \frac{\beta}{1-\rho} c_r, \qquad c_r = \frac{\binom{2r}{r}}{2\,4^r}. \tag{164}$$

If controls are instead unconditional uniform candidates, the exact factor is βc_r*.*

Proof. Let $p(s)$ be the unconditioned law of the Rademacher sum S. Since $\mathbb{E}S = 0$ and $\Pr(A_C = 1) = \rho$, the case and nonedge-control laws are

$$p_+(s) = p(s)\left(1 + \frac{\beta}{r} s\right), \qquad p_-(s) = p(s)\left(1 - \frac{\rho\beta}{r(1-\rho)} s\right). \tag{165}$$

For IID unconditioned copies S, T, write $H(S,T) = \mathbf{1}\{S > T\} + \frac{1}{2}\mathbf{1}\{S = T\}$ and $C_r = \mathbb{E}\{SH(S,T)\}$. Exchangeability gives

$$\mathbb{E}H = \frac{1}{2}, \qquad \mathbb{E}\{TH(S,T)\} = -C_r, \qquad \mathbb{E}\{STH(S,T)\} = 0.$$

The last identity follows by multiplying $H(S,T) + H(T,S) = 1$ by ST and taking expectations. Expanding the two weights in (165) now yields

$$\mathrm{AUROC}_\rho(S) = \frac{1}{2} + \frac{\beta}{r(1-\rho)} C_r. \tag{166}$$

Furthermore,

$$C_r = \frac{1}{4}\mathbb{E}|S - T|.$$

Writing $S = 2K - r$ and $T = 2L - r$ for independent $K, L \sim \mathrm{Bin}(r, 1/2)$ gives

$$\frac{C_r}{r} = \frac{1}{2r}\mathbb{E}|K - L|.$$

Now $K - L$ has the law of $M - r$ for $M \sim \mathrm{Bin}(2r, 1/2)$, and the binomial telescoping identity

$$\mathbb{E}|M - r| = \frac{2}{4^r} \sum_{m=r+1}^{2r} (m - r) \binom{2r}{m} = \frac{r}{4^r} \binom{2r}{r}$$

follows from $(m - r)\binom{2r}{m} = r\{\binom{2r-1}{m-1} - \binom{2r-1}{m}\}$. Thus $C_r/r = \binom{2r}{r}/(2\,4^r) = c_r$, proving (164). Setting the control weight in (165) to one gives the unconditional-control formula. □

Proof of Theorem F.8. Set $\rho_N = N^{-\alpha}$ with $1 < \alpha < r - 1$. Since $N\rho_N \to 0$, Lemma F.9 implies that an independent case and nonedge control both have empty rooted face incidence with probability $1 - o(1)$. Conditional on the trained parameters, every score in Definition F.7 assigns the same value to these two roots. The pair therefore ties except on an event of probability $o(1)$, and its AUROC is $1/2 + o(1)$. On the other hand,

$$\Lambda_N = \binom{N-1}{r-1} \rho_N = \Theta(N^{r-1-\alpha}) \longrightarrow \infty.$$

Lemma F.10 recovers all but an $o_p(1)$ fraction of signs. The case law has Radon–Nikodym weight $1 + (\beta/r)S \in [1 - \beta, 1 + \beta]$ relative to a uniform candidate, and the nonedge-control weight in (165) is also uniformly bounded for large N. Hence a fixed-size case or control candidate contains a misrecovered sign with probability $o(1)$. The degree score $\widehat{S}_C = \sum_{i \in C} \widehat{Z}_i$ agrees with the oracle S_C with probability $1 - o(1)$ under both laws. Lemma F.11 and $\rho_N \to 0$ therefore give

$$\mathrm{AUROC}(\widehat{S}) = \frac{1}{2} + \beta c_r + o(1).$$

Substitution gives $c_3 = \binom{6}{3}/(2 \cdot 4^3) = 5/32$ and $c_4 = \binom{8}{4}/(2 \cdot 4^4) = 35/256$. The inequalities $1 < \alpha < r - 1$ are simultaneously feasible exactly when $r \geq 3$; for graphs ($r = 2$) there is no face–node separation interval. □

The area under the receiver operating characteristic curve (AUROC) statement is a population case–control result and does not assert concentration for an empirical AUROC computed from dependent candidates in one hypergraph. The candidate-face-incidence lower bound also excludes degree-aware, spectral, global, and lower-order-backoff methods; those exclusions are exactly what permits the positive degree construction.

G Finite-Sample Risk Bounds and the Frozen-Baseline Chord Wrapper

This section makes Algorithm 1 operational under the observation model in Section 3 and keeps three claims distinct. The first certifies excess Brier risk relative to the Bayes rule given H and abstains if no pair survives. The second compares frozen order-restricted candidates with the same frozen highest-order predictor under an assumed inequality that bounds the target paired Brier-risk difference by a convex combination of source differences plus an allowance; otherwise it returns that predictor unchanged. Without this bound on a natural stream, Section G.9 uses earlier validation windows and reserves the final window for one evaluation, without a finite-sample OOD or causal certificate. In the two finite-sample procedures, source sites expose aggregate moment equations, whereas a deployment candidate exposes only an order-q view. No central step receives record-level paired full-state and outcome records.

G.1 Order-restricted learned representations

Write the complete prediction input as $H = (G_{<t}, S)$, where $G_{<t}$ is the observed history and S is the candidate hyperedge. The order map restricts both components:

$$Z_q = T_q(G_{<t}, S), \qquad \sigma(Z_q) \subseteq \sigma(Z_{q+1}), \qquad T_{q_{\max}}(H) = H. \tag{167}$$

Both the historical structure and the candidate are transformed by the same declared T_q. Every order-q message, incidence embedding, and score must be recomputed from Z_q, or be proved $\sigma(Z_q)$-measurable. In particular, caching an embedding after an encoder has consumed full-order incidences and then masking only the candidate remains a full-information predictor. A backbone is compatible only if each attempted order can be recomputed from the declared $T_q(G_{<t}, S)$ without a hidden full-incidence cache.

Let $\mathcal{F}_{\mathrm{fit}}$ contain the data and randomness used to fit a backbone. Before any source-moment, support, specification, or paired-loss *certification* release, or any empirical order-selection window, is opened, freeze every attempted score

$$s_{qK} = f_{\widehat{\omega},q}(Z_q, K), \qquad g_{q,K} = \mathrm{clip}\{h_{qK}(s_{qK}), 0, 1\}, \tag{168}$$

including its arity-conditioned head h_{qK}. Conditional on $\mathcal{F}_{\mathrm{fit}}$ these are fixed functions, as required by Lemma E.4; a searched backbone, order, or arity must instead be covered by the simultaneous event. For each K, the highest predeclared order Q_K retains the same frozen baseline $g_{Q_K,K}$. A paired or validation-based comparison may replace it only through its predeclared rule below; otherwise that same predictor is returned unchanged.

G.2 Three deployment data-access settings

For each order, fix one of these data-access settings before inspecting source label moments.

Direct, target-agnostic setting D_q. There is a declared factorization

$$\eta(H) = \phi_q(Z_q)^\top \theta_q, \qquad \chi_q(H) := \phi_q\{\Pi_q(H)\}, \qquad p_q = \dim(\chi_q). \tag{169}$$

Consequently the target conditional feature mean is known pointwise, $a_{\star,q}(z) = \phi_q(z)$, its estimation error is zero, and $V_{\star,q} = 0$. The last equality is accepted only with the structural factorization in (169) and the declared source–target intervention coverage; a small empirical residual alone is not a lumpability certificate. This setting uses no full target contexts and is target-agnostic.

Target-adaptive setting T_q. Here the declared full-state feature map $\chi(H) \in \mathbb{R}^p$ satisfies $\eta(H) = \chi(H)^\top \theta$. An explicitly allowed, unlabeled sample of full target contexts is used *offline* to estimate

$$a_{\star,q}(z) = \mathbb{E}_\star[\chi(H) \mid Z_q = z], \qquad \sup_{z \in \mathcal{Z}_q^{\mathrm{cov}}} \|\widehat{a}_{\star,q}(z) - a_{\star,q}(z)\|_2 \le \epsilon_{a,q}. \tag{170}$$

The fitted conditional-feature-mean map and its simultaneous radius are stored before candidate evaluation. At candidate time the scorer receives only $Z_q = z$ and evaluates $\widehat{a}_{\star,q}(z)$; it never receives H. Target labels are not used in this data-access setting. This is target adaptation, and its risk relative to the Bayes rule given H contains the projection-variance term $V_{\star,q}$, which is positive exactly when the target gives mass to merged states with different outcome means.

The column dimensions remain setting-specific: the direct operator has $p_q = \dim(\phi_q)$ columns and parameter θ_q, whereas the target-adaptive full-state operator has $p = \dim(\chi)$ columns and parameter θ. We write $d_q = p_q$ under D_q and $d_q = p$ under T_q only when a dimension-generic statement is useful. If (169) is not certified and the offline conditional-feature-mean map in (170) is unavailable, then the procedure certifying excess risk relative to the Bayes rule given H does not identify a point score for an order q that merges states with different outcome means. It may report the robust interval over compatible projected means, but must reject that order and ultimately abstain if no other order survives. Validation-based selection may still compare a frozen order-q candidate with the same baseline, but it receives no causal or finite-sample OOD certificate from this argument.

Approximate direct setting. An optional $\delta^{\mathrm{rep}}_{q,\ell}$-approximate predictive extension is described below; it is admitted only by a one-sided error upper certificate, is charged for that error, and is never reported as causal lumpability. When the requested setting is exact, this extension is disabled.

G.3 Aggregate source releases without central record pairs

For source environment e, the source moment feature $\psi_{q,\ell}(H) \in \mathbb{R}^{s_{q,\ell}}$ is outcome-free and available only at the source. Put $\widetilde{\chi}_q = \phi_q \circ \Pi_q$ under D_q and $\widetilde{\chi}_q = \chi$ under T_q. Every setting releases

$$\begin{aligned} A^e_{q,\ell} &= \mathbb{E}_e[\psi_{q,\ell}\widetilde{\chi}_q^\top], & b^e_{q,\ell} &= \mathbb{E}_e[\psi_{q,\ell}Y], \\ b^e_{q,\ell} &= A^e_{q,\ell}\widetilde{\theta}_q, & &\text{exact } \mathsf{D}_q \text{ or } \mathsf{T}_q, \\ b^e_{q,\ell} &= A^e_{q,\ell}\vartheta^{\mathrm{rep}}_{q,\ell} + r^{e,\mathrm{mom}}_{q,\ell}, & &\text{approximate direct setting.} \end{aligned} \tag{171}$$

Here $\widetilde{\theta}_q = \theta_q$ for exact direct and $\widetilde{\theta}_q = \theta$ for exact target-adaptive settings. No zero residual identity is assumed in the approximate direct setting: the reference is frozen from training releases and the validation/specification checks bound $r^{e,\mathrm{mom}}_{q,\ell}$ one-sidedly. Thus the target-adaptive setting uses exactly $A^e_{q,\ell} = \mathbb{E}_e[\psi_{q,\ell}\chi^\top]$ with p columns; only the direct-factorization setting substitutes the deployable feature map $\phi_q(Z_q)$ and has p_q columns. Sites release estimates of these two objects from independent cohorts, or release the objects and dependence-valid covariance/radius summaries directly. In particular, $\widetilde{\chi}_q(H)$ and Y are never released as a paired record. This prevents the estimator from bypassing aggregate identification by directly regressing Y on the full-state feature map. Environment tags are part of the release design: without them, all sites contribute one pooled moment row, so heterogeneity cannot be credited with additional identification rank.

To display one admissible construction, let $U_{e,k}$ be an unlabelled context fold and $L_{e,k}$ an independent label-moment fold. With a fold-external row transformation $\widehat{W}^{(-k)}_{e,q,\ell}$, form

$$\begin{aligned} \widehat{A}^{(k)}_{q,\ell} &= \mathrm{stack}^E_{e=1} \frac{1}{|U_{e,k}|} \sum_{i \in U_{e,k}} \widehat{W}^{(-k)}_{e,q,\ell} \psi_{q,\ell}(H_i)\widetilde{\chi}_q(H_i)^\top, \\ \widehat{b}^{(k)}_{q,\ell} &= \mathrm{stack}^E_{e=1} \frac{1}{|L_{e,k}|} \sum_{j \in L_{e,k}} \widehat{W}^{(-k)}_{e,q,\ell} \psi_{q,\ell}(H_j)Y_j. \end{aligned} \tag{172}$$

The sets $U_{e,k}$ and $L_{e,k}$ are disjoint and need not contain matched units. The final $(\widehat{A}_{q,\ell}, \widehat{b}_{q,\ell})$ averages their cross-fitted blocks. Centering, nuisance fitting, and whitening are learned outside the contributing fold. The same row transformation is applied to A, b, validation residuals, covariance estimates, and radii; whitening only A changes the equation. If sites expose aggregates only, these operations happen at the provider and CHORD consumes their released summaries. Candidate feature maps, source moment features, anchors, and split roles are frozen before any b release is opened.

The target-adaptive sample is separate from both source cohorts. It is split into conditional-feature-mean estimation, support, and projection-variance roles whenever the same contexts would otherwise be reused. Neither target labels nor a paired source $(\widetilde{\chi}_q, Y)$ sample is introduced by this split.

G.4 Support, specification, and moment-validity checks

Source release support and target deployment support are distinct. For a pair (q, ℓ), let $n_{\mathrm{eff}}\{U_e; \psi_{q,\ell}\widetilde{\chi}_q^\top\}$ and $n_{\mathrm{eff}}\{L_e; \psi_{q,\ell}Y\}$ be dependence-adjusted effective counts for the two source aggregates. Define

$$\underline{n}^{\mathrm{src}}_{q,\ell} := \min_e \left[n_{\mathrm{eff}}\{U_e; \psi_{q,\ell}\widetilde{\chi}_q^\top\} \wedge n_{\mathrm{eff}}\{L_e; \psi_{q,\ell}Y\} \right]. \tag{173}$$

The source-support check $\mathsf{S}^{\mathrm{src}}_{q,\ell}$ requires this count to exceed its simultaneous, dimension-adjusted threshold. It controls the release radii for A and b and can shrink when ℓ increases. It says nothing about whether an order-q target cell will occur at deployment.

Separately, let $C_{\star,q}(F)$ be the number of observable extensions of a target q-cell F in a predeclared collection $\mathcal{F}_{\star,q}$. Fix a target quantile level $\alpha_{\mathrm{tar}} \in (0,1)$, an overall failure probability $\delta \in (0,1)$, the number $N_\star$ of target support units, and thresholds $c_{\mathrm{tar}} > 0$ and $\tau_{\mathrm{hole}} \in [0,1]$. Using simultaneous lower and upper confidence bounds (LCB and UCB), set

$$\underline{n}^{\mathrm{tar}}_q := \mathrm{LCB}\big[\mathrm{Quantile}_{\alpha_{\mathrm{tar}}}\{C_{\star,q}(F) : F \in \mathcal{F}_{\star,q}\}\big], \qquad \overline{\pi}^{\mathrm{hole}}_q := \mathrm{UCB}[P_\star\{C_{\star,q}(F) = 0\}]. \tag{174}$$

The target-support check $\mathsf{S}_q^{\mathrm{tar}}$ requires $\underline{n}_q^{\mathrm{tar}} \geq c_{\mathrm{tar}} \log(N_\star/\delta)$ and $\overline{\pi}_q^{\mathrm{hole}} \leq \tau_{\mathrm{hole}}$. Setting T_q estimates these quantities from its reserved unlabelled target contexts. Setting D_q instead uses a predeclared target-support uncertainty set or generative lower bound. We also allow a separately declared hybrid: its scorer remains the direct $\phi_q(Z_q)^\top \vartheta$, but unlabelled target contexts are consumed only by the deployment-support check. We label this selection target-adaptive support monitoring. Variable arity is checked in every deployment stratum.

For later risk accounting let Δ_q^{sup} be a simultaneous upper bound on the Brier contribution of unsupported and declared rare *target* cells. Since a clipped Brier loss is at most one, one valid construction is

$$\Delta_q^{\mathrm{sup}} = \overline{\pi}_q^{\mathrm{hole}} + \epsilon_q^{\mathrm{rare}}(\underline{n}_q^{\mathrm{tar}}, \delta), \tag{175}$$

where the second term is the appropriate bounded, cluster-, or mixing-valid target rare-cell radius at effective count $\underline{n}_q^{\mathrm{tar}}$ and failure level δ. It is not assigned an independent and identically distributed (IID) rate under overlapping contexts. The source count $\underline{n}_{q,\ell}^{\mathrm{src}}$ never enters this target-mass term; it enters $\eta_{q,\ell}$ below. Without a valid bound on one side, that side's support claim remains diagnostic.

Source-only representation-error check. Passing an independent source-validation residual check does not certify sufficient order. For each attempted pair, freeze a candidate coefficient $\vartheta_{q,\ell}^{\mathrm{rep}}$ on its training releases and define $u_{q,\ell}(H) = \eta(H) - \widetilde{\chi}_q(H)^\top \vartheta_{q,\ell}^{\mathrm{rep}}$. Let the reserved source-only specification feature be $\psi_{q,h}^{\mathrm{spec}}(H) \in \mathbb{R}^{s_{q,h}^{\mathrm{spec}}}$; its order h may exceed q and ℓ because it is evaluated only inside source sites. On a declared residual class assume the quantitative completeness inequality

$$\|u_{q,\ell}\|_{L_2(P_e)} \leq c_{e,q,\ell}^{-1} \|\mathbb{E}_e[\psi_{q,h}^{\mathrm{spec}} u_{q,\ell}]\|_2, \qquad c_{e,q,\ell} > 0. \tag{176}$$

This is the required separation condition: without a certified $c_{e,q,\ell}$, a moment residual can miss an orthogonal high-order direction. The independent source-validation aggregate

$$\widehat{r}_{e,q,\ell}^{\mathrm{spec}} = \widehat{\mathbb{E}}_e[\psi_{q,h}^{\mathrm{spec}} \{Y - \widetilde{\chi}_q^\top \vartheta_{q,\ell}^{\mathrm{rep}}\}]$$

and a simultaneous radius $\epsilon_{e,q,\ell}^{\mathrm{spec}}$ give the computable source upper certificate. A provider need release only this pair-specific scalar certificate; reusable specification moments are unnecessary:

$$\overline{\mathcal{E}}_{q,\ell}^{\mathrm{src}} := \max_e c_{e,q,\ell}^{-2} \{\|\widehat{r}_{e,q,\ell}^{\mathrm{spec}}\|_2 + \epsilon_{e,q,\ell}^{\mathrm{spec}}\}^2 \geq \max_e \mathbb{E}_e[u_{q,\ell}(H)^2]. \tag{177}$$

An alternative proper-score estimate is allowed only if it directly supplies the same one-sided upper confidence guarantee. A p-value or non-rejection alone provides no such guarantee.

To transport this source certificate, require the operational coverage inequality

$$\mathbb{E}_\star[u_{q,\ell}(H)^2] \leq \kappa_q^{\mathrm{rep}} \max_e \mathbb{E}_e[u_{q,\ell}(H)^2] + \tau_q^{\mathrm{rep}}, \tag{178}$$

where κ_q^{rep} and τ_q^{rep} come from a known design, a certified density-ratio bound, or the declared target uncertainty set. Here $\kappa_q^{\mathrm{rep}} \geq 0$ multiplies the largest source residual risk and $\tau_q^{\mathrm{rep}} \geq 0$ is an additive target allowance. Define the target representation-error upper bound and admissibility indicator:

$$\begin{aligned} \mathrm{repUCB}_{q,\ell} &:= \kappa_q^{\mathrm{rep}} \overline{\mathcal{E}}_{q,\ell}^{\mathrm{src}} + \tau_q^{\mathrm{rep}}, \\ \mathsf{G}_{q,\ell} &= \mathbf{1}\{\text{structural exactness}\} \\ &\quad \vee \mathbf{1}\{\text{approximate setting enabled}, \ \mathrm{repUCB}_{q,\ell} \leq \delta_{q,\ell}^{\mathrm{rep}}\}. \end{aligned} \tag{179}$$

The tolerance $\delta_{q,\ell}^{\mathrm{rep}} > 0$ is fixed before seeing validation releases. With finite samples, a positive uncertainty radius cannot certify exact equality to zero. Therefore $\overline{V}_{q,\ell} = 0$ is used only when factorization is proved structurally; an empirically screened direct setting is explicitly approximate and carries $\mathrm{repUCB}_{q,\ell}$ into its risk bound. Its reference $\vartheta_{q,\ell}^{\mathrm{rep}}$ must also lie in the approximate confidence equivalence set used to form the deployment interval. If (176) or (178) is unavailable, $\mathsf{G}_{q,\ell}$ is diagnostic and cannot admit the pair. Higher-order specification features serve only the source-side representation-error check. The deployed feature map ϕ_q and its target conditional feature mean remain order-q quantities. For T_q, this check assesses the full-state feature specification, while the risk certificate separately retains the projection variance $V_{\star,q}$.

A common moment equation is checked before rank. For environment e, let $\widehat{\mathfrak{E}}_{q,\ell}^{(-e)}$ be the confidence equivalence set formed without its label-moment release. An independent aggregate validation block gives

$$R_{e,q,\ell}^{\mathrm{stab}} = \inf_{\vartheta \in \widehat{\mathfrak{E}}_{q,\ell}^{(-e)}} \left\| \widehat{\Sigma}_{e,q,\ell}^{-1/2} \{ \widehat{b}_{e,q,\ell}^{\mathrm{val}} - \widehat{A}_{e,q,\ell}^{\mathrm{val}} \vartheta \} \right\|_2 . \tag{180}$$

The independent source-validation moment check $\mathsf{M}_{q,\ell}^{\mathrm{val}}$ passes only when every equivalence set is nonempty and the maximum residual is below its simultaneous validation radius plus a predeclared tolerance. It is used only for falsification. Full rank under a false moment equation identifies only a pseudo-parameter. Temporal blocks use strictly earlier events and dependence-valid clusters; future-positive filtering is forbidden. Failure of this check may reject a pair, but passing is not positive evidence that a zero restriction is exact. Admissibility still needs either a structurally specified exact setting or the one-sided representation upper bound stated here and below; $\mathsf{M}_{q,\ell}^{\mathrm{val}} = 1$ alone can never select a lower order.

G.5 Simultaneous uncertainty and strict identification

All strict statements are conditional on the simultaneous event

$$\begin{aligned} \|\widehat{A}_{q,\ell} - A_{q,\ell}\|_{\mathrm{op}} \le \epsilon_{A,q,\ell}, \qquad \|\widehat{b}_{q,\ell} - b_{q,\ell}\|_2 \le \epsilon_{b,q,\ell}, \\ \|\widetilde{\theta}_q\|_2 \le B \qquad \text{under an exact setting,} \\ \|\vartheta_{q,\ell}^{\mathrm{rep}}\|_2 \le B \qquad \text{under approximate direct,} \\ \sup_{z \in \mathcal{Z}_q^{\mathrm{cov}}} \|\widehat{a}_{\star,q}(z) - a_{\star,q}(z)\|_2 \le \epsilon_{a,q} \qquad \text{under } \mathsf{T}_q. \end{aligned} \tag{181}$$

For independent-cohort releases, the A and b components are resampled from their own cohorts. If a provider forms both on one local cohort, they are jointly resampled and the resulting dependence-valid joint radius is supplied. A multiplier bootstrap uses the independent unit: an observation for independent samples, a whole overlapping gadget, or a node/time cluster. If only aggregates are shared, the provider must supply a valid covariance or radius; central access to raw records is neither assumed nor reconstructed. Simultaneity over all tried (q, ℓ) and settings is obtained by a maximum statistic or a declared allocation of δ.

For strict certification, let $\epsilon_{A,q,\ell}^{\mathrm{an}}$ and $\epsilon_{b,q,\ell}^{\mathrm{an}}$ be dependence-valid analytic radii, let $Q_{A,q,\ell}^{\mathrm{bs}}$ and $Q_{b,q,\ell}^{\mathrm{bs}}$ be simultaneous multiplier-bootstrap quantiles, fix an inflation $\kappa_{\mathrm{bs}} \ge 1$, and let $\epsilon_{A,q,\ell}^{\mathrm{nuis}}$ and $\epsilon_{b,q,\ell}^{\mathrm{nuis}}$ cover nuisance-estimation error. We use

$$\begin{aligned} \epsilon_{A,q,\ell} &= \max\{\epsilon_{A,q,\ell}^{\mathrm{an}}, \kappa_{\mathrm{bs}} Q_{A,q,\ell}^{\mathrm{bs}}\} + \epsilon_{A,q,\ell}^{\mathrm{nuis}}, \\ \epsilon_{b,q,\ell} &= \max\{\epsilon_{b,q,\ell}^{\mathrm{an}}, \kappa_{\mathrm{bs}} Q_{b,q,\ell}^{\mathrm{bs}}\} + \epsilon_{b,q,\ell}^{\mathrm{nuis}}, \qquad \eta_{q,\ell} := \epsilon_{b,q,\ell} + B\epsilon_{A,q,\ell}. \end{aligned} \tag{182}$$

For an exact setting set $\eta_{q,\ell}^{\mathrm{ctr}} = \eta_{q,\ell}$. An approximate setting must also supply a simultaneous moment-defect upper bound $\beta_{q,\ell}^{\mathrm{mom}} \ge \|b_{q,\ell} - A_{q,\ell}\vartheta_{q,\ell}^{\mathrm{rep}}\|_2$ and uses

$$\eta_{q,\ell}^{\mathrm{ctr}} := \eta_{q,\ell} + \beta_{q,\ell}^{\mathrm{mom}}. \tag{183}$$

One computable choice is the independent validation residual at $\vartheta_{q,\ell}^{\mathrm{rep}}$ plus $\epsilon_{b,q,\ell}^{\mathrm{val}} + B\epsilon_{A,q,\ell}^{\mathrm{val}}$. Without this one-sided defect bound, the reference parameter is not certified to lie in the equivalence set and the approximate direct setting abstains. All these inputs are fixed or calibrated without the sealed outcomes. A bootstrap estimate cannot shrink a valid analytic bound. Without an analytic or justified dependence-aware radius, empirical spectra are reported only as diagnostics. Writing $m_{q,\ell}$ for the number of rows in the stacked moment operator, under bounded moments, $\eta_{q,\ell} = O\{\sqrt{m_{q,\ell}/\underline{n}_{q,\ell}^{\mathrm{src}}}\}$ up to scale and logarithmic factors. This is source release uncertainty, distinct from $\underline{n}_q^{\mathrm{tar}}$ and Δ_q^{sup}.

Recovering all coefficients under an exact data-access setting. Let $\widehat{s}_{q,\ell} = \sigma_{p_q}(\widehat{A}_{q,\ell})$ under D_q and $\widehat{s}_{q,\ell} = \sigma_p(\widehat{A}_{q,\ell})$ under T_q, set to zero when the stacked release has fewer rows than the corresponding column dimension. Weyl's inequality yields

$$\widehat{s}_{q,\ell} > \epsilon_{A,q,\ell} \Longrightarrow \sigma_{d_q}(A_{q,\ell}) \ge \widehat{s}_{q,\ell} - \epsilon_{A,q,\ell} > 0. \tag{184}$$

For exact D_q or T_q, on (181), the full-rank estimator $\widehat{\theta}_{q,\ell} = \widehat{A}^{\dagger}_{q,\ell}\widehat{b}_{q,\ell}$ satisfies

$$\|\widehat{\theta}_{q,\ell} - \widetilde{\theta}_q\|_2 \le C_{q,\ell} := \eta_{q,\ell}/\widehat{s}_{q,\ell}. \tag{185}$$

Full-parameter recovery additionally requires a predeclared upper bound on $C_{q,\ell}$. The approximate direct setting may use the same strict singular-value check to control equation geometry, but it makes no claim of $\|\widehat{\theta} - \widetilde{\theta}_q\|$ for a nonexistent exact parameter; it instead requires reference feasibility under $\eta^{\mathrm{ctr}}_{q,\ell}$ and charges the representation-error upper confidence bound.

Target outcome probability. Full recovery is unnecessary when the relevant row is known or certified. Define

$$\widehat{\mathfrak{E}}_{q,\ell} = \left\{\vartheta : \|\vartheta\|_2 \le B,\ \|\widehat{A}_{q,\ell}\vartheta - \widehat{b}_{q,\ell}\|_2 \le \eta^{\mathrm{ctr}}_{q,\ell}\right\}. \tag{186}$$

For each covered target value z, two second-order-cone programs give

$$\begin{aligned}\underline{T}_{q,\ell}(z) &= \inf_{\vartheta \in \widehat{\mathfrak{E}}_{q,\ell}} \widehat{a}_{\star,q}(z)^{\top}\vartheta - B\epsilon_{a,q},\\ \overline{T}_{q,\ell}(z) &= \sup_{\vartheta \in \widehat{\mathfrak{E}}_{q,\ell}} \widehat{a}_{\star,q}(z)^{\top}\vartheta + B\epsilon_{a,q}.\end{aligned} \tag{187}$$

For D_q, use $\widehat{a}_{\star,q}(z) = \phi_q(z)$ and $\epsilon_{a,q} = 0$. For T_q, the radius is precisely the one certified in (170). Under an exact direct or exact target-adaptive setting, the true projected Bayes score lies in the interval on the simultaneous event. Under the approximate direct setting the interval covers the modeled reference score and the representation-error upper bound pays for its distance to the truth. An approximate target-adaptive setting is outside the declared setting. An empty equivalence set always causes rejection. If the deployment claim allows the bounded mechanism drift δ_{mech} in Appendix B.3, replace these endpoints by $\max\{0, \underline{T}_{q,\ell}(z) - \delta_{\mathrm{mech}}\}$ and $\min\{1, \overline{T}_{q,\ell}(z) + \delta_{\mathrm{mech}}\}$. This increases the score radius by at most δ_{mech}; a sharper interval may use the exact clipped-state program in (32).

For cheap screening, truncate the singular value decomposition (SVD) at $\tau_A \ge \epsilon_{A,q,\ell}$, let $\widehat{P} = \widehat{A}^{\dagger}_{\tau_A}\widehat{A}_{\tau_A}$, and compute

$$\widehat{d}_{q,\ell}(z) = \|(I - \widehat{P})\widehat{a}_{\star,q}(z)\|_2, \qquad \widehat{L}_{q,\ell}(z) = \|(\widehat{A}^{\dagger}_{\tau_A})^{\top}\widehat{P}\widehat{a}_{\star,q}(z)\|_2. \tag{188}$$

Writing $D_{q,\ell}(z)$ for half the exact interval width,

$$D_{q,\ell}(z) \le B\widehat{d}_{q,\ell}(z) + \widehat{L}_{q,\ell}(z)\eta^{\mathrm{ctr}}_{q,\ell} + B\epsilon_{a,q}. \tag{189}$$

The three terms are distance from the identified row space, inverse-information amplification times source-release uncertainty, and offline conditional-feature-mean uncertainty. Target-adaptive certification requires the exact width and leverage to be below their declared thresholds over $\mathcal{Z}^{\mathrm{cov}}_q$. Approximate row-space membership is never relabelled full-parameter recovery.

G.6 Projection variance and the complete risk certificate

The identified-score interval controls error around the projected Bayes mean. It does not control the information lost by projection. Under D_q, the certified factorization gives $\overline{V}_{q,\ell} = 0$. Under T_q, unlabelled target contexts do not by themselves reveal $V_{\star,q}$ because that quantity depends on the unknown θ. They instead identify the positive-semidefinite conditional-covariance operator

$$G_{\star,q} := \mathbb{E}_{\star}[\mathrm{Cov}_{\star}\{\chi(H) \mid Z_q\}], \qquad V_{\star,q}(\theta) = \theta^{\top}G_{\star,q}\theta. \tag{190}$$

On a reserved target-context split, a cross-fitted conditional-feature-mean map gives the empirical operator

$$\widehat{G}_q = \frac{1}{n_{\star,V}}\sum_{i=1}^{n_{\star,V}}[\chi(H_i) - \widehat{a}_{\star,q}(Z_{q,i})][\chi(H_i) - \widehat{a}_{\star,q}(Z_{q,i})]^{\top}. \tag{191}$$

Conditional on a conditional-feature-mean map fitted on another fold,

$$\begin{aligned}&\mathbb{E}_{\star}[\{\chi - \widehat{a}_{\star,q}(Z_q)\}\{\chi - \widehat{a}_{\star,q}(Z_q)\}^{\top}]\\ &\quad = G_{\star,q} + \mathbb{E}_{\star}[\{a_{\star,q}(Z_q) - \widehat{a}_{\star,q}(Z_q)\}\{a_{\star,q}(Z_q) - \widehat{a}_{\star,q}(Z_q)\}^{\top}] \succeq G_{\star,q}.\end{aligned} \tag{192}$$

Hence a cluster- or mixing-valid operator radius $\epsilon_{G,q}$ gives the simultaneous positive-semidefinite (PSD) upper certificate $G_{\star,q} \preceq \overline{G}_q := \widehat{G}_q + \epsilon_{G,q} I$. A tractable mechanism-uniform certificate is

$$\overline{V}_q^{\mathrm{ball}} = \min\left\{\frac{1}{4},\, B^2 \lambda_{\max}(\overline{G}_q)\right\}, \qquad \overline{V}_{q,\ell} := \overline{V}_q^{\mathrm{ball}}. \tag{193}$$

This ball bound is the default final certificate and can screen q before ℓ is chosen. The exact quadratic supremum over $\widehat{\mathfrak{E}}_{q,\ell}$ is generally a nonconvex maximization and is not treated as a second-order cone program (SOCP). A pair may replace the default by $\min\{\overline{V}_q^{\mathrm{ball}}, U_{q,\ell}^{\mathrm{quad}}\}$ only when a global method, such as a valid semidefinite program (SDP) relaxation, certifies $U_{q,\ell}^{\mathrm{quad}} \geq \sup_{\vartheta \in \widehat{\mathfrak{E}}_{q,\ell}} \vartheta^\top \overline{G}_q \vartheta$; otherwise it falls back to the ball bound. These quantities dominate the true projection variance on the simultaneous event because the true Bernoulli mechanism takes values in $[0,1]$. Cross-fitting the conditional-feature-mean map or using a separate variance split is required. If the target contexts or a valid PSD radius are absent, an excess-risk guarantee relative to the Bayes rule given H is unavailable; CHORD does not replace the variance by zero.

Under either exact data-access setting, $\eta(H) = \widetilde{\chi}_q(H)^\top \widetilde{\theta}_q$ pointwise, so

$$\Gamma_{q,\ell}^{\mathrm{stab}}(z) := 0. \tag{194}$$

The source-validation moment residual remains a falsification check; it is not charged again as an unidentified target bias. The only approximate setting offered here is direct factorization: its pair-specific target representation-error bound already bounds the entire discrepancy between the frozen reference scorer and η, so it also sets $\Gamma_{q,\ell}^{\mathrm{stab}} = 0$ and charges that upper confidence bound below. We do not offer an approximate target-adaptive setting; without an exact full-state feature map its target residual would require an additional, separately assumed source–target risk bound.

For the interval-midpoint predictor define the squared score-error bound over supported target values

$$\overline{\mathcal{T}}_{q,\ell} = \sup_{z \in \mathcal{Z}_q^{\mathrm{cov}}} \{D_{q,\ell}(z) + \Gamma_{q,\ell}^{\mathrm{stab}}(z)\}^2. \tag{195}$$

An allowed target-distribution upper confidence average may replace the supremum under T_q; setting D_q uses the predeclared target uncertainty set and remains target-agnostic.

For the approximate direct setting, $\overline{\mathcal{T}}_{q,\ell}$ is formed relative to the certified reference $\vartheta_{q,\ell}^{\mathrm{rep}}$ contained in $\widehat{\mathfrak{E}}_{q,\ell}$, with $\Gamma^{\mathrm{stab}} = 0$ because $\mathrm{repUCB}_{q,\ell}$ already bounds the entire representation residual. The general interval-midpoint scorer uses the transport term below. The frozen-reference case instead deploys $\widehat{g}_{q,\ell}(z) = \mathrm{clip}\{\phi_q(z)^\top \vartheta_{q,\ell}^{\mathrm{rep}}, 0, 1\}$, for which the risk term is directly $\mathrm{repUCB}_{q,\ell}$: projection onto $[0,1]$ cannot increase distance to $\eta(H) \in [0,1]$. We encode this case by $\overline{\mathcal{T}}_{q,\ell} = 0$; Minkowski's inequality is needed only for the general midpoint case. Define the setting-specific risk bound on supported target values

$$\mathcal{B}_{q,\ell} := \begin{cases} \overline{V}_{q,\ell} + \overline{\mathcal{T}}_{q,\ell}, & \text{exact direct or exact target-adaptive,} \\ \{\sqrt{\mathrm{repUCB}_{q,\ell}} + \sqrt{\overline{\mathcal{T}}_{q,\ell}}\}^2, & \delta_{q,\ell}^{\mathrm{rep}}\text{-approximate direct.} \end{cases} \tag{196}$$

Combining the Brier decomposition with the simultaneous certificates gives

$$\boxed{\widehat{U}_{q,\ell} = \mathcal{B}_{q,\ell} + \Delta_q^{\mathrm{sup}}.} \tag{197}$$

This bounds excess risk relative to the Bayes rule given H. In the exact target-adaptive setting, $\mathcal{B}$ separates irreducible information loss from projection from the squared transport radius. In the approximate direct setting, $\mathrm{repUCB}_{q,\ell}$ enters the final risk together with its cross term with estimation error. The support term pays explicitly for target holes and rare cells. Source release support $\underline{n}_{q,\ell}^{\mathrm{src}}$ also enters $\overline{\mathcal{T}}_{q,\ell}$ through $\eta_{q,\ell}$; it is not counted again in Δ_q^{sup}, which covers target-cell mass. Calibration is not part of this pre-selection bound; an optional selected-pair allowance is handled after selection below.

Proposition G.1 (Simultaneous certificate and cost-aware bound selection)**.** *Suppose the release, representation/coverage, conditional-feature-mean, covariance, support, source-validation, and target-bias bounds used above hold jointly over every attempted (q,ℓ) and setting with probability at*

least $1-\delta$. *Let* $\widehat{g}_{q,\ell}$ *be the setting-specified clipped predictor: the interval midpoint in general, or the frozen reference in the stated case above. Then, simultaneously for every strictly certified pair,*

$$\mathcal{R}_\star(\widehat{g}_{q,\ell}) - \mathcal{R}^*_{\star,H} \le \widehat{U}_{q,\ell}. \tag{198}$$

If the admissible set $\mathcal{C}$ *in* (200) *is nonempty, the two-stage rule below satisfies, on the same event,*

$$\begin{aligned}\mathcal{R}_\star(\widehat{g}_{\widehat{q},\widehat{\ell}}) - \mathcal{R}^*_{\star,H} + \lambda_\ell \operatorname{cost}_{\mathrm{src}}(\widehat{\ell}) + \lambda_q \operatorname{cost}_{\mathrm{deploy}}(\widehat{q}) \\ \le \min_{(q,\ell)\in\mathcal{C}} \left\{ \widehat{U}_{q,\ell} + \lambda_\ell \operatorname{cost}_{\mathrm{src}}(\ell) + \lambda_q \operatorname{cost}_{\mathrm{deploy}}(q) \right\}.\end{aligned} \tag{199}$$

where $\widehat{\ell} = \widehat{\ell}(\widehat{q})$. *If* $\mathcal{C} = \varnothing$, *the prescribed output is abstention.*

Indeed, the Brier orthogonal decomposition contributes $V_{\star,q}$; the score interval bounds squared error on supported target values; and the support event bounds the remaining target-mass term. For the general approximate-direct midpoint, apply Minkowski before clipping; projection onto $[0,1]$ can only reduce its distance to η, giving the middle line of (196). For the frozen-reference case, the same projection contraction directly gives $\|\widehat{g}-\eta\|^2_{L_2(P_\star)} \le \mathrm{repUCB}_{q,\ell}$, so no separate transport term is needed. Equation (197) therefore holds for every pair on one simultaneous event. Minimizing its certified upper bound plus the two declared costs proves (199); no post-selection union bound is needed.

G.7 Selecting projection and source-moment orders

Fix nonnegative cost weights λ_ℓ and λ_q, nonnegative functions $\operatorname{cost}_{\mathrm{src}}(\ell)$ and $\operatorname{cost}_{\mathrm{deploy}}(q)$, and an admissible excess-risk radius $\eta_{\mathrm{risk}} \ge 0$, all before opening certification releases. Let C_q denote validity of the chosen deployment setting, $\mathsf{S}^{\mathrm{tar}}_q$ target deployment support, $\mathsf{S}^{\mathrm{src}}_{q,\ell}$ source release support, $\mathsf{G}_{q,\ell}$ source-only representation specification, $\mathsf{M}^{\mathrm{val}}_{q,\ell}$ independent source-validation moment check used only for falsification, and $\mathsf{I}_{q,\ell}$ the requested rank or target-interval certificate. The admissible set is

$$\mathcal{C} = \left\{ (q,\ell) : \mathsf{C}_q \mathsf{S}^{\mathrm{tar}}_q \mathsf{S}^{\mathrm{src}}_{q,\ell} \mathsf{G}_{q,\ell} \mathsf{M}^{\mathrm{val}}_{q,\ell} \mathsf{I}_{q,\ell} = 1,\ \widehat{U}_{q,\ell} \le \eta^2_{\mathrm{risk}} \right\}. \tag{200}$$

At fixed q, choose

$$\widehat{\ell}(q) \in \arg\min_{\ell:(q,\ell)\in\mathcal{C}} \{\widehat{U}_{q,\ell} + \lambda_\ell \operatorname{cost}_{\mathrm{src}}(\ell)\}, \tag{201}$$

then choose

$$\widehat{q} \in \arg\min_{q:(q,\widehat{\ell}(q))\in\mathcal{C}} \{\widehat{U}_{q,\widehat{\ell}(q)} + \lambda_\ell \operatorname{cost}_{\mathrm{src}}\{\widehat{\ell}(q)\} + \lambda_q \operatorname{cost}_{\mathrm{deploy}}(q)\}. \tag{202}$$

Break exact ties by the least order; source moment order and deployed order remain separate choices.

G.8 Paired Brier-risk comparison with a fixed baseline under an assumed risk bound

The preceding certificate asks whether an order has the declared mechanism-factorization or target-adaptive interpretation. The paired comparison asks whether replacing the same frozen highest-order predictor improves Brier risk over the declared shift class. For arity K, let Q_K denote that baseline order and let $\mathcal{Q}^{\mathrm{check}}_K$ contain only $q < Q_K$ that pass the predeclared support, specification, required rank or row-containment, and data-access checks. These checks determine which structural interpretation may be reported; they do not by themselves establish a risk improvement.

All selection equations above are applied within an observable arity stratum when the wrapper is used. For each surviving q, choose $\widehat{\ell}(q,K)$ without opening the paired-loss release and freeze

$$g_{q,K} := g_{q,\widehat{\ell}(q,K),K}. \tag{203}$$

An order with no surviving source moment order is omitted from $\mathcal{Q}^{\mathrm{check}}_K$. The paired comparison therefore selects the deployment order q around a candidate whose source moment order was fixed upstream; it does not discard the joint (q,ℓ) selection.

On an independent certification sample unused for fitting any attempted predictor, each source environment provides the paired loss difference

$$D_{i,e,qK} = \{g_{q,K}(Z_{i,q}) - Y_i\}^2 - \{g_{Q_K,K}(Z_{i,Q_K}) - Y_i\}^2, \qquad \widehat{\Delta}_{e,qK} = \frac{1}{n_{eK}} \sum_{i:K_i=K} D_{i,e,qK}, \tag{204}$$

together with a one-sided, simultaneous, dependence-valid radius $r_{e,qK}$, where n_{eK} is the number of certification examples of arity K in environment e. The two losses use the same labelled example, so the comparison is paired; only the aggregate difference and its radius need leave a source. The set $\mathcal{W}_K$ and allowance $\Gamma_{qK}^{\mathrm{shift}}$ in (107) must be fixed from a known environment design or from a separately allowed unlabelled target release. They may not use target labels or the paired certification outcomes. If that assumed risk bound or a valid radius is unavailable, set $\widehat{\Delta}_{qK}^{+} = +\infty$. Corollary E.6 supplies two direct choices: an exact source mixture uses $\Gamma_{qK}^{\mathrm{shift}} = 0$, while a target within half-L_1 total-variation radius ρ_K of such a mixture uses $\Gamma_{qK}^{\mathrm{shift}} = 2\rho_K$. Under the separately declared uniform mechanism-drift class with radius δ_{mech}, add $2\delta_{\mathrm{mech}}$ to either allowance, as proved in Appendix B.3. No such addition certifies an unrestricted mechanism change.

Calibrating overlap and shift allowances. The following conservative recipe uses no target labels. Suppose full contexts take finitely many values and an independent context release supplies empirical histograms $\widehat{p}_\star, \widehat{p}_1, \ldots, \widehat{p}_E$ together with simultaneous half-L_1 radii $r_\star, r_1, \ldots, r_E$. For a weight set $\mathcal{W}_K$ fixed before opening the release, define

$$\rho_K^U := \inf_{w \in \mathcal{W}_K} \left\{ \mathrm{TV}\left(\widehat{p}_\star, \sum_e w_e \widehat{p}_e\right) + r_\star + \sum_e w_e r_e \right\}. \tag{205}$$

The triangle inequality makes $2\rho_K^U$ a valid choice of $\Gamma_{qK}^{\mathrm{shift}}$ on the simultaneous event. The same construction within a projection cell upper-bounds the fiber total-variation radius. A likelihood-ratio upper bound is obtained by maximizing an upper target cell probability over a positive lower reference probability; if any required lower bound is zero, set $\kappa = \infty$ and use the total-variation-only analysis or abstain. In continuous or high-dimensional contexts, a density-ratio model and its coverage guarantee must be declared assumptions; the data alone provide no distribution-free guarantee. If only Z_q is observed, hidden within-fiber overlap is not identified; $(\rho_{\mathrm{fib}}, \kappa)$ must be shown as a sensitivity curve. If neither a known design nor a valid context bound is available, the paired upper bound is $+\infty$ and no certified switch occurs. When a provider can also release simultaneous full-state outcome intervals $[L_h, U_h]$, a directly auditable approximate-factorization bound is

$$\widehat{\mathrm{osc}}_z^U := \max_{h \in \mathcal{F}_z} U_h - \min_{h \in \mathcal{F}_z} L_h, \qquad \widehat{D}_z^U := 2\rho_{\mathrm{fib}}^U \widehat{\mathrm{osc}}_z^U. \tag{206}$$

It follows from (11); without these intervals the corresponding plot is a sensitivity analysis alone.

The target conditional feature mean is handled separately. A provider may cross-fit $z \mapsto \widehat{a}_{\star,q}(z)$ on unlabelled full contexts and release a simultaneous radius $\epsilon_{a,q}$; Equations (187)–(189) add $B\epsilon_{a,q}$ before squaring. A missing or misspecified map therefore widens the interval or triggers abstention. Support thresholds, window boundaries, and every radius are fixed before the sealed labels are opened. For clusters spanning several environment blocks, the independent resampling unit is the complete vector of their contributions to all (e, q, ℓ) cells. The release must preserve this joint dependence.

For a predeclared margin $\gamma_K \geq 0$ and nonempty $\mathcal{Q}_K^{\mathrm{check}}$, let

$$\widetilde{q}_K \in \arg\min_{q \in \mathcal{Q}_K^{\mathrm{check}}} \widehat{\Delta}_{qK}^{+}, \qquad \widehat{q}_K = \begin{cases} \widetilde{q}_K, & \widehat{\Delta}_{qK}^{+}\Big|_{q=\widetilde{q}_K} < -\gamma_K, \\ Q_K, & \text{otherwise}, \end{cases} \tag{207}$$

When $\mathcal{Q}_K^{\mathrm{check}}$ is empty, set $\widehat{q}_K = Q_K$ directly. The bound is the one in (108). The fallback case reuses the same frozen baseline predictor, including its head and calibration. Theorem E.5 then certifies the declared Brier-risk improvement; it makes no claim for a different loss or an unspecified shift. It also does not establish causal validity of the baseline. Thus insufficient evidence returns that predictor unchanged, whereas abstention remains necessary for an excess-risk certificate relative to the Bayes rule given H.

Algorithm 2 Detailed finite-sample CHORD procedures

Require: Frozen backbone and heads; joint maps T_q; the same frozen highest-order baselines $g_{Q_K,K}$; aggregate source moments and joint radii; declared data-access settings, support information, and tolerances; for paired baseline comparison, paired Brier-loss releases and $(\mathcal{W}_K, \Gamma_{qK}^{\mathrm{shift}})$.

Predeclare arities, orders, settings, failure allocation, margins, and either excess-risk certification relative to the Bayes rule given H or paired comparison under the assumed risk bound.
Verify that each $Z_q = T_q(G_{<t}, S)$ transforms both history and candidate; reject any order using a full-incidence cache not measurable from Z_q.
Fit only on allocated training releases, then freeze all attempted representations and predictors before opening any certification release; retain each $g_{Q_K,K}$ exactly as frozen.
for each supported arity K and $q = 0, \dots, Q_K$ **do**
    Under D_q, set $(\widehat{a}_{\star,q}, \epsilon_{a,q}) = (\phi_q, 0)$ only with the declared structural factorization; otherwise require the approximate prediction setting.
    Under T_q, retrieve the frozen offline conditional-feature-mean map and certify $(\epsilon_{a,q}, \overline{G}_q)$ on its reserved release; continue if unavailable.
    Compute $\underline{n}_q^{\mathrm{tar}}, \Delta_q^{\mathrm{sup}}$; continue if a target-side strict check fails.
    for $\ell = 0, \dots, \ell_{\max}(q)$ **do**
        Compute $\underline{n}_{q,\ell}^{\mathrm{src}}$; continue if $\mathsf{S}_{q,\ell}^{\mathrm{src}} = 0$.
        Stack and jointly whiten aggregate $\widehat{A}_{q,\ell}^e \simeq \mathbb{E}_e[\psi_{q,\ell}\widetilde{X}_q^\top]$ and $\widehat{b}_{q,\ell}^e \simeq \mathbb{E}_e[\psi_{q,\ell}Y]$.
        Freeze the pair's fitted prediction rule before paired-loss certification; obtain its one-sided source specification bound and continue if $\mathsf{G}_{q,\ell} = 0$.
        Run the independent source-validation moment check only for falsification; continue if $\mathsf{M}_{q,\ell}^{\mathrm{val}} = 0$, but never treat a pass as proof.
        Compute simultaneous $(\epsilon_{A,q,\ell}, \epsilon_{b,q,\ell})$ and $\eta_{q,\ell}$; apply the strict rank test or exact target intervals and continue if $\mathsf{I}_{q,\ell} = 0$.
        Form the setting-specific excess-risk bound $\widehat{U}_{q,\ell}$ relative to the Bayes rule given H, including conditional-feature-mean, projection, representation, and support allowances.
    end for
    Select $\widehat{\ell}(q, K)$ among surviving source moment orders.
end for
if using paired comparison under the assumed risk bound **then**
    Open the paired-loss release, form every $\widehat{\Delta}_{qK}^+$, and apply (207).
    If evidence is insufficient, return $g_{Q_K,K}$ unchanged; otherwise return $g_{\widehat{q}_K,K}$ with its paired Brier-risk certificate.
else if the admissible set for that excess-risk certificate is empty **then**
    return equivalence sets, intervals, failed checks, and **abstain**.
else
    Select $(\widehat{q}, \widehat{\ell})$ by (202) and return its excess-risk certificate relative to the Bayes rule given H.
end if

G.9 VALIDATION-BASED TEMPORAL COMPARISON ON NATURAL STREAMS

When the assumed source–target risk bound is unavailable, the paired-risk theorem is not invoked. For each arity K, the validation-based procedure freezes the baseline and every order-restricted candidate passing the predeclared source rank screen before opening the selection interval. It partitions that interval into exactly four chronological validation windows. A window is scored using only history available before that window, and history is updated only after the complete window has been evaluated.

For a candidate $q < Q_K$, let each paired difference be candidate loss minus baseline loss on the same query and candidate set. The predeclared rule retains q only if every loss is finite, every window contains at least 30 evaluated groups, the pooled reciprocal-rank loss difference is negative, the largest window-average reciprocal-rank loss difference is nonpositive, and the pooled Brier-loss difference is nonpositive. Among retained orders it minimizes pooled reciprocal-rank loss, breaking ties by pooled Brier loss and then lower q; if none remains, it returns the same frozen baseline unchanged. The later test interval is opened once after this decision and is never used for selection. These validation-based comparisons carry no finite-sample OOD or causal certificate.

G.10 Detailed finite-sample procedures

If $\mathcal{C}$ is empty, CHORD never selects an arbitrary member of an unresolved equivalence class. In particular, absent both a structural equal-outcome-mean certificate and an offline conditional-feature-mean map, the procedure certifying excess risk relative to the Bayes rule given H rejects an order that merges states with different outcome means. Only when the user has explicitly requested the separate $\delta^{\mathrm{rep}}_{q,\ell}$-approximate prediction setting may the approximate-direct bound in (196) be returned instead, and it is never reported as satisfying the equal-outcome-mean condition. Favorable empirical spectra, validation loss, or apparent support without simultaneous radii remain diagnostics for the finite-sample procedures. Certification relative to the Bayes rule given H differs from paired comparison under the assumed source–target risk bound: in the latter, insufficient evidence returns the same frozen baseline unchanged and supports only the Brier-risk statement in Theorem E.5. Validation-based temporal selection is the separate, non-certified comparison in Section G.9.

G.11 No-leakage deployment and optional calibration

Source sites may use $\psi_{q,\ell}$ and higher-order $\psi^{\mathrm{spec}}_{q,h}$ to create moment and representation certificates. These source moment features influence which data-access setting is selected but are never included in its deployed scorer. Each deployed score first constructs $Z_{\widehat{q}} = T_{\widehat{q}}(G_{<t}, S)$ jointly from the history and candidate. It may not reuse a hidden state computed from full incidences unless that state has been proved measurable with respect to $\sigma(Z_{\widehat{q}})$. Under D_q, candidate-time evaluation is exactly $\mathrm{clip}\{\phi_{\widehat{q}}(Z_{\widehat{q}})^\top \widehat{\vartheta}, 0, 1\}$, using the selected exact estimate or frozen approximate reference. Under T_q, full target contexts are consumed only by the offline conditional-feature-mean and variance procedures; candidate-time evaluation is exactly the $[0,1]$-clipped value of $\widehat{a}_{\star,\widehat{q}}(Z_{\widehat{q}})^\top \widehat{\theta}$ (or of the certified interval midpoint). Thus neither finite-sample data-access setting evaluates a source moment feature, a full target state, a future event, or a target label online. In the paired and validation-based procedures, fallback loads the retained $g_{Q_K,K}$ object without changing its weights, head, threshold, or calibration.

If a separately declared projected-only calibration release is available, an untouched source fold with label Y_i, selected frozen score $\widehat{s}_i$, and observed arity K_i may fit an arity-conditioned map

$$\widehat{P}(Y_i = 1 \mid \widehat{s}_i, K_i) = \widehat{g}(\widehat{s}_i, K_i). \tag{208}$$

It is opened only after $(\widehat{q}, \widehat{\ell})$ is fixed and cannot contain paired $(\widetilde{\chi}_q, Y)$ records. If a selected-pair validation event, with its own confidence allocation and an assumed source–target risk bound certified separately for calibration, certifies an additive allowance $\mathcal{R}_\star(\widehat{g}^{\mathrm{cal}}) - \mathcal{R}_\star(\widehat{g}) \le \Delta^{\mathrm{cal}}_{\mathrm{sel}}$, then the final bound is $\widehat{U}_{\widehat{q},\widehat{\ell}} + \Delta^{\mathrm{cal}}_{\mathrm{sel}}$. This post-selection term is not used in either argmin and is not part of Proposition G.1. Without those ingredients calibration is diagnostic and the identity map is used. Conditioning on observable set size avoids the error from pooling across set sizes in Proposition F.2. For the paired Brier-risk switching guarantee, any calibration included in $g_{q,K}$ must instead be fit and frozen before the paired release is opened. A post-selection calibration is a different predictor and is covered only by its separately stated additive allowance; it is never applied when the frozen baseline is returned unchanged.

G.12 Complexity and reported implementation choices

For one pair, let n_A and n_b be the provider-side sample counts for the operator and response aggregates, let $s_{q,\ell}$ and d_q be their row- feature and coefficient dimensions, and let $m_{q,\ell}$ be the number of rows in the stacked released operator. Providers accumulate the two aggregates in

$$O(n_A s_{q,\ell} d_q + n_b s_{q,\ell}) \text{ time}, \qquad O(m_{q,\ell} d_q) \text{ released storage}. \tag{209}$$

The central dense SVD costs $O\{\min(m_{q,\ell} d_q^2, m_{q,\ell}^2 d_q)\}$ time and $O(m_{q,\ell} d_q)$ memory. Nested source moment features append rows; nested direct feature maps append columns, permitting QR updates with orthogonal Q and upper-triangular R, alongside SVD updates. With $B_{\mathrm{bs}} \ge 1$ multiplier-bootstrap draws, provider-side resampling is linear in B_{bs} and embarrassingly parallel. Each exact target interval uses two second-order-cone solves; factored screening needs only matrix solves. The default projection-variance certificate requires one leading-eigenvalue computation. If the optional pair-specific SDP relaxation is used, its solver and tolerance must be

declared and its cost reported; it is not included in the SOCP count. The representation-error upper confidence bound adds $O(n_{\mathrm{spec}} s_{q,h}^{\mathrm{spec}} d_q)$ provider-side work for each reserved specification family; this cost is source-only and does not affect candidate latency, where n_{spec} is the reserved specification-sample count and $s_{q,h}^{\mathrm{spec}}$ is the specification-feature dimension.

Under T_q, fitting a dense offline conditional-feature-mean map on $n_{\star,a}$ unlabeled target contexts costs $O(n_{\star,a} p)$ feature accumulation plus the chosen conditional-mean regressor; this is a one-time cost. Writing nnz for the number of nonzero entries, candidate-time cost is $O(\mathrm{nnz}\, \phi_{\widehat{q}})$ for D_q and the cost of evaluating the stored $Z_{\widehat{q}} \mapsto \widehat{a}_{\star,\widehat{q}}$ map for T_q. It is independent of $\widehat{\ell}$ and of the order of every source-only specification feature.

Large feature maps may replace full rank by a predeclared restricted singular value over sparse differences and an ℓ_1-regularized moment fit. Random sparse probes provide diagnostics only. Signed support is reported only under a beta-min condition exceeding twice the certified coordinate error; otherwise CHORD returns coefficient intervals.

The reproducibility record contains only consequential choices: the candidate data-access settings; $q_{\max}$ and $\ell_{\max}(q)$; full-state feature maps, conditional-feature-mean model class, and which target contexts are allowed; environment construction and independent release roles; specification completeness constants, $(\kappa_q^{\mathrm{rep}}, \tau_q^{\mathrm{rep}}, \delta_{q,\ell}^{\mathrm{rep}})$, the resulting $\mathrm{repUCB}_{q,\ell}$; support and projection-variance bounds; B, δ; dependence unit, bootstrap draws, and inflation; specification, moment, interval, leverage, and risk tolerances; deployment/source costs; calibration rule; and random seeds. For paired baseline comparison it additionally records Q_K, the frozen baseline identifier, γ_K, $\mathcal{W}_K$, $\Gamma_{qK}^{\mathrm{shift}}$, paired-loss radii, and whether a switch occurred. For validation-based temporal comparison it records the four window boundaries, group counts, all paired validation losses, tie-breaking rule, and the sealed test boundary. Finite-sample results report $(\widehat{q}, \widehat{\ell})$, the chosen setting, all check outcomes, $\overline{V}_{q,\ell}$, the squared transport radius, the support term, and the applicable certificate. Empirical results report the selected order, every validation criterion, equality to the saved baseline predictions, and sealed-test metrics, and label the comparison as empirical.

H EXPERIMENTAL PROTOCOLS AND IMPLEMENTATION DETAILS

Each controlled protocol isolates a theorem quantity with an exactly computable population target. Its grid was fixed before aggregation and target labels are evaluation-only. The public temporal benchmarks perform validation-based order selection on an earlier labelled period and reserve the final period for one held-out evaluation. Means and standard deviations use the stated independent seeds or trials. We report the Brier score for probabilistic predictions (Brier, 1950), balanced test error, and the area under the receiver operating characteristic curve (AUROC) for sparse labels (Hanley & McNeil, 1982). Observations are bounded Bernoulli or Rademacher variables, except in the explicitly Gaussian experiment for Theorem E.9.

Outcome and score scope. Let H be the complete candidate context and $g(H) \in [0,1]$ its predicted outcome probability. The reported link-prediction tasks use a binary outcome $Y \in \{0,1\}$ and Brier loss $(g(H) - Y)^2$. The transport and rank arguments depend on the conditional mean, so the same scalar formulas apply to a bounded continuous outcome $Y \in [0,1]$ under squared loss. For C classes, encode the outcome as a one-hot vector $Y \in \{e_1, \ldots, e_C\}$, define $\eta(H) := \mathbb{E}[Y \mid H]$, and use multiclass Brier loss $\|g(H) - Y\|_2^2$, where e_c is the standard basis vector for class c. Constancy inside a projection cell then means equality of the entire vector $\eta(H)$, and water filling gives a sharp interval for every declared linear contrast $u^\top \eta(H)$ for a fixed contrast vector u. The scalar interval midpoint remains minimax for a binary decision or one fixed contrast; a joint vector predictor uses the Euclidean minimum-enclosing-ball center of the attainable probability set.

Certification and real-stream evaluation. The controlled joint order-selection experiment instantiates finite-sample certification of excess Brier risk relative to the full-state Bayes rule under predeclared mechanism and target-support requirements; its dependence-valid radii and target-risk bounds are known by construction. The real temporal study instead freezes each checkpoint before the selection release, uses 40–60% for action calibration and rank screening, 60–80% for arity-specific order selection, and opens 80–100% once for later-window evaluation. We reserve certificate claims for that controlled order-selection design. The real-data tables compare a frozen highest-order baseline predictor with the same checkpoint plus CHORD as validation-based temporal selection; when no alternative is selected, the same saved baseline prediction is returned unchanged.

What must be observable before deployment. The certified route separates two data-access settings. Here H is the complete context, $\chi(H) \in \mathbb{R}^d$ is its mechanism feature, and $\theta \in \mathbb{R}^d$ is the coefficient in the

conditional-mean model. In the direct setting, $\chi(H)$ factors through the order-q projected state $Z_q = z$, so $\chi(H) = \phi_q(z)$ and its target conditional mean is known exactly. In the target-adaptive setting, an independent unlabelled sample of full contexts estimates $a_{\star,q}(z) := \mathbb{E}_\star[\chi(H) \mid Z_q = z]$ by cross-fitting and supplies a simultaneous error radius $\epsilon_{a,q}$. If $\|\widehat{a}_{\star,q}(z) - a_{\star,q}(z)\|_2 \le \epsilon_{a,q}$ and the mechanism coefficient satisfies $\|\theta\|_2 \le B$, the target-score radius increases by the declared term $B\epsilon_{a,q}$ before it is squared in the Brier-risk bound.

An independent context-only target sample also estimates the projected-cell mass $p_{\star,q}(z) := P_\star(Z_q = z)$ and, when full contexts are available, the conditional distribution inside that projection cell. Simultaneous multinomial intervals give an upper bound ρ^U_{fib} on the within-cell total-variation change and an upper likelihood-ratio cap κ^U; a zero lower reference probability sets $\kappa^U = \infty$. The water-filling calculation uses these upper bounds, while the confidence interval for $p_{\star,q}(z)$ contributes to the target-support term. With only Z_q observed, hidden within-cell change remains unidentified, so the paper reports $(\rho_{\mathrm{fib}}, \kappa)$ as a sensitivity grid and the certified route returns no switch without a separately justified bound. The temporal benchmark route requires no hidden-state estimate and makes the distinct, validation-based claim stated above.

Public temporal hypergraphs and frozen scorers. We use the official Cornell triplets for AskUbuntu and Math StackExchange (Benson et al., 2018): they contain respectively 3,029/1,629 nodes and 271,233/822,059 timestamped events. Excluding singleton events, which fall outside the link-prediction candidate definition, leaves 219,076/558,272 interactions. A fixed node dictionary is built from each complete triplet, so the experiment evaluates temporal shift in a transductive node universe; event labels and future incidences remain split chronologically. The frozen scorers use either Deep Sets (Zaheer et al., 2017) or a generic set self-attention encoder (Lee et al., 2019). Independent order blocks are encoded and summed, and the complete-event baseline view alone is used for fitting. This view is the sentinel $Q_K = \mathrm{full}$: it includes every projected subset block and an additional exact-event incidence block. Each finite action $q \le K$ is a distinct projection, including at $q = K$. Each block is constructed from its own projected history; action q never instantiates blocks above q or the full-event block. In this additive implementation that restriction is algebraically equivalent to a zero visibility mask, applied to both history and candidate, so a lower-order action cannot read a cached full incidence.

Each seed uses 0–40% for highest-order baseline fitting, four chronological environments in 40–60% for the affine-score rank screen and calibration, four fixed validation windows in 60–80% for order selection, and four held-out test windows in 80–100%. Within each observable arity, an eligible order is selected only when its paired reciprocal-rank loss has negative pooled mean and nonpositive mean in every validation window, while its pooled paired Brier difference is nonpositive; every window must contain at least 30 finite-loss groups. Ties use pooled Brier difference and then lower order. Otherwise the same saved highest-order baseline predictions are returned unchanged. Each positive is paired with five strict-online one-node corruptions of the same arity and past-degree bin. We macro-average mean reciprocal rank (MRR), area under the precision–recall curve (AUPRC), and Brier score over arity and report the minimum window MRR; order selection uses only the preceding intervals.

The four source environments are equal-event chronological blocks, with no topic or label partition. For observed event arity K and candidate order q, environment e contributes the affine-score row $(1, \bar{s}_{e,q,K})$, where $\bar{s}_{e,q,K}$ is that block's mean frozen sigmoid score. The two-column screen has rank two exactly when at least two block means differ. All 132 candidate screens across the two streams, two scorers, three seeds, and eligible (K, q) pairs have empirical rank two; their smallest singular values range from 0.00145 to 0.0633. This is the predeclared eligibility diagnostic for the temporal comparison. The controlled study below supplies the simultaneous finite-sample rank certificate and includes duplicated-row controls.

Each order uses the same arity-conditioned logistic probability calibrator, fitted on 40–60% before order selection. A strictly monotone recalibration within an arity preserves MRR, so calibration is held fixed as a control for the primary ranking comparison. The separate arity study tests how conditioning on observed event size changes the resulting calibrated outcome probabilities.

A three-seed robustness audit on AskUbuntu with the Deep Sets scorer replaces the logistic map by action-specific isotonic regression (Niculescu-Mizil & Caruana, 2005). Each monotone nonparametric map is fitted only on 40–60%, separately by candidate arity, with one pooled map for arities lacking both outcome classes. The audit reruns the same seed-specific training protocol and retains the original 60–80% selection and 80–100% evaluation intervals. It compares CHORD with pooled temporal cross-validation within each newly paired run; it is an additional same-protocol audit, distinct from the saved predictions in Table 1.

Order-rule and out-of-distribution-training comparisons. On the same frozen predictions, we compare CHORD with always-q deployment (q capped by observable arity) and three validation-only rules that minimize, per arity, pooled reciprocal-rank loss, pooled loss plus ten times its four-window variance, or worst-window loss. The full view is a zero-difference candidate, so each rule retains it unless a strictly better validation objective exists; these diagnostics omit the source-rank and Brier checks and use only data through the 80% boundary. We separately fit full-view Deep Sets checkpoints on four strictly chronological environments in 0–40% using environment-balanced empirical risk minimization (ERM), variance risk extrapolation (V-REx; penalty 10 after five warm-up epochs) (Krueger et al., 2021), or group distributionally robust optimization (Group DRO;

exponentiated-weight step 0.01) (Sagawa et al., 2020). All use 25 epochs of the AdamW optimizer (Loshchilov & Hutter, 2019), batch size 256, learning rate 10^{-3}, and weight decay 10^{-5}. Each checkpoint is frozen before the unchanged 40–60% screen/calibration, 60–80% selection, and 80–100% held-out evaluation; the paired baseline and wrapped model therefore differ only in deployed order. These comparisons test order selection after OOD-oriented training; the training objective and its compute budget are identical within every baseline–wrapper pair.

Octahedral-trade projection test and overlap-constrained shifts. Let r be full hyperedge order and $q < r$ the retained projection order. We enumerate the two orientations in (14) for $r \in \{2, 3, 4, 5, 6\}$ and compare every complete labelled q-shadow. In the mixture experiment, c is the outcome-probability center, γ is its signed orientation contrast, and a is the magnitude of the hidden-mixture shift, so the two orientation weights are $(1 \pm a)/2$. We use $r = 4$, $c = 1/2$, $\gamma = 0.3$, and $a = 0.35$, hence source/target weights $(0.675, 0.325)$ and $(0.325, 0.675)$. We draw 50–2500 cells per environment in 500 trials. For the $(r, q) = (4, 2)$ random-hypergraph experiment, 30 cells are independently embedded in $N = 1000$ vertices across 500 trials; the support diagnostic compares coupled whole-cell retention with independent-edge visibility at edge-retention probability $\rho = 0.08$.

Source-matrix rank, target risk, and joint order selection. The rank audit draws independent bounded product-Rademacher environments and multiaffine quadratic responses in feature dimension $d \in \{4, 8, 12\}$. For each d, E denotes the number of source environments and $E^{\mathrm{gen}}_{\min}$ denotes the smallest generic full-rank environment count in Theorem E.10; we use $E = E^{\mathrm{gen}}_{\min} - 1, E^{\mathrm{gen}}_{\min}, E^{\mathrm{gen}}_{\min} + 1$, generic, duplicated, and collinear mean designs, $n \in \{256, 1024, 4096, 16384\}$ samples per environment, and three seeds. Forty-eight bootstrap replicates estimate operator and response radii. Stacking environment moment equations yields the source operator A and response vector b. The rank certificate is strictly $\widehat{\sigma}_{\min}(A) > \epsilon_A$, where ϵ_A is the simultaneous operator-error radius; A and b estimated from the same observations use a Bonferroni dependence-valid joint radius and only these aggregates centrally.

The canonical Gaussian audit fixes the allowed source-nullspace coefficient radius $R = 0.4$, noise standard deviation $\sigma = 0.3$, source rank eight, smallest nonzero singular value 0.12, and crosses four source sample sizes with components of the target conditional feature mean in the source nullspace. Each of 36 settings uses 6000 Monte Carlo repetitions to test Equation (118).

The executable joint projection/source-moment-order experiment uses the $r = 4$ octahedral orientation $S \in \{-1, 1\}$ from Proposition E.2. The projection orders are $q \in \{r - 1, r\}$ with feature maps 1 and $(1, S)$; the source moment orders are $\ell \in \{0, r\}$ with moment features 1 and $(1, S)$. Bernoulli mechanisms $\Pr(Y = 1 \mid S) = 1/2 + \gamma S$ use $\gamma \in \{0, 0.22\}$, respectively the cases where the outcome mean is constant or varies within projection classes. We cross generic distinct orientation means and duplicated mean 0.25, $E \in \{1, 2, 3\}$, $n \in \{256, 1024, 4096\}$ observations per environment, dense or 0.05 high-order target-measurement retention, and three seeds, for 216 cells. Source releases remain fully observed; only the unlabelled target sample enters the target-adaptive support check in Section 3.

Each pair freezes its source moment estimate constrained to the coefficient ball $\{\theta : \|\theta\|_2 \le B\}$ with $B = 0.75$. A source-only reference distribution with uniform S releases the pair-specific scalar residual upper confidence bound (UCB). The evaluation target has $\Pr_\star(S = 1) = 0.57$, giving exact representation-risk transport multiplier $\kappa^{\mathrm{rep}}_q = 1.14$ and additive allowance $\tau^{\mathrm{rep}}_q = 0$. The transported squared residual UCB supplies the risk bound; a disjoint validation half supplies the fixed-reference moment-defect UCB.

For $q = r$, each of the two orientation cells must pass in each of two target splits: a simultaneous 0.98 Hoeffding lower bound must be positive, at least 16 observations must be retained, and the known upper bound on unobserved and rare-cell mass has upper bound at most 0.01. The representation-error check requires $\kappa^{\mathrm{rep}}_q\{\widehat{\Delta}_{q,\ell} + \mathrm{rad}(\widehat{\Delta}_{q,\ell})\}^2 \le 0.03$; we set the admissible risk threshold $\eta^2_{\mathrm{risk}} = 0.03$. We use the 0.98 bootstrap quantile with 64 replicates, source/deployment increments 0.01 and 0.005, and zero tie slack. Pairwise bootstrap radii and the simultaneous support radius instantiate the predeclared upper-confidence-bound procedure.

Calibration and dependence robustness. The overlap audit uses a balanced two-state projection cell with $\gamma = 0.3$, $\rho_{\mathrm{fib}} \in \{0, 0.02, 0.05, 0.1, 0.2, 0.35, 0.5\}$, and $\kappa \in \{1, 1.05, 1.1, 1.25, 1.5, 2, 4, \infty\}$. The conditional-feature-mean audit uses an eight-dimensional full-rank operator, $B = 0.5$, source supports 64–4096, six errors $\epsilon_a \in [0, 0.2]$, aligned, orthogonal, and random directions, and 8,000 repetitions per seed. For approximate factorization we cross four sample sizes 100–8000, oscillations 0–0.1, and $\rho_{\mathrm{fib}} \in \{0.05, 0.2, 0.4\}$ at tolerance 0.02, with 5,000 repetitions. The dependence audit uses 12 environment rows with 256 observations each, equicorrelation $0, 0.25, 0.5, 0.75, 0.9$, the exact joint Gaussian covariance, and 20,000 repetitions per seed. All grids use three predeclared seeds and exclude the held-out temporal labels. A separate release-only sensitivity replay crosses $\Gamma^{\mathrm{shift}} \in \{0, 0.001, 0.0025, 0.005, 0.01, 0.02, 0.05\}$, one, two, or four chronological validation windows, minimum support counts 4–64, mass tolerances 0.005–0.05, and conditional-feature-mean errors up to 0.12. It accepts only predeployment action records and has no interface for held-out metrics; thresholding and selection use data through the 80% boundary.

Pure r-way dependence and sparse additive models. For the pure r-way construction we take $r \in \{3, 4\}$, $N \in \{512, 2048, 8192\}$, interaction strength $\tau = 0.7$, five seeds, and 12,000 candidates per class. Exact enumeration checks all proper marginal identities before simulation. The sparsity exponent α crosses the theoretical boundary $1/r$ and extends into the regime where lower-order faces are well observed but r-way information vanishes. We report both the full likelihood ratio and $W = \sum_j \prod_i X_{ji}$, the number of outside nodes for which all r replacement edges occur.

The sparse additive experiment uses $r \in \{3, 4\}$, N up to 1024, five seeds, and a predeclared sparsity grid. It compares an $(r-1)$-face score, aligned-face statistic, node degree, and a degree-based fallback selected from observed support. This setting matches Theorem F.8; the pure-interaction regime is evaluated separately. The arity experiment uses five seeds and 30,000 training observations per arity. It trains on $K \in \{2, 3, 4, 5\}$ and tests exact population risks on the unseen arities $K \in \{6, 7, 8, 9, 10\}$. In the decision-reversal regime, $\pi = 0.60$ and the member error probability is $\varepsilon = 0.075$. Exact discrete enumeration evaluates a shared set-size-agnostic logistic head and the fitted set-size-conditioned Bayes posterior from (147); calibration error uses the two bins $T_K \in \{-1, +1\}$.

I Additional quantitative evidence

I.1 Fixed-budget spectral and certification checks

These experiments use the Gaussian population-gradient acquisition channel of Section 6. They isolate the structured inverse problem from representation fitting. We use the explicit weighted frame, equal environment allocation, and exact target coefficient loss. Covariance quantities are computed directly from the declared information matrix; simulation is confined to the finite-sample power check. The value channel has the constant-factor comparison proved in Lemma C.1.

Changing the environment count at fixed budget. For $d = 36$, set $m \in \{0, 2, 4, 6, 8\}$, $E = d - m + 1$, and $w = \max\{1, m^{1/4}\}$. In Frobenius-orthonormal coordinates for zero-diagonal Q, form the exact information matrix of the map $Q \mapsto QV_{\mathrm{env}}$. If its eigenvalues are λ_j, then the gradient-only covariance Σ_Q has eigenvalues $\sigma^2(d-m)/(N\alpha^2\lambda_j)$. Define the normalized covariance $C_Q = (N\alpha^2/\sigma^2)\Sigma_Q$. Table 3 reports $\operatorname{tr} C_Q$ and $\|C_Q\|_F$, valid for every declared $\alpha \le 1/2$. Here Σ_Q uses full Frobenius coordinates of Q. Prediction-energy coordinates multiply both by $1/2$. The linear coefficients add an independent covariance $\sigma^2 I_d/N$ and do not alter the leading rates.

Table 3: Same total budget, different environment spans at $d = 36$. The two columns measure integrated estimation variance and dense-energy testing scale in the same normalized quadratic channel.

Environments E	Unspanned m	Weight w	$\operatorname{tr} C_Q$	$\|C_Q\|_F$
29	8	1.682	111822.73	28032.66
31	6	1.565	56402.97	13041.23
33	4	1.414	31336.58	4086.14
35	2	1.189	24025.70	1021.07
37	0	1.000	22680.00	903.59

The experiment makes the task dependence visible at one fixed dimension. Moving from 29 to 33 environments reduces estimation variance by a factor 3.57 and the testing scale by 6.86, with the same acquisition budget. The asymptotic distinction is established by the matching theorems; these finite-dimensional values expose its source in the actual information matrix.

Equal learning order, different certification order. At the algebraic boundary $d = m(m+1)/2$, we compare the unweighted frame $w = 1$ with $w = m^{1/4}$ for $m \in \{4, 6, 8, 12, 16, 24, 32\}$. Table 4 shows four representative dimensions; every dimension in the declared grid is used in the calculation. If a_j are the singular values of D_U, the normalized inverse second sum remains bounded for both designs, while the inverse fourth sum grows in the unweighted design and stays bounded after weighting. These deterministic spectral diagnostics evaluate the theorem's two normalizations directly at every declared dimension.

Table 4: Critical-count inverse spectra: the weighted design attains both optimal spectral orders.

		$m^{-4}\sum_j a_j^{-2}$		$m^{-6}\sum_j a_j^{-4}$	
m	d	$w=1$	$w=m^{1/4}$	$w=1$	$w=m^{1/4}$
4	10	0.707	0.552	2.467	1.248
8	36	0.942	0.554	8.534	1.908
16	136	1.087	0.500	18.738	1.925
32	528	1.166	0.438	36.261	1.731

Finite-sample certificate power. We also simulate the exact gradient-channel Gaussian coefficient statistic at $d = 15$, $m = 5$, $w = 5^{1/4}$ with known $b = 0$. Write $F = \|\Sigma_Q\|_F$, $M = \|\Sigma_Q\|_{\mathrm{op}}$, and $T = \|\widehat{Q}\|_F^2 - \operatorname{tr}\Sigma_Q$. Set $\delta = 0.05$, $\beta = 0.10$, and accept when $T \le t_\beta := \sqrt{2/\beta}\,F$. The declared tolerance $\varepsilon > t_\beta$ is the larger solution of $\delta(\varepsilon - t_\beta)^2 = 2F^2 + 4M\varepsilon$. This analytic calibration guarantees null power at least 0.90 and false acceptance probability at most 0.05 whenever $\|Q\|_F^2 \ge \varepsilon$; the same conversion by $1/2$ gives target prediction-energy tolerance. In the covariance eigenbasis, alternatives have coordinates $\sqrt{v}\lambda_j/F$, with $v/\varepsilon \in \{0, 0.10, 0.25, 0.50, 1, 2\}$. Each uses 20 000 independent Gaussian repetitions with a fixed random seed. At the null, acceptance is 0.9845; at the declared unsafe boundary and twice that energy, there are no acceptances in either set of 20 000 trials. The intermediate acceptance rates are 0.26845, 0.00380, and 0. These sampled directions check the certificate's implementation; the uniform safety statement follows from the theorem, independently of these trials. Rescaling the covariance and tolerance together places the entire experiment inside the bounded Bernoulli-valid coefficient class without changing any acceptance probability.

Temporal-benchmark order-selection audit. Table 5 compares order rules on exactly the frozen predictions used in Table 1. Fixed low-order rules lower mean reciprocal rank (MRR) in every stream–scorer setting. Pooled temporal cross-validation (CV) and its mean–variance version can improve their aggregate mean while still making harmful seed-level changes; the four-window rule in CHORD produces the same held-out decisions as worst-window validation on this grid.

Table 5: Validation-only order rules on the four real stream–scorer settings. For each method, ΔMRR is wrapped minus full-view baseline. The interval spans the four three-seed setting means; "changed" counts observable-arity–seed decisions out of 48, and "harmful" counts seed-level MRR below the paired frozen baseline out of 12.

Order rule	changed	ΔMRR range	harmful
CHORD	7	$[+0.0008, +0.0046]$	0/12
Always $q = 1$	48	$[-0.4156, -0.3179]$	12/12
Always $q = 2$	48	$[-0.0644, -0.0333]$	12/12
Always $q = 3$	48	$[-0.0234, -0.0148]$	10/12
Pooled temporal CV	15	$[+0.0002, +0.0054]$	2/12
Mean $+10$ Var CV	15	$[+0.0002, +0.0054]$	2/12
Worst-window validation	7	$[+0.0008, +0.0046]$	0/12

Table 6 lists the seven seed runs containing a switch. Each run changes only the $K = 2$ decision from the complete-event view to $q = 2$; arities $K = 3, 4, 5$ replay the baseline exactly. The table reports run-level macro changes, all favorable before rounding on the held-out period.

Table 6: Held-out run-level changes for all seven seed runs containing a CHORD switch. "Ask" denotes AskUbuntu, "Math" Math StackExchange, and Attn the self-attention scorer. MRR, AUPRC, and Brier are equally macro-averaged over $K \in \{2, 3, 4, 5\}$; only $K = 2$ changes. Worst MRR is the minimum, over four test windows, of each window's macro MRR. Each Δ is CHORD minus its paired run-level baseline; Brier gain reverses that subtraction.

Stream	scorer/seed	ΔMRR	ΔAUPRC	Brier gain	Δworst MRR
Ask	Deep Sets/1	$+0.0062$	$+0.0056$	$+0.0017$	$+0.0098$
Ask	Deep Sets/2	$+0.0052$	$+0.0010$	$+0.0005$	$+0.0021$
Ask	Attn/1	$+0.0060$	$+0.0107$	$+0.0007$	$+0.0073$
Ask	Attn/2	$+0.0079$	$+0.0094$	$+0.0007$	$+0.0105$
Math	Deep Sets/0	$+0.0023$	$+0.0006$	$+0.0012$	$+0.0036$
Math	Attn/0	$+0.0036$	$+0.0018$	$+0.0008$	$+0.0031$
Math	Attn/1	$+0.0022$	$+0.0024$	$+0.0008$	$+0.0026$

Selected actions and cost. The temporal-benchmark audit selects $q = 2$ only where its evidence is strongest: 7 observable-arity-2 strata change and the remaining 41 exactly replay the full view. The controlled grid separately

realizes coarse, full-order, and abstaining actions under higher-order mechanisms. After out-of-distribution (OOD)-oriented training, 5 AskUbuntu strata deploy $q = 2$ and 67 of 72 strata replay the baseline, including all 36 Math StackExchange strata. Across the four frozen held-out configurations, backbone training takes 828–856 seconds and the phase computing CHORD plus every diagnostic order rule takes 95.6–101.9 seconds; peak process memory is 1492–1626 MiB. A held-out full-view pass takes 20.34–20.44 seconds versus 13.63–13.81 for $q = 2$. Exhaustively evaluating all three lower-order views for the offline audit costs 1.95–$1.97\times$ one full-view pass, whereas deployment executes only the chosen view. In a separate single-threaded remote-CPU audit with seven repetitions per size, exact water filling over $J = 32{,}768$ full states, where J is the projection-cell cardinality, takes 336.60 ms median; a dense singular value decomposition of a 4096×128 source matrix takes 25.09 ms; and selecting among 65,536 certified order–moment candidates takes 25.53 ms. These measurements match the linear water-filling and candidate-scan costs and the dense matrix-factorization cost stated in Appendix G.

Compatibility with OOD training. Table 7 pairs the same wrapper with full-view Deep Sets trained by empirical risk minimization (ERM), variance risk extrapolation (V-REx), or group distributionally robust optimization (Group DRO). On AskUbuntu, every objective has positive mean MRR, Brier gain, and worst-window MRR after wrapping; on Math StackExchange, all three objectives invoke exact fallback in every seed. Across these and the frozen-scorer study, all 30 paired runs pass the predeclared checks for dataset identity, strict temporal construction, unchanged checkpoint, selection restricted to the first 80% of each stream, and exactly equal fallback predictions.

Table 7: Full-view OOD-trained baseline → the same checkpoint plus CHORD. Empirical risk minimization (ERM), variance risk extrapolation (V-REx), and group distributionally robust optimization (Group DRO) use the same training budget. Values are mean changes over three seeds; changed/fallback counts observable-arity–seed decisions out of 12. The Δ columns are wrapped minus baseline, and Brier gain is baseline minus wrapped score.

Stream	trainer	chg. fallback	ΔMRR	Brier gain	Δ worst MRR
AskUbuntu	ERM	2/10	+0.0036	+0.0006	+0.0033
AskUbuntu	Group DRO	2/10	+0.0033	+0.0005	+0.0031
AskUbuntu	V-REx	1/11	+0.0003	+0.0002	+0.0004
Math StackEx.	ERM	0/12	0.0000	0.0000	0.0000
Math StackEx.	Group DRO	0/12	0.0000	0.0000	0.0000
Math StackEx.	V-REx	0/12	0.0000	0.0000	0.0000

Convergence for the octahedral-trade construction. Here r is full hyperedge order and q is projected order. All 15 nontrivial (r, q) pairs, together with five $q = 0$ checks, have equal complete collections of proper q-subsets (their q-shadows) and disjoint full-order edges. All 500 random embeddings preserve both simple and multihypergraph shadows without collision. Table 8 converges to projected gap 0.21 and transfer regret 0.0441.

Table 8: Projected score-gap convergence across 500 trials with n sampled projection cells per environment. The gap is mean ± standard deviation; remaining columns are trial averages.

n / environment	score gap	reversal recovered	transfer regret
50	0.202 ± 0.097	0.824	0.0419
100	0.212 ± 0.067	0.968	0.0448
250	0.213 ± 0.045	1.000	0.0452
1000	0.211 ± 0.021	1.000	0.0444
2500	0.210 ± 0.014	1.000	0.0442

The rank transition is geometric. In the quadratic audit, d is mechanism dimension and E is the number of source environments. All 27 generic designs obey Equation (119); duplicated and collinear designs have the predicted additional deficiency. The sharp generic thresholds are $E = 3, 6, 9$ for $d = 4, 8, 12$, while counting scalar equations suggests only $E = 3, 5, 7$. All 252 rank-deficient runs reject full rank, with 0.981 bootstrap coverage. Across 36 Gaussian settings, the ratio of empirical to exact minimax risk averages 1.004 and ranges from 0.963 to 1.044.

Table 9 gives the smallest example used by the selector. For full order r, the mechanism feature is $\chi = (1, S)^\top$, where the orientation $S \in \{-1, 1\}$. A constant source moment of order $\ell = 0$ contributes row $(1, \mu_e)$ in environment e, with $\mu_e := \mathbb{E}_e S$; stacking the rows forms the source matrix $A_{r,0}$. One row or repeated values of μ_e have rank one; two distinct means have rank two and identify both mechanism coordinates.

Table 9: Environment-rank transition in the dense $r = 4$ controlled study at $n = 4096$ source observations per environment and mechanism contrast $\gamma = 0.22$. The final column reports the selected source moment order over the indicated environment counts and three seeds.

Source design	E	$\mathrm{rank}(A_{r,0})$	selected ℓ
One generic environment	1	1	r in $3/3$
Distinct environment means	$2, 3$	2	0 in $6/6$
Repeated mean $\mu_e = 0.25$	$1, 2, 3$	1	r in $9/9$

Why the excess-risk certificate uses one-sided upper bounds. Let H denote the complete candidate context. All 76 selections made by the procedure certifying excess Brier risk relative to the Bayes rule given H satisfy the structural and target-support requirements, and their upper confidence bounds cover exact target excess in 76/76 cases. A comparator that selects upon passing an independent validation-residual test makes 29/216 selections that violate those requirements (all with $n = 256$ source observations per environment and an outcome mean that varies within projection classes, always at $q = r - 1$). The largest admitted upper confidence bound is 0.02979 and maximum exact excess 0.00202 at tolerance 0.03.

At $n = 4096$, all 36 settings whose outcome mean is constant within projection classes select $(q, \ell) = (r-1, 0)$; all 18 dense settings whose outcome mean varies within classes select $q = r$; and the same procedure abstains in all 18 sparse cells. Dense cases choose $\ell = r$ for generic $E = 1$ (3/3) and duplicates (9/9), but $\ell = 0$ for generic $E \in \{2, 3\}$ (6/6). Both sparse cells appear in the retained sample, yet the known 0.95 unmeasured target mass fails the predeclared support requirement— the distinction checked by Algorithm 1.

The same grid separates in-distribution order value from OOD deployability. When $\gamma = 0.22$, revealing S gives a strict full-order Bayes gain, so a stepwise order-value analysis (Bian et al., 2025) favors the richer view. CHORD retains it in all 18 dense cells where source rank and target support also suffice, but abstains in all 18 sparse cells with the same high-order mechanism; when $\gamma = 0$, it safely coarsens. Thus predictive value alone is insufficient to establish identification and support for transfer.

Calibration and dependence robustness. We audit the quantities that enter the population bounds. In a balanced two-state projection cell with mechanism contrast $\gamma = 0.30$, endpoint reweighting attains the exact score width and midpoint Brier regret in all 56 combinations of within-cell total-variation radius ρ_{fib} and likelihood-ratio cap κ. The source score gap between projected predictors B and A is 0.265; its worst target value remains positive at $(\rho_{\mathrm{fib}}, \kappa) = (0.20, 2.0)$, with worst target gap 0.020, and reverses at $(0.20, 4.0)$, with gap -0.015. This transition agrees with the exact water-filling boundary.

For the target conditional-feature mean, let ϵ_a bound its Euclidean estimation error and let B bound the mechanism-coefficient norm. All 216 cells obey the analytic risk upper bound obtained in this orthogonalized audit by adding $B^2\epsilon_a^2$ to the declared sampling term, with $B = 0.5$; aligned errors attain the bound, and Monte Carlo risk is within 3.09% of its analytic value. A simultaneous 95% upper bound covers the true outcome-probability variation inside a projection cell in all 240,000 draws and makes 0 unsafe certifications in 60,000 unsafe trials. With 12 equicorrelated environment releases, the dependence-aware interval maintains 94.88–95.06% coverage. At correlation 0.9, treating releases as independent gives only 44.54% coverage and the valid radius is $3.30\times$ larger.

Predeployment sensitivity yields conservative fallback. Let $\Gamma^{\mathrm{shift}} \geq 0$ be the declared allowance for source-to-target change in paired Brier risk. Replaying only the validation record, the number of changed temporal-benchmark strata is $7, 7, 5, 2, 0, 0, 0$ as $\Gamma^{\mathrm{shift}} = 0, 0.001, 0.0025, 0.005, 0.01, 0.02, 0.05$; every rejected change is an exact full-view fallback. Using one pooled, two, or four chronological validation windows gives respectively $12, 9, 7$ changes, so finer temporal resolution makes the stability check stricter. Across 24 combinations of minimum support count 4–64 and mass tolerance 0.005–0.05, the support audit retains the same 76 accepted actions from the 216-setting controlled study, with 76/76 risk-bound coverage. Increasing conditional-feature-mean error ϵ_a from 0 to 0.12 reduces certified pairs from 76 to 11 while maintaining 100% conditional coverage. Each sensitivity rule uses only the predeployment record.

The pure high-order information boundary. The maximum proper-marginal discrepancy is 2.22×10^{-16}. For $r = 3$ and $N = 512$ outside nodes, the information scale $I_N := N\tau^2\{\rho(1-\rho)\}^r$ equals 1.148, 0.328, and 0.00770 at sparsity exponents $\alpha = 0.25$, $1/3$, and 0.55; the corresponding optimal errors are 0.1403 ± 0.0017, 0.2794 ± 0.0011, and 0.4837 ± 0.0026 over five seeds. Here $\rho = N^{-\alpha}$ is edge density and $\tau = 0.7$ is interaction strength. At $N = 8192$, every proper face remains observable with probability at least 0.9999, yet the optimal error and the all-one-row-count error approach $1/2$ as I_N vanishes.

Set-size conditioning and the sparse additive model. On unseen arities, conditioning on observed set size improves (Brier score, classification error, two-bin calibration error) from $(0.246, 0.428, 0.096)$

to $(0.233, 0.385, 0.002)$. At full interaction order $r = 4$, $N = 1024$ outside nodes, and edge density $\rho_N = N^{-\alpha}$ with sparsity exponent $\alpha = 2$, face incidence gives 0.501 ± 0.0004 area under the receiver operating characteristic curve (AUROC), compared with 0.571 ± 0.0034 for node degree and 0.571 ± 0.0033 for its support-selected fallback, as predicted by Theorem F.8.

Nonparametric calibration with the same temporal protocol. An additional three-seed AskUbuntu/Deep Sets audit uses isotonic calibration (Niculescu-Mizil & Caruana, 2005), fitted only on the 40–60% segment. Each seed trains an additional checkpoint under the same fixed protocol and then freezes it for both order rules. Every comparison below is within that run and shares its isotonic-calibrated baseline; no cross-checkpoint change is attributed to calibration. CHORD changes one of the 12 arity–seed decisions, replaying the other 11 exactly, with mean MRR gain 0.001999 and Brier-score reduction 0.000192. Pooled cross-validation changes 6 decisions and produces 2 harmful seed-level MRR changes; CHORD produces none. Table 10 reports every seed, keeping this calibration comparison separate from the original four-setting benchmark.

Table 10: Order selection after isotonic calibration on AskUbuntu with Deep Sets. Each entry is a held-out change relative to the same calibrated baseline in that seed. MRR is mean reciprocal rank, ΔMRR is selected minus baseline MRR, and Brier gain is baseline minus selected Brier score.

	Pooled cross-validation		CHORD	
Seed	ΔMRR	Brier gain	ΔMRR	Brier gain
0	−0.000731	+0.000159	0	0
1	−0.002058	−0.002339	+0.005997	+0.000577
2	+0.002175	+0.000164	0	0
Mean	−0.000205	−0.000672	+0.001999	+0.000192

Misspecified overlap and valid upper envelopes. We separately test deliberate underestimation of the shift class in two three-state projection cells. Reference weights and outcome means are $\mu_A = (0.55, 0.30, 0.15)$, $\eta_A = (0.15, 0.45, 0.85)$, $\mu_B = (0.20, 0.50, 0.30)$, and $\eta_B = (0.20, 0.60, 0.90)$, giving source gap $m_0(B) - m_0(A) = 0.265$. True radii range over $\rho \in \{0.03, 0.075, 0.15, 0.275, 0.425\}$ and $\kappa \in \{1.05, 1.20, 1.40, 1.75, 3, 7\}$. Declared values are drawn from $\{0, 0.01, 0.025, 0.05, 0.10, 0.20, 0.35, \rho\}$ and $\{1, 1.05, 1.10, 1.25, 1.50, 2, 4, \kappa\}$, respectively, retaining only coordinatewise underestimates with at least one strict inequality. This gives 930 deterministic cases. We compare their exact intervals with those obtained by rounding each true parameter upward to the next element of $\{0.05, 0.10, 0.20, 0.35, 0.50\}$ and $\{1.10, 1.25, 1.50, 2, 4, 10\}$, respectively. These upper envelopes are deliberately known to contain the true ambiguity classes.

The upper-envelope intervals contain both true score intervals in all 930 cases and make 533 ranking certifications, all valid. Narrow intervals formed from the underestimated values make 913 certifications, of which 204 permit an actual reversal. For example, true $(\rho, \kappa) = (0.275, 3)$ gives worst gap -0.0975, whereas declaring $(0, 1)$ reports $+0.265$; the containing envelope $(0.35, 4)$ returns -0.1875 and rejects certification. This finite enumeration tests robustness conditional on valid parameter upper bounds; it does not estimate overlap from a real stream or treat the 930 cases as independent statistical trials.