CONSTRAINED POLYNOMIAL PROBABILISTIC ZONOTOPES: WHEN CAN ONE COMPRESSED MODEL SERVE MANY TASKS?

Anonymous authors
Paper under double-blind review

ABSTRACT

When can one compressed probability model meet several prediction requirements simultaneously? We introduce Constrained Polynomial Probabilistic Zonotopes (CPPZs), which combine polynomial constraint geometry, prescribed intrinsic probability laws, and explicit observation semantics. We study shared finite-coordinate continuous representations for declared tasks under worst-case squared prediction risk. For a class of elliptic task families of common order with scalar principal symbols, task-mixture constraints exactly characterize the jointly attainable leading-order risk region, without a common eigenbasis. For the stated nonlinear CPPZ family with smooth, strictly convex geometry, requiring one shared reconstruction with its prescribed intrinsic law preserves the optimal leading-order risk region attainable by unrestricted task-specific decoders of shared codes. This holds for each fixed finite task family differing only in context distribution and every fixed source radius strictly inside the geometric validity domain; each target admits a shared code. The construction separates coarse geometry from residual information and compiles recovery into a valid polynomial probability object. Within the same family, one geometry-dependent measurement update preserves the leading performance of deterministic adaptive linear acquisition, with sublinear additional queries. For two compact linear tasks on a real Hilbert source ball, continuous, linear, and deterministic adaptive-linear encoding have identical attainable risk regions at every coordinate budget, without ellipticity assumptions. Yet complete pairwise information does not determine collective feasibility: two CPPZ task triples with a common finite-context detector can have identical pairwise risk regions at every budget while differing on whether one code meets all three tasks' requirements at the same budget.

1 INTRODUCTION

Accurate predictions do not by themselves specify a common probability model. For a surface sampled uniformly by area, geometry also determines the distribution of normal directions: these are not independent choices. Separate task decoders need not enforce this compatibility. If a compressed representation must recover one such object, does consistency require more coordinates or higher risk?

We introduce Constrained Polynomial Probabilistic Zonotopes (CPPZs), coupling polynomial constraint geometry, prescribed intrinsic laws and explicit observation semantics. Both comparison classes use a continuous shared code, charged by its real coordinates and evaluated by worst-input squared risks. They differ in decoding: separate task forecasts or one legal CPPZ serving all tasks. Codes may depend on declared tasks and risk requests, but not the realized task.

Our central result lifts joint predictive capability to common-object recovery without an extra leading-order risk penalty. In the nonlinear family studied here, requiring one prescribed-law CPPZ preserves the entire optimal leading-order risk region, including rates and leading constants as the coordinate budget grows. This holds for any fixed finite task family differing only in context distributions, at each fixed source radius in the smooth, strictly convex geometric interior, under the stated regularity and smoothing-detector assumptions. Each target has its own shared code (Figure 1).

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.

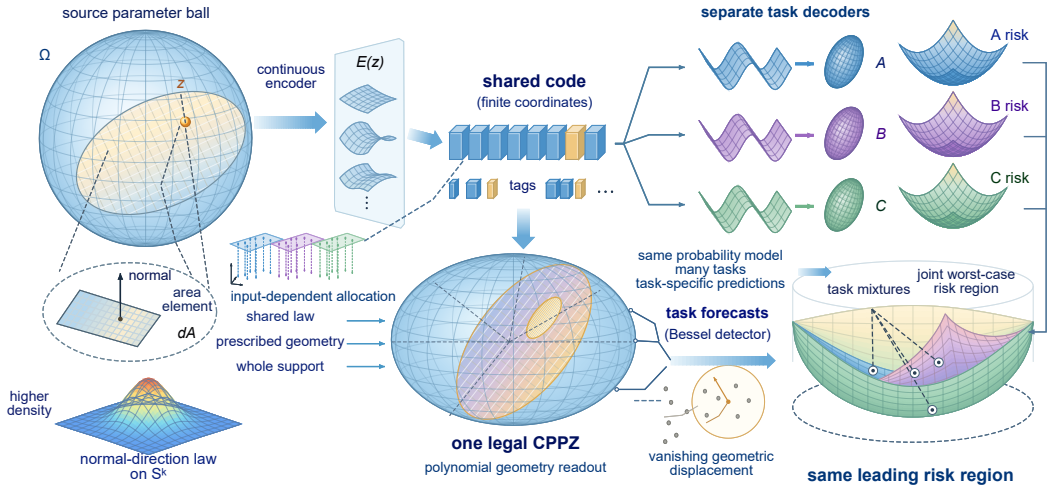


Figure 1: **One law, many risk requirements.** Separate task decoders and one prescribed-law CPPZ attain the same leading risk region in Theorem 4.1's native family at fixed interior radii; each target has a shared code under either decoding constraint. Insets show allocation and relative-tail recovery. Surfaces, densities and risk bowls are schematic.

Two obstacles make this lifting nontrivial. Bounds for weighted task combinations exclude impossible requests but do not construct one code meeting them all. For the stated elliptic tasks, we make these bounds jointly sufficient without a common eigenbasis. A shared decomposition supports input-dependent coordinate allocation, including its description in the budget. To recover a nonlinear object, we first resolve coarse geometry. The residual's geometric displacement then vanishes even when its source norm does not, permitting polynomial reconstruction with the prescribed law without shrinking the source ball.

Preserving this region does not make measurement choices or pairwise task checks interchangeable. For two compact linear tasks on a real Hilbert source ball, every feasible request has a fixed projection at the same budget. In our nonlinear family, two batches match arbitrary-depth deterministic adaptive linear acquisition to leading order with sublinear additional queries, yet fixed measurements can fail. Separate constructions have identical complete pairwise risk regions but different collective requirements. These boundaries distinguish common-object consistency from measurement sufficiency and collective task feasibility.

2 Related Work

Prescribed-law recovery is an output requirement. CPPZs combine polynomial constraint-and-readout syntax (Kochdumper & Althoff, 2023), geometric reference measures (Radul & Alexeev, 2021) and explicit observation semantics. Unlike learning a manifold distribution (Caterini et al., 2021), recovery must maintain the law prescribed by reconstructed geometry. Here the question is the capacity cost of enforcing that law, not only representational expressiveness.

Shared compression imposes different requirements. Goal-oriented low-rank approximation in linear Gaussian inverse problems includes compatibility with an approximate parameter posterior (Spantini et al., 2017). Universal rate–distortion–perception coding fixes an encoder across requests (Zhang et al., 2021); TAMC learns task groups and entropy dependencies (Guo et al., 2025). Our request-dependent codes use real-coordinate budgets for simultaneous worst-input guarantees through one geometry-prescribed law.

Joint attainability requires more than weighted objectives. Hu et al. (2023) study weighted-loss Pareto points in linear multi-task regression; our mixture constraints bound information requirements, with sufficiency proved by a common-code construction. Stable approximation imposes Lipschitz bounds on encoding and recovery (Cohen et al., 2022); best-term approximation permits

source-dependent term selection (Nguyen & Nguyen, 2022). Our continuous message includes its allocation, with budget-dependent stability bounds.

Measurement restrictions distinguish information models. Fair and multi-criteria PCA optimize trace-based group objectives, with exact two-criterion solutions (Samadi et al., 2018; Tantipongpipat et al., 2019). Our two-task result equates worst-input risk regions on a Hilbert ball across fixed-linear, continuous and deterministic adaptive-linear information, using classical matrix-pencil geometry (Johnson & Rodman, 1985). Frozen-path adaptive-information lower bounds are also classical (Krieg et al., 2025). Our nonlinear two-batch construction handles changing background geometry and an infinite residual at fixed radius, charges coarse queries and recovers one valid probability object.

3 CPPZ MODELS AND PREDICTION RISK

A CPPZ is $\mathcal{C} = (\mathcal{S}, \{(\mathfrak{R}_s, \Pi_s, F_s)\}_{s\in\mathcal{S}}, \omega, \{K_t\}_{t=1}^T, \mathsf{v})$, coupling polynomial geometry, prescribed laws and observations across named strata and discrete branches (Definition A.1). Our nonlinear common-law recovery uses the regular pure subclass: a regular carrier $M = \{z \in \Omega : Q(z) = 0\}$, metric g, prescribed positive density ρ, and polynomial readout X, with

$$d\Pi = Z^{-1}\rho\, d\mathcal{H}^k_g|_M, \qquad P = X_{\#}\Pi, \qquad Z = \int_M \rho\, d\mathcal{H}^k_g.$$

Admissibility requires regularity, boundary-nullness, finite required moments and positive finite structural and evidence normalizers. The ambient density version is specified, so restriction to a lower-dimensional carrier is well-defined. For an embedding X, weighting by its area Jacobian gives uniform surface probability. Appendix A.2 establishes representation boundaries, strict affine extension and admissible composition.

3.1 NATIVE GEOMETRY, TASKS AND INFORMATION

On S^k, $k \geq 2$, let σ be normalized area, g the metric and Δ the nonnegative Laplace–Beltrami operator. Fix a smooth support $h_0 > 0$ with $A_0 = \nabla^2 h_0 + h_0 g \succeq \kappa g$, $\kappa > 0$, and set

$$E_r = (I + \Delta)^{-r/2}, \quad h_z = h_0 + E_r z, \quad A_h = \nabla^2 h + hg,$$
$$J_h = \det_g A_h, \quad p_h = J_h/Z_h, \quad Z_h = \int J_h\, d\sigma.$$

The real source space is $H = \{z \in L^2(S^k) : E_r z \perp \mathrm{span}(1, u_1, \ldots, u_{k+1})\}$, with source ball $B_\delta(H)$. This fixes mean support and removes translations. The inverse Gauss map $X_h(u) = h(u)u + \nabla h(u)$ sends $p_h\sigma$ to normalized physical area (Schneider, 2013). Input support functions need not be polynomial; common-law recovery must yield a legal CPPZ with the surface law prescribed by its reconstructed geometry.

Assume $\beta > k$, an integer $s > \beta$, and $r > s + 2 + k/2$. For $\tau_s > 0$, the Bessel operator $K_\beta = (I + \tau_s\Delta)^{-\beta/2}$ has kernel k_β; fix $M_\beta > \sup k_\beta$. A task draws a normal direction $U \sim p_{h_z}\sigma$ and an independent context $V \sim \omega_\tau\sigma$, and observes $Y \mid U, V \sim \mathrm{Bernoulli}(k_\beta(V, U)/M_\beta)$. Declared task weights are smooth, strictly positive and normalized. Its forecast and half-Brier excess are

$$F(z) = M_\beta^{-1} K_\beta p_{h_z}, \qquad \mathrm{Reg}_{z,\tau}(a) = \tfrac{1}{2}\|\sqrt{\omega_\tau}(a - F(z))\|_2^2.$$

Tasks change context distributions, not the object's law; after compression, their optimal forecasts for the same message need not coincide. Write $F_\tau = \sqrt{\omega_\tau}F$ and $T_{b,\tau} = DF_\tau(b)$. This reference-to-decision (R2D) derivative includes geometry, normalization and sensing; its order is $-\mu$, where $\mu = r + \beta - 2$ and $p = k/(2\mu)$.

For compact linear tasks $T_\tau : H \to Y_\tau$ between real Hilbert spaces, unrestricted task decoders, and a continuous code $E : B_\delta(H) \to \mathbb{R}^m$, define

$$\rho^{\tan}_{m,\tau}(E) = \inf_{D_\tau} \sup_{z\in B_\delta(H)} \tfrac{1}{2}\|T_\tau z - D_\tau(Ez)\|^2.$$

Replacing $T_\tau z$ by $F_\tau(z)$ defines nonlinear risks. The upper region $\mathcal{R}_m$ contains targets met by one shared code. For $\mathcal{R}^{\mathrm{nl,CPPZ}}_m$, *one object decoder must meet all inequalities simultaneously*; separate

coordinatewise optimization over object decoders is not allowed. Coordinates, allocation tags and coarse information all count toward m. Public descriptions, polynomial degree and computation are separate resources. Quantization is a subsequent channel, not a continuous finite-alphabet encoder.

4 A COMMON PROBABILITY LAW PRESERVES SHARED CAPACITY

We state the nonlinear result, then its joint encoding mechanism and fixed-radius realization.

4.1 THE COMMON-LAW CAPACITY THEOREM

Fix a finite nonempty task family Θ, with probability simplex Δ_Θ. Let S_r and C_r be the source-to-support and source-to-directional-curvature evaluation norms. They have the harmonic-sum expressions in Appendix A.7. With $h_- = \min h_0$ and κ_- the minimum eigenvalue of A_0, define $\delta_{\mathrm{geom}} = \min\{h_-/S_r, \kappa_-/C_r\}$. Every input in the closed source ball has positive support and curvature exactly for $\delta < \delta_{\mathrm{geom}}$. This is a whole-ball geometric boundary, not a prediction impossibility boundary for individual inputs.

For a background $b \in B_\delta(H)$, the derivative $T_{b,\tau}$ has positive principal energy $a_{b,\tau} = \omega_\tau t_b^2$ on the ordinary metric cosphere, where

$$t_b(u,\xi) = \frac{\tau_s^{-\beta/2}}{M_\beta Z_b}\,\mathrm{cof}_g(A_{h_b}(u))(\xi^\sharp,\xi^\sharp), \qquad |\xi|_g = 1.$$

Here $h_b = h_0 + E_r b$, $Z_b = Z_{h_b}$, and Λ is Liouville measure. Define the background mixture constant

$$\mathsf{h}_b(\pi) = \left[\frac{1}{k(2\pi)^k}\int_{S^*S^k}\left(\sum_{\tau\in\Theta}\pi_\tau\omega_\tau t_b^2\right)^p d\Lambda\right]^{1/p}.$$

The remaining energy fraction is $v_\delta(b) = 1 - \|b\|^2/\delta^2$. Put

$$L_\delta(\pi) = \sup_{\|b\|\le\delta} v_\delta(b)\mathsf{h}_b(\pi), \qquad C_\delta = \{r \ge 0 : \pi\cdot r \ge L_\delta(\pi)\ \forall \pi \in \Delta_\Theta\}. \tag{1}$$

Theorem 4.1 (Common-law capacity in the geometric interior)**.** *Under Section 3's native assumptions, fix a finite nonempty task family and $0 < \delta < \delta_{\mathrm{geom}}$. Then*

$$(2/\delta^2)m^{1/p}\mathcal{R}_m^{\mathrm{nl}} \longrightarrow C_\delta, \qquad (2/\delta^2)m^{1/p}\mathcal{R}_m^{\mathrm{nl,CPPZ}} \longrightarrow C_\delta.$$

Every bounded convergent sequence of normalized feasible profiles has its limit in C_δ. Conversely, each $r \in C_\delta$ admits one continuous task-shared encoder sequence and one task-independent legal pure CPPZ per code, without a side channel, satisfying $\limsup_m (2/\delta^2)m^{1/p}\rho_{m,\tau}^{\mathrm{nl,CPPZ}}(E_m) \le r_\tau$ for every task. Clean source recovery stays in $B_\delta(H)$; complete polynomial filtering adds only lower-order forecast error.

The equality concerns decoder constraints, not task-conditioned coding, every pre-existing code, or exact finite-budget risks.

4.2 THE JOINT CAPACITY CONSTRUCTION

The nonlinear construction builds on joint attainability for linear tasks, here stated beyond the native sphere. Let M be a closed connected smooth k-manifold with positive smooth density and $H = G^\perp \subset L^2(M)$, for finite-dimensional smooth G. Fix a finite nonempty task set Θ. Each Gram $A_\tau = T_\tau^* T_\tau$ is the compression of a positive injective real scalar classical elliptic operator $\widetilde{A}_\tau$ of order -2μ, $\mu > 0$. For a same-class reference $\widetilde{A}_*$, require

$$C_\tau = \widetilde{A}_*^{-1/2}\widetilde{A}_\tau\widetilde{A}_*^{-1/2}, \qquad c_- I \preceq C_\tau \preceq c_+ I, \quad 0 < c_- \le c_+ < \infty.$$

No commutation or common eigenbasis is assumed. Let $p = k/(2\mu)$, and let a_τ denote the positive real even principal energies on the cosphere of $D = \widetilde{A}_*^{-1/(2\mu)}$, with Liouville measure Λ. For π in

the probability simplex Δ_Θ, put

$$h(\pi) = \left[\frac{1}{k(2\pi)^k} \int_{S^*M} \left(\sum_{\tau \in \Theta} \pi_\tau a_\tau \right)^p d\Lambda \right]^{1/p}, \qquad \mathcal{R}_\infty = \{ r \geq 0 : \pi \cdot r \geq h(\pi) \, \forall \pi \in \Delta_\Theta \}. \tag{2}$$

Theorem 4.2 (Shared capacity for elliptic tasks)**.** *Fix this family and $\delta > 0$ before $m \to \infty$. Then*

$$(2/\delta^2) m^{1/p} \mathcal{R}_m^{\tan} \longrightarrow \mathcal{R}_\infty. \tag{3}$$

Every bounded convergent sequence of normalized upper profiles has its limit in $\mathcal{R}_\infty$. Conversely, each $r \in \mathcal{R}_\infty$ has one shared continuous encoder sequence satisfying $\limsup_m (2/\delta^2) m^{1/p} \rho_{m,\tau}^{\tan}(E_m) \leq r_\tau$ for every task.

The full-symbol-compact extension (Appendix B) supplies joint recovery across frozen smooth backgrounds for Section 4.3's nonlinear construction.

Necessity is not a construction. On the radius-δ sphere in the leading $m+1$ directions of $A_\pi = \sum_\tau \pi_\tau A_\tau$, Borsuk–Ulam gives equal-code antipodes. Each task's average squared decoding error is at least its squared signal norm, giving

$$\sum_\tau \pi_\tau \rho_{m,\tau}^{\tan}(E) \geq \frac{\delta^2}{2} \lambda_{m+1}(A_\pi).$$

Weyl asymptotics yield (2). Each inequality concerns a different mixture; none selects the common code required for sufficiency.

Controlling all tasks without a common eigenbasis. One real orthogonal decomposition compares all task errors with common reference energies. At tolerance $\varepsilon > 0$, Lemma B.3 gives $H = F_\varepsilon \oplus V_g \oplus \bigoplus_{j=1}^R V_j$, with finite F_ε recovered exactly. Put $B_j = P_{V_j} \widehat{A}_* |_{V_j}$, including the guard $j = g$. Core eigenvalues satisfy $\lambda_\ell(B_j) \sim C_j \ell^{-1/p}$; the guard satisfies $\limsup_\ell \ell^{1/p} \lambda_\ell(B_g) \leq C_g$. Positive task-dependent coefficients $d_{\tau j}$, continuous in the task, give, for $e = \sum_j e_j + e_g$,

$$\langle A_\tau e, e \rangle \leq (1+\varepsilon) \sum_{j=1}^R d_{\tau j} \langle B_j e_j, e_j \rangle + d_{\tau g} \langle B_g e_g, e_g \rangle, \qquad \sup_\tau d_{\tau g} C_g \leq \varepsilon.$$

The same decomposition preserves every mixture constant: with $\bar{d}_j(\pi) = \sum_\tau \pi_\tau d_{\tau j}$,

$$\sup_{\pi \in \Delta_\Theta} \left| \left[\sum_{j=1}^R (C_j \bar{d}_j(\pi))^p + (C_g \bar{d}_g(\pi))^p \right]^{1/p} - h(\pi) \right| \leq \varepsilon.$$

The first bound controls cross-branch errors; the second preserves leading spectral counts. Thin shells and even-symbol localization give orthogonal cores. Directions with excessive edge or guard energy enter the allocated guard, controlling cross terms without discarding their cost. Exact retained ranks recover the Weyl constants. Cutoffs are fixed before spectral limits, then refined.

Turning mixture bounds into one allocation. Relabel the cores and guard as $K = R+1$ branches, with input-energy and coordinate fractions $y, s \in \Delta_K$. Move a finite guard prefix into the exact block so its remaining spectrum is bounded by $(C_g + \upsilon)\ell^{-1/p}$. The positive coefficients $b_{\tau j}$ include spectral allowances and the factor $1+\varepsilon$. For $f_\tau(y,s) = \sum_j b_{\tau j} y_j s_j^{-1/p}$ and $\bar{b}_j = \sum_\tau \pi_\tau b_{\tau j}$, Lemma B.1 gives

$$\max_{y \in \Delta_K} \min_{s \in \Delta_K} \sum_\tau \pi_\tau f_\tau(y,s) = \left(\sum_j \bar{b}_j^p \right)^{1/p}.$$

Zero numerators contribute zero; positive numerators over zero give infinity. At fixed input-energy fractions y, the upper-risk image is closed and convex. An infeasible target strictly above all mixture bounds would admit a nonnegative separating mixture, contradicting this identity. Uniform slack absorbs localization allowances and gives a continuous allocation with a positive floor, meeting all tasks at once.

Encoding the allocation within the same budget. The allocation must be recovered from the same message. Write $z = b + f + x$, with $b = 0$ for linear recovery, and send

$$E(z) = \big(b, f, AQ_{s(b,y)}x, \|x\|s(b,y)\big), \qquad d_b + d_f + q(M,K) + K \le m, \quad M/m \longrightarrow 1.$$

Here b, f are coarse and exact blocks of dimensions d_b, d_f, and $y_j = \|x_j\|^2/\|x\|^2$ away from zero. Ramped prefixes Q_s and fixed A, conditioned on their admissible intersections, give stable recovery. The K tags describe the allocation and sum to $\|x\|$; both residual blocks vanish at zero, with no extra energy message. A slow stage schedule charges all coarse, exact and guard data before spectral limits. Prefix recovery preserves continuity with $q(M,K) = M + O(M/\log M)$ (Appendices B.3–B.4, where R counts all K branches).

For two tasks, Theorem 5.3 and uniform Weyl convergence suffice without purification. For $p \ge 1$, $\mathcal{R}_\infty$ is the orthant above the single-task optima; for $0 < p < 1$, it is rectangular exactly when principal energies are proportional. At equal single-task constants, nonproportionality raises the worst-task constant above task-conditioned coding; off balance, equality need not imply rectangularity (Appendix B.4). Native regularity gives $p < 1/5$.

4.3 Why the theorem holds at a fixed radius

Why the source radius need not shrink. A linearization at one fixed base leaves a relative error proportional to δ, which does not vanish as m grows. Instead, let $b = \Pi_J z$ be the transmitted coarse harmonic background and let $u, w \in H_J = H \cap \ker \Pi_J$, with $b + u, b + w \in B_\delta(H)$. Lemma A.8 supplies the uniform estimate

$$\|F_\tau(b+u) - F_\tau(b+w) - T_{b,\tau}(u-w)\|_2 \le \eta_J \|T_{b,\tau}(u-w)\|_2, \qquad \eta_J = C\delta(J+1)^{-\gamma} \to 0,$$

for fixed $0 < \gamma < r - s - 2 - k/2$ and sufficiently large J. The comparison is relative to prediction error, not merely an absolute small remainder. The tail's geometric displacement vanishes under E_r although its source norm need not, and δ remains fixed.

A common high-frequency cutoff absorbs the lower-order elliptic remainder and gives uniform tail coercivity. Cofactor and normalizer variation then control relative derivatives; integration along the source segment gives the estimate. No inverse on the whole source space or full-symbol compactness of its infinite-dimensional family is assumed.

One construction across backgrounds. Freeze J first. The smooth finite-dimensional background ball, together with the finite task family, is full-symbol compact after tail-Gram completion. The compact-family common decomposition therefore gives branch bases independent of b and τ, with a continuous background-dependent allocation. The charged code reconstructs $b + f + Q_s x$ inside the source ball, with the same b. Opposite tail inputs yield the constraints in (1); the common allocation attains them. Each decomposition, margin and selector is fixed before its spectral budget limit, then refined by one slow sequence. The onset is target-dependent, not uniform (Appendix C).

The output is a probability object, not a forecast vector. Filter the complete recovered support, including the smooth base, with one degree per budget, not per message. Uniform C^2 error $o(\delta m^{-\mu/k})$ preserves positive support and curvature. For the resulting polynomial H, Lemma A.4 gives a polynomial normal cofactor $\widehat{J}_H$, equal to J_H on the sphere. The inverse Gauss readout is $X_H(u) = H(u)u + (I - uu^\top)\nabla H(u)$. If $J_H \ge j_- > 0$ and $M_J \ge \sup_{[-1,1]^{k+1}} \|\nabla \widehat{J}_H\|$, then on the source cube

$$R_H(u) = \widehat{J}_H(u) + \frac{M_J^2}{2j_-}(\|u\|^2 - 1)^2 \ge j_-/2.$$

Radial projection gives $\widehat{J}_H(u) \ge j_- - M_J|\|u\|^2 - 1|$; completing the square proves positivity. The correction vanishes on the sphere. Structural normalization yields exactly $J_H\sigma/Z_H$, whose pushforward by X_H is uniform physical area. Polynomial moments give finite exact normalizers and prior forecasts (Lemma A.4), without a separately fitted density.

Clean recovery and the relative estimate precede filtering, whose support may leave the original affine source family. Rounding can also change b; it is handled by source-Lipschitz control, not the same-background estimate. The construction has global encoding/recovery Lipschitz bounds $Cm\log(e+m)$ and $Cm\log^2(e+m)$, and $\lceil(\mu/k+2)\log_2(m+1)\rceil$ bits per coordinate preserve the leading risks. This does not give bit-optimality or efficient decoding.

5 WHAT MUST THE REPRESENTATION MEASURE?

5.1 GEOMETRY-DEPENDENT ACQUISITION

A deterministic adaptive-linear policy chooses bounded linear functionals from earlier answers, without the realized task. Query rules can be discontinuous, so we assume no general inclusion between this class and continuous coding. Superscripts lin and alin denote fixed and adaptive linear measurements; $\Lambda_m^{I,\mathrm{type}} = \inf_{N\in I} \max_\tau \rho_{m,\tau}^{\mathrm{type}}(N)$.

Theorem 5.1 (Geometry-first acquisition)**.** *Under Theorem 4.1, every m-query deterministic adaptive-linear policy N has, for sufficiently large m, a two-batch replacement using $M = m + d_m$ unit-norm queries, $d_m = O(\log m/\log\log m)$, and one legal common CPPZ, such that $e_\tau(\widetilde{N}) \le (1+\epsilon_m)e_\tau(N) + o(\delta M^{-\mu/k})$, where $e_\tau = \sqrt{2\rho_\tau^{\mathrm{nl}}}$ and $\epsilon_m \to 0$. All coarse queries are charged. Synchronized $O(\log M)$-bit answer rounding preserves the comparison, with no separate branch or residual-energy message.*

Mechanism. Run N at an interior coarse catalog background c and freeze its at most m query functionals. For every unit u in their common tail kernel K_c, opposite inputs $c \pm \delta\sqrt{v_\delta(c)}u$ give identical answers. Their forecast separation yields $\delta\sqrt{v_\delta(c)}(1-\eta_J)\|T_{c,\tau}|_{K_c}\| \le e_\tau(N)$.

The second batch measures $H_J \ominus K_c$, of dimension at most m. Charged coarse data and inward routing from received words control background and residual energy; relative tails and common-law compilation finish the comparison. The frozen-kernel argument requires no continuity of the policy. Appendix D.2 matches normalized liminf and limsup separately; Theorem D.3 characterizes positive leading-risk regions for every finite task family. Neither result supplies an efficient optimizer.

Proposition 5.2 (Finite-budget geometry adaptation)**.** *There is a four-mode subfamily of the native spherical model with $(r,\beta,\tau_s,M_\beta) = (8,3,1,8)$ on $B_{300}(\mathbb{R}^4)$ in which two fixed queries followed by one geometry-dependent query, each rounded to 32 bits, recover one legal CPPZ with*

$$U_{\mathrm{two},32}(3) \le \tfrac{501}{500}\rho_0 < \tfrac{27}{25}\rho_0 \le \Lambda_3^{\mathrm{lin,nl}}, \qquad \rho_0 = \frac{49\cdot 300^2}{8\cdot 31^{11}}.$$

The lower bound allows arbitrary fixed linear measurements, exact answers and arbitrary nonlinear task decoders.

Two answers identify the expensive residual direction; the third measures it. Linearization has optimum ρ_0 in all three classes, while the nonlinear gap persists for one unweighted task. This is a geometry effect, not task incompatibility. Appendix D.3 gives the modes, weights and interval derivation. Proposition C.1 gives a separate asymptotic separation whose budget limit precedes its small-radius expansion.

5.2 THE TWO-TASK BOUNDARY AND COLLECTIVE REQUIREMENTS

To isolate task multiplicity from nonlinear geometry, we first return to linear tasks.

Theorem 5.3 (The exact two-task case)**.** *For compact linear $T_i : H \to Y_i$, $i = 1,2$, on real Hilbert spaces with separable H, every $m \ge 0$ and $\delta > 0$ satisfy, without ellipticity assumptions,*

$$\mathcal{R}_m^{\mathrm{cont}} = \mathcal{R}_m^{\mathrm{lin}} = \mathcal{R}_m^{\mathrm{alin}} = \left\{ r \ge 0 : \pi\cdot r \ge \tfrac{\delta^2}{2}\lambda_{m+1}(\pi_1 T_1^*T_1 + \pi_2 T_2^*T_2)\ \forall \pi \in \Delta_2 \right\}. \tag{4}$$

Every feasible target has a common orthogonal source projection of rank at most m.

For sufficiency, real matrix-pencil geometry gives a subspace of codimension at most m where both $T_i^*T_i - 2r_iI/\delta^2$ are nonpositive. Measuring its orthogonal complement meets both risks (Appendix D.4 handles degeneracies and infinite sources). The projection depends on the request; sharing need not be free. Combined with Theorem 5.1, this also identifies C_δ for two-batch common-law recovery of one or two native nonlinear tasks.

Corollary 5.4 (Pairwise sufficiency is not collective sufficiency)**.** *A smooth three-task family on S^2 at $p = 1$, with weights in $[1/4, 5/2]$ and identical complete single-task spectra, has free leading continuous sharing but an adaptive-linear/continuous worst-risk ratio with liminf greater than $6/5$. Every pair has exact linear sufficiency at every finite budget.*

This compares information classes for one family (Appendix D.6). Appendix E gives a separate native adaptive-linear/continuous gap; it cannot be chained with Proposition 5.2's fixed/adaptive gap for one instance.

Exact subproblems can still leave the group unresolved.

We now compare task families, rather than information classes within one family. For $q \geq 3$, index $N = 2^{q-1}$ coordinates by sign vectors of fixed parity, with diagonal Grams $A_i = \mathrm{diag}(1 + cs_i)$, $0 < c < 1$. Flipping an omitted sign permutes coordinates and matches every proper-subset problem between parity classes. At budget $N - 1$, the class containing the all-negative coordinate discards it; the other discards a direction equally supported on the singly-positive cells. Projection uppers and uniform-mixture lowers match at $(1-c)/2$ and $(1-c+2c/q)/2$. A request between them needs $N-1$ versus N coordinates, despite identical proper-subset regions at every budget (Proposition D.7; Figure 2). Triples amplify to exact budgets $3M$ versus $4M$, $M \geq 2$, with request-dependent finite precision (Proposition D.6). Odd harmonics and a radial homeomorphism realize pure CPPZs with one positive risk scale (Appendix D.6). Their shared finite-context Bernoulli detector is not native Bessel; their coordinates need not be linear source queries.

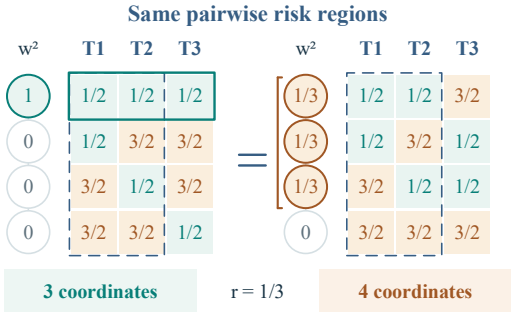


Figure 2: **Identical pairs, different joint budgets.** On $B_1(\mathbb{R}^4)$, $c = 1/2$. Cells: Gram energies; circles: squared discarded-direction weights. At $m = 3$, exact half-risks $1/4, 5/12$ make request $1/3$ require three/four coordinates.

6 Finite Risk Decisions and Shared Recovery

Finite-model studies demonstrate common-law recovery meeting shared requests, a collective minimum missed by pairwise sizing, and query savings from geometry adaptation. Figure 3 distinguishes floating-point uniform bounds, exact rational certificates and rational intervals. Sampled searches only exhibit violations; these studies do not establish the general asymptotic onset.

6.1 Common-Law Risk Requirements

A finite native model uses 32 real harmonic coordinates of degrees two through five, with $(r, \beta, \tau_s, M_\beta) = (8, 3, 1, 8)$, spherical base and opposing positive context weights. The request selects a projection before source access. Rounded coordinates reconstruct one source, polynomial support, prescribed law and inverse Gauss map. All 15 frozen requests at 1.02 times the tangent

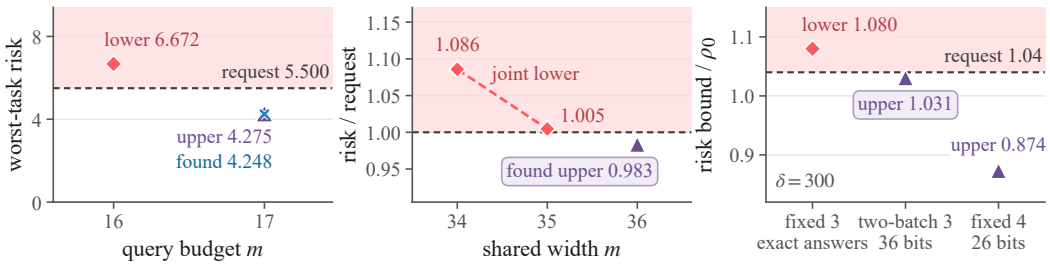


Figure 3: **Three finite budget decisions.** Left: 16 queries excluded; 17 suffice (floating-point bounds, 10^{-15} half-Brier units). Middle: pairs fit at 34; 35 jointly excluded; 36 suffice (rational certificates, fixed image-derived maps). Right: 3 queries, fixed vs. two-batch (rational intervals, native four-mode model). Bit counts (sufficient, not optimal) exclude public descriptions.

boundary, over budgets 8, 12 and 16 and five request ratios, meet nonlinear uniform upper bounds; the largest radial upper/information-lower ratio is about 1.00843 (Appendix F.1).

At 16 queries, opposing feasible requests have a jointly excluded coordinatewise minimum; 17 queries meet it (Figure 3). The same subspace retains this decision at six radii from 1 to 100, with radius-scaled quantization, recomputed compiler margins and refined bounds (Appendix F.1, Table 2). The bounds are floating-point evaluations, not interval certificates. Small absolute risks and global source queries do not establish deployment savings.

On 432 further nonspherical backgrounds, 1,618 requests pass the leading-order screen. Refinement gives sufficient numerical uniform bounds for 1,232; 325 remain unresolved (neither sufficient bounds nor violating inputs found); 61 have inputs violating the tested recovery, not every code. All outputs carry the prescribed law. Appendix F.8 separates the frozen study, finite-radius exclusions and additional-compute refinement; no equal-compute solver advantage is established.

6.2 A COLLECTIVE WIDTH WITH MATCHING CERTIFICATES

Sources are covariance ellipsoids from MNIST, Fashion-MNIST, KMNIST and gray CIFAR-10; tasks linearly reconstruct region-weighted images, not class labels. In one six-task MNIST family, every pair has a feasible 34-coordinate projection, every continuous 35-coordinate code is excluded, and a 36-coordinate projection meets all six requests. This determines the joint minimum (Figure 3).

For frozen dyadic maps $A_i = T_i^\top T_i$, exact integer arithmetic verifies a positive 36-dimensional restriction of a shifted mixture, excluding width 35 by min–max and the antipodal bound. Positivity of $2r_i I - A_i + \tau_i Q Q^\top$ certifies each proposed projection on $\ker Q^\top$. Optimization proposes the matrices; exact checks cover every pair and the joint upper, not covariance or hardware error (Appendix F.4).

All 192 prespecified task families have feasible projections. The frozen protocol determines 110 minimum widths; follow-up search closes one more. The other 81 have feasible codes but undetermined minima. Pairwise sizing is insufficient in 38 families. These frequencies describe the task generator, not classification. On a separate three-cell linear source, the exact linear optimum $11/1500$ exceeds an ideal continuous-code upper $103/16000$ at $m = 111$, including all tags (Appendix F.5); this is not nonlinear CPPZ recovery. Appendices F.2–F.5 retain the complete benchmarks, grouping and trained-code diagnostics, including numerical failures.

6.3 ACQUISITION AND FINITE PRECISION

For Proposition 5.2's family at radii 200–300, fixed-three queries are excluded while two-batch-three and fixed-four suffice. At radius 300, precisions $(13, 13, 10)$ and $(11, 11, 2, 2)$ give sufficient payloads of 36 and 26 bits: fewer queries need not mean fewer bits. The request lies between known bounds, not at an application tolerance. Appendix F.6 retains the full failed/unresolved precision grid. Separate spectral and catalog heuristics show no consistent gain over fixed references at tested budgets (Appendices F.8–F.9); they do not implement the asymptotic replacement construction.

7 DISCUSSION AND LIMITATIONS

For the native family, requiring one legal prescribed-law CPPZ does not reduce the leading risk region attainable by shared codes with separate task decoders. Joint attainability lifts prediction guarantees to a common probability object; measurement restrictions and collective task requirements remain distinct.

The characterization uses the stated common-order scalar elliptic assumptions; its nonlinear realization concerns finite declared tasks and smooth strictly convex geometries at fixed interior radii. Neither asymptotic onset nor geometric-boundary control is uniform. Coordinate and query budgets are separate from polynomial degree, description length and computation. Acquisition uses deterministic linear queries; answer rounding does not model sensor noise or hardware arithmetic. The finite studies establish scoped feasibility and separation, not a uniform finite-budget advantage. Efficient learned recovery and informative finite-budget regimes remain open.

AI USE STATEMENT

In this work, we used generative AI tools for theoretical development, mathematical claims and proofs, experiment design and implementation, synthetic data generation, data preparation, interpretation of results, and translation. Qualitative and thematic data analysis is not applicable to this work. Additionally, we used generative AI tools for code, figures, literature search, and drafting and editing the paper. We have reviewed all AI-assisted work and checked it for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work studies theoretical limits and constructions for shared probabilistic representations. The experiments use synthetic models and existing public image datasets.

REPRODUCIBILITY STATEMENT

Appendices A–E provide the assumptions, constructions and proofs supporting the theoretical results. Appendix F describes data processing, experimental configurations, evaluation protocols, resource budgets and uncertainty estimates. It distinguishes exact and interval certificates from floating-point bounds and sampled errors. The code will be made publicly available upon acceptance.

REFERENCES

M. Althoff, O. Stursberg, and M. Buss. Safety assessment for stochastic linear systems using enclosing hulls of probability density functions. In *European Control Conference (ECC)*, pp. 625–630, 2009.

A. Athalye, N. Carlini, and D. Wagner. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples. In *International Conference on Machine Learning (ICML)*, volume 80 of *Proceedings of Machine Learning Research*, pp. 274–283, 2018.

M. Sh. Birman and M. Z. Solomyak. Asymptotics of the spectrum of variational problems on solutions of elliptic equations. *Siberian Mathematical Journal*, 20(1):1–15, 1979.

A. L. Caterini, G. Loaiza-Ganem, G. Pleiss, and J. P. Cunningham. Rectangular flows for manifold learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pp. 30228–30241, 2021.

S.-Y. Cheng and S.-T. Yau. On the regularity of the solution of the n-dimensional Minkowski problem. *Communications on Pure and Applied Mathematics*, 29(5):495–516, 1976.

T. Clanuwat, M. Bober-Irizar, A. Kitamoto, A. Lamb, K. Yamamoto, and D. Ha. Deep learning for classical Japanese literature. arXiv preprint arXiv:1812.01718, 2018.

A. Cohen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Optimal stable nonlinear approximation. *Foundations of Computational Mathematics*, 22(3):607–648, 2022.

F. Croce and M. Hein. Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks. In *International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pp. 2206–2216, 2020.

R. A. DeVore, R. Howard, and C. Micchelli. Optimal nonlinear approximation. *Manuscripta Mathematica*, 63(4):469–478, 1989.

P. Diaconis, S. Holmes, and M. Shahshahani. Sampling from a manifold. In G. Jones and X. Shen (eds.), *Advances in Modern Statistical Theory and Applications: A Festschrift in Honor of Morris L. Eaton*, volume 10 of *Institute of Mathematical Statistics Collections*, pp. 102–125. Institute of Mathematical Statistics, 2013.

S. Dirksen and T. Ullrich. Gelfand numbers related to structured sparsity and Besov space embeddings with small mixed smoothness. *Journal of Complexity*, 48:69–102, 2018.

S. Dyatlov. *Lecture notes for 18.155: Distributions, elliptic regularity, and applications to PDEs*. MIT, 2026. URL `https://math.mit.edu/%7Edyatlov/18.155/155-notes.pdf`. Version of September 16, 2026.

C. Fifty, E. Amid, Z. Zhao, T. Yu, R. Anil, and C. Finn. Efficiently identifying task groupings for multi-task learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pp. 27503–27516, 2021.

R. J. Gardner, L. Parapatits, and F. E. Schuster. A characterization of Blaschke addition. *Advances in Mathematics*, 254:396–418, 2014.

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

S. Guo, J. Chen, Z. Hu, Z. Chen, W. Yang, Y. Lin, X. Jiang, and L.-Y. Duan. Which tasks should be compressed together? A causal discovery approach for efficient multi-task representation compression. In *International Conference on Learning Representations (ICLR)*, pp. 15264–15287, 2025.

Y. Hu, R. Xian, Q. Wu, Q. Fan, L. Yin, and H. Zhao. Revisiting scalarization in multi-task learning: A theoretical perspective. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, pp. 48510–48533, 2023.

C. R. Johnson and L. Rodman. Convex sets of Hermitian matrices with constant inertia. *SIAM Journal on Algebraic and Discrete Methods*, 6(3):351–359, 1985.

N. Kochdumper and M. Althoff. Constrained polynomial zonotopes. *Acta Informatica*, 60(3):279–316, 2023.

A. V. Kolesnikov and E. Milman. Local L^p-Brunn–Minkowski inequalities for $p < 1$. *Memoirs of the American Mathematical Society*, 277(1360):1–78, 2022.

D. Krieg, E. Novak, and M. Ullrich. On the power of adaption and randomization. *Forum of Mathematics, Sigma*, 13:e152, 2025.

A. Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009.

A. Laptev and Yu. Safarov. Szegö type limit theorems. *Journal of Functional Analysis*, 138(2): 544–559, 1996.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

C.-K. Li and N.-S. Sze. Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations. *Proceedings of the American Mathematical Society*, 136(9):3013–3023, 2008.

A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations (ICLR)*, 2018.

J. Matoušek. *Using the Borsuk–Ulam Theorem: Lectures on Topological Methods in Combinatorics and Geometry*. Universitext. Springer, 2003.

E. Milman. Centro-affine differential geometry and the log-Minkowski problem. *Journal of the European Mathematical Society*, 27(2):709–772, 2025.

A. Modares, Z. Zhang, T. Charalambous, A. Alanwar, and H. Modares. Reachability analysis with probabilistic zonotopes: Learning realized disturbances and refining aleatory uncertainty. arXiv preprint arXiv:2607.01352, 2026.

V. K. Nguyen and V. D. Nguyen. Best n-term approximation of diagonal operators and application to function spaces with mixed smoothness. *Analysis Mathematica*, 48(4):1127–1152, 2022.

R. Ponge. Connes' integration and Weyl's laws. *Journal of Noncommutative Geometry*, 17(2): 719–767, 2023.

A. Radul and B. Alexeev. The base measure problem and its solution. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 130 of *Proceedings of Machine Learning Research*, pp. 3583–3591, 2021.

M. D. Reid and R. C. Williamson. Information, divergence and risk for binary experiments. *Journal of Machine Learning Research*, 12(22):731–817, 2011.

L. Rodman. Isotropic subspaces for pairs of Hermitian matrices. *Electronic Journal of Linear Algebra*, 17:508–517, 2008.

S. Samadi, U. Tantipongpipat, J. Morgenstern, M. Singh, and S. Vempala. The price of fair PCA: One extra dimension. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 31, 2018.

R. Schneider. *Convex Bodies: The Brunn–Minkowski Theory*, volume 151 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, second edition, 2013.

E. Schneidman, S. Still, M. J. Berry II, and W. Bialek. Network information and connected correlations. *Physical Review Letters*, 91(23):238701, 2003.

R. T. Seeley. Complex powers of an elliptic operator. In A. P. Calderón (ed.), *Singular Integrals*, volume 10 of *Proceedings of Symposia in Pure Mathematics*, pp. 288–307. American Mathematical Society, 1967.

A. Spantini, T. Cui, K. Willcox, L. Tenorio, and Y. Marzouk. Goal-oriented optimal approximations of Bayesian linear inverse problems. *SIAM Journal on Scientific Computing*, 39(5):S167–S196, 2017.

T. Standley, A. Zamir, D. Chen, L. Guibas, J. Malik, and S. Savarese. Which tasks should be learned together in multi-task learning? In *International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pp. 9120–9132, 2020.

U. Tantipongpipat, S. Samadi, M. Singh, J. Morgenstern, and S. Vempala. Multi-criteria dimensionality reduction with applications to fairness. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32, 2019.

M. E. Taylor. *Pseudodifferential Operators and Nonlinear PDE*, volume 100 of *Progress in Mathematics*. Birkhäuser, 1991.

R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. In Y. C. Eldar and G. Kutyniok (eds.), *Compressed Sensing: Theory and Applications*, pp. 210–268. Cambridge University Press, 2012.

H. Wang and I. H. Sloan. On filtered polynomial approximation on the sphere. *Journal of Fourier Analysis and Applications*, 23(4):863–876, 2017.

H. Xiao, K. Rasul, and R. Vollgraf. Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. arXiv preprint arXiv:1708.07747, 2017.

G. Zhang, J. Qian, J. Chen, and A. Khisti. Universal rate-distortion-perception representations for lossy compression. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pp. 11517–11529, 2021.

Z. Zhang and A. Alanwar. Hybrid probabilistic zonotopes for identifiable and refinable predictive uncertainty. arXiv preprint arXiv:2608.05454, 2026.

A CPPZ REPRESENTATION AND THE NATIVE PROBABILITY COMPILER

A.1 SOURCE REGISTRIES, ADMISSIBILITY AND OBSERVATIONS

Definition A.1 (Constrained Polynomial Probabilistic Zonotope)**.** For a finite set of named strata $\mathcal{S}$, a CPPZ is the tuple $\mathcal{C} = (\mathcal{S}, \{(\mathfrak{R}_s, \Pi_s, F_s)\}_{s\in\mathcal{S}}, \omega, \{K_t\}_{t=1}^T, \mathsf{v})$. The registry $\mathfrak{R}_s$ declares binary atoms $\beta \in \mathcal{B}_s = \{-1,+1\}^{n_b(s)}$, continuous sources in $\Omega_s = [-1,1]^{n_d(s)} \times \mathbb{R}^{n_g(s)}$, polynomial constraints $Q_{s\beta}$, a metric g_s, pure dimension k_s, positive branch weights $\mathsf{p}_{s\beta}$ and positive continuous ambient-normalized density versions $\rho_{s\beta}$. The readout $F_s(\beta, z)$ is polynomial. Define $M_{s\beta} = \{z \in \Omega_s : Q_{s\beta}(z) = 0\}$ and $d\mu_{s\beta} = \rho_{s\beta}\, d\mathcal{H}^{k_s}_{g_s}|_{M_{s\beta}}$, and

$$\Pi_s = Z_s^{-1} \sum_{\beta\in\mathcal{B}_s} \mathsf{p}_{s\beta}\delta_\beta \otimes \mu_{s\beta}, \qquad Z_s = \sum_\beta \mathsf{p}_{s\beta}\mu_{s\beta}(M_{s\beta}), \tag{5}$$

where $\sum_\beta \mathsf{p}_{s\beta} = 1$, $\sum_s \omega_s = 1$ and $\omega_s > 0$. For likelihood evidence, nonnegative typed K_t give

$$\Pi^0 = \sum_s \omega_s\delta_s \otimes \Pi_s, \qquad P_{\mathcal{C}}^{y_{1:T}} = \widetilde{F}_{\#}\left(\frac{\prod_{t=1}^T K_t(y_t \mid s,\beta,z)}{Z^{\mathrm{ev}}_{1:T}}\,\Pi^0\right), \tag{6}$$

where $Z^{\mathrm{ev}}_{1:t} = \int \prod_{j=1}^t K_j\, d\Pi^0$ and $\widetilde{F}(s,\beta,z) = (s, F_s(\beta,z))$; tags may be forgotten in a common output space. The declaration v distinguishes structural constraints, observational/coarea versions and coupling evidence. Admissibility requires regular pure fibers, boundary-nullness, finite required moments, and finite positive structural and evidence-prefix normalizers. These conditions, polynomial parameters and exact-observation semantics are specified below.

Each stratum has the named-source registry $\mathfrak{R}_s = (\mathcal{B}_s, \Omega_s, g_s, k_s, \{(Q_{s\beta}, \mathsf{p}_{s\beta}, \rho_{s\beta})\}_{\beta\in\mathcal{B}_s}, \mathsf{ord}_s, \prec_s)$; its two orders fix the source names and monomials. Binary sources are atoms, not continuous variables conditioned on $\beta_i^2 = 1$. For $z = (\alpha, \nu)$, Euclidean coordinates allow but do not require Gaussian laws. The finite polynomial data are $Q_{s\beta}(z) = \sum_{j=1}^{J_{s\beta}} A_{s\beta j} z^{R_{s\beta j}} - b_{s\beta}$ and

$$F_s(\beta, z) = c_s + \sum_{\ell=1}^{L_s} G_{s\ell} \prod_{i=1}^{n_b(s)} \beta_i^{B_{si\ell}} \prod_{j=1}^{n_d(s)+n_g(s)} z_j^{E_{sj\ell}},$$

with nonnegative integer exponents and coefficients of the declared dimensions. One bounded-source branch recovers CPZ syntax (Kochdumper & Althoff, 2023, Definition 5), subject to the following geometric and probabilistic conditions. Regularity concerns fibers, not duplicate monomial columns.

Fibers and density versions. Each nonempty interior fiber is a properly embedded C^2 k_s-manifold with closure $M_{s\beta}$ and possibly piecewise-C^2 clipping boundary; every component meets the interior, has full Hausdorff-reference support, and satisfies $\mathcal{H}^{k_s}_{g_s}(M_{s\beta} \cap \partial\Omega_s) = 0$. The metric is smooth and positive definite. The density $\rho_{s\beta}$ is a specified strictly positive continuous version, including boundary values, normalized against ambient metric volume on Ω_s°; its restriction to a lower-dimensional fiber depends on this version, not only on its ambient almost-everywhere class. Required moments are finite and $0 < Z_s < \infty$. Empty branches have zero mass; every stratum has a positive-mass branch. Different dimensions use separate admissible strata; arbitrary singular unions are not included.

Prior and source identity. The prior is defined on the tagged disjoint union, normalizing each stratum before mixing. Its stratum probabilities are ω_s, whereas $\Pi_s(\{\beta\} \times M_{s\beta}) = \mathsf{p}_{s\beta}\mu_{s\beta}(M_{s\beta})/Z_s$ need not equal the pre-constraint weight $\mathsf{p}_{s\beta}$. Tags distinguish overlapping physical images until explicitly forgotten: an atom and a circle's arc-length law, for example, receive their relative masses from ω, not from a comparison of volumes in different dimensions. A shared source name always denotes the same random variable, including under coupling evidence.

Likelihood evidence. Each measurable likelihood K_t uses the observation type and reference declared by v. Equation (6), applied to every admissible prefix, updates branch and stratum probabilities; the empty product and its normalizer equal one. Coupling likelihoods use the declared joint sources.

Exact observations. For a regular C^1 map $O : M \to \mathbb{R}^q$, $q \le k$, on a pure k-carrier, the declared coarea version is $d\Pi_y = Z_y^{-1} J_{O|M}^{-1} \rho \; d\mathcal{H}_g^{k-q}|_{M \cap O^{-1}(y)}$ with $J_{O|M} = \sqrt{\det(DO|_{TM}(DO|_{TM})^*)}$. The adjoint uses the source metric and Euclidean observation metric; full rank and finite positive fiber mass are required (Diaconis et al., 2013, Proposition 2). Coarea determines conditionals almost everywhere; a specified null value additionally requires a chosen version, for example a weakly continuous extension under regular proper-submersion hypotheses. Conditional branch weights use marginal densities against a common observation reference. Structural equality instead uses the declared Hausdorff reference without the inverse Jacobian; it is not conditioning by a null indicator.

A.2 Composition and representation boundaries

Replacing polynomial equations by others defining the same fiber preserves (5) when the metric, density and weights are unchanged. A carrier diffeomorphism ϕ likewise preserves output and evidence laws under metric transport $(\phi^{-1})^* g$ and composition of the density, readout and likelihoods with ϕ^{-1}; it defines a CPPZ transformation only if the resulting data satisfy the polynomial grammar and admissibility conditions. The calculus below uses product measures, polynomial composition and Bayesian updating (Kochdumper & Althoff, 2023; Radul & Alexeev, 2021).

Proposition A.2 (Compositional CPPZ laws)**.** *Admissible CPPZs are closed under tagged finite mixtures, independent products on disjoint source registries, and polynomial output maps, provided all newly required moments are finite. Appending measurable nonnegative likelihoods, including coupling likelihoods on joint sources, preserves the representation if every evidence-prefix normalizer is positive and finite and required posterior moments are finite. For a probability law ν, $U_K\nu = K\nu/\nu(K)$ satisfies $U_L(U_K\nu) = U_{KL}\nu$ and $\nu(KL) = \nu(K)(U_K\nu)(L)$. Polynomial structural refinements remain CPPZs if their declared pure fibers satisfy Section A.1's admissibility conditions.*

Proof. For a mixture, use disjoint stratum tags with weights $a\omega_s$ and $(1-a)\omega'_t$; for conditioned objects, attach each aggregate likelihood divided by its own evidence normalizer, preserving $a, 1-a$. For a product, pair strata and branches and concatenate constraints. The product metric and density give fiber dimension $k_s + k'_t$ and, by Tonelli, $Z_{s,t} = Z_s Z'_t$, $\Pi_{s,t} = \Pi_s \otimes \Pi'_t$ and weight $\omega_s\omega'_t$. Regularity, properness and boundary-nullness pass to the product; independent evidence and its normalizers multiply. Disjoint registries declare independence rather than infer it. A polynomial G changes only the readout to $G \circ F$, including marginals and sums; observables of one source instead use $(F_1(z), F_2(z))$ to preserve dependence. Substitution proves the two update identities. Vanishing likelihoods remain evidence factors, not absorbed into a strictly positive structural density. Additional equations define a new structural fiber with its declared Hausdorff reference; admissibility supplies its dimension, support and normalizer. □

Proposition A.3 (An exact constraint-versus-generator boundary)**.** *Let $M = \{x \in \mathbb{R}^n : g_1(x) = \cdots = g_s(x) = 0\}$ be compact. Suppose a Euclidean source Z assigns zero mass to every zero set of a nonzero polynomial; every source with a Lebesgue density has this property. For any polynomial map $P : \mathbb{R}^m \to \mathbb{R}^n$, either P is constant with value in M, or $\mathbb{P}(P(Z) \in M) = 0$. Thus a finite mixture of such generators cannot represent a nonatomic probability law supported on M.*

Proof. Positive mass on M forces each $g_j \circ P$ to vanish identically, so $P(\mathbb{R}^m) \subseteq M$; compactness makes every coordinate polynomial bounded and therefore constant. The zero-set property for Lebesgue densities follows by induction and Fubini: outside the null common zero set of its coefficient polynomials, a nonzero polynomial has finitely many roots in its last variable. Applied branchwise, the dichotomy shows that any mass on M is concentrated at finitely many points. □

Affine inclusion and its limits. An affine readout $X = c + G_b\beta + G_d\alpha + G_s\nu$, with affine factor constraints, embeds as a degree-one CPPZ once the joint latent distribution and reference and conditioning versions are fixed and matched (Althoff et al., 2009; Modares et al., 2026; Zhang & Alanwar, 2026). Fibers and readouts agree, and the constant positive coarea Jacobian cancels. For example, Modares et al.'s $\{\alpha + \nu : \alpha \in [-1, 1], \nu \sim N(0, 1)\}$ describes uncertain-mean laws until a distribution for α is chosen. Conversely, degree one does not enforce a Gaussian source: identity readout with latent density $\frac{1}{2}e^{-|x|}$ gives a Laplace law.

Strictness has two distinct witnesses. In an affine constrained-source model, fixing the discrete branch leaves affine equalities in a convex source domain, so with a full-support law on this carrier an affine readout has convex output support (the closure of its convex image). The polynomial CPPZ (U, U^2), $U \sim \mathrm{Unif}[-1, 1]$, lies outside finite mixtures of these branches: each positive-mass branch of a mixture supported on the parabolic arc must itself be supported there, and convexity makes it a singleton, since the midpoint of distinct parabola points lies above the arc by $(u - v)^2/4$; finitely many singletons cannot carry the nonatomic law. The conclusion requires convex latent supports, not merely affine readouts with arbitrary source laws. Proposition A.3 does *not* apply to this clipped noncompact variety, which itself has a polynomial generator. It does apply to S^k: sphere constraint, identity readout and constant ambient density on $[-1, 1]^{k+1}$ define its uniform law as a regular pure CPPZ, which no finite mixture of the proposition's unconstrained polynomial generators can realize. This is an exact boundary, not a finite-precision lower bound; it permits approximation, rational or trigonometric maps, normalization layers and manifold-valued sources.

Prescribed intrinsic laws. For a compact regular carrier, a polynomial embedding $X : M \to N$ with $J_X > 0$, and continuous $\psi : N \to \mathbb{R}$, assume $e^{\psi \circ X} J_X$ has a positive continuous integrable ambient extension with the required moments. By the area formula, applied to a test function times e^{ψ} and to the normalizer, $X_{\#}$ maps the law $\propto e^{\psi \circ X} J_X \, d\mathcal{H}^k_g$ on M to the law $\propto e^{\psi} \, d\mathcal{H}^k|_N$. The native compiler constructs this extension for its area law. General scores, products and mixed strata remain representation capabilities; sharp capacity and acquisition guarantees concern only the stated native pure family.

A.3 The native task norm

The identity $K_\beta = \mathbb{E} e^{-\tau_s T \Delta}$, $T \sim \mathrm{Gamma}(\beta/2, 1)$, shows that the kernel k_β is positive and Markov; $\beta > k$ makes it bounded and continuous. For the Bessel detector of Section 3 with support function h and context weight ω, completing the conditional square proves $F_h = M_\beta^{-1} K_\beta p_h$ and $\mathrm{Reg}_h(a) = \frac{1}{2} \|\sqrt{\omega}(a - F_h)\|_2^2$; any positive bounded Markov kernel may replace k_β. Positive upper/lower weight bounds preserve any unweighted secant comparison, and clipping any forecast to $[0, 1]$ cannot increase risk. This is the fixed contextual task, not a loss chosen after the result (Gneiting & Raftery, 2007; Reid & Williamson, 2011).

A.4 Local comparison for perturbative consequences

At a sufficiently small source radius, the single-base comparison is

$$\|F(z) - F(\widetilde{z}) - T_0(z - \widetilde{z})\|_2 \le C_{\mathrm{sec}} \delta \, \|T_0(z - \widetilde{z})\|_2 \,. \tag{7}$$

With the notation and assumptions of Section 3, impose at Sobolev order t the transverse slice $X_0^t = \{\phi \in H^t : \langle \phi, 1 \rangle = \langle \phi, u_i \rangle = 0\}$, with distributional pairings at negative orders.

Translations satisfy $\nabla^2(a \cdot u) + (a \cdot u) g = 0$, while $L_0 h_0 = k J_0$. With $\widetilde{L}_{h_0} f = k^{-1} A_0^{-1} : (\nabla^2(h_0 f) + h_0 f g)$, we have $J_0^{-1} L_0 (h_0 f) = k \widetilde{L}_{h_0} f$, and the linearized Minkowski kernel theorem identifies the eigenspace $-(\widetilde{L}_{h_0} - I) f = f$ as $(a \cdot u)/h_0$ (Cheng & Yau, 1976; Kolesnikov & Milman, 2022; Milman, 2025). Hence $\ker L_0 = \mathrm{span}\{u_i\}$ and $\ker(Q_0 L_0) = \mathrm{span}\{h_0, u_i\}$. The mean removes h_0 because $\int h_0 > 0$, and first moments remove translations. A parametrix estimate and Rellich compactness therefore imply

$$\|\phi\|_{H^{2-\beta}} \le C_0 \|Q_0 L_0 \phi\|_{H^{-\beta}} \quad (\phi \in X_0^{2-\beta}). \tag{8}$$

Indeed a contrary unit sequence has a convergent subsequence in $H^{1-\beta}$; the parametrix estimate upgrades this to $H^{2-\beta}$ and leaves a nonzero transverse kernel element. No surjectivity is used.

For $\phi = E_r(z - \widetilde{z})$, determinant integration gives $J_{h_z} - J_{h_{\widetilde{z}}} = L_{z,\widetilde{z}} \phi$, with the cofactor averaged along the exact support segment, and normalization gives the exact identity $p_{h_z} - p_{h_{\widetilde{z}}} = Q_{z,\widetilde{z}} L_{z,\widetilde{z}} \phi$, where $Q_{z,\widetilde{z}} f = Z_z^{-1}(f - p_{h_{\widetilde{z}}} \langle f, 1 \rangle)$. Since $E_r : L_2 \to W^{s+2,\infty}$, the cofactor, scalar normalizer and density coefficient vary by $O(\delta)$ in the required $W^{s,\infty}$ norms, so the multiplier condition $s > \beta$ gives $\|L_{z,\widetilde{z}} - L_0\|_{H^{2-\beta} \to H^{-\beta}} + \|Q_{z,\widetilde{z}} - Q_0\|_{H^{-\beta} \to H^{-\beta}} \le C\delta$. This includes the rank-one term, since pairing with 1 is bounded on $H^{-\beta}$. By (8) this is a relative perturbation of $Q_0 L_0 \phi$, and $\|K_\beta f\|_2 \asymp \|f\|_{H^{-\beta}}$ proves (7).

Complex powers (Seeley, 1967) make T_0 elliptic of order $-\mu$ modulo finite rank. Its principal amplitude $t_{-\mu}$ is homogeneous of degree $-\mu$ and equals the amplitude t_b of Appendix C at $b = 0$ on the metric unit cosphere. Negative-order Weyl theory (Birman & Solomyak, 1979; Ponge, 2023) gives $s_j(T_0) \sim C_W j^{-\mu/k}$ with $C_W = [k^{-1}(2\pi)^{-k} \int_{S^*M} |t_{-\mu}|^{k/\mu}\, d\Lambda]^{\mu/k}$; bounded positive weights preserve the order $j^{-\mu/k}$.

A.5 Finite probability compilation and its reusable interface

Lemma A.4 (One finite polynomial probability compiler)**.** *Let $h = H|_{S^k} > 0$ be a uniformly convex polynomial support with $d = \deg H$ and $J_h \geq j_- > 0$. Define the polynomial readout, normal cofactor and density by*

$$P_u = I - uu^\top, \qquad X_H(u) = H(u)u + P_u \nabla H(u),$$
$$B_H(u) = \nabla^2 H(u) + (H(u) - u^\top \nabla H(u))I, \qquad \widehat{J}_H(u) = u^\top \operatorname{adj}(B_H(u))u,$$
$$R_H(u) = \widehat{J}_H(u) + \frac{M_J^2}{2j_-}(\|u\|^2 - 1)^2, \qquad \rho_H = \frac{R_H}{\int_{[-1,1]^{k+1}} R_H(u)\, du},$$

where $M_J \geq \sup_{[-1,1]^{k+1}} \|\nabla \widehat{J}_H\|$. On the regular sphere carrier in $[-1,1]^{k+1}$, readout X_H and density ρ_H form a legal pure CPPZ with structural law $J_h\sigma/Z_h$, density degree at most $\max\{kd+2, 4\}$, and finite exact normalizers and prior Bessel forecast.

Proof. The tangent block of B_H is A_h, so its normal cofactor equals J_h on the sphere without ambient invertibility, and its minors give $\deg \widehat{J}_H \leq kd + 2$. Radial projection to the sphere obeys $\|u - u/\|u\|\| \leq t = |\|u\|^2 - 1|$; at zero use any unit point. Hence $R_H \geq j_- - M_J t + M_J^2 t^2/(2j_-) \geq j_-/2$, and the added term vanishes on the carrier. Odd moments vanish, even cube moments $\int_{[-1,1]^{k+1}} u^{2a}\, du$ are $\prod_i 2/(2a_i+1)$, and normalized spherical moments are $\prod_i (1/2)_{a_i}/((k+1)/2)_{|a|}$ by Gaussian radial decomposition. Ambient normalization cancels in the structural law.

The sphere equation is regular and its cube-boundary intersection is a finite null set. The readout is the inverse Gauss map with area Jacobian J_h, so $X_{H\#}(J_h\sigma/Z_h)$ is normalized physical area. The declared Bernoulli likelihood $(k_\beta(v,u)/M_\beta)^y(1 - k_\beta(v,u)/M_\beta)^{1-y}$ has positive finite prefix masses. Write $J_h = \sum_{\ell \leq kd+2, j} c_{\ell j} Y_{\ell j}$ with $Y_{00} = 1$; then $Z_h = c_{00} > 0$ and

$$F_h(v) = \frac{1}{M_\beta c_{00}} \sum_{\ell \leq kd+2} (1 + \tau_s \ell(\ell + k - 1))^{-\beta/2} \sum_j c_{\ell j} Y_{\ell j}(v).$$

The coefficients are finite sphere-moment sums. This finite formula is the prior forecast used here, and every later reconstruction invokes the same object. □

Lemma A.5 (Complete support interface)**.** *Suppose a Borel source decoder $\widehat{z}_m$ takes values in a fixed positive uniformly convex source ball $B_{C_{\mathrm{src}}\delta}(H)$, with C_{src} independent of m. There is one code-independent filtered degree $2L(m)$ such that $\widehat{h}_{m,c} = V_L h_0 + E_r V_L \widehat{z}_m(c)$ satisfies $\sup_c \|\widehat{h}_{m,c} - h_{\widehat{z}_m(c)}\|_{C^2} = o(\delta m^{-\mu/k})$. The compiled probability is legal, with fixed smaller geometric margins. It is Borel in the code and continuous when the source decoder is. With a clean forecast residual of $O(\delta m^{-\mu/k})$, filtering changes the risk by $o(\delta^2 m^{-1/p})$.*

Proof. For $2 + k/2 < t < r$, Sobolev embedding gives $\|(I - V_L)E_r u\|_{C^2} \leq C_t L^{-(r-t)}\|u\|$. The filter has multipliers in $[0,1]$, equals one through L, vanishes above $2L$, preserves the gauge and commutes with E_r (Wang & Sloan, 2017). Filter the smooth base as well: for any $t_0 > q > 2 + k/2$, its C^2 tail is at most $CL^{-(t_0-q)}\|h_0\|_{H^{t_0}}$. Choose L so both errors are lower order. Determinant secants and positive normalizers give, on fixed C^2-bounded convex families,

$$\|p_h - p_{\widetilde{h}}\|_\infty + \|M_{\sqrt{\omega}}(F_h - F_{\widetilde{h}})\|_2 \leq C\|h - \widetilde{h}\|_{C^2}. \tag{9}$$

Small errors preserve margins, and Lemma A.4 gives the exact filtered-support law. Equation (9) and $E_r : H \to C^2$ also give a source-to-forecast Lipschitz constant C_F. Thus a clean root error e_τ, a

source perturbation of at most d within the same ball, and filter forecast error ϵ_{pol} give compiled risk at most $\frac{1}{2}(e_\tau + C_F d + \epsilon_{\mathrm{pol}})^2$. For $e_\tau = O(\delta m^{-\mu/k})$ and $d, \epsilon_{\mathrm{pol}} = o(\delta m^{-\mu/k})$, both the perturbation square and cross term are lower order. No compact code image is required. Apply secants to clean sources *before* filtering: the filtered support need not belong to the original affine source ball. □

A.6 Conditional recovery without an output-degree cap

Fixing a received message separates three costs: choosing a common forecast, restricting its legal output class, and fitting one particular output. The following consequence of classical area-measure geometry identifies a case where the second cost vanishes when decoding a fixed message; it does not construct the shared encoder of Theorem 4.1. Let $\mathcal{L}_{\mathrm{poly}}$ contain all positive strictly convex polynomial supports on S^k, $k \geq 2$, compiled by Lemma A.4. There is no common degree cap or restriction to the original affine source dictionary.

Proposition A.6 (Conditional law closure)**.** *Let $\mathcal{C}$ be a nonempty C^2-compact set of smooth positive strictly convex supports. For one fixed smooth positive task weight ω,*

$$\inf_{H \in \mathcal{L}_{\mathrm{poly}}} \sup_{h \in \mathcal{C}} \tfrac{1}{2}\|F_h - F_H\|_\omega^2 = \inf_{a \in L^2(\omega\sigma)} \sup_{h \in \mathcal{C}} \tfrac{1}{2}\|F_h - a\|_\omega^2. \tag{10}$$

Consequently, for a fixed finite-alphabet encoder whose nonempty cell closures satisfy these hypotheses, the infimum over legal decoders equals the unrestricted one-task risk without changing any transmitted word. This is equality of infima, not attainment at a prescribed degree.

Proof. First consider a finite mixture $\bar{p} = \sum_i \alpha_i p_{h_i}$. The divergence theorem gives $\int u p_{h_i}(u)\, d\sigma = 0$, so $\bar{p}$ is smooth, strictly positive, normalized and has zero first moment. The smooth Minkowski theorem, with prescribed curvature $1/\bar{p}$, gives a smooth strictly convex support h satisfying $J_h = \bar{p}$ (Cheng & Yau, 1976, Theorem 1); this is the normalized area-measure form of Blaschke addition (Gardner et al., 2014, Section 2). Translating the body to contain the origin in its interior makes $h > 0$ without changing J_h. Spherical polynomial approximants $H_n \to h$ in C^2 preserve positive support and curvature for large n. Determinant continuity, a positive normalizer and boundedness of K_β imply $F_{H_n} \to \sum_i \alpha_i F_{h_i}$ in every fixed finite collection of task norms. Lemma A.4 makes each approximant one legal object; no independently fitted density is used.

The compact forecast set $K = \{F_h : h \in \mathcal{C}\}$ has a minimum-radius center in the task Hilbert space: for $r(a) = \sup_{f \in K} \|f - a\|_\omega$, the parallelogram identity gives $r((a+b)/2)^2 + \|a - b\|_\omega^2/4 \leq [r(a)^2 + r(b)^2]/2$, so a minimizing sequence is Cauchy and its limit is a center. Projection onto $\overline{\mathrm{conv}}K$ increases no distance to K, so a center is a limit of finite convex combinations of forecasts. The first step approximates each combination by a legal polynomial forecast; since $|r(a) - r(b)| \leq \|a - b\|_\omega$, their radii tend to the unrestricted optimum. The reverse inequality is immediate. For finitely many words, one approximant per cell at a common tolerance gives finite maximum degree at that tolerance. □

Common centers and task contrast. For finitely many tasks, the same argument applies if their optimal centers coincide in the closed convex hull of K. In particular, if $Q \geq 0$ on $[-R, R]$ in (58) for every task, the exact fiber's endpoint midpoint f_* is such a center. General rounded cells need not have this property: $\max_\tau \inf_{a_\tau}$ cannot be interchanged with $\inf_a \max_\tau$. More generally, if $a\omega_0 \leq \omega_\tau \leq b\omega_0$ with $0 < a \leq b$, the common legal infimum and the taskwise free infimum satisfy

$$R^*_{\mathrm{free}} \leq R^*_{\mathcal{L}_{\mathrm{poly}}} \leq (b/a) R^*_{\mathrm{free}}. \tag{11}$$

Indeed, approximating the ω_0-optimal center by the proposition gives risk at most $bR_0^* + o(1)$ in every task, where R_0^* is the free optimum for ω_0, whereas every free task optimum is at least aR_0^*. These statements control neither decoder storage, degree as tolerance decreases, numerical stability nor computation, and do not replace continuous shared recovery with a sharp leading constant.

Law mixing is not support averaging. On S^2, with axial coordinate $t = u_3$, put $Y(t) = \sqrt{5}(3t^2 - 1)/2$ and $h_\pm = 1 \pm Y/20$. The two support-Hessian eigenvalues are positive, bounded below by $1 - \sqrt{5}/8$. For an axisymmetric support, $J_h = [(1-t^2)h'' - th' + h][-th' + h]$. Substitution gives both normalizers $199/200$ and $(p_{h_+} + p_{h_-})/2 - p_{(h_+ + h_-)/2} = (135t^4 - 30t^2 - 17)/1592 \neq 0$, so the mixture requires a different geometric reconstruction, which averaging the two supports does not implement.

A.7 THE GEOMETRIC INTERIOR AND UNIFORM RELATIVE TAILS

Keep Section 3's source norm, gauge, smooth positive base, detector and finite task family. These estimates replace base-point smallness in the capacity and acquisition proofs; they do not extend the small-radius separation expansions. A radius is fixed before the budget tends to infinity. For $\lambda_\ell = \ell(\ell+k-1)$ and multiplicity $N_{k,\ell}$, with harmonics orthonormal for normalized area σ, define $S_r^2 = \sum_{\ell\ge 2} N_{k,\ell}(1+\lambda_\ell)^{-r}$ and

$$C_r^2 = \sum_{\ell\ge 2} \frac{N_{k,\ell}[3\lambda_\ell^2 - (4k+2)\lambda_\ell + k(k+2)]}{k(k+2)(1+\lambda_\ell)^r}. \tag{12}$$

For $h_- = \min h_0$ and $\kappa_- = \min_{x,|v|_g=1} A_0(x)(v,v)$, put $\delta_{\mathrm{geom}} = \min\{h_-/S_r, \kappa_-/C_r\}$.

Lemma A.7 (Exact whole-ball geometric radius)**.** *Every input in the closed source ball $B_\delta(H)$ has positive support and positive-definite support Hessian if and only if $0 \le \delta < \delta_{\mathrm{geom}}$. In this interior, $h_z \ge h_- - \delta S_r > 0$ and $A_{h_z} \succeq (\kappa_- - \delta C_r)g > 0$. At equality an input has zero support or a zero support-Hessian eigenvalue. This is a whole-ball admissibility boundary, not an impossibility statement for prediction or for individual inputs outside the ball.*

Proof. For a real orthonormal basis $Y_{\ell j}$, the addition formula gives $\sum_j Y_{\ell j}(x)^2 = N_{k,\ell}$, so the Riesz norm on H of $z \mapsto E_r z(x)$ is exactly S_r at every x. For $B_f = \nabla^2 f + fg$, the trace and integrated Bochner identities give $\operatorname{tr} B_{Y_{\ell j}} = (k-\lambda_\ell)Y_{\ell j}$ and $\sum_j \|B_{Y_{\ell j}}(x)\|_{\mathrm{HS}}^2 = N_{k,\ell}[\lambda_\ell^2 - (k+1)\lambda_\ell + k]$; the second holds pointwise because the sum is rotation invariant and integrates to this value. Rotational invariance in the tangent plane and $\mathbb{E}_v B(v,v)^2 = ((\operatorname{tr} B)^2 + 2\|B\|_{\mathrm{HS}}^2)/(k(k+2))$ then imply, for every unit $v \in T_x S^k$, $\sum_j B_{Y_{\ell j}}(x)(v,v)^2 = N_{k,\ell}[(k-\lambda_\ell)^2 + 2(\lambda_\ell^2 - (k+1)\lambda_\ell + k)]/(k(k+2))$. This gives the exact Riesz norm C_r of every directional support-Hessian evaluation. The series converge since $r > k/2 + 2$; Cauchy–Schwarz proves the uniform margins.

At a minimizer of h_0, the negative evaluation representer of norm h_-/S_r makes support zero; at a minimizing point and unit direction of A_0, the negative directional representer of norm κ_-/C_r makes that quadratic form zero. Both belong to H, since only degrees at least two enter. At the smaller radius the preceding bounds keep support and all curvature forms nonnegative, so a vanishing curvature form is a zero eigenvalue. This proves sharpness and excludes every larger closed ball. □

Write $L_b\phi = \operatorname{cof}_g(A_{h_b}) : (\nabla^2\phi + \phi g)$, $Q_b f = Z_b^{-1}(f - p_{h_b}\langle f, 1\rangle)$, $T_{b,\tau} = M_{\sqrt{\omega_\tau}} M_\beta^{-1} K_\beta Q_b L_b E_r$ and $H_J = H \cap \ker \Pi_J$, where $M_{\sqrt{\omega_\tau}}$ multiplies by $\sqrt{\omega_\tau}$ and Π_J is the source harmonic projector through degree J.

Lemma A.8 (Uniform relative tails)**.** *Fix $0 < \delta < \delta_{\mathrm{geom}}$ and $0 < \gamma < r - s - 2 - k/2$. A common finite cutoff $J_{\min}$ and positive constants satisfy*

$$c_*\|u\|_{H^{-\mu}} \le \|T_{b,\tau}u\|_2 \le C\|u\|_{H^{-\mu}}, \quad b \in B_\delta(H),\ u \in H_{J_{\min}}. \tag{13}$$

For $b, c \in B_\delta(H)$ and $u \in H_{J_{\min}}$,

$$\|(T_{b,\tau} - T_{c,\tau})u\|_2 \le C\|E_r(b-c)\|_{W^{s+2,\infty}} \|T_{c,\tau}u\|_2. \tag{14}$$

For $J \ge J_{\min}$, $b \in \Pi_J H$, and $u, w \in H_J$ with $b+u, b+w \in B_\delta(H)$,

$$\|F_\tau(b+u) - F_\tau(b+w) - T_{b,\tau}(u-w)\|_2 \le \eta_J \|T_{b,\tau}(u-w)\|_2, \quad \eta_J = C\delta(J+1)^{-\gamma} \le \tfrac{1}{2}. \tag{15}$$

Forecasts are uniformly Lipschitz in the source norm. Constants and $J_{\min}$ need not remain bounded as $\delta \uparrow \delta_{\mathrm{geom}}$.

Proof. Finite regularity and an invertible shift. The strict embedding margin makes $E_r : L_2 \to C^{s+2}$ compact, so the weakly compact source ball has compact image in this strong topology. Cofactors and determinants have uniform C^s bounds, while Lemma A.7 gives positive support, curvature and normalizer margins. No full classical-symbol compactness of this infinite-dimensional family is used.

Put $h = h_b$, $L_h = L_b$, $Q_h = Q_b$ and $B_h = \operatorname{cof}_g(A_h)$. The tensor A_h is Codazzi on the sphere; differentiating the alternating-determinant expression for its cofactor and interchanging the Codazzi

derivatives cancels its divergence, and smooth approximation gives the same identity at the available regularity. Hence $L_h\phi = \operatorname{div}(B_h\nabla\phi) + (\operatorname{tr} B_h)\phi$ is symmetric. Choose $c_0 > 1 + \sup_h \|\operatorname{tr} B_h\|_\infty$. Then $P_h = -L_h + c_0 I$ has the uniformly coercive form $\langle P_h f, f\rangle = \int [B_h(\nabla f, \nabla f) + (c_0 - \operatorname{tr} B_h) f^2] d\sigma \geq c_1 \|f\|_{H^1}^2$ and a uniform inverse $H^{-1} \to H^1$. Local difference quotients of the weak equation, tested against a cutoff backward difference of a forward difference, yield $\|f\|_{H^2} \leq C(\|P_h f\|_2 + \|f\|_{H^1})$: coefficient terms contain one bounded derivative of B_h and first derivatives of f, Cauchy–Schwarz and ellipticity absorb the highest order, and a fixed chart cover makes the constant uniform. Differentiating j times gives commutators containing coefficient derivatives and derivatives of f, both through order $j+1$; induction and the energy estimate give $\|P_h^{-1}\psi\|_{H^{j+2}} \leq C_j \|\psi\|_{H^j}$ for $0 \leq j \leq s-2$. Smooth coefficient approximations with the same finite bounds and weak limits justify this for C^s coefficients. Interpolation at $j = \beta - 2 \in [0, s-2]$ gives $P_h^{-1} : H^{\beta-2} \to H^\beta$ uniformly. This is the finite-regularity version needed here; the smooth compact-manifold estimate is classical (Dyatlov, 2026, Section 15.1, Theorem 15.7). Symmetry and duality imply

$$\|\phi\|_{H^{2-\beta}} \leq C\|P_h\phi\|_{H^{-\beta}} \leq C(\|L_h\phi\|_{H^{-\beta}} + \|\phi\|_{H^{-\beta}}) \tag{16}$$

by pairing with $\psi \in H^{\beta-2}$ and using $\langle \phi, \psi\rangle = \langle P_h\phi, P_h^{-1}\psi\rangle$, first for smooth functions and then by density. Invertibility of L_h or $Q_h L_h$ has not been inferred.

Normalization and removal of the low-order remainder. The divergence identity gives $\int L_h\phi = \langle \phi, \operatorname{tr} B_h\rangle$, and $s > \beta$ bounds $\operatorname{tr} B_h$ in H^β. Insert $L_h\phi = Z_h Q_h L_h\phi + p_h \int L_h\phi$ in (16); with $\phi = E_r u$, the Bessel isomorphism and positive task weights give $\|u\|_{H^{-\mu}} \leq C(\|T_{b,\tau}u\|_2 + \|u\|_{H^{-\mu-2}})$. On H_J the second norm is at most $(1+\lambda_{J+1})^{-1}\|u\|_{H^{-\mu}}$, so one uniform finite cutoff absorbs it; multiplication estimates give the upper bound. This proves (13) without a whole-source spectral gap.

Relative variation and the true nonlinear tail. On the bounded coefficient sets multiplication is bounded on $H^{-\beta}$ since $s > \beta$. Cofactors, densities and reciprocal normalizers are Lipschitz in the $W^{s+2,\infty}$ support norm. Expanding $Q_b L_b - Q_c L_c = (Q_b - Q_c)L_b + Q_c(L_b - L_c)$ gives the corresponding $H^{2-\beta} \to H^{-\beta}$ bound, including normalization. Compose with E_r, K_β and apply tail coercivity for (14). Spectral Sobolev embedding gives $\|E_r u\|_{W^{s+2,\infty}} \leq C(J+1)^{-\gamma}\|u\|_2$ on H_J. The segment from $b+w$ to $b+u$ stays in the source ball; its displacement from b is a tail of norm at most δ. Integrating the full derivative proves (15). Enlarge the fixed cutoff until $\eta_J \leq 1/2$. The non-relative derivative bound and $\|u\|_{H^{-\mu}} \leq \|u\|_2$ prove source Lipschitzness. □

For $k = 2$, $r = 8$, $h_0 = 1$, $393.7701316584 < \delta_{\mathrm{geom}} < 393.7701316585$. Rational partial sums through degree 128 have squared-support and squared-curvature tails at most $3/(14\,128^{14})$ and $9/(20\,128^{10})$, respectively, from summand bounds $3\ell^{-15}$ and $(9/2)\ell^{-11}$; rational square comparisons give the endpoints, with curvature active. Radius 100 has full-source curvature margin greater than $.7460$. These radii use the declared smoothed source metric, not relative physical deformation, and do not establish onset at the tested budgets.

B SHARED CAPACITY: LOWER SUPPORTS, PURIFICATION AND RECOVERY

Retain Section 4's manifold, source, scalar elliptic operators and common relative form bounds, but allow a compact metric task space Θ. Require $\tau \mapsto C_\tau$ to be continuous and compact in the full Fréchet topology of Ψ^0_{cl}; operator-norm compactness alone is insufficient, as the final counterexample shows. In (2), replace the finite sum by $\int_\Theta a_\tau \, d\pi(\tau)$ for Borel probability mixtures. Now $\mathcal{R}_m^{\tan}$ consists of continuous nonnegative upper profiles, and $\mathcal{R}_\infty = \{r \in [0,\infty]^\Theta : \pi \cdot r \geq h(\pi)$ for every finite task mixture $\pi\}$.

Theorem 4.2: full compact-task form. Fix the family and $\delta > 0$ before $m \to \infty$. For every finite nonempty $F \subset \Theta$, coordinate restriction P_F satisfies

$$P_F\big[(2/\delta^2)m^{1/p}\mathcal{R}_m^{\tan}\big] \longrightarrow P_F\mathcal{R}_\infty. \tag{17}$$

Convergence is local sequential: bounded convergent projected upper-profile sequences have limits in this set, and every finite target has a recovering sequence. A continuous nonnegative r with $\int r\,d\pi \geq h(\pi) + \eta$ for every Borel mixture and some $\eta > 0$ has one shared encoder sequence

satisfying $\sup_\tau[(2/\delta^2)m^{1/p}\rho^{\tan}_{m,\tau}(E_m) - r_\tau]_+ \to 0$. The same holds for each $r \in \mathcal{R}_\infty \cap C(\Theta)_+$ along a target-specific slow diagonal. The proof below covers both this statement and its finite-task main-text specialization.

Completions are spectral tools, not extra source information. With $D = \widetilde{A}_*^{-1/(2\mu)}$, so $\widetilde{A}_* = D^{-2\mu}$, use the cosphere $\Sigma_D = \{d_1 = 1\}$, where d_1 is the principal symbol of D.

B.1 Lower supports and finite-branch allocation

Equal Gramians give equal task risks for every encoder: polar decomposition transports decoders between closed operator ranges, and orthogonal projection never increases error, so T_τ may be replaced by $A_\tau^{1/2}$. Root radii obey $|\sqrt{2\rho_\tau(E)} - \sqrt{2\rho_{\tau'}(E)}| \le \delta\|A_\tau^{1/2} - A_{\tau'}^{1/2}\|$, so profiles are continuous. For $C_\pi = \int C_\tau\, d\pi$, the completed mixture $\widetilde{A}_*^{1/2} C_\pi \widetilde{A}_*^{1/2}$ has compressed Gram $A_\pi = \int A_\tau\, d\pi$, and negative-order Weyl theory gives

$$\sup_\pi \left| m^{1/p}\lambda_{m+1}(A_\pi) - h(\pi) \right| \longrightarrow 0. \tag{18}$$

The barycenters form a compact full-symbol family between $c_- I$ and $c_+ I$; finite relative-form nets, min–max sandwiching and finite-codimension interlacing make the pointwise Weyl law uniform. Finite seminorm nets also show that weak convergence of mixtures gives full-symbol convergence, so h is weakly continuous (Birman & Solomyak, 1979; Ponge, 2023).

For a finitely supported mixture, the direct-sum map $(\sqrt{\pi_\tau}T_\tau)_\tau$ has Gram A_π. Restrict a continuous encoder to the radius-δ sphere of the direct-sum map's leading $m+1$ right singular directions. Borsuk–Ulam (Matoušek, 2003) gives equal-code antipodes, and the parallelogram identity gives

$$\sum_\tau \pi_\tau \rho_{m,\tau}(E) \ge \tfrac{\delta^2}{2}\lambda_{m+1}(A_\pi). \tag{19}$$

Leading singular coordinates give the singleton upper bound (DeVore et al., 1989).

Two tasks do not require purification. For two tasks, let $\varepsilon_m = \sup_{\pi\in\Delta_2} |m^{1/p}\lambda_{m+1}(A_\pi) - h(\pi)|$. If $r \in \mathcal{R}_\infty$ has finite coordinates, then $\pi \cdot (r + \varepsilon_m \mathbf{1}) \ge m^{1/p}\lambda_{m+1}(A_\pi)$ for every $\pi \in \Delta_2$, so Theorem 5.3 supplies a rank-m projection meeting $(\delta^2/2)m^{-1/p}(r + \varepsilon_m\mathbf{1})$. Since $\varepsilon_m \to 0$ by (18), this is the recovering sequence; (19) gives the converse for convergent profiles (the proof of Theorem 5.3 uses only Lemma D.2 and the antipodal bound, not Lemma B.3). Appendix F.2 measures how fast finite budgets approach this limit for a noncommuting pair.

Lemma B.1 (One finite-branch allocation calculation)**.** *For positive continuous coefficients $b_{\tau j}$, set $f_\tau(y,s) = \sum_j b_{\tau j} y_j s_j^{-1/p}$, $y, s \in \Delta_R$, using $0s^{-1/p} = 0$ and infinite value for positive numerator over zero. For fixed y, its upper image is closed and convex in the risk profile. With $\bar b_j = \sum_\tau \pi_\tau b_{\tau j}$, its mixture support and its worst energy split are*

$$\inf_s \sum_\tau \pi_\tau f_\tau(y,s) = \left[\sum_j \left(y_j \bar b_j\right)^{p/(p+1)}\right]^{(p+1)/p}, \qquad \max_y \inf_s \sum_\tau \pi_\tau f_\tau(y,s) = \left[\sum_j \bar b_j^{\,p}\right]^{1/p}.$$

A strict target above all the latter supports is feasible for every y. On a compact finite-dimensional parameter set this feasibility has a Lipschitz selector with a common positive coordinate floor and retained strict margin.

Proof. Convexity is in s, not in y; compactness of the simplex and lower semicontinuity prove closedness. Lagrange multipliers give the first formula and Hölder the second. A separating positive functional on continuous task profiles is a probability measure, so the support inequalities imply feasibility even for compact task sets. Mixing a feasible s with a fixed, sufficiently small weight θ of the uniform allocation costs at most the factor $(1-\theta)^{-1/p}$ and enforces a floor. Tighten the margin, choose finitely many parameter balls, with centers w_i and radii r_i, on which constant feasible allocations s_i remain feasible and whose half-radius balls cover the parameter set. The weights $\phi_i(w) = (r_i - \|w - w_i\|)_+$ have sum bounded below, so $s(w) = \sum_i \phi_i(w)s_i / \sum_i \phi_i(w)$ is Lipschitz and remains feasible by convexity. The scalar minimization has classical approximation antecedents (Nguyen & Nguyen, 2022); the shared recovery still needs a code. □

B.2 SCALAR MICROLOCAL LOCALIZATION

Lemma B.2 (Thin-cluster weighted Szegő input)**.** *For the positive selfadjoint real order-one reference D above, $F = F^* \in \Psi^0_{\mathrm{cl}}$ with even principal symbol f, $B \in \Psi^0_{\mathrm{cl}}$ with principal symbol b and $\Pi = 1_{[L,L+W]}(D)$, if $\log(2+L) = o(W)$ and $W/L \to 0$, then*

$$\frac{\operatorname{tr}[\Pi B \Pi \varphi(\Pi F \Pi)]}{L^{k-1}W} \longrightarrow \frac{1}{(2\pi)^k} \int_{\Sigma_D} b\varphi(f)\, d\Lambda.$$

This holds for continuous φ with $\varphi(0) = 0$, uniformly on fixed compact symbol families. It also holds for interval indicators with fixed continuity-set endpoints. Uniform indicator limits require boundary-nullness for every limiting symbol; windows are frozen at each refinement stage.

Proof. For unit-band projectors P_i and every fixed $S \in \Psi^0$, bounded iterated commutators give $\|P_i S P_j\| \le C_N(1+|i-j|)^{-N}$, and the weighted cumulative Weyl law gives $\operatorname{rank} P_i = O((1+i)^{k-1})$ (Laptev & Safarov, 1996, Lemma 2.3). At separation d, $O(d+1)$ band pairs cross the two cluster edges; summing their squared Hilbert–Schmidt norms, with the partial endpoint bands charged in full, yields $\|(I-\Pi)S\Pi\|^2_{\mathcal{S}_2} = O_S(L^{k-1})$.

On $\operatorname{ran}\Pi$, put $Q = \Pi F \Pi$ and $R_\ell = \Pi F^\ell \Pi - Q^\ell$. The identity $R_\ell = R_{\ell-1}Q + \Pi F^{\ell-1}(I-\Pi)F\Pi$ and Schatten Cauchy–Schwarz give $\|R_\ell\|_{\mathcal{S}_1} = O_\ell(L^{k-1})$. A second return across the edge handles the extra weight:

$$\begin{aligned}|\operatorname{tr}(\Pi B \Pi Q^\ell) - \operatorname{tr}(\Pi B F^\ell \Pi)| &\le \|B\|\|R_\ell\|_{\mathcal{S}_1} \\ &\quad + \|\Pi B(I-\Pi)\|_{\mathcal{S}_2}\|(I-\Pi)F^\ell\Pi\|_{\mathcal{S}_2} = O(L^{k-1}).\end{aligned}$$

Subtracting the cumulative Weyl laws at $L+W$ and L gives normalized error $O(W/L + 1/W)$, sufficient under the stated conditions. This supplies the weighted thin-shell step beyond the unweighted compression estimate of Laptev & Safarov (1996, Theorem 1.2). Uniform polynomial approximation proves the continuous case. For indicators, continuous upper and lower cutoffs control the unweighted boundary count, $|\operatorname{tr}(\Pi B \Pi T)| \le \|B\| \operatorname{tr}|T|$ handles a nonpositive weight, and fixed endpoints and compactness make these boundary errors uniform. No uniform estimate for arbitrarily moving windows is used. □

Lemma B.3 (Uniform scalar microlocal purification)**.** *For every $\epsilon > 0$, the compact family specified in the full statement above has an orthogonal real decomposition $H = F_\epsilon \mathbin{\widehat{\oplus}} V_g \mathbin{\widehat{\oplus}}_{j=1}^{R} V_j$, with finite F_ϵ encoded exactly. There are positive constants C_j, C_g and continuous $d_{\tau j}, d_{\tau g}$ such that, for all $e_j \in V_j$ and $e_g \in V_g$,*

$$\lambda_\ell(P_{V_j}\widetilde{A}_* P_{V_j}|_{V_j}) = C_j \ell^{-1/p}(1+o(1)), \quad \limsup_\ell \ell^{1/p}\lambda_\ell(P_{V_g}\widetilde{A}_* P_{V_g}) \le C_g,$$

$$\left\langle A_\tau(\sum_j e_j + e_g), \sum_j e_j + e_g \right\rangle \le (1+\epsilon)\sum_j d_{\tau j}\left\langle \widetilde{A}_* e_j, e_j\right\rangle + d_{\tau g}\left\langle \widetilde{A}_* e_g, e_g\right\rangle, \tag{20}$$

$$\sup_\tau d_{\tau g} C_g \le \epsilon.$$

Moreover, uniformly over all Borel mixtures,

$$\sup_\pi \left| \left[\sum_j \left(C_j \int d_{\tau j}\, d\pi \right)^p + \left(C_g \int d_{\tau g}\, d\pi \right)^p \right]^{1/p} - h(\pi) \right| \le \epsilon.$$

Real and complexified branch dimensions agree. For a fixed finite task family, the same cores can also satisfy a two-sided comparison: with $B_j = P_{V_j}\widetilde{A}_*|_{V_j}$ and $b_{\tau j} = c_\tau(\theta_j)$, where c_τ is the relative principal symbol and θ_j lies in the corresponding cell, for any prescribed $0 < \nu < 1$, the refinement and the finite exact complement can be chosen so that

$$\left| \left\langle A_\tau \sum_j u_j, \sum_j u_j \right\rangle - \sum_j b_{\tau j}\langle B_j u_j, u_j\rangle \right| \le \nu \sum_j b_{\tau j}\langle B_j u_j, u_j\rangle \quad (u_j \in V_j). \tag{21}$$

This comparison excludes the guard; the full recovery bound above still charges it. Core volumes determine the C_j^p as in the proof below.

Proof. **1. Remove the gauge once, preserving orthogonality.** Let $K = 1_{[L_0,\infty)}(D)L_2$. Smooth finite-dimensional G gives $\|P_K P_G P_K\| = O(L_0^{-N})$ for every N, so for large L_0, $U = P_H P_K (P_K P_H P_K)^{-1/2} : K \to H$ is a real isometry with finite-codimensional range. Its finite-rank smoothing correction from the inclusion tends to zero in every Sobolev seminorm. Hence $\sup_\tau \|\widetilde{A}_*^{-1/2} P_K (U^* A_\tau U - \widetilde{A}_\tau) P_K \widetilde{A}_*^{-1/2}\| \to 0$, and so does the corresponding reference-form difference. One global U preserves cross-shell orthogonality; encode $H \ominus U(K)$ exactly.

We use classical $(1,0)$ symbol composition and sharp Gårding (Taylor, 1991, Propositions 0.3.C and 0.7.A; Eq. (0.7.9)), localized to the compact manifold. For $T \in \Psi^0$ with principal symbol t_0, apply Gårding after order reduction to the nonnegative principal symbol of $\|t_0\|_\infty^2 I - T^*T$. The order-$-1$ remainder is at most $C_T \|u\|_{H^{-1/2}}^2$, hence $O_T(L^{-1})\|u\|^2$ on the tail, giving

$$\|T 1_{[L,\infty)}(D)\| \le \|t_0\|_\infty + o_T(1), \tag{22}$$

uniformly on compact symbol families. Cutoff derivatives affect onset, not the leading norm coefficient.

2. Separate windows, then remove guard and edge energy. At a frozen refinement stage n, partition the projective cosphere into R_n small positive-volume cells with null boundaries. Choose disjoint compact cores $\Gamma_{n,j} \Subset \Gamma_{n,j}^+ \Subset U_{n,j}$, with disjoint collar closures and all excess in a guard G_n of volume γ_n. Real even cutoffs $b_{n,j} \in [0,1]$ are supported in $U_{n,j}$ and equal one near $\Gamma_{n,j}^+$. A real even label f_n has distinct nonzero constant values on the enlarged cores and transitions only in the guard. Choose disjoint nested windows $I_{n,j}^- \Subset I_{n,j}^+$ away from zero, with continuity-set endpoints and $f_n^{-1}(I_{n,j}^+) \subset \Gamma_{n,j}^+ \cup G_n$, and pick $\chi_{n,j} \in C_c^\infty(I_{n,j}^+)$ equal to one on $I_{n,j}^-$. An even $g_n \in [0,1]$ supported in the guard satisfies $(1-b_{n,j})(1-g_n)\chi_{n,j}(f_n) = 0$ and $\int g_n^2 \, d\Lambda \le \gamma_n$.

Use shell width $W(L) = L^\alpha$, edge width $H(L) = L^\beta$ and separating gaps $G(L) = L^\gamma$, where $0 < \beta < \gamma < \alpha < 1$. Write $J_q = [L_q, L_q + W(L_q)]$, $\Pi_q = 1_{J_q}(D)$, $F = \mathrm{Op}(f_n) = F^*$ and $Q_q = \Pi_q F \Pi_q$, and let $E_q \le \Pi_q$ project onto the two edge bands of width $H(L_q)$. The window projectors $W_{q,j} = 1_{I_{n,j}^-}(Q_q)$ are exactly orthogonal and $\chi_{n,j}(Q_q) W_{q,j} = W_{q,j}$.

Purify each window twice. First retain the spectral subspace of $W_{q,j}\Pi_q \,\mathrm{Op}(g_n)^* \,\mathrm{Op}(g_n)\Pi_q W_{q,j}$ below η_n; by Lemma B.2 and Markov's bound the total removed shell fraction is $O(\gamma_n/\eta_n) + o_q(1)$ *without* an R_n multiplier, since compressed traces sum to at most the full trace. Then retain the spectral subspace below θ_n of E_q compressed to this retained subspace, removing $O(H(L_q)/(\theta_n W(L_q)))$ of the shell. Thus the projector $P_{q,j}$ onto the twice-retained subspace satisfies $P_{q,j} \le W_{q,j} \le \Pi_q$, $\|\mathrm{Op}(g_n) P_{q,j}\| \le \sqrt{\eta_n}$, $\|E_q P_{q,j}\| \le \sqrt{\theta_n}$ and $\chi_{n,j}(Q_q) P_{q,j} = P_{q,j}$; the last identity uses only range inclusion. Removed vectors join the guard.

3. Localize the purified spaces without assuming invariance. Freeze polynomial approximants $\mathfrak{p}$ to each $\chi_{n,j}$ with uniform error κ_n on a common interval containing the spectra and the range of f_n. If $B_{\mathrm{out}} \in \Psi^0$ has $|b_{\mathrm{out}}| \le 1$ and $b_{\mathrm{out}} \chi_{n,j}(f_n) = 0$, then

$$\|B_{\mathrm{out}} \chi_{n,j}(Q_q) P_{q,j}\| \le 2\kappa_n + A_n \sqrt{\theta_n} + \omega_n(L_q), \qquad \omega_n(L) \to 0. \tag{23}$$

Split $B_{\mathrm{out}} \chi_{n,j}(Q_q) P_{q,j}$ via $\chi_{n,j}(Q_q) = (\chi_{n,j} - \mathfrak{p})(Q_q) + (\mathfrak{p}(Q_q) - \mathfrak{p}(F)) + \mathfrak{p}(F)$. The first term stays on the shell before B_{out} acts and costs $\kappa_n + o_n(1)$ by (22). The last has principal symbol $b_{\mathrm{out}} \mathfrak{p}(f_n)$ of magnitude at most κ_n, so has the same bound, using only polynomial symbolic composition. For the middle term, telescope compressed and uncompressed powers, e.g. $Q_q^\ell - F^\ell = \sum_{a<\ell} Q_q^{\ell-1-a}(Q_q - F)F^a$ with $Q_q - F = -(I-\Pi_q)F - \Pi_q F(I-\Pi_q)$. The edge part costs $A_n \|E_q P_{q,j}\|$; on the central shell, repeated commutators give $\|(I-\Pi_q)F^a(I-E_q)\Pi_q\| \le C_{n,a,N} H(L_q)^{-N}$ for each fixed a. All central and negative-order errors enter ω_n; A_n is fixed before θ_n.

Put $\widehat{B}_j = \mathrm{Op}(b_{n,j})$ and $R_j = \mathrm{Op}((1-b_{n,j})(1-g_n))$. Symbolic composition yields $I - \widehat{B}_j = R_j + (I - \widehat{B}_j)\,\mathrm{Op}(g_n) + K_j$ with $K_j \in \Psi^{-1}$. Apply (23) to R_j; split $\mathrm{Op}(g_n) P_{q,j}$ at $L_q/2$ and use (22) and commutator decay. Hence $\|(I - \widehat{B}_j) P_{q,j}\| \le 2\kappa_n + \sqrt{\eta_n} + A_n\sqrt{\theta_n} + \omega_n(L_q)$. With $C = C_\tau$, disjoint collars give $\widehat{B}_i^* C \widehat{B}_j \in \Psi^{-1}$ for $i \ne j$. Inserting $\widehat{B}_i P_{q,i}$ and $\widehat{B}_j P_{q,j}$ into $P_{q,i} C P_{q,j}$ costs, per replacement, the localization error times a uniformly bounded factor, and the remaining

$P_{q,i}\widehat{B}_i^* C \widehat{B}_j P_{q,j}$ tends to zero by its negative order. Thus, uniformly in the relative task family,

$$\|P_{q,i}\Pi_q C \Pi_q P_{q,j}\| \le C_0(\kappa_n + \sqrt{\eta_n}) + A_n\sqrt{\theta_n} + \omega_n(L_q), \tag{24}$$

where C_0 uses only the relative upper bound, not global cutoff derivatives.

4. Transfer to reference energy and sum over shells. Task forms act on $D^{-\mu}e$, which need not stay purified. For $e \in \operatorname{ran} P_{q,j}$, subtracting $L_q^{-\mu}I$ on the thin shell gives $\|(I - P_{q,j})D^{-\mu}e\| \le C(W(L_q)/L_q)\|D^{-\mu}e\|$ and bounds $\mathrm{Op}(g_n)D^{-\mu}e$ by $[\sqrt{\eta_n} + o_n(1) + O(W/L_q)]\|D^{-\mu}e\|$. This transfers the detector and off-diagonal estimates to reference energies. Sharp Gårding applied to the upper and lower symbol bounds for $\widehat{B}_j^*(C_\tau - b_{\tau j}I)\widehat{B}_j$ controls diagonal errors by the cell oscillation plus a vanishing remainder. Cross errors in (24) multiply geometric means of reference energies; summing over branches costs at most R_n.

For distinct whole shells, bounded iterated commutators with D and the separated-spectrum Sylvester identity give $\|\Pi_q C_\tau \Pi_{q'}\| \le C_N(1 + \operatorname{dist}(J_q, J_{q'}))^{-N}$, uniformly in τ. Above a cutoff L_0, successive gaps are at least $g_0 \asymp L_0^\gamma$, so for $N > 1$, $\sup_q \sum_{q' \ne q} \|\Pi_q C_\tau \Pi_{q'}\| \le 2C_N \sum_{\ell \ge 1}(1 + \ell g_0)^{-N} \to 0$. This row-sum bound controls the cross-shell quadratic form on $w_q = D^{-\mu}e_q \in \operatorname{ran}\Pi_q$ by $o(1)\sum_q \|w_q\|^2$ and does not require the purified subspaces to be invariant under D or Q_q. Since $\min_{\tau,j} b_{\tau j} \ge c_- > 0$, choose the combined cell, diagonal, cross and gauge errors below νc_-; increasing the finite cutoff then gives (21). Upper cell suprema instead give the one-sided bound for recovery. The parameter order below makes these choices noncircular.

5. Count exact branch dimensions and charge the guard. Detector bounds alone supply only rank lower bounds. Choose $\delta_n > 0$ below half of every core volume, with $R_n\delta_n \to 0$ and $\gamma_n/\eta_n = o(\delta_n)$. Beyond a fixed threshold keep exactly $N_{q,j} = \lfloor L_q^{k-1}W(L_q)[\Lambda_D(\Gamma_{n,j}) - \delta_n]/(2\pi)^k \rfloor$ real orthonormal core vectors and assign the surplus to the guard; these subspaces inherit both detector bounds. Since $(L_q + W(L_q))^{-2\mu}I \preceq \widetilde{A}_* \preceq L_q^{-2\mu}I$ on each shell, counting gives the branch constants $C_j^p = [\Lambda_D(\Gamma_{n,j}) - \delta_n]/[k(2\pi)^k]$, where $\Lambda_D(\Gamma_{n,j})$ is the full-cosphere volume of both antipodal lifts. Real even quantizations make each branch conjugation invariant, so its real dimension equals its complexified dimension; there is no additional factor of two. Apply the one global U to the shell sums; it preserves constants, exact orthogonality and the tolerated form bounds.

The guard contains collars, removals, surplus, gaps and all remaining high vectors. Its counting bound is $C_g^p \lesssim \gamma_n/\eta_n + R_n\delta_n + \limsup_q H(L_q)/(\theta_n W(L_q)) + \limsup_q G(L_q)/W(L_q)$.

6. Choose parameters before taking spectral limits. Core–guard Cauchy–Schwarz costs factors $1+\zeta_n$ and $1+\zeta_n^{-1}$. Take $t_n, \zeta_n \downarrow 0$ and fix cells and R_n. Choose $R_n\delta_n \le t_n\zeta_n^p$ and $R_n(\kappa_n + \sqrt{\eta_n}) \le t_n$; then shrink collars so $\gamma_n/\eta_n \le t_n \min\{\delta_n, \zeta_n^p\}$. Freeze labels, windows, cutoffs and polynomial approximants, hence A_n; only then take $R_nA_n\sqrt{\theta_n} \le t_n$. Finally increase the spectral threshold until all fixed-stage remainders meet tolerance. The radial limsups vanish, $\gamma_n/\eta_n + R_n\delta_n = o(\zeta_n^p)$, and $d_{\tau g} \le (1 + \zeta_n^{-1})c_+ + o(1)$ gives $\sup_\tau d_{\tau g}C_g \to 0$.

Choose $d_{\tau j} = \sup_{U_{n,j}} c_\tau$ plus a vanishing common allowance for summed errors. Uniform task oscillation, exact core counts and vanishing guard support show that the finite-branch functional tends uniformly to (2), proving all parts of the lemma. □

B.3 Stable implementation of the common allocation

For $M, R \ge 1$ define ramps $\alpha_{j\ell}(s) = \min\{1, \max\{0, Ms_j - \ell + 1\}\}$, $v = Q_s x = (\alpha_{j\ell}x_{j\ell})_{j\le R, \ell \le M}$. Every support, and every intersection of supports, is a union of branch prefixes of total size at most $d = M + R - 1$, and there are at most $(M+1)^R$ such nonempty supports.

Lemma B.4 (Prefix measurements and a division-free tag)**.** *If $A : \mathbb{R}^{RM} \to \mathbb{R}^q$ has $\|A\| \le H_A$ and $\sigma_{\min}(A_I) \ge c > 0$ on every such prefix, then $\mathcal{E}_A(x,s) = (AQ_sx, ts)$, $t \ge \|x\|$, determines Q_sx with inverse Lipschitz constant $L_{M,R,A} = c^{-1} + 2M(1+H_A/c)(1+\sqrt{R})$. One can choose $\|A\| \le 1$ with $q = M + O(M/\log M)$ and $L_{M,R,A} \le CM(1+R)\log(M+2)$ whenever $R \le \log(M+2)$ and M is large. Dyadic entries with $O(\log M)$ bits suffice.*

Proof. For two codes, split $v - v' = d_0 + e$ over the common prefix and symmetric difference. Lipschitz ramps give $\|e\| \le M(t+t')\|s - s'\|_\infty$, and the conditioned intersection gives $\|v - v'\| \le$

$c^{-1}\|A(v-v')\| + (1+H_A/c)\|e\|$. For $u = ts, u' = t's'$, $t = \sum u_i$ and $(t+t')\|s-s'\|_\infty \le 2\|u-u'\|_\infty + 2|t-t'| \le 2(1+\sqrt{R})\|u-u'\|_2$. This includes zero and empty intersections. At $t = \|x\|$, the code is continuous at zero even if the allocation is only continuous away from zero.

For existence put $K = (M+1)^R$, $\epsilon = 1/\log(M+2)$, $t_* = \sqrt{2\log(8K)}$, $t_0 = \sqrt{2\log 8}$ and $q = \lceil((1+\epsilon)\sqrt{d} + t_*)^2\rceil$. Gaussian extreme-singular-value bounds (Vershynin, 2012, Corollary 5.35) and a union bound give, with positive probability, every prefix lower singular value at least $\epsilon\sqrt{d}$ and whole-matrix norm at most $\sqrt{q} + \sqrt{RM} + t_0$. Dividing by the latter deterministic bound gives $c = \epsilon\sqrt{d}/(\sqrt{q} + \sqrt{RM} + t_0)$, and expanding q gives the stated estimates. Entry rounding by $c/(8\sqrt{qRM})$ perturbs the operator norm by at most $c/8$; dividing by two keeps the norm at most one and singular values at least $7c/16$. Only prefixes/intersections are conditioned, not all sparse differences. At $q < 2M$ two disjoint full branches already have an untagged measurement kernel, so the tag is substantive. □

For an orthogonal decomposition $z = b + f + x$ and the floored selector of Lemma B.1, use

$$E(z) = (b, f, AQ_{s(b,y)}x, \|x\|s(b,y)), \qquad y_i = \|x_i\|^2/\|x\|^2, \tag{25}$$

with zero tail code at $x = 0$. The four transmitted blocks are the coarse source, the finite exact complement, mixed ramped coordinates and R allocation tags. If their first two dimensions are d_b, d_f, the total requirement is

$$d_b + d_f + q(M,R) + R \le m. \tag{26}$$

Here R includes the guard branch. The tag entries sum to $\|x\|$, so recovery needs no residual-energy message; the decoder uses the received tag and measurements in Lemma B.4. The energy map satisfies $\min(\|x\|, \|x'\|)\|y - y'\|_2 \le 4\|x - x'\|$, since on unit vectors it is 2-Lipschitz into ℓ_1 and $\min(\|x\|, \|x'\|)\|x/\|x\| - x'/\|x'\|\| \le 2\|x - x'\|$. For selector constant L, the ramp difference on the smaller residual is at most $(1+4ML)\|x-x'\| + M\delta L\|b-b'\|$; the tag has the same estimate with $M = 1$. Thus $\mathrm{Lip}(E) \le C_\delta[1 + M(1+L)]$ and $\|E(z)\|^2 \le \|b\|^2 + \|f\|^2 + 2\|x\|^2 \le 2\delta^2$, giving stable encoding and decoding; singular vectors need not vary continuously with a task parameter.

At stage j, put $D_j = d_b + d_f$ and let R_j count all branches, including the guard. Freeze the partition, spectral tolerances, target margin and selector constant L_j. Choose thresholds so every budget beyond them has $R_j \le \log(M+2)$, $L_j \le \log(e+m)$ and $D_j \le m/\log(e+m)$, and all stage errors small. Use the largest M satisfying (26); maximality and $q = M + O(M/\log M)$ from Lemma B.4 give $m - M = O(m/\log m)$, hence $M/m \to 1$. Zero-pad unused coordinates; before the first stage threshold use the zero code and enlarge the finite-budget constants. The clean source $S_m E_m z = b + f + Q_s x$ contracts the source and preserves the encoded background. The inverse tag and orthonormal synthesis give $\mathrm{Lip}(S_m) \le Cm\log^2(e+m)$ on the compact code closure; Kirszbraun extension and source-ball projection give it globally without changing clean recovery.

B.4 Completing the capacity and stability proofs

Choose a purification stage with support error below a fraction of the strict target margin, and move a finite leading guard prefix into the exact block so the remaining guard eigenvalues are at most $(C_g + \upsilon)\ell^{-1/p}$, with υ sufficiently small. After diagonalizing each compressed reference Gram, the normalized truncation bound has the form in Lemma B.1, with $b_{\tau j} = d_{\tau j}C_j$ and $b_{\tau g} = d_{\tau g}(C_g + \upsilon)$; apply the selector and code (25). The guard is allocated, not discarded. The stage sequence charges every block and gives $M/m \to 1$; all target risks follow from (20), and decreasing margins give boundary targets.

The lower set inclusion is (19) plus (18). For finite F, a feasible profile extends by $+\infty$ outside F, and recovery for F gives its finite target. For continuous upper profiles on all of Θ, add to the continuous actual risk profile nonnegative Urysohn functions that equal the Kronecker data on F. This gives (17), not convergence of arbitrary raw risks. For a continuous strict target on Θ, the preceding stage errors and selector are uniform in the task; target-specific thresholds and then decreasing margins give the asserted uniform recovery, including boundary targets.

Put $c_\tau = h(\delta_\tau)$ and $H_* = \sup_\pi h(\pi)$. Every worst-task risk dominates every mixture; recovery of $(H_* + \epsilon)\mathbf{1}$ and singleton Weyl convergence give worst-task leading constant H_* and price $\Pi_\Theta =$

$H_*/\sup_\tau c_\tau$ relative to task-conditioned coding. Dirac constraints give $\mathcal{R}_\infty \subseteq \{r \geq c\}$. For $p \geq 1$, Minkowski gives $h(\pi) \leq \int c_\tau \, d\pi$, hence $\mathcal{R}_\infty = \{r \geq c\}$. Recovery of the corner $r = c$ and the uniform singleton lower bound give the full taskwise conclusion $\sup_\tau |(2/\delta^2)m^{1/p}\rho^{\tan}_{m,\tau}(E_m) - c_\tau| \to 0$.

For $p < 1$, reverse Minkowski gives the reverse inequality; equality for a positive two-task mixture forces proportional positive symbols, so rectangularity is equivalent to one common positive ray. Off balance, concavity of $\Phi(a) = \int a^p$ and $D\Phi(a)[v] = p\int a^{p-1}v$ show that $\Pi_\Theta = 1$ exactly when a hardest task τ_* satisfies $\int_{S^*M} a_{\tau_*}^{p-1} a_\tau \, d\Lambda \leq \int_{S^*M} a_{\tau_*}^p \, d\Lambda$ for every τ. Under capacity balance (equal $\int a_\tau^p \, d\Lambda$), strict concavity makes unit price equivalent to equality of all principal energies a.e. Finally, $s_j(\widetilde{A}_*^{1/2})^2 \asymp j^{-1/p}$ gives the Hilbert–Schmidt boundary, with divergence at $p = 1$.

For precision, quantize (25) on $[-\sqrt{2}\delta, \sqrt{2}\delta]^m$ to $b_m = \lceil (\mu/k+2)\log_2(m+1) \rceil$ bits per coordinate. Then $\|S_m(Q_m E_m z) - S_m(E_m z)\| \leq C\delta m^{3/2}\log^2(e+m)2^{-b_m} = o(\delta m^{-\mu/k})$. Bounded tangent maps preserve the leading risks; for native nonlinear recovery use Lemma A.5's source-perturbation bound. Here Q_m is a separate discontinuous quantization channel, not a continuous finite-alphabet encoder. Onset is target-dependent, and the Lipschitz constants grow rather than remain fixed.

Necessary scope boundaries. Full-symbol compactness cannot be replaced by negative-order operator-norm compactness. On mean-zero real S^1, let $\widetilde{A}_* = (I - \partial_x^2)^{-\mu}$ and let P_N select $N < |n| \leq 2N$. For $c > 2^{2\mu}$, set $\widetilde{A}_N = \widetilde{A}_*^{1/2}[I + (c-1)P_N]\widetilde{A}_*^{1/2}$. Then $\|\widetilde{A}_N - \widetilde{A}_*\| = O(N^{-2\mu})$ and all principal symbols agree, yet budgets comparable to N have an order-one normalized spectral change. The relative smoothing blocks are not compact in full Ψ^0.

C Fixed-radius capacity and the need to adapt to geometry

Keep the native assumptions of Section 3 (hence $p < 1/5$), a finite family of task weights and $0 < \delta < \delta_{\mathrm{geom}}$. At background b, write $A_b = A_{h_b}$ and use L_b, Q_b and the task derivative $T_{b,\tau}$ of Appendix A.7. On the ordinary metric unit cosphere the positive principal amplitude, up to the overall Hessian sign, is $t_b(u,\widehat{\xi}) = \tau_s^{-\beta/2}\operatorname{cof}_g(A_b(u))(\widehat{\xi}^\sharp, \widehat{\xi}^\sharp)/(M_\beta Z_b)$, and $a_{b,\tau} = \omega_\tau t_b^2$. Its integral is continuous on the entire source ball through $E_r : L_2 \to C^2$ and the positive margins. Full-symbol calculus is used below only on frozen finite-dimensional smooth backgrounds.

C.1 Relative tail geometry and the exact nonlinear body

Lemma A.8 supplies the common cutoff $J_{\min}$, tail H_J and relative secant (15); below, $J \geq J_{\min}$ and $D_J = \operatorname{rank}\Pi_J \asymp J^k$. Coercivity is used only on this tail, whose relative secant error $\eta_J \leq 1/2$ tends to zero at the fixed radius; the coarse block is charged, not assumed known to the decoder.

Proof of Theorem 4.1. Lower inclusion. Fix a finite-harmonic b with $\|b\| < \delta$ and $J \geq J_{\min}$ containing it, and let $a_b = (\delta^2 - \|b\|^2)^{1/2}$. In the leading $m+1$ singular directions of $(\sqrt{\pi_\tau}T_{b,\tau}|_{H_J})_\tau$, Borsuk–Ulam gives a unit w with $b \pm a_b w$ equally encoded. For deterministic adaptive linear queries, run the policy at b, freeze its at most m functionals and intersect the same singular space with their common kernel. Every input $b + u$ with u in that intersection follows the same answer path, including internal rounding, so the same antipodal inequality holds without continuity of the query rule (Krieg et al., 2025):

$$\sum_\tau \pi_\tau \rho^{\mathrm{nl}}_{m,\tau} \geq \frac{a_b^2}{2}(1-\eta_J)^2\lambda_{m+1}\left(P_{H_J}\sum_\tau \pi_\tau T_{b,\tau}^* T_{b,\tau}|_{H_J}\right). \tag{27}$$

At fixed (b, J), Weyl's constant is $\mathsf{h}_b(\pi)$. Let $m \to \infty$, then $J \to \infty$; harmonic approximation and continuity of the cofactor integral extend the constraint to all b. Boundary backgrounds contribute nonnegativity. This proves the C_δ outer constraints for all three decoder/acquisition settings, independently of any encoder-class inclusion.

One decomposition for all backgrounds and tasks. Freeze $J \geq J_{\min}$ and $B_J = B_\delta(H) \cap \Pi_J H$. With $G_J = H_J^\perp$, complete each tail Gram by $\widetilde{A}_{b,\tau,J} = P_{H_J}T_{b,\tau}^*T_{b,\tau}P_{H_J} + P_{G_J}(I+\Delta)^{-\mu}P_{G_J}$.

By (13), the completions have common positive relative form bounds and are injective. At fixed J, cofactor and normalizer dependence gives a compact full-symbol family on $B_J \times F$, as required by the compact-task formulation in Appendix B. Lemma B.3 supplies one orthogonal tail decomposition and reference eigenbases *independent of* b, τ, with continuous coefficients $d_{b,\tau,i}$ and mixture supports converging uniformly to $\mathsf{h}_b(\pi)$.

One allocation and one charged code. For $r \in C_\delta$, take the target margin $r^+ = r + \epsilon\mathbf{1}$ and purify until $\pi \cdot r^+ \geq v_\delta(b) h_\epsilon(b,\pi) + \epsilon/2$ uniformly, where $h_\epsilon(b,\pi)$ is the mixture support of the purification at tolerance ϵ. Apply Lemma B.1 to $f_{b,\tau}(y,s) = v_\delta(b) \sum_i d_{b,\tau,i} C_i y_i s_i^{-1/p}$; at $v_\delta(b) = 0$ it is zero. Convex upper epigraphs and the last two claims of that lemma make the purified target feasible for every (b,y), and the lemma's floored Lipschitz selector $s(b,y)$ on the compact parameter set of (b,y) meets every task simultaneously. Write $z = b + f + \sum_i x_i$ with f in the charged exceptional block and code by (25). Its tail inverse (Lemma B.4) recovers $\widehat{z} = b + f + \sum_i Q_{s,i} x_i$, with $\Pi_J \widehat{z} = b$ and $\|\widehat{z}\| \leq \|z\|$, under the budget (26) with $d_b = D_J$ and $d_f = \dim F_\epsilon$. Branch spectra and ramps give $M^{1/p} \|T_{b,\tau}(z - \widehat{z})\|^2 \leq \delta^2 (r_\tau^+ + o(1))$, since $\sum_i \|x_i\|^2 \leq \delta^2 v_\delta(b)$; neither the branch bases nor the recovered source depends on the realized task.

Clean nonlinear comparison and the budget limit. Both clean endpoints $z, \widehat{z}$ lie in B_δ and have the same coarse background b, so (15) adds only $(1 + \eta_J)^2$ to the tangent risk bound. Let $J \to \infty$ and target margins tend to zero along the stage schedule of Appendix B.3, with J, each purification, exact block and selector frozen per stage before the threshold controlling all errors and charges at every larger integer budget; as there, $M/m \to 1$. This produces one recovery sequence, not separately chosen encoders at each background. Uniform onset is not assumed.

Rounding and one legal probability object. The global extension and projection of Appendix B.3 preserve clean recovery in B_δ. Section B.4's bit depth gives source error $o(\delta m^{-\mu/k})$. Lemma A.5, with $C_{\mathrm{src}} = 1$, compiles the complete support and its exact law with risk change $o(\delta^2 m^{-1/p})$. Its source-Lipschitz bound applies even when rounding changes b; the same-background secant would not. No enlarged source ball is needed. The same sequence therefore gives legal recovery and the stated Lipschitz and precision bounds; the upper-profile extension and singleton limits of Section B.4 give set convergence and the main-text constants. □

C.2 When geometry-dependent measurements are necessary

A fixed measurement kernel must serve difficult inputs at every background geometry; adaptive measurements need only serve the background seen in the first batch. We quantify this for native task pairs with the prescribed law and detector. A fixed linear encoder satisfies a joint-background constraint: for any fixed finite catalog of backgrounds and (possibly repeated) task labels (b_i, τ_i) with probabilities ν_i,

$$\liminf_m \frac{2m^{1/p}}{\delta^2} \sum_i \nu_i \rho^{\mathrm{nl}}_{m,\tau_i}(L_m) \geq \mathcal{W}\left(\sum_i \nu_i v_\delta(b_i) a_{b_i,\tau_i} \right), \tag{28}$$

where $\mathcal{W}(a) = [k(2\pi)^k]^{-1/p} (\int_{S^* M} a^p \, d\Lambda)^{1/p}$. For finite-harmonic backgrounds choose $J \geq J_{\min}$ with every $b_i \in \Pi_J H$. On H_J, the leading $(m+1)$-dimensional space of the Gram weighted by $\nu_i v_\delta(b_i)$ meets $\ker L_m$ in a unit vector w. Orthogonality places each indistinguishable pair $b_i \pm \delta \sqrt{v_\delta(b_i)} w$ in B_δ; different pairs need not share codes. The parallelogram inequality and (15) bound the weighted task risks below by the weighted quadratic form on w. Weyl asymptotics and a growing background cutoff give (28), and continuity of the cofactor integral extends it to arbitrary backgrounds. The argument uses one kernel for all catalog entries, so it gives no lower bound for adaptive queries.

Proposition C.1 (A task class requiring geometry adaptation)**.** *Keep the native assumptions of Section 3, with spherical base $h_0 = R > 0$ and two fixed distinct weights ω_1, ω_2. Assume antipodal evenness $\omega_i(-u) = \omega_i(u)$ and equal single-task capacities $\int \omega_1^p \, d\sigma = \int \omega_2^p \, d\sigma$. For $\omega_\pi = \pi\omega_1 + (1-\pi)\omega_2$, the function $c(\pi) = \mathsf{h}_{b=0}(\pi, 1-\pi)$ has a unique maximizer $\pi_* \in (0,1)$. Put $\omega_* = \omega_{\pi_*}$, $I_* = \int \omega_*^p \, d\sigma$, $c_* = c(\pi_*)$, and define $d_i = \frac{2(k-1)c_*}{kRI_*} P_H E_r(k - \Delta)(\omega_*^{p-1} \omega_i)$,*

$\kappa_{\mathrm{ad}} = \frac{\pi_*(1-\pi_*)}{4c_*^2}\|d_1 - d_2\|_H^2 > 0$. *For all sufficiently small fixed positive radii,*

$$\liminf_{m\to\infty} \frac{\Lambda_m^{\mathrm{lin,nl}}(\delta)}{\Lambda_m^{\mathrm{alin,nl}}(\delta)} \geq 1 + \kappa_{\mathrm{ad}}\delta^2 + o(\delta^2). \tag{29}$$

The remainder concerns $\delta \downarrow 0$ after the budget limit, for a fixed task pair. The denominator's leading constant, attained by quantized two-batch legal recovery, is the continuous CPPZ constant.

Proof. Write $d = \omega_1 - \omega_2$ and $I(\pi) = \int \omega_\pi^p \, d\sigma$. At the sphere c is a positive constant times $I^{1/p}$, so

$$c''(\pi) = \frac{(p-1)c(\pi)}{I(\pi)^2}\left[I(\pi)\int \omega_\pi^{p-2} d^2 \, d\sigma - \left(\int \omega_\pi^{p-1} d \, d\sigma\right)^2\right] < 0.$$

Cauchy–Schwarz is strict unless d is proportional to ω_π, which would make the normalized weights equal. Balance gives equal endpoint values, hence the unique interior π_*. Its stationarity and degree-one homogeneity of $\mathcal{W}$ imply $D\mathcal{W}(a_*)[a_{0,1}] = D\mathcal{W}(a_*)[a_{0,2}] = c_*$ for $a_* = \pi_* a_{0,1} + (1-\pi_*)a_{0,2}$.

For $v \in H$, let $H_v = \nabla^2 E_r v + E_r v\, g$. Differentiating the actual cofactor and normalizer at the sphere gives $D_b \log[\mathrm{cof}_g(A_b)(\widehat{\xi},\widehat{\xi})]_{b=0}[v] = \frac{\mathrm{tr}\, H_v - H_v(\widehat{\xi},\widehat{\xi})}{R}$ and $DZ_0[v] = R^{k-1}\int (k-\Delta)E_r v \, d\sigma = 0$ by the gauge. Normalized averaging over unit cotangent directions gives $\int H_v(\widehat{\xi},\widehat{\xi})\, d\widehat{\xi} = \mathrm{tr}\, H_v/k$, so differentiation of $\mathcal{W}$ and integration by parts yield $g_\pi := \nabla_H \mathsf{h}_b(\pi, 1-\pi)|_{b=0} = \frac{2(k-1)c(\pi)}{kRI(\pi)} P_H E_r(k-\Delta)(\omega_\pi^p)$. In particular $g_* := g_{\pi_*} = \pi_* d_1 + (1-\pi_*)d_2$. More precisely, $D\mathcal{W}(a_*)[D_b a_{b,i}|_{b=0}[v]] = \langle d_i, v\rangle$: varying only background i in the mixed functional contributes $\pi_i\langle d_i, v\rangle$, with $(\pi_1,\pi_2) = (\pi_*, 1-\pi_*)$.

Positive geometric margins give $\mathsf{h}_b(\pi) = c(\pi) + \langle g_\pi, b\rangle + O(\|b\|^2)$ uniformly in π. On a sufficiently small source ball, choose a common positive lower bound $c_{\min}$ for $\mathsf{h}_b(\pi)$ and a derivative bound L; comparison with $b = 0$ confines backgrounds maximizing $v_\delta(b)\mathsf{h}_b(\pi)$ to $0 \leq L\|b\| - c_{\min}\|b\|^2/\delta^2$, hence $\|b\| = O(\delta^2)$. Completing the square gives, uniformly in π, $L_\delta(\pi) = c(\pi) + \delta^2\|g_\pi\|^2/(4c(\pi)) + O(\delta^4)$, and uniqueness of π_* then implies $H_\delta := \max_\pi L_\delta(\pi) = c_* + \frac{\delta^2}{4c_*}\|\pi_* d_1 + (1-\pi_*)d_2\|^2 + o(\delta^2)$. The positive-radius maximizing mixture need not remain π_*; the two-task evaluation in Appendix D identifies H_δ as the normalized adaptive-linear leading risk.

For fixed information, choose $b_i = \delta^2 d_i/(2c_*)$ and weights π_i in (28); these backgrounds are admissible for all sufficiently small δ. Stationarity above makes the radial penalty $-c_*\sum_i \pi_i\|b_i\|^2/\delta^2$, and expanding the catalog witness gives $\Phi_\delta := \mathcal{W}(\sum_i \pi_i v_\delta(b_i)a_{b_i,i}) = c_* + \frac{\delta^2}{4c_*}\sum_i \pi_i\|d_i\|^2 + O(\delta^4)$. As maximum task risk bounds weighted risk, the optimal fixed-information leading risk is at least Φ_δ. Dividing by H_δ and using $\sum_i \pi_i\|d_i\|^2 - \|\sum_i \pi_i d_i\|^2 = \pi_*(1-\pi_*)\|d_1 - d_2\|^2$ proves (29). For strictness, put $q = \omega_*^{p-1}(\omega_1 - \omega_2)$. Stationarity gives $\int q\, d\sigma = 0$ and evenness removes degree-one harmonics, so $0 \neq q \in H$. Since $k \geq 2$ and $E_r(k-\Delta)$ is injective on this gauge, $d_1 - d_2 \neq 0$. □

The hypotheses are sufficient, not necessary. The weights $\omega_{1,2} = 1 \pm a u_1$, $0 < |a| < 1$, are positive, normalized and balanced but not even; then $\pi_* = 1/2$ and $k - \Delta$ kills $q = 2a u_1$, so this second-order test vanishes without establishing fixed-information optimality at higher orders. Also, (29) is a lower bound, not a matching expansion of the fixed optimum, and its remainder and admissible radius depend on the task pair.

The pair of Appendix F.1 ($\eta = 1/4$), used in the two-batch pilots of Appendix F.9, is a special case: take $k = 2$, $\beta = 3$, $s = 4$, $r = 8$, $Y = \sqrt{15}(u_1^2 - u_2^2)/2$ and $\omega_\pm = (1 \pm \eta Y)^9/N_\eta$, where $0 < |\eta| < 2/\sqrt{15}$ and $N_\eta = \int (1+\eta Y)^9\, d\sigma$. Exchange symmetry gives balance and $\pi_* = 1/2$. Since $\Delta Y = 6Y$ and $\|Y\|_2 = 1$, $\kappa_{\mathrm{ad}} = 324\eta^2/(R^2 7^8) + O(\eta^4)$. For the simpler weights $1 \pm aY$, $0 < |a| < 2/\sqrt{15}$, the exact coefficient is $4a^2/(R^2 7^8)$, only 4.34×10^{-8} at $R = 1$, $a = 1/4$. Strict asymptotic necessity therefore does not predict a useful finite-budget gain or a uniform onset.

D ACQUISITION: A SHARP PAIRWISE BOUNDARY AND ONE GEOMETRY UPDATE

D.1 LINEAR INFORMATION AND ITS UNOBSERVED KERNEL

For bounded linear $L : H \to \mathbb{R}^m$, let $K = \ker L$ and $P = P_{K^\perp}$. Equal-code antipodes and recovery of Pz from Lz give the exact identity

$$\rho^{\tan}_{m,\tau}(L) = \tfrac{\delta^2}{2}\|T_\tau(I-P)\|^2, \qquad \operatorname{rank} P \le m. \tag{30}$$

Conversely any such projection is measured by orthonormal coordinates. An adaptive scheme's zero-input path fixes at most m functionals, and all vectors in their common kernel follow that same path. The same antipodal bound holds taskwise, and projection onto the kernel complement attains it, so deterministic adaptive and nonadaptive linear upper regions coincide. This is the classical zero-answer mechanism (Krieg et al., 2025; Dirksen & Ullrich, 2018); randomization is not covered.

D.2 FROM ANY ADAPTIVE POLICY TO TWO BATCHES

Fix $0 < \delta < \delta_{\mathrm{geom}}$ (Lemma A.7), write $v = v_\delta$ and use H_J, D_J from Appendix C. All cutoffs below satisfy $J \ge J_{\min}$ and $\eta_J < 1/2$ (Lemma A.8); the charged coarse block includes these fixed initial directions, and constants and onset may depend on the fixed geometric margins. More precisely, Theorem 5.1 gives $\epsilon_m = O((\log m/\log\log m)^{-\gamma/k})$ for fixed $0 < \gamma < r - s - 2 - k/2$, using at most m residual spaces. Synchronized rounding on $[-\delta,\delta]$ to $b_M = \lceil(\mu/k{+}1)\log_2(M{+}1)\rceil$ bits per answer preserves its comparison; the proof below counts the coarse queries within M. Lemma A.8 gives a uniform source-to-forecast Lipschitz constant C_F. In (14), Sobolev embedding bounds the relative factor by $L\|b-c\|$; its tail estimate also gives $\|T_{b,\tau}u\| \le (1+\kappa_J)\|T_{\Pi_J b,\tau}u\|$ on $H_{J_{\min}}$, with $\kappa_J = C_1\delta(J+1)^{-\gamma}$.

Lemma D.1 (Whole frozen kernels)**.** *Run an adaptive policy $\mathcal{N}$ at $c \in \Pi_J H$, $\|c\| < \delta$, querying functionals ℓ_i^c ($i \le m$). Let $K_c = H_J \cap \bigcap_{i\le m}\ker \ell_i^c$ and $V_c = H_J \ominus K_c$. Then $\dim V_c \le m$ and*

$$\|T_{c,\tau}|_{K_c}\| \le \frac{e_\tau(\mathcal{N})}{\delta\sqrt{v(c)}(1-\eta_J)}. \tag{31}$$

Proof. For unit $u \in K_c \subset H_J$, the inputs $c \pm \delta\sqrt{v(c)}u$ have the same answers as c at every query, and half their forecast separation lower-bounds one decoder error. Equation (15) and then the supremum over u give (31), also for stopping rules and rounded paths. No whole-source relative coercivity is used. □

Proof of Theorem 5.1. Inward net. Fix $J \ge J_{\min}$ and $0 < \theta \le 1/4$, and put $h = \delta\theta^2/16$. A maximal h-separated set $\{c_j\}_{j=1}^K$ in $(1-\theta)B_J$ is an h-net, with $K \le (1+32/\theta^2)^{D_J}$ by disjoint radius-$h/2$ balls. If the received coarse value $\widetilde{b} \in B_J$ satisfies $\|\widetilde{b} - b\| \le h$, choose the first center $c_{s(\widetilde{b})}$ within h of $(1-\theta)\widetilde{b}$. Since $v((1-\theta)b) \ge 2\theta - \theta^2 \ge \theta$, $v(b) \le v((1-\theta)b)$ and v is $2/\delta$-Lipschitz,

$$\|b - c_j\| \le 2\delta\theta, \qquad v(b)/v(c_j) \le (1-\theta/4)^{-1}. \tag{32}$$

An ordinary nearest-center net does not control this energy ratio near $v = 0$.

Frozen branch spaces. At every center freeze the reference policy and form $V_j = V_{c_j}$ from Lemma D.1. Measure first the D_J coarse coordinates, then an orthonormal basis of the selected $V_j \subset H_J$. The clean source $z^* = b + P_{V_j}(z-b)$ stays in $B_\delta(H)$, preserves b and has $z - z^* \in K_{c_j}$ with norm at most $\delta\sqrt{v(b)}$. Relative backgrounds, (31), (32) and the true-background tail secant give $\|F_\tau(z) - F_\tau(z^*)\|_2 \le B_{J,\theta}e_\tau(\mathcal{N})$, $B_{J,\theta} = (1+\eta_J)(1+2L\delta\theta)/\big[(1-\eta_J)\sqrt{1-\theta/4}\big]$. No reference decoder or unmeasured residual energy is used.

Noisy routing and one legal law. For coordinate errors bounded by ν with $\sqrt{D_J}\nu \le h$, acquisition and decoding both route safely using the received coarse vector projected onto B_J, through the common interface $\widetilde{b} = P_{B_J}(\text{received coarse})$, $j = s(\widetilde{b})$, $y = \text{received } V_j^* z$, $\widehat{z} = P_{B_\delta(H)}(\widetilde{b} + V_j y)$. Compare $\widehat{z}$ with the clean z^* in *this selected branch*, not the noiselessly selected one. Orthonormal

synthesis and nonexpansive projection give source error at most $(\sqrt{D_J}+\sqrt{m})\nu$. First-index routing is Borel, so Lemma A.5's perturbation bound gives one legal law with $e_\tau(\widetilde{\mathcal{N}};\nu) \le B_{J,\theta}e_\tau(\mathcal{N}) + C_F(\sqrt{D_J}+\sqrt{m})\nu + \epsilon_{\mathrm{pol}}$, where ϵ_{pol} is its filter forecast error.

Schedule. Choose $\theta_J = \min\{1/4, (J+1)^{-\gamma}\}$ and, for all large m, the largest $J_m \ge J_{\min}$ with $D_J \log(1 + 32/\theta_J^2) \le \log m$. Then $J_m \asymp (\log m/\log\log m)^{1/k}$, $d_m = D_{J_m} = O(\log m/\log\log m)$, $K \le m$ and $B_{J_m,\theta_{J_m}} = 1 + O(J_m^{-\gamma})$; the fixed initial cutoff changes only onset, not these rates. The stated bit count b_M gives $\nu \le C\delta 2^{-b_M}$, so ν eventually meets the coarse clearance and adds $O(\delta M^{-\mu/k-1/2})$; choose $\epsilon_{\mathrm{pol}} = O(\delta M^{-\mu/k-1})$. By (31) at zero with a fixed admissible J and the singular-value minimax principle, each reference policy has $e_\tau(\mathcal{N}) \ge c\delta(m+1)^{-\mu/k}$, so the additive term is relatively lower order; its squared/cross terms are lower order on rate-optimal policies.

Budget comparison. For limsup comparison at every total budget M, simulate a near-minimizer at $q = M - D_{J_M}$; $q/M \to 1$. For liminf, simulate a near-minimizing subsequence m_j at $m_j + D_{J_{m_j}}$. Optimization slack is $o(\delta^2 m^{-1/p})$, and class inclusion gives the reverse inequalities. The argument also holds for synchronized quantized schemes and does not assume an optimal limit exists. Ignored unit-norm queries can pad budgets. The schedule counts candidate spaces only, not their vector descriptions or the cost of finding the reference policy, and gives no useful finite onset; for a prescribed inflation ϵ it instead permits $D_J = O(\epsilon^{-k/\gamma})$ and $K \le \exp[O(\epsilon^{-k/\gamma}\log(1/\epsilon))]$. □

D.3 Finite three-query adaptation in the native model

We prove Proposition 5.2 on one four-dimensional source ball for both information classes. The observed geometry rotates the expensive residual direction; this mechanism does not use a small-radius secant approximation.

Model and legal reconstruction. For $u \in S^2$ and $r_u^2 = |u|^2$, define the solid harmonics $Y_1 = \sqrt{15}(u_1^2 - u_2^2)/2$, $Y_2 = \sqrt{15}u_1u_2$ and $Y_{3,4} = \sqrt{165}\,u_{1,2}(21u_3^4 - 14u_3^2r_u^2 + r_u^4)/8$. Their degrees are $2, 2, 5, 5$, and they are orthonormal in $L_2(\sigma)$. For $z = (b, x) \in B_\delta(\mathbb{R}^2 \oplus \mathbb{R}^2)$, set

$$\delta = 300, \qquad h_z = 1 + 7^{-4}(b_1Y_1 + b_2Y_2) + 31^{-4}(x_1Y_3 + x_2Y_4). \tag{33}$$

This is $1 + E_8 z$ in the original orthonormal source metric, with all other directions fixed to zero. Let $F(z) = K_3 p_{h_z}/8$ and $\omega_\pm = 1 \pm (u_1^2 - u_2^2)/1000$; the weights are positive, smooth and normalized. The kernel of K_3 is at most $1 + 3\sum_{\ell\ge1}\ell^{-2} = 1 + \pi^2/2 < 8$. Thus the detector and half-Brier risk are those of Section 3, with $k = 2, \beta = 3, s = 4, r = 8, \tau_s = 1, M_\beta = 8$. The full degree-2 and degree-5 addition identities bound the four-member subset. For $S = (5/7^8 + 11/31^8)^{1/2}$ and $C_H^2 = 35/7^8 + 11(3\cdot30^2 - 10\cdot30 + 8)/(8\cdot31^8)$, Cauchy–Schwarz gives $h_z \ge 1 - \delta S > 0.7206$ and $A_{h_z} \succeq (1 - \delta C_H)g > 0.2605g$ on the entire closed ball, using $\sum_j(\nabla^2 Y_{\ell j}(v,v) + Y_{\ell j})^2 = (2\ell+1)(3\lambda_\ell^2 - 10\lambda_\ell + 8)/8$ for unit tangent v, where $\lambda_\ell = \ell(\ell+1)$. In particular $J_{h_z} > 1/16$. For the solid support H_z, the coefficient ℓ_1 norm is at most $1 + 300(8/7^4 + 117/31^4) < 3$, and the entries of the matrix B_{H_z} of Lemma A.4 have coefficient norm below 78. Its normal cofactor $\widehat{J}_z = \widehat{J}_{H_z}$ has degree at most 12 and $\sup_{[-1,1]^3}|\nabla\widehat{J}_z| \le 12\cdot18\cdot78^2 < 2^{21}$. Lemma A.4 with $j_- = 1/16$ and $M_J = 2^{21}$ therefore gives the single uniform extension

$$R_z(u) = \widehat{J}_z(u) + 2^{45}(|u|^2 - 1)^2 \ge 1/32. \tag{34}$$

Cube-normalizing R_z, the sphere constraint, inverse-Gauss readout and structural normalization give exactly the law $J_{h_z}\sigma/Z_{h_z}$. This compiler applies to every recovered source in the ball; the factored penalty vanishes on the carrier and is not an independent density correction. These geometric estimates concern the restricted four-mode ball, not the full infinite source.

Exact response identities. Polarizing the determinant and subtracting its mean gives

$$F(z) - F(0) = \frac{Bz + C(z,z)}{s(z)}, \quad s(z) = 1 - d_b|b|^2 - d_x|x|^2, \quad d_b = 2/7^8, \quad d_x = 14/31^8. \tag{35}$$

There is no Taylor remainder. Here $B = (B_b, B_x)$ includes the detector and C is symmetric bilinear. The blocks of B satisfy $B_b^* B_b = \lambda_b I$ and $B_x^* B_x = \lambda I$, with $\lambda_b = 1/(4 \cdot 7^{11})$ and $\lambda = 49/(4 \cdot 31^{11})$. The terms $B_b b, C(b,b), C(x,x)$ are even, and $B_x x + 2C(b,x)$ is odd. For $J(b)x = 2C((b,0),(0,x))$ and $A(b) = B_x + J(b)$, exact sphere moments give

$$M(b) = -\begin{pmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{pmatrix}, \qquad A(b)^* A(b) = \lambda I + aM(b) + q|b|^2 I, \quad J(b)^* J(b) = q|b|^2 I,$$
$$a = \frac{25\sqrt{15}}{97111198698059264082}, \qquad q = \frac{1162618225}{126129235502560901588915302783 6}. \tag{36}$$

The eigenvalues are $\lambda \pm a|b| + q|b|^2$. The Frobenius norms c_b, c_x, c_T of the low–low, high–high and complete flattened tensors of C bound their bilinear operator norms, and

$$c_b^2 = \frac{6071}{53620402951995264672}, \qquad c_T^2 = c_b^2 + c_x^2 + 2q,$$
$$c_x^2 = \frac{1769422129040731837}{3103999287749715033744102997047032061684216}. \tag{37}$$

Orthonormality, (36) and (37) are reproducible by finite polynomial arithmetic in rationals and $\sqrt{15}$. For solid harmonics P_i, P_j of degrees ℓ_i, ℓ_j, determinant polarization on the sphere is the restriction of $\mathcal{Q}(P_i, P_j)$, where

$$2\mathcal{Q}(P_i,P_j) = (\ell_i{-}1)(\ell_j{-}1)(\ell_i{+}\ell_j{+}2)P_iP_j + 2(\ell_i{-}1)(\ell_j{-}1)|u|^2 \nabla P_i \cdot \nabla P_j - |u|^4 \operatorname{tr}(\nabla^2 P_i \nabla^2 P_j),$$

which follows by polarizing $2\det T = (\operatorname{tr} T)^2 - \operatorname{tr} T^2$ on the tangent plane. Its mean on a unit degree-ℓ harmonic is $1 - \ell(\ell+1)/2$, yielding d_b, d_x and the constant removed from C. For homogeneous P of degree n, the top harmonic part is $\sum_j \alpha_j |u|^{2j} \Delta_E^j P$, where $\Delta_E = \sum_i \partial_i^2$, $\alpha_0 = 1$ and $\alpha_j = -\alpha_{j-1}/[2j(2n-2j+1)]$; subtract it, divide by $|u|^2$, and repeat. Odd sphere moments vanish and $\int u_1^{2i} u_2^{2j} u_3^{2l}\, d\sigma = (2i{-}1)!!(2j{-}1)!!(2l{-}1)!!/[2(i{+}j{+}l){+}1]!!$. After decomposition, multiply each degree-ℓ squared inner product by $[64(1+\ell(\ell+1))^3]^{-1}$.

Three received answers: uniform upper bound. Let $N = 2^{32} - 1$ and $\nu = \delta/N$. Measure b_1, b_2, round to 2^{32} uniform levels in $[-\delta, \delta]$, and project the received pair onto $B_\delta(\mathbb{R}^2)$ to obtain $\widehat{b}$, so $|b - \widehat{b}| \le e_b = \sqrt{2}\nu$. Choose a unit eigenvector $u(\widehat{b})$ of $M(\widehat{b})$ for $-|\widehat{b}|$, with a fixed tie rule, and a perpendicular unit $v(\widehat{b})$. The third query is $v(\widehat{b})^\top x$, rounded identically to $\widehat{p}$. The decoder reconstructs v from the first two words, synthesizes $(\widehat{b}, v\widehat{p})$, projects onto $B_\delta(\mathbb{R}^4)$, and compiles (34). All queries have unit norm; no branch, norm or source-dependent side message is sent.

Compare z with $y = (b, pv)$ in this received branch, where $p = v^\top x$ and $x = pv + eu$; both y, z are in the ball. Put $s_- = 1 - d_b\delta^2$, $s_x = 1 - d_x\delta^2$, $d = d_b - d_x$ and $t = |b|$. Since $\|M(b) - M(\widehat{b})\| = |b - \widehat{b}|$, we have $u^\top M(b) u \le -t + 2e_b$. Using $s(z) \ge s_x - dt^2$ gives $|e|^2 \|A(b)u\|^2 / s(z)^2 \le \lambda\delta^2/s_x^2 + 2a\delta^2 e_b / s_-^2$, because, after dropping the favorable $-at$ term,

$$\frac{q\delta^2}{\lambda} + \frac{2d\delta^2}{s_x} \le 1 \tag{38}$$

implies $(1-w)(1+(q\delta^2/\lambda)w) \le (1-(d\delta^2/s_x)w)^2$ for $w = t^2/\delta^2 \in [0,1]$. The exact denominator correction is $s(y) - s(z) = d_x e^2$, and $\|xx^\top - (pv)(pv)^\top\|_F^2 = 2p^2e^2 + e^4 \le \delta^4$, so the odd and even root errors are at most

$$O = \left(\frac{\lambda\delta^2}{s_x^2} + \frac{2a\delta^2 e_b}{s_-^2}\right)^{1/2} + \frac{d_x\delta^3\sqrt{\lambda + a\delta + q\delta^2}}{s_-^2},$$
$$E = \frac{c_x\delta^2}{s_-} + \frac{d_x\delta^2}{s_-^2}\left[\sqrt{\lambda_b}\delta + (c_b + c_x)\delta^2\right].$$

Parity bounds the unweighted root error by $\sqrt{O^2 + E^2}$. Differentiating (35) bounds the global Lipschitz constant by $L_F = (\sqrt{\lambda_b} + 2\delta c_T)/s_- + 2\delta^2 d_b(\sqrt{\lambda_b} + \delta c_T)/s_-^2$. Orthonormal synthesis and final nonexpansive projection bound the received-source error by $\sqrt{3}\nu$. Since $\omega_\tau \le 1.001$,

$$U_{\mathrm{two},32}(3) \le \frac{1.001}{2}\left[\sqrt{O^2 + E^2} + L_F\sqrt{3}\nu\right]^2 < \frac{501}{500}\rho_0, \qquad \rho_0 = \frac{\lambda\delta^2}{2}. \tag{39}$$

No continuity across an eigenvector-routing switch is used.

Every fixed measurement space: lower bound. Fix any rank-at-most-three $L : \mathbb{R}^4 \to \mathbb{R}^3$, with optimal worst-task half-risk $R(L)$, and unit kernel vector $u = (u_b, u_x)$, and write $s^2 = |u_b|^2$, $r_x^2 = 1 - s^2$. A task's minimax half-risk on any indistinguishable pair is at least one eighth of its squared forecast separation, for arbitrary nonlinear decoders. Since $(\omega_+ + \omega_-)/2 = 1$, maximum task risk bounds the unweighted chord risk. For $z_\pm = \pm\delta u$, the quadratic numerator cancels and the equal normalizers are at most one:

$$R(L) \geq \rho_0[1 + (\lambda_b/\lambda - 1)s^2]. \tag{40}$$

This is at least $(27/25)\rho_0$ when $s^2 \geq x_0 = (2/25)/(\lambda_b/\lambda - 1)$. Otherwise put $w = u_x/r_x$ and $e = -(w_1^2 - w_2^2, 2w_1w_2)$, so $M(e)w = w$, and set $\beta_0 = 100$, $\alpha = e \cdot u_b$, $b^\circ = \beta_0((e, 0) - \alpha u)$, $T = (\delta^2 - \|b^\circ\|^2)^{1/2}$ and $z_\pm = b^\circ \pm Tu$. The background is perpendicular to u, has norm at most β_0, and $T \geq \sqrt{\delta^2 - \beta_0^2}$; both inputs are in the ball and have the same measurements. Let $h_\beta = \lambda + a\beta_0 + q\beta_0^2$. Their odd numerator half-difference and mean are $N_d = r_xT[A(\beta_0 e)w - 2\beta_0\alpha J(u_b)w]$, $N_a = r_x[-\beta_0\alpha A(\beta_0 e)w + (T^2 + \beta_0^2\alpha^2)J(u_b)w]$. Write their denominators $s_0 \pm \nu_0$. Then $0 < s_0 \leq 1$, $s_0^2 - \nu_0^2 \geq s_-^2$ and $|\nu_0| \leq 2T(d_b - d_x)\beta_0 s$. The odd forecast half-difference is $(s_0N_d - \nu_0N_a)/(s_0^2 - \nu_0^2)$. Using $s_0/(s_0^2 - \nu_0^2) \geq 1$, $\|A(\beta_0 e)w\| = \sqrt{h_\beta}$ and $\|J(u_b)w\| = \sqrt{q}\, s$, its norm is at least

$$\sqrt{\delta^2 - \beta_0^2}\sqrt{1 - s^2}(\sqrt{h_\beta} - Ks^2), \quad K = 2\beta_0\sqrt{q} + \frac{2(d_b - d_x)\beta_0}{s_-^2}\left[\beta_0\sqrt{h_\beta} + \sqrt{q}(\delta^2 + \beta_0^2)\right].$$

The factor $\sqrt{h_\beta} - Ks^2$ is positive through $s^2 = x_0$, and the squared expression decreases there. Dropping the even output is safe; hence

$$\frac{R(L)}{\rho_0} \geq \left(1 - \frac{\beta_0^2}{\delta^2}\right)(1 - x_0)\frac{(\sqrt{h_\beta} - Kx_0)^2}{\lambda} > 1.089708745 > \frac{27}{25}. \tag{41}$$

Together with (40), this covers every fixed map, including mixed low/high measurements.

Arithmetic and scope. All constants use exact sphere moments. Enclose positive square roots by rational endpoints on a 10^{-70} grid, check their squared endpoints with integer arithmetic, and propagate rational intervals. This gives the left side of (38) below 0.234516, $x_0 < 3.050616 \times 10^{-7}$, $\sqrt{h_\beta} - Kx_0 > 0$, and upper and lower ratios below 1.001144513 and above 1.089708745. The conservative gap is at least $1 - (501/500)/(27/25) = 13/180$, not an exact optimal ratio. For the full response with $\omega = 1$, the factor 1.001 disappears, the upper ratio is below 1.000144369, and the same lower proof applies. Thus the gap does not require task disagreement and does not contradict linear-target information results (Krieg et al., 2025). The scale is $\rho_0 \simeq 2.16955 \times 10^{-11}$. The proposition assumes exact algebraic query directions, nearest-level answer rounding and mathematical decoding; its 96 answer bits exclude model and compiler storage, and it does not certify every floating-point operation, sensor noise or hardware arithmetic. It does not resolve the 32-coordinate pilots of Appendix F.9, their five-query exclusions (Table 11) or the full infinite-source finite-budget question. Appendix F.6 extends these bounds to other radii and answer precisions.

Linearized control: spectral imbalance without an acquisition gap. Keep the same ball $B_\delta(\mathbb{R}^4)$, task weights and three-coordinate budget, but replace the forecast by $F_{\tan}(z) = F(0) + Bz$. Parity and (35) give $B^*B = \mathrm{diag}(\lambda_b, \lambda_b, \lambda, \lambda)$ with $\lambda_b > \lambda$. Since $(\omega_+ + \omega_-)/2 = 1$, the equal-weight task mixture has this Gramian, with fourth eigenvalue λ, so (19) bounds every continuous three-coordinate code below by ρ_0. For deterministic adaptive linear queries, the zero-input-path argument after (30) gives the same lower bound. The fixed measurements $b_1, b_2, (x_1 - x_2)/\sqrt{2}$ leave only $w = (0, 0, 1, 1)/\sqrt{2}$ unresolved. Since $Bw = -\sqrt{\lambda}(Y_3 + Y_4)/\sqrt{2}$, both task energies equal λ: exchanging u_1, u_2 reverses $u_1^2 - u_2^2$ and preserves $(Y_3 + Y_4)^2$, so the weight perturbation contributes zero. Decoding the orthogonal projection of z onto $w^\perp$ attains ρ_0 for both tasks. Thus, with exact real answers, $\Lambda_3^{\mathrm{lin,tan}} = \Lambda_3^{\mathrm{alin,tan}} = \Lambda_3^{\mathrm{cont,tan}} = \rho_0$. Together with (39)–(41), this excludes low/high-mode imbalance alone as the explanation of the native gap. Linearization removes both quadratic response and nonlinear normalization; it neither isolates one of these terms nor preserves the full nonlinear area law.

D.4 THE TWO-TASK EXCEPTION, INCLUDING REAL BOUNDARY CASES

Lemma D.2 (Two real forms)**.** *For real symmetric $Q_1, Q_2 \in \mathbb{R}^{N\times N}$, a d-dimensional real subspace nonpositive for both exists exactly when $n_{\le 0}(\theta Q_1 + (1-\theta)Q_2) \ge d$ for all $0 \le \theta \le 1$.*

Proof. Necessity is the inertia variational principle. *Generic pairs.* First take nonsingular forms with $Q_1^{-1}Q_2$ having distinct complex eigenvalues, a dense set since the excluded determinants/discriminants are nonzero polynomials. Selfadjointness in the Q_1 bilinear form makes distinct generalized eigenspaces orthogonal for both forms. A generalized eigenvector z with $z^\top Q_1 z = 0$ would be orthogonal to the entire eigenbasis, contradicting nonsingularity. Real eigenvalues therefore give scalar blocks (a, b) with $ab \neq 0$. A nonreal pair gives, in a real basis, $Q_1 = \mathrm{diag}(1, -1)$, $Q_2 = \left(\begin{smallmatrix} \alpha & \beta \\ \beta & -\alpha \end{smallmatrix}\right)$, $\beta \neq 0$, with pencil determinant $-[(\theta + (1-\theta)\alpha)^2 + (1-\theta)^2\beta^2] < 0$. Each such block contributes one negative eigenvalue throughout and has a common strictly negative direction $(\cos t, \sin t)$: choose $(\cos 2t, \sin 2t)$ opposite the sum of its two distinct nonopposite unit normals.

Scalar blocks and matching. Scalar $(--)$ blocks contribute one common direction and $(++)$ none. Mixed types $P : a > 0 > b$ and $N : a' < 0 < b'$ have roots $u = -b/(a-b)$ and $v = b'/(b'-a')$, and their pencil values are negative for $\theta < u$ and $\theta > v$, respectively. A pair gives the common negative vector $e_P + \sqrt{t}e_N$ exactly when $a/(-a') < t < (-b)/b'$, and such t exists exactly when $v < u$. In the bipartite graph joining P to N when $v < u$, maximum matching size equals $\min_\theta \big(\#\{P : u > \theta\} + \#\{N : v < \theta\}\big)$, and each threshold set is a vertex cover. Conversely any cover leaves all uncovered P roots below all uncovered N roots, so a threshold between them gives a contained cover (use an endpoint for an empty side). Matching–cover equality follows by alternating paths from unmatched P vertices: if Z_P, Z_N are the P and N vertices these paths reach, $(P \setminus Z_P) \cup Z_N$ covers all edges and contains exactly one vertex per matched edge because no augmenting path exists.

Assembly and general pairs. Choose one direction per matched pair, negative scalar block and nonreal block. Their supporting blocks are mutually orthogonal for both forms, so the common negative dimension is $\min_\theta n_-(\theta Q_1 + (1-\theta)Q_2)$. For a general pair, subtract ϵI from both and perturb by less than $\epsilon/2$ to a generic pair; its common d-plane has original forms at most $3\epsilon I/2$. Grassmann compactness as $\epsilon \downarrow 0$ gives a common nonpositive plane, including singular pencils and repeated roots. □

This step uses classical common-sign and pencil geometry (Johnson & Rodman, 1985). Complex scalar compression (Li & Sze, 2008, Theorem 2.2) and common neutral subspaces (Rodman, 2008, Section 1) impose equalities, whereas risk feasibility here requires real nonpositive inequalities. No priority is claimed for the matrix primitive; its role is to identify all three information-class risk regions in (4).

Proof of (4). For $q_i = 2r_i/\delta^2$, mixture lower bounds follow from the same antipodal argument as (19), using only compactness. In finite dimension $N > m$, apply Lemma D.2 to $Q_i = A_i - q_i I$, $d = N - m$, to obtain a common nonpositive d-plane. The projection P onto its orthogonal complement has $\mathrm{rank}\, P \le m$ and $\|T_i(I-P)\|^2 \le q_i$; decode $T_i P z$. If $N \le m$, measure the whole source.

For infinite separable H, compress to increasing finite-dimensional H_N with dense union. Compression cannot increase a pencil's $(m+1)$st eigenvalue, so choose measured m-planes $V_N \subset H_N$. Along a subsequence their orthonormal bases converge weakly, $v_{N,j} \rightharpoonup v_j$; put $V = \mathrm{span}\{v_j\}_{j\le m}$. For $x \perp V$, $x_N = P_{H_N}x - P_{V_N}P_{H_N}x \to x$ strongly because the subtracted V_N component has finitely many coefficients $\langle x, v_{N,j}\rangle$, all tending to zero. In the limit both residual inequalities hold on $V^\perp$, proving the claim. No orthonormality of weak limits or compactness of rank-m projections is assumed. The kernel argument identifies adaptive linear and nonadaptive regions, and linear encoding is continuous, giving all three equalities, zero targets included. □

D.5 STATIC PROFILES AND THE TWO-TASK CONSEQUENCE

Fix a finite native task family F and $0 < \delta < \delta_{\mathrm{geom}}$. For $r \in (0,\infty)^F$, put

$$G_m(r) = \sup_{b \in B_\delta(H)} \inf_{\substack{P=P^*=P^2 \\ \operatorname{rank} P \le m}} \max_{\tau \in F} \frac{m^{1/p} v(b) \|T_{b,\tau}(I-P)\|^2}{r_\tau}. \tag{42}$$

The projection may depend on the background, but not on the residual input or realized task; boundary backgrounds contribute zero. Let $\mathcal{A}_\delta^{\mathcal{I}}$ denote the positive profiles admitting policies in class $\mathcal{I}$ with $\limsup_m (2/\delta^2) m^{1/p} \rho_{m,\tau}^{\mathrm{nl}} \le r_\tau$ for every task, using the indicated recovery restriction.

The coarse prefix can be charged without changing the limiting criterion. For an orthogonal Q, set $V = P_{H_J} \operatorname{ran} Q$ and $P^{[J]} = \Pi_J + P_V$. Then $\operatorname{rank} P^{[J]} \le D_J + \operatorname{rank} Q$ and $\ker P^{[J]} = H_J \cap (\operatorname{ran} Q)^\perp$, so every task residual norm for $P^{[J]}$ is at most that for Q. If $G_m^{[J]}$ adds $\Pi_J H \subseteq \operatorname{ran} P$ to (42), this proves $G_m(r) \le G_m^{[J]}(r) \le (m/(m-D_J))^{1/p} G_{m-D_J}(r)$ for $m > D_J$. For any P containing $\Pi_{J_{\min}} H$, truncate b to $\Pi_J b$ and select its inward center c_j as in Section D.2. With $\alpha_{J,\theta} = (1+\kappa_J)^2 (1+2L\delta\theta)^2/(1-\theta/4) \to 1$ as $J \to \infty$, $\theta \to 0$, relative variation and (32) give

$$v(b)\|T_{b,\tau}(I-P)\|^2 \le \alpha_{J,\theta} v(c_j) \|T_{c_j,\tau}(I-P)\|^2. \tag{43}$$

The prefix condition is essential: all relative comparisons apply to tail residuals, not arbitrary low modes.

Theorem D.3 (Static characterization)**.** *Under the native assumptions, for every fixed finite F and $0 < \delta < \delta_{\mathrm{geom}}$,*

$$\mathcal{A}_\delta^{\mathrm{alin}} = \mathcal{A}_\delta^{\mathrm{2batch,CPPZ}} = \{r > 0 : \limsup_m G_m(r) \le 1\} \subseteq C_\delta.$$

The second class uses one common legal CPPZ and the synchronized answer precision of Theorem 5.1. No convexity or closed-form solution of the static multi-task problem is asserted.

Proof. For necessity at static budget m, take $J = J_m$ from the two-batch schedule and run a recovering policy $\mathcal{N}_q$ with $q = m - D_J$ queries. At $c = \Pi_J b$ its frozen kernel supplies V_c and $P = \Pi_J + P_{V_c}$ of rank at most m. For $\|c\| < \delta$, $v(b) \le v(c)$, relative variation and (31) imply $v(b)\|T_{b,\tau}(I-P)\|^2 \le (1+\kappa_J)^2 e_\tau(\mathcal{N}_q)^2/[\delta^2(1-\eta_J)^2]$; if $\|c\| = \delta$ then $b = c$ and the left side vanishes. Multiply by $m^{1/p}/r_\tau$, take the supremum over b and let $q/m \to 1$, $\kappa_J, \eta_J \to 0$ to obtain $\limsup_m G_m(r) \le 1$.

Conversely, at total budget M let $J = J_M$ and $n = M - D_J$, and choose at each inward center a rank-n orthogonal Q_j with $v(c_j)\|T_{c_j,\tau}(I-Q_j)\|^2 \le n^{-1/p}(G_n(r)+1/n) r_\tau$ for all τ. Measure Π_J and then $V_j = P_{H_J} \operatorname{ran} Q_j$, using at most M queries; the prefix conversion preserves these residual bounds. Equations (43) and (15) give squared forecast error at most $\delta^2 n^{-1/p}(1+\eta_J)^2 \alpha_{J,\theta_J}(G_n(r)+1/n) r_\tau$. The received-word routing, nonexpansive source projection and polynomial compiler of the two-batch proof add lower-order error at its stated precision without changing the branch at the decoder. Since $n/M \to 1$, this attains r through one legal CPPZ. Class inclusion gives the adaptive-linear inner inclusion, and (27) the outer inclusion in C_δ. □

One or two tasks. For $|F| \le 2$ and $r \in C_\delta$, the singleton bounds make $r > 0$. Fix a margin $\xi > 0$, cutoff J and inward catalog before the spectral limit. The constraints $\pi \cdot r \ge v(c_j) \mathsf{h}_{c_j}(\pi)$ and mixture-uniform Weyl convergence on H_J give the finite pencil inequalities for $r + \xi \mathbf{1}$. Theorem 5.3 supplies common tail projections at every center; finite coarse deletion leaves the Weyl constants unchanged. Prepend the charged Π_J and apply (43). Letting $J \to \infty$ and $\theta, \xi \to 0$ along fixed-stage thresholds with $D_J = o(m)$ gives $\limsup_m G_m(r) \le 1$. Theorem D.3 then proves the nonlinear assertion of Theorem 5.3, including common legal recovery and quantization.

D.6 COMMON INFORMATION AND COLLECTIVE OBSTRUCTION

No restriction to cellwise, spectral or nonadaptive measurement spaces is imposed.

Lemma D.4 (Common residual budget). *Let $A_\tau = \operatorname{diag}(d_{\tau l})$ be a fixed finite nonempty family of positive compact diagonal operators on a real Hilbert space; finite dimension is allowed. Set $L_l = \min_\tau d_{\tau l}$ and $U_l = \max_\tau d_{\tau l}$. For $q > 0$, every rank-at-most-m orthogonal projection with $\|A_\tau^{1/2}(I-P)\|^2 \le q$ for all tasks satisfies*

$$m \ge \sum_l f_q(U_l, L_l), \qquad f_q(U,L) = \begin{cases} 1, & L > q, \\ (U-q)/(U-L), & L \le q < U, \\ 0, & U \le q. \end{cases} \tag{44}$$

In particular $f_q(U,q) = 1$ when $U > q$, whereas $f_q(q,q) = 0$.

Proof. First take q outside the countable set of diagonal values. Let $C = \operatorname{span}\{e_l : L_l > q\}$, $J = C^\perp$ and $K = \ker P$; compactness makes C finite dimensional. All $A_\tau - qI$ are positive on C and nonpositive on K, hence $K \cap C = 0$. The projected kernel $K_J = P_J K$ is closed because $K + C$ is closed. Deleting the C component decreases each excess form, so every task remains at most qI on K_J. Counting in the finite-dimensional quotient gives $\operatorname{codim}_J K_J = \operatorname{codim}_H K - \dim C \le m - \dim C$: the common high-energy space is charged, not discarded.

On J, $B = qI - \operatorname{diag}(L_l)$ is positive and boundedly invertible, since compactness and generic $q > 0$ keep its spectrum away from zero. The congruence $u = B^{1/2}x$ preserves codimension and sends K_J to K', on which $D'_\tau = \operatorname{diag}((d_{\tau l} - L_l)/(q - L_l))$ has compression at most I; equivalently, ${D'_\tau}^{1/2}(I - P'){D'_\tau}^{1/2} \preceq I$ for $P' = P_{(K')^\perp}$. Testing e_l and maximizing over tasks gives $\langle P' e_l, e_l\rangle \ge \left[1 - \frac{q - L_l}{U_l - L_l}\right]_+$ when $U_l > L_l$; equal coordinates have no excess on J. Summing the diagonal of P' and adding $\dim C$ proves (44). For arbitrary $q > 0$, take generic $q' \downarrow q$; only finitely many U_l exceed $q/2$, so termwise convergence proves the stated threshold conventions. □

For the three-cell spectrum $d_{\tau,(j,n)} = w_{\tau j}\lambda_n$, with own-cell weight one and other-cell weight (leakage) $\epsilon \in (0,1)$, the lemma gives the finite bound used in Figure 14 (right):

$$m \ge 3\sum_n \phi_\epsilon(q/\lambda_n), \qquad \phi_\epsilon(t) = \min\{1, [(1-t)/(1-\epsilon)]_+\}. \tag{45}$$

For a positive normalized target, put $U = \max_\tau a_\tau/r_\tau$ and $L = \min_\tau a_\tau/r_\tau$. The all-linear necessary condition proved below is

$$\frac{1}{k(2\pi)^k}\int_{S^*M} \Psi_p(U,L)\, d\Lambda \le 1, \qquad \Psi_p(U,L) = \int_0^1 [(1-t)L + tU]^p\, dt. \tag{46}$$

For fixed $0 < L \le U$ and $\lambda_n \sim Cn^{-1/p}$,

$$\sum_n f_q(U\lambda_n, L\lambda_n) \sim C^p q^{-p}\Psi_p(U,L), \qquad q \downarrow 0. \tag{47}$$

For exact powers, set $x = n(q/C)^p$ and split the Riemann integral at L^p, U^p; it equals $[U^{p+1} - L^{p+1}]/[(p+1)(U-L)]$, with value U^p at $U = L$. For asymptotic spectra, remove a finite prefix, sandwich the tail between $(C \pm \varepsilon)n^{-1/p}$, and let $\varepsilon \downarrow 0$.

Lemma D.5 (Common-information rank condition). *Under Section 4's scalar-principal assumptions, fix a finite nonempty task family. Suppose rank-at-most-m projections obey $m^{1/p}\|T_\tau(I - P_m)\|^2 \le r_{m,\tau}$ with $r_m \to r < \infty$. Then $r_\tau > 0$ and (46) holds. For native nonlinear deterministic adaptive linear queries with $\limsup_m (2/\delta^2) m^{1/p}\rho^{\mathrm{nl}}_{m,\tau} \le r_\tau < \infty$, the same condition holds at every background $b \in B_\delta(H)$ after replacing a_τ by $v_\delta(b)a_{b,\tau}$. This assertion uses the ordinary metric cosphere and holds for each fixed $0 < \delta < \delta_{\mathrm{geom}}$.*

Proof. The positive singleton Weyl constants imply $r_\tau > 0$. Fix $\xi > 0$ and normalize each Gram by $r_\tau + \xi$; its squared residual is eventually at most $q_m = m^{-1/p}$. Apply (21) to this finite normalized family, with $b_{\tau j} = a_\tau(\theta_j)/(r_\tau + \xi)$ on the reference cosphere. On $V = \widehat{\bigoplus}_j V_j$ the measured kernel has codimension at most m, and the lower form comparison gives simultaneous residual at most q_m for the diagonal model with branch energies $(1-\nu)b_{\tau j}B_j$. Lemma D.4 and (47) therefore imply

$1 \geq (1-\nu)^p \sum_j C_j^p \Psi_p(\max_\tau b_{\tau j}, \min_\tau b_{\tau j})$ after the fixed-stage budget limit. The exact branch counts in Lemma B.3 give cosphere Riemann sums. Letting cell oscillation, excluded volume and ν vanish, then $\xi \downarrow 0$, continuity of Ψ_p (including $U = L$) proves (46). Restricting inputs excludes the guard only from this lower bound; no upper construction omits it. Scalar localization is essential; the original operators need not commute.

The nonlinear assertion reduces to this linear result through the frozen kernel. At a finite-harmonic interior b and an admissible J containing it, squaring (31) at $c = b$ with $e_\tau^2 = 2\rho_{m,\tau}^{\mathrm{nl}}$ gives, for each candidate policy, a kernel $K_b \subset H_J$ of codimension at most m with $\delta^2 v_\delta(b)(1-\eta_J)^2 \|T_{b,\tau}|_{K_b}\|^2 \leq 2\rho_{m,\tau}^{\mathrm{nl}}$, and the linear result applies on this tail to the normalized smooth derivatives, since finite codimension leaves their principal symbols unchanged. Take budget limits, then $J \to \infty$; finite-harmonic approximation extends the bound to all interior b, since source smoothing makes cofactors and normalizers converge uniformly, and boundary backgrounds ($v_\delta(b) = 0$) impose only the zero bound. □

The condition restores the singleton bound when $U = L$. For two tasks, $\Psi_p(U, L) = \int_0^1 [(1-t)a_1/r_1 + ta_2/r_2]^p dt$ is an average of necessary mixture inequalities, not a replacement for their full two-task criterion. Nor is the finite count sufficient: take three two-dimensional cells with eigenvalues $10, 30/7$, own-cell weight one and other-cell weight $1/10$. At $m = 5, q = 3/2$, (45) reads $5 = 3(17/18 + 13/18)$, yet the averaged task form has smallest eigenvalue $(2/5)(30/7) = 12/7 > 3/2$, excluding every such projection; a residual line spread equally over the three low modes attains $12/7$. Nor can arbitrary matrix diagonals be substituted: $\left(\begin{smallmatrix} 1 & .9 \\ .9 & 1 \end{smallmatrix}\right)$ has rank-one residual .1 though both diagonal entries exceed .2.

Proof of Corollary 5.4. Use normalized area on S^2, $H = \mathrm{span}\{1, u_1, u_2, u_3\}^\perp$ and $T_0 = (I + \Delta)^{-1/2}|_H$. For three equal longitude sectors C_j, put $\omega_{j,t} = 1/4 + (9/4)e^{-t\Delta} 1_{C_j}$ and $T_j = M_{\sqrt{\omega_{j,t}}} T_0$. At any fixed $t > 0$ the weights are smooth, normalized and in $[1/4, 5/2]$, with pointwise average one, and rotation conjugates the complete Gramians, including the finite gauge. At $p = 1$ every mixture has Weyl constant C_0 and shared continuous constant C_0. With $U_t = \max_j \omega_{j,t}$ and $L_t = \min_j \omega_{j,t}$, dominated heat convergence gives $\int \Psi_1(U_t, L_t)\, d\sigma \to (5/2 + 1/4)/2 = 11/8$; fix $t > 0$ with this integral above $6/5$. Lemma D.5 gives a linear constant exceeding $6C_0/5$: apply it to convergent subsequences of normalized near-optimal diagonal profiles; unbounded ones already exceed the bound. Equation (30) extends it to adaptive linear queries, while Theorem 5.3 gives exact sufficiency for every pair. The family and smoothing time are fixed before $m \to \infty$. □

Mixture bounds versus coding risks. For $A_\pi = \sum_i \pi_i A_i$, let s_{pair} maximize $\lambda_{m+1}(A_\pi)$ over mixtures supported on at most two tasks, and let s_{all} maximize it over all mixtures. Both are necessary lower bounds; their ratio need not measure an advantage of continuous over linear encoding. On $B_1(\mathbb{R}^2)$ with $m = 1$, let $A_i = I + u_i u_i^\top$, where $\|u_i\| = 1$ and $\arg u_i \in \{0, \pi/3, 2\pi/3\}$. Then $s_{\mathrm{pair}} = 5/4$, $s_{\mathrm{all}} = 3/2$ and $\Lambda^{\mathrm{lin}} = \Lambda^{\mathrm{cont}} = 7/4$, with risks stated before the half-risk factor. Each pair's smaller eigenvalue is maximized at the midpoint; every full mixture has trace three, and the uniform mixture is $3I/2$. A continuous scalar encoder identifies antipodes $v, -v$ on the unit circle, and some task line is within $\pi/6$ of v, giving worst-task squared error at least $1 + 3/4$. A fixed projection whose kernel has angle $\pi/6$ attains $7/4$ by (30). Thus stronger mixture bounds alone cannot establish the separation of Corollary 5.4.

Complete pairwise information and collective budgets. The parity construction compares two families, not two information classes of one family. Its combinatorial device is classical (Schneidman et al., 2003); its consequence here concerns entire pairwise risk regions and exact common-law reconstruction.

Proposition D.6 (Pairwise equivalence does not determine collective budgets)**.** *For every integer $M \geq 2$, choose $(3M-3)/(3M-1) < c < 1$ and*

$$\frac{1-c}{2} < r < \frac{1-c/3}{2M}. \tag{48}$$

Two triples of positive diagonal tasks on $B_1(\mathbb{R}^{4M})$ have orthogonally equivalent corresponding two-task problems at every coordinate budget, yet their minimum continuous-code budgets for the

common half-risk request r are exactly $3M$ and $4M$. They admit exact regular pure CPPZ realizations with positive polynomial supports on S^2, prescribed area laws and a common finite-context Bernoulli detector. All risks are multiplied by the same positive factor. For every pair, common legal recovery preserves the full continuous-code region.

Proof. Linear tasks. Let $S_\pm = \{s \in \{-1,1\}^3 : s_1 s_2 s_3 = \pm 1\}$ and index coordinates by (s,n), $1 \le n \le M$. Set $A_\tau^\pm = \mathrm{diag}((1+cs_\tau)/n)$ and $T_\tau^\pm = (A_\tau^\pm)^{1/2}$. Apply Proposition D.7 below with $k=3$, $p=1$, so $N = 4M$ and S_+ is its class $S_{-(-1)^k}$. By part (i), corresponding pairs agree after a coordinate permutation, which transports every encoder and both decoders. By part (ii) and (48), $\Lambda^+_{4M-1} = (1-c/3)/(2M) > r$, so S_+ needs $4M$ coordinates, which recover the source exactly. In S_-, discard the all-negative cell and measure the remaining $3M$ coordinates, giving worst-task half-risk $(1-c)/2 < r$. At the uniform mixture the retained cells have energies $(1+c/3)/n$. Since $M(1-c) < 1 - c/3 < 1 + c/3$, all M discarded coordinates lie below all retained ones, so the $(3M)$th eigenvalue is $(1+c/3)/M$. Its half exceeds r, and (19) excludes $3M-1$ coordinates.

Common-law realization. Choose the odd degree $\ell = 2M+1$, let $\lambda_\ell = \ell(\ell+1)$ and $d_\ell = 2\ell+1$, and take $4M$ real degree-ℓ harmonics Y_j orthonormal for normalized area σ. Put $\alpha = 1/(4\sqrt{d_\ell})$, $\phi_z = \sum_j z_j Y_j$ and $h_z = 1 - \alpha\phi_z/(\lambda_\ell - 2)$ for $\|z\| \le 1$. For a full degree-ℓ basis, the addition and Hilbert–Schmidt Bochner identities in the proof of Lemma A.7, on S^2, give $\sum_j Y_j^2 = d_\ell$ and $\sum_j \|\nabla^2 Y_j + Y_j g\|_F^2 = d_\ell(\lambda_\ell - 1)(\lambda_\ell - 2)$. These bound the same sums over the chosen subset and give $h_z \ge 1 - 1/[4(\lambda_\ell - 2)] > 0$ and $A_{h_z} \succeq \big(1 - \frac{1}{4}\sqrt{(\lambda_\ell - 1)/(\lambda_\ell - 2)}\big) g > 0$. Lemma A.4 compiles every support with its exact law. For $B_z = \nabla^2\phi_z + \phi_z g$, the two-dimensional determinant gives $J_{h_z} = 1 + \alpha\phi_z + \frac{\alpha^2 \det_g B_z}{(\lambda_\ell - 2)^2}$, $Z_{h_z} = 1 - \zeta\|z\|^2$, $\zeta = \frac{\alpha^2}{2(\lambda_\ell - 2)} < 1$. Here $\int \det_g B_z \, d\sigma = -(\lambda_\ell - 2)\|z\|^2/2$ by $2\det B = (\mathrm{tr}\, B)^2 - \|B\|_F^2$ and Bochner. Since ℓ is odd and the quadratic determinant is even, $\int Y_j J_{h_z} \, d\sigma = \alpha z_j$ exactly.

Let $S_M = 4\sum_{n=1}^M 1/n$ and let $a_{\tau,j} = (1+cs_\tau)/n$ be the diagonal entry of $A_\tau^\pm$ at $j = (s,n)$. Task τ draws context j with probability $a_{\tau,j}/S_M$, independently of the area-law source, and both families use the same fixed likelihood $K_j(u) = 1/2 + \kappa Y_j(u) \in [1/4, 3/4]$, $\kappa = 1/(4\sqrt{d_\ell})$. The forecast is $\eta_j(z) = 1/2 + \kappa\alpha z_j/(1-\zeta\|z\|^2)$. The radial map $v(z) = (1-\zeta)z/(1-\zeta\|z\|^2)$ is a homeomorphism of B_1 onto itself, since its radial derivative is positive and it fixes the boundary. Half-Brier excess is $c_\alpha \sum_j a_{\tau,j}(v_j - \widehat{v}_j)^2/2$ with $c_\alpha = (\kappa\alpha)^2/[S_M(1-\zeta)^2] > 0$, and every continuous code in z corresponds to one in v, so the lower bounds transfer exactly.

For the projection P of either collective upper construction, encode orthogonal coordinates of $Pv(z)$, recover $\widehat{v} = Pv(z)$, invert by $\widehat{z} = 2\widehat{v}/\big[(1-\zeta) + \sqrt{(1-\zeta)^2 + 4\zeta\|\widehat{v}\|^2}\,\big] \in B_1$ and compile $h_{\widehat{z}}$ once for all forecasts. For each pair and every budget, Theorem 5.3 supplies such a projection for every feasible target. This proves the common-law pairwise statement without a transmitted source norm or branch message. □

A rule given only these pairwise regions must either reserve $4M$ coordinates for both families, overprovisioning the easier one by M, or fail on the harder family at a smaller budget. Contrast and physical risk scale depend on M. No deployment claim is made.

Finite-precision realization. Fix M, c, r as above and let $\eta = r - (1-c)/2 > 0$. Round each retained coordinate of $Pv(z)$ to the nearest of 2^b uniform levels in $[-1,1]$ and project onto $B_1 \cap \mathrm{ran}\, P$; by nonexpansiveness the result $\widetilde{v}$ has $\|Pv(z) - \widetilde{v}\| \le \sqrt{m}/(2^b - 1)$. Both constructions in Proposition D.6 use coordinate projections, so P commutes with every diagonal A_τ and the discarded and rounding errors are orthogonal in every task norm. Since $\|A_\tau\| \le 1 + c$, the compiled worst-task half-risks satisfy $R_{-,b}/c_\alpha \le (1-c)/2 + 3M(1+c)/[2(2^b-1)^2]$ and $R_{+,b}/c_\alpha \le 4M(1+c)/[2(2^b-1)^2]$. Thus $2^b - 1 > \sqrt{2M(1+c)/\eta}$ suffices to meet the same physical request $c_\alpha r$ with $3M$ and $4M$ coordinates, respectively. Invert the same radial map and compile once; projection preserves legality and needs no transmitted source norm. The continuous-coordinate lower bounds still apply to a continuous encoder followed by rounding, since its arbitrary decoder absorbs the rounding, so these are also the minimum coordinate budgets in that restricted finite-precision class. Payloads are $3Mb$ and $4Mb$ bits, excluding the public model, dictionary and compiler; the sufficient b depends on M, c, r and is not a bit-optimality claim for arbitrary discontinuous encoders.

Full budget gaps versus distinct risk regions. The contrast condition only makes the interval in (48) nonempty; it is not necessary for the two collective risk regions to differ, since Proposition D.7(ii) with $k = 3$, $p = 1$ separates $\Lambda^{\pm}_{4M-1}$ for every $0 < c < 1$. For example, $M = 2$, $c = 1/2$, $r = 1/6$ give minimum budgets 7 and 8 although $c < 3/5$ (Appendix F.3, Figure 9).

Every order at once. The same parity device extends to every order $k \geq 3$.

Proposition D.7 (No proper subset of the tasks determines the whole)**.** *Fix integers $k \geq 3$, $M \geq 1$, a real $p \geq 1$ and a contrast $0 < c < 1$. For $\varepsilon \in \{-1, 1\}$ put $S_\varepsilon = \{s \in \{-1, 1\}^k : \prod_\tau s_\tau = \varepsilon\}$, index coordinates by (s, n) with $s \in S_\varepsilon$ and $1 \leq n \leq M$, and let T^ε be the family of k positive diagonal tasks on $B_1(\mathbb{R}^N)$, $N = 2^{k-1}M$, with $A^\varepsilon_\tau = \operatorname{diag}\big((1 + cs_\tau)/n^p\big)$. Then (i) For every proper subset $J \subsetneq \{1, \dots, k\}$ the sub-families $T^+|_J$ and $T^-|_J$ agree after a permutation of coordinates, hence have identical attainable risk regions at every budget; and (ii) With $\Lambda^\varepsilon_{N-1}$ the optimal worst-task half-risk of continuous $(N-1)$-coordinate codes,*

$$\Lambda^{(-1)^k}_{N-1} = \frac{1-c}{2M^p}, \qquad \Lambda^{-(-1)^k}_{N-1} = \frac{1-c(k-2)/k}{2M^p}, \tag{49}$$

and these differ for every $0 < c < 1$ and every k. Consequently every request r with $\Lambda^{(-1)^k}_{N-1} \leq r < \Lambda^{-(-1)^k}_{N-1}$ is met by one family at $N-1$ coordinates and by the other only at N.

Proof. (i). Fix J with $|J| = j < k$. A coordinate (s, n) enters task $\tau \in J$ only through s_τ, so its weight vector for the sub-family depends on $(s|_J, n)$ alone. Given $\sigma \in \{-1, 1\}^J$, the cells of S_ε with $s|_J = \sigma$ correspond to the sign patterns t on the complement with $\prod t = \varepsilon \prod \sigma$; since $k - j \geq 1$, exactly 2^{k-j-1} of the 2^{k-j} patterns have each product. So every σ occurs 2^{k-j-1} times in both classes and the two restricted weight multisets are equal. The permutation matching them transports every encoder and every decoder, not only the mixture spectra.

(ii). At budget $N-1$ the unmeasured space is a line, so (19) and the matching projection give

$$\Lambda^\varepsilon_{N-1} = \tfrac{1}{2} \max_{\pi \in \Delta_k} \min_{s \in S_\varepsilon,\, n \leq M} \frac{1 + c\langle \pi, s\rangle}{n^p} = \frac{1}{2M^p} \max_{\pi \in \Delta_k} \min_{s \in S_\varepsilon} \big(1 + c\langle \pi, s\rangle\big),$$

the inner minimum over n being at $n = M$. The all-negative cell has $\prod_\tau s_\tau = (-1)^k$ and $\langle \pi, s\rangle = -1$ for every π, and $\langle \pi, s\rangle \geq -1$ always, so on $S_{(-1)^k}$ the inner minimum is identically $1 - c$. On the other class every cell carries an odd number of positive signs, at positions P, and $\langle \pi, s\rangle = 2\sum_{\tau \in P} \pi_\tau - 1$ with $|P|$ odd; as $\pi \geq 0$, the minimum over odd P is attained at a singleton, so the inner minimum is $1 + c(2 \min_\tau \pi_\tau - 1)$, maximized by the uniform mixture at $1 - c(k-2)/k$. Both are attained: discard the all-negative cell's coordinate at level M, on which every task has energy $(1-c)/M^p$; and discard the unit direction equally supported on the k singly-positive cells at level M, on which every task has energy $[(1+c) + (k-1)(1-c)]/(kM^p) = (1 - c(k-2)/k)/M^p$. Finally $1 - c < 1 - c(k-2)/k$ because $(k-2)/k < 1$. □

The common-law realization of Proposition D.6 carries over with N in place of $4M$, any odd degree ℓ with $2\ell + 1 \geq N$, and the normalizer $S = 2^{k-1} \sum_{n \leq M} n^{-p}$ in place of S_M: fixing s_τ leaves 2^{k-2} completions of each sign, so $\sum_j a_{\tau,j} = S$ for every task and both classes. The context law $a_{\tau,j}/S$ generally depends on τ, though not on the latent source, and the Bernoulli detector is shared. Nothing else in the argument uses the number of tasks; it needs only orthonormality and the common odd degree of the chosen harmonics.

The separation $1 - (1-c)/(1-c(k-2)/k)$ decreases in k: as tasks are added, the other class's cheapest compromise approaches the all-negative cell, so the claim gains reach but loses size (Figure 11, left).

A scoped perturbation guarantee. On a fixed radius-δ ball, with unrestricted decoders and an unchanged information class, the optimal worst-task root error $e_m(T) = \sqrt{2\Lambda_m(T)}$ obeys $|e_m(T) - e_m(\widetilde{T})| \leq \delta\varepsilon$ whenever $\max_\tau \|T_\tau - \widetilde{T}_\tau\| \leq \varepsilon$ (triangle inequality for each information/decoder pair, then infima and exchange). Hence, for $M = 2$, $c = 4/5$, $r = 7/50$, $\delta = 1$, independent task-map perturbations of norm at most $3/40$ preserve the exact budgets 6 and 8: in

root error, $\sqrt{1/5}+3/40 < \sqrt{7/25} < \sqrt{11/30}-3/40$ separates the easy six-coordinate upper bound from the hard seven-coordinate lower bound, the easy five-coordinate lower bound stays above the request since $\sqrt{19/30}-3/40 > \sqrt{7/25}$, and full-coordinate recovery is exact. This is operator-norm robustness of the linear-task budget gap, not of exact pairwise equivalence or of the subset discrepancy of Appendix F.3; no perturbed CPPZ realization is asserted, and away from a matching upper construction mixture bounds alone do not determine minimum budgets.

E A STRICT ADAPTIVE-LINEAR/CONTINUOUS GAP WITHIN NATIVE CPPZ

Proposition E.1 (Native three-task separation)**.** *For the native model with $h_0 = 1$, $k = 2$, $\beta = 3$, $s = 4$, $r = 9$ ($p = 1/10$), there are fixed smooth, normalized positive task weights and a fixed $\delta > 0$ in the geometric interior such that*

$$\liminf_{m\to\infty} \frac{\Lambda_m^{\mathrm{alin,nl}}}{\Lambda_m^{\mathrm{cont,nl,CPPZ}}} > \frac{6}{5}.$$

The same lower separation holds with synchronized quantized two-batch legal acquisition in the numerator. The three tangent tasks have identical complete single-task spectra.

Proof. Let C_j be three equal longitude sectors, ignoring the null axis. Fix $\epsilon = (4/5)^{10}$, $a = 3/(1+2\epsilon)$, and, for $t > 0$, $\omega_{j,t} = a[\epsilon+(1-\epsilon)e^{-t\Delta}\mathbf{1}_{C_j}]$ and $T_{j,t} = M_{\sqrt{\omega_{j,t}}}T_0$. The smooth positive weights have mean one and average pointwise to one. As in the proof of Corollary 5.4, rotations conjugate the complete tangent Gramians, including the finite gauge. If $s_n(T_0)^2 \sim C_0 n^{-1/p}$, Jensen and the uniform mixture give $H_*(t) = C_0$; Theorem 4.2 gives the continuous tangent constant. For $U_t = \max_j \omega_{j,t}$ and $L_t = \min_j \omega_{j,t}$ put $G(t) = [\int \Psi_p(U_t, L_t)\, d\sigma]^{1/p}$. Lemma D.5, applied as in that proof, gives $\liminf_m m^{1/p} q_m(t) \geq C_0 G(t)$, where $q_m(t) = \inf_{P=P^*=P^2,\,\mathrm{rank}\,P\leq m} \max_j \|T_{j,t}(I-P)\|^2$. Dominated heat convergence yields

$$\lim_{t\downarrow 0} G(t) = \frac{3}{1+2(4/5)^{10}}\left[\frac{10[1-(4/5)^{11}]}{11[1-(4/5)^{10}]}\right]^{10} > \frac{241}{200} > \frac{6}{5}, \tag{50}$$

where the strict inequality is a rational comparison. Fix $t > 0$ with $G(t) > 241/200$ before the budget limit. For sufficiently small fixed δ, the weighted relative secant (7) has factor $\vartheta(\delta) \to 0$. Freezing any adaptive policy at zero and using kernel antipodes gives $\Lambda_m^{\mathrm{alin,nl}} \geq (\delta^2/2)(1-\vartheta)^2 q_m(t)$ by (30). Theorem 4.1 gives denominator constant $H_\delta = \max_\pi L_\delta(\pi) \leq (1+\vartheta)^2 C_0$: differentiating the relative secant bounds each background derivative by $(1+\vartheta)T_0$ in the task norms, and $v_\delta \leq 1$. Thus the ratio's liminf is at least $[(1-\vartheta)/(1+\vartheta)]^2 G(t)$; choose δ after t so this exceeds $6/5$. Quantized two-batch acquisition is a deterministic adaptive linear scheme with a restricted decoder, hence obeys the same lower bound. Theorem 5.1 also preserves the adaptive-linear numerator's optimal normalized extremal limits. □

F EXPERIMENTAL METHODS AND SUPPORTING RESULTS

Sections F.1–F.9 collect finite-model bounds and diagnostics for Section 6's common-law, collective-feasibility and acquisition studies. Sections F.10–F.11 provide complementary prediction and shared-recovery checks under radial mean risk. The four-mode bounds in Section F.6 use rational intervals; the diagonal-family values marked exact in Sections F.3 and F.5 are exact rationals, as are the certificates for the fixed width example in Section F.4. Other experimental bounds use floating point. Input searches do not certify global maxima, and no experiment establishes the general microlocal construction's asymptotic onset.

F.1 REQUESTED-RISK COMPILATION: MODEL AND UNIFORM BOUNDS

This model supplies the uniform bounds that decide the risk requests of Section 6, a finite-budget illustration of Theorems 5.3 and 4.1. Fix $R = 1$, $r = 8$, $\beta = 3$, $\tau_s = 1$, $Y = \sqrt{15}(u_1^2 - u_2^2)/2$, $\omega_\pm = (1 \pm Y/4)^9/N_{1/4}$ and $M_\beta = 8 > \sup k_3$ (Appendix D.3). The source has 32 real harmonic coordinates of degrees 2–5 and radius $\delta = 1$, with uniform curvature deviation at

most $\left[\sum_{\ell=2}^{5} \frac{(2\ell+1)(\lambda_\ell^2-3\lambda_\ell+2)}{(1+\lambda_\ell)^8}\right]^{1/2} \le .004288$, which verifies finite-model convexity. A requested risk profile is public metadata fixed before acquisition, with no source-dependent or realized-task information.

Finite map and bounds. With probability-normalized harmonics, $h_z = 1 + f_z$, $f_z = \sum_j d_j z_j Y_j$, $d_j = (1+\lambda_j)^{-4}$, $\lambda_j = \ell_j(\ell_j+1)$ and $\|z\| \le 1$. On the two-dimensional tangent plane, $H_z = \nabla_S^2 f_z + f_z I$ and the integrated Bochner identity give $J_z = 1 + \operatorname{tr} H_z + \det H_z$, $Z_z = 1 + z^\top D z$, $D_{jj} = (1-\lambda_j/2)d_j^2 < 0$. Indeed, the identity and integration by parts give $\int \det(\nabla_S^2 f + fI) = \int (f^2 - |\nabla_S f|^2/2)$; harmonic orthogonality diagonalizes this quadratic form, while the linear term has zero mean. With $S = M_\beta^{-1} K_\beta$, let $Bz = S \operatorname{tr} H_z$ and polarize $S(\det H_z - (z^\top D z)\mathbf{1})$ into the symmetric bilinear map $C(z,z)$. Then

$$F(z) = F(0) + \frac{Bz + C(z,z)}{s(z)}, \qquad s(z) = 1 + z^\top D z \ge s_* := 1 - \|D\| > .9999996. \tag{51}$$

Set $B_i = M_{\sqrt{\omega_i}} B$, $C_i = M_{\sqrt{\omega_i}} C$ and $A_i = B_i^* B_i$. For a positive ray $v = (v_+, v_-)$ put

$$\ell_m(v) = \frac{1}{2} \max_{0 \le t \le 1} \lambda_{m+1}\left((1-t)A_+/v_+ + tA_-/v_-\right). \tag{52}$$

Every deterministic adaptive m-query policy, and by Borsuk–Ulam every continuous real encoder, satisfies $\max_i \rho_i/v_i \ge \ell_m(v)$. For a policy, freeze its zero-answer path: a unit vector in the common kernel and the leading $(m+1)$-dimensional mixture space gives equally encoded z and $-z$, whose half-separation $B_i z/s(z)$ is odd by (51), with $s(z) \le 1$; the parallelogram bound, summed with the ray-scaled mixture, proves the assertion. No such assertion is made for arbitrary discontinuous encoders or randomized queries. For the $m-c$ adaptive queries that follow a prescribed c-dimensional coarse space W_c, the same argument on $H_c = W_c^\perp$ gives $L_c(m) = \frac{1}{2}\max_{0\le t\le 1} \lambda_{m-c+1}\left(((1-t)A_+ + tA_-)|_{H_c}\right) \le \max_i \rho_i$, which covers all later rounds after that prefix, with $c = 0$ the unrestricted adaptive-linear bound and $c = 5$ the five-query prefix of the two-batch pilots (Section F.9).

For $Q \in \mathbb{R}^{32\times m}$ with orthonormal columns, $P = QQ^\top$ and K with orthonormal columns spanning $\ker P$, write $b_i = \|B_i\|$, $b_{iK} = \|B_i K\|$, $c_i = \|C_i\|$ and $c_{iK} = \|C_i(\cdot, K\cdot)\|$, the bilinear norms bounded by flattened tensor operator norms. For a radius-d source ball retaining positive support and curvature, set $s_d = 1 - d^2\|D\| > 0$, $b_m = \lceil 5.5 \log_2(m+1)\rceil$ and $\nu_m = (2^{b_m}-1)^{-1}$:

$$\begin{aligned} u_i(P;d) &= \frac{d(b_{iK} + 2dc_{iK})}{s_d} + \frac{2d^3\|DK\|(b_i + dc_i)}{s_d^2}, \\ L_i(d) &= \frac{b_i + 2dc_i}{s_d} + \frac{2d^2\|D\|(b_i + dc_i)}{s_d^2}, \\ U_i(Q;d) &= \frac{1}{2}\left[u_i(P;d) + L_i(d)\sqrt{m}\, d\nu_m\right]^2. \end{aligned} \tag{53}$$

The decoder rounds $Q^\top z$ on $[-d,d]$ to b_m bits, synthesizes with Q, projects onto B_d and compiles by Lemma A.4 with positivity constants for that whole ball. Received words alone give one support, density and inverse-Gauss readout with task risk at most $U_i(Q;d)$. With $y = Pz$, $e = z - y$ and $N(z) = Bz + C(z,z)$, the exact identity $F(z) - F(y) = \frac{Be + C(z+y,e)}{s(z)} + \frac{N(y)[s(y)-s(z)]}{s(z)s(y)}$, with $\|e\| \le d$, $\|z+y\| \le 2d$, $\|N_i(y)\| \le d(b_i + dc_i)$ and $|s(y) - s(z)| \le 2d\|DK\|\|e\|$, yields $u_i(P;d)$; differentiation gives the source Lipschitz bound $L_i(d)$, nonexpansive ball projection changes y by at most $\sqrt{m}\, d\nu_m$ under rounding, and adding root errors before squaring retains the cross term. Coefficient and orthogonality roundoff are checked separately, not silently included in this exact-arithmetic bound. Unless stated otherwise $d = \delta = 1$, and its argument is omitted.

Projection-aware refinement. The bound can be tightened without changing the code. Write $t = \|z - Pz\|$, $a = (d^2 - t^2)^{1/2}$, and use subscripts Q, K for restricted linear and bilinear norms, for example $c_{iQK} = \|C_i(Q\cdot, K\cdot)\|$. Put $s_Q = 1 - d^2\|Q^\top D Q\| > 0$, $d_{KK} = \|K^\top D K\|$ and

$d_{QK} = \|Q^\top DK\|$. The same difference identity gives the root-error envelope

$$\frac{b_{iK}t + \min\{c_{iK}t\sqrt{4d^2 - 3t^2},\, 2c_{iQK}at + c_{iKK}t^2\}}{s_d} + \frac{(b_{iQ}a + c_{iQQ}a^2)(2d_{QK}at + d_{KK}t^2)}{s_d s_Q}. \tag{54}$$

Indeed, $\|z + Pz\|^2 \le 4d^2 - 3t^2$, while the alternative expands $C_i(z + Pz, z - Pz)$ into its two blocks; the second fraction bounds the normalizer change. Maximize this envelope over $0 \le t \le d$, add the same rounding term as in (53), and square with factor $1/2$. We upper-bound each of 10^5 scalar intervals using the extrema of t, a, at and $t\sqrt{4d^2 - 3t^2}$, rather than using sampled maxima. Spectral norms and coefficients remain floating-point evaluations, not outward-rounded certificates.

For finite tangent Gramians A_1, A_2 and positive requirements r_i, Theorem 5.3 directly gives the minimal budget $m_{\min}(r) = \max_{0\le t\le 1} n_+((1-t)(A_1 - q_1 I) + t(A_2 - q_2 I))$, $q_i = 2r_i/\delta^2$, where n_+ counts strictly positive eigenvalues: feasibility requires at most m positive directions in every pencil. This tangent consequence does not replace the nonlinear bounds (52) and (53).

Risk requests. For $m \in \{8, 12, 16\}$, $a \in \{1/4, 1/2, 1, 2, 4\}$ and $v = (a^{-1/2}, a^{1/2})$, freeze $r = 1.02\ell_{\rm hi}(v)v$ before encoding, where $\ell_{\rm hi}$ brackets the finite pencil from above. The solver checks both residual inequalities and the whole mixture interval, without assuming concavity of ordered eigenvalues. All 15 requests are met (Table 1). Seven supporting mixtures are interior, with both constraints active within 10^{-5} and risks at least 1.01 times their singleton optima; eight are endpoint-controlled. These rays sample, rather than characterize, the nonlinear Pareto body.

Table 1: All 15 frozen requests, with risks in units of 10^{-15}. Last columns normalize the radial nonlinear upper bound by the request and information lower bound (floating-point evaluation).

m	a	r_+	r_-	Upper/request	Upper/lower
8	1/4	1103.2941	275.8235	.988557	1.008328
8	1/2	551.6471	275.8235	.988656	1.008429
8	1	332.5435	332.5435	.988008	1.007768
8	2	275.8235	551.6471	.988656	1.008429
8	4	275.8235	1103.2941	.988557	1.008328
12	1/4	31.9352	7.9838	.986769	1.006505
12	1/2	15.9676	7.9838	.987658	1.007411
12	1	10.4921	10.4921	.988080	1.007841
12	2	7.9838	15.9676	.987658	1.007411
12	4	7.9838	31.9352	.986769	1.006505
16	1/4	17.4428	4.3607	.985610	1.005323
16	1/2	10.9990	5.4995	.985470	1.005180
16	1	6.8059	6.8059	.987757	1.007513
16	2	5.4995	10.9990	.985470	1.005180
16	4	4.3607	17.4428	.985610	1.005323

The code meets sv for $s \ge s_{\rm meet} := \max_i U_i/v_i$, while $s < s_{\rm imp} := \ell_m(v)$ is excluded; either bound may be loose, so the interval between is unresolved, not an intrinsic sharing gap, and neither conclusion requires solver optimality. Scaling requests among $10^{-15}, 10^{-14}, 10^{-13}$ preserves the selected spaces to printed precision. Across 31 budgets and 65 ratios in $[1/16, 16]$, the 2,015 threshold ratios have median 1.006381 and maximum 1.051811 at $m = 1$; the 1,820 cells with $m \ge 4$ have maximum 1.011993 (Figure 4).

For the opposing $m = 16$ requests in Figure 3, simultaneous compliance requires their coordinate-wise minimum, $(5.4995, 5.4995) \times 10^{-15}$. The mixture $t = .2659$ gives lower bound $6.6723983 \times 10^{-15}$, excluding 16 queries. At $m = 17$, the complement of the common 15-dimensional nonpositive space of $A_i - 2(.9)(5.4995 \times 10^{-15})I$ gives tangent half-risks $4.2482281 \times 10^{-15}$ and 23-bit nonlinear upper bounds $4.2747701 \times 10^{-15}$ for both tasks. Thus 17 linear queries suffice; this does not prove bit optimality. On 64 sphere inputs and two residual witnesses, received-word compilation gives maximum sampled risk 4.2485×10^{-15}, forecast discrepancy 3.06×10^{-16} against direct geometry, and decoded norm below $.8789$. Its 391 answer bits occupy 49 padded bytes. These exploratory checks concern common-law recovery, not independent confirmation or solver superiority.

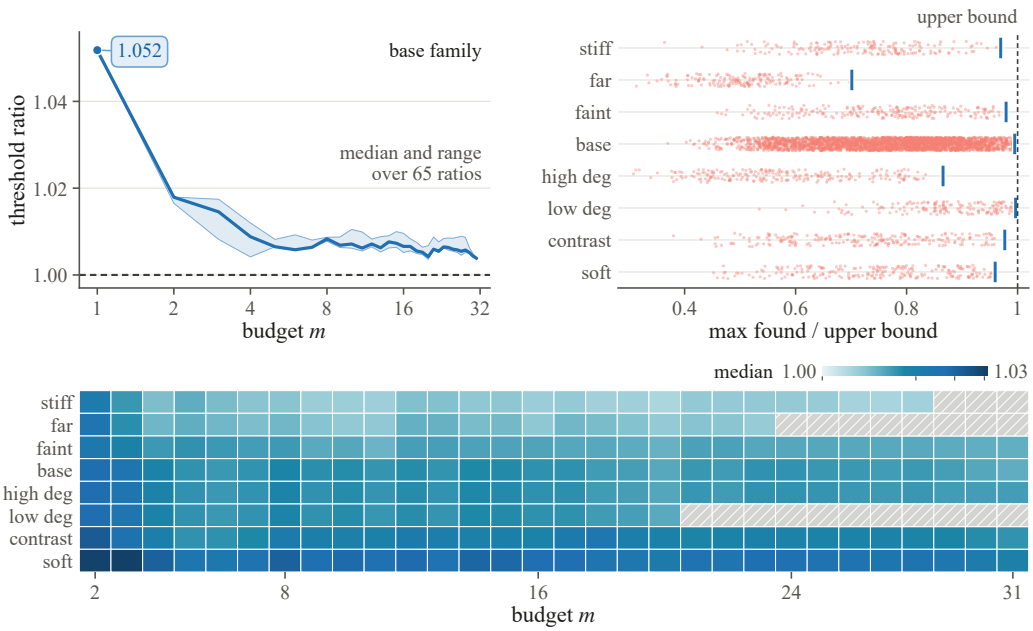


Figure 4: **Tight feasibility bounds across budgets and families.** *Top left*: base-family median and full range of $s_{\text{meet}}/s_{\text{imp}}$ over 65 requested ratios. *Bottom*: family medians of the same ratio from $m = 2$; gray cells are unsupported budgets. *Top right*: sampled search maxima divided by the upper bound, all below one; blue ticks mark family maxima. The 1,911 unsupported map cells and 147 unsupported attack cells are excluded.

Resource scope. The $144/252/368$ answer bits of the $m = 8, 12, 16$ requests ($18/21/23$ bits per answer) exclude $2/3/4$ KiB query matrices, $9.71/14.01$ MB shared geometry/model arrays and 1,166 scalar coefficients per decoded object; array counts omit container overhead. The known-model global-query oracle and absolute risks 10^{-15}–10^{-12} do not establish deployment speedup or robustness to noisy sampling.

Finite-radius validity. The unchanged 17-query plan meets $r_d = 5.4995 \times 10^{-15} d^2$ at all six radii in Table 2, while 16 queries remain excluded. Tasks and sensor are fixed; the 23-bit quantizer spans $[-d, d]$ and compiler positivity is recomputed on B_d, within the geometric domain. The refined bound (54) resolves $d = 64, 100$, where (53) gave upper/request $1.1139, 1.3486$ (Figure 5). This strengthens the guarantee for the same decoder, rather than correcting observed failures. At $d = 32$, support deviation is at most .02996, $s_d \geq .9996447$ and $r_d = 5.631488 \times 10^{-12}$. Numerical checks use 66 reused inputs, not independent confirmation: compiled/rational forecasts agree within 3.47×10^{-16}; quadrature refinement changes sampled risks/original uppers by at most 2.19×10^{-9}/2.9×10^{-15} relatively. Refined root errors change by less than 5.01×10^{-6} under doubled scalar resolution and 7×10^{-15} under quadrature refinement.

Table 2: The same quantized CPPZ at larger finite radii. U_d is the refined upper bound; Max. is the largest sampled loss. The final column ranges over both tasks and all 66 inputs, using the same recovered source for nonlinear and tangent loss.

d	Curvature bound	U_d/r_d	Max./r_d	Nonlinear/tangent
1	.00429	.7728	.7725	$[.9987, 1.0009]$
8	.03430	.7744	.7728	$[.9893, 1.0076]$
16	.06860	.7770	.7732	$[.9787, 1.0153]$
32	.13720	.7852	.7739	$[.9579, 1.0309]$
64	.27440	.8128	.7754	$[.9182, 1.0638]$
100	.42875	.8604	.7770	$[.8762, 1.1035]$

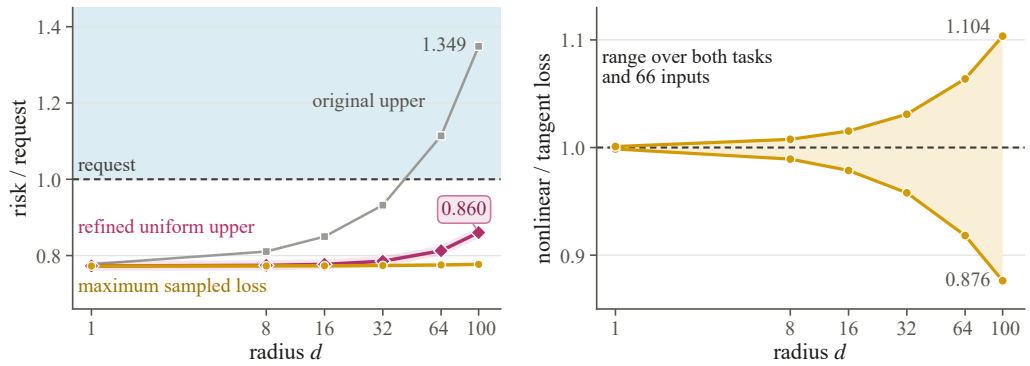


Figure 5: **A sharper bound resolves larger-radius requests.** Both panels use the plan and inputs of Table 2. *Left*: risk divided by the request $r_d \propto d^2$; the original and refined uniform bounds apply to the same quantized code. *Right*: nonlinear/tangent loss range over both tasks and 66 inputs, opening to about eleven percent at $d = 100$.

Variation and numerical checks. Geometrically valid single-factor variants test degree ranges 2–4/2–5/2–6, damping exponents $3.5, 4, 5$, contrasts $.1, .25, .5$ and sensor scales $8, 16$; the combined variant uses degrees 2–6, exponent 5, contrast .5 and scale 16. Sensor rescaling duplicates base ratios, leaving eight distinct settings. Across 16,224 cells, median threshold ratios span 1.001843–1.016303 and maxima for $m \geq 4$ span 1.004046–1.031729 (Figure 4); damping, not dimension alone, orders tightness. Residual-kernel, mixed and rounding-offset searches with ball/boundary controls reach at most .996433 of the upper: .945–.996 in dimensions 21/32 versus .632–.866 in dimension 45. All 200 diagnostic inputs satisfy the evaluated bounds (108 reused, 32 fresh, 60 witnesses; not independent confirmation). Density/weight degrees 12/18 give degree-42 quadrature exactness in exact arithmetic. Refinement changes bounds by at most 1.97×10^{-14} relatively; independent forecasts agree within 3.61×10^{-16}, without directed-rounding certification.

Finite-dimensional numerical reference. A 45-coordinate reference uses degrees 2–6, damping $(1 + \ell(\ell + 1))^{-3/2}$ and radius .25, with curvature deviation .3105. Central differences form the tangent Gramians, normalized by their mean trace. At $m = 1, 2, 4, 6, 8$, the two-task construction matches the numerical mixture lower value within 6.4×10^{-9} relatively. This finite linearized reference does not satisfy the infinite native regularity assumptions or certify competing solvers' convergence limits.

F.2 The two-task construction on generated and image Gram pairs

Constructive recovery. We implement Lemma D.2 through the pencil $Q_2 z = \lambda Q_1 z$: retain jointly negative scalar and nonreal-block directions, pair mixed blocks by threshold matching, and choose each nonreal direction at the midpoint of its negative-arc overlap. Five prespecified Gram families test random, isospectral, nearly commuting (10^{-3}–10^{-1} radians), repeated-spectrum (multiplicities 2–6) and rank-$n/2$ pairs. Projection risks are evaluated by residual eigenvalues.

Across 153,000 pairs ($n = 8$–256, random budgets and boundary targets), 152,631 meet 10^{-9} relative slack and 152,926 meet 10^{-3}. Among 368 rank-deficient misses, 202 targets lie outside the boundary due to coarse search and 166 below $1.5 \times 10^{-6}\lambda_1$. The remaining miss, a nearly commuting $n = 256$ pair, meets 10^{-6}. Preconditioning of all 369 misses resolves many refined targets, but stored-Gram/source-factor risks disagree near rounding resolution; the reported benchmark retains the original requests and counts. At $n = 256$, mean-Gram PCA, target-weighted PCA and the maximizing-mixture eigenspace have median worst risk/target $1.66, 1.33, 1.51$ (Figure 6); their best still misses 104,946 targets by $> 1\%$. At $n = 512, 1024$, the construction meets 39/40 at 10^{-9} in median 3/25 s; the remaining rank-deficient target (2.5×10^{-11}) meets 10^{-6}. On 300 pairs at $n = 16, 32, 64$, eight-restart Grassmann search meets 147 at 10^{-3} and misses 71 by $> 1\%$ (maxi-

mum 64%), versus 298 at 10^{-9} for the construction; median times are 6.2 s versus 1.2 ms at these optimization budgets.

Image-derived requests. Sources are covariance ellipsoids from MNIST (LeCun et al., 1998), Fashion-MNIST (Xiao et al., 2017) and KMNIST (Clanuwat et al., 2018), pooled to 14×14, and gray CIFAR-10 (Krizhevsky, 2009) to 16×16. Tasks reconstruct images with tenfold weights on left/right, top/bottom or center/periphery regions. At budgets 8–64 and 48 target directions per pair, the construction meets 2,303/2,304 boundary requests; restricting to the covariance range resolves the MNIST common-kernel miss. Data PCA misses all targets, by a median 46%; the best of it and the three spectral codes still misses 1,671 by over 1%. At full 28×28 resolution on MNIST and Fashion-MNIST, using 24 directions and covariance-range restriction, all 576 targets are met; PCA's median miss is 37%, and the best spectral code misses 394 by over 1%. By (4), the optimal worst relative risk is $\max_\pi \lambda_{m+1}(\sum_k \pi_k A_k/\lambda_{m+1}(A_k))$. Its median over the four datasets, three pairs and four budgets is 1.148, versus 1.674 for PCA (Figure 7). For MNIST halves it grows from 1.00 at $m = 8$ to 1.88 at $m = 64$: later components differ more between tasks. At full resolution the medians are 1.069 and 1.58. Changing region contrast from threefold to thirtyfold on MNIST and gray CIFAR-10 raises the optimal median from 1.055 to 1.151, versus 1.302 to 1.898 for PCA.

Unweighted image operators provide a separate control: identity, one-pixel Gaussian blur, its residual detail, and horizontal/vertical differences, each Gram normalized to unit norm. For image/differences, blur/detail and blur/differences pairs, 1,147/1,152 boundary targets are met within 10^{-9}; covariance-range restriction resolves all five misses. PCA misses 1,104 targets by over 1% and the best spectral code misses 732. Optimal median relative risk is 1.10, versus 1.62 for PCA. Unlike region tasks, these operators show no collective exclusion in 256 pairwise-feasible targets: excess is at most 3×10^{-15}. Their median relative commutator is .29, versus .35 for region tasks, so near-commutation does not account for this negative result.

Relation to fair dimensionality reduction. The Fantope relaxation used in fair/multi-criteria PCA (Samadi et al., 2018; Tantipongpipat et al., 2019) replaces the unmeasured projection by $0 \preceq X \preceq I$, $\operatorname{tr} X = n - m$. For worst-input risk its constraints are $\frac{1}{2} A_k^{1/2} X A_k^{1/2} \preceq tr_k I$, unlike average loss $\operatorname{tr}(A_k X)$. Already for one task, pinching gives a diagonal water-filling solution below $\frac{1}{2}\lambda_{m+1}(A)$: median ratio .48 on 7×7 pooled region tasks at $m = 4, 8, 16$. On 216 two-task boundary targets, feasible Fantope points give relaxation values at most .85 (median .55), although the achievable optimum is one. On 80 generated pairs at $n = 16, 32$, 77 values are at most .91; three rank-deficient solves do not converge. Rounding misses 136/216 image targets by over 1%. Fair PCA's own trace-loss relaxation returns a projection in all 216 cases, consistent with its guarantee, but worst-input ratios have medians 1.35 with equal and 1.13 with target-matched weights; 216 and 131 targets are missed by over 1%. The two-form construction meets all 216. The difference is the risk criterion, not failure of fair PCA's average-loss guarantee.

The limit region of Theorem 4.2 at finite budgets. On the torus $\mathbb{T}^2$ with the irrational metric $\xi_1^2 + \sqrt{2}\,\xi_2^2$, which removes the lattice coincidences of the square torus, take the reference $\widetilde{A}_* = (1 + \xi_1^2 + \sqrt{2}\,\xi_2^2)^{-\mu}$ and two tasks $\widetilde{A}_j = \widetilde{A}_*^{1/2} W_j \widetilde{A}_*^{1/2}$, where W_1 and W_2 multiply by $e^{a\cos x_1}$ and $e^{a\cos x_2}$. They share no eigenbasis (relative commutators .14–.63), and their principal energies are the weights themselves. For two tasks, Theorem 5.3 characterizes the exact finite-budget region by the mixture bounds $\pi \cdot r \geq g_m(\pi)$, where $g_m(\pi) = m^{1/p}\lambda_{m+1}(A_\pi)$ in the normalized units of (3); the prediction $h(\pi)$ of (2) comes from the weights by quadrature, with no spectrum and no fitted constant. With Fourier modes $|k_i| \leq 40$, and a budget kept only where modes $|k_i| \leq 32$ agree to 10^{-4}, $\max_\pi |g_m/h - 1|$ falls from 11–44% at $m = 16$ to 1.2–4.6% at $m = 256$ and .4–1.8% at the six resolved $m = 1024$ cases, for $p = 2, 1, 1/2$ and $a = 1/2, 1, 2$ (Figure 8). The two tasks are balanced with nonproportional energies, so the worst-task price of sharing has limit one for $p \geq 1$ and a limit strictly above one for $p < 1$. For $p = 2$ it equals one at every budget; for $p = 1$ it lies within 1.3% of one from $m = 256$ on, except 2.7% at the largest amplitude. For $p = 1/2$ it approaches the predicted strict price: at $m = 1024$, 1.188 against 1.183 for $a = 2$, 1.050 against 1.057 for $a = 1$, and 1.013 against 1.015 for $a = 1/2$.

These weights vary only in position, so the cosphere integral in (2) is trivial for them. To exercise it, give each task a preferred frequency direction: $B_j = W_j^{1/2} M_j W_j^{1/2}$ with M_j the Fourier multiplier

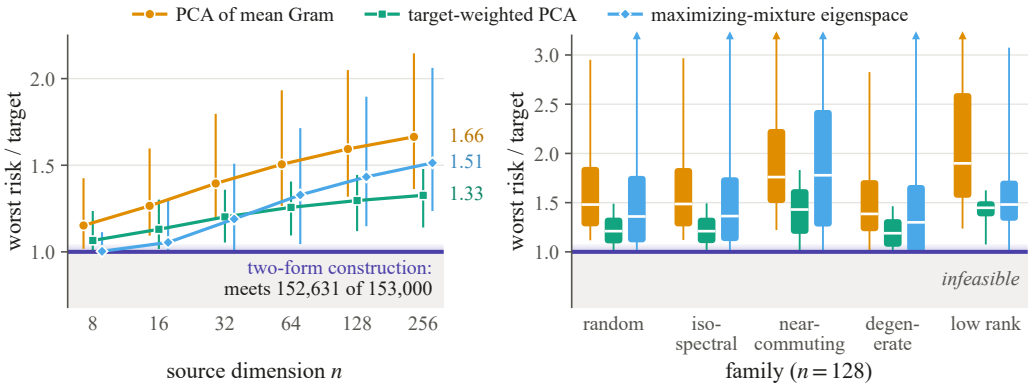


Figure 6: **Spectral codes against the two-form boundary benchmark.** Ratios use frozen numerical requests targeting (4). The unit line marks request satisfaction. The gray band indicates infeasibility for ideal exact-boundary targets; numerical target-search errors and construction misses remain in the reported benchmark. *Left*: worst risk/target of three spectral codes against dimension, median and interquartile range over five families. *Right*: by family at $n = 128$, 5–95% whiskers.

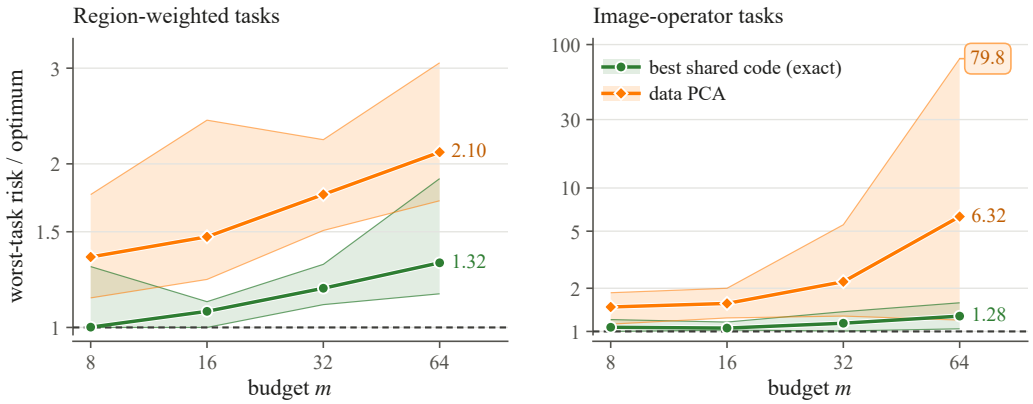


Figure 7: **One shared code costs little; data PCA can cost much more.** Worst task risk relative to its own optimum, $\max_k \rho_k/\rho_k^*$, for the best shared code (exact by (4), attained by the construction) and for data PCA: median and range over four datasets and three task pairs (log axes, separate y ranges). *Left*: region-weighted tasks. *Right*: image, blur, detail and difference tasks.

$1 \pm c\cos 2\theta_\xi$, so that $a_j(x,\theta) = w_j(x)(1 \pm c\cos 2\theta)$, with $a = 1, 2$ and $c = .5, .9$. The prediction then integrates over angles, $h(\pi) = [(8\pi^2)^{-1} \iint a_\pi^p \, e(\theta)^{-1} d\theta \, dx]^{1/p}$ with $e = \cos^2\theta + \sqrt{2}\sin^2\theta$, and reduces to the formula above when $c = 0$ to 10^{-15}. The angular multipliers couple more modes, so fewer budgets pass the resolution check, but at each family's largest resolved budget the deviation is 1.2–3.6%. The price is one for $p = 2$, one for $p = 1$ in all but one family (1.030 at $a = 2$, $c = .9$), and for $p = 1/2$ it is 1.042, 1.111, 1.167 and 1.202 for $(a, c) = (1, .5), (1, .9), (2, .5), (2, .9)$, against predicted 1.046, 1.100, 1.156 and 1.185.

F.3 Pairwise and collective feasibility without a construction

Exact enumeration for the parity triples. For the triples of Proposition D.6, exact rational enumeration at $M = 2$, $c = 1/2$ covers the entire budget axis (Figure 9). All six two-task problems, three from each triple, have the same optimal half-risk at every budget $m = 0, \ldots, 8$; as Proposition D.7(i) requires, no pairwise measurement distinguishes the families at any budget. The collective values differ at $m = 2, 3, 4, 5, 7$. Which family is cheaper depends on the request, not on the parity class: at $r = 1/2$ the minimum collective budgets are 3 for S_- and 2 for S_+, while at $r = 1/6$ they are 7 and 8. The obstruction is therefore existential over requests.

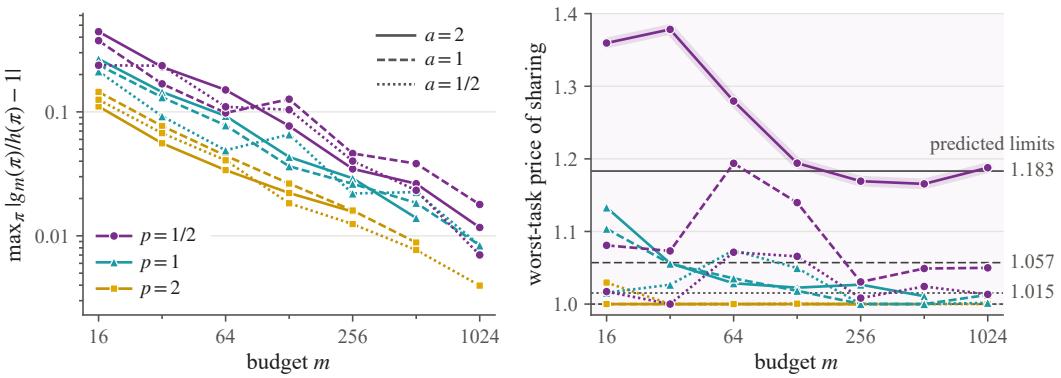


Figure 8: **Finite-budget mixture bounds approach their Weyl limits.** Two noncommuting tasks on an irrational torus; the mixture bounds characterize the exact region by Theorem 5.3. *Left*: largest relative deviation of the finite-budget mixture spectral bound from $h(\pi)$, computed from the symbols alone. *Right*: worst-task price of sharing; thin lines are the predicted limits, one for $p \geq 1$ and strictly above one for $p < 1$. Only budgets converged in resolution are shown.

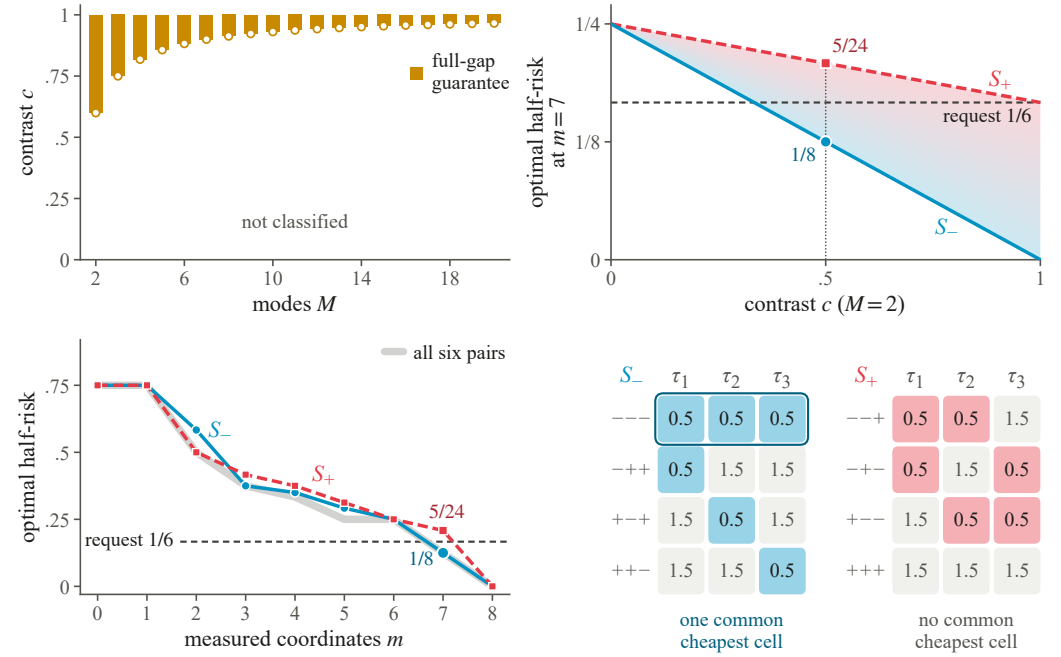


Figure 9: **Identical pairwise information, different collective budgets.** Triples on $B_1(\mathbb{R}^{4M})$ have energies $(1 + cs_\tau)/n$. *Top left*: amber gives minimum budgets $3M$ and $4M$ for the requests in (48); the boundary (open circles) is excluded, and the region below it is unclassified, not equivalence. *Top right*: exact half-risks at $M = 2$, $m = 7$. *Bottom left*: at $c = 1/2$ all six pairwise curves coincide at every budget, yet request $1/6$ needs seven coordinates on S_- and eight on S_+, outside the top-left panel's full-gap regime. *Bottom right*: sign cells (rows) by tasks (columns), with energies $1 + cs_\tau$ at $c = 1/2$; only S_- has a jointly cheapest cell (boxed), while S_+ mixes the three singly-positive cells uniformly. Values come from exact rational arrangement enumeration, not fitting or sampling; CPPZ realization scales every risk by c_α.

Pairwise and collective budgets on random diagonal triples. The gap between pairwise and collective feasibility that Proposition D.6 builds also appears without a construction (Section 5.2). On $B_1(\mathbb{R}^n)$ with three diagonal tasks and budget $m = n - 1$, the optimal worst-task half-risk is ex-

actly $\frac{1}{2}\min_{p\in\Delta}\max_\tau d_\tau\cdot p$ for continuous, fixed linear and adaptive linear classes alike: measuring the orthogonal complement of $w=\sqrt{p^*}$ attains it, and (19) matches it, so the program is an optimum and not a further unattained bound. This program is a matrix game, $\min_{p\in\Delta_n}\max_\tau d_\tau\cdot p=\max_{\lambda\in\Delta_3}\min_j(D^\top\lambda)_j$, so a feasible primal/dual pair (p,λ) whose two sides agree proves optimality in rational arithmetic; floating point only proposes a support. The certificate is rebuilt and rechecked in exact arithmetic, so a wrong proposal is recorded as a failure and cannot become a wrong value, and the measurement is checked against ten matrices whose values are known by hand.

All 1,344 families have exact primal/dual certificates. Each cell of $n\in\{8,12,24\}$ and lognormal spread $\sigma\in\{.35,.7\}$ contains 96 independent-spectrum triples, 96 independently permuted copies of one spectrum and 32 identical-task controls. Collective risk exceeds the worst pair in 48–63, 52–70 and 76–82 of 96 families at $n=8,12,24$, respectively (Figure 10); the smallest positive relative gap is 1.8×10^{-5}. All 192 identical-task controls give exact equality. Fixed requests confound the gap with changing risk scales: at $n=24$, no family is pairwise feasible at $(\sigma,r)=(.35,1/4)$, whereas all are jointly feasible at $(.7,3/8)$. The scale-free window $(\Lambda^{\mathrm{triple}}-\Lambda^{\mathrm{pair}})/\Lambda^{\mathrm{triple}}$ instead has cell means from 2.0% at $(8,.35)$ to 9.4% at $(24,.7)$. Twelve redraws per cell separate only the ends of the dimension range cleanly. The obstruction's frequency and size here describe generated diagonal tasks, not natural task populations; no random family is compiled to a CPPZ.

Geometry-induced tasks. The twelve confirmation backgrounds of Section F.8 each carry eight context tasks whose Grams $A_\tau=DF(0)^*M_{\omega_\tau}DF(0)$ come from the geometry, the law and the detector. They do not commute: their relative commutators $\|[A_i,A_j]\|/(\|A_i\|\|A_j\|)$ run from .047 to .392. Write $P_m=\max_{\{i,j\}}\max_{\pi\in\Delta_2}\frac{\delta^2}{2}\lambda_{m+1}(\pi A_i+(1-\pi)A_j)$, which Theorem 5.3 makes the *exact* threshold for two tasks, and $L_m=\max_{\pi\in\Delta_3}\frac{\delta^2}{2}\lambda_{m+1}(\sum_\tau\pi_\tau A_\tau)$, a lower bound on the collective optimum by (19). Every $r\in[P_m,L_m)$ is then met by every pair but by no code serving all three at that budget. Any evaluated mixture bounds L_m from below, and Courant–Fischer on subintervals of π encloses P_m from above: for any m-dimensional U, $\lambda_{m+1}(M)\le\lambda_{\max}(M|_{U^\perp})$, and on $U^\perp$ the family is affine in π while $\lambda_{\max}$ is convex, so a subinterval's maximum sits at an endpoint; U is taken from the subinterval's midpoint. A direct Weyl bound is too coarse here, since $\|A_i-A_j\|$ is set by the top of the spectrum near 10^{-6} while λ_{m+1} is near 10^{-11}.

Across 56 triples per background and $m=1,\dots,31$, all pair enclosures close and 14,849/20,832 cases (71.3%) have a numerical exclusion window. Every background contributes; the window occupies on average 6.4% of $[0,L_m]$, rising to 10.8% at $m=31$. Three identical tasks give $|L_m-P_m|\le3\times10^{-22}$. Thus pairwise sizing can under-provision operators generated by the CPPZ geometry and detector, not just diagonal constructions. These floating-point inequalities are not outward-rounded certificates. Since L_m is only a collective lower bound, the windows establish resource shortfalls, not continuous–linear separations.

At $m=5$, five of 672 triples have a positive numerical exclusion window, compared with 56% at $m=4$ and 31% at $m=6$; all five remain part of the result. This coincides with closer leading subspaces: the largest pairwise principal-angle sine is at most .202 at $m=5$ and at least .9999 at $m=3,4,6,7,8$. These spaces are approximately aligned, not identical, and their proximity alone does not explain relative risk; the unresolved eigenvalues must be compared with the spectral spread. Indeed, let P_τ be a top-m projector of $A_\tau\succeq0$ and P a proposed common rank-m projector. With $\alpha_\tau=\lambda_{m+1}(A_\tau)$, $\beta_\tau=\lambda_1(A_\tau)$ and $\epsilon_\tau=\|P_\tau(I-P)\|$, spectral decomposition gives $\|(I-P)A_\tau(I-P)\|\le\alpha_\tau+(\beta_\tau-\alpha_\tau)\epsilon_\tau^2$. For unit $x\in\ker P$, the components of x along P_τ and its orthogonal complement contribute at most $\beta_\tau\|P_\tau x\|^2$ and $\alpha_\tau(1-\|P_\tau x\|^2)$, which proves the bound. Exact agreement on a top space makes the common projection attain the hardest singleton risk, but small angles require a residual-scale comparison. An exact example shows why. For $0<t^2<2/3$, put $u_j=(t\cos(2\pi j/3),t\sin(2\pi j/3),\sqrt{1-t^2})$, $A_j=(1+t^2)I-u_ju_j^\top$, $j=0,1,2$. On $B_1(\mathbb{R}^3)$ with two coordinates, each pair's optimal continuous-code half-risk is $7t^2/8$ and the triple's is t^2. The midpoint pair Gram has minimum eigenvalue $7t^2/4$, and discarding $(u_i+u_j)/\|u_i+u_j\|$ attains its half-risk for both tasks; the uniform triple Gram has minimum eigenvalue $2t^2$, and discarding e_3 attains its half-risk for all three; mixture antipodes give the matching lower bounds. Yet the top-two spaces' angle sine is $\sqrt{1-(1-3t^2/2)^2}\to0$ while the relative gap stays $1/8$.

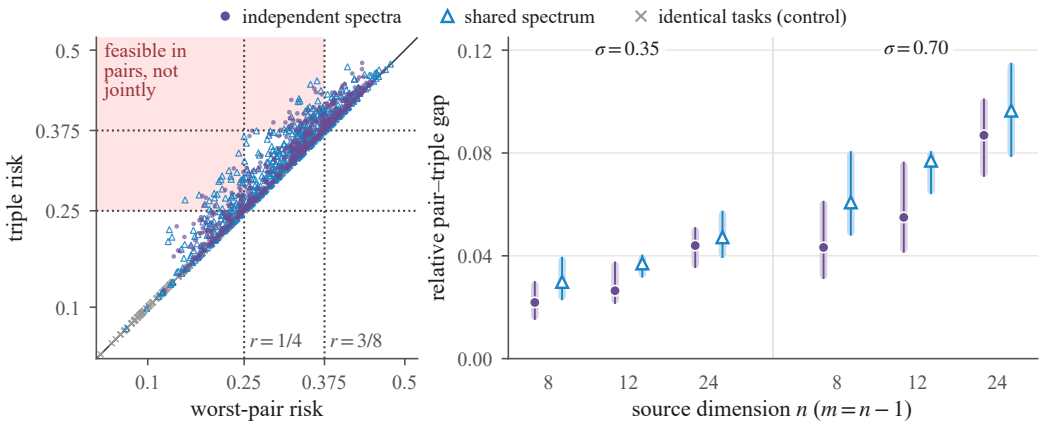


Figure 10: **Pairwise evidence can underestimate collective risk.** Random diagonal triples at $m = n - 1$, every family exactly certified. *Left*: all 1,344 families, worst pair against triple, with the diagonal and the two declared requests (dotted); identical-task controls lie on the diagonal, the other two populations above it. *Right*: the scale-free effect against dimension for both spreads σ; bars span twelve independent redraws of each cell.

The condition number $(1 + t^2)/t^2$ diverges, identifying the missing residual scale rather than a bounded-conditioning obstruction.

Image-derived collective obstructions. On the four covariance sources of Section F.2, use three or four quadrant tasks and two region triples, each with tenfold region weight. Scale random requests to their exact pairwise feasibility threshold. The collective mixture bound excludes 200/2,688 direction–budget pairs: 81 on MNIST, 50 on KMNIST, 35 on gray CIFAR-10 and 34 on Fashion-MNIST, with maximum excesses $29\%, 13\%, 5\%, 8\%$. Each excluded request has feasible pairwise projections but no shared continuous code; every excluding mixture weights at least three tasks. The smallest excess is 5.8×10^{-6}; for all 27 below 10^{-3}, pairwise searches reproduce the boundary within 4×10^{-15}. Contrast matters: among 384 MNIST/CIFAR-10 quadrant targets, threefold weighting gives 7 exclusions (all MNIST, at most 3% excess), whereas thirtyfold weighting gives 111 (64/47 by dataset, at most 68%). The unweighted image operators give none in 256 targets. The claim is therefore a possible, task-dependent collective obstruction, not universal failure of pairwise grouping.

Subset information and collective feasibility. Let $\Lambda^{(j)}$ be the largest exact game value over all j-task subsets at $m = n - 1$. Figure 11 shows diminishing but nonzero increments from complete subset information. At $n \le 16$, the final step is positive in $59\%, 40\%, 25\%, 12\%$ of sampled $k = 3, 4, 5, 6$ families. A 2,856-family replication at $(k, n, \sigma) = (4, 12, .7)$ gives positive-step rates $99.3\%, 86.6\%, 38.6\%$, with 95% intervals no wider than .036. For $k = 6$, increasing n to 24, 32 and 48 raises final-step incidence to $24\%, 42\%, 51\%$; at $n = 48$ mean increments are $.413, .179, .081, .034, .0082$. Both dimension and subset order affect the observed increments. An entrywise-dominated final task gives a zero final step even when earlier steps are positive.

For a random positive diagonal triple, take P_m and L_m as above with $\delta = 1$, now with P_m the largest exact pairwise risk. Exact vertex enumeration supplies L_m but no jointly attaining code at general budgets, so $m_{\min}(r) \ge m_L(r) := \min\{m : L_m \le r\}$, $\widehat{m}(r) := \min\{m : P_m \le r\}$, and $m_L(r) - \widehat{m}(r)$ lower-bounds the pairwise sizing deficit. Exclusion windows occur in 95.6% of sampled families at $n = 8$ and at least 99.9% from $n = 16$. Their union, normalized by L_0, has mean length 6.4% at $n = 8$ and 17.0% at $n = 40$, with certified deficits reaching 2 and 6 coordinates. Singleton sizing gives window shares 37–42% and deficits reaching 10. Replications with 2,880 families per cell give 99.93–100% window incidence at $n = 16$ and 99.3–99.97% at $n = 12$. Empty detected windows remain inconclusive; identical-task controls have exact zero windows and increments.

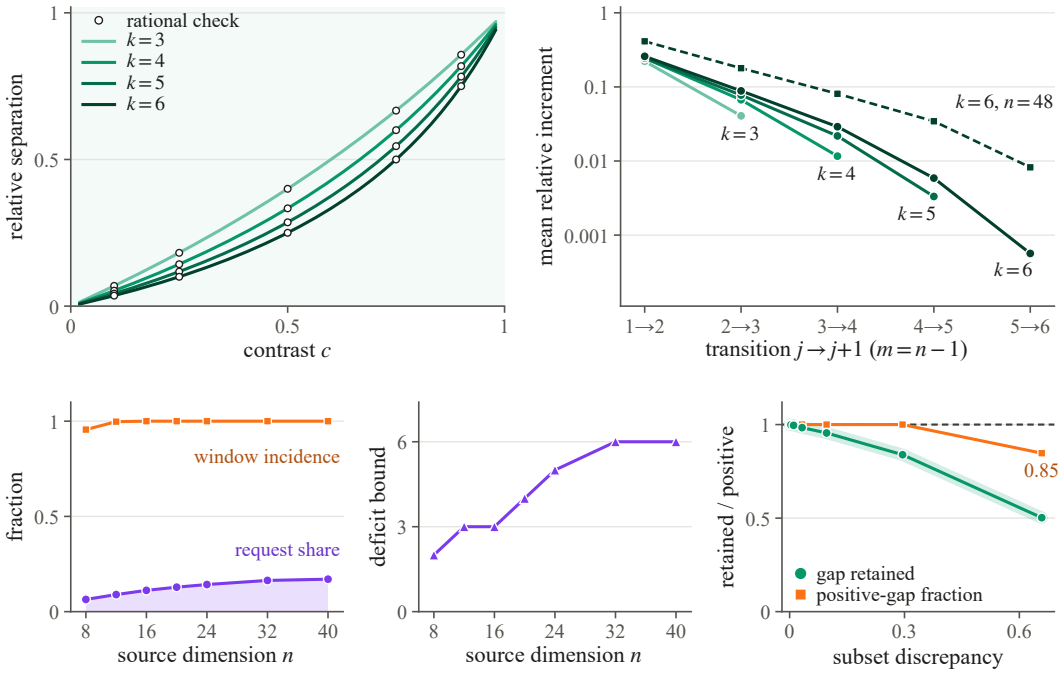


Figure 11: **Subset information leaves collective requirements unresolved.** *Top left*: the exact relative separation $1-(1-c)/(1-c(k-2)/k)$ between the two values in (49), against contrast c, at each order k; markers are rational checks. *Top right*: mean relative increments between complete j-wise and $(j+1)$-wise values at $m = n-1$ on random diagonal families, for $n \le 16$ (solid) and $n = 48$ (dashed, $k = 6$); both order and dimension affect the increments. *Bottom left*: detected exclusion-window shares across all budgets and fractions of families with a detected window. *Bottom middle*: lower bounds on the largest coordinate deficit; neither panel shows exact collective budget curves. *Bottom right*: the collective gap after multiplicative perturbation, against the restricted weight-multiset discrepancy. Incidences describe sampled populations, not natural task distributions.

Multiplicative perturbations of Proposition D.7 break exact subset agreement, measured by optimal matching of restricted weight multisets. The collective gap stays positive in every draw through 29% discrepancy, retaining 96% of its size at 10% and 84% at 29%. At 66% discrepancy it remains positive in 85% of draws and retains half its size. Thus the gap persists away from exact equivalence.

F.4 Grouping, trained encoders and width brackets on image statistics

Setup. On the four covariance ellipsoids of Section F.2 ($\delta = 1$), tasks reconstruct images with tenfold weight on a random axis-parallel rectangle covering 15–45% of the image, or a randomly chosen fixed region (halves, center, periphery or quadrants; ten choices). Requests are $r_k = s_k \rho_k^*$, where $\rho_k^* = \frac{1}{2}\lambda_{m+1}(A_k)$ and the s_k are independent uniform draws from the ranges below. Equation (19) bounds every shared continuous code's worst risk/request below by $\kappa_{\mathrm{LB}}(m) = \max_{\pi \in \Delta_q} \frac{\frac{1}{2}\lambda_{m+1}\left(\sum_k \pi_k A_k\right)}{\pi \cdot r}$, equivalent to Section F.8's request-normalized bound after reweighting π. For these linear tasks it holds at each budget. Local mixture search, initialized at every pair's maximizer, can only underestimate the bound.

Pairwise grouping plans. A pairwise plan repeatedly removes the largest group with all pairs feasible by (4), a worst-case analogue of learned task-grouping affinities (Standley et al., 2020; Fifty et al., 2021). A group with $\kappa_{\mathrm{LB}} > 1$ is collectively infeasible despite its feasible pairs. Figure 12 reports the resulting lower bounds on failure frequency. For six rectangle tasks, 160 families at each $m = 8, 16, 32$ give an overall rate of 3.5% for $s_k \in [1.1, 2]$ (10.0% at $m = 32$), versus

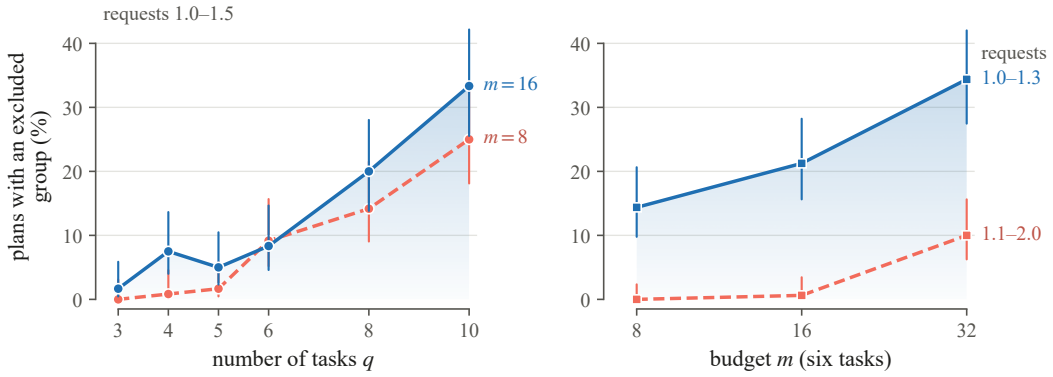


Figure 12: **Pairwise-feasible groups that no shared code can serve.** Share of random-rectangle families whose pairwise plan contains a group of three or more tasks with $\kappa_{\mathrm{LB}} > 1$. *Left*: against q, requests 1–1.5 times each task's own optimum, $m = 8$ (coral, dashed) and 16 (blue, solid), 120 families per point. *Right*: six tasks against m, requests 1.1–2 (coral, dashed) and 1–1.3 (blue, solid) times that optimum, 160 families per point. Bars are 95% Wilson intervals. The shares are lower bounds for these sampled populations and request ranges.

23.3% for $s_k \in [1, 1.3]$ (by budget: $14.4\%, 21.3\%, 34.4\%$). For $s_k \in [1, 1.5]$, $q = 3, 4, 5, 6, 8, 10$ and $m = 8, 16$, the pooled rate rises from .8% to 29.2% across the task-count range. Adding the necessary collective check $\kappa_{\mathrm{LB}} \leq 1$ raises code use from 3.26 to 3.42 per family at $m = 32$, $s_k \in [1, 1.3]$.

Verifying collective failure. Explicit two-form projections verify every pair in all 281 excluded groups (17 and 112 in the six-task studies; 152 in the task-count study). The largest pair ratio is .999996, the smallest recomputed collective bound 1.000102, and stored/recomputed bounds differ by at most 1.3×10^{-4}. Conversely, passing this necessary test does not imply feasibility (Appendix D.6): in the odd-indexed half-sample, projection search meets 174/240 admitted requests for $s_k \in [1.1, 2]$ and 72/223 for $s_k \in [1, 1.3]$; the rest remain unresolved.

Trained encoders on collectively excluded requests. We test whether input searches expose the predicted collective obstruction in learned codes, rather than asking those codes to meet an excluded target. We use the 45 excluded targets of Section F.3 ($\kappa_{\mathrm{LB}} > 1.05$): 33 MNIST, 7 KMNIST, 4 Fashion-MNIST and one CIFAR-10; 34 triples, 11 quadruples, with 40 at $m = 64$. A 2% request inflation makes every pair strictly feasible but leaves collective bounds in $[1.030, 1.261]$. Dataset/task/budget-matched controls have $\kappa_{\mathrm{LB}} = 1/1.02$; 45 draws with replacement give 42 distinct controls. Passing this necessary screen alone does not certify the controls' feasibility. Two width-1,024 GELU encoder layers and one-hidden-layer task decoders train for 5,000 steps on worst task ratios, with a persistent pool of 4,096 projected-gradient inputs (Madry et al., 2018). All 45 groups have violations: searched risk$/\kappa_{\mathrm{LB}} = 1.009$–$1.169$ (minimum 1.0008 across 374 distinct group, pair and control networks). All 80 width/depth/training/decoder variants also violate their requests (up to width 4,096, depth four and 20,000 steps, or linear decoders; minimum ratio 1.0047). Five combined-large-model runs diverge to 62–707 times the bound and are excluded. Closed-form projections meet 21/42 controls but no excluded group (median ratio 1.38).

Bound-guided worst-input search. Weak input searches can overstate robustness (Madry et al., 2018; Athalye et al., 2018; Croce & Hein, 2020). Here κ_{LB} supplies both a threshold and a search space. On the maximizing mixture's top $(m + 1)$-eigensphere V, Borsuk–Ulam gives equal-code antipodes (Matoušek, 2003); the parallelogram identity in (19) forces one to have ratio at least κ_{LB}. We minimize $\|E(z) - E(-z)\|^2$ on this sphere and evaluate both signs (median smallest gap 1.3×10^{-7} on trap networks).

Table 3 shows that search geometry matters more than additional random starts: Adam remains below the bound on 37/45 traps with 65,536 starts, whereas APGD and both eigensphere-constrained

objectives reach all 45. The latter need 64 starts and under three seconds, versus about 27 s for random-start APGD. Merely initializing on V is insufficient; a random subspace reaches none (median $.34\kappa_{\mathrm{LB}}$). The benefit is therefore not exclusive to the antipodal objective. APGD finds violations on 15/42 controls, including the antipodal method's 10; the remaining 27 are unresolved. Across 120 initial, capacity-sweep and independently retrained networks, Adam (8,192 random starts plus the training pool, 300 steps) reports 66 compliant and misses the bound on 100, versus no compliant reports for antipodal search. It also misses the bound on all 42 retrained controls. Two repeated, nonindependent sweeps give corresponding counts 67 and 101.

Table 3: Worst-input searches on 87 trained shared encoders, one per distinct image target: 45 traps ($\kappa_{\mathrm{LB}} > 1$) and 42 controls ($\kappa_{\mathrm{LB}} < 1$). "Compliant": no violated request found; "at bound": maximum-found ratio at least $(1 - 10^{-3})\kappa_{\mathrm{LB}}$. Time: median seconds per trap network on one GPU shared by twelve processes. A compliant report is not a certificate.

Search	Starts×steps	Traps (of 45) compliant	Traps (of 45) at bound	Controls (of 42)	Time (s)
Adam, random starts	$1{,}024 \times 300$	33	7	42	3.0
	$8{,}192 \times 300$	32	7	42	12.3
	$65{,}536 \times 300$	31	8	42	92.1
	$8{,}192 \times 3{,}000$	18	12	42	124.5
Adam, training pool	$4{,}096 \times 300$	33	7	42	7.0
Adam, started on the sphere of V	64×600	33	7	42	2.5
	$4{,}096 \times 600$	28	8	42	13.9
APGD, random starts	$8{,}192 \times 1{,}000$	0	45	27	27.0
	$65{,}536 \times 300$	0	45	27	67.5
APGD, one run per task	$2{,}048 \times 1{,}000$	0	45	27	29.8
APGD, started on the sphere of V	$4{,}096 \times 1{,}000$	0	45	27	15.2
On the sphere of V, antipodal	64×100	13	16	42	.4
	64×600	0	45	32	1.9
	$4{,}096 \times 600$	0	45	32	14.1
On the sphere of V, worst ratio	64×600	0	45	32	2.6
	$4{,}096 \times 600$	0	45	31	14.1
Antipodal, uniform mixture	64×600	3	38	42	1.9
Antipodal, random subspace	64×600	45	0	42	1.9

Width brackets before training. An excluding mixture and a feasible projection bracket the minimum width: $m_L \le m_{\min} \le m_{\mathrm{lin}}$, where exclusion holds at $m_L - 1$ and m_{lin} is the first feasible width found. Equality determines the minimum; an open bracket cannot distinguish lower-bound slack from search failure. Explicit pair projections at $\widehat{m}$ make $m_L > \widehat{m}$ a witness against pairwise sizing.

The complete cohort comprises two prespecified families in each of 96 dataset/region-type/task-count/reference-width cells, with $q = 3, 4, 6, 8$, $m_0 = 8, 16, 32$ and $r_k = s_k \lambda_{m_0+1}(A_k)/2$, $s_k \overset{\mathrm{iid}}{\sim} \mathrm{Unif}[1, 1.6]$. All families are retained; the covariance sources, not new image samples, are reused. Mixture search combines locally refined pairwise grids and 300 random simplex candidates; only exhibited mixtures exclude a width. Projection search tests $m_L, \ldots, m_L + 8$, stopping at first feasibility, with eight starts (mixture, request-weighted mean, individual eigenspaces and random frames) and at most 4,000 Adam steps per width. Full double-precision spectra check the six-vector Rayleigh–Ritz objective initially and every 500 steps. Exclusion requires ratios above $1 + 10^{-7}$; feasibility requires ratios below $1 - 10^{-7}$. All 192 families have feasible projections and pairs; 110 brackets close under this protocol (57.3%; 39, 33, 22 and 16 of 48 by task count). The 82 open brackets have widths $1, 2, 3, 4, 5, 7$ with respective counts $47, 17, 9, 5, 2, 2$. Pairwise sizing fails for 38 (19.8%; 4, 7, 12 and 15 by task count), with one-, two- and three-coordinate deficits in 32, 5 and 1 families. Counts by m_0 are 0, 13 and 25 of 64 (Figure 13).

A concrete width decision. One six-task MNIST family needs 36 coordinates jointly although all 15 pairs fit at 34. Mixture ratios are 1.085759 at 34 and 1.004594 at 35; the 36-coordinate

projection's ratio is 0.983427. This member illustrates the cohort, rather than forming a separate confirmation test. Exact rational certificates verify these decisions for the fixed numerical model, beyond the displayed floating-point ratios.

Keep its requests, mixture weights and found projections fixed. Replay the frozen covariance-to-task construction and regard the finite entries of its maps T_i as exact dyadic rationals. Form $A_i = T_i^\top T_i$ exactly, rather than treating rounded, rank-deficient Gram products as exact positive matrices. An exact absolute row-sum bound puts each difference from the stored Gram below 10^{-12} in operator norm. The unit source ball and all risk requests are unchanged. For nonnegative mixture weights w_i, a 36-column matrix V, and each found projection basis Q, exact integer arithmetic verifies

$$V^\top \left[\sum_i w_i (A_i - 2r_i I) \right] V \succ 0, \qquad 2r_i I - A_i + \tau_i Q Q^\top \succ 0 \quad (\tau_i > 0). \tag{55}$$

The first inequality and min–max give a strictly positive 36th eigenvalue of the shifted mixture, excluding every 35-coordinate continuous code by the antipodal bound. The second gives $x^\top A_i x < 2r_i \|x\|^2$ on $\ker Q^\top$, so the exact orthogonal projection onto $\operatorname{ran} Q$ meets the task request. We check all six tasks for the stored 36-column basis and both tasks for each of the 15 stored 34-column pair bases. In all 37 checks a dyadic congruence makes the tested matrix strictly diagonally dominant with positive diagonal, certified by integer row sums. Numerical eigensolvers only propose the congruences; they do not decide positivity.

For the full cohort, residual-map singular values agree with stored risks within 1.5×10^{-14}; a separate symmetric eigensolver checks all $m_L - 1$ lower witnesses within 1.9×10^{-14}. The 654 pair projections in the 38 undersized families also pass the singular-value check. These are numerical consistency checks, not rational certificates.

Sensitivity to search and cohort selection. Full-spectrum refinement of the 82 open confirmation brackets closes one four-task CIFAR-10 request at width 31, reducing its ratio from 1.001063 to .999306 (independent spectral checks agree within 1.1×10^{-15}). The resulting 111 closed/81 open brackets do not replace the frozen counts in Figure 13.

The earlier 2,375-family cohort uses the same request distribution but excludes slow completions. With three starts and 600 steps, closure is 42.1% ($81.9\%, 38.4\%, 4.8\%$ by m_0). In its 288-family subset, eight starts and 4,000 steps on 169 open brackets raise closure from 39.9% to 59.7%; four families without codes are not re-searched. Rates by task count are $94.4\%, 72.2\%, 40.3\%, 31.9\%$; the complete cohort's three-task rate is instead 81.3%, demonstrating selection sensitivity. Even sixteen starts and 16,000 steps leave four eight-task families unresolved at $m_{\mathrm{lin}} - 1$: upper ratios $1.0349, 1.0209, 1.0248, 1.0351$ versus lower bounds $.99, .95, .97, .90$.

In the earlier cohort, 390/2,375 pairwise widths are excluded (16.4%; 75 deficits of at least two, 14 of at least three; minimum mixture excess 4.9×10^{-5}). Restriction to 1,824 complete-block families changes rates by at most 2.2 percentage points. Although 67.7% need more coordinates than the hardest singleton, the 2,352 found shared codes average 72% fewer than separate codes. These are search-dependent budget brackets, not exact collective budget curves.

Scope. The population consists of generated worst-input linear reconstruction tasks, not classification or natural task distributions. Sampled searches on one architecture/optimizer family establish violations, not all-input compliance or learned CPPZ recovery.

F.5 Pairwise and collective compression on one finite source

On one finite source, the ideal continuous code for three linear tasks has worst-task risk below the all-linear lower bound at every tested budget, although each pair's fixed sign-matching projection (Theorem 5.3) meets its mixture bracket within 5.7×10^{-8} relatively at the chosen worst-task target, not every target in the exact pairwise body (Section 6, Figure 14).

Let $z \in B_1(\mathbb{R}^{3\times 256})$ and $w_{\tau j} = 1$ for $j = \tau$, .1 otherwise, with task half-risk $\frac{1}{2}\sum_{j,n} w_{\tau j}(z_{j,n} - \widehat{z}_{j,n})^2/n$; pairs and triple receive the same full source. The triple code sends Av and three tags $\|z\| s$, where $s_j = (\|z_j\|^2/\|z\|^2 + .01)/1.03$ and $v_{j,n} = [Ms_j - n + 1]_0^1 z_{j,n}$ for $n \le M$, zero otherwise; the code is zero at $z = 0$. The decoder recovers $\|z\|$ by summing the tags, obtains s, inverts A on

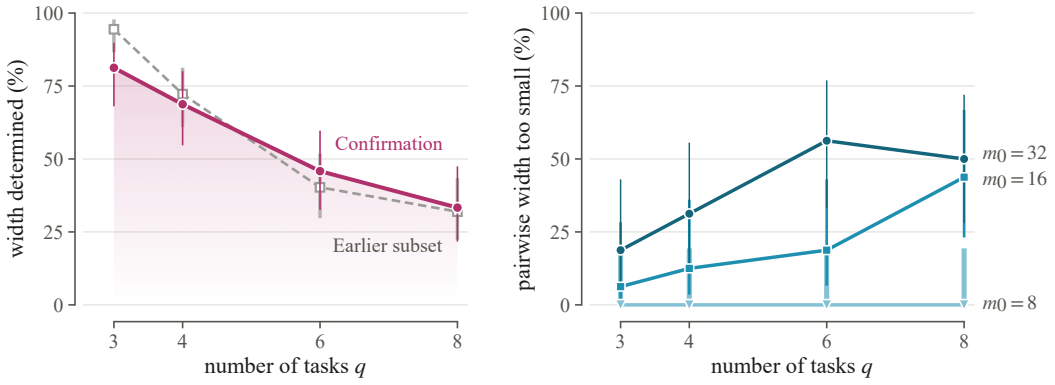


Figure 13: **Width brackets from the bound and a found projection.** *Left*: bracket closure in the earlier second-pass subset (72 families per task count) and the complete confirmation cohort (48 per task count). *Right*: confirmation families with $m_L > \widehat{m}$, by reference width m_0 (16 per point). All 192 confirmation families are included. Bars give approximate 95% Wilson intervals; the sampling distribution is the specified task generator. Open brackets do not establish an information-class separation.

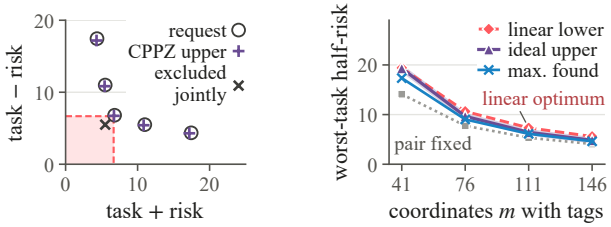


Figure 14: **Joint requests and common-code separation.** Left: the two opposing $m = 16$ requests and their CPPZ upper bounds from Appendix F.1; shading denotes joint exclusion, in 10^{-15} half-Brier units. Right: the separate three-cell linear model, with all-linear lower, ideal continuous upper, maximum-found and pair-projection half-risks (10^{-3} units); the budget includes tags. The two panels concern different task models.

its prefix support and projects onto the source ball. The fixed matrix A has $M + 2 + \lceil M/10 \rceil$ rows; counting the tags, $M = 32, 64, 96, 128$ gives budgets $m = 41, 76, 111, 146$. Left-invertibility of A on all 16,468 maximal admissible prefix unions was checked, with left-inverse Frobenius residual at most 1.444×10^{-10} and singular-value floor 2.986×10^{-4} under conservative IEEE allowances, not directed rounding.

Uniform upper and all-linear lower bounds. Since $(1 - [Ms_j - n + 1]_0^1)^2/n \le 1/(Ms_j)$, including the unencoded tail, and $\|z_j\|^2/\|z\|^2 \le 1.03 s_j$, the ideal exact-arithmetic code has worst-task half-risk at most $U_M = 1.03(1 + .1 + .1)/(2M) = .618/M$. By Lemma D.4, which applies directly to this finite spectrum and requires neither a power law nor an asymptotic limit, the unique root q_m of $m = 3\sum_{n=1}^{256} \phi_{.1}(qn)$ in (45) gives the exact rational bounds

$$\Lambda_m^{\mathrm{lin}} \ge L_m := q_m/2, \qquad (L_{41}, L_{76}, L_{111}, L_{146}) = \left(\frac{25}{1288}, \frac{19}{1785}, \frac{11}{1500}, \frac{37}{6615}\right), \tag{56}$$

valid, by the kernel and zero-path arguments, for arbitrary nonlinear task decoders and deterministic adaptive linear queries, including cross-cell and cross-level measurements. In particular, the common-budget bound resolves $m = 41$. Each pair's projection comes from the real-form sign-matching construction, which searches the mixture interval, checks endpoints and interval bounds and uses target slack 10^{-8} relatively. The triple's fixed comparator pairs low and higher coordinates cyclically across cells and pads the budget remainder; Table 4 compares all four budgets. At $m = 111$ this cyclic cross-cell comparator is exactly optimal among linear codes: its 111 unit vec-

tors $\sqrt{(75-l)/75}\,e_{j,l}+\sqrt{l/75}\,e_{j+1,75-l}$, $j=1,2,3$ cyclically, $1\le l\le 37$, have disjoint supports, complementary-line task energies at most $1.1/75$ and tail energy at most $1/75$ for $n\ge 75$. Thus $\Lambda^{\mathrm{lin}}_{111}=11/1500$ and $\Lambda^{\mathrm{cont}}_{111}\le 103/16000$, a ratio at least $352/309$; linear optimality is not claimed at other budgets.

Table 4: Finite-source pair/triple comparison, with risks in units of 10^{-3}. The rational common-budget lower bound is attained at $m=111$; U_M concerns the ideal continuous code. Maximum-found losses on 1,153 inputs are numerical diagnostics.

m	Pair fixed	Triple fixed	L_m	U_M	Max. found	L_m/U_M
41	14.1026	20.3704	19.4099	19.3125	17.3479	1.0050
76	7.7465	10.7843	10.6443	9.6563	9.0278	1.1023
111	5.3398	7.3333	7.3333	6.4375	6.1250	1.1392
146	4.0741	5.6701	5.5933	4.8281	4.6142	1.1585

Evaluation. The 128 exploratory and 192 held-out sources use Gaussian, $n^{-1/2}$-weighted, four-sparse and Dirichlet energies (concentrations $.3, 1, 3$; norms $1, .5$). Together with 760 residual maximizers over an energy grid, 48 fixed-code maximizers, the zero source and 24 support-boundary probes, they form the 1,153 inputs. All losses use actual encode–decode outputs; relative reconstruction error is at most 1.473×10^{-14} in double precision.

Isospectral noncommuting perturbations. Replacing e_i by $\cos\theta\, e_i+\sin\theta\, e_{i+1}$, cyclically, in $B_i=.1I+.9e_ie_i^\top$ and $A_i=B_i\otimes \mathrm{diag}(1/n)_{n=1}^{256}$ gives $\widetilde{B}_i,\widetilde{A}_i$ with the same complete task spectra, noncommuting at the tested nonzero angles. For $\eta=\max_i\|B_i^{-1/2}\widetilde{B}_iB_i^{-1/2}-I\|$, $(1-\eta)A_i\preceq\widetilde{A}_i\preceq(1+\eta)A_i$ bounds the same ideal code by $(1+\eta)U_M$, and (30) transports (56) to $\widetilde{\Lambda}^{\mathrm{lin}}_m\ge(1-\eta)L_m$ (Table 5), again covering arbitrary task decoders and deterministic adaptive linear queries. Separation therefore holds whenever $\eta<(L_m-U_M)/(L_m+U_M)$; failure of this sufficient test leaves the comparison unresolved. The perturbed tasks use the same four code matrices and 1,153 inputs, and all evaluated losses lie below the transported bounds. This concerns linear tasks on a finite source ball and the stability of their construction, not the smooth-sphere construction or the general noncommuting theorem.

Table 5: Transported all-linear lower/ideal-code upper ratios at all tested budgets and angles. Each row preserves the individual spectra; values at most one are unresolved, not counterexamples.

θ	η	$m=41$	$m=76$	$m=111$	$m=146$
0	0	1.0050	1.1023	1.1392	1.1585
.001	.002850	.9993	1.0961	1.1327	1.1519
.003	.008575	.9880	1.0836	1.1198	1.1388
.005	.014332	.9766	1.0712	1.1070	1.1258
.010	.028868	.9486	1.0405	1.0752	1.0935

Mixture scan. A separate exploratory mixture scan does not support a frequency claim for continuous–linear separation (Appendix D.6, 'Mixture bounds versus coding risks'). Interior ascent missed the 401×401 grid maximum in two of twelve checks, by up to 2.8%; edge-search error was unquantified. These errors bias the ratio oppositely, without a continuous-domain enclosure, so neither detected nor absent mixture gaps determine optimal shared risk.

F.6 Four-mode risk-constrained acquisition

The four-mode model of Proposition 5.2 tests whether three adaptive or four fixed queries meet the same risk request (Figure 15). Use $\delta\in\{50,100,150,200,250,300\}$ and $j\in\{8,10,11,12,14,16,18,20,24,32\}$ answer bits on $[-\delta,\delta]$. The target is $1.04\rho_0(\delta)$; radii are source norms in the smoothed parametrization.

Radius and precision dependence. Keep the four modes and tasks fixed, but take $0 < \delta \le 300$ and j bits per answer. All geometric and compiler bounds above remain valid. In the upper-bound proof replace N by $2^j - 1$ in every rounding term, including $e_b = \sqrt{2}\delta/N$ in the received-direction error. Provided (38) holds, the expression in (39) gives a bound denoted $U_{\mathrm{two}}(\delta, j)$.

For the fixed measurements $b_1, b_2, (x_1 - x_2)/\sqrt{2}$, the unresolved unit high-mode direction u obeys $\|A(b)u\|^2 \le \lambda + at + qt^2$, where $t = |b|$. Its squared coefficient is at most $\delta^2 - t^2$, so completing the square gives, when $q\delta^2 < \lambda$,

$$(1 - t^2/\delta^2)(1 + (a/\lambda)t + (q/\lambda)t^2) \le 1 + (a/\lambda)t - (\delta^{-2} - q/\lambda)t^2 \le 1 + \frac{(a/\lambda)^2\delta^2}{4(1 - q\delta^2/\lambda)}.$$

This bound and $s(z) \ge s_-$ give the leading odd error

$$O_{\mathrm{fix}} = \frac{\sqrt{\lambda}\,\delta}{s_-}\sqrt{1 + \frac{(a/\lambda)^2\delta^2}{4(1 - q\delta^2/\lambda)}} + \frac{d_x\delta^3\sqrt{\lambda + a\delta + q\delta^2}}{s_-^2}.$$

The same even error E, normalizer correction and nonexpansive recovery give the fixed-three upper below. Measuring all four coordinates instead leaves only a source rounding error $2\delta/N$:

$$U_{\mathrm{fix},3} = \frac{1.001}{2}[\sqrt{O_{\mathrm{fix}}^2 + E^2} + L_F\sqrt{3}\delta/N]^2, \qquad U_{\mathrm{fix},4} = \frac{1.001}{2}(2L_F\delta/N)^2. \tag{57}$$

For a fixed-three lower bound $\gamma\rho_0(\delta)$ with $\gamma > 1$, set $x_0 = (\gamma - 1)/(\lambda_b/\lambda - 1)$ and choose any $0 < \beta_0 < \delta$. Equation (40) covers $|u_b|^2 \ge x_0$; the same translated chord covers the remainder whenever $x_0 < 1$, $\sqrt{h_\beta} - Kx_0 > 0$ and the left expression in (41) exceeds γ, with h_β, K evaluated at this β_0, δ. Table 7 lists witnesses verified by rational intervals. These are public proof parameters, not uncharged source information. An upper bound above a request leaves that construction unresolved; only a lower bound above it excludes the whole fixed information class.

Unequal precision and projection-aware rounding. Let the first two answers have j_b bits and the third j_p bits, $e_b = \sqrt{2}\delta/(2^{j_b} - 1)$, $e_p = \delta/(2^{j_p} - 1)$ and $\eta = (e_b^2 + e_p^2)^{1/2}$. The received-direction argument above still applies with this e_b. For $H = (\lambda + a\delta + q\delta^2)^{1/2}$ and $N_E = \sqrt{\lambda_b}\delta + (c_b + c_x)\delta^2$, differentiating the even and odd parts of (35) gives partial Lipschitz bounds

$$E_b = (\sqrt{\lambda_b} + 2c_b\delta)/s_- + 2d_b\delta N_E/s_-^2, \qquad E_x = 2c_x\delta/s_- + 2d_x\delta N_E/s_-^2,$$
$$O_b = \sqrt{q}\,\delta/s_- + 2d_b\delta^2 H/s_-^2, \qquad O_x = H/s_- + 2d_x\delta^2 H/s_-^2.$$

Here $\|J(\Delta b)x\| \le \sqrt{q}\|\Delta b\|\|x\|$ follows from (36). If $w = (\widehat{b}, \widehat{p}v)$ lies in the source ball, the squared unweighted error is at most $V_0 = (O + O_b e_b + O_x e_p)^2 + (E + E_b e_b + E_x e_p)^2$. Otherwise let $r = \delta w/\|w\|$. Since $\|w\|^2 \le \delta^2 - e^2 + 2\delta\eta + \eta^2$, projection implies $e^2 \le 2\delta\eta + \eta^2$; moreover $1 - \delta/\|w\| \le (2|b|e_b + 2|p|e_p + \eta^2)/(2\delta^2)$. Using $|b||p| \le \delta^2/2$ bounds the block displacements by $\beta_b = 2e_b + e_b^2/\delta + e_p/2 + e_b e_p/\delta + \frac{\eta^2}{2\delta}(1 + e_b/\delta)$, $\beta_x = \sqrt{2\delta\eta + \eta^2} + e_p + \eta$. The segment from z to r lies in the ball, giving $V_1 = (E_b\beta_b + E_x\beta_x)^2 + (O_b\beta_b + O_x\beta_x)^2$. Thus $1.001\max\{V_0, V_1\}/2$ bounds both task risks. This retains the parity split and charges radial projection without assuming weighted nonexpansivity.

For a fixed four-query comparison, round both b coordinates to j_b bits and both x coordinates to j_x bits. Preserve the ball-projected coarse pair and project the fine pair to its remaining radius before compilation. Its block errors are at most e_b and $e_x + \sqrt{2\delta e_b + e_b^2}$, where $e_x = \sqrt{2}\delta/(2^{j_x} - 1)$: the latter follows from nonexpansivity in the fine ball and the difference of its squared radii. Substitution in the partial Lipschitz bounds yields a uniform risk bound by the same parity argument. This is a different legal decoder from the full radial projection in (57).

Uniform decisions and answer costs. Rational intervals with exact sphere moments and outward square roots verify the conditions and bounds above at all radii. Table 7 reports the 18-bit slice; Figure 15 reports all 60 equal-precision cells using the better of the original and projection-aware uppers. The refinement resolves nine previously open cells; 24 remain unresolved by either three-query upper. Fixed-four suffices at every tested $j \ge 11$, but remains unresolved at $j = 8, 10$. At

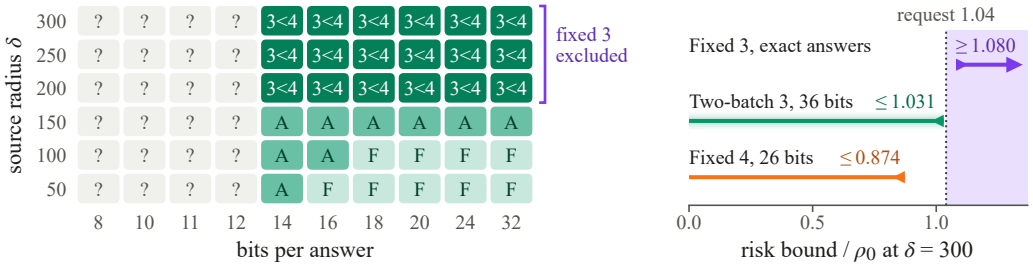


Figure 15: **Geometry adaptation changes the required query budget.** At request $1.04\rho_0(\delta)$ (dotted, right), fixed-three is excluded for $\delta = 200$–300. *Left*: rational-interval decisions (F, fixed-three recovery; A, two-batch-three recovery with fixed-three unresolved; $3 < 4$, three two-batch versus four minimum fixed queries; ?, unresolved). *Right*: refined bounds at $\delta = 300$ with unequal precision and coarse-preserving fixed-four recovery. The 36/26-bit sufficient counts are not bit optima and exclude public descriptions.

$\delta = 300$, the projection-aware bound of Appendix F.6 resolves equal-16-bit adaptive recovery: its upper is $1.002033\rho_0$, below the request. Unequal precision reduces the sufficient payload further (Table 6). The fixed-four comparison receives the same precision-allocation freedom and uses fewer bits despite more queries. Counts exclude public model/query/compiler descriptions and prove neither optimal bit complexity nor a minimum adaptive query count.

Table 6: Refined rational upper bounds at $\delta = 300$, with request $1.04\rho_0$. The final column is rounded upward; all three rows pass. The unequal-precision choices are exploratory sufficient allocations.

Measurements	Bits per answer	Total bits	Upper/ρ_0
Two-batch, 3	$(16, 16, 16)$	48	1.002033
Two-batch, 3	$(13, 13, 10)$	36	1.030504
Fixed, 4	$(11, 11, 2, 2)$	26	.873187

Table 7: Uniform four-mode bounds in units of $\rho_0(\delta)$; the request is 1.04. Public (β_0, γ) witnesses establish the lower bound for all fixed three-query maps with exact answers; upper columns use 18 bits per answer and are rounded upward.

δ	β_0	All fixed 3: lower γ	Fixed 3: upper	Two-batch 3: upper
50	2.5	1.002	1.012899	1.001033
100	10	1.010	1.027208	1.001061
150	22.5	1.023	1.051494	1.001099
200	40	1.042	1.087183	1.001150
250	60	1.060	1.136577	1.001216
300	100	1.080	1.203291	1.001300

Implementation and controls. Across 180 configurations, all decoded losses lie below their uniform bounds on 3,968 shared evaluation rows (512 fresh ball inputs plus boundary grids; 3,244 distinct points). At $\delta = 300$, the 11-bit two-batch code violates the request on 128 rows (maximum ratio 1.128049). At 16 bits the refined analytic bound, not the absence of sampled violations, establishes feasibility. Conversely, the tested 33-bit allocation $(12, 12, 9)$ remains unresolved despite a sampled maximum below the request. Direct geometric and tensor forecasts agree within 2.08×10^{-16} on 64 decoded geometries. At $\delta = 300$ and 18 bits, replacing the prescribed normal law by a uniform law produces 3,584/3,968 violations, versus none with the correct law. This deliberately inadmissible control checks probability semantics, not superiority to another legal model.

F.7 CONDITIONAL RECOVERY AND QUERY DIRECTION

We distinguish two conditional questions in the four-mode model: recovery from exact answers and recovery from a rounded transcript. For exact answers, b_1, b_2 select a least-eigenvalue direction u of $A(b)^*A(b)$, followed by $p = v^\top x$ with $v \perp u$. The consistent sources are $z(e) = y + e\bar{u}$, where $y = (b, pv)$, $\bar{u} = (0, u)$ and $|e| \le R = (\delta^2 - \|y\|^2)^{1/2}$. Neither conditional experiment replaces the whole-ball guarantees of Section F.6 or the asymptotic theorems. Both exact-fiber and direction studies use (35) and (34). Quadrature reproduces their constants within 1.5×10^{-14} relatively; over 200 backgrounds, the Gram identity (36) and eigenvalues agree within 4.7×10^{-16} and 2.8×10^{-16}, respectively.

Legal reconstruction on an exact fiber. Replace the degree-5 pair in (33) by the normalized real order-one degree-L pair, smoothed by $[1 + L(L+1)]^{-4}$, for $L \in \{3, 5, 7, 9\}$, recomputing B, C and $d_x = [L(L+1)/2 - 1]/[1 + L(L+1)]^8$ and keeping the sensor, task weights and compiler. For $\delta \le 300$ the argument of Appendix D.3, with degree-L constants, gives $h_z > .719$ and $A_{h_z} \succeq (6/25)g$. The solid-support matrix has entrywise coefficient ℓ_1 norm at most 3 and $\deg \widehat{J}_z \le 20$, so $\sup_{[-1,1]^3} |\nabla \widehat{J}_z| \le 20 \cdot 18 \cdot 3^2 = 3240$, $\widehat{J}_z \ge 36/625$ on the sphere and the extension in (34) is at least $36/625 - 3240^2/(4 \cdot 2^{45}) > 1/32$ on the cube; every reconstructed y therefore compiles to the prescribed legal law. For a fixed task, compare the projection forecast $F(y)$ with the least conditional half-risk over arbitrary forecasts from the same answers, $R_* = \inf_f \sup_{|e| \le R} \frac{1}{2}\|F(z(e)) - f\|_\tau^2$. Put $G = F - F(0)$, $s_0 = s(y)$, $d = d_x$, $d_R = s_0 - dR^2 > 0$, $A = B\bar{u} + 2C(y, \bar{u})$ and $H = C(\bar{u}, \bar{u}) + dG(y)$; then (35) gives $F(z(e)) - F(y) = (Ae + He^2)/(s_0 - de^2)$. With $\alpha = \|A\|_\tau^2$, $\eta = \langle A, H \rangle_\tau$ and $\chi = \|H\|_\tau^2$, define

$$Q(e) = (\chi s_0^2 - \alpha d^2 R^2)e^2 + 2\eta s_0 d_R e + s_0^2(\alpha - \chi R^2). \tag{58}$$

The endpoint forecasts have midpoint $f_* = F(y) + HR^2/d_R$ and half-separation squared $\alpha R^2/d_R^2$, and direct expansion yields $\frac{\alpha R^2}{d_R^2} - \|F(z(e)) - f_*\|_\tau^2 = \frac{(R^2 - e^2)Q(e)}{(s_0 - de^2)^2 d_R^2}$. If $Q \ge 0$ on $[-R, R]$, this midpoint attains the two-endpoint lower bound, hence $R_* = \alpha R^2/(2d_R^2)$; it suffices to check the endpoints and any interior minimum of the quadratic Q. For the projection decoder, maximizing first over the sign of e leaves a numerator increasing in $|e|$ over a decreasing positive denominator, so its worst case is an endpoint, giving $R_{\mathrm{proj}} = (\alpha R^2 + 2|\eta| R^3 + \chi R^4)/(2d_R^2)$.

Across 960 fibers (four degrees; radii 50, 150, 300; five background fractions, four angles and four third-answer fractions), all three task norms satisfy (58), with minimum $Q/(\alpha s_0^2) = .99766$. Stationary-point checks reproduce the extrema; independent overhead evaluations agree within 1.2×10^{-6} relatively. The largest overheads are 2.34×10^{-3}, 3.72×10^{-5}, 1.95×10^{-6} and 1.98×10^{-7} at $L = 3, 5, 7, 9$ (Figure 16). These floating-point checks measure the chosen legal decoder on exact fibers, not the minimum legal penalty or rounded-cell risk. Degree varies while the answer budget stays at three.

Selected-object overhead on a received codeword. Rounded words name a cell with thickness, not a line: in the rotated coordinates the transcript fixes two coarse slabs and the third query's slab, and the consistent sources are $\mathcal{C}_c = B_\delta(\mathbb{R}^4) \cap \{\ell_i(z) \in I_i(c)\}_{i=1}^3$, a ball met by three slabs. On that cell we compare one legal CPPZ with forecasts chosen freely for each task; both receive the same words and public model, not the source. Four radii $\delta = 150, 200, 250, 300$ and seven precisions from 4 to 24 bits per answer give 448 cells. Each radius has sixteen sources, eight at radius $.2\delta$ to $.8\delta$ and eight on the sphere, where the ball cuts the slab box to a sliver; the same sources serve every precision, and all were fixed before any result. For a selected fit $\widehat{h}$ and a specified legal output class $\mathcal{L}$, distinguish

$$R_{\mathrm{free}}^* = \max_\tau \inf_{a_\tau} \sup_{z \in \mathcal{C}_c} \tfrac{1}{2}\|F(z) - a_\tau\|_\tau^2,$$

$$R_{\mathcal{L}}^* = \inf_{h \in \mathcal{L}} \max_\tau \sup_{z \in \mathcal{C}_c} \tfrac{1}{2}\|F(z) - F_h\|_\tau^2, \qquad R_{\mathrm{fit}} = \max_\tau \sup_{z \in \mathcal{C}_c} \tfrac{1}{2}\|F(z) - F_{\widehat{h}}\|_\tau^2.$$

If $\widehat{h} \in \mathcal{L}$, then $R_{\mathrm{free}}^* \le R_{\mathcal{L}}^* \le R_{\mathrm{fit}}$; the experiment bounds the selected fit, not the infimum over legal outputs. Bounds $L_{\mathrm{free}} \le R_{\mathrm{free}}^* \le U_{\mathrm{free}}$ and $L_{\mathrm{fit}} \le R_{\mathrm{fit}} \le U_{\mathrm{fit}}$ with $L_{\mathrm{free}} > 0$ give

$$\max\left\{0, \frac{L_{\mathrm{fit}}}{U_{\mathrm{free}}} - 1\right\} \le \frac{R_{\mathrm{fit}}}{R_{\mathrm{free}}^*} - 1 \le \frac{U_{\mathrm{fit}}}{L_{\mathrm{free}}} - 1. \tag{59}$$

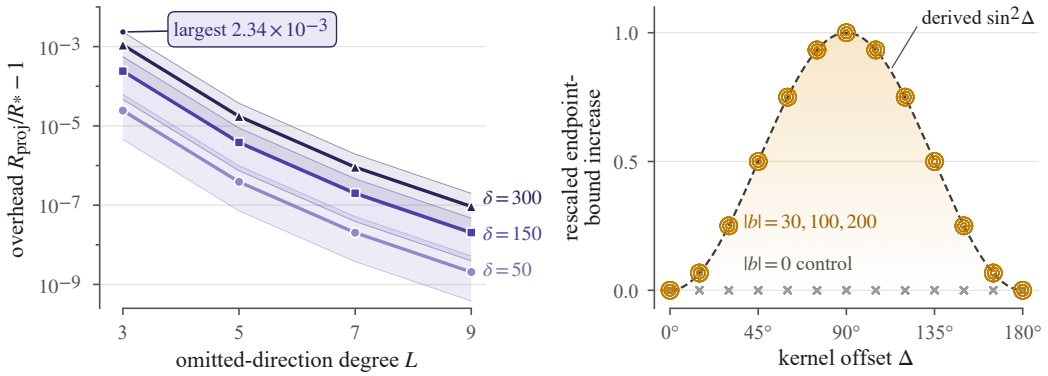


Figure 16: **Conditional recovery and direction sensitivity.** Answers are exact and unquantized. *Left*: this legal decoder's overhead relative to the conditional optimum, median and full range across fibers and task norms at each (L, δ), 960 fibers in total. *Right*: the rescaled endpoint-bound increase at third answer zero as the unmeasured direction rotates by Δ at fixed background and source ball, the indistinguishable endpoints rotating with the query kernel; Equation (60) predicts $\sin^2 \Delta$ without fitting, the $|b| = 30, 100, 200$ series coincide and at $b = 0$ the angular term vanishes.

Lower ends use feasible witnesses re-encoded to the cell's transcript, including its coarse-word-dependent branch: continuous suprema permit closing the true cell, but not importing a neighboring cell's witness. For any probability weights, half their weighted forecast variance bounds every center's half-risk below, without optimality of the finite-witness dual: for witnesses $z_i \in \mathcal{C}_c$ with $f_i = F(z_i)$, weights $\lambda_i \geq 0$ summing to one and $\bar{f} = \sum_i \lambda_i f_i$, $\frac{1}{2}\sum_i \lambda_i \|f_i - a\|_\tau^2 = \ell_\tau + \frac{1}{2}\|a - \bar{f}\|_\tau^2$, $\ell_\tau = \frac{1}{2}\left(\sum_i \lambda_i \|f_i\|_\tau^2 - \|\bar{f}\|_\tau^2\right)$, so $L_{\text{free}} = \max_\tau \ell_\tau$. Upper ends use remainder-free box covers of the numerator $P_a(z) = Bz + C(z,z) - a\, s(z)$ of the centered response $G(z) = [Bz + C(z,z)]/s(z)$, $s(z) = 1 + z^\top D z$, for a centered candidate a, at each chosen task center for U_{free} and at $F_{\hat{h}}$ for U_{fit}. With $E_{ij}^a = C_{ij} - aD_{ij}$, on a box with center c and half-widths r_i the quadratic identity has no remainder and gives $\|P_a(z)\|_\tau \leq N_{a,\tau} := \|P_a(c)\|_\tau + \sum_i r_i \|\partial_i P_a(c)\|_\tau + \sum_{ij} r_i r_j \|E_{ij}^a\|_\tau$, so the half-risk is at most $N_{a,\tau}^2/(2s_{\mathcal{B}}^2)$ whenever $s_{\mathcal{B}} > 0$ bounds s below on the box–cell intersection; here $s \geq 1 - d_b\delta^2 > 0$. Since $D = -\operatorname{diag}(d_b, d_b, d_x, d_x)$ in the rotated coordinates, this bound remains valid for a box protruding outside the ball; a box is discarded only when its nearest point is outside the ball, and surviving and budget-capped boxes retain their upper bounds, so coverage is preserved. The requested relative stopping gap is 10^{-6}.

Cutting-plane refinement alternates each cover's worst feasible witness, free-center dual optimization and legal fitting. Bounds remain valid at each step; their endpoints are evaluated in floating point, not outward rounded. In Figure 18, 256-witness fits suggest 3.2% and 131% overhead in an eleven-bit boundary cell and six-bit sliver; refinement gives raw intervals $[-.5, .8] \times 10^{-6}$ and $[.04, 2.1] \times 10^{-6}$. In 245 cells, refined uppers fall below initial lowers: fit quality improves within the same legal output class.

Selected-fit overhead is below .0035% from six bits (Table 8); 338 numerical lowers are positive, and 432 cells have no additional adverse witness at the selected tolerance. Neither certifies nonconvex fit optimality. The contrast bound $2/999 \simeq .2002\%$ for $.999 \leq \omega_\tau \leq 1.001$ ((11)) applies to arbitrary degree and does not explain the smaller four-mode values. Median absolute overhead, normalized by $\rho_0(\delta)$, falls from $[.4174, .4217]$ at four bits to $[1.701, 2.504] \times 10^{-6}$ at 24 bits; the free-risk lower median falls from 2.41×10^3 to .373. This establishes no slab-width power law. Boundary-cell overhead is no larger and, from eleven bits, at most one quarter of interior overhead; shorter fibers are a possible, unisolated explanation. At 18 bits only 27/64 covers close, widening intervals rather than establishing greater overhead. An independent implementation at tolerance 10^{-4} encloses all refined intervals, preserves lower/upper ordering and agrees on legal uppers within 9.9×10^{-5} for its 385 closed cells. This checks consistency, not strict positivity.

Table 8: Selected-object overhead on 448 received-word cells: numerical evaluations of (59), in units of 10^{-6}, with negative raw lower estimates shown before clipping at zero. "Closed" counts covers meeting the numerical tolerance and "Lower > 0" positive lower estimates.

Bits	Half-width/δ	Median enclosure	Largest upper	Closed	Lower > 0
4	6.7×10^{-2}	$[172.9, 174.7]$	340.5	63/64	64/64
6	1.6×10^{-2}	$[9.3, 11.8]$	19.0	64/64	64/64
8	3.9×10^{-3}	$[-.3, 1.3]$	5.7	64/64	9/64
11	4.9×10^{-4}	$[2.8, 4.2]$	17.9	64/64	49/64
16	1.5×10^{-5}	$[3.6, 5.4]$	26.7	59/64	54/64
18	3.8×10^{-6}	$[1.4, 8.3]$	33.1	27/64	40/64
24	6.0×10^{-8}	$[3.6, 5.2]$	20.9	64/64	58/64

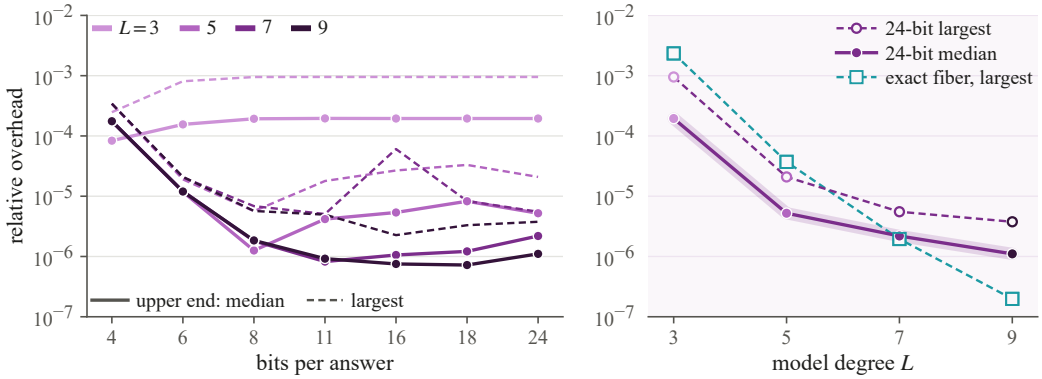


Figure 17: **The same 448 cells under four model degrees.** *Left*: median (solid) and largest (dashed) numerical upper end of the overhead per precision. *Right*: the 24-bit overhead on received codewords against the largest exact-fiber overhead of Figure 16, by degree; from $L = 7$ the codeword values sit at the enclosure's resolution.

Degree dependence. The same 448 cells are evaluated at $L = 3, 5, 7, 9$ (Figure 17), fitting in units of free risk. At $L = 7$, unscaled optimization stops at objects 24 and 55 times worse than the free forecast in two coarse cells; rescaling reduces overhead to 2.2 and 2.5×10^{-4}. At $L = 5$, rescaling changes no upper end by more than 3×10^{-7}. Largest upper estimates are 9.5×10^{-4} at $L = 3$ and 3.4×10^{-4} at higher degrees. The 24-bit median decreases from 1.9×10^{-4} to 5×10^{-6} and then reaches numerical resolution (2×10^{-6} at $L = 7, 9$). This resembles the exact-fiber trend but does not resolve the limiting rounded-cell penalty.

Query alignment and the endpoint bound. In Proposition 5.2 the first two answers choose which high-mode direction the third query leaves unmeasured; (60) gives the endpoint cost of rotating that choice. In the degree-five model, fix the source ball and exact background $b = t(\cos\varphi, \sin\varphi)$, leave $u_\Delta = (\cos\psi, \sin\psi)$ with $\psi = \varphi/2 + \Delta$ unmeasured and measure its perpendicular. At third answer zero the indistinguishable endpoints are $z_\pm(\Delta) = (b, \pm R u_\Delta)$, $R^2 = \delta^2 - t^2$; these inputs change with Δ, so this is not a fixed-input loss comparison. With common normalizer $s_{\mathrm{end}} = 1 - d_b t^2 - d_x R^2$ and $u_\Delta^\top M(b) u_\Delta = -t + 2t\sin^2\Delta$, (35)–(36) give the endpoint lower bound

$$\|A(b)u_\Delta\|^2 = \lambda - at + qt^2 + 2at\sin^2\Delta,$$
$$\mathcal{L}_b(\Delta) := \tfrac{1}{8}\|F(z_+(\Delta)) - F(z_-(\Delta))\|^2 = \frac{R^2\|A(b)u_\Delta\|^2}{2s_{\mathrm{end}}^2}. \tag{60}$$

The Gram spectrum is unchanged. For $t > 0$ the rescaled increase $s_{\mathrm{end}}^2[\mathcal{L}_b(\Delta) - \mathcal{L}_b(0)]/(atR^2)$ is therefore $\sin^2\Delta$, with no fitted parameter. The 676 configurations use norms $0, 30, 100, 200$, thirteen background angles and thirteen offsets. Tensor and formula energies agree within 9.2×10^{-16} relative to λ; chord risks agree within $8 \times 10^{-16}\rho_0$, and the normalized increase matches

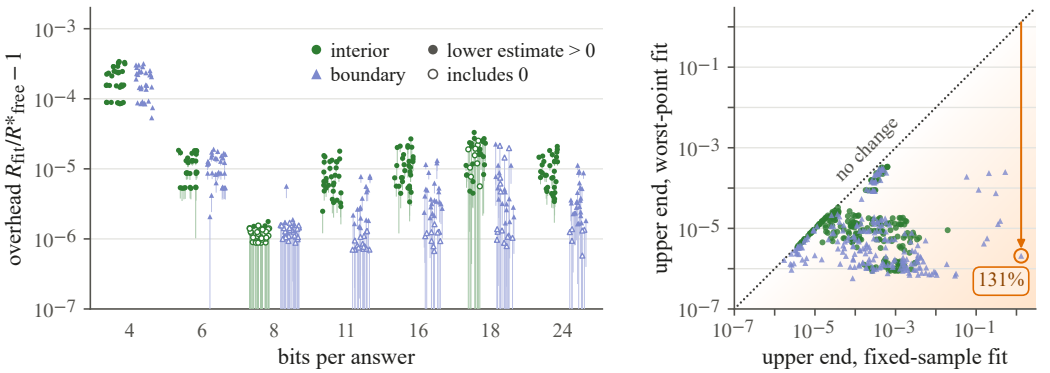


Figure 18: **Conditional overhead of selected legal objects.** *Left*: every cell's enclosure (59), from its numerical lower to its numerical upper end, with interior cells left and sphere-boundary cells right of each precision; filled markers have positive lower estimates and hollow markers include zero numerically, and lower ends below 10^{-7} are drawn to the axis. *Right*: each cell's upper end from a single fit on 256 fixed witnesses against the cutting-plane refinement (dotted: no change). Every cell falls; the six-bit boundary sliver that read 131% (labeled) falls furthest.

$\sin^2 \Delta$ within 4.8×10^{-15} (Figure 16). A 180-degree offset leaves the query line unchanged; zero background and zero-background linearization remove the angular increment, with residuals below $6.0 \times 10^{-16}\rho_0$.

Geometry-dependent direction selection. At fixed background, compare three choices of unmeasured direction: the least eigenvector of $M(\widehat{b})$, its greatest eigenvector, and the mean over 48 uniformly spaced fixed directions. Each uses the same three-query budget but has its own kernel and endpoint inputs; the score is (60), not complete recovery risk. Uniform averaging over background angles makes $u^\top M(b)u = -t\cos(2\psi - \varphi)$ vanish, so the uninformed mean equals the best fixed direction's grid-average score, not its risk at each input. For six radii 50–300, six background fractions $0, .1, .25, .5, .75, .9$ and 24 angles, the uninformed/geometry ratio averages 1.04 at fraction .1 and 1.49 at .9; worst-direction ratios are 1.08 and 1.98 (Figure 19). At radius 300 and fraction .9 they reach 1.96 and 2.92. Background rounding uses the effectively exact 33-bit grid with step $\delta/(2^{32} - 1)$. Zero-background unweighted controls have spread 6.7×10^{-16}. Across 192 task-weighted configurations, pooled ratios change by less than 5×10^{-8} relatively, but the weighted null has residual 7.7×10^{-4}: unweighted isotropy does not imply weighted isotropy. An unrelated background gives ratio 1.26 versus 1.25 for an uninformed direction, checking removal of geometric information rather than equal-information decoding.

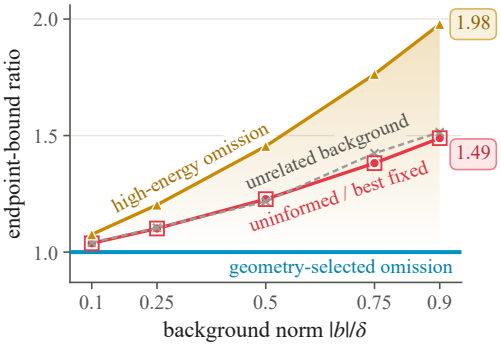


Figure 19: **Geometry changes which direction can be left unmeasured.** Ratios of the conditional endpoint lower bound (60). The geometry rule measures the perpendicular to its selected low-energy direction. The unrelated-background control removes geometry information. The selected/high-energy relative tensor–formula discrepancies are at most $1.93 \times 10^{-15}/9.66 \times 10^{-16}$.

Direction precision versus recovery precision. A B-bit answer on $[-\delta, \delta]$ has step $2\delta/(2^B - 1)$ and maximum rounding error $\nu = \delta/(2^B - 1)$. Selecting half the received background angle gives direction error of order $\nu/|b|$ away from zero. Across radii 100/300, fractions $.1, .25, .5, .9$, sixteen angles and ten precisions from 3 to 32 bits, four bits are the first tested precision retaining at least

90% of the 32-bit direction advantage in every fraction's mean; six retain at least 99.67%. A grid with half that step requires $B + 1$ bits and is not an equal-bit comparator.

Direction accuracy does not ensure recovery accuracy. At $\delta = 300$, fractions $.1, .9$ and 24 angles, encode the endpoint sources $(b, \pm\sqrt{\delta^2 - |b|^2}u(b))$ using 3, 6, 11, 18 or 32 background bits and a 32-bit third answer. Complete ball-projected common-law recovery over these 480 cases reaches $5252.6\rho_0$ at fraction .1 with three background bits and $128.3\rho_0$ at .9 with six bits, despite small direction errors. Independent cofactor forecasts agree within 2.1×10^{-16}. Coarse-geometry error therefore remains distinct from direction error; these sampled risks do not replace the uniform bounds of Section F.6.

F.8 Nonspherical recovery: request feasibility and code reuse

Bases $h_0(u) = 1 + u^\top Qu$ have random orientations and centered Gaussian eigenvalues scaled to $\|Q\|_{\mathrm{op}} \sim \mathrm{Unif}[.02, .12]$. Six development and twelve independent confirmation backgrounds are fixed without performance selection. With 32 degree-2–5 modes, $(r, \beta, M_\beta) = (8, 3, 8)$ and $\delta = .4(1 - 3\|Q\|_{\mathrm{op}})/C_{32}$ (Section F.1), support and curvature on the whole ball exceed $.824$ and $.384$. Nested $K = 2, 4, 8$ tasks use $\omega_j(u) = .1 + .9\,6e^{6v_j^\top u}/\sinh 6$ with independent uniform axes. Budgets $m = 8, 16, 24$ use 20-bit answers and requests $r_{j,m} = \lambda_{m+1}(\delta^2 DF(0)^* M_{\omega_j} DF(0))$, twice singleton tangent optima rather than application tolerances.

Joint request feasibility. With $G_k = \delta^2 DF(0)^* M_{\omega_k} DF(0)/r_{k,m}$, the leading risk/request is bounded below by $\kappa_{\mathrm{LB}} = \max_\pi \frac{1}{2}\lambda_{m+1}(\sum_k \pi_k G_k)$. Search can only understate this bound. It exceeds one in 68/108 settings: 3/36 with two tasks, 29/36 with four and all 36 with eight. In 62 settings the excess exceeds 4%, larger than the measured nonlinear correction (at most 3.44%). Corrections are positive in all 162 development and confirmation settings. Singleton-based requests therefore account for many observed violations. Numerical finite-radius bounds below also exclude 60 of these requests for every deterministic adaptive-linear scheme, quantized or not. The table's counts are therefore not failure rates on certified feasible requests; passing the necessary screen is not a feasibility certificate.

Recovery rules and observed risks. The fixed-study rules in Table 9 all compile the prescribed law. Balanced fixed uses the request-normalized mean Gram; two-batch updates the tail from five degree-two answers, while its matched control keeps the zero-background tail. Both preserve the coarse answers and project the tail to the remaining radius; other fixed codes project radially. Optimized fixed is a local Grassmann solution trained on sphere inputs and 64 signed axes. Evaluation pools 512 fresh sphere/interior inputs, those 64 axes, 32 coarse/tail stress inputs and 80 shared attack witnesses (688 per setting). Optimization uses the rational response; evaluation uses byte decoding and compiled cofactor density. This pooled set is not an independent confirmation sample.

Table 9: Nonspherical common-law recovery. Median, across 108 settings, of the maximum-found task risk/request; the last column counts settings with a violating input, and the remaining settings are unresolved, not certified feasible. The last row, added after the frozen study, is the request-aware common projection; 68 of its 73 violations are requests that no shared code meets at leading order. Absolute maximum-found half-risks range from 3.36×10^{-13} to 7.08×10^{-9}.

Measurements	Median risk/request	Violating settings
Harmonic prefix	2.995	108/108
Balanced fixed	1.494	78/108
Matched fixed prefix	1.828	90/108
Two-batch spectral	1.801	90/108
Optimized fixed	1.504	79/108
Request-aware common projection	1.243	73/108

For these spectral rules, two batches give no consistent gain: median matched-fixed/two-batch risk is 1.00462, versus $.90163$ for the post-evaluation best-fixed reference; only 4/108 settings beat the latter by $.1\%$. These spectral rules do not implement the general capacity optimizer.

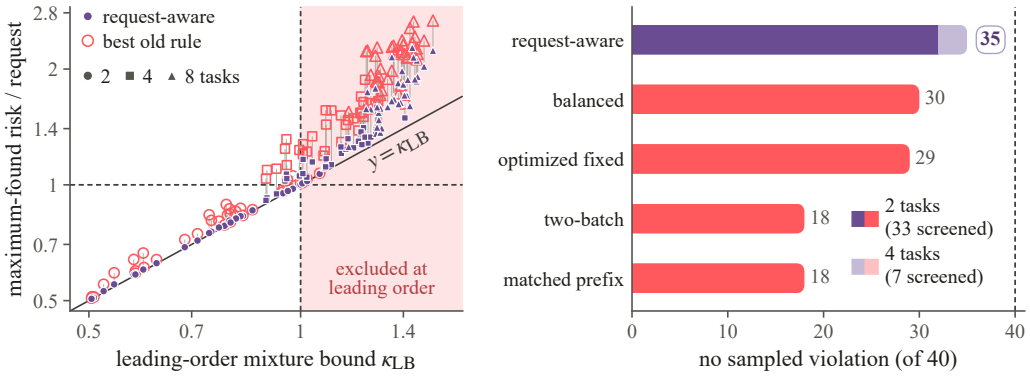


Figure 20: **Leading-order screening and sampled common-law recovery.** *Left*: each confirmation setting's leading-order mixture bound κ_{LB} versus the maximum-found risk/request of the best original rule (red ring) and the request-aware common projection (purple dot), evaluated on the same pooled inputs; marker shape gives the task count. The shaded region denotes leading-order exclusion, not an independently certified finite-radius exclusion for every plotted setting. *Right*: among 40 requests passing the screen, the counts with no sampled violation; shading distinguishes the 33 two-task and seven four-task requests. Passing the screen and finding no violation do not certify feasibility. The comparison precedes the three subsequent recoveries reported in the text.

Finite-radius exclusion bounds. The rational response $f(x) = (Ax + C(x,x))/s(x)$ has exact odd part $f(x) - f(-x) = \alpha(x)\,2Jx - e(x)$, with $\alpha = s_0(s_0+q)/((s_0+q)^2 - \ell^2)$ and $e = 2C(x,x)\ell/((s_0+q)^2 - \ell^2)$, where $q = x^\top Dx$ and $\ell = d^\top x$. On the unit sphere of the maximizing mixture's top $(m+1)$-eigenspace V, $\alpha \geq s_0/(s_0 + \lambda_{\max}(V^\top DV))$ and $\|e\|$ is at most the norm of the restricted quadratic tensor times $\|V^\top d\|$, which gives the lower bound $(\alpha_{\min}\sqrt{\kappa_{\mathrm{LB}}} - e_{\max})^2$. A second sphere removes the remainder instead of bounding it: on an $(m+1)$-plane in $d^\perp$, $\ell \equiv 0$, so $e \equiv 0$ and $\alpha = s_0/(s_0+q)$; compressing to the 31-dimensional $d^\perp$ costs at most one eigenvalue by interlacing, and the plane is the top eigenspace of the compressed maximizing mixture. Either sphere bounds the risk from below at any antipodes with identical words. Such antipodes exist by Borsuk–Ulam for continuous codes and, for adaptive linear schemes, in the unmeasured space of the all-zero answer path, which has codimension at most m and therefore meets the plane; the table's rules are quantized and so covered. The better bound, evaluated in floating point, covers all 23 infeasible requests at $m = 8$, 16 of 20 at $m = 16$ and 21 of 25 at $m = 24$; the other 8 remain leading-order statements.

A post-study common-projection rule minimizes worst normalized tangent risk, attaining κ_{LB} within 4×10^{-10} on all 54 two-task development/confirmation settings by Lemma D.2. For four/eight tasks it uses heuristic Grassmann search, not Theorem 4.2's construction. Encoding, precision and decoding are unchanged; sixteen added witnesses are replayed across all rules (704 inputs), without increasing original maxima. Of the 40 requests passing the leading-order screen, 35 have no sampled violation (32/33 two-task), versus 30 for balanced fixed, 29 for optimized fixed and 18 each for matched/two-batch recovery (Figure 20). Nonlinear corrections are .08–3.4%; the two-task miss has only .02% tangent slack. Among balanced/optimized fixed violations, 68 are excluded at leading order and 60 at finite radius. Additional mean-Gram preconditioning, six full-spectrum restarts and within-space mean-Gram rotation resolve three four-task requests among the five remaining, with uniform upper/request ratios .999660, .991171, .987171. Two retain violations; the original 35 retain their sampled success.

Fresh-background evaluation. Across 432 new backgrounds (3,888 settings), the leading-order screen excludes 2,270 requests (58%); 2,115 exceed one by more than the largest measured nonlinear correction (4.03%). The two-sphere test excludes 1,998 at finite radius, leaving 272 leading-order exclusions unresolved by that test. All witnesses are replayed across rules. For the 1,618 requests passing the screen, common projections meet 1,456 on sampled inputs (90%), versus 1,217 for balanced fixed, 1,147 for optimized fixed, 638 for two-batch and 631 for matched recovery. This

comprises 1,215/1,216 two-task and 241/398 four-task requests; no original rule meets more than 89 four-task requests. Of the 162 misses, 133 already violate tangent requirements; nonlinear corrections up to 2.5% explain 29 others, including one two-task case. Two-task constructions approach the bound within 4.2×10^{-10}; median ratios decrease from 1.39 to 1.18 for four tasks and 2.17 to 1.73 for eight. These are sampled, not uniform, feasibility results.

Uniform bounds and search sensitivity. After refinement, the 1,618 screened requests divide into 1,232 with sufficient numerical uniform bounds, 325 unresolved without found violations, and 61 with violations of the tested recovery—not exclusion of every code. Of the 1,456 original sampled successes, 1,179 have uniform bounds without changing their projections; stronger searches on the other 277 reveal four violations. Reoptimization of the 162 original misses recovers 105 sampled successes and adds 53 uniform bounds, using old witnesses, longer taskwise attacks and 2,048 fresh sources. Median projection-refinement time is 11.5 s; the additional computation precludes a matched-budget superiority claim over the fixed-study rules.

The bounds use the exact rational response above, not a sampled supremum. After task weighting by $1/\sqrt{2r_i}$, let $N(z) = Bz + C(z,z)$, $s(z) = s_0 + d^\top z + z^\top Dz$ and $z = Qu + Kv$, with $t = \|v\|$, $a = \sqrt{1-t^2}$. For $d_Q = \|Q^\top d\|$, $d_K = \|K^\top d\|$, $d_{QQ} = \|Q^\top DQ\|$ and the restricted tensor norms of Section F.1, set $s_Q(t) = s_0 - d_Q a - d_{QQ}a^2$, $s_z(t) = s_Q(t) - d_K t - 2d_{QK}at - d_{KK}t^2$ and $\Delta_s(t) = d_K t + 2d_{QK}at + d_{KK}t^2$. When these lower denominators are positive, the triangle numerator bound in (54) gives the first sufficient test. For unresolved refined plans, retain cross terms instead: $\|N(z) - N(Qu)\| \le t\sqrt{b_K^2 + 4ah_Q + 2th_K + (2ac_{QK} + tc_{KK})^2}$, where h_Q, h_K bound the unit-vector contractions $|\langle BKv, C(Qu, Kv)\rangle|$ and $|\langle BKv, C(Kv, Kv)\rangle|$. Symmetrizing repeated v indices and taking matrix-unfolding norms bounds these contractions; their minimum with $b_K c_{QK}, b_K c_{KK}$ remains valid. Divide this numerator bound by $s_z(t)$ and add $(b_Q a + c_{QQ}a^2)\Delta_s(t)/(s_z(t)s_Q(t))$.

For rounding, put $\nu = (2^{20}-1)^{-1}$ and $\eta = \sqrt{m}\nu$. A global Lipschitz bound times η applies after ball projection. If $t^2 \ge 2\eta - \eta^2$, every rounded word is already in the ball. In that regime the smaller restricted bound is also valid:

$$\frac{\nu \max_{\epsilon\in\{-1,1\}^m} \|BQ\epsilon\| + 2c_{QQ}\eta}{s_0 - d_Q - d_{QQ}} + \frac{(b_Q + c_{QQ})(d_Q + 2d_{QQ})\eta}{(s_0 - d_Q - d_{QQ})^2}.$$

We enumerate corners for $m \le 16$ and otherwise upper-bound their squared norm by the sum of absolute entries of $Q^\top B^\top BQ$. Combining the root errors before squaring retains the cross term. Intervalwise scalar envelopes use 20,000–100,000 intervals; all tensor norms, coefficients and orthogonality checks use floating point, not rational certification.

Code reuse and cost. Eight-task codes serve 64 new convex mixtures with unchanged bytes and law. Mixed-context risks agree with weighted original risks within 4.8×10^{-15} relatively, checking integral linearity without source access. Polynomial/rational forecasts agree within 4.4×10^{-16}; quadrature refinement changes task Grams by less than 1.2×10^{-14} relatively. Payloads are 20/40/60 bytes at $m = 8, 16, 24$; full 32-query recovery is an 80-byte endpoint control. On an RTX 5090, warm batches of 64 objects take about 2.2 ms each for two-batch encoding/decoding, versus .16/.19 ms for balanced fixed, plus about 1.1 ms for common density normalization and prediction. Payloads exclude public models, query descriptions and preprocessing; timings exclude cold starts, sensing and transmission.

F.9 Finite-budget tests of spherical two-batch heuristics

These pilots test gains from a fixed coarse prefix and spectral routing, not the asymptotic emulation construction of Theorem 5.1. That theorem does not promise strict improvement over fixed acquisition. We compare catalog-routed, matched-fixed and full-fixed acquisition on Section F.1's model. Five rounded, ball-projected degree-two answers select the nearest center in $\{0, \pm e_i/2 : 1 \le i \le 5\}$ for the remaining queries, with fixed ties and no branch tag. Matched fixed retains this prefix and the zero-background tail; full fixed uses only zero-background queries. All use identical quantization and recovery.

At $m = 16, 24$, residual spectra validate spaces failing direct dimension/residual checks; grid-optimality corrections of 4.58/126 times the grid lower bound are inconclusive. Full-fixed up-

per/lower differences are below 4×10^{-9} on both quadratures. Disjoint validation and test sets each contain 24 unit sources: eight random and eight each at coarse norms $.75, .25$. Twelve stress inputs combine $0, \pm e_1/2$ with signed endpoint tails. No test inputs select parameters. Table 10 reports mean per-input maximum risk, not worst-input risk.

Table 10: Two-batch acquisition on 24 held-out sources: mean per-input maximum task risk and its ratio to matched- or full-fixed recovery (below one favors two batches). Bits exclude query and catalog storage; full-fixed recovery at $m = 5$ was not recomputed.

m	Bits/answer	Risk	Matched ratio	Full ratio
5	15	$9.3712\,10^{-14}$	1.000000	—
8	18	$2.9648\,10^{-14}$	.999529	1.015391
12	21	$2.9391\,10^{-15}$	1.000007	1.562011
16	23	$8.7693\,10^{-16}$	1.000411	.938665
24	26	$7.5212\,10^{-17}$	1.000561	1.815983

Plan quality and routing resolution. Neither original nor preconditioned plans yield consistent cross-budget gains. All 60 preconditioned tangent plans lie within 3.8×10^{-7} of their numerical mixture brackets. At $m = 24$, preconditioning changes the two-batch/full-fixed mean-risk ratio from 1.816 to $.9382$ on the original 24 inputs; on 512 fresh inputs it is $.9915$ ($.9952$ against matched fixed). Table 10 retains the original protocol. Its per-input matched ratios span $.991$–1.027 (median one), with 11/24 nondefault routes. The catalog's $.25$ routing threshold cannot resolve the two witnesses of Proposition C.1, whose coarse norms are below 6.953×10^{-4}, although the answer grid can. Over 384 replays, relative risk/quadrature discrepancies are below 1.7×10^{-9}/5.4×10^{-9}.

Table 11: Finite bounds for a five-query prefix. $L_5/U > 1$ excludes matching the fixed-code upper bound regardless of later feedback. Values are not interval-certified.

m	$L_0(m)$	U	U/L_0	L_5/U
8	$3.2602303\,10^{-13}$	$3.2855555\,10^{-13}$	1.007768	.992292
12	$1.0286364\,10^{-14}$	$1.0367022\,10^{-14}$	1.007841	1.688089
16	$6.6724179\,10^{-15}$	$6.7225454\,10^{-15}$	1.007513	1.270289

Prefix limitation and direct updates. For $U = \max_i U_i(Q)$, Table 11 bounds any adaptive-linear improvement by $1 - L_0/U < .8\%$. At $m = 12, 16$, $L_5 > U$ excludes matching this fixed-code upper after the five-query prefix; $m = 8$ and other designs remain unresolved. These are worst-input, not mean-risk comparisons. Catalog-free derivative/projection updates from the same five answers also fail the predefined criterion of improvement over both fixed references at more than one budget. Among 108 inputs (24 reused, 24 fresh, 12 stress, 48 witnesses), fresh maximum-found direct/matched ratios are $1.0008063, 1.0001220, 1.0007728$ at $m = 8, 12, 16$; pooled direct/full-fixed ratios are $1.000000349, 2.230872116, 3.261497606$, despite gains on some reused subsets. Searched maxima are not minimax risks; the asymptotic and small-radius results do not guarantee gains at $\delta = 1$.

Common-law implementation. Factored cofactors and exact moments evaluate $R_H = \widehat{J}_H + C(\|u\|^2 - 1)^2$, $C \approx 3.28 \times 10^{10}$, avoiding cancellation; forecasts agree with direct geometry within 3.06×10^{-16}. Answer-bit counts exclude acquisition infrastructure.

F.10 Quartic model and independent prediction checks

These complementary checks compare predictions with independent labels and test sensitivity to model shift. They use quartic constraints and radial mean risk, not the native Bessel worst-input model:

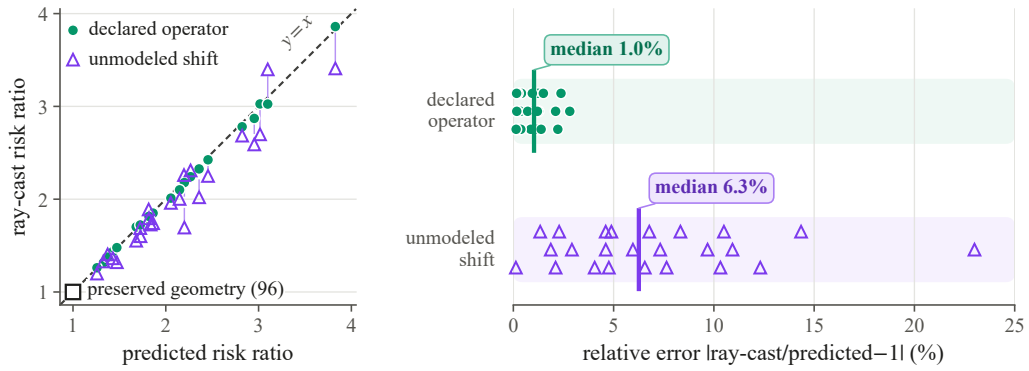


Figure 21: **Independent labels check prediction and model mismatch.** *Left*: ray-cast against predicted risk ratios for the declared operator and under an unmodeled shift. *Right*: their relative errors and medians. The 96 preserved-geometry cells coincide at one despite 15–19% weight changes (subspace change 2.3×10^{-16} or less). These are prediction checks, not minimax or rectangularity tests.

$$\mathcal{M}_{b,z} = \left\{ x : x^\top Q_b x + \lambda_b \sum_{j=1}^{3} (R_b x)_j^4 + \epsilon_b \sum_{|\alpha|=4} z_\alpha c_\alpha x^\alpha = 1 \right\}, \tag{61}$$

with positive triaxial Q_b, non-axis-aligned R_b, $\lambda_b > 0$, 15 normalized degree-four monomial deformations and $\epsilon_b = .5\delta_{\max,b}$, where $\delta_{\max,b}$ is the smallest of .06, a grid positivity bound for the leading quartic form and a global Hessian convexity bound; over 21 bases the radii lie in $[.00163, .00327]$ and the minimum Hessian margin is .6253. At radial intersection r a fixed detector has $Y \mid b, z, u \sim \text{Bernoulli}[\Phi((r_{b,z}(u) - r_{b,0}(u) - \sigma_b)/\sigma_b)]$, with $\sigma_b = .02\,\text{median}_u\, r_{b,0}(u)$, and vMF contexts u have concentration five.

Evaluation uses six held-out bases from a fixed 12/3/6 disjoint-base partition, with 96 unseen directions, 512 deformations per base and $m = 2, 4, \ldots, 12$. These compilers are task-conditioned; the prediction checks do not compare shared and task-conditioned coding. Held-out oracle modes are evaluation-only. The dense 15×15 operator is not a scalability test; root and compiled-oracle discrepancies below 3.7×10^{-14} check numerical fidelity, not model accuracy.

Independent labels and model shift. Independent Bernoulli samples (292,263/242,129) give half-Brier intervals $[6.020, 9.519], [5.834, 9.216]$ for analytic values $7.004, 6.771$ at $m = 4, 8$ (all $\times 10^{-5}$), using $\widehat{r}(a) = \frac{1}{2}[(a-Y)^2 - (\eta - Y)^2]$. Counts follow analytic effect/variance; labels never select representations. Across 48 ray-casting cells (six held-out constraints; $m = 2, 4, 8, 12$), predicted ratios 1.26–3.83 agree with measured ratios within median/maximum $1.0/2.8\%$. Unmodeled directional shift raises the median to 6.3% (Figure 21); absolute budget errors are -4.1 to -7.2% (worst -23%), versus per-base declared-operator errors $-.93$ to $-.59\%$. Independent ray casting has second-order tessellation convergence, no missed rays and fixed/full Möller–Trumbore agreement to 10^{-12}. A separate proportional-weight control, distinct from the preserved-geometry cells, has price/subspace errors $0, -1.8 \times 10^{-16}$; these check identities, not a phase transition.

Secant bracket. Secant constants span $[.229, .300]$ over a nineteenfold radius range; linear fiber-diameter error is 1.8×10^{-15}. All 210 in-radius cells of 294 satisfy $(1-\vartheta)^2 \le \rho^{\text{nl}}/\rho^{\tan} \le (1+\vartheta)^2$ for (7)'s factor ϑ (ratios $1.014844, 1.000019$ at $\delta = .75, .04$). Beyond δ_{cb}, $\vartheta > 1$ makes the lower bound vacuous. This checks the bracket, not the exact nonlinear body.

F.11 Request-dependent shared recovery under mean risk

On Section F.10's quartic family and radial detector, we test request-dependent shared plans under mean half-Brier excess. The sender knows z; the receiver knows the base, tasks, detector

and plan. Neither encoding nor common-law recovery uses the realized task. This tests neither worst-input capacity nor an area-law advantage. Sources are uniform on $B_1(\mathbb{R}^{15})$. Six further independent bases, with uniform $[.82, 1.28]$ axes, Gaussian-QR rotations and uniform $[.055, .14]$ quartic strength, are disjoint from the one-base pilot run and from the 21 bases of Section F.10. They keep quartic coefficients and carrier Hessian positive on the whole ball, independently of risk, through $\epsilon_b = \frac{1}{2}\min\{\lambda_b/(3B_\phi),\, 2q_{\min}/B_H,\, .06\}$, where, in that section's monomial normalization, $B_\phi^2 = \sum_{|\alpha|=4} c_\alpha^2 \prod_j (\alpha_j/4)^{\alpha_j}$ (with $0^0 = 1$) is the global directional bound of the coefficients c_α of (61) and $B_H^2 = R^4 \sum_{ij,\alpha} [c_\alpha \alpha_i(\alpha_j - \mathbf{1}_{i=j})]^2$ is their coefficientwise Frobenius Hessian bound on B_R, with $R = q_{\min}^{-1/2}$ and $q_{\min} = \lambda_{\min}(Q_b)$.

For context axes $(1, 0, 0)$ and $(0, 0, 1)$, request ratio $a \in \{1/4, 1/2, 1, 2, 4\}$, $v = (a^{-1/2}, a^{1/2})$ and tangent task Gramians $A_i = J^\top W_i J$, with J the derivative of the finite probability map at zero, the request is $r = 1.10\, t_m(v)v$, $t_m(v) = \frac{1}{34}\max_{s\in[0,1]} \sum_{j>m} \lambda_j((1-s)A_1/v_1 + sA_2/v_2)$, where $34 = 2(15+2)$ combines half-Brier loss and uniform-ball covariance; t_m normalizes tangent mean risk, not a nonlinear lower bound. Request-spectral recovery uses the maximizing mixture's leading m directions. Comparators share its decoder: a nonlinear-mean-risk-optimized rank-m plan, reuse of that plan at $a = 1$, and unweighted radial-Gram projection. Optimization minimizes log-sum-exp normalized risks from request-spectral, balanced-spectral and random starts; validation chooses a start or checkpoint by maximum normalized task risk.

Independent training/validation/test sets contain $128/512/4{,}096$ sources per base. An orthonormal $Q \in \mathbb{R}^{15\times m}$, selected without test data, sends unsigned 16-bit $Q^\top z$ on $[-1, 1]$. The receiver synthesizes with Q, projects onto B_1 and decodes one quartic fiber with identity readout and area law $dA = \frac{r(u)^2 \|\nabla f(r(u)u)\|}{\nabla f(r(u)u)\cdot u}\, dA_{S^2}(u)$. Both forecasts use this reconstruction and its numerical normalizer. For $P = QQ^\top$, reflection $Rz = (2P - I)z$ preserves the source law and $Pz = PRz$; every decoder of these measurements, rounded or not, therefore obeys $\frac{1}{2}\mathbb{E}\|F_i(z) - a_i(Pz)\|^2 \geq \frac{1}{8}\mathbb{E}\|F_i(z) - F_i(Rz)\|^2$. Average the losses at z and Rz and drop the nonnegative midpoint error; rounding cannot add information.

Table 12: **Balanced-reuse/request-spectral normalized radial mean-risk ratios.** Each method's task risks are normalized by that base's frozen request before pooling. Intervals are simultaneous over all 15 cells (critical value 2.6864 of the maximum standardized log-ratio statistic), from 2,000 paired hierarchical bootstrap draws of the six bases (the independent units) and then sources, preserving all method, request and task pairings. These are not minimax prices.

m	$a = 1/4$	$a = 1/2$	$a = 1$	$a = 2$	$a = 4$
4	1.421	1.329	1.012	1.361	1.535
	[1.280,1.578]	[1.251,1.413]	[.983,1.042]	[1.295,1.431]	[1.422,1.658]
8	1.775	1.389	1.000	1.359	1.732
	[1.707,1.846]	[1.356,1.424]	[1.000,1.000]	[1.323,1.397]	[1.627,1.845]
12	1.763	1.320	1.000	1.275	1.649
	[1.561,1.991]	[1.223,1.426]	[1.000,1.000]	[1.243,1.309]	[1.547,1.758]

Request-spectral and request-optimized codes meet all 90 test-mean requests, versus 18 balanced requests for reuse and none for geometry projection (Table 12). Every off-balance reuse cell has a reflection lower estimate exceeding a request. Request-spectral risk/request spans .8948–.9462, relative risk to the optimized plan .9885–1.0000, and mean absolute task risk 1.10×10^{-6}–4.11×10^{-5}. This supports request-specific planning, not superior CPPZ subspaces, nonlinear encoding or all-source compliance. Quadrature refinement (1,536 to 3,072 directions) changes risks by at most .085%, below the 5% unresolved threshold. Estimates are neither interval-certified nor independently labeled: task 1 at $m = 8, a = 4$ needs 6,252,012 labels, exceeding the five-million cap (effect 3.6838×10^{-6}, variance 1.3803×10^{-6}; normal-approximation 95% interval, 25% relative half-width).

Resource scope. Answers occupy $8/16/24$ bytes; query matrices $480/960/1{,}440$, public base data 296 and decoded deformations 120, excluding implementation overhead. Matrix-plus-answer

cost beats fifteen 16-bit coefficients at $22/69/241$ objects per base/request/plan ($m = 12$ ties at 240). These are not bit-optimality or noisy-sensing guarantees.