*Statistics and Computing*

# Consecutive Failure Risk and Volume Frontiers of Multivariate Prediction Regions at Fixed Joint Coverage

**Online Resource 1 (Supporting Information)**

Yixiang Wang* Haoyu Jiang

School of Information Engineering, Jiangxi University of Technology, Nanchang 330098, China

*Corresponding author: Yixiang Wang (Wangyixiang0712@outlook.com).

---

**Summary.** This document accompanies the main manuscript. It provides complete proofs for the all-shape risk–volume frontier, finite-partition error bounds, smooth counterexamples, and small-budget boundary expansions. It also covers finite-order flatness rates, parametric and global rectangular derivations with interval certificates, finite-history transfer bounds, and dataset-level tables.

**Guide to appendices and code.** Appendix A covers Theorem 1, Proposition 5, and Proposition 1. It also presents exact rational box search and the Fourier infimum construction. Appendix B verifies Theorem 2, the all-budget logistic solution, general odd-order boundary rates, and an infinite-order flat-density example. Appendix C derives Theorem 3, Lemma 1, Corollary 1, Theorem 4, Proposition 2, Theorem 5, Proposition 6, and the common-factor model. Appendix D details convex interval certificates and a frontier jump example. Appendix E verifies Lemma 2 for simultaneous validation bounds. Appendix F verifies Proposition 7 and the minimax lower bound for cross-period dependence. Appendix G lists per-dataset frequencies in Table S2. All scripts and frozen outputs appear in the accompanying electronic supplementary source package (`ESM`).

**Computational setup and precision checks.** All scripts in the `ESM` package were verified under `Python 3.9.13`, `NumPy 2.0.2`, `SciPy 1.13.1`, `Matplotlib 3.9.4`, and `mpmath 1.3.0`. One-dimensional all-shape frontiers use exact rational arithmetic in the standard library. Third-order boundary quadrature uses 65- and 85-digit precision, and three-dimensional non-Gaussian rectangles use 60-digit outward-rounded decimal interval arithmetic. All randomized simulations and block bootstraps fix random seeds for exact reproduction.

---

## A Approximation of the all-shape frontier and exact coverage

*Proof of Theorem 1.* Rewrite volume as a cost with respect to marginal mass. Take the Lebesgue decomposition $\mu = f\,dx + \mu_s$, and pick a zero-volume Borel set $S$ carrying $\mu_s$. Restricting regions to $E = S \cup \{f > 0\}$ preserves mass while saving volume. On $E$, volume equals $\int_B a\,d\mu$, where $a = 0$ on $S$ and $a = 1/f$ on $\{f > 0\} \setminus S$. Order the atomless part of $\mu$ by increasing unit-mass cost $a$, breaking ties via a fixed Borel injection $\mathbb{R}^M \to (0, 1)$. That injection induces a continuous law on every positive-mass level set of an atomless measure. Any prescribed mass can thus be collected cleanly, and the ordering stays consistent across all partition cells and refinements.

Inside each cell, $D_i(m)$ is the initial segment of mass $m$ under this ordering. Its volume $v_i(m)$ integrates the ascending cost quantile function over $[0, m]$. Swapping higher-cost points for equal-mass lower-cost points lowers volume. The function $v_i$ is convex and lower semicontinuous, allowing $+\infty$ at the right endpoint. Nesting yields $\mu(D_i(m)\triangle D_i(m')) = |m - m'|$. The feasible set of (4) is compact and the risk is continuous in $m$, so the minimum is attained. Replacing any originally feasible region cell by cell yields a feasible vector $m$, so the constraint set is nonempty. Upon refinement, old initial segments restrict to initial segments on subcells, making $U_n$ nonincreasing.

Take any feasible set $B$. Form $B_n$ by taking the initial segment $D_i(m_i)$ matching $m_i = \mu(B \cap E_i)$ in each cell. Inside cell $i$, the symmetric difference has mass at most twice $\min(m_i, \mu_i - m_i)$, giving

$$\mu(B_n \triangle B) \leq 2 \sum_i \min(m_i, \mu_i - m_i) \longrightarrow 0.$$

The limit follows from conditional expectation. Along partitions generating the Borel $\sigma$-algebra, $\mathbb{E}_\mu[\mathbf{1}_B \mid \mathcal{P}_n] \to \mathbf{1}_B$ in $L^1(\mu)$. The integral of its distance to $\{0, 1\}$ equals the sum above and vanishes. Applying the risk-difference bound $|Q(B_n) - Q(B)| \leq 2\mu(B_n \triangle B)$ yields $\limsup_n U_n \leq Q(B)$. Taking the infimum over all feasible $B$ completes the atomless case.

When $\mu$ has atoms, list them as $z_j$ with masses $b_j > 0$. At level $n$, partition the atomless part as above and choose binary indicators $e_j \in \{0, 1\}$ for the first $n$ atoms. The attainable mass set of remaining atoms,

$$T_n = \left\{ \sum_{j>n} b_j e_j : e_j \in \{0, 1\} \right\},$$

is compact because the series tail converges uniformly to zero. Impose $\sum_i m_i + \sum_{j \leq n} b_j e_j + t = c$ with $t \in T_n$. Minimize $Q(A_n)$, where $A_n$ collects the cell segments and selected leading atoms. Appending the tail atoms that realize $t$ produces a feasible exact-coverage set with risk at most $Q(A_n)$. Its risk differs from $Q(A_n)$ by at most $2 \sum_{j>n} b_j$. Feasibility stays compact. Refinement plus one extra atom preserves prior feasible extensions, so upper bounds stay nonincreasing. Applying the cell replacement and tail bound to any feasible $B$ yields the same limit $\mathcal{R}(H, c)$. Knowledge of $T_n$ is part of the exact-coverage specification. An arbitrary sequence of real atomic masses carries no general finite membership algorithm. □

**Checking coverage and volume feasibility.** Split $\mu$ into its point atoms and its atomless part. Write $T = \{\sum_j b_j e_j : e_j \in \{0, 1\}\}$ for achievable atomic sums, $a_0$ for total atomless mass, and $v(m)$ for the minimal volume needed to collect mass $m$ from the atomless part (possibly $+\infty$). To reach total coverage $c$, atoms must supply at least $\max(0, c - a_0)$ and at most $c$. If $T \cap [\max(0, c - a_0), c]$ is empty, no exact-coverage set exists. Otherwise that intersection is compact with largest element $s_*$. Because atoms carry zero volume and $v$ is nondecreasing, the minimum volume for coverage $c$ is $H_0(c) = v(c - s_*)$. A finite budget $H$ is feasible if and only if $H_0(c) \leq H$. Purely discrete laws reduce to checking $c \in T$, while atomless laws reduce to checking $v(c) \leq H$.

**Fixing single-period coverages separately.** When an application imposes both $\mu_-(B) = c_-$ and $\mu_+(B) = c_+$, the same cell-wise argument characterizes the frontier. Track both marginal masses in each cell and pick a minimum-volume cell representative attaining that mass pair. Replacing a feasible set cell by cell preserves both coverages without increasing volume. Average cell masses match as well, so the symmetric-difference and risk bounds stay intact. Level-$n$ feasible infima converge to the doubly constrained frontier (taking running minima when needed for monotonicity). Both constraints require their actual attainable mass-pair sets instead of the one-dimensional average-coverage sort.

To verify that a minimum-volume cell representative exists for any finite-volume attainable mass pair, separate atoms from the atomless part inside the cell. Along a volume-minimizing sequence of candidates, pass to a subsequence where every atomic indicator converges. Using separability of $L^1(\mu)$, extract a further subsequence whose atomless indicators converge weak-star in $L^\infty(\mu)$ to a measurable function $0 \le u \le 1$. Both marginal integrals against $d\mu_\pm/d\mu$ pass to the limit. Lower semicontinuity of volume follows by integrating $\min(a, K)$ and letting $K \uparrow \infty$. Applying Lyapunov's purification theorem for atomless vector measures (Dvoretzky et al. 1950) on each cost band $\{j \le a < j+1\}$ replaces $u$ by a set indicator. That indicator preserves all three integrals against $d\mu_-/d\mu$, $d\mu_+/d\mu$, and $a$. Combining all bands with the limiting atomic choices gives the required deterministic representative. Solving the doubly constrained cell subproblem in practice requires a separate routine.

## A.1 Where the all-shape frontier is continuous

**Proposition 5.** *If the average marginal $\mu$ is atomless, the frontier* (3) *is continuous for all $H \ge H_0$, where $H_0$ is the minimum feasible volume. For $H_2 > H_1 \ge H_0$,*

$$0 \le \mathcal{R}(H_1, c) - \mathcal{R}(H_2, c) \le 2\min(c, 1-c)\frac{H_2 - H_1}{H_2 - H_0}.$$

*Proof.* Let $A$ be a minimum-volume exact-coverage set of volume $H_0$, built as the lowest-cost initial segment. For any feasible set $B$ at budget $H_2$, set $t = (H_2 - H_1)/(H_2 - H_0)$. Restrict the atomless vector measure $(\mu, dx)$ to the finite-volume union $A \cup B$. By range convexity (Dvoretzky et al. 1950), we can delete a subset of $B \setminus A$ carrying fraction $t$ of both its mass and its volume, while adding a subset of $A \setminus B$ carrying fraction $t$ of both its mass and its volume. Because $B \setminus A$ and $A \setminus B$ have equal mass, the modified set keeps coverage $c$ and has volume $(1-t)|B| + tH_0 \le H_1$. Only deleted points can raise consecutive failures, so risk increases by at most $2t\mu(B \setminus A) \le 2t\min(c, 1-c)$. Taking the infimum over $B$ yields the stated inequality and continuity for every $H > H_0$.

Right-continuity at $H_0$ follows directly. Working on the effective support $E$, choose a $c$-quantile $b$ of the unit cost $a$, and set $L = \{a < b\}$, $G = \{a = b\}$, and $U = \{a > b\}$. Minimum-volume sets contain all of $L$ and fill coverage $c$ from $G$. For any set $B$ of mass $c$,

$$|B| - H_0 = \int_{L \setminus B} (b - a)\, d\mu + \int_{U \cap B} (a - b)\, d\mu.$$

Hence the missed mass in $L$ plus excess mass in $U$ is at most $\mu\{0 < |a - b| \le \eta\} + (|B| - H_0)/\eta$ for any $\eta > 0$. Repairing both parts and balancing the net mass difference inside $G$ yields a set of volume $H_0$ whose symmetric difference from $B$ has mass at most twice that bound. Letting $|B| \downarrow H_0$ and then $\eta \downarrow 0$ with the risk-difference bound yields right-continuity at $H_0$. □

Atoms can create genuine jumps in the all-shape frontier. Let $Z_t$ equal 0 with probability $2/5$ and otherwise enter a continuous regime with probability $3/5$ where both periods are independent uniform draws on $[1, 2]$, sharing the regime indicator across the pair. At exact coverage $c = 2/5$, any set containing the origin can add only zero-probability continuous points and has risk $3/5$. Dropping the origin requires length $2/3$ inside $[1, 2]$, reducing risk to $2/5 + (3/5)(1/3)^2 = 7/15$. The frontier jumps at budget $2/3$, while remaining covered by Theorem 1.

## A.2 Cell-averaged error relative to the original joint law

*Proof of Proposition 1.* First solve the problem under the cell-averaged law $P_{\mathcal{P}}$, which places $d_i\, \mu(dx)$ on the diagonal and $K_{ij}\, \mu(dx)\mu(dy)$ off the diagonal. Any region with cell masses $m_i$ has joint hit probability matching the two sums in (5). Cell-wise minimum-volume assembly realizes every feasible vector $m$, and every feasible region produces a feasible mass vector. Continuity of

the objective and lower semicontinuity of $v_i$ make the feasible set compact, so $p_{\mathcal{P}}$ is the exact minimum risk under $P_{\mathcal{P}}$.

Next compare $P$ and $P_{\mathcal{P}}$. For $\mu$-almost every $x$, marginal consistency gives $d(x) + \int K(x,y)\,\mu(dy) = 1$, and averaging over cells preserves both marginals. Write $b = \mathbf{1}_B - 1/2$; the difference in joint hit probabilities then equals $\int b(x)b(y)\,d(P - P_{\mathcal{P}})$. On the diagonal, $b(x)^2 = 1/4$, and total diagonal mass matches before and after cell averaging, so the diagonal difference cancels identically. Off the diagonal, $|b(x)b(y)| = 1/4$ bounds the absolute difference by $e_{\mathcal{P}}$ uniformly over every feasible set $B$. Taking infima gives (6), and evaluating the cell-optimal assembly back under $P$ adds at most another $e_{\mathcal{P}}$. Finally, $K_{\mathcal{P}}$ is the conditional expectation of $K$ given the cell-pair partition, so integrability gives $L^1$ convergence to zero along generating refinements. □

**Actionable global bounds from the finite program.** Suppose $\mu$ is atomless and remaining cell masses inside box bounds $[l, u]$ can be merged and sorted by unit cost to hit coverage $c$ at minimal volume. Then (5) admits an explicit stopping criterion. Given $0 < \eta \le 1$, split each $[0, \mu(E_i)]$ into $\lceil 1/\eta \rceil$ equal subintervals of width at most $\eta\mu(E_i)$. Inside each mass box $[l, u]$, take the cheapest mass $l_i$ in every cell. Merge available increments from $l_i$ to $u_i$ in ascending cost order until total mass reaches $c$. Discard boxes where $c$ is unattainable or minimum volume exceeds $H$. Every surviving box yields an exact feasible candidate, and every box containing a feasible point produces at least one candidate. Let $\widehat{p}$ be the smallest cell-model objective among these candidates. The partial derivative $\partial/\partial m_i$ of the objective equals $d_i + 2\sum_j K_{ij} m_j \in [0, 2]$. Shifting every partial derivative by $-1$ leaves directional derivatives along $\sum_i m_i = c$ unchanged. Any two exact-coverage points in one box differ in objective by at most $\sum_i |m_i - m_i'| \le \eta$. Checking at most $\lceil 1/\eta \rceil^J$ boxes certifies

$$\widehat{p} - \eta - e_{\mathcal{P}} \le \mathcal{R}(H, c) \le \widehat{p} + e_{\mathcal{P}},$$

including at zero extra budget, at a cost growing with cell count $J$. Cell error $e_{\mathcal{P}}$ and finite-solver error $\eta$ enter both sides cleanly. Under numerical quadrature, integration error has to be included in both endpoints. Boxes may be pruned only when a certified volume lower bound exceeds $H$, and upper bounds require a verified feasible region. When an uncertain volume interval straddles $H$, one refines further, increases precision, or retains the box's lower bound. Exact rational arithmetic below handles boundary equalities without rounding ambiguity.

**Exact rational box search and independent checks.** We implemented exact rational box search for (5) on models with constant within-cell unit volume costs. Starting from the full box, the solver tightens coordinate bounds via the coverage equation. It fills coverage greedily by ascending unit cost to obtain a feasible candidate $x$, and writes any other allocation in the box as $x + z$. The objective shift equals $(d + 2Kx)^\top z + z^\top K z$ exactly. Adding a nonnegative volume multiplier bounds the linear part by sorting adjusted slopes over the equal-coverage slice. The quadratic part keeps nonnegative diagonal terms and bounds each cross term by the minimum endpoint product. Repeatedly bisecting the box with lowest global lower bound along its longest edge terminates once the global gap meets tolerance. Every arithmetic step stays in exact rational arithmetic.

Two models verified six budgets against an independent script enumerating polygon vertices, edge stationary points, and interior stationary points in the two-dimensional coverage plane. For the three-block model of Section 2.3 at budgets 1, 1.005, and 1.01, inspecting 11, 25, and 51 boxes shrinks the gap below $10^{-4}$, bracketing the exact risks 853/52900, 61/5290, and 1/100.

A second model tests a full-rank continuous kernel combined with diagonal mass. Three cells each carry marginal mass 1/3 with unit-mass volumes $(1, 2, 4)$ at coverage 1/2. Diagonal weights

are $d = (1/4, 1/2, 1/4)$, and the cell kernel is

$$K = \tfrac{1}{4}\begin{pmatrix} 3 & 5 & 1 \\ 5 & 0 & 1 \\ 1 & 1 & 7 \end{pmatrix}.$$

Here $d_i + \sum_j K_{ij}/3 = 1$ and cell approximation error is zero. At minimum feasible volume $2/3$, 43 boxes bracket the exact risk $7/18$ within $10^{-4}$. At absolute volume 0.9, the exact risk drops to $889/2880$. Setting tolerances $10^{-2}$, $10^{-3}$, and $10^{-4}$ inspects 31, 51, and 55 boxes (counting the root and every feasibility-checked child). The resulting actual interval widths are 0.0075955, 0.0004264, and 0.0000710. Exact rational endpoints and feasible allocations are saved in the source archive.

**A smooth model attaining only an infimum.** We construct a smooth two-period model whose optimal risk is approached arbitrarily closely by deterministic regions without being attained by any deterministic set. Normalize support to $(0, 1)$ with marginal density proportional to $\exp\{-1/[x(1-x)]\}$ on $(0, 1)$ and zero outside: smooth, symmetric around $1/2$, and unimodal. Fix $c = 1/2$ and length cap $H = 1/2$. The centered half-length interval holds mass strictly exceeding $1/2$, so $H$ sits strictly above minimum length and the minimum-volume endpoints have nonzero density slope. Let $F$ be the marginal distribution function, and in rank coordinates $u = F(x)$, $v = F(y)$ take the Poisson kernel

$$k_r(u - v) = \frac{1 - r^2}{1 - 2r\cos(2\pi(u - v)) + r^2} = 1 + 2\sum_{j\geq 1} r^j \cos(2\pi j(u - v)), \qquad 0 < r < 1.$$

Both rank marginals are uniform, and the series converges uniformly and absolutely. For any feasible set $B$, put $A = F(B \cap (0, 1))$ with $|A| = 1/2$; termwise integration yields

$$Q(B) = \tfrac{1}{4} + 2\sum_{j\geq 1} r^j \left| \int_A e^{2\pi i j u}\, du \right|^2 > \tfrac{1}{4}.$$

Strict inequality holds by completeness of the Fourier basis: vanishing of all nonconstant coefficients would force $\mathbf{1}_A = 1/2$ almost everywhere.

Now set $B_N = \{0 < x < 1 : \sin(2\pi N F(x)) > 0\}$ for integer $N \geq 1$. Reflection $x \mapsto 1 - x$ exchanges the positive and negative sets while preserving both marginal mass and Lebesgue length, so every $B_N$ has exact mass $1/2$ and exact length $1/2$. Because $\mathbf{1}_{F(B_N)}$ has period $1/N$, only multiples of frequency $N$ can be nonzero; Parseval's identity gives $0 < Q(B_N) - 1/4 \leq r^N/4$. Thus $\mathcal{R}(1/2, 1/2) = 1/4$: every positive accuracy is achieved by a feasible set, whereas zero gap has no deterministic minimizer.

# B Proofs for the frontier over all measurable regions

## B.1 Square-root leading term for regular boundaries

*Proof of Theorem 2.* Exact coverage converts the volume cap into a weighted $L^1$ bound on set edits. For any feasible $B$, write $D = \mathbf{1}_B - \mathbf{1}_A$ and $\Delta V = |B| - V_A$. Combining $\int fD = 0$ with the sign of $D$ on $A$ and $A^c$ yields

$$\lambda \Delta V = \int |f - \lambda|\, |D|\, dx. \tag{20}$$

This also confirms minimality of $V_A$, with equality only up to null sets. Compactness and regularity of $\Gamma$ give a uniform tubular level band parameterized by $s \in \Gamma$ and $z = f(x) - \lambda$, with

volume element $(g(s)^{-1} + O(|z|))\, dS(s)\, dz$ along the flow of $\nabla f/\|\nabla f\|^2$. Thus $|\{|f-\lambda| \le t\}| = O(t)$, and splitting (20) at $t = \sqrt{\delta}$ yields

$$|A \triangle B| \le O(t) + \frac{\lambda V_A \delta}{t} = O(\sqrt{\delta}).$$

Expanding the two-period hit probability around $A$ (using the continuous representative of $w$ near $\Gamma$, which leaves volume integrals unchanged) gives

$$R_2(B) - R_2(A) = \int wD + \iint p(x,y)D(x)D(y)\, dx\, dy.$$

The quadratic remainder is bounded in magnitude by $\|p\|_\infty |A \triangle B|^2 = O(\delta)$ uniformly over feasible $B$. Putting $k = \bar{w}$ and replacing $w$ with $w - k$ in the linear term alters the integral by $k\Delta V = O(\delta)$. Controlling quadratic deficits at fixed mass relates to quantitative bathtub inequalities (Mazari and Ruiz-Balet 2022), and boundary variations of nonlocal set functionals appear in Gaussian stability (Heilman et al. 2016); here we bound the linear term directly across all measurable perturbations.

Introduce a multiplier $t > 0$ to charge (20) against $t\lambda V_A \delta$, and optimize the inclusion decision pointwise:

$$\begin{aligned} R_2(A) - R_2(B) \le\ & t\lambda V_A \delta + O(\delta) \\ & + \int_A [w - k - t(f-\lambda)]_+\, dx + \int_{A^c} [k - w - t(\lambda - f)]_+\, dx. \end{aligned}$$

Boundedness of $w$ follows from $0 \le w \le 2f$ (both for equal and average marginals). For large $t$, both integrands vanish outside $|z| = O(t^{-1})$. Replacing $w(x)$ by $w(s)$ via local Lipschitz continuity and replacing the volume element by $g(s)^{-1}\, dS\, dz$ each incur $O(t^{-2})$ error. Integrating $|z|$ on each side using $\int_0^\infty [a - tz]_+\, dz = [a]_+^2/(2t)$ sums to $K/(2t) + O(t^{-2})$. When $K > 0$, setting $t = \sqrt{K/(2\lambda V_A \delta)}$ yields the upper bound in (7). When $K = 0$, continuity forces $w = k$ everywhere on $\Gamma$, so $|w - k| \le C|f - \lambda|$ both inside and outside the band, reducing the linear gain via (20) to $O(\delta)$.

To attain the leading term when $K > 0$, move $\Gamma$ normally by $\varepsilon h(s) + a$, where $h = (k - w)/g$ satisfies $\int_\Gamma h\, dS = 0$. On a small tubular neighborhood, the coverage derivative with respect to $a$ at the origin is $\lambda \int_\Gamma dS > 0$, while the first derivative in $\varepsilon$ vanishes. The implicit function theorem supplies a shift $a(\varepsilon) = O(\varepsilon^2)$ restoring coverage $c$. Since $|f - \lambda| = g(s)|v| + O(v^2)$ and the normal volume element is $(1 + O(|v|))\, dS\, dv$, we obtain

$$\lambda \Delta V = \frac{\varepsilon^2}{2} \int_\Gamma g h^2\, dS + O(\varepsilon^3) = \frac{\varepsilon^2 K}{2} + O(\varepsilon^3).$$

Meanwhile, risk drops by $-\varepsilon \int_\Gamma wh\, dS + O(\varepsilon^2) = \varepsilon K + O(\varepsilon^2)$. Choosing $\varepsilon = \sqrt{2\lambda V_A \delta / K} - L\delta$ for a sufficiently large constant $L$ stays inside the budget while attaining the leading coefficient. The construction uses only continuity and boundedness of $h$. □

**Independent check on a non-Gaussian marginal.** Take the standard logistic distribution $F$ with density $f$, and set $h_0(u) = u^2 - 1/3$. Define $p(x,y) = f(x)f(y)[1 + h_0(F(x))h_0(F(y))]$. This density is positive with common marginal $f$ because $\int_0^1 h_0(u)\, du = 0$. At $c = 0.90$, write $u = (1-c)/2$ and $v = 1 - u$. Shifting the centered minimum-length rank interval $[u, v]$ rightward by $t$ keeps coverage fixed. Its length is $\mathrm{logit}(v+t) - \mathrm{logit}(u+t)$ (with $\mathrm{logit}(a) = \log(a/(1-a))$), and its risk is $(1-c)^2 + \{\int_{u+t}^{v+t} h_0(z)\, dz\}^2$. Direct expansion of both scalar functions and substitution of both endpoints into Theorem 2 (with counting measure $dS$) both yield coefficient 0.00311656. Shrinking $\delta$ from $10^{-4}$ to $10^{-8}$ along this branch raises the ratio of risk reduction to $\sqrt{\delta}$ from 0.00300005 to 0.00311541.

**All-shape solution for the smooth logistic model across all length budgets.** The same logistic example admits an exact all-shape frontier at every finite budget. Fix $c = 0.90$, map any measurable set $B$ to its rank set $A = F(B)$, and write $a(u) = 1/[u(1-u)]$ for the quantile derivative. Coverage equals $|A|$, original length equals $\int_A a(u)\,du$, and risk equals $(1-c)^2+(\int_A h)^2$ with $h(u) = u^2 - 1/3$. For a rank interval $A = [q, q+c]$,

$$V(q) = \log\frac{(q+c)(1-q)}{q(1-q-c)}, \qquad m(q) = \int_q^{q+c} h(u)\,du = c\{q^2 + cq + (c^2-1)/3\}.$$

Symmetry and unimodality of $f$ make $q_b = (1-c)/2$ the unique minimum-length interval. Setting $q_s = (\sqrt{(4-c^2)/3} - c)/2 \in (q_b, 1-c)$ solves $m(q_s) = 0$, and both $V$ and $m$ increase strictly on $[q_b, q_s]$.

To verify optimality over all measurable sets for $q_b < q \le q_s$, set

$$\mu = \frac{h(q+c) - h(q)}{a(q+c) - a(q)} > 0, \qquad s_\mu(u) = h(u) - \mu a(u).$$

Then $s_\mu(q) = s_\mu(q+c)$, and $s_\mu$ is unimodal. Its derivative is positive on $(0, 1/2]$. On $(1/2, 1)$, vanishing of $s'_\mu$ is equivalent to $\mu = 2u^3(1-u)^2/(2u-1)$, whose right-hand side decreases strictly from $+\infty$ to 0 because $-8u^2+9u-3 < 0$. Thus $[q, q+c]$ collects the highest-scoring set of length $c$. For any $A$ with $|A| = c$ and $\int_A a \le V(q)$, comparing $\int_A s_\mu$ and adding back $\mu\int_A a$ yields $\int_A h \le m(q) \le 0$. This gives the global lower bound on risk, which the interval itself attains.

At any cap $H \in [V(q_b), V(q_s)]$, solving $V(q) = H$ on the right branch and evaluating $(1-c)^2 + m(q)^2$ gives the global frontier. Beyond $V(q_s)$, the optimum stays at $q_s$. Risk hits 0.01 at about 1.73225% extra length, and at 1% extra length the minimum risk is 0.01001119. Expanding $V$ and $m$ at $q_b$ reproduces the square-root coefficient 0.00311656. Unlike the piecewise density in Section 2.3, a single interval stays globally optimal along this entire frontier.

## B.2 Budget exponents for finite-order boundaries

*Proof of* (8). Follow (20) with $r = \delta^{1/(k+1)}$. In tubular coordinates, the volume element is $(1 + O(|v|))\,dv\,dS$, $|f - \lambda| \ge c_0|v|^k$ inside a small fixed tube, and $|f - \lambda|$ is bounded away from zero outside. Hence $|A \triangle B| = O(r) + O(\delta/r^k) = O(r)$, the quadratic risk remainder is $O(r^2)$, and centering $w$ by a constant $b$ adds $O(\delta)$.

Pick the unique constant $b$ minimizing $D_k$ (unique since $q > 1$), and apply multiplier $t > 0$ to the volume deficit as in Appendix B. For large $t$, both positive-part integrands vanish outside $|v| = O(t^{-1/k})$. Replacing $w(s + vn)$ by $w(s)$, $|f - \lambda|$ by $a(s)|v|^k$, and the volume element by $dv\,dS$ each incur $O(t^{-2/k})$ error. Integrating along inner and outer normals via

$$\int_0^\infty [z - tav^k]_+\,dv = \frac{k}{k+1}(ta)^{-1/k}[z]_+^{(k+1)/k}$$

sums to $(k/(k+1))t^{-1/k}D_k$. When $D_k > 0$, adding $t\lambda V_A\delta$ and choosing $t = [D_k/((k+1)\lambda V_A\delta)]^{k/(k+1)}$ yields the upper bound in (8).

To attain this leading term, set $h = a^{-1/k}\operatorname{sgn}(b-w)|b-w|^{1/k}$. First-order optimality in $b$ yields $\int_\Gamma h\,dS = 0$ and $-\int_\Gamma wh\,dS = \int_\Gamma a|h|^{k+1}\,dS = D_k$. Applying normal shift $\varepsilon h$ plus an $O(\varepsilon^2)$ constant shift restores coverage $c$ (since the coverage derivative in the constant shift equals $\lambda|\Gamma| > 0$). Tubular expansion gives $\lambda\Delta V = \varepsilon^{k+1}D_k/(k+1) + O(\varepsilon^{k+2})$ and risk reduction $\varepsilon D_k + O(\varepsilon^2)$. Subtracting a sufficiently large fixed multiple of $\varepsilon_0^2$ from $\varepsilon_0 = [(k+1)\lambda V_A\delta/D_k]^{1/(k+1)}$ keeps the region within budget while attaining the leading coefficient.

When $D_k = 0$, we have $w = b$ identically on $\Gamma$. Inside $|v| \le r$, the centered linear term is $O(r^2)$; elsewhere in the tube, $|v| \le r^{1-k}|v|^k$ combined with (20) gives $O(r^{1-k}\delta) = O(r^2)$; outside the tube the contribution is $O(\delta)$. Adding the quadratic remainder leaves only $O(\delta^{2/(k+1)})$. □

**Smooth density with third-order crossing.** Define $f(x) = L\exp[-(x^2-1)^3]$ on $\mathbb{R}$, with normalizing constant $L > 0$. Because $f$ is even and strictly decreasing in $|x|$, threshold $\lambda = L$ yields the unique minimum-length set $A = [-1, 1]$ at coverage $c = \int_A f$, with outward normal expansion $L - 8Lv^3 + O(|v|^4)$ at both endpoints. Pairing marginal $F$ with $h_0(u) = u^2 - 1/3$ in $p(x, y) = f(x)f(y)[1 + h_0(F(x))h_0(F(y))]$ preserves both marginals. Since $h_0$ differs at the two endpoint ranks and $\int_A h_0(F(x))f(x)\,dx = -c(1-c^2)/12 \neq 0$, we obtain $w(-1) \neq w(1)$. Equation (8) therefore yields a nonzero $\delta^{1/4}$ leading term. This example keeps a unique minimum-volume set and smooth joint density while changing only the crossing order from 1 to 3.

High-precision quadrature confirms the fourth-root coefficient from both ends. Here $c = 0.81986898$, and (8) gives coefficient 0.00750221. Shift the left endpoint right by $10^{-2}$, $10^{-4}$, and $10^{-6}$, adjusting the right endpoint to keep exact mass. The ratio of risk reduction to $\delta^{1/4}$ rises through 0.00718515, 0.00749905, and 0.00750218, with relative length budget near $2.0 \times 10^{-24}$ at the finest step.

**Smoothness alone does not fix a power law.** Set $\psi(z) = z\exp(-1/z^2)$ ($\psi(0) = 0$) and take $f(x) = L\exp[-\psi(x^2-1)]$, normalized to unit integral. Here $\psi$ increases strictly with all derivatives vanishing at zero. The unique minimum-length set remains $[-1, 1]$, and within distance $r$ of either endpoint $|f - L| \leq C\exp(-b/r^2)$ for constants $C, b > 0$. Retaining the same rank kernel, move the left endpoint from $-1$ to $-1 + t$ ($t > 0$) and the right endpoint to $1 + s(t)$ to restore coverage. Equal endpoint densities $L > 0$ yield $s(t)/t \to 1$, and (20) bounds $0 < \delta(t) \leq C't\exp(-b'/t^2)$ (positive by uniqueness of $[-1, 1]$). Because $w(-1) > w(1)$, risk drops by $[w(-1) - w(1)]t + o(t)$, which grows faster than $\delta^\alpha$ for any fixed $\alpha > 0$. Theorem 1 continues to characterize the finite-budget frontier.

# C Derivations for parametric regions and rectangles

## C.1 Curvature and Gaussian probabilities at fixed coverage

*Proof of Theorem 3.* Taylor expansion in local coordinates $u$ gives

$$L(u) = \tfrac{1}{2}u^\top Hu + O(\|u\|^3), \qquad S(u) = S(0) + b^\top u + O(\|u\|^2).$$

Positive definiteness of $H$ lets us choose $U$ small enough that $L(u) \geq k\|u\|^2$ for some $k > 0$. Writing $s = \log(1 + \delta)$, every feasible $u$ obeys $\|u\| = O(\sqrt{s})$ and $u^\top Hu \leq 2s + O(s^{3/2})$.

When $b \neq 0$, Cauchy–Schwarz bounds $-b^\top u \leq \sqrt{b^\top H^{-1}b}\sqrt{u^\top Hu}$ uniformly. Setting $v = -H^{-1}b/\sqrt{b^\top H^{-1}b}$ achieves equality with $v^\top Hv = 1$. Along $u = tv$, we have $L(tv) = t^2/2 + O(t^3)$ with derivative $t + O(t^2)$, yielding a positive root $t = \sqrt{2s} + O(s)$ whose risk matches the upper bound to leading order; substituting $s = \delta + O(\delta^2)$ gives (9). Any cluster point of $u_\delta/\sqrt{2s}$ minimizes $b^\top v$ over $v^\top Hv \leq 1$, uniquely selecting the limit direction.

When $b = 0$, expanding to second order gives $S(u) - S(0) = u^\top Ju/2 + o(\|u\|^2)$. Rescaling $u$ by $\sqrt{2s}$ converges to the ellipsoid $v^\top Hv \leq 1$, whose interior points are feasible for small $s$ and approximate every boundary point. Minimizing $v^\top Jv$ on that ellipsoid (including the origin) gives $\min\{0, \lambda_{\min}(H^{-1/2}JH^{-1/2})\}$, and multiplying by $s$ yields (10). □

*Proof of Lemma 1.* Differentiation gives $g'(x) = A(e^x)$ with $A(a) = aF'(a)/F(a) > 0$, and $A'(a)/A(a) = (1 - A(a))/a - a$. Integrating $(x\phi(x))' = (1 - x^2)\phi(x)$ over $[0, a]$ shows $a\phi(a) > (1 - a^2)\int_0^a \phi(x)\,dx$, so $A(a) > 1 - a^2$, $A'(a) < 0$, and $g''(x) < 0$.

Replacing $(X, Y)$ by $(X, -Y)$ flips $r$ to $-r$ without changing absolute values, so $Q_{-r} = Q_r$. Assume $r > 0$ and set $K_a(x) = \mathbb{P}(|Y| \leq a \mid X = x)$. Because $Y \mid X = x \sim N(rx, 1 - r^2)$, the probability inside $[-a, a]$ decreases strictly with $|x|$, yielding $B_r(a)/F(a) = \mathbb{E}[K_a(X) \mid |X| \leq a] > K_a(a)$. Differentiating across all four square edges gives $B_r'(a) = 4\phi(a)K_a(a)$, hence

$$Q_r'(a) = \frac{4\phi(a)}{F(a)^2}\left[K_a(a) - \frac{B_r(a)}{F(a)}\right] < 0.$$

Under $r = 0$, $B_0 = F^2$ and $Q_0 = 1$. □

*Proof of Corollary 1.* At the equal-width cube, the tangent space to $\sum_i g(x_i) = \log c$ is $\{z : \sum_i z_i = 0\}$. Log-volume equals a constant plus $\sum_i x_i$, with Lagrange multiplier $1/g'(x_0)$, giving second-order volume form $h_0\|z\|_2^2$ on the tangent space. Differentiating (11) along $z$ gives $P_0 \sum_i d_i z_i = P_0(d^\circ)^\top z$, and Theorem 3 supplies $\gamma$. Fixed coverage bounds every half-width below by $F^{-1}(c) > 0$, while bounded volume gives a common upper bound, so small-budget feasible sets stay compact and contract to the unique cube, making the local expansion valid globally. □

*Proof of Theorem 4.* For any fixed first half-width $a \geq q$, the remaining $M-1$ coordinates must reach product coverage $c/F(a)$. By strict concavity in Lemma 1, setting all remaining half-widths equal to $v(a)$ uses the least volume, so checking whether $a$ fits within budget reduces to testing that minimal volume.

Differentiating along the equal-coverage branch gives

$$\frac{d}{da}\log\left[2^M a v(a)^{M-1}\right] = \frac{1}{a}\left[1 - \frac{A(a)}{A(v(a))}\right].$$

For $a > q$, we have $v(a) < q < a$ and $A' < 0$, making the derivative strictly positive. Volume increases continuously from $(2q)^M$ to $+\infty$ as $a \to \infty$ (since $v(a) \to F^{-1}(c^{1/(M-1)}) > 0$). Each $\delta > 0$ thus yields one unique root $a_\delta > q$; larger $a$ violates the budget, while smaller $a$ incurs strictly higher risk. □

*Proof of Proposition 2.* Take polar coordinates with $\Theta \sim \mathrm{Unif}[0, 2\pi)$ and define $D_s = \cosh(2s) - \sinh(2s)\cos(2\Theta)$. The boundary radius of $E_s$ along angle $\Theta$ is $q(s)/\sqrt{D_s}$. Tail probabilities of the bivariate standard normal radius give

$$\mathbb{E}_\Theta \exp\{-q(s)^2/(2D_s)\} = 1 - c.$$

The derivative with respect to $q$ is strictly negative at $q_*$, so $q(s)$ is smooth near 0 and even by axis symmetry. Set $k = q_*^2/2$ and $L(s) = \log(q(s)^2/q_*^2)$. Expanding with $\mathbb{E}\cos(2\Theta) = 0$ and $\mathbb{E}\cos^2(2\Theta) = 1/2$ yields $e^{-k}\{1 - kL(s) + k^2 s^2 + O(s^4)\} = e^{-k}$. Thus $L(s) = ks^2 + O(s^4)$.

Next write $X_t$ for coordinate 1 and $P_0(x) = \mathbb{P}(Z_{t-1} \in E_0 \mid X_t = x)$. Given $X_t = x$, prior coordinate 1 is $N(rx, 1-r^2)$ independently of standard normal prior coordinate 2. Each horizontal slice of $E_0$ is a centered interval whose Gaussian mass drops strictly as the conditional mean moves away from zero. Averaging over coordinate 2 makes $P_0(x)$ even and strictly decreasing in $|x|$.

At polar angle $\theta$, the outward normal speed is $q_* \cos(2\theta)$, arc element is $q_*\, d\theta$, and boundary density is $e^{-k}/(2\pi)$. Time-reversal symmetry doubles the single-boundary derivative, giving

$$\beta_E = 2q_*^2 e^{-k}\, \mathbb{E}_\Theta[\cos(2\Theta) P_0(q_* \cos\Theta)].$$

With $U = |\cos\Theta|$, $2U^2 - 1$ increases strictly while $P_0(q_* U)$ decreases strictly, so their covariance is strictly negative. Since $2U^2 - 1$ has mean zero, $\beta_E < 0$. Substituting into Theorem 3 finishes the proof. □

## C.2 Finite-budget allocation across multiple coordinates

*Proof of Theorem 5.* Minimizing risk is equivalent to minimizing $\sum_i f_i(x_i)$. Each $f_i$ has first derivative matching the first term of (16) and strictly positive second derivative $\lambda_i(1-\lambda_i)e^{x_i}/(1-\lambda_i+\lambda_i e^{x_i})^2$. Because the budgeted feasible set is compact and convex, the minimizer is unique. On the closed coverage simplex without a volume cap, equating derivatives across positive coordinates and bounding zero coordinates below by the common value yields (15); if any $\widehat{x}_i = 0$, the corresponding half-width is infinite and unattainable at any finite budget.

Restoring the volume cap, $x^0$ is strictly feasible for $\delta > 0$, so the Karush–Kuhn–Tucker conditions are necessary and sufficient. Multiplier $\mu = 0$ occurs only when $\widehat{x}$ is feasible; otherwise complementary slackness forces $\mu > 0$ and $W(x) = k$. For fixed multipliers, the left-hand side of (16) increases strictly in $x_i$, uniquely determining the allocation jointly with the sum and volume constraints.

In a common scale family, swapping $x_i$ and $x_j$ preserves both coverage and volume. If $\lambda_i > \lambda_j$, then $f_i'(x) > f_j'(x)$ so $f_i - f_j$ increases strictly. If an allocation had $x_i > x_j$, exchanging them would strictly lower the objective; the optimum therefore satisfies $x_i \le x_j$. □

### C.3 Independent coordinates, cross-period dependence, and global bounds

Keeping single-period coordinates independent allows joint coverage to factor while letting cross-period dependence enter freely through the full joint hit probability $J$. At budget $K = e^k$, feasible half-widths satisfy

$$m_i = F_i^{-1}(c) \le \tau_i \le T_i = \frac{K}{2^M \prod_{j \ne i} m_j},$$

forming a compact search box $\prod_i [-\log F_i(T_i), s]$, with $e^{-x}$ applied coordinate-wise in $J(e^{-x})$.

**Proposition 6** (Bracketing the finite-budget frontier)**.** *Assume each $w_i$ is continuous and convex and $\delta > 0$. Partition the search box into subboxes $B = [l, u]$ of edge length at most $h$. Discard boxes that do not intersect $\sum_i x_i = s$ or that satisfy $W(u) > k$. In each remaining box, pick $z \in B$ with $\sum_i z_i = s$. Put $y = z$ if $W(z) \le k$, and otherwise set*

$$\alpha = \frac{W(z) - k}{W(z) - W_0}, \qquad y = (1 - \alpha)z + \alpha x^0.$$

*Compute*

$$L_B = \max\{0, 1 - 2c + J(e^{-u})\}, \qquad U_B = 1 - 2c + J(e^{-y}).$$

*With $R^\star$ denoting the finite-budget minimum risk, $L = \min_B L_B \le R^\star \le U = \min_B U_B$, and $U - L \to 0$ as $h \downarrow 0$. Hence any tolerance $\varepsilon > 0$ is met after finitely many refinements.*

*Proof.* Because each $w_i$ decreases in $x_i$, $W(u)$ lower-bounds log-volume on $B$, so no feasible box is pruned. Monotonicity of $J$ makes $L_B$ lower-bound risk on $B$. Meanwhile $\sum_i y_i = s$ and convexity gives $W(y) \le k$, so $y$ is globally feasible even when interpolation exits $B$, certifying $U_B \ge R^\star$.

Let $\omega(h)$ be the modulus of continuity of $W$ in $\|\cdot\|_\infty$ over the compact search box; $\omega(h) \to 0$ as $h \downarrow 0$. On any surviving box, $W(z) - k \le W(z) - W(u) \le \omega(h)$, giving $\alpha \le \omega(h)/\log(1 + \delta)$ (with $\alpha = 0$ if $z$ is feasible) and $\|y - z\|_1 \le 2s\alpha$. Applying a union bound across all $2M$ uniform ranks yields $|J(p) - J(p')| \le 2\sum_i |p_i - p_i'|$, and since $|(e^{-x})'| \le 1$ on $x \ge 0$, every surviving box satisfies

$$U_B - L_B \le 2Mh + \frac{4s\,\omega(h)}{\log(1 + \delta)}.$$

Evaluating at the box attaining $L$ controls $U - L$ and justifies finite-step termination. □

Structure speeds up pruning: under retain-and-redraw dynamics, monotonicity of each $f_i$ makes $1 - 2c + c^2 \exp(\sum_i f_i(l_i))$ an additional box lower bound, and any box whose lower bound exceeds the incumbent $U$ is discarded. Exact bounds assume exact function values; certified numeric implementations incorporate quadrature, quantile, and rounding errors into both endpoints, using floating-point solvers to propose tight candidates.

At zero budget ($\delta = 0$), strict convexity of $W$ leaves $x^0$ as the sole feasible point. Suppose $W$ is differentiable and convex with common slope $\eta = w_i'(x_i^0)$. The zero-budget feasible set is

then the intersection of $\sum_i x_i = s$ with the coordinate intervals $\{t : w_i(t) = w_i(x_i^0) + \eta(t - x_i^0)\}$. Subdivision applies directly on this set with gap $2Mh$.

Negative magnitude dependence fits the same framework. Take two independent coordinates where coordinate 1 alternates via $U_{t,1} = 1 - U_{t-1,1}$ from a uniform start, while coordinate 2 redraws independently. Then $D_1(p) = \max(2p - 1, 0)$. For $c > 1/2$, every feasible $p_1 \geq c$ gives $J = c^2(2p_1 - 1)/p_1^2$. Because this expression increases strictly in $p_1$, lowering risk narrows coordinate 1 instead. Proposition 6 also permits cross-coordinate lags when contemporaneous coordinates stay independent, simply by evaluating the full joint hit probability $J$.

## C.4 How a common factor changes the local allocation direction

A common-factor model shows how contemporaneous and temporal dependence interact when computing local allocation directions. Set

$$Z_{i,t} = a_i G_t + \sqrt{1 - a_i^2}\, \epsilon_{i,t}, \qquad i = 1, 2.$$

Here $|a_i| < 1$ are factor loadings, and $G_t$ is a stationary standard Gaussian AR(1) process with lag-one correlation $\rho$. The innovations $\epsilon_{i,t}$ are independent standard normals independent of $G$. Contemporaneous correlation is $\eta = a_1 a_2$, and coordinate lag-one autocorrelations are $a_i^2 \rho$. Setting $(a_1, a_2) = (0.95, 0.10)$ and $\rho = 0.8$ matches the lag-one autocorrelation 0.722 of coordinate 1 in Table 1. Table 1 keeps that temporal correlation while setting contemporaneous correlation to zero, isolating the role of the persistent axis.

To find this model's minimum-area rectangle, fix area via half-widths $u = qe^s$ and $v = qe^{-s}$ at fixed $q$. Differentiating the bivariate normal integral over the rectangle with respect to $s$ yields (up to a positive factor)

$$\int_0^1 \cosh\left(\frac{\eta u v z}{1 - \eta^2}\right) \left[\exp\left(-\frac{u^2 + v^2 z^2}{2(1 - \eta^2)}\right) - \exp\left(-\frac{u^2 z^2 + v^2}{2(1 - \eta^2)}\right)\right] dz,$$

where $\eta = a_1 a_2$. For $s > 0$, $u > v$ and $(u^2 - v^2)(1 - z^2) > 0$ on $[0, 1)$, making the derivative strictly negative; symmetry handles $s < 0$. Thus the square maximizes coverage at every fixed area and uniquely minimizes area at every fixed coverage.

A far-from-minimal swap shows why benchmarking near minimum area matters. Under $(a_1, a_2) = (0.95, 0.10)$ and $\rho = 0.8$, fixing one half-width at 6 requires the other to be $q \approx 1.64485$ for 0.90 coverage, at area 39.48 (far above the minimum 15.17). Swapping sides preserves coverage and area because the single-period bivariate law is exchange-symmetric, yet assigning the narrow side to coordinate 1 gives risk 0.04125 versus 0.01000 when assigned to coordinate 2. Comparing near-minimal regions therefore isolates the genuine effect of small volume budgets.

Conditioning on the factor simplifies two-period integration. For $\tau = (u, v)$, define

$$T_\tau(g) = \prod_{i=1}^{2} \left[\Phi\left(\frac{\tau_i - a_i g}{\sqrt{1 - a_i^2}}\right) - \Phi\left(\frac{-\tau_i - a_i g}{\sqrt{1 - a_i^2}}\right)\right].$$

Then $C = \mathbb{E}T_\tau(G_t)$ and $R_2 = 1 - 2C + \mathbb{E}[T_\tau(G_{t-1})T_\tau(G_t)]$, requiring only one- and two-dimensional Gaussian quadrature despite contemporaneous correlation.

Table S1 compares three configurations using root finding for exact coverage, Gauss–Hermite quadrature for $R_2$, and step-$10^{-4}$ central differences for slope. Increasing nodes per axis from 144 to 224 leaves all displayed digits unchanged. Heterogeneous loadings give a negative slope toward coordinate 1. Equal loadings make risk even in $s$ so the first-order slope vanishes, though 5% extra area still yields a small second-order drop. Setting $\rho = 0$ makes risk identically 0.01 along the entire equal-coverage curve. These three benchmarks let finite-sample experiments separate directional identification, symmetry, and temporal independence.

**Table S1** Population allocation in the two-dimensional common-factor model at fixed coverage 0.90. Slope is the risk derivative with respect to log-stretch $s$ along the equal-coverage curve at the square; the final column shows the feasible risk at 5% extra area with $s > 0$.

| Loadings $(a_1, a_2)$ | Factor $\rho$ | Min. area | Square $R_2$ | Slope | $R_2$ at +5% area |
|---|---|---|---|---|---|
| $(0.95, 0.10)$ | 0.8 | 15.17476 | 0.024134 | −0.076218 | 0.015613 |
| $(0.60, 0.60)$ | 0.8 | 14.94106 | 0.016798 | 0 | 0.016327 |
| $(0.95, 0.10)$ | 0 | 15.17476 | 0.010000 | 0 | 0.010000 |

# D Computational details for finite-budget frontiers

**Accuracy and cost in the three-dimensional non-Gaussian examples.** Convexity yields directly verifiable upper and lower bounds on global risk. Keep $W(x) = \sum_i \log[2F^{-1}(e^{-x_i})]$ and $k = \log K$ for absolute volume cap $K$, and set $f(x) = \sum_i \log[D_i(e^{-x_i})/e^{-2x_i}]$ so that $R_2 = 1 - 2c + c^2 e^{f(x)}$. Under Laplace marginals $W$ is convex, and $f$ is convex for the retain-and-redraw model. For the smooth rank-transition model, writing $z = e^{-x_i} \geq c$ gives the $i$-th summand $\log[1 + \theta_i(1-z)^2]$, whose second derivative in $x_i$ is

$$\frac{2\theta_i z\{2z - 1 - \theta_i(1-z)^2\}}{[1 + \theta_i(1-z)^2]^2}.$$

Because $0 \leq \theta_i \leq 1$, the numerator factor $2z-1-\theta_i(1-z)^2$ increases with $z$. If $2c-1-\theta_i(1-c)^2 \geq 0$, the summand is convex across $[c, 1]$. At $c = 0.90$, strict inequality holds. The verification script checks this condition per coordinate before running convex certification.

Given a budget-feasible candidate $a$ and any multiplier $\mu \geq 0$, set $g = \nabla(f + \mu W)(a)$. Supporting hyperplanes of convex functions, minimized over $\{x_i \geq 0,\ \sum_i x_i = s\}$, yield

$$f(a) + \mu[W(a) - k] + s \min_i g_i - g^\top a \leq \min_{\sum_i x_i = s,\, W(x) \leq k} f(x) \leq f(a).$$

This bound removes any need to trust solver stopping tolerances as accuracy certificates. We compute candidate points and multipliers from the first-order conditions, then blend slightly with $x^0$ for strict volume slack. Every arithmetic, exponential, and logarithmic step uses outward-rounded 60-digit decimal interval arithmetic, and final risk bounds are outward-rounded to 12 decimals. At 5% budget, the heterogeneous retain-and-redraw model gives [0.036278095832, 0.036278095836]. The smooth transition model gives [0.010862990181, 0.010862990182]. Both intervals include optimization gap and rounding error.

An independent box search confirms the joint allocations without reading convex solutions or multipliers. Eliminating $x_3 = s - x_1 - x_2$ and bisecting the $(x_1, x_2)$ rectangle along its longest edge prunes boxes that fail coverage or exceed the volume cap. On surviving boxes, monotonicity gives lower bounds both from $D_i(e^{-x_i})$ at right endpoints and from $D_i(e^{-x_i})/e^{-2x_i}$ at left endpoints. Feasible box centers and $x^0$ update the upper bound. Tightening tolerance from $10^{-4}$ to $10^{-5}$ in double precision inspects 4,279 and 15,483 boxes for heterogeneous retain-and-redraw at 1% budget. At 5% budget, it inspects 1,703 and 5,809 boxes for retain-and-redraw, and 201 and 915 boxes for smooth transition. All three tighter intervals agree with the decimal convex certificates. In a single local run, the tighter searches took about 0.056, 0.023, and 0.004 seconds, respectively. Counts include the root box and every child checked for feasibility after bisection.

**Finite-budget frontiers can have jumps.** A two-coordinate example shows that the left-continuity condition in Proposition 3 is needed for upper-bound convergence. Take $c = e^{-1}$ and $x_i = -\log p_i$, so coverage $c$ reads $(x_1, x_2) = (x, 1-x)$ for $0 < x < 1$. Let $P$ be the piecewise-linear function through $(0.10, 3)$, $(0.25, 1)$, $(0.50, 3)$, $(0.75, 0)$, and $(0.90, 3)$, held constant at 3 outside $[0.10, 0.90]$. Put $d(t) = \max\{0, \log(0.1/t)\}$ for $t > 0$, and define

$$w_1(t) = P(t) + 25(1-t) + d(t), \qquad w_2(t) = -25t + d(t), \qquad F_i^{-1}(e^{-t}) = \tfrac{1}{2}e^{w_i(t)}.$$

Because the steepest slope of $P$ has magnitude 20, both $w_i$ are continuous and strictly decreasing from $+\infty$ to $-\infty$, defining valid continuous marginals supported on $[0, \infty)$. Let the two coordinate processes be independent, with coordinate 1 retaining its prior uniform magnitude rank with probability 1/2 and redrawing uniformly otherwise, while coordinate 2 redraws independently each step. Then $Q(x) = 1 - 2c + c^2(1 + e^x)/2$ increases strictly in $x$, and the equal-coverage log-volume equals $P(x) + d(x) + d(1 - x)$, minimized at $x = 3/4$ with value 0. As $H \uparrow e$, the smallest feasible $x$ approaches 2/3 along the segment from 3/4. Exactly at $H = e$, the point $x = 1/4$ enters the feasible set, jumping the optimal allocation from $(2/3, 1/3)$ to $(1/4, 3/4)$ and dropping risk by $c^2(e^{2/3} - e^{1/4})/2 > 0$.

# E Coverage and consecutive-failure estimation for fixed rectangles

We estimate coverage and consecutive-failure risk simultaneously across all centered rectangles on one validation sequence. Let $Z_t \in \mathbb{R}^M$ be stationary with continuous magnitude marginals $|Z_{t,i}|$. For half-widths $a \in [0, \infty]^M$, put $B(a) = \{z : |z_i| \le a_i,\ 1 \le i \le M\}$ and $I_t(a) = \mathbf{1}\{Z_t \notin B(a)\}$, keeping population quantities $C(a)$ and $R_2(a)$ from (1). From $T$ consecutive observations, compute

$$\widehat{C}_T(a) = 1 - \frac{1}{T}\sum_{t=1}^{T} I_t(a), \qquad \widehat{R}_n(a) = \frac{1}{n}\sum_{t=1}^{n} I_t(a) I_{t+1}(a), \qquad n = T - 1.$$

**Lemma 2** (Simultaneous error bounds for fixed rectangles)**.** *Let $\beta(k)$ be the absolute regularity coefficient of $(Z_t)$. It equals the total-variation distance $\sup_E |P(E) - Q(E)|$ between the joint law of $(Z_s)_{s \le 0}$ and $(Z_s)_{s \ge k}$ and the product of their marginals. Fix integers $K \ge 1$, $2 \le q \le n$, and any $u > 0$, and define $m$ and $p_{K,q}(u)$ as follows.*

$$m = \lfloor n/q \rfloor,$$
$$p_{K,q}(u) = 4q(K+1)^M \exp(-2mu^2) + (n - q)\beta(q - 1).$$

*Outside an event of probability at most $p_{K,q}(u)$, both bounds below hold simultaneously for every $a \in [0, \infty]^M$.*

$$|\widehat{C}_T(a) - C(a)| \le u + \frac{M}{K} + \frac{1}{T}, \qquad |\widehat{R}_n(a) - R_2(a)| \le u + \frac{2M}{K}.$$

*Proof.* Bracket every coordinate half-width between adjacent population quantiles. On coordinate $i$, collect the $j/K$ quantiles of $|Z_{t,i}|$ for $1 \le j < K$ together with 0 and $+\infty$, forming a grid of $(K+1)^M$ rectangles. Each $B(a)$ satisfies $B^- \subseteq B(a) \subseteq B^+$ for adjacent grid rectangles. Marginal continuity bounds each coordinate's probability strip by $1/K$, so $C(B^+) - C(B^-) \le M/K$. Changing the consecutive-failure indicator requires at least one of the two adjacent errors to land in $B^+ \setminus B^-$. By stationarity, that event has probability at most $2M/K$.

Partition the pairs $(Z_t, Z_{t+1})$ into $q$ residue classes by $t$ mod $q$. Consecutive pairs inside one class are separated by $q - 1$ steps from the end of one pair to the start of the next. Replacing the joint law of the $m_r$ pairs in class $r$ sequentially by the product of independent pair marginals incurs total-variation cost at most $(m_r - 1)\beta(q - 1)$ via the independent-blocks method (Yu 1994). Under the product law, apply Hoeffding's inequality across all $(K+1)^M$ grid rectangles. The probability that either the first-step coverage average or the pair-risk average deviates by more than $u$ is at most $4(K+1)^M \exp(-2m_r u^2)$. Because total variation applies to the entire grid event at once, the mixing cost enters only once per residue class.

Weighting the $q$ classes by size recovers the full averages over all $n$ pairs, using $m_r \ge m$ and $\sum_r (m_r - 1) = n - q$. A union bound across classes yields $p_{K,q}(u)$. Monotonicity of coverage and

risk in rectangle size brackets the empirical averages of every $B(a)$ between $B^-$ and $B^+$, adding the population grid gaps $M/K$ and $2M/K$. Extending the empirical coverage average from the first $n$ origins to all $T$ observations shifts it by at most $1/T$. □

The lemma also applies to a data-selected vector of half-widths. For a random selection $a$, $C(a)$ and $R_2(a)$ denote evaluating that vector inside the stationary population probability functions. Provided each candidate uses one fixed vector of half-widths across the validation block, model selection and comparison fall on the same high-probability event of Lemma 2. This yields a uniform non-asymptotic convergence rate in sample size $T$ and absolute regularity coefficient $\beta(q-1)$ for both coverage and consecutive-failure frequency.

# F Propagating estimation error to the entire risk frontier

We measure how precisely finite history pins down the population frontier using two-period distribution error. Let $P$ be the target two-period joint law, $C_P(B)$ the average two-period coverage of $B$, and $Q_P(B) = P(B^c \times B^c)$ its consecutive-failure risk. Write $\mathcal{R}_P(H)$ for the infimum of $Q_P(B)$ over measurable $B$ with $C_P(B) = c$ and $|B| \le H$. Define $C_{\widehat{P}}$, $Q_{\widehat{P}}$, and $\mathcal{R}_{\widehat{P}}(H)$ analogously for a history-based estimate $\widehat{P}$. Total-variation distance is $d_{\mathrm{TV}}(P, \widehat{P}) = \sup_E |P(E) - \widehat{P}(E)|$ over measurable two-period events. When both period marginals match, average coverage reduces to single-period coverage.

**Proposition 7** (Error bound for the exact-coverage frontier)**.** *Assume both $P$ and $\widehat{P}$ have average marginals with densities, $d_{\mathrm{TV}}(P, \widehat{P}) \le e$, and a finite-volume reserve set $D$ satisfies $\widehat{C}(D) := C_{\widehat{P}}(D) > c + e$. Set*

$$s = \frac{e|D|}{\widehat{C}(D) - c - e}.$$

*Then at every feasible budget $H$ of the target problem,*

$$\mathcal{R}_{\widehat{P}}(H + s) - 3e \le \mathcal{R}_P(H) \le \mathcal{R}_{\widehat{P}}(H - s) + 3e, \tag{21}$$

*taking the infimum over an empty set as $+\infty$. If $P$ and $\widehat{P}$ share the same average marginal, then $|\mathcal{R}_{\widehat{P}}(H) - \mathcal{R}_P(H)| \le e$ at every feasible $H$ without requiring $D$.*

*Proof.* For any fixed set $B$, total-variation distance $e$ bounds the error of each single-period coverage event and the consecutive-failure event by $e$, so exact coverage under one law stays within $e$ of $c$ under the other. Write $C$ for the marginal needing coverage repair; both marginals satisfy $C(D) \ge \widehat{C}(D) - e > c$.

If $C(B) > c$, delete mass $C(B) - c$ from $B$ (feasible since $C$ has a density). Volume shrinks, and risk rises by at most $2[C(B) - c]$ because at least one of the two periods has to hit the deleted piece. If $C(B) < c$, add mass $d = c - C(B)$ from $D \setminus B$, which holds mass $m \ge C(D) - C(B) > d$. Collecting highest-density points in $D \setminus B$ (splitting ties by volume) attains average density at least $m/|D \setminus B|$. Added volume is thus at most $d\,|D \setminus B|/m \le s$, while expanding $B$ does not increase consecutive failures.

Applying this repair to any $\widehat{P}$-feasible region at budget $H - s$ produces a $P$-feasible region at budget $H$. Swapping laws shifts risk by at most $e$ and repair adds at most $2e$. Taking infima gives the upper bound. Reversing the roles of $P$ and $\widehat{P}$ at budget $H + s$ yields the lower bound. Under equal average marginals, feasible sets coincide and no repair is needed. □

The same repair controls relative-budget denominators. Let $V_P$ and $V_{\widehat{P}}$ denote minimum coverage-$c$ volumes; then $|V_P - V_{\widehat{P}}| \le s$. Evaluating (21) at $\widehat{H} \pm (2 + \delta)s$ with $\widehat{H} = (1 + \delta)V_{\widehat{P}}$ brackets $\mathcal{R}_P((1 + \delta)V_P)$. For any plugged-in deterministic set $B$ with $C_{\widehat{P}}(B) = c$, we have $|C_P(B) - c| \le e$, while exact population coverage requires target marginal information.

Above minimum volume, Proposition 5 converts budget slack directly into risk error. If $H - s \geq V_{\widehat{P}}$ and $H > V_{\widehat{P}}$, then

$$|\mathcal{R}_P(H) - \mathcal{R}_{\widehat{P}}(H)| \leq 3e + \frac{2\min(c, 1-c)s}{H - V_{\widehat{P}}}.$$

At any fixed positive volume allowance with bounded reserve volume and surplus mass bounded above zero, total-variation consistency of $\widehat{P}$ yields consistent estimation of the frontier value.

**Evaluating a history-selected region on a future error pair.** When $\widehat{B} = \widehat{B}(W)$ is chosen from history $W$, two quantities differ. First, $C_P(\widehat{B})$ plugs $\widehat{B}$ into a fresh independent draw from $P$. Second, deployment evaluates the actual linked future pair $Y = (Z_-, Z_+)$. Write $P = \mathcal{L}(Y)$ and $\rho = d_{\mathrm{TV}}(\mathcal{L}(W, Y), \mathcal{L}(W) \otimes P)$. If membership is jointly measurable, integrating bounded indicators against joint and product laws gives

$$\left| \mathbb{E} \frac{\mathbf{1}_{\widehat{B}}(Z_-) + \mathbf{1}_{\widehat{B}}(Z_+)}{2} - \mathbb{E} C_P(\widehat{B}) \right| \leq \rho.$$

The same bound $\rho$ separates realized consecutive-failure probability from $\mathbb{E}Q_P(\widehat{B})$. If $\widehat{B}$ has exact coverage under $\widehat{P}$ and $d_{\mathrm{TV}}(P, \widehat{P}) \leq e$ with probability at least $1 - \alpha$, actual average coverage stays within $e + \alpha + \rho$ of $c$. A gap between history and target lets mixing coefficients control $\rho$. Bounding $d_{\mathrm{TV}}(\mathcal{L}(Y \mid W), \widehat{P}(W))$ directly yields conditional coverage bounds on the corresponding history events.

A sharp example separates the two evaluations even with exact marginals. Let history $U \sim \mathrm{Unif}[0, 1]$, set both future errors equal to $U$, and take $\widehat{P} = P$. At coverage and length cap $1/2$, both halves $[0, 1/2]$ and $[1/2, 1]$ are population optimal. Choosing whichever half contains $U$ keeps plugged-in coverage $C_P(\widehat{B}) = 1/2$ while actual realized coverage equals 1. Uniform validation bounds (Appendix E) justify plugging random half-widths into stationary population functions, whereas deployment probabilities require controlling $\rho$.

**Estimating the joint law from adjacent observation pairs.** A cell-based estimator separates sampling error, within-cell approximation error, and distribution drift. Suppose training holds $n$ consecutive overlapping pairs $(Z_t, Z_{t+1})$ with average law $\bar{P} = n^{-1} \sum_{t=1}^{n} \mathcal{L}(Z_t, Z_{t+1})$, allowing non-stationary pairs. Partition the single-period space into $K$ measurable cells $E_i$, and assign a density-bearing fill distribution $\nu_i$ inside each cell. Estimate cell-pair masses $\pi_{ij} = \bar{P}(E_i \times E_j)$ by empirical frequencies $\widehat{\pi}_{ij}$ and set

$$P_{\mathcal{G}} = \sum_{i,j} \pi_{ij}\, \nu_i \otimes \nu_j, \qquad \widehat{P} = \sum_{i,j} \widehat{\pi}_{ij}\, \nu_i \otimes \nu_j.$$

Suppose $d_{\mathrm{TV}}(\bar{P}, P_{\mathcal{G}}) \leq b$ and $d_{\mathrm{TV}}(P, \bar{P}) \leq d$. For $2 \leq q \leq n$ and $u > 0$, we obtain (22).

$$\mathbb{P}\Big\{ d_{\mathrm{TV}}(P, \widehat{P}) > b + d + u \Big\} \leq q\, 2^{K^2} \exp\big\{ -2\lfloor n/q \rfloor u^2 \big\} + (n - q)\beta(q - 1). \tag{22}$$

Here $\beta(k)$ upper-bounds the absolute regularity coefficient across all split points of the training block under the same total-variation convention as Lemma 2.

To verify (22), group pairs modulo $q$ and replace the $m_r$ pairs in group $r$ by independent draws at total-variation cost at most $(m_r - 1)\beta(q - 1)$ (Yu 1994). For each of the $2^{K^2}$ cell-pair unions, Hoeffding's bound controls upward frequency deviation by $\exp(-2m_r u^2)$, and taking complements handles downward deviations. Weighting groups by size bounds the total-variation distance between empirical and population cell-mass vectors. Because fill distributions share support per cell and cells are disjoint, $d_{\mathrm{TV}}(P_{\mathcal{G}}, \widehat{P})$ matches the discrete cell-vector distance identically. Adding $b$ and $d$ completes the bound.

Grid resolution trades off approximation against estimation. Fill cells by conditioning a reference probability measure $\nu$ on bounded support. If the density of $\bar{P}$ relative to $\nu \otimes \nu$ is $L$-Lipschitz on cell pairs of diameter at most $h$, cell-averaging error is at most $Lh/2$. Finer grids reduce this approximation term while increasing $K$ in (22). Drift $d$ depends on the relation between training and target periods.

When adjacent errors take equal values with positive probability $\bar{P}(\Delta) = \rho > 0$ on the diagonal $\Delta = \{(x, x)\}$, continuous product fills assign zero mass to $\Delta$ and suffer total-variation error at least $\rho$. Record diagonal observations separately across $K$ diagonal classes (filling class $i$ via $(X, X)$ with $X \sim \nu_i$), and record off-diagonal pairs across $K^2$ product cells. All $K^2 + K$ fill laws are then mutually singular. Bound (22) therefore holds with $2^{K^2+K}$ in place of $2^{K^2}$.

For a fixed $\bar{P}$ whose diagonal part has an integrable density against $\nu$ and off-diagonal part has an integrable density against $\nu \otimes \nu$, take $\nu_i = \nu(\cdot \mid E_i)$ on positive-mass cells. Within-cell error then vanishes in $L^1$ as nested partitions generate the Borel $\sigma$-algebra. Consistent estimation of the target frontier still requires controlling sampling error, drift $d$, and the budget-repair conditions of Proposition 7.

**Necessity of cross-period dependence information (minimax lower bound).** Two processes that are identically distributed on any arbitrarily long training history show that knowing the single-period marginal does not identify the future consecutive-failure frontier. In both processes, training errors are independent standard logistic draws with distribution $F$ and density $f$, independent of the future target pair. Put $\psi(x) = 2F(x) - 1$ and give the target pair the two joint densities

$$p_{\pm}(x, y) = f(x)f(y)[1 \pm \psi(x)\psi(y)].$$

Both densities are smooth, strictly positive, and share marginal $f$ since $\int \psi f = 0$. For any fixed set $B$ of coverage $c = 1/2$, consecutive-failure risk equals $Q_{\pm}(B) = 1/4 \pm m(B)^2$, where $m(B) = \int_B \psi(x)f(x)\,dx$.

Set absolute length cap $H = \log(35/3)$. In rank coordinate $u = F(x)$, the interval $[1/4, 3/4]$ has mass $1/2$, length $2\log 3 < H$, and $m = 0$. The interval $[3/8, 7/8]$ has mass $1/2$, exact length $H$, and $m = \int_{3/8}^{7/8}(2u-1)\,du = 1/8$. Under $p_+$, the optimal risk is $1/4$. Under $p_-$, the optimal risk is at most $1/4 - 1/64$, separating the two frontiers by at least $1/64$. Any estimator reading only training data has identical distribution under both laws. For any exact-coverage feasible shape $\widehat{B}$ chosen from training data, $Q_+(\widehat{B}) + Q_-(\widehat{B}) = 1/2$ pointwise. The sum of expected excess risks across the two laws therefore equals the frontier gap. At least one of the two processes suffers expected excess risk at least $1/128$, even though complete single-period marginal knowledge is available and only future cross-period dependence differs.

# G Consecutive-failure frequencies by dataset

Risk differences plotted in Figure 4 of the main text are computed from the per-dataset frequencies in Table S2, alongside the post-hoc matched-to-reference volume ratios.

**Table S2** Consecutive-failure frequencies for one-step ridge forecasts by dataset. Equal-volume uses the identical per-origin volume as the rolling reference; matched misses scales each panel by a post-hoc constant to match reference failure counts exactly.

| Dataset | Rolling ref. | Equal-volume adj. | Matched misses | Matched vol. ratio |
|---|---|---|---|---|
| ETTh1 | .02079 | .01357 | .01617 | .697 |
| ETTh2 | .01617 | .01357 | .01501 | .855 |
| ETTm1 | .02501 | .01782 | .01976 | .845 |
| ETTm2 | .01588 | .01279 | .01452 | .882 |
| Electricity | .01991 | .01857 | .01940 | .877 |
| Traffic | .03124 | .02895 | .03153 | .722 |
| Six-dataset mean | .02150 | .01754 | .01940 | .813 |

# References

Dvoretzky, A., Wald, A., Wolfowitz, J.: Elimination of randomization in certain problems of statistics and of the theory of games. Proc. Natl. Acad. Sci. 36, 256–260 (1950). https://doi.org/10.1073/pnas.36.4.256

Heilman, S., Mossel, E., Neeman, J.: Standard simplices and pluralities are not the most noise stable. Isr. J. Math. 213, 33–53 (2016). https://doi.org/10.1007/s11856-016-1320-y

Mazari, I., Ruiz-Balet, D.: Quantitative stability for eigenvalues of Schrödinger operator, quantitative bathtub principle, and application to the turnpike property for a bilinear optimal control problem. SIAM J. Math. Anal. 54, 3848–3883 (2022). https://doi.org/10.1137/21M1393121

Yu, B.: Rates of convergence for empirical processes of stationary mixing sequences. Ann. Probab. 22, 94–116 (1994). https://doi.org/10.1214/aop/1176988849