MINIMAX LIMITS OF ADAPTIVE AVERAGING AND BATCHING IN SGD

Anonymous authors
Paper under double-blind review

ABSTRACT

Averaging SGD checkpoints can improve on the last iterate, while larger batches reduce gradient noise but consume more samples. We ask how much these choices can improve the worst-case guarantee under prescribed learning rates and the fixed SGD update rule. To characterize these limits, we study convex quadratics with curvature at most L and additive IID noise, using convex averages with the same averaging weights in every coordinate. We prove that, for fixed batch sizes, choosing averaging weights from the observed trajectory has the same minimax value as fixing them optimally before training. For joint averaging and batching, integer batch sizes may also adapt to training feedback, with a total sample budget and a maximum batch size per update enforced on every run. The minimax value then equals that of the best randomized plan specifying all batch sizes and averaging weights before training, independently of the instance and samples. These randomized plans can strictly improve on deterministic plans. Both equalities hold for any positive learning-rate sequence, with the worst case taken over all finite dimensions. For inverse-root rates $\gamma_t = c/\sqrt{t+1}$, $0 < c \le 1/(2L)$, we further prove that the optimal additional risk above the exact noiseless minimax value for the same rates is $\Theta(C^{-1/2}T^{-3/4} + \log(T)/T)$, with T updates, qT samples, and maximum batch size $C \ge q$. Here $q > 1$ is a fixed integer, and initialization and noise budgets are fixed, positive, and finite. We conclude that these exact values provide benchmarks for comparing learning-rate sequences and bounding the improvement available beyond feasible averaging and batching plans.

1 INTRODUCTION

Learning-rate schedules control how SGD balances optimization progress and gradient noise. Warmup-stable-decay (WSD) uses a constant-rate phase followed by cooldown (Wen et al., 2024), while WSqD replaces the constant phase with a shifted inverse-root schedule and retains the final cooldown (Ma & Chen, 2026). Recent theory gives last-iterate guarantees for cooldown (Schaipp et al., 2025; Ma & Chen, 2026). Its benefits also depend on the noise model and update rule (Singh et al., 2026).

The returned model need not be the last iterate. Tail averaging and exponential moving averages can improve the returned model by combining checkpoints (Jain et al., 2018; Au & Block, 2026). Their benefits depend on the averaging window and weights, while batch sizes determine how much gradient noise enters each update. Larger batches reduce this noise but consume more samples. Guarantees for particular rules do not, by themselves, identify the exact optimum over all admissible averaging and batching rules. With prescribed learning rates and the fixed SGD update rule, we therefore ask what the best worst-case guarantee is, and whether training feedback improves on choices made before training.

To characterize these limits, we study convex quadratics with additive IID noise. The returned model is a convex average of checkpoints, with the same averaging weights in every coordinate. Averaging weights may depend on the full observed trajectory, and integer batch sizes may depend on past observations. Every run must respect a total sample budget and a maximum batch size per update. We call the smallest worst-case expected objective gap over these rules the minimax value, taking the worst case over the stated model class in all finite dimensions.

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.

We answer this question with two exact minimax characterizations. With fixed batch sizes, choosing averaging weights from the observed trajectory has the same minimax value as fixing them optimally before training. For joint averaging and batching, a plan specifies all integer batch sizes and averaging weights before training. We prove that the best randomized plan, chosen independently of the instance and samples, attains the same minimax value as rules using training feedback. Such randomized plans can strictly improve on deterministic plans. Both equalities hold for any prescribed positive learning-rate sequence.

These characterizations also support sharp resource bounds and certified comparisons of learning-rate sequences. Our contributions are:

- **Exact minimax values for adaptive averaging and batching.** We show that the optimal fixed averaging weights and optimal randomized plans attain the respective minimax values over all admissible rules (Theorems 1 and 2).
- **Sharp additional risk under prescribed inverse-root rates.** Matching bounds quantify the additional risk above the exact noiseless minimax value for the same rates, separating the effects of the maximum batch size and total sample budget (Theorem 3). This isolates improvements hidden by the leading total-risk order.
- **Certified comparisons at finite budgets.** For a finite set of prescribed learning-rate sequences, we certify reversals of rankings based on the last iterate with uniform batches and bound the improvement available beyond feasible plans. One feasible plan has lower worst-case risk than every admissible policy with nondecreasing averaging weights on post-update checkpoints (Section 5).

Local LM continuations provide supplementary evidence for the effect of averaging weights on sample allocation (Section 6).

1.1 Related work

Luner & Grimmer (2025, Theorem 3.1) optimize nonnegative averaging weights that sum to one and are fixed before training. For smooth convex GD under an endpoint-risk condition, one Huber instance bounds every average. Hagedorn & Jarre (2024) derive finite-step errors and tune parametric weights for strongly convex stochastic quadratics. Wei et al. (2026) study weighted-average central limit theorems and adaptive weighting with nonasymptotic guarantees. These results do not identify the exact value over all history-dependent convex averages in our additive model. Our proof must account for the dependence between the averaging weights and the trajectory used to choose the weights.

Minibatching and tail averaging have problem-dependent benefits in least-squares regression with random sample Hessians (Jain et al., 2018; Mücke et al., 2019). Balles et al. (2017) study the coupling between batch sizes and learning rates. Wang et al. (2026) optimize sampling and training duration. Li et al. (2026) optimize batches for prescribed learning rates. Our comparison covers all admissible averaging and batching rules, with integer batch sizes and budget constraints that hold on every run. Appendix H.2 gives further comparisons. Local quadratic models of LM pretraining (Meterez et al., 2026b) motivate the supplementary mechanism test.

2 Setting and scope

Instances and trajectories. For $L > 0$ and $D, \sigma_* \geq 0$, let $\mathcal{Q}(L, D, \sigma_*)$ contain the centered quadratics

$$f_\xi(x) = \tfrac{1}{2} x^\top H x + \xi^\top x, \qquad 0 \preceq H \preceq LI, \qquad \mathbb{E}\xi = 0, \qquad \mathbb{E}\|x_0\|^2 \leq D^2, \quad \operatorname{tr}\operatorname{Cov}(\xi) \leq \sigma_*^2.$$

The population objective is $F = \mathbb{E}_\xi f_\xi$, with minimum $F^* = 0$. Future samples are IID and independent of x_0. For a fixed horizon T and prescribed positive rates $\gamma_0, \ldots, \gamma_{T-1}$, SGD follows

$$x_{t+1} = x_t - \gamma_t (H x_t + \xi_t). \tag{1}$$

A batch replaces ξ_t by the mean of fresh individual noises. The Hessian is fixed across samples. Gaussian replication is used for the general lower bounds; fixed-plan risk formulas require only the stated second moments. The resource law in Section 4 assumes $\gamma_t = c/\sqrt{t+1}$ with $c \leq 1/(2L)$.

Checkpoint averaging and sampling plans. The returned model (a scalar convex output) is $\widehat{x}_T = \sum_{t=0}^{T} w_t x_t$, where $w_t \geq 0$ and $\sum_t w_t = 1$. Each averaging weight w_t applies to every coordinate of checkpoint x_t. Fixed weights are chosen before sampling; adaptive weights may use the full history and independent auxiliary randomness. Denote the adaptive output class by $\mathcal{W}_T^{\mathrm{adaptive}}$. An allocation specifies batch sizes $B_0, \ldots, B_{T-1}$. History-dependent batching uses only past observations to select each size before drawing the batch. Every run must use integer batch sizes and respect the total sample budget and maximum batch size (cap). For fixed rates, a complete plan is the pair $p = (w, B)$. The horizon is deterministic.

Risk and quantifiers. The *class risk* of a policy is its worst expected objective gap over the stated instance class. A *minimax value* minimizes class risk over the allowed policies. Every class supremum includes all finite dimensions. Equality of minimax values permits adaptive gains on particular instances. The *additional risk* is the minimax value minus the exact noiseless minimax value for the same prescribed rates.

Shared scalar weights couple the output choice across spectral directions. The dimension supremum permits noise replication at fixed total budgets, which produces instances where observed fluctuations cannot improve the class guarantee.

3 Exact minimax characterizations

3.1 Fixed-plan risk and adaptive outputs

The final output is a linear function of the initialization and each noise injection. At curvature λ, let $a_\lambda(w)$ be the initialization coefficient and $b_{\lambda,j}(w)$ the coefficient of the noise at update j. Writing $p_{s,t}(\lambda) = \prod_{k=s}^{t-1}(1 - \lambda\gamma_k)$,

$$a_\lambda(w) = \sum_t w_t p_{0,t}(\lambda), \qquad b_{\lambda,j}(w) = \gamma_j \sum_{t>j} w_t p_{j+1,t}(\lambda). \tag{2}$$

The empty product equals one.

Lemma 1 (Fixed-output quadratic class risk)**.** *For deterministic scalar convex weights and prescribed positive rates,*

$$\mathfrak{R}(\gamma, w) := \sup_{\mathcal{Q}} \mathbb{E}F(\widehat{x}_w) = \underbrace{\frac{D^2}{2} \max_{0\leq\lambda\leq L} \lambda a_\lambda(w)^2}_{\textit{initialization}} + \underbrace{\frac{\sigma_*^2}{2} \max_{0\leq\lambda\leq L} \lambda \sum_{j<T} b_{\lambda,j}(w)^2}_{\textit{noise}}. \tag{3}$$

Two separate eigendirections attain the two maxima. The risk is convex in w for fixed γ.

Appendix A proves the identity.

Let Δ_T be the simplex on the $T+1$ states and set $\mathcal{V}(\gamma) = \min_{w\in\Delta_T} \mathfrak{R}(\gamma, w)$.

Theorem 1 (Adaptive convex-output minimax value)**.** *Fix $T \geq 1$, $L > 0$, $D, \sigma_* \geq 0$, and any prescribed $\gamma \in (0,\infty)^T$. Over $\mathcal{Q}(L, D, \sigma_*)$ in arbitrary finite dimensions,*

$$\boxed{\inf_{W\in\mathcal{W}_T^{\mathrm{adaptive}}} \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q[F_Q(\widehat{x}_T^W) - F_Q^*] = \mathcal{V}(\gamma).} \tag{4}$$

The same value equals $\sup_{Q\in\mathcal{Q}} \mathbb{E}_Q \min_{w\in\Delta_T}[F_Q(\sum_t w_t x_t) - F_Q^*]$. A deterministic minimizing weight vector attains the upper bound.

Even an oracle that observes the complete trajectory and its objective before choosing weights has the same minimax value. Section 5 uses the exact values to certify finite-budget comparisons.

3.2 Sampling feedback and data-independent randomization

The individual-sample class and its two noise budgets. Allowing batches to change the stochastic path, keep the rates prescribed. Let $\Sigma = \mathrm{Cov}(\xi)$. The individual-sample class $\mathcal{Q}_\Delta$ retains the Hessian and initialization constraints and imposes

$$\Sigma \ker H = \{0\}, \qquad \operatorname{tr}\Sigma \leq \sigma_{\mathrm{ind}}^2, \qquad \Delta_\star := \tfrac{1}{2}\operatorname{tr}(H^\dagger\Sigma) \leq \bar{\Delta}. \tag{5}$$

The kernel condition excludes noise in null directions. The inverse-curvature load $\Delta_\star$ equals $F^* - \mathbb{E}\inf_x f_\xi(x)$, the optimal objective difference of Loizou et al. (2021). Both noise budgets are total quantities and remain fixed as dimension grows. For example, if $\Sigma = \tau^2 H$, then $\Delta_\star = \tau^2 \operatorname{rank}(H)/2$; fixed total load does not keep the noise scale per direction fixed. Each update averages B_j fresh samples. Fix integers $q > 1$ and $C \geq q$, with total sample budget $N = qT$ and maximum batch size C.

The joint objective. For finite $\bar{\Delta}$, define the feasible spectral measures

$$\mathcal{M}_{\bar{\Delta}} = \left\{ \nu \geq 0 \text{ on } [0, L] : \int d\nu \leq \bar{\Delta}, \quad \int 2\lambda \, d\nu \leq \sigma_{\text{ind}}^2 \right\}.$$

Using the transfers $a_\lambda, b_{\lambda,j}$ from (2), a fixed plan has exact class risk

$$\mathfrak{R}_{\Delta,B}(w) = \underbrace{\frac{D^2}{2} \max_{0 \leq \lambda \leq L} \lambda a_\lambda(w)^2}_{\text{initialization}} + \underbrace{\max_{\nu \in \mathcal{M}_{\bar{\Delta}}} \int \lambda^2 \sum_{j<T} \frac{b_{\lambda,j}(w)^2}{B_j} \, d\nu(\lambda)}_{\mathcal{N}_{\Delta,B}(w)\text{: noise under both budgets}}. \tag{6}$$

The factor $1/B_j$ reduces the variance of the jth injection. Both moment constraints act on the same measure ν, with covariance mass $d\mu = 2\lambda \, d\nu$. The joint objective therefore uses both budgets together.

Sampling policy classes and their relaxation. Let $\mathcal{B}_C = \{B \in [1, C]^T : \sum_j B_j = qT\}$. The real relaxation and deterministic integer values are

$$V_T^{\text{rel}} = \min_{w \in \Delta_T,\, B \in \mathcal{B}_C} \mathfrak{R}_{\Delta,B}(w), \qquad V_T^{\text{det}} = \min_{w \in \Delta_T,\, B \in \mathcal{B}_C \cap \mathbb{Z}^T} \mathfrak{R}_{\Delta,B}(w). \tag{7}$$

For a fixed integer batch sequence, adaptive averaging weights have the same minimax value as fixed weights (Proposition C.2).

A data-independent randomized plan selects the complete pair (w, B) using a seed independent of the instance and samples; its batch coordinates need not be independent. The adversary knows the distribution and chooses the instance before the seed is realized. Write V_T^{mix} for the resulting minimax value.

A history-dependent (predictable) policy chooses B_j from preceding observations before drawing batch j, and chooses the final weights from the full observed history. It may also use independent auxiliary randomness. Every run uses integer batch sizes and respects the sample budget and batch cap. Write V_T^{pred} for its minimax value. The real-batch value V_T^{rel} in (7) is a mathematical benchmark.

Write $\ell_Q(w; B) = F_Q(\sum_t w_t x_t) - F_Q^*$. For $\bar{\Delta} = \infty$, replace the noise term by $\frac{\sigma_{\text{ind}}^2}{2} \max_{\lambda \leq L} \lambda \sum_j b_{\lambda,j}(w)^2 / B_j$.

Theorem 2 (Sampling randomness and history feedback)**.** *For every prescribed $\gamma \in (0, \infty)^T$, integers $q > 1, C \geq q$, and $\bar{\Delta} \in [0, \infty]$, under the pathwise budget $N = qT$,*

$$\boxed{V_T^{\text{rel}} \leq V_T^{\text{pred}} = V_T^{\text{mix}} \leq V_T^{\text{det}}.} \tag{8}$$

The randomized-plan objective keeps one instance outside the seed average, namely $\sup_Q \mathbb{E}_{\text{seed},Q} \ell_Q(w; B)$. Correlated rounding gives a randomized integer plan within $9/8$ of V_T^{rel} (Appendix C.3). Data-independent randomization can strictly improve on deterministic plans. For $T = 3$, $\gamma = (1, 1, 1)$, $q = 2, C = 3$ and $L = D = \sigma_{\text{ind}} = \bar{\Delta} = 1$, certified bounds give $V_T^{\text{mix}} \leq 0.104162655 < 0.104999793 \leq V_T^{\text{det}}$ (Appendix G.2).

3.3 Why one adversary controls all choices

A fixed-weight risk identity does not apply after substituting weights chosen from the same noisy trajectory. Nor does a separate bad instance for each fixed plan give a lower bound for adaptive choices. We first use minimax duality to construct a finite spectral adversary common to all feasible plans, with risk $g_A(w, B)$.

Replicate each noise atom into r independent Gaussian coordinates while preserving total variance and both noise budgets. Denote the resulting instance by Q_r. Form $G_i \in \mathbb{R}^{r\times N}$ from the $N = qT$ individual samples at atom i. For any $\varepsilon, \delta \in (0,1)$, sufficiently large finite r gives one event of probability at least $1-\delta$ on which

$$G_i^\top G_i/r \succeq (1-\varepsilon)I_N, \qquad \ell_{Q_r}(w;B) \geq (1-\varepsilon)g_A(w,B)$$

for every atom and every feasible integer plan (w, B). The event controls all real coefficient vectors and consecutive sample partitions simultaneously. Weights and partitions may therefore be chosen from the observations without escaping the bound. Nonnegativity off the event, the dimension supremum, and $\varepsilon, \delta \downarrow 0$ yield the adaptive lower bounds. Fixed weights and data-independent randomized plans attain the matching upper bounds (Appendices B and C.3). Figure 1 illustrates the loss of an observation-dependent advantage as noise directions are replicated.

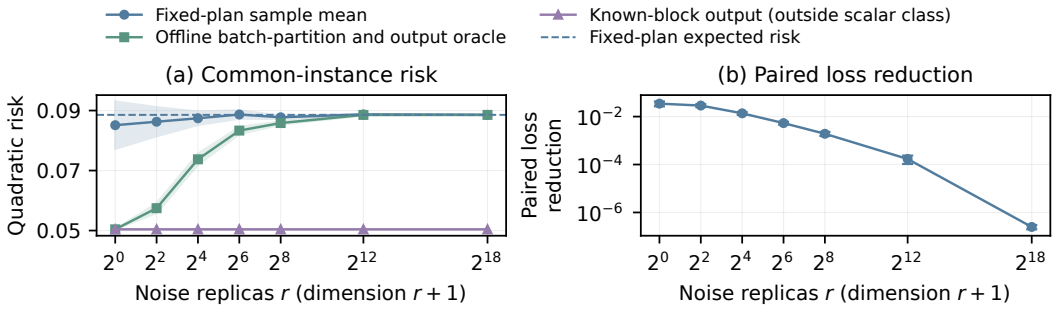


Figure 1: Common-instance risks and paired loss reduction under sequence A, with $T = 8, N = 16, C = 3, H_r = \mathrm{diag}(5/16, I_r)$, $x_0 = (1, 0_r)$ and noise $(0, z/\sqrt{r})$, $z \sim N(0, I_r)$. Points are means over 64 streams. Shaded bands in (a) and error bars in (b) show descriptive 95% confidence intervals for the means. The fixed-plan expected risk is about 0.088556054, from a guarded Float64 enclosure. Paired reduction is fixed-plan minus offline batch-partition and output oracle loss on the same stream. Partitions use consecutive samples without reordering. The known-block output lies outside the shared-scalar class (Appendix G.3).

3.4 Dimension and the output set

For $T = 1, \gamma_0 = 1$ in $\mathcal{Q}(1,1,1)$, the output set changes the minimax value. In dimensions at most d, Proposition B.1 gives

$$V_d^{\mathrm{scalar}} = \begin{cases} \frac{1}{4}, & d = 1, \\ \frac{2^{2/3}}{6}, & d \geq 2, \end{cases} \qquad V_d^{\mathrm{spectral}} = \frac{1}{4}.$$

The scalar value covers fixed weights, adaptive weights and the full path oracle. The spectral class knows H and permits different convex weights in different eigendirections. The scalar oracle also knows H, so the gap comes from the larger output set.

In the two-dimensional witness, both noise signs give the same loss for every scalar weight. Observing the sign adds no benefit, while expanding the output set does (Appendix B.6).

For a general horizon, let $V_d(\gamma)$ restrict the adaptive-output minimax problem to dimension at most d. Appendix B.4 gives $(1-\delta)(1-\varepsilon)\mathcal{V}(\gamma) \leq V_d(\gamma) \leq \mathcal{V}(\gamma)$ at an explicit sufficient dimension, for any $\delta, \varepsilon \in (0,1)$. These sufficient dimensions do not characterize sampling feedback at a fixed small dimension for general horizons.

4 A sharp additional-risk law for inverse-root rates

Consider $\gamma_t = c/\sqrt{t+1}$, $0 < c \leq 1/(2L)$, and $N = qT$. Every convex output has initialization transfer at least that of the endpoint. All four values therefore share the exact noiseless minimax

value V_0, attained by the endpoint. For $s \in \{\mathrm{rel}, \mathrm{det}, \mathrm{mix}, \mathrm{pred}\}$, define

$$V_0 := V_T^{\mathrm{det}}(C, 0) = \frac{D^2}{2} \max_{0 \le \lambda \le L} \lambda p_{0,T}(\lambda)^2, \qquad E_T^s(C) := V_T^s(C, \bar{\Delta}) - V_0. \tag{9}$$

At $C = q$, the common additional-risk value is denoted $E_T(q)$.

Theorem 3 (Additional-risk law and sampling saturation)**.** *Fix positive finite $L, D, \sigma_{\mathrm{ind}}, \bar{\Delta}$, c, with $c \le 1/(2L)$, and an integer $q > 1$. For $\gamma_t = c/\sqrt{t+1}$ and $N = qT$, there are $a, A > 0, T_0$, depending only on the fixed parameters, such that for every $T \ge T_0$, for every $s \in \{\mathrm{rel}, \mathrm{det}, \mathrm{mix}, \mathrm{pred}\}$ and all integer caps $C \ge q$,*

$$\boxed{a\Psi_T(C) \le \underbrace{V_T^s(C, \bar{\Delta}) - V_0}_{\textit{additional risk}} \le A\Psi_T(C),} \qquad \Psi_T(C) = \frac{1}{\sqrt{C}\, T^{3/4}} + \frac{\log T}{T}. \tag{10}$$

The exact noiseless value satisfies $V_0 = \Theta(T^{-1/2})$. A deterministic integer allocation and uniform averaging over the last $K = \max\{1, \lceil T^{3/4}/\sqrt{C} \rceil\}$ checkpoints attain the upper order without knowing $\bar{\Delta}$. The class supremum includes all finite dimensions.

Three costs determine the averaging window. For fixed positive parameters, the upper-bound construction balances

$$\underbrace{K/T^{3/2}}_{\text{initialization}} + \underbrace{1/(CK)}_{\text{recent noise under the cap}} + \underbrace{\log(T)/T}_{\text{total sample budget}}.$$

Longer averaging increases the initialization cost but reduces recent noise. Balancing the first two terms gives $K \asymp T^{3/4}/\sqrt{C}$. The third term dominates when C reaches order $\sqrt{T}/(\log T)^2$. The orders are sharp in T, C with q and the other positive model parameters fixed; the cap scale is asymptotic.

Proposition 1 (Finite-horizon guarantee for the explicit tail construction)**.** *Let $\gamma_j = c/\sqrt{j+1}$, $0 < c \le 1/(2L)$, $T \ge 4$, and integers $q > 1, C \ge q, 1 \le K \le \lfloor T/2 \rfloor$. Let $w^{(K)}$ average the last K post-update states and let B be an integer completion of* (D.21) *with $B_j \ge B_j^0$, $1 \le B_j \le C$, and $\sum_j B_j = qT$. For finite $\bar{\Delta} \ge 0$ and $h_T = \sum_{t=1}^T 1/t$,*

$$\mathfrak{R}_{\Delta,B}(w^{(K)}) - V_0 \le A_b \frac{K}{T^{3/2}} + \bar{\Delta}\left[\frac{4}{CK} + A_q \frac{h_T}{T}\right], \tag{11}$$

where A_b depends only on D, c, and A_q only on q. Their explicit values are given in (D.26)*.*

This explicit allocation establishes the finite-horizon upper bound and its asymptotic order. Section 5 separately certifies numerically configured plans under the specified class budgets. Algorithm 1 gives the full construction.

Why the logarithmic term is unavoidable. Let $\ell = \sum_t w_t(T - t)$ be the mean lookback. If $\ell \ge \sqrt{T}\log T$, initialization alone adds order $\log T/T$ above the exact noiseless baseline. Otherwise, most output mass lies near the endpoint. At large horizons there are $\Theta(\log T)$ disjoint update blocks at different time scales, each with a feasible noise curvature whose effect survives to those outputs. A block receiving n_s samples gives a lower bound proportional to $1/n_s$ on its corresponding instance. The blocks share only qT samples, so one receives at most $qT/\Theta(\log T)$. Each scale defines a separate feasible adversary. The resulting lower bound is uniform over relaxed plans and transfers to the adaptive classes through Theorem 2 (Appendix D.3).

For each fixed cap, $O(qT)$ dimensions suffice for the additional-risk order at large horizons, with constants and onset allowed to depend on the cap. Bounds uniform over caps hold for $d \ge d_{\mathrm{grow}}(T) = O(qT\log T)$, with constants and onset independent of cap and dimension (Corollary E.1).

The noiseless $\Theta(T^{-1/2})$ baseline limits the total-risk order of the prescribed inverse-root sequence. Constant-step averaging attains $(LD^2 + \bar{\Delta}/q)/T$ under the same model and budgets (Proposition E.1). Faster total-risk order therefore requires changing the rates or another assumption that determines the baseline.

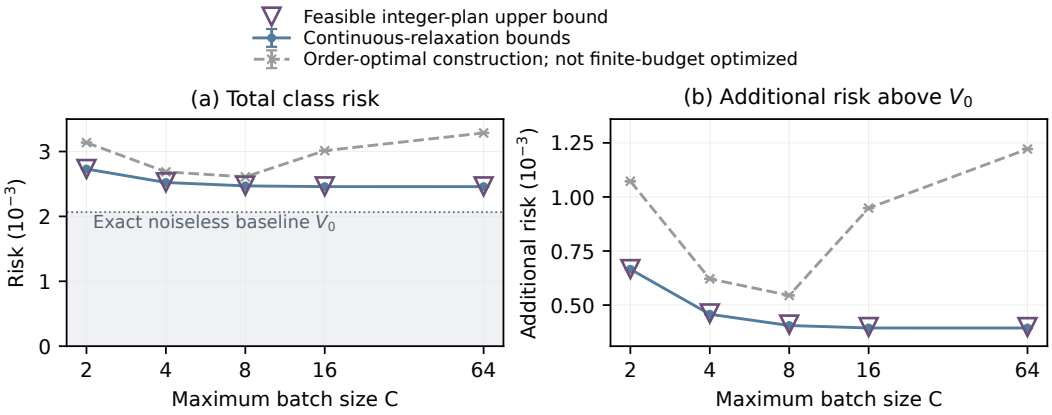


Figure 2: Total and additional risk versus maximum batch size for $\gamma_t = 1/(2\sqrt{t+1})$, $T = 2048$, $q = 2$ and $L = D = \sigma_{\text{ind}} = \bar{\Delta} = 1$. Both panels use the same exact noiseless baseline V_0. Larger caps enlarge the feasible set for the optimum; the rising gray curve is an order-optimal construction without finite-budget optimization. At $C = 16$, the integer-plan upper bound is within 0.212% of the same-cap relaxed lower bound in total risk. Bars are guarded Float64 bounds. Narrow intervals may overlap at plotting resolution. Appendix G.6.1 gives all plan comparisons.

5 Certified Consequences and Quadratic Evidence

5.1 Reversing the Ranking at Identical Contraction

Consider two permutations A and B of the same eight rates in $0 < \gamma_t \le 1/(2L)$, with $L = D = \sigma_{\text{ind}} = \bar{\Delta} = 1$, $q = 2$, and $C = 3$. They have the same noiseless endpoint transfer polynomial. Uniform-batch endpoint evaluation favors A. Yet B admits the integer allocation $B = (3, 3, 3, 1, 1, 1, 1, 3)$, returning $(x_7 + x_8)/2$, with class risk below every admissible policy for A. The certified comparison is $R_{A,\text{end},2} < R_{B,\text{end},2}$ and $V_B^{\text{pred}} \le V_B^{\text{det}} \le R_B^{\text{plan}} < 0.0851448 < 0.0884794 < V_A^{\text{pred}}$. In particular, $V_A^{\text{pred}}/V_B^{\text{det}} > 1.0391$. Appendix G.1 gives both sequences, endpoint intervals, exact-rational witnesses, and a neighborhood of the reversal. A finite-dimensional construction also gives a strict expected-risk separation (Appendix G.3).

5.2 Remaining Improvement of Selected Plans

Write $\theta = (T, L, D, \sigma_{\text{ind}}, \bar{\Delta}, q, C)$ and $R_\theta(p)$ for the class risk of a joint plan $p = (\gamma, w, B)$. The library Γ contains constant, inverse-root, plateau–linear and cosine rate sequences at 18 amplitudes. Its benchmark is $V_\Gamma^{\text{pred}}(\theta) = \min_{\gamma \in \Gamma} V_T^{\text{pred}}(\gamma; \theta)$, with the rate sequence selected deterministically before sampling (Appendix G.4).

Certified gaps for selected plans. The simple-output search uses all integer tail lengths, finite grids of initialized EMA and decreasing geometric weights, and interval averages with optional initial and terminal mass. A full-simplex convex-weight plan permits all $w \in \Delta_T$. Both searches configure integer batches under the same budgets. Table 1 reports six settings with $L = D = \sigma_{\text{ind}} = 1, q = 2, C = 8$.

The final column compares two selected feasible plans. Corollary 1 separately certifies separation from the complete nondecreasing-weight policy class.

5.3 Separation from a Complete Averaging-Policy Class

Let $\mathcal{C}_\uparrow = \{w \in \Delta_T : w_1 \le \cdots \le w_T\}$, with w_0 free. The set contains every tail length, every initialized EMA decay in $[0, 1]$, and their convex mixtures. Write $V_{\gamma,\uparrow}^{\text{pred}}$ for history-dependent integer batching and output selection with realized weights in $\mathcal{C}_\uparrow$. Independent auxiliary randomization is allowed; cap and sample budget hold on every path.

T	$\bar{\Delta}$	Gap to full policy-class value Upper bound (%) Selected simple	Selected full-simplex	Risk reduction between selected plans (%) Lower bound
64	1/4	≤ 19.22	≤ 1.44	≥ 14.83
64	1	≤ 6.77	≤ 2.39	≥ 4.03
64	4	≤ 3.08	≤ 1.23	≥ 1.74
256	1/4	≤ 16.21	≤ 3.43	≥ 10.94
256	1	≤ 6.49	≤ 2.20	≥ 3.97
256	4	≤ 9.63	≤ 1.64	≥ 7.22

Table 1: Separately selected joint plans on the 72-rate library. Rates, weights, and batches may differ between the two plans. The first-row full-simplex entry uses constant rate $31/16$; the simple entry uses plateau–linear rates with amplitude $15/8$. The gap columns upper-bound $100[R_\theta(p)/V_\Gamma^{\mathrm{pred}}(\theta) - 1]$. The final column lower-bounds $100[1 - R_\theta(p_F)/R_\theta(p_S)]$ for the selected simple and full-simplex plans p_S, p_F. Bounds use Arb arithmetic. Appendix G.4 gives the risk intervals and uniform-batch comparisons.

Corollary 1 (Nondecreasing averaging weights are insufficient)**.** *For $T = 64, q = 2, C = 8, L = D = \sigma_{\mathrm{ind}} = 1, \bar{\Delta} = 1/4$ and the 72-rate library Γ, a feasible full-simplex integer plan p_F with constant rate $31/16$ satisfies*

$$R_\theta(p_F) \leq \underbrace{3.003532 \times 10^{-3}}_{U_F} < \underbrace{3.057313 \times 10^{-3}}_{L_{\uparrow,72}} \leq \min_{\gamma \in \Gamma} V_{\gamma,\uparrow}^{\mathrm{pred}}. \tag{12}$$

The rate sequence is selected deterministically before sampling. The instance supremum includes all finite dimensions.

The certified reduction is at least 1.75%. Proposition C.3 transfers the common-adversary certificates to history-dependent choices. Figure 3 shows the witness; Appendix G.4.1 gives the proof. Randomized rate selection is outside this comparison. Across nine prespecified settings, separately optimized feasible plans have certified lower bounds on relative risk reduction of 1.6397%–1.8303%. A reference witness with rate $15/8$ retains strict separation at all three prespecified loads for $T = 64$ (Appendix G.5).

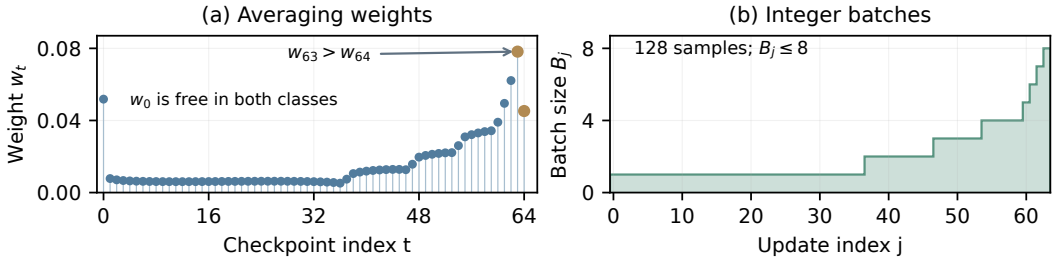


Figure 3: The constant-rate $31/16$ plan from Corollary 1, with certified bounds $U_F = 0.003003532 < L_{\uparrow,72} = 0.003057313$. All 65 weights and 64 integer batches are shown. The initial weight is free in both classes. The orange points mark the decrease $w_{63} > w_{64}$, which violates the post-update constraint. Batch B_j maps x_j to x_{j+1}. The certificates establish the risk separation; the shape identifies a feasible averaging rule outside the class.

6 Local tests in language-model training

We test how prescribed averaging weights change the effect of sample allocation. For a fixed local quadratic model with additive noise, expected output loss has the form

$$\mathcal{L}_{\mathrm{local}}(w, B) = \beta(w) + \sum_{j<T} \frac{v_j(w)}{B_j}. \tag{13}$$

Here $v_j(w)$ is the expected loss contribution of noise injection j before division by batch size; $\beta(w)$ is independent of allocation. There is no instance supremum. For the Tail-64 and Tail-32 outputs, averaging states 65–128 and 97–128 of a 128-update continuation, and their respective allocations,

$$I_{\text{local}} = \sum_{j<T} [v_j(w_{64}) - v_j(w_{32})] \left(\frac{1}{B_{32,j}} - \frac{1}{B_{64,j}} \right). \tag{14}$$

Different noise-response coefficients make the same allocation change affect the two outputs differently.

We test calibrated and triangular allocations in a 124M-parameter plain-SGD LM, using two fixed pretrained anchors and 16 independent streams per anchor. Each 128-update continuation uses 1024 contexts and returns Tail-64 and Tail-32 averages. Calibrated allocations use a fixed generalized Gauss–Newton fit; triangular allocations use no gradients. Allocations and audit predictions are frozen before the streams. If $F_{u,v}$ is evaluation cross-entropy for the last-u output under the last-v allocation, the observed interaction is $I = F_{64,32} + F_{32,64} - F_{64,64} - F_{32,32}$. It compares summed losses of crossed and matched pairs, not the gain of one returned model.

		Interaction I, in 10^{-5} nats/token		
Anchor	Allocation	Frozen prediction	Observed I [98.75% interval]	Underprediction (%)
2	Calibrated	21.680	24.803 [22.402, 27.203]	12.59
2	Triangular	11.518	13.611 [12.319, 14.903]	15.38
3	Calibrated	20.168	24.886 [22.558, 27.214]	18.96
3	Triangular	10.856	13.638 [11.642, 15.634]	20.40

Table 2: Frozen predictions evaluated on 16 streams per anchor. Individual 98.75% Student t_{15} intervals are Bonferroni-adjusted for four comparisons to target 95% simultaneous coverage, conditional on the fixed anchors, fits and evaluation set. Underprediction is $100(\text{observed mean} - \text{frozen prediction})/\text{observed mean}$. Component contrasts are in Table I.2.

All four interactions are positive under the prespecified simultaneous procedure. All four frozen predictions satisfy the half-gain criterion of less than 50% relative error. They underestimate the observed means by 12.59%–20.40%, each falling below its primary interval. Inference conditions on the anchors, fits and evaluation set (Appendix I.2). Matched-output mean gains from calibration are $(1.10\text{–}2.52) \times 10^{-5}$ nats/token; all four descriptive 95% intervals are positive. These secondary comparisons do not establish efficiency including fitting costs. Appendix I.3 reports the prospective fixed-output test. In the separate output-selection experiment, the fitted objective selected endpoint plans that underperformed the fixed tail and EMA outputs (Appendix J). The fixed-output experiments test how averaging weights change the value of sample allocation in local LM continuations.

7 Discussion and implications

The minimax values bound the improvement available beyond a feasible plan. Suppose U is an upper bound on the plan's class risk and $L > 0$ is a lower bound on the minimax value. The plan's relative suboptimality is at most $U/L - 1$. The bound accounts for all admissible rules, including rules not tested in the experiments. Section 5.1 uses the bounds to compare learning-rate sequences. Corollary 1 certifies separation from the entire nondecreasing-weight class.

Table G.9 gives a concrete example. At $\gamma_t = 1/(2\sqrt{t+1})$, $T = 2048, q = 2$, and $L = D = \sigma_{\text{ind}} = \bar{\Delta} = 1$, a cap-16 integer plan is within 0.218484% of the cap-64 minimax total risk. No admissible averaging and batching rule with cap at most 64 can reduce that plan's worst-case total risk by 1%, even with training feedback. A 1% improvement therefore requires a change outside the stated class, such as different learning rates or a larger output set.

The minimax value of training feedback at fixed small dimension remains open. Random sample Hessians and larger output sets also require further analysis. The LM tests examine how fixed averaging weights change the value of sample allocation, conditional on the anchors and fits.

AI USE STATEMENT

We used AI tools mainly to polish the manuscript and check the text for clarity and consistency. The authors take responsibility for the final manuscript.

REPRODUCIBILITY STATEMENT

The supplementary package opens with `README.md` and `EVIDENCE_INDEX.md`, which list the build commands, certificate checks and statistical recomputation scripts. Saved plans, certificate inputs and per-context losses support CPU verification without model weights. The verification scripts independently recompute the reported quadratic oracle risks, LM contrasts and intervals. Rerunning the LM continuations from scratch additionally requires the recorded model checkpoints, raw data and GPU runtime; full weights and corpora are not embedded in the package. Appendices A–E give the proofs. Appendix F distinguishes Float64 bounds, exact rational witnesses and Arb certificates, and Appendix G reports the numerical results supporting the main claims. Appendices I and J give the LM protocols, unresolved comparisons, calibration checks and costs.

REFERENCES

Naman Agarwal, Surbhi Goel, and Cyril Zhang. Acceleration via fractal learning rate schedules. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 87–99. PMLR, 2021. URL https://proceedings.mlr.press/v139/agarwal21a.html.

Kwok Chun Au and Adam Block. Training for the model you return: Improving optimization for iterate-averaged language models. *arXiv preprint arXiv:2606.25086v2*, 2026. URL https://arxiv.org/abs/2606.25086v2.

Francis Bach and Eric Moulines. Non-strongly-convex smooth stochastic approximation with convergence rate $O(1/n)$. *arXiv preprint arXiv:1306.2119*, 2013. URL https://arxiv.org/abs/1306.2119.

Lukas Balles, Javier Romero, and Philipp Hennig. Coupling adaptive batch sizes with learning rates. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, 2017. URL https://www.auai.org/uai2017/proceedings/papers/141.pdf.

Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, and Manish Purohit. Adaptive weighted averaging. In *Proceedings of Thirty Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 684–707. PMLR, 2026. URL https://proceedings.mlr.press/v336/bhaskara26a.html.

Jaeyoung Cha, Jaewook Lee, and Chulhee Yun. Tighter lower bounds for shuffling SGD: Random permutations and beyond. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 3855–3912, 2023. URL https://proceedings.mlr.press/v202/cha23a.html.

DKYoon. SlimPajama-6B. Hugging Face dataset, 2023. URL https://huggingface.co/datasets/DKYoon/SlimPajama-6B. Revision b5f90f419b7489cdba26fdbc8c022fcb5562f968.

Rong Ge, Sham M. Kakade, Rahul Kidambi, and Praneeth Netrapalli. The step decay schedule: A near optimal, geometrically decaying learning rate procedure for least squares. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://arxiv.org/abs/1904.12838v2.

Melinda Hagedorn and Florian Jarre. Optimized convergence of stochastic gradient descent by weighted averaging. *Optimization Methods and Software*, 39(4):699–724, 2024. doi: 10.1080/10556788.2024.2306383. URL https://doi.org/10.1080/10556788.2024.2306383.

Prateek Jain, Sham M. Kakade, Rahul Kidambi, Praneeth Netrapalli, and Aaron Sidford. Parallelizing stochastic gradient descent for least squares regression: Mini-batching, averaging, and model misspecification. *Journal of Machine Learning Research*, 18(223):1–42, 2018. URL https://jmlr.org/papers/v18/16-595.html.

Guy Kornowski and Ohad Shamir. Gradient descent's last iterate is often (slightly) suboptimal. *arXiv preprint arXiv:2604.13870*, 2026. URL https://arxiv.org/abs/2604.13870.

Jiaxiang Li, Zhiqi Bu, and Shiyun Xu. Towards joint scaling laws with optimal batch size schedules. *arXiv preprint arXiv:2607.27731*, 2026. URL https://arxiv.org/abs/2607.27731.

Nicolas Loizou, Sharan Vaswani, Issam Hadj Laradji, and Simon Lacoste-Julien. Stochastic Polyak step-size for SGD: An adaptive learning rate for fast convergence. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 1306–1314. PMLR, 2021. URL https://proceedings.mlr.press/v130/loizou21a.html.

Alan Luner and Benjamin Grimmer. On averaging and extrapolation for gradient descent. *Mathematics of Operations Research*, 2025. doi: 10.1287/moor.2024.0449. URL https://arxiv.org/abs/2402.12493. Theorem numbering follows arXiv version 4.

Jianhao Ma and Yuxin Chen. WSqD: A horizon-free learning rate schedule for large model training. *arXiv preprint arXiv:2607.10959*, 2026. URL https://arxiv.org/abs/2607.10959.

Alexandru Meterez, Pranav Ajit Nair, Depen Morwani, Cengiz Pehlevan, and Sham Kakade. Anytime pretraining: Horizon-free learning-rate schedules with weight averaging. *arXiv preprint arXiv:2602.03702*, 2026a. URL https://arxiv.org/abs/2602.03702v2. Version 2, August 25, 2026.

Alexandru Meterez, Pranav Ajit Nair, Depen Morwani, Cengiz Pehlevan, Sham Kakade, and Alex Damian. A defense of the quadratic model. *arXiv preprint arXiv:2607.21716*, 2026b. URL https://arxiv.org/abs/2607.21716v1.

Nicole Mücke, Gergely Neu, and Lorenzo Rosasco. Beating SGD saturation with tail-averaging and minibatching. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://papers.neurips.cc/paper_files/paper/2019/hash/4d0b954f0bef437c29dfa73fafdf3fa5-Abstract.html.

Fabian Schaipp, Alexander Hägele, Adrien Taylor, Umut Simsekli, and Francis Bach. The surprising agreement between convex optimization theory and learning-rate scheduling for large model training. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 53267–53294, 2025. URL https://proceedings.mlr.press/v267/schaipp25a.html.

Subham Singh, Ashutosh Mishra, and Subha Raut. Same loss, same noise, opposite schedules: Noise structure and optimizer normalization jointly determine whether learning-rate cooldown helps. *arXiv preprint arXiv:2607.12360*, 2026. URL https://arxiv.org/abs/2607.12360.

Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958. URL https://msp.org/pjm/1958/8-1/pjm-v8-n1-p14-p.pdf.

Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In Yonina C. Eldar and Gitta Kutyniok (eds.), *Compressed Sensing: Theory and Applications*, pp. 210–268. Cambridge University Press, 2012. URL https://arxiv.org/abs/1011.3027.

Jinbo Wang, Binghui Li, Zhanpeng Zhou, Mingze Wang, Yuxuan Sun, Jiaqi Zhang, Xunliang Cai, and Lei Wu. Fast catch-up, late switching: Optimal batch size scheduling via functional scaling laws. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=PXWgzUkVwo.

G. W. Wasilkowski and H. Woźniakowski. Can adaption help on the average? *Numerische Mathematik*, 44(2):169–190, 1984. doi: 10.1007/BF01410103. URL https://doi.org/10.1007/BF01410103.

Ziyang Wei, Wanrong Zhu, and Wei Biao Wu. General weighted averaging in stochastic gradient descent: CLT and adaptive optimality. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 1198–1206, 2026. URL https://proceedings.mlr.press/v300/wei26b.html.

Kaiyue Wen, Zhiyuan Li, Jason Wang, David Hall, Percy Liang, and Tengyu Ma. Understanding warmup-stable-decay learning rates: A river valley loss landscape perspective. *arXiv preprint arXiv:2410.05192*, 2024. URL https://arxiv.org/abs/2410.05192.

Difan Zou, Jingfeng Wu, Vladimir Braverman, Quanquan Gu, and Sham M. Kakade. Benign overfitting of constant-stepsize SGD for linear regression. *Journal of Machine Learning Research*, 24(326):1–58, 2023. URL https://jmlr.org/papers/v24/21-1297.html.

GUIDE TO CLAIMS AND EVIDENCE

The proofs concern prescribed learning rates on additive-noise quadratics. The LM studies test a local sampling mechanism. The entries below give the verification path for each main claim; theorem statements specify the full assumptions.

Exact minimax values. Theorems 1 and 2 are proved in Appendices A–C. The dimension supremum is essential to the general equalities. Appendices B.4 and B.6 give finite-dimensional approximation and exact one-step comparisons.

Resource law and finite construction. Appendix D.3 proves Theorem 3 and Proposition 1. Appendix E gives the finite-dimensional scope and constant-step comparison. Appendix G.6 reports the cap comparison, finite construction costs and conservative analytic constants.

Certified finite-budget comparisons. Appendix F gives the continuous-curvature verification method and distinguishes guarded Float64 bounds from exact-rational and Arb certificates. Appendices G.1 and G.3 give the ranking reversal and its finite-dimensional certificate. Appendix G.2 gives the strict randomization witness, with the supplement's `check_randomized_witness.py` as its CPU check. Appendix G.4.1 proves Corollary 1 using the selected $31/16$ plan and all 72 lower certificates. Appendix G.5 reports the nine-point confirmation and the separate $15/8$ reference witness's load check. The supplement's `check_core.py`, `check_confirmation.py` and `replay_selected_witness.py` provide the corresponding CPU checks.

LM protocol, results and limits. Appendix I.1 defines the model, data, allocations and predictions. Appendix I.2 gives the sixteen-stream results in Table 2, including prediction underestimation. Appendix I.3 reports the prospective fixed-output test. Sampling controls and reused-stream calibration checks are in Appendix I.4; earlier accuracy failures and negative restart results are in Appendix I.5. Appendix J reports the failed output selector. These results condition on the reported anchors, fits and evaluation sets and do not establish end-to-end efficiency.

The supplementary `EVIDENCE_INDEX.md` maps each result to its data and verification script. Historical experimental records remain in the supplement, outside the current paper's evidence sequence.

A FIXED-OUTPUT CLASS RISK AND CONTROLLED INNOVATIONS

How a fixed output retains shared noise. On an additive quadratic, let $\Phi_{t,s} = \prod_{k=s}^{t-1}(I - \gamma_k H)$, with empty product I, and write $\|v\|_H^2 = v^\top H v$. For deterministic weights set $U_w = \sum_t w_t \Phi_{t,0}$ and $K_{j,w} = \gamma_j \sum_{t>j} w_t \Phi_{t,j+1}$. Independent centered innovations with covariance Σ_j give, conditionally on x_0,

$$\begin{aligned} \widehat{x}_w &= U_w x_0 - \sum_{j<T} K_{j,w}\xi_j, \\ \mathbb{E}F(\widehat{x}_w) &= \tfrac{1}{2}\|U_w x_0\|_H^2 + \tfrac{1}{2}\sum_{j<T} \operatorname{tr}(H K_{j,w}\Sigma_j K_{j,w}^\top). \end{aligned} \tag{A.1}$$

The unrolled recurrence explains why averaging correlated checkpoints need not remove an early innovation: its transfers add *before* squaring. The rate sequence controls these transfers; the covariance budget controls their allowed allocation. Weights selected after observing these innovations do not satisfy this fixed-kernel covariance formula. Our adaptive-output lower bounds instead use a common event for the entire path.

Proof of (A.1) **and Lemma 1.** Unroll $x_{t+1} = (I - \gamma_t H)x_t - \gamma_t \xi_t$ and interchange the two finite sums. This gives $\widehat{x}_w = U_w x_0 - \sum_{j<T} K_{j,w}\xi_j$. Independence, centering and independence from initialization eliminate all cross-innovation and initialization–innovation terms in the conditional second moment, proving (A.1). For random initialization, average its displayed first term over x_0. This finite expansion uses no sign or contraction bound on the transfer products.

Diagonalize $H = V\,\mathrm{diag}(\lambda_i)V^\top$, and put $z = V^\top x_0$, $S = V^\top \Sigma V$. The fixed weights and stationary covariance of $\mathcal{Q}$ give

$$\mathbb{E}F(\widehat{x}_w) = \frac{1}{2}\sum_i \lambda_i a_{\lambda_i}(w)^2 \mathbb{E}z_i^2 + \frac{1}{2}\sum_i S_{ii}\lambda_i \sum_{j<T} b_{\lambda_i,j}(w)^2.$$

Off-diagonal covariance entries do not enter this trace. The constraints $\sum_i \mathbb{E}z_i^2 \le D^2$ and $\sum_i S_{ii} \le \sigma_*^2$ give the two upper maxima in (3). Their continuous polynomial integrands attain maxima on $[0, L]$. Choose their respective maximizers λ_b, λ_n and take

$$H = \mathrm{diag}(\lambda_b, \lambda_n), \quad x_0 = (D, 0), \quad \xi_j = (0, \sigma_* Z_j), \qquad Z_j \overset{\text{iid}}{\sim} \mathcal{N}(0,1).$$

This is one admissible two-dimensional instance attaining their sum. The directions may have equal curvature; the initialization and noise budgets remain separate. If either budget is zero its term vanishes; zero curvature and constant outputs cause no exception.

For fixed γ, each a_λ and $b_{\lambda,j}$ is linear in w. Each nonnegative weighted squared norm is convex; pointwise maxima and their sum preserve convexity. This proves the stated convexity on the simplex. It establishes neither joint convexity in (γ, w) nor the corresponding identity for weights depending on future innovations. The same identity also covers the raw endpoint $w_T = 1$.

Why the order of optimization matters. For a specified task and rate sequence, the deterministic optimum $\min_w \mathbb{E}F(\widehat{x}_w)$ chooses weights before sampling; the population path oracle $\mathbb{E}\min_w F(\widehat{x}_w)$ can inspect the full path and the true objective. The class benchmark $\min_w \sup_{P\in\mathcal{Q}} \mathbb{E}F_P(\widehat{x}_w)$ instead selects weights against every allowed instance. Neither an oracle on one task nor the fixed-weight covariance calculation alone proves the adaptive minimax identity. Appendix B supplies a common spectral instance and replication argument for that equality.

B General rate sequences and adaptive convex-output risk

We prove Theorem 1. Throughout this appendix, $T < \infty$ and $\gamma \in (0,\infty)^T$ are fixed, and the supremum permits all finite dimensions. This dimension quantifier is essential to the argument. The construction uses independent additive Gaussian noise; the upper bound covers the entire class $\mathcal{Q}$.

Finite transfers do not require stepwise contraction. For every fixed rate, $\sup_{\lambda\in[0,L]} |1-\lambda\gamma_t| = \max\{1, |1-L\gamma_t|\} < \infty$. Thus all finite transfer polynomials are continuous and bounded. Finite second moments and the diagonalization in Appendix A remain valid even when products change sign. The matrices in (B.1) are positive semidefinite because they are outer products or Gram matrices, not because their entries are nonnegative. Likewise, the event (B.6) controls every vector in $\mathbb{R}^T$, regardless of sign. These are the only properties of the rates used in the minimax and replication steps below; their bounds need not be uniform over rate sequences.

B.1 A common finite spectral adversary

Write $n = T+1$, and let $\mathbf{a}_\lambda \in \mathbb{R}^n$ and $\mathbf{B}_\lambda \in \mathbb{R}^{T\times n}$ be the linear maps with $a_\lambda(w) = \mathbf{a}_\lambda^\top w$ and $b_\lambda(w) = \mathbf{B}_\lambda w$. Define the two matrix curves

$$A_\lambda = \lambda \mathbf{a}_\lambda \mathbf{a}_\lambda^\top, \qquad C_\lambda = \lambda \mathbf{B}_\lambda^\top \mathbf{B}_\lambda, \qquad 0 \le \lambda \le L. \tag{B.1}$$

Their entries are continuous polynomials for this finite rate sequence. Their convex hulls $\mathcal{A}, \mathcal{C}$ are compact convex subsets of the finite-dimensional space of symmetric n-by-n matrices. Indeed, Carathéodory's theorem represents every point of either convex hull using at most $n(n+1)/2+1$ curve points; the set of these representations is a continuous image of a compact simplex times a compact product of curvature intervals.

For $w \in \Delta_T$, the objective defining $\mathcal{V}$ is

$$\max_{A\in\mathcal{A},\, C\in\mathcal{C}} G(w; A, C), \qquad G(w; A, C) = \frac{D^2}{2} w^\top A w + \frac{\sigma_*^2}{2} w^\top C w. \tag{B.2}$$

The function is continuous, convex in w, and affine in (A, C). The compact convex minimax theorem (Sion, 1958) therefore gives

$$\mathcal{V}(\gamma) = \max_{A\in\mathcal{A}, C\in\mathcal{C}} \min_{w\in\Delta_T} G(w; A, C). \tag{B.3}$$

All extrema here are attained. In particular, continuity of the minimum over the compact simplex and compactness of $\mathcal{A} \times \mathcal{C}$ give an optimal pair (A_*, C_*). Choose finite representations

$$A_* = \sum_{i=1}^{k_b} p_i A_{\lambda_i}, \qquad C_* = \sum_{j=1}^{k_n} q_j C_{\nu_j}, \qquad p \in \Delta_{k_b-1},\ q \in \Delta_{k_n-1}. \tag{B.4}$$

This is one pair of spectral distributions, chosen before the output rule and the innovations. Selecting a different worst curvature after each realized output would not establish (B.3) or the following bound.

B.2 Replicating coordinates without changing either budget

For an integer $r \geq 1$, construct a diagonal quadratic with $k_b + rk_n$ coordinates. Bias coordinate i has curvature λ_i, initial value $D\sqrt{p_i}$, and no noise. The r coordinates in noise block j have curvature ν_j, zero initialization, and independent Gaussian innovations of variance $\sigma_*^2 q_j/r$ each. Innovations are also independent across updates and blocks. This gives exactly the initialization budget D^2 and covariance trace σ_*^2; the sample Hessian is the same fixed positive-semidefinite matrix on every draw and is at most LI. Thus each construction is an admissible finite-dimensional member Q_r of $\mathcal{Q}$.

Let $Z_{j,r}$ be its r-by-T standard Gaussian matrix of innovations in block j. On every realized path and for every scalar convex weight vector, the actual quadratic loss is

$$F_{Q_r}(\widehat{x}_w) - F_{Q_r}^* = \frac{D^2}{2}\sum_i p_i\lambda_i a_{\lambda_i}(w)^2 + \frac{\sigma_*^2}{2}\sum_j q_j\nu_j \frac{\|Z_{j,r}b_{\nu_j}(w)\|^2}{r}. \tag{B.5}$$

This is a pathwise identity for the constructed instance; it does not substitute random weights into the fixed-weight expectation (3). The same scalar weights act in all coordinates.

Fix $0 < \varepsilon, \delta < 1$. For sufficiently large finite r,

$$\Pr\left\{r^{-1}Z_{j,r}^\top Z_{j,r} \succeq (1-\varepsilon)I_T \text{ for all } j\right\} \geq 1 - \delta. \tag{B.6}$$

An elementary finite bound suffices. The diagonal entries of $r^{-1}Z^\top Z - I_T$ have second moment $2/r$; off-diagonal entries have second moment $1/r$. Thus

$$\mathbb{E}\|r^{-1}Z^\top Z - I_T\|_F^2 = \frac{T(T+1)}{r}.$$

The operator norm is at most the Frobenius norm. Markov's inequality and a union bound establish (B.6) whenever $r \geq k_n T(T+1)/(\delta\varepsilon^2)$. This explicit sufficient dimension bound is conservative; it is not a claim about the replication needed in an experiment. Only a finite r enters each instance, without increasing either problem budget.

On this event, (B.5) is at least $(1-\varepsilon)G(w; A_*, C_*) \geq (1-\varepsilon)\mathcal{V}$ simultaneously for every $w \in \Delta_T$. Hence the inequality holds for weights chosen after observing the entire path, the objective, or auxiliary randomness. Nonnegativity of risk on the complementary event yields

$$\mathbb{E}_{Q_r} \min_{w\in\Delta_T} [F_{Q_r}(\widehat{x}_w) - F_{Q_r}^*] \geq (1-\delta)(1-\varepsilon)\mathcal{V}(\gamma). \tag{B.7}$$

The pathwise minimum is measurable: it is the minimum of a quadratic function on a compact simplex, or equivalently the infimum over its countable dense rational subset. It is integrable, being nonnegative and at most the loss of the deterministic initial point.

B.3 Equality of the minimax values

Every output rule has risk at least the pathwise minimum. Since (B.7) uses a common instance for all rules,

$$\inf_W \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q \ell_Q(W) \geq \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q \min_{w\in\Delta_T} \ell_Q(w) \geq \mathcal{V}(\gamma), \tag{B.8}$$

where the last step takes $\varepsilon, \delta \downarrow 0$. For the reverse bound, choose a deterministic minimizer of (3). It exists by continuity on the compact simplex and has worst-case risk $\mathcal{V}$. It belongs to the adaptive output class as well. For every fixed task its risk also upper-bounds the expected pathwise minimum. This proves both equalities in (B.8) and the theorem.

If a budget is zero, omit its coordinates; if $\mathcal{V} = 0$, the upper bound and nonnegativity prove the statement directly. The proof does not assert a single finite-dimensional maximizer attaining the supremum or a dimension-uniform onset of replication. It also does not assert that adaptive outputs fail to help on a fixed low-dimensional task.

The noiseless endpoint is a separate claim. Without positive contraction, the noiseless class value is still $\min_{w\in\Delta_T}(D^2/2)\max_{\lambda\in[0,L]}\lambda a_\lambda(w)^2$, but need not be attained by the endpoint. On the noiseless fixed instance $H = LI$, one step of size $2/L$ gives $x_1 = -x_0$, so $(x_0 + x_1)/2 = 0$ while the endpoint has positive loss for $x_0 \neq 0$. This is an instance-level illustration, not a zero-risk statement about the entire class. The endpoint formula (9), inverse-root rate bounds and monotone curvature envelopes use their stated positive-contraction assumptions. Prescribed sequences with locally amplifying steps, such as those studied by Agarwal et al. (2021), are therefore within the information equalities' scope; those equalities alone imply no stochastic acceleration guarantee.

Remaining class-risk improvement. For $\mathcal{V} > 0$, define

$$\mathrm{Gain}_{\mathcal{Q}}(\gamma) = \frac{\mathfrak{R}(\gamma, e_T)}{\mathcal{V}(\gamma)} \geq 1.$$

Here e_T places all averaging weight on the endpoint, so the numerator is its exact class risk by (3); the denominator is now the best risk even with full-history scalar convex outputs. The ratio is an improvement factor between two class suprema, not the largest achievable improvement on every individual task. Evaluating the denominator is a convex variational problem.

B.4 A QUANTITATIVE SUFFICIENT DIMENSION

Fix $\varepsilon, \delta \in (0, 1)$. The elementary replication bound above can be sharpened using standard Gaussian singular-value concentration (Vershynin, 2012, Corollary 5.35). For an r-by-T standard Gaussian matrix and $r \geq T$, $s_{\min}(Z) \geq \sqrt{r} - \sqrt{T} - u$ except on an event of probability at most $2e^{-u^2/2}$. For $k_n \geq 1$, a union bound with $u = \sqrt{2\log(2k_n/\delta)}$ gives (B.6) whenever

$$r \geq \left\lceil \left(\frac{\sqrt{T} + \sqrt{2\log(2k_n/\delta)}}{1 - \sqrt{1-\varepsilon}} \right)^2 \right\rceil. \tag{B.9}$$

Let $V_d(\gamma)$ be the adaptive minimax value with dimension at most d, and take k_b, k_n from an optimal common spectral pair in (B.4). Embedding its replicated construction by zero coordinates gives

$$(1-\delta)(1-\varepsilon)\mathcal{V}(\gamma) \leq V_d(\gamma) \leq \mathcal{V}(\gamma), \qquad d \geq k_b + k_n r. \tag{B.10}$$

The upper bound uses the same deterministic output as in (4). If the noise term is absent, no replication is needed. This is a sufficient dimension for approximation; it is not a necessary dimension or an exact fixed-dimensional characterization. In particular, numerical spectral supports need not be supports of the continuous problem's exact optimizer. Applying the bound to those supports only lower-bounds their own common-instance value. Gaussian concentration is an established tool here, not a separate novelty claim.

The cost comes from controlling every coefficient direction at once. For an r-by-N matrix, $r < N$ implies $\mathrm{rank}(G^\top G) \leq r < N$, so a positive multiple of I_N cannot be a lower bound. For sampling feedback, $N = qT$ individual samples replace the T innovations here. This rank obstruction concerns the uniform Gram event, not a necessary dimension for the minimax equality: the reachable coefficient set may be smaller than $\mathbb{R}^N$. Corollary E.1 preserves the smaller additional risk at sufficient finite dimensions; it does not identify adaptive gains at an arbitrary fixed dimension.

B.5 A COMMON ADVERSARY AND THE SHARED-WEIGHT RESTRICTION

Why one instance constrains all choices. Compact minimax first gives a common spectral adversary: its bias and noise moments evaluate every fixed plan, hence every independent mixture.

Represent those moments by finitely many atoms. Place the bias in noiseless coordinates and replicate each noise atom into r coordinates, dividing its variance by r; both noise budgets stay fixed.

For sampling feedback, expose the $N = qT$ individual samples through standard Gaussian matrices $G_i \in \mathbb{R}^{r\times N}$, one per noise atom. For sufficiently large finite r, with probability at least $1-\delta$, simultaneously for all atoms,

$$G_i^\top G_i/r \succeq (1-\varepsilon)I_N, \qquad \|z_i(w,B)\|^2 = \sum_{j<T} b_{\lambda_i,j}(w)^2/B_j.$$

Here z_i assigns coefficient $b_{\lambda_i,j}(w)/B_j$ to each individual sample in batch j. The Gram event controls every coefficient vector, including one chosen from the observations. It therefore bounds all outputs and legal sample partitions at once, including those generated by predictable sampling rules. No independence assumption is made after conditioning on an adaptive batch allocation. On that event, pathwise loss is at least $(1-\varepsilon)\min_{w\in\Delta_T,\, B\in\mathcal{B}_C\cap\mathbb{Z}^T} g_A(w,B)$, where g_A is the common adversary's fixed-plan risk. Nonnegativity off the event and the mixed-plan minimax identity give the feedback lower bound as $\varepsilon,\delta\downarrow 0$. Fixed batches give the output equality by the same argument.

This limit uses increasing dimension, with no change to the budgets. Appendix B.4 already quantifies a sufficient dimension for approximating the output value. Corollary E.1 separately preserves the smaller additional risk without losing a fraction of the noiseless baseline. Finite polynomial transfers and squared-form convexity allow arbitrary positive rates in both equalities. The rate laws and monotone numerical envelopes retain their stated stability assumptions.

Why sharing weights matters. In the inverse-root construction of Appendix E.1, initialization lies in one noiseless bias coordinate of maximizing curvature λ_*, while all noise coordinates start at zero. Thus $x_0 = (D, 0)$ and $x_T = (Dp_{0,T}(\lambda_*), z_T)$. A rule told this block structure could choose the endpoint for the bias block and the initial state for the noise block, returning

$$\widehat{x}_{\mathrm{block}} = (x_T^{\mathrm{bias}}, x_0^{\mathrm{noise}}) = (Dp_{0,T}(\lambda_*), 0),$$
$$F(\widehat{x}_{\mathrm{block}}) = \tfrac{1}{2}D^2\lambda_* p_{0,T}(\lambda_*)^2 = V_0,$$

the exact noiseless value. The bias block favors later contraction; the noise block favors its zero initialization. Shared scalar weights force both to use the same time combination, even with full history. This observation concerns this particular hard instance. It neither identifies the minimax value of the larger coordinatewise output class nor supplies a method for finding such blocks on an unknown task.

B.6 Exact one-step dimension and output-set values

Proposition B.1 (An exact one-step boundary)**.** *In $\mathcal{Q}(1,1,1)$, let $T = 1, \gamma_0 = 1$, and restrict instances to dimension at most d. Fixed scalar convex outputs, full-history adaptive scalar outputs, and the path convex-hull oracle have the same minimax value*

$$V_1^{\mathrm{scalar}} = 1/4, \qquad V_d^{\mathrm{scalar}} = 2^{2/3}/6 \quad (d \geq 2). \tag{B.11}$$

If the output may use H and return $(I-W)x_0 + Wx_1$, with $0 \preceq W \preceq I$ and $WH = HW$, the minimax value is $V_d^{\mathrm{spectral}} = 1/4$ for every $d \geq 1$.

Two directions already force the unrestricted scalar value, even when the rule knows the objective. Spectral weights strictly improve the value for $d \geq 2$. This compares output sets with known H, not unknown-curvature selection. Appendix B.6 proves the result; the general-horizon equality still uses arbitrary finite dimensions.

Enlarging the model or policy class. Write $V_{\mathcal{S}}(\mathcal{Q}) = \inf_{\pi\in\mathcal{S}} \sup_{Q\in\mathcal{Q}} R_Q(\pi)$. For the same loss and policies, $\mathcal{Q} \subseteq \mathcal{Q}'$ implies $V_{\mathcal{S}}(\mathcal{Q}') \geq V_{\mathcal{S}}(\mathcal{Q})$. The shared-scalar lower bounds remain valid. Upper bounds and equalities need not. For fixed $\mathcal{Q}$, $\mathcal{S} \subseteq \mathcal{S}'$ reverses the inequality. Existing methods retain their upper bounds, but old lower bounds need not cover the added actions. Enlarging both sets gives no general one-sided comparison.

We prove Proposition B.1 for the variance-only class $\mathcal{Q}(1,1,1)$. The dimension is at most d, the single update is $x_1 = (I-H)x_0 - \xi$, and x_0 is independent of the centered noise. For scalar outputs,

write $w \in [0,1]$ for the weight on x_1. The path oracle minimizes the true quadratic over the realized segment. Its worst expected risk is a lower bound on the adaptive value, which is a lower bound on the fixed-weight value.

One dimension. For the fixed midpoint, independence and the two second-moment bounds give

$$\mathbb{E}F\big((x_0+x_1)/2\big) \le \max_{0\le\lambda\le 1} \frac{\lambda}{2}\left[(1-\lambda/2)^2 + 1/4\right] = \frac{1}{4}.$$

The derivative of the expression is $(3\lambda-5)(\lambda-1)/8 \ge 0$ on $[0,1]$. For a matching lower bound, take $H=1, x_0=1$ and $\xi \in \{-1,1\}$ with equal probabilities. Then $x_1=-\xi$. On the event $\xi=-1$, both states equal 1 and every point in their convex hull has loss $1/2$. This event has probability $1/2$; loss is nonnegative otherwise. Even the path oracle therefore has risk at least $1/4$.

Two or more dimensions. Set $h=2^{-1/3}$, $H=\mathrm{diag}(h,1)$, $x_0=(1,0)$, and $\xi=(0,Z)$, with Z a symmetric Rademacher variable. This is a legal instance. On either noise outcome,

$$F((1-w)x_0+wx_1) = \frac{h}{2}(1-hw)^2 + \frac{w^2}{2}.$$

Its minimum on $[0,1]$ is attained at $w_* = h^2/(1+h^3) = 2^{1/3}/3$, with value $h/[2(1+h^3)] = 2^{2/3}/6$. The lower bound applies to every pathwise choice of w, including one with access to H, x_0, ξ. In any finite dimension, the fixed-weight risk formula gives

$$\sup_{Q\in\mathcal{Q}(1,1,1)} \mathbb{E}F((1-w_*)x_0+w_*x_1) = \frac{1}{2}\max_{0\le\lambda\le 1}\lambda(1-w_*\lambda)^2 + \frac{w_*^2}{2} = \frac{2}{27w_*} + \frac{w_*^2}{2} = \frac{2^{2/3}}{6}.$$

The bias maximum is at $\lambda = 1/(3w_*) = h$. Padding the two-dimensional witness with zero coordinates proves the claim for every $d \ge 2$. These one-step witnesses use Rademacher noise, which is allowed by the original second-moment assumptions; they do not require the Gaussian replication used in the general proof.

Spectral-coordinate outputs with known H. The enlarged output class may choose a measurable matrix W from the history and H, subject to $0 \preceq W \preceq I$ and $WH = HW$. A deterministic legal choice is $W=(I+H)^{-1}$, which returns $\widehat{x} = (I+H)^{-1}(x_0-\xi)$. Since $H(I+H)^{-2} \preceq I/4$, independence yields

$$\mathbb{E}F(\widehat{x}) \le \frac{1}{8}\big(\mathbb{E}\|x_0\|^2 + \mathrm{tr}\,\Sigma\big) \le \frac{1}{4}.$$

The one-dimensional Rademacher witness, embedded if necessary, still has identical states on an event of probability $1/2$. Every matrix output then equals that state, so the lower bound is $1/4$. Thus $V_d^{\mathrm{spectral}} = 1/4$ in every dimension, strictly below V_d^{scalar} for $d \ge 2$. Giving the scalar oracle the same H information leaves its two-dimensional lower bound intact. The separation is therefore due to the output set, not merely access to the Hessian. No claim for general T or unknown H follows.

C Sampling allocation, randomization and feedback

C.1 The full adaptive minimax identity with minibatches

Let $\mathcal{Q}_{\mathrm{ind}} = \mathcal{Q}(L, D, \sigma_{\mathrm{ind}})$ describe the individual-sample quadratic loss distribution. For fixed B, define

$$\mathfrak{R}_B(\gamma, w) = \frac{D^2}{2}\max_{0\le\lambda\le L}\lambda a_\lambda(w)^2 + \frac{\sigma_{\mathrm{ind}}^2}{2}\max_{0\le\lambda\le L}\lambda\sum_{j<T}\frac{b_{\lambda,j}(w)^2}{B_j}, \qquad \mathcal{V}_B = \min_{w\in\Delta_T}\mathfrak{R}_B(\gamma,w). \tag{C.1}$$

Proposition C.1 (Joint output and sample-allocation value)**.** *For any prescribed $\gamma \in (0,\infty)^T$ and fixed integer allocation B, in arbitrary finite dimension,*

$$\inf_W \sup_{Q\in\mathcal{Q}_{\mathrm{ind}}} \mathbb{E}_Q \ell_Q(W;B) = \sup_{Q\in\mathcal{Q}_{\mathrm{ind}}} \mathbb{E}_Q \min_w \ell_Q(w;B) = \mathcal{V}_B. \tag{C.2}$$

For integer $N \geq T$, let $\mathcal{B}_N = \{B \in \mathbb{N}^T : B_t \geq 1, \sum B_t = N\}$. Then

$$\inf_{B \in \mathcal{B}_N} \inf_{W} \sup_{Q} \mathbb{E}_Q \ell_Q(W; B) = \min_{B \in \mathcal{B}_N, w \in \Delta_T} \mathfrak{R}_B(\gamma, w). \tag{C.3}$$

Its continuous relaxation over $B \in [1, \infty)^T$, $\sum B = N$, is jointly convex in (w, B). If its value is V_{rel} and the integer value is V_{int}, then $V_{\mathrm{rel}} \leq V_{\mathrm{int}} \leq 2V_{\mathrm{rel}}$.

Proof. The average innovation at time j has covariance Σ/B_j. The fixed-weight diagonalization proving (3) therefore gives (C.1). For the adaptive lower bound in Appendix B, replace each noise transfer map by $\mathrm{diag}(B_j^{-1/2})\mathbf{B}_\lambda$. The two continuous Gram curves remain positive semidefinite, so the same compact minimax argument yields one common finite spectral pair. Implement each noise atom using IID individual Gaussian innovations with variance $\sigma^2_{\mathrm{ind}} q_i / r$ per replicated coordinate. Averaging B_j of them gives the exact pathwise noise term

$$\frac{\sigma^2_{\mathrm{ind}}}{2} \sum_i q_i \nu_i \frac{\| Z_{i,r} \, \mathrm{diag}(B_j^{-1/2}) b_{\nu_i}(w) \|^2}{r}.$$

The individual-sample covariance trace remains σ^2_{ind}. The same Gram event controls all weights simultaneously and proves (C.2). The finite outer allocation set and its pre-instance quantifier give (C.3); no exchange with an adaptive sampling rule is made.

For a fixed curvature, $b_{\lambda,j}(w)^2/B_j$ is the perspective of a square of a linear function. Its joint convexity follows directly from the convex quadratic-over-linear function, for either sign of $b_{\lambda,j}$. Sums and curvature maxima preserve convexity, as does the bias term. The feasible sets are compact, with $1 \leq B_t \leq N - T + 1$, so a relaxed minimizer (w^*, B^*) exists. Round each B_t^* down, then distribute the remaining integer samples. The rounded allocation satisfies $\widetilde{B}_t \geq \lfloor B_t^* \rfloor \geq B_t^*/2$. Keeping w^* multiplies every noise term by at most two and leaves bias unchanged; nonnegativity gives $\mathfrak{R}_{\widetilde{B}}(\gamma, w^*) \leq 2V_{\mathrm{rel}}$. The converse inequality is relaxation. At $N = T$, only $B = \mathbf{1}$ is feasible and the equality is exact. □

This result identifies a class optimum given budgets and fixed rates. It does not solve unknown-task output selection. The relaxation is semi-infinite in curvature, not an asserted $O(T)$ exact algorithm; on a finite curvature grid it has a second-order-cone representation. The integer problem itself is not convex. Deterministic allocation is essential to the stated scope; history-adaptive batch decisions are not covered by this identity.

C.2 An exact variational value with both covariance constraints

The following identities apply to any prescribed $\gamma \in (0, \infty)^T$; they require neither inverse-root rates nor stability. For fixed B and finite $\bar{\Delta}$, use the exact objective (6) stated in the main text. Its noise term is $\mathcal{N}_{\Delta,B}(w)$. The kernel condition and load constraint also imply $\Sigma \preceq 2\bar{\Delta}H$: write $G = H^{\dagger/2} \Sigma H^{\dagger/2} \succeq 0$, so $\mathrm{tr}\, G \leq 2\bar{\Delta}$ and $G \preceq 2\bar{\Delta} I$. The converse is not dimension-free: $\Sigma = \tau^2 H$ has $\Delta_\star = \tau^2 \, \mathrm{rank}(H)/2$. The mass $d\nu$ corresponds to covariance mass $d\mu = 2\lambda \, d\nu$; an atom at zero contributes neither covariance nor loss and can be removed. The exact risk (6) is not obtained by taking the smaller of two separate curvature suprema. Both constraints act on the same noise distribution.

Proposition C.2 (Exact joint output value)**.** *For any prescribed $\gamma \in (0, \infty)^T$ and integer allocation B,*

$$\inf_{W} \sup_{Q \in \mathcal{Q}_\Delta} \mathbb{E} \ell_Q(W; B) = \sup_{Q \in \mathcal{Q}_\Delta} \mathbb{E} \min_{w} \ell_Q(w; B) = \min_{w \in \Delta_T} \mathfrak{R}_{\Delta,B}(w). \tag{C.4}$$

Minimizing the last expression over prescribed capped integer allocations gives (7)*. Its continuous relaxation $1 \leq B_j \leq C$, $\sum_j B_j = qT$ is jointly convex in (w, B), and its value is between one half and one times the integer value. For $\bar{\Delta} = \infty$, replace the noise term by that in* (C.1)*. This exact identity concerns prescribed allocations. Theorem 2 identifies the predictable value with independently randomized plans, without equating it to the integer or continuous deterministic program.*

Proof. Diagonalization gives (6) for fixed weights; initialization and noise extrema can be realized in orthogonal blocks. For finite $\bar{\Delta}$, the feasible nonnegative measures form a weakly compact convex set on the compact interval. Their matrix moments $\int \lambda^2 \mathbf{B}_\lambda^\top \operatorname{diag}(B_j^{-1}) \mathbf{B}_\lambda \, d\nu$, together with mass and first moment, have finite representations by Carathéodory's theorem. Compact convex minimax therefore produces one common finite bias/noise pair exactly as in (B.3). For an atom of mass ν_i at $\lambda_i > 0$, put individual variance $2\lambda_i \nu_i / r$ in each of r replicated coordinates. Both covariance budgets are unchanged by replication. Batch averaging introduces $B_j^{-1/2}$ into the transfer map; the common Gaussian Gram event then proves (C.4) by the argument (B.5)–(B.8). Zero noise budgets are handled by omitting noise blocks; the unrestricted case is (C.2). Excluding kernel noise changes no objective-risk value in that unrestricted case.

Each integrand is a nonnegative sum of quadratic-over-linear perspectives in (w, B), regardless of the signs of the transfers. Taking its feasible measure supremum preserves convexity. To round a relaxed optimum, first floor each batch and distribute the remaining samples among entries below the integer cap. There is enough capacity, and every resulting $\widetilde{B}_j \geq B_j/2$. Thus the same weights have at most twice every feasible noise integral, with unchanged nonnegative bias, proving the factor-two bracket. □

For positive finite noise budgets, the noise moment problem also has the semi-infinite dual

$$\mathcal{N}_{\Delta,B}(w) = \inf_{\alpha,\beta \geq 0} \left\{ \sigma_{\mathrm{ind}}^2 \alpha + \bar{\Delta}\beta : 2\lambda\alpha + \beta \geq \lambda^2 \sum_j b_{\lambda,j}(w)^2 / B_j \quad (0 \leq \lambda \leq L) \right\}. \tag{C.5}$$

This is ordinary finite-moment linear-programming duality: compact moment sets give no gap (strict feasibility of the budget inequalities holds at the zero measure). At zero budgets use (6) directly. Finite curvature grids yield perspective SOCP relaxations. A grid value alone is never a continuous-curvature upper bound.

C.3 Sampling randomization, feedback, and finite budgets

This section proves Theorem 2 by equating the minimax values of predictable policies and randomized plans, then comparing this value with the joint convex relaxation. Write $\mathcal{B}_{\mathrm{rel}} = \{B \in [1, C]^T : \sum_j B_j = qT\}$, $\mathcal{K} = \Delta_T \times \mathcal{B}_{\mathrm{rel}}$, and $V_T^{\mathrm{rel}} = \min_{(w,B)\in\mathcal{K}} \mathfrak{R}_{\Delta,B}(w)$. For unrestricted inverse-curvature load, use the variance-measure expression in (C.1). Fix any $\gamma \in (0, \infty)^T$ throughout this proof. At finite T, all transfer features are continuous and bounded; their signs are unrestricted.

One adversary for weights and allocations jointly. Let p be a probability measure on $[0, L]$, and let ν obey the two constraints in (6). For $A = (p, \nu)$, define

$$g_A(w, B) = \frac{D^2}{2} \int \lambda a_\lambda(w)^2 \, dp(\lambda) + \int \lambda^2 \sum_{j<T} \frac{b_{\lambda,j}(w)^2}{B_j} \, d\nu(\lambda). \tag{C.6}$$

For $\bar{\Delta} = \infty$, replace the second term by $\frac{1}{2} \int \lambda \sum_j b_{\lambda,j}(w)^2 / B_j \, d\mu(\lambda)$, where $\mu([0, L]) \leq \sigma_{\mathrm{ind}}^2$. Each g_A is continuous and jointly convex in (w, B), by the quadratic-over-linear perspective for real-valued transfers, and affine in A. The strategy set $\mathcal{K}$ and the feasible measure sets are compact and convex (with the weak topology for measures). Compact convex minimax gives

$$V_T^{\mathrm{rel}} = \min_{\mathcal{K}} \max_{A} g_A = \max_{A} \min_{\mathcal{K}} g_A. \tag{C.7}$$

The maximizer can be chosen with finite support. To see this without fixing B in advance, write $a_\lambda(w) = \mathbf{a}_\lambda^\top w$ and $b_{\lambda,j}(w) = \mathbf{b}_{\lambda,j}^\top w$. Preserve the bias moment of $\lambda \mathbf{a}_\lambda \mathbf{a}_\lambda^\top$. For the noise measure, preserve mass, first moment, and *each* of the T matrix moments $\lambda^2 \mathbf{b}_{\lambda,j} \mathbf{b}_{\lambda,j}^\top$ separately. These are finitely many continuous features on a compact interval, so Carathéodory's theorem gives a finite representation preserving both budgets and (C.6) for every (w, B) at once. In the unrestricted case use features $\frac{1}{2}\lambda \mathbf{b}_{\lambda,j} \mathbf{b}_{\lambda,j}^\top$ and total mass. Zero-curvature noise atoms can be discarded.

An independent mixture of complete integer plans. Let $\mathcal{B}_{\rm int} = \mathcal{B}_{\rm rel} \cap \mathbb{Z}^T$ and $\mathcal{K}_{\rm int} = \Delta_T \times \mathcal{B}_{\rm int}$. A randomized plan is a Borel probability law π on this compact set. It is fixed before the instance is selected. Its independent seed, drawn after the instance is fixed, selects both w and B; neither can use oracle observations or the realized initialization. The adversary knows π, not its realized seed. In particular, the risk order is $\inf_\pi \sup_Q \mathbb{E}_{\pi,Q}\ell$, with the expectation over the seed *inside* the instance supremum. Predictable policies may use such independent auxiliary randomness.

For each fixed plan, its expected quadratic risk is (C.6) for the instance's bias and noise spectral measures. Conversely, the finite representation above realizes every feasible collection of these moments by a finite-dimensional additive Gaussian instance, with bias and noise in separate coordinates. This representation preserves the risk of *every* plan, hence of every mixture. Therefore

$$\begin{aligned} V_T^{\rm mix} &= \min_{\pi \in \mathcal{P}(\mathcal{K}_{\rm int})} \max_A \int g_A(w,B)\, d\pi(w,B) \\ &= \max_A \min_{(w,B)\in\mathcal{K}_{\rm int}} g_A(w,B). \end{aligned} \tag{C.8}$$

The space of probability laws on the compact metric plan set is compact and convex in the weak topology. The displayed payoff is continuous and affine in both the law and the adversary; compact convex minimax applies. Minimizing a linear functional over laws equals minimizing over individual plans. The integer set itself need not be convex. A maximizing adversary has finite support by the same simultaneous matrix-moment representation used above.

One sample event for every batch policy and output. Fix a finite feasible A, with positive-curvature noise atoms λ_i and individual variances $s_i = 2\lambda_i\nu_i$ (or $s_i = \mu_i$ in the unrestricted case). Put each bias atom in its own noiseless coordinate with initialization $D\sqrt{p_i}$. Replicate noise atom i into r zero-initialized coordinates of curvature λ_i and variance s_i/r. This defines one finite-dimensional additive Gaussian instance Q_r, preserving the initialization budget, covariance trace, and inverse-curvature trace for every r.

Draw all $N = qT$ individual samples in advance for analysis, with independent $r \times N$ standard Gaussian matrices G_i. For any $\varepsilon, \delta \in (0,1)$, the event

$$\mathcal{E} = \{G_i^\top G_i/r \succeq (1-\varepsilon) I_N \text{ for every noise block } i\} \tag{C.9}$$

has probability at least $1-\delta$ for sufficiently large finite r. Explicitly, $\mathbb{E}\|G_i^\top G_i/r - I_N\|_F^2 = N(N+1)/r$; Markov's inequality and a union bound over m noise atoms give $r \geq mN(N+1)/(\delta\varepsilon^2)$. When $m = 0$, take the sure event and omit the noise blocks.

Any realized legal policy partitions the samples into consecutive batches of sizes $B_0, \ldots, B_{T-1}$. In noise block i, each individual sample in batch j has output coefficient $b_{\lambda_i,j}(w)/B_j$. Call this length-N coefficient vector $z_i(w,B)$; then

$$\|z_i(w,B)\|^2 = \sum_{j<T} \frac{b_{\lambda_i,j}(w)^2}{B_j}. \tag{C.10}$$

On $\mathcal{E}$, the noise loss of this block is at least $(1-\varepsilon)(\lambda_i s_i/2)\|z_i(w,B)\|^2$. This holds simultaneously for every partition and every weight vector, including weights and batches depending on observed samples: (C.9) controls all real coefficient vectors, including sign-changing ones. The noiseless bias is exactly its term in (C.6). Hence, pathwise on the same event, for every predictable policy and full-history output,

$$\ell_{Q_r}(W;B) \geq (1-\varepsilon) g_A(w,B) \geq (1-\varepsilon) \min_{\mathcal{K}_{\rm int}} g_A. \tag{C.11}$$

Here w, B are the realized choices; no fixed-weight expectation formula is applied to random choices. Predictability ensures that the instance is sampled by the stated oracle. Uniformity of (C.9) is what permits the choices in (C.11).

Nonnegativity off $\mathcal{E}$ gives an expected lower bound $(1-\delta)(1-\varepsilon)\min_{\mathcal{K}_{\rm int}} g_A$ for every policy. The dimension-unrestricted supremum and then $\varepsilon, \delta \downarrow 0$, followed by (C.8), imply $V_T^{\rm pred} \geq V_T^{\rm mix}$. The event is uniform in the auxiliary seed as well as in the observations. A plan drawn before training is a legal predictable policy, so the reverse inequality holds. Also, (C.7) and $\mathcal{K}_{\rm int} \subseteq \mathcal{K}$ give $V_T^{\rm rel} \leq V_T^{\rm mix} = V_T^{\rm pred}$. Prescribed policies give $V_T^{\rm pred} \leq V_T^{\rm det}$, and the floor-and-fill argument following (C.4) gives $V_T^{\rm det} \leq 2V_T^{\rm rel}$. These limiting arguments do not assert that one fixed finite dimension attains the supremum, or that a continuous allocation is an executable batch policy.

Correlated rounding costs at most $9/8$. Take a minimizer (w^*, B^*) of the continuous problem. Write $B_j^* = m_j + \theta_j$, with integer $m_j \geq 1$ and $0 \leq \theta_j < 1$. The sum of the fractional parts is an integer. There is a random vector $Z \in \{0,1\}^T$ with $\mathbb{E}Z_j = \theta_j$ and $\sum_j Z_j = \sum_j \theta_j$ on every path. For completeness, repeatedly choose two fractional coordinates u, v. Increase the first and decrease the second by $a = \min\{1-u, v\}$ with probability $b/(a+b)$, or make the opposite change of size $b = \min\{u, 1-v\}$ otherwise. Each step preserves the sum and conditional expectations and makes at least one coordinate integral. A single fractional coordinate cannot remain because the sum is integral. Thus this procedure terminates and has the stated marginals.

Set $\widetilde{B}_j = m_j + Z_j$. Each batch lies between one and C, since a fractional B_j^* has ceiling at most the integer cap. The total remains exactly qT. Moreover,

$$B_j^* \mathbb{E}\frac{1}{\widetilde{B}_j} = (m_j + \theta_j)\left(\frac{1-\theta_j}{m_j} + \frac{\theta_j}{m_j+1}\right) = 1 + \frac{\theta_j(1-\theta_j)}{m_j(m_j+1)} \leq \frac{9}{8}. \tag{C.12}$$

Keep the deterministic weights w^*. Every noise term of (C.6) is a nonnegative square times a reciprocal batch, even for signed transfers. Thus, for every feasible adversary A,

$$\mathbb{E}g_A(w^*, \widetilde{B}) \leq \tfrac{9}{8} g_A(w^*, B^*). \tag{C.13}$$

The bias term is unchanged. Taking the supremum over A proves $V_T^{\mathrm{mix}} \leq (9/8)V_T^{\mathrm{rel}}$, completing (8). This is a randomized-plan guarantee, not a claim that feedback gains over deterministic plans are at most 12.5%. No tightness assertion is needed.

Reusing finite common-adversary certificates. The argument before (C.7) also shows $\min_{\mathcal{K}} g_A \leq V_T^{\mathrm{pred}}$ for *every* feasible finite A, without optimizing over adversaries. At a feasible $x = (w, B)$, joint convexity gives the computable bound

$$L_A := \max\{0, g_A(x) - \mathrm{gap}_A(x)\} \leq \min_{\mathcal{K}} g_A \leq V_T^{\mathrm{rel}} \leq V_T^{\mathrm{pred}},$$
$$\mathrm{gap}_A(x) = \max_{y \in \mathcal{K}} \langle \nabla g_A(x), x - y\rangle. \tag{C.14}$$

The linear minimization separates into a simplex vertex and a capped batch allocation. Thus the existing lower certificates require no new spectral optimization. A feasible integer pair with risk at most $U^\star$ gives $L^\star \leq V_T^{\mathrm{rel}} \leq V_T^{\mathrm{pred}} \leq V_T^{\mathrm{det}} \leq U^\star$. If a fixed rule has risk in $[R^-, R^+]$, the same interval $[R^-/U^\star, R^+/L^\star]$ encloses its ratio to either policy class's optimum, provided $L^\star > 0$. Numerical guards and floating-point verification of these certificates are described in Appendix G.6.1; (C.14) is the mathematical justification, not a claim of directed-rounding arithmetic.

C.4 Lower bounds for history-dependent outputs in a restricted set

The uniform event used in Appendix C.3 also applies when the realized averaging weights must lie in a fixed subset of the simplex. The learning rates remain prescribed throughout.

Proposition C.3 (A restricted-output lower bound)**.** *Fix $\gamma \in (0,\infty)^T$, the finite-load class $\mathcal{Q}_\Delta$, and a nonempty closed set $\mathcal{W} \subseteq \Delta_T$. Let $V_{\gamma,\mathcal{W}}^{\mathrm{pred}}$ be the minimax value of history-dependent integer batching and output selection with $w \in \mathcal{W}$ on every realization. Independent auxiliary randomization is allowed, and $1 \leq B_j \leq C$ and $\sum_j B_j = qT$ hold on every path. For any feasible finite spectral adversary A in* (C.6)*,*

$$V_{\gamma,\mathcal{W}}^{\mathrm{pred}} \geq \min_{\substack{w \in \mathcal{W},\ B \in [1,C]^T \\ \sum_j B_j = qT}} g_A(w, B). \tag{C.15}$$

The instance supremum ranges over all finite dimensions.

Proof. Denote the right-hand side by L_A. Replicate the noise atoms of A as in Appendix C.3, preserving the initialization and both noise budgets. For any $\varepsilon, \delta \in (0,1)$, a sufficiently large finite replication count gives one instance Q_r and the event in (C.9), of probability at least $1-\delta$. On that event, (C.10) and (C.11) give simultaneously

$$\ell_{Q_r}(w; B) \geq (1-\varepsilon) g_A(w, B) \geq (1-\varepsilon) L_A$$

for every $w \in \mathcal{W}$ and every legal integer grouping. The second inequality holds because integer batches are feasible real batches. The first controls all real coefficient vectors, so transfer signs do not require a stability assumption.

Substituting measurable choices from the observations and an independent seed preserves the pathwise inequality. The adversary Q_r is fixed before both are observed. Loss is nonnegative off the event, so every admissible policy has expected loss at least $(1-\delta)(1-\varepsilon)L_A$ on this same instance. Taking the instance supremum, then the policy infimum, and letting $\varepsilon, \delta \downarrow 0$ proves the result. □

A certified supporting-plane lower bound over $\mathcal{W} \times \mathcal{B}_C$ therefore bounds every history-dependent policy whose realized weights lie in $\mathcal{W}$. Appendix G.4.1 applies the proposition to nondecreasing post-update weights. Different prescribed rate sequences may use different adversaries. Taking their minimum covers deterministic rate selection before sampling; it does not cover randomized rate selection, which would require an adversary common to the rate distribution.

D Exact baseline and the additional-risk boundary

D.1 A half-tail upper bound at uniform batches

Write $M_t = \sum_{j<t} \gamma_j$ and $h_T = \sum_{k=1}^{T} 1/k$. For inverse-root rates with $c \le 1/(2L)$, take $n = \lfloor T/2 \rfloor$, $w_t = 1/(T-n+1)$ for $n \le t \le T$, and $B_j = q$. Then

$$R_T^{\mathrm{uniform,half}} \le \frac{D^2}{4eM_n} + \frac{4\bar{\Delta}h_T}{qT}. \tag{D.1}$$

Indeed, positivity gives $a_\lambda(w) \le e^{-\lambda M_n}$, so its bias is at most $D^2/(4eM_n)$. A geometric sum, using $\gamma_j \ge c/\sqrt{T}$, gives

$$0 \le b_{\lambda,j}(w) \le \gamma_j, \qquad \lambda b_{\lambda,j}(w) \le \frac{\gamma_j}{(T-n+1)c/\sqrt{T}} \le \frac{2\gamma_j}{c\sqrt{T}}.$$

Thus $\lambda^2 \sum_j b_{\lambda,j}^2/q \le 4h_T/(qT)$. Integrating against the inverse-curvature measure in (6) proves the noise bound. The argument extends to zero curvature by continuity. For positive fixed parameters this is $O(T^{-1/2})$; the exact noiseless lower bound established below gives the matching order. This is the half-tail estimate needed for the additional risk proof, without using any additional total-risk regime theorem.

D.2 Optimal additional risk above the exact noiseless class value

This uniform-batch slice supplies the fixed-cap corollary and the lookback estimates used in the joint theorem.

Lemma D.1 (Uniform-batch additional risk)**.** *Fix positive finite $L, D, \sigma_{\mathrm{ind}}, \bar{\Delta}, c$, with $cL \le 1/2$, and integer $q > 1$. At $C = q$, inverse-root steps have $E_T(q) = \Theta(T^{-3/4})$. The last $\lceil T^{3/4} \rceil$ states attain this order; endpoint and half-tail each have additional risk $\Theta(T^{-1/2})$.*

Constants and the starting horizon may depend on the fixed parameters. Write $V_0 = V_T^{\mathrm{det}}(q,0)$, $\mathrm{Bias}(w) = D^2 \max_{\lambda \le L} \lambda a_\lambda(w)^2/2$, and $\mathcal{N}(w) = \mathcal{N}_{\Delta,q}(w)$. Since $B_j \le q$ and $\sum_j B_j = qT$ on each path, $B_j = q$ identically, including for predictable policies. Proposition C.2 gives the exact reduction

$$E_T = \min_{w \in \Delta_T} \{\mathrm{Bias}(w) - V_0 + \mathcal{N}(w)\}. \tag{D.2}$$

This reduction, rather than a fixed-weight formula applied to random weights, supplies the adaptive-output and oracle lower bounds.

1. Exact baseline and the cost of looking back. All products are positive and decrease with the state index. Thus $a_\lambda(w) \ge p_{0,T}(\lambda)$, and the endpoint minimizes every noiseless curvature

simultaneously, proving (9). For $M_T > 1/(2L)$, its maximizing curvature λ_T is interior and unique: the derivative has the sign of

$$1 - 2\lambda \sum_{j<T} \frac{\gamma_j}{1-\lambda\gamma_j}, \qquad \frac{1}{4M_T} \le \lambda_T \le \frac{1}{2M_T}. \tag{D.3}$$

The displayed expression is strictly decreasing, positive at zero and negative at L. The bounds use $1/2 \le 1 - \lambda\gamma_j \le 1$. Also $\log(1-u) \ge -2u$ on $[0, 1/2]$, so $p_{0,T}(\lambda_T)^2 \ge e^{-2}$ and $V_0 \ge D^2/(8e^2 M_T) \ge D^2/(16e^2 c\sqrt{T})$. For every t,

$$\frac{p_{0,t}(\lambda_T)}{p_{0,T}(\lambda_T)} = \prod_{j=t}^{T-1} (1-\lambda_T\gamma_j)^{-1} \ge 1 + \lambda_T \sum_{j=t}^{T-1} \gamma_j \ge 1 + \lambda_T \frac{c}{\sqrt{T}}(T-t).$$

Consequently, with $\ell = \sum_t w_t(T-t)$,

$$\mathrm{Bias}(w) - V_0 \ge 2V_0\lambda_T \frac{c}{\sqrt{T}}\ell \ge a_b \frac{\ell}{T^{3/2}}, \qquad a_b = \frac{D^2}{64e^2c} > 0. \tag{D.4}$$

The maximizer in (D.3) is that of the exact terminal risk; replacing it by an arbitrary approximately worst curvature would leave an uncontrolled subtraction error.

2. Shared noise constrains every short lookback. Use a noise atom at the fixed positive curvature L, of inverse-curvature mass $\rho = \min\{\bar{\Delta}, \sigma_{\mathrm{ind}}^2/(2L)\} > 0$. It satisfies both noise budgets and gives $\mathcal{N}(w) \ge (\rho L^2/q)\sum_j b_{L,j}(w)^2$. Suppose $\ell \le T/8$. Set

$$m = \max\{1, \lceil 4\ell \rceil\}, \qquad r = \left\lceil \frac{\sqrt{T}}{Lc} \right\rceil, \qquad k = T - m - r.$$

For a sufficiently large horizon, $r \le T/4$ and $k \ge 0$. Markov's inequality gives $\sum_{t=T-m}^{T} w_t \ge 3/4$. For those states, the last r or more updates after k imply $p_{k,t}(L) \le e^{-1}$. The telescoping identity $L\sum_{j=k}^{t-1} \gamma_j p_{j+1,t}(L) = 1 - p_{k,t}(L)$ therefore yields

$$L\sum_{j=k}^{T-1} b_{L,j}(w) = \sum_{t>k} w_t[1 - p_{k,t}(L)] \ge \tfrac{3}{4}(1-e^{-1}).$$

Cauchy–Schwarz on these $m + r$ coefficients proves

$$\mathcal{N}(w) \ge \frac{a_n}{4\ell + d\sqrt{T}}, \qquad a_n = \frac{9\rho(1-e^{-1})^2}{16q}, \qquad d = 2 + \frac{1}{Lc}. \tag{D.5}$$

Here $m + r \le 4\ell + 2 + \sqrt{T}/(Lc) \le 4\ell + d\sqrt{T}$. The bias and noise bounds add in (D.2), since their spectral adversaries occupy independent orthogonal blocks of the same class.

If $\ell > T/8$, (D.4) already gives $E(w) \ge a_b/(8\sqrt{T})$, where $E(w) = \mathrm{Bias}(w) - V_0 + \mathcal{N}(w)$. Otherwise put $u = 4\ell + d\sqrt{T}$. Equations (D.4)–(D.5) and arithmetic–geometric mean give

$$E(w) \ge \frac{a_b u}{4T^{3/2}} + \frac{a_n}{u} - \frac{a_b d}{4T} \ge \sqrt{a_b a_n}\, T^{-3/4} - \frac{a_b d}{4T}. \tag{D.6}$$

For example, the last expression is at least $\frac{1}{2}\sqrt{a_b a_n} T^{-3/4}$ once $T^{1/4} \ge a_b d/(2\sqrt{a_b a_n})$. Together with the preceding case this proves a positive lower constant uniform in all weights, and hence the desired minimax lower bound.

3. A sublinear tail attains the matching upper order. Let $T \ge 4$, $1 \le K \le T/2$, and average states $n = T - K + 1, \dots, T$. Write $S = M_T - M_n$. For each curvature, positivity and $1 - \prod_j (1-u_j)^2 \le 2\sum_j u_j$ give

$$\frac{D^2\lambda}{2}[p_{0,n}(\lambda)^2 - p_{0,T}(\lambda)^2] \le D^2 S\lambda^2 e^{-2\lambda M_n} \le \frac{D^2 S}{e^2 M_n^2}.$$

Subtracting the exact maximum at time T, and using $S \le \sqrt{2}cK/\sqrt{T}$ and $M_n \ge c\sqrt{n}$, proves

$$0 \le \mathrm{Bias}(w^{(K)}) - V_0 \le \frac{2\sqrt{2}D^2}{e^2 c}\frac{K}{T^{3/2}}. \tag{D.7}$$

This is a bound on the difference of the two maxima, obtained from a uniform pointwise difference, not from their separate asymptotics.

Split noise indices at $j_0 = \lfloor T/4 \rfloor$. For $j < j_0$, the contraction before the first averaged state gives $b_{\lambda,j} \le \gamma_j e^{-\lambda c\sqrt{T}/4}$. Hence

$$\lambda^2 \sum_{j<j_0} b_{\lambda,j}^2 \le c^2 h_T \lambda^2 e^{-\lambda c\sqrt{T}/2} \le \frac{16 h_T}{e^2 T}. \tag{D.8}$$

For $j \ge j_0$, set $g_- = c/\sqrt{T}$ and $z = 1 - \lambda g_-$. Then $\gamma_j \le 2g_-$ and

$$\lambda b_{\lambda,j} \le 2\sum_{s\ge 0}(1-z)z^s w_{j+1+s}^{(K)}.$$

Extend weights by zero outside their support. The geometric coefficients sum to one for $\lambda > 0$; Jensen's inequality followed by summation over all integer j bounds the squared right side by $4\sum_t (w_t^{(K)})^2 = 4/K$. At zero curvature the claimed inequality holds directly. Thus, using the mass budget in (6),

$$\mathcal{N}(w^{(K)}) \le \frac{\bar{\Delta}}{q}\left(\frac{4}{K} + \frac{16 h_T}{e^2 T}\right). \tag{D.9}$$

Equations (D.7) and (D.9) are finite-horizon bounds, valid before balancing. With $K = \lceil T^{3/4} \rceil$, the support condition holds for all sufficiently large T; both leading terms have order $T^{-3/4}$, and $h_T/T = o(T^{-3/4})$. This proves the upper side of Lemma D.1. All constants depend only on the fixed parameters, while the attaining weights do not use the noise budget.

4. Why endpoint and half-tail have larger additional risk. For the endpoint $\ell = 0$, so (D.5) gives an $\Omega(T^{-1/2})$ noise cost, and its initialization risk above the baseline is zero. For late indices the same geometric estimate gives

$$\lambda^2 \sum_{j\ge j_0} b_{\lambda,j}(e_T)^2 \le \frac{4c^2\lambda^2/T}{1-(1-\lambda c/\sqrt{T})^2} \le \frac{4cL}{\sqrt{T}}.$$

Together with (D.8) this proves its $O(T^{-1/2})$ upper bound. The half-tail output on $\lfloor T/2 \rfloor, \ldots, T$ has $\ell = (T - \lfloor T/2 \rfloor)/2 \ge T/4$, so (D.4) gives its matching lower order; (D.1) gives the upper order. These comparisons concern each fixed output's class risk minus the same V_0. They do not identify its noise term with E_T.

The uniform-batch lemma is an asymptotic statement for positive fixed parameters. It does not cover $D = 0$, a zero noise budget, noise budgets growing with T, or a dimension fixed before taking the class supremum. The general value theorem supplies the unrestricted finite-dimensional reduction; this does not claim a two-dimensional version of the refined equality.

Corollary D.1 (Fixed caps change constants, not the additional risk exponent)**.** *For every integer $C \ge q$,*

$$\frac{q}{C}E_T(q) \le E_T^{\mathrm{pred}}(C) \le E_T^{\mathrm{det}}(C) \le E_T(q). \tag{D.10}$$

Under the fixed positive parameters of Theorem 3, every fixed cap therefore has additional risk $\Theta(T^{-3/4})$.

For the finite-budget inequality, $\mathrm{Bias}(w) - V_0 \ge 0$ and $B_j \le C$ give $\mathcal{N}_{\Delta,B}(w) \ge (q/C)\mathcal{N}_{\Delta,q}(w)$. Minimize and apply the sampling-class comparison (8). The uniform-batch slice and growing-cap proof are in Appendices D.2 and D.3.

D.3 Additional risk under sampling caps and a fixed sample budget

We prove Corollary D.1 and Theorem 3. Throughout, the positive finite $L, D, \sigma_{\mathrm{ind}}, \bar{\Delta}, c$ and integer $q > 1$ are fixed, with $cL \le 1/2$. Constants and the starting horizon below depend only on these parameters, never on C or T. Write $V_0 = D^2 \max_{\lambda \le L} \lambda p_{0,T}(\lambda)^2/2$, $E_T^{\mathrm{rel}} = V_T^{\mathrm{rel}} - V_0$, and

$$e(w, B) = \mathrm{Bias}(w) - V_0 + \mathcal{N}_{\Delta,B}(w), \qquad E_T^{\mathrm{rel}} = \min_{\Delta_T \times \mathcal{B}_{\mathrm{rel}}} e(w, B). \tag{D.11}$$

The second equality follows from (6) and the definition of the relaxation. The exact noiseless value is independent of all batch choices: the endpoint minimizes every noiseless curvature at once. In particular $\mathrm{Bias}(w) - V_0 \ge 0$. By Theorem 2,

$$E_T^{\mathrm{rel}} \le E_T^{\mathrm{pred}} = E_T^{\mathrm{mix}} \le E_T^{\mathrm{det}}. \tag{D.12}$$

We bound the left side from below and exhibit one integer deterministic plan for the right side. This transfers the result to feedback without substituting a fixed-output expectation for a random policy.

The rounding bounds also apply to the additional risk itself: $E_T^{\mathrm{mix}} \le (9/8)E_T^{\mathrm{rel}}$ and $E_T^{\mathrm{det}} \le 2E_T^{\mathrm{rel}}$. To see this, keep an optimal relaxed output fixed. Correlated rounding (C.12) multiplies each reciprocal-batch noise term by at most $9/8$; the deterministic rounding preceding (C.12) multiplies it by at most two. Initialization is unchanged, and its additional risk over V_0 is nonnegative. This proves the assertions directly; subtracting V_0 from a multiplicative bound on total risks alone would not prove them. For a particular feasible real pair, the sharper mixed-plan upper bound is $\mathrm{Bias}(w) - V_0 + (9/8)\mathcal{N}_{\Delta,B}(w)$.

1. The immediate fixed-cap comparison. For every real feasible B and every w, positivity gives

$$\mathcal{N}_{\Delta,B}(w) \ge \frac{q}{C}\mathcal{N}_{\Delta,q}(w), \qquad e(w, B) \ge \frac{q}{C}e(w, q).$$

The first inequality holds under the same two-constraint spectral maximum, term by term. The second uses $q/C \le 1$ and the nonnegative initialization risk above the baseline. Minimizing and using (D.12), while choosing the feasible uniform integer allocation for the upper bound, proves

$$\frac{q}{C}E_T(q) \le E_T^{\mathrm{rel}}(C) \le E_T^{\mathrm{pred}}(C) \le E_T^{\mathrm{det}}(C) \le E_T(q). \tag{D.13}$$

This finite-budget comparison also holds at zero load. For positive fixed load and fixed C, the uniform-batch result in Appendix D.2 proves the stated $\Theta(T^{-3/4})$ corollary. Its constants may depend on that fixed cap; (D.13) alone is not sharp for a growing cap.

Additional-risk and total-risk ratios. For any $C_T \to \infty$, the theorem and fixed-cap corollary give $E_T(q)/E_T^s(C_T) \to \infty$, because $C_T^{-1/2} + \log T/T^{1/4} \to 0$. In the saturated regime $C_T = \Omega(\sqrt{T}/(\log T)^2)$, the more precise ratio is

$$\frac{E_T(q)}{E_T^s(C_T)} \asymp \frac{T^{1/4}}{\log T} \longrightarrow \infty. \tag{D.14}$$

At the same time, for every admissible cap and each policy class,

$$0 \le \frac{V_T^{\mathrm{det}}(q)}{V_T^s(C)} - 1 = \frac{E_T(q) - E_T^s(C)}{V_0 + E_T^s(C)} = O(T^{-1/4}) \longrightarrow 0.$$

The growing additional risk separation therefore coexists with a vanishing relative total-risk advantage. These are differences of optimal class values, not an oracle's variance term or a taskwise speedup.

Lemma D.2 (Disjoint blocks share one sample budget)**.** *For any prescribed learning rates, fix a legal scalar output w and positive batches B with $\sum_j B_j = N$ and $B_j \le C$. Let I_k be disjoint nonempty update blocks, and let each $\rho_k\delta_{\lambda_k}$, $\lambda_k > 0$, separately satisfy the spectral noise constraints. Put $c_k(w) = \lambda_k \sum_{j \in I_k} |b_{\lambda_k,j}(w)|$. Then*

$$\mathcal{N}_{\Delta,B}(w) \ge \max\left\{ \frac{1}{N}\sum_k \rho_k c_k(w)^2, \max_k \frac{\rho_k c_k(w)^2}{C|I_k|} \right\}. \tag{D.15}$$

Proof. Write $n_k = \sum_{j \in I_k} B_j$. For each individually feasible atom, Cauchy–Schwarz gives

$$\mathcal{N}_{\Delta,B}(w) \geq \rho_k \lambda_k^2 \sum_{j \in I_k} \frac{b_{\lambda_k,j}(w)^2}{B_j} \geq \frac{\rho_k c_k(w)^2}{n_k}.$$

Thus $\rho_k c_k(w)^2 \leq \mathcal{N}_{\Delta,B}(w) n_k$. Sum over k and use $\sum_k n_k \leq N$ for the first bound; use $n_k \leq C|I_k|$ for the second. □

The atoms need not be jointly feasible as one mixture. Each is an alternative adversary against the same allocation. The bound explains why several response scales compete for one budget, even without nonnegative state transfers. It concerns fixed w, B; adaptive conclusions require a lower bound uniform over the relaxed plans, followed by the class-value comparison in (D.12). It is not a pathwise substitution of random averaging weights.

2. A cap still limits the dilution of recent noise. Let $\ell = \sum_t w_t(T-t)$. The exact-baseline argument (D.4) is independent of B, and gives

$$e(w,B) \geq a_b \ell / T^{3/2}, \qquad a_b = D^2/(64e^2 c) > 0. \tag{D.16}$$

Set $\rho = \min\{\bar{\Delta}, \sigma_{\mathrm{ind}}^2/(2L)\} > 0$, $a'_n = 9\rho(1-e^{-1})^2/16$, and $d = 2 + 1/(Lc)$. Replacing q by the bound $B_j \leq C$ in the fixed-curvature argument (D.5) yields, for $\ell \leq T/8$,

$$\mathcal{N}_{\Delta,B}(w) \geq \frac{a'_n}{C(4\ell + d\sqrt{T})}.$$

Thus arithmetic–geometric mean, as in (D.6), gives

$$e(w,B) \geq \frac{\sqrt{a_b a'_n}}{\sqrt{C}\, T^{3/4}} - \frac{a_b d}{4T}. \tag{D.17}$$

For $\ell > T/8$, (D.16) is at least $a_b/(8\sqrt{T})$. Since $C \geq q$, after increasing an onset independent of C this also dominates the right side of (D.17). Hence (D.17) holds for all outputs and all caps at sufficiently large horizons. Its negative T^{-1} remainder must be handled before concluding a uniform growing-cap lower bound.

3. Multiple noise time scales compete for the total budget. We prove a second lower bound

$$e(w,B) \geq a_f \frac{\log T}{T}, \qquad a_f > 0, \tag{D.18}$$

uniform over the same real allocations and outputs. If $\ell \geq \sqrt{T}\log T$, (D.16) gives (D.18) with constant a_b. Otherwise set

$$m = \max\{1, \lceil 4\ell \rceil\}, \qquad s_0 = \max\{2m, \lceil \sqrt{T}/(8cL) \rceil, 1\}.$$

There is a constant A_0, depending only on c, L, such that $s_0 \leq A_0\sqrt{T}\log T$ for $T \geq e$. Increase the starting horizon until $4A_0\sqrt{T}\log T \leq T^{3/4}$. For each integer $s = 2^k s_0 \leq T/4$, take the disjoint block of update indices

$$I_s = \{T-2s, \ldots, T-s-1\}, \qquad \lambda_s = \frac{\sqrt{T}}{8cs} \leq L, \qquad n_s = \sum_{j \in I_s} B_j.$$

There are $J = 1 + \lfloor \log_2(T/(4s_0)) \rfloor \geq (\log T)/(4\log 2)$ such blocks. Markov's inequality gives mass at least $3/4$ on states $t \geq T - m$. For $j \in I_s$ and any such state, $t > j$, all intervening updates have index at least $T/2$, and their number is at most $2s$. Therefore

$$\sum_{r=j+1}^{t-1} \gamma_r \leq \frac{4cs}{\sqrt{T}}, \qquad p_{j+1,t}(\lambda_s) \geq \exp\left(-2\lambda_s \sum_{r=j+1}^{t-1} \gamma_r\right) \geq e^{-1}.$$

Here $\log(1-u) \geq -2u$ is valid because $\lambda_s \leq L$ and $\lambda_s \gamma_j \leq 1/2$ at every update. Since $\gamma_j \geq c/\sqrt{T}$,

$$\lambda_s b_{\lambda_s,j}(w) \geq \frac{3}{32es}, \qquad \mathcal{N}_{\Delta,B}(w) \geq \rho \sum_{j \in I_s} \frac{\lambda_s^2 b_{\lambda_s,j}(w)^2}{B_j} \geq \frac{9\rho}{1024e^2 n_s}. \tag{D.19}$$

The last step is weighted Cauchy–Schwarz on the s entries. For each scale separately, a single noise atom of inverse-curvature mass ρ is feasible: its trace is $2\lambda_s\rho \le \sigma_{\mathrm{ind}}^2$, and its load is at most $\bar{\Delta}$. We take the maximum over these alternative adversaries, not their sum; the noise budget is never charged to J atoms of mass ρ simultaneously.

Because the blocks are disjoint, $\sum_s n_s \le qT$. Some block has $n_s \le qT/J$, proving (D.18), for example with

$$a_f = \min\left\{a_b, \frac{9\rho}{4096e^2 q \log 2}\right\} > 0.$$

Only the relaxed deterministic value is used at this stage. The common-adversary minimax argument underlying (D.12) then transfers this bound to arbitrary predictable sampling and adaptive outputs. In particular, choosing a difficult scale after inspecting a realized feedback path is not a step of this proof.

Combining (D.17) and (D.18), each with weight $1/2$, and increasing the onset so that $a_b d/4 \le (a_f/2)\log T$, gives

$$e(w,B) \ge \frac{\sqrt{a_b a_n'}}{2\sqrt{C}\,T^{3/4}} + \frac{a_f \log T}{4T}. \tag{D.20}$$

Minimizing proves the uniform lower side of the theorem for every value in (D.12). The logarithm counts response time scales competing for the same qT samples.

4. An integer allocation for a length-K tail. Fix $1 \le K \le T/2$, put $h_T = \sum_{t=1}^{T} 1/t$, and set

$$B_j^0 = 1 + \left\lfloor \min\left\{C-1, \frac{(q-1)T}{h_T(T-j+K)}\right\}\right\rfloor. \tag{D.21}$$

Since $\sum_{j<T}(T-j+K)^{-1} \le h_T$, its total is at most qT. Fill remaining samples from the last update backwards, without exceeding C. The capacity $CT \ge qT$ makes completion possible. The resulting integer allocation satisfies

$$\frac{1}{B_j} \le \frac{1}{C} + \frac{h_T(T-j+K)}{(q-1)T}, \qquad B_j \ge 1. \tag{D.22}$$

Indeed $1 + \lfloor \min\{C-1, u\}\rfloor \ge \min\{C, u\}$, and completion only increases batches. In fact, every feasible integer completion with $B_j \ge B_j^0$ preserves this inequality. Filling latest updates first is one choice, not a condition of the order bound; the proof does not identify the completion with the best finite constants.

Construction. Input T, q, C. Set $K = \max\{1, \lceil T^{3/4}/\sqrt{C}\rceil\}$, $h_T = \sum_{t=1}^{T} 1/t$, and

$$B_j = 1 + \left\lfloor \min\left\{C-1, \frac{(q-1)T}{h_T(T-j+K)}\right\}\right\rfloor.$$

Let $R = qT - \sum_j B_j$. For $j = T-1, T-2, \ldots, 0$, add $\min\{R, C-B_j\}$ to B_j and subtract it from R. Run the prescribed updates with these batches and return the uniform average of $x_{T-K+1}, \ldots, x_T$.

Algorithm 1: Explicit integer construction. The total is exactly qT and every batch respects C. Completion preserves the batch lower bounds needed for the order proof; it does not guarantee good finite-budget constants.

Proof of Proposition 1. Average states $n = T-K+1, \ldots, T$. Equation (D.7) gives

$$\mathrm{Bias}(w^{(K)}) - V_0 \le \frac{2\sqrt{2}D^2}{e^2 c}\frac{K}{T^{3/2}}.$$

For noise, split at $j_0 = \lfloor T/4\rfloor$. On early indices use $B_j \ge 1$ and (D.8) directly:

$$\lambda^2 \sum_{j<j_0} \frac{b_{\lambda,j}^2}{B_j} \le \frac{16 h_T}{e^2 T}. \tag{D.23}$$

Using the second term of (D.22) on these early indices instead would introduce an unnecessary second harmonic factor.

For late indices put $x = \lambda c/\sqrt{T}$, $z = 1 - x$. When $\lambda > 0$, $0 \le z < 1$; as in (D.9), $\lambda b_{\lambda,j} \le 2\sum_{r\ge 0}(1-z)z^r w^{(K)}_{j+1+r}$. Consequently

$$\lambda^2 \sum_{j\ge j_0} b^2_{\lambda,j} \le 4/K. \tag{D.24}$$

We additionally need the remaining-time-weighted bound

$$\lambda^2 \sum_{j\ge j_0} (T - j + K) b^2_{\lambda,j} \le 20. \tag{D.25}$$

For $j \ge n-1 = T-K$, there are K indices, $\lambda b_{\lambda,j} \le 2/K$, and $T - j + K \le 2K$; their contribution is at most 8. For $j < n-1$, let $d' = n - 1 - j \ge 1$. The same geometric sum gives

$$\lambda b_{\lambda,j} \le \frac{2}{K} z^{d'}(1 - z^K), \qquad T - j + K = 2K + d'.$$

Extending the sum over d' to all positive integers bounds this contribution by

$$\frac{4(1-z^K)^2}{K^2}\left(\frac{2K}{1-z^2} + \frac{1}{(1-z^2)^2}\right) \le 4\min\{1, Kx\}^2 \left(\frac{2}{Kx} + \frac{1}{(Kx)^2}\right) \le 12.$$

The final inequality follows by considering $Kx \le 1$ and $Kx \ge 1$. At $\lambda = 0$, (D.24)–(D.25) hold directly. Combining (D.22) and (D.23)–(D.25), then applying the total inverse-curvature mass constraint, proves the finite upper bound

$$\mathfrak{R}_{\Delta,B}(w^{(K)}) - V_0 \le \frac{2\sqrt{2}D^2}{e^2 c}\frac{K}{T^{3/2}} + \bar{\Delta}\left[\frac{4}{CK} + \left(\frac{16}{e^2} + \frac{20}{q-1}\right)\frac{h_T}{T}\right]. \tag{D.26}$$

Both covariance constraints still define the risk; using just the mass constraint in an upper bound is valid.

5. Balancing the window, and the saturation threshold. For every $C \ge q$, take $K = \max\{1, \lceil T^{3/4}/\sqrt{C}\rceil\}$. For sufficiently large T, uniformly over all caps, $K \le T/2$. Moreover

$$\frac{K}{T^{3/2}} \le \frac{1}{\sqrt{C}\,T^{3/4}} + T^{-3/2}, \qquad \frac{1}{CK} \le \frac{1}{\sqrt{C}\,T^{3/4}}.$$

Since $h_T \le 1 + \log T$, (D.26) proves the matching upper side of the theorem, without requiring the noise budget to choose K or B. Even caps larger than the usable $(q-1)T + 1$ cause no difficulty: the construction remains legal, and the $\log T/T$ term already dominates.

Finally $(\sqrt{C}\,T^{3/4})^{-1}$ is at most a fixed multiple of $(\log T)/T$ exactly when $C = \Omega(\sqrt{T}/(\log T)^2)$. Together with (D.20) this proves the necessary-and-sufficient cap scale for attaining the budget floor in *order*. It is not a prediction of the finite cap at which a numerical curve becomes flat, nor a uniqueness claim for optimal weights. At fixed load, total risk remains $\Theta(T^{-1/2})$ at every cap because V_0 has that order. The improving additional risk and the unchanged contraction baseline are separate consequences of one joint value comparison.

Load dependence of the additional risk guarantee. Theorem 3 fixes positive $\bar{\Delta}$ before taking large horizons. The following consequence of the same value formula explains why its attaining rule need not have a load-uniform competitive ratio. Fix positive $L, D, \sigma_{\mathrm{ind}}, c$, inverse-root rates with $c \le 1/(2L)$, and a finite T for which its tail has $1 < K \le T$. Set $C = q$, so every legal allocation is uniform. Let $\lambda_T \in (0, L]$ maximize the exact noiseless endpoint value. All products are positive, and for every $t < T$, $p_{0,t}(\lambda_T) > p_{0,T}(\lambda_T)$. The tail assigns positive mass to such a state, hence

$$d_T := \mathrm{Bias}(w^{(K)}) - V_0 \ge \frac{D^2\lambda_T}{2}[a_{\lambda_T}(w^{(K)})^2 - p_{0,T}(\lambda_T)^2] > 0.$$

This gap is independent of $\bar{\Delta}$. The endpoint has zero initialization risk above the baseline, and (6) bounds its noise contribution by

$$E_T(q,\bar{\Delta}) \le \Gamma_T \bar{\Delta}, \qquad \Gamma_T := \frac{1}{q} \max_{0\le\lambda\le L} \lambda^2 \sum_{j<T} b_{\lambda,j}(e_T)^2 \in (0,\infty).$$

Here using only the total inverse-curvature mass constraint gives a valid upper bound under both covariance constraints. If $E_T^{\mathrm{explicit}} := \mathfrak{R}_{\Delta,q}(w^{(K)}) - V_0$, then

$$\frac{E_T^{\mathrm{explicit}}}{E_T(q,\bar{\Delta})} \ge \frac{d_T}{\Gamma_T \bar{\Delta}} \longrightarrow \infty \qquad (\bar{\Delta} \downarrow 0). \tag{D.27}$$

This fixed-horizon limit does not contradict the fixed-positive-load, large-horizon theorem. The fixed-positive-load quantifier is essential. In particular, substituting a horizon-dependent $\bar{\Delta}_T$ into Theorem 3 does not give a uniform result: its constants and onset may depend on the load. This restriction is specific to the resource law; the information equalities also include the variance-only class and $\bar{\Delta} = \infty$.

The bias mechanism is not specific to a uniform tail. At a fixed horizon, any fixed, load-independent convex output assigning positive mass before T has $a_{\lambda_T}(w) > p_{0,T}(\lambda_T)$ in this nonnegative-transfer regime. Its positive noiseless bias increment is independent of every batch choice. Meanwhile the uniform-batch endpoint remains feasible for every $C \ge q$ and has additional risk at most $\Gamma_T\bar{\Delta}$. Reallocating samples alone therefore cannot remove this vanishing-load relative gap in additional risk. Removing that obstruction would require changing the output or the type of guarantee, not just improving the allocation.

E Finite dimensions, scope and the same-budget comparison

E.1 Finite dimension while preserving the exact additional risk

All values in this subsection use the individual-sample class (5), the same pathwise budget qT, integer cap C, and scalar convex outputs. The subscript d restricts instances to dimension at most d. It does not restrict a policy's observation information. The noiseless value remains exactly V_0 in every $d \ge 1$: the endpoint attains it, and a one-dimensional noiseless maximizing curvature forces it pathwise.

Exact thresholds in Corollary E.1. The dimension conditions use

$$r_T(m) = \left\lceil \frac{(\sqrt{qT} + \sqrt{2\log(4m)})^2}{(1-1/\sqrt{2})^2} \right\rceil,$$
$$s_0 = \lceil (10 + (8cL)^{-1})\sqrt{T}\log T \rceil, \qquad J_T = 1 + \lfloor \log_2(T/(4s_0)) \rfloor.$$

At the sufficiently large horizons specified below, define

$$d_{\mathrm{fix}}(T) = 1 + r_T(1), \qquad d_{\mathrm{grow}}(T) = 1 + (J_T+1) r_T(J_T+1). \tag{E.1}$$

These are the exact sufficient thresholds, of orders $O(qT)$ and $O(qT\log T)$, respectively. For fixed cap the constants and onset may depend on C; the growing-cap conclusion is simultaneous in C, d, with constants and onset depending only on the fixed parameters of Theorem 3.

Additional risk at an explicit finite dimension. Let $V_{T,d}^{\mathrm{pred}}(C)$ restrict the same finite-load class and feedback policies to dimension at most d, and put $E_{T,d}^{\mathrm{pred}} = V_{T,d}^{\mathrm{pred}} - V_0$. Approximating total risk within a factor $1-\eta$ would only give $(1-\eta)E_T - \eta V_0$, which can erase the smaller additional risk.

Corollary E.1 (Finite-dimensional additional risk)**.** *Under the fixed positive parameters of Theorem 3, use the explicit thresholds $d_{\mathrm{fix}}(T)$ and $d_{\mathrm{grow}}(T)$ in* (E.1)*. For every fixed $C \ge q$, $E_{T,d}^{\mathrm{pred}}(C) = \Theta(T^{-3/4})$ whenever $d \ge d_{\mathrm{fix}}(T) = O(qT)$, at a sufficiently large horizon; the constants and onset may depend on this fixed C. For growing caps, there are $a, A > 0, T_0$, depending only on the fixed parameters and independent of C, d, such that for all $T \ge T_0, C \ge q$ and $d \ge d_{\mathrm{grow}}(T) = O(qT\log T)$,*

$$a\Psi_T(C) \le E_{T,d}^{\mathrm{pred}}(C) \le A\Psi_T(C). \tag{E.2}$$

The proof keeps the exact baseline in a noiseless coordinate and replicates only noise (Appendix E.1). These are sufficient dimensions, not a fixed-dimensional value equality. The conservative onset conditions in Appendix E.2 do not determine the horizons of the separate numerical certificates.

Lemma E.1 (Replication above a pathwise baseline)**.** *Fix $0 < \gamma_j \le 1/(2L)$, $q > 1$ and $C \ge q$ integers, and positive finite $L, D, \sigma_{\mathrm{ind}}, \bar{\Delta}$. Choose $\lambda_* \in \arg\max_{\lambda\in[0,L]} \lambda p_{0,T}(\lambda)^2$ and let $b_*(w) = D^2\lambda_* a_{\lambda_*}(w)^2/2$. For any feasible finite noise measure $\nu = \sum_{i=1}^m \nu_i\delta_{\lambda_i}$, with $\nu_i > 0$, $0 < \lambda_i \le L$, $m \ge 1$, write*

$$n_\nu(w, B) = \sum_{i=1}^m \nu_i\lambda_i^2 \sum_{j<T} b_{\lambda_i,j}(w)^2/B_j, \qquad \underline{E}_\nu = \min_{\Delta_T\times\mathcal{B}_C} \{b_*(w) - V_0 + n_\nu(w, B)\}. \tag{E.3}$$

Here $\mathcal{B}_C$ permits real batches, and feasibility means $\sum_i \nu_i \le \bar{\Delta}$ and $\sum_i 2\lambda_i\nu_i \le \sigma_{\mathrm{ind}}^2$. For $\varepsilon, \delta \in (0, 1)$, let

$$r = \left\lceil \frac{(\sqrt{qT} + \sqrt{2\log(2m/\delta)})^2}{(1 - \sqrt{1-\varepsilon})^2} \right\rceil. \tag{E.4}$$

If $d \ge 1 + mr$, then

$$E_{T,d}^{\mathrm{pred}}(C) \ge (1 - \delta)(1 - \varepsilon)\underline{E}_\nu. \tag{E.5}$$

Proof. Put initialization D in one noiseless coordinate of curvature λ_*. Positivity of the contractions gives $a_{\lambda_*}(w) \ge p_{0,T}(\lambda_*)$ for every convex w, so $b_*(w) \ge V_0$ for each realized output. Replicate noise atom i in r zero-initialized coordinates, each with curvature λ_i and individual noise variance $2\lambda_i\nu_i/r$. These independent Gaussian coordinates and the bias coordinate form one fixed instance of dimension $1 + mr$. Both noise budgets and the initialization bound hold exactly.

Let $N = qT$. Expose the individual noises as $r \times N$ standard Gaussian matrices G_i. Gaussian singular-value concentration (Vershynin, 2012, Corollary 5.35) and a union bound give

$$\Pr(\mathcal{G}) \ge 1 - \delta, \qquad \mathcal{G} = \bigcap_{i=1}^m \{G_i^\top G_i/r \succeq (1 - \varepsilon)I_N\},$$

using (E.4). Indeed, $s_{\min}(G_i) \ge \sqrt{r} - \sqrt{N} - \sqrt{2\log(2m/\delta)} \ge \sqrt{r(1-\varepsilon)}$ on this event.

For every legal sample partition and output, the coefficient of an individual sample in batch j is $b_{\lambda_i,j}(w)/B_j$. Thus its vector has squared norm $\sum_j b_{\lambda_i,j}(w)^2/B_j$, also when both choices use observations and an independent policy seed. The event controls all such vectors simultaneously. Writing the realized noise loss as $\widetilde{n}_\nu$, on $\mathcal{G}$ we have

$$\ell_Q(w; B) - V_0 = b_*(w) - V_0 + \widetilde{n}_\nu(w, B) \ge (1 - \varepsilon)[b_*(w) - V_0 + n_\nu(w, B)] \ge (1 - \varepsilon)\underline{E}_\nu.$$

Off the event the left side is still nonnegative, because the noiseless coordinate alone contributes at least V_0. Taking expectations proves (E.5) for every policy on the same instance, hence for its minimax value. Embed by zero coordinates when $d > 1 + mr$. The minimum over real batches only weakens the lower bound on realized integer choices. □

The right side uses the specified adversary's $\underline{E}_\nu$. It is not identified with the unrestricted optimal additional risk. In contrast, applying a total-risk approximation and then subtracting V_0 leaves the uncontrolled term $-\eta V_0$.

A single noise atom: fixed cap. We prove the first part of Corollary E.1. Use the inverse-root rates of Theorem 3, $\lambda_* = \lambda_T$ from (D.3), and $\rho = \min\{\bar{\Delta}, \sigma_{\mathrm{ind}}^2/(2L)\} > 0$. For this particular bias coordinate, the derivation of (D.4) gives, for every w,

$$b_*(w) - V_0 \ge a_b\ell/T^{3/2}, \qquad \ell = \sum_t w_t(T - t), \quad a_b = D^2/(64e^2c). \tag{E.6}$$

Take $\nu = \rho\delta_L$. The proof of (D.5), with $B_j \le C$, gives for $\ell \le T/8$

$$n_\nu(w, B) \ge \frac{a'_n}{C(4\ell + d_0\sqrt{T})}, \qquad a'_n = \frac{9\rho(1 - e^{-1})^2}{16}, \quad d_0 = 2 + (Lc)^{-1}. \tag{E.7}$$

Adding (E.6) and (E.7) and minimizing the resulting expression in $4\ell + d_0\sqrt{T}$ gives

$$b_*(w) - V_0 + n_\nu(w, B) \geq \frac{\sqrt{a_b a'_n}}{\sqrt{C}\, T^{3/4}} - \frac{a_b d_0}{4T}. \tag{E.8}$$

If $\ell > T/8$, the initialization risk above the baseline is at least $a_b/(8\sqrt{T})$. For sufficiently large T, uniformly over $C \geq q$, it also dominates the right side of (E.8). For each fixed C, the negative remainder is smaller than half the leading term at a sufficiently large horizon. Lemma E.1 with $m = 1$ and $\varepsilon = \delta = 1/2$ proves the finite-dimensional lower order with $d \geq 1 + r_T(1)$. Uniform integer batches and the tail from Lemma D.1 give the matching upper. The constants and onset in this fixed-cap statement may depend on C.

A fixed mixture: all caps simultaneously. We now specify one noise measure before seeing any weights, batches, or observations. Put

$$A_0 = 10 + (8cL)^{-1}, \quad s_0 = \lceil A_0\sqrt{T}\log T\rceil, \quad s_k = 2^k s_0 \leq T/4, \quad J = 1 + \lfloor \log_2(T/(4s_0))\rfloor.$$

Choose a starting horizon depending only on the fixed parameters, large enough that $T \geq e$, $4s_0 \leq T^{3/4}$, and all preceding lookback bounds hold. Then $J \geq (\log T)/(4\log 2)$. Set

$$\lambda_k = \frac{\sqrt{T}}{8cs_k}, \qquad \nu = \frac{\rho}{2}\delta_L + \frac{\rho}{2J}\sum_{k=0}^{J-1}\delta_{\lambda_k}. \tag{E.9}$$

Its mass is ρ, each $\lambda_k \leq L$, and its covariance trace is at most $2L\rho$. Hence both noise budgets hold; the scales share the budget rather than each receiving it. There are at most $J + 1$ positive noise atoms.

If $\ell \geq \sqrt{T}\log T$, (E.6) already gives a lower bound $a_b \log T/T$. Otherwise $m_w = \max\{1, \lceil 4\ell\rceil\}$ satisfies $2m_w \leq s_0$. The fixed, disjoint blocks $I_k = \{T - 2s_k, \ldots, T - s_k - 1\}$ therefore obey the same transfer bound as (D.19), uniformly for every such output. With $n_k = \sum_{j\in I_k} B_j$,

$$n_\nu(w, B) \geq \frac{9\rho}{2048e^2 J}\sum_{k=0}^{J-1}\frac{1}{n_k} \geq \frac{9\rho J}{2048e^2 qT} \geq \frac{9\rho}{8192e^2 q\log 2}\frac{\log T}{T}. \tag{E.10}$$

The middle step is Cauchy–Schwarz and $\sum_k n_k \leq qT$. This argument uses a fixed mixture, not a scale chosen after a feedback path is observed.

Let $a'_f = \min\{a_b, 9\rho/(8192e^2 q\log 2)\} > 0$. Equation (E.10) and the large-lookback alternative give $\underline{E}_\nu \geq a'_f \log T/T$. The atom at L also gives (E.8) with a'_n replaced by $a'_n/2$, uniformly over caps. Average these two bounds and increase the same cap-independent onset until $a_b d_0/(8T) \leq a'_f \log T/(4T)$. It follows that

$$\underline{E}_\nu \geq \frac{\sqrt{a_b a'_n/2}}{2\sqrt{C}\, T^{3/4}} + \frac{a'_f}{4}\frac{\log T}{T}. \tag{E.11}$$

Apply Lemma E.1 at $\varepsilon = \delta = 1/2$, using at most $J+1$ atoms. Since $J = O(\log T)$ and $r_T(J+1) = O(qT)$, the displayed dimension in Corollary E.1 is sufficient. The deterministic construction from Theorem 3 supplies the upper bound for every finite-dimensional subclass. All constants and this onset are independent of C, d. This proves the corollary without a V_0 subtraction error.

E.2 Conservative sufficient onset and finite-budget certificates

The preceding proof chooses $s_0 = \lceil A_0\sqrt{T}\log T\rceil$, $A_0 = 10 + (8cL)^{-1}$, and requires $4s_0 \leq T^{3/4}$, in addition to the lookback and remainder conditions stated there. For $c = 1/2, L = 1$, dropping the ceiling gives the simpler condition $41\log T \leq T^{1/4}$. With natural logarithms its larger boundary root is approximately 1.79045×10^{12}. This is only the boundary of the simplified inequality. The original ceiling and other sufficient conditions must still be checked; this number is not claimed to be a complete T_0.

These are conservative sufficient conditions for the asymptotic finite-dimensional order, not necessary horizons or evidence of failure below them. No fixed-rank value equality follows. Our finite-budget experiments evaluate finite-horizon risk formulas with independent spectral and arithmetic certificates; they do not use this sufficient onset to choose their budgets. No optimization of proof constants is needed for either type of conclusion.

E.3 THE CONSTANT-STEP COMPARISON

Removing noise penalties does not change the noiseless baseline. Even removing every noise term leaves $V_0 = \Theta(T^{-1/2})$. Batches cannot change this noiseless trajectory. A different prescribed learning-rate sequence can cross that boundary.

Proposition E.1 (A same-budget constant-step baseline)**.** *On $\mathcal{Q}_\Delta$ with finite $\bar{\Delta}$, take $0 < \eta \le 1/(2L)$, $B_j = q$, and $\bar{x}_T = T^{-1}\sum_{t=1}^{T} x_t$. Then*

$$R_T^{\mathrm{const}} := \sup_{Q\in\mathcal{Q}_\Delta} \mathbb{E}_Q[F_Q(\bar{x}_T) - F_Q^*] \le \frac{D^2}{2\eta T} + \frac{\bar{\Delta}}{qT}. \tag{E.12}$$

In particular, $\eta = 1/(2L)$ gives $R_T^{\mathrm{const}} \le (LD^2 + \bar{\Delta}/q)/T$.

Constant-step averaging has classical fast-rate precedents (Bach & Moulines, 2013); the proof below uses our budgets. For fixed positive parameters and finite $\bar{\Delta}$, the exact noiseless lower bound (D.3) gives $V_T^{\mathrm{pred}}/R_T^{\mathrm{const}} = \Omega(\sqrt{T})$. Both rate sequences have T updates, qT samples and the same cap; their cumulative step mass is free to differ. This is a separation of class guarantees, not taskwise dominance.

We prove Proposition E.1. Set $r_\lambda = 1 - \eta\lambda \in [1/2, 1]$. The average includes the post-update states $1, \dots, T$, not x_0. For $\lambda > 0$, its initialization and injection transfers are

$$a_\lambda = \frac{1}{T}\sum_{t=1}^{T} r_\lambda^t \le \min\{1, (\eta\lambda T)^{-1}\}, \qquad b_{\lambda,j} = \frac{1 - r_\lambda^{T-j}}{\lambda T}, \quad 0 \le j < T.$$

Here the j-th noise enters x_{j+1}, so there are exactly $T - j$ terms in its geometric sum. It follows that $\lambda a_\lambda^2 \le 1/(\eta T)$ and $\lambda^2 \sum_{j<T} b_{\lambda,j}^2 \le 1/T$. Let s_i be the individual covariance diagonal in an eigenbasis of H. Independent minibatches have covariance Σ/q. The fixed-output decomposition therefore gives

$$\begin{aligned}\mathbb{E}[F(\bar{x}_T) - F^*] &= \tfrac{1}{2}\sum_i \lambda_i a_{\lambda_i}^2 \mathbb{E}x_{0,i}^2 + \frac{1}{2q}\sum_i s_i\lambda_i \sum_{j<T} b_{\lambda_i,j}^2 \\ &\le \frac{D^2}{2\eta T} + \frac{1}{qT}\sum_{i:\lambda_i>0}\frac{s_i}{2\lambda_i} \le \frac{D^2}{2\eta T} + \frac{\bar{\Delta}}{qT}.\end{aligned}$$

Zero-curvature coordinates contribute no objective risk; the kernel condition in (5) excludes noise there. Off-diagonal covariance entries do not enter this quadratic expectation. The initialization and future oracle samples are independent, so no cross term remains.

For fixed positive parameters, the noiseless bound in (D.3) gives $V_T^{\mathrm{pred}} \ge D^2/(16e^2c\sqrt{T})$ at sufficiently large T, uniformly in C. With $L = D = 1$, $\eta = 1/2$,

$$\frac{V_T^{\mathrm{pred}}}{R_T^{\mathrm{const}}} \ge \frac{1}{16e^2c(1 + \bar{\Delta}/q)}\sqrt{T}. \tag{E.13}$$

Positivity follows by taking a noiseless positive-curvature instance with nonzero initialization. The denominator is the worst-case risk of this one implementable output, not an unrestricted training minimax value. The gap is between class risks; the two worst cases need not be the same instance. Indeed, the numerator's noiseless lower bound already comes from limited contraction. Equating ηT to the inverse-root cumulative mass would remove the freedom being tested and would not establish (E.13).

F NUMERICAL BOUNDS AND EVIDENCE VERIFICATION

Continuous-curvature verification. The numerical bounds cover the entire curvature interval, including zero. A feasible spectral adversary gives a lower bound for a fixed plan. Its supporting-plane minimum gives a lower bound for the full output and sampling class. Feasible plans give upper bounds once both the bias and noise are bounded over all curvatures.

Signed transfers. The final rate library includes $1/L < \gamma_t < 2/L$. For a curvature interval $I = [\ell, u]$, each transfer factor lies in $[1 - u\gamma_t, 1 - \ell\gamma_t]$. Interval multiplication and addition enclose

$$p_0 = 1, \quad p_{t+1} = (1 - \lambda\gamma_t)p_t, \quad a_\lambda = \sum_{t=0}^{T} w_t p_t, \\ h_{T-1} = w_T, \quad h_j = w_{j+1} + (1 - \lambda\gamma_{j+1})h_{j+1}. \tag{F.1}$$

Since $b_{\lambda,j} = \gamma_j h_j$, squaring these enclosures bounds a_λ^2 and $Q_B(\lambda) = \sum_j b_{\lambda,j}^2/B_j$, regardless of transfer signs. Write their upper bounds as A_I, Q_I. The bias is at most $D^2 u A_I/2$. If $\alpha, \beta \geq 0$ satisfy $2\ell\alpha + \beta \geq u^2 Q_I$ on every interval, then $\sigma_{\mathrm{ind}}^2\alpha + \bar{\Delta}\beta$ bounds the noise integral by (C.5). No division by a transfer product or monotonicity of squared transfers is used. Appendix G.4 gives the signed-interval protocol and independent Arb replay for the final comparison.

Nonnegative transfers as a special case. The inverse-root studies and restricted-rate controls satisfy $0 < \gamma_t \leq 1/L$. Their products are nonnegative and nonincreasing on $[0, L]$, as are a_λ^2 and $Q_B(\lambda)$ for nonnegative weights. For either function g,

$$\sup_{\ell \leq \lambda \leq u} \lambda g(\lambda) \leq u g(\ell). \tag{F.2}$$

Those studies use this tighter endpoint envelope and impose $2\ell\alpha + \beta \geq u^2 Q_B(\ell)$ for the noise dual. This shortcut is not used for rates with signed transfers. The inverse-root resource law retains the stronger step-size assumption in its theorem.

Subtracting a noiseless interval $[L_0, U_0]$ from a total-risk interval $[L_\Delta, U_\Delta]$ gives the additional risk interval $[\max\{0, L_\Delta - U_0\}, U_\Delta - L_0]$. Study-specific refinement rules are stated with each experiment. Float64 calculations with guards guide optimization; they are not directed-rounding certificates. Arb or exact rational replay is identified separately for each certified comparison.

G Complete numerical results

G.1 Equal contraction does not determine the policy-class ranking

This application uses the general prescribed-rate comparison in Theorem 2; it adds no assumption or new minimax theorem. All values refer to the same additive-quadratic class with $L = D = \sigma_{\mathrm{ind}} = \bar{\Delta} = 1$, $T = 8$, $q = 2$, $C = 3$, and hence 16 individual samples. For sequence S, write V_S^{det} for (7), V_S^{pred} for its predictable-policy value and V_S^{rel} for the joint real-batch relaxation. Let $R_S(w, B)$ denote a fixed pair's worst-case class risk.

An exact contraction control. The two prescribed rate sequences are

$$\gamma^{(A)} = (1/4, 1/4, 1/50, 1/2, 1/2, 1/50, 2/25, 2/25), \\ \gamma^{(B)} = (1/2, 1/4, 1/4, 2/25, 1/50, 1/50, 2/25, 1/2).$$

Both have nonnegative transfers and the same rate multiset, step mass and squared-step mass. More strongly, their noiseless endpoint polynomial is identical on the entire curvature interval:

$$p(\lambda) = (1 - \lambda/2)^2(1 - \lambda/4)^2(1 - \lambda/50)^2(1 - 2\lambda/25)^2.$$

Since every partial product is at least its endpoint product, every convex output obeys $a_\lambda(w) \geq p(\lambda) \geq 0$. Returning the endpoint attains equality simultaneously for all curvatures. Thus both exact optimal noiseless class values equal $V_0 = \frac{1}{2}\max_{[0,1]} \lambda p(\lambda)^2$, regardless of sampling choices. The intermediate noiseless states, and the bias for a general average, need not agree. Permutation invariance of the quadratic noiseless endpoint and order-dependent stability are established observations (Agarwal et al., 2021, Remark 2 and Section 3.3). Here they provide a control for the full-policy comparison.

A lower bound on every policy for A. Choose initialization mass one at curvature $5/16$, and individual noise variance one at curvature 1, in orthogonal directions. The inverse-curvature noise load is $1/2 \le \bar{\Delta}$. The corresponding common-adversary objective in (C.6) is

$$f(w,B) = \frac{5}{32} a_{5/16}(w)^2 + \frac{1}{2}\sum_{j=0}^{7} \frac{b_{1,j}(w)^2}{B_j}, \quad (w,B) \in K := \Delta_8 \times \{B \in [1,3]^8 : \textstyle\sum_j B_j = 16\}. \tag{G.1}$$

It is jointly convex. At a feasible rational point $z = (w^*, B^*)$, let g_w, g_B be its gradients. The supporting-plane bound is

$$\min_K f \;\ge\; L_A := f(z) - g_w^\top w^* - g_B^\top B^* + \min_{0\le t\le 8} g_{w,t} + \min_{B\in[1,3]^8,\,\sum B=16} g_B^\top B. \tag{G.2}$$

The last linear minimum starts with one sample per update and assigns two extra samples to each of the four smallest gradient coordinates. The rational point, with every entry expressed over 10^9, is

$$\begin{aligned} 10^9 w^* &= (0,0,0,141232006,60220161,0,0,0,798547833),\\ 10^9 B^* &= (1854218277,2472290681,1000000000,3000000000,\\ &\qquad 3000000000,1000000000,1760214367,1913276675). \end{aligned}$$

Exact rational evaluation gives $L_A > 0.0884794$, with $f(z) - L_A < 4.13 \times 10^{-10}$. Because this adversary is feasible for every pair in K, $V_A^{\mathrm{rel}} \ge \min_K f \ge L_A$. Theorem 2 then gives $V_A^{\mathrm{pred}} \ge V_A^{\mathrm{rel}} > 0.0884794$. This transfer uses the existing replicated common-adversary argument over arbitrary finite dimensions, uniform over weights and sample partitions. The two spectral atoms are not a claim of a two-dimensional adaptive lower bound.

Continuum bounds for the endpoints and B's integer plan. For a fixed pair put $Q_B(\lambda) = \sum_j b_{\lambda,j}(w)^2/B_j$. Both a_λ^2 and $Q_B(\lambda)$ are nonnegative and nonincreasing on $[0,1]$. Divide this interval into 1024 equal rational cells $[\ell, u]$, including the cell starting at zero. The initialization contribution lies between $\max_{\lambda\in\mathcal{G}} \lambda a_\lambda^2/2$ and $\max_{[\ell,u]} u a_\ell^2/2$, where $\mathcal{G} = \{i/1024 : 0 \le i \le 1024\}$. A single feasible noise atom at λ can have variance $\min\{1, 2\lambda\}$. Thus a lower noise bound is $\max_{\lambda\in\mathcal{G}} \min\{1,2\lambda\}\lambda Q_B(\lambda)/2$. For an upper bound we relax the inverse-curvature constraint and retain covariance trace at most one, giving $\max_{[\ell,u]} u Q_B(\ell)/2$. Initialization and noise occupy orthogonal blocks, so the respective lower bounds can be added; the same holds for the upper bounds. These envelopes cover the continuum, rather than just grid values.

Apply them to both endpoints with $B_j = 2$, and to B's legal plan $B = (3,3,3,1,1,1,1,3)$, $w_7 = w_8 = 1/2$. The outward-rounded endpoint intervals are $[0.1090483, 0.1092843]$ for A and $[0.1218116, 0.1220255]$ for B. In particular,

$$R_{A,\mathrm{end},2} < R_{B,\mathrm{end},2}, \qquad V_A^{\mathrm{pred}} > 0.0884794 > 0.0851448 > R_B^{\mathrm{plan}} \ge V_B^{\mathrm{det}} \ge V_B^{\mathrm{pred}}. \tag{G.3}$$

The printed bounds alone imply $V_A^{\mathrm{pred}}/V_B^{\mathrm{det}} > 1.0391$. No different baseline is subtracted: this is a strict total-risk ordering reversal with identical optimal noiseless contraction.

A four-parameter neighborhood of the reversal. The strict gaps above imply an explicit neighbourhood result without another optimization. Put $\theta = (a,b,c,d)$, $\theta_0 = (1/4, 1/50, 1/2, 2/25)$, and define

$$\gamma^{(A)}(\theta) = (a,a,b,c,c,b,d,d), \qquad \gamma^{(B)}(\theta) = (c,a,a,d,b,b,d,c).$$

Retain the same class budgets, $T = 8, q = 2, C = 3$, and B's fixed integer plan above. For every $\|\theta - \theta_0\|_\infty \le \varepsilon := 1/20000$ with $0 < a,b,c,d \le 1/2$, both orderings remain strict:

$$\begin{aligned} R_{B,\mathrm{end},2}(\theta) - R_{A,\mathrm{end},2}(\theta) &> 0.0106273,\\ V_A^{\mathrm{pred}}(\theta) - R_B^{\mathrm{plan}}(\theta) &> 0.0014346. \end{aligned} \tag{G.4}$$

Their common noiseless endpoint polynomial is always $(1-a\lambda)^2(1-b\lambda)^2(1-c\lambda)^2(1-d\lambda)^2$. Hence they retain the same exact noiseless class value. The box is intersected with the stated rate domain; at the base point its c-direction is one-sided.

To verify the margins, consider any two length-eight rate sequences in $(0, 1/2]^8$ at sup distance at most ε. Extend $R_\gamma(w, B)$ to real batches by (6). For the same $w \in \Delta_8$, $B \in [1, 3]^8$, $\sum_j B_j = 16$, telescoping the positive products gives, for $0 \le \lambda \le 1$,

$$|a'_\lambda - a_\lambda| \le 8\lambda\varepsilon, \qquad |b'_{\lambda,j} - b_{\lambda,j}| \le \left(1 + \frac{7-j}{2}\right)\varepsilon \quad (0 \le j < 8).$$

Indeed, each factor changes by at most $\lambda\varepsilon$; in $b_{\lambda,j}$ there are at most $7 - j$ subsequent factors, multiplied by $\gamma_j \le 1/2$. Also $0 \le a_\lambda, a'_\lambda \le 1$ and $0 \le b_{\lambda,j}, b'_{\lambda,j} \le 1/2$. The initialization term of (6) changes by at most 8ε. Since every feasible measure obeys $\int \lambda^2 d\nu \le \int \lambda d\nu \le 1/2$, the noise change is at most

$$\frac{\varepsilon}{2} \sum_{j=0}^{7} \left(1 + \frac{7-j}{2}\right) = 11\varepsilon.$$

Taking the curvature and measure suprema preserves these uniform bounds, so

$$|R_{\gamma'}(w, B) - R_\gamma(w, B)| \le 19\varepsilon.$$

Taking the infimum over the same real-batch feasible set also gives $|V^{\mathrm{rel}}_{\gamma'} - V^{\mathrm{rel}}_\gamma| \le 19\varepsilon$. Thus Theorem 2 and the rational certificate imply

$$V^{\mathrm{pred}}_A(\theta) \ge V^{\mathrm{rel}}_A(\theta) > 0.0884794 - 19\varepsilon, \qquad R^{\mathrm{plan}}_B(\theta) < 0.0851448 + 19\varepsilon.$$

Their gap is greater than $0.0033346 - 38\varepsilon = 0.0014346$. The endpoint gap is greater than $0.1218116 - 0.1092843 - 38\varepsilon = 0.0106273$, which proves (G.4). The argument concerns the full feedback class through the general minimax theorem, not a fixed-dimensional adaptive value. This small radius establishes local persistence, not typical-task prevalence; the base point's 3.86% ratio is not asserted throughout the box. The margins follow from the rational base-point bounds and the uniform product estimate above.

Verification and interpretation. Exact rational arithmetic checks the rate multisets, polynomial coefficients, all budgets, (G.2), and all 1024 continuum envelopes. An independent forward covariance recursion checks the backward transfer computation at each grid node. Floating-point optimization is used only to discover a feasible rational witness. The inequalities above are verified at that witness without numerical guards. These finite checks invoke the general minimax theorem; they do not replace its proof.

This constructed example establishes the use of the general theorem: a candidate endpoint can favor a rate sequence whose best admissible policy class is strictly worse. It neither estimates the prevalence of such reversals on training tasks nor proves that both output and sampling optimization are necessary for any reversal. No longer-sequence search is used to support the claim.

G.2 Strict benefit from independent plan randomization

Theorem 2 distinguishes observing the history from randomizing a complete plan before sampling. For $T = 3$, $\gamma = (1, 1, 1)$, $q = 2, C = 3$ and $L = D = \sigma_{\mathrm{ind}} = \bar{\Delta} = 1$, Table G.1 gives an exact four-component mixture. Each component obeys the integer sample budget and batch cap. The mixture averages the plan-specific squared transfer kernels before maximizing over the common spectral adversary. Since the instance is chosen before the seed is realized, all components are evaluated on the same instance. The certificates give

$$V^{\mathrm{mix}}_T \le 0.104162655 < 0.104999793 \le V^{\mathrm{det}}_T. \tag{G.5}$$

The certified reduction is at least 0.7972%. The mixture supplies an upper bound, not the exact mixed minimax value.

Why the mixture changes batch sizes. Table G.1 uses only two batch arrays. Conditional on each array, replace its component weights by their probability-weighted mean. Risk is convex in the averaging weights at a fixed batch array, so the resulting two-plan mixture has no greater risk on any instance and inherits the displayed certified upper bound. The witness therefore needs only a randomized choice between $(1, 2, 3)$ and $(2, 2, 2)$, with fixed averaging weights attached to each array. Here $1 - \gamma_t\lambda = 1 - \lambda \ge 0$ for all $\lambda \in [0, 1]$. Strict randomization benefit therefore occurs with nonnegative transfer factors.

k	m_k	$B^{(k)}$	$u^{(k)}$
1	36679935317944884	$(1,2,3)$	$(18042807570357185, 959526663743765,$ $4627968988472108, 26369696777426940)$
2	12474828136329467	$(2,2,2)$	$(29044948936743020, 18047442183161935,$ $17439146398313482, 35468462481781565)$
3	23873208226143405	$(1,2,3)$	$(360828447776450100, 19170206574212017,$ $92471457772035610, 527529887877302300)$
4	26972028319582314	$(2,2,2)$	$(29039438741828800, 18050548424983334,$ $17434351868623282, 35475660964564587)$

Table G.1: The exact four-component mixture for $T = 3$, $\gamma = (1,1,1)$, $q = 2, C = 3, L = D = \sigma_{\mathrm{ind}} = \bar{\Delta} = 1$. Probabilities are $p_k = m_k/100000000000000070$. Each output uses $w^{(k)} = u^{(k)}/\sum_{t=0}^{3} u_t^{(k)}$. The integers encode the saved certified plan exactly, without decimal rounding.

Lower certificates and continuous-curvature verification. The seven legal batch arrays are the six permutations of $(1,2,3)$ and $(2,2,2)$. For each array, a feasible common spectral adversary and a supporting-plane bound cover all convex averaging weights. Their minimum exceeds the displayed deterministic threshold. For the mixture, the interval verifier bounds the averaged squared bias and noise kernels on a complete cover of $[0,1]$, using the two-constraint noise dual in Appendix F. The supplement's `EVIDENCE_INDEX.md`, under `Randomized witness`, identifies the exact component arrays, all seven lower witnesses and the upper cover. `check_randomized_witness.py` replays these certificates at 384 bits using division-free forward covariance propagation. The component table establishes feasibility; the lower witnesses and continuous-curvature cover establish the strict risk comparison.

The full prespecified grid contains 66 settings with $T \in \{2,3,4\}$, rational rates, $q = 2, C = 3$, $L = D = \sigma_{\mathrm{ind}} = 1$ and $\bar{\Delta} \in \{1/8, 1/4, 1\}$. Eight have certified strict separation; the remaining 58 are unresolved. The supplement retains all settings and the original eight-case summary. The existing independent check replays 678 fixed-batch lower bounds and every certified mixture cover, including zero contractions.

G.3 Noise replication and complete sample-grouping oracles

The experiment evaluates the common instance underlying the sequence-A lower bound in Appendix G.1. The horizon, rates and sample budget are fixed. Only the number of replicated noise coordinates changes. For $r \in \{1, 4, 16, 64, 256, 4096, 262144\}$, set

$$H_r = \mathrm{diag}(5/16, I_r), \qquad x_0 = (1, 0_r), \qquad \xi = (0, z/\sqrt{r}), \quad z \sim N(0, I_r).$$

The dimension is $r+1$, total noise variance is 1, inverse-curvature load is $1/2$, and initialization norm is 1 for every r. Every run uses $T = 8$, $N = 16$, $C = 3$ and sequence A in Appendix G.1.

Complete action set and paired samples. There are exactly 1107 vectors in $\mathcal{B}_{\mathrm{int}} = \{B \in \{1,2,3\}^8 : \sum_j B_j = 16\}$. A plan divides 16 individual samples into consecutive batches without reordering. On a stream ω, the grouping-and-output oracle returns

$$\Omega_r(\omega) = \min_{B \in \mathcal{B}_{\mathrm{int}}} \min_{w \in \Delta_8} F_{Q_r}\left(\sum_{t=0}^{8} w_t x_t(B;\omega)\right). \tag{G.6}$$

The oracle sees all samples before choosing both B and w. For any predictable policy, including an independently randomized one, the realized loss is at least $\Omega_r(\omega)$. The empirical mean of numerical oracle values is not an analytic lower bound on expected risk.

We generate 64 independent streams for each r using PCG64 and `SeedSequence([20260923063, r, stream])`, with stream indices 0 through 63. The full grid, seeds, code and tolerance were frozen before generating samples. The same stream supplies the fixed plan, uniform-batch output oracle and grouping-and-output oracle. The fixed plan

is chosen before those streams by replacing the normalized noise Gram matrix with I_{16}. Its integer batches are $(2, 2, 1, 3, 3, 1, 2, 2)$, and its only nonzero weights, on states 3, 4 and 8, are approximately $(0.129447389630, 0.059654802094, 0.810897808276)$. Its expected risk lies in the guarded FP64 interval $[0.088556053822, 0.088556053825]$, for every r. The optimum is specific to the common instance, rather than the class values V^{det} or V^{mix}.

Quadratic programs and numerical checks. Let $G \in \mathbb{R}^{r\times 16}$ contain the independent standard Gaussian samples. For each grouping, an independently checked response matrix $M_B \in \mathbb{R}^{9\times 16}$ maps samples to states. With $v_t = p_{0,t}(5/16)$, the path risk is

$$f_B(w) = w^\top Q_B w, \qquad Q_B = \tfrac{5}{32} v v^\top + \tfrac{1}{2} M_B (G^\top G/r) M_B^\top.$$

Each program optimizes the complete nine-state simplex. A feasible w gives upper value $f_B(w)$; convexity gives lower value

$$f_B(w) + \min_t [\nabla f_B(w)]_t - \nabla f_B(w)^\top w.$$

Both endpoints receive a Float64 guard $10^{-12}(1 + \max_{i,j} |(Q_B)_{ij}|)$. Taking minima across all groupings brackets the path oracle numerically. The guard is a floating-point diagnostic, not an Arb certificate. All 495936 random-path programs meet the frozen 10^{-7} tolerance, with at most 5000 iterations per program. The maximum program width is 3.024×10^{-8}; the maximum grouping-and-output-oracle width is 3.148×10^{-12}. Independent product-form responses reproduce every retained supporting plane, with maximum risk discrepancy 1.39×10^{-16}. Full-vector forward SGD checks three specified groupings on stream 0 at each r, and a separate SLSQP calculation checks three programs. All saved programs, including their bounds, weights and convergence records, remain in the supplement.

Table G.2: All seven common-instance settings, each with 64 streams. Oracle columns use the saved feasible upper endpoints. Fixed B count records how many streams select the fixed plan's grouping under the grouping-and-output oracle.

r	Fixed plan	Uniform-batch output oracle	Grouping-and-output oracle	Paired gain	Fixed B count
1	0.085117160	0.071608939	0.050409391	0.03470777	0
4	0.086271279	0.083818827	0.057456792	0.02881449	0
16	0.087432636	0.095351162	0.073749042	0.01368359	0
64	0.088671447	0.098268991	0.083318313	0.005353134	2
256	0.087751502	0.097838403	0.085834595	0.001916907	7
4096	0.088707421	0.098543193	0.088537151	0.0001702696	31
262144	0.088545941	0.098463940	0.088545693	2.480373e-07	64

Interpretation of the common-instance experiment. The paired oracle gain decreases to about 2.48×10^{-7}, with 95% interval $[1.976, 2.985] \times 10^{-7}$, at $r = 262144$. The gain remains positive; the experiment does not show equality at a fixed dimension. At the largest r, all 64 streams select the fixed plan's integer grouping, so the remaining gain comes from adapting the scalar averaging weights. The fixed-plan sample mean can differ from its population expectation because of Monte Carlo variation. Paired gains therefore compare the two sample risks, rather than subtracting a population value from an oracle sample mean.

The known-block control returns the endpoint on the bias coordinate and the initialization on the noise coordinates. Its risk is exactly $(5/32)p_{0,8}(5/16)^2$, approximately 0.050406941795. The control lies outside the shared-scalar output class. Its risk is the bias-coordinate endpoint risk on the common instance, not the full-class noiseless minimax value V_0.

A finite-dimensional exclusion of all admissible policies. The finite-dimensional result uses an analytic concentration bound and exact rational witnesses, separately from the simulation. Set $r = 2^{18}$, $d = r+1 = 262145$. For a standard Gaussian $G \in \mathbb{R}^{r\times 16}$, Vershynin (2012, Corollary 5.35) with deviation parameter 4 gives, with probability at least $1 - 2e^{-8}$,

$$s_{\min}(G) \geq \sqrt{r} - \sqrt{16} - 4 = 504, \qquad G^\top G/r \succeq (63/64)^2 I \succeq (31/32) I.$$

One event controls all averaging weights and all sample groupings. On that event, nonnegative bias and noise imply $\ell_A(w, B) \geq (31/32)f(w, B)$, where f is the common objective in (G.1). Its rational supporting-plane lower bound is L_A from (G.2). Off the event, loss is nonnegative. Since $2e^{-8} < 1/1024$, every admissible predictable policy satisfies

$$\inf_{\pi \in \Pi_{\mathrm{pred}}} \mathbb{E}_{Q_r} \ell_A(\pi) \geq \frac{1023}{1024}\frac{31}{32} L_A > 0.08563074 > 0.08514479 > U_B \geq R_B^{\mathrm{plan}}. \tag{G.7}$$

Here U_B is the exact rational continuum upper bound for sequence B's fixed plan in Appendix G.1. The inequality holds uniformly over auxiliary policy randomness because the Gram event controls every action. Exact arithmetic gives margin $(1023/1024)(31/32)L_A - U_B \approx 4.85960337 \times 10^{-4}$. The finite Taylor sum $\sum_{k=0}^{39} 8^k/k! > 2048$ verifies the probability comparison using rational arithmetic. The Gaussian theorem is an external mathematical input. The verifier checks the rational witness, all 1024 curvature cells and the final inequality. The dimension is sufficient, not minimal; the result does not assert a general fixed-dimensional minimax equality or the inverse-root asymptotic exponent.

G.4 Matched comparison with signed transfers

This comparison uses the six settings in Section 5. All methods use the same set Γ of prescribed learning-rate sequences. For $t = 0, \ldots, T-1$, set $\gamma_t = a f_t$, where

$$f_t \in \left\{1,\ (t+1)^{-1/2},\ \min\{1, 2(1-t/T)\},\ \frac{1+\cos(\pi t/T)}{2}\right\},$$

and use every amplitude in

$$\mathcal{A} = \{1/32, 1/16, 1/8, 3/16, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1,\\ 5/4, 3/2, 7/4, 9/5, 15/8, 31/16, 63/32\}.$$

There are 72 sequences and 432 rate-scenario records. Rates are selected deterministically before sampling. This finite comparison allows negative transfer factors and keeps the prescribed horizon, noise and sampling budgets.

		Uniform batches		Configured batches		Certified gap
		Selected simple		Selected simple		
T	$\bar{\Delta}$	plan U	Policy value $[L, U]$	plan U	Policy value $[L, U]$	upper bound (%)
64	1/4	3.575	[3.325, 3.400]	3.530	[2.960, 3.004]	19.22
64	1	8.686	[8.463, 8.532]	8.637	[8.089, 8.283]	6.77
64	4	22.098	[21.271, 21.294]	21.367	[20.728, 20.983]	3.08
256	1/4	0.904	[0.846, 0.865]	0.878	[0.755, 0.782]	16.21
256	1	2.276	[2.241, 2.262]	2.276	[2.136, 2.184]	6.49
256	4	7.600	[7.224, 7.252]	7.600	[6.932, 7.046]	9.63

Table G.3: Matched comparisons on one 72-rate library. Risk bounds are scaled by 10^3. Selected simple entries are feasible-plan upper bounds. Policy-value entries bracket the complete policy-class value. The final column is $100(U_{\mathrm{simple,configured}}/L_{\mathrm{full,configured}} - 1)$. All reported bounds are Arb-certified and independently replayed.

Simple output proposals. The search includes every integer tail length $K = 1, \ldots, T$, with weights $1/K$ on $x_{T-K+1}, \ldots, x_T$. Initialized EMA proposals use $\rho_{\mathrm{prop}} = 2^{-d/T}$, with $d \in \{64, 32, 16, 8, 4, 2, 1\}$. Decreasing geometric proposals use $r_{\mathrm{prop}} = \exp(-a/T)$, with

$$a \in \{0, 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 1, 2, 4, 8, 16,\\ 32, 64, 128, 256, 512, 1024\},$$

plus $r_{\mathrm{prop}} = 0$. The certified weights use the rational decay parameters specified below. The search retains the endpoint, full averaging on $x_0, \ldots, x_T$, step-proportional weights on $x_0, \ldots, x_{T-1}$, half-tail averaging on $x_{\lfloor T/2 \rfloor}, \ldots, x_T$, initialized EMA with $d = 8$, and ordinary post-update averaging.

The interval template is defined by

$$\widehat{x}_{\mathrm{seg}} = \alpha x_0 + \beta x_T + \frac{1-\alpha-\beta}{u-s+1}\sum_{t=s}^{u} x_t, \quad \alpha,\beta \geq 0, \quad \alpha+\beta \leq 1.$$

The 45 pairs $s \leq u$ use the knots $\{1, T/8, T/4, T/2, 3T/4, 7T/8, 15T/16, 31T/32, T-1\}$. There is no mixing across intervals, which would enlarge the output class. The template has four output parameters; its batch array still has T entries.

Joint-target screening of the simple outputs. A uniform-batch shortlist supplies a feasible incumbent, but it cannot exclude an output that benefits more from reallocation. The configured comparison screens every fixed candidate and interval support defined above under the same rates, budgets and cap. The simple-uniform column uses the uniform-batch proposals.

For a fixed rate and output, a feasible common spectral adversary A defines a convex objective

$$g_A(w,B) = c_A(w) + \sum_j \frac{a_{A,j}(w)}{B_j}, \qquad a_{A,j}(w) \geq 0.$$

Its relaxed minimum satisfies

$$\ell_{\gamma,w} := \min_{B \in \mathcal{B}_8^{\mathrm{rel}}} g_A(w,B) \leq \min_{B \in \mathcal{B}_8^{\mathrm{int}}} R_\theta(\gamma,w,B).$$

Let U_{inc} be the initial configured simple plan's certified upper bound in that setting. We exclude a candidate only when a certified lower bound for $\ell_{\gamma,w}$ is at least U_{inc}. A whole rate sequence can first be excluded by its certified full-class lower bound under the same test. Uniform-batch ranking is not an exclusion criterion.

For fixed w, the relaxed batch minimizer has capped square-root form. Instead of trusting a floating-point optimum, we certify the dual lower bound, for a proposed $\eta \geq 0$,

$$c_A(w) - 2T\eta + \sum_j \min_{1 \leq b \leq 8} \{a_{A,j}(w)/b + \eta b\}.$$

Each summand is $a+\eta$ when $a \leq \eta$, $a/8 + 8\eta$ when $a \geq 64\eta$, and $2\sqrt{a\eta}$ otherwise. Interval arithmetic verifies the value and the comparison to the incumbent. At an uncertain branch, the unconstrained minimum remains a valid lower.

For each interval support, all three masses and batches remain free. Joint convexity gives a supporting-plane lower bound over the entire mass simplex and batch polytope. Minimizing its mass part checks the three vertices, and minimizing its batch part is a capped linear allocation. A sampled mass triple is never used to exclude an interval. The feasible common adversary is checked independently before use.

Every unscreened candidate receives the same bounded spectral-exchange optimization as below, or reuses its identical cached joint solve. Every resulting plan is rounded to integer batches using largest remainders with index tie breaking, preserving $1 \leq B_j \leq 8$ and $\sum_j B_j = 2T$. We keep the initial incumbent as a fallback and do not retune weights after rounding. The surviving interval problems optimize masses continuously. Selected feasible plans receive full curvature-cover certificates. Solver brackets and unselected proposals are retained; this procedure does not prove the best integer plan in a template family. All tail lengths are enumerated, whereas EMA and geometric parameters cover only the stated finite grids.

Bounds for all admissible policies. Both the full-output uniform-batch program and the joint program are solved on every rate. Spectral exchange starts with 20 geometric nodes from 10^{-3} to 1, the midpoint, and the zero of the largest transfer when it lies in $[0,1]$. It stops at a 0.75% real-program bracket or 12 exchanges. Signed 16384-cell envelopes guide this search. Every remaining bracket is retained rather than triggering an outcome-dependent extension. The selected feasible plans are reevaluated with 65536 cells, doubled at most twice to target 0.1% fixed-plan width.

For each rate γ, a feasible spectral adversary supplies $\ell_\gamma \leq V_\theta^{\mathrm{pred}}(\gamma)$. The full library lower is $L_{\mathrm{full}} = \min_{\gamma \in \Gamma} \ell_\gamma$, including every signed-rate candidate. Separate lowers apply when batches are

fixed at two. Every simple and full-simplex upper in Table G.3 is a legal integer plan. For a simple plan, $U_{\text{simple}}/L_{\text{full}} - 1$ is a sufficient upper bound on its relative gap. A value above 5% leaves sufficiency at that tolerance unresolved; it is not evidence that every member of the template family is worse.

T	$\bar{\Delta}$	Batches	Shape	a	Simple output
64	$1/4$	uniform	constant	$15/8$	interval $[1, 63]$
64	$1/4$	configured	plateau–linear	$15/8$	interval $[1, 60]$
64	1	uniform	constant	$9/5$	interval $[1, 63]$
64	1	configured	constant	$15/8$	interval $[1, 63]$
64	4	uniform	constant	$3/8$	interval $[1, 56]$
64	4	configured	plateau–linear	$5/8$	interval $[1, 16]$
256	$1/4$	uniform	constant	$15/8$	interval $[1, 255]$
256	$1/4$	configured	constant	$9/5$	$d = 8$ EMA
256	1	uniform	constant	$15/8$	interval $[1, 255]$
256	1	configured	constant	$15/8$	interval $[1, 255]$
256	4	uniform	constant	$3/2$	interval $[1, 255]$
256	4	configured	constant	$3/2$	interval $[1, 255]$

Table G.4: Selected feasible simple plans. The configured choices follow the joint-target screen; uniform choices are retained. Neither selection asserts a template-family minimum.

Selected-plan comparison certificates. Let R_S and R_F be the risks of the selected configured simple and full-simplex integer plans. Table G.5 gives their fixed-plan intervals and the complete-library lower bound. The fixed-plan lowers evaluate feasible spectral witnesses at the selected weights and batches. They imply

$$1 - \frac{R_F}{R_S} \geq 1 - \frac{U_F^{\text{plan}}}{L_S^{\text{plan}}}, \qquad \frac{R_F}{V_\Gamma^{\text{pred}}(\theta)} - 1 \leq \frac{U_F^{\text{plan}}}{L_{\text{full}}} - 1. \tag{G.8}$$

The first comparison uses a lower bound for the selected simple plan, not its upper bound or a minimum over its template family.

T	$\bar{\Delta}$	$[L_S^{\text{plan}}, U_S^{\text{plan}}]$	$[L_F^{\text{plan}}, U_F^{\text{plan}}]$	L_{full}	Gain $\geq$	Gap $\leq$
64	$1/4$	[3.526823, 3.529755]	[2.990013, 3.003532]	2.960919	14.83%	1.44%
64	1	[8.630473, 8.636377]	[8.277501, 8.282283]	8.089174	4.03%	2.39%
64	4	[21.354385, 21.366924]	[20.970198, 20.982568]	20.728489	1.74%	1.23%
256	$1/4$	[0.876962, 0.877506]	[0.780563, 0.781023]	0.755145	10.94%	3.43%
256	1	[2.274140, 2.275324]	[2.182295, 2.183729]	2.136809	3.97%	2.20%
256	4	[7.594037, 7.599234]	[7.040588, 7.045305]	6.932190	7.22%	1.64%

Table G.5: Selected-plan certificates for Table 1. Risk bounds are scaled by 10^3 and rounded outward. Gain is $1 - U_F^{\text{plan}}/L_S^{\text{plan}}$ and gap is $U_F^{\text{plan}}/L_{\text{full}} - 1$, computed before rounding. The first row uses the constant-rate $31/16$ full-simplex plan. Its two risk endpoints come from the same plan certificate. Identical selected plans have gain zero.

Certified numerical inputs. The rate formulas at the start of this subsection define the certified sequences exactly, with rational amplitudes. Arb evaluates the square root, π, and cosine within interval arithmetic. The Float64 rate approximations guide optimization; they do not define the certified rates.

Output parameters use a different representation. The EMA decay is the exact rational represented by the shortest decimal string of the Float64 proposal. With this fixed $\widetilde{\rho}$, the certificate uses

$$w_0 = \widetilde{\rho}^T, \qquad w_t = (1 - \widetilde{\rho})\widetilde{\rho}^{T-t} \quad (1 \leq t \leq T).$$

For the selected $T = 256, d = 8$ EMA, $\widetilde{\rho} = 9785720620877001/10^{16}$. Thus $d = 8$ labels the proposal; it does not assert $\widetilde{\rho} = 2^{-1/32}$. Geometric weights likewise use the rational decay $\widetilde{r}$ and exact normalization of $(1, \widetilde{r}, \ldots, \widetilde{r}^T)$. Tail and uniform weights are exact rationals. Interval weights

use the rational endpoint masses and equal interior weights. Full-simplex and step-proportional proposals use exact rational normalization of the proposed decimal weights. Step-proportional weights need not be exactly proportional to the analytic learning rates.

Decimal representations of spectral atoms and masses are treated as exact rationals. Bias masses are normalized exactly; noise masses are rescaled when needed to satisfy both noise constraints before certification. The final batch arrays are exact integers. Table G.5 reports the resulting intervals for the 12 selected configured plans. The center full-simplex plan uses the certificate in Appendix G.4.1; the remaining selected plans and all 432 joint lower witnesses use the verification below. Gain and gap percentages use outward rounding. The center-row percentages in Table 1 also follow directly from its printed conservative endpoints.

Independent verification. Arb at 256 bits verifies every per-rate supporting-plane lower and all selected feasible upper bounds using the signed recurrences in Appendix F. The fixed upper cushions are 10^{-8} for the bias target and noise-dual intercept and are included in every reported ratio. Every batch vector is checked against its integer budget. An independent 320-bit implementation uses forward covariance recurrences for the upper covers and their gradients for the common lowers; it may subdivide a cover further without changing the risk target. Both implementations use Arb.

The additional joint-target exclusions are replayed at 384 bits with a separate derivative recurrence. The same forward-covariance calculation checks the selected upper certificates.

These are class-risk calculations with supplied budgets, not stochastic training experiments or a selection rule for unknown budgets.

G.4.1 Separation from the complete nondecreasing-weight class

Fix $T = 64, \bar{\Delta} = 1/4, q = 2, C = 8, L = D = \sigma_{\mathrm{ind}} = 1$ and the 72-rate library Γ from Appendix G.4. The selected full-simplex integer plan uses constant rate $31/16$, with

$$U_F = \frac{1877207160100019}{6250000000000000000} = 0.0030035314561600304. \tag{G.9}$$

The upper certificate and rational weights and batches are saved in `core_confirmation_20260924/T64_D0.25/upper_A1.9375/`. The same plan supplies the first-row full-simplex entries of Tables 1 and G.5, Corollary 1 and Figure 3.

Selected plan and reference witness. The selected plan has $w_0 \approx 0.051830$, $w_{63} \approx 0.078190 > w_{64} \approx 0.045232$, and integer batches between 1 and 8 summing to 128. In both output classes, $w_0 \geq 0$ participates in the unit-sum constraint but is exempt from the post-update ordering constraint. The post-update decrease places the output outside $\mathcal{C}_\uparrow$; the risk separation follows from the certificates below.

The last noise injection enters the output with coefficient $b_{\lambda,T-1}(w) = \gamma_{T-1} w_T$. Reducing the final weight lowers this contribution, while redistributing weight also changes the initialization and earlier-noise transfers. The certificate evaluates the resulting joint tradeoff. The rate $31/16$ permits negative transfer factors at high curvature. The selected witness does not determine whether the same output-class separation occurs when $0 < \gamma_t \leq 1/L$.

A distinct reference witness uses constant rate $15/8$, with $U_{\mathrm{ref}} = 61078788576018519/20000000000000000000$. That witness fixed the stopping threshold for the complete-family lower-bound search. Its unchanged weights and batches also support the three-load check in Appendix G.5. The selected $31/16$ plan comes from the prespecified upper search in that appendix. No additional optimization is needed for the comparison.

Continuous output and sampling coverage. Let $u^{(K)}$ be uniform on the last K post-update states. Every $w \in \mathcal{C}_\uparrow$ has the representation

$$w = w_0 e_0 + \sum_{s=1}^{T} (T - s + 1) d_s u^{(T-s+1)}, \qquad d_1 = w_1, \quad d_s = w_s - w_{s-1}\ (s \geq 2), \tag{G.10}$$

whose coefficients are nonnegative and sum to one. Initialized EMA has $w_0 = \beta^T$ and $w_t = (1-\beta)\beta^{T-t}$, so every $\beta \in [0, 1]$ belongs to this cover, including the two endpoint limits. Thus the

computation covers continuous parameters and deterministic mixtures on one rate sequence. No EMA parameter grid supplies the lower bound.

For a feasible common spectral adversary, let its joint convex risk be $\Phi(w, B)$. At a feasible anchor (w^a, B^a), let g_w, g_B be its gradients. A global lower bound is

$$\Phi(w^a, B^a) - \langle g_w, w^a\rangle - \langle g_B, B^a\rangle + \min\left\{ g_{w,0}, \min_{1\le K\le T} \frac{1}{K} \sum_{t=T-K+1}^{T} g_{w,t} \right\} + \min_{\substack{B\in[1,8]^T \\ \sum_j B_j = 128}} \langle g_B, B\rangle. \tag{G.11}$$

The last minimum assigns the 64 extra samples to the smallest coefficients, up to seven extras per update. Minimizing the resulting certified lower bounds over all 72 rate sequences gives $L_{\uparrow,72}$. Real batches enlarge the comparator class, so the lower bound also applies to every legal integer allocation. Signed or zero transfer factors do not affect the convexity or this supporting-plane argument.

All rate sequences and the stopping rule. Existing lower certificates for all admissible policies already exceed U_{ref} for 69 rate sequences. Only the three constant rate sequences with amplitudes $15/8, 31/16, 63/32$ require monotone joint programs. The frozen rule allows at most 36 spectral exchanges per remaining rate sequence and stops once its certified lower bound exceeds $1.001U_{\mathrm{ref}}$, or its real-batch search bracket reaches 0.5%. All three rate sequences meet the first criterion after one solve. Their lower bounds, rounded downward, are 0.003057313, 0.003067468, and 0.003113347. The smallest reused lower bound is 0.003058675, rounded downward, at constant rate $9/5$. Consequently the global lower bound is $0.003057313033237668\ldots$, attained among the certificates at constant rate $15/8$, and $U_{\mathrm{ref}} < L_{\uparrow,72}$. No fallback tail enumeration or continuous EMA subdivision is needed.

This closes the strict family comparison, not the family's optimal value. The three Float64 relaxed search brackets retain relative widths of 2.00%, 7.23%, and 24.17%, rounded upward. The early stopping rule does not require those widths to meet 0.5% after strict separation. The reference witness has certified reduction at least 0.1103%. The selected $31/16$ plan strengthens the reduction to at least 1.75%. The result excludes history-dependent nondecreasing post-update averaging weights, including adaptive choices of tails and initialized EMA. The certificates concern the prescribed finite rate set and the nondecreasing-weight constraint. They do not exclude every output with few parameters or establish that all weights need separate optimization. Earlier sensitivity studies on a restricted rate set do not establish robustness to arbitrary budget changes. Randomized rate selection remains outside the comparison.

Directed rounding and independent verification. The three constant-rate lower certificates use 320-bit Arb arithmetic and exact rational feasible adversaries. Reused lower certificates retain their original directed rounding. A separate 384-bit check propagates forward state and output covariances without dividing by contraction products, covering exact zero and negative factors. It checks every lower witness, all output and sampling constraints, and the reference witness's continuous curvature cover, with 8450 checked cells. Both verification runs pass.

Proof of Corollary 1. For 69 rate sequences, the saved certificates lower-bound the full policy class and hence every restricted history-dependent class. The smallest displayed lower bound among these is 0.003058675. For constant rates $15/8, 31/16, 63/32$, the remaining certificates give feasible finite spectral adversaries A_γ with $g_{A_\gamma}(w, B) \ge L_\gamma$ for all $w \in \mathcal{C}_\uparrow$ and all real capped allocations. Their downward-rounded bounds are 0.003057313, 0.003067468, and 0.003113347. Proposition C.3 transfers each bound to $V^{\mathrm{pred}}_{\gamma,\uparrow}$, using the all-coefficient Gram event in Appendix C.3. The three rates need not satisfy the smaller-rate assumption in Appendix E.1. Taking the minimum over the 72 deterministically selected rates and combining with the feasible-plan upper bound proves the corollary. The displayed rational endpoints imply $100(1 - U_F/L_{\uparrow,72}) > 1.75\%$. The 384-bit replay verifies all 72 lower certificates and the 65-vertex output representation. A separate forward-covariance replay checks 154994 cells for the selected plan's upper certificate and its fixed-plan lower witness. The adaptive conclusion also uses Proposition C.3.

G.5 PRESPECIFIED STRUCTURAL CONFIRMATION

We repeat the complete output-family comparison on the fixed grid $T \in \{56, 64, 72\}$, $\bar{\Delta} \in \{9/40, 1/4, 11/40\}$, with $L = D = \sigma_{\mathrm{ind}} = 1, q = 2, C = 8$. At each point, lower certificates cover all 72 rate sequences, the entire simplex of nondecreasing post-update weights with free w_0, and real capped batches. Proposition C.3 extends those bounds to history-dependent integer batches and outputs.

The upper search uses the three prespecified constant rates $15/8, 31/16, 63/32$, integer rounding and fixed-batch output refinement. The selected feasible plan uses $31/16$ at every point. Table G.6 reports strict separation at all nine points. The certified lower bounds on relative risk reduction range from 1.6397% to 1.8303% for the separately optimized plans. The sample budget changes with T, so absolute risks across horizons do not compare efficiency at fixed sample count.

T	$\bar{\Delta}$	Family lower bound	Integer-plan upper bound	Reduction $\geq$ (%)
56	$9/40$	0.003267901	0.003213295	1.6710
56	$1/4$	0.003481228	0.003424029	1.6430
56	$11/40$	0.003694039	0.003633331	1.6434
64	$9/40$	0.002870452	0.002817913	1.8303
64	$1/4$	0.003057313	0.003003532	1.7591
64	$11/40$	0.003243050	0.003187431	1.7150
72	$9/40$	0.002554114	0.002508873	1.7713
72	$1/4$	0.002721020	0.002674615	1.7054
72	$11/40$	0.002886790	0.002839453	1.6397

Table G.6: Prespecified structural confirmation with $L = D = \sigma_{\mathrm{ind}} = 1, q = 2, C = 8$. Each lower bound covers all 72 rate choices and the complete nondecreasing-weight policy class. The selected upper plans use constant rate $31/16$. Bounds and reductions are rounded outward.

A separate calculation uses a reference witness with rate $15/8$, and fixed rational weights and integer batches at $T = 64$. Table G.7 certifies separation at all three prespecified loads. The sufficient condition for a continuous load interval fails. The upper bound at load $11/40$, 0.003239761805, exceeds the lower bound at load $9/40$, 0.002870452399. The certificates establish separation only on the prespecified finite grid.

$\bar{\Delta}$	Family lower bound	Reference-plan upper bound	Reduction $\geq$ (%)
$9/40$	0.002870452	0.002868118	0.0813
$1/4$	0.003057313	0.003053940	0.1103
$11/40$	0.003243050	0.003239762	0.1013

Table G.7: The reference $T = 64$, constant-rate $15/8$ witness with unchanged weights and integer batches, evaluated at three prespecified loads. Only the load changes.

The supplement retains all 648 lower certificates, 30 integer upper records and their independent replay receipts. The verification script replays all lower certificates at 384-bit precision and checks the upper receipts against exact feasibility and file hashes. The saved upper receipts use independent forward-covariance replay. These numerical checks verify the stated inequalities; the adaptive extension relies on the analytic proposition.

G.6 SAMPLING CAPACITY AND FINITE CONSTRUCTION COSTS

G.6.1 THE JOINT CAP TRANSITION AT FIXED NOISE LOAD

This study examines the benefit of optimizing output and sampling together, and the saturation of that benefit. Before computation, we fixed $T \in \{1024, 2048\}$, $C \in \{2, 4, 8, 16, 64\}$, $L = D = \sigma_{\mathrm{ind}} = \bar{\Delta} = 1, c = 1/2, q = 2$. Every comparison has the same rates, updates, individual-sample budget and two covariance constraints. The uniform-batch values are reference comparisons; the earlier $T \leq 512$ cap study is exploratory, not independent confirmation of this boundary.

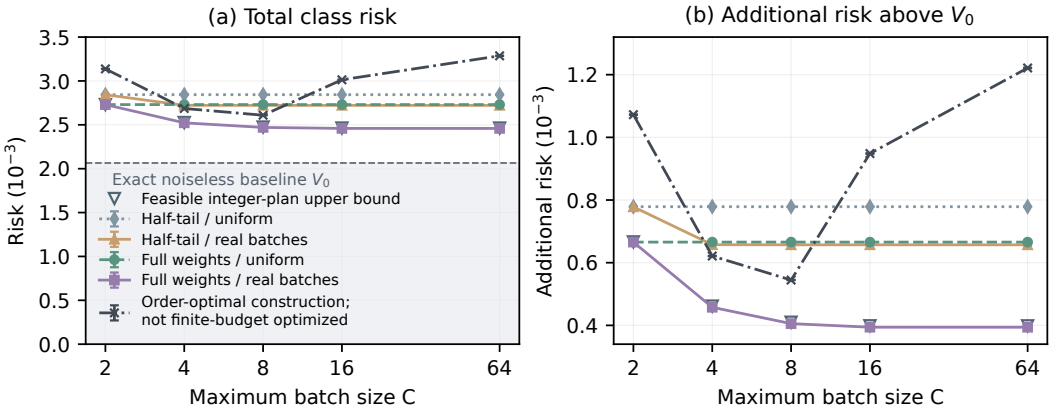


Figure G.1: Total risk and additional risk for $\gamma_t = 1/(2\sqrt{t+1})$, $T = 2048$, $q = 2$ and $L = D = \sigma_{\mathrm{ind}} = \bar{\Delta} = 1$. Both panels use the same saved bounds and exact noiseless baseline V_0. The curves compare half-tail or optimized convex weights with uniform or optimized real batches and the explicit order-attaining construction. Open triangles mark optimized integer-plan upper bounds. Bars are Float64 bounds with guards, not statistical intervals or Arb certificates. Appendix G.6.1 gives the full comparisons.

The three choices and the construction's constants. Figure 2 summarizes the cap comparison. The reference output is uniform half-tail on $\lfloor T/2 \rfloor, \ldots, T$, with uniform batches. We compare all convex averaging weights with uniform batches, real batches optimized with half-tail fixed, and all weights and real batches optimized together. The last value is E_T^{rel}, not the optimum over integer batch policies. Separate integer witnesses and policy brackets appear in Table G.8. Figure G.1 gives the extended output and sampling comparisons. The plotted bounds use Float64 guards, not Arb certificates or statistical intervals.

For finite constants, we evaluate the integer construction in (D.21), with latest-first completion and $K = \max\{1, \lceil T^{3/4}/\sqrt{C} \rceil\}$. We also fix a finite comparison family comprising the unique rounded-up values of 12 geometrically spaced lengths from 1 to $\lfloor T/2 \rfloor$, together with this explicit K. In each of the ten cases these are 13 distinct lengths. For every member, we optimize real batches at its fixed uniform tail output. All 130 members are retained. The reported best family value is enclosed by the minimum of their lower endpoints and the minimum of their upper endpoints; the selected integer witness rounds batches from the member with the smallest feasible upper endpoint. This finite family includes the endpoint. It is not a search over all possible tail lengths or a uniqueness claim about optimal weights.

Small additional risks require their own precision scale. We subtract the independently bracketed exact V_0, not a leading-order approximation. For every joint or fixed-tail convex problem, a feasible spectral adversary and the exact supporting-plane minimum over its strategy set give a lower bound. For the full program, the linear minimum splits into a simplex vertex and a capped budget allocation, as in (C.14). For a fixed tail, only the batch term varies. The continuous upper bound covers $[0, 1]$ with 65536 intervals, using positive-transfer monotonicity and the two-constraint dual envelope. The frozen rule permits doubling interval resolution only when the fixed-pair additional risk precision fails; spectral exchange adds current bias/noise witness atoms, with at most 12 exchanges.

The joint solver uses sparse transfer recurrences and quadratic-over-linear cones for individual injections. For fixed weights, the spectral dual is concave; for each dual measure its minimizing batches have the capped square-root allocation given by water filling. Neither solver's status nor its raw objective is a bound. Each accepted optimum and fixed-pair interval has positive additional risk lower endpoint and width at most 1% of it. An independent implementation propagates forward state/output covariances and computes gradients with prefix/suffix covariance sums. Across all 200 evaluated pairs and convex optima, the largest relative width is 0.990%, and the largest risk-component discrepancy is 7.81×10^{-18}. All 150 common adversaries and 50 integer pairs pass

their feasibility checks. The 10^{-11} guards are Float64 margins, not directed-rounding certificates or sampling confidence intervals.

Transferring numerical bounds to integer policies. Let $[L, U]$ enclose the relaxed additional risk, and let U_I be the smallest feasible integer upper witness among the rounded joint plan, rounded half-tail plan, rounded selected family member and explicit construction. Then both E_T^{pred} and E_T^{det} are in $[L, U_I]$. The correlated-rounding argument in Appendix D.3 can additionally bound the mixed and predictable value by $\mathrm{Bias}(w) - V_0 + (9/8)\mathcal{N}_{\Delta,B}(w)$. Here the evaluated deterministic witness is tighter, so the two policy columns have the same displayed brackets. This does not prove equality of their optima. The widest integer-policy bracket is 2.019% of its lower endpoint; it includes possible integrality cost and is distinct from the 1% numerical tolerance for continuous optima.

For the selected feasible integer pair itself, its worst-case risk minus V_0 is at most U_I, whereas each optimal policy additional risk is at least $L > 0$. Hence its additional risk competitive ratio to any of those classes is at most U_I/L. The maximum is 1.020187774 over all ten unrounded pairs; outward-rounded printed endpoints give at most 1.020306. The witnesses are chosen among the four already evaluated integer constructions, with no new optimization or selection across unreported cases. Their offline optimization uses the supplied class budgets, including the load.

Table G.8: Joint additional risk and executable-policy brackets, in units of 10^{-3}. Real-valued batches define the relaxed convex optimum; the other columns bracket the integer deterministic and predictable/mixed class values. The same exact noiseless value is subtracted throughout.

T	C	E_T^{rel}	$E_T^{\mathrm{pred}} = E_T^{\mathrm{mix}}$	E_T^{det}
1024	2	$[1.1147, 1.1174]$	$[1.1147, 1.1174]$	$[1.1147, 1.1174]$
1024	4	$[0.7978, 0.8044]$	$[0.7978, 0.8140]$	$[0.7978, 0.8140]$
1024	8	$[0.7331, 0.7357]$	$[0.7331, 0.7443]$	$[0.7331, 0.7443]$
1024	16	$[0.7200, 0.7222]$	$[0.7200, 0.7308]$	$[0.7200, 0.7308]$
1024	64	$[0.7200, 0.7222]$	$[0.7200, 0.7308]$	$[0.7200, 0.7308]$
2048	2	$[0.6648, 0.6666]$	$[0.6648, 0.6666]$	$[0.6648, 0.6666]$
2048	4	$[0.4571, 0.4583]$	$[0.4571, 0.4616]$	$[0.4571, 0.4616]$
2048	8	$[0.4048, 0.4063]$	$[0.4048, 0.4102]$	$[0.4048, 0.4102]$
2048	16	$[0.3935, 0.3948]$	$[0.3935, 0.3988]$	$[0.3935, 0.3988]$
2048	64	$[0.3933, 0.3947]$	$[0.3933, 0.3986]$	$[0.3933, 0.3986]$

Bounding the remaining total-risk improvement. Fix rates, class budgets and total samples. Let a feasible integer plan at cap $C \le C_{\max}$ have class risk R_C^{plan}, with $R_C^{\mathrm{plan}} - V_0 \le \overline{E}_C^{\mathrm{plan}}$. If $\underline{E}_{C_{\max}} \le E_T^{\mathrm{pred}}(C_{\max})$, then

$$0 \le \frac{R_C^{\mathrm{plan}}}{V_T^{\mathrm{pred}}(C_{\max})} - 1 \le \frac{\overline{E}_C^{\mathrm{plan}} - \underline{E}_{C_{\max}}}{V_0 + \underline{E}_{C_{\max}}} \le \frac{\overline{E}_C^{\mathrm{plan}} - \underline{E}_{C_{\max}}}{L_0 + \underline{E}_{C_{\max}}}. \tag{G.12}$$

Here $L_0 \le V_0$, with a positive denominator. The plan is also feasible at $C_{\max}$, which gives the first inequality. For the second, increase its numerator risk to $V_0 + \overline{E}_C^{\mathrm{plan}}$ and decrease the optimal denominator risk to $V_0 + \underline{E}_{C_{\max}}$. The numerator difference is nonnegative, so replacing V_0 by its lower endpoint gives the final bound. Thus a common adversary at the larger cap limits the improvement of every admissible output and feedback choice, beyond a particular smaller-cap plan.

The selected cap-16 integer plan and cap-64 lower witness at $T = 2048$ have the directed-rounding verification below. The ten-case Float64 comparisons remain in the supplementary data.

A rigorous check of the cap-16/cap-64 comparison. The $T = 2048$ cap comparison has a directed-rounding certificate. The witnesses are the selected cap-16 joint-rounded integer plan and cap-64 common spectral adversary; no optimization is repeated. Every rate is enclosed as the mathematical number $1/(2\sqrt{t+1})$, rather than treated as its Float64 approximation. Arb ball arithmetic controls rounding throughout the 256-bit computation; higher precision alone is not the certificate.

The proposed decimal weights are made exactly feasible by rational normalization. For the lower-bound tangent point, batches are clipped to $[1, 64]$ and the exact rational budget residual is redistributed

within that box. Bias masses are normalized to sum one, and noise masses are scaled down if needed to satisfy both spectral budgets. The integer upper plan's batches are unchanged. These rational feasibility corrections define the certified plans; no rounded solver objective is used as a bound.

For the lower bound, evaluate the common-adversary objective and both gradients with balls. The supporting-plane constant is enclosed by ball arithmetic. Lower each gradient coordinate to its exact dyadic lower endpoint, then minimize that linear function over the output simplex and capped batch-budget simplex. Since all strategy coordinates are nonnegative, this remains a valid supporting-plane lower bound over the complete relaxed strategy set. Theorem 2 transfers it to the predictable class.

For the integer upper plan, retain a feasible dual pair $\alpha, \beta \geq 0$, increasing the original β by 10^{-10}, and set a bias target 10^{-10} above its previous upper endpoint. Adaptive dyadic subdivision gives 2057 cells covering $[0, 1]$ without gaps. On every cell $[\ell, u]$, the verifier proves

$$\tfrac{1}{2} u a_\ell^2 \leq \text{bias target}, \qquad u^2 Q_B(\ell) \leq 2\ell\alpha + \beta. \tag{G.13}$$

Monotonicity of the positive transfers makes these continuum upper bounds. The dual objective $\alpha + \beta$ bounds the noise risk. A separate ball-validated sign bisection of $1 - 2\lambda \sum_t \gamma_t/(1 - \lambda\gamma_t)$ brackets the unique noiseless maximizer and V_0, checking the printed additional risk thresholds as well as the total-risk comparison.

Table G.9: One Arb-certified remaining-room comparison for the mathematical inverse-root rates, $T = 2048$. Decimal endpoints are rounded outward.

Certified quantity	Bound
Cap-16 integer-plan total risk	≤ 0.002463546486
Cap-64 relaxed class total risk	≥ 0.002458175774
Remaining ratio to cap-64 output/feedback optimum	$\leq 0.218484\%$

An independent 320-bit checker uses forward covariance identities for all 4097 gradient coordinates and every cell. It checks rational feasibility, the noiseless bracket and the final ratio. The two numerical implementations share the Arb library; neither formally verifies the general minimax theorem. The total-risk bounds in Table G.9 imply $R_{16}^{\mathrm{plan}}/V_T^{\mathrm{pred}}(64) - 1 \leq 0.218484\% < 0.23\%$, without relying on Float64 guards. The other nine configurations and the sweep-wide 2.019% additional risk claim retain their numerical-envelope qualification.

What the finite cap sweep establishes. At $C = 16$, $TE_T^{\mathrm{rel}}/\log T$ is enclosed by $[0.1063, 0.1067]$ for $T = 1024$ and $[0.1057, 0.1061]$ for $T = 2048$. From $C = 16$ to 64, the additional decrease in the true relaxed value is at most 0.290% and 0.360%, using monotonicity and the outward bounds. This is measurable saturation in the tested cap range. The asymptotic threshold $\sqrt{T}/(\log T)^2$ has parameter-dependent constants and onset; it is not used to predict a finite threshold of 16.

Sampling-only improvement at half-tail essentially ends by cap 4 in these cases, while joint optimization continues to reduce additional risk through cap 16. The remaining error thus cannot be assessed from sampling efficiency with one output held fixed. Conversely, output optimization at uniform batches misses the improvement that jointly redistributing samples permits. The selected joint outputs' mean lookbacks decrease from 182.3 to 13.5 at $T = 1024$, and from 352.5 to 31.7 at $T = 2048$, between caps 2 and 16. These describe feasible optimized weights; they are not uniform-window estimates.

The explicit order-attaining construction attains the optimal order but loses finite efficiency as cap becomes large. Its additional risk at $C = 64$ is 2.84–3.11 times the relaxed optimum. By contrast, the best of the 13 frozen tail lengths, with sampling optimized, is within 1.4% of the full program there. Both choose the endpoint at cap 64 within this finite family, although the full convex program still has a positive mean lookback. Thus neither the theorem nor these data identify a universal finite window or make the explicit plan a practical selector. They do distinguish an output-family limitation from the finite cost of its allocation construction.

A crossed diagnosis of the explicit allocation's finite cost. At $T = 2048$, the selected cap-8 and cap-64 rules have windows of 108 and 39 states. Denote their weights and allocations by (w_8, B_8)

and (w_{64}, B_{64}). We reuse these two diagonals and compute only (w_8, B_{64}) and (w_{64}, B_8). All four pairs are evaluated at cap 64, with the same 4096 individual samples, rates and unit class budgets. A pair using B_{64} is not asserted feasible at cap 8. These are post-hoc fixed-pair calculations, with no new trajectories or joint optimization. Forward state/output covariances independently check the continuum envelopes; the bounds retain Float64 guards.

Table G.10: Post-hoc diagnosis of the order-attaining allocation at $T = 2048$, in units of 10^{-3}. All four combinations use cap 64 and 4096 samples. The two crossed pairs isolate the effects of changing the window and allocation. Bounds retain Float64 guards.

Window	Allocation	Initialization above V_0	Noise	Total additional risk
w_8	B_8	$[0.02804, 0.02879]$	$[0.51605, 0.51639]$	$[0.54409, 0.54517]$
w_8	B_{64}	$[0.02804, 0.02879]$	$[1.37066, 1.37163]$	$[1.39870, 1.40041]$
w_{64}	B_8	$[0.00980, 0.01054]$	$[0.52200, 0.52234]$	$[0.53181, 0.53288]$
w_{64}	B_{64}	$[0.00980, 0.01054]$	$[1.21075, 1.21161]$	$[1.22056, 1.22215]$

Table G.10 shows that switching from B_8 to B_{64} raises additional risk at either output, whereas shortening the window lowers additional risk with either allocation. The allocation change explains the worsening in this control. The cap-64 plan puts 1865 samples in the final 39 updates, versus 312 for its cap-8 plan, leaving fewer samples for earlier injections that still affect the output. This diagnoses the combined allocation construction; it does not isolate the initial harmonic allocation from its latest-first budget completion. Larger caps retain every smaller-cap plan, so the optimal class value cannot increase. Retaining a smaller-cap plan prevents this worsening, but selection by class risk still uses the supplied budgets and gives no load-uniform guarantee.

G.6.2 Finite constants in the explicit construction

Proposition 1 separates the initialization, cap/window and total-budget terms of the finite-horizon bound. The supplement's `EVIDENCE_INDEX.md`, under *Finite constant decompositions*, gives all ten prespecified cap settings using the windows $K = \lceil T^{3/4}/\sqrt{C} \rceil$. Every window satisfies $1 \le K \le \lfloor T/2 \rfloor$. The analytic terms were recomputed with Arb arithmetic.

At $T = 2048, C = 64, K = 39$, the three terms are 0.0003221482, 0.0016025641 and 0.0887705400. Their sum, 0.0906952523, is about 74 times the explicit plan's additional risk, whose guarded Float64 enclosure is $[0.0012205668, 0.0012221470]$. The harmonic term accounts for most of the difference. The analytic guarantee explains the asymptotic window tradeoff, but its conservative constants do not establish finite-budget near-optimality. The separate risk certificates answer that question.

G.6.3 Uniform-batch reference and frozen confirmation

This frozen study supplies the uniform-batch reference for the cap comparisons. The fixed slice is $L = D = \sigma_{\mathrm{ind}} = 1, c = 1/2, q = C = 2$. With no allocation freedom, (C.4) is precisely the convex program for the full adaptive-output and oracle class value. The exploratory five $\bar{\Delta} = 1$ points at $T = 32, 64, 128, 256, 512$ suggested the candidate exponent and are exploratory. The confirmation design fixes $T = 1024, 2048$, $\bar{\Delta} = 0.25, 1, 4$, the endpoint, half-tail, and $K = \lceil T^{3/4} \rceil$ tail before running these six configurations. The coefficient in K is one at every budget; it is not fitted to either the exploratory or confirmation values.

A precision target for the smaller quantity. Each reported class interval $[L_\Delta, U_\Delta]$ combines a feasible common spectral adversary with a feasible deterministic output. The lower bound subtracts the simplex linearization gap of the common quadratic objective. The upper bound uses 65536 intervals covering $[0, 1]$, with monotone positive transfers and the same two-moment dual envelope as in Appendix F. The exact noiseless reference is enclosed independently. We require

$$E^- := L_\Delta - U_0 > 0, \qquad \frac{(U_\Delta - L_0) - E^-}{E^-} \le 0.01. \tag{G.14}$$

The same precision criterion applies to each named fixed output's additional risk. All six class values and all 24 output comparisons pass; the largest relative width is 0.556%. These are Float64 enclosures

with explicit 10^{-11} risk guards, not formal directed-rounding certificates. No sampling confidence level is attached.

The SOCP introduces the transfer recurrence $h_{\lambda,j} = w_{j+1} + (1 - \lambda\gamma_{j+1})h_{\lambda,j+1}$ and one noise norm constraint per curvature. This avoids dense time-by-time transfer matrices. The initial curvature set has 20 geometric points in $[10^{-3}, 1]$, the exact noiseless maximizer, and the covariance-budget kink $\min\{1, 1/(2\bar{\Delta})\}$. The frozen exchange rule adds the current bias/noise witness atoms until (G.14) holds, with a maximum of ten exchanges. Five cases require no exchange; $T = 2048, \bar{\Delta} = 0.25$ requires one. The common lower witness and continuum upper envelope, rather than the solver's status or raw objective, determine acceptance. Independent forward state/output moments and prefix/suffix covariance sums check every spectral objective, linearization gap and interval dual inequality. Their largest risk-component discrepancy is 1.13×10^{-17}.

Table G.11: Prespecified fixed-batch class comparisons. The scaled additional risk concerns the full adaptive-output optimum. Total-risk improvement is relative to the same-sequence half-tail. Intervals are numerical enclosures, with outward decimal rounding.

T	$\bar{\Delta}$	K	$T^{3/4}E_T$	E_T/V_0	Half-tail gain (%)
1024	0.25	182	$[0.0934, 0.0937]$	$[0.175, 0.177]$	$[5.46, 5.54]$
1024	1	182	$[0.2017, 0.2023]$	$[0.379, 0.381]$	$[4.63, 4.75]$
1024	4	182	$[0.3886, 0.3908]$	$[0.730, 0.735]$	$[8.10, 8.37]$
2048	0.25	305	$[0.0947, 0.0951]$	$[0.150, 0.152]$	$[5.79, 5.87]$
2048	1	305	$[0.2024, 0.2030]$	$[0.322, 0.323]$	$[3.92, 4.02]$
2048	4	305	$[0.3895, 0.3907]$	$[0.619, 0.622]$	$[7.15, 7.29]$

Magnitude and output tradeoffs. The optimized class additional risk is about 15%–74% of the exact noiseless value over these six configurations. Relative to half-tail, optimizing the full output class reduces total class risk by 3.92%–8.37%, with the per-case enclosures in Table G.11. The confirmation $\bar{\Delta} = 1$ scaled values continue the exploratory pattern, while the other loads have distinct constants. Two confirmation horizons do not identify a limiting constant or establish the asymptotic rate independently of Theorem 3.

The supplement's `EVIDENCE_INDEX.md`, under *Uniform-batch output decompositions*, retains the 24 output decompositions and mean lookbacks. At $\bar{\Delta} = 1$, the $K = \lceil T^{3/4} \rceil$ tail has larger finite risk than half-tail at both confirmation horizons. At the other two loads it has smaller risk. The theorem gives an asymptotic scale with parameter-dependent constants and onset, not finite-budget dominance of prefactor one.

H Related information and model assumptions

H.1 Comparison with classical adaptive information

Wasilkowski & Woźniakowski (1984, Theorems 2.1 and 5.1) compare adaptive and nonadaptive inner-product information for a continuous linear operator between Hilbert spaces, under an orthogonally invariant input measure. They minimize average squared reconstruction error over reconstruction maps. Our output is instead constrained to the realized SGD path's scalar convex hull, even if the output oracle knows the population objective. The risk averages over samples and takes a supremum over the allowed Hessians, initializations, covariance budgets and finite dimensions. Batch sizes are integers chosen before each fresh batch and must satisfy the cap and sample budget on every run.

Theorems 1 and 2 use a common spectral adversary and noise replication. One Gram event controls all convex weights and legal consecutive sample partitions. The resulting lower bound therefore covers choices based on observed samples and compares feedback with independent randomization of complete integer plans. The classical information result alone does not provide the output constraint, spectral minimax quantifiers or integer-budget comparison required for these equalities.

H.2 Related assumptions and policy classes

Luner & Grimmer (2025, Theorem 3.1) optimize all fixed convex averages for smooth convex GD under an endpoint-risk condition. Their common Huber instance already provides a lower bound for every fixed average. Our additional step controls weights and sample partitions chosen from the same stochastic trajectory on which loss is evaluated.

Hagedorn & Jarre (2024) derive finite-step errors for strongly convex stochastic quadratics with IID additive noise and constant batches, and tune rates together with polynomial weights. In our notation, their averaging weights are $w_0 = 0$ and $w_t = t^\beta / \sum_{j=1}^{T} j^\beta$, with $\beta \geq 0$. These weights belong to $\mathcal{C}_\uparrow$. Corollary 1 separates a feasible plan from all admissible policies in that class over the specified 72 rate sequences at $T = 64, \bar{\Delta} = 1/4$. The comparison uses our instance class and budgets; it does not establish dominance in their model or full rate-tuning range. Wei et al. (2026) study weighted-average central limit theorems and adaptive weighting with nonasymptotic guarantees. Our question is the exact minimax value over all history-dependent scalar convex outputs.

The fixed Hessian and additive IID noise make the transfer coefficients deterministic. Random sample Hessians instead introduce state-dependent noise $(H_\xi - H)x_t$, as in least-squares analyses of minibatching and tail averaging (Jain et al., 2018; Mücke et al., 2019). Fixed-plan upper formulas require only the stated second moments; Gaussian replication supplies the lower instances. Enlarging the model preserves lower bounds but need not preserve matching upper bounds. For $\Sigma = \tau^2 H$, the load is $\Delta_\star = \tau^2 \operatorname{rank}(H)/2$, so fixed total load differs from fixed noise per direction as rank grows.

Batch and learning-rate coupling is established (Balles et al., 2017). Wang et al. (2026) optimize sampling and training duration under a data budget; Li et al. (2026) optimize batches for prescribed rates using a continuous endpoint-loss functional. We include all history-dependent convex outputs and integer batching, characterize the value of feedback relative to randomized plans, and measure additional risk above the exact noiseless minimax value for the same rates. Tail-averaging tradeoffs are established (Mücke et al., 2019; Zou et al., 2023); the joint cap and sample-budget boundary concerns this finer risk difference.

Results on last iterates, random reshuffling and anytime methods (Ge et al., 2019; Cha et al., 2023; Kornowski & Shamir, 2026) impose different output, sampling or objective assumptions. Controllers that change the updates (Au & Block, 2026) are outside our prescribed-rate comparison. Improvements from adaptive weighting on favorable instances (Bhaskara et al., 2026) are compatible with unchanged worst-case values. Regression source/capacity assumptions (Meterez et al., 2026a) are not nested with the stated quadratic class.

I Local validation in language-model training

The current tests fix averaging weights and sampling allocations before drawing future training streams. The local model predicts output-dependent sampling effects; the experiments measure those effects in nonlinear LM evaluation loss. Earlier diagnostics and failures are summarized after the current results, with full records in the supplement.

I.1 Shared protocol and local model

Model and training. The causal Transformer has 12 layers, width 768, 12 attention heads, feed-forward width 2048, vocabulary size 50257 and context length 512. The Llama-style architecture ties input and output embeddings and has 123551232 unique parameters. All parameters are trained by plain SGD with identity preconditioning, zero momentum, dampening and weight decay, and no clipping, dropout or gradient normalization. Each pretraining update uses eight contexts with 512 target tokens each. Each initialization runs for 131072 updates, or 536870912 processed training tokens, on one GPU. For zero-based global update k,

$$\gamma_k = 0.1 \begin{cases} (k+1)/8192, & k < 8192, \\ 0.01 + 0.99\{1 + \cos[\pi(k - 8192)/122879]\}/2, & k \geq 8192. \end{cases}$$

A bounded pilot on initialization 1 selected the peak rate before any averaging result. Initializations 2 and 3 reuse the schedule without retuning. Their states after update 114688, or 469762048

tokens, are Anchor 2 and Anchor 3. The 128-update continuation uses the prescribed rate slice from 0.005279053746931335 to 0.0052139230912116065, without restart or warmup. Each path consumes 1024 contexts, or 524288 target tokens, with integer batches between 5 and 32. The earlier anchor at update 98304 was not used to select a better confirmation result.

Data and evaluation reuse. We use SlimPajama-6B (DKYoon, 2023) and the `openai-community/gpt2` tokenizer. Their respective revisions are `b5f90f419b7489cdba26fdbc8c022fcb5562f968` and `607a30d783dfa663caf39e06633721c8d4cfcd7e`. Documents are separated by EOS token 50256. Each context starts at a 512-token offset and includes one lookahead token. Training samples are drawn uniformly with replacement from 3060000 contexts. The token budget counts processed tokens, not distinct documents.

The prospective test constructs 16384 distinct evaluation contexts from the first training shard, unused in preceding experiments. Exact tokenized-document and target-context overlap with the existing training and evaluation pools is excluded in source order, without model evaluation or CE filtering. The retained contexts cover 8181 documents and have no exact duplicates; semantic near duplicates are not excluded. This is a held-out part of the train split, not the official test split. The four-stream and sixteen-stream studies, sampling controls and repeated calibration all reuse this set. Earlier diagnostics used a 2048-context development set and a separate 18827-context evaluation set that had already been inspected. Every reported interval conditions on the stated evaluation set.

Outputs and fixed curvature. Tail-64 and Tail-32 average states 65–128 and 97–128. EMA-16 uses $m_0 = \theta_{\mathrm{ref}}$ and $m_t = 2^{-1/16} m_{t-1} + (1 - 2^{-1/16})\theta_t$, retaining anchor mass 2^{-8} at update 128. Let $g_{\mathcal{B}} = \nabla \ell_{\mathcal{B}}(\theta_{\mathrm{ref}})$, where $\ell_{\mathcal{B}}$ averages token CE over a batch. The fixed generalized Gauss–Newton (GGN) operator $\widehat{G}$ averages 256 calibration contexts at the anchor. For logits z_θ and softmax probabilities p, each token contributes

$$(\nabla_\theta z)^\top [\mathrm{diag}(p) - pp^\top](\nabla_\theta z).$$

The operator is positive semidefinite. Its action is computed by logit JVP/VJP without forming a dense matrix; omitted logit second derivatives distinguish GGN from the parameter Hessian. The 256 context indices are reused across anchors, but the operator is recomputed at each anchor. The fixed-curvature recurrence is

$$\delta_0 = 0, \qquad \delta_{t+1} = \delta_t - \gamma_t(g_{\mathcal{B}_t} + \widehat{G}\delta_t). \tag{I.1}$$

The anchor need not be stationary. This local model does not verify constant population curvature or the minimax instance assumptions for the nonlinear LM.

Shared local model and four sampling allocations. For the fixed GGN above and convex averaging weights, the response to a gradient perturbation at update j is

$$K_{j,w}(G) = \gamma_j \sum_{t>j} w_t \prod_{k=j+1}^{t-1} (I - \gamma_k G). \tag{I.2}$$

For zero-mean single-context gradient noise z, define $v_j(w) = \frac{1}{2}\mathbb{E}[(K_{j,w}(\widehat{G})z)^\top \widehat{G} K_{j,w}(\widehat{G})z]$. The expected local objective is $\beta(w) + \sum_j v_j(w)/B_j$, where the mean-trajectory term $\beta(w)$ is independent of batching. The fitted coefficients $\widehat{v}_j(w)$ replace the expectation with independent gradient-difference samples. The operator and coefficients are fixed model quantities; no spectral adversary is optimized.

The calibrated B_{64}, B_{32} allocations minimize the fitted noise score for their respective outputs (Appendix I.3). For $k \in \{65, 97\}$ and updates $t = 1, \ldots, 128$, the triangular allocations use

$$h_t(k) = \begin{cases} (t-1)/(k-1), & t \le k, \\ (128-t)/(128-k), & t > k, \end{cases} \qquad \widetilde{B}_t = 5 + \frac{384 h_t(k)}{\sum_v h_v(k)}.$$

Rational arithmetic determines the continuous values. We take floors and allocate the remaining samples by decreasing fractional part, with earlier updates breaking ties. Both allocations obey the

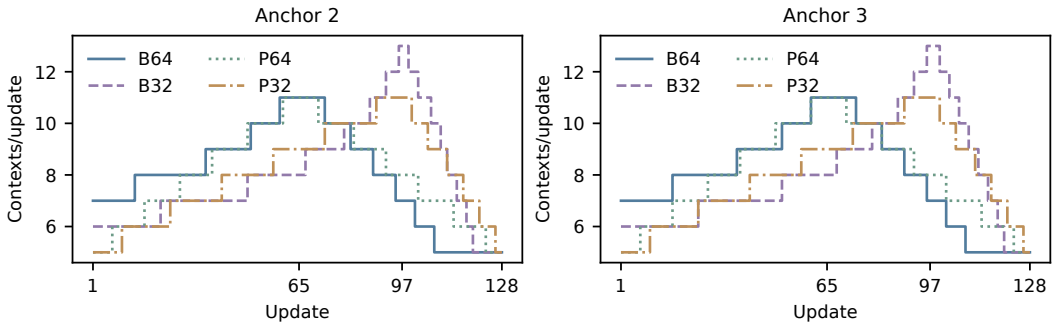


Figure I.1: All four frozen sampling allocations. The first averaged states are 65 for Tail-64 and 97 for Tail-32.

integer budget and cap. P_{64} is also the triangular allocation used in the sampling controls. The allocations do not use gradients or evaluation outcomes.

Fit and audit each use 16 independent gradient-difference directions, with curvature shared between fit and audit. The primary prediction uses audit coefficients; no GGN or allocation is refitted using audit data. For either allocation family, let $D_{64} = F_{64,32} - F_{64,64}$ and $D_{32} = F_{32,64} - F_{32,32}$. Their sum is the interaction I in Section 6. Prediction units are local training quadratic risk, while LM units are nonlinear evaluation CE. Agreement between the objectives is measured empirically.

For the calibrated family, the local interaction is

$$I_{\text{local}} = \sum_j \big(v_j(w_{64}) - v_j(w_{32})\big) \left(\frac{1}{B_{32,j}} - \frac{1}{B_{64,j}}\right). \tag{I.3}$$

The mean-trajectory term cancels because the output is held fixed within each sampling comparison. A training stream is the paired statistical unit; two outputs from one path are not independent samples. The sixteen-stream cohort excludes the earlier four-stream observations. Repeated calibration reuses two control streams per fit.

I.2 Sixteen-stream independent confirmation

The confirmation uses 16 independent streams at each of the two existing pretrained anchors. It reuses the four allocations, learning rates, averaging weights, curvature estimates, fit and audit spectra, and 16384-context evaluation set from Appendix I.1. There is no new calibration, pretraining initialization or evaluation set. Observations from the four-stream study are excluded from this analysis.

Design and paired sampling. A root seed and PCG64 child streams are fixed before generating the ordered arrays. Each stream contains 1024 contexts sampled with replacement from 3060000 training contexts, with 512 targets per context. The four allocations restart from the same anchor and share the same ordered stream. They change only its grouping into 128 updates. The two returned outputs average states 65–128 and 97–128. Each stream therefore yields eight evaluation cells. Allocation order and device assignment are fixed before execution. The 32 ordered arrays differ from each other and from the eight four-stream study arrays.

Primary inference and frozen prediction accuracy. The four interactions in Table 2 use paired stream means and individual 98.75% Student t_{15} intervals, Bonferroni-adjusted for four comparisons to target 95% simultaneous coverage. The coverage relies on the Student approximation for stream means. All four lower endpoints are positive. The same intervals test the prespecified half-gain criterion by strict containment in $(2\widehat{I}/3, 2\widehat{I})$, for the fixed positive audit prediction $\widehat{I}$. All four pass. This rule was fixed before the sixteen-stream cohort, unlike the four-stream post-hoc containment analysis in Appendix I.5.

All four frozen audit predictions lie below their primary intervals and underestimate the observed interaction means by 12.59%–20.40%. The predictions meet the prespecified relative-error tolerance

despite the underprediction. No scale correction or refitting is applied. Table I.1 retains the fixed predictions and the prespecified sensitivity analysis with 20000 whole-stream bootstrap resamples. Each resample selects all eight cells together within an anchor. Bootstrap intervals supplement the primary Student intervals and do not replace them.

Anchor	Allocation	Prediction	Acceptance interval	Bootstrap interval	Error (%)
2	Calibrated	21.680	[14.453, 43.360]	[22.835, 26.931]	12.59
2	Triangular	11.518	[7.679, 23.036]	[12.505, 14.694]	15.38
3	Calibrated	20.168	[13.446, 40.337]	[23.038, 26.986]	18.96
3	Triangular	10.856	[7.237, 21.712]	[12.101, 15.530]	20.40

Table I.1: Frozen predictions, the prespecified half-gain acceptance intervals and whole-stream bootstrap sensitivity intervals. CE units are 10^{-5} nats/token. Each primary interval in Table 2 lies strictly inside its acceptance interval. Bootstrap intervals use level 98.75%; relative absolute error divides by the observed interaction mean.

Components and direct allocation comparisons. Table I.2 reports the two component contrasts with descriptive 95% intervals. Direct comparisons instead match the output and compare calibrated with triangular allocation. Write $\Delta_{64} = F_{64,P_{64}} - F_{64,B_{64}}$ and $\Delta_{32} = F_{32,P_{32}} - F_{32,B_{32}}$, where P denotes a triangular allocation and B a calibrated allocation. Their positive means favor calibrated allocation. All four descriptive 95% intervals are positive (Table I.3); these secondary results do not define a further familywise success criterion.

Anchor	Allocation	D_{64} [95% interval]	D_{32} [95% interval]
2	Calibrated	9.838 [8.125, 11.552]	14.964 [12.437, 17.492]
2	Triangular	6.532 [5.507, 7.556]	7.079 [5.499, 8.659]
3	Calibrated	8.859 [7.953, 9.766]	16.027 [14.272, 17.782]
3	Triangular	5.393 [4.789, 5.997]	8.245 [6.870, 9.621]

Table I.2: Secondary component contrasts in the independent cohort, in 10^{-5} nats/token. Intervals are descriptive Student t_{15} intervals.

Anchor	Matched output	Mean [95% interval]
2	Δ_{64}	2.524 [1.849, 3.200]
2	Δ_{32}	1.453 [0.852, 2.055]
3	Δ_{64}	1.727 [1.268, 2.185]
3	Δ_{32}	1.103 [0.611, 1.595]

Table I.3: Direct calibrated-versus-triangular comparisons in 10^{-5} nats/token. Positive values favor calibrated allocation. The four intervals are descriptive and do not define another familywise success criterion.

The intervals quantify training-stream variation conditional on the two anchors, frozen fits and reused evaluation set. They exclude uncertainty from refitting or new pretraining runs. Direct gains do not establish end-to-end efficiency including calibration costs. Appendix I.6 gives implementation checks and aggregate costs.

I.3 Prospective fixed-output sampling test

The prospective experiment fixes the output before selecting its sampling allocation. Tail-64 is fixed to states 65–128 of the same 128-update window. The mean term $\beta(w)$ cancels between allocations (Appendix I.1); no mean gradient is estimated. The independent calibration, frozen choices and independent confirmation samples make this a prospective sampling test, unlike the matched-path forecasts. The two pretrained anchors are fixed checkpoints, not previously untouched tasks or eight independent pretraining initializations.

Calibration and integer allocation. At each anchor, 16 independent context pairs define $z_\ell = (g_{a_\ell} - g_{b_\ell})/\sqrt{2}$. Using the fixed 256-context GGN and the transfer operator in (I.2), estimate

$$\widehat{v}_j(w) = \frac{1}{32}\sum_{\ell=1}^{16}(K_{j,w}(\widehat{G}_{256})z_\ell)^\top \widehat{G}_{256}(K_{j,w}(\widehat{G}_{256})z_\ell), \qquad B_{\mathrm{cal}} \in \arg\min_{\substack{B_j\in\{5,\dots,32\}\\ \sum_j B_j=1024}} \sum_j \frac{\widehat{v}_j(w)}{B_j}. \tag{I.4}$$

The gradient differences estimate single-context covariance, not batch-8 covariance. All gradients and GGN products are evaluated at the anchor. Starting with every $B_j = 5$, assign each of the remaining 384 samples to the uncapped position with largest $\widehat{v}_j(w)/[B_j(B_j+1)]$, resolving ties by index. Decreasing marginal gains make this exact for the estimated separable integer objective, not for the LM risk or the full minimax class.

The simple comparator minimizes the same fit score among Uniform8, B_{test} defined below, its time reversal, and the two allocations with batch 11 in one half and 5 in the other. All have the same budget and cap. Both confirmation anchors select Uniform8. Other templates are not all executed on the new LM streams, so the comparison is against a calibration-selected template, not the true-CE-best member of the five-template family. A further 16 gradient pairs per anchor audit the predictions without refitting the output, allocation or curvature. Every noise direction uses 32-step Lanczos quadrature and a 16-step check. The largest relative coefficient discrepancy over all 96 directions is 3.46×10^{-12}, below the frozen 10^{-3} tolerance. The normalization is the maximum depth-32 coefficient of that direction; this is a numerical diagnostic, not a population spectral certificate.

The development batch-5 stress path passes the prespecified numerical stability and predicted-gain checks. At the confirmation anchors, only numerical failure could stop the protocol; an unfavorable predicted gain could not remove an anchor or stream. Neither selected allocation reaches cap 32, so the test does not examine cap saturation.

The comparator B_{test} is the unchanged development-anchor endpoint allocation from Appendix J. The supplement's `EVIDENCE_INDEX.md`, under `B_test`, identifies its exact 128-entry integer array. The array has minimum 5, maximum 23 and sum 1024; the prospective test reuses its entries without refitting.

Frozen choices and paired samples. Fit and audit independently draw from the training pool with replacement, using PCG64 seed sequences $(2026092191, s, 0)$ and $(2026092191, s, 1)$ at initialization s. No disjointness is enforced. All averaging weights, allocations, predictions and analysis rules are fixed before generating the confirmation seed and eight ordered streams, four per anchor. The selector never reads future samples or losses. Each allocation regroups the same ordered 1024 contexts from the same anchor under update-indexed rates. Uniform and the selected simple allocation coincide and share a new path, giving 16 unique paths. No historical uniform path is reused. Evaluation uses the held-out 16384-context set described in Appendix I.1. The supplementary manifests retain seeds, sample indices and frozen arrays.

Frozen comparison and uncertainty. The primary difference is $\Delta = F_{\mathrm{eval}}(\text{Tail-64-}B_{\mathrm{simple}}) - F_{\mathrm{eval}}(\text{Tail-64-}B_{\mathrm{cal}})$. Within each initialization, all four paired gains and their mean are reported. The primary family contains the two initialization-specific means. Each receives a two-sided 97.5% Student-t interval with 3 degrees of freedom, giving the prescribed Bonferroni correction. These intervals are conditional on the existing anchor and fitted allocation; four streams provide limited information and require a distributional approximation. They are not intervals over new pretraining tasks or repeated calibration draws.

For fixed returned models, a separate paired block bootstrap uses 128 consecutive contexts per block, 128 blocks, 20000 resamples and PCG64 seed 2026092192. The four-stream mean contrast is averaged contextwise before resampling blocks. Its 97.5% percentile interval measures conditional evaluation uncertainty. Single-stream 95% evaluation intervals are descriptive. The frozen success condition requires both initializations to have positive lower bounds for the training-stream and evaluation intervals, at least three positive streams out of four, and a mean gain exceeding ten times the largest paired microbatch-4/8 mean shift on the first 128 evaluation contexts. This criterion is separate from the earlier fixed-GGN (256) half-gain accuracy test.

Table I.4: Prospective predictions and four-stream results, in 10^{-5} nats/token. Fit/audit predictions concern the local training quadratic; the LM mean and intervals concern nonlinear held-out CE. Training-stream and conditional evaluation intervals quantify different uncertainties.

Anchor	Fit prediction	Audit prediction	LM mean	Stream 97.5%	Eval. 97.5%
2	7.941	7.162	7.633	[2.645, 12.621]	[6.400, 9.038]
3	6.936	6.552	9.676	[6.008, 13.345]	[8.666, 10.746]

Every primary gain is positive, and both initializations pass every frozen condition. The maximum paired numerical shifts are 2.29×10^{-8} and 1.22×10^{-8} nats/token. These are empirical diagnostics, not upper bounds on all floating-point error. Secondary four-stream improvements over Uniform EMA-16 are 7.0845×10^{-5} and 7.5083×10^{-5}; those over the uniform endpoint are 5.2468×10^{-4} and 5.2796×10^{-4}. The endpoint comparison includes averaging and does not isolate sampling. No secondary result replaces the primary comparison. The prospective test supports a fixed-output sampling decision. The separate sampling controls study adds low-cost controls and crosses two prescribed outputs and allocations (Appendix I.4).

I.4 Sampling controls and repeated calibration

The earlier controls use four independent future streams at each confirmation anchor, with the shared 128-update protocol and evaluation set. Rates, outputs, allocations and analysis rules are frozen before drawing streams. Each allocation regroups the same ordered 1024 contexts. The controls compare Uniform8, the target-anchor Tail-64 fit B_{64}, the fixed triangular allocation P_{64}, and the development-anchor Tail-64 allocation B_{dev}. The transferred allocation retains the original development calibration cost. The crossed comparison also uses B_{32} and both Tail-64 and Tail-32 outputs.

Let $G_b = F_{64,U} - F_{64,b}$ and $D_{\mathrm{peak}} = G_{B_{64}} - G_{P_{64}}$. All 24 stream gains over Uniform8 are positive. The two primary D_{peak} intervals contain zero, so the controls do not resolve an incremental gain from target-anchor fitting over triangular allocation. The transferred allocation has a slightly larger mean gain at both anchors. Its descriptive difference from the target-anchor fit is unresolved at Anchor 2 and favors transfer at Anchor 3. For the calibrated crossed outputs, both mean interactions are positive, although Anchor 2's descriptive D_{64} interval contains zero.

The original control and interaction families each use two-sided 97.5% Student t_3 intervals for two anchor-specific means. Table I.5 retains the supplementary 98.75% intervals for all four comparisons together. Secondary intervals are descriptive. Conditional evaluation intervals use 20000 paired resamples of 128-context blocks and hold the returned models fixed. They do not replace training-stream intervals; a positive evaluation interval at Anchor 3 does not resolve its primary control comparison.

Table I.5: Four primary sampling-control means in 10^{-5} nats/token, with supplementary Bonferroni intervals for the four-comparison family.

Anchor	Primary contrast	Mean	98.75% stream interval
2	D_{peak}	0.8784	[-4.6848, 6.4416]
2	I	25.4016	[22.3757, 28.4274]
3	D_{peak}	1.1346	[-2.4331, 4.7022]
3	I	28.6691	[24.9617, 32.3765]

Repeated-calibration sensitivity. Two further fits per anchor independently draw 256 curvature contexts and 16 gradient pairs. Each fit determines Tail-64 and Tail-32 allocations, then reuses control streams 1 and 2. There is no new independent audit or future stream. Across the four fits, two-stream mean Tail-64 gains lie between 5.9481 and 7.9397, and interactions between 24.0383 and 28.5071, in 10^{-5} nats/token. All means are positive. The descriptive 95% Student t_1 intervals use only two reused streams per fit; the fits provide sensitivity checks, not additional independent samples or a population interval over tasks. Full per-stream contrasts, intervals and calibration records remain in the supplement.

I.5 Earlier diagnostics and limitations

Four-stream prediction study. The earlier four-stream cohort uses the same anchors, allocations and frozen predictions as the sixteen-stream confirmation. Its original accuracy procedure passes only two of four predictions under the frozen 12-contrast correction. A later confidence-set containment analysis passes four of four, but is post-hoc. The sixteen-stream study fixes the containment criterion before its independent cohort and does not pool the earlier observations. The original failed tests remain in the supplement alongside the reanalysis.

Development and matched-path diagnostics. Initialization 1 selects the 128-update, 256-context fixed-GGN model after comparing curvature approximations and window lengths. The frozen-gradient model fails, and a permitted increase from 128 to 256 curvature contexts follows a development discrepancy. These choices are development outcomes. Matched-path forecasts use the LM's realized future samples and therefore differ from prospective predictions. On the two confirmation anchors, fixed-GGN predicts the resolved averaging and sampling gains, but fails two endpoint context-loss accuracy checks. One of four sampling accuracy checks on the reused development set is unresolved. The selected windows also fail to resolve the denominator for progress-normalized accuracy. Relative-effect agreement therefore does not establish whole-trajectory accuracy or reproduce the progress-based criterion of Meterez et al. (2026b).

Negative restart result. A separate test restarts fixed-diagonal-preconditioned SGD from an AdamW-derived post-cooldown checkpoint. It uses one future order and three 32-update paths with uniform, front-loaded and back-loaded batches. Every output improves over the anchor by 2.18×10^{-5}–2.81×10^{-5} nats/token, but the endpoint outperforms both tail and EMA on every path. All six sampling contrasts and both interactions remain unresolved under the original paired intervals. The restart and plain-SGD protocols differ, so the contrast does not identify the optimizer as the cause.

I.6 Implementation, numerical checks and costs

The current continuations use RTX 5090 GPUs, Python 3.11.14, PyTorch 2.10.0+cu128, CUDA 12.8 and NumPy 1.26.4. Parameters and forward/backward operations use FP32. TF32, autocast and compiled execution are disabled; attention uses the math backend and deterministic algorithms are active. Single-context gradients are accumulated in FP64 and rounded once before each FP32 SGD update. Returned averages accumulate anchor-relative displacements in FP64. Evaluation uses microbatch 8 with microbatch-4 checks on the first 128 contexts.

An independent forward impulse-response calculation reproduces all 64 shared fit/audit direction summaries, with maximum coefficient discrepancy 8.14×10^{-20}. The prospective calibration's 32-step and 16-step Lanczos checks differ by at most 3.46×10^{-12} relative to each direction's maximum coefficient. All 256 loss arrays in the sixteen-stream study reproduce the contrasts and intervals to within 6×10^{-20} nats/token of the saved analysis. Microbatch changes shift interaction means by at most 6.22×10^{-8}. These are numerical checks, not bounds on all floating-point error or population-model approximation.

The sixteen-stream confirmation has 128 paths, 67,108,864 training tokens and 2,147,483,648 evaluation tokens. Recorded costs are 1.7682 GPU-hours for training, 5.0282 for evaluation and 7.0626 total worker GPU-hours, with no new curvature fit. The prospective test's two confirmation fits cost 1.6544 and 1.6393 GPU-hours; their audits cost 1.2630 and 1.2616 GPU-hours. Development fit/audit adds 3.3095 GPU-hours, and the 16 prospective LM workers cost 0.8648 GPU-hours. These timing scopes include different operations and are not a time-to-target comparison. Equal continuation tokens do not imply equal end-to-end compute.

The supplement retains frozen allocations, ordered sample indices, per-context losses, calibration summaries, seeds and hardware records. CPU scripts reconstruct statistics from the saved losses. Full model weights remain in experiment storage; the package includes their hashes and the prior tensor-audit receipts. Recomputing statistics does not repeat that tensor audit. Historical per-task costs and statistical sensitivity checks are retained outside the paper PDF.

J FAILURE OF THE FITTED EXPECTED-OBJECTIVE SELECTOR

Objective and input estimates. Let $\widehat{g}$ be the mean of 32 single-context gradients at the anchor, and let $z_i = (g_{a_i} - g_{b_i})/\sqrt{2}$ for eight independent pairs of single-context gradients. Write $\widehat{\Sigma} = \frac{1}{8}\sum_{i=1}^{8} z_i z_i^\top$. With $\widehat{G} = \widehat{G}_{256}$, define $\mu_w = -\sum_{j<T} K_{j,w}(\widehat{G})\widehat{g}$. The fitted objective, omitting its plan-independent anchor loss, is

$$J(w, B) = \widehat{g}^\top \mu_w + \tfrac{1}{2}\mu_w^\top \widehat{G}\mu_w + \tfrac{1}{2}\sum_{j<T} \frac{\operatorname{tr}[\widehat{G}K_{j,w}(\widehat{G})\widehat{\Sigma}K_{j,w}(\widehat{G})^\top]}{B_j}. \tag{J.1}$$

We evaluate the mean- and noise-weighted spectral forms using Lanczos quadrature of depth 32. Depth 16, a 128-context curvature pool and each half of the eight noise pairs provide frozen input-sensitivity checks; none is retuned against LM outcomes. This J is not the exact class risk or a held-out CE estimator with a proved finite-sample error bound.

The fixed family contains 78 named simple outputs, including endpoint, tails, initialized EMA, decreasing geometric weights and intervals with initial/terminal mass. The search optimizes the full output simplex at uniform batches, assigns capped square-root allocations to the simple outputs, and alternates output and continuous allocation updates before checking feasible integer plans. These searches provide feasible plans, not a certificate of global integer optimality for the LM.

At the initialization-1 anchor, the selector chooses the endpoint for every optimized output. Its predicted Tail-64 and EMA-16 gains over the uniform endpoint are -0.219951 and -0.156792, while their observed two-order mean gains are $+0.000572623$ and $+0.000555211$ nats/token. The selected nonuniform endpoint improves over the uniform endpoint but remains worse than both fixed averages on each order. The jointly selected output is also the endpoint, making the proposed output–sampling interaction zero. The experiment does not establish useful output selection. Matched-path GGN forecasts use realized samples and are distinct from the fitted objective J.

A fixed-candidate mean-estimation bias. For fixed averaging weights w and curvature G independent of the mean sample, write $S_w = \sum_{j<T} K_{j,w}(G)$ and $M_w = -S_w + S_w G S_w/2$. If $\mathbb{E}\widehat{g} = g$ and $\operatorname{Cov}(\widehat{g}) = \Sigma/m$, expansion of the centered mean error gives

$$\mathbb{E}[\widehat{g}^\top M_w \widehat{g}] = g^\top M_w g + \frac{1}{m}\operatorname{tr}(M_w\Sigma). \tag{J.2}$$

For $0 \le \gamma_t\lambda \le 1$, $M_w(\lambda) = (a_\lambda(w)^2 - 1)/(2\lambda)$ at positive eigenvalues and $M_w(0) = -S_w(0)$. Positive contraction gives $M_{e_T} \preceq M_w$, so the endpoint-minus-average mean-estimation bias is nonpositive. GGN positive semidefiniteness alone does not imply the contraction condition. The identity also cannot be applied directly to weights selected using the same mean sample.

A post-hoc trace correction retains both wrong gain signs. A cross score using the two independent 16-context halves of the mean sample recovers both signs but overpredicts the gains by factors of about 2.12 and 1.85. These are diagnostics for three fixed candidates and one split, not a corrected selector or new training experiment. For fixed w, changing only batches leaves M_w unchanged, so the mean term cancels exactly. The prospective test uses that fixed-output comparison without claiming accurate output selection.