

Predetermined gradient descent: horizon, dimension, and step signs

Logan Bell

September 2026

Abstract

We bound the minimax terminal objective gap of smooth convex gradient descent when the step schedule is fixed in advance. For nonnegative steps, the lower bound has the silver exponent $s = \log_2(1 + \sqrt{2})$, a factor $(\log N)^{-\kappa_0}$ with $\kappa_0 \approx 0.40998$, and an explicit dimension penalty of order $\exp[-0.34560 \dots (\log N)^2/d]$. For arbitrary real steps, a lower bound with exponent 1.5570595 holds in dimension $N + 1$; we also quantify its dependence on smaller dimensions. The bounds give exponential gains over the scalar N^{-2} lower bound when dimension grows more slowly than $\log N$. Constructive positive schedules give an upper bound $0.211556239176(N + 1)^{-s}$ for $N \geq 2097151$ in every dimension. Negative steps do not improve the scalar minimax value at any horizon, or the minimax value in any dimension through five steps.

1 Problem and results

Let $\mathcal{F}_d$ be the differentiable convex functions on $\mathbb{R}^d$ with 1-Lipschitz gradient and an attained minimum. A predetermined schedule $h \in \mathbb{R}^N$ generates

$$x_{i+1} = x_i - h_{i+1} \nabla f(x_i), \quad 0 \leq i < N.$$

Its worst terminal gap and the two minimax values are

$$W_{N,d}(h) = \sup_{\substack{f \in \mathcal{F}_d, \ x_* \in \arg \min f \\ \|x_0 - x_*\| \leq 1}} f(x_N) - f(x_*), \quad V_{N,d}^+ = \inf_{h \geq 0} W_{N,d}(h), \quad V_{N,d}^{\mathbb{R}} = \inf_{h \in \mathbb{R}^N} W_{N,d}(h).$$

The schedule may depend on N and d . Throughout, $N, d \geq 1$, logarithms are natural unless a base is displayed, and

$$K = 2N + 1, \quad s = \log_2(1 + \sqrt{2}) \approx 1.2715533032.$$

The normalization is restored by multiplying objective bounds by LR^2 and dividing the steps by L , when the gradient Lipschitz constant is L and the initial radius is R .

A fixed schedule has the same worst case in all dimensions $d \geq N + 1$. Indeed, restrict an instance to the affine space through a minimizer spanned by the initial displacement and the N update gradients. Orthogonal projection preserves every update and function value. Conversely, an instance embeds in any larger space by making the function constant in the added coordinates. This proves stabilization at dimension $N + 1$ and monotonicity in dimension; see also [7].

The nonnegative bound, proved below, takes the form

$$V_{N,d}^+ \geq cN^{-s}(\log N)^{-\kappa_0} \exp \left[-a_+ \frac{(\log N)^2}{d} - C \frac{(\log N)^3}{d^2} \right], \quad (1)$$

for $N \geq N_0$ and $d \geq C_d \log N$, with fixed positive constants c, C, C_d, N_0 , where

$$D = 2\sqrt{2}(1 + \sqrt{2}) \log 2, \quad \kappa_0 = \frac{\pi^2}{4Ds} \approx 0.4099765758, \quad a_+ = \frac{s+2}{2D} \approx 0.3456032876.$$

For $d \geq c_0(\log N)^2$, the dimension factor is bounded below by a positive constant. For $d/\log N \rightarrow \infty$, this bound and the constructive upper bound give $V_{N,d}^+ = N^{-s+o(1)}$. Theorem 3.12 gives the bounds for dimensions comparable to or smaller than $\log N$.

For real schedules we prove

$$V_{N,N+1}^{\mathbb{R}} \geq \frac{1}{16(N+1)^{1.5570595}} \quad (N \geq 1). \quad (2)$$

The exponent $p_{\dagger}$ defined in Section 4.5 satisfies $1.5570593 < p_{\dagger} < 1.5570595$. With $a_{\dagger}$ defined in (69), we have $a_{\dagger} < 8/9$ and

$$V_{N,d}^{\mathbb{R}} \geq \frac{N^{-p_{\dagger}}}{3600} \exp \left[-a_{\dagger} \frac{(\log N)^2}{d} - C \left(\frac{(\log N)^{3/2}}{d} + \frac{(\log N)^3}{d^2} \right) \right] \quad (3)$$

for sufficiently large N and $d \geq C_d \log N$. If $d \rightarrow \infty$ and $d = o(\log N)$, the two sign domains satisfy

$$\log(N^2 V_{N,d}^+) \geq g_* d - O(\sqrt{d}), \quad g_* \approx 0.2546088726, \quad \log(N^2 V_{N,d}^{\mathbb{R}}) \geq \frac{90207}{1880000} d - O(\sqrt{d}).$$

All implied constants are independent of N, d in the stated regimes.

Every dimension also admits the common comparison

$$\frac{c_Q}{(N+1/2)^2 + 4c_Q/3} \leq V_{N,d}^{\mathbb{R}} \leq V_{N,d}^+ \leq C_N(N+1)^{-s}, \quad c_Q = 0.304194775980, \quad (4)$$

where $C_N = (1 + \sqrt{2})/8$ is valid for all $N \geq 1$ and $C_N = 0.211556239176$ is valid for $N \geq 2097151$. The lower bound follows from the exact quadratic–Huber constant in [12]. The upper schedules and their finite certificates are given in Section 8.

Theorem 6.2 proves $W_{N,1}(h) \geq W_{N,1}(h_+)$, where $(h_+)_i = \max(h_i, 0)$. Consequently $V_{N,1}^{\mathbb{R}} = V_{N,1}^+$ for every horizon. Theorem 7.2 gives the same minimax equality in every dimension for $N \leq 5$. Exact scalar enumeration in [10] also gives $V_{3,1}^{\mathbb{R}} = V_{3,1}^+ \leq 1250000/30372173$.

1.1 Relation to known bounds

The silver schedules of Altschuler–Parrilo [9] attain the exponent s . The composition theorems of Grimmer–Shu–Wang [4] provide the upper constructions used here. Concurrent work of Zhang–Jiang [5, Section 4.1, Theorem 4.13, and Remark 5.9] obtains the same optimized schedules by concatenation and proves an asymptotic objective bound with exponent s . Ye–Liu [11, Theorem 1.1] proved a lower bound $N^{-s} \exp[-O(\sqrt{\log N \log \log N})]$ for nonnegative schedules in unrestricted dimension. Equation (1) gives a logarithmic loss and explicit dimension dependence. For real schedules, (2) improves Ye–Liu’s $\Omega(N^{-1.6342})$ bound [3, Theorem 1.1 and Appendix D].

2 Scalar polynomial estimates

In this section let $w = W_{N,d}(h)$ and $b_i = \max(h_i, 0)$.

Lemma 2.1 (Imaginary-axis bound). *For every $y > 0$,*

$$\sum_i \log(1 + b_i y^2) \leq K \operatorname{arsinh} y + \log(\sqrt{2w}/y). \quad (5)$$

In particular, if $r_i \leq b_i$ are nonnegative and $s_i = k^2 r_i / K^2$, then $\sum_i \log(1 + s_i / k^2) \leq 1 + \log(K\sqrt{2w})$.

Proof. The scalar quadratics give $\|z \prod_i (1 - h_i z^2)\|_{[-1,1]} \leq \sqrt{2w}$. Each negative step contributes a factor at least one on this interval. Deleting those factors gives $P(z) = z \prod_i (1 - b_i z^2)$ with $M := \|P\|_{[-1,1]} \leq \sqrt{2w}$. The polynomial $\zeta^K P((\zeta + \zeta^{-1})/2)$ has modulus at most M on the unit circle, hence in its disk. At $z = iy$, the inverse root inside the disk has modulus $e^{-\operatorname{arsinh} y}$. Therefore $|P(iy)| \leq M e^{K \operatorname{arsinh} y}$. Since $|P(iy)| = y \prod_i (1 + b_i y^2)$, this proves (5). Take $y = 1/K$ and use $K \operatorname{arsinh}(1/K) < 1$ for the final assertion. $\square$

Lemma 2.2 (Ordered tail bound). *Write $b_1 \geq \dots \geq b_N \geq 0$. If $1 \leq k \leq N$ and $w \leq k^2/(2K^2)$, then*

$$\sum_{j>k} b_j \leq \frac{27K^2}{64k}.$$

Proof. In (5) take $y = 8k/(3K)$. Its right side is at most $8k/3$. The vector $u_j = \log(1 + b_j y^2)$ belongs to the ordered simplex $u_1 \geq \dots \geq u_N \geq 0$, $\sum u_j \leq 8k/3$. Its vertices have their first r coordinates equal to $8k/(3r)$ and the rest zero. Convexity bounds $\sum_{j>k} (e^{u_j} - 1)$ by $\max_{r>k} (r - k)(e^{8k/(3r)} - 1)$. For $0 \leq v \leq 1$, $(1 - v)(e^{8v/3} - 1) \leq 3v$: the derivative of $\log((1 + 2v)/(1 - v))$ is at least $8/3$, since $(1 + 2v)(1 - v) \leq 9/8$. Put $v = k/r$ and divide by y^2 . $\square$

Finally, for $0 < \eta \leq 1$,

$$W_{N,d}(h) \geq \eta W_{N,d}(\eta h). \quad (6)$$

An instance f under ηh and the instance ηf under h have the same iterates, and the latter remains 1-smooth.

3 Nonnegative schedules

We adapt the backward-trajectory and adjacent-coordinate construction of Ye–Liu [11, Lemmas 2.2, 3.1, and 3.2; Appendix A], which develops the checkpoint constructions of Ma–Chen [2, Theorem 4.1] and Jung–Cho–Yun [1, Section 2]. We use an arcsine profile with explicit turning cost and allow its parameters to vary between checkpoints. A selection of q steps, called checkpoints, gives a lower-bound instance in $q + 1$ coordinates. The unselected steps between consecutive checkpoints form a gap; its mass is their sum. The local construction and the selection estimate below prove the following bound, with s, D, κ_0, a_+ as defined in Section 1.

Theorem 3.1. *There are positive constants c, C, C_d, N_0 such that, for $N \geq N_0$ and $d \geq C_d \log N$, writing $L = \log N$,*

$$V_{N,d}^+ \geq c N^{-s} L^{-\kappa_0} \exp \left\{ -a_+ \frac{L^2}{d} - C \frac{L^3}{d^2} \right\}. \quad (7)$$

Also, for all $N \geq 1$,

$$V_{N,N+1}^+ \geq c N^{-s} [\log(N + 2)]^{-\kappa_0}. \quad (8)$$

3.1 A local arc and its smooth convex realization

Lemma 3.2 (Discrete turning). *Let P be a positive smooth function on $[0, 1]$, with $P(1) = \epsilon > 0$. Put*

$$v(q) = (P(q), -q), \quad M(q) = -P'(q), \quad \omega(q) = P(q)M(q) - q.$$

Suppose $M > 0$, $\omega > 0$, and $\omega' \geq 0$. For every finite list of nonnegative steps there is a backward trajectory ending at $v(1)$ whose gradient at each point is its projection onto $\text{conv}(0, v([0, 1]))$. The parameters q decrease along the backward trajectory. If H_{turn} is the sum of steps whose projected point has $q > 0$, then

$$\log(1 + H_{\text{turn}}) \leq I := \int_0^1 \frac{1 + \omega'(q) + M(q)^2}{\omega(q)} dq. \quad (9)$$

Proof. The identity $PM' = 1 + M^2 + \omega'$ implies $M' > 0$ and $P'' < 0$. The vector $n(q) = (1, -M(q))$ supports the gradient body at $v(q)$: the assertion against other arc points is the tangent inequality for P , and its support gap against zero is $P(q) + qM(q) > 0$.

Suppose the current backward point is $r = v(Q) + \tau_Q n(Q) = (u, -y)$, where $Q > 0$ and $\tau_Q \geq 0$. For $0 \leq q \leq Q$, define

$$h(q) = \frac{(Q - q) + (P(q) - P(Q))M(q) + \tau_Q(M(Q) - M(q))}{\omega(q)}.$$

This continuous nonnegative function has $h(Q) = 0$. For a prescribed step $0 \leq h < h(0)$, choose a root $h(q) = h$. The identity $h\omega(q) = y - uM(q) + \omega(q)$ gives $r + hv(q) - v(q) = \tau_{\text{new}}n(q)$, where

$$\tau_{\text{new}} = u + (h - 1)P(q) \geq \tau_Q.$$

Indeed, $h\omega(q) \geq (P(q) - P(Q))M(q)$ and $\omega(q) \leq P(q)M(q)$ imply $hP(q) \geq P(q) - P(Q)$. Thus the required projection is valid. If $h \geq h(0)$, use $q = 0$. The inequality is then $y \leq M(0)[u + (h - 1)P(0)]$; the residual is a nonnegative combination of $(1, -M(0))$ and $(1, 0)$. Once $q = 0$ is reached, all earlier steps use this endpoint. Starting at $(\epsilon, -1)$, the backward coordinates satisfy $u \geq \epsilon$, $y \geq 1$, and $y/u \leq 1/\epsilon$.

At a curved step, the projection equation yields

$$h\omega(q) = u[M(Q) - M(q)] - \omega(Q) + \omega(q), \quad \frac{u_{\text{new}}}{u} \leq 1 + \frac{P(q)}{\omega(q)}[M(Q) - M(q)].$$

Regarding P, ω as functions of $m = M(q)$, choose a positive function B with $B'/B = P/\omega$. Direct differentiation gives

$$\frac{B''}{B} = \frac{P^2 M' - M\omega - P\omega'}{\omega^2 M'} = \frac{P + qM}{\omega^2 M'} > 0.$$

The tangent inequality for B bounds the preceding ratio by $B(M(Q))/B(M(q))$. These ratios telescope over all curved steps. Their first-coordinate increase is at least ϵH_{turn} , so

$$1 + H_{\text{turn}} \leq \frac{B(M(1))}{B(M(0))}, \quad \log \frac{B(M(1))}{B(M(0))} = \int_0^1 \frac{PM'}{\omega} dq = I.$$

□

Lemma 3.3 (Arcsine transfer). *For $0 < \epsilon, \gamma \leq 1/2$, put*

$$c_{\text{arc}} = \sqrt{1 + 2\gamma + \epsilon^2}, \quad a_{\text{arc}} = 8 \exp \left\{ \frac{\pi^2(1 + \epsilon^2)}{4\gamma} + 1 \right\} \epsilon^{-4}.$$

For any nonnegative gap of total mass S , a following step $\beta \geq 1 + \epsilon^{-2}$, and $C_{\text{off}} \geq c_{\text{arc}} a_{\text{arc}}/2$, there is a nonnegative convex 1-smooth function on $\mathbb{R}^2$ whose trajectory from $(1, 0)$ ends at $(U, Y + G)$, where

$$G = \frac{\beta}{c_{\text{arc}} S + \epsilon \beta + C_{\text{off}}}, \quad U, Y \geq 0.$$

The function is zero on both rectangles $\{z_1 \leq 0, z_2 \geq 0\}$ and $\{z_1 \leq U, z_2 \geq Y\}$. During the gap the first coordinate decreases and remains at least U , while the second increases from zero to Y .

Proof. Write $C = 1 + \epsilon^2$, $A_\epsilon = \arcsin(C^{-1/2})$, and use

$$\omega(q) = \frac{\gamma}{A_\epsilon \sqrt{C - q^2}}, \quad P(q)^2 = C - q^2 + \frac{2\gamma}{A_\epsilon} [A_\epsilon - \arcsin(q/\sqrt{C})].$$

Then $P(1) = \epsilon$, $P(0) = c_{\text{arc}}$, $-PP' = q + \omega$, $\omega' \geq 0$, and $\int_0^1 \omega = \gamma$. The hypotheses of Lemma 3.2 hold. Since $M = (q + \omega)/P$ and $P^2 \geq C - q^2$,

$$\int_0^1 \left(\frac{1}{\omega} + \frac{q^2}{\omega P^2} \right) dq \leq \frac{C A_\epsilon^2}{\gamma}, \quad \int_0^1 \frac{\omega'}{\omega} dq = \log \frac{\sqrt{C}}{\epsilon}, \quad \int_0^1 \frac{2q + \omega}{P^2} dq \leq \log \frac{c_{\text{arc}}^2}{\epsilon^2}.$$

Here $c_{\text{arc}}^2 \sqrt{C} \leq 9\sqrt{5}/8 < e$. Consequently

$$I < \frac{\pi^2(1 + \epsilon^2)}{4\gamma} + 1 + 3 \log(1/\epsilon).$$

Apply the backward construction to the gap. If its projected gradients are $v(q_j)$, write $P_{\text{sum}} = \sum h_j P(q_j)$ and $Q_{\text{sum}} = \sum h_j q_j$. Then $P_{\text{sum}} \leq c_{\text{arc}} S$ and $Q_{\text{sum}} \leq H_{\text{turn}}$. Set

$$\rho = (c_{\text{arc}} S + \epsilon \beta + C_{\text{off}})^{-1}, \quad r_0 = (\epsilon + P_{\text{sum}}, -1 - Q_{\text{sum}}), \quad z_o = (1, 0) - \rho r_0.$$

Let $\mathcal{K}$ be the convex hull of zero and the finitely many gradients used, including $v(1)$, and define

$$f(z) = \max_{g \in \rho \mathcal{K}} \left\{ \langle g, z - z_o \rangle - \frac{1}{2} \|g\|^2 \right\}.$$

This is the Moreau envelope of the translated support function. Its maximizing vector is $\Pi_{\rho \mathcal{K}}(z - z_o)$. The projection variational inequalities imply $\|\Pi x - \Pi y\|^2 \leq \langle \Pi x - \Pi y, x - y \rangle$; thus f is convex and 1-smooth. The backward projection equations give exactly its prescribed gradient-descent trajectory.

Put $C_{\text{req}} = \epsilon + (1 + H_{\text{turn}})/\epsilon$. The bound on I shows $c_{\text{arc}} a_{\text{arc}}/2 \geq C_{\text{req}}$. For every retained gradient,

$$\langle v(q), z_o \rangle \geq \rho P(q) \left[C_{\text{off}} - \epsilon - \frac{1 + H_{\text{turn}}}{\epsilon} \right] \geq 0.$$

Since its components have signs $(+, -)$, every supporting affine form is nonpositive on the first stated rectangle; there $f = 0$. Immediately before the checkpoint the point is $(1 - \rho P_{\text{sum}}, \rho Q_{\text{sum}})$. After it, set

$$U = 1 - \rho(P_{\text{sum}} + \epsilon \beta) \geq 0, \quad Y = \rho Q_{\text{sum}}, \quad G = \rho \beta.$$

For every retained gradient,

$$\langle v(q), (U, Y) - z_o \rangle = \rho[(1 - \beta)\epsilon P(q) + q] \leq 0,$$

because $q/P(q) \leq 1/\epsilon$. This gives the second zero rectangle. The coordinate monotonicity follows from the signs of the gradients. $\square$

Fix $C_0 = \pi^2/4$ and $B_* = \exp(4096)$. For $b \geq B_*$, define

$$c(b) = 1 + \frac{C_0}{\log b} + \frac{24C_0 \log \log b}{(\log b)^2}, \quad e(b) = \frac{2}{(\log b)^4}, \quad K_b(S) = \frac{c(b)(a + S)}{b} + e(b). \quad (10)$$

The fixed offset a will be chosen below.

Lemma 3.4 (Kernel and coordinate count). *For every $a \geq 1$, every nonnegative schedule, and every selection of q steps $b_i \geq B_*$, let S_i be the intervening gap masses and T the final gap. There is an admissible function on $\mathbb{R}^{q+1}$ with initial point e_1 , minimizer zero, and terminal value*

$$\frac{1}{4(1+T)} \prod_{i=1}^q K_{b_i}(S_i)^{-2}. \quad (11)$$

For $q = 0$ the empty product is one. The coefficient functions c, e decrease, $c \leq 5/4$, and, writing $\Delta(b) = e(b)b/c(b)$ and $B(b) = b - \Delta(b)$,

$$0 < \Delta(b) < b/50, \quad \Delta(b) \leq \frac{3B(b)}{[\log B(b)]^4}. \quad (12)$$

Proof. For $t = \log b \geq 4096$, use $\epsilon = t^{-4}$ and $\gamma = C_0/(t - 22 \log t)$ in Lemma 3.3. Since $22 \log t/t < 1/20$,

$$a_{\text{arc}} = 8e \frac{b}{t^6} \exp\{(t - 22 \log t)t^{-8}\} < 24b/t^6.$$

Also

$$c(b) - 1 - \gamma - \epsilon^2/2 = \frac{C_0 \log t}{t^2} \left(24 - \frac{22}{1 - 22 \log t/t} \right) - \frac{1}{2t^8} > 0.$$

Hence $c_{\text{arc}} \leq c(b) \leq 5/4$ and $c_{\text{arc}} a_{\text{arc}} + \epsilon b \leq 2b/t^4 = e(b)b$. The condition $b \geq 2(1 + t^8)$ holds at $t = 4096$ and persists, since $e^t/(1 + t^8)$ increases for $t \geq 8$. The estimates on Δ follow from $\Delta(b)/b \leq 2/t^4$ and $B(b) \geq 49b/50$. The formulas directly give decreasing c, e .

Apply the local construction using half the original steps: $S = S_i/2$, $\beta = b_i/2$, and

$$C_{\text{off},i} = \frac{1}{2}[c(b_i)(a + S_i) + e(b_i)b_i - c_{\text{arc},i}S_i - \epsilon_i b_i].$$

This is at least $c_{\text{arc},i}a_{\text{arc},i}/2$. Its exact transfer is $G_i = K_{b_i}(S_i)^{-1}$. Let f_i, U_i, Y_i be its local data and define

$$A_1 = 1, \quad \ell_1 = 0, \quad A_{i+1} = A_i G_i, \quad \ell_{i+1} = A_i Y_i, \quad F_i(x) = A_i^2 f_i \left(\frac{x_i - \ell_i}{A_i}, \frac{x_{i+1}}{A_i} \right).$$

Let H_δ be zero on $(-\infty, 0]$, equal to $z^2/2$ on $[0, \delta]$, and equal to $\delta z - \delta^2/2$ above δ . With $\delta = A_{q+1}/(1 + T)$, set

$$F(x) = \frac{1}{2} \sum_{i=1}^q F_i(x) + \frac{1}{2} H_\delta(x_{q+1} - \ell_{q+1}).$$

Before a block's checkpoint, its output is at most its activation offset ℓ_{i+1} , so the next block lies in its first zero rectangle. At the checkpoint that next block starts at normalized input $(1, 0)$. During its evolution its first coordinate stays at least its offset, keeping the preceding block in its second zero rectangle. Induction shows that precisely the prescribed block is active. This also proves the assertion when the arcs differ between blocks.

Each F_i is 1-smooth on two adjacent coordinates. After its coefficient $1/2$ is included, the sum of all quadratic upper remainders, including the Huber term, is at most $\|v\|^2/2$: each coordinate occurs in at most two components. Thus F is globally 1-smooth and convex. Every component is nonnegative and vanishes at zero, since $\ell_i \geq 0$. The initial distance from that minimizer is one. During the tail the final coordinate above its offset ends at $A_{q+1} - \delta T/2 = \delta(1 + T/2) \geq \delta$. All partial tail sums are at most T , so this coordinate stays on the affine Huber branch. Its final contribution is $A_{q+1}^2/[4(1 + T)]$, proving (11). The single Huber function gives the case $q = 0$. $\square$

3.2 The critical splitting potential

For $0 \leq \theta \leq \log 4$, define

$$\begin{aligned} \chi_0 &= e^\theta, \quad z_0 = \sqrt{1 + \chi_0}, \quad \rho = 1 + z_0, \quad p = p_\theta = \log_2 \rho, \quad \nu = 1/p, \\ D_{\chi_0} &= 2z_0\rho \log 2, \quad A = A_\theta = \frac{C_0\chi_0}{D_{\chi_0}p^2}, \quad \kappa_\theta = p_\theta A_\theta. \end{aligned}$$

Here and below these parameters range over the indicated compact interval. Fix $H_0 = e^{1024}$ and write $\ell(z) = \log(H_0 + z)$. Set

$$\begin{aligned} \phi_\theta(z) &= z^\nu \ell(z)^{-A} \exp \left\{ \frac{16(\log \ell(z) + 1)}{\ell(z)} \right\}, \\ \psi_\theta(z) &= C_3 \phi_\theta(z) \ell(z)^{-4\nu}, \quad \phi_\theta(0) = \psi_\theta(0) = 0. \end{aligned} \tag{13}$$

The positive constant C_3 will be fixed before a .

Write $\hat{c}(t) = 1 + C_0/t + 24C_0 \log t/t^2$ and $\delta_* = -\hat{c}'(4096)/\hat{c}(4096) > 0$. Extend the coefficient to every $B > 0$ by

$$\chi_\theta(B) = \begin{cases} e^\theta c(B), & B \geq B_*, \\ e^\theta c(B_*)(B_*/B)^{\delta_*}, & 0 < B < B_*. \end{cases} \tag{14}$$

Lemma 3.5 (Coefficient and potential properties). *The functions ϕ_θ, ψ_θ are strictly increasing and concave. Their logarithmic derivatives $zf'(z)/f(z)$ exceed $1/2$, the ratio $\phi_\theta(z)/z^\nu$ decreases, and*

$$-\frac{z\psi_\theta''(z)}{\psi_\theta'(z)} \leq 1 - \frac{9}{16}. \tag{15}$$

The function $E(B) = 1 - d \log \chi_\theta(B)/d \log B$ is at least one and nonincreasing. Also

$$e^\theta [c(b) - e(b)] \leq \chi_\theta(B(b)). \tag{16}$$

One can fix C_3 , uniformly in θ , so that

$$\phi_\theta(\Delta(b)) \leq \psi_\theta(B(b)) \quad (b \geq B_*). \tag{17}$$

Proof. First $5/4 < p < 7/4$, since $(1 + \sqrt{2})^4 > 2^5$ and $(1 + \sqrt{5})^4 < 2^7$. Thus $4/7 < \nu < 4/5$. We also have

$$\frac{10\chi_0}{D_{\chi_0}p^2} < 2. \quad (18)$$

Indeed, with $z = \sqrt{1 + \chi_0}$ the left side is $5 \log 2(1 - 1/z)/\log(1 + z)^2$. For $z \leq 7/4$, the bounds $\log 2 < 7/10$ and $\log(1 + \sqrt{2}) > 22/25$ make it less than $1875/968 < 2$. For $z \geq 7/4$, use $\log(11/4) > 1$ and $1 - 1/\sqrt{5} < 553/1000$ to obtain $3871/2000 < 2$. These estimates follow from the exponential series and

$$\log(1 + \sqrt{2}) > 2 \left(u + \frac{u^3}{3} + \frac{u^5}{5} \right) = \frac{254\sqrt{2} - 346}{15} > \frac{22}{25}, \quad u = \sqrt{2} - 1.$$

In particular $A < \pi^2/20 < 1/2$.

Both potentials, apart from a positive constant, have the form

$$f(z) = z^\nu \ell(z)^{-\alpha} \exp \left\{ \frac{16(\log \ell(z) + 1)}{\ell(z)} \right\}, \quad \alpha \in \{A, A + 4\nu\} \subset (0, 4).$$

Put $v = z/(H_0 + z)$ and $h(J) = \alpha/J + 16 \log J/J^2$. For $J \geq 1024$, $0 < h < 1/200$, $h' < 0$, and $-h' < 1/200$. If $q = zf'/f$, direct differentiation gives

$$q = \nu - vh(\ell), \quad zq' = -v(1 - v)h(\ell) - v^2h'(\ell).$$

It follows that $1/2 < q < 4/5$, $zq' < 1/200$, and $q^2 - q + zq' < -4/25 + 1/200 < 0$. These prove strict increase and concavity; continuity at zero extends concavity to the closed half-line. For ψ_θ , the bound $zq' \geq -h/4$ also gives

$$q + \frac{zq'}{q} > \frac{4}{7} - \frac{1}{200} - \frac{1}{400} > \frac{9}{16}.$$

This is (15). The formula for q proves the assertion about $\phi_\theta(z)/z^\nu$.

For $t \geq 4096$,

$$\hat{c}' = -\frac{C_0}{t^2} + \frac{24C_0(1 - 2 \log t)}{t^3} < 0, \quad \hat{c}'' = \frac{2C_0}{t^3} + \frac{24C_0(6 \log t - 5)}{t^4},$$

and $|\hat{c}'| \leq 2C_0/t^2$. Consequently $\hat{c}''\hat{c} - (\hat{c}')^2 \geq 2C_0/t^3 - 4C_0^2/t^4 > 0$. This proves that $E = 1 - \hat{c}'/\hat{c}$ is at least one and decreasing above B_* . The extension in (14) matches its value and first derivative and has constant E below B_* . Monotonicity of c , $B(b) \leq b$, and the downward extension give (16).

Finally (12) and the explicit potentials imply, uniformly in θ ,

$$\frac{\phi_\theta(\Delta(b))}{\phi_\theta(B(b))} \sim 2^\nu [\log B(b)]^{-4\nu}.$$

To verify uniformity, note that $\Delta(b)/B(b) \sim 2/(\log B(b))^4$, $\log \Delta(b)/\log B(b) \rightarrow 1$, and both exponents range in compact intervals. The final exponential factor in (13) tends to one. Choose C_3 for all sufficiently large b ; the remaining compact interval is covered by a finite enlargement, using positivity and continuity. $\square$

Lemma 3.6 (Critical scalar inequality). *For the fixed C_3 above, there is a fixed $a \geq 3B_*$ such that, uniformly in θ ,*

$$x, y \geq a, \quad B > 0, \quad w = x + y + B - a, \quad wB = \chi_\theta(B)xy \implies \phi_\theta(w) + \psi_\theta(B) \leq \phi_\theta(x) + \phi_\theta(y). \quad (19)$$

Proof. We use the following parameterized form of Ye–Liu’s silver inequality [11, Lemma 3.9 and Appendix B.6]; its constant-coefficient splitting analogue appears in Jung–Cho–Yun [1, Lemma C.4]. If $p = \log_2(1 + \sqrt{1 + \chi_0})$, then

$$t^p + (1 - t)^p + \chi_0[t(1 - t)]^p \leq 1 \quad (0 \leq t \leq 1), \quad (20)$$

with equality in the interior only at $t = 1/2$. On $(0, 1/2)$ the derivative has the sign of $g(t) = (1 - t)^{1-p} - t^{1-p} + \chi_0(1 - 2t)$. This function is strictly concave, tends to $-\infty$ at zero, and has $g(1/2) = 0$ and $g'(1/2) = 2^{p+1}(p + 1 - 2^p) < 0$. Thus the left side of (20) first decreases and then increases between its equal endpoint values at 0 and $1/2$. Symmetry proves the claim and its strictness.

Assume $x \leq y$. The scalar relation gives

$$\frac{B}{\chi_\theta(B)} \leq x \leq \frac{1 + \sqrt{1 + \chi_\theta(B)}}{\chi_\theta(B)} B.$$

Since $\chi_\theta \geq 1$, it forces $B \geq a/(1 + \sqrt{2})$. Increasing a therefore makes B exceed any required fixed threshold. Thereafter $\chi_\theta(B) \leq 5$ and $B/5 \leq x \leq (1 + \sqrt{2})B$. If $y/x \geq 10^5$, concavity and the decreasing slowly varying factor give

$$\phi(w) - \phi(y) \leq \phi'(y)(x + B) \leq 6\nu \frac{x}{y} \phi(y) \leq 6\nu(y/x)^{\nu-1} \phi(x) < \frac{1}{2} \phi(x).$$

Also $\phi(x) \geq \phi(B)/5$. This leaves at least $\phi(B)/10$, whereas $\psi(B)/\phi(B) \rightarrow 0$ uniformly.

It remains to treat the compact ratio range $1 \leq y/x \leq 10^5$. Write $T = \log B$, $X = x/B$, $Y = y/B$, and

$$Z = \sqrt{(X + Y)^2 + 4\chi_0 XY}, \quad W = \frac{X + Y + Z}{2}.$$

The positive root w_0 of $w_0^2 = (x + y)w_0 + \chi_\theta(B)xy$ is at least w . Uniform Taylor expansion on the compact set $1/5 \leq X \leq 1 + \sqrt{2}$, $X \leq Y \leq 10^5 X$ gives

$$\frac{\phi(x) + \phi(y) - \phi(w_0)}{\phi(B)} = F_0 + \frac{F_1}{T} + \frac{\log T}{T^2} F_2 + O(T^{-2}), \quad (21)$$

where

$$F_0 = X^\nu + Y^\nu - W^\nu, \quad S = W^\nu \log W - X^\nu \log X - Y^\nu \log Y, \quad J = \chi_0 \nu W^{\nu-1} \frac{XY}{Z},$$

$$F_1 = AS - C_0 J, \quad F_2 = 16S - 24C_0 J.$$

For completeness, the two expansions giving this identity are

$$\frac{w_0}{B} = W + \frac{C_0 \chi_0 XY}{TZ} + \frac{24C_0 \chi_0 XY \log T}{T^2 Z} + O(T^{-2}),$$

$$\frac{\phi(Bt)}{\phi(B)} = t^\nu \left[1 - \frac{A \log t}{T} - \frac{16 \log t \log T}{T^2} + \frac{A(A+1)(\log t)^2}{2T^2} + O\left(\frac{\log T}{T^3}\right) \right]$$

for t in any fixed compact positive interval. Replacing $\log(H_0 + Bt)$ by $T + \log t$ costs $O(H_0/B)$. All root, logarithm, and power derivatives are bounded on the compact ranges, so the remainders are uniform.

Inequality (20) implies $F_0 \geq 0$, with equality exactly when $X = Y$. To see the equivalence, set $R = (X^\nu + Y^\nu)^{1/\nu}$ and $t = X^\nu/(X^\nu + Y^\nu)$ in (20); it gives $R^2 \geq (X + Y)R + \chi_0 XY$. Near equality write $X = Me^{-r}$, $Y = Me^r$. Then

$$F_0 = M^\nu G(r), \quad G(r) = 2 \cosh(\nu r) - \left(\cosh r + \sqrt{\cosh^2 r + \chi_0} \right)^\nu,$$

$$G(0) = G'(0) = 0, \quad G''(0) = 2\nu(\nu - 1/z_0) > 0.$$

Indeed $\log_2(1+z) < z$ for $z > 1$, by differentiating $z \log 2 - \log(1+z)$. Compactness gives a uniform lower bound $F_0 \geq cr^2$ in a fixed neighborhood of balance. At balance,

$$S = 2M^\nu \log \rho, \quad J = \frac{\chi_0 \nu M^\nu}{z_0 \rho}, \quad F_1 = 0.$$

The function F_1 is smooth and even in r ; hence $|F_1| \leq Cr^2$ there. Moreover

$$F_2(M, 0) = 2M^\nu \log \rho \left(16 - \frac{24C_0\chi_0}{D_{\chi_0}p^2} \right) > 0,$$

uniformly, since (18) bounds the subtracted term by $24C_0/5 = 6\pi^2/5 < 12$. Reduce the balanced neighborhood so that F_2 remains uniformly positive. For large T , $F_0 + F_1/T \geq 0$ there, and its positive $\log T/T^2$ term dominates the $O(T^{-2})$ remainder. Outside this neighborhood, F_0 has a positive compact minimum. Together the two regions give

$$\phi(x) + \phi(y) - \phi(w) \geq c \phi(B) \frac{\log \log B}{(\log B)^2}.$$

Finally $4\nu > 16/7 > 2$, so this dominates $\psi(B)$ uniformly. Choose a , after C_3 , large enough that the forced lower bound on B exceeds the common threshold. This proves (19). $\square$

3.3 Weighted splitting and fixed gaps

For a chronological list of candidates, a selection T partitions the unselected candidates and fixed gaps into masses $s_i(T)$ preceding the selected candidates and a final mass $s_{\text{tail}}(T)$. The next lemma adapts the tight-set splitting argument of Jung–Cho–Yun [1, Section 3 and Appendix C] and its weighted form in Ye–Liu [11, Lemma 3.7 and Appendix B.4] to a size-dependent coefficient and concave potentials.

Lemma 3.7 (Weighted splitting). *Let ϕ, ψ be increasing concave functions, continuous and zero at the origin and twice differentiable on $(0, \infty)$, with positive first derivatives. Suppose $-z\psi''(z)/\psi'(z) \leq 1 - r$ for some $r > 1/2$. Let $\chi > 0$ be continuously differentiable and piecewise twice differentiable, with $E(B) = 1 - d \log \chi(B)/d \log B \geq 1$ nonincreasing, and with $B/\chi(B)$ increasing from zero to infinity. Assume the scalar implication (19). Fix $g_0, \dots, g_m \geq 0$, candidate sizes $B_1, \dots, B_m > 0$, and $\Lambda \geq a$. Define*

$$C_\Lambda(T) = (\Lambda + s_{\text{tail}}(T)) \prod_{i \in T} \frac{\chi(B_i)(a + s_i(T))}{B_i}, \quad W = C_\Lambda(\emptyset) = \Lambda + \sum_{j=0}^m g_j + \sum_{i=1}^m B_i.$$

If the empty set minimizes this cost, then

$$\phi(W) + \sum_{i=1}^m \psi(B_i) \leq \sum_{j=0}^{m-1} \phi(a + g_j) + \phi(\Lambda + g_m). \quad (22)$$

Proof. We induct on m , with equality for $m = 0$. For fixed gaps maximize the left side of (22) over candidate sizes subject to $C_\Lambda(T) \geq W$ for every T . The singleton constraint, with $x_i = a + \text{mass before } i$ and $y_i = \Lambda + \text{mass after } i$, is $WB_i \leq \chi(B_i)x_i y_i$ and $W = x_i + y_i + B_i - a > y_i$. It bounds $B_i/\chi(B_i) < x_i$ successively in chronological order. Multiplying each subset constraint by $\prod_{i \in T} B_i/\chi(B_i)$ extends it continuously to zero sizes. Constraints containing a zero size are then automatic; the remaining constraints are those of the reduced problem. The feasible set is compact. At a zero size, deletion merges its neighboring fixed gaps; subadditivity and monotonicity of ϕ make the reduced right side no larger than the original one. The induction hypothesis treats this boundary.

At an interior maximum, some nonempty set is tight: otherwise increasing one size slightly increases the objective and preserves all inequalities. We give the needed tight-set property directly. Inserting a candidate of size B into a gap with left mass ℓ and right terminal mass $y \geq a$ multiplies the cost by

$$\frac{\chi(B)(a + \ell)y}{B(\ell + B + y)}. \quad (23)$$

This ratio strictly increases in both ℓ and y . Enlarging the background selection reduces those two masses. Multiplying successive insertion ratios therefore gives

$$C_\Lambda(A \cup B)C_\Lambda(A \cap B) \leq C_\Lambda(A)C_\Lambda(B).$$

If A, B are disjoint and nonempty, the inequality is strict: the first insertion from B has a strictly shorter left or right gap when A is present than when the background is empty. Thus tight sets are closed under union and intersection, and two nonempty tight sets cannot be disjoint. An inclusion-minimal nonempty tight set is contained in every nonempty tight set.

Put $L(B) = E(B)/B = 1/B - (\log \chi)'(B)$. Suppose this common set contains two unequal sizes $u < v$. Perturb them by $(-t, qt)$, leaving the other sizes fixed. Every nonempty tight cost has derivative $W[L(u) - qL(v)]$, whereas the empty cost has derivative $q - 1$. All tight constraints gain slack if

$$q < A_* := \frac{L(u) + 1/W}{L(v) + 1/W}.$$

The objective increases if $q > D_* := [\phi'(W) + \psi'(u)]/[\phi'(W) + \psi'(v)]$. The hypotheses imply

$$A_* \geq \frac{v(W + u)}{u(W + v)} > \sqrt{v/u} > (v/u)^{1-r} \geq \frac{\psi'(u)}{\psi'(v)} \geq D_*.$$

For the strict middle comparison use $W > u + v$; after setting $z = \sqrt{v/u} > 1$, the remaining numerator is $(z - 1)(z^2 - z + 1) > 0$. The derivative ratio follows by integrating the hypothesis on ψ''/ψ' . A choice $D_* < q < A_*$ contradicts maximality.

For two equal sizes b , perturb to $b - t$ and $b + t + t^2/(2b)$. Since $L' \leq -L/B$, the quadratic increase of their log factor product has coefficient $-L(b)/(2b) - L'(b) \geq 1/(2b^2)$. Every tight cost minus W gains at least $(W - b)t^2/(2b^2) + o(t^2)$, while the objective coefficient is

$$\frac{\phi'(W)}{2b} + \frac{\psi'(b)}{2b} + \psi''(b) > 0.$$

The integrated inequality for L gives the same lower expansion at a joining point of χ . All nontight inequalities retain their slack for small perturbations, since there are finitely many. This again contradicts maximality.

A tight singleton therefore exists. At that index the left and right empty subproblem costs $x, y \geq a$ satisfy $W = x + y + B - a$ and $WB = \chi(B)xy$. Both subproblems have empty-set optimality: replacing either side in the tight singleton selection would otherwise lower the full cost. Apply the scalar inequality and the induction hypothesis to the two sides. Their fixed-gap sums partition the right side of (22); their correction terms, together with $\psi(B)$, give its left side. $\square$

We use the change of variables and cancellation of moved checkpoint mass from Ye–Liu [11, Lemmas 3.6 and 3.8; Appendices B.3 and B.5], with the coefficient and potentials defined above.

Proposition 3.8 (Selection with fixed gaps). *Fix C_3 as in Lemma 3.5, and then fix $a \geq 3B_*$ large enough for Lemma 3.6. For every $0 \leq \theta \leq \log 4$, every $m \geq 0$, every fixed $g_0, \dots, g_m \geq 0$, and every $\lambda \geq a$, set*

$$\Psi_\lambda^\theta(T) = (\lambda + s_{\text{tail}}(T)) \prod_{i \in T} e^\theta K_{b_i}(s_i(T)),$$

where only candidates $b_i \geq B_*$ may be selected. If V_λ is the supremum over all nonnegative candidate sizes of the minimum cost, then

$$\phi_\theta(V_\lambda) \leq \sum_{j=0}^{m-1} \phi_\theta(a + g_j) + \phi_\theta(\lambda + g_m). \quad (24)$$

Proof. The induction is simultaneous over all fixed-gap vectors and all terminal weights $\lambda \geq a$. It is immediate for $m = 0$. For an eligible inserted size b , the original insertion ratio is

$$\frac{e^\theta c(b)(a + \ell + \Delta(b))y}{b(\ell + b + y)}, \quad y \geq a. \quad (25)$$

It increases strictly in each adjacent mass: the derivative signs are $b + y - a - \Delta(b) > 0$ and $\ell + b > 0$. If an added checkpoint changes the next selected checkpoint, its new right terminal offset is $a + \Delta(b_{\text{new}})$; the old right mass included the full $b_{\text{new}} > \Delta(b_{\text{new}})$. Thus right mass still decreases when the background selection is enlarged. The tight-set conclusions in the proof of Lemma 3.7 hold for these costs as well. Also $K_b(S)$ strictly decreases in its own size b .

For $b \leq a/2$, the ratio in (25) is at least $a^2/[b(a + b)] \geq 4/3$, so such a size never occurs in a minimum. Unselectable sizes have the same property by definition. Deleting a size $b \leq a$ and merging its two gaps is covered by

$$\phi(a + g_L + g_R + b) \leq \phi(a + g_L) + \phi(a + g_R);$$

at the terminal gap replace the second a by λ . These follow from subadditivity. Thus the induction treats the eligibility boundary, and the minimum cost is continuous across it.

It remains to consider sizes $b_i > a/2$. The supremum on this region is finite. To see this, select a largest candidate. Its preceding mass is at most the total fixed mass plus $m - 1$ times that size, so its factor is uniformly bounded. The induction hypothesis bounds the minimum on its suffix. If some sizes tend to infinity, pass to a subsequence and select the first divergent one. Its prefix remains bounded and its factor tends to zero, while its suffix minimum is uniformly bounded by induction. The full minimum then tends to zero. Consequently a violation of (24) would have an interior maximum with all sizes greater than $a/2$.

At such a maximum both the empty set and a nonempty set are minimizers. If only the empty set were minimizing, increasing a size would improve the minimum. If every minimizing set were nonempty, their common candidate could be decreased: its selected factor strictly increases in every minimizing cost, again improving the minimum. Finite-set continuity justifies both perturbations. Write W for the empty cost at this maximum and choose a nonempty tight selection T .

For each $i \in T$, move $\Delta_i = \Delta(b_i)$ into its preceding gap and replace b_i by $B_i = b_i - \Delta_i$. This preserves every subset's total skipped mass and changes its selected factor to

$$\frac{e^\theta(c(b_i) - e(b_i))}{B_i}(a + s_i + \Delta_i).$$

Inflating its coefficient to $\chi_\theta(B_i)$ preserves empty-set optimality by (16). Treat the original unselected candidates as part of the fixed gaps and apply Lemma 3.7. If u_i are the original masses before the selected indices and u_{tail} the final mass, this gives

$$\phi(W) + \sum_{i \in T} \psi(B_i) \leq \sum_{i \in T} \phi(a + \Delta_i + u_i) + \phi(\lambda + u_{\text{tail}}).$$

Original tightness supplies an induction problem in each original interval between selected indices. The next selected factor is a fixed positive multiple of $a + \Delta_i + s_i$, so the terminal weight in that interval is exactly $a + \Delta_i$. Its empty selection is optimal; otherwise T could be improved. The final interval uses terminal weight λ . Apply the induction hypothesis to these smaller original problems. Their original fixed gaps partition $g_0, \dots, g_m$. At index i the additional term is

$$\phi(a + g + \Delta_i) - \phi(a + g) \leq \phi(\Delta_i) \leq \psi(B_i).$$

These increments are bounded by the terms $\psi(B_i)$ on the left, so $\phi(W)$ is at most the right side of (24), contradicting the violation. $\square$

Lemma 3.9 (Inverse potential and prefix bounds). *For $k \geq 1$ define*

$$\lambda_{k,\theta} = \phi_\theta^{-1}((k+1)\phi_\theta(a+1)). \quad (26)$$

For $k \geq 1$ and $0 \leq \theta \leq \log 4$,

$$ak^{p_\theta} \leq \lambda_{k,\theta} \leq (a+1)(k+1)^2.$$

As $k \rightarrow \infty$, uniformly in θ ,

$$\log \lambda_{k,\theta} = p_\theta \log k + \kappa_\theta \log \log k + O(1),$$

and $|\partial_\theta \log \lambda_{k,\theta}| = O(\log k)$. If there are at most k candidates and their fixed gaps have total mass at most $27k/64$, then every original candidate prefix has

$$\min_T \Psi_\mu^\theta(T) \leq 4 \max(\lambda_{k,\theta}, \mu) \quad (\mu \geq a). \quad (27)$$

Proof. The bound $z\phi'(z)/\phi(z) > 1/2$ implies $\phi(4t) \geq 2\phi(t)$ and $\lambda \leq (a+1)(k+1)^2$. The decreasing ratio $\phi(t)/t^{1/p}$ gives $\lambda \geq (a+1)(k+1)^p \geq ak^p$. These bounds first give $\log \lambda = \Theta(\log k)$. Taking logarithms of (26) and substituting this comparison proves the displayed expansion, uniformly

in θ . The exponential correction in (13) is bounded and tends to one. If $q_\theta(z) = z\phi'_\theta(z)/\phi_\theta(z)$, implicit differentiation gives

$$\partial_\theta \log \lambda = \frac{1}{q_\theta(\lambda)} \left[\frac{p'_\theta}{p_\theta^2} \log \frac{\lambda}{a+1} + A'_\theta \log \frac{\ell(\lambda)}{\ell(a+1)} \right],$$

which proves the derivative bound.

Pad the fixed-gap list to length $k+1$ with zeros. Concavity gives $G := \sum_j \phi(a+g_j) \leq (k+1)\phi(a+27k/[64(k+1)]) \leq \phi(\lambda)$. For every prefix, Proposition 3.8 at terminal weight μ has right side at most $G+\phi(\mu)$: use subadditivity for the terminal term. This is at most $2\phi(\max(\lambda, \mu))$, hence (27). $\square$

3.4 A finite insertion theorem

For $C > E \geq 0$, $E < 1$, let $x > e$ solve $x(\log x - 1) = C - E$ and put $u(C, E) = 2C/x$. The tangent inequality for the logarithm gives

$$2\log(Cz + E) \leq u(C, E)(1 + z) \quad (z \geq 0). \quad (28)$$

Indeed, the tangent at x has constant term $2\log x + 2E/x - 2 = 2C/x$. Equality holds at $z = (x - E)/C > 0$. For $H > 64a$, define

$$C_H = e^\theta c(H) \frac{H-a}{H}, \quad E_H = e^\theta e(H), \quad u_H = u(C_H, E_H). \quad (29)$$

Our fixed choice $a \geq 3B_*$ ensures $C_H > E_H$, $E_H < 1$. Also $C_H \geq 63/64$, so (28) at $z = 3$, or its equality characterization, gives $u_H \geq \frac{1}{2} \log(189/64) > 1/2$.

Theorem 3.10 (Finite checkpoint insertion). *Use the fixed constants and functions above. Let $0 \leq \theta \leq \log 4$, $N \geq k \geq 2$, $t = \log k$, $p = p_\theta$, $\lambda = \lambda_{k,\theta}$, and $K = 2N + 1$. Suppose an integer $M \geq 1$ and a real H satisfy*

$$H \geq k^2, \quad H > 64a, \quad M \log(1 + H/k^2) \geq 1 + \frac{2-p}{2}t.$$

If $d \geq M + 1$ and

$$2\theta(d-1) \geq \log(\lambda/a^2) + \log H + u_H \frac{MH}{H-a}, \quad (30)$$

then

$$V_{N,d}^+ \geq \frac{k^2}{128\lambda(2N+1)^2}. \quad (31)$$

Proof. Fix a nonnegative schedule with worst-case gap w strictly below the proposed bound and scale its steps by $\zeta = k^2/K^2$. Write $s_j = \zeta h_j$. Lemma 2.2 supplies at most k positive candidates whose complementary fixed-gap mass is at most $27k/64$. Its hypothesis holds because $w < k^2/(128\lambda K^2) < k^2/(2K^2)$. Using Lemma 2.1 at $y = 1/K$, and $\lambda \geq ak^p$, gives

$$\sum_j \log(1 + s_j/k^2) \leq K \operatorname{arsinh}(1/K) + \log(K\sqrt{2w}) < 1 + \frac{2-p}{2}t =: B_t.$$

Let $\mathcal{E} = \{j : s_j > H\}$, $m = |\mathcal{E}|$, and let R be the sum of the other steps. Concavity of $\log(1+x/k^2)$ on $[0, H]$ and monotonicity above H imply

$$\log(1 + x/k^2) \geq \frac{\min(x, H)}{H} \log(1 + H/k^2).$$

Consequently

$$R + mH \leq MH. \quad (32)$$

Every exceptional step is a candidate, since its size $H \geq k^2 > 27k/64$ exceeds the total omitted candidate mass.

Choose a minimizing selection I at terminal weight λ . Lemma 3.9 gives $\Psi_\lambda^\theta(I) \leq 4\lambda$. Write $A(I) = \prod_{i \in I} K_{b_i}(s_i)^{-1}$ and let T be its last gap. Since $\Psi_\lambda^\theta(I) = e^{\theta|I|}(\lambda + T)/A(I)$ and

$$\frac{(\lambda + T)^2}{4(1 + T)} \geq \lambda - 1,$$

its realized objective is at least $1/(32\lambda)$ whenever $|I| \leq d - 1$. This contradicts the assumption on w after rescaling by (6).

Otherwise $|I| \geq d$. Let I_r be the first r selected checkpoints. The integer $|I_r \cup \mathcal{E}|$ starts at $m \leq M \leq d - 1$, eventually exceeds $d - 1$, and increases by zero or one. Choose its largest fitting prefix $P = I_r$. There is a following selected checkpoint, and with $j = |\mathcal{E} \setminus P|$,

$$r + j = d - 1. \quad (33)$$

If the next selected size is b_{next} , put $\mu = a + \Delta(b_{\text{next}})$. Its factor is a fixed multiple of $\mu + u_P$, where u_P is the mass following P in the original candidate prefix. Replacing that prefix in I proves exact prefix optimality at terminal weight μ . Therefore (27) gives

$$A(P) \geq \frac{e^{\theta r}}{4} \frac{\mu}{\max(\lambda, \mu)} \geq \frac{e^{\theta r} a}{4\lambda}. \quad (34)$$

The last inequality also holds when P is empty.

Insert the j missing exceptional checkpoints in chronological order. Their preceding unselected masses v_i are ordinary and pairwise disjoint: every earlier exceptional step is already selected when a later one is inserted. The final tail is ordinary and disjoint from these intervals, so

$$\sum_i v_i + T \leq R. \quad (35)$$

Old selected factors decrease when their preceding gaps shrink. For each inserted $b_i > H$, decreasing c, e give

$$e^\theta K_{b_i}(v_i) \leq C_H \frac{a + v_i}{H - a} + E_H.$$

Apply (28), sum, and use (32)–(35):

$$2\theta j + 2 \sum_i \log K_{b_i}(v_i) \leq u_H \frac{jH + \sum_i v_i}{H - a} \leq u_H \frac{MH - T}{H - a}.$$

For every $T \geq 0$,

$$(1 + T)e^{-u_H T/(H-a)} \leq 1 + \frac{H - a}{eu_H} \leq H,$$

using $u_H > 1/2$ and $H > 64a \geq 128$. Combining this with (33), (34), and the exact objective (11), the selection $P \cup \mathcal{E}$ has value at least

$$\frac{e^{2\theta(d-1)} a^2}{64\lambda^2 H \exp\{u_H MH/(H - a)\}} \geq \frac{1}{64\lambda}.$$

It has exactly $d - 1$ checkpoints and uses at most d coordinates. Rescaling contradicts $w < \zeta/(128\lambda)$. $\square$

Define the positive Lambert function by $W(z)e^{W(z)} = z$, and set

$$u_\theta = 2W(e^{\theta-1}), \quad \eta_\theta = \sqrt{2(2-p_\theta)u_\theta}. \quad (36)$$

Corollary 3.11 (Asymptotic insertion bound). *There are fixed $t_0, C > 0$ such that, uniformly for $0 \leq \theta \leq \log 4$, the bound (31) holds whenever $N \geq k$, $t = \log k \geq t_0$, $d \geq M + 1$, where*

$$M = \left\lceil \sqrt{\frac{1 + (2-p_\theta)t/2}{u_\theta}} \right\rceil,$$

and

$$2\theta(d-1) \geq (p_\theta + 2)t + \kappa_\theta \log t + \eta_\theta \sqrt{t} + C. \quad (37)$$

Proof. Put $B_t = 1 + (2-p_\theta)t/2$ and $H = k^2 \exp(B_t/M)$. Then $M \log(1 + H/k^2) > B_t$ and $H > 64a$ for large fixed t_0 . The defining equation for $u(C, E)$ is smooth on a neighborhood of the compact set $(C, E) = (e^\theta, 0)$. Since $c(H) = 1 + O(1/\log H)$ and $e(H) = O((\log H)^{-4})$,

$$u_H \frac{H}{H-a} = u_\theta + O(1/\log H)$$

uniformly. Here $M = \Theta(\sqrt{t})$, so

$$\frac{B_t}{M} + u_H \frac{MH}{H-a} \leq 2\sqrt{B_t u_\theta} + u_\theta + O(M/t) = \eta_\theta \sqrt{t} + O(1).$$

Insert the uniform inverse expansion from Lemma 3.9 into (30). $\square$

Proof of Theorem 3.1. First take all N steps as candidates, all fixed gaps zero, and $\theta = 0$. Choose λ from $\phi_0(\lambda) = N\phi_0(a)$. Proposition 3.8 and $\phi_0(4t) \geq 2\phi_0(t)$ give a selection with $\Psi_\lambda^0 \leq 4\lambda$. The objective calculation in the proof of Theorem 3.10 gives $1/(32\lambda)$ in at most $N + 1$ coordinates. Inverting the explicit potential,

$$\lambda \asymp N^s [\log(N+2)]^{sA_0}, \quad sA_0 = \frac{\pi^2}{4Ds} = \kappa_0.$$

This proves (8); adjusting a fixed constant covers the initial horizons.

For the joint bound set $k = N$ in Corollary 3.11. Choose the small positive solution of

$$2\theta(d-1) = (p_\theta + 2)L + \kappa_\theta \log L + \eta_\theta \sqrt{L} + C. \quad (38)$$

For a sufficiently large fixed C_d and all large N this solution exists uniquely in a fixed small interval: the derivative of the right side is $O(L)$, while that of the left side is $2(d-1)$. The difference is negative at zero and positive at the fixed upper endpoint. Also $\theta = O(L/d)$ and $d \geq M + 1$. Direct differentiation gives

$$p'_\theta = \frac{1 - (1 + e^\theta)^{-1/2}}{2 \log 2}, \quad p_\theta = s + \theta/D + O(\theta^2).$$

The smooth functions $\kappa_\theta, \eta_\theta$ and (38) yield

$$\theta L = \frac{s+2}{2}X + O\left(\frac{X}{\sqrt{L}} + \frac{X^2}{L}\right).$$

Here $X \log L/L$ is absorbed in $X/\sqrt{L}$. Lemma 3.9 now gives

$$\log \lambda_{N,\theta} = sL + \kappa_0 \log L + a_+ X + O\left(1 + \frac{X}{\sqrt{L}} + \frac{X^2}{L}\right).$$

Apply (31) and $(2N+1)^2 \leq 9N^2$. Finally $X/\sqrt{L} \leq 1 + X^2/L$ gives (7) after changing c, C . $\square$

3.5 Logarithmic and sublogarithmic dimensions

For $0 < \theta \leq \log 4$, put

$$\beta_\theta = 2 - p_\theta, \quad b_\theta = \frac{2\theta}{p_\theta + 2}, \quad g_\theta = \beta_\theta b_\theta.$$

Theorem 3.12 (Optimized dimension curve). *There is a unique $\theta_* \in (0, \log 4)$ satisfying*

$$4 - p_{\theta_*}^2 = 4\theta_* p'_{\theta_*}. \quad (39)$$

Write $g_* = g_{\theta_*}$ and $c_* = 1/b_{\theta_*}$, and define

$$E_+(c) = \begin{cases} 2 - g_* c, & 0 < c \leq c_*, \\ p_{\theta_c}, & c \geq c_*, \end{cases} \quad 2c\theta_c = p_{\theta_c} + 2, \quad 0 < \theta_c \leq \theta_*.$$

If $d_N / \log N \rightarrow c \in (0, \infty)$, then $V_{N,d_N}^+ \geq N^{-E_+(c)-o(1)}$. For $d = \lceil c \log N \rceil$ the stronger statement is

$$V_{N,d}^+ \geq N^{-E_+(c)} \exp\{-O_c(\sqrt{\log N})\}.$$

Moreover, for every fixed $\theta \in (0, \log 4]$, as $d \rightarrow \infty$ with $d = o(\log N)$,

$$\log(N^2 V_{N,d}^+) \geq g_\theta d - q_\theta \sqrt{d} - \frac{4\kappa_\theta}{p_\theta + 2} \log d - O_\theta(1), \quad q_\theta = \frac{2\beta_\theta \sqrt{\beta_\theta \theta u_\theta}}{(p_\theta + 2)^{3/2}}. \quad (40)$$

In particular the leading sublogarithmic gain is $g_* \approx 0.2546088726$.

Proof. Fix $\theta > 0$ and put $u = \min(\log N, b_\theta d)$. For large u , choose $k = \lfloor \exp(u - C_\theta \sqrt{u}) \rfloor$. A sufficiently large fixed C_θ makes (37) hold; its positive margin of order $\sqrt{u}$ absorbs the shortening and logarithmic terms. Also $d \geq M + 1$ eventually. The finite theorem gives

$$\log(N^2 V_{N,d}^+) \geq \beta_\theta \min(\log N, b_\theta d) - O_\theta(\sqrt{u}). \quad (41)$$

We optimize this expression explicitly. Its derivative formulas are

$$p'_\theta = \frac{1 - (1 + e^\theta)^{-1/2}}{2 \log 2}, \quad p''_\theta = \frac{e^\theta}{4 \log 2 (1 + e^\theta)^{3/2}} > 0.$$

The sign of g'_θ is the sign of $F(\theta) = 4 - p_\theta^2 - 4\theta p'_\theta$. Since $F' = -(2p_\theta + 4)p'_\theta - 4\theta p''_\theta < 0$, there is at most one critical point. We have $F(0) > 0$. At $\theta = \log 4$, $p_\theta > 3/2$ and $4\theta p'_\theta = 4(1 - 1/\sqrt{5}) > 2$, so $F(\log 4) < 0$. This proves (39) and unique maximality. The function b_θ strictly increases, because $\theta p'_\theta \leq 1 < p_\theta + 2$ on this interval. Before saturation, the gain in (41) is cg_θ ; after saturation it is β_θ , which strictly decreases. If $c \leq c_*$ the maximum is at θ_* ; if $c > c_*$ it is at the unique saturation point θ_c . This proves both assertions for logarithmic dimensions. For the finer sublogarithmic statement, abbreviate $p = p_\theta$, $\beta = \beta_\theta$, $\kappa = \kappa_\theta$. Choose an integer k with

$$\log k = b_\theta d - \frac{\eta_\theta \sqrt{b_\theta}}{p + 2} \sqrt{d} - \frac{\kappa}{p + 2} \log d - O_\theta(1).$$

Substitution in (37) proves it for a sufficiently large constant in the last term; all remaining errors are bounded. Here $k \leq N$ and $d \geq M + 1$ eventually. The finite theorem and inverse expansion give $\log(N^2 V_{N,d}^+) \geq \beta \log k - \kappa \log \log k - O_\theta(1)$. The square-root coefficient is $\beta \eta_\theta \sqrt{b_\theta} / (p + 2) = q_\theta$ and the logarithmic coefficient is $\kappa + \beta \kappa / (p + 2) = 4\kappa / (p + 2)$. $\square$

4 Real schedules: construction and selection

We use the increasing kernel

$$\kappa(z) = \begin{cases} \sqrt{z}(1+z), & 0 < z \leq 1, \\ 2z, & z \geq 1, \end{cases} \quad k(z) = \log \kappa(z). \quad (42)$$

Its logarithm is strictly concave, and $1/2 \leq z\kappa'(z)/\kappa(z) \leq 1$. In particular,

$$\kappa(y)/\kappa(x) \geq y/x \quad (0 < y \leq x), \quad \kappa(z) \geq \sqrt{z}. \quad (43)$$

4.1 A smooth interpolant for every selected subset

We record a finite conjugate construction based on the smooth convex interpolation framework of Taylor–Hendrickx–Glineur [8, Sections 2.3–2.5], in a form that also controls the gradient range.

Lemma 4.1 (Finite conjugate interpolation). *Let $v_0 = 0, v_1, \dots, v_m$ be vectors, and let $a_0 = 0, a_1, \dots, a_m \geq 0$. On $C = \text{conv}\{v_j\}$ put*

$$q(v) = \min \left\{ \sum_j \lambda_j a_j : \lambda_j \geq 0, \sum_j \lambda_j = 1, \sum_j \lambda_j v_j = v \right\},$$

$$f(x) = \max_{v \in C} \{ \langle x, v \rangle - \|v\|^2/2 - q(v) \}.$$

Then f is convex and 1-smooth, $f(0) = \min f = 0$, and its gradient range is contained in C . If

$$a_j \geq a_i + \langle x - v_i, v_j - v_i \rangle \quad (0 \leq j \leq m), \quad (44)$$

then $\nabla f(x) = v_i$ and $f(x) = \langle x, v_i \rangle - \|v_i\|^2/2 - a_i$.

Proof. The affine function in (44) lies below every specified height and hence below their lower convex envelope. At v_i it equals a_i , proving $q(v_i) = a_i$ and $x - v_i \in \partial q(v_i)$. The function $q + \|\cdot\|^2/2$, extended by $+\infty$ off C , is 1-strongly convex. Its conjugate is differentiable with 1-Lipschitz gradient, and the displayed subgradient condition identifies the maximizer. Nonnegativity of q and the vertex 0 give the assertions at the origin. $\square$

The next construction follows the selected-block geometry of Ma–Chen [2, Theorem 4.1] and Ye–Liu [3, Appendix A]. We assemble the translated local blocks in one conjugate interpolant, with damping factors that give the numerator $b_i - 1$.

Lemma 4.2 (An interpolant for a selected subset). *Let $s_1, \dots, s_N \geq 0$. Select any checkpoints $b_1, \dots, b_q > 2$ in their temporal order, with intervening unselected masses $S_1, \dots, S_q$ and final mass T . For every offset $a \geq 1$ there is a convex 1-smooth function in $q + 1$ coordinates, with minimizer 0 and initial point of norm 1, whose terminal value is at least*

$$Q_a(I) = \frac{1}{4(a+T)} \prod_{i \in I} \kappa \left(\frac{a + S_i}{b_i - a} \right)^{-2}, \quad (45)$$

provided $b_i > a$. Every queried gradient g , including the terminal gradient, satisfies

$$\langle g, v \rangle \leq \|g\|^2 \quad \text{for every } v \in \nabla f(\mathbb{R}^{q+1}). \quad (46)$$

Proof. For one block choose $\epsilon > 0$ with $b \geq 1 + \epsilon^{-2}$, and set

$$g = (\epsilon, -1), \quad H = \max\{\epsilon(S+b), (\epsilon + \epsilon^{-1})(S+1)\}.$$

For $0 \leq \delta < 1$ put

$$\rho = (1 - \delta)/H, \quad z_0 = (1, 0) - \rho(1 + S)g, \quad U = 1 - \rho\epsilon(S+b), \quad Y = \rho S.$$

The conjugate of the quadratic on the segment $[0, \rho g]$, with linear height $z_0 \cdot v$, has gradient ρg along the entire gap path $(1, 0) - \rho u g$, $0 \leq u \leq S$. After the checkpoint its point is $(U, Y + \rho b)$. It vanishes on both rectangles

$$\{z_1 \leq \delta, z_2 \geq 0\}, \quad \{z_1 \leq U, z_2 \geq Y\}. \quad (47)$$

Indeed $z_0 \cdot g \geq \epsilon\delta$, and $((U, Y) - z_0) \cdot g = \rho((1-b)\epsilon^2 + 1) \leq 0$. Also $U \geq \delta$. At every active query the residual $z - \rho g$ has first coordinate at least δ , second coordinate at most $Y + \rho$, and $(z - z_0 - \rho g) \cdot \rho g \geq 0$.

For the successive blocks take $\delta_1 = 0$, $\delta_i = 1/b_{i-1}$ for $i > 1$, and define

$$D_1 = 1, \quad \ell_1 = 0, \quad D_{i+1} = D_i \rho_i b_i, \quad \ell_{i+1} = D_i Y_i.$$

Embed the block gradient as $v_i = D_i \rho_i (\epsilon_i e_i - e_{i+1})$, with height

$$a_i = (\ell_i e_i + D_i z_{0,i}) \cdot v_i.$$

Here $z_{0,i}$ is embedded in coordinates $i, i+1$. These heights are nonnegative by the first rectangle. Block i starts at normalized coordinates $((x_i - \ell_i)/D_i, x_{i+1}/D_i) = (1, 0)$. Earlier first coordinates are frozen at their post-checkpoint values; more distant future coordinates are zero.

We check (44) at an active query. The own-vertex comparison with zero is the nonnegative local residual just proved. Every other vertex is inactive at the residual $x - v_i$. For the next block its first residual coordinate is at most $\ell_{i+1} + D_i \rho_i = \ell_{i+1} + D_{i+1}/b_i$, the boundary of its first rectangle. For the previous block its first coordinate is frozen and its second residual coordinate is at least ℓ_i , placing it in its second rectangle. More distant blocks are in the corresponding zero rectangles as well. Inactivity says $a_j \geq (x - v_i) \cdot v_j$, while the own comparison says $(x - v_i) \cdot v_i - a_i \geq 0$; together they imply (44).

Write $D = D_{q+1}$ and $\ell = \ell_{q+1}$. If $q > 0$, add a terminal vertex ce_{q+1} of height zc , where

$$z = \ell + D/b_q, \quad E = D(1 - 1/b_q), \quad c = E/(1 + 2T).$$

For $q = 0$ use $z = 0, E = 1, c = (1 + 2T)^{-1}$. At queries of the last block the terminal residual coordinate is at most z . During the terminal gap the gradient is ce_{q+1} ; previous blocks are inactive because $E - c(1 + u) \geq 0$ for $0 \leq u \leq T$. Thus the same interpolation inequalities hold at every query and at the terminal point. Lemma 4.1 gives one global interpolant, with terminal value

$$\frac{E^2}{2(1 + 2T)} = \frac{1}{2(1 + 2T)} \prod_{i=1}^q \left(\frac{b_i - 1}{H_i} \right)^2. \quad (48)$$

The damping factors telescope to the numerators $b_i - 1$.

For $z_i = (a + S_i)/(b_i - 1)$, choose $\epsilon_i = \sqrt{z_i}$ when $z_i \leq 1$ and $\epsilon_i = 1$ otherwise. Eligibility follows from $a \geq 1$ and $b_i > 2$. Directly,

$$\frac{H_i}{b_i - 1} \leq \inf_{\epsilon > 0} \max\{\epsilon(1 + z_i), (\epsilon + \epsilon^{-1})z_i\} = \kappa(z_i).$$

Since $b_i - a \leq b_i - 1$ and $2(1 + 2T) \leq 4(a + T)$, this proves (45). Finally, different embedded block gradients have nonpositive inner products, and the terminal vertex has the same property. Each vertex therefore satisfies $v_i \cdot v \leq \|v_i\|^2$ on the whole gradient hull. $\square$

The next lemma is Ye-Liu's running-maximum lifting argument [3, Lemma D.1], expressed under the gradient-support condition needed for our interpolant.

Lemma 4.3 (Transfer to real schedules). *For a real schedule h , set*

$$H_t = \sum_{j \leq t} h_j, \quad M_t = \max(0, H_1, \dots, H_t), \quad r_t = M_t - M_{t-1}, \quad \Delta_t = M_t - H_t.$$

If the gradients queried by the schedule r , including its terminal gradient, satisfy (46), then the trajectories under h and r , from the same initial point, satisfy

$$x_t^h = x_t^r + \Delta_t \nabla f(x_t^r), \quad f(x_N^h) = f(x_N^r) + \Delta_N \|\nabla f(x_N^r)\|^2.$$

Consequently (45) applies to the record increments of every real schedule. A witness for ηr , $0 < \eta \leq 1$, gives the lower bound ηQ_a for h .

Proof. If $g = \nabla f(x)$ and $g \cdot v \leq \|g\|^2$ for every gradient v , convexity and integration of directional derivatives give $f(x + tg) = f(x) + t\|g\|^2$ and $\nabla f(x + tg) = g$ for $t \geq 0$. The stated identities follow by induction: at a new cumulative record the trajectories meet, and otherwise the trajectory under r stays fixed while the deficit changes. For the scaled assertion apply the construction to ηr and use the function ηf , whose smoothness is at most one. $\square$

4.2 Suffix minima and their scalar bounds

For $u, r > 0$ define

$$\gamma(u) = \sup_{z > 0} \{2k(z) - uz\}, \quad g(r) = \gamma(2r)/2. \quad (49)$$

The maximizer is $2/u$ for $u \leq 2$, and for $u > 2$ it is $z = 2/(\sqrt{u^2 - 2u + 9} + u - 3)$. Hence

$$\gamma(u) = \begin{cases} \log 16 - 2 \log u - 2, & u \leq 2, \\ \log z + 2 \log(1 + z) - uz, & u > 2. \end{cases} \quad (50)$$

It is convex and strictly decreasing, with $\gamma(u) \geq -\log u - 1$. The maximizer of $k(t/r) - t$ is 1 for $r \leq 1$. For $r > 1$ it is the unique $m \in (1/2, 1)$ satisfying $r = m(3 - 2m)/(2m - 1)$.

Lemma 4.4 (Suffix profiles). *Let $F(t) = \min_j a_j \kappa((t + C_j)/u_j)$, with a finite nonempty index set, $a_j, u_j > 0$, $C_j \geq 0$, and at least one $C_j = 0$. Then $\log F(t) - t$ has a unique maximizer $m \in (0, 1]$. If its maximum is M , then*

$$M - \log F(t) + t \geq \frac{3}{8}(m - t)_+^2. \quad (51)$$

Every bound on $\kappa((m(u) + A)/u)/\kappa(m(u)/u)$, uniform in $u > 0$, also bounds $F(m + A)/F(m)$.

Proof. A finite minimum of strictly concave functions is strictly concave. The zero-shift branch makes $\log F(t) - t$ tend to $-\infty$ at zero; the same holds at infinity. Each branch derivative is at most $1/t$, so $m \leq 1$. Before the maximizer every active right derivative is at least one. For an active branch this puts $x = t + C$ before its single-factor maximizer. On the mixed branch the negative logarithmic curvature is $1/(2x^2) + 1/(u + x)^2$; at its maximizer $q \in (1/2, 1)$ this equals

$1 - 1/q + 3/(4q^2) \geq 3/4$. The pure branch and the case $u \leq 1$ give the same lower bound. Concavity makes derivative jumps at switches nonpositive. Integrating this curvature bound backwards from m proves (51).

At m choose an active branch with derivative at most one. Its argument $x = m + C$ is at least its single-factor maximizer $m(u)$. Concavity of k yields

$$\frac{F(m+A)}{F(m)} \leq \frac{\kappa((x+A)/u)}{\kappa(x/u)} \leq \frac{\kappa((m(u)+A)/u)}{\kappa(m(u)/u)}.$$

The arguments use one-sided derivatives and include switches of the minimizing branch. $\square$

Lemma 4.5 (Growth bounds). *For every suffix profile and $A \geq 2/5$,*

$$F(m+A)/F(m) \leq 19/20 + 7A/6. \quad (52)$$

Put $R = 69809/100000$. There are also the two chord bounds

$$F(m+A)/F(m) \leq a_g + b_g A,$$

obtained from the endpoints

$$(2/5, 17/12), (1+R, 2/R), \quad \text{or} \quad (11/10, 109423/50000), (1+R, 2/R), \quad (53)$$

on the respective closed intervals between those endpoints.

Proof. For $0 < u \leq 1$ the normalized single-factor growth is $1 + A$. For $u > 1$ write $m = m(u)$. After the mixed-to-pure junction it is $C(m)(m+A)$, where

$$C(m) = \frac{\sqrt{(2m-1)(3-2m)}}{m} \leq \frac{2}{\sqrt{3}}.$$

Before the junction it is $\sqrt{1+A/m}(1+A-A/(2m))$. At fixed A this increases in m , as differentiation with respect to $x = A/m$ gives the sign of $A - 3x/2 < 0$. The mixed expression therefore increases up to its junction with the pure expression. Both are bounded by

$$M(A) = \max_{1/2 \leq m \leq 1} C(m)(m+A).$$

The function M is convex. The derivative of single-factor growth in A is at most $2/\sqrt{3} < 7/6$: on a mixed segment its maximum is at an endpoint, since κ' first decreases and then increases; the endpoint derivatives are 1 and $C(m)$.

To bound $M(2/5)$ by $17/12$, for $m \leq 3/4$ its square is at most $529/300 < 289/144$. For $m \geq 3/4$ subtract its squared numerator from $(17m/12)^2$, obtaining

$$P(m) = 4m^4 - \frac{24}{5}m^3 - \frac{2711}{3600}m^2 + \frac{28}{25}m + \frac{12}{25}.$$

Here $P'' \geq 7009/1800$, and at $m_0 = 23/25$,

$$P(m_0) - \frac{P'(m_0)^2}{2(7009/1800)} = \frac{120427647}{136894531250} > 0.$$

This proves (52). For the other endpoints use

$$P_{A,U}(m) = U^2 m^2 - (2m-1)(3-2m)(m+A)^2.$$

Its third derivative is $96m + 48(A-1)$. The following exact rational checks imply positivity on $[1/2, 1]$ by strong convexity:

(A, U)	m_0	lower bound on P''	$P(m_0) - P'(m_0)^2 / (2 \inf P'')$
$(1 + R, 2/R)$	821/1000	19	$> 1/100000$
$(11/10, 109423/50000)$	849/1000	4	$> 1/100000$

The arithmetic implementation accompanying the paper evaluates these fractions exactly. Convexity of M supplies both chords, and Lemma 4.4 transfers them to every suffix profile. $\square$

For positive normalized gaps s_i and checkpoint excesses r_i , let $C_q^{[\tau]}$ be the minimum, over subsets containing q , of $\exp(\tau|I|/2)$ times the product of the kernel factors after all omitted checkpoints are merged into their gaps. For fixed future gaps it is a suffix profile: if the first selected checkpoint is j , its first factor has shift $\sum_{i=2}^j s_i + \sum_{i=1}^{j-1} r_i \geq 0$; choosing checkpoint 1 gives shift zero. The exact recursion is

$$C_q^{[\tau]}(s_1, s_2, \dots) = \min\{e^{\tau/2} \kappa(s_1/r_1) C_{q-1}^{[\tau]}(s_2, \dots), C_{q-1}^{[\tau]}(s_1 + r_1 + s_2, \dots)\}. \quad (54)$$

All future gaps in the two suffixes agree. The last checkpoint is retained in both branches.

Lemma 4.6 (Deletion test). *Let $0 < r \leq 1$ and $c_L > 0$. Suppose $k(s/r) - s$ has maximum $g(r)$ at $s = 1$ and loss at least $c_L(1-s)_+^2/2$. Let F be a suffix profile, and let m and M be the maximizer and maximum of $\log F(t) - t$. Let $\tau, e(r) \geq 0$ and set $v = \sqrt{(2/c_L + 8/3)e(r)}$. If a valid suffix growth line is $a_g + b_g A$ at $A = 1 + r - v > 0$, and*

$$\tau/2 + \log \frac{2}{r[a_g + b_g(1 + r - v)]} - v - e(r) \geq 0, \quad (55)$$

then

$$\begin{aligned} \sup_{s,t>0} [\min\{\tau/2 + k(s/r) + \log F(t), \log F(s + r + t)\} - s - t] \\ \leq g(r) + \tau/2 + M - e(r). \end{aligned}$$

For $e(r) = 0$ the retained branch alone proves this conclusion.

Proof. If the retained branch exceeds the asserted bound, its two losses sum to less than $e(r)$. Weighted Cauchy gives $(1-s)_+ + (m-t)_+ < v$. The deleted argument exceeds $m + 1 + r - v$. Since $\log F(x) - x$ decreases after m , the deleted branch is at most $M + \log(a_g + b_g(1 + r - v)) - 1 + v$ after subtracting $s + t$. Comparison with $g(r) = \log(2/r) - 1$ is (55). $\square$

4.3 The critical scalar line

Throughout this subsection set

$$\begin{aligned} R &= \frac{69809}{100000}, \quad a = \frac{95703}{50000}, \quad b = \frac{103}{80}, \quad L(u) = a - bu, \\ e(r) &= \begin{cases} [\log(2/r) - 1 - a/2 + br]_+, & r < R, \\ 0, & r \geq R. \end{cases} \end{aligned} \quad (56)$$

Lemma 4.7 (Critical deletion inequality). *The unweighted recursion satisfies*

$$\sup_{s_i > 0} \left(\log C_q^{[0]} - \sum_i s_i \right) \leq \sum_i g(r_i) - \sum_{i < q} e(r_i), \quad r_i \geq 41/100. \quad (57)$$

Proof. We first prove an auxiliary bound on the low range. Let $D_R(r) = r/R - 1 - \log(r/R)$ for $r < R$ and zero otherwise, and $k_0 = 7/4$. For $r < R < \sqrt{3}/2$ the left premaximum curvature is at least one: it is at least 1 on the pure branch and at least $3/(4r^2) > 1$ on the mixed branch. Thus Lemma 4.6 uses $v = \sqrt{14k_0 D_R(r)/3}$.

Use the first chord in (53), and write $c = b_g R/2 = 1213247/3115416$, $\delta = 1 - r/R$. The directed deficit for loss $k_0 D_R$ is

$$\delta - v - \log(1 - c(R\delta + v)) - (k_0 - 1)D_R.$$

Here $0 \leq \delta \leq \delta_* = 28809/69809$ when $r \geq 41/100$. The continuous quotient $q(\delta) = D_R/\delta^2$ is bounded by

$$Q(\delta) = \sum_{j=0}^{11} \frac{\delta^j}{j+2} + \frac{\delta^{12}}{14(1-\delta)}.$$

For rational $x > 0$ define $\text{su}(x) = 2^{-40}(1 + \lfloor \sqrt{[2^{80}x]} \rfloor) > \sqrt{x}$ and put $u_* = \text{su}(14k_0 Q(\delta_*)/3)$. Exact fractions give

$$c\delta_*(R + u_*) < 1 - c, \quad 1 + 41/100 - \delta_* u_* > 2/5.$$

The deficit divided by δ decreases in $u = v/\delta$ on this rectangle. On each of the 128 equal rational cells $[l, h]$ of $[0, \delta_*]$, set $u_h = \text{su}(14k_0 Q(h)/3)$. The positive logarithm series gives the uniform rational lower bound

$$1 - u_h + \sum_{j=1}^{12} \frac{c^j l^{j-1} (R + u_h)^j}{j} - (k_0 - 1)hQ(h) > 1/1000. \quad (58)$$

The certificate evaluates every fraction in (58); at $l = 0$ the $j = 1$ term is interpreted as usual. This proves the auxiliary deletion loss $k_0 D_R$ on its entire stated domain.

Put $\rho_0 = 61/100$ and

$$a_0 = \log 4 - \frac{7}{2} \log R + \frac{3}{2} \log \rho_0, \quad b_0 = \frac{7}{4R} - \frac{3}{4\rho_0}.$$

The function $\gamma(u) - 2k_0 D_R(u/2)$ is concave below $2R$; its tangent at $2\rho_0$ is $a_0 - b_0 u$. The exact checks $b > b_0$ and $a - a_0 - (11/10)(b - b_0) > 1/20000$ prove that L dominates this tangent for $41/100 \leq r \leq 11/20$. Thus $e(r) \leq k_0 D_R(r)$ there.

For the remaining range set $r_s = 11/20$. The signed expression in (56) decreases on $[r_s, R]$, since $bR < 1$, and has a unique zero r_z there. On $[r_s, r_z)$ write $v = \sqrt{14e/3}$ and use the second chord in (53). Its coefficients are

$$a_g = \frac{197121741540463}{208760324050000}, \quad b_g = \frac{4722579586}{4175206481}, \quad c_g = b_g R/2 = \frac{2361289793}{5980900000}.$$

For $d(r) = \frac{1}{2}(1 - br)^2 - e(r)$, one has $d'(r) = (1 - br)^2/r \geq 0$ and $d(r_s) > 1/2000$. It follows from $v^2 = 14e/3$ that $v' < 0$ and $v'' \leq 0$. Moreover $1 + r_s - v(r_s) > 11/10$, so every growth argument is in the new chord's interval. Put $H = a_g + b_g(1 + r - v)$ and $F = \log(2/(rH)) - v - e$. Direct differentiation gives

$$F' = -b - b_g/H + (b_g/H - 1)v', \quad F'' = b_g^2(1 - v')^2/H^2 + (b_g/H - 1)v'' \geq c_g^2. \quad (59)$$

Here $b_g < 109423/50000 \leq H \leq 2/R$. At $r_0 = 6207/10000$ the interval certificate proves $F(r_0) - F'(r_0)^2/(2c_g^2) > 1/10000$. Strong convexity proves the directed test over the whole active interval. The zero-loss range uses retention alone.

Apply this one-step inequality with each fixed future gap before taking its supremum. Induction in the exact recursion (54) proves (57). The base $q = 1$ has no deletion loss, as displayed. $\square$

Let U be the upper solution of $L(U) = \gamma(U)$. It satisfies

$$1.7974064665351477 < U < 1.7974064665351478, \quad 2R < U < 2, \quad bU > 2. \quad (60)$$

Define

$$\phi(u) = \begin{cases} a - bu, & u \leq U, \\ \gamma(u), & u \geq U, \end{cases} \quad E(A) = \sup_{u \geq A} (\gamma(u) - \phi(u))_+. \quad (61)$$

The upward derivative jump at U makes ϕ convex and decreasing. It majorizes $\gamma(u) - 2e(u/2)$; on $[2R, U]$ this follows from concavity of $L - \gamma$ and its endpoint values. Since $U < 2$,

$$E(A) = \begin{cases} \max\{\gamma(A) - a + bA, 0\}, & A < U, \\ 0, & A \geq U. \end{cases} \quad (62)$$

In particular E is nonnegative and nonincreasing.

4.4 Normalization and subset selection

We adapt Ye–Liu’s cutoff choice and convex majorization argument [3, Section 3.1 and Lemma 3.1] to the kernel and deletion bound above. The following statement includes the empty selection and retains a bound for the final gap.

Proposition 4.8 (Subset-selection bound). *Fix a nonnegative schedule and a prescribed set of candidates exceeding 2. Use offset $a_{\text{off}} = 1$ or 2 in (45). For $\alpha, \mu > 0$ and $\tau \geq 0$, arrange the candidate excesses $b_i - a_{\text{off}}$ in decreasing order $B_1 \geq \dots \geq B_r$, and put*

$$D_j = a_{\text{off}} + \sum_t s_t - \sum_{i=1}^j B_i, \quad q \in \arg \min_{0 \leq j \leq r} D_j(j+1)^\alpha.$$

Suppose the weighted recursion obeys

$$2 \log C_q^{[\tau]} - 2 \sum_i s_i \leq \sum_i \gamma(2r_i) + \tau q - 2 \sum_{i < q} e(r_i),$$

and ϕ is convex and decreasing with $\phi(u) \geq \gamma(u) - 2e(u/2)$. Put $E(A) = \sup_{u \geq A} (\gamma(u) - \phi(u))_+$ and

$$J(\alpha, \mu) = \mu + \int_0^1 \phi(\mu \alpha t^{-1-\alpha}) dt.$$

Assume the directed inequality holds for $r_i \geq \mu\alpha/2$, $E(\mu\alpha) < \infty$, and $\phi(u) = \gamma(u)$ for $u \geq \mu\alpha 2^{1+\alpha}$. If $J \leq -b_*$, $b_* \geq 0$, and $b_* \leq 1 + \alpha + E(\mu\alpha) + \log \alpha$, an actual subset I with at most q checkpoints satisfies

$$e^{-\tau|I|} Q_{a_{\text{off}}}(I) \geq \frac{\exp(b_*(q+1) - \tau q)}{4\alpha \exp(1 + \alpha + E(\mu\alpha)) D_q(q+1)^\alpha}. \quad (63)$$

For $q > 0$ it contains the last of the q cutoff checkpoints, and the same competitor satisfies

$$a_{\text{off}} + T_I \leq D_q \leq D_r(r+1)^\alpha. \quad (64)$$

For the critical line the deletion inequality applies throughout the required domain, and $\phi = \gamma$ above $\mu\alpha 2^{1+\alpha}$, when $\alpha \geq 1/2$, $\mu\alpha \geq 41/50$, and $\tau = 0$.

Proof. Suppose first that $q \geq 1$, and retain the q largest excesses in their temporal order. Write $m = q + 1$, $D = D_q$, and normalize

$$x_i = \frac{m(a_{\text{off}} + S_i)}{D}, \quad w_i = \frac{m(b_i - a_{\text{off}})}{D}, \quad \Delta = \frac{m(a_{\text{off}} + T)}{D}, \quad \sum_i x_i + \Delta = m.$$

The minimizing property of q , including comparison with $j = 0$, gives for $0 \leq j < q$

$$\sum_{i=j+1}^q w_i^\downarrow \geq m \left[\left(\frac{m}{j+1} \right)^\alpha - 1 \right].$$

These are the lower-tail sums of $\bar{w}_i = m \int_{i/m}^{(i+1)/m} \alpha t^{-1-\alpha} dt$, $1 \leq i \leq q$. In particular every $w_i \geq \alpha$. Weak majorization for a convex decreasing function, followed by Jensen on each cell, gives

$$\sum_i \phi(\mu w_i) \leq \sum_i \phi(\mu \bar{w}_i) \leq m \int_{1/m}^1 \phi(\mu \alpha t^{-1-\alpha}) dt. \quad (65)$$

For completeness, the majorization implication follows by expressing a convex decreasing function on a compact interval containing the weights as an affine decreasing function plus nonnegative multiples of $(c - u)_+$, then using the lower-tail inequalities for each hinge.

Apply the directed recursion to gaps $\mu x_i/2$ and excesses $\mu w_i/2$. The missing last subtraction is at most $E(\mu w_q) \leq E(\mu \alpha)$, so

$$2 \log C_q^{[\tau]} \leq \mu(m - \Delta) + \sum_i \phi(\mu w_i) + \tau q + E(\mu \alpha).$$

The last checkpoint and its terminal mass have not changed. On the omitted interval $0 < t \leq 1/m$, the argument is at least $\mu \alpha m^{1+\alpha} \geq \mu \alpha 2^{1+\alpha}$, so $\phi = \gamma$ by hypothesis. Using $\gamma(u) \geq -\log u - 1$ gives, with $p = 1 + \alpha$,

$$-m \int_0^{1/m} \gamma(\mu \alpha t^{-p}) dt \leq p \log m + \log(\mu \alpha) + p + 1.$$

Substitute (65) and use $e^{\mu \Delta}/\Delta \geq e\mu$ in the factor $1/(4T_a)$, where $T_a = D\Delta/m$. This yields (63). The identity for merged gaps is $(a_{\text{off}} + S_i) + (b_i - a_{\text{off}}) + (a_{\text{off}} + S_{i+1})$; hence the subset minimum is an actual schedule selection.

When $q = 0$, the empty envelope is $1/(4D_0)$ and the stated empty condition gives the same bound. Finally $a_{\text{off}} + T_I \leq D_q$, and cutoff minimality gives $D_q(q+1)^\alpha \leq D_r(r+1)^\alpha$, proving (64). $\square$

For an unweighted recursion, $J \leq -b_*$ and $|I| \leq q$ imply $e^{-b_*|I|} Q_{a_{\text{off}}}(I) \geq [C_0 D_r(r+1)^\alpha]^{-1}$ whenever $C_0 \geq 4\alpha \exp(1 + \alpha + E(\mu \alpha))$. The factor $e^{b_*(q+1)}$ therefore bounds the penalty $e^{b_*|I|}$ for this subset.

4.5 Critical parameter and full-dimensional bound

For the line (56) put

$$A_{\dagger} = \frac{a}{1+b} = \frac{31901}{38125}, \quad G(\alpha) = J(\alpha, A_{\dagger}/\alpha). \quad (66)$$

Thus $\phi(A_{\dagger}) = A_{\dagger}$ and $A_{\dagger} > 5/6$. Integration by parts gives

$$p\mu J_{\mu} = J + A - \phi(A), \quad A = \mu\alpha, \quad \left(\frac{\alpha}{p}G\right)' = \frac{\alpha}{p}J_{\alpha} < 0. \quad (67)$$

The tail is logarithmic, so boundary terms at $t = 0$ vanish and the differentiated integrals are dominated by a constant times $1 + |\log t|$. Strict convexity in μ follows from the interval near zero where both arguments lie in the strictly convex part of γ . At a zero of G , (67) gives $J_{\mu} = 0$.

Proposition 4.9 (Scalar certificate). *There is a unique zero $\alpha_{\dagger}$ of G in the interval*

$$0.5570593 < \alpha_{\dagger} < 0.5570595. \quad (68)$$

At this zero, $\mu = A_{\dagger}/\alpha_{\dagger}$ is the unique minimizer of $J(\alpha_{\dagger}, \mu)$. Also

$$4.01664 < B_{\dagger} := -J_{\alpha} < 4.01666, \quad a_{\dagger} := \frac{3 + \alpha_{\dagger}}{B_{\dagger}} < \frac{8}{9}. \quad (69)$$

The coefficient in (63) is below 20 on a neighborhood of this point. The full-dimensional prefactor with offset one is above $1/16$.

Proof. Here are the integrals and finite error bounds used in the certificate. For $\alpha > 0$, $p = 1 + \alpha$, and $0 < A < 2$, define

$$\ell(c) = \frac{1}{p} \log \frac{Ac^2(1+c^2)}{1+3c^2}, \quad I(\alpha, A) = \int_0^1 \frac{2ce^{\ell(c)}}{(1+c^2)^2} dc, \quad L_1(\alpha, A) = \int_0^1 \frac{2ce^{\ell(c)}\ell(c)}{(1+c^2)^2} dc.$$

A change of variable using the mixed-branch maximizer in (50) gives

$$J_{\gamma}(\alpha, A/\alpha) = A/\alpha + \gamma(A) - 2p + 2pI(\alpha, A).$$

For $0 < A < U \leq 2$ and a line joined to γ at U , put $z = A/U$, $T = z^{1/p}$, $\beta = bU$. Direct integration gives the signed correction

$$B(\alpha) = \int_0^1 (\gamma - \phi)(At^{-p}) dt = \frac{\beta}{\alpha}(T - z) + (\beta + 2p)(T - 1) - 2 \log z. \quad (70)$$

It includes the interval where $\phi > \gamma$. Thus $G = J_{\gamma} - B$ and $-G' = A/\alpha^2 + 2 - 2I + 2L_1 + B'$.

The accompanying programs `signed_bounds.py` and `signed_moments.py` evaluate these formulas with 256-bit Arb arithmetic. On $[\delta, 1]$, $\delta = 1/2048$, they use 2048 midpoint panels and bound the error on each panel by its width cubed times the supremum of the absolute second derivative divided by 24. On $[0, \delta]$, writing $r = 1/p$, the respective absolute tail bounds are

$$T_I = \frac{A^r \delta^{2r+2}}{r+1}, \quad T_L = T_I r \left[-\log A - 2 \log \delta + \log(1 + 3\delta^2) + \frac{1}{r+1} \right].$$

The integrand of L_1 is nonpositive. Differentiating the displayed integrands supplies the second derivatives explicitly in the code. All these calculations use outward interval bounds.

The results are

$$G(0.5570593) > 5.23 \cdot 10^{-7}, \quad G(0.5570595) < -2.11 \cdot 10^{-7}.$$

Three moment calls, the last over the whole alpha bracket, prove (69). They also prove, with $\eta = E(A_{\dagger}) = \gamma(A_{\dagger}) - A_{\dagger}$,

$$\eta < 0.292307, \quad E(5/6) < \log 10 - 2, \quad 4\alpha e^{1+\alpha+\eta} < 20, \quad \frac{e^{-\eta}}{4} \left(\frac{1+\alpha}{2e\alpha} \right)^{1+\alpha} > \frac{1}{16} \quad (71)$$

throughout the bracket. The deletion-domain and $q = 0$ conditions hold there. Equation (67) proves uniqueness and multiplier optimality. Formula (70) also makes J twice continuously differentiable near the critical point. $\square$

Theorem 4.10 (Full dimension). *For every $N \geq 1$,*

$$V_{N,N+1}^{\mathbb{R}} \geq \frac{1}{16(N+1)^{1.5570595}}. \quad (72)$$

The exact exponent $p_{\dagger} = 1 + \alpha_{\dagger}$ satisfies the same bound with a prefactor greater than $1/16$.

Proof. Use the record increments and offset $a_{\text{off}} = 1$, taking every step above 2 as a candidate. If there are r such steps and $H = N + 1$, then $D_r \leq 2H - (r + 1)$. Proposition 4.8 at $\alpha = \alpha_{\dagger}$ and $J = 0$ has denominator at most

$$4\alpha e^{1+\alpha+\eta} \max_{0 < m \leq H} (2H - m)m^{\alpha}.$$

The maximum is at $m = 2\alpha H/(1 + \alpha)$ and gives the prefactor in (71). Lemma 4.2 uses at most $N + 1$ coordinates, and Lemma 4.3 gives the real-step conclusion. Rounding the exponent upward proves (72). $\square$

5 Dimension bounds for real schedules

Call (α, τ, C_0) admissible if, for every nonnegative schedule and every prescribed list of candidates exceeding 2, some subset I satisfies

$$e^{-\tau|I|} Q_2(I) \geq \frac{1}{C_0 D_r (r+1)^{\alpha}}, \quad 2 + T_I \leq D_r (r+1)^{\alpha}. \quad (73)$$

Here D_r and r are as in Proposition 4.8, with offset 2. Both inequalities must hold for the same subset.

Theorem 5.1 (Finite insertion bound). *Let (α, τ, C_0) be admissible, where $0 < \alpha < 1$, $\tau > 0$, and $C_0 \geq 20$. Put $p = 1 + \alpha$, $\beta = 1 - \alpha$, and let u_{τ} solve $\gamma(u_{\tau}) + \tau = u_{\tau}$. For integers $N \geq k \geq 2$, $M \geq 1$, set $r_k = k^2/(k^2 - 2)$. If $d \geq M + 1$ and*

$$\tau(d-1) \geq \left(p + 2 + \frac{\beta}{2M} \right) \log k + u_{\tau} r_k M + \log(5/2), \quad (74)$$

then

$$V_{N,d}^{\mathbb{R}} \geq \frac{k^{\beta}}{20C_0(2N+1)^2}. \quad (75)$$

Proof. Write $K = 2N + 1$. Suppose a schedule has worst case $w < k^\beta / (20C_0K^2)$, and put $s = (k^2/K^2)r(h)$. Lemma 2.2 gives $\sum_{j>k} s_j^\downarrow \leq 27k/64$. Choose the largest k steps of s above 2, or all such steps if fewer exist, as candidates. Their residual mass is at most $2(k+1) + 27k/64 \leq (283/64)k$, so $D_r(r+1)^\alpha$ is at most

$$D_* := 10k^p. \quad (76)$$

The same bound applies to every original temporal prefix of the list: both its residual mass and its candidate count can only decrease.

By Lemma 2.1 and $r_t(h) \leq \max(h_t, 0)$,

$$\prod_t (1 + s_t/k^2) \leq eK\sqrt{2w} < k^{\beta/2},$$

where $e/\sqrt{10C_0} < 1$. Set $H = k^{2+\beta/(2M)} \geq 4$. Concavity of $\log(1 + x/k^2)$ on $[0, H]$ gives $\min(x, H) \leq H \log(1 + x/k^2) / \log(1 + H/k^2)$. Thus, with $E = \{t : s_t > H\}$, $m = |E|$, and ordinary mass $R_o = \sum_{t \notin E} s_t$,

$$m \leq M, \quad R_o + mH \leq MH. \quad (77)$$

Every step in E is a prescribed candidate: its size exceeds 2 and the total mass $27k/64$ outside the largest k steps.

Choose I maximizing $e^{-\tau|I|}Q_2(I)$ over all subsets of the prescribed list. If $|I| \leq d-1$, (73) already gives a witness in at most d coordinates. Otherwise $|I| \geq d$. Let P be the largest actual prefix of I for which $|P \cup E| = d-1$. It exists because the cardinality starts at $m \leq d-1$, increases by zero or one, and ends at least d . Write $r = |P|$ and $j = |E \setminus P| = d-1-r$. There is a next selected checkpoint v .

Let $A(P)$ denote the product of the inverse kernel factors in P , and let x_P be its gap weight before v . Choose a subset I_c satisfying (73) on the same temporal prefix, with terminal weight $x_c = 2 + T_c \leq D_*$. Prefix optimality gives

$$A(P)^2 \geq \frac{4e^{\tau r} x_c}{C_0 D_*} \left[\frac{\kappa(x_P/(v-2))}{\kappa(x_c/(v-2))} \right]^2 \geq \frac{16e^{\tau r}}{C_0 D_*^2}.$$

The last step uses (43) and $x_P, x_c \geq 2$; it also covers the empty prefix.

Insert the j missing exceptional steps in chronological order. Their preceding unselected gaps have ordinary masses u_i . These gaps are disjoint, and the final terminal gap of mass T lies after them, so $\sum_i u_i + T \leq R_o$. Old inverse kernel factors increase under insertion: if a gap S is split into S_1, v, S_2 , its new preceding mass is $S_2 \leq S$, and newly selected steps also shorten later gaps.

By definition of u_τ , $2 \log \kappa(z) + \tau \leq u_\tau(1+z)$. Also $u_\tau > A_0 > 2/e$, where $\gamma(A_0) = A_0$: $\gamma(2/e) = \log 4 > 2/e$. Therefore

$$\begin{aligned} \tau j + 2 \sum_i \log \kappa \left(\frac{2 + u_i}{v_i - 2} \right) &\leq u_\tau \frac{jH + \sum_i u_i}{H-2} \leq u_\tau \frac{MH - T}{H-2}, \\ (2+T)e^{-u_\tau T/(H-2)} &\leq 2 + \frac{H-2}{eu_\tau} \leq H. \end{aligned}$$

Thus the final set $P' = P \cup E$, which has $d-1$ checkpoints, satisfies

$$Q_2(P') \geq \frac{4e^{\tau(d-1)}}{C_0 D_*^2 H} \exp \left(-u_\tau \frac{MH}{H-2} \right).$$

Condition (74), together with $H/(H-2) \leq r_k$, makes this at least $1/(C_0 D_*)$. Scaling by k^2/K^2 as in Lemma 4.3 contradicts the assumed value of w . This proves (75). $\square$

Theorem 5.2 (Joint horizon–dimension bound). *Let $p_{\dagger}, B_{\dagger}, a_{\dagger}$ be defined in (68)–(69). There are constants $C, C_d, N_0 > 0$ such that, with $L_N = \log N$,*

$$V_{N,d}^{\mathbb{R}} \geq \frac{N^{-p_{\dagger}}}{3600} \exp \left[-a_{\dagger} \frac{L_N^2}{d} - C \left(\frac{L_N^{3/2}}{d} + \frac{L_N^3}{d^2} \right) \right] \quad (N \geq N_0, d \geq C_d L_N). \quad (78)$$

In particular, if $d_N/L_N^2 \rightarrow c > 0$, then $\liminf N^{p_{\dagger}} V_{N,d_N}^{\mathbb{R}} \geq e^{-a_{\dagger}/c}/3600$.

Proof. Near the critical zero, take $\alpha = \alpha_{\dagger} + \delta$ and $\mu = A_{\dagger}/\alpha$. Smoothness and the strict constant checks supply fixed $K_0, \delta_0 > 0$ and admissible parameters with $C_0 = 20$ and $\tau = B_{\dagger}\delta - K_0\delta^2 \geq 0$ for $0 \leq \delta \leq \delta_0$. Set $k = N$ and choose $M = \lceil \sqrt{(1 - \alpha_{\dagger})L_N/(2A_0 r_N)} \rceil$. The inequality $u_{\tau} \leq A_0 + \tau$ follows from monotonicity of γ and $\gamma(A_0) = A_0$. Put

$$\begin{aligned} D &= d - 1 - r_N M, \\ G_0 &= \left(3 + \alpha_{\dagger} + \frac{1 - \alpha_{\dagger}}{2M} \right) L_N + A_0 r_N M + \log(5/2), \\ Z &= B_{\dagger} D - \left(1 - \frac{1}{2M} \right) L_N. \end{aligned}$$

The smaller root $\delta = 2G_0/(Z + \sqrt{Z^2 - 4K_0 D G_0})$ satisfies (74). For sufficiently large fixed C_d it is real, positive, and at most δ_0 whenever $d \geq C_d L_N$. Writing $X = L_N^2/d$, expansion of this root gives

$$\delta L_N = a_{\dagger} X + O\left(\frac{X}{\sqrt{L_N}} + \frac{X^2}{L_N} \right).$$

Indeed $G_0 = (3 + \alpha_{\dagger})L_N + O(\sqrt{L_N})$, $D = d - O(\sqrt{L_N})$, and $Z = B_{\dagger}d - O(L_N + \sqrt{L_N})$; the quadratic remainder contributes $O(L_N^3/d^2)$ after multiplying by L_N . Theorem 5.1 gives $N^{1-\alpha_{\dagger}-\delta}/[400(2N+1)^2]$. Using $2N+1 \leq 3N$ proves (78). $\square$

5.1 A variational curve for every logarithmic dimension ratio

For $\alpha \in [\alpha_{\dagger}, 1]$ define

$$\widehat{b}(\alpha) = - \min_{A \geq 5/6} \left\{ \frac{A}{\alpha} + \int_0^1 \phi(At^{-1-\alpha}) dt \right\}. \quad (79)$$

The minimum exists and is unique. Convexity follows from ϕ , with strictness supplied by its unchanged tail. Uniform coercivity follows from the lower bound $A - \log A - 3 - E(5/6)$, so all minimizing A lie in a fixed compact interval. Continuity follows by taking convergent subsequences of minimizers. Equation (67) gives $\widehat{b}(\alpha_{\dagger}) = 0$ and $\widehat{b}(\alpha) > 0$ for $\alpha > \alpha_{\dagger}$; it is strictly increasing, since at fixed μ the integral decreases with α and the admissible set $\mu \geq 5/(6\alpha)$ expands.

The floor $\phi(u) \geq -\log u - 1 - E(A)$ for $u \geq A$ gives $\widehat{b}(\alpha) \leq 1 + \alpha + \log \alpha + E(A)$ at its minimizer. Thus the empty-cutoff condition holds. Since $4\alpha e^{1+\alpha+E(A)} < 40$, Proposition 4.8 gives admissible parameters $(\alpha, \widehat{b}(\alpha), 40)$.

Theorem 5.3 (All logarithmic ratios). *For $c > 0$ put*

$$\mathcal{E}(c) = 2 - \max_{\alpha \in [\alpha_{\dagger}, 1]} (1 - \alpha) \min \left\{ 1, \frac{\widehat{b}(\alpha)}{3 + \alpha} \right\}. \quad (80)$$

If $d_N/\log N \rightarrow c > 0$, then $\limsup [-\log V_{N,d_N}^{\mathbb{R}}]/\log N \leq \mathcal{E}(c)$. Moreover

$$\mathcal{E}(c) = p_{\dagger} + \frac{a_{\dagger}}{c} + O(c^{-2}) \quad (c \rightarrow \infty). \quad (81)$$

Proof. The maximum exists by continuity. Fix an interior α and take $M = \lfloor (\log N)^{1/3} \rfloor$ in Theorem 5.1. Choose $\log k / \log N$ arbitrarily close from below to $\min\{1, \widehat{cb}(\alpha)/(3 + \alpha)\}$. Every other term in (74) is $o(\log N)$, which proves the corresponding gain. Maximizing over fixed parameters gives (80).

Near the critical point the minimizing multiplier is interior and depends smoothly on α ; hence $\widehat{b}(\alpha_{\dagger} + \delta) = B_{\dagger}\delta + O(\delta^2)$. Both $\widehat{b}(\alpha)/(3 + \alpha)$ and $(1 - \alpha)\widehat{b}(\alpha)/(3 + \alpha)$ increase on a fixed right neighborhood of $\alpha_{\dagger}$. In that neighborhood the unsaturated gain increases and the saturated gain decreases. Their junction $\widehat{cb}(\alpha) = 3 + \alpha$ is the unique maximum. For large c it is also the global maximum: outside that neighborhood the gain stays below $1 - \alpha_{\dagger}$ by a fixed amount. Expansion at the junction proves (81). $\square$

5.2 Weighted selection for sublogarithmic dimensions

Set

$$\begin{aligned} \beta_w &= \frac{5271}{5000}, \quad \tau_w = \frac{30069}{40000}, \quad \alpha_w = \frac{19}{25}, \quad \mu_w = \frac{287783}{200000}, \\ \phi_w(u) &= \begin{cases} \log 4 - 2 + \beta_w(2 - u), & u \leq 2, \\ \gamma(u), & u \geq 2. \end{cases} \end{aligned} \quad (82)$$

Proposition 5.4 (Weighted selection bound). *The triple $(\alpha_w, \tau_w, 20)$ is admissible.*

Proof. Put $R_w = 9/10$ and $e_0(r) = \beta_w(r - 1) - \log r$. It decreases on $[1/2, R_w]$, is positive at $1/2$ and negative at R_w . Use $e = (e_0)_+$ below R_w and zero above. The function ϕ_w is convex and decreasing, and majorizes $\gamma(2r) - 2e(r)$ for $r \geq 1/2$. Between the lower zero of e_0 and 1, this follows from convexity of e_0 and its endpoint values.

The left premaximum curvature for $r < R_w$ is at least $3/(4R_w^2)$, so Lemma 4.6 uses $v = \sqrt{Ce_0(r)}$, $C = 8(1 + R_w^2)/3 = 362/75$. Use the global growth line $19/20 + 7A/6$ and set

$$H = 19/20 + (7/6)(1 + r - v), \quad F = \tau_w/2 + \log(2/(rH)) - v - e_0(r).$$

At $r = 1/2$, the shifted argument is greater than $2/5$ and $H > 7/6$. Both inequalities persist on the active interval. The arithmetic certificate encloses F on all 2048 equal cells of $[1/2, 41/50]$ and proves every enclosure positive. Beyond $41/50$, $d(r) = \frac{1}{2}(1 - \beta_w r)^2 - e_0(r)$ is positive and increasing, so $v'' \leq 0$. Formula (59) now gives $F'' \geq 49/361$, since $H \leq 19/6$. At $r_0 = 1751/2000$, the certificate gives $F(r_0) - F'(r_0)^2/(2(49/361)) > 1/10000$. This proves the weighted one-step inequality everywhere.

Induction in (54) contributes $\tau_w q$ after squaring. Let $E_w(A) = \sup_{u \geq A} (\gamma(u) - \phi_w(u))_+$. For $A < 2$ it is $\max\{\gamma(A) - \phi_w(A), 0\}$. At $A = \mu_w \alpha_w$ we have $1 < A < 2$, so the directed domain and omitted-cell condition hold. Formula (70) with $U = 2$, $\beta = 2\beta_w$, and one certified moment proves $J_w(\alpha_w, \mu_w) < -\tau_w$. It also proves

$$4\alpha_w \exp(1 + \alpha_w + E_w(A) - \tau_w) < 20, \quad \tau_w < 1 + \alpha_w + E_w(A) + \log \alpha_w.$$

Proposition 4.8, with $b_* = \tau = \tau_w$, proves the claim. The factor $e^{-\tau_w}$ in the coefficient comes from $q = (q + 1) - 1$, and the same competitor supplies the terminal bound. $\square$

Corollary 5.5 (Growing sublogarithmic dimension). *If $d_N \rightarrow \infty$ and $d_N = o(\log N)$, then*

$$\log(N^2 V_{N, d_N}^{\mathbb{R}}) \geq \frac{90207}{1880000} d_N - O(\sqrt{d_N}). \quad (83)$$

For $d_N / \log N \rightarrow c > 0$, the bound in Theorem 5.3 may additionally be combined with the gain $(1 - \alpha_w) \min\{1, c\tau_w/(3 + \alpha_w)\}$ by taking the larger gain.

Proof. Use Proposition 5.4, take M of order $\sqrt{d_N}$, and choose $\log k = \tau_w d_N / (3 + \alpha_w) - C\sqrt{d_N}$ with sufficiently large fixed C . Then $k \rightarrow \infty$, $k \leq N$, and (74) holds. Its insertion terms are $O(\sqrt{d_N})$. The gain is exactly $(1 - \alpha_w)\tau_w / (3 + \alpha_w) = 90207/1880000$. The fixed-ratio statement follows by the same cutoff argument as Theorem 5.3. $\square$

6 Deleting negative steps in one dimension

Lemma 6.1 (A terminal-point inclusion). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have continuous derivative, a minimizer x_* , and $(x - x_*)f'(x) \geq 0$ for all x . Replace one negative step of any finite schedule by zero. Every terminal point attainable by the modified schedule from $[x_* - R, x_* + R]$ is attainable by the original schedule from the same interval, using the same function.*

Proof. Write $T_b(x) = x - bf'(x)$. If the negative step is $-a$, $a > 0$, write the original terminal map as $S \circ T_{-a} \circ P$, where P and S are the compositions of the steps before and after $-a$. Fix an initial point x_0 , and set $z = P(x_0)$ and $Q = T_{-a} \circ P$. The map Q is continuous and fixes x_* . Moreover, $Q(x_0) = z + af'(z)$ lies at or beyond z on the ray from x_* through z . Hence z lies between $Q(x_*)$ and $Q(x_0)$. The intermediate value theorem gives a point u on the segment from x_* to x_0 with $Q(u) = z$. This point satisfies $|u - x_*| \leq |x_0 - x_*|$, and the original schedule ends at $S(Q(u)) = S(P(x_0))$. $\square$

Theorem 6.2 (Scalar sign equivalence). *For every real schedule h , $W_{N,1}(h) \geq W_{N,1}(h_+)$. Consequently $V_{N,1}^{\mathbb{R}} = V_{N,1}^+$ at every horizon.*

Proof. Apply Lemma 6.1 successively to the negative entries and take the supremum over objectives and initial points. Then $W_{N,1}(h) \geq W_{N,1}(h_+) \geq V_{N,1}^+$. Taking infima gives one inequality between the minimax values. The inclusion $[0, \infty)^N \subset \mathbb{R}^N$ gives the reverse inequality. $\square$

The proof also applies coordinatewise to objectives separable in an orthonormal coordinate system. The replacement initial point lies no farther from the chosen minimizer in each coordinate, and hence in Euclidean distance. This includes all convex quadratics.

7 Sign equivalence through five steps

We combine scalar sign deletion with lower bounds for nonnegative schedules of at most four steps and positive upper bounds at the next horizon.

7.1 Three elementary lower bounds

For a nonnegative schedule with step sum S , take the Huber function

$$f_\delta(x) = \begin{cases} x^2/2, & |x| \leq \delta, \\ \delta|x| - \delta^2/2, & |x| \geq \delta, \end{cases} \quad \delta = 1/(1 + 2S).$$

Starting at 1, every iterate lies on the positive affine branch: $x_i = 1 - \delta \sum_{j=1}^i h_j \geq \delta$. Hence

$$W_{N,1}(h) \geq \frac{1}{2(1 + 2S)}. \tag{84}$$

The quadratic of curvature $\lambda \in [0, 1]$ gives

$$W_{N,1}(h) \geq \frac{\lambda}{2} \prod_i (1 - \lambda h_i)^2. \quad (85)$$

The following bound, which depends on the order of the steps, is due to Tsai–Fatkhullin–Zhang–He [6, Lemma 3.1]. Their lemma is stated for positive steps; the scalar construction below also permits zero steps. The bound is trivial when the selected step is at most one. If b is one selected step, P is the sum of the steps preceding it, and T is the sum of the steps following it, then

$$W_{N,1}(h) \geq \frac{(b-1)_+^2}{2(1+P)^2(1+2T)}. \quad (86)$$

For $b > 1$, take $f'(x) = \min\{\delta, \max\{-\epsilon, x\}\}$, the asymmetric Huber derivative with constant values $-\epsilon$ and δ , where $\delta = (1+P)^{-1}$ and $\epsilon = \delta(b-1)/(1+2T)$. From 1, the prefix reaches δ in the positive affine region. The selected step reaches $-\delta(b-1)$; all subsequent queries stay in the negative affine region. The endpoint is $-\delta(b-1) + \epsilon T \leq -\epsilon$, giving (86).

Proposition 7.1 (Lower bounds for short nonnegative schedules). *Every nonnegative scalar schedule of two, three, or four steps has worst case strictly greater than, respectively,*

$$\frac{9}{200}, \quad \frac{1}{28}, \quad \frac{1}{41}.$$

Zero steps are allowed.

Proof for two steps. Suppose $w = W_{2,1}(a, b) \leq 9/200$ for $a, b \geq 0$, and put $S = a + b$ and $T = ab$. Equation (84) gives $S \geq 91/18$, and (85) at $\lambda = 1$ gives $T \leq S - 7/10$. At $\lambda = 1/2$, therefore, $1 - S/2 + T/4 \leq 33/40 - S/4 \leq -79/180$. This forces $w \geq 6241/129600 > 9/200$, a contradiction. The three- and four-step proofs are in Appendices A and B. $\square$

The scalar lower bounds and the positive upper bounds of Grimmer–Shu–Wang [4, Sections 3.2.2–3.2.3] give the following comparisons in every dimension. Section 8 specifies the schedules and derives the upper bounds.

N	Lower bound with a negative step	Positive upper bound
3	$> 9/200$	$(3 - 2\sqrt{2})/4 < 9/200$
4	$> 1/28$	$< 1/32$
5	$> 1/41$	$< 5/206 < 1/41$

Theorem 7.2 (Short-horizon sign equivalence). *For $d \geq 1$ and $N \leq 5$, $V_{N,d}^{\mathbb{R}} = V_{N,d}^+$. For $N = 3, 4, 5$, schedules containing a negative step are separated strictly from a feasible positive upper bound.*

Proof. A schedule of length N containing a negative entry has at most $N - 1$ positive entries. Scalar embedding and Theorem 6.2 bound its worst case below by that of the scalar schedule h_+ . Delete zeros and pad to $N - 1$ entries with zeros. For $N = 3, 4, 5$, Proposition 7.1 and the positive comparisons prove the claim. For $N \leq 2$, every negative entry is at a boundary. A final negative step cannot lower the objective, by convexity. For a leading negative step $-a$, initialize instead at $\text{prox}_{af}(x_0)$; its distance to a minimizer is no greater and the negative step then lands exactly at x_0 . This reproduces the zero-step trajectory. Successive deletion proves equality for these horizons too. $\square$

A strict advantage from negative steps would require $d \geq 2$ and $N \geq 6$. It would also require at least five strictly positive entries: with at most four, retaining the positive entries and padding with zeros gives a four-step scalar schedule with worst case greater than $1/41$, whereas the five-step schedule in Section 8, padded with zeros, has worst case less than $1/41$.

8 The constructive upper comparison

The composition theorems of Grimmer–Shu–Wang [4, Theorems 1 and 3] give the following recursion. Each schedule has a type, S or F, and a positive parameter, equal to the reciprocal rate in their terminology. The empty schedule has both types and parameter one. An S-join takes two type S schedules with parameters x, y and gives a type S schedule. An F-join takes a type S schedule of parameter x followed by a type F schedule of parameter y and gives a type F schedule. The output parameters are, respectively,

$$J_S(x, y) = \frac{x + y + \sqrt{x^2 + 6xy + y^2}}{2}, \quad J_F(x, y) = \frac{4x + y + \sqrt{y^2 + 8xy}}{2}.$$

For input schedules a, b , each join forms $[a, \mu, b]$, where

$$\mu_S = 1 + \frac{\sqrt{x^2 + 6xy + y^2} - x - y}{2}, \quad \mu_F = 1 + \frac{\sqrt{y^2 + 8xy} - y}{4}.$$

The functions J_S, J_F are increasing and homogeneous of degree one. Every type F schedule constructed by this recursion has terminal gap at most $1/(2z)$ in every dimension, where z is its parameter. A binary tree records a sequence of joins: each leaf is an empty schedule, and each internal vertex inserts one positive step. A tree with m leaves therefore gives a schedule with $m - 1$ steps.

8.1 Schedules for three, four, and five steps

The following positive schedules and upper bounds in every dimension are given by Grimmer–Shu–Wang [4, Sections 3.2.2–3.2.3]. Put $\rho = 1 + \sqrt{2}$ and define

$$x = (\rho + 1 + \sqrt{\rho^2 + 6\rho + 1})/2.$$

N	Schedule	Upper bound
3	$(\sqrt{2}, \rho, 3/2)$	$(3 - 2\sqrt{2})/4 < 9/200$
4	$(\sqrt{2}, (\sqrt{10 + 8\sqrt{2}} - \sqrt{2})/2, \sqrt{1 + 2x}, 3/2)$	$< 1/32$
5	$(\sqrt{2}, 2, \sqrt{2}, \sqrt{7 + 4\sqrt{2}}, 3/2)$	$< 5/206 < 1/41$

The type S schedules $(\sqrt{2})$, $(\sqrt{2}, x - \rho)$, and $(\sqrt{2}, 2, \sqrt{2})$ have parameters ρ , $x = J_S(\rho, 1)$, and $\rho^2 = J_S(\rho, \rho)$, respectively; here $x - \rho = (\sqrt{10 + 8\sqrt{2}} - \sqrt{2})/2$. Joining a type S schedule of parameter u to the type F schedule $(3/2)$, of parameter 4, inserts the step $\sqrt{1 + 2u}$. The three resulting type F schedules therefore have parameters $J_F(\rho, 4)$, $J_F(x, 4)$, and $J_F(\rho^2, 4)$. For the four-step bound, $x > 4$ implies $J_F(x, 4) > 16$. For the five-step bound, $\rho^2 > 29/5$ and $\sqrt{1 + 2\rho^2} > 7/2$ imply $J_F(\rho^2, 4) > 103/5$. The scalar supplement verifies these radical comparisons by squaring positive quantities and comparing rational numbers.

8.2 A finite certificate for all horizons

The certificate below refines the numerical upper comparison for the basic composition family of Grimmer–Shu–Wang [4, Section 4.2 and Theorem 5]. Their displayed coefficient is $0.42312/2 = 0.21156$ in our normalization, valid for $N \geq 2^{18} - 1$. We certify 0.211556239176 for $N \geq 2097151$.

Put $Q_0 = 2^{80}$. The supplement records one selected S-tree and one selected F-tree for each leaf count $1 \leq m \leq 4194303$. A split specifies two smaller children. Starting from numerator Q_0 at each empty tree, replace each square root in the join formula by its integer square root and round the parent numerator down. Monotonicity proves inductively that the exact tree parameter is at least the resulting integer A_m/Q_0 . Equivalently, every parent comparison has the form

$$2c - B \leq 0 \quad \text{or} \quad (2c - B)^2 \leq R,$$

with $(B, R) = (a + b, a^2 + 6ab + b^2)$ for S and $(B, R) = (4a + b, b^2 + 8ab)$ for F. The joining steps themselves are defined by the exact child parameters.

The finite seed certificate proves

$$\frac{Q_0(m+1)^s}{2A_m} \leq \frac{211556239176}{10^{12}}, \quad 2097152 \leq m \leq 4194303, \quad (87)$$

where A_m denotes the F numerator. Here is a fully rational enclosure procedure for the powers in this comparison. Use precision 2^{128} , bound logarithms by

$$\log x = 2 \sum_{j=0}^{79} \frac{t^{2j+1}}{2j+1} + R, \quad t = \frac{x-1}{x+1}, \quad 0 \leq R \leq \frac{2t^{161}}{161(1-t^2)},$$

and bound the exponential by its first 81 Taylor terms and the geometric tail beginning at degree 81. Every operation is rounded outward. This encloses $s = \log(1 + \sqrt{2})/\log 2$ and the powers at 16385 equally spaced nodes of $[1, 2]$. Since $1 < s < 2$, the graph of x^s lies below each chord between upper node values. Outward multiplication by $(1 + \sqrt{2})^{21}$ therefore bounds the powers throughout $[2^{21}, 2^{22}]$. For each integer m in (87), substitute $m+1$ in this chord and clear positive denominators. The certificate comprises 8388604 selected joins and 2097152 seed inequalities. The programs and split data are in `supplement/joint-horizon-dimension`.

To extend the seed bound, we use the doubling construction of Grimmer–Shu–Wang [4, Lemma 10] in terms of composition trees. Replace each empty S leaf by $\lfloor \sqrt{2} \rfloor$ and each empty F leaf by $\lfloor 3/2 \rfloor$. Their parameters are $1 + \sqrt{2}$ and 4, respectively. The joins are increasing and homogeneous, so this replacement multiplies the root parameter by at least $\rho = 1 + \sqrt{2}$ while doubling the number of leaves. For $n = N + 1 \geq 2^{21}$, choose $j \geq 0$ such that $m = \lfloor n/2^j \rfloor \in [2^{21}, 2^{22} - 1]$. After j replacements, append unit steps to reach horizon N . The descent inequality ensures that padding cannot increase the gap. Because $n < 2^j(m+1)$ and $\rho = 2^s$, the resulting gap is at most

$$\frac{Q_0}{2A_m \rho^j} \leq \frac{211556239176}{10^{12}} (N+1)^{-s}.$$

A smaller certificate supplies the all- N constant in (4). It verifies a bound of $3/10$ for $2 \leq N \leq 62$ and $Q_0(m+1)^s/(2A_m) \leq 219/1000$ for $64 \leq m \leq 127$. The same leaf replacement covers all $N \geq 63$. These comparisons use $s < 159/125$, proved by expanding $(1 + \sqrt{2})^{125}$ and comparing it with 2^{159} by integer squares. Both rational constants are smaller than $(1 + \sqrt{2})/8$. For $N = 1$, the step $3/2$ has gap $1/8$, giving equality in the claimed all-horizon upper bound.

9 Open questions

The optimal rates in fixed dimension remain unknown. The nonnegative bound leaves a logarithmic gap at large dimensions, and the real-step lower exponent remains above the constructive exponent s . Sign equivalence is open for nonseparable objectives when $d \geq 2$ and $N \geq 6$. The exact quadratic–Huber constant determines a lower bound; nonlinear scalar objectives may require larger errors.

A The three-step lower bound

Suppose $a, b, c \geq 0$ and $w = W_{3,1}(a, b, c) \leq 1/28$, put $S = a + b + c$, and $r = 1/\sqrt{14}$. The witnesses give

$$S \geq 13/2, \quad |(a-1)(b-1)(c-1)| \leq r, \quad (a-1)_+ \leq r\sqrt{1+2b+2c},$$

$$b \leq 1 + r(a+1)\sqrt{1+2c}, \quad c \leq 1 + \frac{r}{1+r}S.$$

The first step inequality implies $14a^2 - 26a + 13 \leq 2S$; it also holds for $a \leq 1$, since its left side is at most 13. Use $r < 27/100$, $r/(1+r) < 3/14$ and $2uv \leq u^2 + v^2$ to obtain

$$b \leq 1 + \frac{27}{200}((a+1)^2 + 1 + 2c) \leq 1 + \frac{27}{200} \left(\frac{27a}{7} + \frac{15}{14} + \frac{S}{7} + 2c \right).$$

Adding $a + c$ and substituting $c \leq 1 + 3S/14$ gives $1984S \leq 6761 + 4258a$. Combining this upper bound on S with $14a^2 - 26a + 13 \leq 2S$ yields $13888a^2 - 30050a + 6135 \leq 0$. This polynomial is positive and increasing for $a \geq 39/20$, so

$$36/25 < 6135/4258 \leq a < 39/20, \quad S < 38/5.$$

The lower bound on a gives $a - 1 > 11/25$; the bound $c \leq 1 + 3S/14$ gives $c < 92/35$. Finally, $b = S - a - c$ and $S \geq 13/2$ give $b - 1 > 129/140$. If $c > 1$, the quadratic product bound now gives $c - 1 < (27/100)/[(11/25)(129/140)] = 315/473 < 2/3$; if $c \leq 1$, the conclusion $c < 5/3$ already holds. Returning to the middle step inequality and using $\sqrt{13/3} < 25/12$ gives $b < 851/320$. Thus $13/2 \leq S < 39/20 + 851/320 + 5/3 = 1205/192 < 13/2$, the required contradiction.

B The four-step covering certificate

Suppose $h_i \geq 0$ and $w = W_{4,1}(h) \leq 1/41$. Put $S = \sum h_i$, and let P_i and T_i be the sums of the steps before and after h_i . Equations (84) and (86) give

$$S \geq 39/4, \quad (h_i - 1)_+^2 \leq \frac{2}{41}(1 + P_i)^2(1 + 2T_i). \quad (88)$$

The quadratics of curvatures 1 and $4/5$ also give

$$\prod_i |h_i - 1| \leq 221/1000, \quad \prod_i |h_i - 5/4| \leq 603/1000. \quad (89)$$

These are valid outward relaxations, since $(221/1000)^2 > 2/41$ and $(603/1000)^2 > 5^9/(82 \cdot 4^8)$.

We first bound S . For the odd polynomial $p(t) = t \prod_i (1 - h_i t^2)$, use the positive nodes $z = (1, 15/16, 3/4, 1/2, 1/6)$. Define

$$A_j = \frac{1}{z_j} \prod_{k \neq j} \frac{-z_k^2}{z_j^2 - z_k^2}, \quad \sigma_j = \sum_{k \neq j} z_k^{-2}, \quad t = \min_j \sigma_j.$$

Comparing the linear and cubic coefficients in odd Lagrange interpolation gives $1 = \sum A_j p(z_j)$ and $S = \sum \sigma_j A_j p(z_j)$. Equation (85) with $\lambda = s^2$, $s \in [-1, 1]$, gives $\|p\|_{[-1,1]} \leq \sqrt{2/41} < 221/1000$. Subtracting t times the first interpolation identity from the second, we obtain

$$S \leq t + \frac{221}{1000} \sum_j (\sigma_j - t) |A_j| = \frac{1781}{225} + \frac{221}{1000} \frac{26660981}{543375} = \frac{10193191801}{543375000} < 19.$$

Combining this with (88) gives the initial box

$$0 \leq h_1 \leq 5/2, \quad 0 \leq h_2 \leq 6, \quad 0 \leq h_3 \leq 8, \quad 0 \leq h_4 \leq 9/2.$$

Use $T_1 < 19$ for the first bound, $P_4 < 19 - h_4$ for the last, then $P_2 < 5/2, T_2 < 19$ for the second, and $P_3 < 5/2 + 6, T_3 < 9/2$ for the third. In each case, substituting these bounds into (88) and squaring positive quantities gives the stated rational bound.

The certificate `supplement/scalar-schedules/certificates/four-step-cover.json` excludes this box by three interval rules. The sum bounds $39/4 \leq S < 19$ bound each coordinate from below using the other coordinates' upper endpoints, and from above using their lower endpoints. Upper bounds $\bar{P}_i, \bar{T}_i$ in (88) give $h_i \leq 1 + (221/1000)(1 + \bar{P}_i)\sqrt{1 + 2\bar{T}_i}$. For either product in (89), let m_i be the minimum distance from coordinate interval i to the corresponding root, 1 or $5/4$. A product $\prod m_i$ above the threshold rejects the box; otherwise, if $\prod_{j \neq i} m_j > 0$, the distance of h_i from that root is at most the threshold divided by that product.

Endpoints are integer multiples of 10^{-6} , rounded outward, so the resulting intervals contain every real schedule in the original box. Integer square roots and exact squared comparisons enclose each root. An unresolved box is bisected into two closed children whose union covers it. The retained complete binary tree has ten splits and eleven rejected leaves: 21 boxes, maximum depth five. Each rejection follows from the necessary inequalities above. The verifier reconstructs the entire tree, checks completeness, and compares it with the supplied certificate. This proves the four-step assertion in Proposition 7.1.

References

- [1] M. Jung, H. Cho, and C. Yun. Stronger lower bounds for (non-)anytime acceleration of gradient descent. arXiv:2609.04032v2 (2026).
- [2] J. Ma and Y. Chen. A lower bound for stepsize-based acceleration of gradient descent. arXiv:2608.10418v1 (2026).
- [3] Y. Ye and K. Liu. Improved gradient descent lower bounds beyond Nesterov. arXiv:2609.02855v2 (2026).
- [4] B. Grimmer, K. Shu, and A. L. Wang. Composing optimized stepsize schedules for gradient descent. arXiv:2410.16249v2 (2025).

- [5] Z. Zhang and R. Jiang. Accelerated gradient descent by concatenation of stepsize schedules. arXiv:2410.12395v2 (2026).
- [6] C.-E. Tsai, I. Fatkhullin, L. Zhang, and N. He. Lower bounds for anytime acceleration of gradient descent. arXiv:2607.02053v1 (2026).
- [7] L. Bell. An $(N + 1)$ -dimensional bound for the terminal objective of gradient descent. Companion note (2026).
- [8] A. B. Taylor, J. M. Hendrickx, and F. Glineur. Smooth strongly convex interpolation and exact worst-case performance of first-order methods. *Mathematical Programming* 161 (2017), 307–345. arXiv:1502.05666.
- [9] J. M. Altschuler and P. A. Parrilo. Acceleration by stepsize hedging: silver stepsize schedule for smooth convex optimization. *Mathematical Programming* 213 (2025), 1105–1118. doi:10.1007/s10107-024-02164-2.
- [10] L. Bell. Three-step dimension gaps for gradient descent. Companion paper (2026).
- [11] Y. Ye and K. Liu. Silver rate is (almost) optimal for gradient descent acceleration. arXiv:2609.09152v1 (2026).
- [12] L. Bell. An elliptic constant in lower bounds for gradient descent. Companion paper (2026).