COMPARATOR-UNIFORM DYNAMIC REGRET UNDER HEAVY-TAILED NOISE

Anonymous authors
Paper under double-blind review

ABSTRACT

Existing high-probability guarantees for online convex optimization with heavy-tailed gradients fix the comparator before the interaction, pay an adaptation logarithm in the horizon, and require the horizon as an input. We address these three issues on a bounded convex set with adaptively chosen, non-clairvoyant convex Lipschitz losses. The learner receives one stochastic gradient per round, and the noise is assumed only to have a bounded conditional p-th central moment for $p \in (1, 2]$, so its variance may be infinite. Our algorithm aggregates clipped online-gradient-descent experts on a grid of learning rates anchored at the static statistical scale. Each expert uses a threshold tied to its step size. A per-expert bias offset preserves the supermartingale property of the aggregate wealth, while an auxiliary descent sequence used only in the analysis handles the noise residuals. This sequence turns the comparator-dependent noise term into a martingale along a predictable direction without requiring an additional sample. Consequently, one event of probability $1-\delta$ covers every prefix and every comparator path, including paths chosen from the complete feedback record. The resulting regret combines a square-root term in the horizon and path-confidence complexity with a heavy-tail term scaling as $T^{1/p}$ and the same complexity to exponent $(p-1)/p$. The bound contains no additional horizon, dimension, or path-adaptation logarithm. A doubling wrapper gives an anytime version with an iterated-logarithm cost, while the noiseless case recovers the deterministic path-length rate. Self-contained lower bounds match the heavy-tail exponents in path length and confidence and identify the remaining gap in the deterministic term.

1 INTRODUCTION

Modern learning systems often operate while the data distribution changes. Recommendation and pricing engines track shifting demand, controllers respond to changing loads, and forecasters update on drifting streams. Online convex optimization (OCO) is the standard abstraction for these systems (Zinkevich, 2003; Cesa-Bianchi & Lugosi, 2006; Hazan, 2016; Orabona, 2019), and *dynamic regret*, the excess loss relative to an arbitrary comparator sequence $u_1, \ldots, u_T$ rather than a single fixed point, is the standard yardstick for a non-stationary environment. Two properties of deployed systems create the main difficulty. Their stochastic gradients can be heavy-tailed: estimated tail indices of gradient noise in trained networks fall below two, so the variance need not be finite (Şimşekli et al., 2019), which is precisely why clipping is a default ingredient of large-scale training (Zhang et al., 2020; Gorbunov et al., 2020). A deployed system is also run once, so an expected-regret guarantee does not describe its realized trajectory. Under infinite variance, the gap between the mean and the typical observations is substantial. We therefore seek a high-probability bound for the realized run, against a comparator sequence chosen in hindsight, without requiring the run length in advance.

Prior work addresses these requirements separately. Zinkevich (2003) introduced OCO together with projected online gradient descent and the first path-length dynamic bound, of order $A_T\sqrt{T}$ for $A_T = 1 + P_T/D$ with P_T the comparator's total variation, at the price of tuning the step size to P_T. Subsequent work removed several forms of that prior knowledge and refined the models of environmental movement. This includes tracking and fixed-share schemes (Herbster & Warmuth, 1998), dynamic mirror descent (Hall & Willett, 2013), comparators following a known dynamics

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Table 1: Representative dynamic-regret and heavy-tailed results. $A_T = 1 + P_T/D$ is path complexity, $h = \log(4/\delta)$, and $q = (p-1)/p$; rates suppress constants and G, D. *Uniform comp.* means one event covers all comparator sequences, including hindsight choices. *Horizon free* means that no design horizon is supplied; N/A marks an unavailable comparison.

Ref.	Noise	Guarantee	Uniform	Horizon	Dynamic regret
Zinkevich (2003)	bounded	determ.	×	×	$A_T\sqrt{T}$
Zhang et al. (2018)	bounded	determ.	×	×	$\sqrt{TA_T}$
Zhao et al. (2024)	bounded	determ.	×	×	$\sqrt{TA_T}$
Zhang & Cutkosky (2022)	p-th mom.	high prob.	✓	×	static only
Liu (2026a)	p-th mom.	expect.	×	✓	classical algorithms
Liu (2026b)	p-th mom.	high prob.	N/A	×	offline, nonsmooth
Yang et al. (2025)	p-th mom.	high prob.	×	×	distributed network
Aggarwal (2026)	p-th mom.	expect.	×	×	$\sqrt{TA_T} + \sigma T^{1/p} A_T^q$
This paper	p-th mom.	high prob.	✓	✓	$\sqrt{T(A_T+h)} + \sigma T^{1/p}(A_T+h)^q$

(Jadbabaie et al., 2015), and drift-budgeted stochastic models (Besbes et al., 2015). The guarantees remained deterministic or in expectation. Zhang et al. (2018) removed the need to tune the step size to P_T: a grid of step sizes run as experts and combined by a meta learner attains the minimax $\sqrt{TA_T}$ rate with one gradient query per round and no knowledge of P_T, a two-layer template underlying essentially all later refinements (Zhao et al., 2020; 2024). These analyses assume bounded gradients or bounded variance. In offline stochastic optimization, clipping and martingale control yield high-probability rates under only a p-th moment, $p \in (1, 2]$ (Gorbunov et al., 2020; Cutkosky & Mehta, 2021; Vural et al., 2022; Sadiev et al., 2023). Clipping bounds the observations at the cost of a bias term. Online learning followed suit: Zhang & Cutkosky (2022) obtained parameter-free high-probability *static* regret under heavy tails, Liu (2026a) revisited classical OCO algorithms under heavy tails largely in expectation, Liu (2026b) refined the high-probability analysis of clipped methods for nonsmooth convex problems, Yang et al. (2025) proved high-probability dynamic bounds in a distributed setting against per-round minimizers, and Aggarwal (2026) obtained parameter-free dynamic regret in expectation with the correct exponents in noise and path variation. Three limitations remain. (i) *The comparator is frozen before the interaction.* Existing high-probability dynamic bound events are indexed by a fixed comparator sequence, whereas hindsight selection requires one event for all paths; the uncountable set $\mathcal{X}^T$ rules out a direct union bound. The obstruction is concrete: treating the noise-induced term $\langle e_t, x_t - u_t\rangle$ as a martingale difference is invalid once u_t depends on the data, because $x_t - u_t$ is then not predictable. (ii) *The unknown path is paid for twice*: the aggregation charges an adaptation logarithm inside the confidence level at the horizon. (iii) *The horizon is an input*, because the step-size grid is anchored to it. Heavy tails additionally break the mechanism that makes expert aggregation work at all: the tilted-weight identity that cancels the true gradient across experts presumes that all experts see the same feedback, and experts clipping at different thresholds do not.

We propose **BC-MW-OGD**, which aggregates clipped online gradient descent experts with step-size-dependent thresholds and bias compensation. The wealth offset preserves the supermartingale property across heterogeneous clipping levels. An auxiliary descent sequence, used only in the analysis, converts the comparator-dependent noise into a predictable martingale plus a deterministic path potential. This yields one high-probability event for every prefix and comparator path, including paths selected from the full feedback record. A geometric grid anchored at the static statistical scale removes the adaptation logarithm, and a deterministic doubling wrapper removes the horizon input at an iterated-logarithm cost. Table 1 situates the result. Our contributions are as follows.

- We propose BC-MW-OGD and prove a dynamic regret bound under heavy-tailed gradient noise whose high-probability event is independent of the comparator. On one event of probability $1 - \delta$, every prefix and every comparator sequence, including sequences selected in hindsight from the whole feedback record, is covered by the stated bound. Its rate combines a square-root path-confidence term with a heavy-tail term of order $T^{1/p}$, with no hidden horizon, dimension, or path-adaptation logarithm.

- We remove the horizon and the confidence level from the inputs: a deterministic doubling wrapper keeps the same rate for every $n \geq 1$ at an iterated-logarithm price, and the noiseless case recovers the deterministic path-length rate without a confidence penalty.

- We prove self-contained lower bounds, valid for any gradient-only algorithm, showing that the noise term must combine path budget and confidence under the same heavy-tail exponent as the upper bound. The comparison leaves only the deterministic path-confidence term as a gap.

2 Related work

2.1 Dynamic regret and comparator-adaptive aggregation

Dynamic regret measures an online learner against a moving comparator, with difficulty indexed by the comparator's total variation. Zinkevich (2003) gave the first path-length bound; tracking and fixed-share schemes (Herbster & Warmuth, 1998), dynamic mirror descent (Hall & Willett, 2013), comparators following a known dynamics (Jadbabaie et al., 2015), and drift-budgeted stochastic models (Besbes et al., 2015) progressively relaxed what the learner must know about how the environment moves. A geometric grid of step sizes, run as experts and combined by a meta learner, attains the minimax $\sqrt{T(1+P_T)}$ rate without knowing P_T (Zhang et al., 2018). Because the losses are linearized at the meta point, one stochastic gradient per round suffices for the whole grid. The meta layer is usually a multiplicative-weight or Prod-style update with learning-rate tilting (Gaillard et al., 2014), which is also the mechanism behind comparator-adaptive and parameter-free online learning (Orabona & Pál, 2016; Cutkosky & Orabona, 2018), and later refinements make the rate problem-dependent rather than worst case (Zhao et al., 2020; 2024). These methods either rely on a hard gradient bound or control their second-order terms in expectation. Those arguments do not apply to an infinite-variance oracle. A single sample per round creates a second issue: different step sizes require different clipping scales, so the experts no longer see identical feedback. We retain the one-query grid architecture and change its analysis. The wealth process remains a supermartingale under heterogeneous clipping, and the grid anchor and prior charge aggregation to the path-confidence complexity rather than to the number of experts.

2.2 Heavy-tailed noise, clipping, and high-probability rates

A parallel line weakens the noise model from sub-Gaussian or bounded-variance to a bounded p-th moment, $p \in (1, 2]$, beginning with robust stochastic approximation (Nemirovski et al., 2009) and crystallizing around clipping: the clipped gradient has bounded increments, at the price of a bias of order $\sigma^p \tau^{1-p}$ at threshold τ (Zhang et al., 2020; Gorbunov et al., 2020). Combining clipping with Freedman-type arguments yields high-probability rates under unbounded variance for stochastic convex and non-convex problems (Cutkosky & Mehta, 2021; Vural et al., 2022; Sadiev et al., 2023), with a recent refinement for nonsmooth convex objectives (Liu, 2026b). Online counterparts followed: parameter-free high-probability static regret under heavy tails (Zhang & Cutkosky, 2022), classical OCO algorithms revisited under heavy tails largely in expectation (Liu, 2026a), high-probability dynamic bounds in a distributed network against per-round minimizers (Yang et al., 2025), parameter-free dynamic regret in expectation with the correct noise–path exponents (Aggarwal, 2026), and noise-adaptive high-probability bounds under sub-Gaussian noise and strong convexity (Zhang et al., 2026). Most existing analyses tune one threshold to the horizon and confidence level, so the clipping bias is a single scalar. Under aggregation, thresholds vary across experts and the bias becomes expert-dependent. It then breaks the drift identity used by the meta layer. We exhibit a legitimate p-th moment distribution for which the uncompensated aggregate wealth has strictly positive conditional drift. Heavy-tailed expert problems can also contain lower-order terms involving the largest observed loss (Moulin et al., 2025); our meta layer therefore sees only the clipped surrogates. The bias is subtracted inside the wealth recursion, and the clipping lemma keeps the deterministic G^2 contribution separate from the statistical $\sigma^p \tau^{2-p}$ contribution. This yields separate terms at scales $G\sqrt{T}$ and $\sigma T^{1/p}$.

2.3 TIME- AND COMPARATOR-UNIFORM GUARANTEES

Uniform guarantees depend on quantifier order. Time-uniform bounds use nonnegative supermartingales and maximal inequalities, while static comparator-uniform bounds use wealth processes that introduce the comparator only through deterministic duality (Orabona & Pál, 2016; Cutkosky & Orabona, 2018; Zhang & Cutkosky, 2022). Dynamic regret is harder because a data-dependent path appears inside $\sum_t \langle e_t, x_t - u_t \rangle$, making $x_t - u_t$ non-predictable. Discretizing $\mathcal{X}^T$ incurs a dimension-dependent covering factor, whereas restricting comparators to predictable paths excludes hindsight selection. Our auxiliary descent sequence replaces this stochastic dependence with a predictable martingale direction and a deterministic path potential, yielding one event for all comparator paths. A deterministic epoch schedule with summable confidence allocation extends the event over time without using the horizon or δ as input. We also prove a joint path-confidence noise lower bound for gradient-only feedback; the remaining discrepancy is $\sqrt{A_T + h}$ versus $\sqrt{A_T}$ in the deterministic term.

3 SETTING: DYNAMIC REGRET WITH HEAVY-TAILED GRADIENTS

3.1 NOTATION

We work in $\mathbb{R}^d$ with the Euclidean inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. The decision set is $\mathcal{X} \subseteq \mathbb{R}^d$, $D > 0$ is a known upper bound on its diameter, $\Pi_{\mathcal{X}}$ is the Euclidean projection onto $\mathcal{X}$, and $C_\tau(z) = z \min\{1, \tau/\|z\|\}$ is radial clipping at radius $\tau > 0$ (with $C_\tau(0) = 0$). Rounds are indexed by t, prefixes by n, and T denotes a *design* horizon supplied to the finite-horizon algorithm. The filtration $(\mathcal{F}_t)_{t \geq 0}$ carries the interaction, every decision is $\mathcal{F}_{t-1}$-measurable, and $\mathbb{E}_{t-1}[\,\cdot\,] = \mathbb{E}[\,\cdot \mid \mathcal{F}_{t-1}]$. At round t the loss is $f_t : \mathcal{X} \to \mathbb{R}$, the target subgradient is g_t with known bound $G \geq 0$, and the single observation is $\widehat{g}_t = g_t + \xi_t$ with centred noise ξ_t of order $p \in (1, 2]$ and scale $\sigma \geq 0$; we write $q = (p-1)/p \in (0, 1/2]$ and $H = G + \sigma$. A comparator path is $u_{1:n} = (u_1, \ldots, u_n) \in \mathcal{X}^n$, with total variation and path complexity

$$P_n = \sum_{t=2}^{n} \|u_t - u_{t-1}\|, \qquad A_n = 1 + \frac{P_n}{D} \in [1, n], \qquad R_n(u) = \sum_{t=1}^{n} \big[f_t(x_t) - f_t(u_t) \big], \quad (1)$$

where $A_n \leq n$ because each displacement is at most D. Finally $\delta \in (0, 1)$ is a failure probability and $h = \log(4/\delta)$; neither is an input to our algorithms. All logarithms are natural except $\log_2$.

3.2 INTERACTION PROTOCOL AND EVALUATION CRITERION

At each round $t = 1, 2, \ldots$, (i) the learner commits to an $\mathcal{F}_{t-1}$-measurable decision $x_t \in \mathcal{X}$; (ii) the environment fixes a convex loss f_t and a target subgradient g_t, both $\mathcal{F}_{t-1}$-measurable, so it may adapt to the entire past but cannot see the current round's noise; (iii) the learner incurs $f_t(x_t)$ and receives exactly one stochastic sample $\widehat{g}_t$ evaluated at x_t. Nothing else is revealed: no loss values, no second query, and no gradient at any other point. The algorithm's inputs are the scale parameters (D, G, σ, p), plus T for the finite-horizon version; its output is the played sequence $x_1, x_2, \ldots$; and it is judged by the dynamic regret $R_n(u)$ of equation 1 against an *arbitrary* path.

The criterion we adopt is the uniform one: we ask for a single event of probability at least $1 - \delta$ on which the bound holds for every prefix and every path at once,

$$\mathbb{P}\Big(\forall\, n \leq T,\ \forall\, u_{1:n} \in \mathcal{X}^n :\ R_n(u) \leq \mathcal{B}_{T,n}(u, \delta) \Big) \ \geq\ 1 - \delta. \quad (2)$$

The uniform criterion is stronger than the per-path criterion $\forall u_{1:n} : \mathbb{P}(R_n(u) \leq \mathcal{B}) \geq 1 - \delta$ used by prior high-probability dynamic analyses. It permits substitution of a path that is itself a function of the whole feedback record: the hindsight-optimal path, the solution of the offline budget-constrained problem $\inf\{\sum_t f_t(u_t) : P_T(u) \leq \overline{P}\}$, or the path revealed at an arbitrary finite random time. A per-path event cannot support these substitutions; the uncountability of $\mathcal{X}^n$ rules out a direct union bound. The event in equation 2 is measurable: for each realization, both $R_n(u)$ and our bound are continuous in $u_{1:n}$ on the compact set $\mathcal{X}^n$, so non-positivity of their difference on a countable dense subset extends to all of $\mathcal{X}^n$. Two elementary consequences of the setting are used throughout: convexity gives the linearization $R_n(u) \leq \sum_{t \leq n} \langle g_t, x_t - u_t \rangle$, and boundedness gives the saturation

$R_n(u) \le GDn$. Note also that $R_n(u)$ may be negative, so no argument applying a Markov-type inequality to a nonnegative random variable is available.

3.3 Assumptions

Assumption 1 (Bounded convex domain)**.** $\mathcal{X} \subset \mathbb{R}^d$ is nonempty, closed, bounded and convex with $\|x - y\| \le D$ for all $x, y \in \mathcal{X}$, and the exact Euclidean projection $\Pi_{\mathcal{X}}$ is computable.

Assumption 2 (Non-clairvoyant convex Lipschitz losses)**.** For every t, the map $(\omega, x) \mapsto f_t(\omega, x)$ is $\mathcal{F}_{t-1} \otimes \mathcal{B}(\mathcal{X})$-measurable, f_t is convex on $\mathcal{X}$ with $|f_t(x) - f_t(y)| \le G\|x - y\|$, and there exists an $\mathcal{F}_{t-1}$-measurable g_t with $\|g_t\| \le G$ and $f_t(u) \ge f_t(x_t) + \langle g_t, u - x_t\rangle$ for all $u \in \mathcal{X}$.

Assumption 3 (One conditionally centred sample with a p-th moment)**.** The round-t feedback is $\widehat{g}_t = g_t + \xi_t$, queried at x_t and satisfying, almost surely,

$$\mathbb{E}_{t-1}[\xi_t] = 0, \qquad \mathbb{E}_{t-1}\|\xi_t\|^p \le \sigma^p, \qquad p \in (1, 2],\ \sigma \ge 0. \tag{3}$$

Assumption 4 (Known scales, unknown horizon and confidence)**.** The finite-horizon algorithm receives (D, G, σ, p, T), while the anytime algorithm receives (D, G, σ, p). The confidence level δ is used only to state the probability of the event.

Assumptions 1 and 2 define the standard bounded, nonsmooth OCO class for path-length dynamic regret (Zinkevich, 2003; Hazan, 2016; Zhang et al., 2018; Orabona, 2019; Zhao et al., 2024). The analysis uses only convexity and Lipschitzness, and applies to non-unique minimizers. Measurability of f_t with respect to $\mathcal{F}_{t-1}$ allows an adaptive but non-clairvoyant adversary: the environment cannot react to the noise it is about to generate. Without this restriction, no sublinear bound is possible. We bound only the subgradient queried by the oracle, not the whole subdifferential, so arbitrary normal-cone components at the boundary need not be bounded.

Assumption 3 is the standard heavy-tailed model (Nemirovski et al., 2009; Gorbunov et al., 2020; Zhang & Cutkosky, 2022; Sadiev et al., 2023; Liu, 2026a; Aggarwal, 2026): $p = 2$ gives bounded variance, while $p < 2$ allows infinite variance, without requiring symmetry, coordinate independence, or exponential moments. The moment bound is conditional; an unconditional bound would not control predictable subsequences (for example, take $\xi = \sigma\rho^{-1/p}B\epsilon$ with B Bernoulli(ρ) and independent sign ϵ). The oracle supplies one sample per round, and the auxiliary sequence is used only in the proof. The known scales (D, G, σ, p) determine the clipping thresholds (Gorbunov et al., 2020; Sadiev et al., 2023; Liu, 2026b). The same fixed algorithm works for every δ; δ enters only through the probability statement. Measurable conditional versions are fixed on one common probability-one event. When $D = 0$ or $G = 0$, regret is zero; when $H = G + \sigma = 0$, the algorithm returns a fixed point.

4 Algorithm: Bias-Compensated Multi-Rate Clipped OGD

The algorithm runs N copies of projected clipped online gradient descent on a geometric step-size grid and plays a weighted barycenter of their iterates. It uses the two-layer dynamic-regret template of Zhang et al. (2018), with its parameters determined by the heavy-tailed analysis.

4.1 Grid, prior, and coupled clipping thresholds

The grid is anchored at the static statistical scale and covers $[\eta_-, \eta_+]$ with $\eta_- = D/(2C_T)$ and $\eta_+ = D/(2H)$, where $H = G + \sigma$ and $C_T = \max\{H, G\sqrt{T}, \sigma T^{1/p}\}$ are the smallest and largest statistical scales the problem can produce at horizon T. Since $T \ge 1$ gives $C_T \le HT$, the grid has $N \le 1 + \lfloor \log_2 T \rfloor$ cells and satisfies $\eta_J \le \eta_+ < 2\eta_J$, so it dyadically covers the whole admissible range. It is indexed upwards, so i measures how much larger an expert's step size is than the statistically optimal static step size. The prior decays geometrically in that index, $\pi_i = 2^{-(i+1)}$ for $i < J$ with the residual mass $\pi_J = 2^{-J}$ on the last cell. This prior removes the horizon's adaptation logarithm. Whatever the unknown comparator does, the step size it calls for sits at index $i \le p^{-1} \log_2 B$ with $B = A_n + h$ the path-confidence complexity, so the cost of encoding it obeys $\log(1/\pi_i) \le \log B + \log 2 \le B$: the aggregation is charged to a quantity already present in the bound, not to $\log N = \Theta(\log T)$. Expert i clips at $\tau_i = D/\eta_i$, that is, $\eta_i\tau_i = D$. This coupling

Algorithm 1 BC-MW-OGD: bias-compensated multi-rate aggregation of clipped OGD

Require: design horizon $T \geq 1$, diameter bound $D > 0$, gradient bound $G \geq 0$, noise scale $\sigma \geq 0$, moment order $p \in (1, 2]$, projection oracle $\Pi_{\mathcal{X}}$, any $z_\circ \in \mathcal{X}$. *No δ, no path length, no comparator.*

$H \leftarrow G + \sigma$; **if** $H = 0$ **then** play $z_\circ$ at every round and stop
$C_T \leftarrow \max\{H, G\sqrt{T}, \sigma T^{1/p}\}$; $\eta_- \leftarrow D/(2C_T)$; $J \leftarrow \lfloor \log_2(C_T/H) \rfloor$; $N \leftarrow J + 1$
for $i = 0, \ldots, J$ **do** ▷ ascending grid, geometric prior, coupled thresholds
  $\eta_i \leftarrow 2^i \eta_-$; $\tau_i \leftarrow D/\eta_i$; $\lambda_i \leftarrow \eta_i/(4D^2)$; $b_i \leftarrow 2^{p-1} D \sigma^p \tau_i^{1-p}$
  $\pi_i \leftarrow 2^{-(i+1)}$ if $i < J$ else 2^{-J}; $z_{1,i} \leftarrow z_\circ$; $w_{1,i} \leftarrow \pi_i$
end for
for $t = 1, 2, \ldots, T$ **do**
  $x_t \leftarrow \sum_i \alpha_{t,i} z_{t,i}$ where $\alpha_{t,i} = \lambda_i w_{t,i} / \sum_j \lambda_j w_{t,j}$ ▷ rate-tilted barycenter
  play x_t, incur $f_t(x_t)$, and query the oracle *once* to obtain $\widehat{g}_t$
  for $i = 0, \ldots, J$ **do**
    $v_{t,i} \leftarrow C_{\tau_i}(\widehat{g}_t)$ ▷ shared sample, own threshold
    $r_{t,i} \leftarrow \langle v_{t,i}, x_t - z_{t,i} \rangle$ ▷ excess gain at the *pre-update* iterate
    $z_{t+1,i} \leftarrow \Pi_{\mathcal{X}}(z_{t,i} - \eta_i v_{t,i})$ ▷ expert step
    $w_{t+1,i} \leftarrow w_{t,i}\big[1 + \lambda_i (r_{t,i} - b_i)\big]$ ▷ Prod wealth, bias-compensated
  end for
end for

normalizes wealth increments across the grid: $|r_{t,i}| \leq D\tau_i$ while $\lambda_i = 1/(4D\tau_i)$, so $|\lambda_i r_{t,i}| \leq 1/4$ uniformly in i and the multiplicative update remains in the range where $\log(1+s) \geq s - s^2$ holds. It also makes the two costs of clipping coincide: the bias contribution $D\beta(\tau_i)$ and the inflated second-moment contribution $\eta_i \sigma^p \tau_i^{2-p}$ both reduce to $\eta_i^{p-1} D^{2-p} \sigma^p$ up to a factor $2^{p-1} \leq 2$, so one step size optimizes both at once and no separate threshold schedule is required.

4.2 One sample, many thresholds: tilted wealth and its compensation

Every expert receives the same sample, keeping the oracle cost to one query per round. Expert i accumulates wealth by the Prod rule with increment $\lambda_i(r_{t,i} - b_i)$, and the played point is the $\lambda_i w_{t,i}$-weighted barycenter of the expert iterates. Because x_t is this barycenter, the identity

$$\sum_i \lambda_i w_{t,i} \big\langle g_t,\, x_t - z_{t,i} \big\rangle = 0 \tag{4}$$

holds exactly, the true gradient drops out of the aggregate drift, and the total wealth $W_t = \sum_i w_{t,i}$ is left as a nonnegative supermartingale. A single maximal inequality controls all N experts, with no union bound over the grid.

Equation 4 cancels the *true* gradient. Under heterogeneous thresholds its sampled counterpart fails: $\mathbb{E}_{t-1} v_{t,i}$ departs from g_t by as much as $\beta(\tau_i)$ in norm, the discrepancy is expert-dependent, and nothing telescopes. Subtracting $b_i = D\beta(\tau_i)$ restores $\mathbb{E}_{t-1}[r_{t,i}] - b_i \leq \langle g_t, x_t - z_{t,i} \rangle$, after which equation 4 applies to the right-hand side and the supermartingale property returns. We also exhibit a legitimate p-th-moment distribution and a pair of expert states for which the uncompensated aggregate wealth has strictly positive conditional drift. Two further details matter. The excess $r_{t,i}$ must be formed from the pre-update iterate, because it is exactly $x_t - z_{t,i}$ that the tilted weights annihilate and exactly that difference which is predictable. And no expert is ever queried at its own iterate: the decomposition $\langle v_{t,i}, x_t - u_t \rangle = r_{t,i} + \langle v_{t,i}, z_{t,i} - u_t \rangle$ reduces each expert's task to the linearized problem generated at the meta point, and the projection inequality used for the second term is valid for arbitrary input vectors, not only for subgradients at $z_{t,i}$.

4.3 Removing the Horizon

Algorithm 1 uses T only through C_T. Algorithm 2 removes it by partitioning time into epochs $I_k = \{2^k, \ldots, 2^{k+1} - 1\}$ and running a fresh instance with design horizon $T_k = 2^k$ in epoch k. The schedule is deterministic and oblivious; it uses no change-point detection or estimation. The analysis allocates failure probability $\delta_k = \delta/((k+1)(k+2))$ to epoch k, which telescopes to δ over

Algorithm 2 ANYTIME BC-MW-OGD: deterministic epoch schedule, no horizon input

Require: D, G, σ, p, projection oracle $\Pi_{\mathcal{X}}$, any $z_\circ \in \mathcal{X}$. *No horizon, no δ, no restart detection.*

for $k = 0, 1, 2, \ldots$ **do**
  instantiate a fresh BC-MW-OGD (Algorithm 1) with design horizon $T_k \leftarrow 2^k$ and initial point $z_\circ$
  for $t \in I_k = \{2^k, \ldots, 2^{k+1} - 1\}$ **do**
    obtain x_t from that instance, play it, query the oracle once, and return $\widehat{g}_t$ to it
  end for
  discard the instance and all its state
end for

the infinite horizon. Each restart resets the potential, so each epoch contributes a local complexity $1 + P_k/D$ whose constant survives the sum. The resulting confidence term contains the iterated logarithm $h_n = \log\big(4m_n(m_n+1)/\delta\big)$ with $m_n = 1 + \lfloor \log_2 n \rfloor$.

5 MAIN RESULTS: FROM CLIPPING BIAS TO A COMPARATOR-FREE EVENT

Proofs for this section appear in the appendix, in the order used below. The appendix is self-contained: Appendix A proves from scratch the projection, supermartingale, and moment tools used by the arguments. We write $v_{t,i} = C_{\tau_i}(\widehat{g}_t)$, $e_{t,i} = \mathbb{E}_{t-1} v_{t,i} - v_{t,i}$ and $Q_{n,i} = \sum_{t \le n} \|v_{t,i}\|^2$ for the clipped feedback, its centred residual, and its accumulated square, and $\beta(\tau) = 2^{p-1}\sigma^p \tau^{1-p}$ for the clipping bias at radius τ, so that $b_i = D\beta(\tau_i)$.

The proof combines clipping bounds that separate deterministic gradients from heavy-tailed noise, an analysis-only descent sequence that restores a predictable noise direction, and a bias-compensated wealth aggregate over the step-size grid. The resulting event is uniform over prefixes and comparator paths; the detailed lemmas and proofs are given in Appendix B.

5.1 THE MAIN BOUND

Theorem 1 (Uniform high-probability dynamic regret)**.** *Run Algorithm 1 under Assumptions 1–4 and let $h = \log(4/\delta)$. With probability at least $1 - \delta$, all prefixes $n \le T$ and all comparator paths satisfy simultaneously*

$$R_n(u) \le \min\Big\{GDn,\ 146GD\sqrt{T(A_n + h)} + 160\sigma D T^{1/p}(A_n + h)^q\Big\}, \tag{5}$$

and in particular $R_T(u) \le \min\{GDT,\ 146GD\sqrt{T(A_T + h)} + 160\sigma D T^{1/p}(A_T + h)^q\}$ for every $u_{1:T} \in \mathcal{X}^T$ at once.

The bound has no hidden logarithm: T, d and the path complexity enter only where displayed, and the absence of a $\log\log T$ is a consequence of the prior, since the index the unknown path selects satisfies $\log(1/\pi_i) \le \log B + \log 2 \le B$ with $B = A_n + h$, so the encoding cost is absorbed by the complexity itself rather than charged separately. The constants are explicit and stay bounded as $p \downarrow 1$, because $1/(p-1)$ never appears. The step-size comparison is raised to the power $p-1$ rather than divided by it. The linear truncation matters when the statistical term exceeds GDn, since the regret is then saturated and any lower bound must respect the same ceiling. Note also that δ is not an algorithm input, so one fixed trajectory satisfies equation 5 at every confidence level, with a separate event for each. The non-truncated part of equation 5 uses the design horizon T even at prefixes $n < T$; obtaining a genuinely n-dependent statement requires a different algorithm, which is the subject of Theorem 2. The same event also yields the noiseless, static-comparator, and hindsight consequences recorded in Appendix B.

5.2 REMOVING THE HORIZON

Theorem 2 (Anytime, all prefixes, all paths)**.** *Let $m_n = 1 + \lfloor \log_2 n \rfloor$ and $h_n = \log\big(4m_n(m_n + 1)/\delta\big)$. Run Algorithm 2. There is a single event of probability at least $1 - \delta$ on which every integer*

$n \geq 1$ *and every* $u_{1:n} \in \mathcal{X}^n$ *satisfy*

$$R_n(u) \leq \min\Big\{GDn,\ 960GD\sqrt{n(A_n + h_n)} + 960\sigma D n^{1/p}(A_n + h_n)^q\Big\}. \tag{6}$$

The anytime algorithm and its bound use the elapsed time n rather than a declared T. The constants increase from $146/160$ to 960, and the confidence term contains an iterated logarithm through m_n. Each epoch keeps its local complexity $1+P_k/D$; summing these terms works because epoch lengths double and the local complexities sum to at most $A_n - 1$.

5.3 Lower bounds and tightness

The lower bounds are proved from first principles for arbitrary, possibly randomized algorithms that observe gradients only, and the hard instance is fixed after averaging over the learner's randomization, so it never depends on the realized seed.

The construction behind equation 67 reveals the gradient only rarely: within each block the oracle returns a large value with probability $\rho = 1/(2m)$ and zero otherwise, which keeps the conditional p-th moment inside budget while leaving the learner uninformed about the block's sign for the first half of the block in expectation. This is also why the result is stated for gradient-only feedback: revealing exact loss values at a non-zero point would disclose the sign directly. The lower bound therefore applies to the gradient-only feedback model.

Theorem 3 (Joint path–confidence noise complexity)**.** *Let* $e^{-T}/4 \leq \delta \leq 1/32$*,* $h_- = \log(1/(4\delta))$ *and* $\overline{P} \in [0, D(T-1)]$*. For any gradient-only algorithm there exist fixed losses, a fixed comparator path with* $P_T \leq \overline{P}$*, and a legitimate oracle such that*

$$\mathbb{P}\Big[R_T(u) \geq \tfrac{1}{320}\min\big\{GDT,\ GD\sqrt{T\overline{A}} + \sigma D T^{1/p}(\overline{A} + h_-)^q\big\}\Big] \geq 2\delta. \tag{7}$$

The displayed interval is the admissible confidence range for this statement.

At $n = T$, the upper and lower bounds match in the heavy-tail term, including the joint dependence on path budget and confidence. Their only quantitative gap is the deterministic factor $\sqrt{A + h}$ versus $\sqrt{A}$; the rates coincide up to constants when $h = O(A)$ or when the noise term dominates. The necessity of the offset is also explicit: a two-threshold heavy-tailed example gives positive uncompensated wealth drift, whereas subtracting the per-expert bias restores the supermartingale property.

6 Numerical experiments

6.1 Experimental setup

The decision set is the Euclidean ball of diameter $D = 2$ in $\mathbb{R}^5$. Losses are $f_t(x) = G\|x - \theta_t\|$ with $G = 1$, so the instantaneous minimizer is the comparator and the recorded quantity is exactly $R_n(u)$ in equation 1. The learner receives one stochastic gradient per round. Noise is centred Pareto with tail index 1.6, moment $p = 1.5$ and scale $\sigma = 1$, hence infinite variance, and is drawn independently of the (non-clairvoyant) loss. Every method is given the same (T, D, G, σ, p) and the same sample. We compare Algorithm 1 with projected OGD tuned to $H = G + \sigma$, the same descent clipped at the static scale C_T of Section 4, and Ader (Zhang et al., 2018); Table 2 also includes Ader tuned to G only and an ablation that drops the offsets b_i. The score is the 95th-percentile dynamic regret over 1000 independent realizations (500 on the T-sweep), which is the empirical counterpart of the high-probability criterion equation 2.

6.2 Horizon and path-complexity scaling

Theorem 1 predicts a square-root horizon term together with a heavy-tail term. At fixed $A = 8$, Figure 1(a,b) shows that BC-MW-OGD is best across $T \in \{2^{12}, \dots, 2^{17}\}$ and along individual trajectories; the log-log slope is 0.48–0.52, consistent with the square-root term. Figure 1(c) and Table 2 vary the path complexity at $T = 2^{16}$. BC-MW-OGD remains best under both noise variants, while a fixed clipping threshold and the no-offset ablation degrade as A grows; unclipped Ader is consistently less robust to the heavy tails.

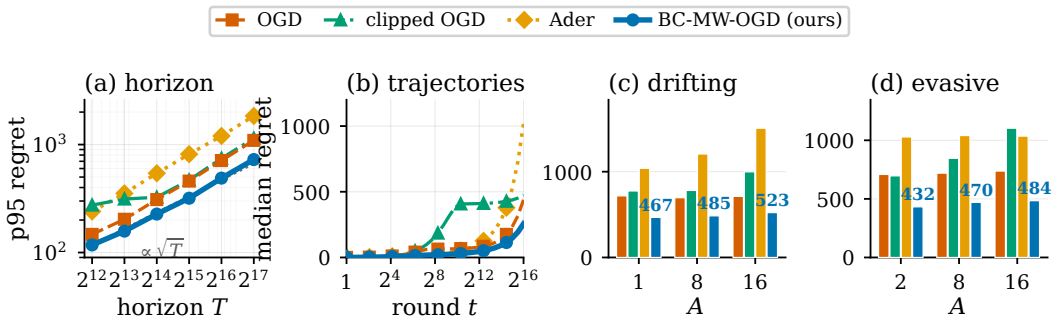


Figure 1: Heavy-tailed dynamic regret, $p = 1.5$, $\sigma = 1$ (infinite variance). (a) 95th-percentile regret versus T at $A = 8$; the dotted guide is $\propto \sqrt{T}$. (b) Median trajectories (log time) with the interquartile band of BC-MW-OGD. (c) Path-complexity sweep on an oblivious drifting target. (d) The same ranking against an evasive target that reacts to the committed x_t. One sample per round; BC-MW-OGD is lowest in every panel.

Table 2: High-probability dynamic regret on a drifting target ($T = 2^{16}$, $d = 5$, $p = 1.5$, $\sigma = 1$, infinite variance). Entries are mean / 95th percentile over 1000 runs. BC-MW-OGD is lowest in every column. Clipped OGD collapses at $A = 32$ because a single static threshold cannot track the path.

Noise	Method	$A = 1$	$A = 8$	$A = 16$	$A = 32$
symmetric	OGD	477.6 / 717.9	467.9 / 696.2	471.9 / 711.8	478.5 / 694.4
	clipped OGD	515.1 / 775.1	523.2 / 781.1	636.7 / 998.1	34728 / 35019
	Ader	849.9 / 1039.1	1032.7 / 1206.2	1346.2 / 1506.8	1754.8 / 1916.1
	Ader (G only)	1209.2 / 1372.5	1296.6 / 1459.4	1497.6 / 1656.7	1996.6 / 2174.2
	ours, no offset	662.3 / 816.2	1098.7 / 1227.3	1482.9 / 1600.3	1954.2 / 2058.1
	BC-MW-OGD	**250.4 / 466.7**	**285.6 / 485.1**	**351.1 / 522.8**	**474.6 / 631.0**
asymmetric	OGD	559.0 / 959.4	495.2 / 791.7	490.0 / 766.6	504.0 / 772.1
	clipped OGD	552.2 / 931.4	542.1 / 793.8	655.9 / 1134.8	34734 / 35118
	Ader	983.0 / 1291.9	1020.4 / 1235.9	1304.5 / 1480.6	1694.0 / 1870.1
	Ader (G only)	1356.4 / 1618.3	1229.7 / 1422.9	1409.9 / 1561.2	1820.8 / 2039.6
	ours, no offset	796.2 / 1146.7	1098.5 / 1288.1	1462.0 / 1632.0	1940.8 / 2100.1
	BC-MW-OGD	**325.7 / 780.6**	**299.9 / 554.6**	**351.6 / 584.3**	**474.0 / 685.7**

6.3 Comparator-uniformity and bias compensation

An evasive target chooses its next minimizer after observing the committed decision, so the comparator depends on the feedback history. Figure 1(d) and Table 3 in Appendix D show that BC-MW-OGD retains the best ranking, while the explicit envelope in equation 5 remains valid on these runs. At $A = 8$, its 95th-percentile regret is about 470, compared with about 720 for OGD and 1040 for Ader. The observed ratio $\max_{n \le T} R_n / \mathcal{B}_n$ stays below 0.50. The offset mechanism is tested in Figure 2 in Appendix D: without compensation, the aggregate wealth has positive drift and grows, whereas the offset restores the supermartingale behavior.

7 Conclusion

We studied dynamic regret with conditionally p-th-moment stochastic gradients, including infinite-variance noise. BC-MW-OGD gives one high-probability event for every prefix and comparator path under a one-gradient oracle, with an anytime extension and matching heavy-tail exponents in the lower bounds. The current result uses the scales (D, G, σ, p) and leaves a gap in the deterministic path-confidence term; removing these restrictions is the main direction for future work.

AI USE STATEMENT

Generative AI tools were used to polish the language of this manuscript. All AI-assisted edits were reviewed by the authors, who take full responsibility for the final content of the paper.

REPRODUCIBILITY STATEMENT

The main contributions are theoretical. Sections 3–5 state the problem, assumptions, algorithms, and guarantees, while Appendices A–C provide the auxiliary results, complete upper-bound proofs, and lower-bound constructions, while Appendix D contains the additional experimental table and figure. The synthetic experiments in Section 6 specify the loss and noise models, parameter choices, baselines, repetition counts, and evaluation criteria. The accompanying implementation contains the algorithmic updates and scripts used to generate the tables and figures; it also checks the bias-compensation identity numerically. Together, these details support independent verification of the theoretical and empirical results.

REFERENCES

Vaneet Aggarwal. Parameter-free dynamic regret under heavy-tailed noise. *arXiv preprint arXiv:2607.27073*, 2026.

Omar Besbes, Yonatan Gur, and Assaf Zeevi. Non-stationary stochastic optimization. *Operations Research*, 63(5):1227–1244, 2015. doi: 10.1287/opre.2015.1408.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006. doi: 10.1017/CBO9780511546921.

Umut Şimşekli, Levent Sagun, and Mert Gürbüzbalaban. A tail-index analysis of stochastic gradient noise in deep neural networks. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, volume 97, pp. 5827–5837, 2019.

Ashok Cutkosky and Harsh Mehta. High-probability bounds for non-convex stochastic optimization with heavy tails. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, pp. 4883–4895, 2021.

Ashok Cutkosky and Francesco Orabona. Black-box reductions for parameter-free online learning in Banach spaces. In *Proceedings of the 31st Conference on Learning Theory (COLT)*, volume 75, pp. 1493–1529, 2018.

Pierre Gaillard, Gilles Stoltz, and Tim van Erven. A second-order bound with excess losses. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, volume 35, pp. 176–196, 2014.

Eduard Gorbunov, Marina Danilova, and Alexander Gasnikov. Stochastic optimization with heavy-tailed noise via accelerated gradient clipping. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pp. 15042–15053, 2020.

Eric C. Hall and Rebecca M. Willett. Dynamical models and tracking regret in online convex programming. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, volume 28, pp. 579–587, 2013.

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3–4):157–325, 2016. doi: 10.1561/9781680831719.

Mark Herbster and Manfred K. Warmuth. Tracking the best expert. *Machine Learning*, 32(2): 151–178, 1998. doi: 10.1023/A:1007424614876.

Ali Jadbabaie, Alexander Rakhlin, Shahin Shahrampour, and Karthik Sridharan. Online optimization: Competing with dynamic comparators. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 38, pp. 398–406, 2015.

Zijian Liu. Online convex optimization with heavy tails: Old algorithms, new regrets, and applications. In *Proceedings of the 37th International Conference on Algorithmic Learning Theory (ALT)*, pp. 1–47, 2026a.

Zijian Liu. Clipped gradient methods for nonsmooth convex optimization under heavy-tailed noise: A refined analysis. In *International Conference on Learning Representations (ICLR)*, 2026b.

Antoine Moulin, Emmanuel Esposito, and Dirk van der Hoeven. When lower-order terms dominate: Adaptive expert algorithms for heavy-tailed losses. In *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2025.

Arkadi Nemirovski, Anatoli Juditsky, Guanghui Lan, and Alexander Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009. doi: 10.1137/070704277.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.

Francesco Orabona and Dávid Pál. Coin betting and parameter-free online learning. In *Advances in Neural Information Processing Systems 29 (NeurIPS)*, pp. 577–585, 2016.

Abdurakhmon Sadiev, Marina Danilova, Eduard Gorbunov, Samuel Horváth, Gauthier Gidel, Pavel Dvurechensky, Alexander Gasnikov, and Peter Richtárik. High-probability bounds for stochastic optimization and variational inequalities: The case of unbounded variance. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, volume 202, pp. 29563–29648, 2023.

Nuri Mert Vural, Lu Yu, Krishna Balasubramanian, Stanislav Volgushev, and Murat A. Erdogdu. Mirror descent strikes again: Optimal stochastic convex optimization under infinite noise variance. In *Proceedings of the 35th Conference on Learning Theory (COLT)*, volume 178, pp. 65–102, 2022.

Yuchen Yang, Kaihong Lu, and Long Wang. Online distributed optimization with clipped stochastic gradients: High probability bound of regrets. *Automatica*, 182(5):112525, 2025. doi: 10.1016/j.automatica.2025.112525.

Jingzhao Zhang, Tianxing He, Suvrit Sra, and Ali Jadbabaie. Why gradient clipping accelerates training: A theoretical justification for adaptivity. In *International Conference on Learning Representations (ICLR)*, 2020.

Jiujia Zhang and Ashok Cutkosky. Parameter-free regret in high probability with heavy tails. In *Advances in Neural Information Processing Systems 35 (NeurIPS)*, pp. 8000–8012, 2022.

Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems 31 (NeurIPS)*, pp. 1323–1333, 2018.

Wentao Zhang, Yutong Zhang, and Wentao Mo. Noise-adaptive high-probability regret bounds for online convex optimization. *arXiv preprint arXiv:2606.08028*, 2026.

Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Dynamic regret of convex and smooth functions. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pp. 12510–12520, 2020.

Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Adaptivity and non-stationarity: Problem-dependent dynamic regret for online convex optimization. *Journal of Machine Learning Research*, 25(98):1–52, 2024.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML)*, pp. 928–936, 2003.

A AUXILIARY TOOLS PROVED FROM SCRATCH

The appendix is self-contained. Every inequality used in the main text is proved here or in Appendix B–C, and the auxiliary results are established from first principles. Throughout, $(\mathcal{F}_t)$ is the filtration of Section 3 and $\mathbb{E}_{t-1}[\cdot] = \mathbb{E}[\cdot \mid \mathcal{F}_{t-1}]$.

A.1 EUCLIDEAN PROJECTION

Lemma A.1 (Projection: existence, uniqueness, nonexpansiveness)**.** *Let $C \subset \mathbb{R}^d$ be nonempty, closed, bounded and convex. Every $a \in \mathbb{R}^d$ has a unique projection $\Pi_C(a) = \arg\min_{x \in C} \|x-a\|^2$, and for all $a, b \in \mathbb{R}^d$,*

$$\|\Pi_C(a) - \Pi_C(b)\| \le \|a - b\|, \qquad \textit{so in particular} \qquad \|\Pi_C(a) - u\| \le \|a - u\| \ \ (u \in C). \tag{8}$$

Moreover, radial clipping C_τ is the Euclidean projection onto the ball of radius τ centred at the origin.

Proof. In finite-dimensional Euclidean space a closed and bounded set is compact, and $x \mapsto \|x - a\|^2$ is continuous, so the minimum over the nonempty compact set C is attained; this gives existence. For uniqueness, suppose $x \neq y$ are both minimizers. Convexity gives $(x+y)/2 \in C$, and expanding squares,

$$\Big\|\frac{x+y}{2} - a\Big\|^2 = \frac{1}{2}\|x-a\|^2 + \frac{1}{2}\|y-a\|^2 - \frac{1}{4}\|x-y\|^2.$$

Since $x \neq y$ the last term is strictly negative, so the midpoint has strictly smaller objective value than the minimum, a contradiction.

Write $x = \Pi_C(a)$. For any $z \in C$ and $s \in [0,1]$ convexity gives $x + s(z-x) \in C$, so optimality implies $\|x + s(z-x) - a\|^2 - \|x-a\|^2 \ge 0$, that is $2s\langle x-a, z-x\rangle + s^2\|z-x\|^2 \ge 0$. Dividing by $s > 0$ and letting $s \downarrow 0$,

$$\langle a - x, z - x\rangle \le 0 \qquad \text{for all } z \in C. \tag{9}$$

Let $y = \Pi_C(b)$. Taking $z = y$ in equation 9 gives $\langle a-x, y-x\rangle \le 0$, and applying equation 9 to the projection y with the point x gives $\langle b-y, x-y\rangle \le 0$. Rewriting the former as $\langle a-x, x-y\rangle \ge 0$ and subtracting the latter, $\langle a-b-(x-y), x-y\rangle \ge 0$, whence

$$\|x-y\|^2 \le \langle a-b, x-y\rangle \le \|a-b\|\,\|x-y\|$$

by Cauchy–Schwarz. If $x = y$ the two sides of equation 8 are 0 and a nonnegative number; otherwise divide by $\|x-y\|$. For $u \in C$ we have $\Pi_C(u) = u$, since distance zero is the smallest possible value, so taking $b = u$ yields the second claim.

For the clipping statement, if $\|a\| \le \tau$ then a lies in the ball and has distance zero. If $\|a\| > \tau$ then for every z with $\|z\| \le \tau$, $\|a - z\| \ge \|a\| - \|z\| \ge \|a\| - \tau$, with equality at $z = \tau a/\|a\|$; by uniqueness this is the projection. □

A.2 A MAXIMAL INEQUALITY FOR NONNEGATIVE SUPERMARTINGALES

Lemma A.2 (Crossing probability)**.** *Let $Z_0, \dots, Z_T$ be a nonnegative integrable supermartingale with respect to $(\mathcal{F}_t)$ and $\mathbb{E}Z_0 \le 1$. Then for every $a > 0$, $\mathbb{P}\big(\max_{0 \le t \le T} Z_t \ge a\big) \le 1/a$.*

Proof. Define the first crossing time $\kappa = \inf\{t \in \{0, \dots, T\} : Z_t \ge a\}$, with $\kappa = T+1$ if the set is empty. For each $s \in \{1, \dots, T\}$ the event $\{\kappa \ge s\}$ is determined by $Z_0, \dots, Z_{s-1}$ and therefore belongs to $\mathcal{F}_{s-1}$. On every sample path,

$$Z_{\min(\kappa, T)} = Z_0 + \sum_{s=1}^{T} \mathbf{1}_{\{\kappa \ge s\}}(Z_s - Z_{s-1}). \tag{10}$$

Indeed, if $\kappa = 0$ all indicators vanish; if $1 \le \kappa \le T$ the sum runs to $s = \kappa$ and telescopes; and if $\kappa = T+1$ the sum runs to T. These three cases are exhaustive. By predictability and the tower property,

$$\mathbb{E}\big[\mathbf{1}_{\{\kappa \ge s\}}(Z_s - Z_{s-1})\big] = \mathbb{E}\big[\mathbf{1}_{\{\kappa \ge s\}}\mathbb{E}[Z_s - Z_{s-1} \mid \mathcal{F}_{s-1}]\big] \le 0,$$

so taking expectations in equation 10 gives $\mathbb{E}Z_{\min(\kappa,T)} \le \mathbb{E}Z_0 \le 1$. On $\{\kappa \le T\}$ the stopped variable is at least a, and on the complement it is nonnegative, so $a\,\mathbb{P}(\kappa \le T) \le 1$. Since $\{\kappa \le T\}$ is exactly $\{\max_{t\le T} Z_t \ge a\}$, the claim follows. No external maximal inequality or optional stopping theorem was used. □

A.3 Two exponential-moment lemmas

Lemma A.3 (Bounded nonnegative increments)**.** *Let Y_t be adapted with $0 \le Y_t \le a$ for a deterministic $a > 0$ and $\mathbb{E}_{t-1}[Y_t] \le m$ for a deterministic $m \ge 0$. For every $\varepsilon \in (0,1)$, with probability at least $1-\varepsilon$,*

$$\sum_{t=1}^{n} Y_t \le 2mn + a\log(1/\varepsilon) \qquad \textit{simultaneously for all } n \le T. \tag{11}$$

Proof. The map $s \mapsto e^s$ is convex on $[0,1]$, so writing $s = (1-s)\cdot 0 + s\cdot 1$ and using the definition of convexity, $e^s \le 1 + (e-1)s$ for $0 \le s \le 1$. Taking $s = Y_t/a \in [0,1]$ and conditioning,

$$\mathbb{E}_{t-1}\big[e^{Y_t/a}\big] \le 1 + (e-1)m/a \le \exp\big((e-1)m/a\big),$$

the last step by $1 + r \le e^r$, which follows from convexity of $e^r - 1 - r$ and the vanishing of its derivative at $r = 0$. Define

$$Z_n = \exp\Big(\frac{1}{a}\sum_{t=1}^{n} Y_t - \frac{(e-1)mn}{a}\Big), \qquad Z_0 = 1.$$

Then $\mathbb{E}_{n-1}[Z_n] = Z_{n-1}e^{-(e-1)m/a}\mathbb{E}_{n-1}[e^{Y_n/a}] \le Z_{n-1}$, so (Z_n) is a nonnegative supermartingale. Lemma A.2 with threshold $1/\varepsilon$ shows that, outside an event of probability at most ε, every $n \le T$ satisfies $\sum_{t\le n} Y_t \le (e-1)mn + a\log(1/\varepsilon)$. Finally $e - 1 < 2$: from $e = 2 + \sum_{k\ge 2} 1/k!$ and $k! \ge 2^{k-1}$ for $k \ge 2$, with strict inequality from $k = 3$ onwards,

$$\sum_{k=2}^{\infty} \frac{1}{k!} < \sum_{k=2}^{\infty} 2^{-(k-1)} = 1, \tag{12}$$

so $e < 3$ and $e - 1 < 2$, which gives equation 11. □

Lemma A.4 (Bounded martingale differences, linearized)**.** *Let X_t be adapted and real valued with $\mathbb{E}_{t-1}[X_t] = 0$, $|X_t| \le a$ and $\mathbb{E}_{t-1}[X_t^2] \le v$ for deterministic $a > 0$, $v \ge 0$. For every $\theta \in (0, 1/a]$ and $\varepsilon \in (0,1)$, with probability at least $1-\varepsilon$,*

$$\sum_{t=1}^{n} X_t \le \theta v n + \frac{\log(1/\varepsilon)}{\theta} \qquad \textit{simultaneously for all } n \le T. \tag{13}$$

Proof. For $|s| \le 1$, absolute convergence of the exponential series and equation 12 give

$$e^s = 1 + s + \sum_{k\ge 2}\frac{s^k}{k!} \le 1 + s + \sum_{k\ge 2}\frac{|s|^k}{k!} \le 1 + s + s^2\sum_{k\ge 2}\frac{1}{k!} \le 1 + s + s^2.$$

Since $|\theta X_t| \le 1$, conditioning yields $\mathbb{E}_{t-1}[e^{\theta X_t}] \le 1 + \theta\mathbb{E}_{t-1}[X_t] + \theta^2\mathbb{E}_{t-1}[X_t^2] \le 1 + \theta^2 v \le e^{\theta^2 v}$. Hence $Z_n = \exp(\theta\sum_{t\le n} X_t - \theta^2 vn)$ with $Z_0 = 1$ satisfies $\mathbb{E}_{n-1}[Z_n] = Z_{n-1}e^{-\theta^2 v}\mathbb{E}_{n-1}[e^{\theta X_n}] \le Z_{n-1}$ and is a nonnegative supermartingale. By Lemma A.2, with probability at least $1-\varepsilon$ we have $Z_n \le 1/\varepsilon$ for all $n \le T$; taking logarithms and dividing by $\theta > 0$ gives equation 13. □

We retain the form $\theta vn + \log(1/\varepsilon)/\theta$ instead of optimizing θ. Binding θ to the expert's step size produces the η^{p-1} noise term without a square-root cross term.

A.4 RADEMACHER SUMS

Lemma A.5 (Absolute first moment of a Rademacher sum)**.** *Let $\epsilon_1, \dots, \epsilon_m$ be independent signs, each ± 1 with probability $1/2$, $m \geq 1$, and $S_m = \sum_{r \leq m} \epsilon_r$. Then*

$$\mathbb{E}S_m^2 = m, \qquad \mathbb{E}S_m^4 = 3m^2 - 2m \leq 3m^2, \qquad \mathbb{E}|S_m| \geq \sqrt{m/3}. \tag{14}$$

Proof. Each sign has mean zero and square one, so expanding the square and using $\mathbb{E}[\epsilon_r \epsilon_s] = \mathbb{E}\epsilon_r \mathbb{E}\epsilon_s = 0$ for $r \neq s$ gives $\mathbb{E}S_m^2 = m$. Expanding the fourth power, any term in which some index appears an odd number of times has expectation zero, because that sign factors out and its first and third moments vanish. The surviving terms are of two kinds: the m terms with all four indices equal, each of expectation one, and the terms in which two distinct indices each appear twice, of which there are $6\binom{m}{2}$ since each pair admits $4!/(2!2!) = 6$ arrangements, again each of expectation one. Hence $\mathbb{E}S_m^4 = m + 6\binom{m}{2} = 3m^2 - 2m$.

Let $Z = |S_m| \geq 0$. Since $Z \leq m$ all moments below are finite, and $\mathbb{E}Z^2 = m > 0$. Applying Cauchy–Schwarz to the pair $Z^{1/2}, Z^{3/2}$,

$$\mathbb{E}Z^2 \leq (\mathbb{E}Z)^{1/2} (\mathbb{E}Z^3)^{1/2},$$

and applying it again to the pair Z, Z^2, $\mathbb{E}Z^3 \leq (\mathbb{E}Z^2)^{1/2} (\mathbb{E}Z^4)^{1/2}$. Substituting the second into the first and using $\mathbb{E}Z^2 = m$,

$$m \leq (\mathbb{E}Z)^{1/2} m^{1/4} (\mathbb{E}Z^4)^{1/4}.$$

Dividing by $m^{1/4} > 0$ and squaring gives $m^{3/2} \leq \mathbb{E}Z\, (\mathbb{E}Z^4)^{1/2}$. Cauchy–Schwarz applied to Z^2 and the constant 1 gives $\mathbb{E}Z^4 \geq (\mathbb{E}Z^2)^2 = m^2 > 0$, so we may divide by its square root and insert the fourth-moment bound already computed:

$$\mathbb{E}|S_m| \geq \frac{m^{3/2}}{(\mathbb{E}Z^4)^{1/2}} \geq \frac{m^{3/2}}{\sqrt{3}\, m} = \sqrt{m/3}. \qquad \square$$

A.5 TWO ELEMENTARY REAL INEQUALITIES

Lemma A.6 (Subadditivity of concave powers, and a Hölder-type sum bound)**.** *For $0 < r \leq 1$ and $a, b \geq 0$,*

$$(a+b)^r \leq a^r + b^r. \tag{15}$$

Moreover, for finitely many $a_j, b_j \geq 0$ and $p \in (1, 2]$ with $q = (p-1)/p$,

$$\sum_j a_j^{1/p} b_j^q \leq \Big(\sum_j a_j\Big)^{1/p} \Big(\sum_j b_j\Big)^q. \tag{16}$$

Proof. For equation 15, if $a = 0$ equality holds. If $a > 0$, set $z = b/a \geq 0$ and $F(z) = 1 + z^r - (1+z)^r$, so that $F(0) = 0$ and, for $z > 0$, $F'(z) = r[z^{r-1} - (1+z)^{r-1}] \geq 0$ because $r - 1 \leq 0$ makes $w \mapsto w^{r-1}$ nonincreasing (with $F' \equiv 0$ when $r = 1$). Hence $F(z) \geq 0$, and multiplying by a^r gives equation 15.

For equation 16, if either sum vanishes then the corresponding sequence is identically zero and both sides are zero. Assume both sums positive and set $\alpha = 1/p \in (0, 1)$. We first prove the weighted geometric–arithmetic mean inequality $x^\alpha y^{1-\alpha} \leq \alpha x + (1-\alpha) y$ for $x, y \geq 0$. If $y = 0$ the left side is zero. If $y > 0$, put $z = x/y$; the function $z \mapsto \alpha z + 1 - \alpha - z^\alpha$ vanishes at $z = 1$ and has derivative $\alpha(1 - z^{\alpha-1})$, which is nonpositive on $(0, 1)$ and nonnegative on $(1, \infty)$, so its minimum is zero, the value at $z = 0$ being recovered by continuity; multiplying by y gives the two-variable inequality. Now set $x_j = a_j / \sum_s a_s$ and $y_j = b_j / \sum_s b_s$ and sum the two-variable inequality termwise:

$$\sum_j x_j^{1/p} y_j^q \leq \frac{1}{p} \sum_j x_j + \frac{p-1}{p} \sum_j y_j = 1.$$

Multiplying back by the appropriate powers of the two totals gives equation 16. No constant in this argument diverges as $p \downarrow 1$. $\square$

B PROOFS OF THE UPPER BOUNDS

Lemma 1 (Clipping under a p-th central moment)**.** *Condition on a σ-algebra $\mathcal{H}$, let g be $\mathcal{H}$-measurable with $\|g\| \le G$, and let ξ satisfy $\mathbb{E}[\xi \mid \mathcal{H}] = 0$ and $\mathbb{E}[\|\xi\|^p \mid \mathcal{H}] \le \sigma^p$. For any $\tau \ge 2G$ and $v = C_\tau(g+\xi)$,*

$$\|v\| \le \tau, \qquad \big\|\mathbb{E}[v \mid \mathcal{H}] - g\big\| \le \beta(\tau), \tag{17}$$

$$\mathbb{E}\big[\|v-g\|^2 \mid \mathcal{H}\big] \le 2\sigma^p\tau^{2-p}, \qquad \mathbb{E}\big[\big\|v - \mathbb{E}[v \mid \mathcal{H}]\big\|^2 \mid \mathcal{H}\big] \le 2\sigma^p\tau^{2-p}, \tag{18}$$

$$\mathbb{E}\big[\|v\|^2 \mid \mathcal{H}\big] \le 2G^2 + 4\sigma^p\tau^{2-p}, \qquad \textit{and if also } \tau \ge 2\sigma \textit{ then } \beta(\tau) \le \tau/2. \tag{19}$$

Proof of Lemma 1. The bound $\|v\| \le \tau$ is immediate from the definition of clipping.

Since $\|g\| \le G \le \tau$, projection onto the ball leaves g fixed, so nonexpansiveness (Lemma A.1) gives

$$\|v - g\| = \|C_\tau(g+\xi) - C_\tau(g)\| \le \|\xi\|, \tag{20}$$

while the triangle inequality gives $\|v-g\| \le \|v\| + \|g\| \le \tau + G \le \frac{3}{2}\tau$. If $v = g$ the squared-error bound is trivial, both sides being zero and nonnegative. Otherwise write $\|v-g\|^2 = \|v-g\|^p\|v-g\|^{2-p}$ and use $2-p \ge 0$ together with the two displays above:

$$\|v-g\|^2 \le \|\xi\|^p\big(\tfrac{3}{2}\tau\big)^{2-p}.$$

As $0 \le 2-p < 1$ we have $(3/2)^{2-p} \le 3/2 \le 2$, and taking conditional expectations gives the first bound in equation 18.

For the bias, write $y = g+\xi$ and note that the clipping residual satisfies $\|y - C_\tau(y)\| = (\|y\| - \tau)_+$ where $a_+ = \max\{a, 0\}$. If the residual is nonzero then $\|y\| > \tau$, and $\|y\| \le \|g\| + \|\xi\| \le G + \|\xi\|$ forces $\|\xi\| > \tau - G \ge \tau/2$. Consequently

$$(\|y\| - \tau)_+ \le (\|\xi\| + G - \tau)_+ \le \|\xi\|\mathbf{1}_{\{\|\xi\| > \tau/2\}}, \tag{21}$$

and on the event $\{\|\xi\| > \tau/2\}$ we have $\|\xi\|^{p-1} \ge (\tau/2)^{p-1}$, so

$$\|\xi\|\mathbf{1}_{\{\|\xi\| > \tau/2\}} \le \Big(\frac{2}{\tau}\Big)^{p-1}\|\xi\|^p. \tag{22}$$

Conditional unbiasedness gives $\mathbb{E}[v \mid \mathcal{H}] - g = \mathbb{E}[v - (g+\xi) \mid \mathcal{H}]$, so the triangle inequality for conditional expectations together with equation 21–equation 22 yields

$$\big\|\mathbb{E}[v \mid \mathcal{H}] - g\big\| \le \mathbb{E}\big[\|v - (g+\xi)\| \mid \mathcal{H}\big] \le \Big(\frac{2}{\tau}\Big)^{p-1}\mathbb{E}\big[\|\xi\|^p \mid \mathcal{H}\big] \le 2^{p-1}\sigma^p\tau^{1-p} = \beta(\tau),$$

which is equation 17.

Let $m = \mathbb{E}[v \mid \mathcal{H}]$. Expanding the inner product,

$$\mathbb{E}\big[\|v-g\|^2 \mid \mathcal{H}\big] = \mathbb{E}\big[\|v-m\|^2 \mid \mathcal{H}\big] + \|m-g\|^2 + 2\big\langle \mathbb{E}[v-m \mid \mathcal{H}], m-g\big\rangle,$$

and the last inner product vanishes because $\mathbb{E}[v - m \mid \mathcal{H}] = 0$. Discarding the nonnegative term $\|m-g\|^2$ and using the bound just proved gives the second inequality in equation 18.

For the second moment, $\|v\|^2 = \|g+(v-g)\|^2 \le 2\|g\|^2 + 2\|v-g\|^2$, where the quadratic inequality follows from $2\langle a, b\rangle \le \|a\|^2 + \|b\|^2$, itself obtained by expanding $\|a-b\|^2 \ge 0$. Taking conditional expectations and applying equation 18 gives $\mathbb{E}[\|v\|^2 \mid \mathcal{H}] \le 2G^2 + 4\sigma^p\tau^{2-p}$. Finally, if $\tau \ge 2\sigma$ then $\beta(\tau)/\tau = 2^{p-1}(\sigma/\tau)^p \le 2^{p-1}2^{-p} = 1/2$. □

The proof keeps G^2 separate from $\sigma^p\tau^{2-p}$. Estimating the raw p-th moment of $g+\xi$ by $(G+\sigma)^p$ would conflate the deterministic gradient with the heavy-tailed noise and would not yield the separated form in Theorem 1.

Lemma 2 (Pathwise dynamic descent inequality)**.** *Let $v_1, \dots, v_T$ be arbitrary vectors, $\eta > 0$, $z_1 \in \mathcal{X}$, and $z_{t+1} = \Pi_{\mathcal{X}}(z_t - \eta v_t)$. Then for every $n \le T$ and every $u_{1:n} \in \mathcal{X}^n$, on every sample path,*

$$\sum_{t=1}^{n}\langle v_t, z_t - u_t\rangle \le \frac{D^2 + 2DP_n}{2\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\|v_t\|^2 \le \frac{D^2A_n}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\|v_t\|^2. \tag{23}$$

Proof of Lemma 2. By equation 8 applied with the point $u_t \in \mathcal{X}$,

$$\|z_{t+1} - u_t\|^2 \le \|z_t - \eta v_t - u_t\|^2 = \|z_t - u_t\|^2 - 2\eta\langle v_t, z_t - u_t\rangle + \eta^2\|v_t\|^2,$$

so rearranging and dividing by 2η,

$$\langle v_t, z_t - u_t\rangle \le \frac{\|z_t - u_t\|^2 - \|z_{t+1} - u_t\|^2}{2\eta} + \frac{\eta}{2}\|v_t\|^2. \tag{24}$$

The potential differences obey the exact identity

$$\sum_{t=1}^{n}\big(\|z_t - u_t\|^2 - \|z_{t+1} - u_t\|^2\big) = \|z_1 - u_1\|^2 - \|z_{n+1} - u_n\|^2 + \sum_{t=2}^{n}\big(\|z_t - u_t\|^2 - \|z_t - u_{t-1}\|^2\big), \tag{25}$$

obtained by reindexing $t + 1$ as t in the negative terms, retaining the first and last, and subtracting termwise. The first term is at most D^2 and the negative term is at most zero. For each middle difference set $a = \|z_t - u_t\|$ and $b = \|z_t - u_{t-1}\|$; both are at most D, and the reverse triangle inequality gives $|a - b| \le \|u_t - u_{t-1}\|$, so

$$a^2 - b^2 = (a - b)(a + b) \le |a - b|(a + b) \le 2D\|u_t - u_{t-1}\|.$$

Substituting into equation 25 bounds the total potential by $D^2 + 2DP_n$, and inserting this into the sum of equation 24 gives the first inequality of equation 23. The second follows from

$$\frac{D^2 + 2DP_n}{2\eta} = \frac{D^2}{\eta}\Big(A_n - \frac{1}{2}\Big) \le \frac{D^2 A_n}{\eta}. \qquad \square$$

Lemma 3 (Uniform transfer without comparator predictability)**.** *Fix $\tau \ge 2G$ and $\eta = D/\tau$, let $v_t = C_\tau(\widehat{g}_t)$, $e_t = \mathbb{E}_{t-1}v_t - v_t$, and let (x_t) be any predictable $\mathcal{X}$-valued sequence. Let $\varepsilon \in (0, 1)$ and $\ell = \log(1/\varepsilon)$. There is an event of probability at least $1 - 2\varepsilon$, defined without reference to any comparator, on which*

$$\sum_{t=1}^{n}\langle e_t, x_t - u_t\rangle \le \frac{D^2(A_n + 4\ell)}{\eta} + 3\eta^{p-1}D^{2-p}\sigma^p n \qquad \textit{for all } n \le T \textit{ and all } u_{1:n} \in \mathcal{X}^n. \tag{26}$$

Separately, with probability at least $1 - \varepsilon$, $\sum_{t\le n}\|v_t\|^2 \le 4G^2 n + 8\sigma^p\tau^{2-p}n + \tau^2\ell$ for all $n \le T$; and on the probability-one event of Assumption 3, $\sum_{t\le n}\langle g_t - \mathbb{E}_{t-1}v_t, x_t - u_t\rangle \le D\beta(\tau)n$ for all n and all paths.

Proof of Lemma 3. We first establish bounds involving only e_t. Integrability follows from $\|v_t\| \le \tau$, and the tower property gives $\mathbb{E}_{t-1}e_t = \mathbb{E}_{t-1}v_t - \mathbb{E}_{t-1}v_t = 0$. The triangle inequality for conditional expectations gives $\|\mathbb{E}_{t-1}v_t\| \le \tau$, hence

$$\|e_t\| \le 2\tau, \qquad \text{and by equation 18} \qquad \mathbb{E}_{t-1}\|e_t\|^2 \le 2\sigma^p\tau^{2-p}. \tag{27}$$

Fix any $y_1 \in \mathcal{X}$ and define, inside the analysis only,

$$y_{t+1} = \Pi_{\mathcal{X}}(y_t - \eta e_t). \tag{28}$$

The conditional expectation $\mathbb{E}_{t-1}v_t$ admits an $\mathcal{F}_{t-1}$-measurable version; once such a version is fixed, e_t is $\mathcal{F}_t$-measurable, and the projection is continuous by Lemma A.1, so by induction y_t is $\mathcal{F}_{t-1}$-measurable and lies in $\mathcal{X}$ for every t. This construction does not require the algorithm to know any conditional expectation: y_t is merely a random variable on the existing probability space, not an additional observation.

Define the scalar variable along the predictable direction, $X_t = \langle e_t, x_t - y_t\rangle$. Since $x_t - y_t$ is predictable, $\mathbb{E}_{t-1}X_t = \langle \mathbb{E}_{t-1}e_t, x_t - y_t\rangle = 0$, and by equation 27, the diameter bound and Cauchy–Schwarz,

$$|X_t| \le 2D\tau, \qquad \mathbb{E}_{t-1}X_t^2 \le D^2\mathbb{E}_{t-1}\|e_t\|^2 \le 2D^2\sigma^p\tau^{2-p}.$$

Apply Lemma A.4 with $a = 2D\tau$, $v = 2D^2\sigma^p\tau^{2-p}$ and $\theta = 1/(2D\tau)$, which satisfies $0 < \theta \le 1/a$. Outside an event of probability at most ε, every $n \le T$ satisfies

$$\sum_{t=1}^{n} X_t \le D\sigma^p\tau^{1-p}n + 2D\tau\ell. \tag{29}$$

Next apply Lemma A.3 with $Y_t = \|e_t\|^2$, $a = 4\tau^2$ and $m = 2\sigma^p\tau^{2-p}$, whose hypotheses are verified by equation 27. Outside an event of probability at most ε, every $n \le T$ satisfies

$$\sum_{t=1}^{n} \|e_t\|^2 \le 4\sigma^p\tau^{2-p}n + 4\tau^2\ell. \tag{30}$$

The intersection of the two events has probability at least $1 - 2\varepsilon$ and is defined in terms of x_t, y_t, e_t only; it does not mention u.

Fix an arbitrary realization in that intersection, and only now choose a comparator path $u_{1:n}$, which may depend on the current and even on all future noise. Lemma 2 is a deterministic algebraic fact, so applying it to the recursion equation 28 gives

$$\sum_{t=1}^{n}\langle e_t, y_t - u_t\rangle \le \frac{D^2A_n}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\|e_t\|^2. \tag{31}$$

Using the exact decomposition $\langle e_t, x_t - u_t\rangle = X_t + \langle e_t, y_t - u_t\rangle$ together with equation 29–equation 31,

$$\sum_{t=1}^{n}\langle e_t, x_t - u_t\rangle \le D\sigma^p\tau^{1-p}n + 2D\tau\ell + \frac{D^2A_n}{\eta} + 2\eta\sigma^p\tau^{2-p}n + 2\eta\tau^2\ell.$$

Since $\eta\tau = D$ we have $D\sigma^p\tau^{1-p} = \eta\sigma^p\tau^{2-p} = \eta^{p-1}D^{2-p}\sigma^p$ and $2D\tau\ell + 2\eta\tau^2\ell = 4D^2\ell/\eta$, so the right-hand side equals the right-hand side of equation 26. The path was chosen after the event was fixed, so the conclusion covers all paths on the same event rather than redistributing failure probability across paths.

For the second assertion, apply Lemma A.3 with $Y_t = \|v_t\|^2$, $a = \tau^2$ and $m = 2G^2 + 4\sigma^p\tau^{2-p}$, the range and conditional mean bounds being supplied by Lemma 1; equation 11 then gives $\sum_{t\le n}\|v_t\|^2 \le 2(2G^2 + 4\sigma^p\tau^{2-p})n + \tau^2\ell$, as claimed.

For the third, Lemma 1 and Cauchy–Schwarz bound each inner product by $\|g_t - \mathbb{E}_{t-1}v_t\|\,\|x_t - u_t\| \le \beta(\tau)D$. This is a pointwise norm bound and does not require u_t to be predictable; summing over $t \le n$ completes the proof. □

The argument does not make the post-hoc quantity $\langle e_t, x_t - u_t\rangle$ a martingale difference. It pays for the comparator with the potential term D^2A_n/η, which is of the same form as a term already present, so the exponent A_n^q in the final bound is unaffected.

Theorem 4 (Single clipped descent, uniform over all paths)**.** *Let $G+\sigma > 0$, $0 < \eta \le D/(2(G+\sigma))$, $\tau = D/\eta$, and run $x_{t+1} = \Pi_{\mathcal{X}}\big(x_t - \eta C_\tau(\widehat{g}_t)\big)$. With $\ell_0 = \log(3/\delta)$, with probability at least $1-\delta$ every $n \le T$ and every comparator path satisfy simultaneously*

$$R_n(u) \le \frac{D^2(2A_n + 5\ell_0)}{\eta} + 2\eta G^2 n + 9\eta^{p-1}D^{2-p}\sigma^p n. \tag{32}$$

Proof of Theorem 4. Let $v_t = C_\tau(\widehat{g}_t)$ and $e_t = \mathbb{E}_{t-1}v_t - v_t$. Convexity and the decomposition $g_t - v_t = (g_t - \mathbb{E}_{t-1}v_t) + e_t$ give

$$R_n(u) \le \sum_{t=1}^{n}\langle g_t, x_t - u_t\rangle = \sum_{t=1}^{n}\langle v_t, x_t - u_t\rangle + \sum_{t=1}^{n}\langle g_t - \mathbb{E}_{t-1}v_t, x_t - u_t\rangle + \sum_{t=1}^{n}\langle e_t, x_t - u_t\rangle.$$

Bound the first sum by Lemma 2 applied to the iterates x_t themselves, the second by the pointwise bias bound of Lemma 3, and the third by equation 26. Allocate failure probability $\delta/3$ to each of the three probabilistic events of Lemma 3, namely equation 29, equation 30 and the clipped square-sum event; their intersection has probability at least $1-\delta$ and does not mention the comparator. On that event, with $\ell_0 = \log(3/\delta)$,

$$R_n(u) \le \frac{D^2A_n}{\eta} + \frac{\eta}{2}\big(4G^2n + 8\sigma^p\tau^{2-p}n + \tau^2\ell_0\big) + D\beta(\tau)n + \frac{D^2(A_n + 4\ell_0)}{\eta} + 3\eta^{p-1}D^{2-p}\sigma^p n.$$

Write $s = \eta^{p-1} D^{2-p} \sigma^p$. Since $\tau = D/\eta$,

$$4\eta\sigma^p\tau^{2-p} = 4s, \qquad D\beta(\tau) = 2^{p-1}s \le 2s, \qquad \frac{\eta\tau^2\ell_0}{2} = \frac{D^2\ell_0}{2\eta},$$

so the noise coefficient is at most $4 + 2 + 3 = 9$, the confidence coefficient is $4 + 1/2 \le 5$, and the deterministic gradient term is $2\eta G^2 n$. Substituting gives equation 32. The event was fixed before any prefix or path was selected, so summing over t introduced no additional failure probability. □

Lemma 4 (Well-posedness and the aggregate wealth)**.** *In Algorithm 1 every wealth factor is strictly positive, all x_t and $z_{t,i}$ lie in $\mathcal{X}$, and the total wealth $W_t = \sum_i w_{t,i}$ satisfies $W_1 = 1$ and $\mathbb{E}_{t-1}[W_{t+1}] \le W_t$; hence $(W_{n+1})_{n=0}^T$ is a nonnegative supermartingale. The conclusion holds for any finite positive prior with $\sum_i \pi_i = 1$ and any step sizes $0 < \eta_i \le D/(2H)$.*

Proof of Lemma 4. By construction $\tau_i \ge 2(G + \sigma)$, so the last part of Lemma 1 gives

$$\beta_i \le \tau_i/2. \tag{33}$$

Suppose all wealths at round t are positive and all expert points lie in $\mathcal{X}$; this holds at initialization because every prior mass is strictly positive. Since $\lambda_i > 0$, the denominator in the definition of $\alpha_{t,i}$ is strictly positive, the $\alpha_{t,i}$ are positive and sum to one, and therefore x_t is a convex combination of points of $\mathcal{X}$ and lies in $\mathcal{X}$. By Cauchy–Schwarz, $\|v_{t,i}\| \le \tau_i$ and the diameter bound,

$$|r_{t,i}| \le D\tau_i. \tag{34}$$

Writing $s_{t,i} = \lambda_i(r_{t,i} - b_i)$, inequalities equation 33–equation 34 and $\lambda_i = 1/(4D\tau_i)$ give

$$|s_{t,i}| \le \lambda_i\big(D\tau_i + D\beta_i\big) \le \frac{1}{4D\tau_i}\Big(D\tau_i + \frac{D\tau_i}{2}\Big) = \frac{3}{8}. \tag{35}$$

Hence $1 + s_{t,i} \ge 5/8 > 0$ and the next wealth is strictly positive, while the projection places $z_{t+1,i}$ in $\mathcal{X}$. Induction shows the algorithm is well defined at every round. Projection is continuous by Lemma A.1 and clipping is continuous, and the remaining updates are compositions of finitely many measurable operations, so if the round's inputs are predictable then the next state is $\mathcal{F}_t$-measurable; since the initial point and the parameters are fixed, induction yields the measurability required for the predictions. Moreover equation 35 shows each wealth grows by a factor of at most $11/8$, so $0 < w_{t,i} \le \pi_i(11/8)^{t-1}$ and the wealth is integrable on a finite horizon.

For the supermartingale property, define the analysis-only true excess $a_{t,i} = \langle g_t, x_t - z_{t,i}\rangle$. As $x_t - z_{t,i}$ is predictable, $\mathbb{E}_{t-1}[r_{t,i}] = \langle \mathbb{E}_{t-1}[v_{t,i}], x_t - z_{t,i}\rangle$, so Lemma 1 and the diameter bound give

$$\mathbb{E}_{t-1}[r_{t,i}] - a_{t,i} = \big\langle \mathbb{E}_{t-1}[v_{t,i}] - g_t,\ x_t - z_{t,i}\big\rangle \le \big\|\mathbb{E}_{t-1}[v_{t,i}] - g_t\big\|\,\|x_t - z_{t,i}\| \le \beta_i D = b_i. \tag{36}$$

Consequently, by the wealth update,

$$\mathbb{E}_{t-1}[W_{t+1}] = \sum_i w_{t,i} + \sum_i \lambda_i w_{t,i}\big(\mathbb{E}_{t-1}[r_{t,i}] - b_i\big) \le W_t + \sum_i \lambda_i w_{t,i} a_{t,i}, \tag{37}$$

where the inequality uses $\lambda_i w_{t,i} > 0$. Let $S_t = \sum_i \lambda_i w_{t,i} > 0$. The definition of x_t gives $S_t x_t = \sum_i \lambda_i w_{t,i} z_{t,i}$, whence

$$\sum_i \lambda_i w_{t,i} a_{t,i} = \Big\langle g_t,\ S_t x_t - \sum_i \lambda_i w_{t,i} z_{t,i}\Big\rangle = 0,$$

which is exactly equation 4. Substituting into equation 37 gives $\mathbb{E}_{t-1}[W_{t+1}] \le W_t$, and $W_1 = \sum_i \pi_i = 1$. □

Only linearity of conditional expectation was used in equation 37; at no point did the argument require the experts' feedback to be independent, so $v_{t,0}, \dots, v_{t,J}$ being deterministic functions of one random vector is harmless. The offset b_i is precisely the bridge in equation 36 that converts heterogeneous clipping bias into a conservative wealth update.

Lemma 5 (All experts from one wealth event)**.** *For any $\varepsilon \in (0,1)$, with probability at least $1-\varepsilon$, Algorithm 1 satisfies, simultaneously for all $i \in \{0,\dots,J\}$ and all $n \le T$,*

$$\sum_{t=1}^{n} r_{t,i} \le \frac{4D^2 \log\bigl(1/(\varepsilon\pi_i)\bigr)}{\eta_i} + \frac{5}{4} b_i n + \frac{\eta_i}{2} Q_{n,i}. \tag{38}$$

Proof of Lemma 5. We first prove the logarithmic inequality we need. For $s \in [-1/2, 1/2]$ set $h(s) = \log(1+s) - s + s^2$, whose derivative is

$$h'(s) = \frac{1}{1+s} - 1 + 2s = \frac{s(1+2s)}{1+s}.$$

For $-1/2 \le s \le 0$ the denominator is positive while $s \le 0$ and $1+2s \ge 0$, so $h'(s) \le 0$ and therefore $h(s) \ge h(0) = 0$. For $0 \le s \le 1/2$ the derivative is nonnegative, so again $h(s) \ge h(0) = 0$. Hence

$$\log(1+s) \ge s - s^2 \qquad \text{for } |s| \le 1/2. \tag{39}$$

By equation 35 we have $|s_{t,i}| \le 3/8$, so equation 39 applies to every wealth increment. Taking logarithms in the multiplicative recursion and summing,

$$\log \frac{w_{n+1,i}}{w_{1,i}} = \sum_{t=1}^{n} \log(1 + s_{t,i}) \ge \lambda_i \sum_{t=1}^{n} (r_{t,i} - b_i) - \lambda_i^2 \sum_{t=1}^{n} (r_{t,i} - b_i)^2. \tag{40}$$

By Lemma 4 and Lemma A.2, with probability at least $1-\varepsilon$ we have $W_{n+1} \le 1/\varepsilon$ for all $n \le T$ simultaneously. On that event, since $w_{n+1,i} \le W_{n+1}$ and $w_{1,i} = \pi_i$,

$$\log \frac{w_{n+1,i}}{w_{1,i}} \le \log \frac{W_{n+1}}{\pi_i} \le \log\bigl(1/(\varepsilon\pi_i)\bigr) \tag{41}$$

holds for every expert; no second union bound over experts is required, because all experts share the single total-wealth event. Combining equation 40 and equation 41,

$$\sum_{t=1}^{n} r_{t,i} \le \frac{\log(1/(\varepsilon\pi_i))}{\lambda_i} + b_i n + \lambda_i \sum_{t=1}^{n} (r_{t,i} - b_i)^2. \tag{42}$$

Using $(a-b)^2 \le 2a^2 + 2b^2$ and $r_{t,i}^2 = \langle v_{t,i}, x_t - z_{t,i}\rangle^2 \le D^2 \|v_{t,i}\|^2$,

$$\lambda_i \sum_{t=1}^{n} (r_{t,i} - b_i)^2 \le 2\lambda_i D^2 Q_{n,i} + 2\lambda_i b_i^2 n.$$

By the parameter definitions $\lambda_i^{-1} = 4D^2/\eta_i$ and $2\lambda_i D^2 = \eta_i/2$, while $b_i = D\beta_i$, $\beta_i \le \tau_i/2$ and $\eta_i \tau_i = D$ give

$$2\lambda_i b_i^2 = 2 \cdot \frac{\eta_i}{4D^2} \cdot D^2 \beta_i^2 = \frac{\eta_i \beta_i^2}{2} \le \frac{\eta_i \tau_i \beta_i}{4} = \frac{D\beta_i}{4} = \frac{b_i}{4}.$$

Substituting these three identities into equation 42 gives

$$\sum_{t=1}^{n} r_{t,i} \le \frac{4D^2 \log(1/(\varepsilon\pi_i))}{\eta_i} + b_i n + \frac{b_i n}{4} + \frac{\eta_i}{2} Q_{n,i},$$

which is equation 38. □

Lemma 6 (Uniform over paths, prefixes and step sizes)**.** *Let $L_i = \log\bigl(4/(\delta\pi_i)\bigr)$. With probability at least $1-\delta$, all experts i, all prefixes $n \le T$ and all comparator paths satisfy simultaneously*

$$R_n(u) \le \frac{D^2(2A_n + 9L_i)}{\eta_i} + 4\eta_i G^2 n + 16\eta_i^{p-1} D^{2-p} \sigma^p n. \tag{43}$$

Proof of Lemma 6. Apply Lemma 5 with failure probability $\delta/4$, obtaining an event of probability at least $1 - \delta/4$ on which all experts and all prefixes satisfy

$$\sum_{t=1}^{n} r_{t,i} \le \frac{4D^2 L_i}{\eta_i} + \frac{5}{4} b_i n + \frac{\eta_i}{2} Q_{n,i}. \tag{44}$$

Next, for each expert assign failure probability $\varepsilon_i = \delta\pi_i/4$ to the clipped square-sum event of Lemma 3. Since $\sum_i \pi_i = 1$, the total failure probability of this family is at most $\delta/4$, and on the intersection all i and n satisfy

$$Q_{n,i} \le 4G^2 n + 8\sigma^p \tau_i^{2-p} n + \tau_i^2 L_i. \tag{45}$$

Finally apply the uniform transfer of Lemma 3 to each expert, again with ε_i. Each application consists of two events of failure probability at most ε_i, so the total over all experts is at most $2\sum_i \varepsilon_i = \delta/2$, and all i, n and u simultaneously satisfy

$$M_{n,i}(u) := \sum_{t=1}^{n} \langle e_{t,i}, x_t - u_t \rangle \le \frac{D^2(A_n + 4L_i)}{\eta_i} + 3\eta_i^{p-1} D^{2-p} \sigma^p n. \tag{46}$$

The three families intersect in an event of probability at least $1 - \delta/4 - \delta/4 - \delta/2 = 1 - \delta$, and this event does not depend on any comparator path. The union bounds used come only from the indicator inequality $\mathbf{1}_{E\cup F} \le \mathbf{1}_E + \mathbf{1}_F$ followed by taking expectations; neither independence of the experts' feedback nor independence of the events is needed.

Fix an arbitrary realization in the intersection and then choose i, n and u arbitrarily. Convexity together with the exact decomposition

$$\langle v_{t,i}, x_t - u_t \rangle = r_{t,i} + \langle v_{t,i}, z_{t,i} - u_t \rangle \tag{47}$$

and the splitting of $g_t - v_{t,i}$ as in the proof of Theorem 4 gives

$$R_n(u) \le \sum_{t=1}^{n} r_{t,i} + \sum_{t=1}^{n} \langle v_{t,i}, z_{t,i} - u_t \rangle + \sum_{t=1}^{n} \langle g_t - \mathbb{E}_{t-1} v_{t,i}, x_t - u_t \rangle + M_{n,i}(u).$$

The first sum is bounded by equation 44; the second by Lemma 2, which gives at most $D^2 A_n/\eta_i + (\eta_i/2)Q_{n,i}$; the third by the pointwise bias bound of Lemma 3, which gives at most $b_i n$. Hence

$$R_n(u) \le \frac{D^2(A_n + 4L_i)}{\eta_i} + \eta_i Q_{n,i} + \frac{9}{4} b_i n + M_{n,i}(u).$$

Inserting equation 45 and equation 46,

$$R_n(u) \le \frac{D^2(2A_n + 8L_i)}{\eta_i} + 4\eta_i G^2 n + 8\eta_i \sigma^p \tau_i^{2-p} n + \eta_i \tau_i^2 L_i + \frac{9}{4} D\beta_i n + 3\eta_i^{p-1} D^{2-p} \sigma^p n.$$

Now $\eta_i \tau_i = D$ gives $\eta_i \tau_i^2 L_i = D^2 L_i/\eta_i$, and writing $s_i = \eta_i^{p-1} D^{2-p} \sigma^p$ we have $\eta_i \sigma^p \tau_i^{2-p} = s_i$ and $D\beta_i = 2^{p-1} s_i \le 2 s_i$. The total noise coefficient is therefore at most $8 + \frac{9}{4}\cdot 2 + 3 = \frac{31}{2} \le 16$, the confidence coefficient is $8 + 1 = 9$, and the path coefficient is 2, which is equation 43. As i, n and u were all chosen after the event was fixed, the event covers all three choices simultaneously. □

Proof of Theorem 1. If $G = 0$ then Assumption 2 makes every loss constant on $\mathcal{X}$ and the regret vanishes, so the claim holds. Assume $G > 0$, hence $H > 0$. Take the event of Lemma 6 and fix an arbitrary realization in it, an arbitrary $n \le T$ and an arbitrary comparator path. Write $B = A_n + h \ge 1$ and define the analysis-only target step size

$$\zeta = \frac{D}{2} \min\Big\{ \frac{1}{H},\ \frac{\sqrt{B}}{G\sqrt{T}},\ \Big(\frac{B}{\sigma^p T}\Big)^{1/p} \Big\}, \tag{48}$$

where the third entry is $+\infty$ when $\sigma = 0$. The algorithm computes neither B nor ζ.

We first show that the grid covers the target step size and controls its index cost. Because $B \ge 1$, each entry of the minimum in equation 48 is at least the corresponding entry at $B = 1$, and the minimum at $B = 1$ equals $\min\{1/H, 1/(G\sqrt{T}), 1/(\sigma T^{1/p})\} = 1/C_T$. Hence

$$\eta_- = \frac{D}{2C_T} \le \zeta \le \frac{D}{2H} = \eta_+. \tag{49}$$

Since $1/p \ge 1/2$, we have $\sqrt{B}/(G\sqrt{T}) \le B^{1/p}/(G\sqrt{T})$, $1/H \le B^{1/p}/H$ and $(B/(\sigma^p T))^{1/p} = B^{1/p}/(\sigma T^{1/p})$. Selecting an index attaining the minimum among the three entries at $B = 1$, bounding the corresponding B-dependent entry by the display above, and using that the minimum is no larger than the selected entry, we obtain

$$\zeta \le B^{1/p} \eta_-. \tag{50}$$

Set $i = \lfloor \log_2(\zeta/\eta_-) \rfloor$. By equation 49 the index is nonnegative, and since $\zeta/\eta_- \le C_T/H$, taking logarithms and rounding down gives $i \le J$. The definition of the floor gives $2^i \le \zeta/\eta_- < 2^{i+1}$, that is,

$$\zeta/2 < \eta_i \le \zeta, \tag{51}$$

while equation 50 gives $i \le p^{-1} \log_2 B \le \log_2 B$. Using the prior bound $\log(1/\pi_i) \le (i+1)\log 2$,

$$\log(1/\pi_i) \le (i+1)\log 2 \le \log B + \log 2.$$

For every $B \ge 1$ the function $B - 1 - \log B$ vanishes at $B = 1$ and has derivative $1 - 1/B \ge 0$, so $\log B \le B - 1$; combined with $\log 2 < 1$,

$$\log(1/\pi_i) \le B. \tag{52}$$

The prior removes the adaptation logarithm at the horizon: the encoding cost of the selected expert is controlled by the path–confidence complexity it requires, not by the logarithm of the number of experts.

We next express the grid bound in terms of B. Since $L_i = h + \log(1/\pi_i)$, equation 52 gives

$$2A_n + 9L_i \le 2A_n + 9h + 9B = 11A_n + 18h \le 18(A_n + h) = 18B, \tag{53}$$

so Lemma 6 together with $n \le T$ yields

$$R_n(u) \le \frac{18D^2B}{\eta_i} + 4\eta_i G^2 T + 16\eta_i^{p-1} D^{2-p}\sigma^p T. \tag{54}$$

We now bound the three terms separately. By $\eta_i > \zeta/2$ and equation 48,

$$\frac{18D^2B}{\eta_i} \le \frac{36D^2B}{\zeta} = 72DB\max\Big\{H,\ G\sqrt{T/B},\ \sigma(T/B)^{1/p}\Big\},$$

and since a maximum of nonnegative numbers is at most their sum,

$$\frac{18D^2B}{\eta_i} \le 72DHB + 72GD\sqrt{TB} + 72\sigma DT^{1/p}B^q. \tag{55}$$

By $\eta_i \le \zeta \le D\sqrt{B}/(2G\sqrt{T})$,

$$4\eta_i G^2 T \le 4 \cdot \frac{D\sqrt{B}}{2G\sqrt{T}} \cdot G^2 T = 2GD\sqrt{TB}. \tag{56}$$

If $\sigma = 0$ the third term vanishes. Otherwise, since $p - 1 > 0$ the step-size bound may be raised to the power $p - 1$, giving

$$16\eta_i^{p-1} D^{2-p}\sigma^p T \le 16\Big[\frac{D}{2}\Big(\frac{B}{\sigma^p T}\Big)^{1/p}\Big]^{p-1} D^{2-p}\sigma^p T = \frac{16}{2^{p-1}} D\, B^q (\sigma^p T)^{1/p} \le 16\sigma DT^{1/p}B^q. \tag{57}$$

Note that $1/(p-1)$ never appears, so no constant diverges as $p \downarrow 1$. Adding equation 55–equation 57,

$$R_n(u) \le 72DHB + 74GD\sqrt{TB} + 88\sigma DT^{1/p}B^q. \tag{58}$$

Finally, we handle the boundary term and the linear truncation. If $B \le T$ then $B^2 \le TB$, so $GB \le G\sqrt{TB}$, and also $B = B^q B^{1/p} \le B^q T^{1/p}$. Therefore

$$72DHB = 72D(G + \sigma)B \le 72GD\sqrt{TB} + 72\sigma DT^{1/p}B^q,$$

which upon substitution into equation 58 produces the coefficients $74+72 = 146$ and $88+72 = 160$. If instead $B > T$, the deterministic bound gives $R_n(u) \le GDn \le GDT \le GD\sqrt{TB}$, which already lies below the non-truncated part of equation 5. In both cases the required non-truncated bound holds, and taking the minimum with $R_n(u) \le GDn$ gives equation 5. The realization, the prefix, the comparator path and the step-size index were all selected after the uniform event had been fixed, so data-dependent and post-hoc choices introduce no further probabilistic difficulty. Taking $n = T$ gives the second display of the theorem. □

Corollary 1 (Noiseless feedback, no confidence penalty)**.** *If the algorithm is given $\sigma = 0$, then on an event of probability one, all prefixes and all comparator paths satisfy $R_n(u) \le \min\{GDn,\ 21GD\sqrt{TA_n}\}$.*

Proof of Corollary 1. If $G = 0$ the regret is zero. Assume $G > 0$. From the conditional moment assumption and the tower property, $\mathbb{E}\|\xi_t\|^p = 0$. If a nonnegative random variable Z satisfies $\mathbb{E}Z = 0$ then for every positive integer r we have $0 \ge r^{-1}\mathbb{P}(Z \ge r^{-1})$, so each of these events is null; as $\{Z > 0\}$ is their countable union, $Z = 0$ almost surely. Intersecting over the finitely many rounds shows that all ξ_t vanish on an event of probability one.

On that event $v_{t,i} = g_t$, because $\tau_i \ge 2G$, and moreover $b_i = 0$ and $Q_{n,i} \le G^2 n$. The total wealth then satisfies, round by round,

$$W_{t+1} - W_t = \sum_i \lambda_i w_{t,i} \langle g_t, x_t - z_{t,i} \rangle = 0$$

by the weighted cancellation equation 4, so $W_t = 1$ identically. In the logarithmic derivation of Lemma 5 we may therefore use $w_{n+1,i} \le 1$ directly instead of a crossing event, obtaining

$$\sum_{t=1}^{n} r_{t,i} \le \frac{4D^2 \log(1/\pi_i)}{\eta_i} + \frac{\eta_i}{2} Q_{n,i}.$$

Adding Lemma 2 for the expert iterates, and noting that both the clipping bias and the centred noise vanish,

$$R_n(u) \le \frac{D^2\big[A_n + 4\log(1/\pi_i)\big]}{\eta_i} + \eta_i G^2 n. \tag{59}$$

Let $A = A_n \in [1, T]$. With $\sigma = 0$ we have $C_T = G\sqrt{T}$ and $\eta_- = D/(2G\sqrt{T})$. Choose

$$\zeta_0 = \frac{D}{2G}\sqrt{A/T} = \eta_- \sqrt{A}, \qquad i = \lfloor \log_2 \sqrt{A} \rfloor.$$

Since $A \le T$ this index lies in the grid and $\zeta_0/2 < \eta_i \le \zeta_0$. The prior satisfies

$$\log(1/\pi_i) \le (i+1)\log 2 \le \tfrac{1}{2}\log A + \log 2 \le \log A + \log 2 \le A,$$

the last step being equation 52 for all $A \ge 1$. Hence the right-hand side of equation 59 is at most

$$\frac{5D^2 A}{\eta_i} + \eta_i G^2 T \le \frac{10 D^2 A}{\zeta_0} + \zeta_0 G^2 T = 20GD\sqrt{TA} + \tfrac{1}{2}GD\sqrt{TA} \le 21GD\sqrt{TA}.$$

Taking the minimum with the linear bound completes the proof. The comparator path enters only through deterministic algebra, so all paths are covered simultaneously. □

Corollary 2 (Static comparators and finite variance)**.** *A static comparator has $A_T = 1$, so on the event of Theorem 1, simultaneously for all $u \in \mathcal{X}$,*

$$R_T(u) \le \min\big\{GDT,\ 146GD\sqrt{T(1+h)} + 160\sigma D T^{1/p}(1+h)^q\big\}.$$

When $p = 2$ the two statistical terms merge, and for every path

$$R_T(u) \le \min\big\{GDT,\ (146G + 160\sigma)D\sqrt{T(A_T + h)}\big\}.$$

Proof of Corollary 2. For a static comparator every consecutive difference is zero, so by definition $P_T = 0$ and $A_T = 1$; substituting into Theorem 1 gives the first display. When $p = 2$ we have $1/p = q = 1/2$, so the two non-truncated terms share the factor $D\sqrt{T(A_T + h)}$ and their coefficients add. Both substitutions take place on the same uniform event and create no new failure probability. □

Corollary 3 (Hindsight and budget-constrained comparators)**.** *On the event of Theorem 1 one may substitute any $\mathcal{X}^T$-valued random path $U_{1:T}$, measurable with respect to the entire feedback record, together with its realized complexity $A_T(U)$. Moreover, for any deterministic budget $\overline{P}$ the same event bounds the regret against $\inf\big\{\sum_{t\le T} f_t(u_t) : u_{1:T} \in \mathcal{X}^T,\ P_T(u) \le \overline{P}\big\}$, the infimum being attained.*

Proof of Corollary 3. On each realization of the uniform event, the inequality of Theorem 1 holds for every element of $\mathcal{X}^T$, hence in particular for the element $U_{1:T}(\omega)$ realized at that ω. The time at which the random map was selected plays no role in this pointwise substitution. For the budget-constrained comparison, $P_T(u)$ is continuous, so the feasible set is a closed subset of the compact set $\mathcal{X}^T$; it is nonempty because every static path is feasible; and the loss sum is continuous by Lipschitzness, hence attains its minimum there. Substituting the uniform inequality at any minimizer and using $A_T \le 1 + \overline{P}/D$ together with monotonicity of nonnegative powers gives the claim. The comparison is expressed as the minimum of a continuous function over a compact set, and the algorithm is never required to compute or observe the minimizer. □

Proof of Theorem 2. If $G = 0$ the regret is identically zero, so assume $G > 0$. In the analysis, allocate to epoch k the failure probability

$$\delta_k = \frac{\delta}{(k+1)(k+2)}, \tag{60}$$

which is an analytical allocation and not an input. Since $\frac{1}{(k+1)(k+2)} = \frac{1}{k+1} - \frac{1}{k+2}$, summing over $k = 0, \dots, M$ gives $1 - 1/(M+2)$, and letting $M \to \infty$ shows $\sum_{k \ge 0} \delta_k = \delta$.

Apply Theorem 1 to each epoch. Although the history preceding an epoch is random, the process inside the epoch is adapted to the shifted filtration, the initial total wealth is again one, and the conditional means and conditional moment bounds are unchanged. The proofs of Appendix B therefore give an unconditional failure probability of at most δ_k for that epoch; no independence across epochs is needed and nothing is conditioned on a favourable future event. For the countably many epoch failure events we use a union bound: sum indicators over finitely many epochs and then let the number of epochs grow, invoking continuity of probability along increasing sequences of events, so the total failure probability is at most $\sum_k \delta_k = \delta$. Let E be the intersection of the epoch success events. On E, every epoch satisfies Theorem 1 for all its local prefixes and all paths.

Fix $\omega \in E$ and then choose $n \ge 1$ and $u_{1:n}$ arbitrarily; write $m = m_n$. The epochs visited are $k = 0, \dots, m-1$, the last possibly incomplete. Let I'_k denote the visited part of epoch k, let P_k be the corresponding local path length and $v_k = P_k/D$, so the local complexity is $1 + v_k$. Jumps of the comparator at epoch boundaries are not counted in any P_k, whence

$$\sum_{k=0}^{m-1} v_k \le P_n/D = A_n - 1. \tag{61}$$

This does not omit the cost of restarting, because the constant one in each local complexity is retained. The confidence term of epoch k is

$$h^{(k)} = \log \frac{4}{\delta_k} = \log \frac{4(k+1)(k+2)}{\delta} \le h_n,$$

since $k + 1 \le m_n$ and $k + 2 \le m_n + 1$ for $k \le m - 1$. The design horizon of each local prefix is $T_k = 2^k$ rather than its elapsed length, so applying the non-truncated part of equation 5 in each epoch and summing,

$$R_n(u) \le 146GD \sum_{k=0}^{m-1} \sqrt{T_k(1 + v_k + h_n)} + 160\sigma D \sum_{k=0}^{m-1} T_k^{1/p}(1 + v_k + h_n)^q. \tag{62}$$

We now bound the two sums. Put $c = 1 + h_n > 0$. Since $T_{m-1} = 2^{m-1} \le n$,

$$\sum_{k=0}^{m-1} T_k = 2^m - 1 < 2n, \tag{63}$$

and for any $r \in [1/2, 1)$ the finite geometric series together with $T_{m-1} \le n$ gives

$$\sum_{k=0}^{m-1} T_k^r = T_{m-1}^r \sum_{j=0}^{m-1} 2^{-rj} \le \frac{n^r}{1 - 2^{-r}} \le \frac{n^r}{1 - 2^{-1/2}} = (2 + \sqrt{2})n^r, \tag{64}$$

which covers both $r = 1/2$ and $r = 1/p$ because $1 < p \le 2$. For the deterministic sum, the square-root case of equation 15 gives $\sum_k \sqrt{T_k(c+v_k)} \le \sqrt{c}\sum_k \sqrt{T_k} + \sum_k \sqrt{T_k v_k}$, so by Cauchy–Schwarz and equation 61–equation 64,

$$\begin{aligned}\sum_k \sqrt{T_k(c+v_k)} &\le (2+\sqrt{2})\sqrt{nc} + \sqrt{\Big(\sum_k T_k\Big)\Big(\sum_k v_k\Big)} \\ &\le (2+\sqrt{2})\sqrt{nc} + \sqrt{2n(A_n-1)} \\ &\le (2+2\sqrt{2})\sqrt{n(A_n+h_n)} \le 5\sqrt{n(A_n+h_n)},\end{aligned} \tag{65}$$

where the penultimate step uses $c = 1 + h_n \le A_n + h_n$ and $A_n - 1 \le A_n + h_n$, and the last uses $\sqrt{2} \le 3/2$. For the noise sum, the q-th power case of equation 15 followed by equation 16 gives

$$\begin{aligned}\sum_k T_k^{1/p}(c+v_k)^q &\le c^q \sum_k T_k^{1/p} + \sum_k T_k^{1/p} v_k^q \\ &\le (2+\sqrt{2})n^{1/p}c^q + \Big(\sum_k T_k\Big)^{1/p}\Big(\sum_k v_k\Big)^q \\ &\le (2+\sqrt{2})n^{1/p}c^q + (2n)^{1/p}(A_n-1)^q \\ &\le (4+\sqrt{2})n^{1/p}(A_n+h_n)^q \le 6n^{1/p}(A_n+h_n)^q,\end{aligned} \tag{66}$$

using $2^{1/p} \le 2$ and $4 + \sqrt{2} \le 6$. Substituting equation 65 and equation 66 into equation 62 gives the deterministic coefficient $146 \cdot 5 = 730 \le 960$ and the noise coefficient $160 \cdot 6 = 960$. Taking the minimum with the per-round linear truncation GDn yields equation 6. Both n and the comparator path were chosen arbitrarily on the single event E, so that event covers all times and all paths simultaneously. □

Corollary 4 (Anytime noiseless bound)**.** *With input $\sigma = 0$, on an event of probability one all $n \ge 1$ and all paths satisfy $R_n(u) \le \min\{GDn,\ 105GD\sqrt{nA_n}\}$.*

Proof of Corollary 4. For each round the nonnegative zero-mean argument in the proof of Corollary 1 gives $\xi_t = 0$ almost surely, and intersecting over the countably many rounds leaves an event of probability one on which all noise vanishes. On that event Corollary 1 applied within each epoch gives the local bound $21GD\sqrt{T_k(1+v_k)}$. Setting $c = 1$ and $h_n = 0$ in the computation equation 65 shows that the sum of these local bounds is at most $5\sqrt{nA_n}$ times the constant, so the overall constant is $21 \cdot 5 = 105$. The linear truncation holds at every round, and taking the minimum completes the proof. □

Corollary 5 (Random evaluation times and vanishing positive regret)**.** *The event of Theorem 2 permits substituting any finite random time $\mathfrak{t}$, not necessarily a stopping time, and choosing the comparator after observing it. If a comparator rule satisfies $(A_n + h_n)/n \to 0$, then on that event $(R_n)_+/n \le 960GD\sqrt{(A_n+h_n)/n} + 960\sigma D\big((A_n+h_n)/n\big)^q \to 0$.*

Proof of Corollary 5. The statement for an arbitrary finite random time follows from pointwise substitution of an integer into the uniform event: on each realization $\mathfrak{t}(\omega)$ is one of the integers already covered, so no optional stopping argument is required, and the comparator may likewise be selected after $\mathfrak{t}$ is observed. For the convergence statement, the right-hand side of equation 6 is nonnegative and therefore also bounds the positive part of the regret; dividing by n and using $1/p - 1 = -q$ gives

$$\frac{(R_n)_+}{n} \le 960GD\sqrt{\frac{A_n+h_n}{n}} + 960\sigma D\Big(\frac{A_n+h_n}{n}\Big)^q.$$

Both exponents $1/2$ and q are positive, so the right-hand side tends to zero whenever the ratio does. The argument controls only the positive part and does not imply that the signed quantity R_n/n converges. □

C PROOFS OF THE LOWER BOUNDS

All constructions below use $\mathcal{X} = [-D/2, D/2]$, so one dimension suffices, and all feedback is provided exclusively through the stochastic gradient oracle. Randomized algorithms are allowed, and their internal seed is drawn before the interaction and independently of the environment's randomness. Each proof first averages over a random prior on the instance and then fixes one instance, so that the final hard instance never depends on the learner's realized seed.

Theorem 5 (Deterministic path cost)**.** *For any $D, G > 0$ and integers $1 \le K \le T$ and any online algorithm, there exist a fixed sequence of convex G-Lipschitz losses and a fixed comparator path with $P_T \le D(K-1)$, observed through a noiseless gradient oracle, such that $\mathbb{P}\big(R_T(u) \ge GD\sqrt{TK}/40\big) \ge 1/16$.*

Theorem 6 (Heavy tails interacting with the path)**.** *For any $D, G > 0$, $\sigma \ge 0$, $p \in (1, 2]$, integers $1 \le K \le T$ and any gradient-only algorithm, there exist fixed linear losses, a fixed comparator path with $P_T \le D(K-1)$, and an oracle satisfying Assumption 3, such that*

$$\mathbb{P}\Big(R_T(u) \ge \tfrac{D}{32} \min\big\{GT,\ \sigma T^{1/p} K^q\big\}\Big) \ge \tfrac{1}{8}. \tag{67}$$

Theorem 7 (Combined budget lower bound)**.** *For any path budget $\overline{P} \in [0, D(T-1)]$ with $\overline{A} = 1 + \overline{P}/D$ and any algorithm, there exist a legitimate instance with $P_T \le \overline{P}$ such that*

$$\mathbb{P}\Big[R_T(u) \ge \tfrac{1}{160} \min\big\{GDT,\ GD\sqrt{T\overline{A}} + \sigma D T^{1/p} \overline{A}^{\,q}\big\}\Big] \ge \tfrac{1}{16}. \tag{68}$$

Theorem 8 (Confidence cost is real even for static comparators)**.** *Let $e^{-T}/4 \le \delta \le 1/8$ and $h_- = \log\big(1/(4\delta)\big)$. For any gradient-only algorithm there exist a single fixed linear loss used at every round, a fixed comparator with $P_T = 0$, and an oracle satisfying Assumption 3, such that $\mathbb{P}\big[R_T(u) \ge \frac{D}{4} \min\{GT, \sigma T^{1/p} h_-^{\,q}\}\big] \ge 2\delta$.*

Proof of Theorem 5. Set $m = \lfloor T/K \rfloor \ge 1$ and $T' = Km$. For any $a \ge 1$ we have $\lfloor a \rfloor \ge a/2$: if $1 \le a < 2$ the left side is one, and if $a \ge 2$ then $\lfloor a \rfloor \ge a - 1 \ge a/2$. Hence

$$T' = K\lfloor T/K \rfloor \ge T/2. \tag{69}$$

Split the first T' rounds into K blocks of length m. Draw independent signs $\epsilon_1, \dots, \epsilon_{T'} \in \{-1, +1\}$ and set $f_t(x) = G\epsilon_t x$ for $t \le T'$ and $f_t \equiv 0$ afterwards. The oracle returns the exact slope, so $\xi_t = 0$ and the moment bound holds for every p and σ. For each block b put $S_b = \sum_{t \in b} \epsilon_t$, let $\operatorname{sgn}_+(s) = 1$ if $s \ge 0$ and -1 otherwise, and choose the block-constant comparator

$$u_t = -\frac{D}{2} \operatorname{sgn}_+(S_b), \qquad t \in \text{block } b, \tag{70}$$

extending the last block's value over the remaining $T - T'$ rounds. The comparator is constant within a block and moves by at most D between blocks, so $P_T \le D(K-1)$. On each block the comparator's accumulated loss is $\sum_{t \in b} G\epsilon_t u_t = -\frac{GD}{2}|S_b|$, which also holds when $S_b = 0$, so the sign convention is immaterial. Therefore

$$R_T = GM + \frac{GD}{2} U, \qquad M = \sum_{t=1}^{T'} \epsilon_t x_t, \qquad U = \sum_{b=1}^{K} |S_b|. \tag{71}$$

Average over the sign prior and the algorithm's seed jointly. When choosing x_t the algorithm has not yet seen the independent sign ϵ_t; viewing the seed as generated before the interaction and independent of the signs, x_t is a function of the past signs and the seed, so $\mathbb{E}[\epsilon_t x_t \mid \text{past signs and seed}] = 0$ and $\mathbb{E}M = 0$. By Lemma A.5,

$$\mathbb{E}R_T = \frac{GD}{2} \sum_{b=1}^{K} \mathbb{E}|S_b| \ge \frac{GDK\sqrt{m}}{2\sqrt{3}} = \frac{GD\sqrt{T'K}}{2\sqrt{3}}. \tag{72}$$

To pass from the mean to a constant-probability statement we control the second moment rather than apply Markov's inequality, which is unavailable because the regret may be negative. For $s < t$ the

variable $\epsilon_s x_s x_t$ is measurable with respect to the information preceding round t and the current sign has conditional mean zero, so $\mathbb{E}[\epsilon_s x_s \epsilon_t x_t] = 0$ and hence

$$\mathbb{E}M^2 = \sum_{t=1}^{T'} \mathbb{E}x_t^2 \le \frac{D^2 T'}{4}. \tag{73}$$

For each sign path Cauchy–Schwarz gives $U^2 = (\sum_b |S_b|)^2 \le K \sum_b S_b^2$, so taking expectations and using $\mathbb{E}S_b^2 = m$,

$$\mathbb{E}U^2 \le K \cdot Km = KT'. \tag{74}$$

By $(a+b)^2 \le 2a^2 + 2b^2$ together with equation 71–equation 74,

$$\mathbb{E}R_T^2 \le 2G^2 \mathbb{E}M^2 + \frac{G^2 D^2}{2}\mathbb{E}U^2 \le \frac{G^2 D^2 T'}{2} + \frac{G^2 D^2 KT'}{2} \le G^2 D^2 KT', \tag{75}$$

the last step using $K \ge 1$. Let $a_0 = GD\sqrt{T'K}/(2\sqrt{3})$, $b_0 = a_0/8$ and $F = \{R_T \ge b_0\}$. By equation 72, $\mathbb{E}R_T \ge a_0$, and even though R_T may be negative,

$$\mathbb{E}R_T = \mathbb{E}[R_T \mathbf{1}_{F^c}] + \mathbb{E}[R_T \mathbf{1}_F] \le b_0 + \mathbb{E}[|R_T|\mathbf{1}_F] \le b_0 + \sqrt{\mathbb{E}R_T^2\, \mathbb{P}(F)},$$

the first bound because $R_T < b_0$ on F^c with $b_0 > 0$, and the second by Cauchy–Schwarz. Rearranging, the square root is at least $a_0 - b_0 = 7a_0/8$, so using $a_0^2 = G^2 D^2 T' K/12$ and equation 75,

$$\mathbb{P}(F) \ge \frac{49 a_0^2}{64\, \mathbb{E}R_T^2} \ge \frac{49}{64 \cdot 12} = \frac{49}{768} \ge \frac{1}{16}. \tag{76}$$

By equation 69,

$$b_0 = \frac{GD\sqrt{T'K}}{16\sqrt{3}} \ge \frac{GD\sqrt{TK}}{16\sqrt{6}} \ge \frac{GD\sqrt{TK}}{40},$$

the last inequality because $6 \le 25/4$ gives $\sqrt{6} \le 5/2$ and $16\sqrt{6} \le 40$. So far the probability is joint over the sign prior and the algorithm's randomization. Writing $\psi(e) = \mathbb{P}_{\mathrm{alg}}(R_T \ge GD\sqrt{TK}/40 \mid \epsilon_{1:T'} = e)$, the displays above show that the finite average of the $\psi(e)$ is at least $1/16$; if every $\psi(e)$ were smaller than that number so would be their average, a contradiction. Hence some fixed sign sequence e has $\psi(e) \ge 1/16$, and once it is fixed both the losses and the comparator equation 70 are determined before the interaction and satisfy $P_T \le D(K-1)$. □

Proof of Theorem 6. Again let $m = \lfloor T/K \rfloor$ and $T' = Km \ge T/2$, and set

$$\rho = \frac{1}{2m}, \qquad \mu = \min\Big\{G,\ \frac{\sigma}{2m^q}\Big\}. \tag{77}$$

Fix arbitrary block signs $\theta_1, \dots, \theta_K \in \{-1, +1\}$ for the moment and set, on block b,

$$f_t(x) = \theta_b \mu x, \qquad u_t = -\theta_b D/2, \tag{78}$$

with zero losses on the remaining rounds and the comparator held at its last value. The slope has magnitude $\mu \le G$ and the path length is at most $D(K-1)$. The comparator is the exact per-round minimizer, so every round's regret is nonnegative:

$$f_t(x_t) - f_t(u_t) = \mu\Big(\frac{D}{2} + \theta_b x_t\Big) \in [0, \mu D]. \tag{79}$$

For each round in the block draw an independent Bernoulli B_t with $\mathbb{P}(B_t = 1) = \rho$ and return

$$\widehat{g}_t = \begin{cases} \theta_b \mu/\rho, & B_t = 1, \\ 0, & B_t = 0. \end{cases} \tag{80}$$

The sampling distribution does not depend on the query point. Given the fixed signs and the past, B_t is independent, so $\mathbb{E}_{t-1}\widehat{g}_t = \rho \cdot \theta_b \mu/\rho = \theta_b \mu = g_t$, and the noise $\xi_t = \widehat{g}_t - g_t$ has conditional central moment

$$\mathbb{E}_{t-1}|\xi_t|^p = (1-\rho)\mu^p + \rho\big(\mu(1/\rho - 1)\big)^p = \mu^p\big[(1-\rho) + \rho^{1-p}(1-\rho)^p\big] \le \mu^p\big[1 + \rho^{1-p}\big] \le 2\mu^p \rho^{1-p}, \tag{81}$$

the final step because $\rho \le 1$ and $p > 1$ give $\rho^{1-p} \ge 1$. By equation 77,

$$2\mu^p \rho^{1-p} \le 2\Big(\frac{\sigma}{2m^q}\Big)^p (2m)^{p-1} = 2 \cdot \frac{\sigma^p}{2^p m^{p-1}} \cdot 2^{p-1} m^{p-1} = \sigma^p, \tag{82}$$

so the oracle satisfies Assumption 3 for every fixed sign sequence.

Now place an independent uniform prior on the block signs, independent of both the algorithm's seed and all the Bernoulli variables. Consider round r of a block, $1 \le r \le m$, and define $E_{b,r} = \{\text{the first } r-1 \text{ Bernoullis of block } b \text{ all vanish}\}$. A union bound gives

$$\mathbb{P}(E_{b,r}) \ge 1 - (r-1)\rho \ge 1 - m\rho = \frac{1}{2}. \tag{83}$$

On $E_{b,r}$ all feedback received so far within the block is zero. Fixing the algorithm's seed, all randomness of the other blocks, and the block's earlier Bernoulli values, the learner's information is independent of θ_b, so its current decision is also independent of θ_b under this conditioning. As the uniform sign prior has mean zero,

$$\mathbb{E}\big[\theta_b x_t \mathbf{1}_{E_{b,r}}\big] = 0. \tag{84}$$

Combining round-wise nonnegativity with equation 79, equation 83 and equation 84,

$$\mathbb{E}\big[f_t(x_t) - f_t(u_t)\big] \ge \mathbb{E}\Big[\mu\big(D/2 + \theta_b x_t\big)\mathbf{1}_{E_{b,r}}\Big] = \frac{\mu D}{2}\mathbb{P}(E_{b,r}) \ge \frac{\mu D}{4},$$

and summing over the first T' rounds gives $\mathbb{E}R_T \ge \mu D T'/4$, while equation 79 gives the pathwise bounds $0 \le R_T \le \mu D T'$. If $\mu = 0$ the threshold in equation 67 is zero and $R_T = 0$, so the claim holds; assume $\mu > 0$ and set $a_1 = \mu D T'/8$. From $R_T \le \mu D T'$, $\mathbb{E}R_T \le a_1 + \mu D T'\,\mathbb{P}(R_T \ge a_1)$, so

$$\mathbb{P}(R_T \ge a_1) \ge \frac{\mu D T'/4 - \mu D T'/8}{\mu D T'} = \frac{1}{8}. \tag{85}$$

To express the threshold in terms of T and K, use $m \ge T/(2K)$ we get $GT' = GKm \ge GT/2$ and

$$\frac{\sigma}{2} K m^{1/p} \ge \frac{\sigma}{2} K \Big(\frac{T}{2K}\Big)^{1/p} = \frac{\sigma T^{1/p} K^q}{2^{1+1/p}} \ge \frac{\sigma T^{1/p} K^q}{4},$$

because $1 + 1/p \le 2$. By the definition of μ we have $\mu T' = \min\{GT',\, \sigma K m^{1/p}/2\}$, so

$$a_1 = \frac{D}{8}\min\Big\{GT',\ \frac{\sigma}{2} K m^{1/p}\Big\} \ge \frac{D}{32}\min\big\{GT,\ \sigma T^{1/p} K^q\big\}. \tag{86}$$

Finally fix the block signs. The joint probability equation 85 is an average of conditional probabilities over finitely many sign sequences, so at least one sequence attains at least $1/8$. Fixing it makes the losses and the comparator path deterministic before the interaction, and the conditional mean and moment bounds have been verified sign by sign in equation 81–equation 82. With equation 86 this gives equation 67. □

Remark 1 (On the finite support of the hard oracle). For each finite T the difficult oracle takes only two values. This does not weaken the bound: Assumption 3 defines a class of distributions with bounded conditional p-th moment, which contains these rare-large-feedback laws, and it does not require every member to have infinite variance or a regularly varying tail. The construction shows that a p-th moment budget alone cannot rule out this statistical difficulty.

Proof of Theorem 7. If $G = 0$, take all losses and all noise identically zero with a fixed comparator; both the threshold and the regret vanish and the event has probability one. Assume $G > 0$. For an integer K write

$$M = GDT, \qquad a = GD\sqrt{TK}, \qquad b = \sigma D T^{1/p} K^q,$$

and note $a \le M$ because $K \le T$. Theorem 5 supplies an instance whose regret is at least $a/40$ with probability at least $1/16$, and Theorem 6 supplies another whose regret is at least $\min\{M, b\}/32$ with probability at least $1/8$. Choose the first instance if $a \ge \min\{M, b\}$ and the second otherwise; in either case the selected instance satisfies, with probability at least $1/16$,

$$R_T \ge \frac{1}{40}\max\big\{a,\ \min\{M, b\}\big\}. \tag{87}$$

We claim

$$\max\{a, \min\{M, b\}\} \geq \tfrac{1}{2}\min\{M, a+b\}. \tag{88}$$

Indeed, if $b \geq M$ the left side is at least M while the right side is at most $M/2$; and if $b < M$ the left side equals $\max\{a, b\} \geq (a+b)/2$, which is also at least $\min\{M, a+b\}/2$. This covers all cases, so equation 87–equation 88 give a combined lower bound with constant $1/80$ under the budget $D(K-1)$.

Now take $K = 1 + \lfloor \overline{P}/D \rfloor$. Since $0 \leq \overline{P} \leq D(T-1)$ we have $1 \leq K \leq T$, and moreover

$$D(K-1) \leq \overline{P}, \qquad K \leq \overline{A} < K+1 \leq 2K. \tag{89}$$

Consequently $GD\sqrt{TK} \geq \frac{1}{2}GD\sqrt{T\overline{A}}$ and $\sigma DT^{1/p}K^q \geq \frac{1}{2}\sigma DT^{1/p}\overline{A}^{\,q}$, the latter using $q \leq 1$ so that $2^q \leq 2$; the former could use the better constant $1/\sqrt{2}$, but $1/2$ suffices. Finally, for nonnegative x and M we have $\min\{M, x/2\} \geq \frac{1}{2}\min\{M, x\}$: if $x \leq M$ the two sides are equal; if $M < x \leq 2M$ the left side is $x/2 \geq M/2$; and if $x > 2M$ the left side is $M \geq M/2$. Substituting the two displays into the integer-budget bound therefore costs one further factor of two, giving the constant $1/160$ and hence equation 68. □

Proof of Theorem 8. If $\sigma = 0$, choose the zero loss, zero noise and any fixed comparator; both the threshold and the regret vanish and the stated event has probability one, which is at least 2δ. Assume $\sigma > 0$; together with $G > 0$ this makes the quantity μ below strictly positive, so a non-zero revelation is distinguishable from zero feedback. The hypothesis on δ gives $0 < h_- \leq T$. Put

$$\rho = \frac{h_-}{2T} \in (0, 1/2], \qquad \mu = \min\Big\{G,\ \frac{\sigma\rho^{\,q}}{2^{1/p}}\Big\}. \tag{90}$$

Consider the two candidate fixed losses $f_+(x) = \mu x$ and $f_-(x) = -\mu x$, whose optimal comparators are $u_+ = -D/2$ and $u_- = D/2$. For a sign $\theta \in \{-1, +1\}$ the oracle returns $\theta\mu/\rho$ with probability ρ and zero otherwise, independently across rounds, so its mean is $\theta\mu$. The computation equation 81 applies verbatim and bounds the conditional central moment by $2\mu^p\rho^{1-p}$, and since $\mu^p \leq \sigma^p\rho^{p-1}/2$ by equation 90,

$$2\mu^p\rho^{1-p} \leq 2 \cdot \frac{\sigma^p\rho^{p-1}}{2} \cdot \rho^{1-p} = \sigma^p. \tag{91}$$

Let E_0 be the event that all T feedbacks equal zero; its probability is $(1-\rho)^T$ and does not depend on the sign of the loss. For $r \in [0, 1/2]$ set $k(r) = \log(1-r) + 2r$, so that $k'(r) = 2 - \frac{1}{1-r} = \frac{1-2r}{1-r} \geq 0$ and $k(0) = 0$, whence $\log(1-r) \geq -2r$. Taking $r = \rho$,

$$\mathbb{P}(E_0) = (1-\rho)^T \geq e^{-2\rho T} = e^{-h_-} = 4\delta. \tag{92}$$

On E_0 the two environments produce identical feedback, so with the algorithm's seed fixed it plays the same sequence of decisions in both. Writing $S = \sum_{t \leq T} x_t$ for that zero-feedback decision sequence, the regrets in the two environments are

$$R_+ = \mu\Big(\frac{DT}{2} + S\Big), \qquad R_- = \mu\Big(\frac{DT}{2} - S\Big), \tag{93}$$

whose sum is μDT; hence for every fixed seed at least one environment satisfies $R_\theta \geq \mu DT/2$. Averaging over the two signs and then over the seed,

$$\mathbb{P}_{\theta,\mathrm{alg},\mathrm{oracle}}\big(E_0 \text{ and } R_\theta \geq \mu DT/2\big) \geq \tfrac{1}{2}\mathbb{P}(E_0) \geq 2\delta, \tag{94}$$

so at least one fixed sign has conditional probability at least 2δ; selecting it fixes both the loss and the comparator. Finally, using $q + 1/p = 1$,

$$\begin{aligned}\frac{\mu DT}{2} &= \frac{D}{2}\min\Big\{GT,\ \frac{\sigma T\rho^{\,q}}{2^{1/p}}\Big\} = \frac{D}{2}\min\Big\{GT,\ \frac{\sigma T^{1/p}h_-^{\,q}}{2^{q+1/p}}\Big\} \\ &= \frac{D}{2}\min\Big\{GT,\ \frac{\sigma T^{1/p}h_-^{\,q}}{2}\Big\} \geq \frac{D}{4}\min\big\{GT,\ \sigma T^{1/p}h_-^{\,q}\big\},\end{aligned}$$

which combined with equation 94 completes the proof. □

Within the stated confidence window, the static noise term must carry a factor of order $\log(1/\delta)^q$. The result concerns the static confidence cost only and does not show that the iterated logarithm incurred for an unknown horizon is necessary.

Proof of Theorem 3. Define the deterministic nonnegative scales

$$M = GDT, \quad a = GD\sqrt{T\overline{A}}, \quad b = \sigma DT^{1/p}\overline{A}^{\,q}, \quad c = \sigma DT^{1/p}h_{-}^{\,q}.$$

Theorem 7 provides a legitimate fixed instance with

$$\mathbb{P}\Big(R_T \geq \tfrac{1}{160}\min\{M, a+b\}\Big) \geq \tfrac{1}{16} \geq 2\delta, \tag{95}$$

the last inequality following from $\delta \leq 1/32$. Theorem 8 provides a second fixed instance whose comparator is static, hence also feasible for the present budget, with

$$\mathbb{P}\Big(R_T \geq \tfrac{1}{4}\min\{M, c\}\Big) \geq 2\delta, \tag{96}$$

which is applicable because the present confidence range is contained in the one required there. Select the first instance if $\min\{M, a+b\} \geq \min\{M, c\}$ and the second otherwise; the choice depends only on deterministic problem parameters, never on the learner's realized seed. On the selected instance, with probability at least 2δ,

$$R_T \geq \frac{1}{160}\max\big\{\min(M, a+b),\ \min(M, c)\big\}. \tag{97}$$

For nonnegative x, y, M we have

$$\max\big\{\min(M, x), \min(M, y)\big\} \geq \tfrac{1}{2}\min(M, x+y), \tag{98}$$

since if $x \geq M$ or $y \geq M$ the left side equals M while the right side is at most $M/2$, and if $x < M$ and $y < M$ the left side is $\max(x, y)$, which is at least $(x+y)/2$ and hence at least $\min(M, x+y)/2$; these cases exhaust all nonnegative parameters. Apply equation 98 with $x = a+b$ and $y = c$. By the already proved subadditivity equation 15, $(\overline{A} + h_{-})^q \leq \overline{A}^{\,q} + h_{-}^{\,q}$, so multiplying by $\sigma DT^{1/p}$ gives $a+b+c \geq a + \sigma DT^{1/p}(\overline{A} + h_{-})^q$. Since the minimum is monotone in its second argument, the threshold in equation 97 is at least

$$\frac{1}{320}\min\big\{M,\ a + \sigma DT^{1/p}(\overline{A} + h_{-})^q\big\},$$

and the event still has probability at least 2δ. This is exactly equation 7. □

The two occurrences of the path budget and the confidence level can therefore be placed under a single exponent q, and within the stated window that scale is necessary. The statement is a minimax combination of two independent hard instances; it does not assert that one instance must realize both difficulty mechanisms at the same time.

D Additional experiments

Table 3: Evasive (data-dependent) comparator. The minimizer moves after seeing the committed x_t, so the path is a function of the noise. $T = 2^{16}$, $S = 1000$. On the same runs, $\max_n R_n/\mathcal{B}_n \leq 0.50$ for the explicit bound of Theorem 1.

Noise	Method	$A = 2$	$A = 4$	$A = 8$	$A = 16$
symmetric	OGD	471.7 / 708.6	477.8 / 711.8	484.2 / 719.6	492.8 / 737.6
	clipped OGD	505.3 / 696.5	526.9 / 826.2	548.6 / 846.1	679.6 / 1103.1
	Ader	843.0 / 1028.3	851.1 / 1032.0	853.4 / 1040.5	856.5 / 1035.7
	BC-MW-OGD	**243.0 / 432.0**	**254.0 / 473.8**	**260.7 / 469.9**	**272.1 / 484.5**
asymmetric	OGD	574.8 / 997.3	565.9 / 937.7	575.5 / 1016.3	590.9 / 1008.6
	clipped OGD	544.0 / 898.2	548.2 / 849.5	572.7 / 939.6	722.4 / 1317.0
	Ader	984.9 / 1300.0	973.5 / 1277.1	991.5 / 1311.8	999.2 / 1326.9
	BC-MW-OGD	**330.7 / 753.7**	**328.1 / 744.8**	**337.5 / 834.4**	**352.1 / 799.4**

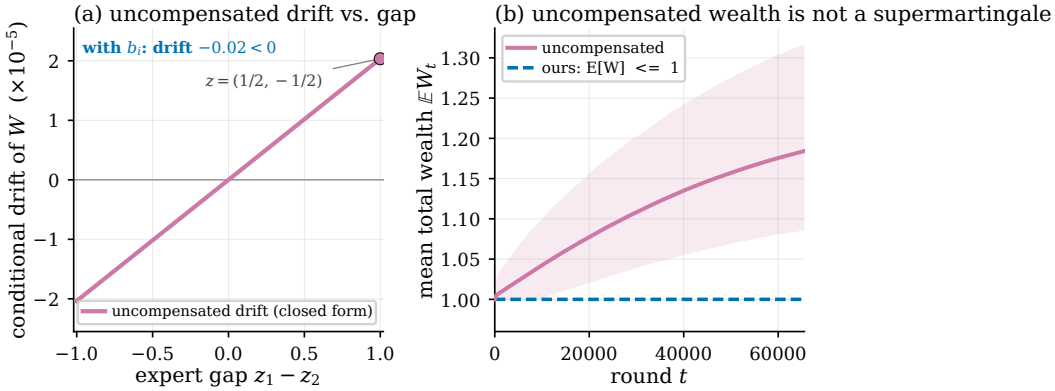


Figure 2: Necessity of the per-expert offset, from the two-point law of Section 5.3. (a) Uncompensated one-step drift of the aggregate wealth equals the displayed identity and is strictly positive whenever $z_1 > z_2$; subtracting b_i makes the drift negative. (b) Started from that state, uncompensated mean wealth (5–95% band) crosses 1 and grows. With the offset, $\mathbb{E}W_T < 10^{-200}$.