DISCOUNTED ONLINE CONVEX OPTIMIZATION WITH AN UNKNOWN DISCOUNT: SHARP RATES AND PREFIX–ENDPOINT EVALUATION

Anonymous authors
Paper under double-blind review

ABSTRACT

We study online convex optimization evaluated with a geometric discount $\lambda \in (0, 1)$ that is unknown to the learner, so one decision sequence must work for every discount. With known horizon and curvature parameter, bounded subgradients, and possibly nonsmooth losses, we use the known-discount benchmark $B_T^{\mathrm{or}}(\lambda)$. This benchmark is attainable and minimax-correct up to absolute constants on a normalized family. We introduce Sharp-BDF, which queries one genuine subgradient and no function values per round, while matching the benchmark's leading constant simultaneously for all λ. The method recycles the quadratic surplus of each base learner through a shared compensation coefficient and transfers the guarantee from a finite forgetting grid to the full discount interval without an ε-net.

The price of not knowing λ depends on the evaluation protocol. The worst-case additive excess is $\Theta(\log N)$ when every prefix must be served, but $\Theta(\log \log N)$ when only a single fixed endpoint is evaluated. Matching lower bounds hold even for randomized full-feedback algorithms and persist after subtracting the true minimax value on the steady-state discount set. A hybrid variant has constant adaptation at both scale extremes.

1 INTRODUCTION

Many online systems forget old information. The forgetting rate is often fixed by the system, while the evaluation window is chosen later. For example, Adam's second-moment estimate with $\beta_2 = 0.999$ has an effective window of about 10^3 steps, its first moment with $\beta_1 = 0.9$ has a window of about 10 (Kingma & Ba, 2015), and discounted reinforcement learning with $\gamma = 0.99$ uses an effective horizon of 10^2. A deployed trajectory may therefore be evaluated using several memory lengths, even though it cannot be rerun for each one. We study the discounted regret $R_T^{\lambda}(u) = \sum_{t \le T} \lambda^{T-t}\big(f_t(x_t) - f_t(u)\big)$, where one decision sequence must be good for every $\lambda \in (0, 1)$, while the learner never sees the value used for evaluation. Our question is the cost of this missing information.

Strong convexity gives fast rates. For $\lambda = 1$, projected gradient descent with step $1/(\mu t)$ pays $\frac{G^2}{2\mu}(1 + \log T)$ instead of the $\sqrt{T}$ of general convex losses (Zinkevich, 2003; Hazan et al., 2007). Discounted losses were first studied through aggregation in the expert setting (Chernov & Vovk, 2009; Chernov & Zhdanov, 2010), where the discount is supplied to the algorithm. For changing environments, a common approach is to run base learners on a grid of forgetting scales or restart lengths and combine them with a meta-aggregator (Hazan & Seshadhri, 2009; Daniely et al., 2015; Zhang et al., 2018). Within this framework, Yuan & Lamperski (2020) use curvature to remove the $\sqrt{\cdot}$ aggregation cost, and Zhang et al. (2024) study adaptive, time-varying discounts chosen by the learner. More recently, Yang et al. (2026) obtain uniform regret over a continuous interval for general convex losses and identify curvature as an open direction; Xie et al. (2026) connect discounted regret to dynamic regret for curved losses.

The problem of adapting to an unknown discount raises three issues. First, the usual $\log H_\lambda$ comparison hides the initialization cost, so it is too coarse for a constant-factor result. Second, preserving the curvature rate for a grid of experts usually requires observing each expert at its own decision.

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

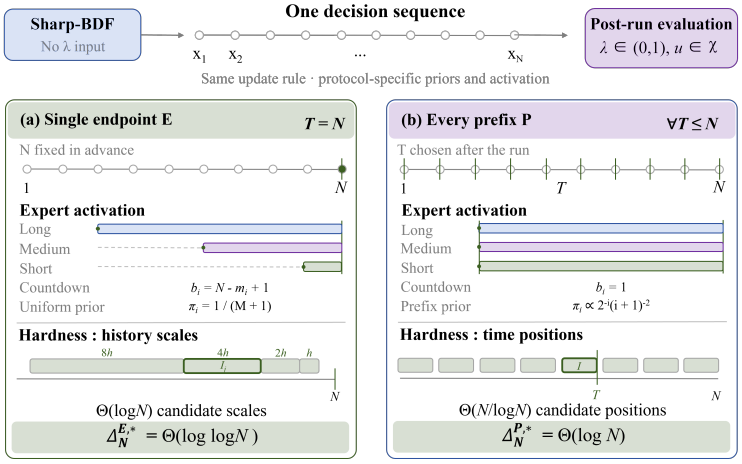


Figure 1: Overview of the two post-run evaluation protocols. A single decision sequence is evaluated either at one fixed endpoint or at every prefix. The corresponding adaptation costs are $\Theta(\log \log N)$ and $\Theta(\log N)$, respectively.

Table 1: Position relative to recent work on curvature, forgetting, and discounting. "Queries" counts genuine loss-oracle calls per round; "values" means evaluations of f_t away from the played point. $H_\lambda = \min\{N, (1-\lambda)^{-1}\}$. The last column is the worst-case *additive* excess over the stated comparator. *E* and *P* denote single-endpoint and prefix-uniform evaluation, respectively.

	Setting	Queries	Excess over comparator
Chernov & Zhdanov (2010)	expert advice; fixed λ	K values	λ supplied
Hazan et al. (2007)	strongly convex; λ=1	1 grad.	no discount
Yuan & Lamperski (2020)	strongly convex; known grid	multi-point	$O(\log)$, static/dynamic mix
Zhang et al. (2024)	convex; learner-chosen λ	1 grad.	λ is a design knob
Yang et al. (2026)	convex; unknown interval	1 grad.	absorbed in $GD\sqrt{H_\lambda \log N}$
Xie et al. (2026)	curved; discount $\to$ dynamic	1 grad.	$O(\log \log)$, dynamic regret
This paper	**strongly convex; unknown λ**	**1 grad., 0 values**	E: $\boldsymbol{\Theta(\log \log N)}$; P: $\boldsymbol{\Theta(\log N)}$

With one observed gradient, the natural surrogate bound is $3G$, which causes a factor of nine. Third, the cost of an unknown discount depends on the evaluation rule. An algorithm evaluated once at a fixed endpoint faces a different problem from one evaluated at every prefix.

We first define the comparison target. The known-discount oracle pays $G^2/(2\mu S_t(\lambda))$ in round t, where $S_t(\lambda) = \sum_{s \le t} \lambda^{t-s}$, so its discounted total is exactly $B_T^{\mathrm{or}}(\lambda) = \frac{G^2}{2\mu} K_T(\lambda)$ with $K_T(\lambda) = \sum_{t \le T} \lambda^{T-t}/S_t(\lambda)$. We show that this benchmark is attainable and minimax-correct up to absolute constants on a normalized family. The key observation behind our method is the identity $\|g_t + \mu(z - x_t)\|^2 = \|g_t\|^2 + 2\mu\ell_t(z)$ for the quadratic surrogate ℓ_t. The extra term produced by a base learner is already tracked by the meta-potential. We can therefore reuse it instead of bounding the surrogate gradient by $3G$. The compensation coefficient $c_{t,i} = 1 - 1/S_{t,i}$ must appear in both the prediction weights and the potential exponent; changing only one breaks the mass bound. A signed identity then extends the guarantees from the finite grid to every λ. We analyze two evaluation protocols: one fixed endpoint and all prefixes. Their lower bounds use different constructions, based on history scales and time positions, respectively. Figure 1 summarizes these protocols and the source of their different adaptation costs. Table 1 compares this problem with recent work. Our main results are:

- We identify $B_T^{\mathrm{or}}(\lambda) = \frac{G^2}{2\mu} K_T(\lambda)$ as the attainable known-discount benchmark, and show that it is minimax-correct up to absolute constants on a normalized family. Sharp-BDF matches its leading constant for every $\lambda \in (0, 1)$ using one subgradient and no function values per round.

- We characterize the price of an unknown discount: the worst-case excess is $\Theta(\log\log N)$ for a fixed endpoint and $\Theta(\log N)$ when all prefixes must be served. Matching lower bounds hold even against randomized full-feedback algorithms and persist after subtracting the true minimax value on the steady-state discount set.
- We give a hybrid algorithm with constant adaptation at both scale extremes and prove a cross-scale constraint that accounts for the remaining intermediate-scale cost.

2 Related Work

2.1 Curvature, Fast Rates, and the Choice of Comparator

For strongly convex losses, projected gradient descent with step $1/(\mu t)$ attains $\frac{G^2}{2\mu}(1+\log T)$. Exp-concavity gives an analogous rate through Online Newton Step (Zinkevich, 2003; Hazan et al., 2007; Cesa-Bianchi & Lugosi, 2006). These rates are reviewed in broader treatments of online convex optimization (Hazan, 2016; Shalev-Shwartz, 2012). They are often summarized by a logarithmic term up to constants. Discounted evaluation requires a more precise quantity: the discounted analogue of $\log T$ is not $\log H_\lambda$. A known-discount oracle pays $G^2/(2\mu S_t(\lambda))$ in round t, so its total is $K_T(\lambda)$. The identity $K_T(\lambda) = 1 + \lambda^T\big(\sum_{t\le T} S_t(\lambda)^{-1} - 1\big)$ separates the forgetting steady state from the initialization residue. The coarser $\log H_\lambda$ term combines these two effects. This is sufficient for order bounds but not for constant-factor comparisons. When $K_T = \Theta(\log T)$, even a multiplicative slack in the comparator exceeds a $\log\log$ additive term. Our claims are additive: the algorithm preserves the oracle's leading constant, and the $\log\log$ versus $\log$ separation is measured *after* subtraction. We therefore prove that $K_T(\lambda)$ is attainable and minimax-correct up to absolute constants on a normalized one-dimensional family.

2.2 Forgetting Mechanisms and Multiscale Aggregation

A line of work studies how a learner should forget. Discounted and windowed aggregation appeared first in the expert setting, where fixed-share methods track a changing best expert (Herbster & Warmuth, 1998; Cesa-Bianchi et al., 2012; Freund et al., 1997; Bousquet & Warmuth, 2002). Variation-aware online optimization provides a complementary route for changing environments (Hazan & Kale, 2010; Chiang et al., 2012; Jadbabaie et al., 2015). Discounted exponential weights were also studied with a fixed discount (Chernov & Vovk, 2009; Chernov & Zhdanov, 2010). For changing environments, a common approach is to run base learners on a geometric grid of window lengths or restart times and combine them with a meta-algorithm (Hazan & Seshadhri, 2009; Adamskiy et al., 2012; Daniely et al., 2015; Zhang et al., 2018; Zhao et al., 2020). Yuan & Lamperski (2020) use curvature to remove the $\sqrt{\cdot}$ cost of the meta layer. Second-order expert aggregators provide related variance-sensitive guarantees (Gaillard et al., 2014; Koolen & Van Erven, 2015; Luo & Schapire, 2015), while MetaGrad applies a related idea to learning rates (van Erven & Koolen, 2016). We use the same multiscale architecture, but heterogeneous discounts create two issues. Mixing can increase the total potential mass, while per-round normalization becomes expert-dependent and breaks telescoping; we therefore budget mass injection explicitly. Moreover, using one observed gradient naively bounds the surrogate gradient by $3G$ and loses a factor of nine. The identity $\|g_t + \mu(z - x_t)\|^2 = \|g_t\|^2 + 2\mu\ell_t(z)$ records this excess in the meta-potential, so it can be recycled when the same compensation coefficient enters both the prediction weights and the potential exponent. A counterexample shows that both uses are necessary, and a signed identity extends the finite-grid guarantees to every λ without an ε-net.

2.3 Unknown Discounts and Adaptation Costs

We next turn to work that treats the discount as an unknown input. The evaluation quantifier affects the cost of adaptation. Zhang et al. (2024) study discounted FTRL with instance-dependent and time-varying discounts, but the discount remains a knob chosen by the algorithm. Xie et al. (2026) reduce discounted to dynamic regret with applications to curved losses, where discounting is an intermediate device and the reported meta cost is an undiscounted cumulative quantity. Closest to us, Yang et al. (2026) demand uniform regret across a continuous interval of unknown discounts, achieve it for general convex losses, and explicitly name curvature as the next step. This work is

also related to methods for unknown problem parameters (Orabona & Pál, 2016; Yang et al., 2024) and lower bounds on the price of adaptivity (Carmon & Hinder, 2024). Our results separate the two evaluation protocols, giving matching $\Theta(\log\log N)$ and $\Theta(\log N)$ bounds whose lower bounds localize the difficulty to disjoint history scales and time positions, respectively. On the steady-state set $\Lambda_T^{\mathrm{st}} = \{\lambda : K_T(\lambda) \le 2\}$, both orders persist after subtracting the true known-discount minimax value. A single-gradient hybrid has constant adaptation at both scale extremes, while a cross-scale constraint accounts for the $\log\log N$ cost at intermediate scales; the projection count remains $O(\log N)$ per round (Yang et al., 2024).

3 Discounted Regret under an Unknown Discount

3.1 Notation

Logarithms are natural unless written $\log_2$, and $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^d$. The decision set $\mathcal{X} \subset \mathbb{R}^d$ is nonempty, compact and convex with diameter $D = \sup_{x,y\in\mathcal{X}} \|x - y\|$; the degenerate case $D = 0$ makes every regret vanish and is excluded. Round t carries a loss $f_t : \mathcal{X} \to \mathbb{R}$; the learner plays $x_t \in \mathcal{X}$ and queries a vector $g_t = g_t(x_t)$ of norm at most $G > 0$, and $\mu > 0$ is the strong-convexity parameter. We write $u \in \mathcal{X}$ for an arbitrary fixed comparator, $N \ge 1$ for the integer budget, and $T \in \{1, \dots, N\}$ for the endpoint of an evaluated prefix. The evaluation discount is $\lambda \in (0, 1)$, with untruncated forgetting scale $h_\lambda = (1 - \lambda)^{-1}$ and truncated scale $H_{T,\lambda} = \min\{T, h_\lambda\}$, abbreviated $H_\lambda = H_{N,\lambda}$. Two discount functionals recur throughout,

$$S_t(\lambda) = \sum_{s=1}^{t} \lambda^{t-s}, \qquad K_T(\lambda) = \sum_{t=1}^{T} \frac{\lambda^{T-t}}{S_t(\lambda)}, \tag{1}$$

the effective accumulated weight and the known-discount benchmark function; the harmonic number $\mathsf{H}_T = \sum_{t\le T} t^{-1}$ is kept notationally distinct from $H_{T,\lambda}$. Our algorithms maintain $M + 1$ base learners on a geometric forgetting grid, $M = \lceil \log_2 N \rceil$, with forgetting rates $\delta_i = \min\{2^i/N, 1\}$, discounts $\gamma_i = 1 - \delta_i \in [0, 1)$, positive priors π_i, iterates $z_{t,i} \in \mathcal{X}$ and unnormalized potential masses $q_{t,i} > 0$. Finally $[a]_+ = \max\{a, 0\}$, and $\mathbb{E}_A$ denotes expectation over the algorithm's internal randomness alone. Endpoint conventions are $S_0 = K_0 = 0$, $S_t(0) = K_T(0) = 1$ with $0^0 = 1$ for the current round's weight, and $S_t(1) = t$, $K_T(1) = \mathsf{H}_T$ as continuous extensions; the evaluated discount itself always lies in the open interval $(0, 1)$.

3.2 Task Definition and Evaluation Protocols

The interaction follows the standard online protocol. In round t, the learner chooses $x_t \in \mathcal{X}$ before observing information about f_t, then receives only the current feedback and cannot use future losses. Its inputs are (N, G, μ), a feasible initial point, and the feedback stream; its output is the sequence $x_1, \dots, x_N$. Performance is measured with the discounted regret

$$R_T^\lambda(u) = \sum_{t=1}^{T} \lambda^{T-t}\big(f_t(x_t) - f_t(u)\big), \tag{2}$$

where (T, λ, u) is selected after the run. Because the discount does not enter any update, every value of λ evaluates the same trajectory. Two evaluation protocols are considered. Under the *prefix-uniform* protocol P, any $T \le N$, $\lambda \in (0, 1)$, and $u \in \mathcal{X}$ may be chosen after the run. Under the *single-endpoint* protocol E, the endpoint is fixed in advance at N, while λ and u are chosen after the run. Protocol E models a deployment audited at a fixed date. Protocol P allows auditing at every prefix, and a guarantee for P also implies one for E. The comparison is with a known-discount oracle, an online algorithm that receives λ before round 1 but does not see future losses. This is different from an offline minimizer. Under curvature, it pays $G^2/\big(2\mu S_t(\lambda)\big)$ in round t. Its discounted total is

$$B_T^{\mathrm{or}}(\lambda) = \frac{G^2}{2\mu} K_T(\lambda), \tag{3}$$

This benchmark is attainable and minimax-correct up to absolute constants, but it is an attainable upper bound rather than a closed form for the minimax value.

Adaptivity is then quantified additively, after subtracting the benchmark of equation 3, on the normalized one-dimensional family $a = G/(2\mu)$, $\mathcal{X} = [-a, a]$, $D = G/\mu$, with $\mathfrak{F}_N$ the deterministic loss sequences on that domain satisfying Assumption 1. For an algorithm A that does not receive λ,

$$\Delta_N^{\mathrm{P}}(A) = \sup\Big[\tfrac{\mu}{G^2}\mathbb{E}_A R_T^{\lambda}(u) - \tfrac{1}{2}K_T(\lambda)\Big]_+, \qquad \Delta_N^{\mathrm{E}}(A) = \sup\Big[\tfrac{\mu}{G^2}\mathbb{E}_A R_N^{\lambda}(u) - \tfrac{1}{2}K_N(\lambda)\Big]_+, \tag{4}$$

the suprema running over $f_{1:N} \in \mathfrak{F}_N$, $u \in \mathcal{X}$ and $\lambda \in (0, 1)$, and additionally over $T \le N$ in the first. Let $\Delta_N^{\mathrm{P},\star}$ and $\Delta_N^{\mathrm{E},\star}$ denote the infima over algorithms, taken over a class that may randomize and may use full feedback. Additive excess is used rather than a ratio because a competitive ratio would conceal the $\log\log N$ term behind the leading constants. The explicit coefficient $1/2$ from equation 3 prevents the adaptation term from absorbing unstated slack. For the lower bounds we fix $D = G/\mu$, while the upper bounds hold on any compact convex $\mathcal{X}$.

3.3 Assumptions

Assumption 1 (Curvature through a subgradient oracle)**.** *$\mathcal{X}$ is nonempty, compact and convex, and every $f_t : \mathcal{X} \to \mathbb{R}$ is real-valued. For each round t and each $x \in \mathcal{X}$ the oracle returns a vector $g_t(x)$ with $\|g_t(x)\| \le G$ such that, for all $y \in \mathcal{X}$,*

$$f_t(y) \;\ge\; f_t(x) + \langle g_t(x), y - x\rangle + \frac{\mu}{2}\|y - x\|^2. \tag{5}$$

Assumption 2 (Causal single-point first-order feedback)**.** *The learner commits x_t before receiving any information about f_t, and then observes exactly one vector, $g_t(x_t)$; it queries neither loss values nor gradients at any other point. We call* full feedback *the strictly richer model in which f_t and g_t may additionally be evaluated at auxiliary points.*

Assumption 3 (Known and unknown quantities)**.** *N, G, μ and a feasible initial point are known. The evaluation discount λ is never revealed and enters no update, and all guarantees indexed by λ refer to one and the same realized decision sequence.*

Assumption 1 is the standard interface used for fast rates (Hazan et al., 2007; Cesa-Bianchi & Lugosi, 2006). It is stated in terms of the vector returned by the oracle, so it also covers nonsmooth losses. The explicit factor $1/2$ fixes the curvature convention used in all constants. Assumption 2 is the classical single-point first-order oracle (Zinkevich, 2003). Existing curvature-preserving grid methods and our baseline use full feedback, whereas Sharp-BDF uses one subgradient. The lower bounds still allow randomization and full feedback, including access to the entire loss function. Hence the separation does not come from limited feedback. The upper bounds hold pathwise, even for adaptive losses; the lower bounds use obliviously randomized sequences. Assumption 3 specifies the information supplied to the learner. Strongly convex OCO already uses μ in the step size $1/(\mu t)$ (Hazan et al., 2007); overestimating it would invalidate our surrogate comparison. The horizon N determines the grid and prior, and the endpoint algorithms also use the number of remaining rounds.

4 Algorithm: Surrogate Recycling with One Gradient

After playing x_t, the learner observes one vector $g_t = g_t(x_t)$. Define the quadratic surrogate

$$\ell_t(x) = \langle g_t, x - x_t\rangle + \frac{\mu}{2}\|x - x_t\|^2, \qquad \ell_t(x_t) = 0. \tag{6}$$

Assumption 1 gives $f_t(x_t) - f_t(u) \le -\ell_t(u)$, so it suffices to control the surrogate regret. Every base learner computes $\nabla\ell_t(z) = g_t + \mu(z - x_t)$ from the same observed gradient. The direct bound $\|\nabla\ell_t(z)\| \le G + \mu D \le 3G$ would give $\frac{9G^2}{2\mu}K_T$, nine times the oracle constant in equation 3.

The factor of nine comes from bounding a quantity that can be expanded. Let $b_{t,i} = \nabla\ell_t(z_{t,i})$. Completing the square gives

$$\|b_{t,i}\|^2 = \|g_t\|^2 + 2\mu\langle g_t, z_{t,i} - x_t\rangle + \mu^2\|z_{t,i} - x_t\|^2 = \|g_t\|^2 + 2\mu\,\ell_t(z_{t,i}). \tag{7}$$

In a step of size $1/(\mu S_{t,i})$, the gradient term is $\|b_{t,i}\|^2/(2\mu S_{t,i})$. Identity equation 7 splits it into the oracle term $\|g_t\|^2/(2\mu S_{t,i}) \le G^2/(2\mu S_{t,i})$ and a residue $\ell_t(z_{t,i})/S_{t,i}$. The exponential-weights potential already tracks this residue, so the update can recycle it instead of bounding $b_{t,i}$ by $3G$.

Algorithm 1 Sharp-BDF: one genuine subgradient per round, zero function values.

Input: budget N; parameters $G, \mu > 0$; feasible $x_\circ \in \mathcal{X}$; projection $\Pi_{\mathcal{X}}$; grid $\{\gamma_i\}_{i=0}^{M}$, priors $\{\pi_i\}$ and activation times $\{b_i\}$ fixed by the protocol

$\alpha_\sharp \leftarrow \mu/(4G^2)$; $\quad q_{0,i} \leftarrow \pi_i,\ S_{0,i} \leftarrow 0,\ z_{b_i,i} \leftarrow x_\circ$ for all i
for $t = 1, \dots, N$ **do**
  $\mathcal{I}_t \leftarrow \{i : b_i \le t\}$ ▷ active experts; $b_0 = 1$, so $\mathcal{I}_t \neq \emptyset$
  for all $i \in \mathcal{I}_t$ **do**
    $S_{t,i} \leftarrow 1 + \gamma_i S_{t-1,i}$; $\quad c_{t,i} \leftarrow \gamma_i S_{t-1,i}/S_{t,i}$ ▷ $= 1 - 1/S_{t,i}$, stable near 0
    $v_{t,i} \leftarrow \pi_i^{1-\gamma_i} q_{t-1,i}^{\gamma_i}$ ▷ each expert forgets on its own clock; never normalize
  end for
  $W_t \leftarrow \sum_{i \in \mathcal{I}_t} v_{t,i} c_{t,i}$
  play $x_t \leftarrow W_t^{-1} \sum_{i \in \mathcal{I}_t} v_{t,i} c_{t,i} z_{t,i}$ if $W_t > 0$, else $x_t \leftarrow x_\circ$ ▷ compensated prediction
  observe one genuine subgradient $g_t \leftarrow g_t(x_t)$ ▷ no loss values, no other gradients
  for all $i \in \mathcal{I}_t$ **do**
    $\ell_{t,i} \leftarrow \langle g_t, z_{t,i} - x_t \rangle + \frac{\mu}{2}\|z_{t,i} - x_t\|^2$ ▷ surrogate; no oracle call
    $q_{t,i} \leftarrow v_{t,i} \exp\big(-\alpha_\sharp c_{t,i} \ell_{t,i}\big)$ ▷ compensated potential
    $z_{t+1,i} \leftarrow \Pi_{\mathcal{X}}\Big(z_{t,i} - \big(g_t + \mu(z_{t,i} - x_t)\big)/(\mu S_{t,i})\Big)$ ▷ base step, curvature-tuned
  end for
  $q_{t,i} \leftarrow q_{t-1,i},\ S_{t,i} \leftarrow S_{t-1,i},\ z_{t+1,i} \leftarrow z_{t,i}$ for $i \notin \mathcal{I}_t$
end for

For expert i active from round b_i, write $n_{t,i} = t - b_i + 1$ for its local round index and set

$$S_{t,i} = S_{n_{t,i}}(\gamma_i), \qquad c_{t,i} = 1 - \frac{1}{S_{t,i}} = \frac{\gamma_i S_{t-1,i}}{1 + \gamma_i S_{t-1,i}} \in [0, 1), \tag{8}$$

Both quantities are computed before the gradient arrives. The potential charges expert i by $c_{t,i}\ell_t(z_{t,i})$, while the master uses weight $v_{t,i}c_{t,i}$. With $r_{t,i} = -\ell_t(z_{t,i})$,

$$r_{t,i} + \frac{\ell_t(z_{t,i})}{S_{t,i}} = \Big(1 - \frac{1}{S_{t,i}}\Big) r_{t,i} = c_{t,i}\, r_{t,i}, \tag{9}$$

the residue cancels and the oracle constant in equation 3 is preserved.

For the compensated update, set $\alpha_\sharp = \mu/(4G^2)$. The compatibility condition $\mu D \le 2G$ gives $|\alpha_\sharp \langle g_t, z - x_t \rangle| \le \frac{1}{2}$, and on that range

$$e^{-\alpha_\sharp c\, \ell_t(z)} \ \le\ 1 - \alpha_\sharp c\, \langle g_t, z - x_t \rangle, \qquad c \in [0, 1],\ z \in \mathcal{X}. \tag{10}$$

The $v_{t,i}c_{t,i}$-weighted prediction cancels the linear terms, so the loss update does not increase total potential mass. Both placements of the compensation are necessary: using $c_{t,i}$ in only one of the prediction weights and potential exponent need not keep the potential bounded.

Algorithm 1 uses one update rule for three protocols; they differ only in π_i and b_i. All variants use the geometric forgetting grid of Section 3, $\delta_i = \min\{2^i/N, 1\}$ and $\gamma_i = 1 - \delta_i$ for $i = 0, \dots, M$. The prefix algorithm $\mathcal{A}_{\mathrm{P}}^\sharp$ activates everything at once, $b_i = 1$, with $\pi_i \propto 2^{-i}(i+1)^{-2}$. The two single-endpoint algorithms instead activate expert i by a countdown,

$$m_i = \min\big\{N, \lceil 8\delta_i^{-1} \log(e\delta_i^{-1}) \rceil\big\}, \qquad b_i = N - m_i + 1, \tag{11}$$

so a short-memory expert runs only while the endpoint's discount can still see it. $\mathcal{A}_{\mathrm{E}}^\sharp$ then takes $\pi_i = 1/(M+1)$, while $\mathcal{A}_{\mathrm{E,loc}}^\sharp$ takes $\pi_i \propto (M - i + 1)^{-3}$, biasing prior mass toward short memories. Since $\delta_0 = 1/N$ forces $m_0 = N$ and $b_0 = 1$, the active set is never empty. Activation uses only the known budget, never a future loss. The grid-to-continuum transfer and the per-round cost are given in Appendix B.

5 Main Results

The appendix gives complete proofs. Each inequality follows from Assumptions 1–3 and elementary calculus. We first establish the known-discount benchmark, then prove the single-gradient upper bounds, derive the lower bounds, and compare the result with the true minimax value.

5.1 What a Known Discount Is Worth

The elementary properties of $S_t(\lambda)$ and $K_T(\lambda)$, the attainability result, and the minimax comparison are stated in Appendix A. They establish the benchmark used in all subsequent additive comparisons.

5.2 One Genuine Gradient, Unchanged Oracle Constant

The supporting exponential, mass, cancellation, and grid-transfer lemmas are collected in Appendix A. They show why the compensated update preserves the oracle constant.

Theorem 1. *Under Assumptions 1–3, Algorithm 1 enjoys the following per-path guarantees, each holding simultaneously for all $u \in \mathcal{X}$ and all $\lambda \in (0,1)$ on a single realized trajectory. Writing $i = i_N(\lambda)$, $C_N = 2 + 8\log(eN)$ and $\mathcal{L}_N = \log\big((M+1)C_N\big)$, the prefix algorithm $\mathcal{A}_{\mathrm{P}}^{\sharp}$ satisfies, simultaneously for all $T \le N$,*

$$R_T^{\lambda}(u) \le \min\Big\{GD\,S_T(\lambda),\ \ B_T^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu}\Big[\frac{1}{2} + 8\log\frac{3}{\pi_i}\Big]\Big\}; \tag{12}$$

the uniform-prior endpoint algorithm $\mathcal{A}_{\mathrm{E}}^{\sharp}$ satisfies

$$R_N^{\lambda}(u) \le \min\Big\{GD\,S_N(\lambda),\ \ B_N^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu}\big(2 + 8\mathcal{L}_N\big)\Big\}, \qquad 2 + 8\mathcal{L}_N \le 50\log(e + \log N); \tag{13}$$

and the scale-local endpoint algorithm $\mathcal{A}_{\mathrm{E,loc}}^{\sharp}$ satisfies

$$R_N^{\lambda}(u) \le \min\Big\{GD\,S_N(\lambda),\ \ B_N^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu}\Big[2 + 8\log\tfrac{57}{2} + 24\log\big(2 + \log_2 H_\lambda\big)\Big]\Big\}. \tag{14}$$

Each variant queries one genuine subgradient and zero function values per round and performs at most $M+1$ Euclidean projections.

The three bounds use the same experts with different activation times and priors. The prefix algorithm works for every T but has a worst-case coding term of $O(\log N)$. The endpoint algorithm gives up early prefixes and reduces this term to $O(\log\log N)$. The scale-local variant is constant when H_λ is bounded and reaches $\log\log N$ only at the largest scales. It is less efficient when $H_\lambda \approx N$. We combine the two expert families in the hybrid algorithm so that the smaller of the two bounds applies to one trajectory.

The hybrid has constant adaptation at both ends of the scale range. The worst-case $\log\log N$ term comes from serving the intermediate scales.

5.3 Lower Bounds: the Price Is Set by the Quantifier

We prove the lower bounds on the normalized family of Section 3, using squared losses $f_t(x) = \frac{\mu}{2}(x - Y_t)^2$, $Y_t = aV_t$, $V_t \in \{\pm 1\}$, and predictions written $x_t = aZ_t$ with $|Z_t| \le 1$. Set $\sigma = \mu a^2 = G^2/(4\mu)$. All bounds permit randomization and full feedback; in fact, the algorithm may see each entire loss function. The proof uses a direct second-moment change of measure.

The block-transfer and static-budget constructions used by the lower bounds are stated in Appendix A.

Theorem 2 (Prefix-uniform protocol: logarithmic price)**.** *If $\log N \ge 16384$, then*

$$\frac{\log N}{16384} \ \le\ \Delta_N^{\mathrm{P},\star} \ \le\ 80\log(eN), \qquad \textit{hence} \qquad \Delta_N^{\mathrm{P},\star} = \Theta(\log N). \tag{15}$$

The single-endpoint protocol needs a different construction, because the prefix argument chooses *where* to evaluate. Fixing the endpoint at N removes that freedom, and the difficulty must instead be hidden among history *scales*.

Theorem 3 (Single-endpoint protocol: doubly logarithmic price)**.** *If $\log\log N \ge 8192$, then*

$$\Delta_N^{\mathrm{E},\star} \ \ge\ \frac{\log\log N}{8192}, \qquad \textit{and with equation 13} \qquad \Delta_N^{\mathrm{E},\star} = \Theta(\log\log N). \tag{16}$$

The discounts used satisfy $K_N(\lambda) \le 2$, and the bound holds against randomized full-feedback algorithms.

The contrast is structural rather than notational. Theorem 2 selects one hard block from among $\Theta(N/\log N)$ time positions and evaluates at its end, which gives $\log N$; Theorem 3 must evaluate at N and can only select among $\Theta(\log N)$ disjoint history scales, which gives $\log\log N$. Both subtract a benchmark with $K \le 2$, so neither is a relabelling of a static logarithmic term. Since both lower bounds are proved against the richer feedback model while Theorem 1 needs only one gradient, the two orders are also exactly the truth in the model we actually use.

Corollary 1 (The separation is not about feedback)**.** *Restrict the infima in equation 4 to algorithms obeying Assumption 2, and write $\Delta_N^{\mathrm{P},\star,G}$ and $\Delta_N^{\mathrm{E},\star,G}$ for the resulting values. Then*

$$\Delta_N^{\mathrm{P},\star,G} = \Theta(\log N), \qquad \Delta_N^{\mathrm{E},\star,G} = \Theta(\log\log N). \tag{17}$$

5.4 Ruling Out Benchmark Slack, and the Shape of the Cost

The appendix also gives the true-oracle comparisons and the cross-scale allocation result.

Let $V_T^F(\lambda)$ be the true known-discount minimax value under full feedback, $v_T(\lambda) = \frac{\mu}{G^2} V_T^F(\lambda)$, and $\Lambda_T^{\mathrm{st}} = \{\lambda \in (0,1) : K_T(\lambda) \le 2\}$ the steady-state set on which a learner told λ needs only constant regret; the minimax comparison in Appendix A gives $v_T \le 1$ there. Define

$$\begin{aligned} \Gamma_N^{\mathrm{E}}(A) &= \sup_{f,\ \lambda\in\Lambda_N^{\mathrm{st}},\ u} \Big[\tfrac{\mu}{G^2}\mathbb{E}_A R_N^{\lambda}(u) - v_N(\lambda)\Big]_+, \\ \Gamma_N^{\mathrm{P}}(A) &= \sup_{f,\ T\le N,\ \lambda\in\Lambda_T^{\mathrm{st}},\ u} \Big[\tfrac{\mu}{G^2}\mathbb{E}_A R_T^{\lambda}(u) - v_T(\lambda)\Big]_+. \end{aligned} \tag{18}$$

The true-oracle comparisons in Appendix A are confined to Λ_T^{st}, where the known-discount problem has constant value. This restriction prevents the growth of the baseline itself from being mistaken for an adaptation cost: subtracting $v_T(\lambda)$ then isolates the price of not knowing λ. Over the full interval, $K_T(\lambda)$ can grow with the effective memory, so the additive results instead subtract the attainable benchmark $B_T^{\mathrm{or}}(\lambda)$; this keeps the oracle's leading constant intact while making the excess comparable across discounts.

6 Numerical Study

6.1 Experimental Setup

The experiments evaluate the achievable side of the theory. The complete setup, instance family, baselines, and supplemental numerical checks are given in Appendix H, including Tables 2–4 and Figure 4. The main text focuses on the two core plots below.

6.2 One Genuine Gradient Preserves the Oracle Constant

Figure 2 evaluates 40 discounts covering $H_\lambda \in [1.05, N]$ at $N = 4096$, $d = 10$, from short-memory to nearly full-history evaluation. Across this range, the hybrid remains close to the known-λ oracle and has the smallest worst-case adaptation gap, 1.71 versus 4.42 for the next best method. BDF-Curv uses the same one-subgradient feedback but incurs a gap of 12.04 and a certificate of 273 versus 78 for the hybrid, illustrating the factor-of-nine cost of the uncompensated surrogate. This comparison isolates the effect of recycling the surrogate surplus rather than changing the feedback model. Fixed-discount baselines work only near their tuned scales and deteriorate away from them. All methods use the same feedback; detailed values appear in Appendix H, Table 2.

6.3 The Price of Adaptivity Splits by Evaluation Quantifier

Figure 3 sweeps N from 2^7 to 2^{14}. The certified adaptation terms follow the scaling predicted by the proofs. A linear fit in $\log N$ has slope 10.1 for the prefix certificate. This is consistent with $8\log 2$ per grid level and the $16\log(M{+}1)$ correction of Lemma 17. The endpoint certificate is linear in $\log\log N$ with slope 24.1, matching the $24\log(2+\log_2 H_\lambda)$ term in equation 14. The realised worst case in panel (b) is one to two orders of magnitude below both certificates and grows more

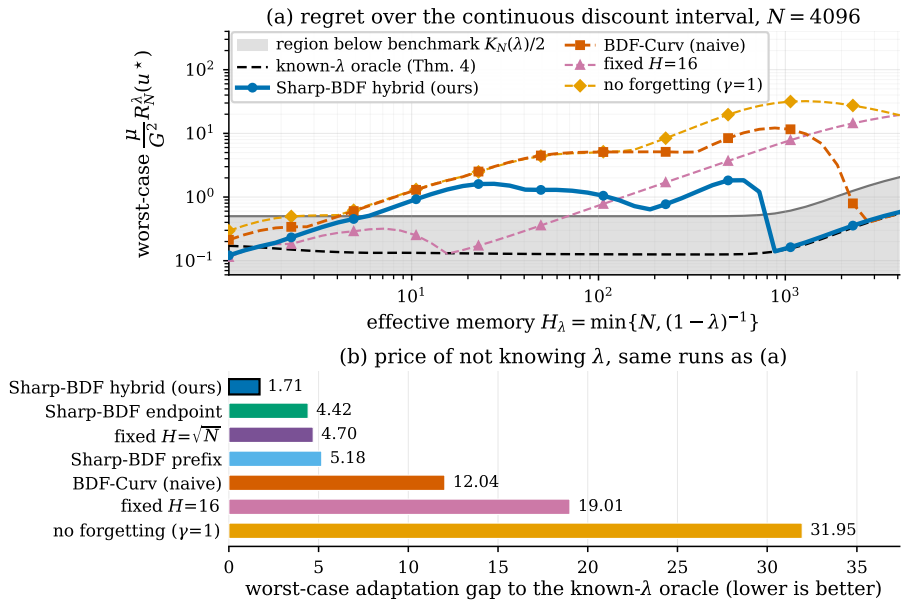


Figure 2: Worst-case results over ten instances at $N = 4096$, $d = 10$. (a) Normalized discounted regret relative to the attainable benchmark and the known-λ oracle. (b) Additive excess from not knowing λ.

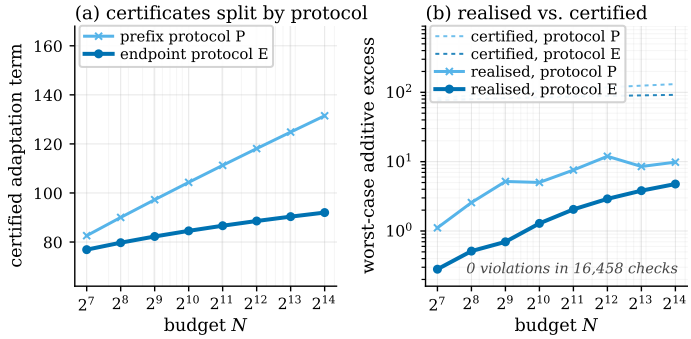


Figure 3: Prefix versus single-endpoint evaluation over ten instances at $d = 5$. (a) Certified adaptation terms. (b) Realised worst-case excess relative to each protocol's certificate.

slowly. It nevertheless has the same ordering: serving every prefix costs two to seven times more than serving one known endpoint at the same N. Because the two protocols use the same instances and the same per-round cost, this ordering is consistent with Corollary 1. The constants in Section 5 are loose at these budgets. The numerical values by horizon are listed in Appendix H, Table 3; the prior comparison and identity checks are also reported there.

7 Conclusion

We studied online convex optimization with an evaluation discount hidden from the learner. Under curvature, the attainable target is $B_T^{\mathrm{or}}(\lambda) = \frac{G^2}{2\mu} K_T(\lambda)$, and Sharp-BDF matches its leading constant for every λ using one genuine subgradient and no function values per round. The cost of not knowing the discount depends on the evaluation rule: the matching excess is $\Theta(\log N)$ when every prefix must be served and $\Theta(\log \log N)$ at a single endpoint. These bounds hold even with randomized full feedback, while a hybrid variant keeps the adaptation cost constant at both scale extremes.

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. The problem setting and assumptions are stated explicitly, and complete proofs of all theoretical results are provided in the appendices. The experiments use a fully specified synthetic instance family. Appendix H records the instance construction, parameter choices, baselines, evaluation protocols, and numerical identity checks. Together with the algorithms and update rules given in the paper, these details allow the theoretical and empirical results to be independently verified and reproduced.

REFERENCES

Dmitry Adamskiy, Wouter M. Koolen, Alexey Chernov, and Vladimir Vovk. A closer look at adaptive regret. In *Algorithmic Learning Theory (ALT)*, volume 7568 of *Lecture Notes in Computer Science*, pp. 290–304. Springer, 2012.

Olivier Bousquet and Manfred K. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3:363–396, 2002.

Yair Carmon and Oliver Hinder. The price of adaptivity in stochastic convex optimization. In *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 772–774, 2024.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Nicolò Cesa-Bianchi, Pierre Gaillard, Gábor Lugosi, and Gilles Stoltz. Mirror descent meets fixed share (and feels no regret). In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012.

Alexey Chernov and Vladimir Vovk. Prediction with expert evaluators' advice. In *Algorithmic Learning Theory (ALT)*, volume 5809 of *Lecture Notes in Computer Science*, pp. 8–22. Springer, 2009.

Alexey Chernov and Fedor Zhdanov. Prediction with expert advice under discounted loss. In *Algorithmic Learning Theory (ALT)*, volume 6331 of *Lecture Notes in Computer Science*, pp. 255–269. Springer, 2010.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pp. 6.1–6.20, 2012.

Amit Daniely, Alon Gonen, and Shai Shalev-Shwartz. Strongly adaptive online learning. In *International Conference on Machine Learning (ICML)*, pp. 1405–1411, 2015.

Yoav Freund, Robert E. Schapire, Yoram Singer, and Manfred K. Warmuth. Using and combining predictors that specialize. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pp. 334–343, 1997.

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

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

Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: Regret bounded by variation in costs. *Machine Learning*, 80(2–3):165–188, 2010.

Elad Hazan and Comandur Seshadhri. Efficient learning algorithms for changing environments. In *International Conference on Machine Learning (ICML)*, pp. 393–400, 2009.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2–3):169–192, 2007.

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

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*, volume 38 of *Proceedings of Machine Learning Research*, pp. 398–406, 2015.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Wouter M. Koolen and Tim Van Erven. Second-order quantile methods for experts and combinatorial games. In *Proceedings of the 28th Annual Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 1155–1175, 2015.

Haipeng Luo and Robert E. Schapire. Achieving all with no parameters: AdaNormalHedge. In *Proceedings of the 28th Annual Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 1286–1304, 2015.

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

Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.

Tim van Erven and Wouter M. Koolen. MetaGrad: Multiple learning rates in online learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 29, 2016.

Yan-Feng Xie, Yu-Jie Zhang, Peng Zhao, and Zhi-Hua Zhou. Dynamic regret via discounted-to-dynamic reduction with applications to curved losses and Adam optimizer. In *International Conference on Machine Learning (ICML)*, 2026.

Wenhao Yang, Yibo Wang, Peng Zhao, and Lijun Zhang. Universal online convex optimization with 1 projection per round. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, pp. 31438–31472, 2024.

Wenhao Yang, Sifan Yang, and Lijun Zhang. Discounted online convex optimization: Uniform regret across a continuous interval. In *International Conference on Learning Representations (ICLR)*, 2026.

Jianjun Yuan and Andrew Lamperski. Trading-off static and dynamic regret in online least-squares and beyond. In *AAAI Conference on Artificial Intelligence*, volume 34, pp. 6712–6719, 2020.

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

Zhiyu Zhang, David Bombara, and Heng Yang. Discounted adaptive online learning: Towards better regularization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 58631–58661, 2024.

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 (NeurIPS)*, volume 33, pp. 12510–12520, 2020.

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

A Deferred Results

Lemma 1. *For $0 < \lambda < 1$, writing $\delta = 1 - \lambda$ and $h = 1/\delta$,*

$$S_t(\lambda) = \frac{1-\lambda^t}{1-\lambda}, \quad S_t(\lambda) = 1 + \lambda S_{t-1}(\lambda), \;\; S_0(\lambda) = 0, \quad \tfrac{1}{2}\min\{t,h\} \le S_t(\lambda) \le \min\{t,h\}. \tag{19}$$

Lemma 2. *For $T \ge 1$ and $0 < \lambda < 1$, with $J_T(\lambda) = \sum_{t \le T} \lambda^{T-t}/t$,*

$$1 \le K_T(\lambda) \le 1 + J_T(\lambda) \le 2 + \log T, \qquad K_T(\lambda) = 1 + \lambda^T\Big(\sum_{t=1}^{T} \frac{1}{S_t(\lambda)} - 1\Big), \tag{20}$$

and consequently $K_T(\lambda) \le 1 + \lambda^T\big((1-\lambda)T + \log T\big)$ as well as $K_T(\lambda) \le 3 + \log H_{T,\lambda}$. For fixed $\lambda < 1$ one has $K_T(\lambda) \to 1$, while for fixed T one has $K_T(\lambda) \to \mathsf{H}_T$ as $\lambda \uparrow 1$.

Lemma 3. *Let ϕ_t be μ-strongly convex on $\mathcal{X}$ with queried subgradients satisfying $\|b_t\| \le L$. For any $\gamma \in [0,1]$, the iteration $\eta_t = 1/\big(\mu S_t(\gamma)\big)$, $z_{t+1} = \Pi_{\mathcal{X}}(z_t - \eta_t b_t)$ satisfies, for every T and every $u \in \mathcal{X}$,*

$$\sum_{t=1}^{T} \gamma^{T-t}\big(\phi_t(z_t) - \phi_t(u)\big) \le \frac{L^2}{2\mu} K_T(\gamma). \tag{21}$$

Theorem 4. *Under Assumptions 1–2, an algorithm handed λ before round 1 satisfies, for all $T \le N$ and all $u \in \mathcal{X}$,*

$$R_T^{\lambda}(u) \;\le\; \min\Big\{GD\, S_T(\lambda),\; B_T^{\mathrm{or}}(\lambda)\Big\}. \tag{22}$$

Theorem 5. *Let $V_T(\lambda)$ be the minimax value on the normalized family of Section 3 when the algorithm knows λ. Then*

$$\frac{G^2}{48\mu} K_T(\lambda) \;\le\; V_T(\lambda) \;\le\; \frac{G^2}{2\mu} K_T(\lambda). \tag{23}$$

Lemma 4. *Let $x, z \in \mathcal{X}$, $\|g\| \le G$, and $\mu D \le 2G$. With $\ell(z) = \langle g, z - x\rangle + \frac{\mu}{2}\|z - x\|^2$ and $\alpha = \mu/(4G^2)$,*

$$e^{-\alpha\ell(z)} \le 1 - \alpha\langle g, z - x\rangle, \qquad \text{and more generally} \qquad e^{-\alpha c\,\ell(z)} \le 1 - \alpha c\,\langle g, z - x\rangle \;\text{ for } c \in [0,1]. \tag{24}$$

Lemma 5. *Fix a round with active set $\mathcal{I}_t$, weights $v_{t,i} > 0$ and coefficients $c_{t,i} \in [0,1]$, and let x_t be the compensated prediction of Algorithm 1. Then the compensated update $q_{t,i} = v_{t,i}\exp(-\alpha_\sharp c_{t,i}\ell_t(z_{t,i}))$ obeys $\sum_{i\in\mathcal{I}_t} q_{t,i} \le \sum_{i\in\mathcal{I}_t} v_{t,i}$. If moreover $v_{t,i} = \pi_i^{1-\gamma_i} q_{t-1,i}^{\gamma_i}$ and inactive experts retain $q_{t,i} = \pi_i$, then*

$$Q_t := \sum_i q_{t,i} \;\le\; 1 + \sum_i \pi_i(1-\gamma_i)\max\{0, t - b_i + 1\}, \tag{25}$$

and for every active expert and every $s \ge b_i$ the log-mass identity

$$\log\frac{q_{s,i}}{\pi_i} = \alpha_\sharp \sum_{t=b_i}^{s} \gamma_i^{s-t} c_{t,i} r_{t,i}, \qquad r_{t,i} = -\ell_t(z_{t,i}), \tag{26}$$

holds exactly.

Lemma 6. *Run the base update and compensation of Algorithm 1 for expert i. Then for every $s \ge b_i$ and every $u \in \mathcal{X}$,*

$$\sum_{t=b_i}^{s} \gamma_i^{s-t}\big[\ell_t(x_t) - \ell_t(u)\big] \le \sum_{t=b_i}^{s} \frac{\gamma_i^{s-t}\|g_t\|^2}{2\mu S_{t,i}} + \sum_{t=b_i}^{s} \gamma_i^{s-t} c_{t,i} r_{t,i} - \frac{\mu}{2} S_{s,i}\|z_{s+1,i} - u\|^2. \tag{27}$$

Consequently, if $Q_s \le C$ then

$$\sum_{t=b_i}^{s} \gamma_i^{s-t}\big[f_t(x_t) - f_t(u)\big] \le \frac{G^2}{2\mu} K_{s-b_i+1}(\gamma_i) + \frac{4G^2}{\mu}\log\frac{C}{\pi_i}. \tag{28}$$

Lemma 7. *For any reals $r_1, \dots, r_T$, with $D_s^\rho = \sum_{t \le s} \rho^{s-t} r_t$ and $D_0^\rho = 0$, and for any $0 \le \gamma \le \lambda < 1$,*

$$D_T^\lambda = D_T^\gamma + (\lambda - \gamma) \sum_{s=1}^{T-1} \lambda^{T-s-1} D_s^\gamma. \tag{29}$$

All coefficients on the right are nonnegative, so per-prefix upper bounds on D_s^γ may be substituted directly.

Lemma 8. *Let $i = i_N(\lambda)$ and $\gamma = \gamma_i$. Then $i \in \{0, \dots, M\}$, $\gamma \le \lambda$, and for every $T \le N$,*

$$1 + (\lambda - \gamma) S_{T-1}(\lambda) \le 2, \qquad 0 \le \frac{1}{S_t(\gamma)} - \frac{1}{S_t(\lambda)} \le \lambda - \gamma, \qquad \sum_{t=1}^{T} \frac{\lambda^{T-t}}{S_t(\gamma)} \le K_T(\lambda) + 1. \tag{30}$$

Theorem 6. *Maintain both expert families of Section 4 at once: the prefix family with priors $\pi_i/2$ and $b_i = 1$, and the endpoint family with priors $\widehat{\pi}_i/2$ and activation equation 11. Run the compensated prediction and update of Algorithm 1 on all $2(M+1)$ experts, querying a single g_t at the common prediction. With $i = i_N(\lambda)$, the resulting trajectory satisfies*

$$R_N^\lambda(u) \le \min\Big\{ GD\, S_N(\lambda),\ \ B_N^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu} \min\Big[\frac{1}{2} + 8 \log \frac{22}{\pi_i},\ \ 2 + 8 \log \frac{22}{\widehat{\pi}_i} \Big] \Big\}. \tag{31}$$

The resulting adaptation term is bounded independently of N when H_λ is bounded and also when N/H_λ is bounded, and is $O(\log \log N)$ in the worst case. The cost is one genuine gradient, zero function values, and at most $2(M+1)$ projections per round.

Lemma 9. *Let $V_t \in \{-1, 1\}$ and let $Z_t \in [-1, 1]$ be any prediction measurable with respect to $V_1, \dots, V_{t-1}$ and an independent seed. Under $\mathbb{P}_0$ the V_t are independent with mean zero; under $\mathbb{P}_1$ they remain independent but have mean $\varepsilon \in (0, 1)$ on a deterministic block I of length m and mean zero elsewhere. Then*

$$\sum_{t \in I} \mathbb{E}_1 Z_t^2 \ \le\ e^{m\varepsilon^2/2} \sqrt{m \sum_{t \in I} \mathbb{E}_0 Z_t^2}. \tag{32}$$

Theorem 7. *Suppose an algorithm satisfies $\mathbb{E}_0 R_N^1(0) \le \sigma B$ in the zero-mean environment, and set $\bar{B} = \max\{1, B\}$ and $Q = N/\bar{B}$. If $Q \ge e^{32}$, define $h = \lfloor \log Q / 16 \rfloor$, $m = 8h$ and $\lambda = 1 - 1/h$. Then there exist a deterministic endpoint $\tau \in [N/2, N]$, an obliviously randomized environment, and the fixed comparator $u = a/2$ with*

$$\mathbb{E}_1 R_\tau^\lambda(a/2) \ \ge\ \frac{G^2}{128\mu} h \ \ge\ \frac{G^2}{4096\mu} \log Q, \qquad \textit{while} \qquad H_{\tau,\lambda} = h, \quad K_\tau(\lambda) \le 2. \tag{33}$$

Lemma 10. *For any integer $h \ge 2$ and $\lambda = 1 - 1/h$ one has $\frac{1}{4} \le \lambda^h \le e^{-1} < \frac{1}{2}$. If in addition $N \ge e$ and $2 \le h \le N/(8 \log N)$, then $K_N(1 - 1/h) \le 2$.*

Lemma 11. *Fix an integer $h \ge 2$ with $2h \le N$ and set $I = \{N - 2h + 1, \dots, N - h\}$, $E_I = \sum_{t \in I} \mathbb{E}_0 Z_t^2$. Let $\mathbb{P}_1$ shift the labels on I to mean $\varepsilon \in (0, 1/2]$. With $\lambda = 1 - 1/h$, $S = S_N(\lambda)$, $W = \sum_{t \in I} \lambda^{N-t}$, $\nu = \varepsilon W / S$, $u = a\nu$ and $b = h\varepsilon^2$, the point $u \in \mathcal{X}$ does not depend on the realized labels and*

$$\frac{1}{\sigma} \mathbb{E}_1 R_N^\lambda(u) \ \ge\ \frac{b}{128} - e^{b/2} \sqrt{b E_I}. \tag{34}$$

Theorem 8. *Consider finitely many experts, expert i active on one interval of length m_i with mass π_i before activation and frozen mass afterwards, and set $\kappa_i = 1 + (1 - \gamma_i) m_i \ge 1$ and $C_\pi = \sum_i \pi_i \kappa_i$, so that $Q_t \le C_\pi$. Given finite nonnegative targets $\mathfrak{a}_i$, a positive prior π with $\alpha_\sharp^{-1} \log(C_\pi / \pi_i) \le \mathfrak{a}_i$ for every i exists if and only if $\sum_i \kappa_i e^{-\alpha_\sharp \mathfrak{a}_i} \le 1$. In particular, within this certificate family the smallest achievable worst-case coding term is exactly*

$$\inf_\pi \max_i \frac{1}{\alpha_\sharp} \log \frac{C_\pi}{\pi_i} = \frac{1}{\alpha_\sharp} \log\Big(\sum_i \kappa_i \Big). \tag{35}$$

Theorem 9. *Fix $N \ge e$ and an integer $4 \le n \le N/(8 \log N)$, put $\lambda_0 = 1 - 1/n$, and choose integer scales $h_1, \dots, h_k \ge 2$ whose annuli $I_j = \{N - 2h_j + 1, \dots, N - h_j\}$ are pairwise disjoint with $2h_j \le n/2$; let $\lambda_j = 1 - 1/h_j$. Suppose an algorithm that does not receive the discount satisfies, for*

$j = 0, \dots, k$ and all losses and comparators, $\frac{\mu}{G^2}\mathbb{E}_A R_N^{\lambda_j}(u) \le \frac{1}{2}K_N(\lambda_j) + A_j$ with $A_j \in [0, \infty)$. If $h_j \ge 4096(1 + A_j)$ for every $j \ge 1$, then necessarily

$$\sum_{j=1}^{k}(1 + A_j)e^{-1024(1+A_j)} \;\le\; 1024(1 + A_0). \tag{36}$$

This holds for arbitrary randomized full-feedback algorithms, not only for potential-based ones.

Corollary 2. *Under the geometry of Theorem 9, if $A_j \le \mathfrak{d}$ for all $j = 0, \dots, k$ and $h_j \ge 4096(1 + \mathfrak{d})$, then $\mathfrak{d} \ge \big(\log k - \log 1024\big)/1024 - 1$.*

Theorem 10. *If $\log\log N \ge 8192$, then for either feedback class $\mathsf{B} \in \{F, G\}$,*

$$\frac{\log\log N}{2048} \;\le\; \inf_{A \in \mathsf{B}} \Gamma_N^{\mathrm{E}}(A) \;\le\; 51\log(e + \log N). \tag{37}$$

Theorem 11. *If $\log N \ge 16384$, then for either feedback class $\mathsf{B} \in \{F, G\}$,*

$$\frac{\log N}{16384} \;\le\; \inf_{A \in \mathsf{B}} \Gamma_N^{\mathrm{P}}(A) \;\le\; 81\log(eN). \tag{38}$$

B Grid Transfer and Computational Cost

Algorithm 1 is analysed at the grid discounts γ_i, for every prefix, and the result is then transported to arbitrary λ by the identity

$$D_T^{\lambda} = D_T^{\gamma} + (\lambda - \gamma)\sum_{s=1}^{T-1}\lambda^{T-s-1}D_s^{\gamma}, \qquad D_s^{\rho} = \sum_{t=1}^{s}\rho^{s-t}r_t, \qquad 0 \le \gamma \le \lambda < 1, \tag{39}$$

which holds for any real sequence $r_1, \dots, r_T$. Its coefficients are nonnegative, so per-prefix upper bounds can be substituted term by term even though discounted regret has no fixed sign. The grid point just below λ is indexed by

$$i_N(\lambda) = \begin{cases} 0, & N(1-\lambda) \le 1, \\ \lceil \log_2\big(N(1-\lambda)\big) \rceil, & N(1-\lambda) > 1, \end{cases} \tag{40}$$

so no ε-net is needed and no probabilistic union over an uncountable parameter set is taken. Every bound is a deterministic algebraic inequality on the realized path and therefore holds for all λ at once.

Per round the algorithm queries one genuine subgradient at x_t and zero function values, performs at most $M + 1 = O(\log N)$ Euclidean projections, and uses $O(d \log N)$ additional vector work and memory; the two-family hybrid introduced later doubles the expert count. One gradient is therefore not one projection: the projection count remains logarithmic, and reducing it is outside the scope of this work.

C Auxiliary Lemmas

This appendix is self-contained. Every inequality is derived from Assumptions 1–3 and elementary calculus; no external lemma is invoked. We collect first the facts used repeatedly later.

Lemma 12. *Under Assumption 1, for all $x, y \in \mathcal{X}$,*

$$|f_t(x) - f_t(y)| \le G\|x - y\|, \qquad \textit{and} \qquad \mu D \le 2G. \tag{41}$$

Proof. Applying equation 5 at x gives

$$f_t(y) - f_t(x) \;\ge\; \langle g_t(x), y - x\rangle + \frac{\mu}{2}\|y - x\|^2 \;\ge\; -G\|y - x\|,$$

where the last inequality uses Cauchy–Schwarz, the bound $\|g_t(x)\| \le G$, and nonnegativity of the quadratic term. Applying equation 5 at y instead gives $f_t(x) - f_t(y) \ge -G\|x - y\|$. Multiplying the second inequality by -1 and combining with the first yields the first claim in equation 41.

For the second claim, add the two strong-convexity inequalities based at x and at y. The function values cancel and we obtain

$$\mu\|x-y\|^2 \;\le\; \langle g_t(x)-g_t(y),x-y\rangle \;\le\; \big(\|g_t(x)\|+\|g_t(y)\|\big)\|x-y\| \;\le\; 2G\|x-y\|.$$

For $x \neq y$, dividing by $\|x-y\|$ gives $\mu\|x-y\| \le 2G$; taking the supremum over pairs gives $\mu D \le 2G$. When $x=y$ no division is needed and both sides vanish. □

Lemma 13. *For any $v \in \mathbb{R}^d$ the Euclidean projection $\Pi_{\mathcal{X}}(v) = \arg\min_{x\in\mathcal{X}} \|x-v\|^2$ exists and is unique. Writing $p = \Pi_{\mathcal{X}}(v)$, for all $u \in \mathcal{X}$,*

$$\|p-u\|^2 \le \|v-u\|^2 - \|v-p\|^2 \le \|v-u\|^2. \tag{42}$$

Consequently, for any $z, u \in \mathcal{X}$, $b \in \mathbb{R}^d$ and $\eta > 0$, the point $z^+ = \Pi_{\mathcal{X}}(z-\eta b)$ satisfies

$$\langle b, z-u\rangle \;\le\; \frac{\|z-u\|^2 - \|z^+-u\|^2}{2\eta} + \frac{\eta}{2}\|b\|^2. \tag{43}$$

Proof. The map $x \mapsto \|x-v\|^2$ is continuous and $\mathcal{X}$ is nonempty and compact, so a minimizer exists. If two distinct minimizers p_1, p_2 existed, convexity would put $(p_1+p_2)/2$ in $\mathcal{X}$, and the parallelogram identity

$$\Big\|\frac{p_1+p_2}{2} - v\Big\|^2 = \frac{1}{2}\|p_1-v\|^2 + \frac{1}{2}\|p_2-v\|^2 - \frac{1}{4}\|p_1-p_2\|^2$$

would make the midpoint strictly better, a contradiction. Hence p is unique.

Fix $u \in \mathcal{X}$. By convexity $p+s(u-p) \in \mathcal{X}$ for $s \in [0,1]$, so $F(s) = \|p+s(u-p)-v\|^2$ is minimized over $[0,1]$ at $s=0$; its right derivative there is therefore nonnegative, i.e. $2\langle p-v, u-p\rangle = F'(0) \ge 0$, which is $\langle v-p, u-p\rangle \le 0$. Expanding

$$\|v-u\|^2 = \|v-p\|^2 + \|p-u\|^2 + 2\langle v-p, p-u\rangle$$

and noting the cross term is nonnegative gives equation 42. Substituting $v = z-\eta b$ into equation 42 and discarding the nonpositive term $-\|v-p\|^2$ yields

$$\|z^+-u\|^2 \le \|z-u\|^2 - 2\eta\langle b, z-u\rangle + \eta^2\|b\|^2,$$

and rearranging and dividing by 2η gives equation 43. □

Lemma 14. *Let ℓ_t be the surrogate equation 6 built from $g_t = g_t(x_t)$. Then for all $u \in \mathcal{X}$,*

$$f_t(x_t) - f_t(u) \;\le\; \ell_t(x_t) - \ell_t(u) = -\ell_t(u), \tag{44}$$

ℓ_t is a μ-strongly convex quadratic, and $\nabla\ell_t(x) = g_t + \mu(x-x_t)$ with $\|\nabla\ell_t(x)\| \le 3G$.

Proof. Taking $x = x_t$ and $y = u$ in equation 5 and rearranging gives

$$f_t(x_t) - f_t(u) \;\le\; -\langle g_t, u-x_t\rangle - \frac{\mu}{2}\|u-x_t\|^2,$$

whose right-hand side is $-\ell_t(u)$; since $\ell_t(x_t) = 0$ this is equation 44. Differentiating equation 6 gives the stated gradient, whose Hessian is μI, so ℓ_t is μ-strongly convex. Finally, by Lemma 12,

$$\|\nabla\ell_t(x)\| \le \|g_t\| + \mu\|x-x_t\| \le G + \mu D \le 3G.$$

□

Lemma 15. *If $e^{-\alpha\phi}$ is concave on a convex set, $p_i \ge 0$ and $\sum_{i=1}^k p_i = 1$, then*

$$\sum_{i=1}^k p_i e^{-\alpha\phi(z_i)} \;\le\; e^{-\alpha\phi(\sum_i p_i z_i)}. \tag{45}$$

Proof. For $k=1$ this is an equality. Assume the claim for $k-1$ points. If $p_k = 1$ all other weights vanish and the claim is an equality. If $p_k < 1$, set $\bar z = \sum_{i<k} \frac{p_i}{1-p_k} z_i$, whose weights are nonnegative and sum to one, so $\bar z$ lies in the set. Two-point concavity and then the inductive hypothesis give

$$e^{-\alpha\phi((1-p_k)\bar z + p_k z_k)} \;\ge\; (1-p_k)e^{-\alpha\phi(\bar z)} + p_k e^{-\alpha\phi(z_k)} \;\ge\; \sum_{i<k} p_i e^{-\alpha\phi(z_i)} + p_k e^{-\alpha\phi(z_k)},$$

completing the induction. □

Lemma 16. *Let $\pi_i > 0$ with $\sum_i \pi_i = 1$, let $q_i > 0$, and let $\gamma_i \in [0,1]$. Then*

$$\sum_i \pi_i^{1-\gamma_i} q_i^{\gamma_i} \;\le\; \sum_i q_i + \sum_i (1-\gamma_i)\pi_i. \tag{46}$$

Proof. Fix $0 < \gamma < 1$ and define $F(r) = \gamma r + (1-\gamma) - r^\gamma$ for $r > 0$. Then $F'(r) = \gamma\big(1 - r^{\gamma-1}\big)$, which is negative for $0 < r < 1$ because $r^{\gamma-1} > 1$ there, and positive for $r > 1$. Hence F attains its global minimum at $r = 1$, where $F(1) = 0$, so $r^\gamma \le \gamma r + 1 - \gamma$; for $\gamma \in \{0, 1\}$ this holds with equality. Taking $r = q_i/\pi_i$ and multiplying by π_i gives

$$\pi_i^{1-\gamma_i} q_i^{\gamma_i} \;\le\; \gamma_i q_i + (1-\gamma_i)\pi_i \;\le\; q_i + (1-\gamma_i)\pi_i,$$

the last inequality using $q_i > 0$ and $\gamma_i \le 1$. Summing over i gives equation 46. □

Lemma 17. *With $M = \lceil \log_2 N \rceil$, $\delta_i = \min\{2^i/N, 1\}$, $\gamma_i = 1 - \delta_i$, $a_i = 2^{-i}(i+1)^{-2}$, $Z_N = \sum_{i=0}^{M} a_i$ and $\pi_i = a_i/Z_N$, one has*

$$1 \le Z_N \le 2, \qquad c_\pi := \sum_i (1-\gamma_i)\pi_i \le \frac{2}{N}, \qquad \log \frac{3}{\pi_i} \le \log 6 + i \log 2 + 2\log(i+1). \tag{47}$$

Proof. Since $a_0 = 1$ and all terms are nonnegative, $Z_N \ge 1$; since $(i+1)^{-2} \le 1$, $Z_N \le \sum_{i \ge 0} 2^{-i} = 2$. Using $\delta_i \le 2^i/N$,

$$c_\pi = \frac{1}{Z_N} \sum_{i=0}^{M} \delta_i \frac{2^{-i}}{(i+1)^2} \le \frac{1}{N Z_N} \sum_{i=0}^{M} \frac{1}{(i+1)^2}.$$

For integers $j \ge 2$ the function $x \mapsto x^{-2}$ is at least j^{-2} on $[j-1, j]$, so $\sum_{j \ge 1} j^{-2} \le 1 + \int_1^\infty x^{-2} dx = 2$; with $Z_N \ge 1$ this gives $c_\pi \le 2/N$. Finally $Z_N \le 2$ yields $3/\pi_i = 3 Z_N 2^i (i+1)^2 \le 6 \cdot 2^i (i+1)^2$, and taking logarithms gives the last claim. □

Lemma 18. *With m_i and b_i as in equation 11, $\widehat{Z}_M = \sum_{j=1}^{M+1} j^{-3}$ and $\widehat{\pi}_i = \big(\widehat{Z}_M (M-i+1)^3\big)^{-1}$, one has $1 \le \widehat{Z}_M \le 3/2$, $\delta_i m_i \le 9(M-i+1)$, $1 + \sum_i \widehat{\pi}_i \delta_i m_i \le 19$, and, for $i = i_N(\lambda)$,*

$$\log \frac{19}{\widehat{\pi}_i} \;\le\; \log \frac{57}{2} + 3\log\big(2 + \log_2 H_\lambda\big). \tag{48}$$

Proof. For an integer $p \ge 2$, monotonicity of x^{-p} on $[1, \infty)$ gives $j^{-p} \le \int_{j-1}^{j} x^{-p} dx$ for $j \ge 2$, hence $\sum_{j=1}^{M+1} j^{-p} \le 1 + \int_1^\infty x^{-p} dx = 1 + \frac{1}{p-1}$. With $p = 3$ this gives $\widehat{Z}_M \le 3/2$, and $\widehat{Z}_M \ge 1$ is immediate; with $p = 2$ it gives $\sum_{j=1}^{M+1} j^{-2} \le 2$.

Put $k_i = M - i$. The grid satisfies $h_i^{\mathrm{grid}} := 1/\delta_i \le 2^{k_i}$: for $i < M$ we have $h_i^{\mathrm{grid}} = N/2^i \le 2^{M-i}$, and for $i = M$ we have $h_M^{\mathrm{grid}} = 1 = 2^0$. Hence $\log(e h_i^{\mathrm{grid}}) \le 1 + k_i \log 2 \le k_i + 1$, and by the ceiling in equation 11,

$$\delta_i m_i \le 8 \log(e h_i^{\mathrm{grid}}) + \delta_i \le 9(k_i + 1) = 9(M - i + 1).$$

Therefore

$$1 + \sum_i \widehat{\pi}_i \delta_i m_i \le 1 + \frac{9}{\widehat{Z}_M} \sum_{j=1}^{M+1} j^{-2} \le 1 + 9 \cdot 2 = 19.$$

For equation 48, take $i = i_N(\lambda)$. If $i = M$ then $k_i + 1 = 1 \le 2 + \log_2 H_\lambda$. If $i < M$ then $2^{M-1} < N$ gives $k_i = M - i < 1 + \log_2(N/2^i) = 1 + \log_2 h_i^{\mathrm{grid}}$. Moreover $h_i^{\mathrm{grid}} \le H_\lambda$: when $N(1-\lambda) \le 1$ we have $i = 0$ and $h_0^{\mathrm{grid}} = N = H_\lambda$, while when $N(1-\lambda) > 1$ the grid gives $\delta_i \ge 1 - \lambda$ and hence $h_i^{\mathrm{grid}} \le (1-\lambda)^{-1} = H_\lambda$. In all cases $k_i + 1 \le 2 + \log_2 H_\lambda$, so using $\widehat{Z}_M \le 3/2$,

$$\log \frac{19}{\widehat{\pi}_i} = \log\big(19 \widehat{Z}_M\big) + 3\log(k_i + 1) \le \log \frac{57}{2} + 3\log\big(2 + \log_2 H_\lambda\big).$$

□

D PROOFS FOR SECTION 5.1

Proof of Lemma 1. Expanding the finite geometric sum gives $\lambda S_t = S_t - 1 + \lambda^t$, hence $(1-\lambda)S_t = 1 - \lambda^t$, which is the closed form. Separating the last term from the rest gives the recursion $S_t = 1 + \lambda S_{t-1}$, and $S_0 = 0$ is the empty-sum convention.

For the upper bound, each of the t terms is at most 1, so $S_t \le t$; and the finite sum is dominated by the infinite geometric series, so $S_t \le h$.

For the lower bound, first suppose $t \le h$. For integers $j \ge 0$ we claim

$$(1-\delta)^j \;\ge\; 1 - j\delta. \tag{49}$$

This follows by induction: equality holds at $j = 0$, and multiplying the j-th instance by the nonnegative number $1-\delta$ gives $(1-\delta)^{j+1} \ge 1 - (j+1)\delta + j\delta^2 \ge 1 - (j+1)\delta$. Summing equation 49 over $j = 0, \dots, t-1$ yields

$$S_t \;\ge\; t - \frac{\delta t(t-1)}{2} \;\ge\; \frac{t}{2},$$

the last inequality using $\delta t \le 1$. Now suppose $t > h$. The function $b \mapsto \log(1-b) + b$ has derivative $-b/(1-b) \le 0$ on $[0, 1)$ and vanishes at 0, so $\log(1-\delta) \le -\delta$ and therefore $\lambda^t \le e^{-\delta t} < e^{-1}$. The closed form then gives $S_t \ge h(1 - e^{-1}) \ge h/2$, using $e > 2$, which follows from the exponential series $e = 1 + 1 + \frac{1}{2!} + \cdots > 2$. Combining the two cases gives the claim. □

Proof of Lemma 2. We first show

$$\frac{1}{S_t(\lambda)} \;\le\; \frac{1}{t} + (1-\lambda). \tag{50}$$

Write $\delta = 1 - \lambda$. Since $\lambda^{-t} = (1+\delta/\lambda)^t \ge 1 + t\delta/\lambda \ge 1 + t\delta$ by the nonnegativity of the higher-order binomial terms, we get $\lambda^t \le (1+t\delta)^{-1}$. Substituting into $S_t = (1-\lambda^t)/\delta$ gives $S_t \ge t/(1+t\delta)$, and taking reciprocals gives equation 50. Hence

$$K_T \;\le\; \sum_{t=1}^{T} \frac{\lambda^{T-t}}{t} + \delta \sum_{t=1}^{T} \lambda^{T-t} = J_T + 1 - \lambda^T \;\le\; J_T + 1.$$

For the lower bound, $K_1 = 1$ and $K_T = \lambda K_{T-1} + 1/S_T$; since $S_T \le 1/\delta$ we have $1/S_T \ge \delta$, so induction from $K_1 = 1$ gives $K_T \ge \lambda + \delta = 1$. Since $\lambda^{T-t} \le 1$,

$$J_T \le \mathsf{H}_T = 1 + \sum_{t=2}^{T} \frac{1}{t} \le 1 + \int_1^T \frac{dx}{x} = 1 + \log T,$$

where the integral comparison uses $1/x \ge 1/t$ on $[t-1, t]$. This proves the chain of inequalities in equation 20.

For the identity, note that termwise

$$\frac{\lambda^{T-t}}{S_t} = \frac{\delta\lambda^{T-t}}{1-\lambda^t} = \delta\lambda^{T-t} + \frac{\delta\lambda^T}{1-\lambda^t} = \delta\lambda^{T-t} + \frac{\lambda^T}{S_t}.$$

Summing from $t = 1$ to T gives $K_T = 1 - \lambda^T + \lambda^T \sum_{t \le T} S_t^{-1}$, which is the identity in equation 20. Combining it with equation 50 and $\mathsf{H}_T \le 1 + \log T$ gives $\sum_{t \le T} S_t^{-1} - 1 \le \delta T + \mathsf{H}_T - 1 \le \delta T + \log T$, hence $K_T \le 1 + \lambda^T\big(\delta T + \log T\big)$.

For the bound by $H_{T,\lambda}$: if $T \le h$ then $H_{T,\lambda} = T$ and the chain already gives the stronger $K_T \le 2 + \log H_{T,\lambda}$. If $T > h$, put $m = \lfloor h \rfloor \ge 1$; for $t \ge m+1$ we have $t > h$, hence $1/t \le 1/h = \delta$, so

$$J_T = \sum_{t=1}^{m} \frac{\lambda^{T-t}}{t} + \sum_{t=m+1}^{T} \frac{\lambda^{T-t}}{t} \le \mathsf{H}_m + \delta \sum_{t=m+1}^{T} \lambda^{T-t} \le 1 + \log m + 1 \le 2 + \log h,$$

and $K_T \le 1 + J_T$ gives $K_T \le 3 + \log H_{T,\lambda}$.

For the limits, fix $\lambda < 1$ so $\delta > 0$. From $0 \le K_T - 1 \le e^{-\delta T}(\delta T + \log T)$ and the series bound $e^{\delta T} \ge (\delta T)^3/6$, together with $\log T \le T$, the right-hand side is at most $6/(\delta^2 T^2) + 6/(\delta^3 T^2)$, which tends to zero; hence $K_T \to 1$. Now fix T and let $\lambda \uparrow 1$: each of the finitely many quantities satisfies $S_t(\lambda) \to t$ and $\lambda^{T-t} \to 1$, so the finite sum converges to $\sum_{t \le T} 1/t = \mathsf{H}_T$. □

Proof of Lemma 3. Write $d_t = \|z_t - u\|$. Strong convexity of ϕ_t gives $\phi_t(z_t) - \phi_t(u) \le \langle b_t, z_t - u\rangle - \frac{\mu}{2} d_t^2$, and applying equation 43 with $\eta_t = 1/(\mu S_t(\gamma))$ yields

$$\phi_t(z_t) - \phi_t(u) \;\le\; \frac{\mu}{2}\big(S_t(\gamma) - 1\big) d_t^2 - \frac{\mu}{2} S_t(\gamma) d_{t+1}^2 + \frac{L^2}{2\mu S_t(\gamma)}. \tag{51}$$

Suppose first $\gamma > 0$. Multiply equation 51 by γ^{T-t} and sum over $t = 1, \dots, T$. Since $S_1 = 1$, the coefficient of d_1^2 vanishes. For $2 \le t \le T$ the net coefficient of d_t^2 is

$$\frac{\mu}{2}\gamma^{T-t}\big(S_t - 1\big) - \frac{\mu}{2}\gamma^{T-(t-1)} S_{t-1} = \frac{\mu}{2}\gamma^{T-t}\big(S_t - 1 - \gamma S_{t-1}\big) = 0$$

by the recursion in equation 19. Only the terminal term $-\mu S_T d_{T+1}^2/2 \le 0$ and the gradient terms remain; discarding the former gives equation 21. The same computation covers $\gamma = 1$.

If $\gamma = 0$, only the weight at $t = T$ is nonzero and $S_T(0) = 1$, so equation 51 directly gives $\phi_T(z_T) - \phi_T(u) \le -\frac{\mu}{2} d_{T+1}^2 + \frac{L^2}{2\mu} \le \frac{L^2}{2\mu}$, and $K_T(0) = 1$ makes this the claim. □

Proof of Theorem 4. Apply Lemma 3 with $\phi_t = f_t$, $\gamma = \lambda$ and $L = G$, which is legitimate by Assumption 1; this gives the bound $\frac{G^2}{2\mu} K_T(\lambda) = B_T^{\mathrm{or}}(\lambda)$. Separately, Lemma 12 gives $f_t(x_t) - f_t(u) \le GD$ for every round; multiplying by the nonnegative weights λ^{T-t} and summing gives $R_T^\lambda(u) \le GD\, S_T(\lambda)$. Both bounds hold for the same algorithm and the same regret, so their minimum is valid. □

We next prove the two lower bounds used in Theorem 5.

Lemma 19. *On the normalized family, let Y_t be independent and uniform on $\{-a, a\}$ and $f_t(x) = \frac{\mu}{2}(x - Y_t)^2$. These losses satisfy Assumption 1, and for every T and λ, even for an algorithm given the entire loss sequence's feedback, the worst-case expected discounted regret is at least $G^2/(16\mu)$.*

Proof. We have $f_t'' = \mu$, and for $x \in [-a, a]$, $|f_t'(x)| = \mu|x - Y_t| \le 2\mu a = G$, so Assumption 1 holds. Write $w_t = \lambda^{T-t}$ and $S = \sum_t w_t$. Given the observations, the offline weighted least-squares minimizer is $u = \big(\sum_t w_t Y_t\big)/S \in [-a, a]$. Expanding the squares,

$$\sum_t w_t (x_t - Y_t)^2 - \sum_t w_t (u - Y_t)^2 = \sum_t w_t\big(x_t^2 - 2x_t Y_t\big) + \frac{\big(\sum_t w_t Y_t\big)^2}{S}.$$

Since x_t depends only on past observations and an independent seed, and Y_t has conditional mean zero, $\mathbb{E}[x_t Y_t] = 0$. Independence and zero mean give $\mathbb{E}\big(\sum_t w_t Y_t\big)^2 = a^2 \sum_t w_t^2$. Hence

$$\mathbb{E} R_T^\lambda(u) = \frac{\mu}{2}\sum_t w_t \mathbb{E} x_t^2 + \frac{\mu a^2}{2}\frac{\sum_t w_t^2}{\sum_t w_t} \;\ge\; \frac{\mu a^2}{2}\cdot\frac{1 + \lambda^T}{1 + \lambda} \;\ge\; \frac{\mu a^2}{4} = \frac{G^2}{16\mu},$$

where the penultimate equality evaluates the ratio of finite geometric sums and the last uses $1 + \lambda^T \ge 1$ and $1 + \lambda \le 2$. A lower bound on average over a randomized environment implies that at least one deterministic observation sequence attains it, since otherwise the weighted average over the finitely many sequences would be strictly smaller. □

Lemma 20. *On the same normalized family, every algorithm satisfies*

$$\sup_{f_{1:T}, u} \mathbb{E}_A R_T^\lambda(u) \;\ge\; \frac{G^2}{24\mu} J_T(\lambda). \tag{52}$$

Proof. Draw P uniformly on $[0, 1]$; given $P = p$, let $Y_t \in \{-a, a\}$ be independent with $\mathbb{P}(Y_t = a \mid P = p) = p$, set $\theta = a(2P - 1) \in [-a, a]$, and keep the squared losses of Lemma 19. Full feedback is equivalent to observing the Y_t; the algorithm never sees P or θ.

We compute the posterior variance directly. For nonnegative integers r, s put $I_{r,s} = \int_0^1 p^r (1-p)^s \, dp$. For $r = 0$, direct integration gives $I_{0,s} = 1/(s + 1)$. For $r \ge 1$, integration by parts gives $I_{r,s} =$

$\frac{r}{s+1} I_{r-1,s+1}$, the boundary terms vanishing because p^r vanishes at 0 and $(1-p)^{s+1}$ vanishes at 1. Iterating,

$$I_{r,s} = \frac{r!\, s!}{(r+s+1)!}. \tag{53}$$

After n observations of which k equal a, Bayes' rule gives posterior density $p^k(1-p)^{n-k}/I_{k,n-k}$, so by equation 53

$$\mathbb{E}[P \mid k, n] = \frac{k+1}{n+2}, \qquad \mathbb{E}[P^2 \mid k, n] = \frac{(k+1)(k+2)}{(n+2)(n+3)},$$

and subtracting,

$$\mathrm{Var}(P \mid k, n) = \frac{(k+1)(n-k+1)}{(n+2)^2(n+3)}. \tag{54}$$

The marginal probability of the count k is $\binom{n}{k} I_{k,n-k} = 1/(n+1)$ for $k = 0, \ldots, n$. We therefore need $\sum_{k=0}^{n}(k+1)(n-k+1)$. Setting $m = n+1$ and $j = k+1$, this is $\sum_{j=1}^{m} j(m+1-j)$. Pairing terms gives $\sum_{j=1}^{m} j = m(m+1)/2$, and summing the identity $j^3 - (j-1)^3 = 3j^2 - 3j + 1$ and substituting the first-power sum gives $\sum_{j=1}^{m} j^2 = m(m+1)(2m+1)/6$. Hence

$$\sum_{j=1}^{m} j(m+1-j) = (m+1)\frac{m(m+1)}{2} - \frac{m(m+1)(2m+1)}{6} = \frac{m(m+1)(m+2)}{6}.$$

Averaging equation 54 over the marginal of k therefore gives $\mathbb{E}\,\mathrm{Var}(P \mid Y_{1:n}) = \frac{1}{6(n+2)}$, and since $\theta = 2aP - a$,

$$\mathbb{E}\,\mathrm{Var}(\theta \mid Y_{1:n}) = \frac{2a^2}{3(n+2)}. \tag{55}$$

Conditioning on the past observations and the independent seed fixes x_t; for any random variable Z, constant b and conditional mean m, expanding the square gives $\mathbb{E}[(b-Z)^2 \mid \mathcal{F}] = (b-m)^2 + \mathrm{Var}(Z \mid \mathcal{F})$. Applying this with $n = t-1$ and equation 55 gives $\mathbb{E}(x_t - \theta)^2 \geq \frac{2a^2}{3(t+1)}$. Given θ and the past, the current label has conditional mean θ, so expanding the two squared losses and taking the conditional expectation over the current observation gives

$$\mathbb{E}\big[f_t(x_t) - f_t(\theta)\big] = \frac{\mu}{2}\mathbb{E}(x_t - \theta)^2 \;\geq\; \frac{\mu a^2}{3(t+1)}.$$

Summing with the discount weights and using $t+1 \leq 2t$,

$$\mathbb{E}R_T^{\lambda}(\theta) \;\geq\; \frac{\mu a^2}{3}\sum_{t=1}^{T}\frac{\lambda^{T-t}}{t+1} \;\geq\; \frac{\mu a^2}{6} J_T(\lambda) = \frac{G^2}{24\mu} J_T(\lambda).$$

Passing from the randomized environment to a worst-case deterministic sequence requires no minimax exchange: for any algorithm the expected regret of each fixed loss sequence and fixed comparator is at most its supremum, so the average under the above joint distribution is also at most that supremum. □

Proof of Theorem 5. The upper bound is Theorem 4. For the lower bound, fix any algorithm. Lemmas 19 and 20 give

$$\sup_{f,u} \mathbb{E}_A R_T^{\lambda}(u) \;\geq\; \max\Big\{\frac{G^2}{16\mu},\; \frac{G^2}{24\mu} J_T\Big\}.$$

Since $\max\{a, b\} \geq (a+b)/2$ for nonnegative a, b, the right-hand side is at least $\frac{G^2}{48\mu}(1 + J_T) \geq \frac{G^2}{48\mu} K_T$, the last inequality by equation 20. As this holds for every algorithm, it survives taking the infimum. □

E PROOFS FOR SECTION 5.2

Proof of Lemma 4. Write $a = \langle g, z - x\rangle$. Cauchy–Schwarz together with $\|g\| \le G$, $\|z - x\| \le D$ and $\mu D \le 2G$ gives

$$|\alpha a| \le \alpha G D \le \frac{\mu}{4G^2} \cdot G \cdot \frac{2G}{\mu} = \frac{1}{2}, \qquad a^2 \le G^2 \|z - x\|^2.$$

We next establish the scalar inequality we need, on the interval where it is used. For $v \in [-\frac{1}{2}, \frac{1}{2}]$ set $F(v) = v + v^2 + \log(1 - v)$, so that

$$F'(v) = 1 + 2v - \frac{1}{1 - v} = \frac{(1 + 2v)(1 - v) - 1}{1 - v} = \frac{v(1 - 2v)}{1 - v}.$$

On $[-\frac{1}{2}, 0]$ the denominator is positive, $v \le 0$ and $1 - 2v \ge 0$, so $F' \le 0$ and hence $F(v) \ge F(0) = 0$. On $[0, \frac{1}{2}]$ all three factors are nonnegative, so $F' \ge 0$ and again $F(v) \ge F(0) = 0$. Therefore

$$-\log(1 - v) \;\le\; v + v^2, \qquad v \in [-\tfrac{1}{2}, \tfrac{1}{2}]. \tag{56}$$

Since $\mu/(2G^2) \ge \alpha$ and $a^2 \le G^2\|z - x\|^2$,

$$\ell(z) = a + \frac{\mu}{2}\|z - x\|^2 \;\ge\; a + \frac{\mu}{2G^2} a^2 \;\ge\; a + \alpha a^2 \;\ge\; -\frac{1}{\alpha}\log(1 - \alpha a),$$

the last equality follows by substituting $v = \alpha a$ into equation 56 and dividing by $\alpha > 0$. Multiplying by $-\alpha$ reverses the inequality, and exponentiating gives the first claim of equation 24. Its right-hand side is at least $\frac{1}{2}$, so no logarithm of a nonpositive number occurs.

For the second claim, fix the real number $y = -\alpha\ell(z)$ and set $H(c) = 1 - c + ce^y - e^{cy}$ for $c \in [0, 1]$. Then $H(0) = H(1) = 0$ and $H''(c) = -y^2 e^{cy} \le 0$. A function with nonpositive second derivative lies above the chord joining its endpoint values: indeed, for $0 \le r \le c \le q \le 1$ we have $H'(r) - H'(q) = -\int_r^q H''(v)\, dv \ge 0$, and integrating this over the rectangle $[0, c] \times [c, 1]$ gives $(1 - c)\int_0^c H'(r)dr \;\ge\; c\int_c^1 H'(q)dq$, i.e. by the fundamental theorem of calculus $H(c) \ge (1 - c)H(0) + cH(1) = 0$. Hence

$$e^{-\alpha c\,\ell(z)} \;\le\; 1 - c + ce^{-\alpha\ell(z)} \;\le\; 1 - c + c\big(1 - \alpha a\big) = 1 - \alpha c a,$$

which is the second claim. □

Proof of Lemma 5. Suppose first $W_t > 0$. Then x_t is a convex combination of feasible points, so $x_t \in \mathcal{X}$, and by construction

$$\sum_{i \in \mathcal{I}_t} v_{t,i} c_{t,i}\big(z_{t,i} - x_t\big) = 0. \tag{57}$$

Multiplying the second inequality of equation 24 by $v_{t,i} > 0$ and summing over active experts gives

$$\sum_{i \in \mathcal{I}_t} q_{t,i} \;\le\; \sum_{i \in \mathcal{I}_t} v_{t,i} - \alpha_\sharp \Big\langle g_t,\, \sum_{i \in \mathcal{I}_t} v_{t,i} c_{t,i}(z_{t,i} - x_t)\Big\rangle = \sum_{i \in \mathcal{I}_t} v_{t,i},$$

by equation 57. Now suppose $W_t = 0$. All $v_{t,i}$ are strictly positive and all $c_{t,i}$ are nonnegative, so every active $c_{t,i}$ must vanish; the compensated update then gives $q_{t,i} = v_{t,i}$ and the claim holds with equality.

For equation 25, Lemma 16 applied expertwise gives $v_{t,i} \le q_{t-1,i} + (1 - \gamma_i)\pi_i$ for active experts, while inactive experts are unchanged. Combining with the previous paragraph,

$$Q_t \;\le\; Q_{t-1} + \sum_{i \in \mathcal{I}_t} (1 - \gamma_i)\pi_i.$$

Iterating from $Q_0 = \sum_i \pi_i = 1$ and noting that expert i has been active exactly $\max\{0, t - b_i + 1\}$ times among the first t rounds gives equation 25.

For equation 26, take logarithms in the compensated update and subtract $\log \pi_i$:

$$\log \frac{q_{t,i}}{\pi_i} = \gamma_i \log \frac{q_{t-1,i}}{\pi_i} + \alpha_\sharp c_{t,i} r_{t,i}.$$

At $t = b_i - 1$ the left-hand side is zero. Unrolling the recursion one round at a time gives the current term first, then multiplies the accumulated expression by γ_i before adding each new term. This yields equation 26 at round s. The computation is valid for $\gamma_i = 0$ as well, using the convention $0^0 = 1$ for the current round's weight. □

Proof of Lemma 6. Write $\xi = z_{t,i} - x_t$ and $b_{t,i} = g_t + \mu\xi$. Expanding the square gives the identity

$$\|b_{t,i}\|^2 = \|g_t\|^2 + 2\mu\langle g_t, \xi\rangle + \mu^2\|\xi\|^2 = \|g_t\|^2 + 2\mu\,\ell_t(z_{t,i}), \tag{58}$$

which is exact and requires no sign condition on $\ell_t(z_{t,i})$.

Put $D_t = \|z_{t,i} - u\|^2$. Since ℓ_t is a μ-strongly convex quadratic with gradient $b_{t,i}$ at $z_{t,i}$,

$$\ell_t(z_{t,i}) - \ell_t(u) \;\le\; \langle b_{t,i}, z_{t,i} - u\rangle - \frac{\mu}{2}D_t,$$

and applying Lemma 13 with learning rate $1/(\mu S_{t,i})$ gives

$$\ell_t(z_{t,i}) - \ell_t(u) \;\le\; \frac{\mu}{2}\big(S_{t,i} - 1\big)D_t - \frac{\mu}{2}S_{t,i}D_{t+1} + \frac{\|b_{t,i}\|^2}{2\mu S_{t,i}}.$$

Substituting equation 58 into the last term splits it as

$$\frac{\|b_{t,i}\|^2}{2\mu S_{t,i}} = \frac{\|g_t\|^2}{2\mu S_{t,i}} + \frac{\ell_t(z_{t,i})}{S_{t,i}}.$$

Because $\ell_t(x_t) = 0$ and $r_{t,i} = -\ell_t(z_{t,i})$, adding $r_{t,i}$ to the expert's regret produces exactly

$$r_{t,i} + \frac{\ell_t(z_{t,i})}{S_{t,i}} = \Big(1 - \frac{1}{S_{t,i}}\Big)r_{t,i} = c_{t,i}r_{t,i},$$

so the single-round bound reads

$$\ell_t(x_t) - \ell_t(u) \;\le\; \frac{\mu}{2}\big(S_{t,i} - 1\big)D_t - \frac{\mu}{2}S_{t,i}D_{t+1} + \frac{\|g_t\|^2}{2\mu S_{t,i}} + c_{t,i}r_{t,i}. \tag{59}$$

For $\gamma_i > 0$, multiply equation 59 by γ_i^{s-t} and sum from b_i to s. In the first active round $S_{b_i,i} = 1$, so the coefficient of the initial distance vanishes. For any interior round $t > b_i$ the net coefficient of D_t is

$$\frac{\mu}{2}\gamma_i^{s-t}\big(S_{t,i} - 1 - \gamma_i S_{t-1,i}\big) = 0$$

because $S_{t,i} = 1 + \gamma_i S_{t-1,i}$. Only $-\mu S_{s,i}D_{s+1}/2$ survives, which is equation 27. For $\gamma_i = 0$ every round has $S_{t,i} = 1$ and $c_{t,i} = 0$, the discounted sum retains only round s, and equation 59 is already equation 27; no division by a vanishing discount occurs.

Finally, equation 44 bounds the true regret by the surrogate regret in each round, and the discount weights are nonnegative, so the inequality is preserved after summation. Discarding the nonpositive terminal term in equation 27, bounding the first sum using $\|g_t\| \le G$ and Lemma 3's benchmark bookkeeping, namely $\sum_{t=b_i}^{s}\gamma_i^{s-t}/S_{t,i} = K_{s-b_i+1}(\gamma_i)$ after relabelling local time, and bounding the second sum via equation 26 and $q_{s,i} \le Q_s \le C$ gives

$$\sum_{t=b_i}^{s}\gamma_i^{s-t}\big[f_t(x_t) - f_t(u)\big] \;\le\; \frac{G^2}{2\mu}K_{s-b_i+1}(\gamma_i) + \alpha_\sharp^{-1}\log\frac{C}{\pi_i},$$

and $\alpha_\sharp^{-1} = 4G^2/\mu$ gives equation 28. □

Proof of Lemma 7. The definition of a discounted sum gives $r_t = D_t^\gamma - \gamma D_{t-1}^\gamma$. Substituting this into the target sum,

$$\begin{aligned} D_T^\lambda &= \sum_{t=1}^{T}\lambda^{T-t}D_t^\gamma - \gamma\sum_{t=1}^{T}\lambda^{T-t}D_{t-1}^\gamma \\ &= D_T^\gamma + \sum_{s=1}^{T-1}\big(\lambda^{T-s} - \gamma\lambda^{T-s-1}\big)D_s^\gamma \\ &= D_T^\gamma + (\lambda - \gamma)\sum_{s=1}^{T-1}\lambda^{T-s-1}D_s^\gamma, \end{aligned}$$

where the term involving D_0^γ vanishes. Since $\lambda \geq \gamma$ the coefficients are nonnegative, so substituting upper bounds preserves the inequality. Nothing in the derivation requires r_t or D_s^γ to have a sign. □

Proof of Lemma 8. Write $\delta = 1 - \lambda$. Since $0 < N\delta < N$, the index in equation 40 never exceeds $\lceil \log_2 N \rceil = M$.

Suppose first $N\delta \geq 1$. Rounding up in binary and truncating at 1 give $\delta \leq \delta_i \leq 2\delta$, hence $\gamma = 1 - \delta_i \leq 1 - \delta = \lambda$ and $0 \leq \lambda - \gamma = \delta_i - \delta \leq \delta$. Since $S_{T-1}(\lambda) \leq 1/\delta$ by equation 19, the first inequality of equation 30 follows. For the comparison of weights, the closed forms give

$$\frac{S_t(\gamma)}{S_t(\lambda)} = \frac{\delta}{\delta_i} \cdot \frac{1-\gamma^t}{1-\lambda^t} \;\geq\; \frac{\delta}{\delta_i} \;\geq\; \frac{1}{2},$$

using $\gamma \leq \lambda$, which gives $1 - \gamma^t \geq 1 - \lambda^t$; the same formula remains valid at $\gamma = 0$.

Now suppose $N\delta < 1$. Then $i = 0$ and $\gamma = 1 - 1/N$, so $\gamma < \lambda$ and $\lambda - \gamma < 1/N$; since $S_{T-1}(\lambda) \leq T - 1 \leq N$, the first inequality of equation 30 again holds. Moreover, by equation 49,

$$S_t(\gamma) = \sum_{j=0}^{t-1} \Big(1 - \frac{1}{N}\Big)^j \;\geq\; t - \frac{t(t-1)}{2N} \;\geq\; \frac{t}{2},$$

while $S_t(\lambda) \leq t$. This argument includes $N = 1$, where only $t = 1$ occurs.

For the Lipschitz bound, write $S_t(\rho) = \sum_{j=0}^{t-1} \rho^j$ and $S_t'(\rho) = \sum_{j=0}^{t-2} (j+1)\rho^j$, the latter being an empty sum equal to zero when $t = 1$. Counting coefficients in the square of a polynomial,

$$S_t(\rho)^2 = \sum_{j=0}^{t-1} (j+1)\rho^j + \sum_{j=t}^{2t-2} (2t-1-j)\rho^j \;:$$

for $0 \leq j \leq t-1$ the pairs (a, b) with $a + b = j$ and $0 \leq a, b \leq t-1$ are $(0, j), \ldots, (j, 0)$, giving coefficient $j + 1$; for $t \leq j \leq 2t - 2$ the first coordinate ranges over $j - (t-1), \ldots, t-1$, giving $2t - 1 - j$ integers. Subtracting termwise,

$$S_t(\rho)^2 - S_t'(\rho) = t\rho^{t-1} + \sum_{j=t}^{2t-2} (2t-1-j)\rho^j \;\geq\; 0 \qquad \text{for all } \rho \in [0, 1].$$

Since $S_t \geq 1$ and $S_t' \geq 0$, it follows that

$$-1 \;\leq\; \frac{d}{d\rho} \frac{1}{S_t(\rho)} = -\frac{S_t'(\rho)}{S_t(\rho)^2} \;\leq\; 0,$$

and integrating over $[\gamma, \lambda]$ gives the second inequality of equation 30. Consequently

$$\sum_{t=1}^{T} \frac{\lambda^{T-t}}{S_t(\gamma)} \;\leq\; K_T(\lambda) + (\lambda - \gamma) S_T(\lambda).$$

If $N(1-\lambda) \geq 1$ we showed $\lambda - \gamma \leq 1 - \lambda$, and $S_T(\lambda) \leq (1-\lambda)^{-1}$, so the extra term is at most 1. If $N(1-\lambda) < 1$ then $\lambda - \gamma < 1/N$ while $S_T(\lambda) \leq T \leq N$, so the extra term is again at most 1. This gives the third inequality of equation 30. □

Proof of Theorem 1. By equation 25 with $b_i = 1$ for all i, the per-round mass injection is $c_\pi = \sum_i (1 - \gamma_i)\pi_i$, so $Q_t \leq 1 + tc_\pi \leq 1 + 2t/N \leq 3$ for $t \leq N$ by Lemma 17. Lemma 6 with $C = 3$ therefore gives, for every expert i and every prefix $s \leq N$,

$$R_s^{\gamma_i}(u) \;\leq\; \frac{G^2}{\mu} \Big[\frac{1}{2} K_s(\gamma_i) + 4 \log \frac{3}{\pi_i} \Big]. \tag{60}$$

Fix T, λ, u and take $i = i_N(\lambda)$, $\gamma = \gamma_i$. Applying Lemma 7 to the possibly signed sequence $r_t = f_t(x_t) - f_t(u)$ gives

$$R_T^\lambda = R_T^\gamma + (\lambda - \gamma) \sum_{s=1}^{T-1} \lambda^{T-s-1} R_s^\gamma,$$

whose coefficients are nonnegative because $\gamma \le \lambda$, so equation 60 may be substituted for each prefix. The benchmark part of the resulting positive combination is exactly

$$\frac{G^2}{2\mu}\Big[K_T(\gamma) + (\lambda-\gamma)\sum_{s=1}^{T-1}\lambda^{T-s-1}K_s(\gamma)\Big] = \frac{G^2}{2\mu}\sum_{t=1}^{T}\frac{\lambda^{T-t}}{S_t(\gamma)} \;\le\; \frac{G^2}{2\mu}\big(K_T(\lambda)+1\big),$$

where the equality is Lemma 7 applied to $r_t = 1/S_t(\gamma)$ (for which $D_s^\gamma = K_s(\gamma)$) and the inequality is the third part of equation 30. The coding part carries the coefficient sum $1+(\lambda-\gamma)S_{T-1}(\lambda) \le 2$, so it contributes at most $8(G^2/\mu)\log(3/\pi_i)$. Recognizing $\frac{G^2}{2\mu}K_T(\lambda)$ as $B_T^{\mathrm{or}}(\lambda)$ and intersecting with the Lipschitz bound $R_T^\lambda(u) \le GD\,S_T(\lambda)$ of Lemma 12 gives equation 12.

By equation 25, expert i is active exactly m_i rounds in total, so $Q_s \le 1 + \sum_i \bar\pi_i\delta_i m_i$ with $\bar\pi_i = 1/(M+1)$. By the ceiling in equation 11, $m_i \le 8\delta_i^{-1}\log(e\delta_i^{-1}) + 1$, hence

$$\delta_i m_i \le 8\log(e\delta_i^{-1}) + \delta_i \le 8\log(eN) + 1,$$

and since the priors sum to one, $Q_s \le 2 + 8\log(eN) = C_N$. Lemma 6 with $C = C_N$ and $\pi_i = \bar\pi_i$ therefore gives, for each local prefix length $r \in \{1,\dots,m_i\}$,

$$\sum_{t=b_i}^{b_i+r-1}\gamma_i^{b_i+r-1-t}\big[f_t(x_t) - f_t(u)\big] \;\le\; \frac{G^2}{\mu}\Big[\frac{1}{2}K_r(\gamma_i) + 4\mathcal{L}_N\Big]. \tag{61}$$

Fix λ, take $i = i_N(\lambda)$ and write $m = m_i$, $b = b_i$. Relabelling local time as $1,\dots,m$ and applying the positive-coefficient transfer used for the prefix case, which requires only these local-prefix bounds, gives

$$\sum_{t=b}^{N}\lambda^{N-t}\big[f_t(x_t) - f_t(u)\big] \;\le\; \frac{G^2}{\mu}\Big[\frac{1}{2}K_m(\lambda) + \frac{1}{2} + 8\mathcal{L}_N\Big]. \tag{62}$$

If $m = N$ this covers the whole history and the constant $\frac{1}{2}$ is below the 2 of equation 13. If $m < N$, Lemma 21 below bounds the discarded prefix's contribution by G^2/μ and gives $K_m(\lambda) \le K_N(\lambda) + 1$; adding these, the constants total $\frac{1}{2} + \frac{1}{2} + 1 = 2$. Intersecting with the Lipschitz bound gives equation 13.

For the numerical form, set $a_N = 1 + \log N \ge 1$. Since $\log 2 = \int_1^2 x^{-1}dx \ge \frac{1}{2}$, the ceiling bound gives $M + 1 \le 2 + \log N/\log 2 \le 2a_N$ and $C_N = 2 + 8a_N \le 10a_N$, whence $\mathcal{L}_N \le \log 20 + 2\log a_N$. Writing $b_N = \log(e + \log N) \ge 1$ and using $a_N \le e + \log N$, so $\log a_N \le b_N$, together with $\log 20 < 4$, verified by $e^4 \ge 1 + 4 + \frac{4^2}{2} + \frac{4^3}{6} > 20$, which gives

$$2 + 8\mathcal{L}_N \le 2 + 8\log 20 + 16b_N \le 34 + 16b_N \le 50b_N,$$

which is the stated numerical bound.

With $\widehat\pi_i$ in place of $\bar\pi_i$, equation 25 and Lemma 18 give $Q_s \le 1 + \sum_i \widehat\pi_i\delta_i m_i \le 19$. Hence the coding term in the per-expert guarantee equation 28 is $4(G^2/\mu)\log(19/\widehat\pi_i)$. The local-prefix transfer multiplies it by at most 2, and the history truncation together with the comparison of local initialization adds the same constant $2G^2/\mu$ as above, so

$$R_N^\lambda(u) \;\le\; B_N^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu}\Big[2 + 8\log\frac{19}{\widehat\pi_i}\Big].$$

Substituting equation 48 and intersecting with the Lipschitz bound gives equation 14.

The quantities $S_{t,i}$, $c_{t,i}$, $v_{t,i}$ and the prediction depend only on past state. After the single g_t is received, each $\ell_t(z_{t,i})$ and each $b_{t,i}$ is computed by vector algebra without calling the loss oracle. Each active expert performs one projection and there are at most $M+1$ experts; state and vector work are $O(d(M+1))$. Every bound above is an algebraic inequality on the given loss path, so all of them hold simultaneously for the continuum of discounts without any probabilistic union. □

The endpoint versions used two facts about truncation, which we isolate.

Lemma 21. *Fix $\lambda \in (0,1)$, let $i = i_N(\lambda)$ and $m = m_i$ as in equation 11. If $m < N$ then*

$$GD\,\lambda^m S_{N-m}(\lambda) \;\le\; \frac{G^2}{\mu} \qquad \textit{and} \qquad K_m(\lambda) \;\le\; 2 \;\le\; K_N(\lambda) + 1. \tag{63}$$

Proof. Since expert 0 activates in round 1, the hypothesis $m < N$ rules out $i = 0$ and therefore forces $N(1-\lambda) > 1$. Put $h = (1-\lambda)^{-1}$, $v = 1/\delta_i$ and $y = \log(ev) \ge 1$. The two-fold covering property of the grid gives $v \le h \le 2v$, and since the ceiling in equation 11 is not truncated by N,

$$8vy \;\le\; m \;\le\; 8vy + 1 \;\le\; 9vy.$$

Set $z = m/h$, so that $z \ge 8vy/(2v) = 4y$. Using $\log y \le y - 1$ for $y \ge 1$, which follows from $\int_1^y u^{-1}du \le \int_1^y 1\,du$, and $\log 9 < 3$,

$$\log m \;\le\; \log 9 + \log v + \log y \;\le\; \log 9 + 2y - 2 \;\le\; 3y \;\le\; z, \tag{64}$$

where the penultimate inequality needs only $\log 9 \le y + 2$, whose right-hand side is at least 3.

For the discarded history, $S_{N-m}(\lambda) \le h$ and $\lambda^m \le e^{-m/h} = e^{-z}$, so

$$\lambda^m S_{N-m}(\lambda) \;\le\; he^{-z} \;\le\; 2v\,e^{-4\log(ev)} = 2e^{-4}v^{-3} \;\le\; 2e^{-4},$$

and Lemma 12 gives $GD \le 2G^2/\mu$, hence $GD\lambda^m S_{N-m}(\lambda) \le 4e^{-4}G^2/\mu \le G^2/\mu$ because $e^4 \ge 1 + 4 > 4$.

For the benchmark, Lemma 2 and equation 64 give

$$K_m(\lambda) \;\le\; 1 + e^{-z}\big(z + \log m\big) \;\le\; 1 + 2ze^{-z} \;\le\; 1 + \frac{2}{e} < 2,$$

since $z \mapsto ze^{-z}$ has derivative $e^{-z}(1-z)$ and hence maximum $1/e$ at $z = 1$, and $e > 2$. Finally $K_N(\lambda) \ge 1$ by equation 20. □

Proof of Theorem 6. Lemmas 5 and 6 apply verbatim to each expert: neither requires all experts to activate simultaneously, nor that the priors come from any particular formula. The prefix family injects total mass $\sum_i \frac{\pi_i}{2}\delta_i N \le 1$ by Lemma 17, and the endpoint family injects $\sum_i \frac{\widehat{\pi}_i}{2}\delta_i m_i \le 9$ by Lemma 18. The initial total mass is 1, so every prefix satisfies $Q_t \le 1 + 1 + 9 = 11$.

For a prefix-family expert the prior is $\pi_i/2$, so equation 28 with $C = 11$ gives, for all $s \le N$,

$$R_s^{\gamma_i}(u) \;\le\; \frac{G^2}{2\mu}K_s(\gamma_i) + \frac{4G^2}{\mu}\log\frac{22}{\pi_i}.$$

The nonnegative-coefficient transfer of Lemma 7 increases the benchmark part by at most $G^2/(2\mu)$ via the third inequality of equation 30 and multiplies the coding term by at most 2, giving

$$R_N^{\lambda}(u) \;\le\; B_N^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu}\Big(\frac{1}{2} + 8\log\frac{22}{\pi_i}\Big).$$

For an endpoint-family expert the prior is $\widehat{\pi}_i/2$, so equation 28 gives each local-prefix bound with coding term $(4G^2/\mu)\log(22/\widehat{\pi}_i)$. Transferring in local time again adds $G^2/(2\mu)$ to the benchmark and doubles the coding term. If $m_i = N$ this already controls the entire history; if $m_i < N$, Lemma 21 contributes a local-initialization cost $G^2/(2\mu)$ and a truncation cost G^2/μ. In both cases

$$R_N^{\lambda}(u) \;\le\; B_N^{\mathrm{or}}(\lambda) + \frac{G^2}{\mu}\Big(2 + 8\log\frac{22}{\widehat{\pi}_i}\Big).$$

Both displayed inequalities bound the *same* $R_N^{\lambda}(u)$, produced by the one common prediction sequence, so the smaller may be taken; intersecting with the Lipschitz bound gives equation 31. No choice of which algorithm to run was made on the basis of the unknown λ.

For the scale interpretation, write $H = H_\lambda$ and $\rho = N/H \ge 1$. If $H = N$ then $i = 0$; if $H < N$ then $i = \lceil \log_2 \rho \rceil$. In either case $i \le 1 + \log_2 \rho$ and $i + 1 \le 2 + \log_2 \rho$, so $Z_N \le 2$ gives

$$\log\frac{22}{\pi_i} \;\le\; \log 44 + i\log 2 + 2\log(i+1) \;\le\; \log 88 + \log\rho + 2\log\big(2 + \log_2\rho\big),$$

while Lemma 18 gives $\log(22/\widehat{\pi}_i) \le \log 33 + 3\log(2 + \log_2 H)$. The first expression is bounded independently of N when N/H is bounded, the second when H is bounded, and since $H \le N$ the second is always $O(\log\log N)$. The complexity claim follows from the expert count $2(M+1)$, all experts sharing the single gradient at the common prediction. □

F Proofs for Section 5.3

Proof of Lemma 9. For $t \in I$ the prediction Z_t depends only on $V_1, \dots, V_{t-1}$ and an independent seed. A single observation $v \in \{-1, 1\}$ has probability $(1 + \varepsilon v)/2$ under the shifted mean and $1/2$ under the null, so the ratio is $1 + \varepsilon v$, and the likelihood ratio of the history visible before predicting Z_t is

$$L_{t-1} = \prod_{r=s}^{t-1} (1 + \varepsilon V_r),$$

an empty product being 1, where $I = \{s, \dots, s + m - 1\}$. The seed's distribution is the same under both environments and independent of the labels, so it contributes no factor. Summing termwise over the finitely many discrete histories gives the change-of-measure identity $\mathbb{E}_1 Z_t^2 = \mathbb{E}_0[L_{t-1} Z_t^2]$; if the seed is not discrete, apply this identity for each fixed seed value and then integrate over the seed.

By independence under $\mathbb{P}_0$,

$$\mathbb{E}_0 L_{t-1}^2 = \prod_{r=s}^{t-1} \mathbb{E}_0 (1 + \varepsilon V_r)^2 = (1 + \varepsilon^2)^{t-s} \;\le\; e^{m\varepsilon^2},$$

the last inequality from $1 + b \le e^b$, which follows for $b \ge 0$ from the exponential series. Cauchy–Schwarz and $|Z_t| \le 1$, hence $Z_t^4 \le Z_t^2$, give

$$\mathbb{E}_1 Z_t^2 \;\le\; \sqrt{\mathbb{E}_0 L_{t-1}^2}\sqrt{\mathbb{E}_0 Z_t^4} \;\le\; e^{m\varepsilon^2/2}\sqrt{\mathbb{E}_0 Z_t^2}.$$

Summing over $t \in I$ and applying Cauchy–Schwarz once more, to the vectors $(1)_{t \in I}$ and $\big(\sqrt{\mathbb{E}_0 Z_t^2}\big)_{t \in I}$, gives equation 32. □

Proof of Theorem 7. In the zero-mean environment the current label is conditionally independent of Z_t and has mean zero, so $\mathbb{E}_0[f_t(aZ_t) - f_t(0)] = \frac{a}{2}\mathbb{E}_0 Z_t^2$. The hypothesis therefore gives

$$\sum_{t=1}^{N} \mathbb{E}_0 Z_t^2 \;\le\; 2B. \tag{65}$$

The assumption $Q \ge e^{32}$ gives $h \ge 2$, and $\bar{B} \ge 1$ gives $\log Q \le \log N$, so $h \le \log N/16$ and $m \le \log N/2$. We also need $N \ge 4m$: writing $L = \log N \ge 32$, the exponential series gives $e^L \ge L^2/2 \ge 2L$, hence $N \ge 2 \log N \ge 4m$. Partition $\{\lfloor N/2 \rfloor + 1, \dots, N\}$ from left to right into disjoint blocks of length m, discarding the incomplete tail; the number k of complete blocks satisfies $k \ge N/(2m) - 1 \ge N/(4m)$. If every complete block had null energy exceeding $2B/k$, the total would exceed $2B$, contradicting equation 65. Hence some deterministic block $I = \{s, \dots, s + m - 1\}$ satisfies

$$\sum_{t \in I} \mathbb{E}_0 Z_t^2 \;\le\; \frac{2B}{k} \;\le\; \frac{8Bm}{N}. \tag{66}$$

This block is selected from the algorithm's expected behaviour under the null environment and does not depend on the labels eventually drawn. Put $\tau = s + m - 1 \in [N/2, N]$.

Let $\mathbb{P}_1$ give the labels in I mean $\varepsilon = \frac{1}{2}$ and leave all others at mean zero, independently. Lemma 9 with equation 66 gives

$$\sum_{t \in I} \mathbb{E}_1 Z_t^2 \;\le\; e^{m/8}\sqrt{m \cdot \frac{8Bm}{N}} = m\, e^{m/8}\sqrt{\frac{8B}{N}} \;\le\; 8\sqrt{8}\, h\, Q^{-1/2} e^h \;\le\; 8\sqrt{8}\, h\, Q^{-7/16},$$

using $m = 8h$ and $B/N \le 1/Q$ in the third inequality and $h \le \log Q/16$ in the fourth. Since $Q \ge e^{32}$,

$$8\sqrt{8}\, Q^{-7/16} \;\le\; 8\sqrt{8}\, e^{-14} < \frac{1}{128},$$

which one checks numerically from $e > 2$: $e^{14} > 2^{14} = 16384$ while $1024\sqrt{8} < 3072 < 16384$. Therefore

$$\sum_{t \in I} \mathbb{E}_1 Z_t^2 \;\le\; \frac{h}{128}. \tag{67}$$

Fix the comparator $u = a\varepsilon = a/2$. For $t < s$ the current label has mean zero, so expanding the squares,

$$\mathbb{E}_1\big[f_t(aZ_t) - f_t(a\varepsilon)\big] = \frac{\sigma}{2}\big(\mathbb{E}_1 Z_t^2 - \varepsilon^2\big) \;\ge\; -\frac{\sigma\varepsilon^2}{2}. \tag{68}$$

For $t \in I$ the label has mean ε, so $\mathbb{E}_1[f_t(aZ_t) - f_t(a\varepsilon)] = \frac{\sigma}{2}\mathbb{E}_1(Z_t - \varepsilon)^2$, and the identity $(z - \varepsilon)^2 - \big(\frac{\varepsilon^2}{2} - z^2\big) = 2\big(z - \frac{\varepsilon}{2}\big)^2 \ge 0$ gives

$$\mathbb{E}_1\big[f_t(aZ_t) - f_t(a\varepsilon)\big] \;\ge\; \frac{\sigma\varepsilon^2}{4} - \frac{\sigma}{2}\mathbb{E}_1 Z_t^2. \tag{69}$$

At the endpoint $\tau = s + m - 1$ the weights inside the block sum to $\sum_{t=s}^{\tau} \lambda^{\tau - t} = h(1 - \lambda^m)$, while those before the block sum to at most $\sum_{t<s} \lambda^{\tau - t} = \lambda^m S_{s-1}(\lambda) \le h\lambda^m$. Weighting equation 68–equation 69, using $\lambda^{\tau - t} \le 1$ on the energy terms together with equation 67, and substituting $\varepsilon = \frac{1}{2}$,

$$\frac{1}{\sigma}\mathbb{E}_1 R_\tau^\lambda(a/2) \;\ge\; \frac{1}{16}h\big(1 - \lambda^m\big) - \frac{1}{2}\cdot\frac{h}{128} - \frac{1}{8}h\lambda^m = h\Big(\frac{15}{256} - \frac{3}{16}\lambda^m\Big).$$

From $\log(1 - 1/h) \le -1/h$ and $m = 8h$ we get $\lambda^m \le e^{-8} < \frac{1}{256}$, so the bracket is at least $\frac{15}{256} - \frac{3}{4096} = \frac{237}{4096} > \frac{1}{32}$, whence $\mathbb{E}_1 R_\tau^\lambda(a/2) \ge \sigma h/32 = G^2 h/(128\mu)$. Since $\log Q/16 \ge 2$ we have $\lfloor \log Q/16 \rfloor \ge \log Q/32$, giving the second inequality of equation 33.

First $\tau \ge N/2 > h$, so $H_{\tau,\lambda} = h$. Put $z = \tau/h$; from $h \le \log N/16$ we get $z \ge 8N/\log N$. Writing $L = \log N \ge 32$, the series bound $e^L \ge L^2/2$ gives

$$\frac{8N}{\log N} = \frac{8e^L}{L} \;\ge\; 4L \;\ge\; L + 1 \;\ge\; \log\tau + 1.$$

By Lemma 2, $K_\tau(\lambda) \le 1 + e^{-z}(z + \log\tau)$. The function $z \mapsto ze^{-z}$ has derivative $(1 - z)e^{-z}$ and so is at most $1/e$ on $z \ge 0$, while $z \ge \log\tau + 1$ gives $(\log\tau)e^{-z} \le (\log\tau)/(e\tau) \le 1/e$ using $\log\tau \le \tau$. Hence $K_\tau \le 1 + 2/e < 2$.

The quantities τ, λ, u are chosen deterministically, and the shifted environment draws its labels only afterwards. If every deterministic label sequence gave expected regret below the right-hand side of equation 33, then so would the average over label sequences, contradicting what we proved. Hence at least one deterministic sequence attains it, and all the squared losses satisfy Assumption 1 by Lemma 19. □

Proof of Theorem 2. Use $\mathcal{A}_{\mathrm{P}}^\sharp$. Subtracting $\frac{1}{2}K_T(\lambda)$ from equation 12 after dividing by G^2/μ gives, for all losses, $T \le N$, λ and u,

$$\frac{\mu}{G^2}R_T^\lambda(u) - \frac{1}{2}K_T(\lambda) \;\le\; \frac{1}{2} + 8\log\frac{3}{\pi_{i_N(\lambda)}}.$$

By Lemma 17 and $i_N(\lambda) \le M$, $\log(3/\pi_{i_N(\lambda)}) \le \log 6 + M\log 2 + 2\log(M + 1)$, and $M\log 2 \le \log N + \log 2$. We also need $M + 1 \le N$: induction gives $2^{n-1} \ge n$ for integers $n \ge 1$, with equality at $n = 1$ and the following inequality from $2^n \ge 2n \ge n + 1$; hence $\log_2 N \le N - 1$ and, rounding up, $M \le N - 1$. Therefore $\log(3/\pi_{i_N(\lambda)}) \le \log 12 + 3\log N$ and

$$\Delta_N^{\mathrm{P}}(A) \;\le\; \frac{1}{2} + 8\log 12 + 24\log N \;\le\; 80(1 + \log N),$$

using $\log 12 < 3$, verified by $e^3 > 1 + 3 + \frac{9}{2} + \frac{27}{6} = 13 > 12$. Taking the infimum over algorithms gives the upper bound. (The same computation with the looser bookkeeping $K_T \le 2 + \log N$ retained in the excess yields the same order.)

Fix any algorithm A not receiving λ and write $d_A = \Delta_N^{\mathrm{P}}(A)$, $L = \log N$. If $d_A \ge L$ then already $d_A \ge L/16384$, so assume $d_A < L$.

For each deterministic squared-loss sequence and $u = 0$, the definition equation 4 gives $\frac{\mu}{G^2}\mathbb{E}_A R_N^\lambda(0) \le \frac{1}{2}K_N(\lambda) + d_A$. Let $\lambda \uparrow 1$: the decision sequence does not depend on the evaluation discount, the finitely many weights converge to 1, and each round's loss difference is bounded by Lemma 12, so the limit may be taken termwise inside the expectation. With Lemma 2 this gives $\mathbb{E}_A R_N^1(0) \le \frac{G^2}{\mu}\big(\frac{1}{2}\mathsf{H}_N + d_A\big)$. Averaging over the zero-mean independent squared-loss environment leaves the right-hand side unchanged, and since $\sigma = G^2/(4\mu)$ this is exactly the hypothesis of Theorem 7 with $B = 2\mathsf{H}_N + 4d_A$. From $\mathsf{H}_N \le 1 + L$ and $d_A < L$ we get $1 \le B \le 2 + 6L \le 8L$, hence $Q = N/B \ge N/(8L)$ and

$$\log Q \;\ge\; L - \log(8L) \;\ge\; L/2, \tag{70}$$

the last inequality because $F(L) = L/2 - \log(8L)$ has derivative $\frac{1}{2} - \frac{1}{L} > 0$ for $L \ge 32$ and $F(32) = 16 - 8\log 2 > 8 > 0$. Since $L \ge 16384$, equation 70 also gives $Q \ge e^{32}$, so Theorem 7 applies and produces τ, λ and $u = a/2$ with $\frac{\mu}{G^2}\mathbb{E}_1 R_\tau^\lambda(u) \ge \frac{\log Q}{4096}$ and $K_\tau(\lambda) \le 2$. The definition equation 4 holds for each deterministic loss sequence, so averaging over this environment forces

$$d_A \;\ge\; \frac{\log Q}{4096} - 1 \;\ge\; \frac{L}{8192} - 1 \;\ge\; \frac{L}{16384},$$

the last inequality being equivalent to $L \ge 16384$. Both cases $d_A \ge L$ and $d_A < L$ therefore give the same bound, and taking the infimum over algorithms gives equation 15. □

Proof of Lemma 10. For the first lower bound, let $x \ge 2$ be real and put $F(x) = x\log(1 - 1/x)$. Then

$$F'(x) = \log\Big(1 - \frac{1}{x}\Big) + \frac{1}{x-1} = -\log\Big(1 + \frac{1}{x-1}\Big) + \frac{1}{x-1} \;\ge\; 0,$$

using $\log(1+u) \le u$ for $u \ge 0$, which follows from $\int_0^u (1+s)^{-1}ds \le u$. Hence $F(h) \ge F(2) = \log\frac{1}{4}$, and exponentiating gives $(1-1/h)^h \ge \frac{1}{4}$. Conversely $\log(1-1/h) \le -1/h$ gives $\lambda^h \le e^{-1}$, and $e > 2$ gives $e^{-1} < \frac{1}{2}$.

For the second claim, put $z = N/h \ge 8\log N$. By Lemma 2,

$$K_N(1 - 1/h) \;\le\; 1 + e^{-z}\big(z + \log N\big) \;\le\; 1 + \frac{9}{8}ze^{-z} \;\le\; 1 + \frac{9}{8e} < 2,$$

where $\log N \le z/8$ was used in the middle inequality, $ze^{-z} \le 1/e$ was verified in the proof of Lemma 21, and $e > 2$ gives $9/(8e) < 9/16 < 1$. □

Proof of Lemma 11. Since $0 \le W \le S$ we have $0 \le \nu \le \varepsilon \le \frac{1}{2}$, so $u = a\nu \in [-a, a]$ and u does not depend on the realized labels.

First the benchmark signal weights. Inside the block the age $N - t$ runs exactly over $h, h+1, \dots, 2h-1$, so

$$W = \lambda^h S_h(\lambda) = h\lambda^h\big(1 - \lambda^h\big).$$

By Lemma 10, $\lambda^h \ge \frac{1}{4}$ and $1 - \lambda^h \ge \frac{1}{2}$, hence

$$W \;\ge\; \frac{h}{8}, \qquad S \le h, \qquad \frac{W^2}{2S} \;\ge\; \frac{h}{128}. \tag{71}$$

Next compare the two environments. Let $\mathcal{F}_{t-1}$ contain the algorithm's seed and the first $t-1$ labels, so Z_t is measurable with respect to it. For $t \in I$ the likelihood ratio of this history under $\mathbb{P}_1$ relative to $\mathbb{P}_0$ is $L_{t-1} = \prod_{s\in I,\, s<t}(1 + \varepsilon V_s)$, because a mean-$\varepsilon$ sign variable takes the value v with probability $(1+\varepsilon v)/2$ against $1/2$ under the null, while the other rounds and the seed are unchanged and contribute no factor. Under the null these labels are independent, so

$$\mathbb{E}_0 L_{t-1}^2 = \big(1 + \varepsilon^2\big)^{|I \cap \{1,\dots,t-1\}|} \;\le\; e^{h\varepsilon^2} = e^b.$$

Cauchy–Schwarz then gives $\mathbb{E}_1|Z_t| = \mathbb{E}_0[L_{t-1}|Z_t|] \le e^{b/2}\sqrt{\mathbb{E}_0 Z_t^2}$, and summing over the h rounds of the block and applying Cauchy–Schwarz to the resulting finite vectors,

$$\sum_{t\in I}\mathbb{E}_1|Z_t| \;\le\; e^{b/2}\sqrt{hE_I}. \tag{72}$$

Finally expand the true loss. Write $m_t = \mathbb{E}_1 V_t = \varepsilon \mathbf{1}\{t \in I\}$. Since V_t is independent of the earlier history and of the algorithm's current randomness,

$$\frac{1}{\sigma}\mathbb{E}_1\big[f_t(x_t) - f_t(u)\big] = \frac{1}{2}\mathbb{E}_1\big[(Z_t - V_t)^2 - (\nu - V_t)^2\big] = \frac{1}{2}\mathbb{E}_1 Z_t^2 - m_t \mathbb{E}_1 Z_t + m_t \nu - \frac{\nu^2}{2}.$$

Multiplying by the weights and summing, using $\sum_t \lambda^{N-t} m_t = \varepsilon W$ and $\nu = \varepsilon W/S$, gives the exact identity

$$\frac{1}{\sigma}\mathbb{E}_1 R_N^{\lambda}(u) = \frac{1}{2}\sum_{t=1}^{N} \lambda^{N-t}\mathbb{E}_1 Z_t^2 - \varepsilon \sum_{t \in I} \lambda^{N-t}\mathbb{E}_1 Z_t + \frac{\varepsilon^2 W^2}{2S}. \tag{73}$$

The first term is nonnegative and may be dropped. The second has absolute value at most $\varepsilon \sum_{t \in I} \mathbb{E}_1 |Z_t|$, since the weights lie in $[0, 1]$. Controlling the last term by equation 71 and the middle one by equation 72, and substituting $b = h\varepsilon^2$,

$$\frac{1}{\sigma}\mathbb{E}_1 R_N^{\lambda}(u) \;\geq\; \frac{\varepsilon^2 h}{128} - \varepsilon e^{b/2}\sqrt{hE_I} = \frac{b}{128} - e^{b/2}\sqrt{bE_I},$$

which is equation 34. Note that the costs of the comparator on the rounds after the block and in the earlier history are retained in equation 73 rather than set to zero. □

Proof of Theorem 3. Fix any algorithm A not receiving the discount, and write $\mathfrak{d} = \Delta_N^{\mathrm{E}}(A)$, $L = \log N$, $\ell = \log L$. If $\mathfrak{d} = +\infty$ the claim holds, and if $\mathfrak{d} \geq \ell$ then $\mathfrak{d} \geq \ell/8192$. So assume

$$0 \leq \mathfrak{d} < \ell, \qquad \ell \geq 8192. \tag{74}$$

Put $n = \lfloor N/(8L) \rfloor$ and $\lambda_0 = 1 - 1/n$. Since L far exceeds 64 we have $N/(8L) \geq 2$ and $n \geq 2$, which also follows from $e^L \geq L^3/6$ for $L \geq 64$. Lemma 10 gives $K_N(\lambda_0) \leq 2$. Averaging the defining guarantee over the zero-mean environment with $u = 0$,

$$\frac{\mu}{G^2}\mathbb{E}_0 R_N^{\lambda_0}(0) \;\leq\; \frac{1}{2}K_N(\lambda_0) + \mathfrak{d} \leq 1 + \mathfrak{d},$$

and since the squared-loss expansion gives $\mathbb{E}_0 R_N^{\lambda_0}(0) = \frac{\sigma}{2}\sum_t \lambda_0^{N-t}\mathbb{E}_0 Z_t^2$ while $G^2/\mu = 4\sigma$, we obtain

$$\sum_{t=1}^{N} \lambda_0^{N-t}\mathbb{E}_0 Z_t^2 \;\leq\; 8(1 + \mathfrak{d}). \tag{75}$$

In the most recent $\lfloor n/2 \rfloor$ rounds the age $k = N - t$ is below $n/2$, so $(1 - 1/n)^k \geq 1 - k/n \geq \frac{1}{2}$ by equation 49. Dropping the other nonnegative terms of equation 75 and bounding these weights below by $\frac{1}{2}$,

$$\sum_{t=N-\lfloor n/2 \rfloor + 1}^{N} \mathbb{E}_0 Z_t^2 \;\leq\; 16(1 + \mathfrak{d}). \tag{76}$$

Put $q = \lceil \log_2(L^2) \rceil$ and $k = \lfloor \log_2(n/2) \rfloor - q$, and for $j = 0, \ldots, k-1$ set $h_j = 2^{q+j}$ and $I_j = \{N - 2h_j + 1, \ldots, N - h_j\}$. We first check there are enough legal blocks. From $n \geq N/(16L)$ and the rounding bounds,

$$k \;\geq\; \log_2(n/2) - 1 - \log_2(L^2) - 1 \;\geq\; \frac{L - \log(32L^3)}{\log 2} - 2.$$

For $L \geq 64$ we have $\log(32L^3) \leq L/2$: the function $F(L) = L/2 - \log 32 - 3\log L$ has derivative $\frac{1}{2} - \frac{3}{L} > 0$ and $F(64) = 32 - 5\log 2 - 18\log 2 > 32 - 23 > 0$, using $\log 2 < 1$. Hence

$$k \;\geq\; \frac{L}{2\log 2} - 2 \;\geq\; \frac{L}{2} - 2 \;\geq\; \frac{L}{4}, \tag{77}$$

while $k \leq \log_2 n \leq L/\log 2 \leq 2L$ since $\log 2 \geq \frac{1}{2}$; in particular k is a positive integer. Each scale satisfies $h_j \geq 2^q \geq L^2 \geq 2$, consecutive age ranges $[h_j, 2h_j - 1]$ and $[2h_j, 4h_j - 1]$ are disjoint,

and the largest age obeys $2h_{k-1} = 2^{q+k} = 2^{\lfloor \log_2(n/2) \rfloor} \le n/2$, so all blocks lie inside the recent window of equation 76. Since the blocks are disjoint, some deterministic index $j_\star$ satisfies

$$E_{I_{j_\star}} \;\le\; \frac{16(1+\mathfrak{d})}{k}, \tag{78}$$

and this index depends only on the algorithm's expected null-environment energy, not on the labels eventually drawn.

Put

$$Q = \frac{k}{16(1+\mathfrak{d})}, \qquad b = \tfrac{1}{2}\log Q, \qquad h = h_{j_\star}, \qquad \varepsilon = \sqrt{b/h}. \tag{79}$$

These are legal. By equation 74 and equation 77, $Q \ge L/\big(64(1+\ell)\big)$, so $\log Q \ge \ell - \log\big(64(1+\ell)\big)$. For $\ell \ge 64$ we have $\log\big(64(1+\ell)\big) \le \ell/2$: the difference of the two sides has derivative $\frac{1}{2} - \frac{1}{1+\ell} > 0$, and at $\ell = 64$ it exceeds $32 - 13 > 0$ because $\log 64 < 6$ and $\log 65 < 7$. Hence

$$\log Q \;\ge\; \ell/2 \;\ge\; 32, \tag{80}$$

so $b \ge 16 > 0$. Also $Q \le k \le 2L$, so $b \le \frac{1}{2}\log(2L) \le L/2 \le h/4$, using $\log(2L) \le L$ for $L \ge 2$. The function $L - \log(2L)$ has derivative $1 - 1/L > 0$ and is positive at $L = 2$ because $e^2 > 4$. Together with $h \ge L^2$ and $L \ge 2$, this gives $0 < \varepsilon \le \frac{1}{2}$.

Shift the labels in $I_{j_\star}$ to mean ε, leaving the rest at mean zero, and apply Lemma 11. By equation 78, $E_I \le 1/Q$, so $e^{b/2}\sqrt{bE_I} \le \sqrt{b}\,Q^{-1/4}$. From $\log Q \ge 32$ and $b \ge 16$ we get $Q^{1/4}\sqrt{b} \ge e^8 \cdot 4 > 256$, since $e > 2$ gives $e^8 > 2^8 = 256$; taking reciprocals and multiplying by b gives $\sqrt{b}\,Q^{-1/4} \le b/256$. Lemma 11 therefore gives, with $\lambda = 1 - 1/h$,

$$\frac{1}{\sigma}\mathbb{E}_1 R_N^\lambda(u) \;\ge\; \frac{b}{128} - \frac{b}{256} = \frac{b}{256}. \tag{81}$$

From $2h \le n/2$ we get $h \le n/4 \le N/(8\log N)$, so Lemma 10 gives $K_N(\lambda) \le 2$. Dividing equation 81 by $G^2/\mu = 4\sigma$ and using equation 79 and equation 80,

$$\frac{\mu}{G^2}\mathbb{E}_1 R_N^\lambda(u) \;\ge\; \frac{b}{1024} = \frac{\log Q}{2048} \;\ge\; \frac{\ell}{4096}.$$

The per-sequence guarantee defining $\mathfrak{d}$, averaged over this environment, therefore forces

$$\mathfrak{d} \;\ge\; \frac{\ell}{4096} - \frac{1}{2}K_N(\lambda) \;\ge\; \frac{\ell}{4096} - 1 \;\ge\; \frac{\ell}{8192},$$

the last inequality using $\ell \ge 8192$. This settles equation 74, and the remaining cases were handled at the outset. The randomized environment is only a device: if every deterministic label sequence violated the bound at this fixed λ and u, so would a finite mixture of them. Taking the infimum over algorithms gives the first claim of equation 16, and combining it with the numerical upper bound in equation 13 gives the stated order. □

Proof of Corollary 1. Algorithms satisfying Assumption 2 form a subset of the full-feedback class. For the same supremum criterion over losses, an infimum over a smaller class cannot be smaller, so $\Delta_N^{\mathrm{P},\star,G} \ge \Delta_N^{\mathrm{P},\star}$ and $\Delta_N^{\mathrm{E},\star,G} \ge \Delta_N^{\mathrm{E},\star}$; Theorems 2 and 3 bound the right-hand sides below by the logarithmic and doubly logarithmic orders. Conversely, equation 12 with Lemma 17 gives a prefix excess of $O(\log N)$, and equation 13 gives an endpoint excess of at most $50\log(e + \log N)$; both are attained by algorithms obeying Assumption 2. □

G Proofs for Section 5.4

Proof of Theorem 8. Each round an active expert's mass grows by at most $(1-\gamma_i)\pi_i$, while inactive and finished experts are unchanged. Accumulating over all rounds,

$$Q_t \;\le\; 1 + \sum_i \pi_i(1-\gamma_i)m_i = \sum_i \pi_i\kappa_i = C_\pi,$$

the equality holding because the initial total mass is 1 and $\sum_i \pi_i = 1$.

Suppose $\alpha_\sharp^{-1}\log(C_\pi/\pi_i) \le \mathfrak{a}_i$ for every i. Exponentiating and rearranging gives $\pi_i \ge C_\pi e^{-\alpha_\sharp \mathfrak{a}_i}$. Multiplying by κ_i and summing,

$$C_\pi = \sum_i \kappa_i \pi_i \;\ge\; C_\pi \sum_i \kappa_i e^{-\alpha_\sharp \mathfrak{a}_i}.$$

Since $C_\pi \ge 1 > 0$, dividing by it gives the stated condition.

Let $Z = \sum_i e^{-\alpha_\sharp \mathfrak{a}_i} > 0$ and put $\pi_i = e^{-\alpha_\sharp \mathfrak{a}_i}/Z$, which is strictly positive because the targets are finite. The assumed condition gives

$$C_\pi = \frac{\sum_i \kappa_i e^{-\alpha_\sharp \mathfrak{a}_i}}{Z} \;\le\; \frac{1}{Z},$$

hence $C_\pi/\pi_i \le e^{\alpha_\sharp \mathfrak{a}_i}$; taking logarithms and dividing by $\alpha_\sharp$ gives the required per-expert bound.

If all targets are at most $\mathfrak{a}$, necessity requires $e^{-\alpha_\sharp \mathfrak{a}} \sum_i \kappa_i \le 1$, i.e. $\mathfrak{a} \ge \alpha_\sharp^{-1} \log \sum_i \kappa_i$. Conversely, with m experts the uniform prior $\pi_i = 1/m$ gives $C_\pi/\pi_i = \sum_j \kappa_j$ for every i simultaneously, attaining that value. Lower bound and attainable upper bound coincide, giving equation 35.

For the prefix grid $m_i = N$, so

$$N+1 \;\le\; \sum_{i=0}^{M} \kappa_i = (M+1) + \sum_{i=0}^{M} \min\{2^i, N\} \;<\; 5N,$$

the lower bound coming from the single term $i = M$, where $\delta_M = 1$ and hence $\kappa_M = 1 + N$, and the upper bound from $M + 1 \le N$ together with $\sum_{i=0}^{M} 2^i = 2^{M+1} - 1 < 4N$. For endpoint activation, Lemma 18 gives $\delta_i m_i \le 9(M - i + 1)$, so

$$M{+}1 \;\le\; \sum_i \kappa_i \;\le\; (M{+}1){+}9\sum_{i=0}^{M}(M{-}i{+}1) = (M{+}1){+}\frac{9}{2}(M{+}1)(M{+}2) \;\le\; 5(M{+}1)(M{+}2),$$

using $M + 2 \ge 2$ in the final inequality. The logarithms are $\Theta(\log N)$ and $\Theta\big(\log(M+1)\big) = \Theta(\log\log N)$.

Take a single expert with $\pi_1 = 1$ and run Algorithm 1. Whenever the compensation coefficient is positive the master prediction coincides with that expert's prediction, so $\ell_t(z_{t,1}) = 0$; whenever it vanishes the exponent vanishes too. Induction from $q_{0,1} = 1$ gives $q_{t,1} = 1$ for all t, so the realized compensated meta-regret is identically zero. Yet if $\gamma_1 < 1$ and $m_1 > 0$ then $C_\pi = 1{+}(1{-}\gamma_1)m_1 > 1$, and the certificate asserts a strictly positive cost. Hence equation 35, although exactly optimal for this certificate family, can overstate the true meta-regret, and necessity must be argued by change of measure instead. □

Proof of Theorem 9. Work in the zero-mean independent squared-loss environment with $x_t = aZ_t$ and $\sigma = G^2/(4\mu)$. Lemma 10 gives $K_N(\lambda_0) \le 2$. Averaging the hypothesis at $j = 0$ with $u = 0$ and using $\mathbb{E}_0 R_N^{\lambda_0}(0) = \frac{\sigma}{2}\sum_t \lambda_0^{N-t}\mathbb{E}_0 Z_t^2$ gives

$$\sum_{t=1}^{N} \lambda_0^{N-t}\mathbb{E}_0 Z_t^2 \;\le\; 8(1 + A_0).$$

In the most recent $\lfloor n/2 \rfloor$ rounds the integer age $r = N - t$ is below $n/2$, so $\lambda_0^r \ge 1 - r/n \ge \frac{1}{2}$ by equation 49. Discarding the other nonnegative terms and dividing by $\frac{1}{2}$,

$$\sum_{t=N-\lfloor n/2 \rfloor+1}^{N} \mathbb{E}_0 Z_t^2 \;\le\; 16(1 + A_0). \tag{82}$$

Every I_j lies in this recent window and they are pairwise disjoint, so writing $E_j = \sum_{t\in I_j} \mathbb{E}_0 Z_t^2$ we get $\sum_j E_j \le 16(1 + A_0)$.

For each j construct a separate alternative environment, shifting only the labels in I_j to mean

$$\varepsilon_j = \sqrt{b_j/h_j}, \qquad b_j = 1024(1 + A_j).$$

The hypothesis $h_j \geq 4096(1 + A_j)$ gives $0 < \varepsilon_j \leq \frac{1}{2}$, so Lemma 11 applies and supplies a fixed comparator u_j with

$$\frac{1}{\sigma}\mathbb{E}_j R_N^{\lambda_j}(u_j) \;\geq\; \frac{b_j}{128} - e^{b_j/2}\sqrt{b_j E_j}.$$

Here $h_j \leq n/4 \leq N/(8\log N)$, so Lemma 10 gives $K_N(\lambda_j) \leq 2$, and averaging the hypothesis over this environment requires

$$\frac{1}{\sigma}\mathbb{E}_j R_N^{\lambda_j}(u_j) \;\leq\; 4(1 + A_j).$$

Since $b_j/128 = 8(1 + A_j)$, comparing the two displays and rearranging gives $e^{b_j/2}\sqrt{b_j E_j} \geq 4(1 + A_j)$. Both sides are nonnegative, so squaring and dividing by $b_j e^{b_j} > 0$ gives

$$E_j \;\geq\; \frac{16(1 + A_j)^2}{1024(1 + A_j)e^{1024(1+A_j)}} = \frac{1}{64}(1 + A_j)e^{-1024(1+A_j)}.$$

Summing these lower bounds over the disjoint blocks and comparing with equation 82 gives equation 36. Shifting a mean merely constructs a different alternative sequence distribution inside the proof; the blocks are never shifted simultaneously, and since the assumed guarantee holds for all sequences it must hold for each alternative environment separately. □

Proof of Corollary 2. For $x \geq 1$ the function $F(x) = xe^{-1024x}$ has derivative $e^{-1024x}(1 - 1024x) < 0$, so it is decreasing; hence $A_j \leq \mathfrak{d}$ gives $(1 + A_j)e^{-1024(1+A_j)} \geq (1 + \mathfrak{d})e^{-1024(1+\mathfrak{d})}$. Substituting these into equation 36 and using $1 + A_0 \leq 1 + \mathfrak{d}$ on the right,

$$k(1 + \mathfrak{d})e^{-1024(1+\mathfrak{d})} \;\leq\; 1024(1 + \mathfrak{d}).$$

Dividing by $1 + \mathfrak{d} > 0$, taking logarithms and rearranging gives the claim. □

Proof of Theorem 10. Use $\mathcal{A}_{\mathrm{E}}^{\sharp}$. On Λ_N^{st} we have $K_N(\lambda) \leq 2$, so equation 13 together with $v_N \geq 0$ gives

$$\frac{\mu}{G^2}R_N^{\lambda}(u) - v_N(\lambda) \;\leq\; 1 + 2 + 8\mathcal{L}_N \;\leq\; 1 + 50\log(e + \log N) \;\leq\; 51\log(e + \log N),$$

the last inequality using $\log(e + \log N) \geq 1$. This algorithm needs only one gradient per round, so it bounds both classes.

Fix any full-feedback algorithm and write $\mathfrak{d} = \Gamma_N^{\mathrm{E}}(A)$, $L = \log N$, $\ell = \log L$. If $\mathfrak{d} \geq \ell$ or $\mathfrak{d} = +\infty$ the claim holds, so assume $0 \leq \mathfrak{d} < \ell$.

Put $n = \lfloor N/(8L) \rfloor$, $q = \lceil \log_2(L^2) \rceil$, $k = \lfloor \log_2(n/2) \rfloor - q$ and $h_j = 2^{q+j}$ for $j = 0, \ldots, k - 1$. Since $L \geq 64$, the series bound $e^L \geq L^3/6$ gives $N/(8L) \geq L^2/48 > 2$, hence $n \geq N/(16L)$ and

$$k \;\geq\; \frac{L - \log(32L^3)}{\log 2} - 2.$$

For $L \geq 64$ the function $L/2 - \log 32 - 3\log L$ has derivative $\frac{1}{2} - \frac{3}{L} > 0$ and value $32 - 23\log 2 > 0$ at $L = 64$, so $\log(32L^3) \leq L/2$ and therefore

$$k \;\geq\; L/2 - 2 \;\geq\; L/4. \tag{83}$$

Each scale satisfies $h_j \geq L^2$; the consecutive age ranges $[h_j, 2h_j - 1]$ and $[2h_j, 4h_j - 1]$ are disjoint; and the largest age obeys $2h_{k-1} = 2^{\lfloor \log_2(n/2) \rfloor} \leq n/2$. All the geometric hypotheses of Theorem 9 therefore hold.

Next the signal-strength hypothesis. We have $\mathfrak{d} < \ell$ and $L^2 = e^{2\ell}$, so for $\ell \geq 128$ the exponential series and $1 + \ell \leq 2\ell$ give

$$e^{2\ell} \;\geq\; \frac{(2\ell)^3}{6} = \frac{4}{3}\ell^3 \;\geq\; 8192\ell \;\geq\; 4096(1 + \ell),$$

the middle inequality being equivalent to $\ell^2 \geq 6144$, valid for $\ell \geq 128$. Hence every $h_j \geq 4096(1 + \mathfrak{d})$.

The anchor discount $\lambda_0 = 1 - 1/n$ and every annulus discount $1 - 1/h_j$ lie in Λ_N^{st} by Lemma 10. By the definition of $\mathfrak{d}$, at each of these discounts

$$\frac{\mu}{G^2}\mathbb{E}_A R_N^{\lambda}(u) \;\le\; v_N(\lambda) + \mathfrak{d} \;\le\; \frac{1}{2}K_N(\lambda) + \mathfrak{d},$$

using $v_N \le K_N/2$ from Theorem 5. We may therefore set every A_j equal to $\mathfrak{d}$ in Theorem 9, and Corollary 2 gives

$$\mathfrak{d} \;\ge\; \frac{\log k - \log 1024}{1024} - 1 \;\ge\; \frac{\ell - \log 4096}{1024} - 1,$$

using equation 83. Since $\ell \ge 8192$ and $\log 4096 = 12\log 2 < 12$,

$$\frac{\ell - \log 4096}{1024} - 1 - \frac{\ell}{2048} = \frac{\ell}{2048} - 1 - \frac{\log 4096}{1024} \;\ge\; 4 - 1 - \frac{12}{1024} > 0,$$

so $\mathfrak{d} \ge \ell/2048$. This holds for every full-feedback algorithm, hence for every algorithm in the smaller single-gradient class; taking infima over each class gives equation 37. □

Proof of Theorem 11. Take $\mathcal{A}_{\mathrm{P}}^{\sharp}$. By equation 12, $K_T \le 2$ on Λ_T^{st}, and $v_T \ge 0$, the excess inside the steady-state set is at most $\frac{3}{2} + 8\log\bigl(3/\pi_{i_N(\lambda)}\bigr)$. The bound $\log(3/\pi_i) \le \log 12 + 3\log N$ verified in the proof of Theorem 2 makes this at most $\frac{3}{2} + 8\log 12 + 24\log N \le 81(1+\log N)$. This algorithm lies in the single-gradient class, so both classes obey the bound.

Fix any full-feedback algorithm, write $\mathfrak{d} = \Gamma_N^{\mathrm{P}}(A)$, and assume $0 \le \mathfrak{d} < L$, the cases $\mathfrak{d} \ge L$ and $\mathfrak{d} = +\infty$ being immediate. Put

$$n = \Bigl\lfloor \frac{N}{8L} \Bigr\rfloor, \quad \lambda_0 = 1 - \frac{1}{n}, \quad Q = \frac{n}{64(1+\mathfrak{d})}, \quad h = \Bigl\lfloor \frac{\log Q}{16} \Bigr\rfloor, \quad m = 8h. \tag{84}$$

These are legal. Since $L \ge 64$, the series bound $N = e^L \ge L^3/6$ gives $N/(8L) \ge L^2/48 > 2$, and $\lfloor x \rfloor \ge x/2$ for real $x \ge 2$ gives $n \ge N/(16L)$, so

$$Q \;\ge\; \frac{N}{1024L(1+L)} \;\ge\; \frac{N}{2048L^2}.$$

For $L \ge 64$ the function $L/2 - \log 2048 - 2\log L$ has derivative $\frac{1}{2} - \frac{2}{L} > 0$ and value $32 - 23\log 2 > 0$ at $L = 64$, so

$$\log Q \;\ge\; L/2 \;\ge\; 32. \tag{85}$$

Hence $h \ge 2$, and $Q \le n \le N$ gives $h \le L/16$ and $m \le L/2$. We also need enough complete blocks: the series bound $N = e^L \ge L^4/24 \ge 64L^2$ for $L \ge 64$ gives $n \ge N/(16L) \ge 4L \ge 8m$.

Lemma 10 gives $K_N(\lambda_0) \le 2$, hence $v_N(\lambda_0) \le 1$ by Theorem 5. In the zero-mean environment the guarantee defining $\mathfrak{d}$ therefore yields, exactly as in equation 75,

$$\sum_{t=1}^{N} \lambda_0^{N-t}\mathbb{E}_0 Z_t^2 \;\le\; 8(1+\mathfrak{d}),$$

and since the most recent $r = \lfloor n/2 \rfloor$ rounds carry weight at least $\frac{1}{2}$ by equation 49,

$$\sum_{t=N-r+1}^{N} \mathbb{E}_0 Z_t^2 \;\le\; 16(1+\mathfrak{d}). \tag{86}$$

Partition the most recent r rounds from left to right into complete blocks of length m, discarding the short tail; the number of complete blocks is $k = \lfloor r/m \rfloor$. Since $n \ge 8m$ and $n \ge 4$ we have $r \ge n/4$ and $r/m \ge 2$, hence $k \ge r/(2m) \ge n/(8m)$. By equation 86 some deterministic block $I = \{s, \dots, s+m-1\}$ satisfies

$$E_I = \sum_{t \in I} \mathbb{E}_0 Z_t^2 \;\le\; \frac{16(1+\mathfrak{d})}{k} \;\le\; \frac{128(1+\mathfrak{d})m}{n} = \frac{2m}{Q}. \tag{87}$$

Put $\tau = s + m - 1$; this endpoint belongs to the prefixes the protocol permits, and $\tau \ge N - r + 1 \ge N/2$. The block's position is determined by the algorithm's expected null-environment behaviour, not by the realized data.

Shift only the labels in I to mean $\frac{1}{2}$, keeping the others independent with mean zero. The likelihood-ratio computation of Lemma 9 gives

$$\sum_{t\in I}\mathbb{E}_1 Z_t^2 \;\le\; e^{m/8}\sqrt{mE_I} \;\le\; \sqrt{2}\,m\,e^{m/8}Q^{-1/2} \;\le\; 8\sqrt{2}\,h\,Q^{-7/16} \;\le\; \frac{h}{128}, \tag{88}$$

the penultimate inequality using $m = 8h$ and $h \le \log Q/16$, and the last following from $\log Q \ge 32$, since $8\sqrt{2}\,Q^{-7/16} \le 8\sqrt{2}\,e^{-14} < 1/128$ because $\sqrt{2} < 2$ and $e^{14} > 2^{14} > 2048$.

Take $\lambda = 1 - 1/h$ and the fixed comparator $u = a/2$. For rounds before the block the expected regret is at least $-\sigma/8$ by equation 68; for rounds inside the block, $(z - \frac{1}{2})^2 \ge \frac{1}{8} - z^2$ gives an expected regret of at least $\frac{\sigma}{16} - \frac{\sigma}{2}\mathbb{E}_1 Z_t^2$. The total weight inside the block is $h(1-\lambda^m)$ and that before it is at most $h\lambda^m$, so with equation 88,

$$\frac{1}{\sigma}\mathbb{E}_1 R_\tau^\lambda(a/2) \;\ge\; \frac{h}{16}\big(1-\lambda^m\big) - \frac{h}{256} - \frac{h}{8}\lambda^m = h\Big(\frac{15}{256} - \frac{3}{16}\lambda^m\Big).$$

Since $\lambda^m \le e^{-8} < 1/256$, the bracket exceeds $\frac{237}{4096} > \frac{1}{32}$, so

$$\frac{\mu}{G^2}\mathbb{E}_1 R_\tau^\lambda(a/2) \;\ge\; \frac{h}{128} \;\ge\; \frac{\log Q}{4096} \;\ge\; \frac{L}{8192}, \tag{89}$$

the middle inequality using $\log Q/16 \ge 2$, so that rounding down loses at most a factor of two.

Put $z = \tau/h$. From $\tau \ge N/2$ and $h \le L/16$ we get $z \ge 8N/L$, and the series bound $N \ge L^2/2$ gives $z \ge 4L \ge \log\tau + 1$. By Lemma 2,

$$K_\tau(\lambda) \;\le\; 1 + e^{-z}\big(z + \log\tau\big) \;\le\; 1 + \frac{1}{e} + \frac{1}{e} < 2,$$

using $ze^{-z} \le 1/e$ and $(\log\tau)e^{-z} \le (\log\tau)/(e\tau) \le 1/e$. Hence $\lambda \in \Lambda_\tau^{\mathrm{st}}$ and $v_\tau(\lambda) \le 1$ by Theorem 5. Comparing equation 89 with the definition of $\mathfrak{d}$,

$$\mathfrak{d} \;\ge\; \frac{L}{8192} - v_\tau(\lambda) \;\ge\; \frac{L}{8192} - 1 \;\ge\; \frac{L}{16384},$$

the last inequality using $L \ge 16384$. The randomized environment converts to a deterministic sequence as before: if every deterministic label sequence made the expected regret strictly smaller than the bound just derived, so would a finite mixture, a contradiction. Taking the infimum over full-feedback algorithms and using the class inclusion gives equation 38 for both feedback models. □

H Additional Experimental Results

H.1 Experimental Setup

The theory makes two limits of the experiments explicit. The statements are worst case over losses, so each reported number is a supremum over an instance family rather than an average. The lower bounds quantify over *all* algorithms, so simulations cannot establish Theorem 2 or 3. We therefore report experiments on the achievable side of Theorems 1 and 6, the shapes of their certificates, and the exactness of the identities the proofs rest on.

We work in the normalized geometry of Section 3 with $\mu = G = 1$ and $\mathcal{X}$ the Euclidean ball of radius $G/(2\mu)$ in $\mathbb{R}^d$. The instance family has ten members: i.i.d. sign targets and a hierarchical unknown-mean instance, which are the constructions behind Lemmas 19 and 20; piecewise-constant targets with segment lengths $8, 32, 128$; period-$4, 16, 64$ oscillation; an instance whose best forgetting scale changes halfway through the run; and a nonsmooth drift $\frac{\mu}{2}\|x - y_t\|^2 + \rho\|x - c_t\|$. Absolute time scales are deliberately independent of N, since an N-dependent family would conflate the growing burden of covering more forgetting scales with the growing intrinsic difficulty of the instance itself. Assumption 1 is verified numerically on every instance: the largest observed $\|g_t\|/G$ is 0.98 and the strong-convexity slack never falls below -10^{-16}.

Regret is measured against the exact maximizer of $u \mapsto R_T^\lambda(u)$ over $\mathcal{X}$. Because a trajectory does not depend on the evaluation discount, each algorithm is run once per instance and scored at every λ, following Section 3. We compare the three variants of Algorithm 1 with four baselines. The

first is BDF-Curv, the natural single-gradient reduction of Section 4. It makes the same oracle calls but uses uncompensated weights at the meta rate $\mu/(9G^2)$. The other baselines are three learners committed to $H \in \{16, \sqrt{N}, \infty\}$ and the known-λ oracle of Theorem 4, rerun separately for each λ. We report normalized regret $\frac{\mu}{G^2} R_T^\lambda(u^\star)$ and the *adaptation gap*, namely the excess over that oracle on the same instance.

This section contains the numerical summaries and diagnostic checks omitted from the main text. Table 2 reports the worst-case numerical values underlying Figure 2. Table 3 gives the corresponding values for the prefix–endpoint comparison in Figure 3. Figure 4 isolates the effect of the two prior families at the short- and long-memory ends. Finally, Table 4 checks the exact identities used by Algorithm 1.

Table 2: Worst case over 10 instances at $N = 4096$, $d = 10$, $\mu = G = 1$, taken over 40 discounts spanning $H_\lambda \in [1.05, N]$. "Gap" is the additive price of not knowing λ: the excess over the known-λ oracle of Theorem 4 run on the same instance. The last column is the certificate of Theorem 1; fixed-discount learners carry none. Lower is better; the hybrid is best on both worst-case columns while still querying one genuine subgradient and zero function values per round.

	per-round cost			worst case over λ		
Algorithm	grad.	values	proj.	$\frac{\mu}{G^2} R_N^\lambda$	gap	certified
Sharp-BDF hybrid (ours)	1	0	20.1	**1.84**	**1.71**	78
Sharp-BDF endpoint	1	0	7.1	4.93	4.42	89
fixed $H = \sqrt{N}$	1	0	1.0	5.21	4.70	—
Sharp-BDF prefix	1	0	13.0	4.64	5.18	118
BDF-Curv (naive)	1	0	13.0	12.16	12.04	273
fixed $H = 16$	1	0	1.0	19.52	19.01	—
no forgetting ($\gamma = 1$)	1	0	1.0	32.09	31.95	—

Table 3: The quantifier separation, worst case over 10 instances, $d = 5$. The certified terms are the additive excesses of Theorem 1: the prefix certificate grows by $8 \log 2 \approx 5.5$ per doubling of N, while the endpoint certificate grows only through $24 \log(2 + \log_2 H_\lambda)$. Across all three experiments the realised excess never exceeded its certificate, with 0 violations in 16,458 checks.

		certified		realised		
$\log_2 N$	N	P	E	P	E	P/E
7	128	82.6	76.9	1.11	0.28	4.0
8	256	90.0	79.7	2.57	0.51	5.0
9	512	97.3	82.3	5.18	0.70	7.4
10	1024	104.3	84.6	5.01	1.29	3.9
11	2048	111.3	86.6	7.57	2.05	3.7
12	4096	118.1	88.6	11.95	2.91	4.1
13	8192	124.8	90.3	8.52	3.82	2.2
14	16384	131.5	92.0	9.83	4.76	2.1

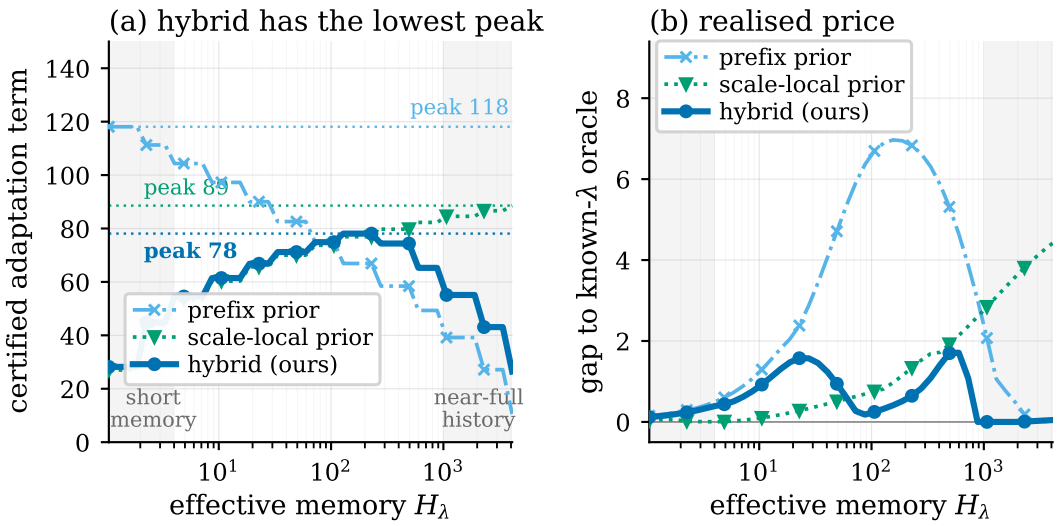


Figure 4: Effect of the prior at $N = 4096$, $d = 10$, worst case over ten instances. (a) Certified adaptation term against effective memory; dotted horizontal lines mark each variant's peak. (b) Realised worst-case gap to the known-λ oracle. Shaded bands mark the short-memory and near-full-history ends.

Table 4: Exactness of the two identities Algorithm 1 relies on, measured at every round of every run in double precision, with the realised total potential mass against its certified budget.

Quantity	observed	asserted by
$\left\lvert \lVert g_t + \mu(z - x_t)\rVert^2 - \lVert g_t\rVert^2 - 2\mu\ell_t(z) \right\rvert$	1.3×10^{-16}	eq. equation 7
$\left\lvert \log(q_{s,i}/\pi_i) - \alpha_\sharp \sum_t \gamma_i^{s-t} c_{t,i} r_{t,i} \right\rvert$	1.9×10^{-12}	eq. equation 26
$\max_t Q_t$, hybrid	1.000	≤ 11, Theorem 6
genuine function values per round	0	Assumption 2