Dynamic Regret in Online Convex Optimization with Indicator Switching Costs

Naram Mhaisen **George Iosifidis**
Faculty of Electrical Engineering, Mathematics and Computer Science
TU Delft, Netherlands

Abstract

We study dynamic regret in online convex optimization with an *indicator switching cost*: a penalty incurred whenever two consecutive decisions differ. This captures startup overheads such as server activation, model deployment, and cache updates, and on a bounded domain it recovers norm-based movement costs as a special case. Existing techniques for indicator costs handle only static comparators, and we show that a straightforward adaptation to dynamic regret provably fails. Instead, We propose a meta-learning framework with a set of randomized FTRL base learners restarted at dyadic time scales and aggregated by a movement-aware master. The master then mixes the proposed densities and samples actions via maximal coupling of the consecutive mixtures. The algorithm achieves strongly adaptive regret of $\tilde{\mathcal{O}}(\sqrt{|I|})$ on every interval I and dynamic regret of $\tilde{\mathcal{O}}(\min\{\sqrt{T(S_T+1)}, T^{2/3}(P_T+1)^{1/3}\})$, both in expectation and including indicator switching costs. Here, S_T counts comparator switches and P_T is the comparator path length. The dynamic regret bound holds for all comparator sequences and requires no prior knowledge of S_T or P_T. It is minimax-optimal (up to logarithmic factors) for tracking piecewise-constant comparators, and also captures frequently moving comparators with small total path length.

1 Introduction

We study Online Convex Optimization (OCO) in non-stationary environments. At each round $t \in [T]$, a learner selects an action $\boldsymbol{x}_t \in \mathcal{X}$ *before* observing the convex loss $f_t(\cdot)$, and incurs $f_t(\boldsymbol{x}_t)$. A central performance criterion in this setting is *dynamic regret*, which compares the learner against a time-varying comparator sequence $\{\boldsymbol{u}_t\}_{t=1}^T$. To capture stability requirements, dynamic regret is often augmented with a *switching cost* $m_t(\boldsymbol{x}_t, \boldsymbol{x}_{t-1})$ that penalizes changes between consecutive actions. The augmented formulation, known as *smoothed* OCO (SOCO), models the central tradeoff between responsiveness to non-stationarity and stability of the decision sequence. While this tradeoff is well understood for switching costs that are norm-based, the regime in which any reconfiguration incurs an overhead independent of the magnitude of the change remains underexplored.

In many practical problems, the switching cost is modeled as an *indicator switching cost* [6, 13, 18]: $\lambda_t \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}$, which charges a (possibly time-varying) penalty λ_t whenever $\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}$, regardless of how large the change is. Such costs arise whenever any reconfiguration triggers a one-shot overhead: server activation in cloud scheduling, cache updates in networks, model redeployment in production ML pipelines, retraining-trigger costs in continual learning, and the per-iteration privacy budget in differentially private learning [2]. The indicator switching cost is, in fact, the *more demanding* primitive. Setting $\lambda_t = \|\boldsymbol{x}_t - \boldsymbol{x}_{t-1}\|$ recovers the norm-based switching cost as a special case, and on a bounded domain $\|\boldsymbol{x}_t - \boldsymbol{x}_{t-1}\| \leq D\,\mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}$ where $D \doteq \operatorname{diam}(\mathcal{X})$ is the diameter of the set $\mathcal{X}$, so any algorithm that controls the indicator switching cost automatically

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

controls the norm-based one (up to the factor D), but not vice versa. In this sense indicator switching costs subsume norm-based ones and impose a strictly stronger notion of stability.

In Learning with Expert Advice (LEA), where the decision set is the simplex and losses are linear, indicator switching costs admit a useful reduction. Under a maximal coupling[1], the expected number of switches is exactly the ℓ_1 movement of the played distributions, so the problem reduces to controlling ℓ_1 norm on the simplex, and both static [4] and dynamic[2] [9] regret bounds follow from standard algorithms such as exponentiated gradient and fixed-share. In general OCO, a line of work has nevertheless made progress on this harder regime via the FTRL framework, obtaining minimax-optimal *static* regret under indicator switching costs [19, 2]. These remain the only available results in OCO, because the other main framework, Online Mirror Descent (OMD), defines each iterate recursively from the previous one via a Bregman divergence term (see, e.g., [14, Sec. 6]). This recursive dependence of each action on its predecessor precludes a closed-form density that can be evaluated pointwise, which is necessary for the maximal coupling used to control indicator switches.

The dynamic-regret case has thus remained open, and the difficulty traces back to the FTRL framework itself. Even without switching costs, vanilla FTRL is suboptimal for dynamic regret [12]; prior work has recovered dynamic-regret guarantees through enhancements such as history pruning [15, 16], but these methods proceed by linearizing the FTRL objective, after which the iterate also becomes a recursive update, inheriting the same difficulty as OMD in deriving the density. Our first contribution shows that even when assisted with perturbations and fixed restarts (which can be viewed as the analogue of history pruning but without linearization) FTRL still fails to deliver (optimal) sublinear dynamic regret. A different architecture beyond a single FTRL run is therefore required.

The natural question is thus: *can one obtain dynamic regret guarantees in OCO when the learner pays indicator switching costs, and if so, against which comparator classes?* We answer this in the affirmative for the two canonical comparator classes in the dynamic regret literature: piecewise-constant comparators that change at most S_T times, and comparators whose total variation is bounded by P_T. As we detail later in the contributions, the bounded-variation class is the larger of the two, and it further admits comparators that change at every round while moving only slightly, sequences against which the switch count S_T is uninformative. Against the first class, we obtain order-optimal dynamic regret of $\tilde{\mathcal{O}}(\sqrt{(S_T+1)\,T})$, matching the minimax rate up to logarithmic factors. Against the second, the same algorithm attains $\tilde{\mathcal{O}}(T^{2/3}P_T^{1/3})$. A single comparator-oblivious algorithm thus handles both canonical forms of non-stationarity simultaneously. The approach section below details how this is achieved through a multiscale meta-learner over restarted lazy randomized FTRL experts.

1.1 Contributions

We first state the standing assumptions used throughout the paper.

Assumption 1. *The set $\mathcal{X} \subset \mathbb{R}^d$ is convex and bounded with diameter $D \doteq \sup_{\boldsymbol{x},\boldsymbol{y}\in\mathcal{X}} \|\boldsymbol{x}-\boldsymbol{y}\|$ and radius $R \doteq \sup_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{x}\|$.*

Assumption 2. *The losses are chosen independently of the learner's randomization (an oblivious adversary). Each $f_t : \mathcal{X} \to \mathbb{R}$ is twice differentiable, convex, G-Lipschitz, and β-smooth; i.e.,*

$$\|\nabla f_t(\boldsymbol{x})\| \le G, \qquad \nabla^2 f_t(\boldsymbol{x}) \preceq \beta I, \qquad \forall t \in [T], \boldsymbol{x} \in \mathcal{X}.$$

Assumption 3. *The switching coefficients $\{\lambda_t\}_{t=2}^T$ satisfy $0 \le \lambda_t \le \lambda$ for a known constant $\lambda > 0$.*

Unlike the losses, the switching coefficients may depend on the learner's current and past actions because our analysis uses only their uniform upper bound λ.

Assumption 4. *There exists $M_f > 0$ such that $0 \le f_t(\boldsymbol{x}) \le M_f, \quad \forall t \in [T], \boldsymbol{x} \in \mathcal{X}$.*

Our performance criterion is the *dynamic regret with indicator switching cost* $\mathcal{R}_T^{\mathbf{1}}$, defined as:

$$\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T) \doteq \sum_{t=1}^T \big(f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t)\big) + \sum_{t=2}^T \lambda_t\, \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\},$$

[1] Specifically, under such a coupling, $\Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) = \frac{1}{2}\|\boldsymbol{x}_t - \boldsymbol{x}_{t-1}\|_1$; see, e.g., [9, Sec. 5.2].

[2] In LEA, Dynamic regret is known as tracking regret. Also, strongly adaptive regret implies tracking regret [10, Sec. 3].

When clear from context, we drop $\{\boldsymbol{u}_t\}_{t=1}^T$ from the argument. $\mathcal{R}_T^{\mathbf{1}}$ compares the learner against a time-varying comparator sequence $\{\boldsymbol{u}_t\}_{t=1}^T = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T)$ while charging the learner for its own switches. We measure the complexity of the comparator sequence through two canonical quantities:

$$S_T(\{\boldsymbol{u}_t\}_{t=1}^T) \doteq \sum_{t=2}^{T} \mathbf{1}\{\boldsymbol{u}_t \neq \boldsymbol{u}_{t-1}\}, \qquad P_T(\{\boldsymbol{u}_t\}_{t=1}^T) \doteq \sum_{t=2}^{T} \|\boldsymbol{u}_t - \boldsymbol{u}_{t-1}\|,$$

These measures induce two comparator classes,

$$\mathcal{U}_S(s) \doteq \{\{\boldsymbol{u}_t\}_{t=1}^T : S_T(\{\boldsymbol{u}_t\}_{t=1}^T) \leq s\}, \qquad \mathcal{U}_P(p) \doteq \{\{\boldsymbol{u}_t\}_{t=1}^T : P_T(\{\boldsymbol{u}_t\}_{t=1}^T) \leq p\}.$$

These classes capture different forms of non-stationarity. $\mathcal{U}_S(s)$ allows the comparator to change at most s times, regardless of the magnitude. $\mathcal{U}_P(p)$ instead places no limit on how often the comparator changes, only on the total magnitude of changes, which cannot exceed p. Since any sequence with s changes has total movement at most Ds, we have $\mathcal{U}_S(s) \subseteq \mathcal{U}_P(Ds)$; but $\mathcal{U}_P$ also contains sequences that, e.g., move at every round by small amounts, for which S_T becomes uninformative. Our guarantees are therefore stated simultaneously for every sequence as functions of S_T and P_T, with the algorithm itself oblivious to which complexity is smaller. The contributions are listed below.

An impossibility for randomized FTRL. We first show that the existing machinery for indicator switching costs, known as FPRLL [18, 19, 2][3], cannot deliver optimal dynamic regret:

(a) Without restarts, FPRLL suffers $\Omega(T)$ dynamic regret even on $\mathcal{U}_S(1)$.

(b) With any fixed restart period k, FPRLL(k) suffers $\Omega\big(T^{2/3}\tau^{1/3}/\log T\big)$ dynamic regret on $\mathcal{U}_S(\tau)$.

Note that the same lower bounds remain valid for the larger class $\mathcal{U}_P(\cdot)$.

Strongly adaptive regret. We introduce a comparator-oblivious algorithm that combines restarted FPRLL base learners at all dyadic time scales, aggregated by the discounted-normal predictor of [9]. For every interval $I = [s, e] \subseteq [T]$ of length L,

$$\mathcal{R}_I^{\mathbf{1}}(\boldsymbol{u}) \doteq \mathbb{E}\left[\sum_{t\in I}\big(f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u})\big) + \sum_{t=s+1}^{e} \lambda_t\, \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right] \;\leq\; \tilde{\mathcal{O}}(\sqrt{L}) \qquad \text{(Theorem 12)}$$

uniformly in $\boldsymbol{u} \in \mathcal{X}$ and in the interval I. To our knowledge, this is the first strongly adaptive regret guarantee under indicator switching costs for OCO, and may be of independent interest.

Optimal dynamic regret for piecewise-constant comparators. A standard reduction along the time segmentation of $\{\boldsymbol{u}_t\}_{t=1}^T$ converts the strongly adaptive bound above into a dynamic regret guarantee:

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}\right] \;\leq\; \tilde{\mathcal{O}}\big(\sqrt{(S_T+1)\,T}\big), \qquad \text{(Corollary 13)}$$

without prior knowledge of S_T. This matches, up to logarithmic factors, the minimax rate for tracking moving comparators [22], despite the learner incurring an additional cost for its indicator switches.

Path-length refinement and a unified bound. The same algorithm is adaptive to path measure:

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}\big] \;\leq\; \tilde{\mathcal{O}}\Big(\max\{P_T^{1/3}T^{2/3},\, \sqrt{T}\}\Big). \qquad \text{(Theorem 14)}$$

Combining the two yields the unified guarantee

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}\big] \;\leq\; \tilde{\mathcal{O}}\Big(\min\Big\{\sqrt{(S_T+1)\,T},\;\; (P_T+1)^{1/3}\,T^{2/3}\Big\}\Big). \qquad (1)$$

Limitation. In norm-based SOCO, the bound $\mathcal{O}(\sqrt{T(1+P_T)})$ is attainable on the larger class $\mathcal{U}_P$, and since $\mathcal{U}_S(s) \subseteq \mathcal{U}_P(Ds)$, that single bound automatically yields the optimal $\mathcal{O}(\sqrt{T(1+S_T)})$ on $\mathcal{U}_S$. Under indicator switching costs we do not preserve this unification: we recover $\tilde{\mathcal{O}}(\sqrt{T(S_T+1)})$ on $\mathcal{U}_S$, but on $\mathcal{U}_P$ our bound degrades to $\tilde{\mathcal{O}}(T^{2/3}P_T^{1/3})$. Whether $\tilde{\mathcal{O}}(\sqrt{T(1+P_T)})$ is achievable under indicator switching remains open. The difficulty is that the learner is charged per change, whereas the comparator is charged only by total movement P_T.

[3]These works use different names for essentially the same algorithm: FTRL with perturbations and a barrier; see the Addendum section of [19].

1.2 Overview of the approach

To streamline presentation, we sketch here the algorithmic structure and analysis strategy. The algorithm has two layers: a base layer of restarted lazy randomized FTRL learners, one per dyadic scale, and a meta layer that aggregates them based on their expected loss and decision change.

Base layer. For each dyadic scale $H \in \mathcal{H} \doteq \{1, 2, 4, \dots, 2^{\lfloor \log_2 T \rfloor}\}$, we run a fresh copy of FPRLL that restarts at the beginning of every block of length H. Each copy maintains a probability density $\mathcal{Q}_t^{(H)}$ over $\mathcal{X}$, induced by a perturbed FTRL objective. Small scales adapt quickly but restart, and hence switch, often; large scales are stable but slow to respond to non-stationarity. The base layer thus spans the full tradeoff between responsiveness and switching control.

Meta layer: At each round t, the master maintains weights $\boldsymbol{v}_t \in \Delta_{\mathcal{H}}$ and forms the mixture density

$$\mathcal{P}_t \doteq \sum_{H \in \mathcal{H}} v_{t,H}\, \mathcal{Q}_t^{(H)}. \tag{2}$$

Action $\boldsymbol{x}_t$ is sampled from $\mathcal{P}_t$ via maximal coupling with $\mathcal{P}_{t-1}$, so that $\Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) = \|\mathcal{P}_t - \mathcal{P}_{t-1}\|_{\mathrm{TV}}$, where $\|\cdot\|_{\mathrm{TV}}$ is total variation distance. After observing f_t, the master scores each scale H by the surrogate loss $g_t^{(H)} \doteq \mathbb{E}_{\boldsymbol{x} \sim \mathcal{Q}_t^{(H)}}[f_t(\boldsymbol{x})] + \lambda \|\mathcal{Q}_t^{(H)} - \mathcal{Q}_{t-1}^{(H)}\|_{\mathrm{TV}}$, which charges each scale for both its expected loss and the switching cost it would have incurred had the learner followed it alone.

We learn the mixing weights using the discounted-normal predictor of [9]. It achieves strongly adaptive regret, hence competing with every sub-interval, while also controlling the number of transitions between experts within it. In our setting, this ensures the master performs nearly as the best restart scale for that interval, without paying an uncontrolled cost for switching between scales.

Analytically, at the base layer, we fix a dyadic scale H and characterize the dynamic regret of FPRLL restarted every H rounds (Theorem 7). The resulting bound exposes the tradeoff between the cost of restarts and the cost of comparator variation. At the meta layer, a total-variation decomposition of the combined actions, and the surrogate loss design, enable a reduction to the "H-experts $\lambda/2$-switching" problem on $\mathcal{H}$ (Lemma 10), which is handled via the discounted-normal predictor to obtain a strongly adaptive guarantee at the meta level. Combining the two layers yields our two branches in (1): for $\mathcal{U}_S$, we partition the horizon into piecewise-constant segments, cover each by its dyadic decomposition, and sum the base regret against the segment's static comparator, giving $\tilde{\mathcal{O}}(\sqrt{(S_T+1)T})$. For $\mathcal{U}_P$, sparsity is uninformative, but some dyadic scale in $\mathcal{H}$ always matches the local rate of comparator drift; the master tracks it, yielding $\tilde{\mathcal{O}}(T^{2/3} P_T^{1/3})$.

2 The FPRLL base learner

Setup. Let $\mathcal{X} \doteq \{\boldsymbol{x} \in \mathbb{R}^d : s_c(\boldsymbol{x}) \geq 0,\ c = 1, \dots, C\}$ with each s_c concave (so $\mathcal{X}$ is convex). Without loss of generality, $\mathbf{0}$ is strictly feasible and $s_c(\mathbf{0}) = 1$. For $\gamma \in (0, 1)$, we define the shrunk domain $\mathcal{X}^\gamma \doteq (1-\gamma)\mathcal{X}$ and the scaled log-barrier:

$$b_\gamma(\boldsymbol{x}) \doteq \frac{b(\boldsymbol{x})}{M_\gamma}, \qquad b(\boldsymbol{x}) \doteq \sum_{c=1}^{C} \log \frac{\bar{s}_c}{s_c(\boldsymbol{x})}, \qquad \bar{s}_c \doteq \sup_{\boldsymbol{z} \in \mathcal{X}} s_c(\boldsymbol{z}), \qquad M_\gamma \doteq \max\Big\{1,\ \sup_{\boldsymbol{z} \in \mathcal{X}^\gamma} b(\boldsymbol{z})\Big\}.$$

This is the standard logarithmic barrier, shifted to make $b \geq 0$ on $\mathcal{X}$ and scaled so that $b_\gamma \leq 1$ on $\mathcal{X}^\gamma$, as in [19]. The barrier confines learner actions to the interior; comparators remain unrestricted in $\mathcal{X}$, and the shrinkage error will be accounted for explicitly. We let $G_c \doteq \sup_{\boldsymbol{x} \in \mathcal{X}^\gamma} \|\nabla s_c(\boldsymbol{x})\|, \forall c$. Next, we define the σ-strongly convex function F_t:

$$F_t \doteq \sum_{\tau=0}^{t} f_\tau, \quad \text{with} \quad f_0 \doteq r + b_\gamma, \quad r(\boldsymbol{x}) \doteq \tfrac{\sigma}{2}\|\boldsymbol{x}\|^2. \tag{3}$$

Minimizers and their densities. Let $\boldsymbol{p} \sim \mathrm{Lap}(\mu)$ have density $\nu(\boldsymbol{p}) = (2\mu)^{-d} \exp(-\|\boldsymbol{p}\|_1/\mu)$. The *fresh-perturbation minimizer* at round t is

$$\tilde{\boldsymbol{x}}_{t+1} \doteq \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} \big\{F_t(\boldsymbol{x}) + \langle \boldsymbol{p}_t, \boldsymbol{x} \rangle\big\}, \tag{4}$$

Algorithm 1 Follow the Perturbed Regularized Lazy Leader (FPRLL)

Require: horizon T, domain $\mathcal{X}$, parameters $\gamma, \sigma, \mu > 0$
Ensure: actions $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$
1: Sample $\boldsymbol{p}_0 \sim \mathrm{Lap}(\mu)$; set $\boldsymbol{x}_1 \leftarrow \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{X}}\{f_0(\boldsymbol{x}) + \langle \boldsymbol{p}_0, \boldsymbol{x}\rangle\}$
2: **for** $t = 1, \dots, T$ **do**
3: Play $\boldsymbol{x}_t$; observe f_t and incur loss $f_t(\boldsymbol{x}_t)$
4: Evaluate $\mathcal{Q}_t(\boldsymbol{x}_t)$ and $\mathcal{Q}_{t+1}(\boldsymbol{x}_t)$ via (5); sample $z \sim \mathrm{Unif}[0,\ \mathcal{Q}_t(\boldsymbol{x}_t)]$
5: **if** $z < \mathcal{Q}_{t+1}(\boldsymbol{x}_t)$ **then**
6: $\boldsymbol{x}_{t+1} \leftarrow \boldsymbol{x}_t$ ▷ retain: switching probability $= 0$ here
7: **else**
8: **repeat** ▷ sample from residual of $\mathcal{Q}_{t+1}$
9: Sample $\boldsymbol{p} \sim \mathrm{Lap}(\mu)$; solve $\boldsymbol{y} \leftarrow \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{X}}\{F_t(\boldsymbol{x}) + \langle \boldsymbol{p}, \boldsymbol{x}\rangle\}$
10: Evaluate $\mathcal{Q}_{t+1}(\boldsymbol{y})$ and $\mathcal{Q}_t(\boldsymbol{y})$ via (5); sample $z' \sim \mathrm{Unif}[0,\ \mathcal{Q}_{t+1}(\boldsymbol{y})]$
11: **until** $z' > \mathcal{Q}_t(\boldsymbol{y})$
12: $\boldsymbol{x}_{t+1} \leftarrow \boldsymbol{y}$
13: **end if**
14: **end for**

with induced distribution $\mathcal{Q}_{t+1}$. For every perturbation, the barrier ensures an interior minimizer of the perturbed objective, and strong convexity makes it unique. Thus $\boldsymbol{x} \mapsto -\nabla F_t(\boldsymbol{x})$ is a bijection from $\mathrm{int}(\mathcal{X})$ onto $\mathbb{R}^d$, and the change-of-variables formula (Lemma 15) gives

$$\mathcal{Q}_t(\boldsymbol{x}) = \nu\big(-\nabla F_{t-1}(\boldsymbol{x})\big)\,\big|\det\big(-\nabla^2 F_{t-1}(\boldsymbol{x})\big)\big|, \qquad \boldsymbol{x} \in \mathrm{int}(\mathcal{X}). \tag{5}$$

The closed form lets us evaluate $\mathcal{Q}_t$ *pointwise*, which the lazy coupling below exploits directly.

Two perspectives. Algorithm 1 can be deployed in two ways. As a standalone algorithm, it lazy-samples an action $\boldsymbol{x}_t \sim \mathcal{Q}_t$ at each round (lines 5–12), *maximally coupling* consecutive draws so that $\Pr(\boldsymbol{x}_{t+1} \neq \boldsymbol{x}_t) = \|\mathcal{Q}_t - \mathcal{Q}_{t+1}\|_{\mathrm{TV}}$. Alternatively, the meta-algorithm of Section 3 consumes Algorithm 1 as a *density oracle*: it queries $\mathcal{Q}_t$ without invoking the sampler, and lazy-samples instead at the mixture level. Either way, the analysis below targets the same per-round quantity,

$$\mathbb{E}_{\boldsymbol{x}\sim\mathcal{Q}_t}[f_t(\boldsymbol{x})] \ + \ \lambda\|\mathcal{Q}_t - \mathcal{Q}_{t-1}\|_{\mathrm{TV}},$$

This upper-bounds the expected loss plus switching cost when the algorithm is run standalone, and is the surrogate score charged to each base learner by the meta-algorithm.

Regret guarantees. The following statements characterize the performance of FPRLL and its restarted variants. The first statement describes dynamic-regret of FPRLL, explicit in all tunable parameters (σ, μ) and the comparator path budget P_T.

Theorem 5 (Dynamic regret of FPRLL)**.** *If* FPRLL *is run with* $\gamma = \frac{1}{\sqrt{T}}$*, then for any* $\sigma, \mu > 0$*,*

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}\right] \le \frac{G^2}{2\sigma}T + R\sqrt{2d}\,\mu + (TG + \sigma R)P_T + \big(CG_cP_T + GR\big)\sqrt{T} + \lambda T\Big(\frac{\beta d}{\sigma} + \frac{\sqrt{d}\,G}{\mu}\Big) + \frac{\sigma}{2}R^2 + 1.$$

In particular, choosing $\mu^\star = \sqrt{\frac{\lambda G T}{R\sqrt{2}}}$, $\qquad \sigma^\star = \sqrt{\frac{(G^2 + 2\lambda\beta d)\,T}{R^2}}$*, yields*

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}\right] \le (R + P_T)\sqrt{(G^2 + 2\lambda\beta d)\,T} + 3\sqrt{\lambda d R G T} + TG\,P_T + \big(CG_cP_T + GR\big)\sqrt{T} + 1.$$

Specializing the above result to a *fixed* comparator on a length-k block and optimizing (γ, σ, μ) accordingly yields an $\mathcal{O}(C_{\mathrm{base}}\sqrt{k})$ static-regret bound with indicator switching cost, which is a per-block primitive used throughout the multiscale analysis.

Corollary 6 (Static regret on a block)**.** *Consider* FPRLL*, and let* $I = \{(j-1)k + 1, \dots, jk\}$ *be any full restart block. On this block, the algorithm is restarted and run with* $\gamma = \frac{1}{\sqrt{k}}, \mu_k^\star = \sqrt{\frac{\lambda G k}{R\sqrt{2}}}, \sigma_k^\star = \frac{\sqrt{(G^2 + 2\lambda\beta d)\,k}}{R}$*. Then, for every fixed comparator* $\boldsymbol{u} \in \mathcal{X}$*,*

$$\begin{aligned}\mathbb{E}\left[\mathcal{R}_I^{\mathbf{1}}(\boldsymbol{u})\right] &\doteq \mathbb{E}\Big[\sum_{t\in I}\big(f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u})\big) + \lambda \sum_{t=(j-1)k+1}^{jk} \mathbf{1}\{x_t \neq x_{t-1}\}\Big] \\ &\le R\sqrt{(G^2 + 2\lambda\beta d)\,k} + 3\sqrt{\lambda d R G\,k} + GR\sqrt{k} + \lambda + 1 \le C_{\mathrm{base}}\sqrt{k},\end{aligned}$$

where $C_{\mathrm{base}} \doteq R\sqrt{G^2+2\lambda\beta d}+3\sqrt{\lambda d R G}+GR+\lambda+1$.

For each k, let FPRLL(k) denote the algorithm that restarts FPRLL every k rounds, and let $\mathcal{R}^{\mathbf{1},k} \doteq \mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right]$ be its expected dynamic regret against $\{\boldsymbol{u}_t\}_{t=1}^T$. The following theorem shows that, among the dyadic restart periods $k \in \mathcal{H}$, there always exists one that achieves a balance between laziness cost and tracking cost.

Theorem 7 (A dyadic restart scale)**.** *Assume for simplicity that T is a power of two, and let $\mathcal{H} \doteq \{1,2,4,\dots,T\}$. For each $k \in \mathcal{H}$, FPRLL(k) uses within each block the tuning $\gamma = \frac{1}{\sqrt{k}}$, $\mu_k^\star = \sqrt{\frac{\lambda G k}{R\sqrt{2}}}$, $\sigma_k^\star = \frac{\sqrt{(G^2+2\lambda\beta d)\,k}}{R}$. Define the constants $A \doteq C_{\mathrm{base}}+\lambda$, $B \doteq \sqrt{G^2+2\lambda\beta d}+CG_c$. Then there exists a dyadic restart scale $k^\dagger \in \mathcal{H}$ such that*

$$\mathcal{R}_T^{\mathbf{1},k^\dagger} \;\le\; \sqrt{2}\min_{1\le k\le T}\left(A\frac{T}{\sqrt{k}}+(B+G)P_T\,k\right).$$

In particular, for $P_T = \mathcal{O}(T^\alpha)$, $\alpha \in [0,1)$, $\mathcal{R}^{\mathbf{1},k^\dagger} = \mathcal{O}\big(T^{2/3}P_T^{1/3}\big)$.

We complement Theorem 7 with a matching lower bound, up to logarithmic factors, for restarted randomized lazy learners of the form in Algorithm 1. The lower bound exposes a tradeoff in the restart period k between tracking $(f_t(\boldsymbol{x}_t)-f_t(\boldsymbol{u}_t))$, and laziness costs $\mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}$.

Tracking cost. A long restart period prevents the learner from reacting to comparator switches that occur inside a block. The next theorem captures this through a single oblivious distribution that is hard simultaneously for all k.

Theorem 8 (Tracking lower bound)**.** *Consider online linear optimization on $\mathcal{X} = [-1,1]$ with losses $f_t(x) = g_t x$, $g_t \in \{-1,+1\}$. Let $\tau \ge 1$ and $T \ge 4\tau$. There exists a distribution $\mathcal{D}$ over oblivious loss sequences such that, for every integer $k \in [4,\, T/\tau]$, FPRLL(k) run against $\mathcal{D}$ satisfies*

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}(u_{1:T})\big] \;=\; \Omega\left(\frac{k\tau}{\log(T/\tau)}\right)$$

for the associated comparator sequence with $S_T(u_{1:T}) = \tau$ and $P_T(u_{1:T}) = 2\tau$.

Thus the tracking term is $\Omega(k\tau/\log(T/\tau))$, which worsens with k.

Laziness cost. A short restart period forces repeated starts from scratch. Adapting [19, Thm. 4], on a length-k block, each algorithm that makes $O(\sqrt{k})$ expected switches must incur $\Omega(\sqrt{k})$ expected regret against a fixed comparator. Replaying this hard instance over the T/k blocks gives $\Omega(T/\sqrt{k})$ regret against a stationary comparator with $S_T = P_T = 0$; see Lemma 19. This term improves with k.

For every fixed k, mixing the tracking adversary with the laziness adversary tuned to that scale gives

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}\big] \;=\; \Omega\left(\frac{k\tau}{\log(T/\tau)}+\frac{T}{\sqrt{k}}\right). \qquad \text{(Proposition 20)}$$

3 Meta-learning framework

We now construct a meta-learner on top of a number of restarted base experts. First, we invoke a geometric cover fact that controls how intervals decompose across dyadic scales. Second, we model each scale $H \in \mathcal{H}$ as an expert that restarts every H rounds. Third, we attach to each expert a surrogate loss that combines its expected loss with its switching cost, aggregate the experts into a mixture $\mathcal{P}_t$, and sample via lazy coupling of the mixtures, reducing the master to online linear optimization on the simplex with an ℓ_1 switching penalty. Finally, we solve this reduced problem with the strongly adaptive algorithm of [9].

Dyadic schedule and a geometric cover. For $j = 0,1,\dots,\lfloor\log_2 T\rfloor$, define the family of level-j dyadic intervals $\mathcal{G}_j \doteq \big\{[(i-1)2^j+1,\, i2^j] : i \in \mathbb{N},\; i2^j \le T\big\}$, and collect

$$\mathcal{G} \doteq \bigcup_{j=0}^{\lfloor\log_2 T\rfloor}\mathcal{G}_j, \qquad \mathcal{H} \doteq \{2^j : 0 \le j \le \lfloor\log_2 T\rfloor\}, \qquad K \doteq |\mathcal{H}| = \lfloor\log_2 T\rfloor+1.$$

Thus $\mathcal{G}$ is the set of all dyadic intervals contained in $[T]$, and $\mathcal{H}$ is the set of dyadic *lengths*. The following standard combinatorial fact shows that every interval in $[T]$ admits a short dyadic cover whose lengths sum to a term of the order $\sqrt{L}$.

Lemma 9 (Dyadic geometric cover)**.** *For every interval $I = [s, e] \subseteq [T]$ of length $L = e - s + 1$, there exist consecutive dyadic intervals $J_1, \dots, J_m \in \mathcal{G}$ partitioning I such that*

$$m \le 2\lceil \log_2 L \rceil + 2 \qquad \textit{and} \qquad \sum_{r=1}^{m} \sqrt{|J_r|} \le C_{\mathrm{gc}} \sqrt{L}, \qquad C_{\mathrm{gc}} \doteq 2 + \sqrt{2}. \tag{6}$$

Restarted FPRLL base at each dyadic scale. For each dyadic scale $H \in \mathcal{H}$, we define a base expert $\mathcal{E}_H$ that runs a fresh FPRLL(H) instance on every consecutive block of length H, restarting at the beginning of each new block (Corollary 6). Let $\mathcal{Q}_t^{(H)}$ denote the marginal law of $\mathcal{E}_H$'s action at round t. We attach to each expert its instantaneous expected loss, TV movement, and surrogate loss:

$$\ell_t^{(H)} \doteq \mathbb{E}_{\boldsymbol{x} \sim \mathcal{Q}_t^{(H)}}[f_t(\boldsymbol{x})], \qquad c_t^{(H)} \doteq \|\mathcal{Q}_t^{(H)} - \mathcal{Q}_{t-1}^{(H)}\|_{\mathrm{TV}}, \qquad g_t^{(H)} \doteq \ell_t^{(H)} + \lambda\, c_t^{(H)}, \tag{7}$$

with the convention $c_1^{(H)} \doteq 0$. We state the main analysis using the exact quantities $\ell_t^{(H)}$ and $c_t^{(H)}$ for clarity. In practice, these quantities need not be available in closed form; a practical implementation can therefore feed the meta-learner a sampled surrogate whose conditional mean is an easily computable upper bound on $g_t^{(H)}$. This detail is deferred to Section 5.1.

The surrogate $g_t^{(H)}$ is the natural expert loss for our problem: it already prices in the switching cost that the expert would incur if the master followed it exclusively. Indeed, if actions were sampled directly from $\mathcal{Q}_t^{(H)}$ under the lazy coupling of Lemma 16, then $\Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) = c_t^{(H)}$, and therefore for every comparator $\{\boldsymbol{u}_t\}_{t=1}^T \in \mathcal{X}^T$,

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) + \lambda \sum_{t=2}^{T} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right] - \sum_{t=1}^{T} f_t(\boldsymbol{u}_t) = \sum_{t=1}^{T} \big(g_t^{(H)} - f_t(\boldsymbol{u}_t)\big). \tag{8}$$

Hence a meta-learner that competes with the best cumulative surrogate automatically inherits both the prediction *and* the switching guarantees of the corresponding expert.

Density-mixture master and the reduction. At round t, the meta-learner plays a weight vector $\boldsymbol{v}_t \in \Delta_K \doteq \left\{\boldsymbol{v} \in \mathbb{R}_+^{\mathcal{H}} : \sum_{H \in \mathcal{H}} v_H = 1\right\}$, and forms the mixture marginal as in Eq. (2) The played action $\boldsymbol{x}_t$ is then drawn from $\mathcal{P}_t$ via the lazy coupling with $\mathcal{P}_{t-1}$, which guarantees $\boldsymbol{x}_t \sim \mathcal{P}_t$ and $\Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) = \|\mathcal{P}_t - \mathcal{P}_{t-1}\|_{\mathrm{TV}}$. The full procedure is summarized in Algorithm 2 below.

Algorithm 2 Density-mixture master (instantiation of the reduction)

Require: base experts $\{\mathcal{E}_H\}_{H \in \mathcal{H}}$; meta-learner $\mathcal{M}$ satisfying Lemma 26 with $N = K$ experts and switching cost $D = \lambda/M$
Ensure: actions $\boldsymbol{x}_1, \dots, \boldsymbol{x}_T$

```
Initialize M and obtain v_1 ∈ Δ_K
Set P_1 ← Σ_{H∈H} v_{1,H} Q_1^{(H)}; sample x_1 ~ P_1
for t = 1, 2, ..., T do
    Play x_t, observe f_t
    For each H ∈ H: update E_H to obtain Q_{t+1}^{(H)}        ▷ restarted FPRLL(H)
    Form surrogate losses g_t^{(H)} = E_{Q_t^{(H)}}[f_t] + λ ||Q_t^{(H)} − Q_{t−1}^{(H)}||_TV
    Feed ℓ_t ≐ g_t/M ∈ [0,1]^K to M; receive v_{t+1} ∈ Δ_K    ▷ Algorithm 6
    Form mixture P_{t+1} ← Σ_{H∈H} v_{t+1,H} Q_{t+1}^{(H)}
    x_{t+1} ← LazySample(x_t, P_t, P_{t+1})                  ▷ Algorithm 3
end for
```

The next lemma is the basis of the analysis. It decomposes the switching probability of the master into an *expert-movement* term (the average TV movement of the dyadic coordinates) and a *weight-movement* term (the ℓ_1 movement of the meta weights).

Lemma 10 (One-step master reduction)**.** *Let* $\boldsymbol{g}_t \doteq (g_t^{(H)})_{H\in\mathcal{H}}$. *For every* $t \geq 2$,

$$\|\mathcal{P}_t - \mathcal{P}_{t-1}\|_{\mathrm{TV}} \;\leq\; \sum_{H\in\mathcal{H}} v_{t,H}\, c_t^{(H)} + \tfrac{1}{2}\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1, \tag{9}$$

and consequently $$\mathbb{E}[f_t(\boldsymbol{x}_t)] + \lambda \Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) \;\leq\; \langle \boldsymbol{v}_t, \boldsymbol{g}_t\rangle + \tfrac{\lambda}{2}\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1. \tag{10}$$

The interpretation of Eq. (10) is immediate: the meta-learner sees expert losses $\boldsymbol{g}_t$ that already price in each expert's own TV movement, and pays only for its own weight movement in ℓ_1. We have therefore reduced the master problem to an *experts-with-switching-costs* problem over the $K = \mathcal{O}(\log T)$ dyadic scales, with switching-cost parameter $\lambda/2$.

Strongly adaptive guarantee for the meta-learner. To realize the reduction, we feed the surrogate losses $\{\boldsymbol{g}_t\}$ into the strongly adaptive algorithm of [9], which competes with the best expert on *every* interval while controlling switches among experts. Its guarantees require a bounded surrogate range, which is ensured by Assumption 4 and Assumption 3. Hence, $g_t^{(H)} \leq M_f + \lambda \doteq M$.

Lemma 11 (Meta-regret)**.** *There exists an online algorithm producing weights* $\boldsymbol{v}_t \in \Delta_K$ *such that for every interval* $[s, e] \subseteq [T]$ *of length* $L = e - s + 1$, *there is a universal constant* $C_{\mathrm{dm}} > 0$ *for which:*

$$\sum_{t=s}^{e} \langle \boldsymbol{v}_t, \boldsymbol{g}_t\rangle + \tfrac{\lambda}{2} \sum_{t=s+1}^{e} \|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 \;\leq\; \min_{H\in\mathcal{H}} \sum_{t=s}^{e} g_t^{(H)} + C_{\mathrm{dm}}\sqrt{M(M+\lambda)\, L \log(KT)}, \tag{11}$$

4 Regret guarantees

We now combine the base-learner analysis of Section 2 with the meta-learner of Section 3 to derive our main regret bounds. The argument exploits the meta-learner's strongly adaptive property in two distinct ways. First, pairing it with the *static*-regret guarantee of the base learner on each restart block yields a strongly adaptive regret (SAR) bound, which in turn implies a dynamic-regret bound against piecewise-constant comparators. Second, pairing it with the *dynamic*-regret guarantee of the base learner at the best dyadic restart scale yields a path-length-based bound against arbitrary comparators.

4.1 Piecewise-constant comparators via strongly adaptive regret

The interval regret with indicator switching penalty on $I = [s, e] \subseteq [T]$ is

$$\mathrm{SAR}_I(\boldsymbol{u}) \doteq \mathbb{E}\left[\sum_{t=s}^{e} f_t(\boldsymbol{x}_t) + \lambda \sum_{t=s+1}^{e} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right] - \sum_{t=s}^{e} f_t(\boldsymbol{u}). \tag{12}$$

Algorithm 2 attains $\mathrm{SAR}_I(\boldsymbol{u}) = \tilde{\mathcal{O}}(\sqrt{L})$ *simultaneously* on all intervals of length L.

Theorem 12 (Strongly adaptive regret)**.** *For every interval* $I = [s, e] \subseteq [T]$ *of length* $L = e - s + 1$,

$$\mathrm{SAR}_I(\boldsymbol{u}) \leq C_{\mathrm{base}} C_{\mathrm{gc}}\sqrt{L} + C_{\mathrm{dm}}\, C_{\mathrm{gc}}\sqrt{M(M+\lambda)L\log(KT)} + \lambda\big(2\lceil \log_2 L\rceil + 1\big) = \tilde{\mathcal{O}}(\sqrt{L}). \tag{13}$$

Proof sketch. The bound combines three ingredients already established: **(i)** *Dyadic cover*: Lemma 9 partitions I into $m = \mathcal{O}(\log L)$ consecutive dyadic blocks $J_1, \ldots, J_m \in \mathcal{G}$ with $\sum_{r=1}^{m}\sqrt{|J_r|} \leq C_{\mathrm{gc}}\sqrt{L}$. **(ii)** *Meta-level interval regret*: applied blockwise, Lemma 11 ensures that on each J_r the meta-learner competes with the unique dyadic expert whose restart period equals $|J_r|$, at an $\tilde{\mathcal{O}}(\sqrt{|J_r|})$ cost. **(iii)** *Base-level static regret*: on its matched restart block, that expert is by definition a fresh run of FPRLL on a horizon of length $|J_r|$, which by Corollary 6 has static regret at most $C_{\mathrm{base}}\sqrt{|J_r|}$.

The one-step bound of Lemma 10 lifts the meta inner-product bound to the master cost, and boundary movement of $\boldsymbol{v}_t$ across blocks is controlled by the simplex bound $\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 \leq 2$. Summing (i)–(iii) yields Eq. (13); the full argument is given in Section E.4. □

A SAR-to-dynamic-regret reduction based on segmenting the comparator at its switch points (paying a non-dominant λS_T term) yields the piecewise-constant branch of our guarantee (see Section E.5):

Corollary 13 (Dynamic regret, S_T branch)**.** *For every comparator sequence, Algorithm 2 attains*

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\big] \;\leq\; \tilde{\mathcal{O}}\left(\sqrt{(S_T+1)\,T}\right) + \lambda\, S_T.$$

4.2 Path-varying comparators

For comparators with bounded path length, segmenting at switch points is no longer informative since $\boldsymbol{u}_t$ may drift by arbitrarily small amounts at every round. Instead, we invoke the base learner's dynamic-regret characterization (Theorem 7), which identifies the optimal dyadic restart scale $k^\dagger \in \mathcal{H}$ matching the comparator's drift. For the meta-learner notation, this is the expert with scale $H = k^\dagger$.

Theorem 14 (Dynamic regret, P_T branch)**.** *For every comparator sequence, Algorithm 2 attains*

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\big] \;\le\; \min_{H\in\mathcal{H}} \left(\sum_{t=1}^T g_t^{(H)} - \sum_{t=1}^T f_t(\boldsymbol{u}_t)\right) \;+\; C_{\mathrm{dm}}\sqrt{M(M+\lambda)\,T\log(KT)}. \quad (14)$$

In particular, for $H = k^\dagger$ from Theorem 7, we obtain

$$\mathbb{E}\big[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\big] \;=\; \tilde{\mathcal{O}}\left(\max\left\{T^{2/3}P_T^{1/3},\,\sqrt{T}\right\}\right). \quad (15)$$

Proof sketch. The one-step bound (Lemma 10), summed over $t = 1, \dots, T$, reduces the master cost to the meta inner-product loss plus a TV movement term. Applying Lemma 11 on the full interval $[1, T]$ yields Eq. (14). Substituting $H = k^\dagger$ and invoking the base-level dynamic regret bound of Theorem 7 gives Eq. (15). See Section E.6 for details. □

5 Implementation and computational cost

The exact surrogates used above involve expectations and total-variation distances that need not have closed forms. In this section, we give the more efficient sampled implementation of Algorithm 2 and account for its computational cost.

5.1 Computable surrogate losses

For each scale H, precompute$\epsilon_H \doteq \beta d/\sigma_H^\star + \sqrt{d}\,G/\mu_H^\star$, using the tuning of Corollary 6 for $\sigma_H^\star$ and $\mu_H^\star$. By Lemma 24, this bounds the expert's TV movement within a restart block. Then, define the deterministic switching bound

$$\hat{c}_t^{(H)} \doteq \begin{cases} 0, & t = 1, \\ \min\{\epsilon_H, 1\}, & t > 1 \text{ and } t-1, t \text{ are in the same block}, \\ 1, & t \in \{H+1, 2H+1, \dots\}. \end{cases} \quad (16)$$

After observing f_t, draw a fresh auxiliary sample $\boldsymbol{y}_t^{(H)} \sim \mathcal{Q}_t^{(H)}$ for each scale, using randomness independent across rounds and scales and of the randomness used to sample the played actions. Set

$$\hat{\ell}_t^{(H)} \doteq f_t(\boldsymbol{y}_t^{(H)}), \qquad \hat{g}_t^{(H)} \doteq \hat{\ell}_t^{(H)} + \lambda\hat{c}_t^{(H)}, \quad (17)$$

and feed $\hat{\boldsymbol{g}}_t/M$ to the meta-learner in place of $\boldsymbol{g}_t/M$ in Algorithm 2. This loss needs one sample and one loss evaluation per scale, and no TV computation.

Let $\mathcal{F}_t^-$ denote the history just before these auxiliary samples are drawn. Then

$$\mathbb{E}[\hat{g}_t^{(H)} \mid \mathcal{F}_t^-] = \bar{g}_t^{(H)} \doteq \ell_t^{(H)} + \lambda\hat{c}_t^{(H)} \ge g_t^{(H)}, \qquad 0 \le \hat{g}_t^{(H)} \le M. \quad (18)$$

Thus the sampled score estimates an upper bound on the surrogate. Its switching term is covered by the bounds already charged in the base-learner analysis, including the λ cost of a restart. Consequently, the strongly adaptive and both dynamic-regret rates in Section 4 hold in expectation for this implementation. For this implementation, Eq. (14) holds with $g_t^{(H)}$ replaced by $\bar{g}_t^{(H)}$ on its right-hand side, so the minimum is over the experts' cumulative mean scores. Section F proves this inequality and shows how the base-learner bounds yield the stated strongly adaptive and dynamic-regret rates.

5.2 Per-round complexity

Our framework adds polylogarithmic expected overhead to the FPRLL base learners. Each base learner supports two operations: (i) drawing a sample from its distribution and (ii) evaluating its density at a point [19]. With $K = \mathcal{O}(\log T)$ experts, each round has three costs.

Updating the master. The discounted-normal predictor updates the expert weights in $\mathcal{O}(K \log T) = \mathcal{O}(\log^2 T)$ arithmetic operations [9].

Scoring the experts. Computing the sampled scores (from the previous subsection) requires one fresh sample and one loss evaluation per expert, giving K of each.

Sampling the played action. By Sherman and Koren [19, Lem. 2], Algorithm 3 uses a constant expected number of samples and pointwise density evaluations of its input distributions. Here the inputs are the mixtures $\mathcal{P}_{t-1}$ and $\mathcal{P}_t$. Sampling from $\mathcal{P}_t$ selects expert H with probability $v_{t,H}$ and draws once from $\mathcal{Q}_t^{(H)}$, requiring one base-sampling call. Evaluating either mixture density requires summing $\mathcal{O}(\log T)$ weighted expert densities. Thus generating the played action requires $\mathcal{O}(1)$ base samples and $\mathcal{O}(\log T)$ expert-density evaluations in expectation.

Section F.2 details the implementation and cost of the two base-learner operations: drawing a sample and evaluating a density at a point.

6 Related work

Our work connects two largely separate lines of research: algorithms with few learner decision *changes*, and algorithms with guarantees against *moving* comparators. Existing results typically provide one of these guarantees, but not both, for general OCO.

Lazy OCO. A line of work studies online learners constrained to switch $o(T)$ times against a *static* comparator. [5] obtained $\sqrt{dT} + dT/S$ regret with S expected switches. [13] studied this setting under the name "consistent" online optimization and used Poisson Process schedules, but obtained a weaker tradeoff in terms of T. [18] and [2] improved the [5] guarantee to $\sqrt{T} + dT/S$, and [3] sharpened this to $\sqrt{T} + \sqrt{d}T/S$ while removing smoothness assumptions via log-concave sampling. [20] obtained the same T/S-type dependence for S switches deterministically, but under a continuous norm-based switching constraint. These works typically present regret and switching count separately; we report their sum, which is equivalent via standard reductions [18, Lem. 17] and directly captures the optimal joint tradeoff. For the *strongly* convex case, [18] obtains better bounds of order $\widetilde{\mathcal{O}}(T/S^2)$. Against *adaptive* adversaries, the problem is provably harder: [7] established a $\Theta(T/\sqrt{S})$ minimax rate for an expected S switches. All of these results target a static comparator across the time horizon.

Moving comparators and switching costs. A second line studies dynamic regret under *norm-based* movement penalties $\|\boldsymbol{x}_t - \boldsymbol{x}_{t-1}\|$ or other Lipschitz switching costs [8, 30, 24, 16]. [24] achieved the optimal $\mathcal{O}(\sqrt{T(1+P_T)})$ dynamic regret by incorporating switching cost into a meta-aggregation loss. For OCO with memory, Zhao et al. [29] (and [28]) also include each expert's norm-based switching cost in the meta-loss for the different objective of bounding dynamic policy regret. Our approach is similar in spirit, but aggregates over total-variation distance and uses different meta- and base-learners. Most recently, [16] showed that FTRL itself can attain dynamic-regret guarantees with additional laziness properties such as staleness. Relatedly, [26] studies unbounded comparators, and [11] extends this setting to moving unbounded comparators with bounds adaptive to individual switching-cost parameters λ_t. However, these costs remain norm-based, and the resulting algorithms may still change decisions every round. Strongly adaptive methods provide another route to dynamic or tracking regret: [10] introduced geometric-covering reductions from static regret to strongly adaptive regret, and [23] showed that strongly adaptive regret implies dynamic-regret bounds in OCO, typically with an additional function-variation term. Zhang et al. [25] use discounted-normal prediction to obtain adaptive regret but again under norm-based switching costs.

Beyond the two lines above, our framework draws on adjacent threads. Differentially private online learning naturally incentivizes few decision changes, since each update consumes privacy budget, and hence algorithms developed in one setting often transfer to the other [2]. On the algorithmic side, Ader and SWORD aggregate grids of step sizes under Hedge or optimistic Hedge masters to adapt to unknown path length without switching costs [22, 27]. Our master/base design follows the meta-learning-over-restarting-experts line of Follow-the-Leading-History and the sleeping-experts perspective [1]. Xu and Zhang [21] use restarted FTPL and expert aggregation for dynamic and adaptive regret, without switching costs. Our contribution is the dynamic analysis of restarted FPRLL (Theorem 7) and the density-mixture/TV reduction (Lemma 10) that connects the discounted-normal master to indicator switching costs.

Acknowledgment

This work was supported by the Dutch National Growth Fund project "Future Network Services", and by the European Commission through Grants No. 101139270 "ORIGAMI" and No. 101192462 "FLECON-6G"."

References

[1] D. Adamskiy, W. M. Koolen, A. Chernov, and V. Vovk. A closer look at adaptive regret. In *Proc. of ALT*, 2012.

[2] N. Agarwal, S. Kale, K. Singh, and A. Thakurta. Differentially private and lazy online convex optimization. In *Proc. of COLT*, 2023.

[3] N. Agarwal, S. Kale, K. Singh, and A. G. Thakurta. Improved differentially private and lazy online convex optimization: lower regret without smoothness requirements. In *Proc. of ICML*, 2024.

[4] J. Altschuler and K. Talwar. Online learning over a finite action set with limited switching. In *Proc. of COLT*, 2018.

[5] O. Anava, E. Hazan, and S. Mannor. Online learning for adversaries with memory: price of past mistakes. In *Proc. of NeurIPS*, 2015.

[6] N. Cesa-Bianchi, O. Dekel, and O. Shamir. Online learning with switching costs and other adaptive adversaries. In *Proc. of NeurIPS*, 2013.

[7] L. Chen, Q. Yu, H. Lawrence, and A. Karbasi. Minimax regret of switching-constrained online convex optimization: No phase transition. In *Proc. of NeurIPS*, 2020.

[8] N. Chen, G. Goel, and A. Wierman. Smoothed online convex optimization in high dimensions via online balanced descent. In *Proc. of COLT*, 2018.

[9] A. Daniely and Y. Mansour. Competitive ratio vs regret minimization: achieving the best of both worlds. In *Proc. of ALT*, 2019.

[10] A. Daniely, A. Gonen, and S. Shalev-Shwartz. Strongly adaptive online learning. In *Proc. of ICML*, 2015.

[11] E. Esposito, A. Jacobsen, H. Qiu, and M. Zhang. Parameter-free dynamic regret: Time-varying movement costs, delayed feedback, and memory. In *Proc. of ICML*, 2026.

[12] A. Jacobsen and A. Cutkosky. Parameter-free mirror descent. In *Proc. of COLT*, 2022.

[13] M. R. K. Jaghargh, A. Krause, S. Lattanzi, and S. Vassilvtiskii. Consistent online optimization: Convex and submodular. In *Proc. of AISTATS*, 2019.

[14] H. B. McMahan. A Survey of Algorithms and Analysis for Adaptive Online Learning. *J. Mach. Learn. Res.*, 18(1):3117–3166, 2017.

[15] N. Mhaisen and G. Iosifidis. On the dynamic regret of following the regularized leader: Optimism with history pruning. In *Proc. of ICML*, 2025.

[16] N. Mhaisen and G. Iosifidis. Partially lazy gradient descent for smoothed online learning. In *Proc. of AISTATS*, 2026.

[17] F. Orabona. A Modern Introduction to Online Learning. *arXiv.1912.13213*, 2022.

[18] U. Sherman and T. Koren. Lazy oco: Online convex optimization on a switching budget. In *Proc. of COLT*, 2021.

[19] U. Sherman and T. Koren. Lazy oco: Online convex optimization on a switching budget. *arXiv 2102.03803*, 2023.

[20] G. Wang, Y. Wan, T. Yang, and L. Zhang. Online convex optimization with continuous switching constraint. In *Proc. of NeurIPS*, 2021.

[21] Z. Xu and L. Zhang. Online non-convex learning in dynamic environments. In *Proc. of NeurIPS*, pages 51930–51962, 2024.

[22] L. Zhang, S. Lu, and Z.-H. Zhou. Adaptive online learning in dynamic environments. In *Proc. of NeurIPS*, 2018.

[23] L. Zhang, T. Yang, Z.-H. Zhou, et al. Dynamic regret of strongly adaptive methods. In *International conference on machine learning*, pages 5882–5891. PMLR, 2018.

[24] L. Zhang, W. Jiang, S. Lu, and T. Yang. Revisiting smoothed online learning. In *Proc. of NeurIPS*, 2021.

[25] L. Zhang, W. Jiang, J. Yi, and T. Yang. Smoothed online convex optimization based on discounted-normal-predictor. In *Proc. of NeurIPS*, 2022.

[26] Z. Zhang, A. Cutkosky, and Y. Paschalidis. Optimal comparator adaptive online learning with switching cost. In *Proc. of NeurIPS*, 2022.

[27] P. Zhao, Y.-J. Zhang, L. Zhang, and Z.-H. Zhou. Dynamic regret of convex and smooth functions. In *Proc. of NeurIPS*, volume 33, 2020.

[28] P. Zhao, Y.-X. Wang, and Z.-H. Zhou. Non-stationary online learning with memory and non-stochastic control. In *Proc. of AISTATS*, 2022.

[29] P. Zhao, Y.-H. Yan, Y.-X. Wang, and Z.-H. Zhou. Non-stationary online learning with memory and non-stochastic control. *J. Mach. Learn. Res.*, 24(206):1–70, 2023.

[30] Y. Zhao, Q. Zhao, X. Zhang, E. Zhu, X. Liu, and J. Yin. Understand dynamic regret with switching cost for online decision making. *ACM Trans. on Intelligent Systems and Technology*, 11(3), 2020.

Appendix

This appendix contains the technical details omitted from the main text. We first derive the change-of-variables formula underlying the fresh perturbed minimizer distribution and recall the lazy-sampling coupling used to preserve the correct marginals while controlling switches. We then prove the lower bounds, the base-learner guarantees, the dyadic restart and interval-covering results, and finally the meta-learner reductions leading to the strongly adaptive and dynamic-regret guarantees.

Contents.

- Section A: change of variables and lazy sampling.
- Section B: lower bounds for restart-based FPRLL.
- Section C: proofs for the base learner.
- Section D: dyadic interval covering.
- Section E: proofs for the meta-learner and reductions.
- Section F: guarantees for sampled surrogates and implementation costs of the base primitives.

A Change of variable and lazy sampling

A.1 Change of variable formula

Lemma 15 (Density of the fresh minimizer)**.** *Let*

$$\tilde{\boldsymbol{x}}_t \doteq \arg\min_{\boldsymbol{x}\in\mathcal{X}}\Big\{F_{t-1}(\boldsymbol{x}) + \langle \boldsymbol{p}_t, \boldsymbol{x}\rangle\Big\}, \qquad \boldsymbol{p}_t \sim \mathrm{Lap}(\mu),$$

where

$$\nu(\boldsymbol{p}) = \frac{1}{(2\mu)^d}\exp\left(-\frac{\|\boldsymbol{p}\|_1}{\mu}\right).$$

Then $\tilde{\boldsymbol{x}}_t$ has density

$$\mathcal{Q}_t(\boldsymbol{x}) = \nu\left(-\nabla F_{t-1}(\boldsymbol{x})\right)\left|\det\left(-\nabla^2 F_{t-1}(\boldsymbol{x})\right)\right|, \qquad \boldsymbol{x}\in\mathrm{int}(\mathcal{X}). \tag{19}$$

Proof. By optimality of $\tilde{\boldsymbol{x}}_t$,

$$\nabla F_{t-1}(\tilde{\boldsymbol{x}}_t) + \boldsymbol{p}_t = 0,$$

hence

$$\boldsymbol{p}_t = -\nabla F_{t-1}(\tilde{\boldsymbol{x}}_t).$$

Define

$$T_t(\boldsymbol{x}) \doteq -\nabla F_{t-1}(\boldsymbol{x}).$$

Since F_{t-1} is strongly convex, its conjugate F_{t-1}^* is differentiable and

$$\nabla F_{t-1}^* = (\nabla F_{t-1})^{-1}.$$

Therefore

$$T_t^{-1}(\boldsymbol{p}) = \nabla F_{t-1}^*(-\boldsymbol{p}).$$

So $\tilde{\boldsymbol{x}}_t = T_t^{-1}(\boldsymbol{p}_t)$.

Applying the change-of-variables formula to the map T_t, the density of $\tilde{\boldsymbol{x}}_t$ is

$$\mathcal{Q}_t(\boldsymbol{x}) = \nu(T_t(\boldsymbol{x}))\left|\det DT_t(\boldsymbol{x})\right|.$$

Using

$$T_t(\boldsymbol{x}) = -\nabla F_{t-1}(\boldsymbol{x}), \qquad DT_t(\boldsymbol{x}) = -\nabla^2 F_{t-1}(\boldsymbol{x}),$$

we obtain

$$\mathcal{Q}_t(\boldsymbol{x}) = \nu\left(-\nabla F_{t-1}(\boldsymbol{x})\right)\left|\det\left(-\nabla^2 F_{t-1}(\boldsymbol{x})\right)\right|,$$

as claimed. □

Algorithm 3 LAZYSAMPLE

Input: current action $\boldsymbol{x}$, current distribution $\mathcal{Q}$, next distribution $\mathcal{P}$
Output: next action $\boldsymbol{y}$
Sample $z \sim \mathrm{Unif}[0,\ \mathcal{Q}(\boldsymbol{x})]$
if $z < \mathcal{P}(\boldsymbol{x})$ **then** ▷ Keep the same action
    Set $\boldsymbol{y} = \boldsymbol{x}$
else ▷ Sample a new action
    repeat
        Sample $\boldsymbol{y} \sim \mathcal{P}$
        Sample $z' \sim \mathrm{Unif}[0,\ \mathcal{P}(\boldsymbol{y})]$
    until $z' > \mathcal{Q}(\boldsymbol{y})$
end if
return $\boldsymbol{y}$

A.2 Lazy sampling

We recall a standard lazy-sampling routine for realizing a maximal coupling between two distributions. We use it only through the marginal-preservation and switching identities in Lemma 16; the results (and their proof) are standard and can be found in [18].

Lemma 16 (Standard lazy-sampling identities)**.** *Let $\boldsymbol{X} \sim \mathcal{Q}$, and let*

$$\boldsymbol{Y} \sim \textsc{LazySample}(\boldsymbol{X}, \mathcal{Q}, \mathcal{P}).$$

Then:

$$\boldsymbol{Y} \sim \mathcal{P}, \tag{20}$$

$$\Pr(\boldsymbol{Y} \neq \boldsymbol{X}) = \|\mathcal{P} - \mathcal{Q}\|_{\mathrm{TV}}. \tag{21}$$

Correspondence with Algorithm 1. Algorithm 1 is precisely LAZYSAMPLE instantiated with the consecutive FPRLL marginals $(\mathcal{Q}_t, \mathcal{Q}_{t+1})$. In this correspondence, the current state is $\boldsymbol{x}_t$, the current and next distributions are $\mathcal{Q}_t$ and $\mathcal{Q}_{t+1}$, and the output is $\boldsymbol{x}_{t+1}$. The "keep" test in Algorithm 1 matches the first branch of Algorithm 3. If that test fails, the algorithm must draw a fresh sample from $\mathcal{Q}_{t+1}$; in our setting this is implemented by drawing $\boldsymbol{p} \sim \mathrm{Lap}(\mu)$ and solving the perturbed optimization problem, whose minimizer has law $\mathcal{Q}_{t+1}$ by definition. Finally, Lemma 15 supplies a closed-form expression for these densities, so the comparisons involving $\mathcal{Q}_t(\cdot)$ and $\mathcal{Q}_{t+1}(\cdot)$ required by lazy sampling can be evaluated pointwise.

B Lower bounds

Throughout the lower-bound analysis, we write centered losses $f_t(x) = g_t x$, with $|g_t| \leq 1$, on $[-1, 1]$ for simplicity. The corresponding admissible losses are $\bar{f}_t(x) = 1 + f_t(x) \in [0, 2]$, satisfying Assumption 4 with $M_f = 2$. This shift adds only an x-independent constant to each cumulative objective. It therefore leaves the minimizers, gradients and Hessians unchanged, and hence also the densities in (5) and the lazy-sampling probabilities. Finally, for every comparator sequence, $\bar{f}_t(x_t) - \bar{f}_t(u_t) = f_t(x_t) - f_t(u_t)$, so the regret and the resulting FPRLL lower bounds are unchanged.

First, we need the following tool, which lower bounds FPRLL dynamic regret without restarts:

Theorem 17. *Consider the one-dimensional problem on $\mathcal{X} = [-1, 1]$ with linear losses*

$$f_t(x) = g_t x, \qquad g_t \in \{-1, +1\}.$$

Let the learner predict at each round

$$x_t \doteq\!= \operatorname*{argmin}_{x \in (-1,1)} \Big\{ \langle G_{t-1} + p_t, x \rangle + r_t(x) + b(x) \Big\}, \tag{22}$$

where $G_{t-1} = G_0 + \sum_{i=1}^{t-1} g_i$, for an arbitrary $G_0 \in \mathbb{R}$,

$$b(x) = -\frac{1}{M_\gamma} \log(1 - x^2)$$

is the logarithmic barrier for $\mathcal{X}$, $r_t(x) = \frac{\sigma_t}{2}x^2$ with $\sigma_t \geq 0$, and p_t are arbitrary symmetric perturbations.

Then, there exists an oblivious loss sequence and a one-switch comparator sequence such that

$$\mathbb{E}\left[\sum_{t=1}^{T}\big(f_t(x_t) - f_t(u_t)\big)\right] \;\geq\; \frac{T}{2}.$$

Proof. Fix $T = 2n$ and consider the oblivious loss sequence

$$g_t = \begin{cases} +1, & t \leq n, \\ -1, & t > n, \end{cases} \qquad u_t = -g_t.$$

Define the cumulative gradients

$$G_t \doteq G_0 + \sum_{i=1}^{t} g_i.$$

Step 1: Structure of the prediction. For each round t, define

$$H_t(x) \doteq r_t'(x) + b'(x) = \sigma_t x + \frac{2x}{M_\gamma(1-x^2)} = x\left(\sigma_t + \frac{2}{M_\gamma(1-x^2)}\right), \qquad x \in (-1,1).$$

Since $\sigma_t \geq 0$,

$$H_t'(x) = \sigma_t + \frac{2(1+x^2)}{M_\gamma(1-x^2)^2} > 0,$$

so H_t is strictly increasing. Moreover, H_t is odd and satisfies

$$\lim_{x\to -1^+} H_t(x) = -\infty, \qquad \lim_{x\to 1^-} H_t(x) = +\infty.$$

Thus H_t is a bijection from $(-1,1)$ onto $\mathbb{R}$, and its inverse H_t^{-1} is odd and strictly increasing.

By first-order optimality of the FTRL objective,

$$H_t(x_t) = -(G_{t-1} + p_t),$$

and therefore

$$x_t = -\,H_t^{-1}(G_{t-1} + p_t).$$

Step 2: Sign of the expectation. Define

$$m_t(c) \doteq \mathbb{E}\big[H_t^{-1}(c + p_t)\big].$$

Since H_t^{-1} is increasing, m_t is increasing. Since H_t^{-1} is odd and p_t is symmetric, m_t is odd:

$$m_t(-c) = -m_t(c).$$

Hence

$$c \geq 0 \;\Rightarrow\; m_t(c) \geq 0, \qquad c \leq 0 \;\Rightarrow\; m_t(c) \leq 0.$$

Using $x_t = -H_t^{-1}(G_{t-1} + p_t)$,

$$\mathbb{E}[x_t] = -\,m_t(G_{t-1}),$$

which implies

$$G_{t-1} \geq 0 \;\Rightarrow\; \mathbb{E}[x_t] \leq 0, \qquad G_{t-1} \leq 0 \;\Rightarrow\; \mathbb{E}[x_t] \geq 0.$$

Step 3: Pairing argument. For each $k \in \{1,\dots,n\}$, pair round k with round $2n-k+1$, and define

$$R_k \doteq (f_k(x_k) - f_k(u_k)) + (f_{2n-k+1}(x_{2n-k+1}) - f_{2n-k+1}(u_{2n-k+1})).$$

Since $g_k = +1$ and $g_{2n-k+1} = -1$,

$$R_k = (1 + x_k) + (1 - x_{2n-k+1}).$$

Both terms are nonnegative since $x_t \in [-1,1]$.

The cumulative gradients before the two rounds are

$$A_k \doteq G_{k-1} = G_0 + k - 1, \qquad B_k \doteq G_{2n-k} = G_0 + k = A_k + 1.$$

Step 4: Lower bounding each pair. We distinguish two cases.

Case 1: $A_k \le 0$**.** Then $\mathbb{E}[x_k] \ge 0$, so

$$\mathbb{E}[1 + x_k] \ge 1,$$

and since $1 - x_{2n-k+1} \ge 0$,

$$\mathbb{E}[R_k] \ge 1.$$

Case 2: $A_k > 0$**.** Then $B_k > 0$, so $\mathbb{E}[x_{2n-k+1}] \le 0$, and thus

$$\mathbb{E}[1 - x_{2n-k+1}] \ge 1.$$

Since $1 + x_k \ge 0$,

$$\mathbb{E}[R_k] \ge 1.$$

In all cases, $\mathbb{E}[R_k] \ge 1$ for every k. Summing over $k = 1, \dots, n$,

$$\mathbb{E}\left[\sum_{t=1}^{T}\big(f_t(x_t) - f_t(u_t)\big)\right] = \sum_{k=1}^{n}\mathbb{E}[R_k] \ge n = \frac{T}{2}.$$

□

Note that starting with positive wave is not necessary, since what matters is the switch, as shown next.

Corollary 18. *Under the assumptions of Theorem 17, let $T = 2n$ and consider the oblivious loss sequence*

$$g_t = \begin{cases} -1, & t \le n, \\ +1, & t > n, \end{cases} \qquad u_t = -g_t.$$

Then

$$\mathbb{E}\left[\sum_{t=1}^{T}\big(f_t(x_t) - f_t(u_t)\big)\right] \;\ge\; \frac{T}{2}.$$

Proof. The proof is identical to that of Theorem 17, except that in the pairing step 3, we have

$$R_k \doteq (f_k(x_k) - f_k(u_k)) + (f_{2n-k+1}(x_{2n-k+1}) - f_{2n-k+1}(u_{2n-k+1})).$$

Since now $g_k = -1$ and $g_{2n-k+1} = +1$,

$$R_k = (1 - x_k) + (1 + x_{2n-k+1}).$$

The cumulative gradients before the two rounds are

$$A_k \doteq G_{k-1} = G_0 - (k-1), \qquad B_k \doteq G_{2n-k} = G_0 - k = A_k - 1.$$

If $A_k \ge 0$, then $\mathbb{E}\left[x_k\right] \le 0$, so $\mathbb{E}\left[1 - x_k\right] \ge 1$. If $A_k < 0$, then $B_k < 0$, so $\mathbb{E}\left[x_{2n-k+1}\right] \ge 0$, and hence $\mathbb{E}\left[1 + x_{2n-k+1}\right] \ge 1$. Since both terms are always nonnegative, in either case $\mathbb{E}\left[R_k\right] \ge 1$. Summing over $k = 1, \dots, n$ yields the claim. □

Now we are ready to prove Theorem 8:

Proof. Let

$$J \doteq \left\lfloor \log_2\left(\frac{T}{4\tau}\right)\right\rfloor, \qquad \mathcal{H}_\tau \doteq \{2^j : j = 0, \dots, J\}.$$

The adversary samples a half-gadget length h uniformly from $\mathcal{H}_\tau$. It then builds the loss sequence by concatenating τ gadgets of length $2h$, with alternating orientation: for odd $m \in \{1, \dots, \tau\}$, the m-th gadget is

$$(\underbrace{+1, \dots, +1}_{h \text{ rounds}}, \underbrace{-1, \dots, -1}_{h \text{ rounds}}),$$

and for even $m \in \{1, \dots, \tau\}$, the m-th gadget is

$$(\underbrace{-1, \dots, -1}_{h \text{ rounds}}, \underbrace{+1, \dots, +1}_{h \text{ rounds}}).$$

Since $h \le T/(4\tau)$, the total length of these τ gadgets satisfies

$$2h\tau \le T/2 \le T,$$

so the construction fits in the horizon. For the remaining rounds $t > 2h\tau$, define

$$g_t \doteq g_{2h\tau}.$$

Finally, set the comparator to oppose the losses (for each realized randomization):

$$u_t \doteq -g_t \qquad \text{for all } t = 1, \dots, T.$$

Comparator path length. Inside each gadget, the comparator switches exactly once at the midpoint. The padded rounds are constant and create no further switches. Therefore

$$P_T(u) = 2\tau.$$

Note that for every round t, since $u_t = -g_t$ and $x_t \in [-1, 1]$,

$$f_t(x_t) - f_t(u_t) = g_t x_t - g_t u_t = g_t x_t + 1 \ge 0.$$

Hence $\mathbb{E}[\mathcal{R}_T] \ge 0$ for every realization of the randomness.

Choose the dyadic scale. Fix any integer restart period k satisfying $4 \le k \le T/\tau$, and define

$$h^\star \doteq 2^{\lfloor \log_2(k/4) \rfloor}.$$

Since $1 \le k/4 \le T/(4\tau)$, we have $h^\star \in \mathcal{H}_\tau$. By construction,

$$\frac{k}{8} < h^\star \le \frac{k}{4}, \qquad 2h^\star \le \frac{k}{2} < k.$$

Thus each gadget is strictly shorter than one restart block.

Let $A \doteq \{h = h^\star\}$. Because h is drawn uniformly from $\mathcal{H}_\tau$,

$$\Pr(A) = \frac{1}{J+1}.$$

Number of uncut gadgets Condition on the event A. The active gadget region has total length

$$2h^\star\tau \le \frac{k\tau}{2}.$$

Since restart boundaries occur every k rounds, the number of restart boundaries falling inside this active region is at most

$$\left\lfloor \frac{2h^\star\tau}{k} \right\rfloor \le \frac{2h^\star\tau}{k} \le \frac{\tau}{2}.$$

Because every gadget has length $2h^\star < k$, a single restart boundary can intersect the interior of at most one gadget, and a single gadget can contain at most one restart boundary. Therefore at most $\tau/2$ gadgets are cut by a restart, and at least

$$\tau - \tau/2 = \tau/2$$

gadgets are fully contained in a single restart block.

Regret from each uncut gadget. Consider any gadget that is not cut by a restart. Over that gadget, the learner runs continuously with no reset. Reindex the gadget as a fresh $2h^\star$-round instance. The state at the beginning of the gadget simply becomes the initial gradient G_0 for that local instance.

For an odd gadget, the loss sequence is precisely the two-phase sequence used in the proof of Thm. 17. Hence that gadget contributes expected dynamic regret at least $\frac{2h^\star}{2} = h^\star$. For an even gadget, the same lower bound follows from Corollary 18.

Summing over the at least $\tau/2$ uncut gadgets yields

$$\mathbb{E}\left[\mathcal{R}_T \mid A\right] \;\geq\; \frac{\tau}{2}\, h^\star \;\geq\; \frac{k\tau}{16}.$$

We expand the unconditional expectation:

$$\mathbb{E}\left[\mathcal{R}_T\right] = \Pr(A)\; \mathbb{E}\left[\mathcal{R}_T \mid A\right] + \Pr(A^c)\; \mathbb{E}\left[\mathcal{R}_T \mid A^c\right].$$

Since $\mathbb{E}\left[\mathcal{R}_T\right] \geq 0$, both conditional expectations are nonnegative, and in particular

$$\Pr(A^c)\; \mathbb{E}\left[\mathcal{R}_T \mid A^c\right] \geq 0.$$

Therefore we may drop this second term and obtain the lower bound

$$\mathbb{E}\left[\mathcal{R}_T\right] \;\geq\; \Pr(A)\; \mathbb{E}\left[\mathcal{R}_T \mid A\right].$$

Using $\Pr(A) = 1/(J+1)$ and the conditional lower bound above gives

$$\mathbb{E}\left[\mathcal{R}_T\right] \;\geq\; \frac{k\tau}{16(J+1)}.$$

Finally, since $T/\tau \geq 4$,

$$J + 1 = 1 + \left\lfloor \log_2\left(\frac{T}{4\tau}\right) \right\rfloor = \Theta(\log(T/\tau)).$$

Hence $\mathbb{E}\left[\mathcal{R}_T\right] = \Omega\left(\frac{k\tau}{\log(T/\tau)}\right)$. This completes the proof. □

Lemma 19 (Sherman branch)**.** *Suppose that a fresh run of FPRLL(k) makes at most $C\sqrt{k}$ expected switches on each length-k block, for a constant $C > 0$. Then there exists an oblivious distribution over T-round loss sequences and a fixed comparator $u^\star \in [-1, 1]$ such that*

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u^\star, \ldots, u^\star)\right] = \Omega\left(\frac{T}{\sqrt{k}}\right).$$

Proof. We use the following consequence of Sherman and Koren [19, Thm. 4]: for any algorithm making $\mathcal{O}(\sqrt{k})$ expected switches in a length-k online linear problem on $[-1, 1]$, there exist a stochastic loss distribution $\mathcal{D}_k$ and a comparator $u^\star \in [-1, 1]$, fixed as a minimizer of the expected loss under $\mathcal{D}_k$, such that this algorithm incurs $\Omega(\sqrt{k})$ expected static regret against $u^\star$. Apply this result to a fresh run of FPRLL(k).

Construct the T-round adversary by drawing, independently on each full restart block, a fresh length-k sample from this same distribution $\mathcal{D}_k$. Although the realized losses differ from block to block, the comparator remains the same $u^\star$, since $u^\star$ is defined by $\mathcal{D}_k$, not by the realized sample. On any remaining rounds $t > k\lfloor T/k \rfloor$, set $f_t \equiv 0$.

Thus every full block contributes $\Omega(\sqrt{k})$ expected static regret. Since $\mathcal{R}_T^{\mathbf{1}}$ only adds the learner's nonnegative indicator switching cost, the same lower bound holds for $\mathcal{R}_T^{\mathbf{1}}$. Therefore

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u^\star, \ldots, u^\star)\right] \geq \left\lfloor \frac{T}{k} \right\rfloor \Omega(\sqrt{k}) = \Omega\left(\frac{T}{\sqrt{k}}\right),$$

where $k \leq T$. The comparator is stationary throughout, so $S_T = P_T = 0$. □

Under the block tuning of Corollary 6, maximal coupling and Lemma 24 bound the expected number of within-block switches by

$$(k-1)\left(\frac{\beta d}{\sigma_k^\star}+\frac{\sqrt{d}\,G}{\mu_k^\star}\right)=\mathcal{O}(\sqrt{k}).$$

Thus Lemma 19 applies.

Proposition 20 (Combined lower bound)**.** *Consider the following two full-horizon adversaries:*

1. *the adversary from Theorem 8, together with comparator sequence $u_{1:T}^{\mathrm{gad}}$, for which*

$$S_T(u_{1:T}^{\mathrm{gad}})=\tau, \qquad P_T(u_{1:T}^{\mathrm{gad}})=2\tau,$$

and

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T}^{\mathrm{gad}})\right]=\Omega\left(\frac{k\tau}{\log(T/\tau)}\right);$$

2. *the adversary from Lemma 19, together with the stationary comparator sequence $u_{1:T}^{\mathrm{sh}}\equiv u^\star$, for which*

$$S_T(u_{1:T}^{\mathrm{sh}})=0, \qquad P_T(u_{1:T}^{\mathrm{sh}})=0,$$

and

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T}^{\mathrm{sh}})\right]=\Omega\left(\frac{T}{\sqrt{k}}\right).$$

For the fixed restart period k, define a mixed adversary that, at time 0, chooses one of these two branches uniformly at random and then plays the chosen branch for the entire horizon. The resulting mixed adversary may depend on k through the Sherman branch. Let $u_{1:T}$ denote the comparator sequence associated with the realized branch. Then

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T})\right]=\Omega\left(\frac{k\tau}{\log(T/\tau)}+\frac{T}{\sqrt{k}}\right).$$

Proof. Conditioning on the branch chosen by the adversary gives

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T})\right]=\frac{1}{2}\,\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T}^{\mathrm{gad}})\right]+\frac{1}{2}\,\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T}^{\mathrm{sh}})\right].$$

Applying Theorem 8 to the first term and Lemma 19 to the second term yields the claim. □

C Proofs for the base learner

To prove Theorem 5, we establish four intermediate lemmas and then combine them.

- Lemma 21 gives the main regret decomposition, splitting the expected dynamic regret into a *static* term ($\mathbf{I}$) and a *dynamic* term ($\mathbf{II}$). This extends the decomposition in [15] to our perturbed, barrier-regularized setting.
- Lemma 22 bounds the static term ($\mathbf{I}$), explicitly accounting for the effect of the Laplace perturbations.
- Lemma 23 bounds the dynamic term ($\mathbf{II}$), capturing the dependence on comparator drift through the barrier geometry.
- Lemma 24 controls the total variation distance between consecutive sampling distributions, which we use to bound the switching-cost contribution.

Among these ingredients, only the total-variation bound in Lemma 24 follows the general approach of [18]; the other three lemmas are specific to our setting.

Lemma 21. *Let $\{f_t(\cdot)\}_{t=1}^T$ be an arbitrary set of functions. Let $\boldsymbol{p}_t \sim \mathrm{Lap}(\mu)$ for $t = 0, \dots, T-1$. Assume that for all t, the fresh perturbed minimizer*

$$\tilde{\boldsymbol{x}}_{t+1} \doteq \operatorname*{argmin}_{\boldsymbol{x}} F_t(\boldsymbol{x}) + \langle \boldsymbol{p}_t, \boldsymbol{x} \rangle$$

is well-defined. Define the loss part of $\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)$ as

$$\mathcal{R}_T(\{\boldsymbol{u}_t\}_{t=1}^T) = \sum_{t=1}^T \left(f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t)\right),$$

where $\boldsymbol{x}_t$ are the output of Algorithm 1. Then, the algorithm that selects the actions $\tilde{\boldsymbol{x}}_{t+1}, \forall t$ achieves the following dynamic regret bound:

$$\mathbb{E}\left[\mathcal{R}_T\right] \le \overbrace{\mathbb{E}\left[\sum_{t=1}^T \big(F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1})\big)\right]}^{(\mathbf{I})} + \underbrace{\sum_{t=1}^{T-1} F_t(\boldsymbol{u}_{t+1}^\gamma) - F_t(\boldsymbol{u}_t^\gamma)}_{(\mathbf{II})} + r(\boldsymbol{u}_1^\gamma) + 1 + G\gamma RT. \quad (23)$$

Proof of Lemma 21.

$$\begin{aligned}
&\sum_{t=1}^T f_t(\tilde{\boldsymbol{x}}_t) - \sum_{t=1}^T f_t(\boldsymbol{u}_t) - \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle \\
&= \sum_{t=1}^T f_t(\tilde{\boldsymbol{x}}_t) - \sum_{t=1}^T f_t(\boldsymbol{u}_t^\gamma) - \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle \; + \; \sum_{t=1}^T \big(\overbrace{f_t(\boldsymbol{u}_t^\gamma) - f_t(\boldsymbol{u}_t)}^{a_t}\big) \\
&= \sum_{t=1}^T \left(F_t(\tilde{\boldsymbol{x}}_t) - F_{t-1}(\tilde{\boldsymbol{x}}_t)\right) - \left(\sum_{t=1}^T \left(F_t(\boldsymbol{u}_t^\gamma) - F_{t-1}(\boldsymbol{u}_t^\gamma)\right) + \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle\right) \; + \; a_{1:T} \\
&= \sum_{t=1}^T F_t(\tilde{\boldsymbol{x}}_t) - \sum_{t=1}^T F_{t-1}(\tilde{\boldsymbol{x}}_t) - \left(\sum_{t=1}^T F_t(\boldsymbol{u}_t^\gamma) + \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle - \sum_{t=1}^T F_{t-1}(\boldsymbol{u}_t^\gamma)\right) + a_{1:T} \\
&= \sum_{t=1}^T F_t(\tilde{\boldsymbol{x}}_t) - \sum_{t=0}^{T-1} F_t(\tilde{\boldsymbol{x}}_{t+1}) - \left(F_T(\boldsymbol{u}_T^\gamma) + \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle + \sum_{t=1}^{T-1} F_t(\boldsymbol{u}_t^\gamma) - \sum_{t=0}^{T-1} F_t(\boldsymbol{u}_{t+1}^\gamma)\right) \\
&\quad + a_{1:T} \\
&\le \sum_{t=1}^T F_t(\tilde{\boldsymbol{x}}_t) - \sum_{t=1}^{T-1} F_t(\tilde{\boldsymbol{x}}_{t+1}) - \left(F_T(\boldsymbol{x}_{T+1}) + \langle \boldsymbol{p}_T, \boldsymbol{x}_{T+1} \rangle + \sum_{t=1}^{T-1} F_t(\boldsymbol{u}_t^\gamma) - \sum_{t=1}^{T-1} F_t(\boldsymbol{u}_{t+1}^\gamma)\right) \\
&\quad + a_{1:T} + f_0(\boldsymbol{u}_1^\gamma) \\
&\le \sum_{t=1}^T F_t(\tilde{\boldsymbol{x}}_t) - \sum_{t=1}^T F_t(\tilde{\boldsymbol{x}}_{t+1}) - \left(\sum_{t=1}^{T-1} F_t(\boldsymbol{u}_t^\gamma) - \sum_{t=1}^{T-1} F_t(\boldsymbol{u}_{t+1}^\gamma)\right) - \langle \boldsymbol{p}_T, \boldsymbol{x}_{T+1} \rangle + a_{1:T} + r(\boldsymbol{u}_1^\gamma) + 1.
\end{aligned}$$

The first inequality uses the update-rule optimality

$$F_T(\boldsymbol{x}_{T+1}) + \langle \boldsymbol{p}_T, \boldsymbol{x}_{T+1} \rangle \le F_T(\boldsymbol{u}_T^\gamma) + \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle,$$

and $-F_0(\tilde{\boldsymbol{x}}_1) \le 0$ (since $r \ge 0$ and $b_\gamma \ge 0$ on $\mathcal{X}$). The second inequality uses $b_\gamma \le 1$ on $\mathcal{X}^\gamma$, giving $f_0(\boldsymbol{u}_1^\gamma) \le r(\boldsymbol{u}_1^\gamma) + 1$. Overall

$$\begin{aligned}
&\sum_{t=1}^T \left(f_t(\tilde{\boldsymbol{x}}_t) - f_t(\boldsymbol{u}_t)\right) - \langle \boldsymbol{p}_T, \boldsymbol{u}_T^\gamma \rangle \\
&\quad \le \sum_{t=1}^T \left(F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1})\right) + \sum_{t=1}^{T-1} \big(F_t(\boldsymbol{u}_{t+1}^\gamma) - F_t(\boldsymbol{u}_t^\gamma)\big) + r(\boldsymbol{u}_1) + 1 + a_{1:T} - \langle \boldsymbol{p}_T, \boldsymbol{x}_{T+1} \rangle.
\end{aligned}$$

Since the terminal perturbation and terminal point are used only for analysis and are never played by the algorithm, we set $\boldsymbol{p}_T = \mathbf{0}$ and choose $\boldsymbol{x}_{T+1} = \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{X}} F_T(\boldsymbol{x})$. Then the preceding optimality inequality remains valid, and both terminal inner-product terms vanish. In addition, by Lipschitzness:

$$a_{1:T} = \sum_{t=1}^{T} f_t(\boldsymbol{u}_t^\gamma) - f_t(\boldsymbol{u}_t) \le \sum_{t=1}^{T} G\|\boldsymbol{u}_t^\gamma - \boldsymbol{u}_t\| \le G\gamma RT. \tag{24}$$

Finally: Note since $\boldsymbol{x}_t$ and $\tilde{\boldsymbol{x}}_t$ have the same density for all t, we have

$$\mathbb{E}\left[f_t(\boldsymbol{x}_t)\right] = \mathbb{E}\left[f_t(\tilde{\boldsymbol{x}}_t)\right] \tag{25}$$

The above property is guaranteed by the lazy sampling procedure: the fresh minimizer and the (potentially) kept one share the same density.

Hence

$$\begin{aligned}\mathbb{E}\left[\mathcal{R}_T\right] &= \mathbb{E}\left[\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t)\right] = \mathbb{E}\left[\sum_{t=1}^{T} f_t(\tilde{\boldsymbol{x}}_t) - f_t(\boldsymbol{u}_t)\right] \\ &\le \sum_{t=1}^{T} \mathbb{E}\left[(F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}))\right] \;+\; \sum_{t=1}^{T-1} \left(F_t(\boldsymbol{u}_{t+1}^\gamma) - F_t(\boldsymbol{u}_t^\gamma)\right) \;+\; r(\boldsymbol{u}_1) + 1 + G\gamma RT.\end{aligned}$$

□

Lemma 22. *Under the same conditions as Lemma 21, if we further assume that F_t is σ-strongly convex for all t, then*

$$(\mathbf{I}) \le \tfrac{G^2}{2\sigma}\, T \;+\; R\sqrt{2d}\,\mu.$$

, where μ is the scale parameter of the Laplace distribution.

Proof of Lemma 22. For each t, define

$$\tilde{F}_t(\cdot) \doteq F_t(\cdot) + \langle \boldsymbol{p}_{t-1}, \cdot\rangle. \tag{26}$$

Then,

$$\begin{aligned}&F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}) && (27)\\ &= F_t(\tilde{\boldsymbol{x}}_t) + \langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_t\rangle - \Big(F_t(\tilde{\boldsymbol{x}}_{t+1}) + \langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1}\rangle\Big) + \langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t\rangle && (28)\\ &= \tilde{F}_t(\tilde{\boldsymbol{x}}_t) - \tilde{F}_t(\tilde{\boldsymbol{x}}_{t+1}) + \langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t\rangle. && (29)\end{aligned}$$

By strong convexity of $\tilde{F}_t$, for

$$\boldsymbol{s}_t \doteq \nabla F_t(\tilde{\boldsymbol{x}}_t) + \boldsymbol{p}_{t-1},$$

we have

$$\tilde{F}_t(\tilde{\boldsymbol{x}}_t) - \tilde{F}_t(\tilde{\boldsymbol{x}}_{t+1}) \;\le\; \langle \boldsymbol{s}_t, \tilde{\boldsymbol{x}}_t - \tilde{\boldsymbol{x}}_{t+1}\rangle - \frac{\sigma}{2}\delta_t^2, \tag{30}$$

where

$$\delta_t \doteq \|\tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t\|.$$

The optimality condition of the update step at $\tilde{\boldsymbol{x}}_t$ gives

$$0 = \nabla F_{t-1}(\tilde{\boldsymbol{x}}_t) + \boldsymbol{p}_{t-1},$$

and therefore

$$\nabla f_t(\tilde{\boldsymbol{x}}_t) = \nabla F_t(\tilde{\boldsymbol{x}}_t) + \boldsymbol{p}_{t-1} = \boldsymbol{s}_t.$$

Thus, substituting into (30) yields

$$\begin{aligned}\tilde{F}_t(\tilde{\boldsymbol{x}}_t) - \tilde{F}_t(\tilde{\boldsymbol{x}}_{t+1}) &\le \langle \boldsymbol{g}_t, \tilde{\boldsymbol{x}}_t - \tilde{\boldsymbol{x}}_{t+1}\rangle - \frac{\sigma}{2}\delta_t^2 && (31)\\ &\le G\delta_t - \frac{\sigma}{2}\delta_t^2 \le \frac{G^2}{2\sigma}, && (32)\end{aligned}$$

where $\boldsymbol{g}_t \doteq \nabla f_t(\tilde{\boldsymbol{x}}_t)$. The second inequality follows from Cauchy–Schwarz and Lipschitzness, and the third from the elementary inequality

$$ax - \frac{b}{2}x^2 \le \frac{a^2}{2b}, \qquad a, b > 0.$$

Combining (29) and (32), and summing over $t = 1, \dots, T$, we obtain the following pathwise bound:

$$\sum_{t=1}^{T} \Big(F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}) \Big) \le \frac{G^2}{2\sigma} T + \sum_{t=1}^{T} \langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t \rangle. \tag{33}$$

Now let $\mathcal{A}$ denote the original perturbation law, under which the vectors $\{\boldsymbol{p}_t\}_{t=0}^{T-1}$ are i.i.d. Laplace. Let $\mathcal{B}$ be any other law on the perturbations such that each $\boldsymbol{p}_t$ has the same marginal distribution as under $\mathcal{A}$.

Recall that $\tilde{\boldsymbol{x}}_t$, by definition, depends only on the randomness of $\boldsymbol{p}_{t-1}$. Therefore, as long as the perturbation marginals coincide, the induced marginal distribution of each $\tilde{\boldsymbol{x}}_t$ is identical under $\mathcal{A}$ and $\mathcal{B}$. Hence,

$$\sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{A}} \left[F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}) \right] = \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{B}} \left[F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}) \right]. \tag{34}$$

We now choose $\mathcal{B}$ to simplify the perturbation term in (33). Namely, let $\mathcal{B}$ be the law defined by

$$\boldsymbol{p}_0 \sim \mathrm{Lap}(\mu), \qquad \boldsymbol{p}_t = \boldsymbol{p}_0 \quad \forall t \in [T{-}1]. \tag{35}$$

This law matches the single-time marginal distribution of each $\boldsymbol{p}_t$ under $\mathcal{A}$, since under $\mathcal{A}$ the perturbations $\boldsymbol{p}_t$ are i.i.d. Laplace.

Taking expectation of (33) under $\mathcal{B}$, and using (34), gives

$$(\mathbf{I}) = \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{A}} \left[F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}) \right] \tag{36}$$

$$= \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{B}} \left[F_t(\tilde{\boldsymbol{x}}_t) - F_t(\tilde{\boldsymbol{x}}_{t+1}) \right] \le \frac{G^2}{2\sigma} T + \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{B}} \left[\langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t \rangle \right]. \tag{37}$$

To bound the remaining sum, we rewrite it as

$$\sum_{t=1}^{T} \langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t \rangle = \langle \boldsymbol{p}_{T-1}, \tilde{\boldsymbol{x}}_{T+1} \rangle - \langle \boldsymbol{p}_0, \tilde{\boldsymbol{x}}_1 \rangle - \sum_{t=2}^{T} \langle \boldsymbol{p}_{t-1} - \boldsymbol{p}_{t-2}, \tilde{\boldsymbol{x}}_t \rangle. \tag{38}$$

Under $\mathcal{B}$, we have $\boldsymbol{p}_{t-1} - \boldsymbol{p}_{t-2} = 0$ for $t = 2, \dots, T$, hence

$$\sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{B}} \left[\langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t \rangle \right] = \mathop{\mathbb{E}}_{\mathcal{B}} \left[\langle \boldsymbol{p}_0, \tilde{\boldsymbol{x}}_{T+1} - \tilde{\boldsymbol{x}}_1 \rangle \right]. \tag{39}$$

Since $\boldsymbol{p}_T = \mathbf{0}$, the terminal point $\tilde{\boldsymbol{x}}_{T+1}$ is deterministic for the fixed loss sequence. Using $\mathbb{E}_{\mathcal{B}}[\boldsymbol{p}_0] = \mathbf{0}$, we obtain

$$\left| \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mathcal{B}} \left[\langle \boldsymbol{p}_{t-1}, \tilde{\boldsymbol{x}}_{t+1} - \tilde{\boldsymbol{x}}_t \rangle \right] \right| = \left| \mathop{\mathbb{E}}_{\mathcal{B}} \left[\langle \boldsymbol{p}_0, \tilde{\boldsymbol{x}}_1 \rangle \right] \right| \tag{40}$$

$$\le R \mathop{\mathbb{E}}_{\mathcal{B}} \left[\|\boldsymbol{p}_0\| \right] \tag{41}$$

$$\le R\sqrt{2d}\,\mu, \tag{42}$$

where the last inequality follows from the bound on the first absolute moment,

$$\mathbb{E}\|\boldsymbol{p}_0\| \le \sqrt{2d}\,\mu,$$

of a d-dimensional Laplace distribution with scale parameter μ.

Substituting this into (37) completes the proof. □

Lemma 23. *Under the same conditions as Lemma 22, we have that*

$$(\mathbf{II}) \leq (TG + \sigma R + CG_c\sqrt{T})\ P_T \tag{43}$$

Proof of Lemma 23. From strong convexity

$$F_t(\boldsymbol{u}_{t+1}^\gamma) - F_t(\boldsymbol{u}_t^\gamma) \leq \|\boldsymbol{q}_t\|\|\boldsymbol{u}_{t+1}^\gamma - \boldsymbol{u}_t^\gamma\| - \frac{\sigma}{2}\|\boldsymbol{u}_{t+1}^\gamma - \boldsymbol{u}_t^\gamma\|^2, \tag{44}$$

where $\boldsymbol{q}_t = \nabla F_t(\boldsymbol{u}_{t+1}^\gamma)$.

To bound $\|\boldsymbol{q}_t\|$:

$$\nabla F_t(\boldsymbol{x}) = \nabla f_{1:t}(\boldsymbol{x}) + \sigma\boldsymbol{x} + \nabla b_\gamma(\boldsymbol{x}). \tag{45}$$

$$= \nabla f_{1:t}(\boldsymbol{x}) + \sigma\boldsymbol{x} - \sum_{c=1}^{C} \frac{1}{M_\gamma s_c(\boldsymbol{x})}\nabla s_c(\boldsymbol{x}). \tag{46}$$

• For $\|\nabla b_\gamma(\boldsymbol{u}_{t+1}^\gamma)\|$:

By definition, any $\boldsymbol{y} \in \mathcal{X}^\gamma$ can be written as $\boldsymbol{y} = (1-\gamma)\boldsymbol{x}$ for some $\boldsymbol{x} \in \mathcal{X}$, hence $\boldsymbol{y} = (1-\gamma)\boldsymbol{x} + \gamma\boldsymbol{0}$. Since s_c is concave,

$$s_c(\boldsymbol{y}) \geq (1-\gamma)s_c(\boldsymbol{x}) + \gamma s_c(\boldsymbol{0}) \geq \gamma,$$

where we used $s_c(\boldsymbol{x}) \geq 0$ for $\boldsymbol{x} \in \mathcal{X}$ and the normalization $s_c(\boldsymbol{0}) = 1$.

The gradient of the scaled barrier is

$$\nabla b_\gamma(\boldsymbol{x}) = -\frac{1}{M_\gamma}\sum_{c=1}^{C}\frac{1}{s_c(\boldsymbol{x})}\,\nabla s_c(\boldsymbol{x}).$$

Since $\|\nabla s_c(\boldsymbol{x})\| \leq G_c$ for all vectors in the shrunk set $\boldsymbol{x} \in \mathcal{X}^\gamma$, then using $s_c(\boldsymbol{x}) \geq \gamma$ on $\mathcal{X}^\gamma$ yields

$$\sup_{\boldsymbol{x}\in\mathcal{X}^\gamma}\|\nabla b_\gamma(\boldsymbol{x})\| \leq \frac{1}{M_\gamma}\sum_{c=1}^{C}\frac{G_c}{\gamma} = \frac{1}{\gamma M_\gamma}\sum_{c=1}^{C}G_c \leq \frac{CG_c}{\gamma}.$$

Finally, setting $\gamma = 1/\sqrt{T}$, we get $\|\nabla b_\gamma(\boldsymbol{u}_{t+1}^\gamma)\| \leq CG_c\sqrt{T}$.

• For $\sigma\|\boldsymbol{u}_{t+1}^\gamma\| \leq (1-\gamma)\sigma R \leq \sigma R$

• For $\|\nabla f_{1:t}(\boldsymbol{u}_{t+1}^\gamma)\|$, we can only do $\|\nabla f_{1:t}(\boldsymbol{u}_{t+1}^\gamma)\| \leq tG \leq TG$.

Combining the three bounds gives

$$\|\boldsymbol{q}_t\| \leq TG + \sigma R + CG_c\sqrt{T}.$$

Substituting this into (44), dropping the negative quadratic term, and summing over $t = 1, \ldots, T-1$, we obtain

$$(\mathbf{II}) \leq (TG + \sigma R + CG_c\sqrt{T})\sum_{t=1}^{T-1}\|\boldsymbol{u}_{t+1}^\gamma - \boldsymbol{u}_t^\gamma\|.$$

Since $\sum_{t=1}^{T-1}\|\boldsymbol{u}_{t+1}^\gamma - \boldsymbol{u}_t^\gamma\| \leq P_T$, the result follows. □

Lemma 24. *Suppose the loss sequence $\{f_t\}$ consists of convex, G-Lipschitz, and β-smooth functions. Then, the total variation distance between $\mathcal{Q}_t$ and $\mathcal{Q}_{t+1}$ is bounded by*

$$\|\mathcal{Q}_{t+1} - \mathcal{Q}_t\|_{TV} \leq \frac{\beta d}{\sigma} + \frac{\sqrt{d}\,G}{\mu}.$$

Proof of Lemma 24. By the change of variables formula, the density ratio satisfies

$$\frac{\mathcal{Q}_{t+1}(\boldsymbol{x})}{\mathcal{Q}_t(\boldsymbol{x})} = \frac{\nu(-\nabla F_t(\boldsymbol{x}))}{\nu(-\nabla F_{t-1}(\boldsymbol{x}))} \cdot \Big|\frac{\det(-\nabla^2 F_t(\boldsymbol{x}))}{\det(-\nabla^2 F_{t-1}(\boldsymbol{x}))}\Big|$$

We bound each component separately.

Determinant ratio. Let $\lambda_{i,t}$ denote the i-th largest eigenvalue of the symmetric matrix $\nabla^2 F_t(\boldsymbol{x})$. Because the newly revealed loss function f_t is β-smooth, the spectral norm (operator norm) of its Hessian is bounded by β. By applying Weyl's inequality, which bounds the change in eigenvalues by the spectral norm of the difference, we obtain

$$|\lambda_{i,t} - \lambda_{i,t-1}| \leq \|\nabla^2 F_t(\boldsymbol{x}) - \nabla^2 F_{t-1}(\boldsymbol{x})\|_{\mathrm{op}} = \|\nabla^2 f_t(\boldsymbol{x})\|_{\mathrm{op}} \leq \beta$$

Furthermore, since F_t is also σ-strongly convex, all eigenvalues satisfy $\lambda_{i,t} \geq \sigma > 0, \forall t$. Therefore,

$$\frac{\det(\nabla^2 F_t(\boldsymbol{x}))}{\det(\nabla^2 F_{t-1}(\boldsymbol{x}))} \leq \prod_{i=1}^{d} \left(\frac{\lambda_{i,t-1} + \beta}{\lambda_{i,t-1}} \right) \leq \prod_{i=1}^{d} \left(1 + \frac{\beta}{\sigma} \right) = \left(1 + \frac{\beta}{\sigma} \right)^d \leq \exp\left(\frac{\beta d}{\sigma} \right).$$

Density ratio. Using the L_1 structure of the Laplace distribution and the reverse triangle inequality,

$$\begin{aligned} \frac{\nu(-\nabla F_t(\boldsymbol{x}))}{\nu(-\nabla F_{t-1}(\boldsymbol{x}))} = \exp\left(\frac{\|\nabla F_{t-1}(\boldsymbol{x})\|_1 - \|\nabla F_t(\boldsymbol{x})\|_1}{\mu} \right) &\leq \exp\left(\frac{\|\nabla F_{t-1}(\boldsymbol{x}) - \nabla F_t(\boldsymbol{x})\|_1}{\mu} \right) \\ &= \exp\left(\frac{\|\nabla f_t(\boldsymbol{x})\|_1}{\mu} \right). \end{aligned}$$

Since f_t is G-Lipschitz,

$$\|\nabla f_t(\boldsymbol{x})\|_1 \leq \sqrt{d}\, \|\nabla f_t(\boldsymbol{x})\|_2 \leq \sqrt{d}\, G.$$

Thus,

$$\frac{\nu(-\nabla F_t(\boldsymbol{x}))}{\nu(-\nabla F_{t-1}(\boldsymbol{x}))} \leq \exp\left(\frac{\sqrt{d}\, G}{\mu} \right).$$

Combining both bounds,

$$\frac{\mathcal{Q}_{t+1}(\boldsymbol{x})}{\mathcal{Q}_t(\boldsymbol{x})} \leq \exp\left(\frac{\beta d}{\sigma} + \frac{\sqrt{d}\, G}{\mu} \right) = \exp(\epsilon),$$

where

$$\epsilon \doteq \frac{\beta d}{\sigma} + \frac{\sqrt{d}\, G}{\mu}.$$

Taking reciprocals gives

$$\frac{\mathcal{Q}_t(\boldsymbol{x})}{\mathcal{Q}_{t+1}(\boldsymbol{x})} \geq \exp(-\epsilon),$$

which implies

$$\mathcal{Q}_{t+1}(\boldsymbol{x}) - \mathcal{Q}_t(\boldsymbol{x}) \leq (1 - e^{-\epsilon})\, \mathcal{Q}_{t+1}(\boldsymbol{x}).$$

Integrating over any measurable set E yields

$$\mathcal{Q}_{t+1}(E) - \mathcal{Q}_t(E) \leq 1 - e^{-\epsilon} \leq \epsilon.$$

To also get $\mathcal{Q}_t(E) - \mathcal{Q}_{t+1}(E) \leq \epsilon$ we start the proof again but with the reciprocal bound $\mathcal{Q}_t(\boldsymbol{x})/\mathcal{Q}_{t+1}(\boldsymbol{x})$, and using $|\lambda_{i,t} - \lambda_{i,t-1}| \leq \beta$, we get

$$\frac{\det(\nabla^2 F_{t-1}(\boldsymbol{x}))}{\det(\nabla^2 F_t(\boldsymbol{x}))} \leq \prod_{i=1}^{d} \left(\frac{\lambda_{i,t} + \beta}{\lambda_{i,t}} \right) \leq \exp\left(\frac{\beta d}{\sigma} \right).$$

The remaining proof steps are the same thereafter, and we get the same bound ϵ. Hence, we have

$$\|\mathcal{Q}_{t+1} - \mathcal{Q}_t\|_{TV} \leq \epsilon.$$

□

Now we are ready to prove the main theorem.

Proof of Theorem 5. By linearity of expectation and $\lambda_t \le \lambda$,

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}\right] \le \mathbb{E}\left[\mathcal{R}_T\right] + \lambda \sum_{t=2}^{T} \mathbb{E}\left[\mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right].$$

The bound on $\mathbb{E}\left[\mathcal{R}_T\right]$ follows by combining Lemma 22 and Lemma 23. For the second term, the lazy sampling procedure implies

$$\mathbb{E}\left[\mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right] = \Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) = \|\mathcal{Q}_t - \mathcal{Q}_{t-1}\|_{\mathrm{TV}}.$$

Applying the bound on $\|\mathcal{Q}_t - \mathcal{Q}_{t-1}\|_{\mathrm{TV}}$ from Lemma 24 yields the claimed result. □

C.1 Proof of Theorem 7

Proof. Fix any dyadic scale $k \in \mathcal{H}$, and partition the horizon into $m = T/k$ consecutive blocks

$$B_i = \{(i-1)k + 1, \ldots, ik\}, \qquad i = 1, \ldots, m.$$

For each block, define the within-block path budget

$$P_i \doteq \sum_{t=(i-1)k+2}^{ik} \|\boldsymbol{u}_t - \boldsymbol{u}_{t-1}\|.$$

Applying Theorem 5 on each block with horizon k and the given choices of (γ, μ, σ), we obtain

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right]$$

$$\le \sum_{i=1}^{m} \left[(R + P_i)\sqrt{(G^2 + 2\lambda\beta d)\,k} + 2^{5/4}\sqrt{\lambda d R G\,k} + GkP_i + \left(CG_cP_i + GR\right)\sqrt{k} + 1\right] + \lambda(m-1),$$

where the $\lambda(m-1)$ term comes from the switching cost at the restart boundaries. Since $2^{5/4} < 3$, $\sum_{i=1}^m P_i \le P_T$, and $m = T/k$, this implies

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le A\,\frac{T}{\sqrt{k}} + B\,P_T\sqrt{k} + GP_Tk.$$

Because $k \ge 1$, we have $\sqrt{k} \le k$, and therefore

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le \phi(k) \doteq A\,\frac{T}{\sqrt{k}} + (B+G)P_T\,k.$$

We now optimize ϕ over $k > 0$. Differentiating gives

$$\phi'(k) = -\frac{1}{2}AT\,k^{-3/2} + (B+G)P_T,$$

so the unique stationary point is

$$\bar{k} = \left(\frac{AT}{2(B+G)P_T}\right)^{2/3}.$$

After clipping to the admissible range $[1, T]$, define

$$k^\star \doteq \min\{T, \max\{1, \bar{k}\}\}.$$

Since $\mathcal{H}$ is dyadic, there exists

$$k^\dagger \doteq 2^{\lfloor \log_2 k^\star \rfloor} \in \mathcal{H}$$

such that

$$\frac{k^\star}{2} \le k^\dagger \le k^\star.$$

Hence

$$\phi(k^\dagger) \le \sqrt{2}\,A\,\frac{T}{\sqrt{k^\star}} + (B+G)P_T\,k^\star,$$

which proves the first claim.

It remains to simplify the three regimes. If $1 \le \bar{k} \le T$, then $k^\star = \bar{k}$, and substituting $\bar{k}$ into the display above yields

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] = \mathcal{O}\left(A^{2/3}(B+G)^{1/3}T^{2/3}P_T^{1/3}\right).$$

If $\bar{k} < 1$, then $AT < 2(B+G)P_T$, so $k^\star = 1$ and

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le \sqrt{2}AT + (B+G)P_T = \mathcal{O}\left((B+G)P_T\right).$$

If $\bar{k} > T$, then $AT > 2(B+G)P_T T^{3/2}$, so $k^\star = T$ and

$$(B+G)P_T T \le \frac{A}{2}\sqrt{T}.$$

Therefore,

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(u_{1:T})\right] \le \sqrt{2}A\sqrt{T} + (B+G)P_T T = \mathcal{O}\left(A\sqrt{T}\right).$$

Finally, if $P_T = T^\alpha$ with $\alpha \in [0,1]$, then

$$P_T = T^\alpha \le T^{(2+\alpha)/3} \qquad \text{and} \qquad \sqrt{T} \le T^{(2+\alpha)/3},$$

so both boundary branches are dominated by the interior rate

$$T^{(2+\alpha)/3} = T^{2/3}P_T^{1/3}.$$

This proves the final claim. □

D Dyadic intervals

D.1 Proof of Lemma 9

To prove the given expression, we need the following geometric covering fact about intervals:

Lemma 25. *Orabona [17, Thm. 15.9] Let $J = [q, r] \subseteq \mathbb{N}$ be an arbitrary interval. Then J can be partitioned into two finite sequences of disjoint and consecutive intervals,*

$$(J_{-k}, \dots, J_0) \subseteq \mathcal{I}|_J, \qquad (J_1, \dots, J_p) \subseteq \mathcal{I}|_J,$$

such that

1. *the intervals $J_{-k}, \dots, J_0, J_1, \dots, J_p$ are consecutive and their union is J;*
2. *$|J_{-i}|/|J_{-i+1}| \le 1/2$ for every $i \ge 1$;*
3. *$|J_i|/|J_{i-1}| \le 1/2$ for every $i \ge 2$.*

Proof. Apply Lemma 25 to $I = [s, e]$. We obtain two sequences

$$(J_{-k}, \dots, J_0) \subseteq \mathcal{I}|_I, \qquad (J_1, \dots, J_p) \subseteq \mathcal{I}|_I,$$

whose union is I. Relabeling from left to right gives a partition $J_1, \dots, J_m \in \mathcal{G}$, where $m = k+p+1$.

Since every interval in $\mathcal{I}$ has dyadic length, and the ratios in Lemma 25 are at most $1/2$, the block lengths on each side lie on distinct dyadic scales. As every block has length at most L, each side contains at most $\lceil \log_2 L \rceil + 1$ blocks. Therefore

$$m \le 2\lceil \log_2 L \rceil + 2,$$

It remains to bound the sum of square roots. For the left side, write

$$a_i \doteq |J_{-i}|, \qquad i = 0, \dots, k,$$

and define the tail sums

$$S_i \doteq \sum_{j=i}^{k} a_j, \qquad i = 0, \dots, k, \qquad S_{k+1} \doteq 0.$$

Because the lengths decrease geometrically,

$$S_{i+1} \le \sum_{r \ge 1} \frac{a_i}{2^r} \le a_i.$$

Hence

$$S_i = a_i + S_{i+1} \le 2a_i$$

Therefore

$$\sqrt{S_i} + \sqrt{S_{i+1}} \le (\sqrt{2}+1)\sqrt{a_i},$$

Multiplying both sides by $(\sqrt{S_i} - \sqrt{S_{i+1}})/\sqrt{a_i}$ allows us to bound $\sqrt{a_i}$ as follows:

$$\sqrt{a_i} \le (1+\sqrt{2})\big(\sqrt{S_i} - \sqrt{S_{i+1}}\big).$$

Summing over $i = 0, \dots, k$ yields

$$\sum_{i=0}^{k} \sqrt{|J_{-i}|} \le (1+\sqrt{2}) \sum_{i=0}^{k} \big(\sqrt{S_i} - \sqrt{S_{i+1}}\big) = (1+\sqrt{2})\sqrt{S_0}.$$

If we set

$$A \doteq \sum_{i=0}^{k} |J_{-i}|, \qquad B \doteq \sum_{i=1}^{p} |J_i|,$$

then $S_0 = A$, and the same argument on the right side gives

$$\sum_{i=1}^{p} \sqrt{|J_i|} \le (1+\sqrt{2})\sqrt{B}.$$

Since $A + B = L$, we conclude

$$\sum_{r=1}^{m} \sqrt{|J_r|} \le (1+\sqrt{2})(\sqrt{A} + \sqrt{B}) \le (1+\sqrt{2})\sqrt{2(A+B)} = (2+\sqrt{2})\sqrt{L}.$$

□

E Proofs for meta-learner

E.1 Proof for the one step

Proof of Lemma 10. Write

$$\mathcal{P}_t - \mathcal{P}_{t-1} = \sum_{H \in \mathcal{H}} v_{t,H} \big(\mathcal{Q}_t^{(H)} - \mathcal{Q}_{t-1}^{(H)}\big) + \sum_{H \in \mathcal{H}} (v_{t,H} - v_{t-1,H})\, \mathcal{Q}_{t-1}^{(H)}.$$

Taking total variation and using the triangle inequality yields

$$\|\mathcal{P}_t - \mathcal{P}_{t-1}\|_{\mathrm{TV}} \le \sum_{H \in \mathcal{H}} v_{t,H} \left\|\mathcal{Q}_t^{(H)} - \mathcal{Q}_{t-1}^{(H)}\right\|_{\mathrm{TV}} + \left\| \sum_{H \in \mathcal{H}} (v_{t,H} - v_{t-1,H}) \mathcal{Q}_{t-1}^{(H)} \right\|_{\mathrm{TV}}.$$

The first term is exactly

$$\sum_{H \in \mathcal{H}} v_{t,H}\, c_t^{(H)}.$$

For the second term, set

$$a_H \doteq v_{t,H} - v_{t-1,H}, \qquad A_+ \doteq \{H \in \mathcal{H} : a_H > 0\}, \qquad A_- \doteq \{H \in \mathcal{H} : a_H < 0\}.$$

Since both $\boldsymbol{v}_t$ and $\boldsymbol{v}_{t-1}$ lie in the simplex, $\sum_{H\in\mathcal{H}} a_H = 0$. Hence

$$\alpha \doteq \sum_{H\in A_+} a_H = -\sum_{H\in A_-} a_H = \frac{1}{2}\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1.$$

If $\alpha = 0$, there is nothing to prove. Otherwise,

$$\sum_{H\in\mathcal{H}} a_H \mathcal{Q}_{t-1}^{(H)} = \alpha(\mu_+ - \mu_-),$$

where

$$\mu_+ \doteq \sum_{H\in A_+} \frac{a_H}{\alpha}\,\mathcal{Q}_{t-1}^{(H)}, \qquad \mu_- \doteq \sum_{H\in A_-} \frac{-a_H}{\alpha}\,\mathcal{Q}_{t-1}^{(H)}$$

are probability distributions (convex combination of probability distributions). Therefore,

$$\left\|\sum_{H\in\mathcal{H}} a_H \mathcal{Q}_{t-1}^{(H)}\right\|_{\mathrm{TV}} = \alpha\,\|\mu_+ - \mu_-\|_{\mathrm{TV}} \le \alpha = \frac{1}{2}\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1.$$

This proves Eq. (9).

To pass from the meta dot product to the master expectation, define

$$\boldsymbol{\ell}_t \doteq \left(\ell_t^{(H)}\right)_{H\in\mathcal{H}}, \qquad \boldsymbol{c}_t \doteq \left(c_t^{(H)}\right)_{H\in\mathcal{H}}, \qquad \boldsymbol{g}_t = \boldsymbol{\ell}_t + \lambda \boldsymbol{c}_t.$$

By the lazy sampling properties (Eq. (20)), we have $\boldsymbol{x}_t \sim \mathcal{P}_t$. Also, using linearity of integration gives:

$$\langle \boldsymbol{v}_t, \boldsymbol{\ell}_t\rangle = \sum_{H\in\mathcal{H}} v_{t,H}\,\ell_t^{(H)} \overset{\text{(LOTUS)}}{=} \sum_{H\in\mathcal{H}} v_{t,H}\int_{\mathcal{X}} f_t(\boldsymbol{x})\,d\mathcal{Q}_t^{(H)}(\boldsymbol{x}) = \int_{\mathcal{X}} f_t(\boldsymbol{x})\,d\Big(\sum_{H\in\mathcal{H}} v_{t,H}\,\mathcal{Q}_t^{(H)}\Big)(\boldsymbol{x}) \tag{47}$$

$$\overset{Eq.\ (2)}{=} \int_{\mathcal{X}} f_t(\boldsymbol{x})\,d\mathcal{P}_t(\boldsymbol{x}) \overset{\text{(LOTUS)}}{=} \mathop{\mathbb{E}}_{\boldsymbol{x}_t\sim\mathcal{P}_t}[f_t(\boldsymbol{x}_t)] \tag{48}$$

Also, again by the lazy coupling identity Eq. (21),

$$\Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) = \|\mathcal{P}_t - \mathcal{P}_{t-1}\|_{\mathrm{TV}}.$$

Combining these two identities with Eq. (9) yields

$$\mathbb{E}[f_t(\boldsymbol{x}_t)] + \lambda\Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) \le \langle \boldsymbol{v}_t, \boldsymbol{\ell}_t\rangle + \lambda\langle \boldsymbol{v}_t, \boldsymbol{c}_t\rangle + \frac{\lambda}{2}\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 = \langle \boldsymbol{v}_t, \boldsymbol{g}_t\rangle + \frac{\lambda}{2}\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1.$$

This proves Eq. (10). □

E.2 The DM strongly adaptive regret result

We use the following result from [9]:

Lemma 26 ([9], Thm. 1)**.** *Fix a horizon T, a number of experts N, a switching cost $D \ge 0$, and an oblivious loss sequence $l_1, \ldots, l_T \in [0,1]^N$. Then there exists a randomized online algorithm choosing experts $I_t \in [N]$ such that for every interval $[s,e] \subseteq [T]$ of length $L = e - s + 1$,*

$$\mathbb{E}\left[\sum_{t=s}^{e} l_t(I_t) + D\sum_{t=s+1}^{e} \mathbf{1}\{I_t \neq I_{t-1}\}\right] \le \min_{i\in[N]} \sum_{t=s}^{e} l_t(i) + C_{\mathrm{dm}}\sqrt{(D+1)L\log(NT)}, \tag{49}$$

where $C_{\mathrm{dm}} > 0$ is universal.

Algorithm 4 Fixed-Share scale learner at scale τ

Require: scale τ, switching cost $D \geq 1$

$\eta \leftarrow \sqrt{\frac{\log(N\tau)}{D\tau}}$
$z_1 \leftarrow (1/N, \ldots, 1/N) \in \Delta_N$
for $t = 1, 2, \ldots, T$ **do**
  output z_t
  observe loss vector $\ell_t \in [0,1]^N$
  if $\tau \geq 16D\log(N\tau)$ **then**
    update

$$z_{t+1}(i) = \frac{z_t(i)e^{-\eta\ell_t(i)} + \frac{1}{N\tau}}{\sum_{j=1}^{N}\left(z_t(j)e^{-\eta\ell_t(j)} + \frac{1}{N\tau}\right)}, \qquad i = 1, \ldots, N$$

  else
    $z_{t+1} \leftarrow z_t$
  end if
end for

E.3 Algorithmic mechanism behind [9, Thm. 1]

We now review the algorithmic mechanism that yields the imported guarantee of [9, Thm. 1]. Since our reduction works on the simplex, we state the construction in its simplex-valued OLO form.

A minor terminology warning is that The Daniely-Mansour mechanism itself is already hierarchical: it has a family of scale-dependent learners, together with an internal combination procedure that merges them across scales. Thus, when we invoke the Daniely-Mansour learner as our meta learner, that meta learner has its own internal base/meta structure. We use the term *scale learner* below for the internal Daniely-Mansour components to avoid confusion.

For simplicity, this section follows the constructive proof convention of Daniely and Mansour [9]: we describe the mechanism for $T = 2^J$ and $D \geq 1$. This restriction is only for the constructive review below; the imported guarantee in Lemma 26 is stated for all $D \geq 0$, and the case $D < 1$ is handled in Daniely and Mansour [9] by reduction to $D = 1$. The construction has three pieces:

1. a scale learner for each dyadic scale τ;
2. a two-way combiner that softly interpolates between two simplex-valued algorithms;
3. a multiscale wrapper that combines the dyadic-scale learners recursively.

Intuitively, the scale-τ learner is tuned to intervals of length about τ, while the combiner ensures that passing from coarse to fine scales preserves control of the switching cost.

Gate function. For parameters $\tau \geq 1$ and $Z \in (0, 1/e]$, let $\tilde{g}_{\tau,Z}$ solve

$$8\tilde{g}'_{\tau,Z}(x) = \frac{x}{\tau}\tilde{g}_{\tau,Z}(x) + Z, \qquad \tilde{g}_{\tau,Z}(0) = 0.$$

Define

$$g_{\tau,Z}(x) := \Pi_{[0,1]}(\tilde{g}_{\tau,Z}(x)), \qquad U_{\tau,Z} := \tilde{g}^{-1}_{\tau,Z}(1),$$

where $\Pi_{[a,b]}$ denotes projection onto $[a,b]$.

Algorithm 4 is the scale-dependent learner used in the proof of Theorem 2.1.

Algorithm 5 is the key internal combination step. It mixes the two input algorithms using the gate value $g_{\tau,Z}(x_t)$, where the scalar state x_t tracks the recent difference between their surrogate losses.

Algorithm 6 maintains one scale learner for each dyadic scale

$$T,\ T/2,\ T/4,\ \ldots,\ 2,$$

and combines them recursively from coarse to fine scales. On any interval I, one of these scales is of the correct order of magnitude for $|I|$, and the recursive combination guarantees that the final

Algorithm 5 Two-way combiner $\mathrm{Combine}_{\tau,Z}(P, Q)$

Require: simplex-valued OLO algorithms P, Q, scale τ, parameter $Z \in (0, 1/e]$, switching cost $D \geq 1$

initialize $x_1 \leftarrow 0$
for $t = 1, 2, \ldots, T$ **do**
  obtain current actions $p_t \in \Delta_N$ from P and $q_t \in \Delta_N$ from Q
  if $\tau \geq 64D \log(1/Z)$ **then**
    $\alpha_t \leftarrow g_{\tau,Z}(x_t)$
    output $r_t \leftarrow (1 - \alpha_t)p_t + (\alpha_t)q_t$
  else
    output $r_t \leftarrow p_t$
  end if
  observe loss vector ℓ_t
  let p_{t+1}, q_{t+1} be the next actions of P, Q
  define the surrogate losses

$$\widetilde{L}_t(P) := \frac{\langle \ell_t, p_t \rangle + D\|p_{t+1} - p_t\|_{\mathrm{TV}}}{3}, \qquad \widetilde{L}_t(Q) := \frac{\langle \ell_t, q_t \rangle + D\|q_{t+1} - q_t\|_{\mathrm{TV}}}{3}$$

  set

$$b_t := \frac{\widetilde{L}_t(P) - \widetilde{L}_t(Q)}{\sqrt{D}}, \qquad x_{t+1} := \Pi_{[-2, U_{\tau,Z}+2]}\left(\left(1 - \frac{1}{\tau}\right) x_t + b_t\right)$$

end for
return algorithm object $\mathcal{C}$ with $\mathcal{C}.\textsc{Query}(\ell_t) \to r_t$

Algorithm 6 Multiscale OLO algorithm underlying Daniely–Mansour Thm. 2.1

Require: horizon $T = 2^J$, number of experts N, switching cost $D \geq 1$

$Z \leftarrow \frac{1}{2T \log T}$
for $u = 0, 1, \ldots, J - 1$ **do**
  $\tau_u \leftarrow 2^{-u}T$
  $A_u \leftarrow \textsc{Fixed-Share}(\tau_u, D)$
end for
$B_0 \leftarrow A_0$
for $u = 1, 2, \ldots, J - 1$ **do**
  $B_u \leftarrow \mathrm{Combine}_{\tau_u,Z}(B_{u-1}, A_u)$
end for
return B_{J-1} ▷ same interface: $B_{J-1}.\textsc{Query}(\ell_t) \to \boldsymbol{v}_t \in \Delta_N$

algorithm inherits that scale's regret bound up to logarithmic overhead. Here J denotes the logarithm of the DM horizon, whereas N denotes the number of experts in the DM subproblem. In our meta-learning application, $N = K = |\mathcal{H}|$.

Remarks.

1. All three algorithms share the same interface: when queried with $\ell_t \in [0, 1]^N$ at round t, each returns a weight vector $\boldsymbol{v}_t \in \Delta_N$. COMBINE is a constructor that takes two such objects and produces a third. DM-MULTISCALE returns the root of the resulting recursion tree; querying it round by round on $(\boldsymbol{g}_t/M)$ is exactly what Algorithm 2 does.
2. The denominator 3 in Algorithm 5 is the specialization of their factor $(M+1)$ to the case $M = 2$, which is the value used in the proof of the multiscale combination theorem.
3. The smallest dyadic scale is 2, not 1, since the construction uses $\tau_u = 2^{-u}T$ for $u = 0, \ldots, J - 1$.
4. The output of Algorithm 6 is a sequence $(p_t)_{t=1}^T \subseteq \Delta_N$. By the OLO/EXP equivalence in Daniely and Mansour [9, Sec 5.2], this simplex-valued algorithm can be converted into a randomized experts algorithm with the same expected loss and switching cost.

Now, we can prove the exact statement of Lemma 11 based on Lemma 26. We use Lemma 26 in its stated form for all $D \geq 0$; the algorithmic description in Section E.3 is included only to recall the constructive mechanism and follows the $D \geq 1$ proof convention of Daniely and Mansour [9].

Proof of Lemma 11. We reduce our setting to the discrete expert problem by defining the number of experts $N \doteq |\mathcal{H}|$. To accommodate the bounded surrogate range M, we define the normalized oblivious loss sequence $l_t \doteq g_t/M \in [0,1]^N$ for all $t \in [T]$, and set the switching cost parameter to $D \doteq \lambda/M$.

Let $I_t \in [N]$ be the randomized expert chosen by the algorithm from Lemma 26. We define our meta-learner's distribution over the simplex as $v_t(i) \doteq \Pr(I_t = i)$ for $i \in [N]$. Because the losses are oblivious, v_t is determined purely online from the past loss history and the algorithm's internal randomness.

For each step t, the expected loss is exactly the inner product:

$$\mathbb{E}[l_t(I_t)] = \sum_{i=1}^{N} v_t(i) l_t(i) = \langle v_t, l_t \rangle.$$

Furthermore, the joint law of (I_{t-1}, I_t) is a coupling of the marginal distributions v_{t-1} and v_t. For any coupling of two discrete distributions, the mismatch probability is lower bounded by their total variation distance. Therefore,

$$\Pr(I_t \neq I_{t-1}) \geq \mathrm{TV}(v_t, v_{t-1}) = \frac{1}{2}\|v_t - v_{t-1}\|_1.$$

Substituting these expected loss and mismatch probability bounds into the expectation on the left-hand side of Eq. (49) yields a lower bound on the discrete algorithm's cost:

$$\sum_{t=s}^{e} \langle v_t, l_t \rangle + \frac{D}{2} \sum_{t=s+1}^{e} \|v_t - v_{t-1}\|_1 \leq \mathbb{E}\left[\sum_{t=s}^{e} l_t(I_t) + D \sum_{t=s+1}^{e} \mathbf{1}\{I_t \neq I_{t-1}\} \right].$$

By chaining this with the upper bound directly from Lemma 26, we obtain the simplex formulation of the regret:

$$\sum_{t=s}^{e} \langle v_t, l_t \rangle + \frac{D}{2} \sum_{t=s+1}^{e} \|v_t - v_{t-1}\|_1 \leq \min_{i \in [N]} \sum_{t=s}^{e} l_t(i) + C_{\mathrm{dm}} \sqrt{(D+1) L \log(NT)}.$$

Finally, we substitute $l_t = g_t/M$ and $D = \lambda/M$ into the above inequality. Multiplying both sides by M scales the square root term as follows:

$$M \cdot C_{\mathrm{dm}} \sqrt{\left(\frac{\lambda}{M} + 1\right) L \log(NT)} = C_{\mathrm{dm}} \sqrt{M^2 \left(\frac{\lambda + M}{M}\right) L \log(NT)} = C_{\mathrm{dm}} \sqrt{M(M+\lambda) L \log(NT)}.$$

This exactly recovers Eq. (11) and completes the proof. □

E.4 Proof of the SAR guarantee

Proof of Theorem 12. Write $J_r = [a_r, b_r]$, so $a_1 = s$, $b_m = e$, and $b_r + 1 = a_{r+1}$ for $r < m$.

By Eq. (48):

$$\mathbb{E}[f_s(\boldsymbol{x}_s)] = \langle \boldsymbol{v}_s, \boldsymbol{\ell}_s \rangle \leq \langle \boldsymbol{v}_s, \boldsymbol{g}_s \rangle .$$

Summing the one step bound in Lemma 10 over $t = s+1, \ldots, e$, and using the above inequality yields

$$\mathbb{E}\left[\sum_{t=s}^{e} f_t(\boldsymbol{x}_t) + \lambda \sum_{t=s+1}^{e} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\} \right] \leq \sum_{t=s}^{e} \langle \boldsymbol{v}_t, \boldsymbol{g}_t \rangle + \frac{\lambda}{2} \sum_{t=s+1}^{e} \|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1. \tag{50}$$

Since Lemma 11 holds simultaneously for every interval under the single run that generates $(\boldsymbol{v}_t)$, we may apply it to each block J_r using the same sequence $(\boldsymbol{v}_t)$.

Since each J_r is a dyadic block, define $H_r \doteq |J_r| \in \mathcal{H}$. The expert with scale H_r restarts at the beginning of J_r, so on J_r it is exactly a fresh FPRLL(H_r) run.

Thus, for every r,

$$\sum_{t=a_r}^{b_r} \langle \boldsymbol{v}_t, \boldsymbol{g}_t \rangle + \frac{\lambda}{2} \sum_{t=a_r+1}^{b_r} \|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 \le \sum_{t=a_r}^{b_r} g_t^{(H_r)} + C_{\mathrm{dm}} \sqrt{M(M+\lambda)|J_r| \log(KT)}. \tag{51}$$

Summing Eq. (51) over $r = 1, \ldots, m$ gives

$$\sum_{r=1}^{m} \sum_{t=a_r}^{b_r} \langle v_t, g_t \rangle + \frac{\lambda}{2} \sum_{r=1}^{m} \sum_{t=a_r+1}^{b_r} \|v_t - v_{t-1}\|_1 \tag{52}$$

$$\le \sum_{r=1}^{m} \sum_{t \in J_r} g_t^{(H_r)} + C_{\mathrm{dm}} \sqrt{M(M+\lambda)\log(KT)} \sum_{r=1}^{m} \sqrt{|J_r|}. \tag{53}$$

Because the blocks are consecutive,

$$\sum_{t=s+1}^{e} \|v_t - v_{t-1}\|_1 = \sum_{r=1}^{m} \sum_{t=a_r+1}^{b_r} \|v_t - v_{t-1}\|_1 + \sum_{r=1}^{m-1} \|v_{a_{r+1}} - v_{a_{r+1}-1}\|_1.$$

Therefore,

$$\sum_{t=s}^{e} \langle v_t, g_t \rangle + \frac{\lambda}{2} \sum_{t=s+1}^{e} \|v_t - v_{t-1}\|_1 \tag{54}$$

$$\le \sum_{r=1}^{m} \sum_{t \in J_r} g_t^{(H_r)} + C_{\mathrm{dm}} \sqrt{M(M+\lambda)\log(KT)} \sum_{r=1}^{m} \sqrt{|J_r|} + \frac{\lambda}{2} \sum_{r=1}^{m-1} \|v_{a_{r+1}} - v_{a_{r+1}-1}\|_1. \tag{55}$$

Since $\boldsymbol{v}_t, \boldsymbol{v}_{t-1} \in \Delta_N$, one has $\|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 \le 2$, so each boundary term is at most λ. Thus

$$\sum_{t=s}^{e} \langle v_t, g_t \rangle + \frac{\lambda}{2} \sum_{t=s+1}^{e} \|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 \le \sum_{r=1}^{m} \sum_{t \in J_r} g_t^{(H_r)} + C_{\mathrm{dm}} \sqrt{M(M+\lambda)\log(KT)} \sum_{r=1}^{m} \sqrt{|J_r|} + \lambda(m-1). \tag{56}$$

Combining Eqs. (50) and (56), we obtain

$$\mathbb{E}\left[\sum_{t=s}^{e} f_t(x_t) + \lambda \sum_{t=s+1}^{e} \mathbf{1}\{x_t \ne x_{t-1}\} \right] \tag{57}$$

$$\le \sum_{r=1}^{m} \sum_{t \in J_r} g_t^{(H_r)} + C_{\mathrm{dm}} \sqrt{M(M+\lambda)\log(KT)} \sum_{r=1}^{m} \sqrt{|J_r|} + \lambda(m-1). \tag{58}$$

By Corollary 6, for any $\boldsymbol{u} \in \mathcal{X}$ and every r,

$$\sum_{t \in J_r} \left(g_t^{(H_r)} - f_t(\boldsymbol{u}) \right) \le C_{\mathrm{base}} \sqrt{|J_r|}. \tag{59}$$

summing over all the dyadic intervals we obtain

$$\sum_{r=1}^{m} \sum_{t \in J_r} g_t^{(H_r)} \le \min_{u \in \mathcal{X}} \sum_{t=s}^{e} f_t(u) + C_{\mathrm{base}} \sum_{r=1}^{m} \sqrt{|J_r|}. \tag{60}$$

Substituting Eq. (60) into Eq. (58), subtracting the comparator term from both sides, and using:

$$\lambda(m-1) \le \lambda(2\lceil \log_2 L \rceil + 1), \qquad \text{by the first part of Eq. (6)},$$

and

$$\sum_{r=1}^{m} \sqrt{|J_r|} \le C_{\mathrm{gc}} \sqrt{L}, \qquad \text{by the second part of Eq. (6)},$$

the lemma statement follows □

E.5 Proof of Corollary 13

Let $\{\boldsymbol{u}_t\}_{t=1}^T \in \mathcal{X}^T$ be an arbitrary comparator sequence, and let

$$S_T \doteq \sum_{t=2}^{T} \mathbf{1}\{\boldsymbol{u}_t \neq \boldsymbol{u}_{t-1}\}.$$

This sequence induces a partition of $[T]$ into $S_T + 1$ maximal consecutive segments on which $\boldsymbol{u}_t$ is constant,

$$\mathcal{S}_i = [s_i, e_i], \qquad i = 1, \ldots, S_T + 1,$$

with $s_1 = 1$, $e_{S_T+1} = T$, and $e_i + 1 = s_{i+1}$ for $i = 1, \ldots, S_T$. Writing $\boldsymbol{u}^{(i)}$ for the common value of $\boldsymbol{u}_t$ on $\mathcal{S}_i$,

$$\sum_{t=1}^{T} f_t(\boldsymbol{u}_t) = \sum_{i=1}^{S_T+1} \sum_{t \in \mathcal{S}_i} f_t(\boldsymbol{u}^{(i)}). \tag{61}$$

Decomposing the switching cost along the partition. The total indicator switching cost splits into within-segment and boundary contributions:

$$\sum_{t=2}^{T} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\} = \underbrace{\sum_{i=1}^{S_T+1} \sum_{t=s_i+1}^{e_i} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}}_{\text{within segments}} + \underbrace{\sum_{i=2}^{S_T+1} \mathbf{1}\{\boldsymbol{x}_{s_i} \neq \boldsymbol{x}_{s_i-1}\}}_{S_T \text{ boundary switches}}. \tag{62}$$

Combining Eqs. (61) and (62) and taking expectation,

$$\begin{aligned}
\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] &\leq \sum_{i=1}^{S_T+1} \left(\mathbb{E}\left[\sum_{t \in \mathcal{S}_i} f_t(\boldsymbol{x}_t) + \lambda \sum_{t=s_i+1}^{e_i} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\} \right] - \sum_{t \in \mathcal{S}_i} f_t(\boldsymbol{u}^{(i)}) \right) \\
&\quad + \lambda \sum_{i=2}^{S_T+1} \Pr\left(\boldsymbol{x}_{s_i} \neq \boldsymbol{x}_{s_i-1}\right) \\
&= \sum_{i=1}^{S_T+1} \mathrm{SAR}_{\mathcal{S}_i}(\boldsymbol{u}^{(i)}) + \lambda \sum_{i=2}^{S_T+1} \Pr\left(\boldsymbol{x}_{s_i} \neq \boldsymbol{x}_{s_i-1}\right),
\end{aligned} \tag{63}$$

where the last equality uses the definition of SAR in Eq. (12). Each boundary probability is at most one, so the second sum is bounded by λS_T.

Applying the SAR guarantee. By Theorem 12, for every segment $\mathcal{S}_i$,

$$\mathrm{SAR}_{\mathcal{S}_i}(\boldsymbol{u}^{(i)}) = \tilde{\mathcal{O}}\left(\sqrt{|\mathcal{S}_i|}\right).$$

Summing over i and applying Cauchy–Schwarz,

$$\sum_{i=1}^{S_T+1} \sqrt{|\mathcal{S}_i|} \leq \sqrt{S_T + 1} \sqrt{\sum_{i=1}^{S_T+1} |\mathcal{S}_i|} = \sqrt{(S_T+1)\,T}. \tag{64}$$

Substituting into Eq. (63),

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \leq \tilde{\mathcal{O}}\left(\sqrt{(S_T+1)\,T}\right) + \lambda S_T,$$

which is the claimed bound. □

E.6 Proof for the DR guarantee

Proof of Theorem 14. Fix any comparator sequence $\{\boldsymbol{u}_t\}_{t=1}^T \subseteq \mathcal{X}$. By definition,

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le \mathbb{E}\left[\sum_{t=1}^T f_t(\boldsymbol{x}_t)\right] - \sum_{t=1}^T f_t(\boldsymbol{u}_t) + \lambda \sum_{t=2}^T \Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}).$$

For $t = 1$, since $c_1^{(H)} = 0$ for every $H \in \mathcal{H}$, we have $g_1^{(H)} = \ell_1^{(H)}$, and therefore

$$\mathbb{E}\left[f_1(\boldsymbol{x}_1)\right] = \langle v_1, g_1 \rangle.$$

For each $t \ge 2$, Lemma 10 gives

$$\mathbb{E}\left[f_t(\boldsymbol{x}_t)\right] + \lambda \Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) \le \langle v_t, g_t \rangle + \frac{\lambda}{2}\|v_t - v_{t-1}\|_1.$$

Summing over $t = 1, \dots, T$, we obtain

$$\mathbb{E}\left[\sum_{t=1}^T f_t(\boldsymbol{x}_t)\right] + \lambda \sum_{t=2}^T \Pr(\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}) \le \sum_{t=1}^T \langle v_t, g_t \rangle + \frac{\lambda}{2}\sum_{t=2}^T \|v_t - v_{t-1}\|_1.$$

Subtracting $\sum_{t=1}^T f_t(\boldsymbol{u}_t)$, we get

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le \sum_{t=1}^T \langle v_t, g_t \rangle + \frac{\lambda}{2}\sum_{t=2}^T \|v_t - v_{t-1}\|_1 - \sum_{t=1}^T f_t(\boldsymbol{u}_t). \tag{65}$$

Now fix any $H \in \mathcal{H}$, and add and subtract $\sum_{t=1}^T g_t^{(H)}$ on the right-hand side of Eq. (65). This yields

$$\begin{aligned}\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le &\left(\sum_{t=1}^T \langle v_t, g_t \rangle + \frac{\lambda}{2}\sum_{t=2}^T \|v_t - v_{t-1}\|_1 - \sum_{t=1}^T g_t^{(H)}\right) \\ &+ \left(\sum_{t=1}^T g_t^{(H)} - \sum_{t=1}^T f_t(\boldsymbol{u}_t)\right).\end{aligned}$$

We now control the two brackets separately. The first is the regret of the meta-learner relative to expert H. Applying Lemma 11 on the full interval $[1, T]$, whose length is T, gives

$$\sum_{t=1}^T \langle v_t, g_t \rangle + \frac{\lambda}{2}\sum_{t=2}^T \|v_t - v_{t-1}\|_1 - \sum_{t=1}^T g_t^{(H)} \le C_{\mathrm{dm}} \sqrt{M(M+\lambda)\, T \log(KT)}.$$

Hence, for every $H \in \mathcal{H}$,

$$\mathbb{E}\left[\mathcal{R}_T^{\mathbf{1}}(\{\boldsymbol{u}_t\}_{t=1}^T)\right] \le \sum_{t=1}^T g_t^{(H)} - \sum_{t=1}^T f_t(\boldsymbol{u}_t) + C_{\mathrm{dm}} \sqrt{M(M+\lambda)\, T \log(KT)}.$$

Taking the minimum over $H \in \mathcal{H}$ proves Eq. (14).

It remains to instantiate the expert term. By the proof of Theorem 7, applied with $\lambda_t \equiv \lambda$ and using $c_t^{(H)} \le 1$ at restart boundaries, there exists a dyadic restart scale $k^\dagger \in \mathcal{H}$ such that

$$\sum_{t=1}^T g_t^{(k^\dagger)} - \sum_{t=1}^T f_t(\boldsymbol{u}_t) \le \sqrt{2}\, A\, \frac{T}{\sqrt{k^\star}} + (B+G) P_T\, k^\star.$$

Substituting $H = k^\dagger$ into Eq. (14) gives Eq. (15). □

F Sampled-surrogate guarantees and oracle accounting

We now explain why using $\hat{\boldsymbol{g}}_t$ from Section 5 preserves the strongly adaptive and dynamic-regret rates in Section 4. The original analysis uses the surrogate vector in three places, which we check below. Section F.2 then details the computational costs.

F.1 Why the regret guarantees still hold

Fix an oblivious loss sequence and use the definitions in Eqs. (16) to (18).

1. The one-step master reduction. We first check that the sampled scores still control the expected cost of the actions actually played. Let $\mathcal{Y}$ denote all the auxiliary samples $\{\boldsymbol{y}_t^{(H)}\}_{t,H}$. For the fixed loss sequence, the base densities are deterministic. Conditional on $\mathcal{Y}$, the simplex weights from Section E.3 and the mixtures $\mathcal{P}_t$ are also fixed, while the randomness used to sample the played actions is independent of $\mathcal{Y}$. Hence Lemma 16 gives the required action marginals and switching probabilities, and Lemma 10 applies conditional on $\mathcal{Y}$.

The weights $\boldsymbol{v}_t$ are chosen before the round-t auxiliary samples are drawn. Thus Eq. (18) gives

$$\mathbb{E}[\langle \boldsymbol{v}_t, \boldsymbol{g}_t \rangle] \le \mathbb{E}[\langle \boldsymbol{v}_t, \bar{\boldsymbol{g}}_t \rangle] = \mathbb{E}[\langle \boldsymbol{v}_t, \hat{\boldsymbol{g}}_t \rangle].$$

Averaging the conditional one-step bound and summing over any interval $I = [s, e]$ therefore yields

$$\mathbb{E}\left[\sum_{t=s}^{e} f_t(\boldsymbol{x}_t) + \lambda \sum_{t=s+1}^{e} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right] \le \mathbb{E}\left[\sum_{t=s}^{e} \langle \boldsymbol{v}_t, \hat{\boldsymbol{g}}_t \rangle + \frac{\lambda}{2} \sum_{t=s+1}^{e} \|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1\right]. \tag{66}$$

At the first round s, we use $\mathbb{E}[f_s(\boldsymbol{x}_s)] = \mathbb{E}[\langle \boldsymbol{v}_s, \boldsymbol{\ell}_s \rangle] \le \mathbb{E}[\langle \boldsymbol{v}_s, \hat{\boldsymbol{g}}_s \rangle]$; no switch into the interval is charged on the left. Thus the master cost is bounded by the sampled scores fed to the meta-learner, together with its usual weight-movement penalty.

2. The meta-regret bound. The meta-learner receives the realized sequence $\hat{\boldsymbol{g}}_1, \dots, \hat{\boldsymbol{g}}_T$. These scores satisfy $0 \le \hat{g}_t^{(H)} \le M$, so Lemma 11 applies to every realization. For $I = [s, e]$ of length $L = e - s + 1$,

$$\sum_{t=s}^{e} \langle \boldsymbol{v}_t, \hat{\boldsymbol{g}}_t \rangle + \frac{\lambda}{2} \sum_{t=s+1}^{e} \|\boldsymbol{v}_t - \boldsymbol{v}_{t-1}\|_1 \le \min_{H \in \mathcal{H}} \sum_{t=s}^{e} \hat{g}_t^{(H)} + C_{\mathrm{dm}} \sqrt{M(M+\lambda) L \log(KT)}.$$

Taking expectations, using $\mathbb{E}[\min_H Z_H] \le \min_H \mathbb{E}[Z_H]$ and $\mathbb{E}[\hat{g}_t^{(H)}] = \bar{g}_t^{(H)}$, and combining with Eq. (66) gives

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=s}^{e} f_t(\boldsymbol{x}_t) + \lambda \sum_{t=s+1}^{e} \mathbf{1}\{\boldsymbol{x}_t \neq \boldsymbol{x}_{t-1}\}\right] \\ &\qquad \le \min_{H \in \mathcal{H}} \sum_{t=s}^{e} \left(\ell_t^{(H)} + \lambda \hat{c}_t^{(H)}\right) + C_{\mathrm{dm}} \sqrt{M(M+\lambda) L \log(KT)}. \end{aligned} \tag{67}$$

On $[1, T]$, subtracting the comparator losses gives Eq. (14) with $g_t^{(H)}$ replaced by $\bar{g}_t^{(H)} = \ell_t^{(H)} + \lambda \hat{c}_t^{(H)}$ on its right-hand side.

3. The best-expert term after the meta-regret bound. It remains to bound the expert cost on the right of Eq. (67). Compared with the original surrogate, its expected hitting loss $\ell_t^{(H)}$ is unchanged; only $c_t^{(H)}$ is replaced by $\hat{c}_t^{(H)}$. This replacement uses the switching bounds already charged in the base-learner proofs: $\hat{c}_t^{(H)} \le \epsilon_H$ within a block, and $\hat{c}_t^{(H)} = 1$ at a restart. Thus, for a restart block J,

$$\lambda \sum_{t \in J} \hat{c}_t^{(H)} \le \lambda + \lambda(|J| - 1)\epsilon_H,$$

with the initial λ omitted when J starts at round 1. The proofs of Corollary 6 and Theorem 7 already use these same within-block and restart charges. In particular, the proof of Corollary 6 gives, on every full restart block J of length H and for every fixed comparator,

$$\sum_{t \in J} \left(\bar{g}_t^{(H)} - f_t(\boldsymbol{u})\right) \le C_{\mathrm{base}} \sqrt{H},$$

including the allowance for a switch at the block's start. Using this bound in Eq. (67) on each matched dyadic block, with the same geometric cover and boundary accounting, proves the strongly adaptive bound Eq. (13). Segmenting the comparator sequence then gives the dynamic-regret bound Corollary 13.

For path-length regret, the proof of Theorem 7 likewise bounds $\sum_{t=1}^{T}(\bar{g}_t^{(H)} - f_t(\boldsymbol{u}_t))$ using the same switching charges. Applying this bound in Eq. (67) on $[1, T]$ and choosing the same dyadic scale $k^\dagger$ gives Eq. (15). Hence all three final regret rates hold in expectation for the sampled implementation.

F.2 Implementation and cost of the base primitives

Section 5 counts calls to two FPRLL operations: sampling (BS) and density evaluation (BE). Here we describe their implementation and dependence on the dimension d. We assume second-order oracle access to the accumulated loss, barrier and regularizer, following the computational model of Sherman and Koren [19, Sec. 3].

Sampling (BS). To sample from $\mathcal{Q}_t^{(H)}$, draw a Laplace perturbation with the parameters of the H-expert and solve the perturbed convex minimization defining its fresh action in Eq. (4). Thus one base-sampling call requires one convex minimization over the barrier domain.

Density evaluation (BE). The closed form in Eq. (5) requires the gradient and Hessian of the base learner's accumulated objective, an evaluation of the Laplace density, and one $d \times d$ determinant. Given the derivatives, the Laplace-density evaluation takes $\mathcal{O}(d)$ arithmetic and the determinant takes $\mathcal{O}(d^3)$ by dense factorization. This is the change-of-variables implementation of Sherman and Koren [19, Lem. 18] and Agarwal et al. [2, Lem. 2].

Overall cost. These base operations admit polynomial-time implementations in the second-order oracle model [19, Sec. 3]. By the counts in Section 5, each round requires $\mathcal{O}(\log T)$ expected calls to each base primitive, $\mathcal{O}(\log T)$ loss evaluations, and $\mathcal{O}(\log^2 T)$ additional arithmetic. Thus our framework adds only polylogarithmic expected overhead and has polynomial per-round dependence on d whenever the base primitives and loss evaluations have polynomial cost. The logarithmic number of base learners comes from covering the unknown restart scale by a dyadic grid.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the paper's theoretical contributions and scope, and the main claims are matched by the theorem statements and appendix proofs.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

Justification: The paper discusses the main assumptions and scope of the results, including convexity, smoothness/Lipschitzness, compact decision sets, and the online-learning setting considered.

Guidelines:

- The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The paper is primarily theoretical. The assumptions are directly at the beginning of contribution section. The full proofs are provided in the appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

Justification: The paper does not include experimental results; the main claims are theoretical.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: The paper does not include experiments requiring code or data to reproduce empirical results.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.

- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does not include experiments, training procedures, or test evaluations.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper does not include experiments, so statistical significance reporting is not applicable.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: The paper does not include experiments requiring computational resources.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The work is theoretical and does not involve human subjects, private data, deployed systems, or high-risk applications; it conforms with the NeurIPS Code of Ethics.

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: The work is foundational online-learning theory and does not introduce a deployed system or application with direct societal impact.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper does not release data, models, or artifacts with high misuse risk.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [N/A]

Justification: The paper does not use existing datasets, pretrained models, or code assets as part of the contribution. Prior theoretical results are cited in the bibliography.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [N/A]

Justification: The paper does not introduce or release new datasets, models, or code assets.

Guidelines:

- The answer [N/A] means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.

- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or human-subject research, so IRB approval is not applicable.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs are not used as part of the core method, theory, proofs, or scientific contribution of the paper.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.