From Non-Convex to Strongly Convex: Curvature-Adaptive FTPL for Online Optimization

Moses Charikar
Stanford University
moses@cs.stanford.edu

Chirag Pabbaraju
Stanford University
cpabbara@cs.stanford.edu

Ambuj Tewari
University of Michigan, Ann Arbor
tewaria@umich.edu

Abstract

Curvature adaptivity is a classical theme in online optimization: for convex Lipschitz losses, adaptive methods interpolate between the optimal $O(\sqrt{T})$ regret for general convex losses and $O(\log T)$ regret under strong convexity. Recent work has shown that Follow-the-Perturbed-Leader (FTPL) achieves optimal $O(\sqrt{T})$ regret even for online non-convex Lipschitz losses, assuming access to an approximate offline-optimization oracle, but these guarantees do not exploit curvature. We show that FTPL can be made curvature-adaptive in the non-convex setting, without knowing in advance how curvature will accumulate over time. Our algorithm replaces the fixed perturbation scale of standard FTPL with a time-varying scale chosen using only past information. We give a simple follow-the-leader tuning rule for this scale and show that it competes, up to constants, with the best choice in hindsight. The resulting method achieves $O(\sqrt{T})$ regret for arbitrary non-convex Lipschitz losses and improves as cumulative curvature grows; with sufficiently accurate oracle calls, it achieves $O(\log T)$ regret when cumulative curvature grows linearly, which includes the classical strongly convex regime. We complement these upper bounds with matching lower bounds for prescribed cumulative-curvature sequences, already for one-dimensional convex losses, showing that the tradeoff between worst-case non-convex regret and curvature-driven fast rates is intrinsic.

1 Introduction

A central question in online optimization is how regret scales with the number of rounds T. Two classical results characterize this dependence: for convex and Lipschitz losses, the optimal regret is $O(\sqrt{T})$ [Zinkevich, 2003], while for strongly convex and Lipschitz losses, it improves to $O(\log T)$ [Hazan et al., 2006]. Both rates are known to be tight [Abernethy et al., 2008].

A natural goal, therefore, is to design algorithms whose guarantees adapt to the curvature of the loss sequence. In a seminal work, Bartlett et al. [2007] introduced adaptive gradient methods for online *convex* optimization whose regret seamlessly interpolates between $O(\sqrt{T})$ and $O(\log T)$ depending on the cumulative strong convexity in the sequence. These ideas have since influenced widely used methods such as Adagrad [Duchi et al., 2011] and Adam [Kingma and Ba, 2015]. More recently, Suggala and Netrapalli [2020] showed that the Follow-the-Perturbed-Leader (FTPL) framework achieves $O(\sqrt{T})$ regret even for *non-convex* Lipschitz losses, assuming access to an oracle that solves perturbed offline problems.

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

These results naturally raise the following question:

> *Can FTPL be made adaptive to curvature, achieving $O(\sqrt{T})$ regret for general non-convex Lipschitz losses while improving to $O(\log T)$ when sufficient strong convexity is present?*

In this paper, we answer this question in the affirmative. Our approach is based on introducing a time-varying noise scale in the FTPL perturbations, chosen adaptively using only past information. We first derive a general regret bound parameterized by this sequence of noise scales. We then show that selecting these parameters can itself be cast as a meta online learning problem, for which a simple FTL approach achieves a constant competitive ratio. The resulting algorithm achieves regret that interpolates between $O(\sqrt{T})$ and $O(\log T)$, without prior knowledge of the curvature of the losses. We further complement our upper bounds with matching lower bounds, showing that the bounds we meta-optimize are intrinsic to the problem and not an artifact of our method.

Beyond its theoretical interest, our setting is motivated by a broad class of problems where the learner encounters composite losses of the form

$$f_t = g_t + r_t,$$

where g_t is a data-dependent term that may be non-convex, and r_t is a strongly convex regularizer. Such composite structures are ubiquitous in modern machine learning and optimization.

One prominent example arises in continual learning [Kirkpatrick et al., 2017], where g_t represents a task-specific loss, typically non-convex due to neural network parameterizations, and r_t is a regularizer that *tethers the current iterate to previously learned models*, thereby mitigating catastrophic forgetting. In contrast, in online robust regression [Maronna et al., 2019], even with a linear model, the loss g_t can be non-convex when using robust losses such as the Tukey biweight function, while r_t is typically a strongly convex shrinkage term (e.g., ℓ_2 regularization) that *tethers the iterate to a fixed reference point*, such as the origin.

In these settings, the effective curvature of the loss sequence depends on the interplay between the non-convex data term and the accumulated regularization. In favorable regimes, the regularizer dominates and the problem behaves like a strongly convex one; in adversarial regimes, the non-convexity prevails. Since this balance is typically unknown a priori, it is essential to design algorithms that automatically adapt, achieving $O(\sqrt{T})$ regret in the worst case while improving to $O(\log T)$ whenever sufficient curvature emerges.

Our key contributions are as follows.

- **Curvature-adaptive FTPL across non-convex to strongly convex regimes.** We design an FTPL algorithm with a time-varying perturbation scale that adapts to the curvature of the loss sequence, achieving regret that interpolates from $O(\sqrt{T})$ for arbitrary non-convex Lipschitz losses to $O(\log T)$ when sufficient strong convexity is present. To our knowledge, this is the first FTPL method that achieves such curvature adaptivity in the non-convex regime, bridging the gap between prior adaptive methods for online convex optimization (e.g., [Bartlett et al., 2007]) and non-convex but non-adaptive FTPL guarantees (e.g., [Suggala and Netrapalli, 2020]).
- **Meta-online tuning and optimality guarantees.** We derive a regret bound parameterized by the perturbation schedule and show that selecting this schedule can be cast as a meta online learning problem. We give a simple constant-competitive strategy for this problem, yielding near-optimal tuning without prior knowledge of the loss sequence. We also prove matching lower bounds, showing that the bounds we meta-optimize are intrinsic to the problem.
- **Empirical validation.** We empirically evaluate our method on representative problems with composite non-convex and strongly convex structure, demonstrating improved performance over non-adaptive FTPL and baseline methods in regimes with varying curvature.

2 Related Works

Oracles in Online Optimization. The FTPL algorithm [Hannan, 1957, Kalai and Vempala, 2005] is especially attractive for efficient online learning over structured decision sets. The main computa-

tional burden in FTPL at each round is solving a perturbed offline problem. If the cost functions and perturbation are both linear, then the perturbed offline problem becomes an instance of the offline optimization problem itself. However, for general decision sets, there is known to be an exponential (in the ambient dimension d) gap in the oracle complexities of online and statistical learning if the learner is only provided access to the offline optimization oracle [Hazan and Koren, 2016]. For decision classes with exponentially many experts in the ambient dimension, the lower bound of Hazan and Koren [2016] implies that no algorithm using only a standard offline optimization oracle and polynomial additional computation can achieve $O(\text{poly}(d)T^{\alpha})$ regret for any $\alpha < 1$. It was therefore surprising when Agarwal et al. [2019] showed that FTPL with linear perturbations, which makes one call per time step to a *linearly perturbed* offline optimization oracle, achieves $O(\text{poly}(d)T^{2/3})$ cumulative regret against *non-convex* Lipschitz functions on a bounded decision set. The $O(\text{poly}(d)T^{2/3})$ rate was later improved to $O(\text{poly}(d)\sqrt{T})$ by Suggala and Netrapalli [2020].

Adaptive Step-Sizes and Regularization. Online gradient algorithms with different step-size schedules achieve $O(\sqrt{T})$ and $O(\log T)$ regret against convex Lipschitz and strongly convex Lipschitz functions respectively [Hazan et al., 2006, Zinkevich, 2003]. Bartlett et al. [2007] designed a single adaptive algorithm that interpolates between these rates and automatically adapts to the actual amount of curvature present in the adversary's *convex* functions. Some related lower bounds can be found in the work of Abernethy et al. [2008]. Subsequent work developed a broader theory of adaptive regularization and adaptive first-order methods, including AdaGrad-style algorithms [Duchi et al., 2011, McMahan and Streeter, 2010] and adaptive FTRL/mirror-descent frameworks [McMahan, 2017]. These ideas later influenced practical optimization methods for deep learning such as Adam [Kingma and Ba, 2015]. While adaptive step-size methods are also widely studied for stochastic non-convex optimization [Li and Orabona, 2019, Ward et al., 2020], their guarantees are typically formulated in terms of convergence to stationary points rather than regret minimization. Also note that, as observed by Agarwal et al. [2019], standard FTRL-style regularization does not necessarily stabilize the minimizers of non-convex objectives, limiting the direct applicability of FTRL-based approaches for provable regret guarantees in the non-convex regime.

Adaptive Perturbation. Adaptive variants of FTPL have previously been studied through adaptive learning-rate schedules, data-dependent perturbation magnitudes, and generalized FTPL frameworks for oracle-efficient online learning [Hutter and Poland, 2005, Wang et al., 2022]. Existing works primarily focus on adapting to favorable data sequences, for example when the loss of the best expert is small or the adversary is stochastic, while retaining the classical $O(\sqrt{T})$ dependence on the horizon. In contrast, our goal is to adapt to curvature itself. To the best of our knowledge, no prior online optimization algorithm, let alone one that is perturbation-based for oracle-efficiency, achieves a full interpolation between the optimal $O(\sqrt{T})$ regret for Lipschitz non-convex losses and the $O(\log T)$ regret achievable for strongly convex losses.

Universal Online Learning. Yan et al. [2023] adapt to unknown curvature in online convex optimization by aggregating learners over a geometric grid of curvature parameters. Their fast-rate analysis exploits convexity and curvature of individual losses, which need not hold in our setting even when cumulative losses are strongly convex. Consequently, their aggregation guarantees do not directly extend: generic expert aggregation can hurt the desired fast rates, while a direct ensemble of FTPL learners generally requires multiple optimization oracle calls per round. Our method instead uses previously observed cumulative curvature to tune a time-varying perturbation schedule, achieving curvature-dependent rates with one oracle call per round.

3 Adaptive FTPL

Problem Setup. The setup comprises of an online game between an *adversary* and a *learner* that proceeds in rounds. At round t, the learner chooses some $x_t \in \mathcal{X}$, where $\mathcal{X}$ is a bounded space; simultaneously, the adversary picks a function $f_t : \mathcal{X} \to \mathbb{R}$. Both the learner's and adversary's choices at round t can depend on the entire history of the game up until round $t-1$. Thereafter, the choices made by the learner and adversary are revealed to each other, and the learner suffers a loss $f_t(x_t)$. The learner's objective is to minimize their *regret* at the end of the game; that is, the learner

aims to minimize the cumulative loss they suffer at the end of T rounds, relative to the single best $x \in \mathcal{X}$ they could have played at each round in hindsight. Formally, define the regret of a sequence of moves $x_{1:T} := x_1, \ldots, x_T$ with respect to a sequence of functions $f_{1:T} := f_1, \ldots, f_T$ as

$$\text{Regret}_T(f_{1:T}; x_{1:T}) := \sum_{t=1}^{T} f_t(x_t) - \inf_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(x). \tag{1}$$

We will assume that $\mathcal{X} \subseteq \mathbb{R}^d$ with ℓ_∞ diameter $D < \infty$. We will also assume that the functions f_t that the adversary can play are guaranteed to be L-Lipschitz with respect to the ℓ_1 norm. That is, for every $x, y \in \mathcal{X}$, and for every t, $|f_t(x) - f_t(y)| \leq L\|x - y\|_1$.

Curvature. We say that a function $f : \mathcal{X} \to \mathbb{R}$ has strong convexity constant $\mu > 0$ (with respect to ℓ_1) if for every $x, y \in \mathcal{X}$, it holds that

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|y - x\|_1^2. \tag{2}$$

We use the convention that $\mu = 0$ if (2) above does not hold for f for any $\mu > 0$.

We note that strong convexity can be defined with respect to many norms/Bregman divergences, and there are more generalized notions, e.g., uniform convexity, that include strong convexity as a special case. These are all valid notions of "curvature". In this paper, we fix one specific notion of strong convexity, namely (2) with the ℓ_1 norm.

The "curvature-adaptivity" that we will be concerned with is not across different notions of curvature, but across the magnitude of curvature under this fixed notion of strong convexity. In particular, we will seek an algorithm that exploits the (cumulative) strong convexity seen so far. This is directly analogous to Bartlett et al. [2007]: their work also seeks to adapt to the cumulative curvature (for an explicitly fixed notion of curvature) observed so far. While Bartlett et al. [2007] establish this phenomenon for online convex optimization, our contribution is to establish the analogous curvature-adaptive phenomenon for online non-convex optimization.

Approximate Optimization Oracle. As in Suggala and Netrapalli [2020], we assume that the learner has access to an "(α, β)-approximate optimization oracle". Given a function $f : \mathcal{X} \to \mathbb{R}$ and a vector $\sigma \in \mathbb{R}^d$, the oracle $\mathcal{O}_{\alpha,\beta}(f, \sigma)$ returns $\hat{x} \in \mathcal{X}$ satisfying

$$f(\hat{x}) - \langle \sigma, \hat{x} \rangle \leq \inf_{x \in \mathcal{X}} \left(f(x) - \langle \sigma, x \rangle\right) + (\alpha + \beta\|\sigma\|_1). \tag{3}$$

From a worst case complexity point of view, most interesting non-convex problems will not have efficient oracles, barring a surprising collapse of complexity classes. However, note that the benchmark term in the regret definition (1) in a non-convex problem *itself* solves a difficult offline (offline because all losses are available to the comparator) optimization problem. When the offline optimization problem is computationally hard, one cannot hope for an online algorithm that is substantially easier in full generality. For this reason, when the offline oracle solves an intractable problem, the standard oracle-efficient viewpoint in online learning is to keep the computational complexity no worse than that of the offline problem (e.g., see Hazan and Koren [2016]) by making sure that the online learner only needs calls to an exact or approximate offline optimization oracle. We nevertheless note that in practice, many worst-case hard problems do admit highly effective heuristic solvers, primary examples being the successes of gradient-based optimization in deep learning and various specialized solvers for robust regression.

We can now state our main Adaptive FTPL algorithm, which takes as input the oracle $\mathcal{O}_{\alpha,\beta}$, and a non-decreasing schedule $\rho_1, \rho_2, \ldots$ of *noise scales* to invoke the oracle with.

Algorithm 1 Adaptive Follow the Perturbed Leader (AdaFTPL)

Input: Approximate optimization oracle $\mathcal{O}_{\alpha,\beta}$, non-decreasing noise scale schedule $\rho_1, \rho_2 \ldots,$
for $t = 1 \ldots T$ **do**
  Generate random vector σ_t such that $\{\sigma_{t,j}\}_{j=1}^{d} \overset{\text{i.i.d.}}{\sim} \text{Exp}(1)$[1]
  Let $F_{t-1} = \sum_{i=1}^{t-1} f_i$; predict x_t as

$$x_t = \mathcal{O}_{\alpha,\beta}\left(F_{t-1},\, \rho_t \sigma_t\right)$$

  Observe loss function f_t
end for

We are now ready to state our main regret guarantee for Adaptive FTPL. We note that throughout the paper, we will use the convention $1/0 = +\infty$.

Theorem 1 (Adaptive FTPL Regret)**.** *Let D be the ℓ_∞ diameter of $\mathcal{X}$, and suppose that the functions $f_1, \ldots, f_T$ encountered by Adaptive FTPL (Algorithm 1) are L-Lipschitz with respect to the ℓ_1 norm. Furthermore, let S_t be the strong convexity constant of $F_t = \sum_{i=1}^{t} f_i$. Then, the algorithm has the following regret guarantee:*

$$\mathbb{E}_\sigma\left[\text{Regret}_T(f_{1:T}; x_{1:T})\right] \le \sum_{t=1}^{T} \min\left\{ \frac{125L^2d^2D}{\rho_t} + \frac{\alpha + \beta\rho_t d}{20} + 2\beta dL,\, \frac{2L^2}{S_t} + 2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}} \right\} + dD\rho_T + \alpha T + d\beta \sum_{t=1}^{T} \rho_t. \tag{4}$$

We give the complete proof of Theorem 1 in Section A, but provide a detailed sketch here.

Proof Sketch. By a standard argument for FTPL (e.g., see Lemma 12 in Hutter and Poland [2005], Chapter 4 in Cesa-Bianchi and Lugosi [2006]), it suffices to study the regret of Algorithm 1 in the setting where the functions $f_1, \ldots, f_T$ are specified upfront by an *oblivious* adversary, and show the regret bound for every fixed $f_1, \ldots, f_T$; so, fix such an oblivious adversary. Now, for an oblivious adversary, expected regret does not change if instead of drawing a fresh σ_t at each time step, we draw a random vector $\sigma = \{\sigma_j\}_{j=1}^{d} \overset{\text{i.i.d.}}{\sim} \text{Exp}(1)$ once at the start and reuse it at every time step. With these considerations, our goal is to upper bound the regret in the left-hand side of (4) for fixed $f_1, \ldots, f_T$, where the expectation is only with respect to σ drawn at the start by the algorithm.

The key object that enables the analysis is a "ghost" move $\bar{x}_{t+1} = \mathcal{O}_{\alpha,\beta}(F_t, \rho_t\sigma)$. Note that $\bar{x}_{t+1}$ is separate from $x_t = \mathcal{O}_{\alpha,\beta}(F_{t-1}, \rho_t\sigma)$ which the algorithm plays at time t, *as well as* $x_{t+1} = \mathcal{O}_{\alpha,\beta}(F_t, \rho_{t+1}\sigma)$ which the algorithm plays at time step $t+1$. We introduce $\bar{x}_{t+1}$ solely for the sake of analysis, and for the reason that it uses a noise vector whose scale *matches* that of the noise vector used in the *previous* time step. This will help us (1) establish a standard decomposition of the regret into a *stability* and *perturbation* term, and (2) directly utilize the bound on the stability term from Suggala and Netrapalli [2020].

Namely, we show that

$$\mathbb{E}_\sigma\left[\text{Regret}_T(f_{1:T}; x_{1:T})\right] \le \underbrace{L\sum_{t=1}^{T} \mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right]}_{\text{stability}} + \underbrace{dD\rho_T + \alpha T + d\beta\sum_{t=1}^{T}\rho_t}_{\text{perturbation}}. \tag{5}$$

This decomposition follows a standard analysis once the ghost object $\bar{x}_{t+1}$ has been defined. The stability term arises directly from Lipschitzness. In order to bound the rest of the perturbation terms, we crucially observe that $\bar{x}_{t+1}$ is a move that the learner would have *liked to play* at time step t, but cannot do so, since it does not know f_t. Thus, we can employ a standard "be-the-leader" analysis for $\bar{x}_{t+1}$ to bound the perturbation terms in the desired form.

[1] A random variable $X \sim \text{Exp}(\eta)$ if for every $s \ge 0$, $\Pr[X \ge s] = \exp(-\eta s)$.

It then remains to bound the stability terms $\mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right]$. We do this in two ways, allowing us to interpolate between non-convexity and strong convexity. First, we observe again that both x_t and $\bar{x}_{t+1}$ use a noise vector having the same scale ρ_t; the only difference is that x_t invokes the oracle with F_{t-1}, whereas $\bar{x}_{t+1}$ invokes the oracle with F_t. But this means that we can directly use the analysis of Suggala and Netrapalli [2020] — in their setting, they upper-bound $\mathbb{E}_\sigma\left[\|x_t - x_{t+1}\|_1\right]$ where both x_t and x_{t+1} use the same noise vector; the same upper bound goes through for us, provided we replace x_{t+1} with $\bar{x}_{t+1}$.

Second, we use a different upper bound on $\|x_t - \bar{x}_{t+1}\|$ arising from the analysis of follow-the-leader (FTL) for strongly convex online optimization. Namely, if we assume that the strong convexity constant of the cumulative function F_t is S_t, then $\|x_t - \bar{x}_{t+1}\|$ can be upper-bounded as $O(L/S_t)$. This gives a different, curvature-cognizant upper bound on the perturbation term. This bound is vacuous when F_t is not strongly convex (given that we are using the convention $1/0 = +\infty$), but it nevertheless always applies.

Finally, given that we have two upper bounds for the stability term, both of which always hold, we can use the *minimum* of these two. This gives us the final regret upper bound in Theorem 1. □

3.1 A meta-online learning problem to choose ρ_t

Observe that the regret bound in Theorem 1 holds for any non-decreasing $\rho_1, \ldots, \rho_T$. We now consider the problem of optimally choosing this sequence. A natural way to do this is to optimize the upper bound that we obtain, as a function of $\rho_1, \ldots, \rho_T$.

Namely, for any horizon T, and error parameters α, β, define the *offline* optimal value of the regret upper bound over all possible non-decreasing noise scale schedules $\rho_1 \leq \cdots \leq \rho_T$ that are lower-bounded by $r \geq 0$ (i.e., $r \leq \rho_1$), as:

$$\mathrm{OPT}_T^{\alpha,\beta}(r) := \inf_{r \leq \rho_1 \leq \cdots \leq \rho_T} \left[\sum_{t=1}^{T} \min\left\{ \frac{125L^2d^2D}{\rho_t} + \frac{\beta\rho_t d}{20} + 2L\beta d + \frac{\alpha}{20}, \frac{2L^2}{S_t} + 2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}} \right\} + dD\rho_T + \alpha T + d\beta \sum_{t=1}^{T} \rho_t \right]. \tag{6}$$

When $\alpha = \beta = 0$, we denote the quantity above simply as $\mathrm{OPT}_T^0(r)$. In this case, a single consistent choice of ρ at all time steps suffices, i.e.,

$$\mathrm{OPT}_T^0(r) = \inf_{\rho \geq r} \left[\sum_{t=1}^{T} \min\left\{ \frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t} \right\} + dD\rho \right]. \tag{7}$$

This follows by noting that, for every feasible $r \leq \rho_1 \leq \cdots \leq \rho_T$, we can simply set $\rho_t = \rho_T$ for every t; this keeps the sequence feasible, and does not increase any individual term inside the summation.

Given this, a natural ρ_t that the algorithm can choose, solely as a function of the information it has seen so far, is $\mathrm{OPT}_{t-1}^0(r)$. This is the value of ρ_t that would realize the best offline regret in hindsight, had the game been stopped before round t, and corresponds to an FTL move itself in the space of noise schedules. This choice of ρ_t turns out to be non-decreasing; more interestingly, we can show that the regret of Algorithm 1 instantiated with this choice of ρ_t is constant-factor competitive with $\mathrm{OPT}_T^{\alpha,\beta}(r)$.

Theorem 2 (FTL for Noise Scale Schedule). *Fix $r > 0$. Let $\rho_1 = r$, and for $t \geq 2$,*

$$\rho_t = \min \arg\inf_{\rho \geq r} \left[\sum_{i=1}^{t-1} \min\left\{ \frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_i} \right\} + dD\rho \right], \tag{8}$$

Then,

1. *$\rho_1 \leq \cdots \leq \rho_T$, and every ρ_t depends only on $S_1, \ldots, S_{t-1}$.*
2. *For this noise scale schedule, Algorithm 1 has the following regret guarantee:*

(a) With $\alpha = \beta = 0$, we have $\mathbb{E}_\sigma\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \le C_1 \cdot \mathrm{OPT}_T^0(r)$.

(b) If $\beta \le c/T$, we have $\mathbb{E}_\sigma\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \le C_2 \cdot \mathrm{OPT}_T^{\alpha,\beta}(r)$.

Here, C_1 and C_2 are constants depending only on d, D, L, r, and c.

We note that ρ_t as given by (8) can be efficiently computed, provided that we are given the values $S_1, \ldots, S_{t-1}$. To see this, for each S_i, consider $b_i = \frac{125L^2d^2D}{2L^2/S_i}$ (where $b_i = 0$ if $S_i = 0$), and sort the b_i values, so that $b_{i_1} \le b_{i_2} \le \cdots \le b_{i_{t-1}}$. Furthermore, set $b_{i_0} = 0$ and $b_{i_t} = \infty$, and consider the intervals $[\tilde{b}_{i_0}, b_{i_1}), [\tilde{b}_{i_1}, b_{i_2}), \ldots, [\tilde{b}_{i_{t-1}}, b_{i_t})$, where $\tilde{b}_{i_j} = \max(b_{i_j}, r)$. For any $\rho \in [\tilde{b}_{i_j}, b_{i_{j+1}})$, observe that the objective function in (8) evaluates to

$$\frac{125jL^2d^2D}{\rho} + dD\rho + \sum_{k<t : b_{i_k} > \rho} \frac{2L^2}{S_i}.$$

Note that the last term is constant for every $\rho \in [\tilde{b}_{i_j}, b_{i_{j+1}})$. Thus, the "unconstrained" minimizer for this interval balances the first two terms, giving

$$\hat{\rho}_j = \sqrt{125jL^2d}.$$

This is a valid value for $\hat{\rho}_j$, provided $\sqrt{125jL^2d} \in [\tilde{b}_{i_j}, b_{i_{j+1}})$. However, if $\sqrt{125jL^2d} \notin [\tilde{b}_{i_j}, b_{i_{j+1}})$, we simply set $\hat{\rho}_j$ to be either $\tilde{b}_{i_j}$ or $b_{i_{j+1}}$, based on whichever of the two evaluates to a smaller objective in (8). Finally, we set ρ_t to be the $\hat{\rho}_j$ value which realizes the smallest objective among all $\hat{\rho}_j$'s.

We give the complete proof of Theorem 2 in Section A, which goes through a be-the-leader analysis again. Namely, it considers the regret of the algorithm, had it chosen ρ_{t+1} instead of ρ_t at time step t (it cannot do so, since it does not know f_t). Thereafter, it bounds the cost of instead having played ρ_t, by using the specific form of the objective function.

Theorem 2 allows us to derive precise regret guarantees of Algorithm 1, with the choice of ρ_t given in (8), for different curvature regimes.

Corollary 3.1 (Curvature-adaptive Regret)**.** *Consider Algorithm 1 instantiated with the noise schedule ρ_t given in (8), and suppose that $\alpha, \beta = O(1/T)$. Then, the following guarantee always holds:*

$$\mathbb{E}_\sigma[\mathrm{Regret}_T(f_{1:T}; x_{1:T})] \le O(\sqrt{T}).$$

Additionally, if the cumulative curvature satisfies $S_t \ge t^\gamma$, we have that

$$\mathbb{E}_\sigma[\mathrm{Regret}_T(f_{1:T}; x_{1:T})] = \begin{cases} O(\sqrt{T}), & \gamma \le \frac{1}{2}, \\ O(T^{1-\gamma}), & \frac{1}{2} < \gamma < 1, \\ O(\log T), & \gamma = 1. \end{cases}$$

The proof of Corollary 3.1 is a calculation upper-bounding $\mathrm{OPT}_T^{\alpha,\beta}(r)$ in each of the regimes, and is given in Section A.

It is instructive to compare the regime-wise guarantees above to the regime-wise guarantees given in Corollary 3.3 in Bartlett et al. [2007] for online *convex* optimization. Specifically, Bartlett et al. [2007] consider the setting where every individual f_t has curvature $t^{-\lambda}$ (which implies that $S_t \approx t^{1-\lambda}$), and do casework on λ. Mapping their cases to our cases above yields identical regret bounds as theirs. In this way, Corollary 3.1 generalizes their guarantees by analyzing different regimes of *cumulative* curvature, which allows non-convex functions too.

Lastly, we comment about the setting of the error parameters α, β in our results. Theorem 2 required no condition on α, but required $\beta = O(1/T)$. On the other hand, Corollary 3.1 required $\alpha = O(1/T)$ as well. Similar conditions on the error parameters were required by Suggala and Netrapalli [2020] as well. In particular, Suggala and Netrapalli [2020] required setting $\alpha = O(1/\sqrt{T})$ and $\beta = O(1/T)$ in order to guarantee $O(\sqrt{T})$ regret. This suffices for us as well for the purposes of getting a uniform $O(\sqrt{T})$ regret guarantee. However, since we are aiming for $o(\sqrt{T})$ regret in the cases where curvature is ample, we require α to be smaller. This is made clear by considering the expression (6) for $\mathrm{OPT}_T^{\alpha,\beta}(r)$, which explicitly contains an αT term.

4 Lower Bound

The previous section showed that the regret of our algorithm, instantiated with the suggested noise scale schedule, is at most $O(\mathrm{OPT}_T^0(r))$. In fact, it holds that $\mathrm{OPT}_T^0(r) \leq \mathrm{OPT}_T^0(0) + dDr$ (see Claim B.3), which means that, holding r, D, L fixed, the regret of our algorithm grows with T as $O(\mathrm{OPT}_T^0(0))$. We will now show that a regret of $\Omega(\mathrm{OPT}_T^0(0))$ is necessary in the worst case for *any* algorithm. This shows that the $O(\mathrm{OPT}_T^0(0))$ regret bound that we obtained above is not just a consequence of our analysis, but is in fact inherent. Moreover, it also shows that our algorithm achieves a minimax optimal regret.

Our lower bound holds for any given sequence of cumulative curvatures satisfying $0 = s_0 \leq s_1 \leq s_2 \leq \cdots \leq s_T$. In particular, this means that an $\Omega(\mathrm{OPT}_T^0(0))$ regret is necessary already for a sequence of *convex* functions. Given such a curvature sequence, we derive a sequence of functions $f_1, \ldots, f_T$ that are L-Lipschitz on $\mathcal{X}$, and have cumulative curvature $S_t = s_t$ for all t.

More concretely, define the minimax regret for a cumulative curvature sequence $s_{1:T} = s_1, \ldots, s_T$ to be

$$\mathfrak{R}_T(s_{1:T}) := \inf_{\mathcal{A}} \sup_{\substack{f_1,\ldots,f_T \ L\text{-Lipschitz on } \mathcal{X} \\ S_t = s_t \text{ for all } t}} \mathbb{E}_{\mathcal{A}} \left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right]. \tag{9}$$

Theorem 3 (Curvature-specific Lower Bound)**.** *Fix the dimension $d = 1$. There is a universal constant $c > 0$, such that for every $T \geq 1$ and every non-decreasing sequence $0 = s_0 \leq s_1 \leq \cdots \leq s_T$ satisfying $s_t - s_{t-1} \leq \frac{L}{2D}$ for all $t \in [T]$,*

$$\mathfrak{R}_T(s_{1:T}) \geq c \cdot \inf_{\rho \geq 0} \left[\sum_{t=1}^{T} \min\left\{\frac{L^2 D}{\rho}, \frac{L^2}{s_t}\right\} + D\rho\right] = \Omega\left(\mathrm{OPT}_T^0(0)\right). \tag{10}$$

As discussed above, since our algorithm always has a regret guarantee of $O(\mathrm{OPT}_T^0(0))$, we thus have that the minimax regret satisfies $\mathfrak{R}_T(s_{1:T}) = \Theta(\mathrm{OPT}_T^0(0))$.

The entire proof of Theorem 3 comprises of several intricate steps, and we defer the details to Section B. Here, we highlight some salient aspects of the construction. Previous work by Abernethy et al. [2008] shows a minimax lower bound of $\Omega(\sqrt{T})$ for generic online convex optimization, as well as a lower bound of $\Omega(\log T)$ for strongly convex online optimization. These constructions use (1) a sequence of random linear functions for the general convex case, and (2) a sequence of quadratics with drifting centers for the strongly convex case.

Neither of these constructions, by themselves, directly extend to our case, where we wish to show a lower bound for an arbitrary sequence of non-negative and non-decreasing cumulative curvatures. Nevertheless, a natural strategy would be to carefully mix in both the above components. That is, we choose every f_t to comprise of a random linear term, together with a quadratic function whose center z_t follows a suitably chosen random process. The curvature constraints imply that the coefficient on the quadratic term ought to be $a_t = s_t - s_{t-1}$. The main technical innovation in the analysis is the construction of fictitious functions $G_t(x)$ that are regularized forms of the cumulative functions $F_t(x)$. These functions enable lower-bounding the regret of any algorithm by relating the best-in-hindsight loss on F_T to the telescoping summation of $G_t(z_t) - G_{t-1}(z_{t-1})$.

Finally, in Section B.1, we complement the regime-specific regret upper bounds given in Corollary 3.1 with matching lower bounds. This shows that the guarantees for our algorithm given in Corollary 3.1 are optimal.

5 Simulations

We perform synthetic experiments on a sequence of quadratic functions, and compare the regret of three algorithms: FTL, FTPL and AdaFTPL. We specifically choose quadratics to have precise control on the cumulative curvatures, and also because the optimization oracle for all the algorithms can be computed exactly in closed form. Concretely, we set $d = 10$, and consider $\mathcal{X} = [-D/2, D/2]^d$.

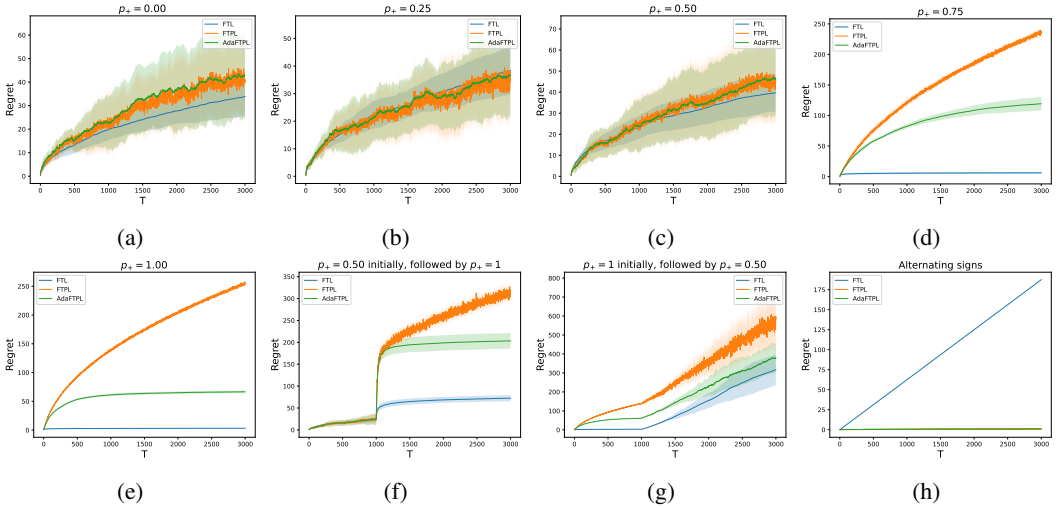


Figure 1: (a)-(e) We gradually increase curvature. (f),(g) Curvature is introduced/taken away in the middle. (h) Adversarial sequence that causes FTL to oscillate heavily and incur linear regret.

Each f_t is of the form

$$f_t(x) := \kappa_t \sum_{i=1}^{d} a_{t,i}(x_i - b_{t,i})^2, \tag{11}$$

where $\kappa_t \in \{+1, -1\}$ is set to be $+1$ with probability p_+, and -1 otherwise. Each $a_{t,i} \sim [0, 1/2D]$, and $b_{t,i} \sim [-D/2, D/2]$ uniformly at random, which ensures that f_t is 1-Lipschitz on $\mathcal{X}$. Note that if $p_+ = 0$, every f_t is non-convex, and as p_+ increases, the cumulative curvature builds up, with all the functions being strongly convex at $p_+ = 1$. We consider horizons ranging from $T = 1$ to $T = 3000$. For each horizon T, we run each of the algorithms on the sequence $f_1, \ldots, f_T$, and compute its final regret. Recall that FTPL as given by Suggala and Netrapalli [2020] is simply AdaFTPL with $\rho_t = \rho$ for all t. We set $\rho = \sqrt{T}$ as suggested by Suggala and Netrapalli [2020], which guarantees an $O(\sqrt{T})$ regret bound. For AdaFTPL, we set ρ_t according to the noise scale schedule in Theorem 2. Figures 1a-1e plot the regret of the algorithms (averaged over 10 random runs) against the time horizon T.

We can observe that when p_+ is small, all the three algorithms perform similarly. As p_+ increases, and curvature starts showing up in the function sequence, AdaFTPL adapts to this, and achieves much better regret than FTPL. In fact, in the cases where $p_+ = 0.75$ and $p_+ = 1$, the regret of AdaFTPL grows similarly to the regret of FTL ($\sim \log T$), whereas the regret of FTPL grows significantly worse ($\sim \sqrt{T}$).

To highlight the adaptive behavior of AdaFTPL, we also consider: (1) a sequence where $p_+ = 0.5$ for the first 1000 functions, and $p_+ = 1$ for the rest of the functions, and (2) a sequence where $p_+ = 1$ for the first 1000 functions, and $p_+ = 0.5$ for the rest of the functions. The coefficients $a_{t,i}$ for the first 1000 functions are scaled down; the idea is to either introduce/take away positive curvature in the middle of the sequence, and see if AdaFTPL adapts to it as expected. We can see in Figures 1f and 1g that this is indeed so: in the first case, beyond $T = 1000$, the regret of AdaFTPL grows akin to FTL, and much slower than FTPL. In the latter case, AdaFTPL exploits the available curvature in the first 1000 steps, and thereafter still maintains its conservative regret guarantee.

Finally, we also include an example (Figure 1h) of an adversarial sequence of functions (with $d = 1$) that has alternating κ_t signs, a standard and well-known failure case for FTL. For this sequence, FTL plays a sequence of moves that oscillates a lot, causing linear regret, whereas both FTPL and AdaFTPL, by virtue of their perturbations, play stable moves that result in sublinear regret. This example highlights how FTL fails *drastically* in the worst case, even if it may work well in other problem instances.

6 Conclusion

In this work, we derived an Adaptive FTPL algorithm that uses time-varying perturbations, and enjoys regret guarantees that interpolate between $O(\log T)$ and $O(\sqrt{T})$ based on the curvature present in the function sequence. We also obtained an intuitive and efficient method to set the noise scale of the perturbation at each round, as a function of the curvatures of the functions seen so far. We give a matching lower bound construction which shows that our regret guarantees are optimal. Our work thus extends adaptive regret rates in the theory of online optimization significantly beyond online convex optimization.

Acknowledgements

This work is supported by Moses Charikar's and Gregory Valiant's Simons Investigator Awards, and a Google PhD Fellowship. ChatGPT was routinely used in the development and understanding of ideas used in the paper

References

Jacob Abernethy, Peter L. Bartlett, Alexander Rakhlin, and Ambuj Tewari. Optimal strategies and minimax lower bounds for online convex games. In *Proceedings of the 21st Annual Conference on Learning Theory*, pages 414–424. Omnipress, 2008. 1, 2, 4

Naman Agarwal, Alon Gonen, and Elad Hazan. Learning in non-convex games with an optimization oracle. In *Conference on Learning Theory*, pages 18–29. PMLR, 2019. 2, 2

Peter Bartlett, Elad Hazan, and Alexander Rakhlin. Adaptive online gradient descent. *Advances in neural information processing systems*, 20, 2007. 1, 2, 3, 3.1

Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2017. B

Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006. 3, A

John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, 12(7), 2011. 1, 2

James Hannan. Approximation to bayes risk in repeated play. In M. Dresher, A. W. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games III*, volume 3 of *Annals of Mathematics Studies*, pages 97–139. Princeton University Press, Princeton, NJ, 1957. 2

Elad Hazan and Tomer Koren. The computational power of optimization in online learning. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 128–141, 2016. 2, 3

Elad Hazan, Adam Kalai, Satyen Kale, and Amit Agarwal. Logarithmic regret algorithms for online convex optimization. In *Conference on Learning Theory*, pages 499–513. Springer, 2006. 1, 2

Marcus Hutter and Jan Poland. Adaptive online prediction by following the perturbed leader. 2005. 2, 3, A

Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005. 2

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015. 1, 2

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwińska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017. 1

Xiaoyu Li and Francesco Orabona. On the convergence of stochastic gradient descent with adaptive stepsizes. In *The 22nd international conference on artificial intelligence and statistics*, pages 983–992. PMLR, 2019. 2

Ricardo A. Maronna, R. Douglas Martin, Victor J. Yohai, and Matias Salibián-Barrera. *Robust Statistics: Theory and Methods (with R)*. Wiley, 2 edition, 2019. 1

H Brendan McMahan. A survey of algorithms and analysis for adaptive online learning. *Journal of Machine Learning Research*, 18(90):1–50, 2017. 2

H Brendan McMahan and Matthew Streeter. Adaptive bound optimization for online convex optimization. *arXiv preprint arXiv:1002.4908*, 2010. 2

Arun Sai Suggala and Praneeth Netrapalli. Online non-convex learning: Following the perturbed leader is optimal. In *Algorithmic Learning Theory*, pages 845–861. PMLR, 2020. 1, 2, 3, 3, 3, 3.1, 5, A, 2, A

Guanghui Wang, Zihao Hu, Vidya Muthukumar, and Jacob D Abernethy. Adaptive oracle-efficient online learning. *Advances in Neural Information Processing Systems*, 35:23398–23411, 2022. 2

Rachel Ward, Xiaoxia Wu, and Leon Bottou. Adagrad stepsizes: Sharp convergence over nonconvex landscapes. *Journal of Machine Learning Research*, 21(219):1–30, 2020. 2

Yu-Hu Yan, Peng Zhao, and Zhi-Hua Zhou. Universal online learning with gradient variations: A multi-layer online ensemble approach. *Advances in Neural Information Processing Systems*, 36: 37682–37715, 2023. 2

Andrew C Yao. Lower bounds by probabilistic arguments. In *24th Annual Symposium on Foundations of Computer Science (sfcs 1983)*, pages 420–428. IEEE, 1983. B

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pages 928–936, 2003. 1, 2

A Proofs from Section 3

We restate and prove Theorem 1.

Theorem 1 (Adaptive FTPL Regret)**.** *Let D be the ℓ_∞ diameter of $\mathcal{X}$, and suppose that the functions $f_1, \dots, f_T$ encountered by Adaptive FTPL (Algorithm 1) are L-Lipschitz with respect to the ℓ_1 norm. Furthermore, let S_t be the strong convexity constant of $F_t = \sum_{i=1}^{t} f_i$. Then, the algorithm has the following regret guarantee:*

$$\mathbb{E}_\sigma \left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \le \sum_{t=1}^{T} \min\left\{ \frac{125L^2d^2D}{\rho_t} + \frac{\alpha+\beta\rho_t d}{20} + 2\beta dL, \frac{2L^2}{S_t} + 2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}} \right\} + dD\rho_T + \alpha T + d\beta \sum_{t=1}^{T} \rho_t. \tag{4}$$

Proof. By a standard argument for FTPL (e.g., see Lemma 12 in Hutter and Poland [2005], Chapter 4 in Cesa-Bianchi and Lugosi [2006]), it suffices to consider the setting where the functions are specified upfront by an *oblivious* adversary, and show the regret bound for every fixed $f_1, \dots, f_T$. With obliviously specified functions, it furthermore suffices in Algorithm 1 to simply draw a random vector $\sigma = \{\sigma_j\}_{j=1}^{d} \overset{\text{i.i.d.}}{\sim} \mathrm{Exp}(1)$ once at the start and reuse it at every time step, instead of drawing a fresh σ_t at each step. With these considerations, our goal is to upper bound the regret in the left-hand side of (4) for fixed $f_1, \dots, f_T$, where the expectation is only with respect to σ drawn at the start of the game.

We begin by defining a key quantity in the analysis: for any $t = 1, \dots, T$, let $\bar{x}_{t+1} = \mathcal{O}_{\alpha,\beta}(F_t, \rho_t\sigma)$. Note that $\bar{x}_{t+1}$ is separate from $x_{t+1} = \mathcal{O}_{\alpha,\beta}(F_t,\ \rho_{t+1}\sigma)$ that the algorithm plays at time step $t+1$. We introduce $\bar{x}_{t+1}$, which uses the noise vector from the *previous* time step, solely for the sake of analysis. This will help us (1) establish a standard decomposition of the regret into a *stability* and *perturbation* term, and (2) directly utilize the bound on the stability term from Suggala and Netrapalli [2020].

We now claim that[2]:

$$\mathbb{E}_\sigma\left[\sum_{t=1}^{T} f_t(x_t) - \inf_{x\in\mathcal{X}}\sum_{t=1}^{T} f_t(x)\right] \le \underbrace{L\sum_{t=1}^{T}\mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right]}_{\text{stability}} + \underbrace{dD\rho_T + \alpha T + d\beta\sum_{t=1}^{T}\rho_t}_{\text{perturbation}}. \quad (12)$$

To see this, fix any $x^\star \in \mathcal{X}$. We have that

$$\begin{aligned}\sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x^\star) &= \sum_{t=1}^{T}(f_t(x_t) - f_t(\bar{x}_{t+1})) + \sum_{t=1}^{T}(f_t(\bar{x}_{t+1}) - f_t(x^\star)) \\ &\le L\sum_{t=1}^{T}\|x_t - \bar{x}_{t+1}\|_1 + \sum_{t=1}^{T}(f_t(\bar{x}_{t+1}) - f_t(x^\star)), \quad (13)\end{aligned}$$

where the inequality follows by the Lipschitzness assumption. We thus proceed to bound the second term in (13). With a view to apply the "be-the-leader" inequality, define

$$g_t(x) := f_t(x) - \langle(\rho_t - \rho_{t-1})\sigma, x\rangle,$$

(where $\rho_0 = 0$), and observe that

$$\sum_{i=1}^{t} g_i(x) = \sum_{i=1}^{t} f_i(x) - \langle\rho_t\sigma, x\rangle = F_t(x) - \langle\rho_t\sigma, x\rangle.$$

So, by the definition of $\bar{x}_{t+1}$, we have that $\bar{x}_{t+1}$ is the (approximate) minimizer of $\sum_{i=1}^{t} g_i(x)$, where the summation crucially includes $i = t$. Namely, for any $t = 1, \ldots, T$,

$$\begin{aligned}\sum_{i=1}^{t} g_i(\bar{x}_{t+1}) &= F_t(\bar{x}_{t+1}) - \langle\rho_t\sigma, \bar{x}_{t+1}\rangle \\ &\le \inf_{x\in\mathcal{X}}\left(F_t(x) - \langle\rho_t\sigma, x\rangle\right) + (\alpha + \beta\|\rho_t\sigma\|_1) = \inf_{x\in\mathcal{X}}\sum_{i=1}^{t} g_i(x) + (\alpha + \beta\|\rho_t\sigma\|_1).\end{aligned} \quad (14)$$

We can thus instantiate the standard "be-the-leader" inequality to obtain that $\sum_{t=1}^{T} g_t(\bar{x}_{t+1}) \lesssim \inf_{x\in\mathcal{X}}\sum_{t=1}^{T} g_t(x)$. In more detail,

$$\begin{aligned}\sum_{t=1}^{T} g_t(\bar{x}_{t+1}) - \inf_{x\in\mathcal{X}}\sum_{t=1}^{T} g_t(x) &\le \sum_{t=1}^{T} g_t(\bar{x}_{t+1}) - \sum_{t=1}^{T} g_t(\bar{x}_{T+1}) + (\alpha + \beta\|\rho_T\sigma\|_1) \\ &\qquad\qquad \text{(plugging } t = T \text{ in (14))} \\ &= \sum_{t=1}^{T-1} g_t(\bar{x}_{t+1}) - \sum_{t=1}^{T-1} g_t(\bar{x}_{T+1}) + (\alpha + \beta\|\rho_T\sigma\|_1) \\ &\le \sum_{t=1}^{T-1} g_t(\bar{x}_{t+1}) - \sum_{t=1}^{T-1} g_t(\bar{x}_T) + (\alpha + \beta\|\rho_{T-1}\sigma\|_1) + (\alpha + \beta\|\rho_T\sigma\|_1) \\ &\qquad\qquad \text{(plugging } t = T-1 \text{ in (14))} \\ &= \sum_{t=1}^{T-2} g_t(\bar{x}_{t+1}) - \sum_{t=1}^{T-2} g_t(\bar{x}_T) + (\alpha + \beta\|\rho_{T-1}\sigma\|_1) + (\alpha + \beta\|\rho_T\sigma\|_1) \\ &\vdots \qquad\qquad \text{(repeating all the way down)} \\ &\le g_1(\bar{x}_2) - g_1(\bar{x}_3) + \sum_{t=2}^{T}(\alpha + \beta\|\rho_t\sigma\|_1)\end{aligned}$$

[2]We note that setting $\rho_1 = \cdots = \rho_T = \frac{1}{\eta}$ recovers the bound in Suggala and Netrapalli [2020].

$$\leq \sum_{t=1}^{T}(\alpha + \beta\|\rho_t\sigma\|_1) \qquad \text{(plugging } t=1 \text{ in (14))}$$

$$= \alpha T + \beta\|\sigma\|_1 \sum_{t=1}^{T}\rho_t. \qquad \text{(each } \rho_t \geq 0)$$

Instantiating the above inequality with $x = x^\star$, we get

$$\sum_{t=1}^{T} g_t(\bar{x}_{t+1}) \leq \sum_{t=1}^{T} g_t(x^\star) + \alpha T + \beta\|\sigma\|_1 \sum_{t=1}^{T}\rho_t$$

$$\implies \sum_{t=1}^{T} f_t(\bar{x}_{t+1}) - \sum_{t=1}^{T}\langle(\rho_t - \rho_{t-1})\sigma, \bar{x}_{t+1}\rangle \leq \sum_{t=1}^{T} f_t(x^\star) - \sum_{t=1}^{T}\langle(\rho_t - \rho_{t-1})\sigma, x^\star\rangle + \alpha T + \beta\|\sigma\|_1 \sum_{t=1}^{T}\rho_t$$

$$\implies \sum_{t=1}^{T}(f_t(\bar{x}_{t+1}) - f_t(x^\star)) \leq \sum_{t=1}^{T}\langle(\rho_t - \rho_{t-1})\sigma, \bar{x}_{t+1} - x^\star\rangle + \alpha T + \beta\|\sigma\|_1 \sum_{t=1}^{T}\rho_t$$

$$\implies \sum_{t=1}^{T}(f_t(\bar{x}_{t+1}) - f_t(x^\star)) \leq \sum_{t=1}^{T}|\rho_t - \rho_{t-1}|\|\sigma\|_1\|\bar{x}_{t+1} - x^\star\|_\infty + \alpha T + \beta\|\sigma\|_1 \sum_{t=1}^{T}\rho_t$$

(Hölder's inequality)

$$\implies \sum_{t=1}^{T}(f_t(\bar{x}_{t+1}) - f_t(x^\star)) \leq D\|\sigma\|_1\rho_T + \alpha T + \beta\|\sigma\|_1 \sum_{t=1}^{T}\rho_t,$$

where we used that $\mathcal{X}$ has ℓ_∞ diameter D, and that $\rho_0, \ldots, \rho_T$ are non-negative and non-decreasing, with $\rho_0 = 0$. Substituting the bound above in (13), and taking expectation with respect to σ, we get

$$\mathbb{E}_\sigma\left[\sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x^\star)\right] \leq L\sum_{t=1}^{T}\mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right] + D\rho_T\,\mathbb{E}_\sigma\left[\|\sigma\|_1\right] + \alpha T + \mathbb{E}_\sigma\left[\|\sigma\|_1\right]\cdot\beta\sum_{t=1}^{T}\rho_t$$

$$= L\sum_{t=1}^{T}\mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right] + dD\rho_T + \alpha T + d\beta\sum_{t=1}^{T}\rho_t,$$

which was the claimed stability-perturbation decomposition in (12).

Now, since x_t and $\bar{x}_{t+1}$ both use the *same* noise vector $\rho_t\sigma$ (whose entries are i.i.d. exponential random variables with parameter $1/\rho_t$), we can directly use the bound on the stability term established in Suggala and Netrapalli [2020, Theorem 1]. Specifically, they show that:

$$\mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right] \leq \frac{125Ld^2D}{\rho_t} + \frac{\beta\rho_t d}{20L} + 2\beta d + \frac{\alpha}{20L}. \tag{NC}$$

On the other hand, we can also bound this term differently by considering the curvature of $F_t = \sum_{i=1}^{t} f_i$. Namely, let S_t be the strong convexity constant of F_t with respect to ℓ_1 (where $S_t = 0$ if F_t is not strongly convex). Then note that the function $\Phi(x) = F_t(x) - \langle\rho_t\sigma, x\rangle$ also has strong convexity constant S_t.

Let $z_t \in \arg\min_{x\in\mathcal{X}}\Phi(x)$. Since $\bar{x}_{t+1}$ also approximately minimizes $\Phi(x)$, we have that $\Phi(\bar{x}_{t+1}) - \Phi(z_t) \leq \alpha + \beta\|\rho_t\sigma\|_1$. Invoking the strong convexity inequality (2) with $y = \bar{x}_{t+1}$ and $x = z_t$, and using that $\Phi(x)$ is minimized at z_t, we get

$$\Phi(\bar{x}_{t+1}) \geq \Phi(z_t) + \frac{S_t}{2}\|\bar{x}_{t+1} - z_t\|_1^2 \implies \frac{S_t}{2}\|\bar{x}_{t+1} - z_t\|_1^2 \leq \Phi(\bar{x}_{t+1}) - \Phi(z_t) \leq \alpha + \beta\|\rho_t\sigma\|_1$$

$$\implies \|\bar{x}_{t+1} - z_t\|_1 \leq \sqrt{\frac{2(\alpha + \beta\|\rho_t\sigma\|_1)}{S_t}}. \tag{15}$$

But note also that x_t approximately minimizes $F_{t-1}(x) - \langle\rho_t\sigma, x\rangle = \Phi(x) - f_t(x)$. That is,

$$\Phi(x_t) - f_t(x_t) \leq \inf_{x\in\mathcal{X}}(\Phi(x) - f_t(x)) + \alpha + \beta\|\rho_t\sigma\|_1 \leq \Phi(z_t) - f_t(z_t) + \alpha + \beta\|\rho_t\sigma\|_1.$$

This means that

$$\Phi(x_t) - \Phi(z_t) \le \alpha + \beta\|\rho_t\sigma\|_1 + f_t(x_t) - f_t(z_t) \le \alpha + \beta\|\rho_t\sigma\|_1 + L\|x_t - z_t\|_1, \tag{16}$$

where we used Lipschitzness in the last inequality. Invoking the strong convexity inequality (2) again with $y = x_t$ and $x = z_t$, and using that $\Phi(x)$ is minimized at z_t, we get

$$\begin{aligned}
\Phi(x_t) \ge \Phi(z_t) + \frac{S_t}{2}\|x_t - z_t\|_1^2 \implies & \frac{S_t}{2}\|x_t - z_t\|_1^2 \le \Phi(x_t) - \Phi(z_t) \\
\implies & \frac{S_t}{2}\|x_t - z_t\|_1^2 \le \alpha + \beta\|\rho_t\sigma\|_1 + L\|x_t - z_t\|_1 \\
& \qquad \text{(using (16))} \\
\implies & \|x_t - z_t\|_1 \le \frac{L + \sqrt{L^2 + 2S_t(\alpha + \beta\|\rho_t\sigma\|_1)}}{S_t} \\
\implies & \|x_t - z_t\|_1 \le \frac{2L}{S_t} + \sqrt{\frac{2(\alpha + \beta\|\rho_t\sigma\|_1)}{S_t}}, \qquad (17)
\end{aligned}$$

where in the second-to-last step, we solved a quadratic inequality in the variable $\|x_t - z_t\|_1$, and in the last step, we used that $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$.

Combining (15) and (17) with the triangle inequality, we get that

$$\|x_t - \bar{x}_{t+1}\|_1 \le \frac{2L}{S_t} + 2\sqrt{\frac{2(\alpha + \beta\|\rho_t\sigma\|_1)}{S_t}}.$$

Taking expectation with respect to σ, we get

$$\begin{aligned}
\mathbb{E}_\sigma\left[\|x_t - \bar{x}_{t+1}\|_1\right] &\le \frac{2L}{S_t} + 2\sqrt{\frac{2}{S_t}} \cdot \mathbb{E}_\sigma\left[\sqrt{\alpha + \beta\|\rho_t\sigma\|_1}\right] \\
&\le \frac{2L}{S_t} + 2\sqrt{\frac{2}{S_t}} \cdot \sqrt{\mathbb{E}_\sigma\left[\alpha + \beta\|\rho_t\sigma\|_1\right]} && \text{(Jensen's inequality)} \\
&= \frac{2L}{S_t} + 2\sqrt{\frac{2(\alpha + \beta\rho_t d)}{S_t}}. && \text{(SC)}
\end{aligned}$$

Combining this strongly-convex stability bound (SC) with the fallback non-convex stability bound (NC) from above, and substituting into the stability-perturbation decomposition (12) of the regret, we get that

$$\begin{aligned}
\mathbb{E}_\sigma\left[\sum_{t=1}^T f_t(x_t) - \inf_{x\in\mathcal{X}}\sum_{t=1}^T f_t(x)\right] \le \sum_{t=1}^T \min\left\{\frac{125L^2d^2D}{\rho_t} + \frac{\alpha + \beta\rho_t d}{20} + 2\beta dL, \frac{2L^2}{S_t} + 2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}}\right\} \\
+ dD\rho_T + \alpha T + d\beta\sum_{t=1}^T \rho_t,
\end{aligned}$$

which is the desired bound. □

We next restate and prove Theorem 2.

Theorem 2 (FTL for Noise Scale Schedule)**.** *Fix $r > 0$. Let $\rho_1 = r$, and for $t \ge 2$,*

$$\rho_t = \min\arg\inf_{\rho \ge r}\left[\sum_{i=1}^{t-1}\min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_i}\right\} + dD\rho\right], \tag{8}$$

Then,

1. *$\rho_1 \le \cdots \le \rho_T$, and every ρ_t depends only on $S_1, \ldots, S_{t-1}$.*
2. *For this noise scale schedule, Algorithm 1 has the following regret guarantee:*

(a) With $\alpha = \beta = 0$, we have $\mathbb{E}_\sigma\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \le C_1 \cdot \mathrm{OPT}_T^0(r)$.

(b) If $\beta \le c/T$, we have $\mathbb{E}_\sigma\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \le C_2 \cdot \mathrm{OPT}_T^{\alpha,\beta}(r)$.

Here, C_1 and C_2 are constants depending only on d, D, L, r, and c.

Proof. Recall that we set $\rho_1 = r$, and for $t \ge 2$,

$$\rho_t = \min \operatorname*{arg\,inf}_{\rho \ge r} \underbrace{\left[\sum_{i=1}^{t-1} \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_i}\right\} + dD\rho\right]}_{:=\Psi_{t-1}(\rho)},$$

where we defined the objective function inside the parentheses to be $\Psi_{t-1}(\rho)$. It is clear that ρ_t depends only on $S_1, \dots, S_{t-1}$. We will now argue that $\rho_1 \le \cdots \le \rho_T$. Note that $r = \rho_1 \le \rho_2$ holds immediately, since the minimization is over $\rho \ge r$.

Now, fix $t \ge 3$, and consider any $y \in [r, \rho_{t-1})$. Since ρ_{t-1} is the *smallest* infimizer of $\Psi_{t-2}(\rho)$, it must be the case that

$$\Psi_{t-2}(y) > \Psi_{t-2}(\rho_{t-1}).$$

Therefore,

$$\begin{aligned}\Psi_{t-1}(y) &= \Psi_{t-2}(y) + \min\left\{\frac{125L^2d^2D}{y}, \frac{2L^2}{S_{t-1}}\right\} \\ &> \Psi_{t-2}(\rho_{t-1}) + \min\left\{\frac{125L^2d^2D}{\rho_{t-1}}, \frac{2L^2}{S_{t-1}}\right\} = \Psi_{t-1}(\rho_{t-1}),\end{aligned}$$

where we used that the function $\min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_{t-1}}\right\}$ is non-increasing in ρ. Thus, any feasible minimizer of Ψ_{t-1} must at least be ρ_{t-1}, implying that $\rho_t \ge \rho_{t-1}$.

Now, define the *clean* offline optimum, namely the best regret upper bound in hindsight without any error terms, to be

$$\mathrm{OPT}_T^0(r) := \inf_{r \le \rho_1 \le \cdots \le \rho_T} \left[\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\rho_t}, \frac{2L^2}{S_t}\right\} + dD\rho_T\right]. \tag{18}$$

We claim that

$$\mathrm{OPT}_T^0(r) = \inf_{\rho \ge r}\left[\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t}\right\} + dD\rho\right] = \inf_{\rho \ge r} \Psi_T(\rho). \tag{19}$$

This follows simply by noting that, for every feasible $r \le \rho_1 \le \cdots \le \rho_T$, we can simply set $\rho_t = \rho_T$ for every t; this keeps the sequence feasible, and does not increase any individual term inside the summation.

Be-the-leader on clean sequence. We will now argue that the sequence of ρ_t's above has small regret on the sequence of clean objective functions, compared to $\mathrm{OPT}_T^0(r)$. This is the standard be-the-leader analysis, and proceeds as follows: let $g_0(\rho) := dD\rho$, and $g_t(\rho) = \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t}\right\}$ for $t \ge 1$. By definition, for any $i \in \{1, \dots, T\}$:

$$\sum_{t=0}^{i} g_t(\rho_{i+1}) = \Psi_i(\rho_{i+1}) = \inf_{\rho \ge r} \Psi_i(\rho), \tag{20}$$

and therefore,

$$\begin{aligned}\sum_{t=0}^{T} g_t(\rho_{t+1}) - \inf_{\rho \ge r} \Psi_T(\rho) &= \sum_{t=0}^{T} g_t(\rho_{t+1}) - \sum_{t=0}^{T} g_t(\rho_{T+1}) \\ &= \sum_{t=0}^{T-1} g_t(\rho_{t+1}) - \sum_{t=0}^{T-1} g_t(\rho_{T+1})\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{t=0}^{T-1} g_t(\rho_{t+1}) - \sum_{t=0}^{T-1} g_t(\rho_T) && \text{(from (20))} \\
&= \sum_{t=0}^{T-2} g_t(\rho_{t+1}) - \sum_{t=0}^{T-2} g_t(\rho_T) \\
&\vdots && \text{(repeating all the way down)} \\
&\leq g_0(\rho_1) - g_0(\rho_2) \\
&= dD(\rho_1 - \rho_2) \\
&= dD(r - \rho_2) \leq 0. && \text{(since } \rho_1 = r\text{)}
\end{aligned}
$$

Thus, we get that

$$
g_0(\rho_1) + \sum_{t=1}^{T} g_t(\rho_{t+1}) - \mathrm{OPT}_T^0(r) \leq 0 \quad \implies \quad \sum_{t=1}^{T} g_t(\rho_{t+1}) \leq \mathrm{OPT}_T^0(r) - dDr. \tag{21}
$$

Actual regret on clean sequence. Note that (21) benchmarks the performance of playing ρ_{t+1} at time step t, but actually, we want to benchmark the performance of playing ρ_t. For this, observe that for any $t \geq 1$,

$$
\begin{aligned}
g_t(\rho_t) - g_t(\rho_{t+1}) &= \min\left\{\frac{125L^2d^2D}{\rho_t}, \frac{2L^2}{S_t}\right\} - \min\left\{\frac{125L^2d^2D}{\rho_{t+1}}, \frac{2L^2}{S_t}\right\} \\
&\leq \frac{125L^2d^2D}{\rho_t} - \frac{125L^2d^2D}{\rho_{t+1}},
\end{aligned}
$$

where we used that $\rho_t \leq \rho_{t+1}$. Summing over t, the sum telescopes, and we get

$$
\sum_{t=1}^{T} \left(g_t(\rho_t) - g_t(\rho_{t+1})\right) \leq \frac{125L^2d^2D}{\rho_1} - \frac{125L^2d^2D}{\rho_{T+1}} \leq \frac{125L^2d^2D}{r}.
$$

Substituting this in (21), we get

$$
\sum_{t=1}^{T} g_t(\rho_t) \leq \mathrm{OPT}_T^0(r) - dDr + \frac{125L^2d^2D}{r}.
$$

Folding in the terminal term $dD\rho_T$ on both sides gives

$$
\begin{aligned}
dD\rho_T + \sum_{t=1}^{T} g_t(\rho_t) &\leq \mathrm{OPT}_T^0(r) - dDr + \frac{125L^2d^2D}{r} + dD\rho_T \\
&\leq \mathrm{OPT}_T^0(r) - dDr + \frac{125L^2d^2D}{r} + dD\rho_{T+1} \\
&\leq \mathrm{OPT}_T^0(r) - dDr + \frac{125L^2d^2D}{r} + dD\rho_{T+1} + \sum_{t=1}^{T} g_t(\rho_{T+1}) \\
&= 2\mathrm{OPT}_T^0(r) - dDr + \frac{125L^2d^2D}{r}. \quad \text{(by definition of } \rho_{T+1} \text{ and } \mathrm{OPT}_T^0(r)\text{)}
\end{aligned}
$$

Using $dDr \leq \mathrm{OPT}_T^0(r)$ further gives

$$
dD\rho_T + \sum_{t=1}^{T} g_t(\rho_t) \leq C \cdot \mathrm{OPT}_T^0(r), \tag{22}
$$

where $C = \max\left(2, 1 + \frac{125L^2d}{r^2}\right)$.

Relating $\mathrm{OPT}_T^0(r)$ to $\mathrm{OPT}_T^{\alpha,\beta}(r)$. Recall the definition of $\mathrm{OPT}_T^{\alpha,\beta}(r)$:

$$\mathrm{OPT}_T^{\alpha,\beta}(r) := \inf_{r\le\rho_1\le\cdots\le\rho_T}\left[\sum_{t=1}^T \min\left\{\frac{125L^2d^2D}{\rho_t}+\frac{\beta\rho_t d}{20}+2L\beta d+\frac{\alpha}{20},\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}}\right\}\right.$$
$$\left.+dD\rho_T+\alpha T+d\beta\sum_{t=1}^T\rho_t\right],$$

and note that for every feasible $\rho_1\le\cdots\le\rho_T$,

$$\sum_{t=1}^T \min\left\{\frac{125L^2d^2D}{\rho_t}+\frac{\beta\rho_t d}{20}+2L\beta d+\frac{\alpha}{20},\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}}\right\}+dD\rho_T+\alpha T+d\beta\sum_{t=1}^T\rho_t$$
$$\ge\sum_{t=1}^T g_t(\rho_t)+dD\rho_T\ge\sum_{t=1}^T g_t(\rho_T)+dD\rho_T\ge\mathrm{OPT}_T^0(r),$$

where we used that g_t is non-increasing. Infimizing over all feasible $\rho_1\le\cdots\le\rho_T$, we get that

$$\mathrm{OPT}_T^{\alpha,\beta}(r)\ge\mathrm{OPT}_T^0(r). \tag{23}$$

Upper-bounding the regret with respect to $\mathrm{OPT}_T^{\alpha,\beta}(r)$. Recall that we upper-bounded the "clean terms" of the regret of our ρ_t schedule in terms of the clean offline optimum in (22). Furthermore, we upper-bounded the clean offline optimum with the true offline optimum in (23). It remains to fold in the "non-clean" terms that appear in the regret of our ρ_t schedule, and upper-bound these in terms of $\mathrm{OPT}_T^{\alpha,\beta}(r)$.

Towards this, for $\rho\ge r$, consider the function

$$M_t(\rho) := \min\left\{\frac{125L^2d^2D}{\rho}+\frac{\beta\rho d}{20}+2L\beta d+\frac{\alpha}{20},\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho d)}{S_t}}\right\}.$$

Recall that $g_t(\rho)=\min\left\{\frac{125L^2d^2D}{\rho},\frac{2L^2}{S_t}\right\}$. Suppose it is the case that $\frac{125L^2d^2D}{\rho}\le\frac{2L^2}{S_t}$. Then,

$$M_t(\rho)\le\frac{125L^2d^2D}{\rho}+\frac{\beta\rho d}{20}+2L\beta d+\frac{\alpha}{20}=g_t(\rho)+\frac{\beta\rho d}{20}+2L\beta d+\frac{\alpha}{20}.$$

On the other hand, suppose that $\frac{125L^2d^2D}{\rho}>\frac{2L^2}{S_t}$. Then,

$$M_t(\rho)\le\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho d)}{S_t}}=g_t(\rho)+2\sqrt{g_t(\rho)(\alpha+\beta\rho d)}.$$

Therefore, the following bound always holds:

$$M_t(\rho)\le g_t(\rho)+\frac{\beta\rho d}{20}+2L\beta d+\frac{\alpha}{20}+2\sqrt{g_t(\rho)(\alpha+\beta\rho d)}.$$

Instantiating this inequality for our noise-scale schedule $\rho_1,\dots,\rho_T$, and summing, we get

$$\sum_{t=1}^T M_t(\rho_t)\le\sum_{t=1}^T g_t(\rho_t)+\frac{\beta d}{20}\sum_{t=1}^T\rho_t+2L\beta dT+\frac{\alpha T}{20}+2\sum_{t=1}^T\sqrt{g_t(\rho_t)(\alpha+\beta\rho_t d)}$$
$$\le\sum_{t=1}^T g_t(\rho_t)+\frac{\beta d}{20}\sum_{t=1}^T\rho_t+2L\beta dT+\frac{\alpha T}{20}+2\sqrt{\left(\sum_{t=1}^T g_t(\rho_t)\right)\left(\sum_{t=1}^T(\alpha+\beta\rho_t d)\right)}.$$
(Cauchy-Schwarz inequality)

Adding in the perturbation terms from (12), we get that

$$\begin{aligned}
&\sum_{t=1}^{T} M_t(\rho_t) + dD\rho_T + \alpha T + d\beta \sum_{t=1}^{T} \rho_t \\
&\leq \sum_{t=1}^{T} g_t(\rho_t) + dD\rho_T + \frac{21\beta d}{20} \sum_{t=1}^{T} \rho_t + 2L\beta dT + \frac{21\alpha T}{20} + 2\sqrt{\left(\sum_{t=1}^{T} g_t(\rho_t)\right)\left(\sum_{t=1}^{T}(\alpha + \beta\rho_t d)\right)} \\
&\leq \underbrace{\sum_{t=1}^{T} g_t(\rho_t) + dD\rho_T}_{\text{clean regret from (22)}} + \frac{21\beta d}{20} \sum_{t=1}^{T} \rho_t + 2L\beta dT + \frac{21\alpha T}{20} + 2\sqrt{\underbrace{\left(\sum_{t=1}^{T} g_t(\rho_t) + dD\rho_T\right)}_{\text{clean regret from (22)}}\left(\sum_{t=1}^{T}(\alpha + \beta\rho_t d)\right)} \\
&\leq C \cdot \mathrm{OPT}_T^{\alpha,\beta}(r) + \underbrace{\frac{21\beta d}{20} \sum_{t=1}^{T} \rho_t + 2L\beta dT + \frac{21\alpha T}{20}}_{:= E_T} + 2\sqrt{C \cdot \mathrm{OPT}_T^{\alpha,\beta}(r)\left(\sum_{t=1}^{T}(\alpha + \beta\rho_t d)\right)} \\
&\leq C \cdot \mathrm{OPT}_T^{\alpha,\beta}(r) + E_T + 2\sqrt{C \cdot \mathrm{OPT}_T^{\alpha,\beta}(r) \cdot E_T}. \qquad (24)
\end{aligned}$$

Bounding E_T. Finally, we aim to upper-bound E_T in terms of $\mathrm{OPT}_T^{\alpha,\beta}(r)$. We have

$$\begin{aligned}
E_T &= \frac{21\beta d}{20} \sum_{t=1}^{T} \rho_t + 2L\beta dT + \frac{21\alpha T}{20} \\
&\leq \frac{21\beta dT\rho_T}{20} + 2L\beta dT + \frac{21\alpha T}{20}. \qquad \text{(since } \rho_t \text{ is non-decreasing)}
\end{aligned}$$

Now, observe that $\alpha T \leq \mathrm{OPT}_T^{\alpha,\beta}(r)$. Furthermore, from (22) and (23),

$$dD\rho_T \leq dD\rho_T + \sum_{t=1}^{T} g_t(\rho_t) \leq C \cdot \mathrm{OPT}_T^{0}(r) \leq C \cdot \mathrm{OPT}_T^{\alpha,\beta}(r).$$

Finally, since it also holds that $dDr \leq \mathrm{OPT}_T^{\alpha,\beta}(r)$, we have that $L\beta dT \leq \frac{L\beta T}{Dr} \cdot \mathrm{OPT}_T^{\alpha,\beta}(r)$. Combining all this, we get

$$E_T \leq \left(\frac{21C\beta T}{20D} + \frac{2L\beta T}{Dr} + \frac{21}{20}\right) \cdot \mathrm{OPT}_T^{\alpha,\beta}(r).$$

Plugging in $\beta = O(1/T)$. Note that $\beta \leq c/T$ suffices to get $E_T \leq C' \cdot \mathrm{OPT}_T^{\alpha,\beta}(r)$, where C' is a constant depending on L, D, d, r and c. Finally, plugging this back in (24), we get

$$\sum_{t=1}^{T} M_t(\rho_t) + dD\rho_T + \alpha T + d\beta \sum_{t=1}^{T} \rho_t \leq C'' \cdot \mathrm{OPT}_T^{\alpha,\beta}(r),$$

where C'' is a constant depending on L, D, d, r and c. Recalling that the quantity on the left-hand side above is precisely the upper-bound on regret shown in Theorem 1 completes the proof. □

Corollary 3.1 (Curvature-adaptive Regret)**.** *Consider Algorithm 1 instantiated with the noise schedule ρ_t given in (8), and suppose that $\alpha, \beta = O(1/T)$. Then, the following guarantee always holds:*

$$\mathbb{E}_\sigma[\mathrm{Regret}_T(f_{1:T}; x_{1:T})] \leq O(\sqrt{T}).$$

Additionally, if the cumulative curvature satisfies $S_t \geq t^\gamma$, we have that

$$\mathbb{E}_\sigma[\mathrm{Regret}_T(f_{1:T}; x_{1:T})] = \begin{cases} O(\sqrt{T}), & \gamma \leq \frac{1}{2}, \\ O(T^{1-\gamma}), & \frac{1}{2} < \gamma < 1, \\ O(\log T), & \gamma = 1. \end{cases}$$

Proof. In view of Theorem 2, it suffices to upper-bound $\mathrm{OPT}_T^{\alpha,\beta}(r)$, where

$$\mathrm{OPT}_T^{\alpha,\beta}(r) := \inf_{r\le\rho_1\le\cdots\le\rho_T}\left[\sum_{t=1}^{T}\min\left\{\frac{125L^2d^2D}{\rho_t}+\frac{\beta\rho_t d}{20}+2L\beta d+\frac{\alpha}{20},\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho_t d)}{S_t}}\right\}+dD\rho_T+\alpha T+d\beta\sum_{t=1}^{T}\rho_t\right].$$

Choose $\rho_t=\rho$ for all t. By taking the non-convex branch in the min, we have that

$$\begin{aligned}\mathrm{OPT}_T^{\alpha,\beta}(r) &\le \sum_{t=1}^{T}\left(\frac{125L^2d^2D}{\rho}+\frac{\beta\rho d}{20}+2L\beta d+\frac{\alpha}{20}\right)+dD\rho+\alpha T+d\beta T\rho\\ &= \frac{125L^2d^2DT}{\rho}+\frac{21\beta\rho dT}{20}+2L\beta dT+dD\rho+\frac{21\alpha T}{20}\\ &= O\left(\frac{T}{\rho}+\rho\right). && \text{(substituting } \alpha,\beta=O(1/T))\end{aligned}$$

Setting $r=1$ and $\rho=\sqrt{T}$ gives $\mathrm{OPT}_T^{\alpha,\beta}(r)\le O(\sqrt{T})$.

Now, suppose that $S_t\ge t^\gamma$. When $\gamma\le 1/2$, we already have the $O(\sqrt{T})$ regret bound from above.

Regime $\frac{1}{2}<\gamma<1$. Choose $r=1$ and $\rho_t=\rho=1$ for all t. By taking the strongly-convex branch in the min, we have that

$$\begin{aligned}\mathrm{OPT}_T^{\alpha,\beta}(r) &\le \sum_{t=1}^{T}\left(\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho d)}{S_t}}\right)+dD\rho+\alpha T+d\beta T\rho\\ &\le O\left(\sum_{t=1}^{T}t^{-\gamma}+T^{-1/2}\sum_{t=1}^{T}t^{-\gamma/2}\right)\\ &\le O\left(T^{1-\gamma}+T^{1/2-\gamma/2}\right)=O(T^{1-\gamma}).\end{aligned}$$

Regime $\gamma=1$. Choose $r=1$ and $\rho_t=\rho=1$ for all t. By taking the strongly-convex branch in the min, we have that

$$\begin{aligned}\mathrm{OPT}_T^{\alpha,\beta}(r) &\le \sum_{t=1}^{T}\left(\frac{2L^2}{S_t}+2L\sqrt{\frac{2(\alpha+\beta\rho d)}{S_t}}\right)+dD\rho+\alpha T+d\beta T\rho\\ &= O\left(\sum_{t=1}^{T}\frac{1}{t}+T^{-1/2}\sum_{t=1}^{T}t^{-1/2}\right)=O(\log T).\end{aligned}$$

□

B Proofs from Section 4

We restate and prove Theorem 3.

Theorem 3 (Curvature-specific Lower Bound)**.** *Fix the dimension $d=1$. There is a universal constant $c>0$, such that for every $T\ge 1$ and every non-decreasing sequence $0=s_0\le s_1\le\cdots\le s_T$ satisfying $s_t-s_{t-1}\le\frac{L}{2D}$ for all $t\in[T]$,*

$$\mathfrak{R}_T(s_{1:T})\ge c\cdot\inf_{\rho\ge 0}\left[\sum_{t=1}^{T}\min\left\{\frac{L^2D}{\rho},\frac{L^2}{s_t}\right\}+D\rho\right]=\Omega\left(\mathrm{OPT}_T^0(0)\right). \tag{10}$$

Setup. Fix $\mathcal{X}$ to be the interval $[-D/2, D/2] \subseteq \mathbb{R}$. Let $0 = s_0 \leq s_1 \leq \cdots \leq s_T$ be the given sequence of cumulative curvatures. For any $t \in [T]$, let

$$a_t := s_t - s_{t-1} \geq 0.$$

The lower bound of Theorem 3 holds for any sequence satisfying that $a_t \leq L/2D$ for all t. As will become clearer, this condition is used solely to ensure that the functions used in the lower bound are all L-Lipschitz on $\mathcal{X}$.

Recall that

$$\mathfrak{R}_T(s_{1:T}) := \inf_{\mathcal{A}} \sup_{\substack{f_1,\ldots,f_T \ L\text{-Lipschitz on } \mathcal{X} \\ S_t = s_t \text{ for all } t}} \mathbb{E}_{\mathcal{A}} \left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right]. \tag{25}$$

and we seek to show that

$$\mathfrak{R}_T(s_{1:T}) \geq c \cdot \inf_{\rho \geq 0} \left[L^2 \sum_{t=1}^{T} \min\left\{ \frac{D}{\rho}, \frac{1}{s_t} \right\} + D\rho \right]. \tag{26}$$

It will be convenient to instead lower bound $\mathfrak{R}_T(s_{1:T})$ by the quantity:

$$\inf_{\gamma \geq 0} \left[D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma} \right]. \tag{27}$$

Indeed, this will be sufficient, due to the following claim:

Claim B.1. *For every $s \geq 0$ and $\gamma \geq 0$,*

$$\frac{1}{2} \min\left\{ \frac{1}{s}, \frac{1}{\gamma} \right\} \leq \frac{1}{s+\gamma} \leq \min\left\{ \frac{1}{s}, \frac{1}{\gamma} \right\},$$

with the convention $1/0 = \infty$.

Proof. The upper bound follows because $s + \gamma \geq s$ and $s + \gamma \geq \gamma$. The lower bound follows because $s + \gamma \leq 2\max\{s, \gamma\}$. □

Hence, if we lower-bound $\mathfrak{R}_T(s_{1:T})$ in terms of the quantity in (27), the desired lower bound in (26) will follow by the change of variable $\gamma = \rho/D$.

We thus proceed to lower-bound $\mathfrak{R}_T(s_{1:T})$ by $\inf_{\gamma \geq 0} \left\{ D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t+\gamma} \right\}$, and divide the proof into 8 steps:

Step 1: Defining the hard distribution for a fixed γ

Fix an arbitrary deterministic learner. Also, fix $\gamma \geq 0$ which satisfies:

1. $s_t + \gamma > 0$ for every t.
2. $H_2(\gamma) := \sum_{t=1}^{T} \frac{1}{(s_t+\gamma)^2} \leq \frac{D^2}{L^2}$

We will later choose an optimal value of γ which satisfies these conditions.

Let $\varepsilon_1, \ldots, \varepsilon_T$ be independent Rademacher random variables in $\{-1, +1\}$. Set

$$g := \frac{L}{8}.$$

The functions $f_1, \ldots, f_T$ will be a sequence of carefully chosen quadratics, where each f_t has curvature a_t (and since $a_t \geq 0$, each quadratic is convex). The centers and linear tilts ℓ_t of each f_t are specified by the following process.

Initialize $z_0 = 0$, and declare the process to be active. For each round $t = 1, 2, \ldots, T$:

1. If the process is active before round t, set $\ell_t = g\varepsilon_t$; if it has been declared inactive, set $\ell_t = 0$.

2. Set
$$f_t(x) := \frac{a_t}{2}(x - z_{t-1})^2 - \ell_t(x - z_{t-1}). \tag{28}$$
3. Update
$$z_t = z_{t-1} + \frac{\ell_t}{s_t + \gamma}. \tag{29}$$
4. If the process was active before round t and $|z_t| > D/4$, declare the process inactive.

Observe that once the process above becomes inactive, all future linear tilts are zero and the center stays constant. Note also that all the functions $f_1, \ldots, f_T$ are determined as a function of γ and the sequences $s_1, \ldots, s_T$ and $\varepsilon_1, \ldots, \varepsilon_T$.

Step 2: Feasibility of centers and functions

We first argue that all the centers $z_0, \ldots, z_T$ belong to $\mathcal{X} = [-D/2, D/2]$. For this, note that assuming the bound on $H_2(\gamma)$ given in Item 2, the magnitude of each update satisfies
$$\left|\frac{\ell_t}{s_t + \gamma}\right| \le g\sqrt{H_2(\gamma)} \le \frac{L}{8} \cdot \frac{D}{L} = \frac{D}{8}.$$
Now observe that $z_0 = 0$, and for any t where the process is active before round t, we have $|z_{t-1}| \le D/4$. Furthermore, if at round t, $|z_t|$ becomes larger than $D/4$ causing the process to become inactive, we have
$$|z_t| \le |z_{t-1}| + D/4 \le D/4 + D/8 = 3D/8 < D/2.$$
Thereafter, the center remains fixed, and hence $z_t \in \mathcal{X}$ for all $t \in \{0, 1, \ldots, T\}$.

We now argue that all the functions are feasible. Observe that even if the process has been declared inactive, the quadratic terms with coefficients a_t are still played. Thus, for every round,
$$f_t''(x) = a_t \ge 0,$$
so f_t is convex. The cumulative loss $F_t = \sum_{i=1}^t f_i$ satisfies
$$F_t''(x) = \sum_{i=1}^t a_i = s_t,$$
so the cumulative curvature is exactly the prescribed value s_t.

Since all the centers lie in $\mathcal{X}$, for every $x \in \mathcal{X}$,
$$|f_t'(x)| = |a_t(x - z_{t-1}) - \ell_t| \le a_t D + |\ell_t|.$$
Since $a_t \le L/2D$ and $|\ell_t| \le g = L/8$,
$$|f_t'(x)| \le \frac{L}{2} + \frac{L}{8} < L. \tag{30}$$
Thus, every f_t is convex, L-Lipschitz on $\mathcal{X}$, and has the required cumulative curvature schedule.

Step 3: The learner's expected cumulative loss is nonnegative

For any t, observe that conditioned on $\varepsilon_1, \ldots, \varepsilon_{t-1}$, the learner's move x_t is determined (since it is deterministic). Furthermore, the center z_{t-1}, as well as whether the process remains active/has been declared inactive before round t is also determined. We consider the two cases separately.

Case 1: The process is active before round t. Then $\ell_t = g\varepsilon_t$, where ε_t is independent of $\varepsilon_1, \ldots, \varepsilon_{t-1}$, and has mean zero. Hence,
$$\mathbb{E}[\ell_t \mid \varepsilon_1, \ldots, \varepsilon_{t-1}] = 0.$$
By definition of f_t (28),
$$\mathbb{E}[f_t(x_t) \mid \varepsilon_1, \ldots, \varepsilon_{t-1}] = \frac{a_t}{2}(x_t - z_{t-1})^2 \ge 0.$$

Case 2: The process has become inactive before round t. Then $\ell_t = 0$, so

$$f_t(x_t) = \frac{a_t}{2}(x_t - z_{t-1})^2 \geq 0,$$

and therefore

$$\mathbb{E}[f_t(x_t) \mid \varepsilon_1, \ldots, \varepsilon_{t-1}] \geq 0.$$

In both cases,

$$\mathbb{E}[f_t(x_t) \mid \varepsilon_1, \ldots, \varepsilon_{t-1}] \geq 0.$$

Thus, we get that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(x_t)\right] = \sum_{t=1}^{T} \mathbb{E}[f_t(x_t)] = \sum_{t=1}^{T} \mathbb{E}_{\varepsilon_1,\ldots,\varepsilon_{t-1}} \mathbb{E}[f_t(x_t) \mid \varepsilon_1, \ldots, \varepsilon_{t-1}] \geq 0. \tag{31}$$

Step 4: A regularized cumulative loss that has an exact quadratic invariant

Define the fictitiously regularized cumulative loss

$$G_t(x) := F_t(x) + \frac{\gamma}{2}x^2.$$

We claim that, for all $t \geq 0$,

$$G_t(x) = G_t(z_t) + \frac{s_t + \gamma}{2}(x - z_t)^2, \tag{32}$$

and for $t \geq 1$,

$$G_t(z_t) = G_{t-1}(z_{t-1}) - \frac{\ell_t^2}{2(s_t + \gamma)}. \tag{33}$$

To see this, observe that for $t = 0$, $G_0(x) = \frac{\gamma}{2}x^2$, and since $z_0 = 0$ and $s_0 = 0$, (32) holds. Now, let $t \geq 1$, and suppose that (32) holds for $t - 1$. Then, we have that

$$\begin{aligned}
G_t(x) &= G_{t-1}(x) + f_t(x) \\
&= G_{t-1}(z_{t-1}) + \frac{s_{t-1} + \gamma}{2}(x - z_{t-1})^2 + \frac{a_t}{2}(x - z_{t-1})^2 - \ell_t(x - z_{t-1}) \\
&\qquad\qquad \text{(by definition of } f_t\text{)} \\
&= G_{t-1}(z_{t-1}) + \frac{s_t + \gamma}{2}(x - z_{t-1})^2 - \ell_t(x - z_{t-1}). \qquad \text{(using that } a_t = s_t - s_{t-1}\text{)} \\
&= G_{t-1}(z_{t-1}) + \frac{s_t + \gamma}{2}\left[(x - z_{t-1} - \frac{\ell_t}{s_t + \gamma}\right]^2 - \frac{\ell_t^2}{2(s_t + \gamma)}. \quad \text{(completing the square)}
\end{aligned}$$

Substituting $x = z_t$, and recalling the center update from (29) gives (33). Thereafter, substituting the expression for $G_t(z_t)$ from (33) back in the above gives (32).

Step 5: Telescoping the regularized cumulative loss

Telescoping (33) gives

$$G_T(z_T) = -\sum_{t=1}^{T} \frac{\ell_t^2}{2(s_t + \gamma)}. \tag{34}$$

Since $z_T \in \mathcal{X}$, and using that $F_T(x) = G_T(x) - (\gamma/2)x^2 \leq G_T(x)$,

$$\inf_{x \in \mathcal{X}} F_T(x) \leq F_T(z_T) \leq G_T(z_T).$$

Therefore,

$$-\inf_{x \in \mathcal{X}} F_T(x) \geq -G_T(z_T) = \sum_{t=1}^{T} \frac{\ell_t^2}{2(s_t + \gamma)}. \tag{35}$$

Taking expectation, and combining (31), (35), we get

$$\mathbb{E}\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \geq \mathbb{E}\left[\sum_{t=1}^{T} \frac{\ell_t^2}{2(s_t + \gamma)}\right]. \tag{36}$$

Note that the sole randomness in the expectations above is from $\varepsilon_1, \ldots, \varepsilon_T$.

Step 6: Process remains active throughout with constant probability

Observe that the centers $z_0, z_1, \ldots, z_T$ follow a random process, and if the process is active before round t,

$$z_t = \sum_{i=1}^{t} \frac{\ell_i}{s_i + \gamma}.$$

Consider a coupling (z_t, m_t), where $m_t = z_t$ up until the time the process becomes inactive, but thereafter, m_t continues to evolve as above even after z_t has become constant. That is, for every t,

$$m_t = \sum_{i=1}^{t} \frac{\ell_i}{s_i + \gamma}.$$

Note that each term in the sum above has mean 0. Furthermore, the variance of m_T is

$$\mathrm{Var}[m_T] = \mathbb{E}[m_T^2] = g^2 \sum_{t=1}^{T} \frac{1}{(s_t + \gamma)^2}.$$

We are now in a position to invoke Kolmogorov's maximal inequality [Billingsley, 2017, Theorem 22.4].

Proposition B.2 (Kolmogorov's maximal inequality)**.** *Let $m_t = \sum_{i=1}^{t} x_i$, where $x_1, \ldots, x_T$ are independent mean-zero random variables with finite variance. Then, for every $r > 0$,*

$$\Pr\left[\max_{1 \le t \le T} |m_t| \ge r\right] \le \frac{\mathbb{E}[m_T^2]}{r^2}.$$

By Proposition B.2 applied to our sequence $m_1, \ldots, m_T$, we get that

$$\Pr\left[\max_{1 \le t \le T} |m_t| \ge D/4\right] \le \frac{16 g^2}{D^2} \sum_{t=1}^{T} \frac{1}{(s_t + \gamma)^2} \le \frac{1}{4}, \tag{37}$$

where we used the assumed bound from Item 2 on $H_2(\gamma)$, and recalled that $g = L/8$. Thus, with probability at least $3/4$, the sequence m_t never exceeds $D/4$ in absolute value. But on this event, m_t coincides with z_t. We therefore conclude that with probability at least $3/4$, the process stays active throughout, i.e., every $\ell_t = g\varepsilon_t$; let A denote this event. Recalling the lower bound on regret from (36), we get

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] &\ge \mathbb{E}\left[\sum_{t=1}^{T} \frac{\ell_t^2}{2(s_t + \gamma)}\right] \\
&= \Pr[A] \cdot \mathbb{E}\left[\sum_{t=1}^{T} \frac{\ell_t^2}{2(s_t + \gamma)} \;\middle|\; A\right] + \underbrace{\Pr[\neg A] \cdot \mathbb{E}\left[\sum_{t=1}^{T} \frac{\ell_t^2}{2(s_t + \gamma)} \;\middle|\; \neg A\right]}_{\ge 0} \\
&\ge \frac{3L^2}{512} \sum_{t=1}^{T} \frac{1}{s_t + \gamma}.
\end{aligned} \tag{38}$$

Step 7: Choosing γ

We now choose γ appropriately, so that both Items 1 and 2 hold. Let

$$\Psi(\gamma) := D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma}, \qquad \gamma \ge 0.$$

We claim that $\inf_{\gamma \ge 0} \Psi(\gamma)$ is attained at some $\gamma_\star \ge 0$, and $\gamma_\star$ satisfies both the required conditions (hence, we will choose $\gamma = \gamma_\star$). To see this, note first that

$$\Psi(2L/D) = 2LD + L^2 \sum_{t=1}^{T} \frac{1}{s_t + 2L/D} < \infty,$$

so let $B := \Psi(2L/D) < \infty$. Note that $B \geq 2LD$. Then, for any $\gamma > \frac{B+1}{D^2}$,

$$\Psi(\gamma) \geq D^2\gamma > D^2\left(\frac{B+1}{D^2}\right) = B+1 > B.$$

Thus, if the infimum $\inf_{\gamma\geq 0}\Psi(\gamma)$ is attained at all, it is attained in $[0, \frac{B+1}{D^2}]$, which includes $2L/D$.

Now note that if every $s_t > 0$, then $\Psi(\gamma)$ is continuous on $[0, \frac{B+1}{D^2}]$. However, if any $s_t = 0$, then the term $\frac{1}{s_t+\gamma}$ blows up at $\gamma = 0$. Nevertheless, note in this case that for any $\gamma > 0$,

$$\Psi(\gamma) \geq \frac{L^2}{s_t+\gamma} = \frac{L^2}{\gamma}.$$

So, if we set $\delta := \frac{L^2}{B+1}$, then for any $0 < \gamma < \delta$,

$$\Psi(\gamma) \geq \frac{L^2}{\gamma} > \frac{L^2}{\delta} = B+1.$$

Furthermore, $\delta = \frac{L^2}{B+1} \leq \frac{L^2}{2LD+1} \leq \frac{L^2}{2LD} < 2L/D$. Thus, in this case, if the infimum $\inf_{\gamma\geq 0}\Psi(\gamma)$ is attained at all, it is attained in $[\delta, \frac{B+1}{D^2}]$. Furthermore, Ψ is continuous on this interval.

So, define the closed and bounded interval $K \subseteq [0,\infty)$ as

$$K = \begin{cases} \left[0, \frac{B+1}{D^2}\right] & \text{if } s_t > 0 \text{ for all } t, \\ \left[\delta, \frac{B+1}{D^2}\right] & \text{otherwise.} \end{cases} \tag{39}$$

Then, it always holds that

- $2L/D \in K$.
- For any $\gamma \geq 0$ such that $\gamma \notin K$, $\Psi(\gamma) > \Psi(2L/D)$.
- Ψ is continuous on K.

Then, by the extreme value theorem, there exists $\gamma_\star \in K$ such that

$$\Psi(\gamma_\star) = \min_{\gamma\in K}\Psi(\gamma) \leq \Psi(2L/D).$$

Together with the properties of K above, we have shown that $\gamma_\star$ attains $\inf_{\gamma\geq 0}\Psi(\gamma)$.

We now check Items 1 and 2 for $\gamma = \gamma_\star$. Note that if $s_t > 0$ for all t, it immediately holds that $s_t + \gamma_\star > 0$ for all t. On the hand, if some $s_t = 0$, then $\gamma_\star > 0$, and we again get that $s_t + \gamma_\star > 0$ for all t. Thus, Item 1 holds.

Now, for $\gamma > 0$, we have that

$$\Psi'(\gamma) = D^2 - L^2\sum_{t=1}^{T}\frac{1}{(s_t+\gamma)^2} = D^2 - L^2 H_2(\gamma).$$

If $\gamma_\star > 0$, by the first-order optimality condition, we have that

$$H_2(\gamma_\star) = \frac{D^2}{L^2}. \tag{40}$$

On the other hand, if $\gamma_\star = 0$, the right derivative of Ψ at $\gamma = 0$ must be non-negative, which gives

$$H_2(\gamma_\star) \leq \frac{D^2}{L^2}$$

Thus, we have shown that Item 2 holds.

Step 8: Putting things together, and concluding

Recalling the regret lower bound (38), and plugging in $\gamma_\star$, we have that

$$\mathbb{E}\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \geq \frac{3L^2}{512} \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star}.$$

We now wish to further lower bound the right-hand side above with $\Psi(\gamma_\star)$. For this, suppose that $\gamma_\star = 0$. In this case,

$$\Psi(\gamma_\star) = L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star}.$$

On the other hand, if $\gamma_\star > 0$, we have by the first-order optimality condition (40) that

$$\begin{aligned} D^2 &= L^2 \sum_{t=1}^{T} \frac{1}{(s_t + \gamma_\star)^2} \\ \implies D^2\gamma_\star &= L^2 \sum_{t=1}^{T} \frac{\gamma_\star}{(s_t + \gamma_\star)^2} = L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star} \cdot \frac{\gamma_\star}{s_t + \gamma_\star} \leq L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star}, \end{aligned}$$

and hence,

$$\Psi(\gamma_\star) = D^2\gamma_\star + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star} \leq 2L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star}.$$

In either case, we are guaranteed that $\Psi(\gamma_\star) \leq 2L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma_\star}$; substituting this in our regret bound above, we get

$$\mathbb{E}\left[\mathrm{Regret}_T(f_{1:T}; x_{1:T})\right] \geq \frac{3}{1024}\Psi(\gamma_\star) = \frac{3}{1024} \cdot \inf_{\gamma \geq 0} \left[D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma} \right]. \tag{41}$$

The bound above holds for every deterministic learning algorithm, over the randomness of $\varepsilon_1, \ldots, \varepsilon_T$. This implies a corresponding lower bound for any randomized algorithm, by the standard Yao argument [Yao, 1983]. Concretely, let $\mathcal{A}$ be any randomized algorithm. Conditioned on any fixed realization r of its randomness, the algorithm $\mathcal{A}_r$ is deterministic. Thus, we can invoke the lower bound in (41) for $\mathcal{A}_r$, to get

$$\mathbb{E}_{\varepsilon_{1:T}}\left[\mathrm{Regret}_{\mathcal{A}_r,T}(f_{1:T}; x_{1:T})\right] \geq \frac{3}{1024} \cdot \inf_{\gamma \geq 0} \left[D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma} \right].$$

Averaging over the randomness r then gives that

$$\mathbb{E}_r\, \mathbb{E}_{\varepsilon_{1:T}}\left[\mathrm{Regret}_{\mathcal{A}_r,T}(f_{1:T}; x_{1:T})\right] = \mathbb{E}_{\varepsilon_{1:T}}\, \mathbb{E}_r\left[\mathrm{Regret}_{\mathcal{A}_r,T}(f_{1:T}; x_{1:T})\right] \geq \frac{3}{1024} \cdot \inf_{\gamma \geq 0} \left[D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma} \right].$$

In particular, this implies the existence of a fixed sequence $\varepsilon_1, \ldots, \varepsilon_T$, and corresponding functions $f_1, \ldots, f_T$, such that

$$\mathbb{E}_r\left[\mathrm{Regret}_{\mathcal{A}_r,T}(f_{1:T}; x_{1:T})\right] \geq \frac{3}{1024} \cdot \inf_{\gamma \geq 0} \left[D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma} \right].$$

Since the lower bound above holds for every randomized learning algorithm $\mathcal{A}$, we get

$$\begin{aligned} \mathfrak{R}_T(s_{1:T}) &\geq \frac{3}{1024} \cdot \inf_{\gamma \geq 0} \left[D^2\gamma + L^2 \sum_{t=1}^{T} \frac{1}{s_t + \gamma} \right] \\ &\geq c \cdot \inf_{\rho \geq 0} \left[L^2 \sum_{t=1}^{T} \min\left\{ \frac{D}{\rho}, \frac{1}{s_t} \right\} + D\rho \right]. \qquad \text{(using Claim B.1)} \end{aligned}$$

This concludes the proof of Theorem 3. □

B.1 Optimality of regime-specific rates

Here, we argue that the rates obtained by our algorithm in the different curvature in Corollary 3.1 are optimal.

Recall from Theorem 3 that the minimax regret is lower bounded (upto constant factors) by $\mathrm{OPT}_T^0(0)$, where

$$\mathrm{OPT}_T^0(0) \geq c \cdot \underbrace{\inf_{\rho \geq 0} \left[\sum_{t=1}^{T} \min\left\{ \frac{1}{\rho}, \frac{1}{S_t} \right\} + \rho \right]}_{:=\Psi(\rho)}.$$

Here, the constant c hides a dependence on d, D, L. It thus suffices to calculate $\Psi(\rho)$ in the regimes where $S_t = t^\gamma$ (for $\gamma \geq 1/2$), and show a matching characterization as that in Corollary 3.1.

Consider

$$\Psi(\rho) = \inf_{\rho \geq 0} \left[\sum_{t=1}^{T} \min\left\{ \frac{1}{\rho}, \frac{1}{S_t} \right\} + \rho \right],$$

for $S_t = t^\gamma$.

Towards this, we first further characterize the precise form of $\Psi(\rho)$. For $0 \leq t < T$, consider the intervals $I_t = [t^\gamma, (t+1)^\gamma)$, and let $I_T = [T^\gamma, \infty)$. For any $\rho \in I_t$, observe that $\Psi(\rho)$ evaluates to

$$\frac{t}{\rho} + \rho + \sum_{t < i \leq T} i^{-\gamma}.$$

Note that the last term is constant for every $\rho \in I_t$. Thus, the "unconstrained" minimizer for this interval balances the first two terms, giving

$$\hat{\rho}_t = \sqrt{t}.$$

This is a valid value for $\hat{\rho}_t$, provided $\sqrt{t} \in I_t$. However, if $\sqrt{t} \notin I_t$, we simply set $\hat{\rho}_t$ to be either t^γ or $(t+1)^\gamma$, based on whichever of the two evaluates to a smaller $\Psi(\rho)$. Thus, we get that

$$\inf_{\rho \geq 0} \Psi(\rho) = \min_{0 \leq t \leq T} \left(\frac{t}{\hat{\rho}_t} + \hat{\rho}_t + \sum_{t < i \leq T} i^{-\gamma} \right). \tag{42}$$

When $\gamma < 1/2$, it may be verified that the minimum above gets achieved at $t = T$ (for large enough T), whereby $\hat{\rho}_T = \sqrt{T}$, giving that the quantity in (42) is at least $T/\sqrt{T} + \sqrt{T} = \Omega(\sqrt{T})$. When $\gamma \geq 1/2$, the minimum in (42) gets achieved at $t = 0$ as $\hat{\rho}_0 \to 0$, whereby the quantity in (42) is at least $\sum_{i=1}^{T} i^{-\gamma}$, which gives the claimed lower bounds on the rates in each regime of γ.

B.2 Relating $\mathrm{OPT}_T^0(r)$ and $\mathrm{OPT}_T^0(0)$

Claim B.3. *Recall the definition of* $\mathrm{OPT}_T^0(r)$*:*

$$\mathrm{OPT}_T^0(r) = \inf_{\rho \geq r} \left[\sum_{t=1}^{T} \min\left\{ \frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t} \right\} + dD\rho \right].$$

For any $r \geq 0$*, it holds that*

$$\mathrm{OPT}_T^0(r) \leq \mathrm{OPT}_T^0(0) + dDr.$$

Proof. For any $\rho \geq 0$, let $\tilde{\rho} = \max(r, \rho)$. Then observe that for any t,

$$\min\left\{ \frac{125L^2d^2D}{\tilde{\rho}}, \frac{2L^2}{S_t} \right\} \leq \min\left\{ \frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t} \right\}.$$

Furthermore,

$$dD\tilde{\rho} \leq dD\rho + dDr.$$

Therefore,

$$\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\tilde{\rho}}, \frac{2L^2}{S_t}\right\} + dD\tilde{\rho} \leq \sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t}\right\} + dD\rho + dDr.$$

Infimizing over $\rho \geq 0$ gives,

$$\begin{aligned}
&\inf_{\rho\geq 0}\left[\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\tilde{\rho}}, \frac{2L^2}{S_t}\right\} + dD\tilde{\rho}\right] \leq \inf_{\rho\geq 0}\left[\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t}\right\} + dD\rho + dDr\right] \\
\implies &\inf_{\rho\geq r}\left[\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t}\right\} + dD\rho\right] \leq \inf_{\rho\geq 0}\left[\sum_{t=1}^{T} \min\left\{\frac{125L^2d^2D}{\rho}, \frac{2L^2}{S_t}\right\} + dD\rho\right] + dDr \\
\implies &\mathrm{OPT}_T^0(r) \leq \mathrm{OPT}_T^0(0) + dDr. \qquad \square
\end{aligned}$$

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 results we derive directly reflect what we state in the abstract.

   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: We are candid about the merits and limitations of our results in the paper.

   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: We state all assumptions necessary of our theoretical results.

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: [Yes]

   Justification: We only perform basic synthetic experiments that are simple to implement, and describe the setup in Section 5.

   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: [No]

Justification: We only perform basic synthetic experiments that are simple to implement, and describe the setup in Section 5.

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: We only perform basic synthetic experiments that are simple to implement, and describe the setup in Section 5.

   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: [Yes]

   Justification: We plot the standard deviation in our figures.

   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: We only perform basic synthetic experiments that are simple to implement, and describe the setup in Section 5.

   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: We have followed the code to the best of our knowledge.

   Guidelines:

   - 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: We don't anticipate any drastic societal impacts of our work.

    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: Our paper is primary theoretical, and we don't anticipate any 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: Not applicable. Please see our responses above regarding code and experiments.

    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: Not applicable. Please see our responses above regarding code and experiments.

    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: Not applicable.

    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: 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: [Yes]

    Justification: We used ChatGPT routinely in the development and verification of some of the technical ideas used in 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.