Optimal Dynamic Regret in Proper Online Learning with Strongly Convex Losses and Beyond

Dheeraj Baby
dheeraj@ucsb.edu

Yu-Xiang Wang
yuxiangw@cs.ucsb.edu

Dept. of Computer Science, UC Santa Barbara

Abstract

We study the framework of *universal dynamic regret* minimization with *strongly convex* losses. We answer an open problem in (Baby and Wang, 2021) by showing that in a *proper learning* setup, Strongly Adaptive algorithms can achieve the near optimal dynamic regret of $\tilde{O}(d^{1/3}n^{1/3}\text{TV}[u_{1:n}]^{2/3} \vee d)$ against any comparator sequence $u_1, \ldots, u_n$ *simultaneously*, where n is the time horizon and $\text{TV}[u_{1:n}]$ is the Total Variation of comparator. These results are facilitated by exploiting a number of *new* structures imposed by the KKT conditions that were not considered in (Baby and Wang, 2021) which also lead to other improvements over their results such as: (a) handling non-smooth losses and (b) improving the dimension dependence on regret. Further, we also derive near optimal dynamic regret rates for the special case of proper online learning with exp-concave losses and an L_∞ constrained decision set.

1 INTRODUCTION

Online Convex Optimization (OCO) (Hazan, 2016) is a powerful learning paradigm for the task of sequential decision making. It is modelled as an interactive game between a learner and adversary as follows: For each time step $t \in [n] := \{1, 2, \ldots, n\}$, the learner plays a point $\boldsymbol{x}_t \in \mathbb{R}^d$. Then the adversary reveals a convex loss $f_t : \mathbb{R}^d \to \mathbb{R}$. A common objective in online learning is to minimize the learner's static regret against a convex set of benchmark points $\mathcal{D} \subset \mathbb{R}^d$: $R_{\text{static}} = \sum_{t=1}^n f_t(\boldsymbol{x}_t) - \inf_{\boldsymbol{w} \in \mathcal{D}} \sum_{t=1}^n f_t(\boldsymbol{w})$.

However, the notion of static regret is not befitting to applications where the environment is non-stationary. To alleviate this issue, one may aim to control the dynamic regret against a sequence of comparators in $\mathcal{D}$ (the comparator sequence may be potentially unknown to the learner):

$$R_{\text{dynamic}}(\boldsymbol{w}_{1:n}) = \sum_{t=1}^n f_t(\boldsymbol{x}_t) - \sum_{t=1}^n f_t(\boldsymbol{w}_t),$$

where we use the shorthand $\boldsymbol{w}_{1:n} := \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$. Here each $\boldsymbol{w}_t \in \mathcal{D}$. Dynamic regret rates are usually expressed in terms of the time horizon n and a regularity measure aka path length that captures the smoothness of the comparator sequence. For example, in (Zhang et al., 2018a) a regularity measure $V_n(\boldsymbol{w}_{1:n}) = \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_2$ is defined. They propose an algorithm that attains a (near) optimal dynamic regret of $\tilde{O}(\sqrt{n(1 + V_n(\boldsymbol{w}_{1:n}))})$ when the losses f_t are convex ($\tilde{O}$ hides factors of $\log n$.) Such dynamic regret rates are sometimes referred as *universal dynamic regret* rates as they are applicable to any comparator sequence $\boldsymbol{w}_{1:n}$.

However, optimal dynamic regret rates in terms of path length of the arbitrary comparator sequence when the loss functions have extra curvature properties such as strong convexity or exp-concavity, have been long eluded in the literature until a recent breakthrough by (Baby and Wang, 2021). They define a path length in terms of the Total Variation (TV) of the comparator sequence as: $\text{TV}(\boldsymbol{w}_{1:n}) = \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1$. They show that when the losses are strongly convex / exp concave and gradient Lipschitz, a Strongly Adaptive (SA) online learner ((Hazan and Seshadhri, 2007; Daniely et al., 2015)) can attain a (near) optimal dynamic regret rate of $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$ [1] against all sequences with

Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS) 2022, Valencia, Spain. PMLR: Volume 151.

[1] $\tilde{O}^*$ hides factors of d and $\log n$; $a \vee b := \max\{a, b\}$.

$\mathrm{TV}(\boldsymbol{w}_{1:n}) \leq C_n$ where C_n is a quantity that may be unknown to the learner. However, this rate is attained using an improper SA algorithm whose decisions can lie outside $\mathcal{D}$. A question that was left open was whether improper learning is strictly necessary to achieve the optimal rates for exp-concave optimization. In this work, we answer this in the negative by showing that a proper version of the SA algorithms can attain the optimal (modulo log factors and dimension dependencies) dynamic regret rates whenever the losses are strongly convex.

We summarize our main contributions below.

- We provide a new analysis that extends the results of (Baby and Wang, 2021) to proper strongly convex online learning to attain the near *optimal* dynamic regret rate of $\tilde{O}(d^{1/3}n^{1/3}C_n^{2/3} \vee d)$ for Strongly Adaptive methods (see Corollary 5). In contrast to (Baby and Wang, 2021), our results imply an important conclusion that improper learning is *not strictly necessary* for attaining such fast rates with general strongly convex losses. To the best of our knowledge, this is the *first* result that achieves near optimal dynamic regret in a setting of proper learning under strongly convex losses.
- For exp-concave losses, we prove an analogous result that Strongly Adaptive algorithms can attain a near optimal dynamic regret of $\tilde{O}^*((n^{1/3}C_n^{2/3} \vee 1))$ in the special case of L_∞ (box) constrained decision set, $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$ (see Theorem 10).
- To facilitate these results we discover and exploit a number of new structures imposed by the KKT conditions that were not considered in (Baby and Wang, 2021), which could be of independent interest.

Notes on scope and relevance. Under exp-concave or strongly convex losses, the important question of finding an optimal (wrt universal dynamic regret) and proper algorithm has remained resistant to attacks in the non-stationary online learning literature for almost two decades since the work of (Zinkevich, 2003). In this work, we take the first steps in addressing this question by showing optimality of proper SA learner in proper learning settings. The fact that a proper version of Strongly Adaptive algorithms can lead to optimal rates was highly unclear from the analysis of (Baby and Wang, 2021). Further, by lifting the gradient smoothness assumption for the revealed losses, we modestly enlarge the applicability of the results when compared to (Baby and Wang, 2021). Though our proof techniques bear some semblance with that of (Baby and Wang, 2021) in terms of the usage of KKT conditions, this similarity is only superficial and we introduce several new non-trivial ideas in the analysis for attaining the new results (see Sections 2.2 and 4.1).

2 A GENTLE START: SQUARED LOSS GAMES

To start with, we consider the following squared loss game which will later play a pivotal role in the generalization to strongly convex losses.

- At time $t \in [n] := \{1, \ldots, n\}$, player predicts $x_t \in [-B, B]$.
- Adversary reveals a label $y_t \in [-G, G]$
- Player suffers loss $(y_t - x_t)^2$.

We make the following assumption.

Assumption A1: We assume that $[-B, B] \subseteq [-G, G]$ with $B \geq 1$ without loss of generality.

Define a class of comparators as:

$$\mathcal{TV}^B(C_n) := \left\{ w_{1:n} : \mathrm{TV}(w_{1:n}) := \sum_{t=2}^{n} |w_t - w_{t-1}| \leq C_n, \right.$$
$$\left. |w_t| \leq B \; \forall t \in [n] \right\}.$$

We are interested in simultaneously controlling the dynamic regret against all sequences in $\mathcal{TV}^B(C_n)$. The main algorithm we use for this task is the Follow-the-Leading-History (FLH) from (Hazan and Seshadhri, 2007) with Online Gradient Descent (OGD) run on the decision set $[-B, B]$ as base learners. This algorithm will be referred as *FLH-OGD* strategy henceforth. We provide a description of FLH in Appendix B for completeness. We have the following performance guarantee.

Theorem 1. *Suppose the labels y_t generated by the adversary belong to $[-G, G]$. Let x_t be the prediction at time t of FLH with learning rate $\zeta = 1/(2(G+B)^2)$, base learners as OGD with step sizes $1/(2t)$ and decision set $[-B, B]$. Then for any comparator sequence $(w_1, \ldots, w_n) \in \mathcal{TV}^B(C_n)$*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - w_t)^2 = \tilde{O}\left(n^{1/3} C_n^{2/3} \vee 1\right),$$

where $\tilde{O}(\cdot)$ hides dependence on logarithmic factors of horizon n, G, B and $a \vee b := \max\{a, b\}$.

Remark 2 **(Adaptivity to C_n and safe (non-stochastic) oracle inequality).** *The FLH-OGD strategy does not require C_n as an input. Further, Theorem 1 has implications in non-parametric regression under safety constraints. When the non-parametric estimator for the $\mathcal{TV}^B$ sequence class is required to obey a safety constraint that the estimator's outputs x_t must also lie in $[-B, B]$, Theorem 1 implies the following oracle inequality:*

$$\sum_{t=1}^{n}(y_t - x_t)^2 + g(x_t)$$
$$\leq \min_{w_{1:n}} \sum_{t=1}^{n}(y_t - w_t)^2 + g(w_t) + \tilde{O}\left(n^{1/3}\mathrm{TV}(w_{1:n})^{2/3} \vee 1\right),$$

where $g(x)$ is a safety constraint such that $g(x) = \infty$ when $|x|> B$ and zero otherwise. This is a strict generalization of Remark 2 in (Baby and Wang, 2021).

2.1 Key insight behind proof of Theorem 1

The insight we used in deriving regret rate in Theorem 1 for a proper learning setup is based on the following idea: Suppose that we need to compete against a comparator sequence that incurs a Total Variation (TV) of C_n. We observe that, this comparator sequence of decisions in hindsight requires to obey the TV constraint while the decisions of the Strongly Adaptive (SA) learner need not obey any such constraints. Consider a time interval I where the comparator sequence assumes a constant value (say v_1) in an *arbitrary* convex decision set D. There could be some other point in D (say v_2) which can incur better cumulative loss within that interval. Note that the comparator sequence may not assume the value v_2 in the interval I due to the global TV constraint. Due to the strongly adaptive property, the regret (against v_1) of the SA learner in interval I is then bounded by the regret (against v_1) of the static point v_2, which is less than or equal to zero, plus an extra log term. The presence of such non-positive terms can delicately offset the effect of the positive log terms when summed across all such intervals to get favorable dynamic regret rates. How small the non-positive terms are, when summed across all intervals, depends on the magnitude of C_n (and indirectly on n).

2.2 Detailed road map for the proof of Theorem 1

In this section, we focus on conveying the main ideas of our proof deferring the formal details to Appendix C. We start by briefly reviewing the proof strategy of (Baby and Wang, 2021) and then intuitively capture the points of similarities and differences in our analysis. Throughout the proof we use the shorthand $[a, b] := \{a, a + 1, \ldots, b\}$ for two natural numbers $a < b$.

We start by characterizing the offline optimal. Define the sign function as $\mathrm{sign}(x) = 1$ if $x > 0$; -1 if $x < 0$; and some $v \in [-1, 1]$ if $x = 0$.

Lemma 3. ***(characterization of offline optimal)*** *Consider the following convex optimization problem (where $\tilde{z}_1, \ldots, \tilde{z}_{n-1}$ are introduced as dummy variables)*

$$\min_{\tilde{u}_1,\ldots,\tilde{u}_n,\tilde{z}_1,\ldots,\tilde{z}_{n-1}} \quad \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2$$
$$\text{s.t.} \quad \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \ \forall t \in [n-1],$$
$$\sum_{t=1}^{n-1}|\tilde{z}_t| \leq C_n, \tag{1a}$$
$$-B \leq \tilde{u}_t \ \forall t \in [n], \tag{1b}$$
$$\tilde{u}_t \leq B \ \forall t \in [n], \tag{1c}$$

Let $u_1, \ldots, u_n, z_1, \ldots, z_{n-1}$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the constraint (1a). Further, let $\gamma_t^- \geq 0, \gamma_t^+ \geq 0$ be the optimal dual variables that correspond to constraints (1b) and (1c) respectively for all $t \in [n]$. By the KKT conditions, we have

- ***stationarity:*** *$u_t - y_t = \lambda(s_t - s_{t-1}) + \gamma_t^- - \gamma_t^+$, where $s_t \in \partial|z_t|$ (a subgradient). Specifically, $s_t = \mathrm{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t|> 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.*
- ***complementary slackness:*** *(a) $\lambda\left(\sum_{t=2}^{n}|u_t - u_{t-1}|-C_n\right) = 0$; (b) $\gamma_t^-(u_t + B) = 0$ and $\gamma_t^+(u_t - B) = 0$ for all $t \in [n]$*

Let the optimal solution constructed by the offline oracle be denoted by $u_{1:n}$ (termed as offline optimal henceforth). In (Baby and Wang, 2021), a partition $\mathcal{P} = \{[i_s, i_t], i \in [M]\}$ of $[n]$ is formed with cardinality $|\mathcal{P}|= M = O(n^{1/3}C_n^{2/3} \vee 1)$. The partition has an additional property that within each bin $[i_s, i_t] \in \mathcal{P}$, we have $C_i := \sum_{j=i_s+1}^{i_t}|u_j - u_{j-1}|\leq B/\sqrt{i_t - i_s + 1}$ (see Lemma 17). Then for each bin, a three term regret decomposition is employed as follows:

$$\underbrace{\sum_{j=i_s}^{i_t}(y_j - x_j)^2 - (y_j - \bar{y}_i)^2}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \bar{y}_i)^2 - (y_j - \bar{u}_i)^2}_{T_{2,i}}$$
$$+ \underbrace{\sum_{j=i_s}^{i_t}(y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{3,i}}, \tag{2}$$

where $\bar{u}_i = \sum_{j=i_s}^{i_t} u_j/(i_t - i_s + 1)$ and $\bar{y}_i = \sum_{j=i_s}^{i_t} y_j/(i_t - i_s + 1)$ and x_j are the predictions of the learner. They use online averaging as base learners for FLH. By strong adaptivity, they show $T_{1,i} = O(\log n)$. They show that $T_{3,i}$ can be $O(\lambda C_i)$ in general where λ is the dual variable arising from the KKT conditions (see Lemma 3) which can be even $\Theta(n)$ in the worst case. Since $\bar{y}_i$ is the static minimizer of $g(x) = \sum_{j=i_s}^{i_t}(y_j - x)^2$, they bound $T_{2,i}$ by a non-positive term which when added to $T_{3,i}$ can diminish into an $O(1)$ quantity. Thus regret within the bin $[i_s, i_t]$ is $T_{1,i} + T_{2,i} + T_{3,i} = O(\log n)$. This regret bound is added across all $O(n^{1/3}C_n^{2/3} \vee 1)$ bins of $\mathcal{P}$ to yield an $\tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$ dynamic regret.

In our protocol of squared loss games, the labels $y_t \in [-G, G] \supseteq [-B, B]$. So we can't use online averages as base learner for constructing a proper learning algorithm. So in this work we use projected OGD as base learners with decision set $[-B, B]$. With such an algorithm, we may attempt to work with a slightly modified version of the three term regret decomposition of (2) as:

$$\underbrace{\sum_{j=i_s}^{i_t}(y_j - x_j)^2 - (y_j - \Pi(\bar{y}_i))^2}_{T'_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \Pi(\bar{y}_i))^2 - (y_j - \bar{u}_i)^2}_{T'_{2,i}} + \underbrace{\sum_{j=i_s}^{i_t}(y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T'_{3,i}}, \tag{3}$$

where $\Pi(x)$ is the projection of $x \in \mathbb{R}$ to the interval $[-B, B]$. Unfortunately while doing so, the term $T'_{2,i}$ can be not negative enough to diminish $T'_{3,i}$ to an $O(1)$ quantity. We provide an empirical demonstration of this phenomenon in Fig.1. At this point, we hope that we have made a clear case on why the analysis of (Baby and Wang, 2021) cannot be directly extended to handle proper learning.

To get around this issue, we first identify two regimes for the dual variable λ. We show that when $\lambda = O(n^{1/3}/C_n^{1/3})$, one can still work with the same partitioning $\mathcal{P}$ of (Baby and Wang, 2021) (see Lemma 17) and use a decomposition similar to Eq.(3) to get the desired regret bound (see Lemma 19).

Before explaining the details of the regime $\lambda = \Omega(n^{1/3}/C_n^{1/3})$, we introduce the following definitions for convenience:

Definition 4.

- *For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 1 if $u_j = u_a \in (-B, B)$ for all $j \in [a, b]$ and $u_b > u_{b+1}$ and $u_a > u_{a-1}$.*
- *For a bin $[a, b] \subseteq \{2, \ldots, n-1\}$, the offline optimal solution is said to assume Structure 2 if $u_j = u_a \in (-B, B)$ for all $j \in [a, b]$ and $u_b < u_{b+1}$ and $u_a < u_{a-1}$.*
- *For a bin $[a, b]$, we define $gap_{min}(\beta, [a, b]) := \min_{j \in [a,b]} |u_j - \beta|$ where $\beta \in \mathbb{R}$.*

Consider the following two conditions.

Condition 1: For a bin $[i_s, i_t] \in \mathcal{P}$, the offline optimal satisfies $\text{gap}_{\min}(-B, [i_s, i_t]) \geq \text{gap}_{\min}(B, [i_s, i_t])$ and within at-least one sub-interval $[r, s] \subseteq [i_s, i_t]$, the offline optimal assumes the form of Structure 2.

Condition 2: For a bin $[i_s, i_t] \in \mathcal{P}$, the offline optimal satisfies $\text{gap}_{\min}(-B, [i_s, i_t]) < \text{gap}_{\min}(B, [i_s, i_t])$ and within at-least one sub-interval $[r, s] \subseteq [i_s, i_t]$, the offline optimal assumes the form of Structure 1.

Define:

$$\mathcal{Q} := \{[i_s, i_t] \in \mathcal{P} : \text{the offline optimal satisfies Condition 1 or 2 in } [i_s, i_t]\}.$$

We refine a bin $[i_s, i_t] \in \mathcal{Q}$ that satisfy Condition 1 into smaller sub-intervals as shown in Fig.2, such that: for a style U sub-interval, the offline optimal takes the form of Structure 2 and for a style V sub-interval, the offline optimal has a non-decreasing section followed by an optional decreasing section. A similar refinement is also performed for bins in $\mathcal{Q}$ that satisfy Condition 2.

Our strategy is to bound:

$$\text{regret in style U sub-intervals} = O(\log n) + \text{a negative term}. \tag{4}$$

This is accomplished by a two term regret decomposition. Suppose $[a, b]$ is a style U sub-interval. We use the decomposition:

$$\underbrace{\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - w)^2}_{T_1} + \underbrace{\sum_{j=a}^{b}(y_j - w)^2 - (y_j - u_j)^2}_{T_2}, \tag{5}$$

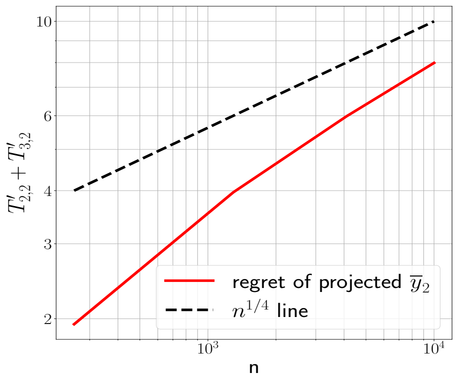

Figure 1: *Plot of $T'_{2,2} + T'_{3,2}$ (see Eq.(3) with $i = 2$) for the Example 15 in Appendix C. In this example, $C_n = O(1/\sqrt{n})$ and the partitioning procedure of (Baby and Wang, 2021) creates a partition $\mathcal{P}$ of $[n]$ containing two bins. We see that $T'_{2,2}+T'_{3,2}$ in the second bin grows roughly as $O(n^{1/4})$. However for applying the analysis of (Baby and Wang, 2021), we require this quantity for each bin in $\mathcal{P}$ to grow as $O(1)$. This makes the direct extension of the techniques in (Baby and Wang, 2021) with $\bar{y}_i$ replaced by $\Pi(\bar{y}_i)$ as in Eq.(3) inapplicable for the proper learning setting we study.*

with $w = \Pi\left(\sum_{j=a}^{b} y_j/(b-a+1)\right)$.

Next, we bound

$$\text{regret in style V sub-intervals} = O(\log n), \quad (6)$$

using a similar two term regret decomposition as in Eq.(5) with w replaced by a carefully chosen $w_j \in [(u_a \wedge \ldots \wedge u_b), (u_a \vee \ldots \vee u_b)]$ such that $\sum_{j=a+1}^{b} \mathbb{I}\{w_j \neq w_{j-1}\} \leq 6$ where $\mathbb{I}\{\cdot\}$ is the indicator function taking values in $\{0, 1\}$. We use the notation $x \wedge y = \min\{x, y\}$

We perform this task of refinement for every interval $[i_s, i_t]$ in $\mathcal{Q}$. Then we bound the regret in the resulting sub-intervals (as per Eq.(4) or (6)) and add the regret bounds across all such sub-intervals. Note that the total number of sub-intervals after refinement can be much larger than $|\mathcal{P}|= O(n^{1/3}C_n^{2/3} \vee 1)$. So if the bound in Eq.(4) is not tight enough, then there is a possibility that the resulting regret bound can be highly sub-optimal. This poses a major challenge in contrast to the analysis of (Baby and Wang, 2021) where they only need to work with a partition of size $O(n^{1/3}C_n^{2/3} \vee 1)$ and bound the regret in each interval of the partition by an $\tilde{O}(1)$ quantity.

To address this issue, we form tight bounds for Eq.(4) by exploiting certain structures in the KKT conditions that were previously unexplored in (Baby and Wang, 2021) via Lemmas 16, 22, 23 and 24. Of particular interest is Lemma 16 which highlights a fundamental way in which the adversary is constrained. Then we prove that if every bin $[i_s, i_t] \in \mathcal{Q}$ satisfies $\text{gap}_{\min}(-B, [i_s, i_t]) \vee \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ where μ_{th} is as defined in Lemma 23, then the culmination of the negative terms in Eq.(4) can gracefully offset the effect of the positive $O(\log n)$ terms in Eq.(4) and Eq.(6) when summed across all refined intervals to obtain an $O(n^{1/3}C_n^{2/3} \vee 1)$ bound overall for $\sum_{[i_s,i_t]\in\mathcal{Q}} \sum_{j=i_s}^{i_t} (y_j - x_j)^2 - (y_j - u_j)^2$ (see proof of Lemma 24).

Further we show in Lemma 23 that when $\lambda = \Omega(n^{1/3}/C_n^{1/3})$ and $C_n = \tilde{O}(n)$, the criterion $\text{gap}_{\min}(-B, [i_s, i_t]) \vee \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ is always satisfied for every bin $[i_s, i_t] \in \mathcal{Q}$. This can be seen informally as follows. Recall that the TV of the offline optimal within the bin $[i_s, i_t]$ is a "small" quantity that is at-most $(B/\sqrt{i_t - i_s + 1}) \leq B$. So if $\text{gap}_{\min}(-B, [i_s, i_t])$ is small, then due to this small TV constraint, we expect the quantity $\text{gap}_{\min}(B, [i_s, i_t])$ to be sufficiently large and vice versa.

Finally, for each bin in $\mathcal{R} := \mathcal{P} \setminus \mathcal{Q}$ we show (by using Lemma 20) that its regret contribution can be bounded by $O(\log n)$. Since $|\mathcal{R}|= O(n^{1/3}C_n^{2/3} \vee 1)$, such regret bounds lead to $\tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$ bound overall when summed across all bins in $\mathcal{R}$.

Before closing this section, we capture the intuition behind the importance of the criterion $\text{gap}_{\min}(-B, [i_s, i_t]) \vee \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ and why it can produce a sufficiently negative term in Eq.(4). Let's consider a style U sub-interval $[a, b]$ obtained by refining a bin $[i_s, i_t] \in \mathcal{Q}$ which satisfy Condition 1. Since $[a, b]$ is style U sub-interval, the offline optimal takes the form of Structure 2 in $[a, b]$. Suppose that $|B + u_a|\geq \text{gap}_{\min}(-B, [i_s, i_t]) \geq \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$. Here the first inequality holds by the definition of $\text{gap}_{\min}(-B, [i_s, i_t])$. Also, note that $u_j = u_a$ for all $j \in [a, b]$ by the definition of Structure 2. Let $\bar{y}_{a\to b} := \sum_{j=a}^{b} y_j/(b-a+1)$. From the KKT conditions it can be shown that $\bar{y}_{a\to b} < u_a$. We provide intuitive explanation for the case $\Pi(\bar{y}_{a\to b}) = -B$. This can happen only when $\bar{y}_{a\to b} \leq -B$. Qualitatively in such a scenario, we expect the decision $-B$ to be much better than playing the decision u_a which is bigger than $-B$. Whenever there is sufficient gap (more formally a gap of at-least μ_{th}) between $-B$ and u_a, one can expect that u_a can be very sub-optimal in comparison to $-B$ ($= \Pi(\bar{y}_{a\to b})$) which makes the term T_2 in Eq.(5) (with $w = -B$ and $u_j = u_a$) sufficiently negative.

When $\bar{y}_{a\to b} \in (-B, B)$, T_2 with $w = \bar{y}_{a\to b}$ can be shown to be sufficiently negative using the arguments of (Baby and Wang, 2021). However, the interplay of this negative term with the sum of regret bounds in all refined intervals is more delicate as described in the proof of Lemma 24.

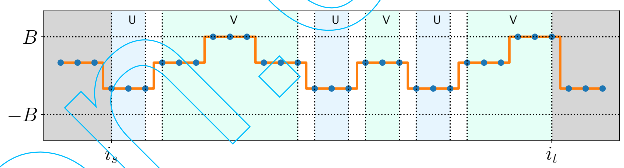

Figure 2: *Refinement of a bin $[i_s, i_t] \in \mathcal{P}$ that satisfy Condition 1 in Section 2.2 into smaller style U and style V sub-intervals. Blue dots represent the optimal sequence*

3 PERFORMANCE GUARANTEES FOR STRONGLY CONVEX LOSSES

In this section, we extend the results on squared error losses to general strongly convex losses.

3.1 Strongly convex losses and box decision set

In this section, we show that the style of analysis presented for squared error losses directly generalizes to strongly convex losses in multi-dimensions whenever the decision set is an L_∞ norm ball. The main idea is to provide a reduction to the uni-variate squared loss games via standard surrogate loss tricks (Hazan et al., 2007) and instantiate FLH-OGD appropriately. All unspecified proofs for this section are deferred to Appendix D. We consider the following protocol:

- At time $t \in [n]$ learner predicts $\boldsymbol{x}_t \in \mathbb{R}^d$ with $\|\boldsymbol{x}_t\|_\infty \leq B$.
- Adversary reveals loss f_t.
- Learner suffers loss $f_t(\boldsymbol{x}_t)$.

We have the following Corollary due to Theorem 1.

Corollary 5. *Let the loss functions f_t be H strongly convex in L_2 norm across the (box) domain $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$. i.e, $f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + \nabla f_t(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}) + \frac{H}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$. Suppose $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G_\infty$ for all $\boldsymbol{x} \in \mathcal{D}$. For each $i \in [d]$, construct surrogate losses $\ell_t^{(i)} : \mathbb{R} \to \mathbb{R}$ as $\ell_t^{(i)}(x) = (x - (\boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H))^2$ where $\boldsymbol{x}_t$ is the prediction of the learner at time t. By running d instances of uni-variate FLH-OGD (Fig.4 in Appendix B) with decision set $[-B, B]$ and learning rate $\zeta = 1/(2(2B + G_\infty/H)^2)$ where instance i predicts $\boldsymbol{x}_t[i]$ at time t and suffers losses $\ell_t^{(i)}$, we have*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right),$$

for any comparator sequence $\boldsymbol{w}_{1:n}$ with $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^{n} \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. $\tilde{O}(\cdot)$ hides the dependence on factors of $\log n, B, H, G_\infty$.

When compared with the information theoretic lower bound of (Baby and Wang, 2021) (Proposition 11 there), we see that the rate of Theorem 1 is optimal (modulo log factors) wrt to n, C_n and d. The dependence of $\tilde{O}(d)$ for low C_n regimes is due to the fact that we only assume $\|\nabla f_t(\boldsymbol{x})\|_\infty = O(1)$ as opposed to assuming $\|\nabla f_t(\boldsymbol{x})\|_2 = O(1)$.

Remark 6. ***(relaxed assumptions & improvements)*** *Unlike (Baby and Wang, 2021), we do not assume gradient Lipschitzness of the losses f_t. Further, for the box decision set, our results attain an optimal $O(d^{1/3})$ dimension dependence on regret in the non-trivial regime of $C_n \geq 1/n$ in comparison to the $O(d^2)$ dependence of (Baby and Wang, 2021) for strongly convex losses.*

Remark 7. *We emphasize that the theory developed in Section 2 is vital for extending the results with the surrogate losses as in Corollary 5. Consider squared losses $\ell_t(x) = (x - y_t)^2$ with labels y_t such that $|y_t| \leq Y$ for all t. (Baby and Wang, 2021) requires that the predictions x_t obey $x_t \in [-Y, Y]$. In our use case with surrogate losses $\ell_t^{(i)}(x) = (x - (\boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H))^2$ such a requirement can be not well defined. Here the labels can be regarded $y_t = \boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H$ which depends on $\boldsymbol{x}_t[i]$. As per the setup of Corollary 5, the i^{th} FLH-OGD instance uses losses $\ell_t^{(i)}$, $t \in [n]$ and its prediction at time t is $\boldsymbol{x}_t[i]$. So constructing a uniform bound Y to contain the predictions $\boldsymbol{x}_t[i]$ requires a uniform bound on the predictions $\boldsymbol{x}_t[i]$ itself for all t which is self conflicting. Hence the strategy of (Baby and Wang, 2021) for squared error losses is incompatible for using the surrogate losses $\ell_t^{(i)}$.*

3.2 Strongly convex losses and general convex decision sets

In this section, we show how to convert an optimal algorithm described in Section 3 for the box decision set to an optimal (modulo factors of $\log n$ and dimensions dependencies) algorithm for any convex decision set via a black box reduction. This reduction is essentially due to the seminal work of (Cutkosky and Orabona, 2018).

We have the following guarantee for the scheme in Fig. 3.

Box to general convex set reduction: Inputs - Decision set $\mathcal{W}$, $G > 0$

1. Let $\mathcal{D}$ be the tightest box that circumscribes $\mathcal{W}$. i.e, $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq \sup_{\boldsymbol{w}\in\mathcal{W}}\|\boldsymbol{w}\|_\infty\}$.
2. Let $\mathcal{A}$ be the algorithm attaining the guarantee in Corollary 5 with decision set $\mathcal{D}$ and $G_\infty = 2G$.
3. At round t, get iterate $\boldsymbol{x}_t$ from $\mathcal{A}$.
4. Play $\hat{\boldsymbol{x}}_t = \Pi_{\mathcal{W}}(\boldsymbol{x}_t) := \mathrm{argmin}_{\boldsymbol{y}\in\mathcal{W}}\|\boldsymbol{x}_t - \boldsymbol{y}\|_1$.
5. Get loss f_t.
6. Construct surrogate loss $\ell_t(\boldsymbol{x}) = f_t(\boldsymbol{x}) + G \cdot S(\boldsymbol{x})$, where $S(\boldsymbol{x}) := \|\boldsymbol{x} - \Pi_{\mathcal{W}}(\boldsymbol{x})\|_1$.
7. Send $\ell_t(\boldsymbol{x})$ to $\mathcal{A}$.

Figure 3: Black box reduction from box to arbitrary convex decision set. This technique is due to (Cutkosky and Orabona, 2018).

Theorem 8. *Assume the notations in Fig. 3. Let the input decision set be $\mathcal{W}$. Let the losses be H strongly convex in L_2 norm across $\mathcal{D}$ and satisfy $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G$ for all $\boldsymbol{x} \in \mathcal{D}$. Then the reduction scheme in Fig. 3 guarantees that*

$$\sum_{t=1}^n f_t(\hat{\boldsymbol{x}}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right),$$

for any comparator sequence $\boldsymbol{w}_{1:n} \in \mathcal{W}$ with $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. $\tilde{O}(\cdot)$ hides the dependence on factors of $\log n, H, G_\infty$.

Proof. We start by listing several observations. First, note that the function $S(\boldsymbol{x})$ is convex and 1-Lipschitz across $\mathbb{R}^d$. (Proposition 1 in (Cutkosky and Orabona, 2018)).

Also, the sub-gradient $\partial S(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^d : \boldsymbol{y}[j] = \mathrm{sign}\left(x[j] - \Pi_{\mathcal{W}}(\boldsymbol{x})[j]\right) j \in [d]\}$ (due to Theorem 4 in (Cutkosky and Orabona, 2018)). Here $\mathrm{sign}(a) = a/|a|$ if $|a| > 0$ and any number between $[-1, 1]$ otherwise.

Finally the surrogate losses ℓ_t are H strongly convex in $L2$ norm across $\mathcal{D}$, as adding a convex function to strongly convex function preserves strong convexity. However, ℓ_t are *not* gradient Lipschitz due to the component $G\|\boldsymbol{x} - \Pi_{\mathcal{W}}(\boldsymbol{x})\|_1$ being not smooth.

We have that for any $\boldsymbol{x} \in \mathcal{D}$,

$$\begin{aligned}\|\nabla \ell_t(\boldsymbol{x})\|_\infty &\leq \|\nabla f_t(\boldsymbol{x})\|_\infty + G\|\partial S(\boldsymbol{x})\|_\infty \\ &\leq 2G,\end{aligned}$$

where the last line is due to the assumption that $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G$ and $\partial S(\boldsymbol{x})$ is just a vector of signs as established before.

Hence we have that the losses ℓ_t sent to algorithm $\mathcal{A}$ satisfy the conditions of Corollary 5 with $G_\infty = 2G$. Hence we have that

$$\sum_{t=1}^n \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right), \quad (7)$$

where $\boldsymbol{w}_{1:n}$ is as mentioned in the theorem statement.

By Taylor's theorem, we have that for some $\boldsymbol{z}$ in the line segment joining $\boldsymbol{x}_t$ and $\hat{\boldsymbol{x}}_t$

$$\begin{aligned}f_t(\hat{\boldsymbol{x}}_t) &= f_t(\boldsymbol{x}_t) + \nabla f_t(\boldsymbol{z})^T(\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t) \\ &\leq f_t(\boldsymbol{x}_t) + G\|\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t\|_1 \\ &= \ell_t(\boldsymbol{x}_t)\end{aligned}$$

where the inequality is due to Holder's inequality and the assumption that $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G$ for all $\boldsymbol{x} \in \mathcal{D}$.

Further for any $\boldsymbol{w}_t \in \mathcal{W}$, we have that $f_t(\boldsymbol{w}_t) = \ell_t(\boldsymbol{w}_t)$. Thus overall we obtain,

$$\sum_{t=1}^n f_t(\hat{\boldsymbol{x}}_t) - f_t(\boldsymbol{w}_t) \leq \sum_{t=1}^n \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t). \quad (8)$$

Combining Eq.(7) and (8) now yeilds the theorem. □

Remark 9. *We emphasize that the removal of gradient smoothness assumption for strongly convex losses (from (Baby and Wang, 2021)) as done in the current work was important to apply the reduction scheme of Fig.3 as the losses ℓ_t are not gradient smooth.*

4 PERFORMANCE GUARANTEES FOR EXP-CONCAVE LOSSES

In this section, we control the dynamic regret with exp-concave and gradient smooth losses when the decision set is an L_∞ ball. All unspecified lemma statements and proofs are deferred to Appendix E. We make the following assumptions:

Assumption B1: The loss functions ℓ_t are α exp-concave in the box decision set $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$ .ie, $\ell_t(\boldsymbol{y}) \geq \ell_t(\boldsymbol{x}) + \nabla \ell_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{\alpha}{2}\left(\nabla \ell_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})\right)^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.

Assumption B2: The loss functions ℓ_t satisfy $\|\nabla \ell_t(\boldsymbol{x})\|_2 \leq G$ and $\|\nabla \ell_t(\boldsymbol{x})\|_\infty \leq G_\infty$ for all $\boldsymbol{x} \in \mathcal{D}$. Without loss of generality, we let $G \wedge G_\infty \wedge B \geq 1$, where $a \wedge b := \min\{a, b\}$.

We consider the following protocol:

- At time $t \in [n]$ learner predicts $\boldsymbol{x}_t \in \mathbb{R}^d$ with $\|\boldsymbol{x}_t\|_\infty \leq B$.
- Adversary reveals the loss function ℓ_t.

In view of Assumption B1, following (Hazan et al., 2007), one can define the surrogate losses:

$$f_t(\boldsymbol{x}) = \left(\sqrt{\alpha/2}\nabla\ell_t(\boldsymbol{x}_t)^T(\boldsymbol{x}-\boldsymbol{x}_t) + 1/\sqrt{2\alpha}\right)^2. \quad (9)$$

It follows that

$$\sum_{t=1}^{n} \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t) \leq \sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t),$$

where $\boldsymbol{x}_t, \boldsymbol{w}_t \in \mathcal{D}$.

Further, we make two useful observations about surrogate losses f_t.

First for $\boldsymbol{x} \in \mathcal{D}$, since $\left|\sqrt{\alpha/2}\nabla\ell_t(\boldsymbol{x}_t)^T(\boldsymbol{x}-\boldsymbol{x}_t) + 1/\sqrt{2\alpha}\right| \leq 2GB\sqrt{\alpha d/2} + 1/\sqrt{2\alpha} := \gamma$, we have that f_t are $1/(2\gamma^2)$ exp-concave over $\mathcal{D}$ (see Section 3.3 in (Cesa-Bianchi and Lugosi, 2006)).

Second, since $\nabla^2 f_t(\boldsymbol{x}) = \nabla\ell_t(\boldsymbol{x}_t)\nabla\ell_t(\boldsymbol{x}_t)^T \preccurlyeq G^2\boldsymbol{I}$, we have that the losses f_t are G^2 gradient Lipschitz over $\mathcal{D}$.

We are interested in controlling the regret:

$$R_n(C_n) := \sup_{\substack{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n \in \mathcal{D} \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n}} \sum_{t=1}^{n} \ell_t(\boldsymbol{x}_t) - \ell_t(\boldsymbol{w}_t),$$

where $\boldsymbol{x}_t$ is the decisions of the algorithm.

We have the following performance guarantee when the losses are exp-concave.

Theorem 10. *Suppose Assumptions B1-B2 are satisfied. Define $\gamma := 2GB\sqrt{\alpha d/2} + 1/\sqrt{2\alpha}$. By using the base learner as ONS with parameter $\zeta = \min\left\{\frac{1}{16GB\sqrt{d}}, 1/(4\gamma^2)\right\}$, decision set $\mathcal{D}$, loss at time t to be f_t and choosing learning rate of FLH as $\eta = 1/(2\gamma^2)$, FLH-ONS (Fig.4 in Appendix B) obeys*

$$\begin{aligned} R_n(C_n) &\leq \sup_{\substack{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n \in \mathcal{D} \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n}} \sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) \\ &= \tilde{O}\Bigg(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha) \times \\ &\qquad (n^{1/3}C_n^{2/3} \vee 1)\Bigg)\mathbb{I}\{C_n > 1/n\} \\ &\quad + \tilde{O}\left(d(8G^2B^2\alpha d + 1/\alpha)\,\mathbb{I}\{C_n \leq 1/n\}\right), \end{aligned}$$

where $\boldsymbol{x}_t$ is the decision of the algorithm at time t and $\tilde{O}(\cdot)$ hides polynomial factors of $\log n$. $\mathbb{I}\{\cdot\}$ is the boolean indicator function assuming values in $\{0, 1\}$.

Remark 11. ***(relaxed assumptions & improvements)*** *In (Baby and Wang, 2021), it is assumed that the losses are gradient Lipschitz and exp-concave over an enlarged set $\mathcal{D}^\dagger = \{\boldsymbol{x} : \|\boldsymbol{x}\|_\infty \leq B + G\}$ where B and G are as in Assumptions B1-B2. While our proper learning results doesn't require gradient Lipschitzness and require exp-concavity to hold in the smaller constraint set $\mathcal{D}$ as in Assumption B1. Further (Baby and Wang, 2021) attains a worse dependence of $O(d^{3.5})$ in the non-trivial regime $C_n \geq 1/n$.*

Further, we show in Appendix F that when the decision set is a polytope satisfying certain conditions, we can reparametrize the original problem into the framework of box constrained online learning with exp-concave losses.

4.1 Road map for the proof of Theorem 10

The proof of Theorem 10 is facilitated by generalising the arguments used for proving Theorem 1. We first form a coarse partition of $[n]$ namely $\mathcal{P}$ in Lemma 26 by a direct extension of Lemma 17. For the regime where dual varaible $\lambda = O(d^{1.25}n^{1/3}/C_n^{1/3})$, we employ a two term regret decomposition for each bin $[i_s, i_t] \in \mathcal{P}$ as follows:

$$\underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_i)}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\check{\boldsymbol{u}}_i) - f_j(\boldsymbol{u}_j)}_{T_{2,i}},$$

where $\boldsymbol{x}_j$ is the prediction of the FLH-ONS algorithm and $\boldsymbol{u}_{1:n}$ is the offline optimal sequence in Lemma 25. We exhibit a choice of $\check{\boldsymbol{u}}_i \in \mathcal{D}$ in Lemma 29 so that $T_{1,i} + T_{2,i}$ when summed across all bins $[i_s, i_t] \in \mathcal{P}$ yield a total regret of $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$.

For handling the alternate regime $\lambda = \Omega(d^{1.25}n^{1/3}/C_n^{1/3})$, we provide a refinement scheme `fineSplit` in Fig.8 in Appendix E. Specifically let $\mathcal{R}$ be the set of all intervals in $\mathcal{P}$ that satisfy the prerequisite of `fineSplit` procedure. Let $\mathcal{S} := \mathcal{P} \setminus \mathcal{R}$.

For each interval in $\mathcal{R}$, we invoke `fineSplit`. This refinement scheme splits the original interval into sub-bins that satisfy either the properties in Lemma 36 (which can be regarded as a generalization of style U sub-bins in Section 2.2) or Lemma 37 (which can be regarded as a generalization of style V sub-bins in Section 2.2). Sub-bins that satisfy condition in Lemma 36 is termed as style U$^+$ sub-bins and those that satisfy condition in Lemma 37 is termed as style V$^+$ sub-bins henceforth for brevity. Sub-bins satisfying conditions of both Lemmas 36 and 37 are regarded as style U^+ sub-bins. For each such sub-bin $[a, b]$, we employ a two term regret decomposition as follows:

$$\underbrace{\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_j)}_{T_1} + \underbrace{\sum_{j=a}^{b} f_j(\check{\boldsymbol{u}}_j) - f_j(\boldsymbol{u}_j)}_{T_2}. \quad (10)$$

We term the sequence $\check{\boldsymbol{u}}_{a:b}$ as the *ghost sequence* as they are fictitious intermediate comparator sequence introduced solely for the purpose of analysis. We provide a mechanical way of generating an appropriate ghost sequence in the `generateGhostSequence` procedure in Fig.7 which satisfies the properties stated in Lemma 31. Of particular interest is how we choose the ghost sequence for style U^+ sub-bins. Suppose for a style U^+ sub-bin $[a, b]$, let $k \in [d]$ be the coordinate where the offline optimal takes the form of Structure 1 or Structure 2 (see Definition 33). Then we set for all $j \in [a, b]$:

$$\check{\boldsymbol{u}}_j[k] = \Pi\left(\boldsymbol{u}_a[k] - \frac{1}{(b-a+1)\beta}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k]\right),$$

where $\Pi(\cdot)$ is the projection to $[-B, B]$ and $\beta := G^2$. This choice is very different from the unprojected gradient descent update used in (Baby and Wang, 2021). It can be viewed as a lazy projected gradient descent like update (with step size $1/((b-a+1)\beta)$) where the update operation is performed only across coordinate k. Note that it is not exactly gradient descent across coordinate k since in the second term above we are using $\nabla f_j(\boldsymbol{u}_j)[k]$ instead of $\nabla f_j(\boldsymbol{u}_a)[k]$.

The choice of $\check{\boldsymbol{u}}_j[k']$ for $k' \neq k$ is more involved and is accomplished by carefully selecting a sequence that switches only $O(1)$ times and assumes values in $[(\boldsymbol{u}_a[k'] \wedge \ldots \wedge \boldsymbol{u}_b[k']), (\boldsymbol{u}_a[k'] \vee \ldots \vee \boldsymbol{u}_b[k'])]$ as mentioned in `generateGhostSequence` procedure in Fig.7 in Appendix E.

Next, by using similar gap criteria used in Section 2.2 and exploiting gradient Lipschitzness, we show that $T_1 + T_2$ in Eq.(10) can be bounded by $O^*(\log n) +$ a negative term for each style U^+ sub-bin obtained by refining bins in $\mathcal{R}$. For each style V^+ sub-bin, the regret is bounded by $O^*(\log n)$ (see Lemma 32). When such bounds are added for all sub-bins generated by invoking `fineSplit` on every interval in $\mathcal{R}$, we show that the negative terms gracefully offset the culmination of $O^*(\log n)$ terms to result in a regret bound of $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$ (see Proof of Lemma 40).

The regret contribution from all bins in $\mathcal{S}$ is bounded by $\tilde{O}^*(n^{1/3}C_n^{2/3} \vee 1)$ using Lemma 32. Finally summing the regret contributions from bins in $\mathcal{R}$ and $\mathcal{S}$ yield the theorem.

5 CONCLUSION AND FUTURE WORK

In this work we presented a new analysis that extends the results of (Baby and Wang, 2021) and showed near optimal universal dynamic regret in proper learning settings for strongly convex losses. Results on the special case of exp-concave losses and box decision set are also derived. Further we relaxed the gradient Lipschitzness assumption for losses revealed and derived regret rates with improved dependence on d.

An important open problem is to extend these results for exp-concave losses with general convex decision sets.

Acknowledgments

The research was partially supported by NSF Award #2007117 and a start-up grant from UCSB CS department.

References

Dmitry Adamskiy, Wouter M. Koolen, Alexey Chernov, and Vladimir Vovk. A closer look at adaptive regret. *Journal of Machine Learning Research*, 2016.

Dheeraj Baby and Yu-Xiang Wang. Online forecasting of total-variation-bounded sequences. In *Neural Information Processing Systems (NeurIPS)*, 2019.

Dheeraj Baby and Yu-Xiang Wang. Adaptive online estimation of piecewise polynomial trends. *Neural Information Processing Systems (NeurIPS)*, 2020.

Dheeraj Baby and Yu-Xiang Wang. Optimal dynamic regret in exp-concave online learning. In *COLT*, 2021.

Dheeraj Baby, Hilaf Hasson, and Yuyang Wang. Dynamic regret for strongly adaptive methods and optimality of online krr, 2021a.

Dheeraj Baby, Xuandong Zhao, and Yu-Xiang Wang. An optimal reduction of tv-denoising to adaptive online learning. *AISTATS*, 2021b.

Omar Besbes, Yonatan Gur, and Assaf Zeevi. Nonstationary stochastic optimization. *Operations research*, 63(5):1227–1244, 2015.

Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, New York, NY, USA, 2006. ISBN 0521841089.

Ting-Jui Chang and Shahin Shahrampour. On online optimization: Dynamic regret analysis of strongly convex and smooth problems. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

Xi Chen, Yining Wang, and Yu-Xiang Wang. Nonstationary stochastic optimization under lp, q-variation measures. 2018.

Ashok Cutkosky. Parameter-free, dynamic, and strongly-adaptive online learning. In *Proceedings of the 37th International Conference on Machine Learning*, 2020.

Ashok Cutkosky and Francesco Orabona. Black-box reductions for parameter-free online learning in banach spaces. In *COLT*, 2018.

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

David L Donoho and Iain M Johnstone. Minimax estimation via wavelet shrinkage. *The annals of Statistics*, 26(3):879–921, 1998.

Pierre Gaillard and Sébastien Gerchinovitz. A chaining algorithm for online nonparametric regression. In *Conference on Learning Theory*, pages 764–796, 2015.

Gautam Goel and Adam Wierman. An online algorithm for smoothed regression and lqr control. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, 2019.

Adityanand Guntuboyina, Donovan Lieu, Sabyasachi Chatterjee, and Bodhisattva Sen. Adaptive risk bounds in univariate total variation denoising and trend filtering. 2017.

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

Elad Hazan and Comandur Seshadhri. Adaptive algorithms for online decision problems. In *Electronic colloquium on computational complexity (ECCC)*, volume 14, 2007.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. 2007.

Ali Jadbabaie, Alexander Rakhlin, Shahin Shahrampour, and Karthik Sridharan. Online optimization: Competing with dynamic comparators. In *Artificial Intelligence and Statistics*, pages 398–406, 2015.

Kwang-Sung Jun, Francesco Orabona, Stephen Wright, and Rebecca Willett. Improved Strongly Adaptive Online Learning using Coin Betting. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017.

Seung-Jean Kim, Kwangmoo Koh, Stephen Boyd, and Dimitry Gorinevsky. ℓ_1 trend filtering. *SIAM Review*, 51(2):339–360, 2009.

Wouter M Koolen, Alan Malek, Peter L Bartlett, and Yasin Abbasi. Minimax time series prediction. In *Advances in Neural Information Processing Systems (NIPS'15)*, pages 2557–2565. 2015.

Wojciech Kotłowski, Wouter M. Koolen, and Alan Malek. Online isotonic regression. In *Annual Conference on Learning Theory (COLT-16)*, volume 49, pages 1165–1189. PMLR, 2016.

Enno Mammen. Nonparametric regression under qualitative smoothness assumptions. *Annals of Statistics*, 19(2):741—759, 1991.

Aryan Mokhtari, Shahin Shahrampour, A. Jadbabaie, and Alejandro Ribeiro. Online optimization in dynamic environments: Improved regret rates for strongly convex problems. *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 7195–7201, 2016.

Francesco Ortelli and Sara van de Geer. Prediction bounds for (higher order) total variation regularized least squares. *arXiv preprint arXiv:1904.10871*, 2019.

Alexander Rakhlin and Karthik Sridharan. Online nonparametric regression. In *Conference on Learning Theory*, pages 1232–1264, 2014.

Ryan J Tibshirani. Adaptive piecewise polynomial estimation via trend filtering. *Annals of Statistics*, 42(1):285–323, 2014.

Sara van de Geer. Estimating a regression function. *Annals of Statistics*, 18(2):907—924, 1990.

Yu-Xiang Wang, Alex Smola, and Ryan Tibshirani. The falling factorial basis and its statistical applications. In *International Conference on Machine Learning (ICML-14)*, pages 730–738, 2014.

Yu-Xiang Wang, James Sharpnack, Alex Smola, and Ryan J Tibshirani. Trend filtering on graphs. *Journal of Machine Learning Research*, 17(105):1–41, 2016.

Tianbao Yang, Lijun Zhang, Rong Jin, and Jinfeng Yi. Tracking slowly moving clairvoyant: optimal dynamic regret of online learning with true and noisy gradient. In *International Conference on Machine Learning (ICML-16)*, pages 449–457, 2016.

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

Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems (NeurIPS-18)*, pages 1323–1333, 2018a.

Lijun Zhang, Tianbao Yang, Zhi-Hua Zhou, et al. Dynamic regret of strongly adaptive methods. In *Inter-*

national Conference on Machine Learning (ICML-18), pages 5877–5886, 2018b.

Peng Zhao and Lijun Zhang. Improved analysis for dynamic regret of strongly convex and smooth functions. *L4DC*, 2021.

Peng Zhao, Y. Zhang, L. Zhang, and Zhi-Hua Zhou. Dynamic regret of convex and smooth functions. *NeurIPS*, 2020.

Peng Zhao, Yu-Xiang Wang, and Zhi-Hua Zhou. Non-stationary online learning with memory and non-stochastic control. 2021.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning (ICML-03)*, pages 928–936, 2003.

A Discussion on Related Work

In this section, we compare and contrast our work with several existing lines of research.

Dynamic regret minimization in non-stationary online learning. Apart from (Baby and Wang, 2021), our work fits into the broad literature of dynamic regret minimization in online learning such as (Zinkevich, 2003; Besbes et al., 2015; Jadbabaie et al., 2015; Yang et al., 2016; Mokhtari et al., 2016; Chen et al., 2018; Zhang et al., 2018a,b; Yuan and Lamperski, 2020; Goel and Wierman, 2019; Baby and Wang, 2019; Zhao et al., 2020; Cutkosky, 2020; Baby and Wang, 2020; Zhao and Zhang, 2021; Baby et al., 2021b; Zhao et al., 2021; Chang and Shahrampour, 2021; Baby et al., 2021a). However, to the best of our knowledge none of these works are known to attain the optimal dynamic regret rate for the proper learning setting in terms of path length of the arbitrary comparator sequence.

Adaptive online learning. There is a complementary body of work on Strongly Adaptive regret minimization such as (Daniely et al., 2015; Jun et al., 2017; Cutkosky, 2020) and Adaptive regret minimization such as (Hazan and Seshadhri, 2007; Adamskiy et al., 2016) (which are in fact Strongly Adaptive wrt exp-concave losses) that aims at controlling the static regret in any local time interval. This work focuses on developing new guarantees for algorithms that are Strongly Adaptive (SA) wrt strongly convex / exp-concave losses. The base learners we use for SA methods are the static regret minimizing algorithms from (Hazan et al., 2007).

Locally adaptive non-parametric regression. Our work is closely related to locally adaptive non-parametric regression literature from the statistics community such as (Mammen, 1991; van de Geer, 1990; Donoho and Johnstone, 1998; Kim et al., 2009; Tibshirani, 2014; Wang et al., 2014, 2016; Guntuboyina et al., 2017; Ortelli and van de Geer, 2019). This work supplements them by removing the statistical assumptions and enabling to go beyond squared error losses for the non-parametric function class of TV bounded functions.

Online non-parametric regression. The results of (Rakhlin and Sridharan, 2014) certifies that the minimax rate for competing against a reference class of TV bounded functions with squared error losses is $O(n^{1/3})$. However this bound doesn't capture the correct dependence on C_n and is arrived via non-constructive arguments. On the other hand we arrive at the optimal dependence on both n and C_n via an efficient algorithm. Further, our results with squared error losses in Section 2 are more general than that of (Baby and Wang, 2021) (see Remark 2). Results on online non-parametric regression against reference class of Lipschitz functions, Sobolev functions and isotonic functions can be found in (Gaillard and Gerchinovitz, 2015; Koolen et al., 2015; Kotłowski et al., 2016) respectively. However as noted in (Baby and Wang, 2019), these classes feature functions that are more regular than TV bounded functions. In fact they can be embedded inside a TV bounded function class. So the minimax optimally for TV class implies minimax optimality for the smoother function classes as well.

We refer the reader to (Baby and Wang, 2021) and references therein for a more elaborate survey on existing literature.

B Preliminaries

For the sake of completeness, we recall the description of Follow-the-Leading-History (FLH) algorithm from (Hazan and Seshadhri, 2007).

FLH enjoys the following guarantee against any base learner.

Proposition 12. *(Hazan and Seshadhri, 2007) Suppose the loss functions are exp-concave with parameter α. For any interval $I = [r, s]$ in time, the algorithm FLH Fig.4 with learning rate $\zeta = \alpha$ gives $O(\alpha^{-1}(\log r + \log|I|))$ regret against the base learner in hindsight.*

Definition 13. *((Daniely et al., 2015)) An algorithm is said to be Strongly Adaptive (SA) if for every contiguous interval $I \subseteq [n]$, the static regret incurred by the algorithm is $O(poly(\log n)\Gamma^*(|I|))$ where $\Gamma^*(|I|)$ is the value of minimax static regret incurred in an interval of length $|I|$.*

It is known from (Hazan et al., 2007) that OGD and ONS achieves static regret of $O(\log n)$ and $O(d \log n)$ for strongly convex and exp-concave losses respectively. Hence in view of Proposition 12 and Definition 13, we can conclude that:

- FLH with OGD as base learners is an SA algorithm for strongly convex losses.

FLH: inputs - Learning rate ζ and n base learners $E^1, \ldots, E^n$

1. For each t, $v_t = (v_t^{(1)}, \ldots, v_t^{(t)})$ is a probability vector in $\mathbb{R}^t$. Initialize $v_1^{(1)} = 1$.
2. In round t, set $\forall j \leq t$, $x_t^j \leftarrow E^j(t)$ (the prediction of the j^{th} bas learner at time t). Play $x_t = \sum_{j=1}^t v_t^{(j)} x_t^{(j)}$.
3. After receiving f_t, set $\hat{v}_{t+1}^{(t+1)} = 0$ and perform update for $1 \leq i \leq t$:
$$\hat{v}_{t+1}^{(i)} = \frac{v_t^{(i)} e^{-\zeta f_t(x_t^{(i)})}}{\sum_{j=1}^t v_t^{(j)} e^{-\zeta f_t(x_t^{(j)})}}$$
4. Addition step - Set $v_{t+1}^{(t+1)}$ to $1/(t+1)$ and for $i \neq t+1$:
$$v_{t+1}^{(i)} = (1 - (t+1)^{-1})\hat{v}_{t+1}^{(i)}$$

Figure 4: FLH algorithm

- FLH with ONS as base learners is an SA algorithm for exp-concave losses. (We treat dimension d as a constant problem parameter and consider minimaxity only wrt n.)

We have the following guarantee on runtime.

Proposition 14. *(Hazan and Seshadhri, 2007) Let ρ be the per round run time of base learners and r_n be the static regret suffered by the base learners over n rounds. Then FLH procedure has a runtime of $O(\rho n)$ per round. To improve the runtime one can use AFLH procedure from (Hazan and Seshadhri, 2007) that incurs $O(\rho \log n)$ runtime overhead per round and suffers $O(r_n \log n)$ static regret in any interval.*

Similar runtime improvements at the expense of blowing up the regret by a factor of $\log n$ can also be obtained from the IFLH algorithm of (Zhang et al., 2018b).

C Proofs for Section 2

We start by characterizing the offline optimal. Define the sign function as $\text{sign}(x) = 1$ if $x > 0$; -1 if $x < 0$; and some $v \in [-1, 1]$ if $x = 0$. We start by presenting a sequence of useful lemmas.

Lemma 3. *(**characterization of offline optimal**) Consider the following convex optimization problem (where $\tilde{z}_1, ..., \tilde{z}_{n-1}$ are introduced as dummy variables)*

$$\begin{aligned}
\min_{\tilde{u}_1, \ldots, \tilde{u}_n, \tilde{z}_1, \ldots, \tilde{z}_{n-1}} \quad & \frac{1}{2}\sum_{t=1}^n (y_t - \tilde{u}_t)^2 \\
\text{s.t.} \quad & \tilde{z}_t = \tilde{u}_{t+1} - \tilde{u}_t \ \forall t \in [n-1], \\
& \sum_{t=1}^{n-1} |\tilde{z}_t| \leq C_n, && \text{(1a)} \\
& -B \leq \tilde{u}_t \ \forall t \in [n], && \text{(1b)} \\
& \tilde{u}_t \leq B \ \forall t \in [n], && \text{(1c)}
\end{aligned}$$

Let $u_1, \ldots, u_n, z_1, \ldots, z_{n-1}$ be the optimal primal variables and let $\lambda \geq 0$ be the optimal dual variable corresponding to the constraint (1a). Further, let $\gamma_t^- \geq 0, \gamma_t^+ \geq 0$ be the optimal dual variables that correspond to constraints (1b) and (1c) respectively for all $t \in [n]$. By the KKT conditions, we have

- ***stationarity:*** $u_t - y_t = \lambda (s_t - s_{t-1}) + \gamma_t^- - \gamma_t^+$, *where $s_t \in \partial |z_t|$ (a subgradient). Specifically, $s_t =$*

$\text{sign}(u_{t+1} - u_t)$ if $|u_{t+1} - u_t| > 0$ and s_t is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $s_n = s_0 = 0$.

- ***complementary slackness:*** *(a) $\lambda\left(\sum_{t=2}^{n}|u_t - u_{t-1}| - C_n\right) = 0$; (b) $\gamma_t^-(u_t + B) = 0$ and $\gamma_t^+(u_t - B) = 0$ for all $t \in [n]$*

Proof. We can form the Lagrangian of the optimization problem as:

$$\mathcal{L}(\tilde{u}_{1:n}, \tilde{z}_{1:n-1}, \tilde{\boldsymbol{v}}, \tilde{\lambda}, \tilde{\gamma}^+_{1:n}, \tilde{\gamma}^-_{1:n}) = \frac{1}{2}\sum_{t=1}^{n}(y_t - \tilde{u}_t)^2 + \tilde{\lambda}\left(\sum_{t=1}^{n-1}|\tilde{z}_t| - C_n\right) + \sum_{t=1}^{n-1}\tilde{v}_t(\tilde{u}_{t+1} - \tilde{u}_t - \tilde{z}_t)$$
$$+ \sum_{t=1}^{n}\tilde{\gamma}_t^-(-B - \tilde{u}_t) + \tilde{\gamma}_t^+(\tilde{u}_t - B),$$

for dual variables $\tilde{\lambda} > 0$, $\tilde{v}_{1:n}$ unconstrained, $\tilde{\gamma}^-_{1:n} \geq 0$ and $\tilde{\gamma}^+_{1:n} \geq 0$. Let $(u_{1:n}, z_{1:n}, v_{1:n}, \lambda, \gamma^-_{1:n}, \gamma^+_{1:n})$ be the optimal primal and dual variables. By stationarity conditions (via the derivative wrt u_t), we have:

$$u_t - y_t + v_{t-1} - v_t - \gamma_t^- + \gamma_t^+ = 0,$$

where we take $v_0 = v_n = 0$. Stationarity conditions via derivative wrt z_t yields

$$v_t = \lambda s_t.$$

Combining the above two equations and the complementary slackness rules yields the lemma. □

Example 15. *We describe the example used to create Fig.1. We adopt the notations of Lemma 3.*

- *$G = 4$ and $B = 2$.*
- *For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $u_j = B - \frac{1}{2n^{3/4}}$ for all $j \in [2kn^{3/4} + 1, (2k+1)n^{3/4}]$.*
- *For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $u_j = B$ for all $j \in [(2k+1)n^{3/4} + 1, (2k+2)n^{3/4}]$.*
- *$y_1 = y_{n^{3/4}} = B - \frac{1}{2n^{3/4}} - \frac{n^{3/4}-2}{n}$. $y_j = B - \frac{1}{2n^{3/4}} - (1 - 2/n)$ for all $j \in [2, n^{3/4} - 1]$.*
- *For each $k \in [1, \frac{n^{1/4}}{2} - 1]$, $y_{2kn^{3/4}+1} = y_{(2k+1)n^{3/4}} = B - \frac{1}{2n^{3/4}} - \frac{n^{3/4}-2}{n}$. $y_j = B - \frac{1}{2n^{3/4}} - (1 - 2/n)$ for all $j \in [2kn^{3/4} + 2, (2k+1)n^{3/4} - 1]$.*
- *For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $y_j = G$ for all $j \in [(2k+1)n^{3/4} + 1, (2k+2)n^{3/4}]$.*
- *$\gamma_j^- = 0$ for all $j \in [n]$.*
- *For each $k \in [0, \frac{n^{1/4}}{2} - 1]$, $\gamma_j^+ = 0$ for all $j \in [2kn^{3/4} + 1, (2k+1)n^{3/4}]$.*
- *For each $k \in [0, \frac{n^{1/4}}{2} - 2]$, $\gamma^+_{(2k+1)n^{3/4}+1} = \gamma^+_{(2k+2)n^{3/4}} = G - B - \frac{n^{3/4}-2}{n}$. $\gamma_j^+ = G - B - 2(1 - 1/n)$ for all $j \in [(2k+1)n^{3/4} + 2, (2k+2)n^{3/4} - 1]$.*
- *$\gamma^+_{n-n^{3/4}+1} = \gamma_n^+ = G - B - \frac{n^{3/4}-2}{n}$.*
- *$\lambda = n^{3/4} - 2$.*
- *$s_t = 1/n + (t-1)\frac{1-2/n}{n^{3/4}-2}$ for $1 \leq t \leq n^{3/4} - 1$. $s_{n^{3/4}} = 1$.*
- *For each $k \in [0, \frac{n^{1/4}}{2} - 2]$, $s_t = 1 - 1/n + (t - 1 - (2k+1)n^{3/4})\frac{2/n-2}{n^{3/4}-2}$ for $(2k+1)n^{3/4} + 1 \leq t \leq (2k+2)n^{3/4} - 1$. $s_{(2k+2)n^{3/4}} = -1$.*
- *For each $k \in [1, \frac{n^{1/4}}{2} - 1]$, $s_t = -1 + 1/n + (t - 1 - 2kn^{3/4})\frac{2-2/n}{n^{3/4}-2}$. $s_{(2k+1)n^{3/4}} = 1$.*

- $s_t = 1 - 1/n + (t - 1 - n + n^{3/4})\frac{2/n-1}{n^{3/4}-1}$ *for* $n - n^{3/4} + 1 \leq t \leq n - 1$. $s_n = 0$.

Terminology. We will refer to the optimal primal variables $u_1, \ldots, u_n$ in Lemma 3 as the *offline optimal solution* in this section. For two natural numbers $a < b$, we denote $[a, b] = \{a, a + 1, \ldots, b\}$.

Definition 4.

- *For a bin* $[a, b] \subseteq \{2, \ldots, n-1\}$*, the offline optimal solution is said to assume Structure 1 if* $u_j = u_a \in (-B, B)$ *for all* $j \in [a, b]$ *and* $u_b > u_{b+1}$ *and* $u_a > u_{a-1}$.
- *For a bin* $[a, b] \subseteq \{2, \ldots, n-1\}$*, the offline optimal solution is said to assume Structure 2 if* $u_j = u_a \in (-B, B)$ *for all* $j \in [a, b]$ *and* $u_b < u_{b+1}$ *and* $u_a < u_{a-1}$.
- *For a bin* $[a, b]$*, we define* $gap_{min}(\beta, [a, b]) := \min_{j \in [a,b]} |u_j - \beta|$ *where* $\beta \in \mathbb{R}$.

The following Lemma plays a central role in the analysis. Qualitatively, it captures a fundamental way in which the adversary is constrained.

Lemma 16. ***($\boldsymbol{\lambda}$-length lemma)*** *Suppose that the offline optimal solution sequence takes the form of Structure 1 or Structure 2 in an interval* $[j, j + \ell - 1]$ *for some* $\ell > 0$ *and* $j \in \{2, \ldots, n - 1\}$*. Then* $\lambda \leq \frac{(B+G)\ell}{2}$.

Proof. We consider the case of Structure 2. Arguments are similar for case of Structure 1. Let the optimal sign assignments be written as $s_{j+k-1} = -1 + \epsilon_k$ where $\epsilon_k \in [0, 2]$ for all $k \in [\ell - 1]$. From the KKT conditions, we have

$$\begin{aligned}
y_j &= u - \lambda\epsilon_1 \\
y_{j+1} &= u - \lambda(\epsilon_2 - \epsilon_1) \\
&\vdots \\
y_{j+\ell-2} &= u - \lambda(\epsilon_{\ell-1} - \epsilon_{\ell-2}) \\
y_{j+\ell-1} &= u - \lambda(2 - \epsilon_{\ell-1})
\end{aligned}$$

Consider a vector $\boldsymbol{z} = [\epsilon_1, \epsilon_2 - \epsilon_1, \ldots, 2 - \epsilon_{\ell-1}]^T$. Note that the condition $\|\boldsymbol{z}\|_\infty > 0$ is always satisfied. Otherwise we must have $2 = \epsilon_{\ell-1} = \ldots = \epsilon_1$. But $\epsilon_1 = 2$ makes $\|\boldsymbol{z}\|_\infty > 0$ yielding a contradiction.

Let k^* be such that $|\boldsymbol{z}[k^*]| = \|\boldsymbol{z}\|_\infty$. Since $\lambda \geq 0$, we can write $\lambda = \frac{|y_{j+k^*-1} - u|}{\|\boldsymbol{z}\|_\infty}$. Since $|y_{j+k^*-1} - u|$ is bounded, a lower bound on $\|\boldsymbol{z}\|_\infty$ will yield an upper bound on λ. To this end, we consider the following optimization problem:

$$\begin{aligned}
\min_{t, \epsilon_1, \ldots, \epsilon_{\ell-1}} \quad & t \\
\text{s.t.} \quad & 0 \leq \epsilon_i \leq 2 \; \forall i \in [\ell - 1], \\
& \epsilon_1 \leq t, \\
& |\epsilon_{i+1} - \epsilon_i| \leq t \; \forall i \in [\ell - 2], \\
& 2 - \epsilon_{\ell-1} \leq t
\end{aligned}$$

We can form the Lagrangian as:

$$\begin{aligned}
\mathcal{L}(t, \epsilon_{1:\ell-1}, a_{1:\ell-1}, b_{1:\ell-1}, c_{1:\ell-2}, d_{1:\ell-2}, e_1, e_{\ell-1}) = t &- \sum_{i=1}^{\ell-1} a_i\epsilon_i + \sum_{i=1}^{\ell-1} b_i(\epsilon_i - 2) \\
&+ \sum_{i=1}^{\ell-2} c_i(-t - \epsilon_{i+1} + \epsilon_i) + \sum_{i=1}^{\ell-2} d_i(\epsilon_{i+1} - \epsilon_i - t) \\
&+ e_1(\epsilon_1 - t) + e_2(2 - \epsilon_{\ell-1} - t)
\end{aligned}$$

Stationarity conditions are:

$$\frac{\partial \mathcal{L}}{\partial t} = 0 \implies 1 + \sum_{i=1}^{\ell-2} -c_i - d_i - e_1 - e_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \epsilon_1} = 0 \implies -a_1 + b_1 + c_1 - d_1 + e_1 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \epsilon_{\ell-1}} = 0 \implies -a_{\ell-1} + b_{\ell-1} - c_{\ell-2} + d_{\ell-2} - e_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial \epsilon_i} = 0 \implies -a_i + b_i - c_{i-1} + c_i + d_{i-1} - d_i = 0, \text{ where } i \in \{2, \ldots, \ell-2\}$$

Complementary slackness conditions are:

$$\begin{aligned} a_i\epsilon_i &= 0,\ i \in [\ell-1] \\ b_i(\epsilon_i - 2) &= 0,\ i \in [\ell-1] \\ c_i(-t - \epsilon_{i+1} + \epsilon_i) &= 0,\ i \in [\ell-2] \\ d_i(\epsilon_{i+1} - \epsilon_i - t) &= 0,\ i \in [\ell-2] \\ e_1(\epsilon_1 - t) &= 0 \\ e_2(2 - \epsilon_{\ell-1} - t) &= 0 \end{aligned}$$

Dual feasibility conditions are $a_i \geq 0,\ b_i \geq 0$ for $i \in [\ell-1]$ and $c_i \geq 0,\ d_i \geq 0$ for $i \in [\ell-2]$ and $e_1 \geq 0,\ e_2 \geq 0$.

Primal feasibility conditions are given by the constraint set of the optimization problem.

Now we form a guess for optimal primal and dual variables as $t = 2/\ell$ and $\epsilon_i = 2i/\ell$ for $i \in [\ell-1]$ and $a_i = b_i = 0$ for $i \in [\ell-1]$ and $c_i = 0$ for $i \in [\ell-2]$ and $e_1 = e_2 = d_1 = \ldots = d_{\ell-2} = 1/\ell$. All the KKT conditions can be readily verified for this solution guess.

Recall that, earlier we defined $\boldsymbol{z} = [\epsilon_1, \epsilon_2 - \epsilon_1, \ldots, 2 - \epsilon_{\ell-1}]^T$ and $\lambda = \frac{|y_{j+k^*-1} - u|}{\|\boldsymbol{z}\|_\infty}$ where k^* is such that $|\boldsymbol{z}[k^*]| = \|\boldsymbol{z}\|_\infty$. By the previous optimization problem we deduce that $\|\boldsymbol{z}\|_\infty \geq 2/\ell$. Since $|y_{j+k^*-1} - u| \leq B + G$, we conclude that $\lambda \leq (B+G)\ell/2$

□

Next, we exhibit a useful partitioning scheme of the interval $[n]$.

Lemma 17. *((Baby and Wang, 2021))(**key partition**) Initialize $\mathcal{P} \leftarrow \Phi$. Starting from time 1, spawn a new bin $[i_s, i_t]$ whenever $\sum_{j=i_s+1}^{i_t+1} |u_j - u_{j-1}| > B/\sqrt{n_i}$, where $n_i = i_t - i_s + 2$. Add the spawned bin $[i_s, i_t]$ to $\mathcal{P}$.*

Let $M := |\mathcal{P}|$. We have $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$.

Notations. For bin $[i_s, i_t] \in \mathcal{P}$ we define: $n_i = i_t - i_s + 1$, $\bar{u}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} u_j$, $\bar{y}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} y_j$, $\Gamma_i^+ = \sum_{j=i_s}^{i_t} \gamma_j^+$, $\Gamma_i^- = \sum_{j=i_s}^{i_t} \gamma_j^-$, $\Delta s_i = s_{i_t} - s_{i_s - 1}$, $C_i = \sum_{j=i_s+1}^{i_t} |u_j - u_{j-1}|$.

For any general bin $[a, b]$ define the quantities $n_{a\to b}, \bar{u}_{a\to b}, \bar{y}_{a\to b}, \Gamma_{a\to b}^+, \Gamma_{a\to b}^-, \Delta s_{a\to b}, C_{a\to b}$ analogously as above.

Next we calculate the static regret guarantee of the FLH-ONS strategy.

Lemma 18. *((Hazan et al., 2007), (Hazan and Seshadhri, 2007)) Consider a bin $[a, b] \subseteq [n]$ and a point $w \in [-B, B]$. Under the setting of Theorem 1 we have*

$$\begin{aligned} \sum_{t=a}^{b} (y_t - x_t)^2 - (y_t - w)^2 &\leq 10(B+G)^2 \log n \\ &= \tilde{O}(1), \end{aligned}$$

where x_t are the predictions of FLH-OGD.

Proof. The losses $(y_t - x)^2$ are strongly convex with parameter 2. Further the gradients are bounded by $2(G + B)$. Hence by Theorem 1 in (Hazan et al., 2007) we have the static regret guarantee of OGD being $4(G+B)^2 \cdot (2\log n)/4 = 2(G+B)^2 \log n$.

The losses $(y_t - x)^2$ are $1/(2(G+B)^2)$ exp-concave. So by applying Theorem 3.2 in (Hazan and Seshadhri, 2007) we have the regret of FLH against any base experts bounded as $8(G+B)^2 \log n$.

Adding these regret bounds yields the lemma.

□

Lemma 19. ***(low λ regime)*** *If the optimal dual variable $\lambda = O\left(\frac{n^{1/3}}{C_n^{1/3}}\right)$, we have the regret of FLH-OGD strategy bounded as*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - u_t)^2 = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1),$$

where x_t is the prediction of FLH-OGD at time t.

Proof. Throughout this proof, the bins $[i_s, i_t]$ we consider belong to the partition $\mathcal{P}$.

Case 1: When the offline optimal solution touches the boundary B within a bin $[i_s, i_t]$. We use a three term regret decomposition as follows.

$$\underbrace{\sum_{j=i_s}^{i_t} (y_j - x_j)^2 - (y_j - B)^2}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} (y_j - B)^2 - (y_j - \bar{u}_i)^2}_{T_{2,i}} + \underbrace{\sum_{j=i_s}^{i_t} (y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{3,i}}$$

Now $T_{1,i} = O(\log n)$ by strong adaptivity of FLH. Observe that due to complementary slackness, $\gamma_j^- = 0$ uniformly within the bin since the TV within the bin is at-most $B/\sqrt{n_i} < 2B$ and hence the solution never touches $-B$ boundary within this bin. By using the KKT conditions, we have $y_j = u_j - \lambda(s_j - s_{j-1}) + \gamma_j^+$. So

$$\begin{aligned} T_{2,i} &= \sum_{j=i_s}^{i_t} (\bar{u}_i - B)^2 + 2(y_j - \bar{u}_i)(\bar{u}_i - B) \\ &= n_i(\bar{u}_i - B)^2 + 2n_i(\bar{y}_i - \bar{u}_i)(\bar{u}_i - B) \\ &\leq_{(a)} B^2 + 2(\bar{u}_i - B)(\Gamma_i^+ - \lambda \Delta s_i) \\ &\leq_{(b)} B^2 + 4\lambda C_i + 2\Gamma_i^+(\bar{u}_i - B) \end{aligned} \tag{12}$$

where in line (a) we used KKT conditions and $|\bar{u}_i - B| \leq B/\sqrt{n_i}$ due to the TV constraint within bin and in line (b) we used: (i) $|\bar{u}_i - B| \leq C_i$ as the optimal solution assumes the value B at some time point in $[i_s, i_t]$ (ii) $|\Delta s_i| \leq 2$.

We have

$$\begin{aligned} T_{3,i} &= \sum_{j=i_s}^{i_t} (u_j - \bar{u}_i)^2 + 2(y_j - u_j)(u_j - \bar{u}_i) \\ &\leq n_i C_i^2 + 2\sum_{j=i_s}^{i_t} (-\lambda(s_j - s_{j-1}) + \gamma_j^+)(u_j - \bar{u}_i) \\ &=_{(a)} n_i C_i^2 + 2\lambda(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)) + 2\lambda C_i + 2\sum_{j=i_s}^{i_t} \gamma_j^+(u_j - \bar{u}_i) \\ &=_{(b)} n_i C_i^2 + 2\lambda(s_{i_s-1}(u_{i_s} - \bar{u}_i) - s_{i_t}(u_{i_t} - \bar{u}_i)) + 2\lambda C_i + 2\Gamma_i^+(B - \bar{u}_i) \\ &\leq_{(c)} B^2 + 6\lambda C_i + 2\Gamma_i^+(B - \bar{u}_i), \end{aligned} \tag{13}$$

where line (a) is obtained by a rearrangement of the sum and line (b) is obtained by the complementary slackness condition which states that $\gamma_j^+ = 0$ if $u_j < B$. Line (c) is obtained by $|u_j - \bar{u}_i| \leq C_i$ for any $j \in [i_s, i_t]$ and by applying triangle inequality.

So overall we can bound the regret within this bin by adding Eq.(12) and (13) with $T_{1,i} = O(\log n)$ as

$$T_{1,i} + T_{2,i} + T_{3,i} \leq O(\log n) + 2B^2 + 10\lambda C_i. \tag{14}$$

Case 2: When the offline optimal solution touches boundary $-B$ within a bin $[i_s, i_t]$. This case can be treated similar to Case 1.

Case 3: When the offline optimal solution doesn't touch either boundaries within a bin $[i_s, i_t]$. Here we use a two term regret decomposition as

$$\underbrace{\sum_{j=i_s}^{i_t} (y_j - x_j)^2 - (y_j - \bar{u}_i)^2}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} (y_j - \bar{u}_i)^2 - (y_j - u_j)^2}_{T_{2,i}}.$$

By following the analysis used in obtaining the bound of Eq.(13) (where we use $\gamma_j^- = \gamma_j^+ = 0$ due to complementary slackness), we obtain

$$T_{1,i} + T_{2,i} \leq O(\log n) + B^2 + 6\lambda C_i \tag{15}$$

By summing up the regret bounds which assumes the form in Eq.(14) (for Case 1 and 2) or Eq.(15) (for Case 3) across all bins in the partition $\mathcal{P}$, we obtain the overall regret as

$$\begin{aligned}\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - u_t)^2 &\leq O(|\mathcal{P}| \log n B^2) + 2B^2 |\mathcal{P}| + 10\lambda C_n \\ &= \tilde{O}(n^{1/3} C_n^{2/3} \vee 1),\end{aligned}$$

where in the last line we used the fact that $|\mathcal{P}| = O(n^{1/3} C_n^{2/3} \vee 1)$ and $\lambda = O((n/C_n)^{1/3})$ by the premise of the lemma. ☐

Lemma 20. ***(monotonic sequence)*** *Consider a bin $[i_s, i_t] \in \mathcal{P}$ such that the offline optimal solution is monotonic within this bin. Then the regret of FLH-OGD strategy within this bin is at-most $31(B+G)^2 \log n = O(\log n)$.*

Proof. When the optimal sequence is monotonic within a bin $[i_s, i_t] \in \mathcal{P}$, it is always possible to form *at-most* 3 bins: $[i_s, r_1]$, $[r_1 + 1, r_2]$, $[r_2 + 1, i_t]$ such that the offline optimal solution is constant within bins $[i_s, r_1]$ and $[r_2 + 1, i_t]$ alongside the condition that the bin $[r_1 + 1, r_2]$ satisfies one of the following properties: a) $s_{r_1} = s_{r_2} = 1$ and the offline optimal solution is non-decreasing within bin $[r_1 + 1, r_2]$ or b) $s_{r_1} = s_{r_2} = -1$ and the offline optimal solution is non-increasing within bin $[r_1 + 1, r_2]$. (see for eg. Fig.5).

Due to Lemma 18, the regret within bins $[i_s, r_1]$ and $[r_2 + 1, i_t]$ is at-most $10(B+G)^2 \log n$ each. Note that this three sub-bin refinement can make sure that the offline optimal solution doesn't touch the boundaries $\pm B$ within the bin $[r_1 + 1, r_2]$. We bound the regret within bin $[r_1 + 1, r_2]$ via a two term regret decomposition as follows.

$$\underbrace{\sum_{j=r1+1}^{r_2} (y_j - x_j)^2 - (y_j - \bar{u}_{r1+1\to r_2})^2}_{T_1} + \underbrace{\sum_{j=r1+1}^{r_2} (y_j - \bar{u}_{r1+1\to r_2})^2 - (y_j - u_j)^2}_{T_2}.$$

We have $T_1 \leq 10(B+G)^2 \log n$. Further due to KKT conditions we have,

$$
\begin{aligned}
T_2 &= \sum_{j=r1+1}^{r_2} (u_j - \bar{u}_{r1+1\to r_2})(2y_j - u_j - \bar{u}_{r1+1\to r_2})\\
&= \sum_{j=r1+1}^{r_2} (u_j - \bar{u}_{r1+1\to r_2})(2y_j - 2u_j + u_j - \bar{u}_{r1+1\to r_2})\\
&= \sum_{j=r1+1}^{r_2} (u_j - \bar{u}_{r1+1\to r_2})^2 + 2\lambda(u_j - \bar{u}_{r1+1\to r_2})(s_{j-1} - s_j)\\
&\le n_i C_i^2 + \sum_{j=r1+1}^{r_2} 2\lambda(u_j - \bar{u}_{r1+1\to r_2})(s_{j-1} - s_j),
\end{aligned}
$$

where in the last line we used $|u_j - \bar{u}_{r1+1\to r_2}| \le C_i$. We also have $n_i C_i^2 \le B^2$ by the construction in Lemma 17. By expanding the second term followed by a regrouping of terms in the summation, we can write

$$
\begin{aligned}
\sum_{j=r1+1}^{r_2} 2\lambda(u_j - \bar{u}_{r1+1\to r_2})(s_{j-1} - s_j) &= 2\lambda\left(s_{r_1}(u_{r_1+1} - \bar{u}_{r1+1\to r_2}) - s_{r_2}(u_{r_2} - \bar{u}_{r1+1\to r_2})\right)\\
&\quad + 2\lambda \sum_{j=r_1+2}^{r_2} |u_j - u_{j-1}|\\
&= 2\lambda C_{r_1+1\to r_2}\\
&\quad + 2\lambda\left(s_{r_1}(u_{r_1+1} - \bar{u}_{r1+1\to r_2}) - s_{r_2}(u_{r_2} - \bar{u}_{r1+1\to r_2})\right). \qquad (16)
\end{aligned}
$$

Since $s_{r_1} = s_{r_2} = 1$ if the offline optimal is non-decreasing in $[r_1+1, r_2]$ or $s_{r_1} = s_{r_2} = -1$ if the offline optimal is non-increasing in $[r_1+1, r_2]$, we have $s_{r_1}u_{r_1+1} - s_{r_2}u_{r_2} = -|u_{r_1+1} - u_{r_2}| = -C_{r_1+1\to r_2}$. Hence we see that the second term exactly cancels with the first term in Eq.(16).

Thus overall we have shown that the total regret in $[i_s, i_t]$ is at-most $31(B+G)^2\log n$. ∎

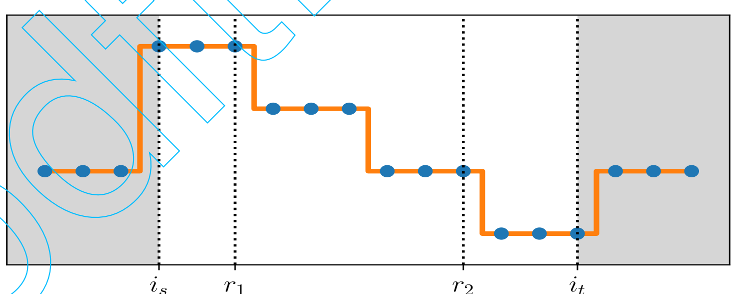

Figure 5: *An example of a configuration referred in the proof of Lemma 20. Here $s_{r_1} = s_{r_2} = -1$ and the sequence is non-increasing within $[r_1+1, r_2]$.*

Lemma 21. *Suppose there exists an interval $[a, b]$ (which may not belong to $\mathcal{P}$) with length ℓ such that the optimal sequence takes the form of Structure 1 or Structure 2 within $[a, b]$. Assume that $\bar{y}_{a\to b} \in [-B, B]$. Then the regret of FLH-OGD within the bin $[a, b]$ at-most $10(B+G)^2\log n - \frac{4\lambda^2}{\ell}$.*

Proof. We use a two term regret decomposition as follows:

$$
\underbrace{\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - \bar{y}_{a\to b})^2}_{T_1} + \underbrace{\sum_{j=a}^{b}(y_j - \bar{y}_{a\to b})^2 - (y_j - u_j)^2}_{T_2}.
$$

By the Definition 4 of Structure 1 and 2, the offline optimal solution is constant within bin $[a, b]$. We denote $u_j = u$ for all $j \in [a, b]$. Further $|\Delta s_{a\to b} = 2|$. We have,

$$\begin{aligned} T_2 &= -\ell(\bar{y}_{a\to b} - u)^2 - 2\sum_{j=a}^{b}(y_j - \bar{y}_{a\to b})(\bar{y}_{a\to b} - u) \\ &= -\ell(\bar{y}_{a\to b} - u)^2 \\ &=_{(a)} -\frac{-\lambda^2(\Delta s_{a\to b})^2}{\ell} \\ &= -\frac{4\lambda^2}{\ell}, \end{aligned}$$

where line (a) is obtained by the KKT conditions $y_j = u - \lambda(s_j - s_{j-1})$ for all $j \in [a, b]$ and hence $\bar{y}_{a\to b} = u - \frac{\lambda \Delta s_{a\to b}}{\ell}$. Due to Lemma 18, we have $T_1 \le 10(B+G)^2 \log n$. Combining both bounds yields the lemma. □

Lemma 22. *Consider a bin $[a, b]$ with length ℓ.*

Case 1: *When offline optimal takes the form of Structure 1 within this bin and $\bar{y}_{a\to b} \ge B$, then*

$$\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - u_j)^2 \le 10(B+G)^2 \log n - \ell(B - u_a)^2,$$

and

Case 2: *When offline optimal takes the form of Structure 2 within this bin and $\bar{y}_{a\to b} \le -B$, then*

$$\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j - u_j)^2 \le 10(B+G)^2 \log n - \ell(B + u_a)^2,$$

where x_j are the predictions of the FLH-OGD algorithm.

Proof. We consider Case 2. Arguments for Case 1 are similar. We employ a two term regret decomposition as follows.

$$\underbrace{\sum_{j=a}^{b}(y_j - x_j)^2 - (y_j + B)^2}_{T_1} + \underbrace{\sum_{j=a}^{b}(y_j + B)^2 - (y_j - u_j)^2}_{T_2}.$$

By Definition 4, the offline optimal solution is constant within bin $[a, b]$. So we have $u_j = u_a$ for all $j \in [a, b]$. From the KKT conditions, we have

$$\begin{aligned} T_2 &= \sum_{j=a}^{b}(u_a + B)^2 + 2(y_j - u_a)(u_a + B) \\ &= \ell(u_a + B)^2 - 2\lambda\Delta s_{a\to b}(u_a + B) \\ &= \ell(u_a + B)^2 - 4\lambda(u_a + B), \end{aligned} \tag{17}$$

where in the last line we used $\Delta s_{a\to b} = 2$ for Structure 2. From the premise of the lemma for Case 2, we have $\bar{y}_{a\to b} \le -B$. Since $\bar{y}_{a\to b} = u_a - 2\lambda/\ell$, we must have

$$\bar{y}_{a\to b} \le -B \implies \lambda \ge \frac{\ell}{2}(u_a + B).$$

Plugging this lower bound to Eq.(17) and noting that $u_a + B \geq 0$, we get

$$T_2 \leq -\ell(u_a + B)^2.$$

By Lemma 18, we have $T_1 \leq 10(B+G)^2 \log n$. Now summing T_1 and T_2 results in the lemma.

$\square$

Lemma 23. ***(large margin bins)*** *Assume that $\lambda \geq \phi \frac{n^{1/3}}{C_n^{1/3}}$ for some constant ϕ that do not depend on n and C_n. Consider a bin $[i_s, i_t] \in \mathcal{P}$ within which the offline optimal solution takes the form of Structure 1 or Structure 2 (or both) for some appropriate sub-intervals of $[i_s, i_t]$. Let $\mu_{th} = \sqrt{\frac{36(B+G)^3 C_n^{1/3} \log n}{\phi n^{1/3}}}$. Then $gap_{min}(-B, [i_s, i_t]) \vee gap_{min}(B, [i_s, i_t]) \geq \mu_{th}$ whenever $C_n \leq \left(\frac{B^2 \phi}{144(B+G)^3 \log n}\right)^3 n = \tilde{O}(n)$.*

Proof. Suppose $\text{gap}_{\min}(-B, [i_s, i_t]) < \mu_{\text{th}}$. Then the largest value of offline optimal attained within this bin $[i_s, i_t]$ is at-most $-B + \mu_{\text{th}} + B/\sqrt{n_i}$ (recall $n_i := i_t - i_s + 1$ and TV within this bin is at-most $B/\sqrt{n_i}$ by Lemma 17). So $\text{gap}_{\min}(B, [i_s, i_t]) \geq 2B - \mu_{\text{th}} - B/\sqrt{n_i}$. Our goal is to show that whenever C_n obeys the constraint stated in the lemma, we must have

$$2B - \mu_{\text{th}} - B/\sqrt{n_i} \geq \mu_{\text{th}}. \tag{18}$$

Let ℓ_i be the length of a sub-interval of $[i_s, i_t]$ where the offline optimal solution assumes the form of Structure 1 or Structure 2. Due to Lemma 16, we have

$$n_i \geq \ell_i \geq \frac{2\lambda}{(G+B)} \geq \frac{2\phi}{(G+B)} \frac{n^{1/3}}{C_n^{1/3}}, \tag{19}$$

where the last inequality follows due to the condition on λ assumed in the current lemma. So a sufficient condition for Eq.(18) to be true is

$$2B \geq 2\left(2\sqrt{\frac{36(G+B)^3 C_n^{1/3} \log n}{\phi n^{1/3}}} \vee B\sqrt{\frac{(G+B) C_n^{1/3}}{2\phi n^{1/3}}}\right).$$

Recall that by Assumption A1 in Section 2, we have $G \geq B \geq 1$ WLOG. So the above maximum will be attained by the first term and can be further simplified as

$$2B \geq 4\sqrt{\frac{36(G+B)^3 C_n^{1/3} \log n}{\phi n^{1/3}}}.$$

The above condition is always satisfied whenever $C_n \leq \left(\frac{B^2 \phi}{144(B+G)^3 \log n}\right)^3 n$.

At this point, we have shown that $\text{gap}_{\min}(-B, [i_s, i_t]) < \mu_{\text{th}} \implies \text{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\text{th}}$ under the conditions of the lemma. Taking the contrapositive yields $\text{gap}_{\min}(B, [i_s, i_t]) < \mu_{\text{th}} \implies \text{gap}_{\min}(-B, [i_s, i_t]) \geq \mu_{\text{th}}$. $\square$

Lemma 24. ***(high $\boldsymbol{\lambda}$ regime)*** *If the optimal dual variable $\lambda \geq \phi \frac{n^{1/3}}{C_n^{1/3}} = \Omega\left(\frac{n^{1/3}}{C_n^{1/3}}\right)$ for some constant $\phi > 0$ that doesn't depend on n and C_n, we have the regret of FLH-OGD strategy bounded as*

$$\sum_{t=1}^{n} (y_t - x_t)^2 - (y_t - u_t)^2 = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1),$$

where x_t is the prediction of FLH-OGD at time t.

Proof. Throughout the proof, we consider only the regime where $C_n \leq \left(\frac{B^2\phi}{84(B+G)^3 \log n}\right)^3 n = \tilde{O}(n)$. In the alternate regime where $C_n = \tilde{\Omega}(n)$, the trivial regret bound of $\tilde{O}(n)$ is near minimax optimal.

Reminiscent to the road-map in Section 2.2, it is useful to define the following condition:

Condition (A): Let a bin $[a, b]$ be given such that $C_{a \to b} \leq B/\sqrt{b-a+1}$. It satisfies at-least one of the following criteria. (i) $\text{gap}_{\min}(B, [a, b]) \geq \text{gap}_{\min}(-B, [a, b])$ and the optimal solution takes the form of Structure 1 in at-least one sub-interval $[r, s] \subseteq [a, b]$; or (ii) $\text{gap}_{\min}(-B, [a, b]) \geq \text{gap}_{\min}(B, [a, b])$ and the optimal solution takes the form of Structure 2 in at-least one sub-interval $[r, s] \subseteq [a, b]$.

Consider a bin $[i_s, i_t] \in \mathcal{P}$ that satisfies Condition (A). We refine $[i_s, i_t]$ into a partition that contains smaller sub-intervals as follows:

$$\mathcal{P}_i := \{[i_s, \underline{i}_1 - 1], [\underline{i}_1, \bar{i}_1], [\underline{i}'_1, \bar{i}'_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}], [\underline{i}'_{m^{(i)}}, \bar{i}'_{m^{(i)}} := i_t]\}, \tag{20}$$

such that:

1. If $\text{gap}_{\min}(B, [i_s, i_t]) > \text{gap}_{\min}(-B, [i_s, i_t])$, then the offline optimal in the intervals $[\underline{i}_j, \bar{i}_j]$, $j \in [m^{(i)}]$ takes the form of Structure 1. Further, let k be the largest value in $[i_s, i_t]$ such that $u_{i_s:k}$ is constant. If $u_{i_s} > u_{i_s - 1}$ and $u_k > u_{k+1}$, then we treat the first sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}_1 = i_s$. Similarly let k be smallest value in $[i_s, i_t]$ such that $u_{k:i_t}$ is constant. If $u_{k-1} < u_k$ and $u_{i_t} > u_{i_t+1}$ then we treat the last sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}'_{m^{(i)}} = i_t + 1$.

2. If $\text{gap}_{\min}(B, [i_s, i_t]) \leq \text{gap}_{\min}(-B, [i_s, i_t)$, then the offline optimal in the intervals $[\underline{i}_j, \bar{i}_j]$, $j \in [m^{(i)}]$ takes the form of Structure 2. Further, let k be the largest value in $[i_s, i_t]$ such that $u_{i_s:k}$ is constant. If $u_{i_s} < u_{i_s - 1}$ and $u_k < u_{k+1}$, then we treat the first sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}_1 = i_s$. Similarly let k be smallest value in $[i_s, i_t]$ such that $u_{k:i_t}$ is constant. If $u_{k-1} > u_k$ and $u_{i_t} < u_{i_t+1}$ then we treat the last sub-interval in $\mathcal{P}_i$ as empty by putting $\underline{i}'_{m^{(i)}} = i_t + 1$.

3. In all sub-intervals $[\underline{i}'_j, \bar{i}'_j]$, $j \in [m^{(i)}]$, the offline optimal sequence can be split into piece-wise monotonic sections with at-most 2 pieces.

An illustration of this refinement scheme is given in Fig.6.

Let there be $m_1^{(i)}$ bins among $\{[\underline{i}_1, \bar{i}_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}]\}$ which satisfy the property in Lemma 21. Let their lengths be denoted by $\{\ell^{(1)}_{1^{(i)}}, \ldots, \{\ell^{(1)}_{m_1^{(i)}}\}$. These bins will be referred as *Type 1* bins henceforth.

Similarly let there be $m_2^{(i)}$ bins among $\{[\underline{i}_1, \bar{i}_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}]\}$ which satisfy either Case 1 or Case 2 in Lemma 22. Let their lengths be denoted by $\{\ell^{(2)}_{1^{(i)}}, \ldots, \{\ell^{(2)}_{m_2^{(i)}}\}$. These bins will be referred as *Type 2* bins henceforth.

Each bin in Type 1 and Type 2 can be paired with one adjacent bin (if non-empty) in $\mathcal{P}_i$ where the optimal sequence displays a piece-wise monotonic behaviour with at-most 2 pieces. (For example the bin $[\underline{i}_1, \bar{i}_1]$ can be paired with $[\underline{i}'_1, \bar{i}'_1]$ where in the later the optimal sequence displays a piece-wise monotonic behaviour. See Fig.6 for example.) To see why this is true, consider the case $\text{gap}_{\min}(B, [i_s, i_t]) \leq \text{gap}_{\min}(-B, [i_s, i_t)$. By construction, the optimal solution must preclude the form of Structure 2 in the bin $[\underline{i}'_k, \bar{i}'_k]$ where $k \in [m^{(i)}]$. This means the offline optimal can either take a non-increasing form in $[\underline{i}'_k, \bar{i}'_k]$ or it can monotonically increase and then optionally monotonically decrease. In both the cases, it can be split into at-most 2 sections where the solution is purely monotonic. Similar arguments apply for the case $\text{gap}_{\min}(B, [i_s, i_t]) > \text{gap}_{\min}(-B, [i_s, i_t)$.

Similarly, if bin $[i_s, \underline{i}_1 - 1]$ is non-empty then the offline optimal must assume a piece-wise monotonic structure with at-most 2 pieces. Then applying Lemma 20 to each of the 2 pieces separately and adding the regret bounds yields

$$\sum_{j=i_s}^{\underline{i}_1 - 1} f_j(x_j) - f_j(u_j) = \tilde{O}(1). \tag{21}$$

Note that $m_1^{(i)} + m_2^{(i)} = m^{(i)}$. Let the total regret contribution from Type 1 bins along with their pairs and Type 2 bins along with their pairs be referred as $R_1^{(i)}$ and $R_2^{(i)}$ respectively.

Since a sub-bin that is paired with a Type 1 or Type 2 bin can be split into at-most 2 sub-intervals where the optimal sequence is purely monotonic (see Fig. 6), we can bound the regret within such sub-bins $[\underline{i}'_k, \overline{i}'_k]$, $k \in [m^{(i)}]$ by at-most $62(G+B)^2 \log n$ by Lemma 20.

For a Type 2 bin $[a,b] \subseteq [i_s, i_t]$, we can have two possible configurations: If $\mathrm{gap}_{\min}(B, [i_s, i_t]) > \mathrm{gap}_{\min}(-B, [i_s, i_t])$ then $B - u_a \geq \mathrm{gap}_{\min}(B, [i_s, i_t]) \geq \mu_{\mathrm{th}}$ where the first inequality follows by the definition of $\mathrm{gap}_{\min}(B, [i_s, i_t])$ and the last inequality follows by Lemma 23. Similarly If $\mathrm{gap}_{\min}(B, [i_s, i_t]) \leq \mathrm{gap}_{\min}(-B, [i_s, i_t])$ then $B + u_a \geq \mathrm{gap}_{\min}(-B, [i_s, i_t]) \geq \mu_{\mathrm{th}}$. With this observation and using the results of Lemma 22, we can bound the regret contribution from any Type 2 bin and its pair as:

$$
\begin{aligned}
R_2^{(i)} &\leq \sum_{j=1^{(i)}}^{m_2^{(i)}} \left(\left(10(G+B)^2 \log n - \ell_j^{(2)} \mu_{\mathrm{th}}^2 \right) + 62(G+B)^2 \log n \right) \\
&\leq 72 m_2^{(i)} (G+B)^2 \log n - \mu_{\mathrm{th}}^2 \left(\sum_{j=1^{(i)}}^{m_2^{(i)}} \ell_j^{(2)} \right).
\end{aligned}
$$

From Eq.(19), we have $\ell_j^{(2)} \geq \frac{2\phi n^{1/3}}{(G+B) C_n^{1/3}}$ for $j \in \{1^{(i)}, \ldots, m_2^{(i)}\}$. So we can continue as

$$
\begin{aligned}
R_2^{(i)} &\leq 72 m_2^{(i)} (G+B)^2 \log n - \mu_{\mathrm{th}}^2 \frac{2\phi n^{1/3}}{(G+B) C_n^{1/3}} m_2^{(i)} \\
&= 0,
\end{aligned}
$$

where the last line is obtained by plugging in the value of μ_{th} from Lemma 23.

So by refining every interval in $\mathcal{P}$ that satisfy Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$
\sum_{i=1}^{M} R_2^{(i)} \leq 0, \tag{22}
$$

where we recall that $M := |\mathcal{P}| = O(n^{1/3} C_n^{2/3} \vee 1)$ and assign $R_2(i) = 0$ for intervals in $\mathcal{P}$ that do not satisfy Condition (A).

For any Type 1 bin, its regret contribution can be bounded by Lemma 21. So we have the regret contribution from Type 1 bins and their pairs bounded as

$$
\begin{aligned}
R_1^{(i)} &\leq \sum_{j=1}^{m_1^{(i)}} \left(\left(10(G+B)^2 \log n - \frac{4\lambda^2}{\ell_{j^{(i)}}^{(1)}} \right) + 62(G+B)^2 \log n \right) \\
&= 72 m_1^{(i)} (G+B)^2 \log n - 4\lambda^2 \sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}}
\end{aligned}
$$

By refining every interval in $\mathcal{P}$ that satisfies Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$
\begin{aligned}
\sum_{i=1}^{M} R_1^{(i)} &\leq 72(G+B)^2 \log n \sum_{i=1}^{M} m_1^{(i)} - 4\lambda^2 \sum_{i=1}^{M} \sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}} \\
&\leq 72(G+B)^2 M_1 \log n - 4\lambda^2 \frac{M_1^2}{n},
\end{aligned} \tag{23}
$$

where in the last line: a) we define $M_1 := \sum_{i=1}^{M} m_1^{(i)}$ with the convention that $m_1^{(i)} = 0$ if the bin $[i_s, i_t] \in \mathcal{P}$ doesn't satisfy Condition (A); b) applied AM-HM inequality and noted that $\sum_{i=1}^{M} \sum_{j=1}^{m_1^{(i)}} \ell_{j^{(i)}}^{(1)} \leq n$.

To further bound Eq.(23), we consider two separate regimes as follows.

Recall that $\lambda \geq \phi \frac{n^{1/3}}{C_n^{1/3}}$. So continuing from Eq.(23),

$$\begin{aligned}\sum_{i=1}^{M} R_1^{(i)} &\leq 72(G+B)^2 M_1 \log n - 4\phi^2 \frac{n^{2/3}}{C_n^{2/3}} \frac{M_1^2}{n} \\ &\leq 0, \end{aligned} \tag{24}$$

whenever $M_1 \geq \frac{18(G+B)^2 \log n}{\phi^2} n^{1/3} C_n^{2/3} = \tilde{\Omega}(n^{1/3} C_n^{2/3})$.

In the alternate regime where $M_1 \leq \frac{18(G+B)^2 \log n}{\phi^2} n^{1/3} C_n^{2/3} = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1)$, we trivially obtain:

$$\sum_{i=1}^{M} R_1^{(i)} = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1) \tag{25}$$

Putting everything together by combining the bounds in Eq.(21), (22), (24) and (25), we can bound the total regret contribution from the bins that satisfy Condition (A) as:

$$\sum_{i=1}^{M} R_1^{(i)} + R_2^{(i)} + \tilde{O}(1) = \tilde{O}(n^{1/3} C_n^{2/3} \vee 1),$$

where we have assigned $R_1^{(i)} = R_2^{(i)} = 0$ for bins that don't satisfy Condition (A).

Throughout the proof till now, we have only considered bins $[i_s, i_t] \in \mathcal{P}$ which satisfy Condition (A). Not meeting this criterion will only make the arguments easier as explained below.

If a bin $[i_s, i_t] \in \mathcal{P}$ doesn't satisfy Condition (A), by taking a logical negation of Condition (A), we conclude that this can only happen if the optimal solution precludes the form of either Structure 1 or Structure 2 (or both) within some sub-interval of $[i_s, i_t]$. Consequently by applying similar arguments we used to handle the bins $[\underline{i}'_k, \overline{i}'_k]$, $k \in [m^{(i)}]$, we can split the offline optimal sequence $u_{i_s:i_t}$ into at-most 2 piece-wise monotonic sections and use Lemma 20 to bound the regret in $[i_s, i_t]$ as $\tilde{O}(1)$. Since $|\mathcal{P}| = O(n^{1/3} C_n^{2/3} \vee 1)$, we conclude that the total regret from all bins that don't satisfy Condition (A) is $\tilde{O}(n^{1/3} C_n^{2/3} \vee 1)$. □

Proof. **of Theorem 1**. The proof is now immediate from Lemmas 19 and 24. □

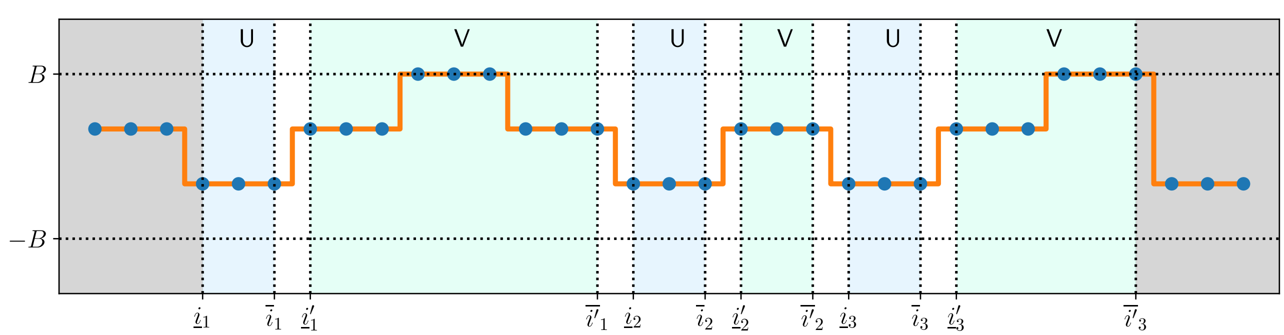

Figure 6: *Refinement of a bin that satisfy Condition (A) in the proof of Lemma 24 with $gap_{min}(-B, [i_s, i_t]) \geq gap_{min}(B, [i_s, i_t])$. Here we assign $i_s = \underline{i}_1$ and $i_t = \overline{i}'_3$ The following pairs are formed in the proof of Lemma 24: $\mathcal{P}_i = ([\underline{i}_1, \overline{i}_1], [\underline{i}'_1, \overline{i}'_1]), ([\underline{i}_2, \overline{i}_2], [\underline{i}'_2, \overline{i}'_2]), ([\underline{i}_3, \overline{i}_3], [\underline{i}'_3, \overline{i}'_3])$. Blue dots represent the optimal sequence*

D Proofs for Section 3

Corollary 5. *Let the loss functions f_t be H strongly convex in L_2 norm across the (box) domain $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_\infty \leq B\}$. i.e, $f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + \nabla f_t(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{H}{2}\|\boldsymbol{y} - \boldsymbol{x}\|_2^2$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$. Suppose $\|\nabla f_t(\boldsymbol{x})\|_\infty \leq G_\infty$ for all $\boldsymbol{x} \in \mathcal{D}$. For each $i \in [d]$, construct surrogate losses $\ell_t^{(i)} : \mathbb{R} \to \mathbb{R}$ as $\ell_t^{(i)}(x) = (x - (\boldsymbol{x}_t[i] - \nabla f_t(\boldsymbol{x}_t)[i]/H))^2$ where $\boldsymbol{x}_t$ is the prediction of the learner at time t. By running d instances of uni-variate FLH-OGD (Fig.4 in Appendix B) with decision set $[-B, B]$ and learning rate $\zeta = 1/(2(2B + G_\infty/H)^2)$ where instance i predicts $\boldsymbol{x}_t[i]$ at time t and suffers losses $\ell_t^{(i)}$, we have*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) = \tilde{O}\left(d^{1/3} n^{1/3} C_n^{2/3} \vee d\right),$$

for any comparator sequence $\boldsymbol{w}_{1:n}$ with $TV(\boldsymbol{w}_{1:n}) := \sum_{t=2}^{n}\|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \leq C_n$. $\tilde{O}(\cdot)$ hides the dependence on factors of $\log n, B, H, G_\infty$.

Proof. Due to strong convexity, we have for any $\boldsymbol{w}_t \in \mathbb{R}^d$,

$$\begin{aligned}
f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t) &\leq -\langle \nabla f_t(\boldsymbol{x}_t), \boldsymbol{w}_t - \boldsymbol{x}_t \rangle - \frac{H}{2}\|\boldsymbol{w}_t - \boldsymbol{x}_t\|^2 \\
&= H\left(\langle \nabla f_t(\boldsymbol{x}_t)/H, \boldsymbol{x}_t - \boldsymbol{x}_t \rangle + (1/2)\|\boldsymbol{x}_t - \boldsymbol{x}_t\|^2)\right) \\
&\quad - H\left(\langle \nabla f_t(\boldsymbol{x}_t)/H, \boldsymbol{w}_t - \boldsymbol{x}_t \rangle + (1/2)\|\boldsymbol{w}_t - \boldsymbol{x}_t\|^2)\right) \\
&= \sum_{i=1}^{d} H\left(\nabla f_t(\boldsymbol{x}_t)[i](\boldsymbol{x}_t[i] - \boldsymbol{x}_t[i])/H + (1/2)(\boldsymbol{x}_t[i] - \boldsymbol{x}_t[i])^2\right) \\
&\quad - H\left(\nabla f_t(\boldsymbol{x}_t)[i](\boldsymbol{w}_t[i] - \boldsymbol{x}_t[i])/H + (1/2)(\boldsymbol{w}_t[i] - \boldsymbol{x}_t[i])^2\right) \\
&= (H/2)\left(\sum_{i=1}^{d} \ell_t^{(i)}(\boldsymbol{x}_t[i]) - \ell_t^{(i)}(\boldsymbol{w}_t[i]),\right) \qquad (26)
\end{aligned}$$

where the last line is obtained by completing the squares. Let $\boldsymbol{u}_t \in \mathbb{R}^d$ for $t \in [n]$ be defined as the offline optimal sequence corresponding to the optimization problem:

$$\begin{aligned}
\min_{\tilde{\boldsymbol{u}}_1,\ldots,\tilde{\boldsymbol{u}}_n,\tilde{z}_1,\ldots,\tilde{\boldsymbol{z}}_{n-1}} \quad & \sum_{t=1}^{n}\sum_{i=1}^{d} \ell_t^{(i)}(\tilde{\boldsymbol{u}}_t[i]) \\
\text{s.t.} \quad & \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+1} - \tilde{\boldsymbol{u}}_t \ \forall t \in [n-1], \\
& \sum_{t=1}^{n-1}\|\tilde{\boldsymbol{z}}_t\|_1 \leq C_n, \\
& \|\tilde{\boldsymbol{u}}_t\|_\infty \leq B \ \forall t \in [n],
\end{aligned}$$

Let $C_n[i] = \sum_{t=2}^{n}|\boldsymbol{u}_t[i] - \boldsymbol{u}_{t-1}[i]|$ be its TV allocated to coordinate i. By Theorem 1, the FLH-OGD instance i with learning rate $\zeta = 1/(2(2B + G_\infty/H)^2)$ attains the regret of $\tilde{O}\left(n^{1/3}(C_n[i])^{2/3} \vee 1\right)$ regret. WLOG, let's assume that FLH-OGD instances for coordinates $i \in [k]$, $k \leq d$ incurs $\tilde{O}\left(n^{1/3}(C_n[i])^{2/3}\right)$ regret wrt losses $\ell_t^{(i)}$ and the regret incurred by FLH-OGD instances for coordinates $k > k'$ is $O(\log n)$. Let $R_n(\boldsymbol{w}_{1:n}) := \sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{w}_t)$

and $R'_n(\boldsymbol{w}_{1:n}) := (H/2)\left(\sum_{t=1}^n \sum_{i=1}^d \ell_t^{(i)}(\boldsymbol{x}_t[i]) - \ell_t^{(i)}(\boldsymbol{w}_t[i])\right)$. From Eq.(26) $R_n(\boldsymbol{w}_{1:n}) \le R'_n(\boldsymbol{w}_{1:n})$. We have,

$$
\begin{aligned}
R_n(\boldsymbol{w}_{1:n}) &\le R'_n(\boldsymbol{w}_{1:n}) \\
&\le \sup_{\substack{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n \in \mathcal{D} \\ \sum_{t=2}^n \|\boldsymbol{w}_t - \boldsymbol{w}_{t-1}\|_1 \le C_n}} R'_n(\boldsymbol{w}_{1:n}) \\
&= R'_n(\boldsymbol{u}_{1:n}) \\
&= (d-k)\tilde{O}(1) + \sum_{i=1}^k \tilde{O}\left(n^{1/3}(C_n[i])^{2/3}\right) \\
&\le (d-k)\tilde{O}(1) + \tilde{O}\left(n^{1/3}(k)^{1/3}\left(\sum_{i=1}^k C_n[i]\right)^{2/3}\right),
\end{aligned}
$$

where the last line follows by Holder's inequality $\boldsymbol{x}^T\boldsymbol{y} \le \|\boldsymbol{x}\|_3\|\boldsymbol{y}\|_{3/2}$, where we treat $\boldsymbol{x}$ as just a vector of ones in $\mathbb{R}^k$. The above expression can be further upper bounded by $\tilde{O}\left(2d \vee 2d^{1/3}n^{1/3}C_n^{2/3}\right)$. □

E Proofs for Section 4

We start by inspecting the KKT conditions.

Lemma 25. ***(characterization of offline optimal)*** *Consider the following convex optimization problem (where $\tilde{\boldsymbol{z}}_1, ..., \tilde{\boldsymbol{z}}_{n-1}$ are introduced as dummy variables).*

$$
\begin{aligned}
\min_{\tilde{\boldsymbol{u}}_1,\dots,\tilde{\boldsymbol{u}}_n,\tilde{\boldsymbol{z}}_1,\dots,\tilde{\boldsymbol{z}}_{n-1}} \quad & \sum_{t=1}^n f_t(\tilde{\boldsymbol{u}}_t) \\
\text{s.t.} \quad & \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{u}}_{t+1} - \tilde{\boldsymbol{u}}_t \;\forall t \in [n-1], \\
& \sum_{t=1}^{n-1} \|\tilde{\boldsymbol{z}}_t\|_1 \le C_n, \qquad (28a) \\
& \|\tilde{\boldsymbol{u}}_t\|_\infty \le B \;\forall t \in [n], \qquad (28b)
\end{aligned}
$$

Let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n, \boldsymbol{z}_1, \dots, \boldsymbol{z}_{n-1} \in \mathbb{R}^d$ be the optimal primal variables and let $\lambda \ge 0$ be the optimal dual variable corresponding to the constraint (28a). Further, let $\boldsymbol{\gamma}_t^+, \boldsymbol{\gamma}_t^- \in \mathbb{R}^d$ with $\boldsymbol{\gamma}_t^+ \ge \boldsymbol{0}$ and $\boldsymbol{\gamma}_t^- \ge \boldsymbol{0}$ be the optimal dual variables that correspond to constraint (28b). Specifically for $k \in [d]$, $\boldsymbol{\gamma}_t^+[k]$ corresponds to the dual variable for the constraint $\boldsymbol{u}_t[k] \le B$ induced by the relation (28b). Similarly $\boldsymbol{\gamma}_t^-[k]$ corresponds to the constraint $-B \le \boldsymbol{u}_t[k]$. By the KKT conditions, we have

- ***stationarity:*** $\nabla f_t(\boldsymbol{u}_t) = \lambda(\boldsymbol{s}_t - \boldsymbol{s}_{t-1}) + \boldsymbol{\gamma}_t^- - \boldsymbol{\gamma}_t^+$*, where $\boldsymbol{s}_t \in \partial|\boldsymbol{z}_t|$ (a subgradient). Specifically,* $\boldsymbol{s}_t[k] = \mathrm{sign}(\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k])$ *if $|\boldsymbol{u}_{t+1}[k] - \boldsymbol{u}_t[k]| > 0$ and $\boldsymbol{s}_t[k]$ is some value in $[-1, 1]$ otherwise. For convenience of notations later, we also define $\boldsymbol{s}_n = \boldsymbol{s}_0 = \boldsymbol{0}$.*
- ***complementary slackness:*** *(a) $\lambda(\sum_{t=2}^n \|\boldsymbol{u}_t - \boldsymbol{u}_{t-1}\|_1 - C_n) = 0$; (b) $\boldsymbol{\gamma}_t^-[k](\boldsymbol{u}_t[k] + B) = 0$ and $\boldsymbol{\gamma}_t^+[k](\boldsymbol{u}_t[k] - B) = 0$ for all $t \in [n]$ and all $k \in [d]$.*

The proof of the above lemma is similar to that of Lemma 3 and hence omitted.

Terminology. We will refer to the optimal primal variables $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ in Lemma 25 as the *offline optimal sequence* in this section. We reserve the term *FLH-ONS* for the instantiation of FLH with ONS as base learners with parameters as in Theorem 10.

Notations. For bin $[i_s, i_t] \in \mathcal{P}$ we define: $n_i = i_t - i_s + 1$, $\bar{\boldsymbol{u}}_i = \frac{1}{n_i}\sum_{j=i_s}^{i_t} \boldsymbol{u}_j$, $\boldsymbol{\Gamma}_i^+ = \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^+$, $\boldsymbol{\Gamma}_i^- = \sum_{j=i_s}^{i_t} \boldsymbol{\gamma}_j^-$, $\Delta\boldsymbol{s}_i = \boldsymbol{s}_{i_t} - \boldsymbol{s}_{i_s-1}$, $C_i = \sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1$.

For any general bin $[a, b]$ define the quantities $n_{a\to b}, \bar{\boldsymbol{u}}_{a\to b}, \boldsymbol{\Gamma}^+_{a\to b}, \boldsymbol{\Gamma}^-_{a\to b}, \Delta\boldsymbol{s}_{a\to b}, C_{a\to b}$ analogously as above.

The following is a direct extension for Lemma 17.

Lemma 26. ***(key partition)*** *Initialize $\mathcal{P} \leftarrow \Phi$. Starting from time 1, spawn a new bin $[i_s, i_t]$ whenever $\sum_{j=i_s+1}^{i_t+1} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 > B/\sqrt{n_i}$, where $n_i = i_t - i_s + 2$. Add the spawned bin $[i_s, i_t]$ to $\mathcal{P}$.*

Let $M := |\mathcal{P}|$. We have $M = O\left(1 \vee n^{1/3} C_n^{2/3} B^{-2/3}\right)$.

Proposition 27. *The losses f_t defined in Eq.(9) are:*

- *G^2 gradient Lipschitz over the domain $\mathcal{D}$ in Assumption B1*
- *Define $\gamma := 2GB\sqrt{\alpha d/2} + 1/\sqrt{2\alpha}$. Then the losses f_t are $\alpha' := 1/(2\gamma^2)$ exp-concave across $\mathcal{D}$.*
- *f_t are $G' := 2\alpha G^2 B\sqrt{d} + G$ Lipschitz in L2 norm across $\mathcal{D}$.*

Proof. The first two statements have been already proved in Section 4. For the last statement we have that

$$\nabla f_t(\boldsymbol{x}) = (\alpha \nabla \ell_t(\boldsymbol{x}_t)^T(\boldsymbol{x} - \boldsymbol{x}_t) + 1)\nabla \ell_t(\boldsymbol{x}_t).$$

So by triangle inequality we obtain that $\|\nabla f_t(\boldsymbol{x})\|_2 \leq 2\alpha G^2 B\sqrt{d} + G$. □

Lemma 28. ***(Strongly Adaptive regret)*** *((Hazan et al., 2007), (Hazan and Seshadhri, 2007)) Consider any bin $[a, b]$ and a comparator $\boldsymbol{w} \in \mathcal{D}$. Under Assumptions B1-2 in Section 4, the static regret of the FLH-ONS with losses f_t obeys*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{w}) \leq 10d(8G^2B^2\alpha d + 1/\alpha)\log n$$

where $\boldsymbol{x}_j$ are predictions of FLH-ONS and γ is as defined in Theorem 10.

Proof. Let $\alpha' = 1/(2\gamma^2)$. The static regret of ONS is $5d(G'D + 1/\alpha')\log n$ for α' exp-concave losses (Theorem 2 in (Hazan et al., 2007)) where D is the diameter of the decision set. We have $D = 2B\sqrt{d}$ for the box decision set. the static regret of ONS in our setting is at-most $5d(2G'B\sqrt{d} + 1/\alpha')\log n$.

The regret of the FLH against any of its base experts is at-most $(4/\alpha')\log n$ for α' exp-concave losses (Theorem 3.2 in (Hazan and Seshadhri, 2007)). Adding both these regret bounds, using Proposition 27 and further upper bounding the sum results in the lemma.

□

Lemma 29. ***(low $\boldsymbol{\lambda}$ regime)*** *If the optimal dual variable $\lambda = O\left(\frac{d^{1.5}n^{1/3}}{C_n^{1/3}}\right)$, we have the regret of FLH-ONS strategy bounded as*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t) = \tilde{O}\left(10d(8G^2B^2\alpha d + 1/\alpha)(n^{1/3}C^{2/3} \vee 1)\right),$$

where $\boldsymbol{x}_t$ is the prediction of FLH-ONS at time t.

Proof. Consider a bin $[i_s, i_t] \in \mathcal{P}$. Note that for any $j \in [i_s, i_t]$ and $k \in [d]$, both $\boldsymbol{\gamma}_j^+[k]$ and $\boldsymbol{\gamma}_j^-[k]$ can't be simultaneously non-zero due to complementary slackness and the fact that $C_i \leq B/\sqrt{n_i} < 2B$ by the construction in Lemma 26. For some fixed $\check{\boldsymbol{u}} \in \mathcal{D}$, we have

$$\underbrace{\sum_{j=i_s}^{i_t} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}})}_{T_{1,i}} + \underbrace{\sum_{j=i_s}^{i_t} f_j(\check{\boldsymbol{u}}) - f_j(\boldsymbol{u}_j)}_{T_{2,i}}.$$

By virtue of Lemma 28, we have $T_{1,i} = \tilde{O}(d^2)$. Due to gradient Lipschitzness in Proposition 27

$$T_{2,i} \leq \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle + \frac{G^2}{2} \|\check{\boldsymbol{u}} - \boldsymbol{u}_j\|_2^2. \tag{29}$$

We construct $\check{\boldsymbol{u}}$ as follows:

- If there exists a $j \in [i_s, i_t]$ and $k \in [d]$ such that $\boldsymbol{u}_j[k] = B$, set $\check{\boldsymbol{u}}[k] = B$.
- If there exists a $j \in [i_s, i_t]$ and $k \in [d]$ such that $\boldsymbol{u}_j[k] = -B$, set $\check{\boldsymbol{u}}[k] = -B$.
- If the optimal solution doesn't touch either boundaries $\pm B$ in $[i_s, i_t]$ across a coordinate, set $\check{\boldsymbol{u}}[k] = \boldsymbol{u}_{i_s}[k]$.

It is easy to see that $\check{\boldsymbol{u}} \in \mathcal{D}$ and $\|\check{\boldsymbol{u}} - \boldsymbol{u}_j\|_2 \leq \|\check{\boldsymbol{u}} - \boldsymbol{u}_j\|_1 \leq C_i$ for all $j \in [i_s, i_t]$. Using this observation along with the KKT conditions, we continue from Eq.(29) as

$$\begin{aligned}
T_{2,i} &\leq G^2 n_i C_i^2 + \sum_{j=i_s}^{i_t} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle, \\
&\leq G^2 n_i C_i^2 + \sum_{j=i_s}^{i_t} \lambda \langle \boldsymbol{s}_j - \boldsymbol{s}_{j-1}, \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle + \langle \boldsymbol{\gamma}_j^- - \boldsymbol{\gamma}_j^+, \check{\boldsymbol{u}} - \boldsymbol{u}_j \rangle \\
&\leq_{(a)} G^2 B^2 + \lambda \langle \boldsymbol{s}_{i_s - 1}, \boldsymbol{u}_{i_s} - \check{\boldsymbol{u}} \rangle - \lambda \langle \boldsymbol{s}_{i_t}, \boldsymbol{u}_{i_t} - \check{\boldsymbol{u}} \rangle + \lambda \sum_{j=i_s+1}^{i_t} \|\boldsymbol{u}_j - \boldsymbol{u}_{j-1}\|_1 \\
&\quad + \sum_{j=i_s}^{i_t} \sum_{k=1}^{d} \boldsymbol{\gamma}_j^-[k](\check{\boldsymbol{u}}[k] - \boldsymbol{u}_j[k]) - \gamma_j^+[k](\check{\boldsymbol{u}}[k] - \boldsymbol{u}_j[k]) \\
&\leq_{(b)} G^2 B^2 + 3\lambda C_i,
\end{aligned}$$

where line (a) is obtained by using that fact that $C_i \leq B/\sqrt{n_i}$ and a rearrangement of the summations and line (b) is obtained by noting that $\boldsymbol{\gamma}_j^-[k] = 0$ when $\boldsymbol{u}_j[k] > -B$ via complementary slackness and $\check{\boldsymbol{u}}[k] - \boldsymbol{u}_j[k]$ is zero when $\boldsymbol{u}_j[k] = -B$ since by construction of $\check{\boldsymbol{u}}$: $\check{\boldsymbol{u}}[k] = -B$ if $\boldsymbol{u}_j[k] = -B$ for some $j \in [i_s, i_t]$. Similar arguments are applied to show the terms including $\boldsymbol{\gamma}_j^+$ also sums to zero. In line (b) we also used the fact that $\langle \boldsymbol{s}_{i_s-1}, \boldsymbol{u}_{i_s} - \check{\boldsymbol{u}} \rangle \leq \|\boldsymbol{s}_{i_s-1}\|_\infty \|\boldsymbol{u}_{i_s} - \check{\boldsymbol{u}}\|_1 \leq C_i$. Similarly $\langle \boldsymbol{s}_{i_t}, \boldsymbol{u}_{i_t} - \check{\boldsymbol{u}} \rangle \leq C_i$

Hence summing $T_{1,i}$ and $T_{2,i}$ across all bins in $\mathcal{P}$ yields

$$\begin{aligned}
\sum_{i=1}^{M} T_{1,i} + T_{2,i} &\leq_{(a)} \tilde{O}(Md^2) + \lambda C_n \\
&\leq \tilde{O}\left(10d(8G^2B^2\alpha d + 1/\alpha)(n^{1/3}C^{2/3} \vee 1)\right),
\end{aligned}$$

where we recall that $M := |\mathcal{P}| = O(1 \vee n^{1/3} C_n^{2/3})$ by Lemma 26 and in line (a) we used $\sum_{i=1}^{M} C_i \leq C_n$ and $\lambda = O(d^{1.5} n^{1/3} / C_n^{1/3})$ by the premise of the current Lemma.

□

Definition 30. *For a bin $[a, b]$, the offline optimal is said to be* ***piece-wise maximally monotonic in*** $[\mathbf{a}, \mathbf{b}]$ ***with m pieces across some coordinate*** $\mathbf{k} \in [\mathbf{d}]$*, if we can split $[a, b]$ into m disjoint consecutive bins $[a_1, b_1], \ldots, [a_m, b_m]$ such that the offline optimal sequence within each $[a_i, b_i]$ is purely monotonic across coordinate k'. Further, right-extending any interval $[a_i, b_i]$ to $[a_i, b_i + 1]$ if $b_i + 1 \in [a, b]$ makes $\boldsymbol{u}_{a_i:b_i+1}[k']$ non-monotonic. The sections $[a_i, b_i]$ for $i \in [m]$ are termed* ***maximally monotonic sections****.*

Lemma 31. *The sequence returned at Step 3 of `generateGhostSequence` in Fig.7 has the following properties:*

`generateGhostSequence`: Inputs- (1) offline optimal sequence (2) two numbers $k_{\text{fix}} \in [d] \cup \{0\}$ and $u_{\text{fix}} \in [-B, B]$ (3) an interval $[a, b] \subseteq [n]$ where the offline optimal is piece-wise maximally monotonic with at-most 4 pieces across any coordinate $k \in [d]$.

1. Initialize $\mathcal{Q} \leftarrow \Phi$.
2. For each coordinate $k \in [d]$:
   (a) If k is same as k_{fix}, then set $\check{\boldsymbol{u}}_t[k] = u_{\text{fix}}$ for all $t \in [a, b]$. Goto Step 2.
   (b) If the optimal solution is constant across coordinate k, set $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_a[k]$ for all $t \in [a, b]$. Goto Step 2
   (c) If the optimal solution monotonically increases (decreases) first across coordinate k, then:
      i. Split $[a, b]$ into at-most 7 sub-bins – $[\underline{r}_i, \bar{r}_i]$, $i \in [7]$ – with the following properties:
         - $\underline{r}_1 = a$. $\underline{r}_2$ is the largest value in $[a, b]$ such that $\boldsymbol{u}_{\underline{r}_1:\underline{r}_2}$ is monotonically increasing (decreasing) and $\boldsymbol{u}_{\underline{r}_2}[k] \underset{(<)}{>} \boldsymbol{u}_{\underline{r}_2-1}[k]$. Set $\bar{r}_1 = \underline{r}_2 - 1$.
         - $\bar{r}_2$ is the largest value in $[a, b]$ such that $\boldsymbol{u}_{\underline{r}_2:\bar{r}_2}[k]$ is constant.
         - If $\bar{r}_2 = b$, then set $[\underline{r}_i, \bar{r}_i]$, $i \in [3, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_3 = \bar{r}_2 + 1$.
         - If $\boldsymbol{u}_{\underline{r}_3:b}[k]$ is a constant, set $\bar{r}_3 = b$. Set $[\underline{r}_i, \bar{r}_i]$, $i \in [4, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_4$ is the largest point in $[\underline{r}_3, b]$ such that $\boldsymbol{u}_{\underline{r}_3:\underline{r}_4}[k]$ is monotonically decreasing (increasing) and $u_{\underline{r}_4}[k] \underset{(>)}{<} u_{\underline{r}_4-1}[k]$. Set $\bar{r}_3 = \underline{r}_4 - 1$.
         - $\bar{r}_4$ is the largest point in $[\underline{r}_4, b]$ such that $\boldsymbol{u}_{\underline{r}_4:\bar{r}_4}[k]$ is constant.
         - If $\bar{r}_4 = b$, Set $[\underline{r}_i, \bar{r}_i]$, $i \in [5, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_5 = \bar{r}_4 + 1$.
         - If $\boldsymbol{u}_{\underline{r}_5:b}[k]$ is constant, then set $\bar{r}_5 = b$ and $[\underline{r}_i, \bar{r}_i]$, $i \in [6, 7]$ to be empty. Goto Step 2(c)(ii).
         - $\underline{r}_6$ is the largest point such that $\boldsymbol{u}_{\underline{r}_5:\underline{r}_6}[k]$ is monotonically increasing (decreasing) and $\boldsymbol{u}_{\underline{r}_6}[k] \underset{(<)}{>} \boldsymbol{u}_{\underline{r}_6-1}[k]$. Set $\bar{r}_5 = \underline{r}_6 - 1$.
         - $\bar{r}_6$ is the largest point in $[\underline{r}_6, b]$ such that $\boldsymbol{u}_{\underline{r}_6:\bar{r}_6}$ is constant.
         - If $\bar{r}_6 = b$, set $[\underline{r}_7, \bar{r}_7]$ as empty. Goto Step 2(c)(ii).
         - Set $\underline{r}_7 = \bar{r}_6 + 1$ and $\bar{r}_7 = b$.

      ii. Assign $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_1}[k]$ for all $t \in [\underline{r}_1, \bar{r}_1]$; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_2}[k]$ for all $t \in [\underline{r}_2, \bar{r}_2]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_3}[k]$ for all $t \in [\underline{r}_3, \bar{r}_3]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_4}[k]$ for all $t \in [\underline{r}_4, \bar{r}_4]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_5}[k]$ for all $t \in [\underline{r}_5, \bar{r}_5]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\underline{r}_6}[k]$ for all $t \in [\underline{r}_6, \bar{r}_6]$ if non-empty; $\check{\boldsymbol{u}}_t[k] = \boldsymbol{u}_{\bar{r}_7}[k]$ for all $t \in [\underline{r}_7, \bar{r}_7]$ if non-empty;
3. Return $\{\check{\boldsymbol{u}}_a, \ldots, \check{\boldsymbol{u}}_b\}$.

Figure 7: *`generateGhostSequence` procedure. If line 2(c) is replaced by "If the optimal solution monotonically decreases first across coordinate k, then", then we propagate that change by replacing the phrases increasing/decreasing and $>$ / $<$ in the lines below 2(c)(i) by the bracketed statements next to it.*

P1 The elements in the sequence changes only at-most 3d times. i.e, $\sum_{j=a+1}^{b} \mathbb{I}(\check{\boldsymbol{u}}_j \neq \check{\boldsymbol{u}}_{j-1}) \leq 7d$, *where* $\mathbb{I}(\cdot)$ *is the indicator function.*

P2 Every member of the sequence lie in the box decision set $\mathcal{D}$.

P3 For any $j \in [a, b]$, $\sum_{\substack{k=1 \\ k \neq k_{fix}}}^{d} |\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]| \leq C_{a \to b}$, *where* $C_{a \to b}$ *is the TV of the offline optimal in bin* $[a, b]$.

Proof. Observe that in the procedure detailed in Fig.7, we split the bin $[a, b]$ into at-most 7 bins across any coordinate. The value of the comparator across that coordinate stays unchanged in each of the new sub-bins. This implies that number of distinct comparators in $\{\check{\boldsymbol{u}}_a, \ldots, \check{\boldsymbol{u}}_b\}$ is at-most $7d$. It is also easy to see that each $\check{\boldsymbol{u}}_j$, $j \in [a, b]$ stays inside the decision set $\mathcal{D}$.

Note that for any $j \in [a, b]$ and any $k \in [d] \setminus \{k_{\text{fix}}\}$, $\check{\boldsymbol{u}}_j[k]$ coincides with the value of $\boldsymbol{u}_{j'}[k]$ for some $j' \in [a, b]$. This implies that $|\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]| \leq C_{a \to b}[k]$ for any $j \in [a, b]$, where $C_{a \to b}[k]$ is the TV of the optimal solution across coordinate k in bin $[a, b]$. So $\sum_{\substack{k=1 \\ k \neq k_{\text{fix}}}}^{d} |\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]| \leq \sum_{\substack{k=1 \\ k \neq k_{\text{fix}}}}^{d} C_{a \to b}[k] \leq C_{a \to b}$. Thus Property 3 is true. $\square$

Lemma 32. *(**monotonic bins**) Consider a bin $[a, b]$ with length ℓ where the offline optimal sequence is piece-wise maximally monotonic in $[a, b]$ across any coordinate with at-most 4 pieces. Let the TV of the optimal solution within bin $[a, b]$ denoted by $C_{a \to b}$ be at-most $B/\sqrt{\ell}$. Then we have the regret of FLH-ONS strategy in this bin bounded as*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha) \log n,$$

where $\boldsymbol{x}_j$ are the predictions of the FLH-ONS lagorithm.

Proof. We first construct a useful sequence of comparators:
$\check{\boldsymbol{u}}_{a:b} = \texttt{generateGhostSequence}\ (\boldsymbol{u}_{1:n}, k_{\text{fix}} = 0, u_{\text{fix}} = 0, [a, b])$.

We remark that as $k_{\text{fix}} = 0 \notin [d]$, the condition in Step 2(a) of Fig.7 is never satisfied.

Next, we employ a two term regret decomposition as follows

$$\underbrace{\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_j)}_{T_1} + \underbrace{\sum_{j=a}^{b} f_j(\check{\boldsymbol{u}}_j) - f_j(\boldsymbol{u}_j)}_{T_2}.$$

By noting that there are only at-most $7d$ change points in the comparator sequence (see Lemma 31), we can sum up the SA regret guarantee from Lemma 28 against each of the constant sections of $\check{\boldsymbol{u}}_{a:b}$ to obtain

$$T_1 \leq 70d^2(8G^2B^2\alpha d + 1/\alpha) \log n.$$

To bound T_2 we use gradient Lipschitzness in Proposition 27 and look at a coordinate-wise decomposition.

$$\begin{aligned} T_2 &\leq \sum_{j=a}^{b} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}}_j - \boldsymbol{u}_j \rangle + \frac{G^2}{2} \|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_2^2 \\ &\leq \frac{\ell G^2 C_{a \to b}^2}{2} + \sum_{k=1}^{d} \sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]), \end{aligned} \tag{30}$$

where in the last line we used that fact that $\|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_2^2 \leq \|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_1^2 \leq C_{a \to b}^2$ by Property 3 of Lemma 31, where $C_{a \to b}$ is the TV of the optimal solution within bin $[a, b]$.

Since $C_{a \to b} \leq B\sqrt{\ell}$, we have the first term in Eq.(30) bounded by $\frac{G^2B^2}{2}$. Next we proceed to bound the second term in Eq.(30) coordinate-wise. Consider a coordinate $k \in [d]$. We have two cases:

Case 1: When the optimal solution across coordinate k in bin $[a, b]$ has a structure described in Step 2(b) of the `generateGhostSequence` procedure of Fig.7. In this case $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_j[k] = \boldsymbol{u}_a[k]$ for $j \in [a, b]$. So

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0.$$

Case 2: When the optimal solution across coordinate k in bin $[a,b]$ has a structure described in Step 2(c) of the `generateGhostSequence` procedure of Fig.7. In this case, we can write

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = \sum_{i=1}^{7} \sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]),$$

where $[\underline{r}_i, \bar{r}_i]$, $i \in [7]$ are as defined in `generateGhostSequence` of Fig.7.

From Step 2(c)(ii) we have for each $i \in \{2,4,6\}$, $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_j[k] = \boldsymbol{u}_{\underline{r}_i}[k]$ for all $j \in [\underline{r}_i, \bar{r}_i]$ if non-empty. So $\sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0$ for each $i \in \{2,4,6\}$.

Next we consider the interval $[\underline{r}_1, \bar{r}_1]$. If within bin $[\underline{r}_1, \bar{r}_1]$, the optimal solution across coordinate k is constant, then $\sum_{j=\underline{r}_1}^{\bar{r}_1} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0$. Otherwise let $[\underline{r}_1, \bar{r}_1] = [\underline{r}_1, p] \cup [p+1, \bar{r}_1]$ such that the optimal solution is constant in $[\underline{r}_1, p]$ and non-decreasing (non-increasing) within $[p+1, \bar{r}_1]$ across coordinate k. Recall from Fig.7 that $\underline{r}_1 = a$. Since $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_a[k]$ for all $j \in [\underline{r}_1, p]$ we get $\sum_{j=\underline{r}_1}^{p} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0$. Further note that due to the presence of bins $[\underline{r}_1, p]$ and $[\underline{r}_2, \bar{r}_2]$ the solution $\boldsymbol{u}_j[k]$ for $j \in [p+1, \bar{r}_1]$ will never touch the boundaries $\pm B$. So by the KKT conditions and using $\check{\boldsymbol{u}}_j[k] = \boldsymbol{u}_a[k]$ for $j \in [p+1, \bar{r}_1]$, we have

$$\begin{aligned}
\sum_{j=p+1}^{\bar{r}_1} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) &= \sum_{j=p+1}^{\bar{r}_1} \lambda(\boldsymbol{s}_j[k] - \boldsymbol{s}_{j-1}[k])(\boldsymbol{u}_a[k] - \boldsymbol{u}_j[k]) \\
&= \lambda \left(\boldsymbol{s}_p[k](\boldsymbol{u}_{p+1}[k] - \boldsymbol{u}_a[k]) - \boldsymbol{s}_{\bar{r}_1}[k](\boldsymbol{u}_{\bar{r}_1}[k] - \boldsymbol{u}_a[k])\right) \\
&\quad + \lambda \sum_{j=p+2}^{\bar{r}_1} |\boldsymbol{u}_j[k] - \boldsymbol{u}_{j-1}[k]| \\
&= 0, \qquad (31)
\end{aligned}$$

where the last line is obtained as follows: Observe that $\boldsymbol{s}_p[k] = \boldsymbol{s}_{\bar{r}_1}[k] = 1(\text{or } -1)$ and $\boldsymbol{s}_p[k]\boldsymbol{u}_{p+1}[k] - \boldsymbol{s}_{\bar{r}_1}[k]\boldsymbol{u}_{\bar{r}_1}[k] = -C_{p+1\to\bar{r}_1}$ due to monotonicity of $\boldsymbol{u}_{p+1:\bar{r}_1}$

By using similar arguments we used to show Eq.(31), it can be proved that

$$\begin{aligned}
\sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) &= \sum_{j=\underline{r}_i}^{\bar{r}_i} \nabla f_j(\boldsymbol{u}_j)[k](\boldsymbol{u}_{\underline{r}_i}[k] - \boldsymbol{u}_j[k]) \\
&= 0,
\end{aligned}$$

for $i \in \{3,5\}$.

Further, by using similar arguments we used to handle $[\underline{r}_1, \bar{r}_1]$, it can be shown that

$$\sum_{j=\underline{r}_7}^{\bar{r}_7} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) = 0.$$

Thus overall by combining Case 1 and 2 and continuing from Eq.(30), we have $T_2 \le G^2B^2/2$. Thus the total regret

$$\begin{aligned}
T_1 + T_2 &\le 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n + G^2B^2/2 \\
&\le 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n,
\end{aligned}$$

which concludes the proof.

□

Definition 33. *We introduce the following definitions for convenience.*

- *For a bin $[a,b] \subseteq \{2,\ldots,n-1\}$, the offline optimal solution is said to assume Structure 1 across coordinate k if $\boldsymbol{u}_j[k] = \boldsymbol{u}_a[k] \in (-B,B)$ for all $j \in [a,b]$ and $\boldsymbol{u}_b[k] > \boldsymbol{u}_{b+1}[k]$ and $\boldsymbol{u}_a[k] > \boldsymbol{u}_{a-1}[k]$.*
- *For a bin $[a,b] \subseteq \{2,\ldots,n-1\}$, the offline optimal solution is said to assume Structure 2 across coordinate k if $\boldsymbol{u}_j[k] = \boldsymbol{u}_a[k] \in (-B,B)$ for all $j \in [a,b]$ and $\boldsymbol{u}_b[k] < \boldsymbol{u}_{b+1}[k]$ and $\boldsymbol{u}_a[k] < \boldsymbol{u}_{a-1}[k]$.*
- *A bin $[r,s]$ is said to contain Structure 1 and Structure 2 if across some coordinate k, the offline optimal solution assumes the form of Structure 1 in an interval $[a,b] \subset [r,s]$ and Structure 2 in some interval $[a',b'] \subset [r,s]$ with $[a,b] \cap [a',b'] = \Phi$.*
- *For a bin $[a,b]$, we define $GAP_{min}(\beta,[a,b])[k] := \min_{j\in[a,b]} |\boldsymbol{u}_j[k] - \beta|$, where $\beta \in \mathbb{R}$.*

Next we provide a lemma analogous to Lemma 21.

Lemma 34. *Consider a bin $[a,b]$ with length ℓ where the TV of the offline optimal obeys $C_{a\to b} \le B/\sqrt{\ell}$. Assume that for some coordinate $k' \in [d]$, $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 1 or Structure 2. Further suppose that across all coordinates, the offline optimal solution is piece-wise maximally monotonic in $[a,b]$ with at-most 4 pieces. If $\left|\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k']\right| \le B$, then*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \le 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{2\lambda^2}{\ell G^2},$$

where $\boldsymbol{x}_j$ are the predictions of FLH-ONS.

Proof. Let $k_{\text{fix}} = k'$ and $u_{\text{fix}} = \boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k']$. Consider a comparator sequence $\check{\boldsymbol{u}}_{a:b} = \texttt{generateGhostSequence}(\boldsymbol{u}_{1:n}, k_{\text{fix}}, u_{\text{fix}}, [a,b])$. We use a two term regret decomposition

$$\underbrace{\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\check{\boldsymbol{u}}_j)}_{T_1} + \underbrace{\sum_{j=a}^{b} f_j(\check{\boldsymbol{u}}_j) - f_j(\boldsymbol{u}_j)}_{T_2}. \tag{32}$$

By Properties 1 and 2 in Lemma 31, we know that the comparator $\check{\boldsymbol{u}}_{a:b}$ changes only at-most $7d$ times and every single point in the sequence belongs to $\mathcal{D}$. Hence by strong adaptivity (Lemma 28), we have

$$T_1 \le 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n. \tag{33}$$

Further via gradient Lipschitzness in Proposition 27,

$$\begin{aligned}
T_2 &\le \sum_{j=a}^{b} \langle \nabla f_j(\boldsymbol{u}_j), \check{\boldsymbol{u}}_j - \boldsymbol{u}_j\rangle + \frac{G^2}{2}\|\check{\boldsymbol{u}}_j - \boldsymbol{u}_j\|_2^2\\
&= \sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2\right)\\
&\quad + \sum_{\substack{k=1\\k\ne k'}}^{d}\sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k]) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k])^2\right)\\
&\le \frac{G^2B^2}{2} + \sum_{\substack{k=1\\k\ne k'}}^{d}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k] - \boldsymbol{u}_j[k])\\
&\quad + \sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2\right),
\end{aligned}$$

where in the last line we have used the facts that $\sum_{\substack{k=1\\k\neq k'}}^{d}(\check{\boldsymbol{u}}_j[k]-\boldsymbol{u}_j[k])^2 \leq \left(\sum_{\substack{k=1\\k\neq k'}}^{d}|\check{\boldsymbol{u}}_j[k]-\boldsymbol{u}_j[k]|\right)^2 \leq C_{a\to b}^2 \leq B^2/\ell$ by Property 3 of Lemma 31 and the TV constraint assumed in the premise of the current lemma.

Since the optimal solution across any coordinate is piece-wise maximally monotonic with at-most 4 pieces, by following the same arguments used in Case 1 and 2 in the proof of Lemma 32, we can write

$$\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k](\check{\boldsymbol{u}}_j[k]-\boldsymbol{u}_j[k]) = 0, \tag{34}$$

for any $k \neq k'$.

Recall that $\boldsymbol{u}_j[k'] = \boldsymbol{u}_a[k'] \in (-B,B)$ for all $j \in [a,b]$. Further by our construction, $\check{\boldsymbol{u}}_j[k'] = \boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b}\nabla f_j(\boldsymbol{u}_j)[k']$, for all $j \in [a,b]$. The key observation is to realize that $(\check{\boldsymbol{u}}_j[k']-\boldsymbol{u}_j[k'])$ stays at a constant value for all $j \in [a,b]$. So we have

$$\begin{aligned}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k']-\boldsymbol{u}_j[k']) &= (\check{\boldsymbol{u}}_a[k']-\boldsymbol{u}_a[k'])\sum_{j=a}^{b}\nabla f_j(\boldsymbol{u}_j)[k']\\ &= \frac{-1}{\ell G^2}\left(\sum_{j=a}^{b}\nabla f_j(\boldsymbol{u}_j)[k']\right)^2. \end{aligned}\tag{35}$$

Further we have,

$$\sum_{j=a}^{b} \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k']-\boldsymbol{u}_j[k'])^2 = \frac{1}{2\ell G^2}\left(\sum_{j=a}^{b}\nabla f_j(\boldsymbol{u}_j)[k']\right)^2. \tag{36}$$

Combining Eq.(35) and (36), we get

$$\begin{aligned}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k']-\boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k']-\boldsymbol{u}_j[k'])^2 &= \frac{-1}{2\ell G^2}\left(\sum_{j=a}^{b}\nabla f_j(\boldsymbol{u}_j)[k']\right)^2\\ &=_{(a)} \frac{-1}{2\ell G^2}\left(\lambda \Delta \boldsymbol{s}_{a\to b}[k']\right)^2\\ &= (b)\frac{-2\lambda^2}{\ell G^2},\end{aligned}$$

where line (a) is due to the KKT conditions and the fact that $\boldsymbol{u}_j[k'] \in (-B,B)$ thus making $\boldsymbol{\gamma}_j^+[k'] = \boldsymbol{\gamma}_j^-[k'] = 0$ and line (b) is due to the fact that $|\Delta \boldsymbol{s}_{a\to b}[k']| = 2$ for Structure 1 and Structure 2.

Hence overall we have shown that $T_2 \leq \frac{G^2B^2}{2} - \frac{2\lambda^2}{\ell G^2}$. Combining with Eq.(33) we conclude that the total regret of the FLH-ONS strategy within the bin $[a,b]$ is bounded by

$$\begin{aligned}T_1 + T_2 &\leq 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n + \frac{G^2B^2}{2} - \frac{2\lambda^2}{\ell G^2}\\ &\leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{2\lambda^2}{\ell G^2}.\end{aligned}$$

□

Lemma 35. *Consider a bin $[a,b]$ with length ℓ where the TV of the offline optimal obeys $C_{a\to b} \leq B/\sqrt{\ell}$. Assume that for some coordinate $k' \in [d]$, $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 1 or Structure 2. Further suppose that across all coordinates, the offline optimal solution is piece-wise maximally monotonic in $[a,b]$ with at-most 2 pieces.*

Case 1: *When $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 1 and $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \geq B$, then*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2,$$

and

Case 2: *When $\boldsymbol{u}_{a:b}[k']$ takes the form of Structure 2 and $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \leq -B$, then*

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}(B + \boldsymbol{u}_a[k'])^2,$$

where $\boldsymbol{x}_j$ are the predictions of FLH-ONS.

Proof. We consider Case 1. The arguments for the alternate case are similar. We proceed in a similar way as in the proof of Lemma 34. Let $k_{\text{fix}} = k'$ and $u_{\text{fix}} = B$. Consider a comparator sequence $\check{\boldsymbol{u}}_{a:b} =$ $\texttt{generateGhostSequence}(\boldsymbol{u}_{1:n}, k_{\text{fix}}, u_{\text{fix}}, [a, b])$. We use a two term regret decomposition as in Eq.(32). Using similar argumets as in the proof of Lemma 34, we have

$$T_1 \leq 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n.$$

Bounding T_2 in a similar fashion as in the proof of Lemma 34, we have

$$T_2 \leq \frac{G^2B^2}{2} + \sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2\right), \tag{37}$$

where we have used Eq.(34) for bounding the cross terms for coordinates $k \neq k'$

The main difference is in how we handle the last term of Eq.(37). Recall that $\check{\boldsymbol{u}}_j[k'] = B$ and $\boldsymbol{u}_j[k'] = \boldsymbol{u}_a[k']$ for all $j \in [a, b]$. So

$$\sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2\right) = \frac{G^2\ell}{2}(B - \boldsymbol{u}_a[k'])^2 - 2\lambda(B - \boldsymbol{u}_a[k']), \tag{38}$$

where the last line is obtained via the KKT conditions and the fact that $\Delta\boldsymbol{s}_{a\to b}[k'] = -2$ for Case 1. (Recall that $|\boldsymbol{u}_a[k']| < B$ by the definition of Structure 1. So by complementary slackness $\boldsymbol{\gamma}_j^+[k'] = \gamma_j^-[k'] = 0$.)

By the premise of the lemma for Case 1, we have $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \geq B$. Again by using the KKT conditions and noting that $\Delta s_{a\to b}[k'] = -2$, we conclude that

$$\lambda \geq \frac{(B - \boldsymbol{u}_a[k'])\ell G^2}{2}.$$

Plugging this lower bound for λ to Eq.(38) and noting that $(B - \boldsymbol{u}_a[k']) \geq 0$, we get

$$\sum_{j=a}^{b}\left(\nabla f_j(\boldsymbol{u}_j)[k'](\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k']) + \frac{G^2}{2}(\check{\boldsymbol{u}}_j[k'] - \boldsymbol{u}_j[k'])^2\right) \leq \frac{-\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2.$$

Hence overall, we conclude that

$$\begin{aligned} T_1 + T_2 &\leq 70d^2(8G^2B^2\alpha d + 1/\alpha)\log n + \frac{G^2B^2}{2} - \frac{\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2 \\ &\leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}(B - \boldsymbol{u}_a[k'])^2. \end{aligned}$$

$\square$

`fineSplit`: Input - (1) offline optimal sequence $\boldsymbol{u}_{1:n}$ (2) an interval $[r, s] \subseteq [n]$. Across some coordinate $k \in [d]$, the offline optimal solution must take the form of both Structure 1 and 2 or either one of them at-least two times within some appropriate sub-intervals of $[r, s]$.

1. Initialize $\mathcal{Q} \leftarrow \Phi$, $\mathcal{Q}' \leftarrow \Phi$.
2. For each coordinate $k \in [d]$ across which the optimal solution takes the form of Structures 1 and 2 or either one of them at-least two times within some appropriate sub-intervals of $[r, s]$:
   (a) if $\text{GAP}_{\min}(B, [r, s])[k] > \text{GAP}_{\min}(-B, [r, s])[k]$ then add intervals $[a, b] \subset [r, s]$ where the offline optimal across coordinate k assumes the form of Structure 1 to $\mathcal{Q}$.
   (b) if $\text{GAP}_{\min}(B, [r, s])[k] \leq \text{GAP}_{\min}(-B, [r, s])[k]$ then add intervals $[a, b] \subset [r, s]$ where the offline optimal across coordinate k assumes the form of Structure 2 to $\mathcal{Q}$.
3. For each bin $[a, b] \in \mathcal{Q}$ if there exists another interval $[p, q] \in \mathcal{Q}$ with $[p, q] \subseteq [a, b]$, then remove $[a, b]$ from $\mathcal{Q}$.
4. Sort intervals in $\mathcal{Q}$ in increasing order of the left endpoints. (i.e $[a, b] < [p, q]$ if $a < p$).
5. Starting from the first bin, for each bin $[a, b] \in \mathcal{Q}$:
   (a) if there exists an interval $[p, q] \in \mathcal{Q}$ such that $a < p$ and $b < q$, then remove $[p, q]$ from $\mathcal{Q}$
6. Add disjoint and maximally continuous intervals that are the subsets of $[r, s] \setminus \{\cup_{[a,b] \in \mathcal{Q}} [a, b]\}$ to $\mathcal{Q}'$ such that the interval $[r, s]$ can be fully covered by disjoint intervals from $\mathcal{Q}$ and $\mathcal{Q}'$.
7. Return $(\mathcal{Q}, \mathcal{Q}')$.

Figure 8: *`fineSplit` procedure.*

Lemma 36. *Suppose `fineSplit` is invoked with input $[r, s]$ such that $C_{r \to s} \leq B/\sqrt{s - r + 1}$. The offline optimal solution within any bin $[a, b] \in \mathcal{Q}$ at Step 7 of `fineSplit` procedure in Fig.8 is piece-wise maximally monotonic in $[a, b]$ with at-most 4 pieces across any coordinate $k \in [d]$. Further there exists a coordinate $k \in [d]$ that satisfy one of the following conditions:*

1. *The offline optimal within bin $[a, b]$ takes the form of Structure 1 across coordinate k and $B - \boldsymbol{u}_a[k] \geq GAP_{min}(B, [r, s])[k] \geq GAP_{min}(-B, [r, s])[k]$.*
2. *The offline optimal within bin $[a, b]$ takes the form of Structure 2 across coordinate k and $B + \boldsymbol{u}_a[k] \geq GAP_{min}(-B, [r, s])[k] \geq GAP_{min}(B, [r, s])[k]$.*

Proof. We start by a basic observation.

FACT 1: Note that $C_{r \to s} \leq B/\sqrt{s - r + 1} \leq B$. So the $\boldsymbol{u}_{r:s}[k']$ cannot touch both B and $-B$ boundaries.

Consider a bin $[a, b] \in \mathcal{Q}$. By the construction of `fineSplit`, there exists a coordinate $k \in [d]$ across which the optimal solution stays constant within $[a, b]$ and assumes the form of Structure 1 or 2. For the sake of contradiction, let's assume that for some $k' \in [d]$, with $k' \neq k$, the optimal solution is maximally monotonic in $[a, b]$ with at-least 5 pieces across the coordinate k'. This can happen only when the optimal solution increases (decreases) then decreases (increases) then increases (decreases) then decreases (increases) and finally increase (decrease) again within bin $[a, b]$ and evolve arbitrarily there on-wards. Combined with FACT 1, such a behaviour can result in one of the following configurations across the coordinate k':

- Both Structure 1 and Structure 2 are formed.
- Only Structure 2 is formed at-least two times. This means that if $[x, y] \subset [a, b]$ is a maximally monotonic section with $\boldsymbol{u}_{x:y}[k']$ increasing, then $\boldsymbol{u}_y[k'] = B$. Then $\text{GAP}_{\min}(-B, [r, s])[k'] > \text{GAP}_{\min}(B, [r, s])[k'] = 0$.

- Only Structure 1 is formed at-least two times. This means that if $[x, y] \subset [a, b]$ is a maximally monotonic section with $\boldsymbol{u}_{x:y}[k']$ decreasing, then $\boldsymbol{u}_y[k'] = -B$. Then $\text{GAP}_{\min}(B, [r, s])[k'] > \text{GAP}_{\min}(-B, [r, s])[k'] = 0$.

In all of the above cases, at-least one sub-interval of $[a, b]$ will be added to $\mathcal{Q}$ at Step 2(a) or 2(b). This would imply that at Step 3, the bin $[a, b]$ is removed from $\mathcal{Q}$ and never added again resulting in a contradiction.

The last statement of the Lemma is immediate from Steps 2(a)-(b) of `fineSplit`. ☐

Lemma 37. *Suppose `fineSplit` is invoked with input $[r, s]$ such that $C_{r \to s} \leq B/\sqrt{s - r + 1}$. The offline optimal solution within any interval $[p, q] \in \mathcal{Q}'$ at Step 7 of `fineSplit` procedure in Fig.8 is piece-wise maximally monotonic in $[p, q]$ with at-most 4 pieces across any coordinate.*

Proof. Consider a coordinate $k \in [d]$ and a bin $[p, q] \in \mathcal{Q}'$. We provide the arguments for the case when $\text{GAP}_{\min}(-B, [r, s])[k] \geq \text{GAP}_{\min}(B, [r, s])[k]$. The arguments for the complementary case are similar. We start by stating two facts.

FACT 1: $\text{GAP}_{\min}(-B, [p, q])[k] > 0$.

To see this, assume for the sake of contradiction that $\text{GAP}_{\min}(-B, [p, q])[k] = 0$. Then this means that $\text{GAP}_{\min}(-B, [p, q])[k] = \text{GAP}_{\min}(B, [r, s])[k] = 0$. So the optimal solution across coordinate k, $\boldsymbol{u}_{r:s}[k]$ must touch both B and $-B$ at distinct time points in $[r, s]$. This would violate the TV constraint that $C_{p \to q} \leq B/\sqrt{s - r + 1} \leq B$, thus yielding a contradiction.

FACT2: It is not the case that there exists two intervals $[p_1, q_1], [p_2, q_2] \subset [p, q]$ within which the offline optimal takes the form of Structure 2 across the coordinate $k \in [d]$.

Let's prove the above fact via contradiction. Assume that there exists $[p_1, q_1], [p_2, q_2] \subset [p, q] \in \mathcal{Q}'$ such that the offline optimal takes the form of Structure 2 within them across the coordinate $k \in [d]$. Then $[p_i, q_i]$ $(i = 1, 2)$ must have been added to $\mathcal{Q}$ in step 2(b) of `fineSplit`. Since intervals in $\mathcal{Q}$ don't overlap with intervals in $\mathcal{Q}'$ due to Step 6, this would mean that the interval $[p_i, q_i]$ $(i = 1, 2)$ got removed from $\mathcal{Q}$ later.

Case 1: Consider the case where $[p_i, q_i]$ $(i = 1, 2)$ has been removed at Step 5(a). This means that there exists an interval $[a, b] \subseteq [r, s]$ where the offline optimal has Structure 1 or 2 across some coordinate $k' \neq k$ and $[p_i, q_i] \cap [a, b] \neq \Phi$. Observe that $[a, b]$ is never removed from $\mathcal{Q}$ since we are processing bins in sorted order at Step 4-5. This would contradict the fact that intervals in $\mathcal{Q}$ don't overlap with intervals in $\mathcal{Q}'$ due to Step 6.

Case 2: Consider the case where $[p_i, q_i]$ $(i = 1, 2)$ has been removed at Step 3. This means that there exists an interval $[x, y] \subseteq [p_i, q_i]$ where the offline optimal assumes Structure 1 or 2 across some coordinate $k' \neq k$. If $[x, y]$ is present in the final $\mathcal{Q}$ in Step 7, then this would again warrant a contradiction to the non-overlapping property between the intervals of $\mathcal{Q}$ and $\mathcal{Q}'$. If $[x, y]$ is removed at a later point through Step 5(a), by using similar arguments as in Case 1 yields a contradiction. Thus we conclude that the FACT 2 is true.

FACT 3: It is not the case that there exists two intervals $[p_1, q_1], [p_2, q_2] \subset [p, q]$ within the offline optimal takes the form of Structure 1 in $[p_1, q_1]$ and Structure 2 in $[p_2, q_2]$ across the coordinate $k \in [d]$.

The above fact can be proven using similar arguments that are used in proving FACT 2.

In light of FACT 1, FACT 2 and FACT 3, we conclude the statement of the lemma.

☐

Next we introduce a structural lemma analogous to Lemma 16.

Lemma 38. ***(λ-length lemma)*** *Consider a bin $[a, b] \subseteq \{2, \ldots, n - 1\}$ with length ℓ. Suppose that within this bin, the offline optimal solution sequence assumes the form of Structure 1 or Structure 2 across some coordinate $k \in [d]$, then $\lambda \leq \frac{G_\infty \ell}{2}$, where G_∞ is as in Assumption B2.*

Proof Sketch. The arguments for this proof are almost identical to that used for proving Lemma 16. We outline the parts where there are differences. We provide the arguments for Structure 2. Structure 1 can be handled similarly. Let the optimal sign assignments across coordinate k be written as $\boldsymbol{s}_j[k] = -1 + \epsilon_j$ where $\epsilon_j \in [0, 2]$

and $j \in [a,b]$. From the KKT conditions, we can write:

$$\begin{aligned}\nabla f_a(\boldsymbol{u}_a)[k] &= \lambda \epsilon_a \\ \nabla f_{a+1}(\boldsymbol{u}_{a+1})[k] &= \lambda(\epsilon_{a+1} - \epsilon_a) \\ &\vdots \\ \nabla f_{b-1}(\boldsymbol{u}_{b-1})[k] &= \lambda(\epsilon_{b-1} - \epsilon_{b-1}) \\ \nabla f_b(\boldsymbol{u}_b)[k] &= \lambda(2 - \epsilon_{b-1})\end{aligned}$$

Define the vector $\boldsymbol{z} = [\epsilon_a, \epsilon_{a+1} - \epsilon_a, \ldots, 2 - \epsilon_{b-1}]^T$. As noted in the proof of Lemma 16, we must have $\|\boldsymbol{z}\|_\infty > 0$. Let j^* be such that $\|\boldsymbol{z}\|_\infty = |\boldsymbol{z}[j^*]|$. Then $\lambda = \nabla f_{a+j^*-1}(\boldsymbol{u}_{a+j^*-1})[k]/\|\boldsymbol{z}\|_\infty$. From the optimization problem considered in the proof of Lemma 16, we have $\|\boldsymbol{z}\|_\infty \geq 2/\ell$. Since $\|\nabla f_j(\boldsymbol{u}_j)\|_\infty \leq G_\infty$ for all $j \in [n]$ by Assumption B2, we have $\lambda = \nabla f_{a+j^*-1}(\boldsymbol{u}_{a+j^*-1})[k]/\|\boldsymbol{z}\|_\infty \leq (G_\infty \ell)/2$.

□

Lemma 39. ***(large margin bins)*** *Assume that $\lambda \geq d^{1.5}\phi \frac{n^{1/3}}{C_n^{1/3}}$ for a constant $\phi = \sqrt{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)}$ that does not depend on n and C_n. Consider a bin $[i_s, i_t] \in \mathcal{P}$ within which the offline optimal solution takes the form of Structure 1 or Structure 2 (or both) across a coordinate $k \in [d]$ for some appropriate sub-intervals of $[i_s, i_t]$. Let $\mu_{th} = \sqrt{\frac{140d^{1.5}(8G^2B^2\alpha+G^2B^2+1/\alpha)G_\infty C_n^{1/3}\log n}{G^2\phi n^{1/3}}}$. Then*
$GAP_{min}(-B,[i_s,i_t])[k] \vee GAP_{min}(B,[i_s,i_t])[k] \geq \mu_{th}$,
whenever $C_n \leq \left(\frac{B^2G^2\phi}{560d^{1.5}(8G^2B^2\alpha+G^2B^2+1/\alpha)G_\infty \log n}\right)^3 n = \tilde{O}(n)$.

Proof. Suppose $\mathrm{GAP}_{\min}(-B,[i_s,i_t])[k] < \mu_{\mathrm{th}}$. Then the largest value of the optimal solution across coordinate k attained within this bin $[i_s,i_t]$ is at-most $-B + \mu_{\mathrm{th}} + B/\sqrt{n_i}$ (recall $n_i := i_t - i_s + 1$ and $C_i \leq B/\sqrt{n_i}$ due to Lemma 26). So $\mathrm{GAP}_{\min}(B,[i_s,i_t])[k] \geq 2B - \mu_{\mathrm{th}} - B/\sqrt{n_i}$. Our goal is to show that whenever C_n obeys the constraint stated in the lemma, we must have

$$2B - \mu_{\mathrm{th}} - B/\sqrt{n_i} \geq \mu_{\mathrm{th}}. \tag{39}$$

Let ℓ_i be the length of a sub-interval of $[i_s,i_t]$ where the offline optimal solution assumes the form of Structure 1 or Structure 2. Due to Lemma 38, we have

$$n_i \geq \ell_i \geq \frac{2\lambda}{G_\infty} \geq \frac{2d^{1.5}\phi n^{1/3}}{G_\infty C_n^{1/3}} \tag{40}$$

where the last inequality follows due to the condition assumed in the current lemma. So a sufficient condition for Eq.(39) to be true is

$$2B \geq 2\left(2\sqrt{\frac{140d^{1.5}(8G^2B^2\alpha + G^2B^2 + 1/\alpha)G_\infty C_n^{1/3}\log n}{G^2\phi n^{1/3}}} \vee B\sqrt{\frac{G_\infty C_n^{1/3}}{2d^{2.5/2}\phi n^{1/3}}}\right).$$

Recall that by Assumption B2, we have $G \wedge G_\infty \wedge B \geq 1$. So the above maximum will be attained by the first term and can be further simplified as

$$2B \geq 4\sqrt{\frac{140d^{1.5}(8G^2B^2\alpha + G^2B^2 + 1/\alpha)G_\infty C_n^{1/3}\log n}{G^2\phi n^{1/3}}}.$$

The above condition is always satisfied whenever $C_n \leq \left(\frac{B^2G^2\phi}{560d^{1.5}(8G^2B^2\alpha+G^2B^2+1/\alpha)G_\infty \log n}\right)^3 n$.

At this point, we have shown that $\mathrm{GAP}_{\min}(-B,[i_s,i_t])[k] < \mu_{\mathrm{th}} \implies \mathrm{GAP}_{\min}(B,[i_s,i_t])[k] \geq \mu_{\mathrm{th}}$ under the conditions of the lemma. Taking the contrapositive yields $\mathrm{GAP}_{\min}(B,[i_s,i_t])[k] < \mu_{\mathrm{th}} \implies \mathrm{GAP}_{\min}(-B,[i_s,i_t])[k] \geq \mu_{\mathrm{th}}$. □

Lemma 40. *(high λ regime) Suppose the optimal dual variable $\lambda \geq d^{1.5}\phi \frac{n^{1/3}}{C_n^{1/3}} = \Omega\left(\frac{n^{1/3}}{C_n^{1/3}}\right)$ for $\phi = \sqrt{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)}$ that does not depend on n and C_n. We have the regret of FLH-ONS strategy bounded as*

$$\sum_{t=1}^{n} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{u}_t) = \tilde{O}\left(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)\mathbb{I}\{C_n > 1/n\} + \tilde{O}\left(d(8G^2B^2\alpha d + 1/\alpha)\,\mathbb{I}\{C_n \leq 1/n\}\right),$$

where $\boldsymbol{x}_t$ is the prediction of FLH-ONS at time t and $\mathbb{I}\{\cdot\}$ is the boolean indicator function taking values in $\{0,1\}$.

Proof. Throughout the proof we assume that $C_n\left(\frac{B^2G^2\phi}{560d^{1.5}(8G^2B^2\alpha+G^2B^2+1/\alpha)G_\infty \log n}\right)^3 n$. Otherwise the trivial regret bound of $\tilde{O}(n)$ is near minimax optimal.

First we consider the regime where $C_n \geq 1/n$. It is useful to define the following annotated condition.

Condition (A): Let a bin $[r,s]$ be given. For some coordinate $k' \in [d]$, there exists disjoint intervals $[r_1,s_1],[r_2,s_2] \subset [r,s]$ that satisfy at-least one of the following: (i) $\boldsymbol{u}_{r_1:s_1}[k']$ has the form of Structure 1 and $\boldsymbol{u}_{r_2:s_2}[k']$ has the form of Structure 2; (ii) Both $\boldsymbol{u}_{r_1:s_1}[k']$ and $\boldsymbol{u}_{r_2:s_2}[k']$ have the form of Structure 1; (iii) Both $\boldsymbol{u}_{r_1:s_1}[k']$ and $\boldsymbol{u}_{r_2:s_2}[k']$ have the form of Structure 2.

The above condition is basically the prerequisite for the `fineSplit` procedure of Fig.8.

Let $[i_s,i_t] \in \mathcal{P}$ be a bin that satisfy Condition (A) for a coordinate $k' \in [d]$. Here $\mathcal{P}$ is the partition obtained in Lemma 26.

Let $(\mathcal{Q},\mathcal{Q}')$ be the collections of intervals obtained by invoking the `fineSplit` procedure with the bin $[i_s,i_t]$ as input. Let's write $\mathcal{Q} \cup \mathcal{Q}' \cup \{\Phi\}$ as a collection of disjoint consecutive intervals as follows:

$$\mathcal{Q} \cup \mathcal{Q}' \cup \{\Phi\} := \{[i_s, \underline{i}_1 - 1], [\underline{i}_1, \bar{i}_1], [\underline{i}'_1, \bar{i}'_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}], [\underline{i}'_{m^{(i)}}, \bar{i}'_{m^{(i)}}]\}, \tag{41}$$

with $\bar{i}'_{m^{(i)}} = i_t$.

Here we follow the convention that the bins $[\underline{i}_p, \bar{i}_p] \in \mathcal{Q}$ and $[\underline{i}'_p, \bar{i}'_p] \in \mathcal{Q}' \cup \{\Phi\}$ for all $p \in [m^{(i)}]$. Similar to the proof of Lemma 24, for enforcing this convention, we may have to set either of the bins $[i_s, \underline{i}_1 - 1]$ or $[\underline{i}'_{m^{(i)}}, \bar{i}'_{m^{(i)}}]$ to be empty. More precisely, if i_s belongs to some interval in $\mathcal{Q}$, then we set the first sub-interval $[i_s, \bar{i}_1 - 1]$ to be empty by setting $\bar{i}_1 = i_s$. Similarly, if i_t belongs to some interval in $\mathcal{Q}$, we treat the sub-interval $[\underline{i}'_{m^{(i)}}, \bar{i}'_{m^{(i)}}]$ as empty by setting $\underline{i}'_{m^{(i)}} = i_t + 1$. Further some of the intervals: $[\underline{i}'_k, \bar{i}'_k]$, $k \in [m^{(i)}]$ can be empty. For example if $\underline{i}_{k+1} = \bar{i}_k + 1$, then $[\underline{i}'_k, \bar{i}'_k]$ is treated as empty.

Note that if the first sub-interval $[i_s, \underline{i}_1 - 1]$ is non-empty then it must belong to $\mathcal{Q}'$ according to our convention. By Lemma 37 and Lemma 32,

$$\sum_{j=i_s}^{\underline{i}_1 - 1} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) = \tilde{O}\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n\right). \tag{42}$$

We proceed to bound the regret in $[\underline{i}_1, \bar{i}'_{m^{(i)}}]$. Let $\mathcal{P}_1^{(i)}$ denote the collection of bins among $\mathcal{Q} = \{[\underline{i}_1, \bar{i}_1], \ldots, [\underline{i}_{m^{(i)}}, \bar{i}_{m^{(i)}}]\}$ which satisfy the property in Lemma 34. Let $|\mathcal{P}_1^{(i)}| := m_1^{(i)}$ and their lengths be denoted by $\{\ell^{(1)}_{1^{(i)}}, \ldots, \ell^{(1)}_{m_1^{(i)}}\}$. These bins will be referred as *Type 1* bins henceforth.

Similarly let $\mathcal{P}_2^{(i)} = \mathcal{Q} \setminus \mathcal{P}_1^{(i)}$ which satisfy either of the properties in Lemma 35. Let $|\mathcal{P}_2^{(i)}| := m_2^{(i)}$ and their lengths be denoted by $\{\ell^{(2)}_{1^{(i)}}, \ldots, \ell^{(2)}_{m_2^{(i)}}\}$. These bins will be referred as *Type 2* bins henceforth. A bin $[a,b] \in \mathcal{P}_2^{(i)}$ satisfy at-least one of the following properties

P1: For some coordinate $k \in [d]$, the offline optimal satisfy the condition of Case 1 in Lemma 35 and $B - \boldsymbol{u}_a[k] \geq \mu_{th}$.

P2: For some coordinate $k \in [d]$, the offline optimal satisfy the condition of Case 2 in Lemma 35 and $B + \boldsymbol{u}_a[k] \geq \mu_{th}$.

To see this, let's inspect the way in which the bin $[a,b]$ has been added to $\mathcal{Q}$ when we invoke `fineSplit` with the input bin $[i_s, i_t]$. If $[a,b]$ has been added via Step 2-(a), then we have $\mathrm{GAP}_{\min}(B, [i_s,i_t])[k] > \mathrm{GAP}_{\min}(-B, [i_s,i_t])[k]$ for a coordinate k. By Lemma 39 it holds that $\mathrm{GAP}_{\min}(B, [i_s,i_t])[k] \geq \mu_{\mathrm{th}}$ under the C_n regime we consider. So $B - \boldsymbol{u}_a[k] \geq \mathrm{GAP}_{\min}(B, [i_s,i_t])[k] \geq \mu_{\mathrm{th}}$ where the first inequality follows by the definition of GAP (see Definition 33). Further, observe that $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] < -B$ is never satisfied, where $\ell = b - a + 1$. Otherwise it will imply that $-\frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] = \frac{2\lambda}{\ell G^2} < -B - \boldsymbol{u}_a[k'] \leq 0$ which is not true as $\lambda \geq 0$. We must also have $\boldsymbol{u}_a[k'] - \frac{1}{\ell G^2}\sum_{j=a}^{b} \nabla f_j(\boldsymbol{u}_j)[k'] \notin [-B, B]$. Otherwise, bin $[a,b]$ would have been already added to $\mathcal{P}_1^{(i)}$ and would have never present in $\mathcal{P}_2^{(i)}$. So we conclude that property P1 follows. Property P2 can also be shown to be true using similar arguments when the bin $[a,b]$ has been added to $\mathcal{Q}$ via Step 2-(b) of `fineSplit`.

Each bin $[\underline{i}_k, \bar{i}_k]$, $k \in [m^{(i)}]$ of Type 1 and Type 2 can be paired with an adjacent bin $[\underline{i}'_k, \bar{i}'_k] \in \mathcal{Q}' \cup \{\Phi\}$, $k \in [m^{(i)}]$ which is either empty or the optimal sequence displays a piece-wise maximally monotonic behaviour in $[\underline{i}'_k, \bar{i}'_k]$ across all coordinates as recorded in Lemma 37.

Note that $m_1^{(i)} + m_2^{(i)} = m^{(i)}$. Let the total regret contribution from Type 1 bins along with their pairs and Type 2 bins along with their pairs be referred as $R_1^{(i)}$ and $R_2^{(i)}$ respectively.

For a bin $[a,b] \in \mathcal{P}_2^{(i)}$, in either of the cases covered by the properties P1 and P2, we have by Lemma 35 that

$$\sum_{j=a}^{b} f_j(\boldsymbol{x}_j) - f_j(\boldsymbol{u}_j) \leq 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell G^2}{2}\mu_{\mathrm{th}}^2,$$

Let $[a',b'] \in \mathcal{Q}' \cup \{\Phi\}$ be the pair assigned to $[a,b]$. If it is non-empty, then due to Lemma 37 and Lemma 32 the regret from the bin $[a',b']$ is at-most $70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n$.

So we can bound $R_2^{(i)}$ as

$$\begin{aligned} R_2^{(i)} &\leq \sum_{j=1^{(i)}}^{m_2^{(i)}} \left(\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{\ell_j^{(2)} G^2}{2}\mu_{\mathrm{th}}^2 \right) + 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n \right) \\ &\leq m_2^{(i)} 140 d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{G^2\mu_{\mathrm{th}}^2}{2}\left(\sum_{j=1^{(i)}}^{m_2^{(i)}} \ell_j^{(2)}\right), \end{aligned}$$

From Eq.(40), we have $\ell_j^{(2)} \geq \frac{2\phi d^{1.5} n^{1/3}}{G_\infty C_n^{1/3}}$ for $j \in \{1^{(i)}, \ldots, m_2^{(i)}\}$ under the regime of λ we consider. So we can continue as

$$\begin{aligned} R_2^{(i)} &\leq 140 m_2^{(i)} d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - G^2\mu_{\mathrm{th}}^2 m_2^{(i)} \frac{\phi d^{1.5} n^{1/3}}{G_\infty C_n^{1/3}} \\ &\leq 140 m_2^{(i)} d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\log n - G^2\mu_{\mathrm{th}}^2 m_2^{(i)} \frac{\phi d^{1.5} n^{1/3}}{G_\infty C_n^{1/3}} \\ &= 0, \end{aligned}$$

where the last line is obtained by plugging in the value of μ_{th} as in Lemma 39.

So by refining every interval in $\mathcal{P}$ (recall that $\mathcal{P}$ is from Lemma 26) that satisfy Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$\sum_{i=1}^{M} R_2^{(i)} \leq 0,$$

where we recall that $M := |\mathcal{P}| = O(n^{1/3}C_n^{2/3} \vee 1)$ and assign $R_2(i) = 0$ for intervals in $\mathcal{P}$ that do not satisfy Condition (A).

For any Type 1 bin, its regret contribution can be bounded by Lemma 34. The regret contribution from its pair can be bounded by Lemma 32 as before. So we have

$$
\begin{aligned}
R_1^{(i)} &\leq \sum_{j=1}^{m_1^{(i)}} \left(\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{2\lambda^2}{\ell_{j^{(i)}}^{(1)} G^2} \right) + 70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n \right) \\
&= 140 m_1^{(i)} d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n - \frac{2\lambda^2}{G^2}\sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}}.
\end{aligned}
$$

So by refining every interval in $\mathcal{P}$ that satisfy Condition (A) and summing the regret contribution from all Type 2 bins and their pairs across all refined intervals in $\mathcal{P}$ yields

$$
\begin{aligned}
\sum_{i=1}^{M} R_1^{(i)} &\leq 140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\log n \sum_{i=1}^{M} m_1^{(i)} - \frac{2\lambda^2}{G^2}\sum_{i=1}^{M}\sum_{j=1}^{m_1^{(i)}} \frac{1}{\ell_{j^{(i)}}^{(1)}} \\
&\leq 140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha) M_1 \log n - \frac{2\lambda^2}{G^2}\frac{M_1^2}{n}, \\
&\leq 140d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha) M_1 \log n - \frac{2\lambda^2}{G^2}\frac{M_1^2}{n} \qquad (43)
\end{aligned}
$$

where in the last line: a) we define $M_1 := \sum_{i=1}^{M} m_1^{(i)}$ with the convention that $m_1^{(i)} = 0$ if the i^{th} bin in $\mathcal{P}$ doesn't satisfy Condition (A); b) applied AM-HM inequality and noted that $\sum_{i=1}^{M}\sum_{j=1}^{m_1^{(i)}} \ell_{j^{(i)}}^{(1)} \leq n$.

To further bound Eq.(43), we consider two separate regimes as follows.

Recall that $\lambda \geq d^{1.5}\phi \frac{n^{1/3}}{C_n^{1/3}}$. So continuing from Eq.(43),

$$
\begin{aligned}
140d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha) M_1 \log n - 2\lambda^2 \frac{M_1^2}{G^2 n} &\leq 140d^3(8G^2B^2\alpha + G^2B^2 + 1/\alpha) M_1 \log n \\
&\quad - 2d^{2.5}\phi^2 \frac{n^{2/3}}{C_n^{2/3}} \frac{M_1^2}{G^2 n} \\
&\leq 0,
\end{aligned}
$$

whenever $M_1 \geq \frac{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\log n}{\phi^2} n^{1/3}C_n^{2/3} = \tilde{\Omega}(n^{1/3}C_n^{2/3})$.

In the alternate regime where $M_1 \leq \left(\frac{70(8G^2B^2\alpha + G^2B^2 + 1/\alpha)\log n}{\phi^2} n^{1/3}C_n^{2/3} \vee 1 \right) = \tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$, we trivially obtain $\sum_{i=1}^{M} R_1^{(i)} = \tilde{O}\left(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)$.

The regret contribution from all sub-bins that starts at i_s $i \in [M]$ which are not paired in Eq.(41) is only at-most $\tilde{O}(d^{2.5}(n^{1/3}C_n^{2/3} \vee 1))$ by adding the bound of Eq.(42) across all $O(n^{1/3}C_n^{2/3} \vee 1)$ bins in $\mathcal{P}$.

Throughout the entire proof we have assumed that $m_1^{(i)}$ and $m_2^{(i)}$ are non-zero for some bin $[i_s, i_t] \in \mathcal{P}$. Not meeting this criterion will only make the arguments easier as explained below.

We have shown that the total regret contribution from the refined bins $\sum_{i=1}^{M} R_1^{(i)} + R_2^{(i)} = \tilde{O}(n^{1/3}C_n^{2/3} \vee 1)$, we trivially obtain $\sum_{i=1}^{M} R_1^{(i)} = \tilde{O}\left(140d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)$ under the conditions of the lemma, where we have taken $R_1^{(i)} = R_2^{(i)} = 0$ if the i^{th} bin $[i_s, i_t] \in \mathcal{P}$ doesn't satisfy Condition (A) across any coordinate.

If a bin doesn't satisfy Condition (A) across any coordinate, then the offline optimal solution within that bin assumes a piece-wise maximally monotonic structure with at-most 4 pieces across any coordinate. By Lemma 32, the regret within such bins is $\tilde{O}\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)\right)$. Since there can be at-most $O(n^{1/3}C_n^{2/3} \vee 1)$ such bins in $\mathcal{P}$, the total regret contribution from those bins is again $\tilde{O}\left(70d^2(8G^2B^2\alpha d + G^2B^2 + 1/\alpha)(n^{1/3}C_n^{2/3} \vee 1)\right)$. Now putting everything together yields the lemma.

If $C_n \leq 1/n$, then we have

$$\begin{aligned}\sum_{t=1}^{n} f_t(\boldsymbol{x}_j) - f_t(\boldsymbol{u}_t) &\leq \sum_{t=1}^{n} f_t(\boldsymbol{x}_j) - f_t(\boldsymbol{u}_1) + \sum_{t=1}^{n} f_t(\boldsymbol{u}_1) - f_t(\boldsymbol{u}_t)\\ &\leq_{(a)} \tilde{O}\left(10d(8G^2B^2\alpha d + 1/\alpha)\log n\right) + GnC_n\\ &= \tilde{O}\left(d(8G^2B^2\alpha d + 1/\alpha\right)\end{aligned}$$

where line (a) follows from the fact that f_t is G Lipschitz. □

Proof. **of Theorem 10**. The proof is immediate from the results of Lemmas 29 and 40.

□

F Reparametrization of certain polytopes to box

Proposition 41. *Consider an online problem with losses f_t that are α exp-concave on the decision set $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{c} \leq \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ such that $\boldsymbol{A}$ is full rank and $\boldsymbol{0} < \boldsymbol{b} - \boldsymbol{c}$.*

We can reparametrize this into an equivalent online learning problem with losses $\tilde{f}_t(\boldsymbol{z}) = f_t(\boldsymbol{A}^{-1}(\boldsymbol{D}^{-1}(\boldsymbol{z}+\boldsymbol{1})+\boldsymbol{c}))$ that are α exp-concave on the decision set $\tilde{\mathcal{D}} = \{\ \boldsymbol{z} \in \mathbb{R}^d : \|\boldsymbol{z}\|_\infty \leq 1\}$, where $\boldsymbol{D} = \mathrm{diag}(2/(\boldsymbol{b}[1] = \boldsymbol{c}[1]), \ldots, 2/(\boldsymbol{b}[d] - \boldsymbol{c}[d]))$ and $\boldsymbol{1}$ is the vector of ones in $\mathbb{R}^d$.

Further if the losses f_t are G Lipschitz in $\mathcal{D}$, then the losses $\tilde{f}_t$ are $\|\boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\|_{op} G$ Lispchitz in $\tilde{\mathcal{D}}$.

Proof. We have,

$$\begin{aligned}&\boldsymbol{c} \leq \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\\ \iff &\boldsymbol{0} \leq \boldsymbol{A}\boldsymbol{x} - \boldsymbol{c} \leq \boldsymbol{b} - \boldsymbol{c}.\end{aligned}$$

Then we have $\boldsymbol{0} \leq \boldsymbol{D}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{c}) \leq (2)\boldsymbol{1}$. This equivalent to $-\boldsymbol{1} \leq \boldsymbol{D}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{c}) - \boldsymbol{1} \leq \boldsymbol{1}$. By putting $\boldsymbol{z} = \boldsymbol{D}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{c}) - \boldsymbol{1}$ we can rewrite the original decision set as $\|\boldsymbol{z}\|_\infty \leq 1$.

Since $\boldsymbol{A}$ is full rank, there is a one-one mapping between the original decision set $\mathcal{D}$ and the new decision set $\tilde{\mathcal{D}} := \{\boldsymbol{z} \in \mathbb{R}^d : \|\boldsymbol{z}\|_\infty \leq 1\}$. Given a $\boldsymbol{z} \in \tilde{\mathcal{D}}$, we can find the corresponding point $\boldsymbol{x} \in \mathcal{D}$ as $\boldsymbol{x} = \boldsymbol{A}^{-1}(\boldsymbol{D}^{-1}(\boldsymbol{z}+\boldsymbol{1})+\boldsymbol{c})$. So the losses in the new parametrization becomes $\tilde{f}_t(\boldsymbol{z}) = f_t(\boldsymbol{A}^{-1}(\boldsymbol{D}^{-1}(\boldsymbol{z}+\boldsymbol{1})+\boldsymbol{c}))$.

Let $B := \boldsymbol{A}^{-1}\boldsymbol{D}^{-1}$ and $\boldsymbol{d} := \boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\boldsymbol{1} + \boldsymbol{A}^{-1}\boldsymbol{c}$ so that $\tilde{f}_t(\boldsymbol{z}) = f_t(\boldsymbol{B}\boldsymbol{z}+\boldsymbol{d})$. Then we have

$$\begin{aligned}\nabla \tilde{f}_t(\boldsymbol{z}) &= \boldsymbol{B}^T \nabla f_t(\boldsymbol{B}\boldsymbol{z}+\boldsymbol{d})\\ &= \boldsymbol{B}^T \nabla f_t(\boldsymbol{x}),\end{aligned}$$

for a point $\boldsymbol{x} = (\boldsymbol{B}\boldsymbol{z}+\boldsymbol{d}) \in \mathcal{D}$.

Similarly

$$\begin{aligned}\nabla^2 \tilde{f}_t(\boldsymbol{z}) &= \boldsymbol{B}^T \nabla^2 f_t(\boldsymbol{B}\boldsymbol{z}+\boldsymbol{d})\boldsymbol{B}\\ &= \boldsymbol{B}^T \nabla^2 f_t(\boldsymbol{x})\boldsymbol{B}.\end{aligned}$$

From the above two equations we can easily verify that $\nabla^2 \tilde{f}_t(\boldsymbol{z}) \succcurlyeq \alpha \nabla \tilde{f}_t(\boldsymbol{z}) \nabla \tilde{f}_t(\boldsymbol{z})^T$ as the functions f_t itself are α exp-concave in $\mathcal{D}$.

Further by Holder's inequality we have $\|\nabla \tilde{f}_t(\boldsymbol{z})\| \leq \|\boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\|_{\mathrm{op}} \|\nabla f_t(\boldsymbol{x})\|_2 \leq \|\boldsymbol{A}^{-1}\boldsymbol{D}^{-1}\|_{\mathrm{op}} G$.

□